Modelling and Identification of Power System Components

Åström, Karl Johan

Published in:
Real-time control of electric power systems

1972

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

Link to publication

Citation for published version (APA):
Åström, K. J. (1972). Modelling and Identification of Power System Components. In E. Handschin (Ed.), *Real-time control of electric power systems: Proceedings of the Symposium on real-time control of electric power systems, Brown, Boveri & Company limited, Baden, Switzerland, 1971* Elsevier.

Total number of authors:
1

Modelling and Identification of Power System Components

K. J. ÅSTRÖM

Division of Automatic Control, Lund Institute of Technology, Lund, Sweden

SUMMARY

There is a continuing tendency to apply many of the powerful results of modern control theory to various industrial processes. Power systems have been indicated as one area where significant progress can be expected. Practically all results of modern control theory require that models of the processes in terms of state equations are available. The need to obtain such models has been a strong motivation for research in the area of modelling and identification. Some progress made in this area is reviewed in this paper. Modelling based on physical equations and on plant experiments is discussed and compared. Particular emphasis is given to parameter estimation techniques like the maximum likelihood method which offer a possibility of combining physical *a priori* knowledge with experimental investigations. The formulation of identification problems is discussed, including the choice of criteria and model structures.

The techniques are illustrated by applications to data obtained from measurements on various components of a power system. The examples include an electric generator, a nuclear reactor and a drum boiler, and serve to illustrate the potentials and limitations of system identification and modelling techniques when they are applied to real data.

1. INTRODUCTION

The design of a control system is frequently divided into two steps: determination of a mathematical model and design of a control strategy. In fact most of the control theory that has been developed postulates that a model of the system and its environment is available. To use much of the existing control theory it is therefore necessary to have techniques to determine suitable models for the processes to be controlled.

Before the advent of modern control theory most results were restricted to linear systems, assuming a model specified by a transfer function. The modelling then reduces to the determination of transfer functions. This is conveniently done experimentally by introducing a sinusoidal variation in the input and measuring amplitude and phase relations between the input and the output. Interesting applications of this technique to power systems were soon made. In Sweden, for example, a group at the ASEA company under the direction of Dr. A. Garde made extensive measurements on power system components for the purpose of designing control systems. A typical study was the determination of the transfer function from power input to frequency variations in the Swedish power net reported by Oja[20].

A characteristic feature of many significant results in control theory that have been developed over the past 20 years is that they require other models than transfer functions. Typically, many of the results of modern control theory assume that the system is described by time domain models like

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \end{aligned} \tag{1}$$

where $\mathbf{u}$ is the input, $\mathbf{y}$ the output, $\mathbf{x}$ the state variable and $\mathbf{v}$ a disturbance.

To apply the results of modern control theory to industrial processes it is therefore necessary to have techniques available to determine models like eqn. (1) for the different processes. In this paper we will outline some progress towards the solution of this problem. The results, which can be applied to many industrial processes, are illustrated by computations on data obtained from experiments on power system components. Two approaches, modelling from physical *a priori* knowledge and modelling based on measurements of inputs and outputs only, are discussed in §2. Possibilities of combining the techniques are also investigated. Parameter estimation methods and formulation of identification problems are covered in §3. The choice of model structures for linear deterministic and linear stochastic systems is reviewed in §4 and §5. Selection of criteria for identification problems is the topic of §6. In §7, §8 and §9 the techniques discussed in the previous sections are applied to the modelling of power system components. The results are based on measurements on real processes. The examples include a power generator, a nuclear reactor and a drum boiler.

2. MODELLING AND IDENTIFICATION

The problem to be considered is thus how to obtain a model like the one given by eqn. (1) for industrial processes. This problem is sometimes called the inverse problem because a solution is given and the problem is to find the equation which has the given solution. Problems of this type do of course arise in many fields: biology, medicine, economy, physics and chemistry. There are certain advantages in the modelling and identification problems originating from the field of automatic control:

There is a specific purpose in doing the modelling (design of control strategies).
It is often fairly easy to do experiments. (Control systems are designed in such a

way that control variables can be manipulated and outputs measured.)

Strictly speaking, the problem may not be so well defined. Even if the design of a control system is the final goal, it is of course often very valuable to have insight and understanding of several of the system properties that do not enter the control design directly. The possibility of making experiments may be severely limited because it may be necessary to experiment under normal operating conditions, and large changes in inputs may be prohibitive for safety and economic reasons.

Process models can be obtained from basic physical laws, from pure input–output experiments or from a combination of these. The different approaches have advantages and disadvantages which are briefly discussed below.

2.1 Modelling from physical principles

The required models can in principle be derived from basic physical laws expressing conservation of mass, momentum and energy, combined with material equations like Boyle's law or Hooke's law. The models obtained in this way have the great advantage of a wide range of validity. Usually, they also provide a good insight into the behaviour of the system. The drawbacks of modelling from physical laws are that: the required knowledge is not always available; the procedure is frequently time-consuming (consider the time required to develop Newtonian mechanics); and it is often difficult to make sensible approximations. A typical difficulty is to find good approximations of distributed parameter systems. Experience has shown that the models developed from basic physical principles tend to be very complex. Since the complexity of the model indirectly implies a complex control strategy and vice versa, then if a system can be successfully controlled by a simple strategy it can probably also be modelled satisfactorily by a simple model. In the area of power systems, dynamics of components like generators, motors, transmission lines and hydroelectric stations are well understood in the sense that models can conveniently be derived from physical principles. On the other hand, components like thermal boilers and nuclear power stations are not sufficiently well known for models suitable for control to be derived from physical principles alone. In thermal boilers we have, for example, the difficulty with the two-phase flow and the uncertainty in heat transfer coefficients. Even if the basic neutron kinetics in nuclear power stations is well known it leads to very complex models if three-dimensional effects are considered. A large portion of nuclear reactor dynamics also consists of thermal dynamics and hydromechanics.

2.2 Input–output modelling

The pure case of input–output modelling consists of the determination of a model from input–output measurements only. This is often called the black box approach. The advantage of this method is that it is usually done fairly quickly. Experience has also shown that it usually leads to fairly simple models. One serious disadvantage is that in most methods it is possible to determine linear models only. This means that the validity of the model is limited. A change in operating condi-

References pp. 23–24

tions, input signals, etc., may thus lead to a different model. Another disadvantage is that available *a priori* knowledge is not used. For example, it is almost impossible to exploit *a priori* knowledge when a transfer function is determined using frequency response methods. However, in some cases the input–output approach may be the only possibility. This may be the case in the characterization of disturbances such as load variations.

Recognizing the advantages and disadvantages of modelling from physical equations and from input–output experiments alone it seems highly desirable to try to exploit both methods in order to solve the modelling problem. In the next section we will discuss techniques that can be used to do this.

3. PARAMETER ESTIMATION

Using available *a priori* knowledge about the system to be controlled, it is frequently possible to arrive at a class of models $\{M\}$ that can represent the system. The class of models can, for example, be the class of all stable linear systems having positive impulse responses, all systems described by equations like eqn. (1), where the functions **f** and **g** depend on a parameter, etc.

The problem is to design an experiment on the real system which makes it possible to select one model in the class which is a good representation of the real system. For a control system the most natural way to select a suitable model would be to compare the performances of the control strategies designed on the models when used to control the real system. Since such a selection is very difficult to do, simpler ways are often chosen. It is common to select the models by comparing the error between the model variables and the corresponding system variables. The comparison is often based on minimization of a loss function, for example of the type

$$V(y, y_m) = \int_0^\infty e^2(t)\mathrm{d}t \tag{2}$$

where y is the system output, y_m the model output and e the error. The error can for example be defined as

$$e = y - y_m \tag{3}$$

or

$$e = \mathrm{A}y - \mathrm{A}y_m \tag{4}$$

where A is some operator. Examples of this are discussed in ref. 5.

Using this formulation, the identification problem reduces to an optimization problem. Select the model in the class $\{M\}$ such that the chosen criterion is as small as possible. It is natural to ask if there is a unique minimum, if there is a natural choice of loss functions and how the results are influenced by the choice of loss functions. Some of these problems will be covered in the following sections. In particular, it will be shown that if auxiliary assumptions are made there are in fact natural loss functions and there will be unique minima.

4. LINEAR DETERMINISTIC SYSTEMS

Consider the case when the model is given by a state equation like (1) where the functions **f** and **g** depend on a set of parameters $\alpha_1, \alpha_2, \ldots, \alpha_m$ which are considered as components of a vector $\boldsymbol{\alpha}$. Also assume that the functions **f** and **g** are linear in **x** and **u** and that disturbances **v** are neglected. We thus have the standard state space description of a linear system

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \tilde{\mathbf{A}}\mathbf{x} + \tilde{\mathbf{B}}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{aligned} \tag{5}$$

where the elements of the matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\mathbf{C}$ and $\mathbf{D}$ depend on the parameter $\boldsymbol{\alpha}$. It will be shown that some of the problems raised in the previous section are non-trivial even in such simple cases. Let the experiment be arranged in such a way that the input **u** is constant over sampling intervals of constant length. The values of the state variables, the inputs and the outputs at the sampling instants are then given by

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{6}$$

where the sampling interval is chosen as the time unit and **A**, **B**, **C** and **D** are constant matrices whose elements depend on the parameter $\boldsymbol{\alpha}$. The matrices **A** and **B** are related to $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ through well-known equations.

Let n, p and r be the dimensions of **x**, **u** and **y** respectively. The descriptions (5) and (6) of the system then contain

$$N_1 = n^2 + n(r + p) + rp \tag{7}$$

parameters. It is well known that the input–output relation for eqns. (6) can be characterized by at most

$$N_2 = n(r + p) + rp \tag{8}$$

parameters. It is thus clear that in an identification experiment where the input **u** of eqns. (5) or (6) is perturbed and the output **y** is observed it is possible to determine at most N_2 parameters. Since the models (5) and (6) have N_1 coefficients, it is clear that it is not possible to determine all coefficients of eqns. (5) or (6) from an identification experiment. We can thus say that the model is not identifiable. The difficulty can be overcome by choosing a canonical structure or by using additional information about the system.

It is well known that if the coordinates in the state space eqns. (6) are changed by the linear transformation

$$\mathbf{z}(t) = \mathbf{T}\mathbf{x}(t) \tag{9}$$

References pp. 23–24

then eqns. (6) transform to

$$\begin{aligned} \mathbf{z}(t+1) &= \mathbf{TAT}^{-1}\mathbf{z}(t) + \mathbf{TBu}(t) \\ \mathbf{y}(t) &= \mathbf{CT}^{-1}\mathbf{z}(t) + \mathbf{Du}(t) \end{aligned} \tag{10}$$

which have the same input–output relation as eqns. (6). It is now natural to ask if there are transformations $\mathbf{T}$ such that the transformed eqns. (10) are characterized by fewer coefficients than eqns. (6). Many transformations with this property are known (see, for example, refs. 1, 18 and 26). One example is given below.

4.1 Canonical structures

Assume that the system (6) is completely controllable and completely observable. Let $\mathbf{A}$ be non-singular and let $\mathbf{C}$ have rank r. Then there exists a transformation $\mathbf{T}$ such that the transformed system (10) becomes

$$\mathbf{z}(t+1) = \begin{bmatrix} a_{11} & a_{12}\dots a_{1n} \\ a_{21} & a_{22}\dots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{r1} & a_{r2}\dots a_{rn} \\ & \mathbf{E} \end{bmatrix} \mathbf{z}(t) + \begin{bmatrix} b_{11} & b_{12}\dots b_{1p} \\ b_{21} & b_{22}\dots b_{2p} \\ \vdots & \vdots \quad \vdots \\ . & . \quad . \\ b_{n1} & b_{n2}\dots b_{np} \end{bmatrix} \mathbf{u}(t)$$

$$\mathbf{y}(t) = \begin{bmatrix} 1 & 0\dots 0\dots 0 \\ 0 & 1\dots 0\dots 0 \\ \vdots & \vdots \quad \vdots \quad \vdots \\ 0 & 0\dots 1\dots 0 \end{bmatrix} \mathbf{z}(t) + \mathbf{Du}(t) \tag{11}$$

where $\mathbf{E}$ is a matrix with one non-zero element in each row. (This element can be chosen as 1.) The proof is straightforward and omitted. There are also duel representations.

The representation (11) contains N_2 parameters and it is thus a minimum parameter representation. There are, however, two difficulties with this representation. The matrix $\mathbf{E}$ reflects the way in which the state variables are coupled to the output. In general, there might be

$$N_3 = \binom{n-1}{r-1} \tag{12}$$

different matrices $\mathbf{E}$. Hence if there is no *a priori* knowledge about the manner in which the outputs are coupled to the states there are (N_3) different models of the type (11). In practice, this means that in order to fit a general linear model (11) to experimental data it is necessary to determine the best values of the parameters a_{ij}, b_{ij} and d_{ij} for all N_3 possible $\mathbf{E}$ matrices. If $r = 1$ then $N_3 = 1$ and there are no difficulties, since there is only one alternative. The case $r = 1$ implies that the system has only one output. Completely observable and completely controllable linear systems with one output thus have only one structural parameter—namely the order of the system. For such systems it is therefore possible to obtain unique canonical representations. To identify linear systems with one output it is sufficient to choose a minimal parameter structure and determine the parameters of models with successively increasing order. For true multivariable systems the problem is much more difficult, since for each model of order n there are N_3 models with different internal couplings. This means that if no structural information is available *a priori* it is necessary to consider N_3 cases for each order of the model. Since N_3 is a fairly large number even for moderate values of r and n, investigation of all possible internal couplings is a significant burden.

There are many canonical structures similar to eqns. (11). They will, however, all suffer from the same difficulty that unless the internal couplings are known there are many different models.

If it is attempted to bypass the difficulties by identifying a multivariable system as an interconnection of single output systems, another difficulty is encountered. Owing to uncertainties, poles of the single output systems originating from the same mode of the multivariable system are easily estimated as being different. This means that the model obtained will be of too high an order due to false modes and that the internal couplings of the multivariable system will also be represented incorrectly.

It is thus crucial to exploit *a priori* information in the identification of a multivariable system.

There is also another difficulty with the representation (11). Due to its peculiar structure it can be shown to be very sensitive to parameter variations. Examples are given in ref. 15.

4.2 *Using* a priori *knowledge*

Having found some difficulties associated with the choice of canonical structures, we will now consider some problems associated with the use of *a priori* physical knowledge. With some knowledge about the physics of the process it may be possible to impose some conditions on the elements of the matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\mathbf{C}$ and $\mathbf{D}$ of eqns. (5) or $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ of eqns. (6). An example is given below.

Example 1
Consider a system described by eqns. (5) where

$$\tilde{\mathbf{A}} = \begin{bmatrix} \alpha_1 & 0 & \alpha_6 & \alpha_{10} & \alpha_{15} \\ \alpha_2 & 0 & \alpha_7 & \alpha_{11} & \alpha_{16} \\ \alpha_3 & 0 & \alpha_8 & \alpha_{12} & \alpha_{17} \\ \alpha_4 & 0 & 0 & \alpha_{13} & 0 \\ \alpha_5 & 0 & \alpha_9 & \alpha_{14} & \alpha_{18} \end{bmatrix}$$

$$\tilde{\mathbf{B}} = \begin{bmatrix} 0 & \alpha_{20} & \alpha_{24} \\ 0 & \alpha_{21} & \alpha_{25} \\ 0 & \alpha_{22} & \alpha_{26} \\ \alpha_{19} & 0 & 0 \\ 0 & \alpha_{23} & \alpha_{27} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

This is in fact a simplified model of a drum boiler (see ref. 9). The state variables and the inputs have the following meanings:

x_1 drum pressure
x_2 drum level
x_3 drum liquid mean temperature
x_4 riser tube mean temperature
x_5 mean value of steam-to-water ratio in drum and risers
u_1 fuel flow
u_2 feedwater flow
u_3 steam flow

The fact that certain elements of the matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\mathbf{C}$ and $\mathbf{D}$ are zero is obtained from physical considerations[9].

If it is attempted to fit a model having the structure given in Example 1 it is of course of interest to know if all the parameters $\alpha_1, \alpha_2, \ldots, \alpha_{27}$ can be determined from input–output experiments. The existence of a neat criterion to decide this is still an open problem. Algorithms which show that the parameters can be determined locally are available (see ref. 6). In this particular example we have $n = 5$, $r = 2$, $p = 3$. Also notice that the matrix $\mathbf{D}$ vanishes identically. The input–output relation can be characterized by 25 parameters. We thus find that the 27 parameters of the model cannot be determined from an input–output experiment in this example.

5. LINEAR STOCHASTIC SYSTEMS

When solving control problems, the characteristics of disturbances are frequently as important as the process dynamics. A significant contribution in modern control theory has been to model disturbances as stochastic processes and to exploit the theory of stochastic processes to obtain control strategies that take into account certain characteristics of the disturbances (see, for example, ref. 2). The following model is frequently used to represent a control system subject to random disturbances:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{v}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{e}(t) \end{aligned} \tag{13}$$

In this model $\{\mathbf{v}(t), t = 0, \pm 1, \pm 2, \ldots\}$ and $\{\mathbf{e}(t), t = 0, \pm 1, \pm 2, \ldots\}$ are sequences of independent equally distributed random vectors with zero mean values and given covariance matrices

$$\begin{aligned} \text{cov}[\mathbf{v}(t), \mathbf{v}(t)] &= \mathbf{R}_1 \\ \text{cov}[\mathbf{e}(t), \mathbf{e}(t)] &= \mathbf{R}_2 \\ \text{cov}[\mathbf{v}(t), \mathbf{e}(t)] &= \mathbf{R}_{12} \end{aligned} \tag{14}$$

The model (13) is fairly general. It can, for example, be used to represent finite dimensional linear systems whose disturbances are weakly stationary random processes with rational spectral densities.

The problem of determining the parameters of the model (13) will now be discussed. If all the elements of the matrices **A**, **B**, **C**, **D**, $\mathbf{R}_1$, $\mathbf{R}_{12}$ and $\mathbf{R}_2$ are considered as parameters the model contains

$$N_4 = n(1.5n + p + 2r + 0.5) + r(p + 0.5r + 0.5) \tag{15}$$

parameters. Apart from the difficulties associated with the determination of the parameters of the deterministic part of the system, *i.e.* **A**, **B**, **C** and **D**, it can be shown that all elements of the matrices $\mathbf{R}_1$, $\mathbf{R}_{12}$ and $\mathbf{R}_2$ cannot be determined from the data of an input–output experiment.

Assume, for example, that the matrix $\mathbf{R}_2$ is positive definite. Then the model given by eqns. (13) is equivalent to the model

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}\boldsymbol{\varepsilon}(t) \\ \mathbf{y}(t) &= \mathbf{C}\hat{\mathbf{x}}(t) + \mathbf{D}\mathbf{u}(t) + \boldsymbol{\varepsilon}(t) \end{aligned} \tag{16}$$

where $\{\boldsymbol{\varepsilon}(t), t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of independent equally distributed random vectors with zero mean values and covariances **R**. The models are equivalent in the sense that the input–output relations are the same and that the stochastic properties of the outputs are also the same[2].

The proof of the statement follows from the Kalman–Bucy filtering theorem. The assumption that $\mathbf{R}_2$ is positive definite can be relaxed. The model (16) is not unique. There are in general many matrices **K** for which the stochastic properties are the same. The representation is unique, however, if it is required that the matrix $\mathbf{A} - \mathbf{KC}$ has all eigenvalues inside the unit circle.

The model (16) contains

$$N_5 = n(n + 2r + p) + r(0.5r + p + 0.5) \tag{17}$$

parameters. It has several interesting properties. The state variable $\hat{\mathbf{x}}$ of eqns. (16) can be interpreted as the best linear estimate of the state **x** of eqns. (13) based on observed outputs. The quantities $\{\boldsymbol{\varepsilon}(t)\}$ are the innovations associated with the

References pp. 23–24

stochastic process $\{\mathbf{y}(t), t = 0, \pm 1, \pm 2, \ldots\}$ of eqns. (13) or (16). The matrix $\mathbf{K}$ can be interpreted as the steady state gain of the Kalman filter associated with the model (13). The Kalman filter for estimating the state of eqns. (13) is thus given by

$$\hat{\mathbf{x}}(t+1) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}[\mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t) - \mathbf{D}\mathbf{u}(t)] \tag{18}$$

This equation is trivially obtained by eliminating $\boldsymbol{\varepsilon}$ in eqns. (16).

Hence if the purpose of the identification is to design a predictor or a regulator based on linear stochastic control theory the model (16) has several advantages over the model given by eqns. (13). If the model (13) is determined it is necessary to solve a Riccati equation in order to obtain the gain of the Kalman filter. It also turns out that the algorithms for the identification problem are simpler for the model structure (16). Notice, however, that even with eqns. (16) there are the difficulties with ambiguities discussed previously for deterministic systems.

In the special case of systems with one output the redundancy in the general model can be reduced by first transforming the state variables so that $\mathbf{TAT}^{-1}$ becomes a matrix on companion form and then using the transformation given by eqns. (16). We will then obtain the following model:

$$\mathbf{z}(t+1) = \begin{bmatrix} -a_1 & 1 & 0\ldots 0 \\ -a_2 & 0 & 1\ldots 0 \\ \vdots & \vdots & \vdots \quad \vdots \\ -a_{n-1} & 0 & 0\ldots 1 \\ -a_n & 0 & 0\ldots 0 \end{bmatrix} \mathbf{z}(t) + \begin{bmatrix} b_{11} & \ldots b_{1p} \\ b_{21} & \ldots b_{2p} \\ \vdots & \vdots \\ b_{n-1,1} & \ldots b_{n-1,p} \\ b_{n1} & \ldots b_{np} \end{bmatrix} \mathbf{u}(t) + \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix} \varepsilon(t)$$

$$y(t) = [1 \quad 0 \quad 0\ldots 0]\mathbf{z}(t) + \varepsilon(t) \tag{19}$$

By eliminating the state variable $\mathbf{z}$ the following input–output relation is obtained:

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = b_{11}u_1(t-1) + \cdots + b_{n1}u_1(t-n) + \cdots$$
$$+ b_{1p}u_p(t-1) + \cdots + b_{np}u_p(t-n) + \varepsilon(t) + c_1\varepsilon(t-1) + \cdots + c_n\varepsilon(t-n) \tag{20}$$

where

$$c_i = a_i + k_i \quad \text{for} \quad i = 1, 2, \ldots, n \tag{21}$$

The model (20) is thus a canonical representation of a linear stochastic system with one output and several inputs.

The model (20) can be written in a slightly more compact form if the polynomials

$$A(z) = z^n + a_1 z^{n-1} + \cdots + a_n$$
$$B(z) = [b_{11}\ldots b_{1p}]z^{n-1} + [b_{21}\ldots b_{2p}]z^{n-2} + [b_{n1}\ldots b_{np}] \tag{22}$$
$$C(z) = z^n + c_1 z^{n-1} + \cdots + c_n$$

and the shift operator q defined by

$$\mathrm{q}x(t) = x(t+1) \tag{23}$$

are introduced. The model (20) then becomes

$$A(\mathrm{q})y(t) = B(\mathrm{q})\mathbf{u}(t) + C(\mathrm{q})\varepsilon(t) \tag{24}$$

6. CRITERIA

Having chosen a model whose parameters can be determined from input–output data it remains to find the parameters of the model such that the model fits the experimental data. The crucial problem is then to find suitable criteria. Having the control application in mind, it would be natural to evaluate the model on the basis of the performance of the control strategies designed from it. This is generally very difficult to do, and it is thus necessary to use other criteria. The criteria can be chosen ad hoc as, for example, in eqn. (2).

If statistical assumptions are made, the parameters can also be determined using statistical parameter estimation techniques. This leads frequently to an optimization problem with a given criterion. It is thus often possible to give statistical interpretations to many criteria.

6.1 Maximum likelihood estimates

Assuming that the process is governed by the model (16) where $\{\boldsymbol{\varepsilon}(t)\}$ is a sequence of independent gaussian random variables, the parameters can be determined using the method of maximum likelihood. It can be shown that the likelihood function for estimating the model parameters is given by

$$-\ln L = \frac{1}{2}\sum_{t=1}^{N} \boldsymbol{\varepsilon}^T(t)\mathbf{R}^{-1}\boldsymbol{\varepsilon}(t) + \frac{N}{2}\ln\det\mathbf{R} + \text{const} \tag{25}$$

where N is the record length and $\mathbf{R}$ the covariance of $\boldsymbol{\varepsilon}(t)$. The identification problem then reduces to the problem of minimizing the function (25) with respect to the unknown parameters. The matrix $\mathbf{R}$ is frequently not known. This means that it is necessary to minimize eqn. (25) with respect to $\mathbf{R}$ also. Notice that this can be done analytically. We have

$$\mathrm{Min}_{\mathbf{R}}\left[\frac{1}{2}\sum_{t=1}^{N} \boldsymbol{\varepsilon}^T(t)\mathbf{R}^{-1}\boldsymbol{\varepsilon}(t) + \frac{N}{2}\ln\det\mathbf{R}\right] = \frac{rN}{2} + \frac{N}{2}\ln\det\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^T(t) \tag{26}$$

The minimum is assumed for

$$\mathbf{R} = \mathbf{R}_0 = \frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^T(t) \tag{27}$$

References pp. 23–24

This follows from the identity

$$\sum_{i=1}^{N} \mathbf{x}_i^T \left(\sum_{j=1}^{N} \mathbf{x}_j \mathbf{x}_j^T \right)^{-1} \mathbf{x}_i = n \tag{28}$$

where n is the dimension of the vectors $\mathbf{x}_i$.

The maximum likelihood identification thus reduces to the minimization of the loss function

$$V = \det \sum_{t=1}^{N} \boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^T(t) \tag{29}$$

and conversely an identification problem with the criterion (29) can be interpreted as a parameter estimation problem. An estimate of the covariance $\mathbf{R}$ is then given by eqn. (27).

In the special case of systems with one output, ε is a scalar and the loss function reduces to

$$V = \sum_{t=1}^{N} \varepsilon^2(t) \tag{30}$$

Under the restrictive assumption that the data were actually generated by a model (16) where $\{\varepsilon(t)\}$ is a sequence of independent gaussian random variables it is possible to pose and answer several statistical problems. With minor additional assumptions it can be shown that:

(a) the estimates converge to the true parameters with probability one as the record length N increases (consistency);

(b) for large N there are no other estimation procedures that give estimates with smaller variances (asymptotic efficiency);

(c) an estimate of the covariance of the estimate $\hat{\boldsymbol{\alpha}}$ is given by

$$\operatorname{cov}[\boldsymbol{\alpha}, \boldsymbol{\alpha}] = \frac{2V}{N} V_{\alpha\alpha}^{-1} \tag{31}$$

(asymptotic normality).

The precise statements of these results are given in ref. 3 for the single output case and in refs. 7 and 27 for the multivariable case.

Many other statistical problems can also be posed. For example, the determination of the order of a model can be approached as a statistical problem. Let V_n denote the loss function obtained for a model with n parameters. The function V_n decreases with increasing n. The problem is to decide if the decrease in V is significant or not. For systems with one output the test quantity

$$F_{n_1,n_2} = \frac{V_{n_1} - V_{n_2}}{V_{n_2}} \frac{N - n_2}{n_2 - n_1} \quad \text{for} \quad n_2 > n_1 \tag{32}$$

where N is the number of observations of the output, can be shown to be asymptotically F-distributed. The quantities F_{n_1,n_2} can thus be used to test whether the loss function is reduced significantly when the number of parameters in the model is increased from n_1 to n_2.

Notice that one can test whether the residuals are gaussian and uncorrelated. It is, however, virtually impossible to assert that the data were actually generated by a model (16) with specific parameters. In practice, this is of course never true, since the model (16) is only an approximation of a complex process. Thus, the results obtained by using the statistical theories must be handled with great care when applied to real data. Notice that when the methods are tried on simulated data it is always possible to assert that the assumptions required by the statistical theory are fulfilled!

6.2 *Other interpretations*

The algorithm obtained by minimizing the criterion (29) can be given a physical interpretation, even in the case when no statistical assumptions are made. Equation (16) can be rewritten as

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}[\mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t) - \mathbf{D}\mathbf{u}(t)] \\ \boldsymbol{\varepsilon}(t) &= \mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t) - \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{33}$$

The quantity $\mathbf{C}\hat{\mathbf{x}}(t) + \mathbf{D}\mathbf{u}(t)$ can be interpreted as a one-step prediction of $\mathbf{y}(t)$ based on $\mathbf{y}(t-1), \mathbf{y}(t-2), \ldots$. The quantity $\boldsymbol{\varepsilon}(t)$ can thus be interpreted as the error in the prediction of the output. To minimize the loss function (29) thus means that the parameters of the model are changed in such a way that the error in predicting the output one step ahead is as small as possible according to the criterion (29). In the case of single output systems the criterion (30) is simply the mean squares prediction error.

7. POWER GENERATOR DYNAMICS

In this section, the techniques described in the previous sections will be applied to power generator modelling. The results of this section are based on ref. 17.

The analysis is based on experiments made by Dr. Stanton[24, 25] on a 50 MW turboalternator. The experiments consisted of recording the variations in the terminal voltage V, the active and reactive components of the armature current, I_r and I_q respectively, and the angular velocity ω during normal operation. The instrumentation used is described in ref. 23. This report also contains a description of the experimental procedure and the difficulties with this type of experiment. Several experiments were performed by Stanton. In this example, we will base computations on an experiment performed with the governor blocked, since we are interested in open loop dynamics. The record length was 600 sec and the sampling interval was 0.5 sec for the angular velocity and 0.125 sec for the other variables.

In this example, the dynamics relating angular velocity to electric torque $M = VI_r/\omega$ will be considered. A plot of these two variables is shown in the upper part of Fig. 1. The generator is frequently described by the simplified model

References pp. 23–24

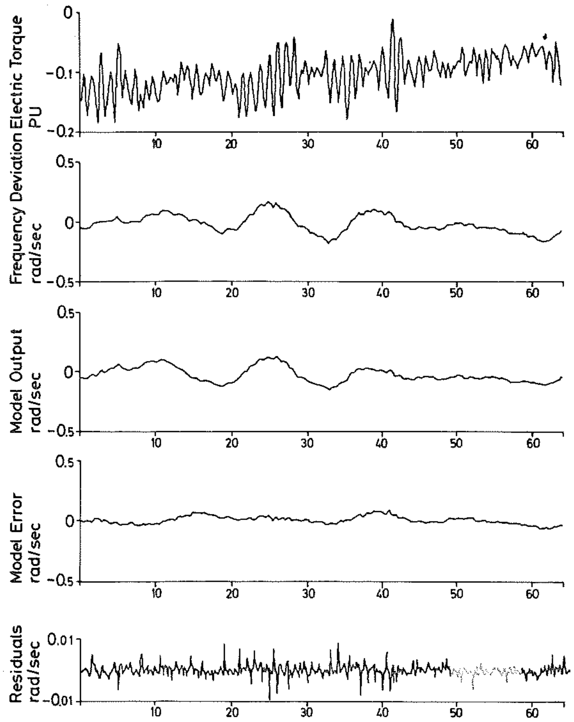


Fig. 1. The results of identification of power generator dynamics. The uppermost curve shows the measured input (electric torque $\delta M = I_r V/\omega$ in per unit and measured output (angular velocity ω in rad/sec). The model output computed from the fifth order model with coefficients given in Table 1 and disturbances neglected, the model error and the residuals $\{\varepsilon(t)\}$ are also shown in the diagram.

$$J\frac{\mathrm{d}\omega}{\mathrm{d}t} + D\omega = M_{\mathrm{m}} - \frac{VI_{\mathrm{r}}}{\omega} \tag{34}$$

where ω is the angular velocity of the rotor, J the moment of inertia of the rotor, D the damping coefficient, V the terminal voltage, I_r the active component of the armature current and M_m the mechanical torque. There are also more elaborate descriptions leading to models of higher order.

Since it is not known *a priori* that a model like eqn. (34) is compatible with the data, one could first attempt to fit general linear models with the canonical structure (20) and different orders n. This will give an indication of the complexity required to describe the data. The loss functions obtained from a maximum likelihood identification of models of different order are shown in Table 1. To save computer time, the computations are based on 1000 input–output pairs only. Since the sampling interval of the angular velocity is 0.5 sec and the sampling interval of the electric power is 0.125 sec, a compatible data set is created by interpolating the angular velocity. The initial state of the model (20) is also determined.

In Table 1 the test quantities (32) for testing the order of the system are also

TABLE 1

MINIMAL VALUES OF THE LOSS FUNCTION FOR MAXIMUM LIKELIHOOD IDENTIFICATION OF MODELS OF DIFFERENT ORDER RELATING ANGULAR VELOCITY ω TO ELECTRIC TORQUE $I_r V/\omega$ FOR THE ELECTRIC GENERATOR
The computations are based on 1000 data points

n	V_n	$F_{n,n+1}$
1	0.01442	126
2	0.009562	26
3	0.008560	73
4	0.006656	17
5	0.006230	2.0
6	0.006180	

TABLE 2

COEFFICIENTS OF FIFTH ORDER MODEL RELATING ANGULAR VELOCITY ω TO ELECTRIC POWER FOR ELECTRIC GENERATOR

N	1000	2400
a_1	-2.628 ± 0.031	-2.581 ± 0.024
a_2	2.290 ± 0.074	2.204 ± 0.049
a_3	-0.715 ± 0.082	-0.640 ± 0.048
a_4	0.078 ± 0.067	0.024 ± 0.044
a_5	-0.025 ± 0.026	-0.004 ± 0.019
b_1	-0.835 ± 0.092	-0.836 ± 0.061
b_2	1.120 ± 0.244	1.016 ± 0.016
b_3	-0.805 ± 0.308	-0.529 ± 0.197
b_4	0.599 ± 0.250	0.198 ± 0.162
b_5	-0.080 ± 0.100	0.150 ± 0.075
c_1	-0.899 ± 0.036	-0.854 ± 0.027
c_2	-0.005 ± 0.039	-0.011 ± 0.024
c_3	-0.013 ± 0.036	-0.027 ± 0.024
c_4	-0.624 ± 0.034	-0.617 ± 0.021
c_5	0.579 ± 0.029	0.541 ± 0.021
λ	0.0025	0.0025

evaluated. Assuming that all assumptions required for the order test are fulfilled, a straightforward application of the order test would thus indicate that a fifth order model is appropriate. The coefficients of the fifth order model are shown in Table 2. The accuracy estimates are obtained from eqn. (31). The results obtained by performing the identification on a longer data set $N = 2400$ are also shown in Table 2. A comparison of the results obtained for $N = 1000$ and $N = 2400$ shows that it is not unreasonable to assume that the system is time-invariant. A comparison of the accuracy estimates also indicates that they decrease as $1/\sqrt{N}$, as can be expected from the theory. Notice in Table 2 that there is a significant difference between the relative accuracies of the parameters b_i and the relative accuracies of a_i and c_i. The results are quite typical of those obtained in other cases when no perturbations in the input are used.

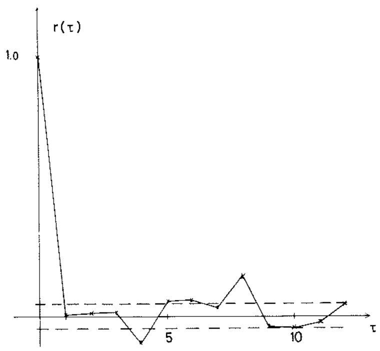
References pp. 23–24

The results of the identification are illustrated in Fig. 1, which shows the measured inputs and outputs, the model output and the difference between the measured output and the model output. The model output is computed from the model (20) with the coefficients given in the first column of Table 2 and disturbances $\{e(t)\}$ neglected. The model output thus explains how much of the actual output that can be explained from the input. Figure 1 reveals that only about half of the observed output can be related to the input. The signal-to-noise ratio is thus fairly low. The situation is quite typical for data obtained during normal operation with no extra perturbations introduced. Also notice from Table 2 that with a record of length 125 sec it is possible to get reasonable parameter estimates of the parameters a_i and c_i.

The residuals can be interpreted as the one-step prediction errors obtained from a predictor determined from the model (20). The residuals thus show how well the output can be predicted one step ahead. The standard deviation of the residuals is $\lambda = 0.0025$, which means that it is possible to predict the angular velocity one sampling interval ahead with a standard deviation of 0.0025 rad/sec using the model obtained.

If the identification procedure is to have a nice statistical interpretation the residuals should be independent stochastic variables. In Fig. 2 the covariance of the residuals is shown. This diagram indicates that the residuals at least are uncorrelated. The residuals are, however, not normally distributed, as is seen by the diagram of cumulative frequencies in Fig. 3. This means, for example, that the results of the order test can be questioned in a case like this.

The results obtained clearly indicate that the simple first order model (34) is not compatible with the data.



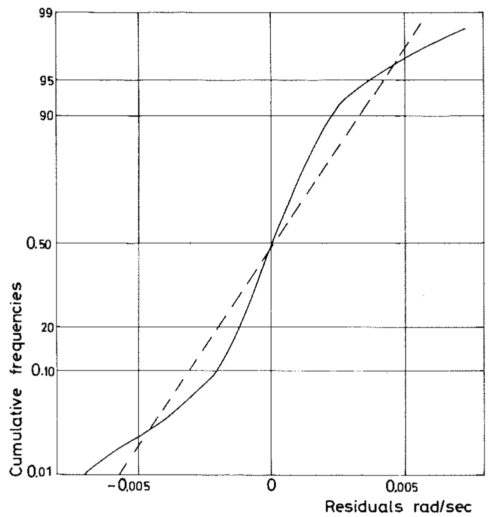


Fig. 2. Sample covariance function for the residuals of the model (20) with parameters according to the first column of Table 2.

Fig. 3. Cumulative frequencies of residuals of the model (20). The scales of the diagram are such that a normal distribution corresponds to a straight line.

8. NUCLEAR REACTOR DYNAMICS

The experiments were done by AB Atomenergi on the ÅGESTA reactor, located in a suburb of Stockholm. A schematic picture of the reactor is shown in Fig. 4. The results of this section are based on ref. 12.

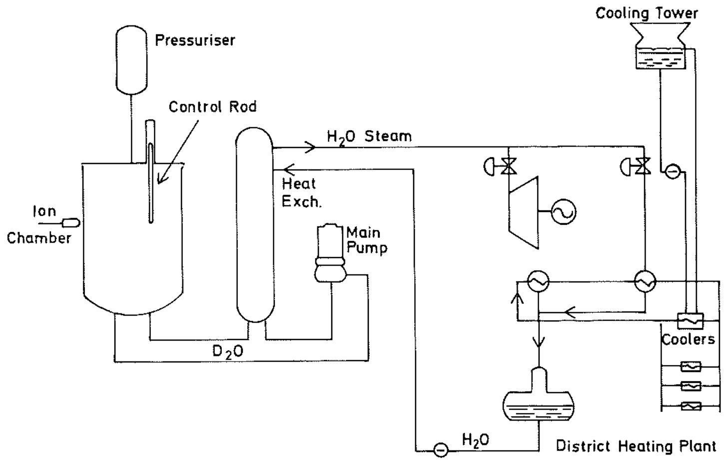


Fig. 4. Schematic diagram of the Ågesta nuclear reactor.

In the experiments the control rod position was perturbed and the nuclear power was measured. The input was chosen as a PRBS* sequence with period 127. The shortest pulse length was 20 sec and the sampling interval was 5 sec. The input–output signals from the experiment are shown in Fig. 5. As an initial attempt, a maximum likelihood identification is carried out using models (20) of different orders. The values of the loss functions obtained are shown in Table 3.

A straightforward application of the order test indicates that the model has to be of at least fourth order. At a risk level of 5% the test limit is 3, and the decrease in the loss function obtained when going from a fourth order system to a fifth order system is therefore not significant. The coefficients of the fourth order model are shown in Table 4.

The results of the identification are illustrated in Fig. 5. It is clear from this diagram that a major part of the observed output is caused by the input. The magnitude of the model error never exceeds 0.8 MW. The residuals have a standard deviation of 0.113 MW, which means that the output can be predicted 5 sec ahead with a standard deviation of 0.113 MW. The large model error and the large residual at time 20.2 min were traced to a malfunction of the control rod servo.

By comparing the results given in Table 4 with those in Table 2 we find that the relative accuracies of the parameters associated with the nuclear reactor dynamics are significantly higher than those associated with the dynamics of the electric

* PRBS = pseudo random binary sequence.

References pp. 23–24

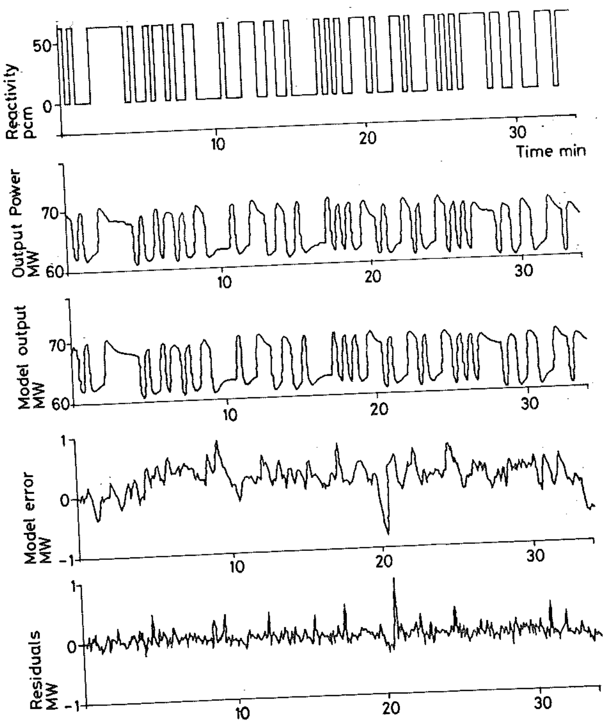


Fig. 5. The identification of nuclear reactor dynamics. The figure shows the measured input, reactivity in pcm (per cent mille = 10^{-5}) (control rod position), the measured output in MW (nuclear power), the model output, the model error and the residuals.

TABLE 3

MINIMAL VALUES OF THE LOSS FUNCTION FOR MAXIMUM LIKELIHOOD IDENTIFICATION OF MODELS OF DIFFERENT ORDER RELATING NUCLEAR POWER TO CONTROL ROD POSITION

The computations are based on 1125 input–output pairs

n	V_n	$F_{n,n+1}$
1	203.344	2502
2	26.378	218
3	16.642	61.2
4	14.284	1.2
5	14.238	0.2
6	14.229	

TABLE 4

PARAMETERS OF A FOURTH ORDER MODEL RELATING NUCLEAR POWER TO CONTROL ROD POSITION

The estimates are based on 1125 input–output pairs with a sampling interval of 5 sec

a_1	-2.442 ± 0.023
a_2	2.096 ± 0.047
a_3	-0.679 ± 0.029
a_4	0.032 ± 0.004
b_1	0.1013 ± 0.0002
b_2	-0.2377 ± 0.0024
b_3	0.1884 ± 0.0045
b_4	-0.0517 ± 0.0023
c_1	-1.479 ± 0.038
c_2	0.502 ± 0.061
c_3	0.122 ± 0.055
c_4	-0.072 ± 0.030
λ	0.113

generator. Compare in particular the coefficients b_i. This is due to the fact that perturbation signals were used in the reactor experiment, while no perturbations were introduced in the generator experiment. Similar observations have been made in many other situations.

9. DRUM BOILER DYNAMICS

In this section, applications to drum boiler modelling are discussed. The results are based on refs. 10 and 4. The experiments were performed on the P16–G16 power plant of the Öresundsverket of Sydkraft AB in Malmö, Sweden. The unit consists of a Steinmüller drum boiler and a Stal–Laval turbine. It has a power of 160 MW. A schematic picture of the boiler is shown in Fig. 6.

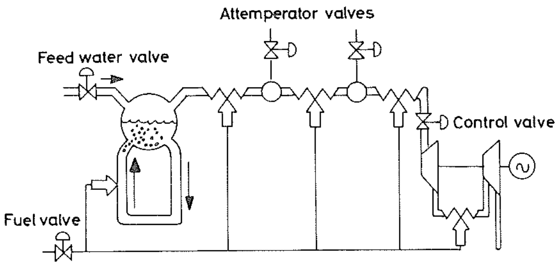


Fig. 6. Schematic diagram of the boiler–turbine unit.

In the experiment, the input variables fuel flow, feedwater flow and control valve position were perturbed and the relevant process variables were recorded. Since the open loop dynamics were of major interest all major regulators except the fuel–air regulator were removed. For safety reasons only one process variable at a time was changed. The determination of linear multivariable models was based on superposition of the results of several experiments.

Many experiments were performed, both for the purpose of determining linear steady state models and for the purpose of determining non-linear models. A detailed discussion of the measurements including identification and modelling is given by Eklund[10].

In this example, the modelling of the drum boiler only will be discussed. The inputs are taken as fuel flow, feedwater flow and steam flow. The output variables are chosen as drum level and drum pressure. A set of experimental data is shown in Fig. 7. The sampling interval was chosen as 10 sec.

9.1 *Linear models*

Maximum likelihood identification of single output models for drum pressure and drum levels leads to models of orders 3 and 2. It is thus not unreasonable that a model like the one given in Example 1, §4, is compatible with the data. The model given in Example 1 was obtained from physical considerations. Since it contains

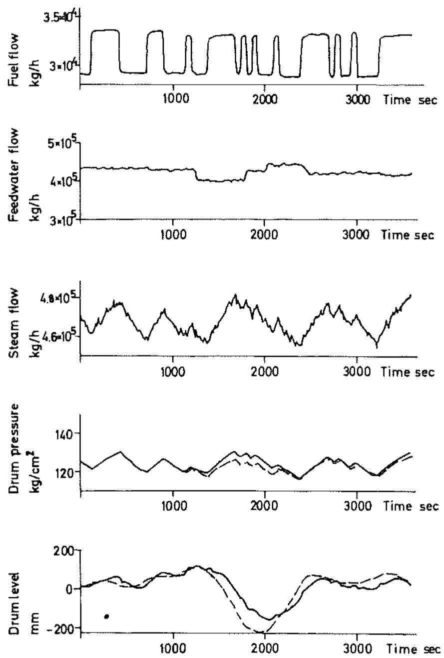


Fig. 7. Results of linear modelling of a drum boiler. The full lines are the measured data and the broken lines are the model outputs.

27 parameters, while at most 25 parameters can be determined from input–output data, some parameters must be fixed. There are disturbances in the data which are not white measurement noise. It is then necessary to choose the model structure (16) which allows for a more general noise model. Thus the 10 elements of the **K**-matrix must also be included as parameters. Since a model (16) was determined from physical considerations initial estimates of all parameters except those of the **K**-matrix were available. Several attempts were made to identify different parameter sets. A typical result obtained is listed in Table 5. Notice the significant reduction in the value of the loss function. The model obtained with the parameters of Table 5 is compared with the measurements in Fig. 7.

The estimated covariance matrix is

$$\hat{\mathbf{R}} = \begin{bmatrix} 0.0240 & 0.016 \\ 0.016 & 12.4 \end{bmatrix}$$

This means that the drum pressure can be predicted 10 sec ahead with a standard deviation of 0.16 bar and the drum level with a standard deviation of 3.5 mm.

9.2 *Non-linear models*

For power system analysis it is highly desirable to have fairly simple models available which give the gross behaviour of the different system components.

TABLE 5

INITIAL GUESSES AND PARAMETER ESTIMATES OBTAINED WHEN FITTING A MODEL WITH THE STRUCTURE OF EXAMPLE 1 TO DRUM BOILER MEASUREMENTS

	Initial	*Identified*
V	0.2567×10^{17}	0.3850×10^{5}
α_1	-0.02835	-0.02052
α_3	0.01382	0.00896
α_5	-0.1341×10^{-3}	-0.0738×10^{-3}
α_{10}	0.0406	0.1678
α_{13}	-0.0454	-0.3295
α_{16}	-0.2266	-0.3622
α_{19}	0.1162×10^{-4}	0.1446×10^{-4}
α_{24}	-0.01575	-0.00496
k_{11}	0	1.271
k_{12}	0	-0.1041×10^{-2}
k_{21}	0	0.01027
k_{22}	0	0.1279×10^{-2}
k_{31}	0	-0.1214
k_{32}	0	-0.06576
k_{41}	0	0.5432
k_{42}	0	0.01398
k_{51}	0	-0.7702×10^{-3}
k_{52}	0	0.3585×10^{-4}

From studies of linear models like the one just discussed a great deal of insight into the behaviour of boiler–turbine units was obtained. It was found that the gross behaviour of a boiler–turbine unit could be described by the equations

$$\begin{aligned} \frac{\mathrm{d}p}{\mathrm{d}t} &= \alpha[-f(p, u_2) + g(u_1, u_3)] \\ P &= f(p, u_2) \end{aligned} \tag{35}$$

where p is the drum pressure, P the output power, u_1 the fuel flow, u_2 the control valve setting and u_3 the feedwater flow. A model of the structure (35) can also be derived from an energy balance if several (crude) approximations are done. Such a derivation will also give the structure of the functions f and g. Details of this are given in ref. 4. It is shown that a possible choice is

$$f(p, u_2) = \alpha_4[u_2 p^{5/8} - \alpha_5] \tag{36}$$

$$\alpha g(u_1, u_3) = \alpha_2 u_1 - \alpha_3 u_3 \tag{37}$$

$$\alpha = \alpha_1/\alpha_4 \tag{38}$$

The boiler–turbine unit can thus be represented by a non-linear model with five parameters $\alpha_1, \alpha_2, \ldots, \alpha_5$. Determining these parameters from several experiments covering a wide operating range the following values are obtained for the unit P16–G16.

References pp. 23–24

$\alpha_1 = 0.00305$
$\alpha_2 = 0.02$
$\alpha_3 = 4.4 \times 10^{-2}$
$\alpha_4 = 11.45$
$\alpha_5 = 8.2$

The performance of the model at two different operating conditions is illustrated in Figs. 8 and 9. Notice in particular that the model agrees well with experiments over a wide operating range.

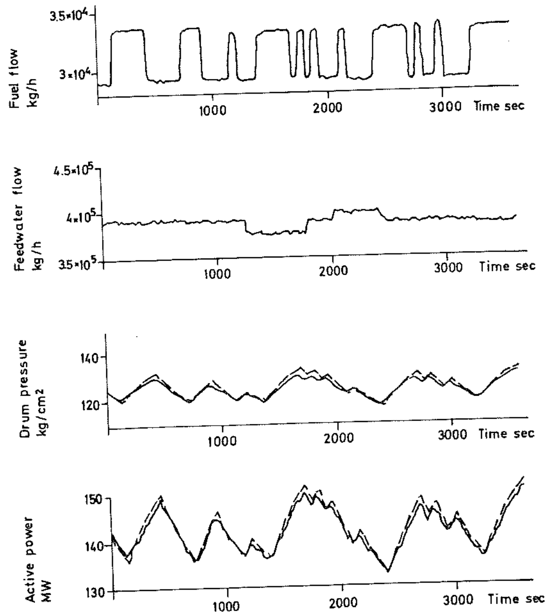


Fig. 8. Comparison of measured boiler–turbine data (full lines) and the responses of the non-linear model (35) (broken lines).

ACKNOWLEDGEMENTS

It is my pleasure to acknowledge my gratitude to several persons and institutions. My research in the field of system identification has been partially supported by the Swedish Board for Technical Development (contract 71-50/U33). Data from the power generator measurement were kindly provided by Dr. K. N. Stanton. The results of experiments on the nuclear reactor were provided by AB Atomenergi, Stockholm, Sweden. The experiments on the thermal power plant were done by Dr. K. Eklund as part of his thesis in collaboration with Sydkraft AB, Malmö, Sweden. The computations reported have been performed by my collaborators at the Lund University, Dr. K. Eklund and Mr. S. Lindahl. Special thanks are due to I. Gustavsson with whom I have enjoyed extensive discussions on applications of system identification over many years.

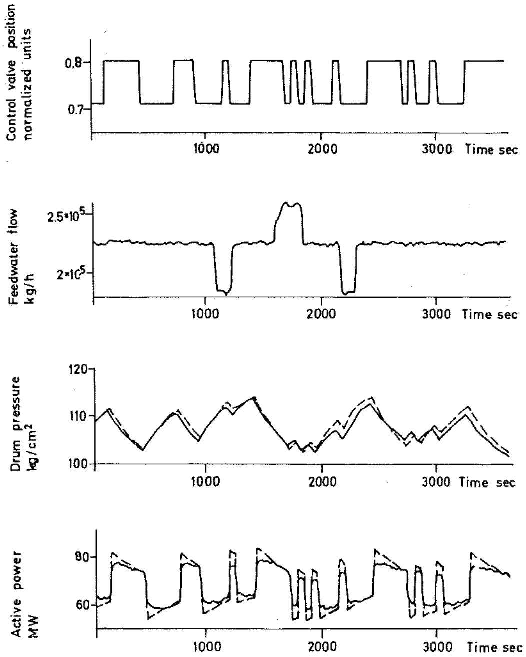


Fig. 9. Comparison of measured boiler–turbine data (full lines) and the responses of the non-linear model (35) (broken lines).

NOTES AND REFERENCES

System modelling and identification is discussed in many recent articles. Many references are given in the recent survey articles[5,19] and in the book *System Identification*[22]. There have also been two IFAC symposia in Prague in 1967 and 1970 entirely devoted to this field. The use of maximum likelihood techniques is discussed in refs. 3, 7, 8, 11, 13, 14, 16 and 27. Canonical structures for linear multivariable systems are treated in refs. 1, 18 and 26. A more detailed discussion of the nuclear reactor identification is given in ref. 13. The same techniques have also been applied to the modelling of the Halden reactor in Norway (see ref. 21). The section on boiler modelling is based on refs. 4 and 10 which include detailed analyses of many measurements.

REFERENCES

1 I. Ackermann, Die minimale Ein-Ausgangs-Beschreibung von Mehrgrössen-Systemen und ihre Bestimmung aus Ein-Ausgangs-Messungen, *Regelungstechnik und Prozessdatenverarbeitung*, **5** (1971) 203–206.
2 K. J. Åström, *Introduction to Stochastic Control Theory*, Academic Press, New York, 1970.
3 K. J. Åström and T. Bohlin, Numerical identification of linear dynamic systems from normal operating records, paper at the IFAC Symposium on Theory of Self-Adaptive Systems, Teddington, England, in P. H. Hammond (Ed.), *Theory of Self-Adaptive Control Systems*, Plenum Press, New York, 1966.

4 K. J. ÅSTRÖM AND K. EKLUND, A simplified nonlinear model of a drum-boiler turbine unit, *Report 7104*, Division of Automatic Control, Lund Institute of Technology, April 1971, to appear in *Intl. J. Control.*

5 K. J. ÅSTRÖM AND P. EYKHOFF, System identification—a survey, *Automatica*, **7** (1971) 123–162.

6 R. BELLMAN AND K. J. ÅSTRÖM, On structural identifiability, *Mathematical Biosciences*, **7** (1970) 329–339.

7 P. E. CAINES, The parameter estimation of state variable models of multivariable linear systems, *Control Systems Centre Report No. 146*, The University of Manchester Institute of Science and Technology, April 1971.

8 J. EATON, Identification for control purposes, *IEEE Winter Meeting, New York, 1967.*

9 K. EKLUND, Linear mathematical models of the drum–downcomer–riser loop of a drumboiler, *Report 6809*, Division of Automatic Control, Lund Institute of Technology, November 1968.

10 K. EKLUND, Linear drum boiler–turbine models, *Report 7117*, Division of Automatic Control, Lund Institute of Technology, November 1971.

11 C. A. GALTIERI, The problem of regulation in the computer control of industrial processes, Part I: Theory, *IBM Res. Rept RJ382*, San Jose, California, 1966.

12 I. GUSTAVSSON, Maximum likelihood identification of dynamics of the Ågesta reactor and comparison with results of spectral analysis, *Report 6903*, Division of Automatic Control, Lund Institute of Technology, February 1969.

13 I. GUSTAVSSON, Parametric identification of multiple input single output dynamic systems, *Report 6907*, Division of Automatic Control, Lund Institute of Technology, July 1969.

14 I. GUSTAVSSON, Comparison of different methods for identification of linear models for industrial processes, *2nd IFAC Symposium on Identification and Process Parameter Estimation, Prague, 1970*, Paper 11.4.

15 P. HAGANDER, C. KÄLLSTRÖM AND K. J. ÅSTRÖM, Real time computing: Implementing linear filters, *Report 7106B*, Division of Automatic Control, Lund Institute of Technology, May 1971.

16 R. L. KASHYAP, Maximum likelihood identification of stochastic linear systems, *IEEE Trans. Autom. Control*, **AC-15** (1970) 25.

17 S. LINDAHL, Identification of power generator dynamics from normal operating data, Term paper in course on System Identification Techniques, Division of Automatic Control, Lund Institute of Technology, 1971.

18 D. G. LUENBERGER, Canonical forms for linear multivariable systems, *IEEE Trans.*, **AC-12** (1967) 290–293.

19 R. E. NIEMAN, D. G. FISHER AND D. E. SEBORG, A review of process identification and parameter estimation techniques, *Intl. J. Control*, **13** (1971) 209.

20 V. OJA, Frequency-response method applied to the study of turbine regulation in the Swedish power system, in R. OLDENBURGER (Ed.), *Frequency Response*, Macmillan, New York, 1955.

21 G. OLSSON, Maximum likelihood identification of the Halden boiling water reactor, unpublished.

22 A. P. SAGE AND J. L. MELSA, *System Identification*, Academic Press, New York, 1971.

23 K. N. STANTON, An incremental wattmeter for use in interconnected power systems, *Trans. Instn Engrs. Aust.*, May 1964, p. 20.

24 K. N. STANTON, Estimation of turboalternator transfer functions using normal operating data, *Proc. IEE*, **122** (1965) 1713–1720.

25 K. N. STANTON, Use of normal operating data from a power system to estimate turboalternator transfer functions, Report, Purdue University, West Lafayette, Indiana, 1969.

26 W. G. TUEL, JR., Canonical forms for linear systems—I, *IBM Res. Rept RJ375*, San Jose, California, 1966.

27 K. T. WOO, Maximum likelihood identification of noisy systems, *2nd IFAC Symposium on Identification and Process Parameter Estimation, Prague, 1970*, Paper 3.1.

Discussion

J. P. WAHA (*INTERCOM, S.A., Brussels, Belgium*)

In non-linear identification it is felt that quasi-linearization is useful. Work done in

Belgium on a thermal model of an 80 MVA 150 kV/70 kV transformer has shown its possibilities. The theoretical model used was of the first order, with non-linear forcing functions of the form U^m. Using the m exponent from constructive data, errors in temperature of more than 10°C were obtained. Identifying by quasi-linearization and using the new coefficients, including a new m exponent, led to an identification error of less than 2°C. The measurement length was approximately one week, the sampling period was of the order of minutes, and the dominant time constant of the transformer was about 4 hours.

What method was used to identify the non-linear model of the drum boiler?

K. J. ÅSTRÖM

The particular values of the coefficients were obtained by curve fitting to several different experiments. The full details are given in ref. 4 of my paper.

L. S. DZUNG (*Brown, Boveri & Co. Ltd., Baden, Switzerland*)

The difference between the measured behaviour and that predicted by the theoretical model may be due to insufficiency of the model, or to errors of measurement or of the predictor. Professor Åström tries to minimize the deviation by successively increasing the number of free parameters of his model. This is certainly in order if the aim is to simulate the system behaviour as accurately as possible. Now, in the case of a small number of model parameters, it is possible to ascribe physical meaning to each parameter, which may then be modified to obtain the desired system behaviour. If the number of parameters is increased, it is not always possible to interpret the proper physical meaning of each parameter. In fact, some of these parameters may have no physical significance. A model with too many parameters may therefore be less useful for the designer of the system, although it may be helpful to the designer of control apparatus. I think there should be some criterion for the optimal number of parameters for this aspect also.

K. J. ÅSTRÖM

The loss function will always decrease with an increasing number of parameters. The decrease in the loss function, however, may not be significant. As discussed in §6 of my paper, I have found it useful to approach the problem of order as a statistical hypothesis test using the test quantity (32). I do not claim that this gives *the optimal* number of parameters. It is, however, one way to get an indication of the order. The possibility of giving a physical interpretation of the model parameters depends on the model structure.

C. J. EAGLEN (*Brown, Boveri & Co. Ltd., Baden, Switzerland*)

How does the sampling interval influence parameter estimation for control and simulation?

K. J. ÅSTRÖM

The selection of the sampling interval is very important. If the sampling interval is h, it is virtually impossible to estimate parameters of modes with time constants less than, say, $0.3h$ with a reasonable accuracy. Hence, the selection of sampling interval puts an upper limit to the frequency range over which the model is valid. This problem is further discussed in the following references.

K. J. ÅSTRÖM, On the choice of sampling rates in parametric identification of time series, *Information Sciences*, **1** (1969) 273–278.
I. GUSTAVSSON, Choice of sampling interval for parametric identification, *Report 7103*, Division of Automatic Control, Lund Inst. of Technology, April 1971.

H. GLAVITSCH (*Brown, Boveri & Co. Ltd., Baden, Switzerland*)
Referring to the first practical case studied, I would like to ask if the power generator supplied an isolated load or if it was connected to an infinite bus.

K. J. ÅSTRÖM
The measurements were done by Dr. Stanton. Since he is in this room I would kindly ask him to answer this question.

K. N. STANTON (*Systems Control, Inc., Palo Alto, California, U.S.A.*)
The power generator data were collected from a 50 MW machine operating normally and interconnected with the rest of the power system through high voltage cables, but the coupling and the data are typical of those experienced in power systems.

H. GLAVITSCH
From Dr. Stanton's answer I gather that the generator was connected to the rest of the system. Hence, the model in this case has to be at least of second order because of the interaction with other generators.

K. J. ÅSTRÖM
The order test does in fact indicate that a model of fifth order would be consistent with the experimental data.

W. B. JERVIS (*Merz and McLellan, Newcastle upon Tyne, England*)
I should like to confine my remarks to the random sampling measurements which the author referred to in the nuclear boiler studies. I note that the author employed a large signal level into the reactor controls which gave an electrical power excursion of approximately 10 MW. This value seems to be an extremely high error level for injection and the effect of non-linearities could lead to serious error if linearized theory was assumed for the model. This influence could be shown by injecting a small level perturbation, *e.g.* 1 MW or smaller, and comparing the results obtained.

What precautions would the author take to ensure that site measurements and model results give comparable results?

Regarding the length of time employed for random sampling, I noticed that the author gave an illustration for a 30 min period. My personal experience with sinusoidal sampling has shown that significant error can be introduced if the length of time employed is sufficient to allow load conditions on the machine and network to change between start and finish of the test. It would seem that short-time sampling should be the aim, provided that accuracy is not lost by losing the effects of the largest time lags. I would like to mention that practical models, particularly for boilers, have an economic incentive to a user to shorten the commissioning time at site by predicting the correct setting for the automatic control settings.

K. J. Åström

The experiments on the nuclear reactor were performed by AB Atomenergi. According to them (see ref. below) the ÅGESTA reactor can be modelled by linear models for the power excursions used. Our models also agree well with those obtained by other techniques and smaller excursions. I do not know enough about nuclear reactors to judge whether the ÅGESTA reactor is non-typical in this respect. I appreciate your remark concerning the use of models to shorten commissioning.

P. Å. Bliselius, H. Vollmer and F. Åkerhjelm, Experimental and theoretical dynamic study of the Ågesta Nuclear Power Station, *Report AE-376*, AB Atomenergi, Stockholm, 1969.

A. H. Glattfelder (*Eidg. Technische Hochschule, Zürich, Switzerland*)

I would like to ask a few questions on the last part of your paper—trying to link modelling and identification (input–output only methods).

Professor Profos has shown that certain characteristic coefficients of boiler dynamics can easily be calculated from geometrical and physical data. What are the main connections between these coefficients and the matrix elements given in your paper? How well do identified and precalculated values agree?

In the normal operation of a power plant, roughly stationary conditions exist for only about 30–60 min. How large will confidence intervals for the matrix elements be after this time? Do you feel that this would be a sufficient basis on which to construct an optimal control scheme giving better results than conventional linear control?

K. J. Åström

The model like the one in Example 1, §4.2, and the non-linear boiler model in §9 are in fact based on physical considerations following work by Professor Profos and other related results. Thus the parameters can also be determined from construction data. A detailed comparison of identified parameter values with construction data is given in Dr. Eklund's thesis (ref. 10 of my paper). If some crucial parameters were drifting so much that the performance of the control scheme deteriorated, I would propose that you tracked the parameters in real time and incorporated these estimates in your control strategy.

G. Quazza (*ENEL, Milan, Italy*)

One difficulty with statistical identification of not entirely stationary systems, as our experience in trying to determine the transfer function $\Delta f/\Delta P$ of an electric power system has clearly shown, is associated with the conflicting requirements of stationariness and low frequency accuracy in the choice of the recording sample length. Reasonably good confidence limits in the transfer function parameters affecting low frequency response could be obtained only after several hours of recording samples—in the case of the electric power system, when spontaneous random disturbances are used as the stimulating input—while power system structure changes are likely to occur and become relevant in 2 or 3 hours, with consequent modifications in the transfer function itself. On the other hand, the typical spectrum of spontaneous random load changes is approximately Poissonian, with decreasing magnitudes for increasing frequencies. Hence, the "stimulation" of high frequencies is scarce, with consequently a poor approximation of the identifiable transfer function in that range.

For these reasons, statistical identification of electrical power systems by spontaneous signals usually yields less accurate results than deterministic identification. However, it has the advantage of not necessarily requiring isolation of the network from interconnected neighbours.

Could the author comment upon such problems and on the comparison of statistical and deterministic identification?

K. J. Åström

The first statistical identification techniques relied on the spontaneous random disturbances as stimulating inputs. This is not, however, a prerequisite for using statistical identification techniques. One great advantage of the methods I have discussed today is that they make few assumptions on the input. In fact I strongly recommend that external perturbations are used. In the examples presented in my paper only the power generator experiment was done using natural perturbations. In all the other cases external perturbations were used. A comparison of the relative accuracies of the parameters of the power generator with, for example, those of the nuclear reactor clearly shows the importance of using external perturbations.