A Course in

TEACHING TIME SERIES TO CHEMICAL ENGINEERS

B. P. GRAHAM AND A. JUTAN
University of Queensland
St. Lucia, Qld. Australia 4067

A GRADUATE COURSE on time series analysis is offered in the department of chemical engineering at the University of Queensland. The course is based on an interactive graphics time series identification and modelling computer package (TSIM). This package is an enhanced version of a similar package obtained from the process control group in the department of chemical engineering at McMaster University in Hamilton, Canada.

The course takes one month (intensive) and is divided into three sections. The first section involves classical identification methods such as impulse testing, frequency methods, and fourier methods. It is followed by an identification project. The second section involves the graduates' attendance at an intensive three-day industrial workshop on time series analysis and its applications. In this workshop the basic theory of time series identification and modelling is presented. The topics covered include data generation, identification, parameter estimation, and diagnostic model checking. The theory is augmented by extensive hands-on experience with the TSIM computer package. As an introduction to time series applications, topics in process control are presented. This includes a unit on minimum variance control. Normally, this second section would be spread out over one semester, but the concurrent industrial workshop with attendees from outside the university make this format more efficient.

The final section of the course, following the industrial workshop, is an extensive project which forms the course assessment. The project requires that the students generate data and then identify and fit a model to it. Optimal input sequences, forecasting using leading indicators, and statistical diagnostic tests for model adequacy are covered. Finally, the students are required to design a minimum variance controller using their model and to develop a simulation to test out the improved control. The graduates are encouraged to use their own experimental data in addition to data from a simulation.

Throughout the course a basic familiarity with statistics is assumed. Most of the graduates have either attended, or are concurrently attending, a course in applied statistics which is also offered in the department.

In this paper we will briefly describe the TSIM computer package and demonstrate how it is used to identify a transfer function plus noise model. The example given is that of the gas furnace data of Box and Jenkins.

The course takes one month (intensive) and is divided into three sections. The first section involves classical identification methods such as impulse testing, frequency methods, and fourier methods.

TIME SERIES ANALYSIS

It is often necessary to gain a better understanding of the dynamic behaviour of a process if

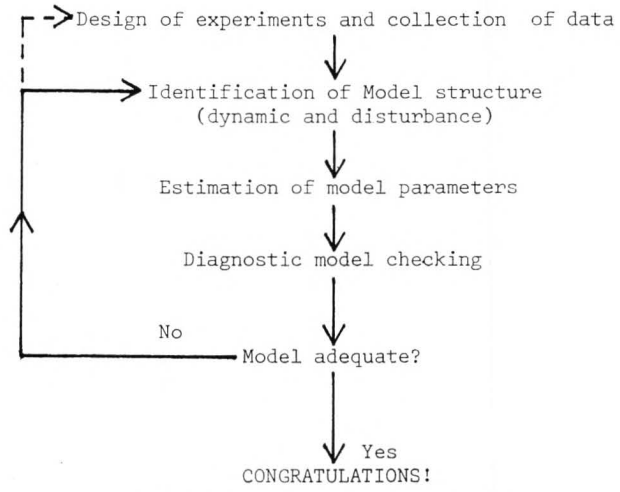
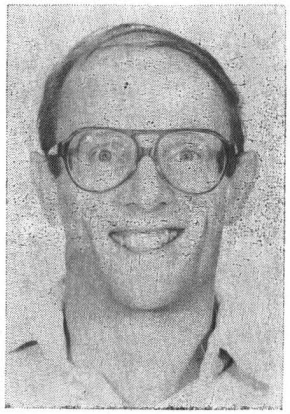


FIGURE 1. Time series identification procedure.



Bruce P. Graham graduated from Flinders University of South Australia in 1980 with an honours degree in applied mathematics. From 1981 to 1982 he worked on real-time computing systems for THORN—E.M.I. Electronics (Australia). He joined the Department of Chemical Engineering at the University of Queensland in 1983 as a PhD student and is currently completing his PhD on fuzzy identification and control. **(L)**

Arthur Jutan is a temporary lecturer in the Department of Chemical Engineering at the University of Queensland. He is currently on leave from Xerox Research Centre of Canada. He obtained his BSc (Eng) from University of Witwatersrand in South Africa in 1970, and his MEng and PhD degrees from McMaster University in 1976. He is a registered professional engineer in Ontario and a member of AIChE. **(R)**

---

one is to make improvements in its design, operation, or control. One approach that could be taken would be to develop a process model from the theoretical unsteady-state mass, energy, and momentum balances for the process. However, the time and effort necessary to develop such a model for anything but a rather simple unit operation is often not economically justifiable. Furthermore, many of the parameters in the mechanistic model would often be unknown and have to be estimated from process and laboratory data. Therefore, a useful alternative in many cases is to obtain the necessary dynamic information directly from data collected from the process. It is this approach, of empirically identifying the dynamic characteristics of a system directly from data collected from it, that constitutes the field of time series analysis. The time series identification procedure consists of a series of steps as shown in Fig. 1.

The procedure aims to identify and estimate values for the parameters of a transfer function plus noise model relating values of the input, x_t, to values of the output, y_t. The model has a deterministic and a stochastic part, shown in block diagram form in Fig. 2.

The transfer function is a deterministic relationship between the input and the output. The noise term is stochastic, and describes random fluctuations that cannot be explained by the transfer function. Noise may be due to such things as measurement errors, catalyst fluctuations, etc. The model is of the form

$$y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + N_t$$

where x_t and y_t are the input to and output from the process at time t, respectively (B = backward shift operator).

THE TSIM COMPUTER PACKAGE

The TSIM computer package consists of five interactive computer programs which, when used in sequence, largely automate the identification and model estimation procedure outlined above. The individual programs are

TSIDENT—identification of a stochastic model for a single time series
TSUNIFIT—parameter estimation for a stochastic model
TSTRANS—transforming and seasonal differencing of a time series pair to generate a new data set
TSXFERID—identification of the transfer function model and generation of noise residuals
TSXFERFIT—simultaneous parameter estimation for the transfer function and noise models, plus diagnostic model checking

Each program in the package is user interactive, asking the user, via questions on the terminal screen, to enter the required data. Most program data is entered directly from the keyboard. The time series are stored in a file in column format.

The results of each program are displayed on the terminal screen and are stored in a file for subsequent perusal or printing by the user. The programs also maintain an optional graphics facility which is available if the user is using a graphics terminal. This facility enables the graphi-

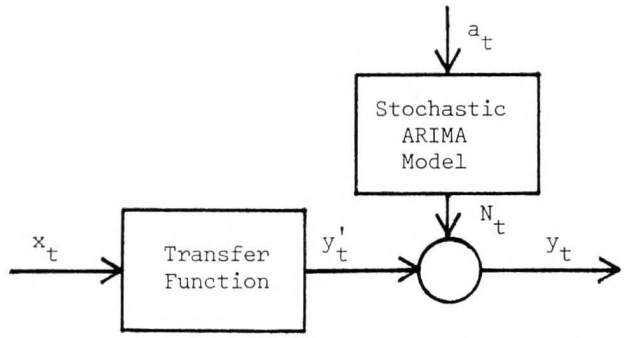


FIGURE 2. Transfer function plus noise model.

cal representation of many of the results from the programs. All subsequent figures in this paper were produced by the computer package.

TRANSFER FUNCTION MODEL IDENTIFICATION OF GAS FURNACE DATA

1. The Gas Furnace Data

It is desired to obtain a transfer function plus noise model relating methane gas feedrate to carbon dioxide concentration in a gas furnace. In the furnace, air and methane combined to form a mixture of gases containing carbon dioxide. The air feed was kept constant but the methane feedrate was varied, and the resulting carbon dioxide concentration in the off-gases was measured. A particular experiment resulted in a time series containing 296 successive pairs of observations (x_t = coded methane feedrate, y_t = carbon dioxide

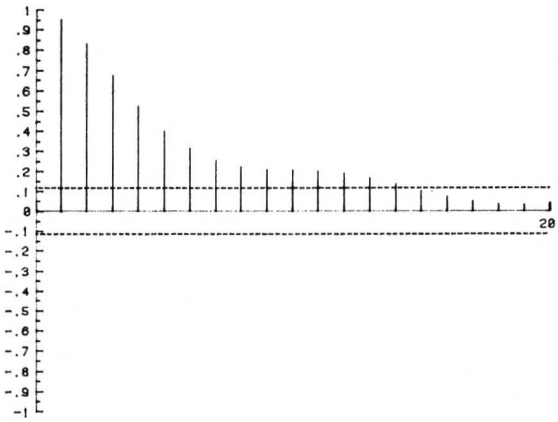


FIGURE 3. Autocorrelations of input time series.

concentration) which were readings at 9 second intervals. This example is from Box and Jenkins (1970).

2. Stochastic Model of the Methane Feedrate, x_t

The first step in the identification procedure is to identify a stochastic ARIMA model for x_t. The model is of the form

$$\nabla^d x_t = \frac{\theta(B)}{\phi(B)} a_t$$

where a_t is white noise (∇ = difference operator).

The program TSIDENT is used to generate the first 20 autocorrelations and partial autocorrelations of the methane feedrate data. These correlations are used to identify the structure of stochastic variations by comparing their patterns

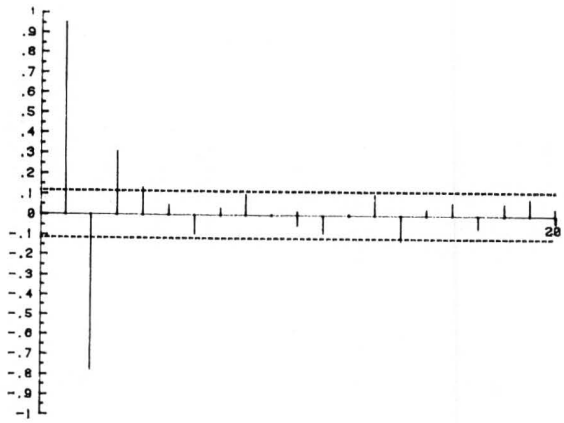


FIGURE 4. Partial autocorrelations of input time series.

(Figs. 3 and 4) with patterns from theoretical stochastic models. Figs. 3 and 4 show that the autocorrelations tail off, and the partial autocorrelations cut off after 3 or 4 points (the fourth point is close to the 95 percent confidence limit). From the theory we can show that the polynomial θ is simply the unit constant, and the polynomial ϕ is of third or fourth order. Also, the input series is stationary, *i.e.* it has a constant mean, hence the degree of differencing required (d) is zero.

Assuming that ϕ is third order, initial estimates for the parameters ϕ_1 to ϕ_3 are obtained from the values of the autocorrelations via the Yule-Walker equations (see Box and Jenkins). These initial estimates are refined using a non-linear least squares algorithm by running the program TSUNIFIT. The final stochastic model for x_t is

$$x_t = \frac{1}{1 - 1.94B + 1.32B^2 - 0.31B^3} a_t$$

The residuals, a_t, produced when using this

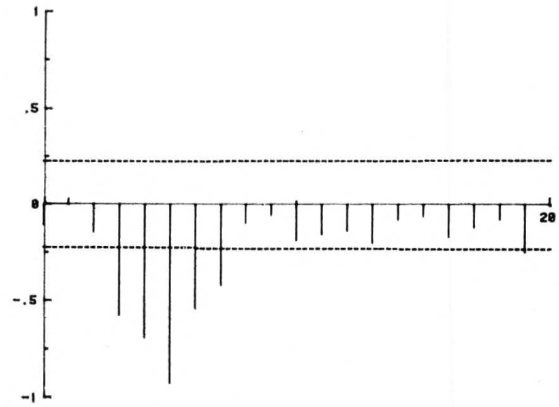


FIGURE 5. Impulse response function.

stochastic model appear to be a good approximation to white noise, since their autocorrelations are small and the chi-squared statistic is well within its allowable range. Hence it is a statistically significant model. We will use this as our stochastic model of the x_t for simplicity (minimum number of parameters).

3. Transfer Function Identification

The program TSXFERID is used to generate the impulse response function for 20 past values of the input x_t. This function gives a model of the form

$$y'_t = V(B)x_t = \hat{v}_0 x_t + \hat{v}_1 x_{t-1} + \ldots + \hat{v}_{20} x_{t-20}$$

The coefficients of this model $(\hat{v}_i)$, and the step response of the model, are shown in Fig. 5 and 6, respectively.

By comparing these impulse response weights with those from known model orders we can tentatively identify our transfer function model order.

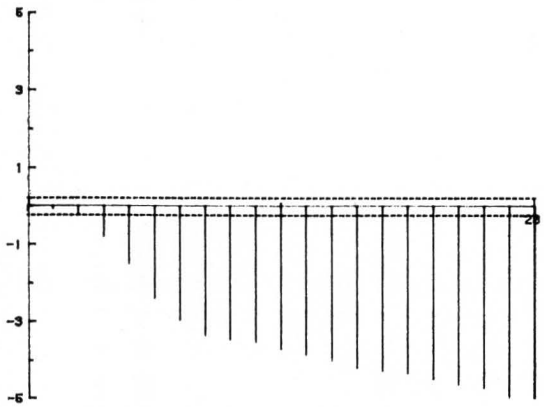


FIGURE 6. Step response function.

In this example the impulse response (Fig. 5) contains a dead-time of three, followed by two odd points, and then an exponential decay. This indicates that the transfer function model has the following form

$$y'_t = \frac{\omega_0 - \omega_1 B - \omega_2 B^2}{1 - \delta_1 B} x_{t-3}$$

Initial estimates for the model parameters are obtained by equating coefficients between the impulse response function and the transfer function. The noise residuals, N_t, are generated as the difference between the impulse response function output and the actual carbon dioxide concentration, *i.e.* $N_t = y_t - y'_t$.

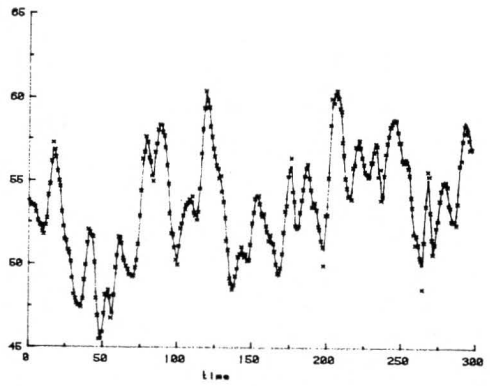


FIGURE 7. Transfer function plus noise model.

4. Stochastic Model of the Noise N_t

Repeating the procedure of section 2 for the noise, N_t, we obtain the stochastic model

$$N_t = \frac{1}{1 - 1.62B + 0.66B^2} a_t$$

Examination of the autocorrelations of the residuals and the chi-squared statistic shows that the noise model is adequate.

5. Transfer Function Plus Noise Model Parameter Estimation

By running the program TSXFERFIT we simultaneously refine the initial parameter estimates for the transfer function and the noise models. The resulting final transfer function plus noise model is

$$y_t = -\frac{(0.52 + 0.4B + 0.51B^2)}{(1 - 0.55B)} x_{t-3} + N_t$$

$$N_t = \frac{1}{(1 - 1.53B + 0.63B^2)} a_t$$

That is

$$y_t - 0.55y_{t-1} = -(0.52x_{t-3} + 0.4x_{t-4} + 0.51x_{t-5}) + N_t$$
$$N_t - 1.53N_{t-1} + 0.63N_{t-2} = a_t$$

Statistical diagnostic tests are applied to the model. These checks show that the above model fits the gas furnace data adequately. A plot of the model-generated y_t's versus the actual carbon dioxide concentration data is shown in Fig. 7.

REFERENCES

Box, G. E. P. and G. M. Jenkins (1970) Time Series Analysis: forecasting and control. Holden Day.

MacGregor, J. F. (1983) Course Notes for Industrial Workshop in Process Control. Mc Master University, Hamilton Canada. □