Monograph:

Billings, S.A. and Voon, W.S.F. (1986) Piecewise Linear Identification of Nonlinear Systems. Research Report. Acse Report 299 . Dept of Automatic Control and System Engineering. University of Sheffield

PIECEWISE LINEAR IDENTIFICATION OF

NONLINEAR SYSTEMS.

by

S. A. Billings B.Eng, Ph.D, C.Eng, MIEE, AFIMA, MInstMC,

W. S. F. Voon, B.Eng, Ph.D.

RESEARCH REPORT No. 299.

August 1986

Department of Control Engineering
The University of Sheffield
Mappin Street
SHEFFIELD S1 3JD

Abstract

Methods of identifying a series of piecewise linear models which approximate to a nonlinear system over some operating range are investigated. Both spatial linear models and models with signal dependent parameters are considered.

1. Introduction

Piecewise linear modelling is the process of developing a series of locally linear models which approximate to a nonlinear system over some defined operating range. The main advantage of the piecewise linear approach is that well known linear algorithms can be used to identify the models and develop control strategies in a manner which utilizes the wealth of knowledge and experience that is available for linear systems. The disadvantage is that the model developed may be a poor approximation to the nonlinear system under study so that simplicity in the analysis is obtained at the expense of an inadequate model and inefficient control.

There are several possible ways in which nonlinear systems can be approximated by locally linear models. White [1971] introduced the term exhaustive linearization to describe a system which can be linearized about each of the admissible inputs and showed that such systems can be represented exactly in terms of their linear behaviour. Unfortunately, to obtain such a representation the response of the nonlinear system to a vanishingly small test signal must be used to identify what White [1971] calls the linear operator. This introduces a considerable conflict because the effect of noise on the measurements causes the S/N ratio to decrease rapidly as the input is made vanishingly small and hence the identification of an accurate linear map becomes extremely difficult.

De Hoff and Rock [1979] introduced the idea of multiple linearization of nonlinear systems where the model form is chosen to allow separate models of the steady state response and off-equilibrium behaviour. This approach is simple and efficient when the nonlinear system has a slow moving input and operates close to the steady state.

Cyrot-Normand and Mien [1980] demonstrated that state-affine models can be constructed by combining several linear models identified from step tests for different levels of plant operation. Several processes were modelled using this approach and the technique was shown to work well for slowly varying levels of plant operation. A similar idea was considered by Haber and Vajk [1982] , Haber and

Keviczky [1985] who derived algorithms for the identification of linear models with signal dependent parameters to represent nonlinear systems. Diekmann and Unbehauen [1985] introduced an on-line algorithm and showed that the different sets of parameters in operating point dependent linear models can be estimated in parallel from one period of measurement.

In time series analysis Tong [1978a, b] introduced a class of models called threshold models. These are linear models where the parameters vary according to the amplitude of a finite number of past values of some variable. Threshold autoregressive (AR) and autoregressive moving average (ARMA) models and associated estimation algorithms have been derived by Tong [1978a, b] and used in the modelling of nonlinear time series. State dependent models (SDM's) were introduced by Priestley [1980] again for the study of time series. SDM's are state variable models which can be interpreted as locally linear ARMA models in which the evolution of the process is governed by a set of coefficients which depend on the system state at a previous time instant.

In the present study both the identification of spatial piecewise linear and signal dependent linear models of nonlinear systems are investigated. In section 2 the linearization of the NARMAX model (Nonlinear AutoRegressive Moving Average model with eXogenous inputs) [Leontaritis and Billings, 1985] is investigated and used as a basis for the development of a spatial piecewise linear model. An estimation algorithm for this model is developed in section 3. In section 4 the estimation of signal dependent linear models is considered and methods of patching or glueing together these models in an attempt to reconstruct an approximation to the nonlinear system description, are described. Simulation results are included for both these techniques and throughout the algorithms are compared and related to those already in the literature.

2. Linearization of the NARMAX Model

If a system is linear then it is finitely realizable and can be represented by the linear difference equation model

$$y(t) = \sum_{i=1}^{n_y} (a_i \; y(t-i) + \sum_{i=1}^{n_u} (b_i \; u(t-i)) \tag{1}$$

If the Hankel matrix of the system has finite rank. When the system is nonlinear a similar representation can be derived [Leontaritis and Billings, 1985] to yield the nonlinear difference equation model

$$y(t) = F^* \left[y(t-1), \ldots y(t-n_y), u(t-d), \ldots u(t-d-n_u+1)\right] \tag{2}$$

where $F^*[.]$ is some nonlinear function. The model of eqn (2) can be shown to exist providing the system under study is finite dimensional and a linearized model would exist if the system were operated close to an equilibrium point.

An equivalent representation can be derived [Leontaritis and Billings, 1985] for stochastic systems to yield the NARMAX model (nonlinear autoregressive moving average model with exogenous inputs)

$$z(t) = F\left[z(t-1), \ldots z(t-n_z), u(t-d), \ldots u(t-d-n_u+1), \varepsilon(t-1), \ldots \varepsilon(t-n_e)\right] + \varepsilon(t) \tag{3}$$

where $\varepsilon(t)$ represents the prediction errors. Because the NARMAX model can be used to represent a wide class of nonlinear systems this representation will be used as a basis for the development of piecewise linear identification algorithms for nonlinear processes.

The effects of noise will at this stage only complicate the analysis and consequently the representation of eqn (2) will be used to represent the nonlinear system.

Assuming that $F^*[.]$ is smooth enough to have a Taylor series representation the linearization of eqn (2) at the operating point Δ_k is

$$\delta y(t) = \sum_{i=1}^{n_y} \left.\frac{\partial F^*[.]}{\partial y(t-i)}\right|_{\Delta_k} \delta y(t-i) + \sum_{i=1}^{n_u} \left.\frac{\partial F^*[.]}{\partial u(t-d-i+1)}\right|_{\Delta_k} \delta u(t-d-i+1) \tag{4}$$

where $\Delta_k = [y(t-1), \ldots, y(t-n_y), u(t-d), \ldots, u(t-d-n_u+1)]$

The operating point can for simplicity, be redefined as

$$\Delta_k = [y_1, y_2, \ldots\ldots, y_{n_y}, u_1, u_2, \ldots\ldots, u_{n_u}] \tag{5}$$

and

$$\delta y(t) = y(t) - y(t)\Big|_{\Delta_k} \quad , \quad \text{for } t=1,2,\ldots,N$$

$$\delta u(t) = u(t) - u(t)\Big|_{\Delta_k} \quad , \quad \text{for } t=1\ ,2,\ldots,N \qquad (6)$$

Substituting equation (6) in equation (4) yields

$$y(t)=y(t)\Big|_{\Delta_k} + \sum_{i=1}^{n_y} \frac{\partial F^*[.]}{\partial y(t-i)}\Big|_{\Delta_k} y(t-i) - \sum_{i=1}^{n_u} \frac{\partial F^*[.]}{\partial y(t-i)}\Big|_{\Delta_k} y(t-i)\Big|_{\Delta_k}$$

$$+ \sum_{i=1}^{n_u} \frac{\partial F^*[.]}{\partial u(t-d-i+1)}\Big|_{\Delta_k} u(t-d-i+1) - \sum_{i=1}^{n_u} \frac{\partial F^*[.]}{\partial u(t-d-i+1)}\Big|_{\Delta_k} u(t-d-i+1)\Big|_{\Delta_k} \qquad (7)$$

which is valid in a small neighbourhood of Δ_k and can be written as

$$y(t) = \Theta o\Big|_{\Delta_k} + \sum_{i=1}^{n_y} \Theta_i\Big|_{\Delta_k} y(t-i) + \sum_{i=1}^{n_u} \Theta_{n_y+i}\Big|_{\Delta_k} u(t-d-i+1) \qquad (8)$$

where

$$\Theta o\Big|_{\Delta_k} = y(t)\Big|_{\Delta_k} - \sum_{i=1}^{n_y} \frac{\partial F^*[.]}{\partial y(t-i)}\Big|_{\Delta_k} y(t-i)\Big|_{\Delta_k}$$

$$- \sum_{i=1}^{n_u} \frac{\partial F^*[.]}{\partial u(t-d-i+1)}\Big|_{\Delta_k} u(t-d-i+1)\Big|_{\Delta_k} \qquad (9)$$

$$\Theta_i\Big|_{\Delta_k} = \frac{\partial F^*[.]}{\partial y(t-i)}\Big|_{\Delta_k} \quad , \ i=1,2,\ldots\ldots,n_y. \qquad (10)$$

$$\Theta_{n_y+i}\Big|_{\Delta_k} = \frac{\partial F^*[.]}{\partial u(t-d-i+1)}\Big|_{\Delta_k} \quad , \ i=1,2,\ldots\ldots,n_u. \qquad (11)$$

The linearized NARMAX model of equation (8) is this a locally linear model which is valid in a small region about a particular operating point Δ_k.

3. Spatial Piecewise Linear Models

Assuming that the system under study is relatively smooth a number of linear models of the form of eqn (8) could be used to provide a spatial piecewise linear approximation to the nonlinear system eqn (2).

In this case the model of eqn (8) can be rewritten as

$$y(t) = \Theta_o^{(k)} + \sum_{i=1}^{n_y} \Theta_i^{(k)} y(t-i) + \sum_{i=1}^{n_u} \Theta_{n_y+i}^{(k)} u(t-d-i+1) \qquad (12)$$

or

$$y(t) = L^{(k)} [1, y(t-1), \ldots y(t-n_y), u(t-d), \ldots u(t-d-n_y+1)] \qquad (13)$$

where $L^{(k)}$ [.] is a locally linear function of its argument about a particular operating point Δ_k and there are m operating points and hence m operating regions which cover the global space of operation. Spatial piecewise linear modelling would mean that the parameters in a large number of linearized models would need to be estimated each valid in a small region of operation. In practice a conflict may arise with this approach. To achieve a greater accuracy in the approximation, the region of operation of each model would need to be reduced. This in turn would mean that the signal to noise ratio also reduces, since the noise free output of the system is reduced but the noise remains unchanged, and hence the models which are estimated become inaccurate. A compromise is therefore required to achieve a reasonable degree of approximation without incurring an unacceptably low S/N ratio.

The spatial piecewise linear model is a global representation of a nonlinear system formulated as a family of locally linear models suspended together over the global space. Each of the individual locally linear models is independent and has no influence on any other locally linear model. The stability criteria developed for linear systems is therefore applicable within each operating region.

Within each region the accuracy of the locally linear model should be comparatively good providing the operating point is constrained to lie within this region. The disadvantages of this approach is the fairly large number of models which may be necessary to characterise the nonlinear system. This number increases exponentially as the number of intervals for each independent variable is increased. For example, a ten variable function where each variable has ten intervals results in 10^{10} linear models. Another problem is the selection of optimal operating regions for the locally piecewise linear models and this can be computationally time consuming.

3.1 Parameter Estimation for the Spatial Model

The estimation of each locally linear model suspended over the global space requires a set of global excitation data. The ideal choice of input would be the uniformly distributed random input within the maximum and minimum input constraints of the nonlinear process such that every subspace is equally weighted during the estimation of the locally linear models. The maximum and minimum values of the input (ie, u_{max}, u_{min}) are divided into a suitable choice of intervals, say m_u. The input intervals are $(u_{min}, u_1]$, $(u_1, u_2]$,..., $(u_{m_u-1}, u_{m_u}]$. Similarly, the maximum and minimum values of the output (i.e. y_{max}, y_{min}) are divided into a suitable choice of intervals, say m_y. The output intervals are $(y_{min}, y_1]$, $(y_1, y_2]$,..., $(y_{m_y-1}, y_{m_y}]$. For simplicity, let equation (12) be a second order linear system with $n_y=n_u=2$, d=1 and $m_u=m_y=3$ say. Then

$$y(t)=\Theta_0^{(k)} +\Theta_1^{(k)}y(t-1) +\Theta_2^{(k)}y(t-2)+\Theta_3^{(k)}u(t-1)+\Theta_4^{(k)}u(t-2) \qquad (14)$$

for the operating region indexed as k and the spatial piecewise linear model can be represented diagramatically as shown in Fig.1.

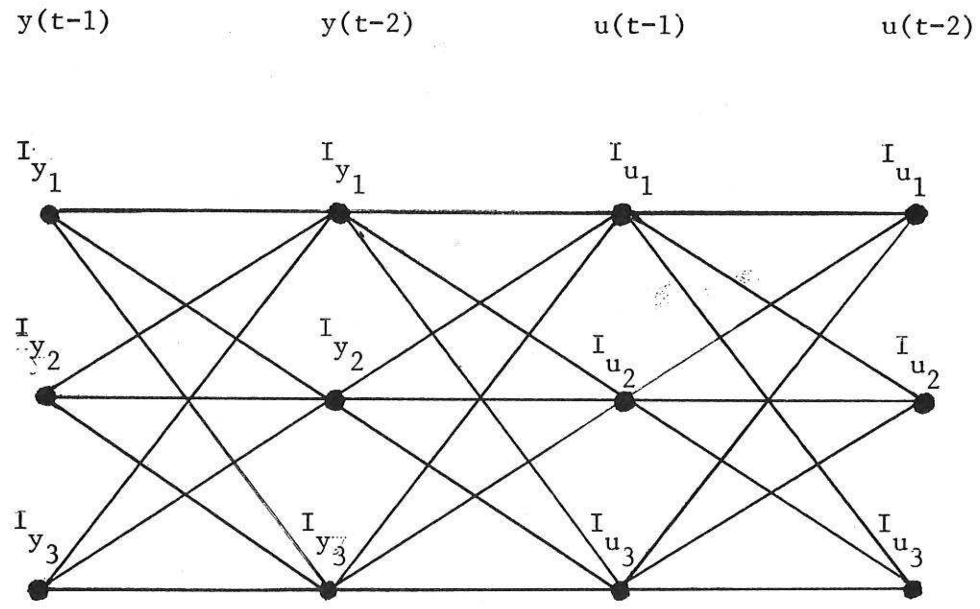

Fig.1. Schematic representation of a Spatial Piecewise Linear Model.

The intervals in Fig.1 are defined by

$$I_{y_1} = (y_{min}, y_1]$$
$$I_{y_2} = (y_1, y_2]$$
$$I_{y_3} = (y_2, y_{max}]$$
$$I_{u_1} = (u_{min}, u_1]$$
$$I_{u_2} = (u_1, u_2]$$
$$I_{u_3} = (u_2, u_{max}]$$

Each of the paths linking all the variables y(t-1), y(t-2), u(t-1), u(t-2) represents a region of operation for a spatial piecewise linear model. For example, the path linking I_{y_1}, I_{y_1}, I_{u_1}, I_{u_1} represents the operating region

$$Y_{min} < y(t-1) \leq y_1$$
$$y_{min} < y(t-2) \leq y_1$$
$$u_{min} < u(t-1) \leq u_1$$
$$u_{min} < u(t-2) \leq u_1$$

and the path linking I_{y_1}, I_{y_2}, I_{u_1}, I_{u_3} represents a model within the region

$$y_{min} < y(t-1) \leq y_1$$
$$y_1 < y(t-2) \leq y_2$$
$$u_{min} < u(t-1) \leq u_1$$
$$u_2 < u(t-2) \leq u_{max}$$

The total number of models or operating region is given by $m=(m_y^{n_y} \times m_y^{n_u})$ and there are 81 for equation (14) thus each operating region k=1,2,...,m is characterized by a locally linear model and can be expressed in matrix form as

$$y(t) = \psi^{(k)^T}(t)\,\Theta^{(k)} \qquad (15)$$

where $\psi^{(k)}(t)=[1,y(t-1),\ldots,y(t-n_y),\ u(t-d),\ldots,u(t-d-n_u+1)]^T$

$$\Theta^{(k)}=[\Theta_0^{(k)},\Theta_1^{(k)},\ldots,\Theta_{n_y}^{(k)},\ \Theta_{n_y+1}^{(k)},\ldots,\Theta_{n_y+n_u}^{(k)}]^T \qquad (16)$$

Parameter estimates for the locally linear models can be evaluated using least squares

$$\hat{\Theta}^{(k)} = [\underline{\Psi}^{(k)T}\underline{\Psi}^{(k)}]^{-1}\underline{\Psi}^{(k)T}\underline{Y}^{(k)} \tag{17}$$

where

$$\underline{\Psi}^{(k)} = [\psi^{(k)}(1),\psi^{(k)}(2),\ldots\ldots,\psi^{(k)}(N^{(k)})]^T$$

$$\underline{Y}^{(k)} = [y(1),y(2),\ldots\ldots\ldots\ldots,y(N^{(k)})]^T$$

and $N^{(k)}$ are the number of data points in region k. If noise is present in the data a parameter estimation algorithm which yields unbiased estimates such as generalised least squares, or a prediction error method, extended least squares etc. [Norton, 1986] should be employed. The bounded operating regions of the spatial piecewise linear model must be taken into consideration when selecting the input-output intervals to ensure that there are sufficient data within each region. However, it is possible that some of the regions are not reachable for a particular nonlinear system and these can obviously be ignored. The global spatial piecewise linear model is therefore obtained by estimating the locally linear model of equation (15) iteratively over the range of reachable operating regions using the least squares estimator of equation (17).

3.2 Simulation Results

Two nonlinear models were simulated to illustrate the spatial piecewise linear modelling technique.

3.2.1. The Wiener Model

A Wiener model represented as

$$y(t) = 2 \tanh\left\{\frac{0.4z^{-1}\,u(t)}{1 - 0.8z^{-1}}\right\} \tag{18}$$

was globally excited by a uniformly distributed random input of amplitude range ±1.0. Five hundred data points were used for the spatial piecewise linear estimation. The maximum and minimum range of the input and output were divided into 4 equal intervals respectively to yield a total of 16 operating regions. A least squares algorithm was used to estimate a first order locally linear model for each region and the results are given in table 1. By inspecting the coefficients $\Theta_1^{(k)}$, the locally linear models are stable in all the operating regions except in 4, 8 and 13 where the poles are outside the unit circle. The predicted output of the spatial piecewise linear model is comparatively good when compared with the Wiener model output figure 2. Both sinusoidal and PRBS inputs of amplitude range ±0.9 were also used to excite the Wiener model and the spatial piecewise linear model to compare their outputs and determine whether the spatial piecewise linear model is input sensitive [Billings and Voon 1984]. These results are illustrated in figures 3 and 4 and clearly indicate that the spatial piecewise linear model is adequate.

3.2.2 An Implicit Model

An implicit model described as

$$y(t) = 0.5y(t-1) + 0.3u(t-1) + 0.3u(t-1)y(t-1) + 0.5u^2(t-1) \qquad (19)$$

was globally excited by an uniformly distributed random input of amplitude range ±1.0 for 500 data points. The maximum and minimum range of the input and output were divided into 2 equal intervals respectively which form a total of 4 operating regions. First order linear models were estimated for each region using a least squares algorithm and the results are tabulated in table 2. Inspection of the coefficients $\Theta_1^{(k)}$ in table 2 shows that all the locally linear models are stable. Figure 5 shows that the predicted output of the spatial piecewise linear model is comparable to the output of the implicit model. In order to check the input sensitivity of the spatial piecewise linear model, a sinusoidal and PRBS input of amplitude ±0.9 were used to excite both the spatial piecewise linear model and the implicit model for comparison and the results are illustrated in figure 6 and 7. The results clearly illustrate that the spatial piecewise linear model adequately characterises the implicit model eqn (9).

4. Models With Signal Dependent Parameters

This section deals with the identification of nonlinear systems based on linear models with signal dependent parameters. It has often been argued that many practical processes can be represented by signal dependent models where the parameters depend on the input signal, output signal, an external signal, operating point etc. Several authors have investigated models of this type including Cyrot-Normand and Mien [1980] , Diekmann and Uubehauen [1985], and Haber et al [1982, 1985] and good results have been obtained.

Consider a signal dependent linear model described as

$$y(t) = \Theta_o(\omega(t)_k) + \sum_{i=1}^{n_y} \Theta_i(\omega(t)_k)\, y(t-i) + \sum_{i=1}^{n_u} \Theta_{n_y+i}(\omega(t)_k)\, u(t-d-i+1) \quad (20)$$

where $\omega(t)$ is the signal upon which the coefficients depend. In the modelling of a power station for example it may be appropriate to select $\omega(t)$ as the power output. The notation $\Theta_i(\omega(t)_k)$ in eqn (20) should be interpreted to mean that the parameters Θ_j, $j=1,2,\ldots (n_u+n_y+1)$ take on different values depending upon which of the operating points, denoted by k, the amplitude of $\omega(t)$ has been located at. The model eqn (20) can therefore be viewed as a linear system for each choice of k and this simplifies the identification and aids interpretation of the results. However, these locally linear models can then be patched together to form an approximate global nonlinear description of the system under investigation. This can be achieved by replacing the coefficients $\Theta_j(\omega(t)_k)$ $j=1,2,\ldots(n_u+n_y+1)$ in eqn (20) with polynomial functions of $\omega(t)$. Fitting a series of linear models thus gives an evaluation of these polynomials at particular values of $\omega(t)$ which can then be used to estimate the polynomial description. This procedure transforms or combines a large number of linear models into a concise nonlinear system representation. To illustrate these ideas consider eqn (20) expressed in matrix form

$$y(t) = \psi^T(t)\Theta(\omega(t)_k) \quad (21)$$

where $\psi(t) = [1, y(t-1), \ldots, y(t-n_y), u(t-d), \ldots, u(t-d-n_u+1)]^T$

$$\Theta(\omega(t)_k) = [\Theta_0(\omega(t)_k), \Theta_1(\omega(t)_k), \ldots, \Theta_{n_y}(\omega(t)_k), \Theta_{n_y+1}(\omega(t)_k), \ldots \ldots\ldots, \Theta_{n_y+n_u}(\omega(t)_k)]^T$$

Assuming that the signal dependent parameter vector $\Theta(\omega(t)_k)$ can be approximated by a finite degree polynomial function of $\omega(t)$ the individual elements of $\Theta(\omega(t)_k)$

can be expressed as

$$\Theta_i(\omega(t)_k) = \beta_i^T W(t)_k, \quad i=0,1,\ldots\ldots,n_y+n_u \tag{22}$$

where $W(t)_k = [1,\omega(t)_k,\omega^2(t)_k,\ldots\ldots,\omega^\ell(t)_k]^T$

$$\beta_i = [\beta_{i0},\beta_{i1},\beta_{i2},\ldots\ldots\ldots\ldots,\beta_{i\ell}]^T$$

Substituting equation (22) in equation (21) gives the model

$$y(t) = \psi^T(t)\,[\beta^T W(t)_k] \tag{23}$$

where

$$\beta = \begin{bmatrix} \beta_{00} & \beta_{10} & \beta_{20} & \cdots & \beta_{n_u+n_y,0} \\ \beta_{01} & \beta_{11} & \beta_{21} & \cdots & \beta_{n_u+n_y,1} \\ \beta_{02} & \beta_{12} & \beta_{22} & \cdots & \beta_{n_u+n_y,2} \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \beta_{0\ell} & \beta_{1\ell} & \beta_{2\ell} & \cdots & \beta_{n_u+n_y,\ell} \end{bmatrix} \tag{24}$$

The model of eqn (23) is thus an alternate representation of the model in eqn (20) where the signal dependent coefficients $\Theta_j(\omega(t)_k)$, $j=1,2,\ldots,(n_u+n_y+1)$ have been expanded as polynomials, evaluated at the operating points $\omega_k(t)$.

If $W(t)_k$ in eqn (23) is not restricted to specific values which define the operating points but is allowed to vary continuously then eqn (23) becomes the global nonlinear model.

$$y(t) = \psi^T(t)\,[\beta^T W(t)] \tag{25}$$

where $W(t) = [1,\omega(t),\ \omega^2(t),\ldots\ \omega^\ell(t)]^T$.

The model of eqn (25) thus provides one means of patching a series of linear signal dependent models together to yield a nonlinear system description of the process under investigation.

4.1 Parameter Estimation for Signal Dependent Parameter Models

The proposed method applies to nonlinear systems that can be linearized about each operating point in which the systems coefficients are dependent on a signal $\omega(t)$. The aim is to combine a series of locally linear models into a single overall nonlinear model eqn (25), capable of characterizing the nonlinear system behaviour. A priori information regarding the signal which the parameters in the model depend upon and the allowable linear region of plant operation about each of the operating points $\omega(t)_k$ is of considerable importance in contructing the most appropriate model of the system.

Parameter estimation for signal dependent models can be achieved in two stages. At the initial stage k is set to one and the nonlinear process is perturbed by a bounded input chosen to ensure the process behaves linearly about the operating point $\omega(t)_1$. This procedure is repeated for each of the m operating points $\omega(t)_k$, $k=1,2,\ldots m$, and in each case N^k data pairs y(k), u(k) are recorded. Linear models are then estimated using standard linear algorithms [Norton 1986] to yield estimates of the parameter vectors $\Theta(\omega(t)_1)$, $\Theta(\omega(t)_2)\ldots\ \Theta(\omega(t)_m))$ in eqn (21). Combining these estimates in the matrix $\Theta\Theta(\omega(t))$

$$\Theta\Theta(\omega(t)) = \begin{bmatrix} \Theta(\omega(t)_1) \\ \Theta(\omega(t)_2) \\ \vdots \\ \Theta(\omega(t)_m) \end{bmatrix}^T$$

$$= \underbrace{\begin{bmatrix} \Theta_o(\omega(t)_1) & \Theta_1(\omega(t)_1) & \ldots & \Theta_{n_y+n_u}(\omega(t)_1) \\ \Theta_o(\omega(t)_2) & \Theta_1(\omega(t)_2) & \ldots & \Theta_{n_y+n_u}(\omega(t)_2) \\ \vdots & \vdots & & \vdots \\ \Theta_o(\omega(t)_m) & \Theta_1(\omega(t)_m) & \ldots & \Theta_{n_y+n_u}(\omega(t)_m) \end{bmatrix}^T}_{(n_y+n_u+1)\text{ number of coefficients}} \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} \begin{matrix} m \\ \text{number of} \\ \text{experiments}\end{matrix} \qquad (26)$$

and defining

$$WW(t) = \begin{bmatrix} W(t)_1 \\ W(t)_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ W(t)_m \end{bmatrix}^T = \underbrace{\begin{bmatrix} 1 & \omega(t)_1 & \omega^2(t)_1 & \cdot\cdot & \omega^\ell(t)_1 \\ 1 & \omega(t)_2 & \omega^2(t)_2 & \cdot\cdot & \omega^\ell(t)_2 \\ \cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ 1 & \omega(t)_m & \omega^2(t)_m & & \omega^\ell(t)_m \end{bmatrix}}_{(\ell+1)\text{ terms for a } \ell\text{'th degree of nonlinearity}} \Bigg\} \begin{array}{l} m \\ \text{number of} \\ \text{experiments} \end{array} \quad (27)$$

gives the relationship

$$\Theta\Theta(\omega(t)) = \beta^T WW(t) \quad (28)$$

where β is defined by eqn (24).

Since all the elements of $\Theta\Theta(\omega(t))$ and $WW(t)$ in eqn (28) have known values once the linear estimation is complete the coefficient vector β can be estimated by a least squares algorithm

$$\hat{\beta} = (WW(t)WW^T(t))^{-1}\ WW(t)\ \Theta\Theta^T(\omega(t)) \quad (29)$$

ssuming that the matrix $[WW(t)\ WW^T(t)]$ is well conditioned. The number of linear models m obtained from m number of experiments must be greater than or equal to the highest degree of nonlinearity plus 1, (i.e. $m \geq (\ell+1)$) in order for equation (29) to have a unique solution for β.

4.2 Simulation Results

A Hammerstein, implicit and general nonlinear models were simulated to illustrate signal dependent parameter modelling as described above.

4.2.1. A Hammerstein Model

A Hammerstein model described as

$$y(t) = 0.8\ y(t-1) + 0.4\ \tanh\ (2u(t-1)) \quad (30)$$

was simulated over the global input range of ±1.0. Eleven bounded first order linearized models of input range ±0.1 about the mean of the input at various operating points were estimated using a least squares algorithm and the results are tabulated in table 3. The signal dependent coefficient vector $\Theta\Theta(\omega(t))$ was assumed to be a function of the input (i.e. $\omega(t)=u(t)$). The coefficient vector β was estimated using eqn (28) for a third degree nonlinearity to give the following form for eqn (27).

$$\Theta\Theta(u(t)) = \begin{bmatrix} -0.0005 & 0.2078 & 0.0094 & 0.1715 \\ 0.7990 & 0.0092 & -0.0010 & -0.0013 \\ 0.6738 & 0.0027 & -0.8411 & -0.0043 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \end{bmatrix}$$

The global nonlinear model eqn (25) therefore becomes

$$y(t)=[1,y(t-1).u(t-1)] \left[\begin{bmatrix} -0.0005 & 0.2078 & 0.0094 & 0.1715 \\ 0.7990 & 0.0092 & -0.0010 & -0.0013 \\ 0.6738 & 0.0027 & -0.8411 & -0.0043 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \end{bmatrix} \right] \quad (31)$$

The linear models specified by the 22 parameters in Table 3 have therefore been combined or patched together to yield the nonlinear model of eqn (31) containing just 12 parameters.

A sinusoidal and PRBS input signal of ±0.9 amplitude range at zero operating point were used to excite both the signal dependent parameter model and the Hammerstein model to compare their outputs and determine whether the model of eqn (31) is input sensitive [Billings and Voon 1984]. These results are illustrated in figures 8 and 9 which clearly indicate that for this example the signal dependent parameter model is adequate.

4.2.2 <u>An Implicit Model</u>

An implicit model described by

$$y(t) = 0.5y(t-1) + 0.3u(t-1) + 0.3y(t-1)u(t-1) + 0.5u^2(t-1) \quad (32)$$

was simulated over the global input range of ±1.0. Eleven bounded first order linearized models of input range ±0.1 about the mean of the input at various operating points were identified using a least squares estimator and the estimates are tabulated in table 4. The signal dependent coefficient vector $\Theta\Theta(\omega(t))$ was assumed to be a function of the input (ie $\omega(t)=u(t)$). The coefficient vector β

was estimated using eqn (29) for a fourth degree nonlinearity and to give the following form for eqn (28)

$$\Theta\Theta(u(t)) = \begin{bmatrix} -0.0019 & 0.0525 & -0.6104 & -0.6156 & -0.3919 \\ 0.4548 & 0.4334 & 0.0580 & -0.1933 & -0.0036 \\ 0.3034 & 1.1151 & 0.3770 & 0.3723 & 0.2398 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \\ u^4(t-1) \end{bmatrix}$$

The global nonlinear model eqn (25) therefore becomes

$$y(t) = [1, y(t-1), u(t-1)]$$

$$x \left[\begin{bmatrix} -0.0019 & 0.0525 & -0.6104 & -0.6156 & -0.3919 \\ 0.4548 & 0.4334 & 0.0580 & -0.1933 & -0.0036 \\ 0.3034 & 1.1151 & 0.3770 & 0.3723 & 0.2398 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \\ u^4(t-1) \end{bmatrix} \right] \quad (33)$$

In order to check whether the signal dependent parameter model is adequate to represent the nonlinear system eqn (32), a sinusoidal and PRBS input signal of ±0.9 amplitude range at zero operating point were used to excite both systems. The outputs are compared in figures 10 and 11 which clearly illustrate that the signal dependent parameter model is adequate.

4.2.3. <u>A General Model</u>

The general model illustrated in Fig.12 was simulated over the global input range of ±1.0.

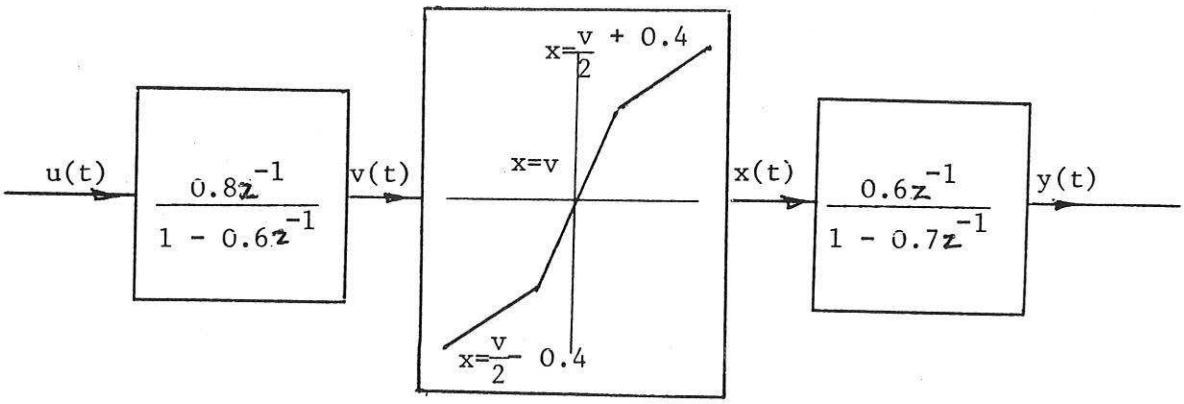

Fig.12 Simulated sample 4.2.3.

Fifteen second order linear models valid over the input range 0.1 about the mean of the input at various operating points were estimated using a least squares algorithm to yield the coefficients in table 5. The signal dependent parameter vector $\Theta\Theta(\omega(t))$ was assumed to be a function of the input, (ie. $\omega(t)=u(t)$). The coefficient vector β was estimated using eqn (29) for a third degree nonlinearity to give the following form for eqn (28).

$$\Theta\Theta(u(t)) = \begin{bmatrix} -0.0019 & 0.155 & -0.0127 & 0.0578 \\ 1.287 & 0.0028 & -0.121 & 0.0145 \\ -0.409 & -0.0032 & -0.079 & 0.011 \\ -0.0002 & 0.0013 & 0.0010 & -0.0024 \\ 0.439 & 0.0004 & -0.319 & -0.0015 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \end{bmatrix}$$

Hence the global nonlinear model can be expressed as

$$y(t) = [1, y(t-1), y(t-2), u(t-1), u(t-2)]$$

$$\times \left[\begin{bmatrix} -0.0019 & 0.155 & -0.0127 & 0.0578 \\ 1.287 & 0.0028 & -0.121 & 0.0145 \\ -0.409 & -0.0032 & 0.079 & -0.011 \\ -0.0002 & 0.0013 & 0.0010 & -0.0024 \\ 0.439 & 0.0004 & -0.319 & -0.0015 \end{bmatrix} \begin{bmatrix} 1 \\ u(t-1) \\ u^2(t-1) \\ u^3(t-1) \end{bmatrix} \right] \quad (34)$$

A sinusoidal and a PRBS input with amplutides of ± 0.9 about a zero operating point were used to excite the signal dependent parameter model and the general model. The outputs are compared in figures 13 and 14. Figure 13 indicates that the signal dependent parameter model is comparable with the general model for a sinusoidal input. However, for the PRBS input, figure 14 shows that the response of the signal dependent parameter model is poor compared with the output of the general model yielding a residual error of approximately 40 per cent of the general model output.

5. Conclusions

Two methods of characterizing nonlinear systems by fitting a series of locally linear models have been described. The spatial piecewise linear model will, providing the nonlinearities are smooth, provide an adequate representation of a nonlinear system but this may only be achieved at the expense of fitting a very large number of linear models each valid in a small region of operation.

Alternatively, if the nonlinearity is assumed to be produced by a measurable system variable a series of signal dependent linear models can be estimated. These models can then be patched together to yield a nonlinear description of the process. It is however important to emphasise that the linear models, or the equivalent nonlinear model obtained by combining them, may only be valid for relatively slow moving inputs. This occurs because the models represent just one possible trajectory over the domain of system operation. The models will therefore produce an excellent prediction of the process output when perturbed by the input used for the identification or by an input which causes the system to traverse slowly from one operating region to another. If however the nonlinearity is not perfectly dependent upon one signal or the models are perturbed by inputs which cause rapid transient changes in the operating point this will produce an output that may be considerably different to the output of the process. This means that the identified model does not provide an adequate representation of the nonlinear system but is only valid for a small class of input signals. The model is input sensitive [Billings and Voon, 1984]. This effect, which is often overlooked, can be severe as indicated by the 40 % discrepancy in the predicted output of the model in example 4.2.3.

The advantages of approximation by linear models can therefore be outweighed by either an excessive linear model set in the case of spatial piecewise linear modelling or by a model which is input sensitive when using a signal dependent representation. The difficulty of choosing the operating points and range of input variation to ensure linear operation can also introduce a severe problem. Methods of detecting nonlinearity in data [Billings and Voon 1983, Haber 1985] can however be employed to aid the investigator in this choice.

The disadvantages described above suggest that it may in some situations be worthwhile fitting a nonlinear model to the process rather than attempting to approximate it by a series of linear models. Recent developments in the field of nonlinear systems identification have shown that parsimonious nonlinear models can be fitted using relatively simple extensions of linear algorithms to yield a representation which provides excellent predictions of the system behaviour over the global operating region [Billings and Voon 1984, 1986]. If a linear description of the system is then required the nonlinear model can be linearized about any chosen operating point. This could be used to produce a series of locally linear models which can be analysed and interpreted by well known linear methods with the added advantage that an optimal choice of operating points can be investigated to minimise the number of linear models and the errors introduced by using the linear description can be readily evaluated. In any event the identification of locally linear models should not be regarded as a panacea for nonlinear systems identification.

6. References

Billings S. A., Voon W. S. F. (1983): Structure detection and model validity tests in the identification of nonlinear systems; Proc IEE, Part D, 130, 193-199.

Billings S.A., Voon W. S. F. (1984): Least squares parameter estimation algorithms for nonlinear systems: Int. J. Systems Science, 15, 601-615.

Billings S. A., Voon W. S. F. (1986): A prediction error and stepwise regression estimation algorithm for nonlinear systems: Int. J. Control 44, 803-822.

Cyrot-Normand D, Mien H. D. V. (1980): Nonlinear state-affine identification methods: Applications to electrical power plants; IFAC Symp on Auto. Control in Power Generation; Distribution and Protection: 449-462.

De Hoff R.L. Rock S.M. (1979): Developments of simplified nonlinear models from multiple linearizations! 17th IEEE Conf. on Decision and Control, 316-318.

Diekmann K., Unbehauen H. (1985): On-line parameter estimation in a class of nonlinear systems via modified least squares and instrumental variable algorithms; 7th IFAC Symp. Ident. and Syst. Par. Est., 149-153.

Haber R. (1985): Nonlinearity tests for dynamic processes; 7th IFAC Symp. Ident. and Syst. Par. Est, 409-414.

Haber R, Keviczky L. (1985): Identification of linear systems having signal dependent parameters: Int. J. Systems Sci, 16, 869-884.

Haber R., Vajk I. (1982): Nonlinear system identification by linear systems having signal dependent parameters: 6th IFAC Symp. Ident. & Syste. Par. Est., 421-426.

Leontaritis I. J., Billings S. A., (1985): Input-Output parametric models for nonlinear systems. Part I Deterministic nonlinear systems. Part II Stochastic nonlinear systems. Int. J. Control, 41, 303-344.

Norton J. P. (1986): Introduction to System Identification: Academic Press.

Priestley M. B. (1980); State-dependent models: A general approach to nonlinear time series analysis; Journal of Time Series Analysis; 1, 47-71.

Tony H. (1978a); On a threshold model; "Pattern Recognition and Signal Processing" (Ed C. H. Chen), Sitjhoff and Noordhoff, Amsterdam.

Tony H. (1978b): A view on nonlinear time series model building; Proc Int. Time Series, Nottingham, U.K.

White J.I. (1971): The representation of nonlinear systems by exhaustive linearization; Proc. IEEE, 98-99.

Table 1. Spatial Piecewise linear estimation for example 3.2.1.

K	Operating Region	$\theta_0^{(k)}$	$\theta_1^{(k)}$ y(t-1)	$\theta_2^{(k)}$ u(t-1)
1	-1.00<u(t-1)≤-0.495 -1.53<y(t-1)≤-0.710	-0.2338	0.6585	0.4852
2	-1.00<u(t-1)≤-0.495 -0.71<y(t-1)≤ 0.110	-0.0744	0.6918	0.6739
3	-1.00<u(t-1)≤-0.495 0.11<y(t-1)≤ 0.930	-0.0131	0.8475	0.7931
4	-1.00<u(t-1)≤-0.495 0.93<u(t-1)≤ 1.750	-0.4313	1.2394	0.7214
5	-0.495<u(t-1)≤0.004 -1.53<y(t-1)≤-0.710	-0.0031	0.8325	0.6037
6	-0.495<u(t-1)≤0.004 -0.710<y(t-1)≤0.110	-0.0095	0.7821	0.7602
7	-0.495<u(t-1)≤0.004 -0.110<y(t-1)≤0.930	-0.01064	0.8446	0.7964
8	-0.495<u(t-1)≤0.004 0.930<y(t-1)≤1.750	-0.2935	1.0995	0.6275
9	0.004<u(t-1)≤0.502 -1.53<y(t-1)≤-0.710	0.1349	0.9650	0.6861
10	0.004<u(t-1)≤0.502 -0.710<y(t-1)≤0.110	0.0038	0.8190	0.7901
11	0.004<u(t-1)≤0.502 0.110<y(t-1)≤0.930	0.0203	0.7820	0.7181
12	0.004<u(t-1)≤0.502 0.930<y(t-1)≤1.75	-0.0537	0.8971	0.5414
13	0.502<u(t-10≤1.00 -1.53<y(t-1)≤-0.710	0.2576	1.1246	0.7711
14	0.502<u(t-1)≤1.00 -0.710<y(t-1)≤0.110	0.0084	0.7896	0.7633
15	0.502<u(t-1)≤1.00 0.110<y(t-1)≤0.930	0.1244	0.6697	0.6199
16	0.502<u(t-1)≤1.00 0.930<y(t-1)≤1.750	0.3648	0.6515	0.3278

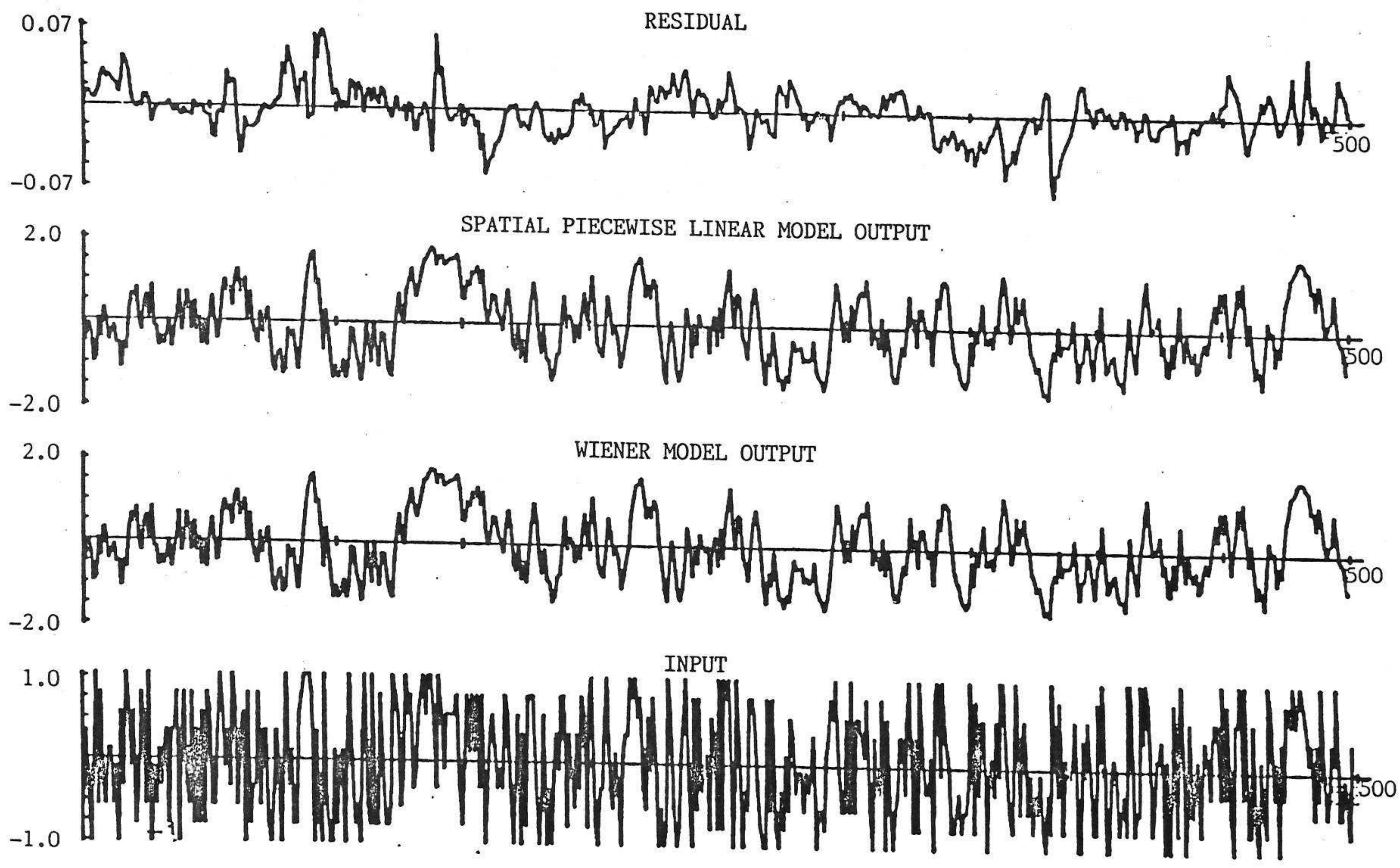

Fig.2. Simulation results for example 3.2.1.

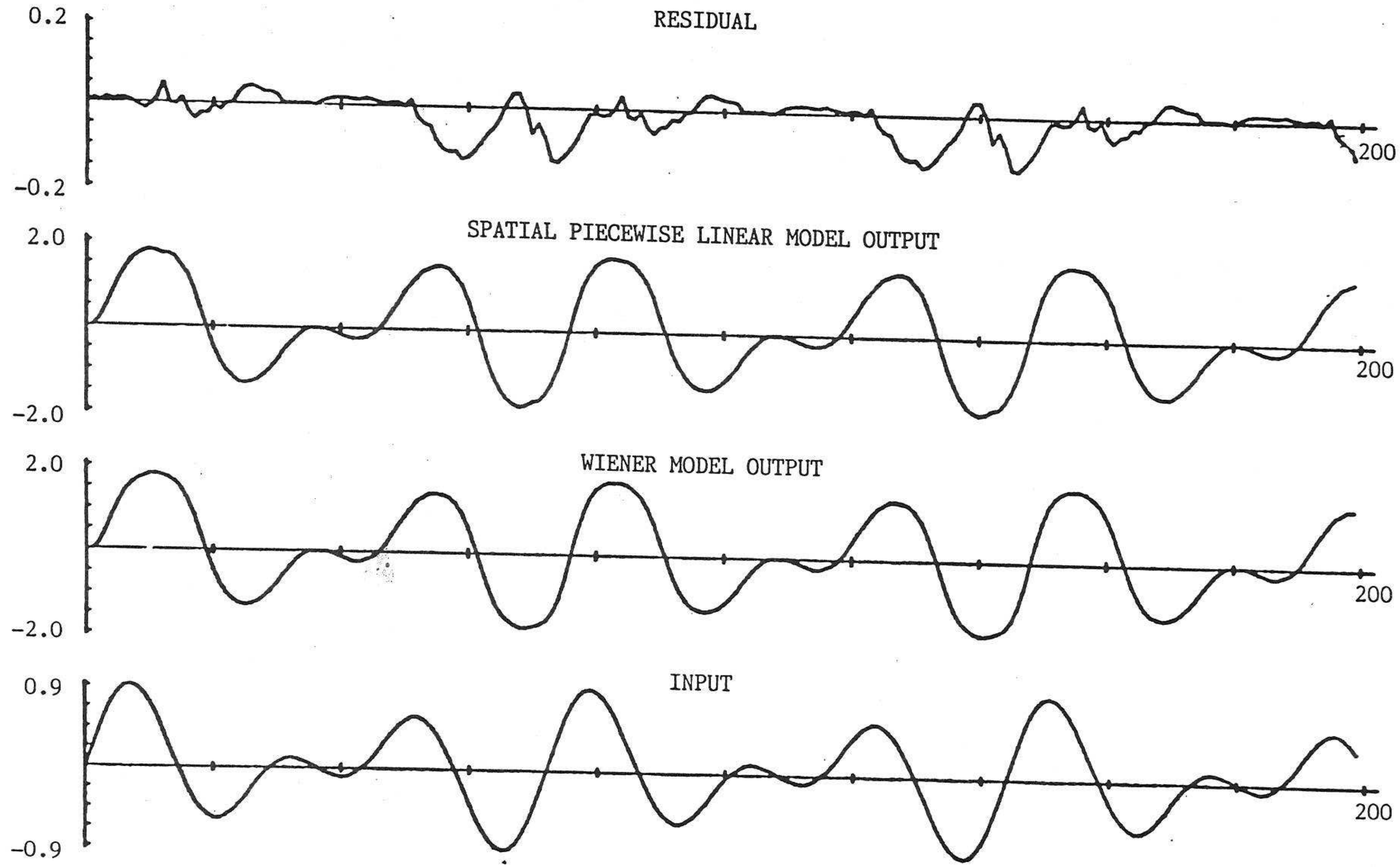

Fig.3. Simulation of example 3.2.1. for a sinusoidal excitation.

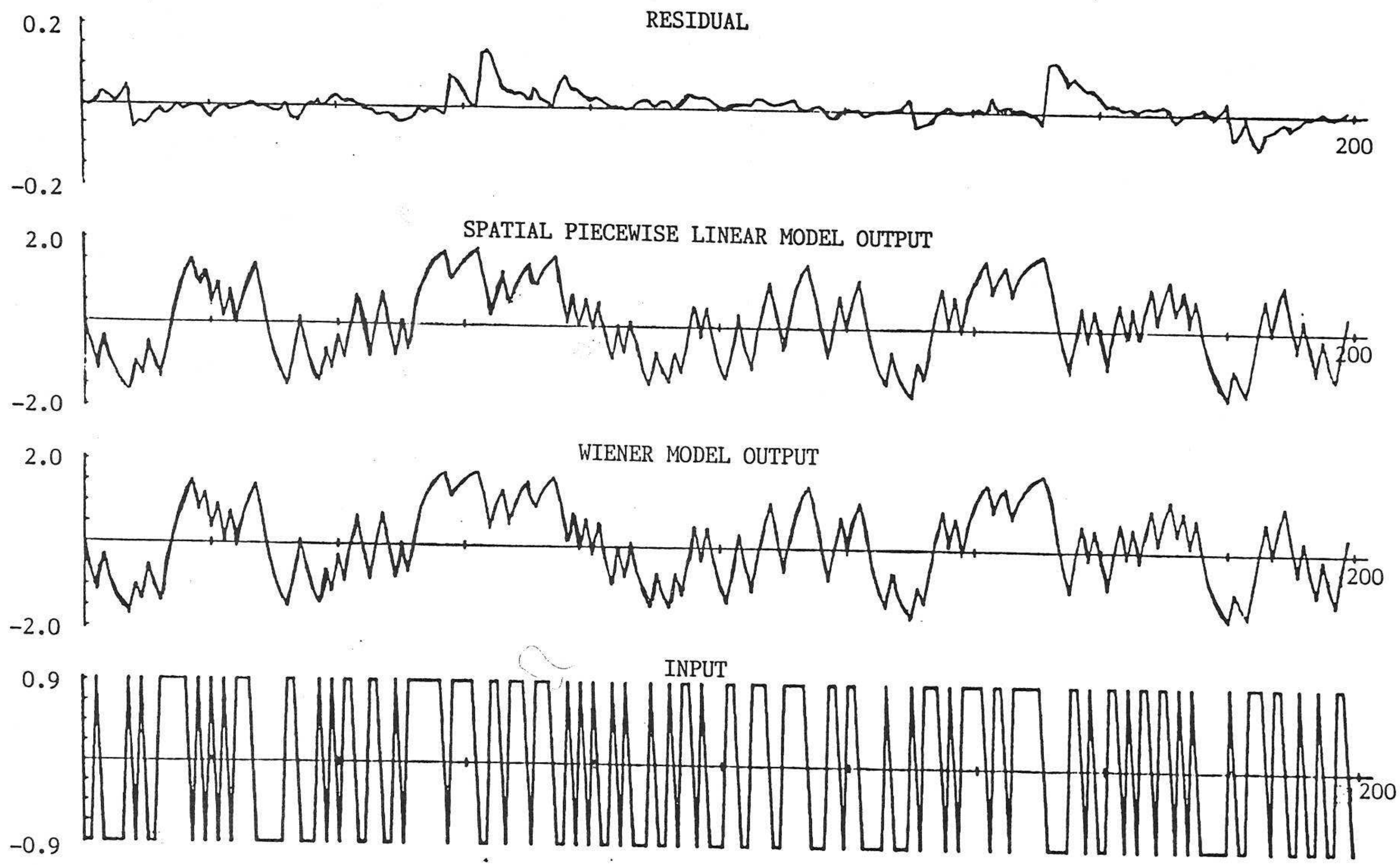

Fig.4. Simulation of example 3.2.1. for a prbs input.

K	Operating region	$\theta_0^{(k)}$	$\theta_1^{(k)}$ y(t-1)	$\theta_2^{(k)}$ u(t-1)
1	-1.00<u(t-1) ≤0.004 -0.059<y(t-1)≤1.157	-0.0568	0.3162	-0.1678
2	-1.00 <u(t-1)≤0.004 1.157<y(t-1)≤2.373	0.1664	0.3161	0.2011
3	0.004<u(t-1)≤1.000 -0.059<y(t-1)≤1.157	-0.1623	0.6682	0.9582
4	0.004<u(t-1)≤1 000 1.157<y(t-1)≤2.373	-0.484	0.6942	1.423

Table 2. Spatial piecewise linear estimation for example 3.2.2.

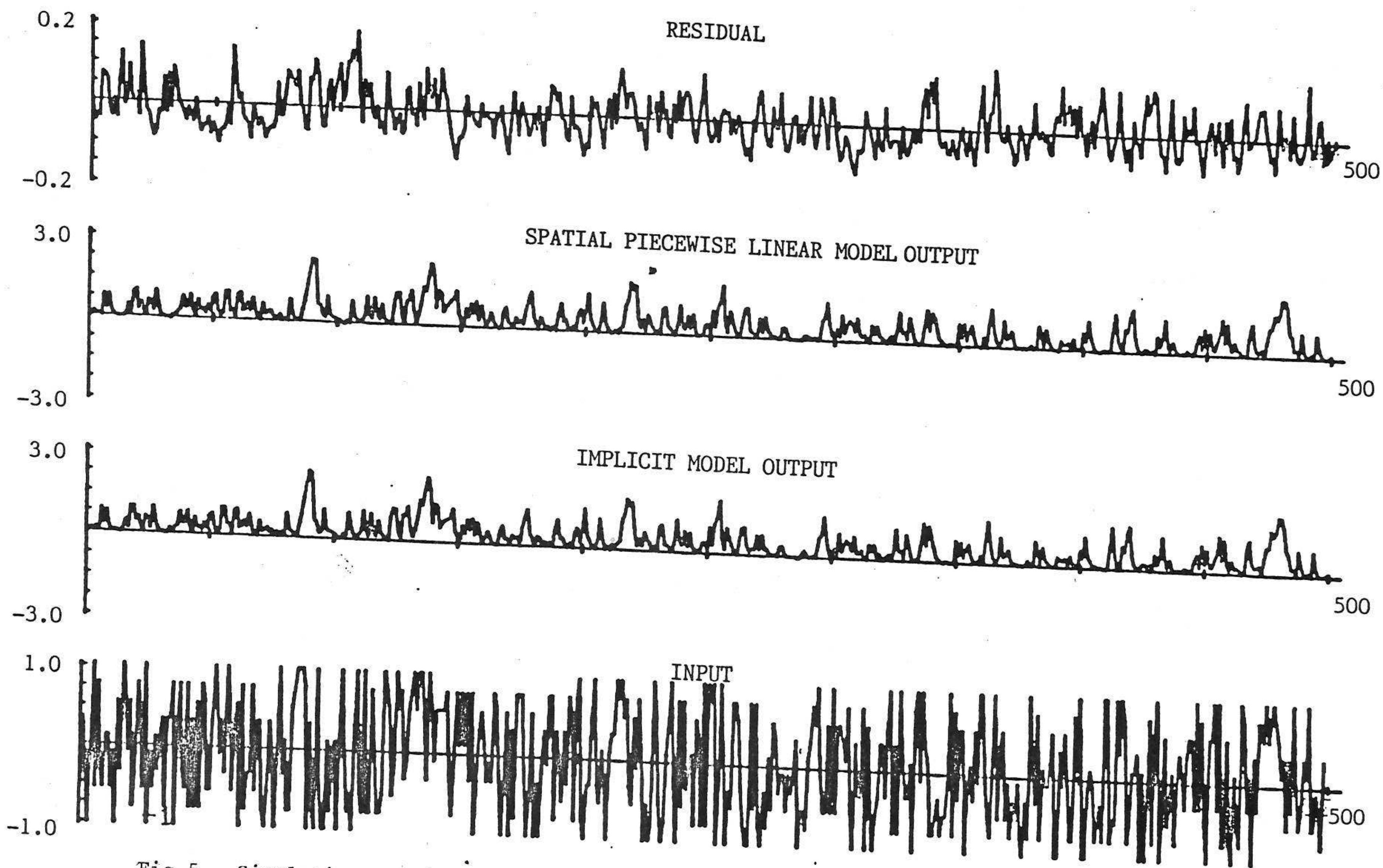

Fig.5, Simulation results for example 3.2.2.

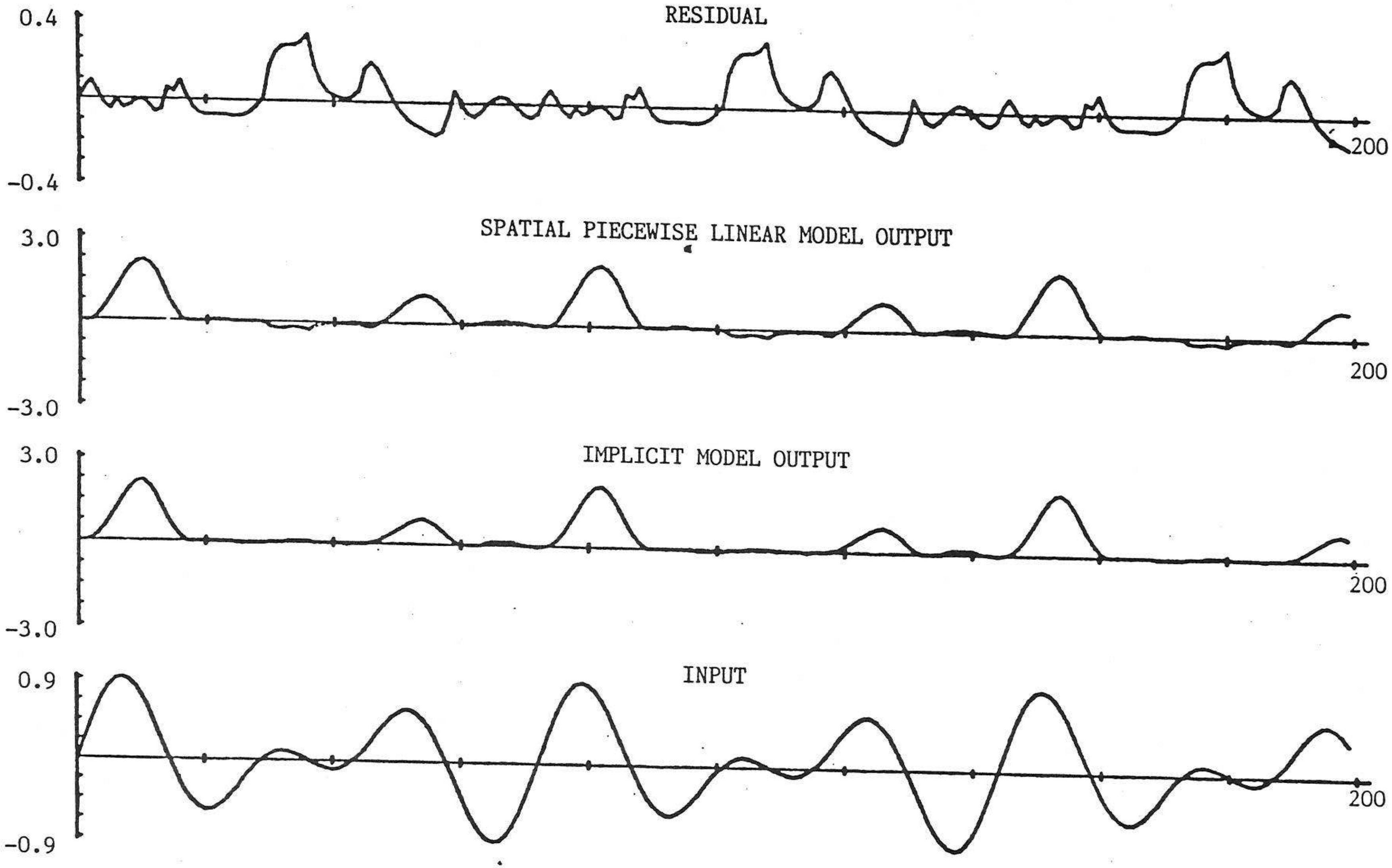

Fig.6. Simulation of example 3.2.2. for a sinusoidal excitation.

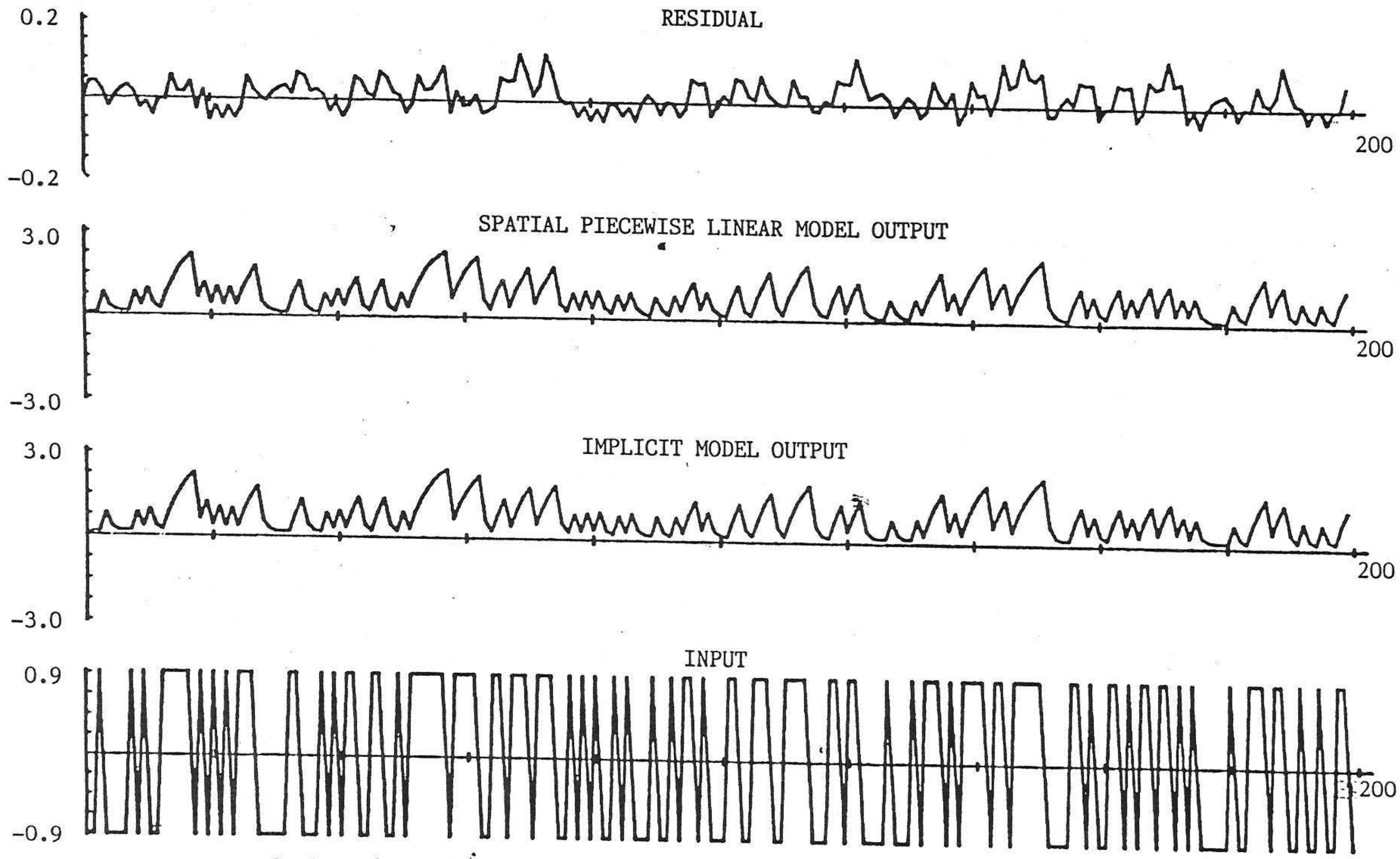

Fig.7. Simulation of example 3.2.2. for a prbs excitation.

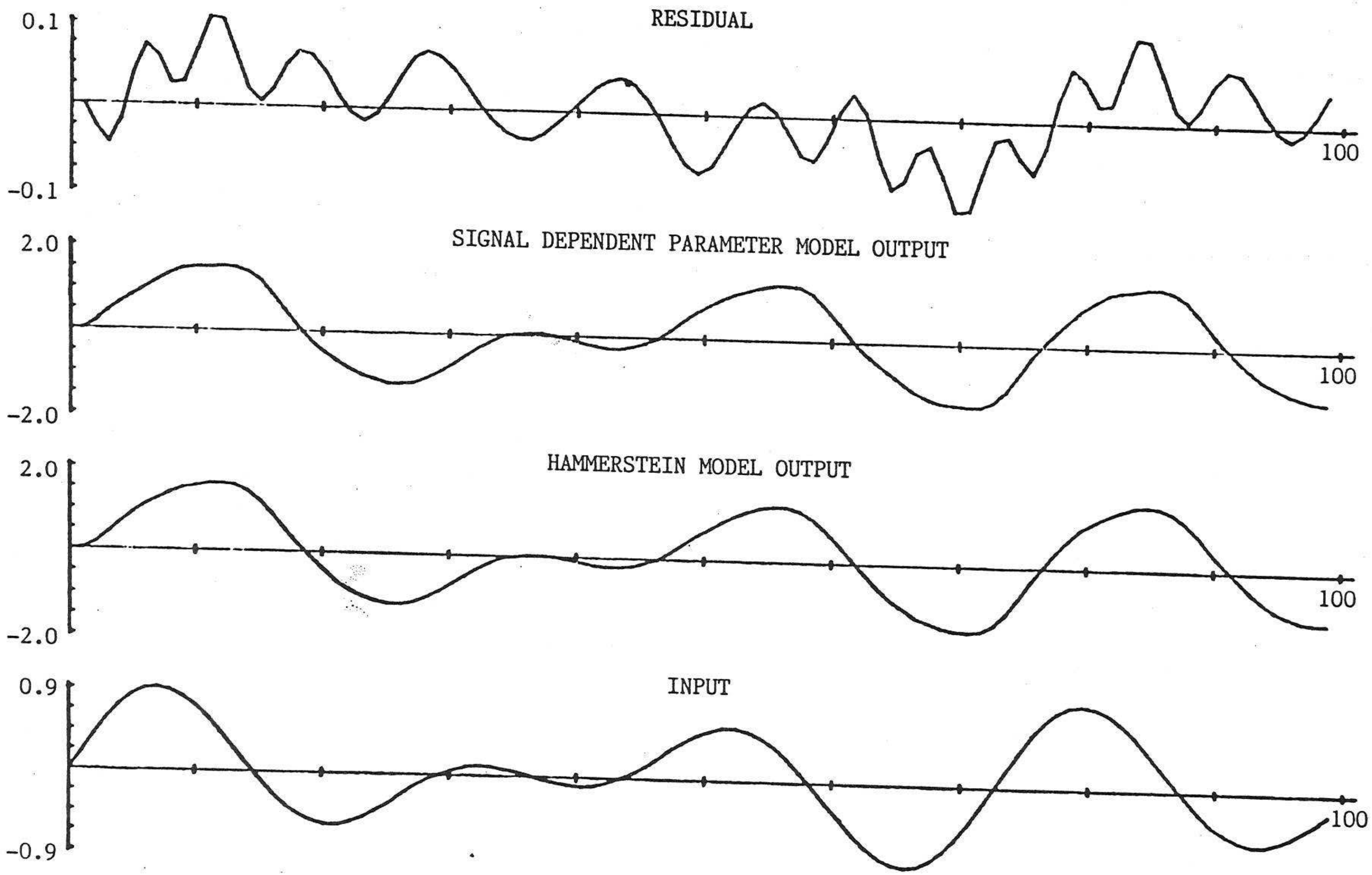

Fig.8. Simulation of example 4.2.1. for a sinusoidal excitation.

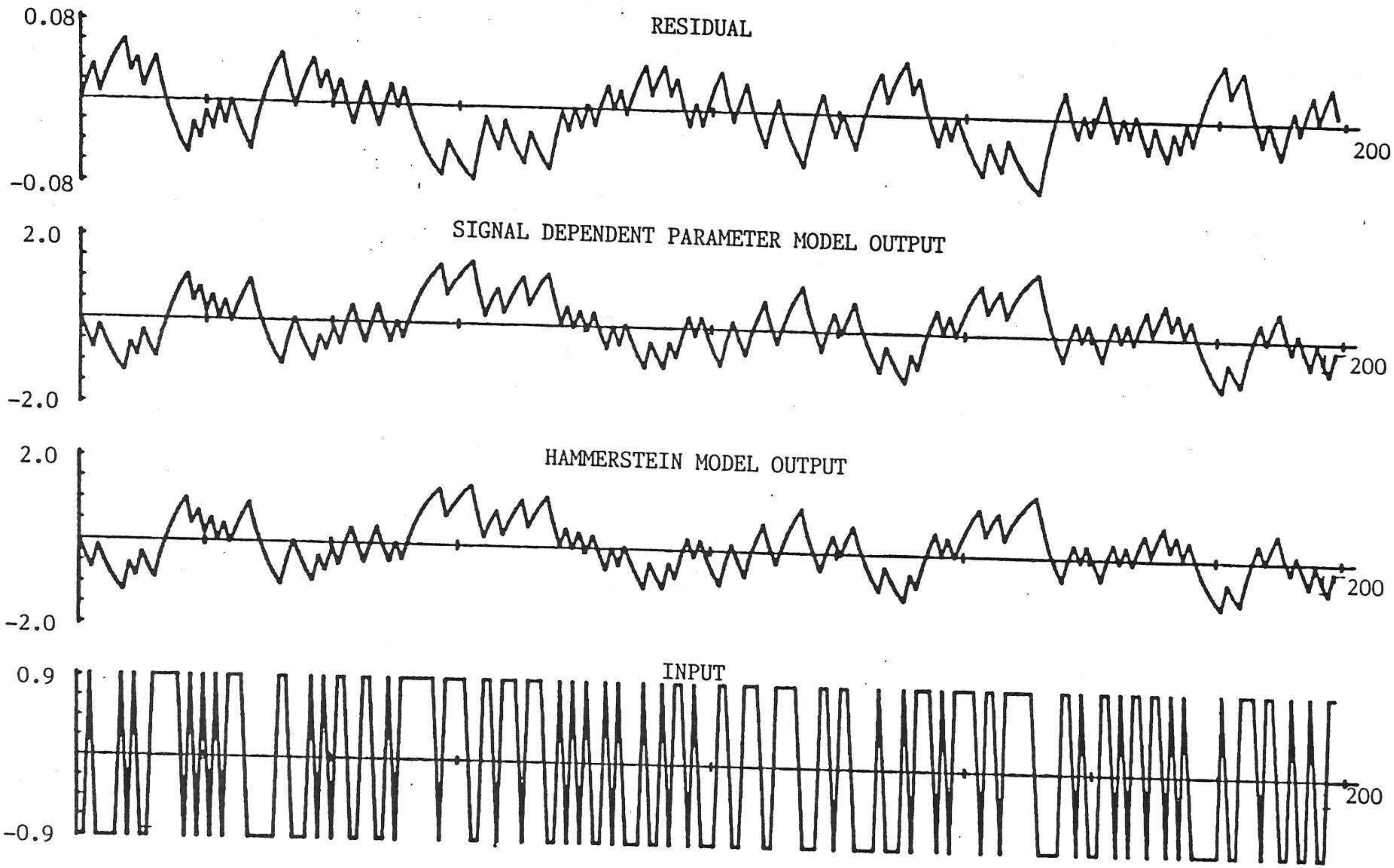

Fig.9. Simulation of example 4.2.1. for a prbs excitation.

Experiment number k	input mean	θ_0 $(\omega(t))_k$	θ_1 $(\omega(t))_k y(t-1)$	θ_2 $(\omega(t))_k u(t-1)$
1	-0.9	-0.2995	0.7999	0.0865
2	-0.7	-0.2292	0.7994	0.1755
3	-0.5	-0.1338	0.7985	0.3358
4	-0.3	-0.0400	0.7987	0.5670
5	-0.1	-0.0005	0.7990	0.7584
6	0.0	0.0000	0.7991	0.7879
7	0.1	0.0004	0.7992	0.7584
8	0.3	0.0395	0.7992	0.5671
9	0.5	0.1311	0.7993	0.3388
10	0.7	0.2286	0.7996	0.1761
11	0.9	0.3031	0.7997	0.0851

Table 3. Linear model estimation for example 4.2.1.

Experiment number k	input mean	θ_0 $(\omega(t))_k$	θ_1 $(\omega(t))_k y(t-1)$	θ_2 $(\omega(t))_k u(t-1)$
1	-0.9	-0.3493	0.2278	-0.5440
2	-0.7	-0.2274	0.2864	-0.3822
3	-0.5	-0.1249	0.3347	-0.2089
4	-0.3	-0.0621	0.1760	-0.0209
5	-0.1	-0.0019	0.4505	0.1885
6	0.0	-0.0051	0.4919	0.3025
7	0.1	-0.0020	0.5253	0.4247
8	0.3	-0.0689	0.5864	0.7008
9	0.5	-0.2344	0.6451	1.0370
10	0.7	-0.5576	0.7028	1.4704
11	0.9	-1.1592	0.7599	2.0770

Table 4. Linear model estimation for example 4.2.2.

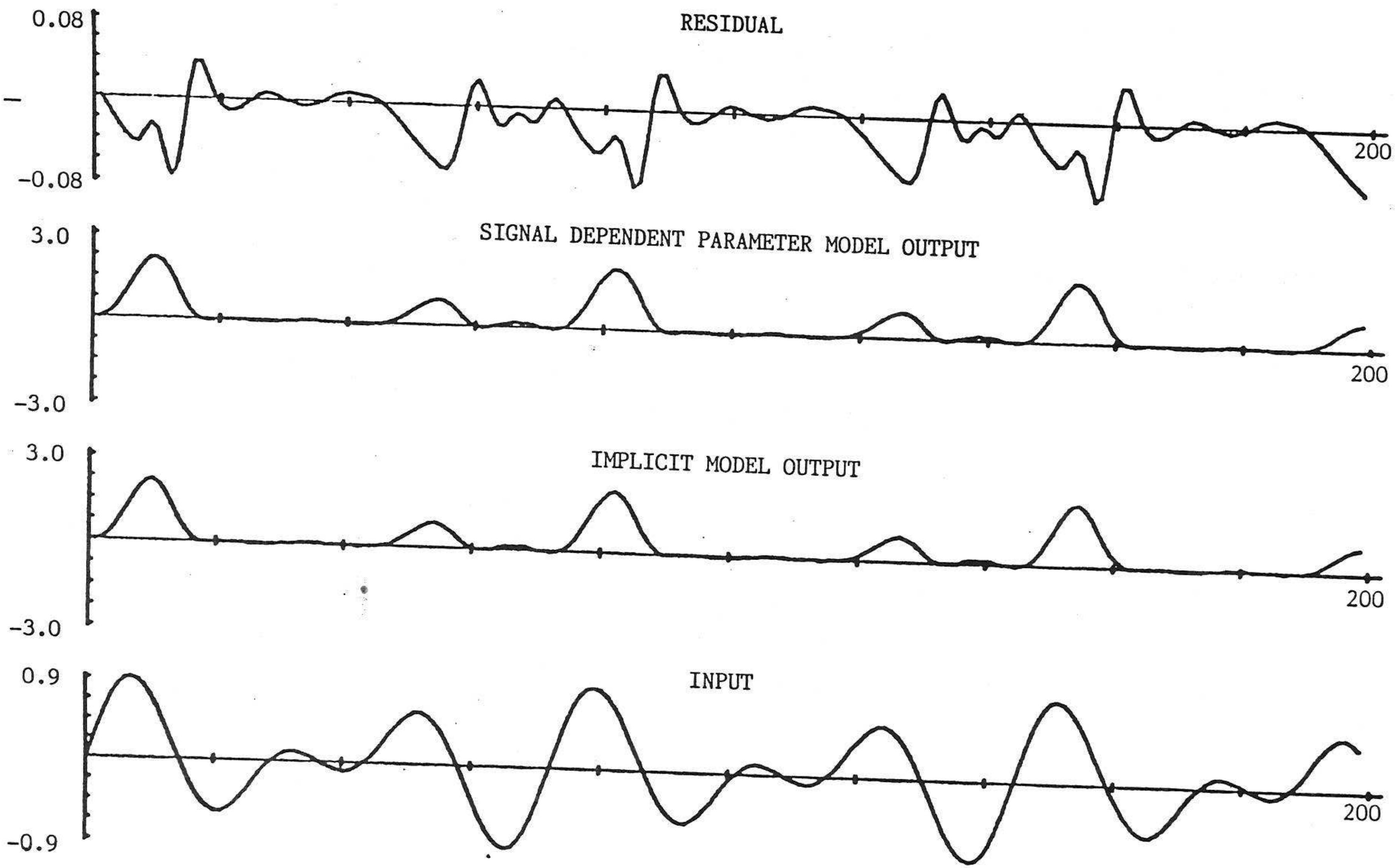

Fig.10 Simulation of example 4.2.2. for a sinusoidal excitation.

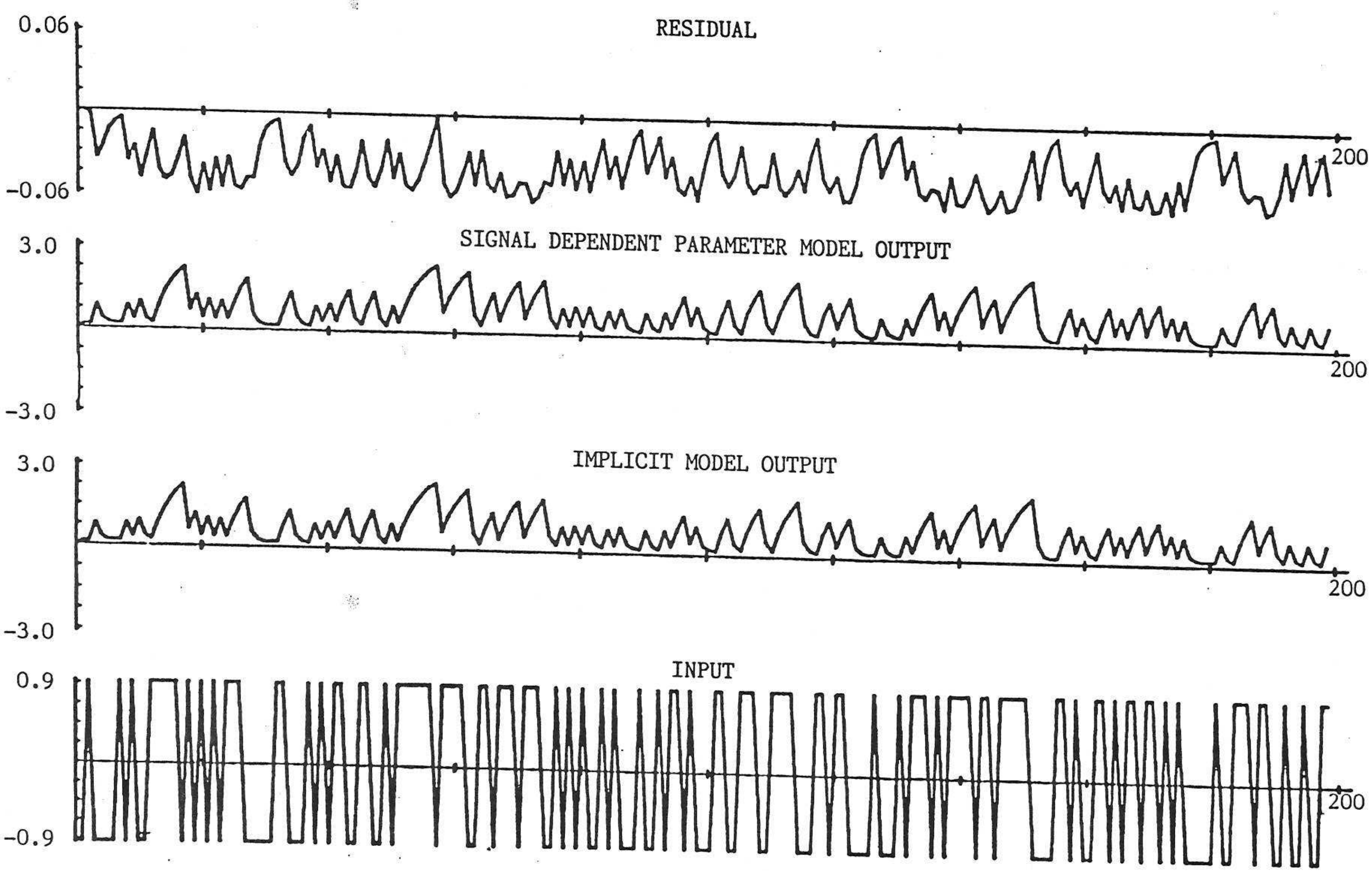

Fig.11 Simulation of example 4.2.2. for a prbs excitation.

Experiment number k	input mean	θ_0 $(\omega(t))_k$	θ_1 $(\omega(t))_k y(t-1)$	θ_2 $(\omega(t))_k y(t-2)$	θ_3 $(\omega(t))_k u(t-1)$	θ_4 $(\omega(t))_k u(t-2)$
1	-0.9	-0.1786	1.18317	-0.3353	0.0012	0.2398
2	-0.8	-0.1665	1.20000	-0.3497	0.0007	0.2398
3	-0.7	-0.1549	1.2166	-0.3634	0.0001	0.2399
4	-0.5	-0.1341	1.2445	-0.3857	0.0008	0.2399
5	-0.4	-0.0576	1.2600	-0.3888	-0.0019	0.3598
6	-0.3	-0.0016	1.2886	-0.4096	-0.0001	0.4789
7	-0.1	-0.0004	1.2879	-0.4086	-0.0001	0.4789
8	0.0	0.0000	1.2878	-0.4083	-0.0001	0.4789
9	0.1	0.0004	1.2877	-0.4083	-0.0001	0.4789
10	0.3	0.0015	1.2879	-0.4088	-0.0002	0.4789
11	0.4	0.0478	1.2629	-0.3862	0.0013	0.3597
12	0.5	0.1111	1.2693	-0.3978	-0.0006	0.2408
13	0.7	0.1364	1.2408	-0.3790	0.0007	0.2389
14	0.8	0.1475	1.2266	-0.3676	0.0004	0.2386
15	0.9	0.1585	1.2120	-0.3553	-0.0002	0.2385

Table 5. Linear model estimation for example 4.2.3.

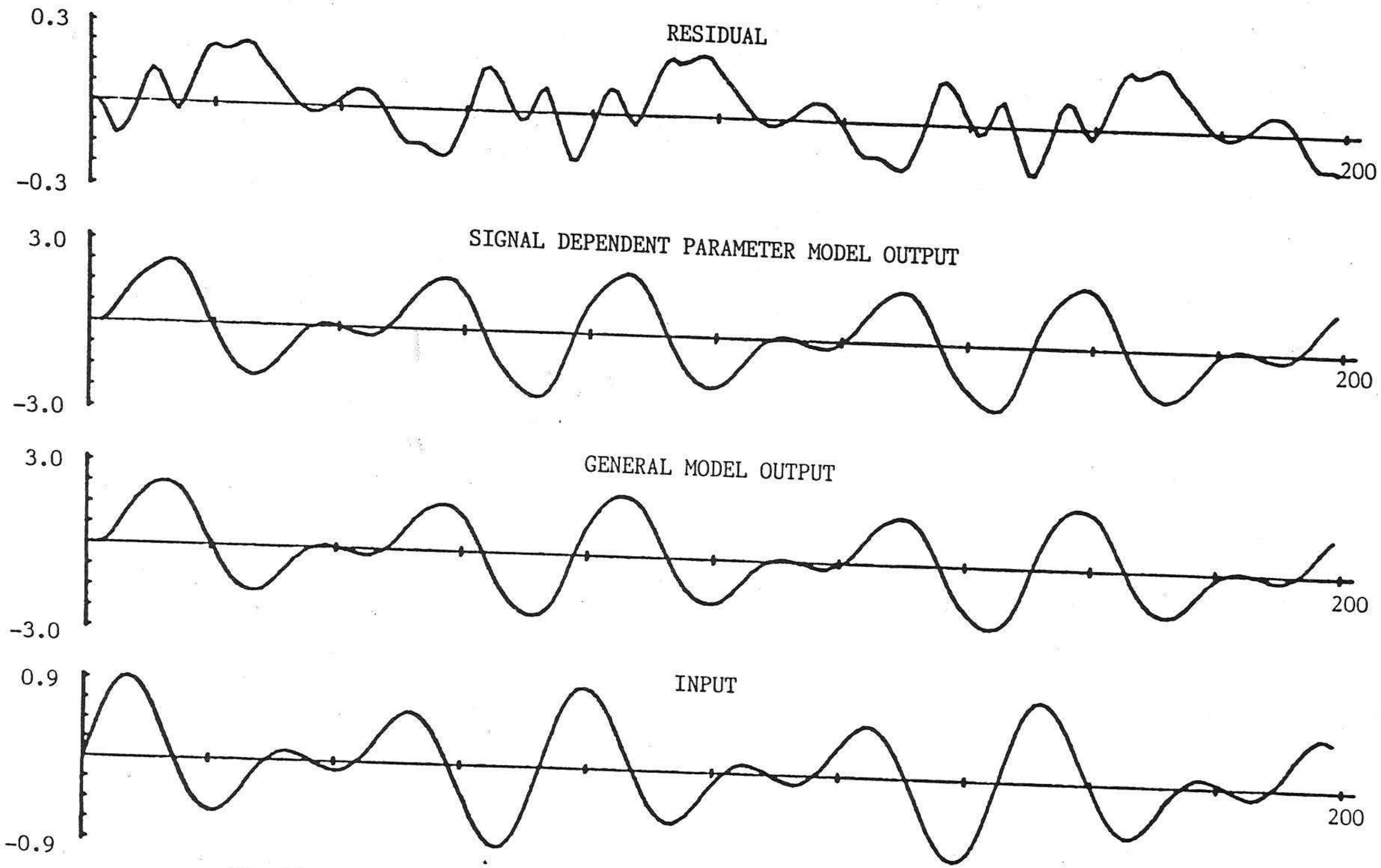

Fig.13 Simulation of example 4.2.3. for a sinusoidal excitation.

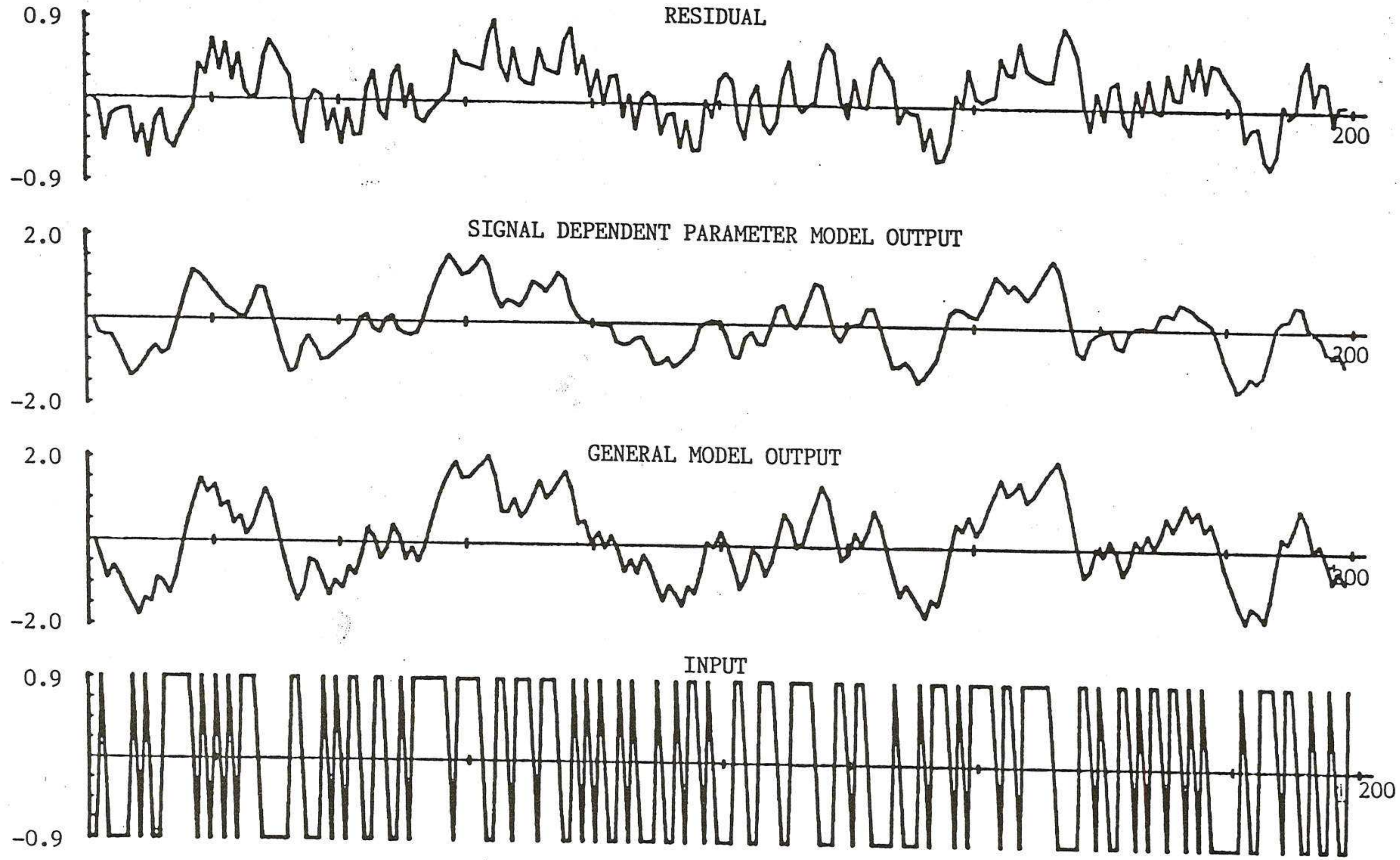

Fig.14 Simulation of example 4.2.3. for a prbs excitation.