6

Applications

This chapter starts with a brief survey of the reported applications of Wiener and Hammerstein models in both system modelling and control. Next, the estimation of parameter changes in the context of fault detection and isolation is considered in Section 6.2. Modelling vapor pressure dynamics in a five stage sugar evaporation station is studied in Section 6.3. Two nominal models of the process, i.e., a linear model and a neural network Wiener model are developed based on the real process data recorded at the Lublin Sugar Factory in Poland. Finally, Section 6.4 summarizes the results.

6.1 General review of applications

A few authors [39, 42, 95, 96, 128, 130, 133, 134, 154] have studied the control of the pH neutralization process with control methods that use Wiener and Hammerstein models.

Based on the Hammerstein model, Fruzzetti *et al.* [39] proposed a model predictive control (MPC) strategy for a pH neutralization process and a binary distillation column. MPC is a control strategy in which a model of the process is used to predict future process outputs [62, 127]. Using the Hammerstein model and an inverse model of the nonlinear element, it is possible to apply a linear MPC controller in the analyzed control scheme. The controlled system is identified using the Hammerstein model, defined as a special form or the NARMAX model, and applying a forward regression orthogonal estimator [18].

A model of the pH neutralization process was employed by Gerkšič *et al.* [42] for testing the MPC scheme. The actual and future control values are calculated minimizing a performance criterion over a certain future horizon. The method of Gerkšič *et al.* uses a two-step procedure for Wiener system identification. In the first step, with a testing input signal slowly varied from the lower to the upper edge of the working range, a polynomial model of the

A. Janczak: Identification of Nonlinear Systems, LNCIS 310, pp. 159–185, 2005.

nonlinear element is determined. Another testing signal composed of a sequence of random steps and a small noise sequence is used in the other step to identify the linear system. The performance of three MPC schemes with different disturbance rejection techniques was compared. Also, comparisons with other control schemes were made showing that MPC is superior to the proportional plus integral (PI) scheme.

The modelling of the pH neutralization process with a polynomial model of the inverse nonlinear element and a frequency response model of the linear element Wiener was studied by Kalafatis *et al.* [95, 96]. In a pilot scale pH plant [95], a continuous stirred tank reactor (CSTR), a strong base (NaOH) reacts with feed streams of a strong acid (HCL) and a buffering solution ($NaHCO_3$). The flow rate of the strong base is the process input and the pH of the exiting stream is the process output.

An experimental pH neutralization process, in which a 0.82 mM HCL solution is added to a 2.0 mM NaOH solution, was examined by Norquay *et al.* [128]. The Wiener model of the process, consisting of a step-response model followed by cubic spline nonlinearity, is determined based on a response to an input signal with a random amplitude and frequency and applied to the MPC scheme. The obtained Wiener model gives an excellent fit to data in comparison with a linear step response model, which captures system dynamics but not the gain.

In [130], Norquay *et al.* discuss two different Wiener system identification procedures and apply Wiener models in MPC of a pH neutralization process. In a two stage procedure, a nonlinear element model is fitted to the steady-state data followed by the estimation of linear model parameters. A single stage procedure calculates the parameters of both models simultaneously with prediction error techniques.

MPC used by Norquay *et al.*, see Fig. 6.1 for the block diagram, is based on the internal model control scheme (IMC) [127]. In Fig. 6.1, G_p represents the controlled process, G_c and G_m represent the controller and the process model, respectively. The inverse model of the nonlinear element that follows the Wiener model cancels its nonlinear gain. Applying another inverse model of the nonlinear element, nonlinear properties of the neutralization process are also cancelled. In this way, the nonlinear predictive control problem reduces to a linear one.

It has also been shown by experimental evaluation and testing that MPC developed by Norquay *et al.* outperforms the proportional plus integral plus derivative (PID) and linear MPC control strategies.

The model reference adaptive control scheme for the control of the pH neutralization process in a continuous flow reactor was studied by Pajunen [133]. In this approach, the nonlinear element is modelled by a piecewise linear function and a pulse transfer model is used as a model of the linear system. The parameters of the controller are calculated recursively with the least squares type adaptation algorithm. The method is useful for systems with unknown and time-varying parameters that can be described by the Wiener model. Global

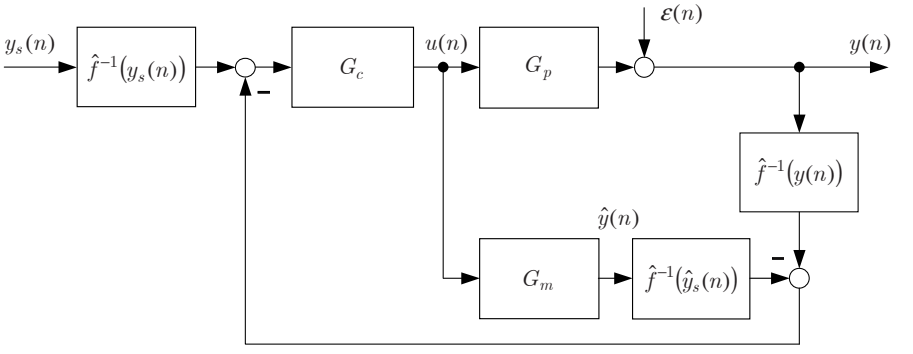

Fig. 6.1. Wiener-model-based prediction control based on the inversion of nonlinearity

stability of the control algorithm is established assuming that the nonlinear element can be exactly represented by a linear spline function for a given set of breakpoints.

Patwardhan *et al.* [134] used a Wiener model-based MPC method to control the pH level in an acid-base neutralization process. The method proposed by Patwardhan *et al.* uses a MIMO Wiener model with two inputs, the acid flow rate and the base flow rate, and two outputs, i.e., the level and pH. The Wiener system was identified with the partial least squares method. The performance of three model predictive controllers, i.e., linear, Hammerstein model-based, and Wiener model-based ones, was compared. Both the Hammerstein and Wiener model-based MPC schemes outperform the linear one. The MPC scheme is also able to meet several set-point changes into nonlinear regions that the other controllers cannot handle.

A control strategy that linearizes nonlinearity using the inverse nonlinear element model was applied by Sung [154] to control a simulated pH neutralization process in the CSTR. The identification of the pH neutralization process is performed by a simple relay-based feedback method that separates the identification of the nonlinear element from the identification of the linear dynamic system. As a result, a nonparametric nonlinear element model, in the form of a look-up table, is obtained and the ultimate gain, i.e., the gain at which the system controlled by the proportional controller oscillates, and the oscillation frequency are determined.

A continuous stirred-tank fermentation process that exhibits both nonlinear and nonstationary features was studied by Roux *et al.* [143]. In this application, the parameters of the Hammerstein model are estimated using the RLS algorithm. Based on three different models, i.e., linear, nonlinear, and Hammerstein ones, adaptive prediction control algorithms are derived and tested. Both one-step-ahead and multi-step-ahead cost functions are applied.

The application of Wiener models in the control of distillation columns was studied in [19, 20, 110, 114, 129, 138, 146, 159]. Bloemen *et al.* [19, 20] used the Wiener model in MPC algorithms based on polytopic descriptions and an inverse of nonlinearity and compared the obtained results with two other algorithms based on a linear state space model and an FIR model. The identified system is a computer simulator of a moderate-high purity distillation column. The Wiener system is identified via an indirect closed loop algorithm based on the subspace identification methods [113, 157, 164].
A pilot binary distillation column with 26 trays was identified by Ling and Rivera [110]. They used a pulse transfer model of the linear system, a polynomial model of the nonlinear element, and a white noise input signal uniformly distributed in $[-0.5, 0.5]$. Four different models, i.e., second order Volterra series, Wiener, Hammerstein, and NARX models were identified. Their performance was evaluated based on their steady-state responses and step responses at different amplitudes. The obtained results can be summarized as follows:

- The NARX model provides the most accurate approximation of column dynamics.
- The second order Volterra series model can capture correctly column dynamics only in a quite limited operation range.
- The Wiener and Hammerstein models can approximate well column dynamics if the input design is selected properly.

Luyben and Eskinat [114] used a continuous-time Hammerstein model for modelling a 20-tray binary distillation column. The identification of the column is performed via consecutive use of two or more relay-feedback tests with different relay heights and different known dynamic elements inserted into the loop. The Hammerstein model was also applied by Luyben and Eskinat to the identification of a steam-water heat exchanger. Exciting the system input with the PRBS signal, the parameters of the Hammerstein model are estimated using the Narendra-Gallman algorithm [120], see Section 5.2.
Considering the identification of nonlinear systems with block-oriented models, Pearson and Pottman [138] illustrated it with the application to a simulated distillation column example. Under the assumption that the nonlinear steady-state characteristics are known a priori, they determined both the Wiener and Hammerstein models and compared their performance with linear and feedback block-oriented models.
A model predictive control of a C2-splitter based on a MIMO Wiener model was considered by Norquay *et al.* [129]. The C2-splitter is a high purity 114 tray distillation column, which separates a feed stream consisting primarily of ethylene and ethane with a small amount of methane into an ethylene product and ethane waste stream. The Wiener model employed in this application is constructed using cubic splines for nonlinear elements and the first order plus dead time models for linear elements. While the reboiler duty (in terms of the flow of the heating stream) and the reflux flow rate are model inputs, the top composition and the bottom temperature of the column are model outputs.

To identify the static nonlinear element, steady-state responses to changes of the process input about a nominal operating point were recorded. Based on the steady state data, a piecewise polynomial model was constructed. To gain some understanding of system dynamics, step tests were performed, and the first order plus dead time models fitted.
Before industrial implementation, the control strategy was tested with a simulated C2-splitter. The applied predictive control algorithm of the IMC type appears successful in the rejection of major disturbances. Comparisons with linear IMC have shown the Wiener model-based approach to be superior.
Sentoni *et al.* [146] used a Wiener model consisting of a set of discrete-time Laguerre systems and a multilayer perceptron model of the nonlinear element to model a simulated binary distillation column. The application of the model was illustrated with a nonlinear model predictive control example.
The chromatographic separation process was modelled with a MIMO neural network Wiener model by Visala *et al.* [159]. The separation unit consists of two interconnected columns and is used for the separation of different compounds from the incoming solution. Separate input-output models are identified for both columns. The model is used for process monitoring. In this way, changes in dynamics and drifts into undesired disturbed states can be observed by on-line simulation. To include changes in process dynamics, the model can be retrained with recent production data.
A two-step identification procedure was used by Cervantes *et al.* to identify the CSTR and the polymerization reactor. First, the linear dynamic system is identified using the correlation method, then the nonlinear element is approximated with a piecewise linear function. On the basis of the obtained Wiener models, an MPC control strategy is developed and tested.
Another two-step identification procedure was used by Rollins and Bhandari [142] for the identification of MIMO Wiener and Hammerstein systems. In the first step, the nonlinear element is identified from the ultimate response data of sequential step tests. The parameters of the linear dynamic system are estimated in the second step under the constraint of the fitted nonlinear element model. A simulation study, based on simulated CSTR, showed higher prediction accuracy of the proposed method in comparison with that of the unconstrained identification procedure.
Nonlinear structure identification of chemical processes based on a nonlinear input-output operator was studied by Menold *et al.* [118]. They introduced a deterministic suitability measure that qualifies the capability of a model class to capture the input-output behavior of a nonlinear system. The suitability measure can be used to select the model structure prior to the actual parameter identification. A strategy for computing the suitability measures with respect to the Hammerstein, the linear, the Wiener, and the diagonal Volterra model, was evaluated with the example of simulated CSTR.
Davide *et al.* [32] applied the Wiener model to the modelling of quartz microbalance polymer-coated sensors used for the detection of mixtures of n-octane and toulene. The Wiener model is composed of two linear dynamic models

followed by a single nonlinear element. Exciting system inputs with two independent white Gaussian signals, linear dynamic systems are identified using the correlation method. Having the linear dynamic systems identified, a polynomial model of the nonlinear element can be fitted.

A radial basis function neural network was applied as a model of a hydraulic pilot plant by Knohl *et al.* [99]. Due to the linear-in-parameters form of the model, the RLS algorithm with a constant trace can be used for parameter estimation. Based on the developed neural network model, an indirect adaptive control strategy was derived and tested.

An inverse nonlinear element model was applied by Knohl and Unbehauen [98] for the compensation of input nonlinearity in an electro-hydraulic servo system. The system is controlled by the standard indirect adaptive control method. In this application, both the nonlinear element and its inverse are modelled with the RBF neural network. As the Hammerstein model of the system which contains a dead zone nonlinearity can be written in the linear-in-parameters form, the RLS algorithm is used for parameter estimation.

A Wiener model describing a pneumatic valve for the control of fluid flow was used by Wigren [165]. In this example, a linear dynamic model is used to describe the dynamic balance between a control signal, a counteractive spring force and friction. A flow through the valve, being the model output, is a nonlinear function of the valve stem position.

Ikonen and Najim [72] modelled the valve pressure difference in a pump-valve pilot system with a MISO Wiener model. The inputs to the model are the pump command signal, the true valve position, and the consistency of the pulp, whereas the pressure difference over the valve is the model output.

Celka *et al.* [25] considered modelling an electrooptical dynamic system running in a chaotic mode. They showed that, under certain assumptions, such a system can be represented by a Wiener model.

An adaptive precompensation of nonlinear distortion in Wiener systems via a precompensator of a Hammerstein model structure was studied by Kang *et al.* [97]. The examined system consists of an adaptive Wiener model-type estimator and an adaptive precompensator that linearizes the overall system. It is assumed that the nonlinear element is modelled by a polynomial. The parameters of the compensator, being ideally the inverse of the Wiener system, are estimated minimizing the mean square error defined as a difference between the delayed input signal and the system output. An adaptive algorithm for updating the parameters of the precompensator uses the stochastic gradient method. The proposed technique can be applied to the compensation of nonlinear distortions in electronic devices or electromechanical components, e.g., high power amplifiers in satellite communication channels, distortion reduction of loudspeakers, active noise cancellation.

The charging process in diesel engines, considered by Aoubi [6], also reveals nonlinear input-output behavior and a strong dependence of its dynamics on the operating point. The SISO generalized Hammerstein model is defined as

$$\hat{A}(q^{-1})\hat{y}(n) = b_0 + \sum_{k=1}^{r} \hat{B}_k(q^{-1})u^k(n). \tag{6.1}$$

Based on real experimental data, two-input one-output generalized Hammerstein model composed of a third order polynomial model of the nonlinear element and a second order linear system is derived. The inputs to the model are the engine speed and the fuel mass, the output is the loading pressure. The model contains 48 adjustable parameters, which are calculated with the least squares technique.

Modelling continuous-time and discrete-time chaotic systems, including the typical Duffing, Henon, and Lozi systems, is another interesting application of Wiener models proposed by Chen *et al.* [28]. In this application, a single hidden layer neural network model of the nonlinear element, trained with the steepest descent algorithm, acts as a neural network controller of the linear part of the Wiener model. The identification of discrete-time chaotic systems using both Wiener and Hammerstein neural network models was also studied by Xu *et al.* [168].

Different control strategies for Hammerstein systems have been considered by several authors, see [13] for adaptive predictive control based on the polynomial Hammerstein model, and [124, 125] for minimum-time dead-beat controllers. The application of the MIMO Hammerstein model to an iron oxide pellet cooling process used in pyrometallurgical installations was studied by Pomerleau *et al.* [140]. For the cooling zone of an induration furnace, they used a two-input two-output Hammerstein model with decoupled models of nonlinear elements and a coupled model of the linear dynamic system. Based on the empirical Hammerstein model, a linear MPC strategy was developed and compared with a phenomenological nonlinear MPC.

The estimation of parameter changes in Wiener and Hammerstein systems in the context of fault detection and isolation is considered in [78, 79, 91, 92], see Section 6.2 for more details. The identification of steam pressure dynamics in a five stage sugar evaporator is studied in [82, 91], more details are given in Section 6.3.

A theoretical model of the muscle relaxation process that expresses the interaction between the drug dose and the relaxation effect can be described as a continuous-time Wiener model [33]. This Wiener model consists of a third order linear dynamic system plus dead time followed by static nonlinearity. The dead time is introduced to model the transport delay of the drug.

Quaglini *et al.* [141] used a Wiener model, composed of the ARX model and a polynomial model of the nonlinear element, for the relaxation function of soft biological tissues. The identification is made by iterative minimization of two different cost functions that are linear in the parameters of the linear dynamic model and the nonlinear element model, respectively.

Wiener and Hammerstein models have also been applied to the modelling of some other biological systems. A brief survey of such applications is given by Hunter and Korenberg [71].

6.2 Fault detection and isolation with Wiener and Hammerstein models

Model-based fault detection and isolation (FDI) methods require both a nominal model of a system, i.e., a model of the system in its normal operation conditions, and models of the faulty system to be available [38, 135, 30]. A nominal model of the system is used in a fault detection step to generate residuals, defined as a difference between output signals of the system and its model. An analysis of these residuals gives an answer to the question whether any fault occurs or not. If a fault occurs, a fault isolation step is performed in a similar way analyzing residual sequences generated with the models of the faulty system – see Fig. 6.2

In the case of complex industrial systems, e.g., a five-stage sugar evaporation station, the above procedure can be even more useful if it is applied not only to the overall system but also to its chosen sub-modules. Designing an FDI system with such an approach may result in both high fault detection sensitivity and high fault isolation reliability.

The problem considered here can be stated as follows: Given system input and output sequences and knowing the nominal models of Wiener or Hammerstein systems, generate a sequence of residuals and process this sequence to detect and isolate all changes of system parameters caused by any system fault. Both abrupt (step-like) and incipient (slowly developing) faults are to be considered as well.

Assume that the nominal models of a Wiener or Hammerstein system defined by the nonlinear function $f(\cdot)$ and the polynomials $A(q^{-1})$ and $B(q^{-1})$ are known. These models describe the systems in their normal operating conditions with no malfunctions (faults). Moreover, assume that at the time k a step-like fault occurred, and caused a change in the mathematical model of the system. This change can be expressed in terms of additive components of pulse transfer function polynomials. The polynomials $\mathcal{A}(q^{-1})$ and $\mathcal{B}(q^{-1})$ of the pulse transfer function $\mathcal{B}(q^{-1})/\mathcal{A}(q^{-1})$ of the faulty system have the form:

$$\mathcal{A}(q^{-1}) = A(q^{-1}) + \Delta A(q^{-1}), \tag{6.2}$$

$$\mathcal{B}(q^{-1}) = B(q^{-1}) + \Delta B(q^{-1}), \tag{6.3}$$

where

$$A(q^{-1}) = a_1 q^{-1} + \cdots + a_{na} q^{-na}, \tag{6.4}$$

$$B(q^{-1}) = b_1 q^{-1} + \cdots + b_{nb} q^{-nb}, \tag{6.5}$$

$$\Delta A(q^{-1}) = \alpha_1 q^{-1} + \cdots + \alpha_{na} q^{-na}, \tag{6.6}$$

$$\Delta B(q^{-1}) = \beta_1 q^{-1} + \cdots + \beta_{nb} q^{-nb}. \tag{6.7}$$

The characteristic of the static nonlinear element $g(\cdot)$ in the faulty state can be expressed as a sum of $f(\cdot)$ and its change $\Delta f(\cdot)$:

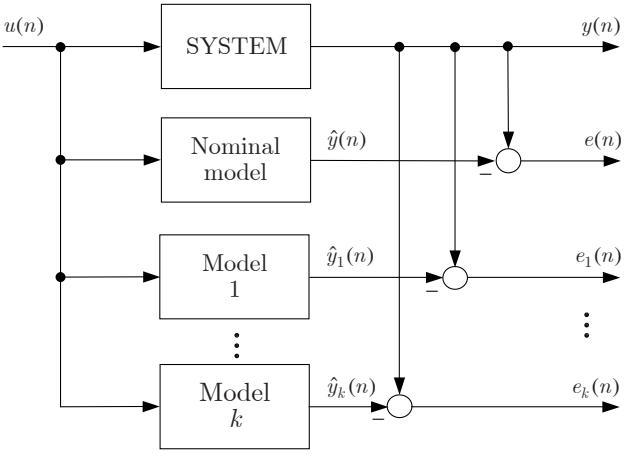

Fig. 6.2. FDI with the nominal model and a bank of models of the faulty system

$$g\big(u(n)\big) = f\big(u(n)\big) + \Delta f\big(u(n)\big), \tag{6.8}$$

where

$$\Delta f\big(u(n)\big) = \mu_0 + \mu_2 u^2(n) + \mu_3 u^3(n) + \cdots . \tag{6.9}$$

For Wiener systems, it is assumed that both $f(\cdot)$ and $g(\cdot)$ are invertible. The inverse nonlinear function of the faulty system can be written as a sum of the inverse function $f^{-1}(\cdot)$ and its change $\Delta f^{-1}(\cdot)$:

$$g^{-1}\big(y(n)\big) = f^{-1}\big(y(n)\big) + \Delta f^{-1}\big(y(n)\big), \tag{6.10}$$

where

$$\Delta f^{-1}\big(y(n)\big) = \gamma_0 + \gamma_2 y^2(n) + \gamma_3 y^3(n) + \cdots . \tag{6.11}$$

Note that $\Delta f^{-1}(\cdot)$ here does not denote the inverse of $\Delta f(\cdot)$ but only a change in the inverse nonlinear characteristic.

6.2.1 Definitions of residuals

The residual $e(n)$ is defined as a difference between the output of the system $y(n)$ and the output of its nominal model $\hat{y}(n)$:

$$e(n) = y(n) - \hat{y}(n). \tag{6.12}$$

Although both series-parallel and parallel Wiener and Hammerstein models can be employed as nominal models for residual generation, we will show that the major advantage of serial-parallel ones is their property of transformation into the linear-in-parameters form. With the nominal series-parallel Hammerstein model (Fig. 6.3) defined by the following expression:

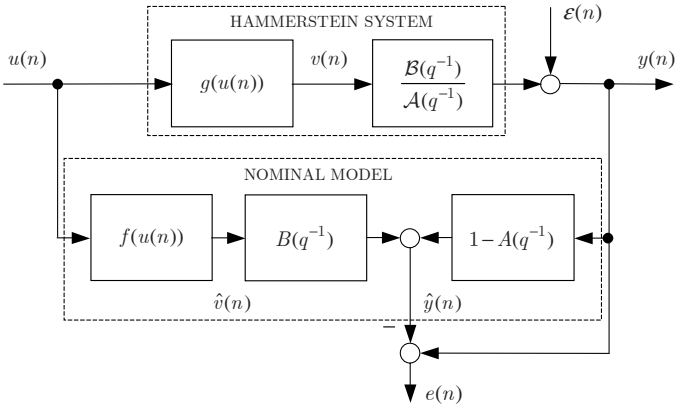

Fig. 6.3. Generation of residuals using the series-parallel Hammerstein model

$$\hat{y}(n) = \big[1 - A(q^{-1})\big]y(n) + B(q^{-1})f\big(u(n)\big), \tag{6.13}$$

we have

$$e(n) = A(q^{-1})y(n) - B(q^{-1})f\big(u(n)\big). \tag{6.14}$$

The output of the faulty Hammerstein system is

$$y(n) = \frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})}g\big(u(n)\big) + \varepsilon(n). \tag{6.15}$$

Thus, taking into account (6.2), (6.3) and (6.8), (6.15) can be written in the following form:

$$\begin{aligned} y(n) = {} & \big[1 - A(q^{-1}) - \Delta A(q^{-1})\big]y(n) + \big[B(q^{-1}) + \Delta B(q^{-1})\big] \\ & \big[f\big(u(n)\big) + \Delta f\big(u(n)\big)\big] + \mathcal{A}(q^{-1})\varepsilon(n). \end{aligned} \tag{6.16}$$

Next, assume that the disturbance $\varepsilon(n)$ is

$$\varepsilon(n) = \frac{\epsilon(n)}{\mathcal{A}(q^{-1})}, \tag{6.17}$$

where $\epsilon(n)$ is a zero-mean white noise. Now, the substitution of (6.17) into (6.16), and (6.13) and (6.16) into (6.12) results in a residual equation expressed in terms of changes of the polynomials $A(q^{-1})$, $B(q^{-1})$ and the function $f\big(u(n)\big)$:

$$\begin{aligned} e(n) = {} & -\Delta A(q^{-1})y(n) + \Delta B(q^{-1})f\big(u(n)\big) + \big[B(q^{-1}) \\ & + \Delta B(q^{-1})\big]\Delta f\big(u(n)\big) + \epsilon(n). \end{aligned} \tag{6.18}$$

Similar deliberations for the parallel Hammerstein model (Fig. 6.4) give

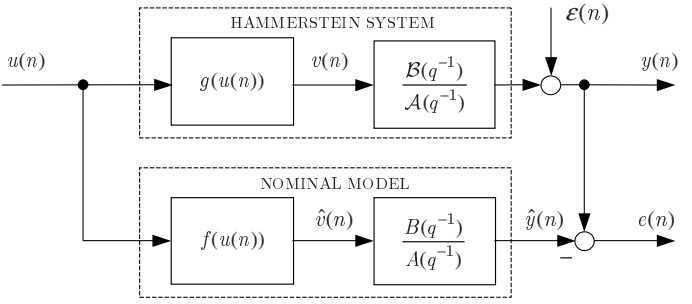

Fig. 6.4. Generation of residuals using the parallel Hammerstein model

$$\begin{aligned} e(n) = {} & \frac{A(q^{-1})\Delta B(q^{-1}) - B(q^{-1})\Delta A(q^{-1})}{A(q^{-1})\big[A(q^{-1}) + \Delta A(q^{-1})\big]} f\big(u(n)\big) \\ & + \frac{B(q^{-1}) + \Delta B(q^{-1})}{A(q^{-1}) + \Delta A(q^{-1})} \Delta f\big(u(n)\big) + \varepsilon(n). \end{aligned} \tag{6.19}$$

In the case of Wiener systems, residual generation with the series-parallel Wiener model is more complicated as both the model of the nonlinear element and the model of the inverse nonlinear element are used (Fig. 6.5). A simpler form of the series-parallel model can be obtained if we assume that the function $f(\cdot)$ is invertible and introduce the following modified definition of the residual [78]:

$$e(n) = f^{-1}\big(y(n)\big) - f^{-1}\big(\hat{y}(n)\big). \tag{6.20}$$

The nominal parallel Wiener model can be expressed as

$$\hat{y}(n) = f\left(\frac{B(q^{-1})}{A(q^{-1})} u(n)\right). \tag{6.21}$$

The output of the faulty Wiener system is

$$y(n) = g\left[\frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})} g\big(u(n)\big) + \varepsilon(n)\right]. \tag{6.22}$$

From (6.21), it follows that the output of the linear dynamic model is

$$f^{-1}\big(\hat{y}(n)\big) = \frac{B(q^{-1})}{A(q^{-1})} u(n). \tag{6.23}$$

Equation (6.23) can be written in the form

$$f^{-1}\big(\hat{y}(n)\big) = \big[1 - A(q^{-1})\big] f^{-1}\big(\hat{y}(n)\big) + B(q^{-1})u(n). \tag{6.24}$$

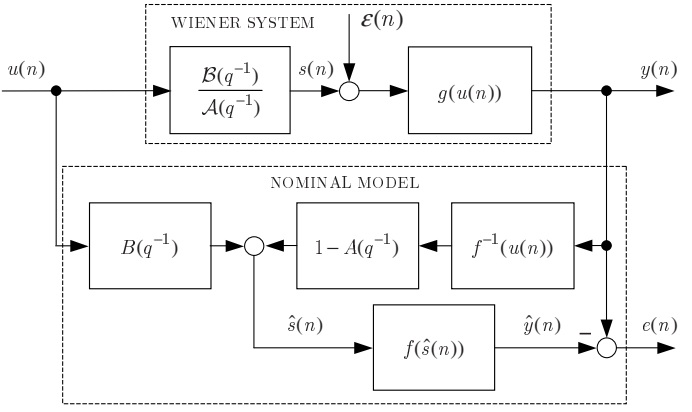

Fig. 6.5. Generation of residuals using the series-parallel Wiener model

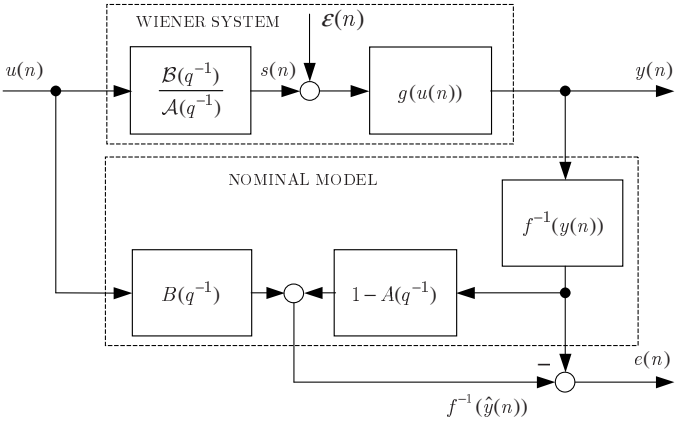

Fig. 6.6. Generation of residuals using the modified series-parallel Wiener model.

Therefore, the modified series-parallel model can be defined as follows:

$$f^{-1}\big(\hat{y}(n)\big) = \big[1 - A(q^{-1})\big] f^{-1}\big(y(n)\big) + B(q^{-1})u(n). \tag{6.25}$$

The generation of residuals with the modified series-parallel Wiener model is illustrated in Fig. 6.6. Taking into account (6.2), (6.3) and (6.10), it follows from (6.22) that the signal $f^{-1}\big(y(n)\big)$ for the faulty Wiener system can be expressed as

$$f^{-1}\big(y(n)\big) = \big[1 - A(q^{-1}) - \Delta A(q^{-1})\big]\big[f^{-1}\big(y(n)\big) + \Delta f^{-1}\big(y(n)\big)\big] + \big[B(q^{-1}) + \Delta B(q^{-1})\big]u(n) - \Delta f^{-1}\big(y(n)\big) + \mathcal{A}(q^{-1})\varepsilon(n). \tag{6.26}$$

Hence, assuming that the disturbance $\varepsilon(n)$ is given by (6.17), the residual (6.20) is

$$e(n) = -\Delta A(q^{-1})f^{-1}\big(y(n)\big) + \Delta B(q^{-1})u(n) - \big[A(q^{-1}) + \Delta A(q^{-1})\big]\Delta f^{-1}\big(y(n)\big) + \epsilon(n). \tag{6.27}$$

A major advantage of both the series-parallel Hammerstein model and the modified series-parallel Wiener model is that their residual equations (6.18) and (6.27), after a proper change of parameterization, can be written in the linear-in-parameters form. Then these redefined parameters can be estimated with linear regression methods.

6.2.2 Hammerstein system. Parameter estimation of the residual equation

Assume that the function $\Delta f(\cdot)$ describing a change of the steady-state characteristic $f(\cdot)$ caused by a fault has the form of a polynomial of the order r:

$$\Delta f\big(u(n)\big) = \mu_0 + \mu_2 u^2(n) + \cdots + \mu_r u^r(n). \tag{6.28}$$

Thus, the residual equation (6.18) can be written in the following linear-in-parameters form:

$$e(n) = -\sum_{j=1}^{na} \alpha_j y(n-j) + \sum_{j=1}^{nb} \beta_j f\big(u(n-j)\big) + d_0 + \sum_{k=2}^{r}\sum_{j=1}^{nb} d_{kj} u^k(n-j) + \epsilon(n), \tag{6.29}$$

where

$$d_0 = \mu_0 \sum_{j=1}^{nb} (b_j + \beta_j), \tag{6.30}$$

$$d_{kj} = \mu_k (b_j + \beta_j). \tag{6.31}$$

Note that (6.29) has $M = na + nb + nb(r-1) + 1$ unknown parameters α_j, β_j, d_{kj} and d_0. In a disturbance-free case ($\epsilon(n) = 0$), these parameters can be calculated performing $N = M$ measurements of the input and output signals and solving the following set of linear equations:

$$e(n) = -\sum_{j=1}^{na} \alpha_j y(n-j) + \sum_{j=1}^{nb} \beta_j f\big(u(n-j)\big) + d_0 + \sum_{k=2}^{r}\sum_{j=1}^{nb} d_{kj} u^k(n-j), \quad n = 1, 2, \ldots, M. \tag{6.32}$$

In practice, it is more realistic to assume that the system output is disturbed by the disturbance (6.17). Now, performing $N \gg M$ measurements of the input and output signals, the parameters of the residual equation (6.29) can be estimated using the least squares method with the parameter vector $\hat{\boldsymbol{\theta}}(n)$ and the regression vector $\mathbf{x}(n)$ defined as follows:

$$\begin{aligned}\hat{\boldsymbol{\theta}}(n) &= \left[\hat{\alpha}_1 \ \ldots \ \hat{\alpha}_{na} \ \hat{\beta}_1 \ \ldots \ \hat{\beta}_{nb} \ \hat{d}_0 \ \hat{d}_{21} \ldots \hat{d}_{2nb} \ldots \hat{d}_{r1} \ldots \hat{d}_{rnb}\right]^T, \\ \mathbf{x}(n) &= \big[-y(n-1) \ldots -y(n-na) \ f\big(u(n-1)\big) \ldots f\big(u(n-nb)\big) \\ &\qquad 1 \ u^2(n-1) \ldots u^2(n-nb) \ldots u^r(n-1) \ldots u^r(n-nb)\big]^T.\end{aligned} \tag{6.33}$$

The vector $\hat{\boldsymbol{\theta}}(n)$ can be calculated on-line using the recursive least squares algorithm:

$$\hat{\boldsymbol{\theta}}(n) = \hat{\boldsymbol{\theta}}(n-1) + \boldsymbol{P}(n)\mathbf{x}(n)\big(e(n) - \mathbf{x}^T(n)\hat{\boldsymbol{\theta}}(n-1)\big), \tag{6.34}$$

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) - \frac{\boldsymbol{P}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\boldsymbol{P}(n-1)}{1 + \mathbf{x}^T(n)\boldsymbol{P}(n-1)\mathbf{x}(n)}, \tag{6.35}$$

with $\hat{\boldsymbol{\theta}}(0) = \mathbf{0}$ i $\mathbf{P}(0) = \alpha\mathbf{I}$, where $\alpha \gg 1$ and $\mathbf{I}$ is an identity matrix. Parameter estimation of the residual equation (6.29) often results in asymptotically biased estimates. That is the case if system additive output disturbances are not given by (6.17) but have the property of zero-mean white noise, i.e., $\varepsilon(n) = \epsilon(n)$. In this case, the residual equation (6.18) has the form

$$\begin{aligned}e(n) = &-\Delta A(q^{-1})y(n) + \Delta B(q^{-1})f\big(u(n)\big) + \big[B(q^{-1}) \\ &+ \Delta B(q^{-1})\big]\Delta f\big(u(n)\big) + \big[A(q^{-1}) + \Delta A(q^{-1})\big]\epsilon(n).\end{aligned} \tag{6.36}$$

The term $[A(q^{-1}) + \Delta A(q^{-1})]\,\epsilon(n)$ in (6.36) makes the parameter estimates asymptotically biased and the residuals $e(n) - \mathbf{x}^T(n)\hat{\boldsymbol{\theta}}(n-1)$ correlated. To obtain unbiased parameter estimates, other known parameter estimation methods, such as the instrumental variables method, the generalized least squares method, or the extended least squares method, can be employed [35, 149]. Using the extended least squares method, the vectors $\hat{\boldsymbol{\theta}}(n)$ i $\mathbf{x}(n)$ are defined as follows:

$$\hat{\boldsymbol{\theta}}(n) = \left[\hat{\alpha}_1 \ \ldots \ \hat{\alpha}_{na} \ \hat{\beta}_1 \ldots \hat{\beta}_{nb} \ \hat{d}_0 \ \hat{d}_{21} \ldots \hat{d}_{2nb} \ldots \hat{d}_{r1} \ldots \hat{d}_{rnb} \ \hat{c}_1 \ldots \hat{c}_{na}\right]^T, \tag{6.37}$$

$$\begin{aligned}\mathbf{x}(n) = &\big[-y(n-1)\ldots-y(n-na) \ f\big(u(n-1)\big)\ldots f\big(u(n-nb)\big) \ 1 \ u^2(n-1) \\ &\ldots u^2(n-nb)\ldots u^r(n-1)\ldots u^r(n-nb) \ \hat{e}(n-1)\ldots\hat{e}(n-na)\big]^T,\end{aligned} \tag{6.38}$$

where $\hat{e}(n)$ denotes the one step ahead prediction error of $e(n)$:

$$\hat{e}(n) = e(n) - \mathbf{x}^T(n)\hat{\boldsymbol{\theta}}(n-1). \tag{6.39}$$

Example 6.1 The nominal Hammerstein model composed of the linear dynamic system

$$\frac{B(q^{-1})}{A(q^{-1})} = \frac{0.63467q^{-1} + 0.48069q^{-2}}{1 - 1.3231q^{-1} + 0.4346q^{-2}} \tag{6.40}$$

and the nonlinear element

$$f\big(u(n)\big) = \tanh(u(n)) \tag{6.41}$$

was used in the simulation example. The faulty Hammerstein system is described as

$$\frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})} = \frac{0.1752q^{-1} + 0.1534q^{-2}}{1 - 1.6375q^{-1} + 0.6703q^{-2}}, \tag{6.42}$$

$$\begin{aligned} g\big(u(n)\big) = {} & \tanh(u(n)) - 0.25u^2(n) - 0.2u^3(n) + 0.15u^4(n) \\ & - 0.1u^5(n) + 0.05u^6(n) - 0.025u^7(n) + 0.0125u^8(n). \end{aligned} \tag{6.43}$$

The steady-state characteristics of both the nominal model and the faulty Hammerstein system are shown in Fig. 6.7. The system input was excited with a sequence of $N = 25000$ pseudorandom numbers of uniform distribution in the interval $(-1.5,\ 1.5)$. The system output was disturbed by the disturbance $\varepsilon(n)$ defined as

$$\varepsilon(n) = \frac{\epsilon(n)}{\mathcal{A}(q^{-1})} \tag{6.44}$$

or

$$\varepsilon(n) = \epsilon(n), \tag{6.45}$$

where $\epsilon(n)$ are the values of a Gaussian pseudorandom sequence $\mathcal{N}(0, 0.05)$. For the system disturbed by (6.44), the parameters of (6.29) were estimated with the RLS method – the RLS_{I} example. In the case of the disturbance (6.45), the parameter estimation was performed using both the RLS and RELS methods – the RLS_{II} and RELS_{II} examples. The obtained parameter estimates are given in Table 6.1. The identification error of $\Delta f\big(u(n)\big)$ is shown in Fig. 6.8. A comparison of estimation accuracy measured by four different indices defining estimation accuracy of the residuals $e(n)$, the function $\Delta f\big(u(n)\big)$, the parameters $\mu_j, j = 2, \ldots, 8$, and α_j i β_j, $j = 1, 2$, is given in Table 6.2. The highest accuracy is obtained in the RLS_{I} example for the system disturbed by (6.44). For the system disturbed with (6.45), a comparison of parameter estimates and indices confirms asymptotical biasing of parameter estimates obtained in the RLS_{II} example.

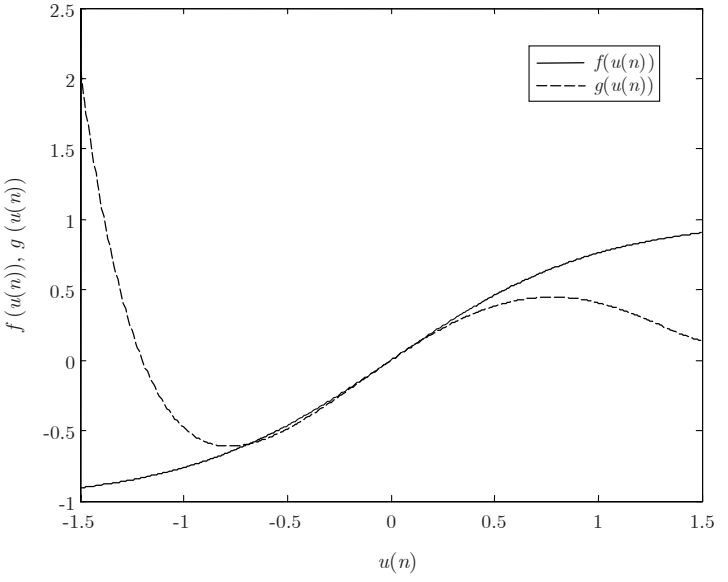

Fig. 6.7. Hammerstein system. Nonlinear functions $f\big(u(n)\big)$ and $g\big(u(n)\big)$

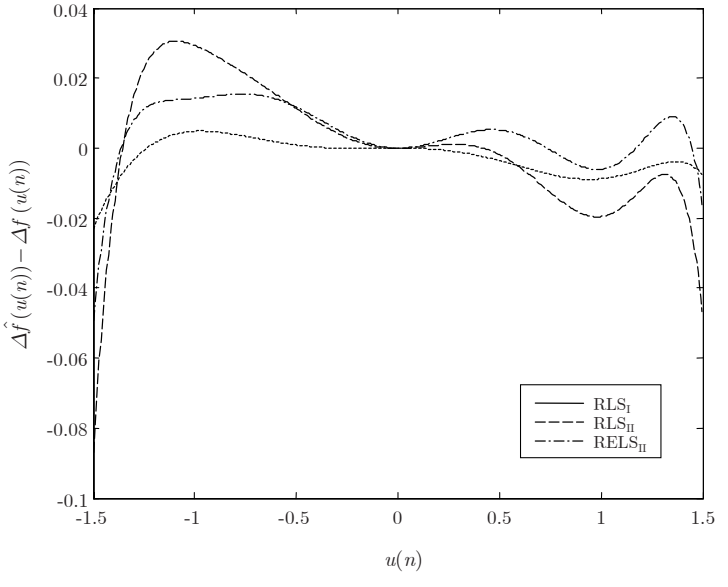

Fig. 6.8. Identification error of the function $\Delta f\big(u(n)\big)$

Table 6.1. Parameter estimates

Parametr	True value	RLS_I $\varepsilon(n) = \frac{\epsilon(n)}{\mathcal{A}(q^{-1})}$	RLS_II $\varepsilon(n) = \epsilon(n)$	RELS_II $\varepsilon(n) = \epsilon(n)$
α_1	-0.3144	-0.3151	-0.1478	-0.3062
α_2	0.2357	0.2369	0.0708	0.2269
β_1	-0.4594	-0.4616	-0.4664	-0.4625
β_2	-0.3273	-0.3305	-0.2972	-0.3278
μ_2	-0.2500	-0.2431	-0.2871	-0.3124
μ_3	-0.2000	-0.1770	-0.1362	-0.1671
μ_4	0.1500	0.1440	0.2431	0.2908
μ_5	-0.1000	-0.1202	-0.1460	-0.1275
μ_6	0.0500	0.0498	-0.0372	-0.0604
μ_7	-0.0250	-0.0210	-0.0183	-0.0202
μ_8	0.0125	0.0138	0.0387	0.0405
c_1	-1.6375			-1.1253
c_2	0.6703			0.1341

Table 6.2. Comparison of estimation accuracy

Index	RLS_I $\varepsilon(n) = \frac{\epsilon(n)}{\mathcal{A}(q^{-1})}$	RLS_II $\varepsilon(n) = \epsilon(n)$	RELS_II $\varepsilon(n) = \epsilon(n)$
$\frac{1}{N}\sum_{i=1}^{N}\left[e(n)-\hat{e}(n)\right]^2$	2.82×10^{-3}	9.62×10^{-3}	4.49×10^{-3}
$\frac{1}{400}\sum_{i=1}^{400}\left[f\big(u(n)\big)-\hat{f}\big(u(n)\big)\right]^2$	3.69×10^{-5}	8.77×10^{-4}	1.37×10^{-4}
$\frac{1}{7}\sum_{j=2}^{8}(\mu_j-\hat{\mu}_j)^2$	3.08×10^{-5}	3.45×10^{-4}	9.87×10^{-5}
$\frac{1}{4}\sum_{j=1}^{2}\left[(a_j-\hat{a}_j)^2+(b_j-\hat{b}_j)^2\right]$	4.07×10^{-6}	1.40×10^{-2}	3.88×10^{-5}

6.2.3 Wiener system. Parameter estimation of the residual equation

Assume that the function $\Delta f^{-1}(\cdot)$ describing a change in the inverse steady-state characteristic caused by a fault has the form of a polynomial of the order r:

$$\Delta f^{-1}\big(y(n)\big) = \gamma_0 + \gamma_2 y^2(n) + \cdots + \gamma_r y^r(n). \tag{6.46}$$

As in the case of the Hammerstein system, the residual equation (6.27) can be expressed in the linear-in-parameters form:

$$\begin{aligned} e(n) = &- \sum_{j=1}^{na} \alpha_j f^{-1}\big(y(n-j)\big) + \sum_{j=1}^{nb} \beta_j u(n-j) + d_0 \\ &+ \sum_{k=2}^{r} \sum_{j=0}^{na} d_{kj} y^k(n-j) + \epsilon(n), \end{aligned} \tag{6.47}$$

where

$$d_0 = -\gamma_0 \Big[1 + \sum_{j=1}^{na} (a_j + \alpha_j)\Big], \tag{6.48}$$

$$d_{kj} = \begin{cases} -\gamma_k, & j = 0 \\ -\gamma_k (a_j + \alpha_j), & j = 1, \ldots, na. \end{cases} \tag{6.49}$$

For disturbance-free Wiener systems, the parameters α_j, β_j, d_{kj} and d_0 can be calculated solving a set of $M = na + nb + (na+1)(r-1) + 1$ linear equations. In the stochastic case, parameter estimation with the parameter vector $\hat{\boldsymbol{\theta}}(n)$ and the regression vector $\mathbf{x}(n)$ defined as

$$\hat{\boldsymbol{\theta}}(n) = \big[\hat{\alpha}_1 \ \ldots \ \hat{\alpha}_{na} \ \hat{\beta}_1 \ \ldots \ \hat{\beta}_{nb} \ \hat{d}_0 \ \hat{d}_{20} \ldots \hat{d}_{2na} \ldots \hat{d}_{r0} \ldots \hat{d}_{rna}\big]^T, \tag{6.50}$$

$$\begin{aligned} \mathbf{x}(n) = \big[&-f^{-1}\big(y(n-1)\big) \ldots - f^{-1}\big(y(n-na)\big) \ u(n-1) \ldots u(n-nb) \\ &1 \ y^2(n) \ldots y(n-na)^2 \ldots y^r(n) \ldots y(n-na)^r\big]^T \end{aligned} \tag{6.51}$$

can be performed with the LS or the RLS method.
An alternative to the above procedure can be parameter estimation with the extended least squares (ELS) algorithm. In this case, the parameter vector $\boldsymbol{\theta}(n)$ and the regression vector are defined as

$$\hat{\boldsymbol{\theta}}(n) = \big[\hat{\alpha}_1 \ \ldots \hat{\alpha}_{na} \ \hat{\beta}_1 \ldots \hat{\beta}_{nb} \ \hat{d}_0 \ \hat{d}_{20} \ldots \hat{d}_{2na} \ldots \hat{d}_{r0} \ldots \hat{d}_{rna} \ \hat{c}_1 \ldots \hat{c}_{na}\big]^T, \tag{6.52}$$

$$\begin{aligned} \mathbf{x}(n) = \big[&- f^{-1}\big(y(n-1)\big) \ldots - f^{-1}(y(n-na)\big) \ u(n-1) \ldots u(n-nb) \\ &1 \ y^2(n) \ldots y(n-na)^2 \ldots y^r(n) \ldots y(n-na)^r \ \hat{e}(n-1) \ldots \\ &\hat{e}(n-na)\big]^T, \end{aligned} \tag{6.53}$$

where $\hat{e}(n)$ denotes the one step ahead prediction error of $e(n)$:

$$\hat{e}(n) = e(n) - \mathbf{x}^T(n)\hat{\boldsymbol{\theta}}(n-1). \tag{6.54}$$

The parameters obtained with the LS method or the ELS method are asymptotically biased. This comes from the well-known property of the least squares

method stating that to obtain consistent parameter estimates, the regression vector $\mathbf{x}(n)$ should be uncorrelated with the system disturbance $\epsilon(n)$, i.e., $E[\,\mathbf{x}(n)\,\epsilon(n)\,] = \mathbf{0}$. Obviously, this condition is not fulfilled for the model (6.47) as the powered system outputs $y^2(n), \ldots, y^r(n)$ depend on $\epsilon(n)$. A common way to obtain consistent parameter estimates in such cases is the application of the instrumental variable method (IV) or its recursive version (RIV). Instrumental variables should be chosen to be correlated with regressors and uncorrelated with the system disturbance $\epsilon(n)$. Although different choices of instrumental variables can be made, replacing $y^2(n), \ldots, y^r(n-na)$ with their approximated values obtained by filtering $u(n)$ through the nominal model is a good choice. This leads to the following estimation procedure:

1. Simulate the nominal model (6.21) to obtain $\hat{y}(n)$.
2. Estimate the parameters of (6.47) using the IV or the RIV method with the instrumental variables

$$\begin{aligned}\mathbf{z}(n) = \big[&- f^{-1}\big(y(n-1)\big) \ldots - f^{-1}\big(y(n-na)\big)\ u(n-1) \ldots u(n-nb) \\ &1\ \hat{y}^2(n) \ldots \hat{y}^2(n-na) \ldots \hat{y}^r(n) \ldots \hat{y}^r(n-na)\big]^T.\end{aligned}$$

Example 6.2. The nominal Wiener model used in the simulation example consists of the linear dynamic system

$$\frac{B(q^{-1})}{A(q^{-1})} = \frac{0.5q^{-1} - 0.3q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \tag{6.55}$$

and the nonlinear element of the following inverse nonlinear characteristic:

$$f^{-1}\big(y(n)\big) = y(n) - \frac{1}{6}y^3(n). \tag{6.56}$$

The faulty Wiener system is defined by the pulse transfer function

$$\frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})} = \frac{0.3q^{-1} - 0.2q^{-2}}{1 - 1.75q^{-1} + 0.85q^{-2}} \tag{6.57}$$

and the inverse nonlinear characteristic (Fig. 6.9) of the form

$$g^{-1}\big(y(n)\big) = \tan\big(y(n)\big). \tag{6.58}$$

The system was excited with a sequence of $N = 50000$ pseudorandom values of uniform distribution in $(-1, 1)$. The system output was disturbed additively with the disturbances (6.45), with $\{\epsilon(n)\}$ – a pseudorandom sequence, uniformly distributed in $(-0.01,\ 0.01)$. Parameter estimation was performed with the RLS and RELS methods for both a disturbance-free case – $\mathrm{RLS_I}$ example, and for the system disturbed by (6.45) – the $\mathrm{RLS_{II}}$ $\mathrm{RELS_{II}}$ examples. The parameter estimates are given in Table 6.4, and Table 6.3 shows a comparison of estimation accuracy. The identification error of the function

$\Delta f^{-1}(\cdot)$ is shown in Fig. 6.10. An assumed finite order of the polynomial (6.46), i.e., $r = 11$ is another source of the identification error as the function $\Delta f^{-1}(\cdot)$ can be represented accurately with the power series. A comparison of the results obtained in the RLS_{II} example and the RELS_{II} example shows the inconsistency of the RLS estimator.

Table 6.3. Comparison of estimation accuracy

Example Index	RLS_{I} $\varepsilon(n) = 0$	RLS_{II} $\varepsilon(n) = \epsilon(n)$	RELS_{II} $\varepsilon(n) = \epsilon(n)$
$\frac{1}{N}\sum_{i=1}^{N}\left[e(n) - \hat{e}(n)\right]^2$	8.07×10^{-6}	1.64×10^{-4}	8.60×10^{-5}
$\frac{1}{400}\sum_{i=1}^{400}\left[\Delta f^{-1}\big(\hat{y}(n)\big) - \Delta\hat{f}^{-1}\big(\hat{y}(n)\big)\right]^2$	1.39×10^{-6}	1.29×10^{-4}	3.68×10^{-5}
$\frac{1}{10}\sum_{j=2}^{11}\left(\gamma_j - \hat{\gamma}_j\right)^2$	1.73×10^{-4}	1.44×10^{-2}	4.31×10^{-3}
$\frac{1}{4}\sum_{j=1}^{2}\left[(a_j - \hat{a}_j)^2 + (b_j - \hat{b}_j)^2\right]$	1.72×10^{-6}	4.09×10^{-4}	5.75×10^{-6}

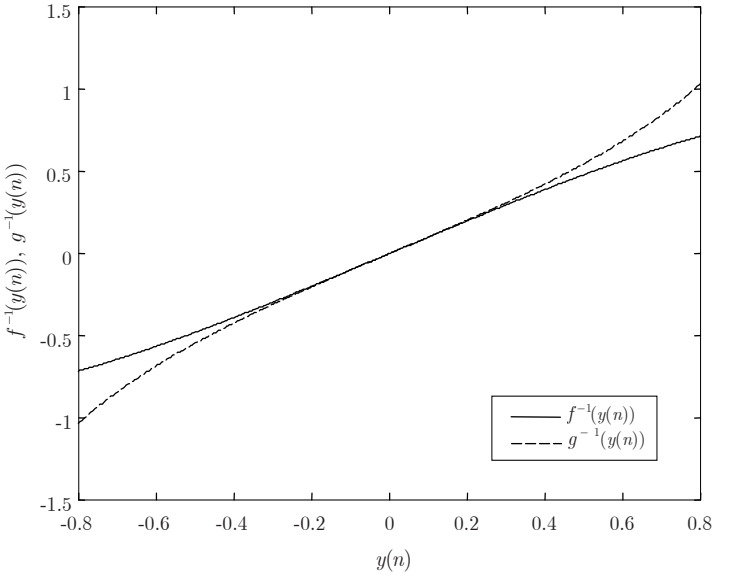

Fig. 6.9. Wiener system. Inverse nonlinear functions $f^{-1}\big(\hat{y}(n)\big)$ and $g^{-1}\big(\hat{y}(n)\big)$

Table 6.4. Parameter estimates

Parameter	True value	RLS_{I} $\varepsilon(n) = 0$	RLS_{II} $\varepsilon(n) = \epsilon(n)$	RELS_{II} $\varepsilon(n) = \epsilon(n)$
α_1	-0.2500	-0.2482	-0.2164	-0.2487
α_2	0.1500	0.1491	0.1293	0.1526
β_1	-0.2000	-0.2006	-0.2060	-0.2032
β_2	0.1000	0.1006	0.1065	0.1021
γ_2	0	-0.0000	-0.0019	-0.0002
γ_3	0.5000	0.4833	0.3445	0.4177
γ_4	0	0.0002	0.0054	-0.0002
γ_5	0.1333	0.1646	0.4240	0.2895
γ_6	0	-0.0004	-0.0057	-0.0011
γ_7	0.0540	0.0372	-0.0835	-0.0277
γ_8	0	-0.0004	-0.0053	-0.0029
γ_9	0.0219	0.0136	-0.0846	-0.0346
γ_{10}	0	0.0010	0.0084	0.0068
γ_{11}	0.0089	0.0199	0.0753	0.0539
c_1	-1.7500			-0.9514
c_2	0.8500			-0.0061

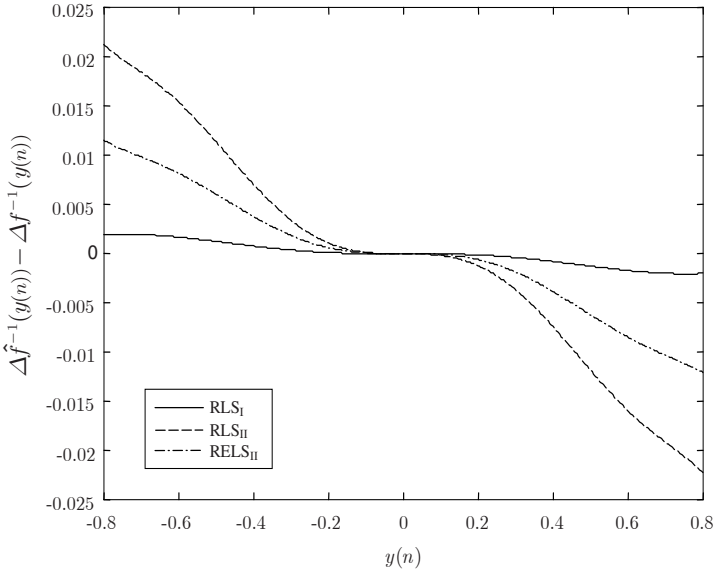

Fig. 6.10. Identification error of the inverse nonlinear function $\Delta f^{-1}\big(u(n)\big)$

6.3 Sugar evaporator. Identification of the nominal model of steam pressure dynamics

The main task of a multiple-effect sugar evaporator is thickening the thin sugar juice from sugar density of approximately 14 to 65–70 Brix units (Bx). Other important tasks are steam generation, waste steam condensation, and supplying waste-heat boilers with water condensate. In a multiple-effect evaporation process, the sugar juice and the saturated steam are fed to the successive stages of the evaporator at a gradually decreasing pressure. Many complex physical phenomena and chemical reactions occur during the thickening of the sugar juice, including sucrose decomposition, the precipitation of calcium compounds, the decomposition of acid amides, etc. A flow of the steam and sugar juice through the successive stages of the evaporator results in a close relationship between temperatures and pressures of the juice steam in these stages. Moreover, juice steam temperature and juice steam pressure depend also on other physical quantities such as the rate of the juice flow or juice temperature.

6.3.1 Theoretical model

A theoretical approach to process modelling relies upon deriving a mathematical model of the process by applying the laws of physics. Theoretical models obtained in this way are often too complicated to be used, for example, in control applications. Despite this, theoretical models are a source of valuable information on the nature of the process. This information can be very useful for the identification of experimental mathematical models based on process the input-output data. Moreover, theoretical models can serve as benchmarks for the evaluation and verification of experimental models.
Modelling the multiple-effect sugar evaporator is complicated by the fact that it consists of a number of evaporators connected both in series and in parallel. All this makes the formulation of a theoretical model a difficult task. Theoretical models are derived based on mass and energy balances for all stages. To perform this, the following assumptions are made [111]:

1. The juice and the steam are in a saturated equilibrium.
2. Both the mass of the steam in the juice chamber and the mass of the steam in the steam chamber are constant.
3. The operation of juice level controllers makes it possible to neglect variations of juice levels.
4. Heat losses to the surrounding environment are negligible.
5. Mixing is perfect.

The model of the dependence of steam pressure in the steam chamber of the stage k on steam pressure in the steam chamber of the stage $k-1$, given by Carlos and Corripio [24], has the form

$$P_k = P_{k-1} - \frac{\gamma_k (O_{k-1})^2}{\rho_k^2 (T_{k-1}, P_{k-1})}, \tag{6.59}$$

$$P_1 = P_0 - \frac{\gamma_1 S^2}{\rho_1^2 (T_0, P_0)}, \tag{6.60}$$

where P_k denotes juice steam pressure in the steam chamber of the stage k, γ_k is the conversion factor, O_k is the steam flow rate from the stage k, ρ_k is the steam density at the stage k, and T_k is steam temperature at the stage k. The models (6.59) and (6.60) describe steady-state behavior of the process. To include dynamic properties of the process, a modified model described with differential equations of the first order was proposed:

$$P_k = \tau_k \frac{dP_{k-1}}{dt} + P_{k-1} - \frac{\gamma_k (O_{k-1})^2}{\rho_k^2 (T_{k-1}, P_{k-1})}, \tag{6.61}$$

$$P_1 = \tau_1 \frac{dP_0}{dt} + P_0 - \frac{\gamma_1 S^2}{\rho_1^2 (T_0, P_0)}, \tag{6.62}$$

where τ_k is the time constant.
Equations (6.59) – (6.62) are a part of the overall model of the sucrose juice concentration process. This model comprises also mass and energy balances for all stages [111].

6.3.2 Experimental models of steam pressure dynamics

In a neural network model (Fig. 6.11), it is assumed that the model output $\hat{P}_2(n)$ is a sum of two components defined by the nonlinear function of the pressure $P_3(n)$ and the linear function of the pressure $P_1(n)$. The model output at the time n is

$$\hat{P}_2(n) = \hat{f}\left[\frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} P_3(n)\right] + \hat{C}(q^{-1}) P_1(n), \tag{6.63}$$

with

$$\hat{A}(q^{-1}) = 1 + \hat{a}_1 q^{-1} + \hat{a}_2 q^{-2}, \tag{6.64}$$

$$\hat{B}(q^{-1}) = \hat{b}_1 q^{-1} + \hat{b}_2 q^{-2}, \tag{6.65}$$

$$\hat{C}(q^{-1}) = \hat{c}_1 q^{-1} + \hat{c}_2 q^{-2}, \tag{6.66}$$

where $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_1$, $\hat{b}_2$, $\hat{c}_1$ i $\hat{c}_2$ denote model parameters. The function $\hat{f}(\cdot)$ is modelled with a multilayer perceptron containing one hidden layer consisting of M nodes of the hyperbolic tangent activation function:

$$\hat{f}(\hat{s}(n)) = \sum_{j=1}^{M} w_{1j}^{(2)} \tanh\left(x_j(n)\right) + w_{10}^{(2)}, \tag{6.67}$$

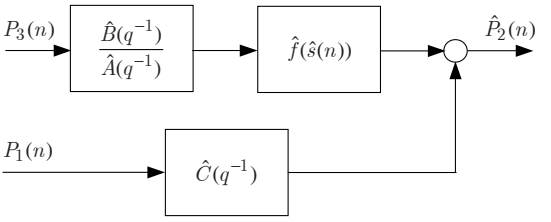

Fig. 6.11. Neural network model of steam pressure dynamics

$$x_j(n) = w_{j1}^{(1)} \hat{s}(n) + w_{j0}^{(1)}, \tag{6.68}$$

where $\hat{s}(n) = \left[\hat{B}(q^{-1})/\hat{A}(q^{-1})\right] P_3(n)$ is the output of a linear dynamic model, and $w_{10}^{(1)}, \ldots, w_{M1}^{(1)}$, $w_{10}^{(2)}, \ldots, w_{1M}^{(2)}$ are the weights of the neural network. The neural network structure contains a Wiener model [76, 75] in the path of the pressure $P_3(n)$ and a linear finite impulse response filter of the second order in the path of the pressure $P_1(n)$. In a linear model of steam pressure dynamics, the Wiener model is replaced with a linear dynamic model:

$$\hat{P}_2(n) = \frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} P_3(n) + \hat{C}(q^{-1}) P_1(n). \tag{6.69}$$

6.3.3 Estimation results

Parameter estimation of the models (6.63) i (6.69) was performed based on a set of 10000 input-output measurements recorded at the sampling rate of 10s. The RLS algorithm was employed for the estimation of the ARX model. The neural network Wiener model of the structure $\mathcal{N}(1-24-1)$ was trained recursively with the backpropagation learning algorithm [81]. For both models, 50000 sequential steps were made, processing the overall data set five times. Then the models were tested with another data set of 8000 input-output measurements.

The results of the estimation and the testing are shown in Figs 6.13 – 6.15, and a comparison of the test results for both models is given in Table 6.5. Lower values of the mean square prediction error of the pressure P_2 obtained for the neural network model for both the training and testing sets confirm the nonlinear nature of the process.

As the analyzed model of steam pressure dynamics is characterized by a low time constant of approximately 20s, fast fault detection is possible. Moreover, as steam pressure is not controlled automatically, there is no problem of closed-loop system identification.

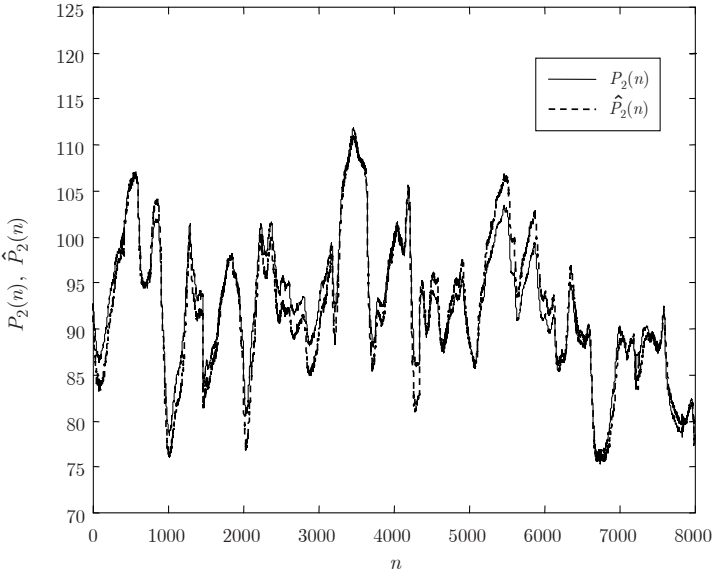

Fig. 6.12. Steam pressure in the steam chamber 2 and the output of the neural network model – the testing set

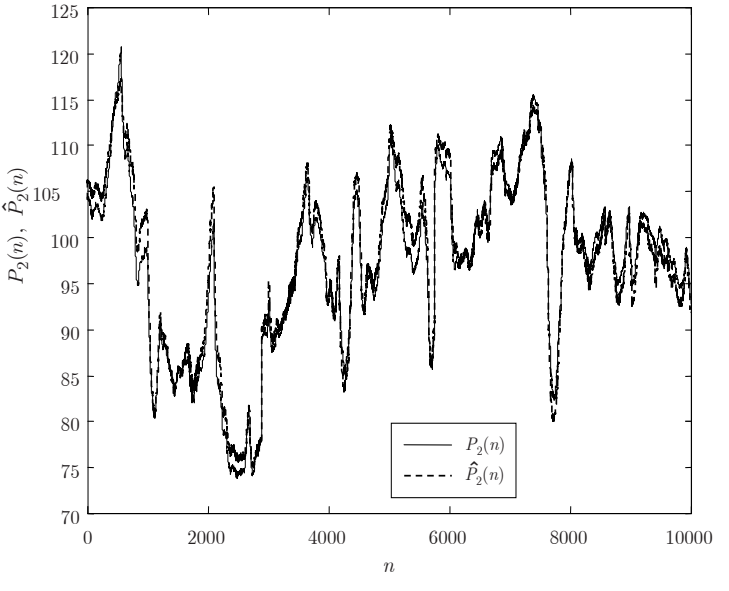

Fig. 6.13. Steam pressure in the steam chamber 2 and the output of the neural network model – the training set

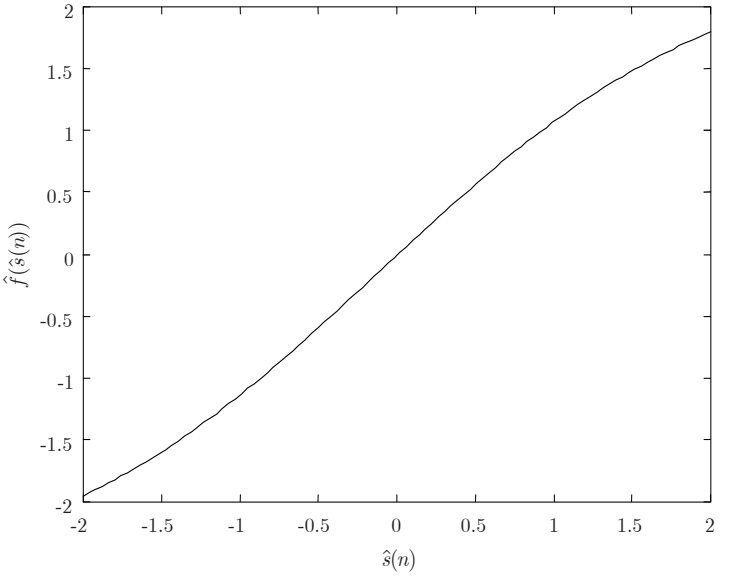

Fig. 6.14. Estimated nonlinear function

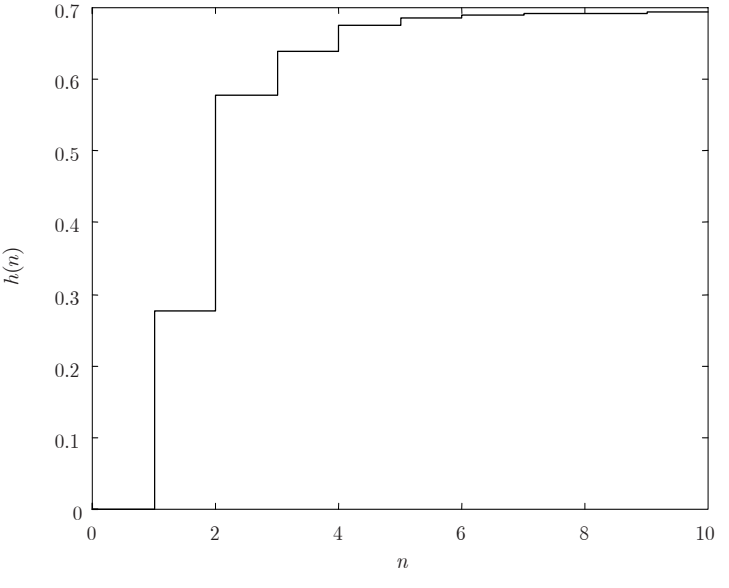

Fig. 6.15. Step response of the linear dynamic model

Table 6.5. Mean square prediction error of P_2 [kPa]

	ARX model	Neural network model
Training set	2.103	1.804
Testing set	2.322	2.084

6.4 Summary

Identification methods of system parameter changes, discussed in this chapter, require a nominal model of the system, the sequences of the system input and output signals and the sequence of residuals to be available. If the system is disturbance free, it is possible to calculate system parameter changes in both Wiener and Hammerstein systems by solving a set of linear equations. In practice, however, it is more realistic to assume that the system output is disturbed by additive disturbances and to estimate system parameter changes using parameter estimation techniques.

In Hammerstein systems with polynomial nonlinearities, system parameter changes can be estimated with the LS method. In this case, consistent parameter estimates are obtained provided that the system output is disturbed additively with (6.17). For other types of output disturbances, the use of the LS method results in inconsistent parameter estimates. Therefore, to obtain consistent parameter estimates, other parameter estimation methods, e.g., the ELS method can be used.

The estimation of system parameter changes in Wiener systems with inverse nonlinear characteristics described by polynomials with the LS method also results in inconsistent parameter estimates. To obtain consistent parameter estimates, an IV estimation procedure has been proposed.

For systems with characteristics described by the power series, the estimation of the parameters of linear dynamic systems and changes of the nonlinear characteristic for Hammerstein systems or the inverse nonlinear characteristic for Wiener systems can be made employing neural network models of the residual equation. Note that neural network models can be also useful in the case when the order of the system polynomial nonlinearity r is high. This comes from the well-known fact that the high order r introduces errors due to noise and model uncertainty, and slows down the convergence rate [9].

Nominal models of systems are identified based on input-output measurements recorded in the normal operation conditions. In spite of the fact that such measurements are often available for many real processes, the determination of nominal models at a high level of accuracy is not an easy task, as has been shown in the sugar evaporator example. The complex nonlinear nature of the sugar evaporation process, disturbances of a high intensity and correlated inputs that do not fulfill the persistent excitation condition are among the reasons for the difficulties in achieving high accuracy of system modelling.