Regression Parameter Estimation with Serially Correlated Errors

C.A. Glasbey

Thesis submitted for the degree of Doctor of Philosophy

University of Edinburgh

1984

DECLARATION

The following record of research work is submitted as a Thesis for the Degree of Doctor of Philosophy at the University of Edinburgh, having been submitted for no other Degree.

The research work was carried out under the supervision of Mr P.R. Fisk and Mr R. Thompson. Except where due acknowledgement is made the work is original.

C.A. Glasbey.

ABSTRACT

Least-squares regression is a standard statistical technique for relating variables by means of functional relationships. On the assumption that the errors are independent of one another and have equal variance, the regression parameter estimators are approximately efficient and the conventional estimators of the variances of these estimators are approximately unbiased. However, if the errors are correlated, then in general neither property holds true.

This thesis is directed towards overcoming these deficiences when errors are serially correlated; arising because a series of measurements is made on a single experimental unit, rather than on a number of separate units. Two types of solution are considered: (1) model the error process, either empirically or mechanistically, and estimate any unknown parameters jointly with the regression parameters (chapters 4 to 8); (2) derive estimators of the variances of least-squares regression parameter estimators which take account of the error correlations (chapter 9). The diversity of models considered in the first solution is facilitated computationally by the specification in chapter 2 of a new class of serially-structured error processes, termed generalized autoregressive-moving average processes. In chapter 3 solutions of linear stochastic difference and differential equations are shown to be in this class.

Empirical models of the error processes are based entirely on the corresponding data sets; no other information being available on the forms the correlations should take. Basically, a regression

function is fitted by least-squares estimation, the residuals are examined and an error model is identified. In the three examples considered (pharmacokinetic data, data on transport systems within a leaf, and the energy demands of a mechanical model of a suckler cow) the error processes are assumed to be stationary and either the sample autocorrelation coefficients or the periodograms are used to identify the appropriate autocorrelation functions or spectral functions in chapters 4 and 5 respectively.

Antithetically, mechanistic models of the error processes are derived from considerations of the methods by which the data sets were generated. Three particular types of model are considered in the thesis. In chapter 6 linear stochastic difference and differential equations are used: first-order models to describe the variations in a cow's milk yield over its lactation; second-order models to describe the growth in an animal's weight. Stochastic compartment systems are used in chapter 7 to model the pharmacokinetic and leaf transport data previously considered in chapters 4 and 5. Then, in chapter 8 a related model appropriate to cumulative counts is used to describe seed germination numbers.

In chapter 9 estimates of the variances of regression parameter estimators are obtained which are conservative, that is never downward biased, for broad classes of serially-structured error processes. The technique arose as a solution to the inadequacies of the empirical modelling strategy, and is applied to the data sets used in chapter 4.

ACKNOWLEDGEMENTS

This work was undertaken as part of the research programme of the Agricultural and Food Research Council's Unit of Statistics. I am grateful to the AFRC for both financial support and facilities to undertake the research presented.

I am indebted to Drs D. Colquhoun, J.E. Dale, J.M. Bruce, D. Neilson, R.B. Thiessen and R.E.L. Naylor for being permitted to use their data.

My thanks are due to my supervisors Mr P.R. Fisk and Mr R. Thompson for their valuable comments and advice and to Professor D.J. Finney and Dr H.D. Patterson for their encouragement and support. I am also grateful for the widespread support I have received from other members of staff of the Unit of Statistics and from my wife. Finally, I would like to thank Miss L.E. Robertson for typing this thesis so conscientiously.

CONTENTS

Conventions and Notation

Character conventions

The character set is restricted to lower-case and upper-case Latin and Greek alphabets. A general convention has been used as follows:

underlined lower-case characters, for example $\underset{\sim}{x}$, denote column vectors;

lower-case characters with single subscripts, for example x_i, denote individual elements in vectors, in this case the ith element in $\underset{\sim}{x}$;

otherwise lower-case characters, for example x, denote scalar constants or variables;

underlined upper-case characters, for example $\underset{\sim}{A}$, denote matrices, the only exceptions being $\underset{\sim}{D}$, $\underset{\sim}{N}$, $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$ which denote vector random variates in order to distinguish them from their realisations (although in some instances a distinction is not made between a random variate and its realisation);

upper-case characters with double subscripts, for example A_{ij}, denote individual elements in matrices, in this example the (ij)th element of $\underset{\sim}{A}$;

otherwise upper-case characters, for example A, are only used for special purposes.

The specific definitions of all characters used in the thesis follow after other notation has been established.

Matrix notation

Standard matrix notation has been used as follows:

$\underset{\sim}{x}^T$ denotes the row vector which is the transpose of the column vector $\underset{\sim}{x}$;

$\underset{\sim}{A}^T$ denotes the transpose of the matrix $\underset{\sim}{A}$;

$|\underset{\sim}{A}|$ denotes the determinant of the square matrix $\underset{\sim}{A}$;

$tr(\underset{\sim}{A})$ denotes the trace of the square matrix $\underset{\sim}{A}$;

$\underset{\sim}{A}^{-1}$ denotes the inverse of the non-singular square matrix $\underset{\sim}{A}$.

Probability notation

Standard notation has been used as follows:

$E(\underset{\sim}{Y})$ denotes the vector of expectations of elements in $\underset{\sim}{Y}$;

$var(\underset{\sim}{Y})$ denotes the matrix of variances and covariances between elements in $\underset{\sim}{Y}$;

$cov(Y_i,Y_j)$ denotes the covariance between Y_i and Y_j ;

$E(\underset{\sim}{Y}|\underset{\sim}{Z})$ denotes the expectation of $\underset{\sim}{Y}$ conditional upon the realisation of $\underset{\sim}{Z}$;

$cov(Y_i,Y_j|\underset{\sim}{Z})$ denotes the covariance between Y_i and Y_j conditional upon the realisation of $\underset{\sim}{Z}$;

$\underset{\sim}{Y} \sim (\underset{\sim}{f},\underset{\sim}{V})$ denotes that $\underset{\sim}{Y}$ is distributed with mean $\underset{\sim}{f}$ and variance matrix $\underset{\sim}{V}$;

$\underset{\sim}{Y} \sim N(\underset{\sim}{f},\underset{\sim}{V})$ denotes that $\underset{\sim}{Y}$ is multivariate normally distributed with mean $\underset{\sim}{f}$ and variance matrix $\underset{\sim}{V}$;

$\hat{\underset{\sim}{\alpha}}$ denotes a vector random variate which is an estimator of the column vector of parameters, $\underset{\sim}{\alpha}$.

Other mathematical notation

Other mathematical notation has been used as follows:

$\bar{x}$ denotes the complex conjugate of x ;

$\ell n(x)$ denotes the natural logarithm of x ;

$|x|$ denotes the absolute value of x ;

δt denotes a small positive increment to t ;

$o(\delta t)$ denotes a term of order less than δt ;

$\sum_{i=j}^{k} x_i$ denotes the summation from the jth to the kth elements in $\underset{\sim}{x}$, and is defined to be zero if $k < j$;

$\prod_{i=j}^{k} x_i$ denotes the product from the jth to the kth elements in $\underset{\sim}{x}$, and is defined to be unity if $k < j$;

$\partial \underset{\sim}{f} / \partial \underset{\sim}{\alpha}^T$ denotes a matrix, the (ij)th coefficient of which is the partial derivative $\partial f_i / \partial \alpha_j$;

$\partial \underset{\sim}{f} / \partial \alpha$ denotes a column vector, the ith coefficient of which is the partial derivative $\partial f_i / \partial \alpha$;

$\partial f / \partial \underset{\sim}{\alpha}$ denotes a column vector, the ith coefficient of which is the partial derivative $\partial f / \partial \alpha_i$;

$\underset{\sim}{1}$ denotes a vector each of whose coefficients is unity;

$\underset{\sim}{0}$ denotes either a vector, or a matrix, of zeros according to context;

$\{x : x \geq 0,\ x \leq 1\}$ denotes the set of all scalars, say x, such that x lies between 0 and 1 ;

$x \in \Omega$ denotes that x is a member of the set Ω;

$\Omega^{\{1\}} \subset \Omega^{\{2\}}$ denotes that every member of the set $\Omega^{\{1\}}$ is also a member of the set $\Omega^{\{2\}}$.

Subscript, superscript and postscript notation

A standard notation has been employed throughout the thesis as follows:

a single subscript, for example x_i, is used to denote an element in a vector, although in some instances this has not been rigorously applied and the index i has been allowed to take non-positive values;

a double subscript, for example A_{ij}, is used to denote an element in a matrix;

a superscript in angular brackets, for example $x^{\langle i \rangle}$, is used to denote the ith derivative of x with respect to time, conventionally denoted by t;

a superscript in curly brackets, for example $x^{\{i\}}$, is used as an index label, thus $x^{\{1\}}$, $x^{\{2\}}$ and $x^{\{3\}}$ each denote a different scalar;

similarly, * superscript, for example x^*, is used to distinguish between x and x^*;

any superscript which does not involve angular or curly brackets, for example x^i, is used in the conventional sense as a power, in this case x raised to the ith power;

a postscript in brackets, for example $x(t)$, is used to denote the value at t of a scalar variable x which is indexed by a continuous variable.

Numbering in thesis

The thesis is divided into chapters, referred to as chapter 2 for example. Chapters are divided into sections and sub-sections,

referred to, for example, as section 2.1 and [2.1.4] respectively. Equations which are referred to elsewhere in the text are labelled, and these are numbered consecutively within sections, and referred to as equation (2.1.1) for example. Tables and figures are also numbered consecutively in sections and referred to, for example, as table 2.1.1 and figure 2.1.1 respectively.

Characters - Latin lower-case

a: an index variable.

$\underset{\sim}{a}$: the n-vector of weight-corrected departures of the vector of observations from the regression vector, that is $\underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})$.

b: an index variable.

$\underset{\sim}{b}(t)$: the p-vector in chapter 7 of expected number of particles that will immigrate to each compartment in time t.

$\underset{\sim}{c}^{\{i\}}$: n-vectors used in the decomposition of $\underset{\sim}{C}$ in chapter 9.

$\underset{\sim}{d}$: the n-vector of spectral coefficients (at harmonic frequencies) of the error process, used in chapter 5.

e: the base of natural logarithms, that is the constant 2.713...

$\underset{\sim}{e}$: the n-vector of departures (either the random variate or its realisation) of the vector of observations from the regression vector, that is $(\underset{\sim}{y}-\underset{\sim}{f})$, except in chapter 9 where the departures are weight-corrected, that is $\underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})$.

$\underset{\sim}{f}$: the n-vector of regression function values.

g: an index variable.

h: an index variable.

i: an index variable.

j: an index variable.

k: an index variable.

ℓ: an index variable.

m: the number of regression parameters.

n: the number of observations.

$\underset{\sim}{n}$: the p-vector in chapter 7 of the expected number of particles in each compartment at time zero.

$o(\delta t)$: denotes term of order less than δt .

p: the non-negative first integer parameter in the specification of a GARMA process, see [2.2.1], which also has specialized meanings in different chapters:

order of difference/differential equations in chapters 3 and 6;

number of Markov processes in chapter 4;

number of compartments in chapter 7.

p(t): the cumulative probability function used in chapter 8.

$\underset{\sim}{p}$: the n-vector of cumulative probabilities used in chapter 8.

q: the non-negative second integer parameter in the specification of a GARMA process, see [2.2.1].

$\underset{\sim}{q}$: the p-vector in chapter 7 of emigration rates from compartments.

r: the degrees of freedom of the conservative estimator of variance, $\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}$, used in chapter 9, defined in [9.2.18].

$\underset{\sim}{r}$: the p-vector in chapter 7 of immigration rates into compartments.

s: general index variable except in chapter 9 where it denotes the index of the regression parameter of interest, that is β_s .

t: an index variable, conventionally representing time, except

in chapter 9 where it denotes the number of vectors $\underset{\sim}{c}^{\{i\}}$ in the decomposition of $\underset{\sim}{C}$.

$\underset{\sim}{t}$: the n-vector of observation times, in ascending order.

u: an index variable.

$\underset{\sim}{u}$: an n-vector, the sth column of $\underset{\sim}{U}$ used in chapter 9.

v: an index variable.

$\underset{\sim}{v}$: a general vector which is defined in a specific context.

w: an index variable.

$\underset{\sim}{w}$: a general vector which is defined in a specific context.

x: an index variable.

$\underset{\sim}{x}$: a general vector which is defined in a specific context.

y: an index variable.

$\underset{\sim}{y}$: the n-vector of observations.

$\underset{\sim}{z}$: the n-vector of model residuals, that is the departures $\underset{\sim}{e}$ corrected to take account of model covariances by $\underset{\sim}{z} = \underset{\sim}{\Theta}^{-1}\underset{\sim}{\Phi}\,\underset{\sim}{e}$.

Characters - Greek lower-case

α: a single model parameter.

$\underset{\sim}{\alpha}$: a vector of model parameters whose length depends on the specific context.

β: a single regression parameter, that is m is equal to 1.

$\underset{\sim}{\beta}$: the m-vector of regression parameters.

$\underset{\sim}{\gamma}$: the vector of autocovariances of the error process.

δ: the non-negative second parameter in the specification of a cGARMA process, see [2.2.7].

δt: a small positive increment to t .

ε: an arbitrarily small positive number.

$\zeta^{\{k\}}(t)$: the first set of functions, indexed by continuous t, used in decomposition of variance matrix of cGARMA process, see [2.2.7].

$\underset{\sim}{\zeta}^{\{k\}}$: the first set of n-vectors used in decomposition of variance matrix of GARMA process, see [2.2.1].

$\eta^{\{k\}}(t)$: the second set of functions, indexed by continuous t, used in decomposition of variance matrix of cGARMA process, see [2.2.7].

$\underset{\sim}{\eta}^{\{k\}}$: the second set of n-vectors used in decomposition of variance matrix of GARMA process, see [2.2.1].

θ: a single parameter in a moving average process of order 1 in discrete time.

$\underset{\sim}{\theta}$: a vector of parameters in a moving average process in discrete time.

ι: the square-root of -1, used in chapter 5.

κ: a single weight associated with an exponential term.

$\kappa(\ell)$: a variable weight used in [4.5.5].

$\underset{\sim}{\kappa}$: a vector of weights associated with exponential terms.

λ: a single rate constant in an exponential term, except in chapter 9 where it denotes an eigenvector.

$\lambda(\ell)$: a variable rate constant used in [4.5.5].

$\underset{\sim}{\lambda}$: a vector of rate constants in exponential terms.

μ: a shift parameter used in chapter 9.

$\mu(t)$: the weight function used to specify a fitter in [2.2.8].

$\underset{\sim}{\mu}$: a general vector which is defined in a specific context, usually to describe weighting constants.

ν: a parameter used to define a filter in [2.2.8].

ξ: a single parameter in an autoregressive process of order 1 in continuous time.

$\underset{\sim}{\xi}$: a vector of parameters in an autoregressive process in continuous time.

π: the constant 3.141.... .

$\rho(t)$: the autocorrelation function of a stationary process in continuous time, indexed by continuous t .

$\underset{\sim}{\rho}$: the vector of autocorrelation coefficients of a stationary process in discrete time, or the correlations between consecutive terms in a Markov process, depending upon the context.

σ^2: a variance term.

τ^2: a variance term, usually the scaling parameter in a GARMA process, see [2.2.1].

υ: the non-negative second integer parameter used in the definition of Ω in chapter 9, see [9.2.11].

ϕ: a single parameter in an autoregressive process of order 1 in discrete time.

$\underset{\sim}{\phi}$: a vector of parameters in an autoregressive process in discrete time.

χ^2_i: the central chi-square probability distribution with i degrees of freedom.

$\underset{\sim}{\psi}$: a vector of parameters in a moving average process in continuous time.

ω: the non-negative first integer parameter used in the definition of Ω in chapter 9, see [9.2.11].

Characters - Latin upper-case

$A^{\{i\}}$: scalar random variates, defined in [3.3.2] and used in the solution of differential equations.

$\underset{\sim}{A}$: a general matrix which is defined in a specific context.

$\underset{\sim}{C}$: the n by n symmetric matrix used in chapter 9.

$\underset{\sim}{D}$: the random vector of length n of the periodogram used in chapter 5.

$\underset{\sim}{G}$: the p by p matrix of right-eigenvectors of $\underset{\sim}{Q}$, used in chapter 7.

$\underset{\sim}{H}$: the n by n complex matrix used in [5.2.2].

$\underset{\sim}{I}$: the identity matrix, the size being specified in each situation.

$\underset{\sim}{K}$: a matrix of weights associated with exponential terms.

L_M: the negative Gaussian log-likelihood function, defined in [2.3.2].

L_P: the negative of the logarithm of the pseudo-likelihood function, used in chapter 5.

L_R: the negative residual Gaussian log-likelihood function, defined in [2.3.2].

L_S: the sum of squares function, defined in [2.3.2].

N: the standard normal probability distribution, except in chapter 8 where it denotes the total number of events in the cumulative counts model.

$\underset{\sim}{N}(t)$: the random vector of length p, used in chapter 7, of the number of particles in each compartment at time t.

$\underset{\sim}{P}(t)$: the p by p matrix used in chapter 7 of transition probabilities of particles between compartment in time t.

$\underset{\sim}{Q}$: the p by p matrix used in chapter 7 of transition probability rates of particles between compartments.

$\underset{\sim}{R}$: the n by n matrix used to specify generalized least-squares estimation in chapter 9, see [9.2.1].

S: the generalized sum of squares, defined in [2.3.2].

T: used as a superscript to denote a vector or matrix transpose.

$\underset{\sim}{U}$: the n by m matrix used in chapter 9, defined in [9.2.1].

$\underset{\sim}{V}$: the n by n symmetric positive-definite variance matrix of observations $\underset{\sim}{y}$, and therefore also of $\underset{\sim}{e}$, possibly after the removal of a scaling parameter.

$\underset{\sim}{W}$: the n by n diagonal matrix of weights.

$\underset{\sim}{X}$: the n by m regression design matrix, defined in [2.3.2], except in chapter 9 where it is weight-corrected as defined in [9.2.1].

Y(t): a random variate indexed by continuous t.

$\underset{\sim}{Y}$: a random vector.

Z(t): a random variate indexed by continuous t, used as input.

$\underset{\sim}{Z}$: a random vector, used as input.

Characters - Greek upper-case

$\underset{\sim}{\Gamma}$: a general n by n lower-triangular band matrix used in [2.4.10].

Δ: a differencing operator acting on the lower-triangle of a matrix, defined in [9.2.11].

$\underset{\sim}{\Theta}$: the n by n lower-triangular matrix with bandwidth (q+1) used as the second matrix in the decomposition of a variance matrix of a GARMA process, defined in [2.2.2].

$\underset{\sim}{\Lambda}$: a general variance matrix.

Π: the product symbol.

$\sum$: the summation symbol.

$\underset{\sim}{T}$: an n by n symmetric matrix used in chapter 9, the lower triangle being defined by $\Delta^2\underset{\sim}{V}$ in [9.2.14].

$\underset{\sim}{\Phi}$: the n by n lower-triangular matrix with bandwidth (p+1) used as the first matrix in the decomposition of a variance matrix of a GARMA process, defined in [2.2.2].

$\underset{\sim}{\Psi}$: an n by n symmetric matrix used in chapter 9, the lower-triangle being defined by $\Delta\underset{\sim}{V}$ in [9.2.13].

Ω: the set of matrices of size n by n considered to be possible variance matrices, used in chapter 9.

1. Prelude

1.1 Introduction

1.1.1 Least-squares regression

A fundamental objective in scientific research is to relate variables by means of functional relationships. If variables are not observed precisely, but subject to experimental errors of various sorts, then no function will fit exactly. A standard statistical technique for fitting functions in this situation is to regress observations of one variable, the dependent variate, against a function of the observations on other explanatory variables. In specific applications, objectives inevitably vary but often include the following:

a) to summarize the data;

b) to predict values of the dependent variate for specific values of the explanatory variables;

and c) to discover and understand fundamental relationships between variates.

Conventionally in this thesis, the observations are denoted by $\underset{\sim}{y}$, a vector of length n, and the values the regression function takes, based on the values of the explanatory variables, are denoted by $\underset{\sim}{f}$ which depends on m unknown parameters $\underset{\sim}{\beta}$. The standard way of estimating these parameters is by the method of ordinary least-squares, that is by minimizing the criterion function

$$(\underset{\sim}{y} - \underset{\sim}{f})^T(\underset{\sim}{y} - \underset{\sim}{f})$$

with respect to $\underset{\sim}{\beta}$, where superscript T denotes the transpose of a vector (or matrix). If $\underset{\sim}{f}$ is a function linear in $\underset{\sim}{\beta}$ then

$$\underset{\sim}{f} = \underset{\sim}{X}\,\underset{\sim}{\beta}$$

for some n by m matrix $\underset{\sim}{X}$ commonly called the regression design matrix, and $\underset{\sim}{\beta}$ is estimated as

(1.1.1) $$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{y} \quad ,$$

provided that $(\underset{\sim}{X}^T\underset{\sim}{X})$ is non-singular, where the superscript -1 denotes a matrix inverse. If $\underset{\sim}{f}$ is non-linear in $\underset{\sim}{\beta}$ it may still be approximated by a linear function, based on a Taylor expansion of $\underset{\sim}{f}$ in terms of $\underset{\sim}{\beta}$ about a value close to the best-fitting value.

Under certain assumptions, this method of estimation is optimal in the sense that it gives the linear unbiased estimator of $\underset{\sim}{\beta}$ with the smallest variance, and the variance of the estimator can be estimated as

(1.1.2) $$\hat{\text{var}}(\hat{\underset{\sim}{\beta}}) = \hat{\tau}^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \quad ,$$

where $$\hat{\tau}^2 = \hat{\underset{\sim}{e}}^T\hat{\underset{\sim}{e}}/(n-m) \quad ,$$

and the vector of residuals is defined as

$$\hat{\underset{\sim}{e}} = \underset{\sim}{y} - \underset{\sim}{X}\,\hat{\underset{\sim}{\beta}} \quad .$$

1.1.2 Reasons for noticeably correlated residuals

Two important conditions for $\hat{\underset{\sim}{\beta}}$ to be optimal, and $\hat{\text{var}}(\hat{\underset{\sim}{\beta}})$ to be a good estimator of $\text{var}(\hat{\underset{\sim}{\beta}})$, are:

a) the functional relationship $\underset{\sim}{f}$ is of approximately the correct form;

b) the individual elements in the vector $(\underset{\sim}{y} - \underset{\sim}{f})$ are distributed independently of one another with a common variance τ^2.

If the data exhibit systematic departures from the fitted regression curve, that is the residuals are noticeably correlated, then it cannot be possible for both these conditions to be true. (The term "noticeably" is important because some correlation is automatically introduced into the residuals by the process of estimating $\underset{\sim}{\beta}$, so that even if conditions (a) and (b) both hold

$$\text{var}(\hat{\underset{\sim}{e}}) = \tau^2(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T). \quad)$$

In most applications of regression it is assumed that condition (a) is violated; so the emphasis is on finding a suitable alternative regression function if systematic departures are encountered. However in some situations, for example when data are collected serially in time from a single experimental unit, assumption (b) is not wholly plausible. For serially-structured data, any errors causing a discrepancy between $\underset{\sim}{y}$ and $\underset{\sim}{f}$ may persist over several observations, thus manifesting themselves as systematic departures. This poses a dilemma: when serially-structured data depart systematically from a fitted regression is it because the function is wrong or is it because the errors are correlated?

The answer is that it is impossible to distinguish between the two alternatives for a single series of serially-structured data. To illustrate this, consider a series of data $(\underset{\sim}{y})$ generated by

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 \sin x_i + e_i \qquad \text{for } i = 1,\ldots,n,$$

where $x_i = i/10$ for $i = 1,\ldots,n$, β_1 and β_2 are constants, β_3 and $\underset{\sim}{e}$ are random terms with

$$\beta_3 \sim N(0,\sigma^2) \text{ independently of } \underset{\sim}{e},$$

and $$e_i \sim N(0,\tau^2) \text{ independently for } i = 1,\ldots,n .$$

Notation is being used here that will reappear throughout the

thesis: y_i denotes the ith element in the vector $\underset{\sim}{y}$ of length n, and $\beta_3 \sim N(0,\sigma^2)$ denotes that β_3 is normally distributed with mean zero and variance σ^2, with the obvious extension to the multidimensional case. Two models of the data:

i) $\underset{\sim}{y} \sim N(\beta_1\underset{\sim}{1} + \beta_2\underset{\sim}{x} + \beta_3 \sin \underset{\sim}{x}\ ,\ \tau^2\underset{\sim}{I})\ ,\ \beta_3 \sim N(0,\sigma^2)\ ,$

ii) $\underset{\sim}{y} \sim N(\beta_1\underset{\sim}{1} + \beta_2\underset{\sim}{x}\ ,\ \sigma^2(\sin \underset{\sim}{x})(\sin \underset{\sim}{x})^T + \tau^2\underset{\sim}{I})\ ,$

are mathematically equivalent because they give identical probability densities for $\underset{\sim}{y}$. (Here $\underset{\sim}{1}$ denotes a vector of 1's, $\sin \underset{\sim}{x}$ denotes an n-vector whose ith element is $\sin x_i$ and $\underset{\sim}{I}$ denotes the identity matrix of size n.) If inference on the model parameters is to be made from a single realisation of $\underset{\sim}{y}$, then either τ^2, β_1, β_2 and the particular realisation of β_3 (ignoring its underlying distribution) may be estimated or τ^2, β_1, β_2 and σ^2 may be estimated. However, the form of the likelihood is different in the two approaches so the maximum likelihood estimators of the common parameters τ^2, β_1 and β_2 will not agree. Also, there are philosophical differences between the two methods which have parallels in certain analyses of variance: in the first case β_3 is treated as a fixed effect, whereas in the second case it is treated as a random effect. Most importantly, in (i), conditioning on β_3, the oscillations about a straight line are modelled by a systematic sinusoid plus white noise error, whereas in (ii) they are modelled by a correlated error structure. Thus there is in some sense an interchangeability or duality between the regression model and the error model.

A sensible compromise would seem to be to require the regression function to describe the long-term trends in the serially-structured data and the correlations in the errors the

short-term fluctuations. Of course these are vague terms open to different interpretations. The final arbiter will be the use to which the model is to be put, for example, for a simple summary of data it may be preferable for the regression function to explain all systematic variability whereas with more fundamental objectives a correlated stochastic component may make more sense.

Predominantly in this thesis the postulated regression function will be assumed correct and attempts to improve the fit of the model will be channelled into different choices of error correlation structure. Because of the duality noted above this is in no way a restrictive choice to make. In the examples considered the regression functions do explain most of the variability in the data, and an alternative justification for concentrating on correlation structures is that they are a more parsimonious description of the residual variability.

1.1.3 Effects of correlated errors on least-squares regression

There are two well understood consequences of correlated errors on least-squares regression estimation. In the first place, the estimator of $\mathrm{var}(\hat{\underset{\sim}{\beta}})$ given in equation (1.1.2) is usually biased because if

$$\underset{\sim}{y} \sim N(\underset{\sim}{f}, \tau^2 \underset{\sim}{V})$$

where $\underset{\sim}{V}$ is an n by n symmetric positive-definite matrix,

then $$\mathrm{var}(\hat{\underset{\sim}{\beta}}) = \tau^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{V}\,\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} ,$$

which will not in general be equal to the expectation of $\hat{\tau}^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}$, both because $\hat{\tau}^2$ will not be an unbiased estimator of τ^2 and because

$$(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \neq (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{V}\,\underset{\sim}{X}\,(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} .$$

(See for example Judge, Griffiths, Hill and Lee (1980, pp113-117).) Typically, if correlations are positive between adjacent observations in the serially observed data and consecutive elements in each column of $\underset{\sim}{X}$ are positively correlated (i.e. the influence of $\underset{\sim}{\beta}$ on $\underset{\sim}{f}$ changes slowly) then $\hat{\text{var}}(\underset{\sim}{\hat{\beta}})$ will be biased downwards and its use will exaggerate the precision with which $\underset{\sim}{\beta}$ is estimated.

The second consequence is that $\underset{\sim}{\hat{\beta}}$ is usually an inefficient estimator of $\underset{\sim}{\beta}$. The best linear estimator when $\underset{\sim}{V}$ is known is the generalized least-squares estimator (Aitken, 1934)

$$(\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

provided that $(\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X})$ is non-singular. The estimator has a variance matrix given by

$$\tau^2(\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1} \quad , \tag{1.1.3}$$

which will never be larger and will usually be smaller than the variance matrix of the ordinary least-squares estimator, $\underset{\sim}{\hat{\beta}}$. Expressing this more precisely, the difference between $\text{var}(\underset{\sim}{\hat{\beta}})$ and equation (1.1.3) is always a positive-semidefinite matrix. Therefore, least-squares estimation does not make the most effective use of the data.

As an aside, it may be noted that although the least-squares estimator of $\text{var}(\underset{\sim}{\hat{\beta}})$ is biased, the least-squares estimator of $\underset{\sim}{\beta}$ is unbiased whatever form the covariances of $\underset{\sim}{y}$ take because, from equation (1.1.1)

$$\begin{aligned} E(\underset{\sim}{\hat{\beta}}) &= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\, E(\underset{\sim}{y}) \quad , \\ &= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X}\,\underset{\sim}{\beta} \quad , \\ &= \underset{\sim}{\beta} \quad . \end{aligned}$$

1.1.4 Solutions

The regression objectives in a particular problem will bear upon the choice of solution to overcome the difficulties of [1.1.3]. Nevertheless, three broad categories of approach to regression parameter estimation when the errors are correlated can be identified as follows:

a) The easiest option, when faced with correlated errors, is to retain the least-squares parameter estimator $\hat{\underset{\sim}{\beta}}$ but discard the biased estimate of its variance matrix. This is most useful when no estimation of precision is required, for example when replicates of experimental units are also available and between unit variability is of prime consideration. This strategy can be justified by arguing that least-squares estimation is often not very inefficient, and is intuitively appealing irrespective of probabilistic assumptions as it gives a simply understood summary of a set of observations. However, no reliable information on precision is available.

b) A second approach is to assume that the errors arose from a distribution with a particular variance matrix $\underset{\sim}{V}$, where $\underset{\sim}{V}$ possibly has a few unknown parameters. Once these parameters have been estimated from the data, $\underset{\sim}{\beta}$ can be estimated, for example by generalized least-squares. More simply, regression and error variance parameters can be estimated jointly by maximizing the log-likelihood, or some variant of it. Thus, provided the correct form of $\underset{\sim}{V}$ is assumed, $\underset{\sim}{\beta}$ can be estimated with asymptotic efficiency and approximately unbiased estimates of precision can be obtained.

c) Because of the intuitive appeal of least-squares estimation and the dependence of approach (b) on the assumed structure for $\underset{\sim}{V}$, a third solution is to make a compromise between (a) and (b). The least-squares estimator of $\underset{\sim}{\beta}$ is retained but an estimate of $\text{var}(\underset{\sim}{\hat{\beta}})$ is obtained which takes some account of the correlations in the errors. This approach has the potential of yielding valid estimates which are less dependent on the assumed structure for $\underset{\sim}{V}$ than those in approach (b).

1.2 Outline of thesis

1.2.1 Structure

The thesis divides into two main parts: chapters 2 to 8 are concerned with solution (b) in [1.1.4], that is modelling the error process; chapter 9 develops a solution in category (c), that is the estimation of the variance of $\underset{\sim}{\hat{\beta}}$ without explicit reference to $\underset{\sim}{V}$. Although, for the sake of generality, section 1.1 dealt with correlations of an arbitrary nature this thesis concentrates solely on serially correlated errors.

The key to the variety of models considered in chapters 2 to 8 is a decomposition of a class of serially-structured variance matrices given in chapter 2. This enables large variance matrices to be manipulated without recourse to handling each element separately. In chapter 3 the solutions are given for linear stochastic difference and differential equations which are used in later chapters. They are also shown to be examples of the processes considered in chapter 2.

Two approaches have been used in modelling the error process, denoted "empirical" and "mechanistic" by Thornley (1976, pp4-6). The empirical method is described by Thornley as consisting of

> "looking at the experimental data, possibly doing some analysis of the data, and trying to make an intelligent guess at a (usually simple) form of equation or set of equations which can be used as a mathematical model and fitted to the data."

In the present context, this means using the structure in the residuals to identify an appropriate error model. In contrast

> "a [mechanistic] model can be constructed by looking at the structure of the system, by dividing the system into components, and by trying to understand the behaviour of the whole system in terms of the behaviour of the individual system components and their interaction with one another."

When the processes are understood by which the errors are generated these can be used to identify a mechanistic model for the data.

The distinction between the two approaches is a little arbitrary as Thornley admits:

> "It needs to be stressed that there is no clearly defined dividing line between the two methods, and it is usual for most modelling exercises to contain both empiricism and mechanism in varying admixtures. It is more a matter of emphasis."

Certainly in this thesis there is a lot of empiricism in the mechanistic models. However, the labels serve as a useful distinction between two groups of chapters in the thesis.

1.2.2 Empirical models

With regard to the empirical approach, it is impossible for a single series of observations to be of any help whatsoever in choosing among all possible error variance matrices because, in the words of Bartlett (1978, p264),

> "dependence has so many more possibilities a priori than independence."

Therefore gross assumptions have to be made to restrict the choice. For serially-structured data it may be reasonable to make the assumption, common in time series analysis, that the error process is stationary, or at least stationary in its first and second moments, in which case the correlations between errors depends solely on the time separation between them. The choice then is between modelling the autocorrelation function or modelling its Fourier transform, the spectrum. In chapter 4, the correlations between errors in a regression model based on serially-observed data are modelled, beyond a particular fixed time separation, by a sum of exponentials. In chapter 5, the spectrum of the error process is modelled instead, an approach inspired by Robinson (1978).

1.2.3 Mechanistic models

There are vast numbers of mechanistic models of which three particular types are considered in chapters 6, 7 and 8. In chapter 6, linear stochastic difference and differential equations are used: a first-order model is used to describe the variations in a cow's milk yield over its lactation, developing from the work of Dhanoa and Le Du (1982); and a second-order model is used to describe the growth in an animal's weight. This aproach to modelling growth was stimulated by Sandland and McGilchrist's (1979) paper in a special review issue of Biometrics. A second paper in the same issue (Matis and Wehrly, 1979) prompted the use of stochastic compartment models in chapter 7. These are used to model pharmacokinetic data and data on transport systems within a leaf. Finally, in chapter 8 a related model is used to describe cumulative seed germination counts.

1.2.4 Data sets

Each of chapters 4 to 9 consists of initial sections of mathematical development followed by applications to real data. The techniques have been illustrated and explored using six data sets. Three sets consist of single series of observations and have been used in chapters 4 and 9:

a) Colquhoun's data, 124 observations on the relaxation current flowing through the end-plate membrane of a muscle fibre following a voltage jump, modelled by a sum of exponentials (Colquhoun, 1978);

b) Dale's data, 138 observations on the radioactive emission from a wheat leaf after it has been fed ^{14}C , modelled by a sum of exponentials (Bauermeister, Dale, Williams and Scobie, 1980);

c) Bruce's data, 200 observations on the daily energy demand of a mechanical model of a suckler cow, modelled by non-linear multiple regression (Burnett and Bruce, 1978; Bruce, 1980).

Also, Colquhoun's data have been used in chapters 5 and 7 and Dale's data in chapter 7. The other three data sets are multiple series of observations:

d) Neilson's data, between 32 and 44 observations on the milk yields of each of 23 cows (Neilson, Whittemore, Lewis, Alliston, Roberts, Hodgson-Jones, Mills, Parkinson and Prescott, 1983);

e) Thiessen's data, between 82 and 133 observations on the weights of each of 5 cattle (Thiessen, Hnizdo, Maxwell, Gibson and Taylor, 1984);

f) Naylor's data, between 7 and 12 observations on the germination counts of seeds in each of 33 petri-dishes (Hunter, Glasbey and Naylor, 1984).

Neilson's and Theissen's data have been used in chapter 6 and Naylor's data in chapter 8.

All the regression functions used are non-linear in their parameters but this is not particularly restrictive since, once correlations are introduced, any linearity in the model is automatically lost. As a consequence, estimation has to proceed iteratively and variances of estimators are at best an approximation.

1.3 Review of literature

1.3.1 Tests for autocorrelation

The standard test for uncorrelated errors against the alternative of serially correlated errors, when fitting regression models, is based on the Durbin-Watson statistic (Durbin and Watson, 1950, 1951, 1971). This is a function

$$\sum_{i=2}^{n} (\hat{e}_i - \hat{e}_{i-1})^2 \Big/ \sum_{j=1}^{n} \hat{e}_j^2$$

of the n-vector of residuals, $\hat{\underset{\sim}{e}}$, and is approximately equal to

$$2(1 - \hat{\rho}_1)$$

where $\hat{\rho}_1$ is the sample autocorrelation of the residuals at lag one (in other words it is the average correlation between adjacent residuals). The test employs small-sample significance levels

which take account of the correlation among the residuals induced by regression fitting, and has been widely used, particularly in econometrics. Johnston (1972, pp249-258) discussed it and alternatives. Weber and Monarchi (1982) reviewed more recent work.

A deficiency of the Durbin-Watson test is that it tests only for correlation at lag one, and thus may miss correlation at other lags. If the sample size is sufficiently large for the effect of the regression fitting to be ignored, then the approximate distribution of the sample autocorrelations for stationary processes is well understood, see for example Box and Jenkins (1976, pp34-36). In particular, if errors are independent then the sample autocorrelations of the residuals are asymptotically independent normal deviates, each with mean zero and variance 1/n. Therefore the assumption of independence can be tested. Various statistics have also been developed to help choose from among a class of linear processes called autoregressive-moving average models (ARMA models). A recent development is to introduce an array of statistics termed generalized partial autocorrelations (Glasbey, 1982). These will be described further in chapter 4.

If the model that is fitted includes lagged values of the dependent variate, then this affects the sampling distribution of the residual autocorrelations. Box and Pierce (1970) derived the asymptotic distribution of the autocorrelations of ARMA residuals and Pierce (1971a) showed the results still to be valid if a regression function is fitted as well as the ARMA model. The portmanteau statistic which they proposed for testing model inadequacy, essentially the sum of squares of the sample

autocorrelations at low lags, was later modified by Ljung and Box (1978). They changed the weights given to different terms in the sum of squares statistic in order to improve the rate of convergence to an asymptotic χ^2 distribution. Ansley and Newbold (1979) recommended instead the examination of autocorrelations at each lag separately.

In a related situation, Ripley (1981, pp98-101) discussed statistics for testing for uncorrelated errors where data are spatially, that is two-dimensionally, structured rather than serially, that is one-dimensionally, structured.

1.3.2 Duality between regression misspecification and correlated errors

Once it has been identified that residuals are autocorrelated the question arises as to how the model is to be changed to accommodate them. The choice is between changing the regression function and specifying a correlated errors structure, but the duality between these alternatives has received little attention in the literature. When noticeably correlated residuals have been encountered in situations where there has been no a priori reason to assume the errors to be independent (such as with serially-structured data) it has been conventional to attribute all lack of fit to correlation in the errors. A noteworthy exception is the attention given by Chatterjee and Price (1977, pp123-142) to the problem. They devoted a chapter to it, gave an example of correlated residuals which were explained by an omitted explanatory variable, and in general advocated the introduction of correlated errors into a model only as an action of last resort.

The duality can appear in many guises, two particular instances are: Blight and Ott (1975) used the duality when they described the departures of a polynomial approximation from an exact function by a first-order autoregressive (AR(1)) process; and Woodward and Gray (1983) showed that an eleventh-order autoregressive (AR(11)) process and a deterministic sinusoid together with a second-order autoregressive error process are almost equivalent models of the famous lynx data. In discussing the approximate duality between the two models they expressed a preference for the AR(11) process because of a lack of any justification for fitting a sinusoid. However, one could equally well argue the lack of justification for fitting an AR(11) process!

1.3.3 Effects of correlated errors on least-squares estimators

If it is decided that correlated residuals are caused by correlated errors then one pertinent question is "How poor are the least-squares estimators?", because if little efficiency has been lost then a more complicated analysis may be unnecessary. The efficiency of regression parameter estimators depends on both the design matrix $\underset{\sim}{X}$ and the error variance matrix $\tau^2\underset{\sim}{V}$. A common assumption is that $\underset{\sim}{V}$ is derived from a first-order autoregressive (AR(1)) process with parameter ϕ, where $|\phi| < 1$ (therefore $V_{ij} = \phi^{i-j}$ if $i \geq j$), and $\underset{\sim}{X}$ is a known single column that follows an independent AR(1) process with parameter x, where $|x| < 1$. The efficiency of the least-squares regression parameter estimator $\hat{\beta}$, that is the ratio of the variance of the generalized least-squares estimator to the variance of the least-squares estimator, is

$$\frac{(1 - \phi^2)(1 - \phi x)}{(1 + \phi^2 - 2\phi x)(1 + \phi x)} ,$$

see for example Malinvaud (1980, pp510-513). Provided ϕ lies between 0 and 0.5 the efficiency remains above 58%, although as ϕ approaches 1 the efficiency drops to 0%.

This basic approach is open to many generalizations, for instance in recent work Chipman (1979) and Krämer (1980) have considered more general design matrices $\underset{\sim}{X}$, but the same AR(1) error process as above. Chipman showed that the ordinary least-squares estimators of the regression parameters in a simple linear trend are never less than 75% efficient provided that ϕ is positive.

More generally, the least-squares estimator is equivalent to the generalized least-squares estimator, and is thus fully efficient, if and only if the design matrix is expressible as

$$\underset{\sim}{X} = \underset{\sim}{A}\,\underset{\sim}{B} ,$$

where $\underset{\sim}{A}$ is an n by m matrix consisting of m of the eigenvectors of $\underset{\sim}{V}$ and $\underset{\sim}{B}$ is a nonsingular m by m matrix. (This was proved in generality by Zyskind, 1967.) However, even when the estimator is efficient, the least-squares estimator of its variance will usually be biased. For example, in the above case of AR(1) processes the relative bias of the variance estimator is

$$[E(\hat{var}(\hat{\beta})) - var(\hat{\beta})]/var(\hat{\beta}) \simeq \frac{1}{n-1}\left[n - \frac{1 + \phi x}{1 - \phi x}\right]\frac{1 - \phi x}{1 + \phi x} - 1 ,$$

where n is the length of the series (Judge, Griffiths, Hill and Lee, 1980, p178). Therefore, in particular, if both ϕ and x are positive the variance estimator is downward biased.

Watson (1955) found the matrices $\underset{\sim}{X}$ which, conditional on $\underset{\sim}{V}$, gave the most extreme biases in least-squares estimators of the variances of the regression parameter estimators and the lowest efficiencies of regression parameter estimators; the bounds were expressible in terms of the eigenvalues of $\underset{\sim}{V}$. (Bloomfield and Watson (1975) gave a corrected proof of one theorem). Watson and Hannan (1956) applied these results to the situation of autoregressive-moving average processes by obtaining approximate expressions for the eigenvalues of $\underset{\sim}{V}$, but unfortunately the bounds they obtained were very wide. Tighter bounds can be derived by making more specific assumptions about $\underset{\sim}{X}$ and $\underset{\sim}{V}$, but there seems to be no end to the possibilities that can be considered. Judge, Griffiths, Hill and Lee (1980, pp174-179), in their book, give a more comprehensive review of the subject.

1.3.4 Empirical models

If errors are correlated, then one way of overcoming the deficiencies of ordinary least-squares estimation is to model the error covariance structure and re-estimate the regression parameters. There is a large literature on regression parameter estimation with empirically chosen error models. Cochrane and Orcutt (1949) proposed an iterative least-squares method for incorporating an AR(1) error process. Initially,

$$\sum_{i=2}^{n} \{(\underset{\sim}{y} - \underset{\sim}{X}\,\underset{\sim}{\beta})_i - \phi(\underset{\sim}{y} - \underset{\sim}{X}\,\underset{\sim}{\beta})_{i-1}\}^2$$

is minimized with respect to $\underset{\sim}{\beta}$ with ϕ set to zero, obtaining $\underset{\sim}{\beta}^{\{1\}}$, where $(\underset{\sim}{y} - \underset{\sim}{X}\,\underset{\sim}{\beta})_i$ denotes the ith coefficient in the n-vector $\underset{\sim}{y} - \underset{\sim}{X}\,\underset{\sim}{\beta}$. Then ϕ is estimated by minimizing the above sum of

squares with $\underset{\sim}{\beta}$ held constant at $\underset{\sim}{\beta}^{\{1\}}$, which gives $\phi^{\{1\}}$, and $\underset{\sim}{\beta}$ can be re-estimated with ϕ held constant at this value. After several repeats of this procedure, values converge to give final estimates of $\underset{\sim}{\beta}$ and ϕ. Anderson (1954) reviewed this and other early work. Rao and Griliches (1969) compared various procedures by Monte Carlo simulation and recommended Durbin's (1960) method which also has the advantage of being able to handle any order of autoregressive process, although Spitzer (1979) cast doubt on their results and found full maximum likelihood estimation preferable.

Developments in computers since the early 1960's have removed the need for ad hoc estimation procedures as an approximation to full maximum likelihood estimation and restrictions to linear regression with autoregressive error models. The result has been a series of papers exploiting different generalizations: Duncan and Jones (1966) allowed for a general stationary error process; Pierce (1971b, 1972) and Harvey and Phillips (1979) used ARMA error models; Gallant and Goebel (1976) and Glasbey (1980) handled non-linear regression functions. Another variation was the use of residual maximum likelihood (Cooper and Thompson, 1977) in which error parameters are estimated only in the sub-space of the data orthogonal to the regression model (this is specified further in section 2.3). In other papers, regression curves were fitted when errors were correlated, but with a particular example in mind. For example, Campbell and Walker (1977) fitted a sinusoid plus an AR(2) error model to the lynx data, Glasbey (1979) fitted generalized logistic curves with AR(1) errors to cattle weights, and Crowder and Tredger (1981) fitted exponentially damped polynomials with AR(1) errors to biological recovery data.

The above techniques draw heavily upon the time series literature, for which the fundamental text on ARMA models is Box and Jenkins (1976). Newbold (1981) recently reviewed the rapidly expanding literature. Nearly all of the empirical error modelling has been with observations equally spaced in time. This reflects a preoccupation in the time series literature because, although some interest has been shown in missing observations, for example Ljung (1982), very little has been done with other types of unequal spacings. Exceptions are the early paper by Quenouille (1958), and recent work on AR(1) processes (Robinson, 1977) and spectral analysis (Gaster and Roberts, 1977). In the case of spectral analysis this offers the oportunity of avoiding the aliasing associated with equally spaced observations. A related topic, observations of a continuous process at discrete times, has received some attention recently and been reviewed by Jones (1981). Phadke and Wu (1974) modelled sunspots data using a continuous analogue of an ARMA(2,1) process. Pandit and Wu (1975) studied the identifiability of parameters in second-order stochastic differential equations.

Another type of empirical error model has been encountered in regression with spatially correlated errors. This is a recent development (see for example Cliff and Ord, 1981, pp231-240) which has been used by Cook and Pocock (1983) in analysing geographical mortality data. Nearest-neighbour adjustment in the analysis of designed experiments is a special case of regression with correlated errors (Atkinson, 1969). There has been a lot of recent interest, see for example the paper by Wilkinson, Eckert, Hancock and Mayo (1983) and the ensuing discussion.

1.3.5 Growth models

One particular situation in which correlated errors are encountered is in the study of growth, which has had a long history, and data have been analysed in many ways. Two predominant strands in the literature have been biological and statistical. In the biological approach, non-linear regression functions have been proposed which supposedly come close to representing the true processes in growth, but the statistical techniques by which these curves are fitted to data are often quite inadequate. Whereas statistical methods have concentrated on good estimation procedures, but have inclined towards linear regression, in particular using polynomial functions which made no biological sense. A large number of references, primarily to the biological approach, are given by Kowalski and Guire (1974). A series of statistical papers on the analysis of multiple series probably reached its apotheosis with the work of Grizzle and Allen (1969). The relationship between biological and statistical models was discussed by Finney (1978).

Another statistical approach has been to model growth of individuals by means of stochastic difference or differential equations. First-order equations were considered by Finney (1958) and by Mitchell (1968). In a review article, Sandland and McGilchrist (1979) tried to unify the statistical and biological strands by incorporating stochastic variability into biologically-realistic differential equations and tying in recent work on population dynamics.

1.3.6 Stochastic compartment models

Deterministic compartment models have been in use for some considerable time, but stochastic generalizations are quite a recent development which have been reviewed from a statistical perspective by Matis and Wehrly (1979) and from a mathematical perspective by Purdue (1979).

The simplest and most commonly chosen form of stochasticity (termed P1 by Matis and Wehrly, 1979) is that arising from having only a finite number of particles in a system, each behaving independently with constant transition probability rates between compartments. The result is a Markov process in continuous time on a discrete state-space, see for example Cox and Miller (1965, pp178-186). Kodell and Matis (1976) considered the fitting of a two-compartment stochastic model which allowed emigration, although not immigration, and estimated the rate constants by generalized least-squares from the number of particles in each compartment at various times. The variance matrix, which took account of the correlations over time generated by the stochastic model, has a particular structure which simplified its inversion. Matis and Hartley (1971) had previously considered the fitting of a larger model by generalized least-squares, but using only the total number of particles in the system at different times. The method was only applicable to small data sets because the variance matrix had to be inverted explicitly. One reason for the use of generalized least-squares is that the exact distributions yield a very complicated likelihood function. An alternative approach, due to Lehoczky and Gaver (1977), is to approximate the stochastic model by a diffusion process, and then fit this to the data by exact maximum likelihood estimation.

There are a multitude of ways in which the above simple form of stochastic compartment model can be generalized, such as: allowing the transition rates to vary with time, removing the Markovian property, introducing dependence between particles and thus losing linearity. Invariably, in all these cases, the mathematical analysis is far in advance of statistical techniques for identifying and fitting appropriate models to observed data sets.

1.3.7 Model misspecification

Whether the error model has been chosen empirically or mechanistically, very little attention has been given to the consequences of the assumed model not being correct. Engle (1974) showed that if errors are from an AR(2) process but are assumed to be from an AR(1) process, then the ordinary least-squares regression estimators can be more efficient than the assumed maximum likelihood estimators. Patterson and Silvey (1980), faced with covariances of a non-serial nature when combining results from series of experiments, commented:

> "errors in the estimation of variance parameters can seriously affect the accuracy of estimated variety means. On the other hand, use of too simple a variance model can result in loss of efficiency and biased estimates of error."

Thus, there is a problem that has been recognised only occasionally and seemingly never solved.

1.3.8 Distribution of estimators

For a single series of correlated observations there is no true replication. Therefore the sample is only of size one and the asymptotic properties of maximum likelihood estimators (as

described, for example, by Cramér, 1946, pp500-506) are not applicable without changes in assumptions. For example, as Bartlett (1978, pp264-271) indicated, there are ways round the problem for stationary processes, such as the approach of Mann and Wald (1943) and Anderson (1971, pp188-211) for linear stochastic difference equations. However, the same properties do not necessary hold for non-stationary processes; for example White (1958) showed that for the explosive AR(1) process the least-squares estimator has a limiting Cauchy distribution.

When a regression term is also present the asymptotic properties depend on the form the regression takes with increasing sample size (Jennrich, 1969); for example, estimators of transient parameters will not converge to their true values as the sample size increases (Crowder, 1976). This highlights the arbitrariness of appealling to asymptotic properties: it is usually a totally artificial question to ask what form a regression function takes as the sample size increases beyond the actual number of observations; I could suppose it did anything that suited my purposes such as genuinely replicating (that is $f_{n+1} = f_1$, $f_{n+2} = f_2$, etc) and so avoid transient parameters. The real question, which has seldom been addressed theoretically, is whether the actual sample size is large enough for asymptotic results to be a good approximation.

So much for a general discussion; the pertinent issue here is the appropriateness of asymptotic theory of parameter estimators and likelihood ratio tests to the preceding models. These may be grouped into three categories:

a) Empirical models consisting of non-linear regressions with stationary error processes fitted by maximum likelihood or

asymptotically equivalent variants. These fall within the cases considered in the above discussion so results are asymptotically valid although, if observations are unequally spaced, their distribution in time must satisfy constraints (Robinson, 1977). For actual sample sizes the approximations may not be so good, and in particular distributions of non-linear parameter estimators may be skew (Ross, 1978).

b) If growth is modelled by non-stationary solutions of stationary differential equations, then long-term solutions will be stationary and so be of the same type as (a) except for start-up parameters which are only transient. But, as argued above, outside the bounds of the data, issues of this sort are artificial and to all intents and purposes we are in situation (a).

c) For stochastic compartment models Gaussian estimators (Whittle, 1961) are used instead of maximum likelihood estimators. There is a form of replication in these types of data because each particle in the system behaves independently. Chiang (1956) showed that regression parameter estimators approach asymptotic efficiency as the number of particles increases even if the number of observation times remains constant. However, these results do not apply to variance parameter estimators whose variances may include fourth order terms (see for example, Fisk, 1967).

2. Regression parameter estimation with generalized autoregressive-moving average errors

2.1 Introduction

In this and the following chapter the ground is prepared before the consideration of specific models in chapters 4 to 8 involving regression with serially correlated errors. Initially, a class of serially-structured variance matrices is defined and some simple mathematical properties are elucidated. Statistical and computational features are then considered for estimating regression parameters with error variances in this class.

In order to estimate regression parameters by maximum likelihood when the errors are correlated it is often necessary to handle a large variance matrix. The evaluation of elements in this matrix and repeated inversion of the matrix in the course of an iterative estimation procedure can be prohibitively expensive in computer time. For this reason a large body of literature has built up expounding methods of evaluating the likelihood for particular types of covariance structure without explicit use of the variance matrix. This chapter is a continuation of the tradition.

In section 2.2 a decomposition into lower-triangular band matrices is found for a large class of serially-structured variance matrices. This enables a wider class of covariance structures to be used than have previously been considered in the literature. The form of decomposition prompts the name "generalized autoregressive-moving average" (GARMA) processes. A large number

of serially-generated stochastic processes are demonstrated to be of this type including discrete realisations of a class of continuous processes. The decomposition is the key which enables the wide range of models considered in chapters 4 to 8 to be fitted to data.

In section 2.3 the use of the decomposition is set within the context of regression parameter estimation with GARMA errors. Choices of optimization criteria and algorithms and approximations to the variances of parameter estimators are also considered in this section.

Finally, in section 2.4 a computer program, REGAME, is described which implements the techniques of section 2.3. This program has been used to fit the models considered in chapters 4 to 8.

2.2 The generalized autoregressive-moving average structure

2.2.1 Definition of a GARMA (p,q) process

I will start with a rather technical definition. Its purpose will become clearer in [2.2.2] and its meaning should be clarified by the examples that follow.

A random vector $\underset{\sim}{e}$ of length n is defined to be a generalized autoregressive-moving average process of orders p and q, denoted GARMA (p,q), for non-negative integers p and q, if

$$\text{var}(\underset{\sim}{e}) = \tau^2 \underset{\sim}{V} ,$$

where τ^2 is a positive scaling parameter, $\underset{\sim}{V}$ is an $n \times n$ real, symmetric positive-definite matrix and there exist, possibly

complex, vectors $\underset{\sim}{\zeta}^{\{k\}}$ and $\underset{\sim}{\eta}^{\{k\}}$ of length n for $k=1,\ldots,p$ such that for $j = 1,\ldots,n$, if $p>0$, then

$$V_{ij} = \sum_{k=1}^{p} \zeta_i^{\{k\}} \eta_j^{\{k\}} \quad \text{for} \quad i > j+q-p+1 , \tag{2.2.1}$$

and if $p=0$, then

$$V_{ij} = 0 \quad \text{for} \quad i > j+q+1 .$$

This statement needs some explanation. The structure of $\underset{\sim}{V}$ may become clearer by taking as an example a GARMA(1,1) process with n equal to 5, then

$$\underset{\sim}{V} = \begin{bmatrix} V_{11} & \zeta_2\eta_1 & \zeta_3\eta_1 & \zeta_4\eta_1 & \zeta_5\eta_1 \\ \zeta_2\eta_1 & V_{22} & \zeta_3\eta_2 & \zeta_4\eta_2 & \zeta_5\eta_2 \\ \zeta_3\eta_1 & \zeta_3\eta_2 & V_{33} & \zeta_4\eta_3 & \zeta_5\eta_3 \\ \zeta_4\eta_1 & \zeta_4\eta_2 & \zeta_4\eta_3 & V_{44} & \zeta_5\eta_4 \\ \zeta_5\eta_1 & \zeta_5\eta_2 & \zeta_5\eta_3 & \zeta_5\eta_4 & V_{55} \end{bmatrix} .$$

The use of the name GARMA process will become apparent in [2.2.4]. It may be noted that diagonal, and some off-diagonal, elements of $\underset{\sim}{V}$ are undefined in equation (2.2.1) when q is greater than or equal to p, although they are implicitly constrained by the restriction that $\underset{\sim}{V}$ is positive-definite. Also, elements close to the diagonal are defined twice in equation (2.2.1) when q is less than $(p-1)$, so that for example

$$V_{i,i+1} = \sum_{k=1}^{p} \zeta_i^{\{k\}} \eta_{i+1}^{\{k\}} ,$$

but because $\underset{\sim}{V}$ is symmetric

$$V_{i,i+1} = V_{i+1,i} = \sum_{k=1}^{p} \zeta_{i+1}^{\{k\}} \eta_i^{\{k\}} ,$$

which places constraints on possible values for $\underset{\sim}{\zeta}^{\{k\}}$ and $\underset{\sim}{\eta}^{\{k\}}$ for $k = 1,\ldots,p$.

2.2.2 Decomposition of $\underset{\sim}{V}$

If $\underset{\sim}{V}$ is the variance matrix of a GARMA (p,q) process then there exist $n \times n$ real, lower-triangular, band matrices $\underset{\sim}{\Phi}$ and $\underset{\sim}{\Theta}$ with bandwidths (p+1) and (q+1) respectively such that

$$\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T = \underset{\sim}{\Theta}\ \underset{\sim}{\Theta}^T \quad .$$

(A lower-triangular band matrix with bandwidth $(\ell+1)$ has zero elements above the diagonal and after ℓ terms below the diagonal.)

To prove this, define $\underset{\sim}{\Phi}$ by:

a) $\Phi_{ii} = 1$ for $i = 1,\ldots,n$;

b) $\Phi_{i,i-\min(p,i-q-1)},\ldots,\Phi_{i,i-1}$ for $i = q+2,\ldots,n$,

are any real solution of the simultaneous linear equations

$$\text{(2.2.2)} \qquad \sum_{j=0}^{\min(p,i-q-1)} \Phi_{i,i-j}\ V_{i-j,i-\ell} = 0 \quad \text{for } \ell = (q+1),\ldots,(i-1);$$

c) $\Phi_{ij} = 0$ otherwise.

One point in this definition which requires clarification is equation (2.2.2) which consists of $(i-q-1)$ equations in $\min(p,i-q-1)$ unknowns, so if $i>p+q+1$ there are more equations than unknowns and a solution does not necessarily exist. However, j is less than or equal to p and ℓ is greater than q, therefore $(i-j)>(i-\ell)+q-p+1$ and equation (2.2.1) can be used to substitute for $V_{i-j,i-\ell}$ in equation (2.2.2) giving

$$\sum_{k=1}^{P} \eta_{i-\ell}^{\{k\}} \sum_{j=0}^{\min(p,i-q-1)} \Phi_{i,i-j}\ \zeta_{i-j}^{\{k\}} = 0 \quad \text{for } \ell=(q+1),\ldots,(i-1),$$

which is satisfied by any solution of

$$\sum_{j=0}^{\min(p,i-q-1)} \Phi_{i,i-j}\ \zeta_{i-j}^{\{k\}} = 0 \quad \text{for } k=1,\ldots,p \ .$$

This consists of p equations in $\min(p,i-q-1)$ unknowns and so does have a, possibly complex, solution which is also a solution of equation (2.2.2). But because $\underset{\sim}{V}$ is real the real part of the solution is also a solution of equation (2.2.2). Therefore a real, not necessarily unique, solution of equation (2.2.2) always exists.

To illustrate the definition of $\underset{\sim}{\Phi}$, consider the example of a GARMA(1,1) process with n equal to 5. $\underset{\sim}{\Phi}$ has a bandwidth of 2 and is defined to be

$$\underset{\sim}{\Phi} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -\zeta_3/\zeta_2 & 1 & 0 & 0 \\ 0 & 0 & -\zeta_4/\zeta_3 & 1 & 0 \\ 0 & 0 & 0 & -\zeta_5/\zeta_4 & 1 \end{bmatrix} .$$

Returning to the proof, from the definition, $\underset{\sim}{\Phi}$ is real and is also lower-triangular with bandwidth $(p+1)$ because

$$\Phi_{ij} = 0 \quad \text{when either } j>i \text{ or } j<i-p , \quad \text{for } i=1,\ldots,p .$$

When $j>q+1$

$$(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)_{i,i-j} = \sum_{\ell=i-n}^{i-1} (\underset{\sim}{\Phi}\,\underset{\sim}{V})_{i,i-\ell}\Phi_{i-j,i-\ell} = 0$$

because $\Phi_{i-j,i-\ell} = 0$ when $\ell<j$, so in particular when $\ell<q$, and $(\underset{\sim}{\Phi}\,\underset{\sim}{V})_{i,i-\ell} = 0$ when $\ell>q$ from (b) above. Therefore, because $(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)$ is also symmetric, it has a bandwidth of $(2q+1)$, meaning that after q terms off the diagonal all elements are zero. It is also positive-definite because $\underset{\sim}{V}$ and $\underset{\sim}{\Phi}$ are positive-definite, so by the Cholesky triangular decomposition there exists a real lower-triangular matrix $\underset{\sim}{\Theta}$ with bandwidth $(q+1)$ such that

$$\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T = \underset{\sim}{\Theta}\ \underset{\sim}{\Theta}^T \quad .$$

(See for example Wilkinson and Reinsch (1971, pp50-56).)

For the example already considered

$$(\underset{\sim}{\Phi}\ \underset{\sim}{V}) =$$

$$\begin{bmatrix}
V_{11} & \zeta_2\eta_1 & \zeta_3\eta_1 & \zeta_4\eta_1 & \zeta_5\eta_1 \\
\zeta_2\eta_1 & V_{22} & \zeta_3\eta_2 & \zeta_4\eta_2 & \zeta_5\eta_2 \\
0 & (\zeta_3\eta_2 - V_{22}\zeta_3/\zeta_2) & (V_{33} - \zeta_3^2\eta_2/\zeta_2) & \zeta_4(\eta_3 - \zeta_3\eta_2/\zeta_2) & \zeta_5(\eta_3 - \zeta_3\eta_2/\zeta_2) \\
0 & 0 & (\zeta_4\eta_3 - V_{33}\zeta_4/\zeta_3) & (V_{44} - \zeta_4^2\eta_3/\zeta_3) & \zeta_5(\eta_4 - \zeta_4\eta_3/\zeta_3) \\
0 & 0 & 0 & (\zeta_5\eta_4 - V_{44}\zeta_5/\zeta_4) & (V_{55} - \zeta_5^2\eta_4/\zeta_4)
\end{bmatrix}$$

which has zero elements after one term below the diagonal, and

$$(\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T) =$$

$$\begin{bmatrix}
V_{11} & \zeta_2\eta_1 & 0 & 0 & 0 \\
\zeta_2\eta_1 & V_{22} & (\zeta_3\eta_2 - V_{22}\zeta_3/\zeta_2) & 0 & 0 \\
0 & (\zeta_3\eta_2 - V_{22}\zeta_3/\zeta_2) & (\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T)_{33} & (\zeta_4\eta_3 - V_{33}\zeta_4/\zeta_3) & 0 \\
0 & 0 & (\zeta_4\eta_3 - V_{33}\zeta_4/\zeta_3) & (\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T)_{44} & (\zeta_5\eta_4 - V_{44}\zeta_5/\zeta_4) \\
0 & 0 & 0 & (\zeta_5\eta_4 - V_{44}\zeta_5/\zeta_4) & (\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T)_{55}
\end{bmatrix}$$

where $(\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T)_{ii} = V_{ii} - 2\zeta_i^2\eta_{i-1}/\zeta_{i-1} + V_{i-1,i-1}\zeta_i^2/\zeta_{i-1}^2$

for i=3,4,5,

which has a bandwidth of 3.

It is convenient at this point to note for future use that

$$|\underset{\sim}{\Phi}| = \prod_{i=1}^{n} \Phi_{ii} = 1 \ ,$$

and

$$|\underset{\sim}{\Theta}| = \prod_{i=1}^{n} \Theta_{ii} \quad .$$

Not only does the definition of a GARMA process in [2.2.1] imply that $\underset{\sim}{V}$ can be decomposed into lower-triangular band matrices, but also if $\underset{\sim}{V}$ can be so decomposed, then it is a GARMA

process, provided $\Phi_{i,i-p}$ is non-zero for $i > p+q+1$. Because this is of no practical importance only a brief sketch of the proof is given below.

If $(\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T)$ has a bandwidth of $(2q+1)$ then, because of the structure of $\underset{\sim}{\Phi}$, for $j = 1,\ldots,n$

$$(\underset{\sim}{\Phi}\ \underset{\sim}{V})_{i+p,j} = 0 \qquad \text{when } i+p > j+q+1 .$$

Therefore

$$V_{ij} = \frac{-1}{\Phi_{i+p,i}} \sum_{k=i+1}^{i+p} \Phi_{i+p,k}\ V_{kj}$$

when $i > \max(j+q-p+1, q+1)$

$$= \frac{-1}{\Phi_{i+p,i}} \left\{ \frac{-\Phi_{i+p,i+1}}{\Phi_{i+p+1,i+1}} \sum_{k=i+2}^{i+p+1} \Phi_{i+p+1,k}\ V_{kj} + \sum_{k=i+2}^{i+p} \Phi_{i+p,k}\ V_{kj} \right\}$$

by substituting for $V_{i+1,j}$,

$$= \sum_{k=n-p+1}^{n} A_{ik}\ V_{kj}$$

after repeated substitution, for some coefficients A_{ik} which do not depend on j . This is a restatement of equation (2.2.1) except for the case when $j < p$ and i lies between $(j+q-p+1)$ and q which requires special treatment and will not be considered here.

2.2.3 Numerical example of decomposition for GARMA(1,1) process

To illustrate the preceding algebra, suppose that $\underset{\sim}{e}$ is a vector of length 5 with a variance matrix

$$\underset{\sim}{V} = \begin{bmatrix} 1.0 & 0.5 & 0.333 & 0.25 & 0.2 \\ 0.5 & 2.0 & 0.667 & 0.5 & 0.4 \\ 0.333 & 0.667 & 3.0 & 0.75 & 0.6 \\ 0.25 & 0.5 & 0.75 & 4.0 & 0.8 \\ 0.2 & 0.4 & 0.6 & 0.8 & 5.0 \end{bmatrix} .$$

This matrix satisfies equation (2.2.1) with p and q equal to 1,

$$\underset{\sim}{\xi}^T = (1.0, \quad 0.5, \quad 0.333, \quad 0.25, \quad 0.2)$$

and $$\underset{\sim}{\eta}^T = (1.0, \quad 2.0, \quad 3.0, \quad 4.0, \quad 5.0).$$

Therefore $\underset{\sim}{e}$ is a GARMA(1,1) process.

From [2.2.2] $\underset{\sim}{\Phi}$ is defined to be

$$\begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -0.667 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.75 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.8 & 1.0 \end{bmatrix} ,$$

from which it follows that

$$(\underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^T) = \begin{bmatrix} 1.0 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.5 & 2.0 & -0.667 & 0.0 & 0.0 \\ 0.0 & -0.667 & 3.0 & -1.5 & 0.0 \\ 0.0 & 0.0 & -1.5 & 4.563 & -2.4 \\ 0.0 & 0.0 & 0.0 & -2.4 & 6.280 \end{bmatrix} ,$$

and

$$\underset{\sim}{\Theta} = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 1.323 & 0.0 & 0.0 & 0.0 \\ 0.0 & -0.504 & 1.657 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.905 & 1.935 & 0.0 \\ 0.0 & 0.0 & 0.0 & -1.24 & 2.177 \end{bmatrix} ,$$

which is a lower-triangular matrix with a bandwidth of 2 .

2.2.4 Interpretation of GARMA property

If $\underset{\sim}{e}$ is a GARMA (p,q) process then, defining an n-vector $\underset{\sim}{z}$ by

(2.2.3) $\quad \underset{\sim}{\Theta}\ \underset{\sim}{z} = \underset{\sim}{\Phi}\ \underset{\sim}{e}$,

it follows that

$$\begin{aligned} \mathrm{var}(\underset{\sim}{z}) &= \tau^2\, \underset{\sim}{\Theta}^{-1}\ \underset{\sim}{\Phi}\ \underset{\sim}{V}\ \underset{\sim}{\Phi}^{T}(\underset{\sim}{\Theta}^{-1})^{T} \\ &= \tau^2 \underset{\sim}{I} \ , \end{aligned}$$

where $\underset{\sim}{I}$ is the identity matrix of size n . Therefore

$$e_i = -\phi_{i,i-1}e_{i-1}\cdots-\phi_{i,i-p}e_{i-p}+\theta_{ii}z_i+\ldots+\theta_{i,i-q}z_{i-q}$$
$$\text{for } i=\max(p+1,q+1),\ldots,n,$$

where coefficients of $\underset{\sim}{z}$ are uncorrelated and have equal variance. When $\underset{\sim}{z}$ is normally distributed, this has the appearance of a finite realisation of an autoregressive-moving average process (see for example, Box and Jenkins, 1976 pp73-78), but with the non-constant parameters, hence the name "generalized autoregressive-moving average process".

I have found the analogy with an ARMA process helpful in understanding the property defined by equation (2.2.1), although in many applications the decomposition in [2.2.2] may be no more than a mathematical trick. Use of the name GARMA process is misleading to the extent that it may neither represent a finite sample from an infinite-dimensional process, as is the case for an ARMA process, nor a true generating mechanism.

If $\underset{\sim}{e}$ is a GARMA(1,0) process, then it has the second-moment property of a non-stationary Markov process because

$$e_i = -\phi_{i,i-1}e_{i-1} + \theta_{ii}z_i \qquad \text{for } i=2,\ldots,n \ ;$$

and if $\underset{\sim}{e}$ is a GARMA(p,q) process, with $q<p-1$, then it may in certain circumstances be represented as a sum of p Markov

processes, the kth of which has variance $\underset{\sim}{V}^{\{k\}}$ with

$$V_{ij}^{\{k\}} = \zeta_i^{\{k\}} \eta_j^{\{k\}} \qquad \text{for } i \geq j$$

provided these matrices are all real and positive-definite. This is because $\underset{\sim}{e}$ can be represented as

$$\underset{\sim}{e} = \sum_{k=1}^{p} \underset{\sim}{e}^{\{k\}}$$

where $\underset{\sim}{e}^{\{k\}}$ is a vector of length n, distributed with a variance matrix of $\underset{\sim}{V}^{\{k\}}$, and is hence a GARMA(1,0) process and a Markov process.

GARMA processes also have similarities with the linear processes used by control engineers. Harvey and Phillips (1979) have used this association in the case of ARMA processes to evaluate the likelihood by means of the Kalman filter, a technique developed by control engineers (see for example, Priestley, 1981b, pp807-815). The technique, which is also applicable to GARMA processes, involves giving a state-space representation of a process in terms of a multidimensional Markov process.

2.2.5 Examples

Any $n \times n$ real, symmetric positive-definite matrix $\underset{\sim}{V}$ can be expressed as

$$V_{ij} = \sum_{k=1}^{n} \zeta_i^{\{k\}} \eta_j^{\{k\}} \qquad \text{for } i \geq j$$

using its eigenvectors, and so is the variance matrix of a GARMA(n,n-1) process. Alternatively, and trivially, $\underset{\sim}{V}$ is the variance matrix of a GARMA(0,n-1) process because there are no elements V_{ij} for which $i > j+n$, so equation (2.2.1) does not apply. The examples below are not of this meaningless type; p and q take values close to zero.

a) In chapter 3 it will be shown that the solutions of linear stochastic difference and differential equations with constant coefficients are GARMA processes.

b) The autocorrelation coefficient at lag ℓ of an ARMA(p,q) process is

$$\rho_\ell = \sum_{k=1}^{p} \kappa_k e^{\ell\lambda_k} \qquad \text{for} \quad \ell > q-p+1 \ ,$$

where $\kappa_1, \ldots, \kappa_p$ and $\lambda_1, \ldots, \lambda_p$ are constants (see chapter 4). Therefore

$$V_{ij} = \rho_{i-j} = \sum_{k=1}^{p} (\kappa_k e^{i\lambda_k})(e^{-j\lambda_k}) \qquad \text{for} \quad i > j+q-p+1 \ ,$$

and so, as already alluded to in [2.2.4], the process is a GARMA(p,q) process. Also, non-stationary ARMA processes, ARMA processes with non-constant variances and ARMA processes with missing values are all GARMA processes.

c) In chapter 7 it will be shown that the solution of a type of stochastic compartment model is a GARMA process.

d) In chapter 8 it will be shown that the cumulative number of random events form a GARMA process.

e) To illustrate the variety of GARMA processes consider regression with random coefficients. If

$$(\underset{\sim}{e} | \underset{\sim}{\beta}) \sim (\underset{\sim}{X}\ \underset{\sim}{\beta},\ \sigma^2\ \underset{\sim}{I}) \ ,$$

which means that conditional upon $\underset{\sim}{\beta}$, $\underset{\sim}{e}$ is distributed with a mean of $\underset{\sim}{X}\ \underset{\sim}{\beta}$ and a variance matrix of $\sigma^2\underset{\sim}{I}$, where $\underset{\sim}{\beta}$ is a p-vector with

$$\underset{\sim}{\beta} \sim (\underset{\sim}{\alpha},\ \underset{\sim}{\Lambda}) \ ,$$

then $\operatorname{cov}(e_i, e_j) = \sum_{k=1}^{p} (X_{ik})\left(\sum_{\ell=1}^{p} \Lambda_{k\ell} X_{j\ell}\right) \qquad \text{for} \quad i > j+1$

and $\underset{\sim}{e}$ satisfies the definition in [2.2.1] and so is a GARMA(p,p) process.

2.2.6 Properties

If $\underset{\sim}{e}$ is a GARMA(p,q) process then many simple variates derived from $\underset{\sim}{e}$ are also GARMA processes. Examples include the following:

a) The moving average of length $\ell+1$, $\sum_{g=i-\ell}^{i} \mu_{i-g} e_g$, is a GARMA(p,q+$\ell$) process because

$$\operatorname{cov}\left(\left(\sum_{g=i-\ell}^{i} \mu_{i-g} e_g\right), \left(\sum_{h=j-\ell}^{j} \mu_{j-h} e_h\right)\right)$$

$$= \tau^2 \sum_{g=i-\ell}^{i} \sum_{h=j-\ell}^{j} \mu_{i-g}\mu_{j-h} \sum_{k=1}^{p} \zeta_g^{\{k\}}\eta_h^{\{k\}} \quad \text{if} \quad i-\ell > j+q-p+1$$

by substituting for the variance of $\underset{\sim}{e}$ from equation (2.2.1),

$$= \tau^2 \sum_{k=1}^{p} \left(\sum_{g=i-\ell}^{i} \mu_{i-g}\, \zeta_g^{\{k\}}\right)\left(\sum_{h=j-\ell}^{j} \mu_{j-h}\, \eta_h^{\{k\}}\right)$$

$$\text{if} \quad i > j+(q+\ell)-p+1.$$

This is a variance matrix of the same form as equation (2.2.1) but with q replaced by $(q+\ell)$, so the new process is a GARMA(p,q+ℓ) process.

b) The sum of elements in $\underset{\sim}{e}$, is a GARMA(p+1,q) process because

$$\operatorname{cov}\left(\left(\sum_{g=1}^{i} e_g\right), \left(\sum_{h=1}^{j} e_h\right)\right)$$

$$= \sum_{g=1}^{j+q-p} \sum_{h=1}^{j} V_{gh} + \tau^2 \sum_{g=j+q-p+1}^{i} \sum_{h=1}^{j} \sum_{k=1}^{p} \zeta_g^{\{k\}}\eta_h^{\{k\}} \quad \text{if} \quad i > j+q-p$$

by substituting for the variance of $\underset{\sim}{e}$ from equation (2.2.1)

$$= \sum_{g=1}^{j+q-p} \sum_{h=1}^{j} V_{gh} + \tau^2 \sum_{k=1}^{p} \left(\sum_{g=1}^{i} \zeta_g^{\{k\}} - \sum_{g=1}^{j+q-p} \zeta_g^{\{k\}}\right)\left(\sum_{h=1}^{j} \eta_h^{\{k\}}\right)$$

$$= \sum_{g=1}^{j+q-p} \sum_{h=1}^{j} \left(V_{gh} - \tau^2 \sum_{k=1}^{p} \zeta_g^{\{k\}}\eta_h^{\{k\}}\right) + \tau^2 \sum_{k=1}^{p} \left(\sum_{g=1}^{i} \zeta_g^{\{k\}}\right)\left(\sum_{h=1}^{j} \eta_h^{\{k\}}\right)$$

$$\text{if} \quad i > j+q-(p+1)+1$$

by algebraic manipulation. This is the sum of one term in j alone and p terms which are products of single terms in i and j, and thus satisfies equation (2.2.1) for a GARMA(p+1,q) process.

c) If $\underset{\sim}{e}$ is a GARMA(p,q) process and $\underset{\sim}{w}$ is a GARMA(u,v) process uncorrelated with $\underset{\sim}{e}$ and of the same length, then $\underset{\sim}{e} + \underset{\sim}{w}$ is a GARMA(p+u,max(q+u,v+p)) process because

$$\mathrm{cov}(e_i+w_i,\ e_j+w_j) = \tau^2 \sum_{k=1}^{p+u} \zeta_i^{\{k\}}\eta_j^{\{k\}} \qquad \text{for} \quad i>j+\max(q-p+1,v-u+1).$$

This meets the conditions of equation (2.2.1) for a GARMA(p+u,max(q+u,v+p)) process.

d) With the same conditions pertaining as in (c) above $(\underset{\sim}{e}^T,\ \underset{\sim}{w}^T)^T$ is a GARMA(max(p,u),max(p,u)+max(q-p,v-u)) process because the covariance between element i and element j is of the form of equation (2.2.1) as the sum of max(p,u) terms when $i>j+\max(q-p,v-u)+1$.

2.2.7 Extension to continuous processes

The preceding results for discrete processes extend naturally to continuous processes. A random variable $Y(t)$, indexed by t over a continuous finite or infinite range, is defined to be a cGARMA(p,δ) process for non-negative integer p and non-negative real δ if

$$(2.2.4) \qquad \mathrm{cov}(Y(s),Y(t)) = \tau^2 \sum_{k=1}^{p} \zeta^{\{k\}}(s)\eta^{\{k\}}(t) \qquad \text{for} \quad s>t+\delta\ ,$$

where $\zeta^{\{k\}}(t)$ and $\eta^{\{k\}}(t)$ for $k=1,\ldots,p$ are, possibly complex, functions over the same continuous range of t as $Y(t)$.

If times $t_1,t_2,\ldots,t_n$ ranked in ascending order are independent of $Y(t)$, then the sequence $Y(t_1),Y(t_2),\ldots,Y(t_n)$

in discrete time forms a GARMA(p,p+ℓ) process where ℓ is the smallest integer such that

$$t_{i+\ell+1} - t_i > \delta \qquad \text{for} \quad i = 1,\ldots,(n-\ell-1),$$

because

$$\operatorname{cov}(Y(t_i),Y(t_j)) = \tau^2 \sum_{k=1}^{p} \zeta^{\{k\}}(t_i)\eta^{\{k\}}(t_j)$$

$$\text{if} \quad i > j+(p+\ell)-p+1$$

and equation (2.2.1) is satisfied. For example, if δ equals zero, then ℓ is equal to -1 for any set of times and the sequence in discrete time is a GARMA(p,p-1) process.

2.2.8 Properties of a cGARMA process

If $Y(t)$ is a cGARMA(p,δ) process then, in a similar manner to the discrete case, many simple derived variates are also cGARMA processes. Examples include the following:

a) The filter of length ν, $\int_{s-\nu}^{s} \mu(s-u)Y(u)du$ is a cGARMA(p,$\delta+\nu$) process because

$$\operatorname{cov}\left(\left(\int_{s-\nu}^{s} \mu(s-u)Y(u)du\right),\left(\int_{t-\nu}^{t} \mu(t-v)Y(v)dv\right)\right)$$

$$= \tau^2 \sum_{k=1}^{p} \left(\int_{s-\nu}^{s} \mu(s-u)\zeta^{\{k\}}(u)du\right)\left(\int_{t-\nu}^{t} \mu(t-v)\eta^{\{k\}}(v)dv\right)$$

$$\text{if} \quad s-\nu > t+\delta ,$$

provided these integrals exist, by substituting for $\operatorname{cov}(Y(u),Y(v))$ from equation (2.2.4). This satisfies equation (2.2.4) for a cGARMA(p,$\delta+\nu$) process.

b) The integral of $Y(t)$, $\int_{o}^{s} Y(u)du$ is a cGARMA(p+1,δ) process because

$$\operatorname{cov}\left(\left(\int_0^s Y(u)du\right),\left(\int_0^t Y(v)dv\right)\right)$$

$$= \int_0^{t+\delta}\left(\int_0^t \operatorname{cov}(Y(u),Y(v))dv\right)du$$

$$+ \tau^2 \int_{t+\delta}^{s}\left(\int_0^t \sum_{k=1}^{p} \zeta^{\{k\}}(u)\eta^{\{k\}}(v)dv\right)du \qquad \text{if } s>t+\delta$$

by substituting for $\operatorname{cov}(Y(u),Y(v))$ from equation (2.2.4) when $u>t+\delta$,

$$= \int_0^{t+\delta}\left\{\int_0^t [\operatorname{cov}(Y(u),Y(v)) - \tau^2 \sum_{k=1}^{p} \zeta^{\{k\}}(u)\eta^{\{k\}}(v)]dv\right\}du$$

$$+ \tau^2 \sum_{k=1}^{p}\left(\int_0^s \zeta^{\{k\}}(u)du\right)\left(\int_0^t \eta^{\{k\}}(v)dv\right) ,$$

by algebraic manipulation. This is a sum of a term in t alone and p terms which are products of single terms in s and t and thus satisfies equation (2.2.4) for a GARMA$(p+1,\delta)$ process.

2.3 Parameter estimation

2.3.1 Model

In this section a very general form of model is considered which includes all the models considered in chapters 4 to 8 as special cases. An n-vector of observations $\underset{\sim}{y}$ is assumed to be a realisation of the model

$$\underset{\sim}{y} \sim (\underset{\sim}{f},\ \tau^2\underset{\sim}{V}) ,$$

where $\underset{\sim}{f}$ is a known n-vector except for an m-vector $\underset{\sim}{\beta}$ of parameters, τ^2 is a scaling parameter which is possibly known and $\underset{\sim}{V}$ is the variance matrix of a GARMA(p,q) process and is known except possibly for a few parameters.

2.3.2 Optimization criteria

If $\underset{\sim}{y}$ is multivariate normally distributed then its negative log-likelihood to within an additive constant is

$$L_M = \frac{n}{2} \ell n\ \tau^2 + \frac{1}{2}\ \ell n|\underset{\sim}{V}| + \frac{S}{2\tau^2} \quad ,$$

where S is the generalized sum of squares, that is

$$S = (\underset{\sim}{y}-\underset{\sim}{f})^T \underset{\sim}{V}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}) \quad .$$

The unknown parameters in the model can be estimated by minimizing L_M . This gives maximum likelihood estimates if $\underset{\sim}{y}$ is multivariate normally distributed, otherwise they have been called Gaussian estimates (Whittle, 1961).

A second estimation method, proposed by Cooper and Thompson (1977), is to minimize the residual negative log-likelihood which, to within an additive constant, is

$$L_R = (\frac{n-m}{2})\ell n\ \tau^2 + \frac{1}{2}\ \ell n(|\underset{\sim}{V}||\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X}||\underset{\sim}{X}^T\underset{\sim}{X}|^{-1}) + \frac{S}{2\tau^2} \quad ,$$

where $\underset{\sim}{X}$ is an n by m matrix defined by

$$\underset{\sim}{X} = \partial \underset{\sim}{f}/\partial \underset{\sim}{\beta}^T \quad .$$

With this approach, the parameters in $\underset{\sim}{V}$ are estimated only from the contrasts in $\underset{\sim}{y}$ orthogonal to the contrasts used to estimate the regression parameters $\underset{\sim}{\beta}$.

A third function that can be minimized is the sum of squares

$$L_S = \frac{n}{2}\ \ell n\ \tau^2 + \frac{S}{2\tau^2}$$

which omits $\ell n|\underset{\sim}{V}|$ totally.

The use of the first two of these optimization criteria will be investigated in subsequent chapters. The third criterion can not be used when $\underset{\sim}{V}$ has parameters additional to those in $\underset{\sim}{f}$ because then L_S may have no minimum.

Other estimation criteria exist, particularly for non-normally distributed data. When only first and second moments of a distribution are known, Wedderburn (1974) proposed optimizing the quasi-likelihood. However, this suffers from the same deficiency as L_S when $\underset{\sim}{V}$ has parameters additional to those in $\underset{\sim}{f}$.

2.3.3 Numerical evaluation of the optimization functions

The decomposition of $\underset{\sim}{V}$ in [2.2.2] can be used to evaluate the optimization functions without explicitly inverting $\underset{\sim}{V}$ or calculating its determinant. If the departures of the data from a regression function are denoted $\underset{\sim}{e}$, where

(2.3.1) $$\underset{\sim}{e} = \underset{\sim}{y} - \underset{\sim}{f} ,$$

then $$S = \underset{\sim}{z}^T \underset{\sim}{z} \quad \text{where} \quad \underset{\sim}{\Theta}\ \underset{\sim}{z} = \underset{\sim}{\Phi}\ \underset{\sim}{e} ,$$

and $\underset{\sim}{z}$ are called the residuals. Other terms required may also be evaluated relatively easily using

$$|\underset{\sim}{V}| = |\underset{\sim}{\Theta}|^2 |\underset{\sim}{\Phi}|^{-2} = \prod_{i=1}^{n} \theta_{ii}^2 ,$$

and $$\underset{\sim}{X}^T \underset{\sim}{V}^{-1} \underset{\sim}{X} = (\underset{\sim}{\Theta}^{-1} \underset{\sim}{\Phi}\ \underset{\sim}{X})^T (\underset{\sim}{\Theta}^{-1} \underset{\sim}{\Phi}\ \underset{\sim}{X}) .$$

This technique for evaluating the likelihood is a generalization of Ansley's (1979) method for ARMA processes.

2.3.4 Optimization algorithms

In general, the regression function will be non-linear in its parameters which can thus only be estimated iteratively. Even if the regression function is linear in its parameters the presence of correlations between observations will usually mean that the log-likelihood can only be optimized iteratively. The sole exception

to this rule is the scale parameter τ^2 , if unknown, which can be estimated directly, conditional on the other parameters, as

$$\hat{\tau}^2 = S/n$$

for L_M and L_S and as

$$\hat{\tau}^2 = S/(n-m)$$

for L_R . So it is necessary only to iterate on $\underset{\sim}{\beta}$ and the unknown parameters in $\underset{\sim}{V}$. Most, but not all, of the optimization functions can be re-expressed as sums of squares for which special optimization routines are available. For example

$$L_M(\hat{\tau}^2) = \frac{n}{2} - \frac{n}{2} \ln n + \frac{n}{2} \ln(S|\underset{\sim}{V}|^{1/n}) ,$$

which can be minimized by minimizing

$$S|\underset{\sim}{V}|^{1/n} = (|\underset{\sim}{V}|^{1/(2n)} \underset{\sim}{z})^T(|\underset{\sim}{V}|^{1/(2n)} \underset{\sim}{z})$$

which is a sum of squares. But L_M , with τ^2 known, cannot be expressed as a sum of squares. In order to apply a single algorithm in all circumstances a general derivative-free numerical optimization routine has been used (see [2.4.8]). Thus, some computational efficiency has been lost in certain problems in order to use the same optimization algorithm in all problems. Bard (1974, pp83-140) gives a general discussion of alternative algorithms.

2.3.5 Variances of parameter estimators

For the general model considered in this section it is not possible to appeal to asymptotic results in order to derive approximate variances of parameter estimators because the sample is only of size one. However, for certain types of correlation structure, for example stationary time series, asymptotic results have been found to be valid, as discussed in [1.3.8]. Approximate

variances of parameter estimators can thus be calculated by inverting the Hessian matrix, that is the matrix of second derivatives of the optimization function. The validity of this approximation for particular models will be discussed as they arise in later chapters.

2.3.6 Prediction of fit based on past observations

With serially correlated data it is of interest, in assessing the goodness-of-fit of a model, not merely to know the predicted value of the ith observation, that is $\hat{f}_i$, but also the predicted value conditional on all the earlier observations, denoted $(\hat{y}_i | y_1,\ldots,y_{i-1})$. This provides an assessment of fit taking account of the error correlation structure. From equations (2.2.3) and (2.3.1)

$$y_i = f_i - \sum_{j=1}^{\min(p,i-1)} \phi_{i,i-j}\, e_{i-j} + \sum_{k=0}^{\min(q,i-1)} \theta_{i,i-k}\, z_{i-k} \qquad \text{for } i = 1,\ldots,n .$$

Conditional upon $(y_i,\ldots,y_{i-1})$, $(e_i,\ldots,e_{i-1})$ is predicted to be $(\hat{e}_i,\ldots,\hat{e}_{i-1})$ and $(z_i,\ldots,z_{i-1})$ is predicted to be $(\hat{z}_i,\ldots,\hat{z}_{i-1})$, where

$$\hat{\underset{\sim}{e}} = \underset{\sim}{y} - \hat{\underset{\sim}{f}}$$

and

$$\underset{\sim}{\theta}\, \hat{\underset{\sim}{z}} = \underset{\sim}{\phi}\, \hat{\underset{\sim}{e}} , \tag{2.3.5}$$

and because z_i is uncorrelated with observations before i its minimum variance unbiased linear predictor is zero. Therefore

$$(\hat{y}_i | y_1,\ldots,y_{i-1}) = \hat{f}_i - \sum_{j=1}^{\min(p,i-1)} \phi_{i,i-j}\hat{e}_{i-j} + \sum_{k=1}^{\min(q,i-1)} \theta_{i,i-k}\hat{z}_{i-k} \qquad \text{for } i=1,\ldots,n .$$

Numerically, this can be derived quite simply by substituting from the ith row of equation (2.3.5) to give

$$(\hat{y}_i | y_1 \ldots y_{i-1}) = \hat{f}_i + \hat{e}_i - \theta_{ii}\hat{z}_i$$

$$= y_i - \theta_{ii}\hat{z}_i \qquad \text{for } i = 1,\ldots,n .$$

Therefore, once a model has been fitted to data the predicted fit, taking account of correlation, can be evaluated straightforwardly.

2.3.7 Simulation

Another useful feature of the GARMA property is that it makes it relatively easy to simulate $\underset{\sim}{y}$ from the model

$$\underset{\sim}{y} \sim N(\underset{\sim}{f}, \tau^2\underset{\sim}{V}) .$$

A random variate $\underset{\sim}{z}$ is generated with

$$\underset{\sim}{z} \sim N(\underset{\sim}{0}, \tau^2\underset{\sim}{I}) ,$$

which is then transformed to $\underset{\sim}{e}$ by

$$\underset{\sim}{\Phi}\,\underset{\sim}{e} = \underset{\sim}{\theta}\,\underset{\sim}{z} ,$$

and

$$\underset{\sim}{y} = \underset{\sim}{f} + \underset{\sim}{e} .$$

This has the desired properties because

$$E(\underset{\sim}{y}) = \underset{\sim}{f} ,$$

$$\text{var}(\underset{\sim}{y}) = \tau^2\, \underset{\sim}{\Phi}^{-1}\underset{\sim}{\theta}\,\underset{\sim}{\theta}^{T}\,(\underset{\sim}{\Phi}^{-1})^{T} = \tau^2\,\underset{\sim}{V} ,$$

and linear transforms of normal variates are also normal.

2.4 Computer program REGAME

2.4.1 Programming philosophy

The computer program REGAME, Regression Estimation with Generalized Autoregressive-Moving Average Errors, implements the methods described in sections 2.2 and 2.3. It is the program used

in chapters 4 to 8 to fit models to the data sets. The program is written in FortranIV with code as similar as possible to the preceding algebra. Some efficiency has been sacrificed in order to improve clarity in the program. For example, variables are transferred between routines in parameter lists wherever possible. An indicator variable can be set to generate the printing of all intermediate results in the program. The program has been designed so that the chances of total failure are minimized and warnings and failures are clearly signalled to the user. All variables are explicitly declared and the code is copiously annotated. The program has a modular structure both for simplicity and so that subroutines REGAMD and REGTRA can be used independently.

2.4.2 Program structure

The user must supply different subroutines REGMOD and REGDER for each model to be fitted. REGMOD evaluates $\underset{\sim}{f}$, and the first (p+q+1) elements on and below the diagonal of $\underset{\sim}{V}$, when supplied with values for the model parameters. Details are given in [2.4.4]. REGDER evaluates $\underset{\sim}{X}$ if the residual maximum likelihood optimization function is to be used. Details are given in [2.4.5]. In REGLIK the optimization function is evaluated using REGAMD to decompose $\underset{\sim}{V}$ into lower-triangular band matrices $\underset{\sim}{\Phi}$ and $\underset{\sim}{\Theta}$ (see [2.4.9], and using REGTRA to transform vectors using $\underset{\sim}{\Phi}$ and $\underset{\sim}{\Theta}$, for example to obtain $\underset{\sim}{z}$ (see [2.4.10]). A NAG library routine E04JBF (Numerical Algorithms Group, 1983), minimizes the optimization function (see [2.4.8]) and REGMON monitors and outputs the result of each iteration. At the

completion of the optimization, control returns to the main subroutine REGAME. REGVAR calculates variances of the parameter estimators by inverting the approximate Hessian matrix of the optimization function. A detailed output of results is produced by REGOUT to a secondary output channel NOUT10, optionally including line printer graphs printed by PLTLNG and PLTACR. Output suitable for input to another program, CEVOPE, which will be described in chapter 9, is produced by REGCVI to channels NOUT8 and NOUT9. Input and output channels are declared in a named common block /NINOUT/. Intermediate diagnostic output can be generated by MATPN1, MATPRO, MATPR1 and MATPR2. A flow chart is shown in figure 2.4.1. Three common blocks /REGCM1/, /REGCM2/ and /REGCM3/ pass parameters between REGAME and REGLIK which cannot be passed directly because E04JBF is in between. /REGCM2/ is also shared with REGMON and REGVAR and /REGCM3/ with REGVAR. The program is listed in Appendix A, excluding the NAG routines and the general output routines /NINOUT/, MATPN1, MATPRO, MATPR1, MATPR2, PLTLNG and PLTACR.

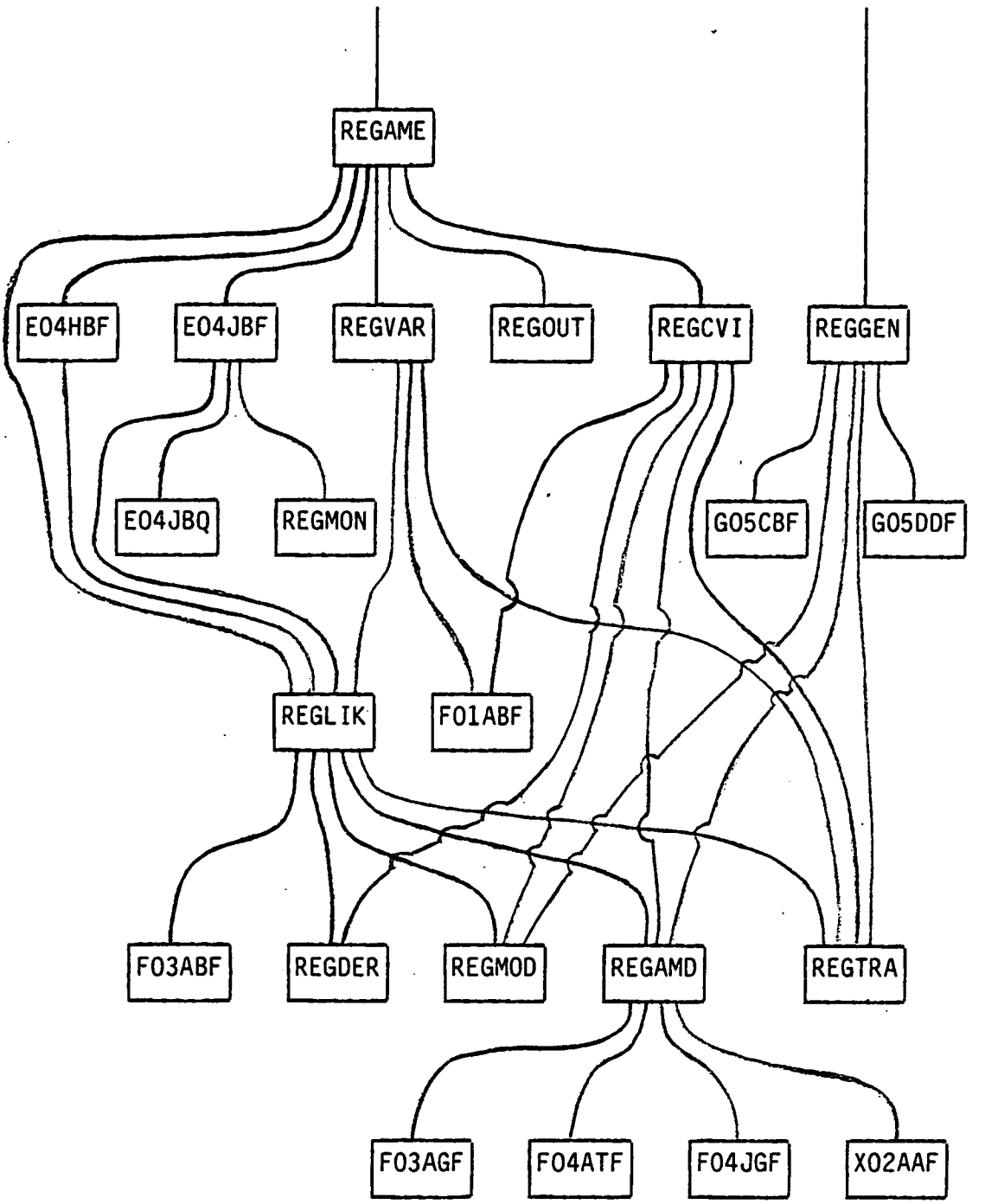
A separate main subroutine REGGEN is available to generate pseudo-random normal deviates from the first and second moments supplied by REGMOD.

2.4.3 Input/Output to REGAME

The user must write a computer program which calls subroutine REGAME. The variables to be transferred in the parameter list are specified below. Input parameters must be assigned values before the subroutine is called. On return to the calling program output parameters will have been assigned values by REGAME.

Figure 2.4.1

REGAME flow diagram (excluding general output routines)



Input

Y(NL)	vector of observations ($\underset{\sim}{y}$).
NL	number of observations (n), 200≥n≥1.
ACCTOL	accuracy to which parameters are to be estimated (explained in [2.4.8]), >0.
NCALMX	maximum number of function evaluations to be allowed.
NLIK	choice of optimization criterion: 1 for L_M, 2 for L_R, 3 for L_S.
NSCALE	set to 1 if τ^2 is to be estimated, otherwise τ^2 set to 1.
NM	number of regression parameters (m), 5≥m≥0 .
NDER	set to 1 if REGDER has been supplied, otherwise $\underset{\sim}{X}$ will be calculated by difference methods.
NVAR	set to 1 if variances of parameter estimators are to be estimated from second differences of the optimization function, otherwise the approximation supplied by E04JBF will be used.
NPARVR	on input this declares the first dimension of PARVAR, NPARVR≥NPAR + NSCALE.
NOUT	controls how detailed the output results are: 0 iterations, error messages and final estimates to NOUT6 (explained in [2.4.6]); 1 and to NOUT10 (explained in [2.4.7]); 2 and the fitted model to NOUT10; 3 and graphs of the fitted model to NOUT10; 4 and intermediate results from REGAME, REGVAR to NOUT6;

5 and intermediate results from REGLIK, REGAMD and REGTRA to NOUT6.

NCVI set to 1 if output required for input to CEVOPE.

Input and output

PARAM(16) vector of model parameters ordered so that regression parameters ($\underset{\sim}{\beta}$) precede other parameters, initial estimates input and final estimates (plus $\hat{\tau}^2$ if NSCALE=1) output.

NPAR number of model parameters input, on output this is increased by 1 if NSCALE=1. On input 15≥NPAR≥1.

SE(16) vector of approximate standard errors of parameter estimates, on input contains initial estimates of the standard errors (SE > 0), except where a parameter is to be held constant in which case SE=0, on output contains final estimates of standard errors including the standard error of $\hat{\tau}^2$ if NSCALE=1.

Output

XLIK final value of optimization function.

Z(NL) vector of residuals ($\underset{\sim}{\hat{z}}$).

PARVAR(NPARVR,NPARVR) matrix of approximate variances of parameter estimators with correlations below the diagonal.

IFAIL failure indicator:

0 optimization and estimation of variances of parameter estimators successful;

1 one of parameters input to REGAME or output by REGMOD or REGDER outside permitted range;

2 there is too much rounding error at the starting value of PARAM for E04HBF to work;

3 the number of function evaluations has exceeded NCALMX;

4 not all the conditions for an optimum have been satisfied but no better point could be found;

5 the problem is ill-conditioned because, in a local search, no optimum could be found;

10 optimization successful but variances of parameter estimators derived from E04JBF approximation because of ill-conditioning in the matrix of second differences or because NVAR $\neq$ 1.

11-15 failure 1-5 above and variances of parameter estimators only an approximation.

2.4.4 Input/Output to REGMOD

REGAME calls subroutine REGMOD with the parameter list specified below and input parameters assigned values. The user must write a subroutine REGMOD which assigns appropriate values to the output parameters.

Input

PARAM(NPAR)	vector of model parameters.
NPAR	number of parameters.
NLMAX	the first dimension of VAR.
NL	number of observations.
NFDIF	set to 1 if only F is required.

Output

F(NL)	vector of fitted values ($\underset{\sim}{f}$).

VAR(NLMAX,11) the first (p+q+1) elements on and below the diagonal of $\underset{\sim}{V}$, with $V_{i,i-\ell}$ in VAR(i,p+q+1-ℓ) for i=1,...,n, ℓ=0,...,min(p+q,i-1).

NPQA set to NP+NQ+1.

NP generalized autoregressive order (p), 5≥p≥0.

NQ generalized moving average order (q), 5≥q≥0.

NTEQ set to 1 if $\underset{\sim}{V}$ is such that all rows of $\underset{\sim}{\Phi}$ after the first (p+q) have the same elements before the diagonal (see [2.4.9]).

Control

NOUT an input to REGAME which controls level of output. If NOUT=0 error messages should only be output to NOUT6.

IFAIL set to 1 on output if $\underset{\sim}{f}$ or $\underset{\sim}{V}$ cannot be evaluated with the parameters set to PARAM.

2.4.5 Input/Output to REGDER

REGAME calls subroutine REGDER with the parameter list specified below and input parameters assigned values. The user must write a subroutine REGDER which assigns appropriate values to the output parameters.

Input

PARAM(NPAR) vector of model parameters.

NPAR number of parameters.

NLMAX the first dimension of DERIV.

NL number of observations.

NM number of regression parameters.

Output

DERIV(NLMAX,NM) matrix of first partial derivatives of F with respect to the first NM parameters in PARAM ($\underset{\sim}{X}$).

Control

NOUT an input to REGAME which controls level of output. If NOUT=0 error messages should only be output to NOUT6.

IFAIL set to 1 on output if $\underset{\sim}{X}$ cannot be evaluated with the parameters set to PARAM.

2.4.6 Output to channel NOUT6

To channel NOUT6 is output a history of the optimization iterations together with a brief summary of the final results. For interactive work this would conventionally be a computer user terminal.

Initially a heading is output, followed at each iteration by:

(i) Number of iterations.

(ii) Number of function evaluations.

(iii) Value of the optimization function.

(iv) Convergence criteria: change in the optimization function from the last iteration,
slope of the optimization function with respect to the parameters,
change in the parameters from the last iteration.

(v) A measure of conditioning: the ratio of largest to smallest elements on the diagonal of a lower-triangular decomposition

of a matrix of approximations to the second derivatives of the optimization function.

(vi) The symbol 'C' is printed if the change in the parameters has been constrained by the maximum size of step permitted.

(vii) Current estimates of the parameters.

After the iterations are completed a brief summary of the results is output:

(i) Value of IFAIL, the failure indicator.

(ii) Final value of the optimization function.

(iii) Final estimates of the parameters.

(iv) An estimate of the standard errors of the parameter estimators.

(v) Triangular array of correlations of parameter estimators.

2.4.7 Output to channel NOUT10

To channel NOUT10 a much more comprehensive set of results can be output. This would normally be directed to a computer disc file or a fast printer such as a line printer.

If NOUT=0 then there is no output to channel NOUT10. If NOUT>1 a heading is output together with settings for the input parameters:

NCALMX, NLIK, NSCALE, NM, NDER, ACCTOL, NVAR, NPEST, SE,

where NPEST is the number of parameters being estimated, in other words the number of non-zero coefficients in SE. The results at each iteration and upon completion are output as described in [2.4.6], together with the vector of first partial derivatives of the optimization function with respect to the parameters and the variances of the parameter estimators (PARVAR) with correlations exceeding 0.9 in magnitude highlighted by the symbol **.

If NOUT≥2, then in addition the following are also output:

$$p,\ q,\ \underset{\sim}{y},\ \underset{\sim}{\hat{f}},\ (\underset{\sim}{\hat{y}}|\text{past}),\ \underset{\sim}{\hat{z}}/\hat{\tau},\ \hat{\tau}^2\text{VAR},$$

with elements $\hat{z}_i/\hat{\tau}$ exceeding 2.5 in magnitude highlighted by the symbol ** .

If NOUT≥3, then line printer graphs are produced of:

(i) y_i, $\hat{f}_i$, and $(\hat{y}_i|y_1,\ldots,y_{i-1})$ plotted against i , for $i=1,\ldots,n$;

(ii) $y_i-\hat{f}_i$, $(\hat{y}_i|y_1,\ldots,y_{i-1})-\hat{f}_i$, and $\hat{z}_i/\hat{\tau}$ plotted against i , for $i=1,\ldots,n$;

(iii) $\hat{z}_i/\hat{\tau}$ plotted against $\hat{f}_i$, for $i=1,\ldots,n$;

(iv) the columns of $\hat{\tau}^2\text{VAR}$ plotted against i , for $i=1,\ldots,n$.

An example of output from REGAME is given in Appendix B, the background to which is described in section 4.3.

2.4.8 Optimization algorithm

A derivative-free quasi-Newton algorithm (Gill and Murray, 1972), implemented in the NAG library as routine E04JBF, is used to minimize the optimization function. A second routine, E04HBF, selects appropriate step lengths for calculating first derivatives by difference methods. The parameters and optimization function are scaled to lie approximately between -1 and 1. To achieve this the parameters are transformed about their initial estimates and scaled by the initial estimates of the standard errors. Similarly the optimization function is transformed about its initial value less the number of parameters and scaled by the number of parameters. If the algorithm is successful then on output the scaled parameter estimates are within a prespecified distance ACCTOL of their minimizing values.

Approximate second derivatives of the optimization function, with τ^2 replaced by $\hat{\tau}^2$, are output from the routine.

2.4.9 Algorithms in REGAMD

a) $\underset{\sim}{\Phi}$ is calculated by finding a solution, for i=q+2,...,n, of the simultaneous equations

$$\sum_{j=1}^{\min(p,i-q-1)} \Phi_{i,i-j} V_{i-j,i-\ell} = -V_{i,i-\ell}$$

for $\ell = (q+1),\ldots,\min(i-1,p+q)$,

using NAG routine F04ATF if the equations are of full rank and F04JGF otherwise. If the equations are of full rank, then the solution is unique and $\underset{\sim}{\Phi}$ satisfies equation (2.2.2). If the equations are not of full rank, then $\underset{\sim}{\Phi}$ may not satisfy equation (2.2.2); a warning is printed but the algorithm continues on the assumption that p has been overspecified. If NTEQ is set equal to 1 then, provided the equations are of full rank, they are only solved for i between (q+1) and (p+q+1), after which every row of $\underset{\sim}{\Phi}$ has identical elements.

b) The elements up to q below the diagonal of $\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T$ are calculated for $i = 1,\ldots,n$ by:

$$(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)_{i,i-\ell} = \sum_{k=0}^{\min(p,i-\ell-1,q-\ell)} (\underset{\sim}{\Phi}\,\underset{\sim}{V})_{i,i-\ell-k}\, \Phi_{i-\ell,i-\ell-k}$$

for $\ell = 0,\ldots,\min(q,i-1)$,

where $(\underset{\sim}{\Phi}\,\underset{\sim}{V})_{i,i-\ell-k} = \sum_{j=0}^{\min(p,i-1)} \Phi_{i,i-j} V_{i-j,i-\ell-k}$

for $\ell+k = 0,\ldots,p+q$.

c) $\underset{\sim}{\Theta}$ is obtained from $(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)$ by use of the NAG library routine F03AGF. This performs the Cholesky triangular

decomposition of a real positive-definite symmetric band matrix (Wilkinson and Reinsch, 1971, pp50-56).

2.4.10 Algorithms in REGTRA

Denote a real positive-definite lower-triangular matrix with band width ℓ by $\underset{\sim}{\Gamma}$ (this represents either $\underset{\sim}{\Phi}$ or $\underset{\sim}{\Theta}$), then transformations of vectors by $\underset{\sim}{\Gamma}$ can be derived as follows:

a) If $\underset{\sim}{w} = \underset{\sim}{\Gamma}\,\underset{\sim}{v}$

$$w_i = \sum_{k=\max(\ell+1-i,1)}^{\ell} \Gamma_{i,i-\ell+k} v_{i-\ell+k} \qquad \text{for } i=1,\ldots,n.$$

b) If $\underset{\sim}{w} = \underset{\sim}{\Gamma}^T \underset{\sim}{v}$

$$w_i = \sum_{j=i}^{\min(i+\ell-1,n)} \Gamma_{ji} v_j \qquad \text{for } i=1,\ldots,n.$$

c) If $\underset{\sim}{w} = \underset{\sim}{\Gamma}^{-1} \underset{\sim}{v}$

$$w_i = \Big(v_i - \sum_{k=\max(\ell+1-i,1)}^{\ell-1} \Gamma_{i,i-\ell+k} w_{i-\ell+k}\Big)/\Gamma_{ii} \qquad \text{for } i=1,\ldots,n.$$

d) If $\underset{\sim}{w} = (\underset{\sim}{\Gamma}^{-1})^T \underset{\sim}{v}$

$$w_i = \Big(v_i - \sum_{j=i+1}^{\min(i+\ell-1,n)} \Gamma_{ji} w_j\Big)/\Gamma_{ii} \qquad \text{for } i=n,n-1,\ldots,1.$$

3. The general solutions of linear stochastic difference and differential equations with constant coefficients

3.1 Introduction

This chapter serves the dual purposes of showing that a wide range of processes are either GARMA or cGARMA processes as defined in chapter 2, and specifying the general solutions of linear stochastic difference and differential equations with constant coefficients, special cases of which will be used in chapters 4 and 6. It is the second preparatory chapter prior to the use of specific models to fit real data in chapters 4 to 8.

The basic building blocks of many mathematical models of processes which evolve in time are difference equations or, their continuous analogues, differential equations. Among them the most mathematically tractable are those which are linear with constant coefficients. The general solutions of these types of equations are well understood. In the case of difference equations Brand (1966, pp373-412), for example, gave a general derivation of the solution when the input is deterministic. This extends in a straightforward manner to stochastic, that is random, input (see for example, Bartlett, 1978, pp162-165). Brand (1966, pp127-160) also gave a derivation of the general solution of deterministic differential equations. The extension to stochastic input is more difficult in this case, requiring the development of a stochastic calculus (see for example, Hoel, Port and Stone, 1972, pp111-170), although the apparent similarity with the deterministic solution is retained.

The general solutions of linear difference and differential equation are stated in sections 3.2 and 3.3 respectively. They are demonstrated to be solutions of the equations that satisfy the initial conditions, although the technicalities of stochastic calculus are not entered into. When the input process is a GARMA or a cGARMA process the output process is also shown to be a GARMA or a cGARMA process. The output could be used as the input to another linear equation for which the output would once again be a GARMA or cGARMA process. Thus, there is a wide range of models available to fit real data which are all GARMA or cGARMA processes. The similarities between sections 3.2 and 3.3 serve to emphasise the close connection between difference and differential equations.

Special cases of the models in this chapter are considered in later chapters: in chapter 4 stationary solutions alone are considered (a common restriction in the time series literature), whilst in chapter 6 first-order and second-order equations are used.

3.2 Difference equations

3.2.1 Model

A random vector $\underset{\sim}{Y}$ representing a finite or infinite sequence of variates $Y_0, Y_1, \ldots$, is defined to have been generated by the pth order linear stochastic difference equation

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + Z_t \qquad \text{for } t \geq p \,, \tag{3.2.1}$$

where $\phi_1,\ldots,\phi_p$ are constants, $\underset{\sim}{Z}$ is a random vector representing a finite or infinite sequence of variates $Z_0,Z_1,\ldots$, with known distribution, and $(Y_0,\ldots,Y_{p-1})$ also have a known distribution and are independent of $\underset{\sim}{Z}$. The starting point has been taken to be $t = 0$, but this does not incur any loss of generality because t could be redefined to be $t-t_0$ in order to initiate the process at t_0.

3.2.2 Solution

The general solution of equation (3.2.1) is

$$(3.2.2)\qquad Y_t = \sum_{i=1}^{p} Y_{i-1} \sum_{j=1}^{p} K_{ij} e^{t\lambda_j} + \sum_{k=p}^{t} Z_k \sum_{\ell=1}^{p} K_{p\ell} e^{(t+p-1-k)\lambda_\ell} \qquad \text{for } t \geq 0,$$

where $\lambda_1,\ldots,\lambda_p$ denote the elements in the p-vector $\underset{\sim}{\lambda}$, $e^{\lambda_1},\ldots,e^{\lambda_p}$ are the roots of the pth order polynomial

$$x^p - \phi_1 x^{p-1} \ldots - \phi_p = 0,$$

which are assumed to all be distinct, $\underset{\sim}{K}$ is a square matrix of size p defined to be the solution of the p^2 linear equations

$$\sum_{j=1}^{p} K_{ij} e^{t\lambda_j} = I_{i,t+1} \qquad \text{for } i = 1,\ldots,p,\ t = 0,\ldots,p-1,$$

and $\underset{\sim}{I}$ is the identity matrix of size p. The generality of this solution when $\underset{\sim}{Z}$ is deterministic has been proved by Brand (1966, pp373-412), who also gave the solution when the polynomial equation has some roots equal. This complication is more of theoretical than of practical concern because it is highly unlikely that roots will be estimated as being equal when real data are being modelled. Below I will confine myself to demonstrating that equation (3.2.2) is self-consistent for $t=0,\ldots,(p-1)$ and is a solution of equation (3.2.1) when $t \geq p$.

First of all, when $t<p$ the second term on the right side of equation (3.2.2) is zero because the upper limit in the k-summation is less than the lower limit, therefore

$$Y_t = \sum_{i=1}^{p} Y_{i-1} \sum_{j=1}^{p} K_{ij} e^{t\lambda_j} \qquad \text{for } t=0,\ldots,p-1 ,$$

$$= \sum_{i=1}^{p} Y_{i-1} I_{i,t+1} = Y_t ,$$

from the definition of $\underset{\sim}{K}$.

Therefore equation (3.2.2) makes sense when $t=0,\ldots,(p-1)$.

Define ϕ_0 to be -1 to simplify the notation then, from equation (3.2.2), for $t>p$ we get

$$-\sum_{g=0}^{p} \phi_g Y_{t-g} = -\sum_{g=0}^{p} \phi_g \sum_{i=1}^{p} Y_{i-1} \sum_{j=1}^{p} K_{ij} e^{(t-g)\lambda_j}$$

$$-\sum_{g=0}^{p} \phi_g \sum_{k=p}^{t-g} Z_k \sum_{\ell=1}^{p} K_{p\ell} e^{(t-g+p-1-k)\lambda_\ell} ,$$

$$= -\sum_{i=1}^{p}\sum_{j=1}^{p} Y_{i-1}K_{ij} e^{(t-p)\lambda_j} \{ \sum_{g=0}^{p} \phi_g e^{(p-g)\lambda_j} \}$$

$$-\sum_{k=p}^{t-p}\sum_{\ell=1}^{p} Z_k K_{p\ell} e^{(t-1-k)\lambda_\ell} \{ \sum_{g=0}^{p} \phi_g e^{(p-g)\lambda_\ell} \}$$

$$-\sum_{k=t-p+1}^{t} Z_k \sum_{g=0}^{t-k} \phi_g \sum_{\ell=1}^{p} K_{p\ell} e^{(t+p-1-g-k)\lambda_\ell} ,$$

by dividing the second term into two at $k=t-p$ and reordering the expressions. The first two terms are zero by the definition of $\underset{\sim}{\lambda}$ because the terms in curly brackets are no more than a re-expression of the polynomial of which $e^{\lambda_1},\ldots, e^{\lambda_p}$ are the roots. An extra term can be added on:

$$- \sum_{k=t-p+1}^{t-1} Z_k \sum_{g=t-k+1}^{p} \phi_g \sum_{\ell=1}^{p} K_{p\ell} e^{(t+p-1-g-k)\lambda_\ell}$$

which is zero by the definition of $K_{p\ell}$ because $p-2 \geq t+p-1-g-k \geq 0$. The result is that

$$- \sum_{g=0}^{p} \phi_g Y_{t-g} = - \sum_{k=t-p+1}^{t-1} Z_k \sum_{\ell=1}^{p} K_{p\ell} e^{(t-1-k)\lambda_\ell} \{ \sum_{g=0}^{p} \phi_g e^{(p-g)\lambda_\ell} \}$$

$$- Z_t \phi_o \sum_{\ell=1}^{p} K_{p\ell} e^{(p-1)\lambda_\ell}$$

$$= Z_t$$

because the first term is zero by the definition of $\underset{\sim}{\lambda}$ as above and the second term simplifies by the definition of $\underset{\sim}{K}$. What remains is a restatement of equation (3.2.1) for $t \geq p$ and thus equation (3.2.2) has been demonstrated to be a solution.

3.2.3 Relation to GARMA property

In this sub-section it will be proved that if the input vector of equation (3.2.1), $\underset{\sim}{Z}$, is a GARMA (a,b) process then the output vector, $\underset{\sim}{Y}$, is a GARMA(p+a,p+b-1) process. From equation (3.2.2)

$$(3.2.3) \quad \mathrm{cov}(Y_s,Y_t) = \sum_{i=1}^{p} \sum_{j=1}^{p} \mathrm{cov}(Y_{i-1},Y_{j-1}) \sum_{k=1}^{p} \sum_{\ell=1}^{p} K_{ik} K_{j\ell} e^{s\lambda_k} e^{t\lambda_\ell}$$

$$+ \sum_{i=p}^{s} \sum_{j=p}^{t} \mathrm{cov}(Z_i,Z_j) \sum_{k=1}^{p} \sum_{\ell=1}^{p} K_{pk} K_{p\ell} e^{(s+p-1-i)\lambda_k} e^{(t+p-1-j)\lambda_\ell} .$$

If $\underset{\sim}{Z}$ is a GARMA(a,b) process then

$$(3.2.4) \quad \mathrm{cov}(Z_i,Z_j) = \sum_{g=1}^{a} \zeta_i^{\{g\}} \eta_j^{\{g\}} \qquad \text{when } i \geq j+b-a+1$$

for appropriately chosen vectors $\underset{\sim}{\zeta}^{\{g\}}$ and $\underset{\sim}{\eta}^{\{g\}}$ of possibly

infinite length for $g = 1,\ldots,a$. Therefore, restricting attention to the case when $s>t+b-a$ we have

$$\sum_{i=p}^{s}\sum_{j=p}^{t} \operatorname{cov}(Z_i,Z_j)\, e^{-i\lambda_k} e^{-j\lambda_\ell}$$

$$= \sum_{i=p}^{t+b-a}\sum_{j=p}^{t} \operatorname{cov}(Z_i,Z_j) e^{-i\lambda_k} e^{-j\lambda_\ell}$$

$$+ \sum_{i=t+b-a+1}^{s}\sum_{j=p}^{t}\sum_{g=1}^{a} \zeta_i^{\{g\}}\eta_j^{\{g\}} e^{-i\lambda_k} e^{-j\lambda_\ell},$$

$$= \sum_{i=p}^{t+b-a}\sum_{j=p}^{t} [\operatorname{cov}(Z_i,Z_j) - \sum_{g=1}^{a} \zeta_i^{\{g\}}\eta_j^{\{g\}}] e^{-i\lambda_k} e^{-j\lambda_\ell}$$

$$+ \sum_{g=1}^{a}\{\sum_{i=p}^{s} \zeta_i^{\{g\}} e^{-i\lambda_k}\}\{\sum_{j=p}^{t} \eta_j^{\{g\}} e^{-j\lambda_\ell}\},$$

by substituting from equation (3.2.4) and reordering the expressions. Substituting for this in equation (3.2.3) when $s>t+b-a$ and reordering gives

$$\operatorname{cov}(Y_s,Y_t) = \sum_{k=1}^{p}\{e^{s\lambda_k}\}\{\sum_{i=1}^{p}\sum_{j=1}^{p} \operatorname{cov}(Y_{i-1},Y_{j-1}) \sum_{\ell=1}^{p} K_{ik}K_{j\ell} e^{t\lambda_\ell}\}$$

$$+ \sum_{k=1}^{p}\{e^{s\lambda_k}\}\{\sum_{i=p}^{t+b-a}\sum_{j=p}^{t}[\operatorname{cov}(Z_i, Z_j)$$

$$- \sum_{g=1}^{a} \zeta_i^{\{g\}}\eta_j^{\{g\}}] \sum_{\ell=1}^{p} K_{pk}K_{p\ell} e^{(p-1-i)\lambda_k} e^{(t+p-1-j)\lambda_\ell}\}$$

$$+ \sum_{g=1}^{a}\{\sum_{i=p}^{s}\sum_{k=1}^{p} \zeta_i^{\{g\}}K_{pk} e^{(s+p-1-i)\lambda_k}\}\{\sum_{j=p}^{t}\sum_{\ell=1}^{p} \eta_j^{\{g\}}K_{p\ell} e^{(t+p-1-j)\lambda_\ell}\}.$$

This is the sum of $(p+a)$ products of a term dependent only on s and a term dependent only on t provided $s>t+(p+b-1)-(p+a)+1$, therefore $\underset{\sim}{Y}$ is a GARMA(p+a, p+b-1) process.

The orders in the GARMA specification above are not necessarily the smallest possible. In particular, if $a = 0$ and $s \geq t+b+1$, then $\mathrm{cov}(Z_s, Z_t) = 0$ and, because Y_t depends only on Z_u for $u \leq t$, $\mathrm{cov}(Z_s, Y_t) = 0$. It follows from equation (3.2.1) that

$$\mathrm{cov}(Y_s, Y_t) = \sum_{i=1}^{p} \phi_i \, \mathrm{cov}(Y_{s-i}, Y_t) \qquad \text{when } s \geq t+b+1 .$$

Therefore there exist $\underset{\sim}{\xi}^{\{g\}}$ and $\underset{\sim}{\eta}^{\{g\}}$ for $g = 1,\ldots,p$ such that

$$\mathrm{cov}(Y_s, Y_t) = \sum_{g=1}^{p} \xi_s^{\{g\}} \eta_t^{\{g\}} \qquad \text{for } s \geq t+b-p+1$$

and $\underset{\sim}{Y}$ is a GARMA(p,b) process. One particular example of this is that when $\underset{\sim}{Z}$ is a stationary moving average process of order q the stationary solution of equation (3.2.1), if one exists, is not only a GARMA(p,q) process but also an ARMA(p,q) process.

3.3 Differential equations

3.3.1 Model

A random variate $Y(t)$, indexed by t over a continuous finite or infinite interval, is defined to have been generated by the linear stochastic differential equation

$$(3.3.1) \qquad Y^{\langle p \rangle}(t) - \sum_{i=1}^{p} \xi_i Y^{\langle p-i \rangle}(t) = Z^{\langle q \rangle}(t) - \sum_{j=1}^{q} \psi_j Z^{\langle q-j \rangle}(t) \qquad \text{for } t \geq 0 ,$$

where the terms in angular brackets denote the order of derivative with respect to t, $\xi_1,\ldots,\xi_p$, $\psi_1,\ldots,\psi_q$ are constants, $Z(t)$ is a random variate with known distribution over a continuous

interval, $(Y^{<0>}(0),\ldots,Y^{<p>}(0))$ also have a known distribution, and $q \leq p-1$. If $Z(t)$ is not stochastically differentiable q times with respect to t , as will often be the case for a random variate, then equation (3.3.1) makes no formal sense. This can be rectified by "integrating" both sides of the equation q times with respect to t (see Hoel, Port and Stone (1972, pp.162) for a specific example). The restriction to q being less than p is not necessary for equation (3.3.1) to define a valid process, but it is necessary for a general solution which does not include derivatives of $Z(t)$. The starting point has been taken to be t=0, but this does not incur any loss of generality because t could be redefined to be $t-t_0$ in order to initiate the process at t_0 .

3.3.2 Solution

The general solution of equation (3.3.1) is

$$(3.3.2) \quad Y(t) = \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij} e^{t\lambda_j} + \int_0^t Z(u) \sum_{k=1}^{p} \kappa_k e^{(t-u)\lambda_k} du$$

for $t \geq 0$

provided the integral exists, where

$$A^{\{i-1\}} = Y^{<i-1>}(0) - \sum_{h=0}^{i-2} Z^{<h>}(0) \sum_{k=1}^{p} \kappa_k \lambda_k^{i-2-h}$$

for $i = 1,\ldots,p$,

$\lambda_1,\ldots,\lambda_p$ denote the elements in the p-vector $\underset{\sim}{\lambda}$ and are the roots of the pth order polynomial

$$x^p - \xi_1 x^{p-1} \ldots - \xi_p = 0 ,$$

which are assumed to all be distinct, $\underset{\sim}{K}$ is a square matrix of size p defined to be the solution of the p^2 linear equations

$$\sum_{j=1}^{p} K_{ij}\lambda_j^t = I_{i,t+1} \qquad \text{for } i=1,\ldots,p \ , \ t=0,\ldots,p-1 \ ,$$

$\underset{\sim}{\kappa}$ is a vector of length p defined to be the solution of the p linear equations

$$\sum_{i=0}^{p-j} \xi_i \sum_{k=1}^{p} \kappa_k \lambda_k^{p-i-j} = \psi_{q+1-j} \qquad \text{for } j = 1,\ldots,p \ ,$$

$\underset{\sim}{I}$ is the identity matrix of size p, ξ_0 and ψ_0 are defined to be -1 and $\psi_\ell = 0$ if $\ell<0$. The same terms $\underset{\sim}{\lambda}$ and $\underset{\sim}{K}$ have been used as in section 3.2 to emphasise the similarity between difference and differential equations, although they are defined differently. The generality of the solution when $Z(t)$ is deterministic has been proved by Brand (1966, pp 127-160). Below I will once again confine myself to demonstrating that equation (3.3.2) is self-consistent in its first $(p-1)$ derivatives when $t=0$ and is a solution of equation (3.3.1) for $t > 0$.

From equation (3.3.2), by applying standard techniques of differentiation

$$Y^{\langle 1\rangle}(t) = \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij}\lambda_j \, e^{t\lambda_j}$$

$$+ Z(t) \sum_{k=1}^{p} \kappa_k + \int_0^t Z(u) \sum_{k=1}^{p} \kappa_k \lambda_k \, e^{(t-u)\lambda_k} \, du \ ,$$

and

$$Y^{\langle 2\rangle}(t) = \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij}\lambda_j^2 \, e^{t\lambda_j}$$

$$+ Z^{\langle 1\rangle}(t) \sum_{k=1}^{p} \kappa_k + Z(t) \sum_{k=1}^{p} \kappa_k \lambda_k + \int_0^t Z(u) \sum_{k=1}^{p} \kappa_k \lambda_k^2 \, e^{(t-u)\lambda_k} \, du.$$

By repeated differentiation

$$Y^{<g>}(t) = \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij}\lambda_j^g e^{t\lambda_j}$$

$$+ \sum_{h=0}^{g-1} Z^{<h>}(t) \sum_{k=1}^{p} \kappa_k \lambda_k^{g-1-h} + \int_o^t Z(u) \sum_{k=1}^{p} \kappa_k \lambda_k^g e^{(t-u)\lambda_k} du$$

for $g = 0,\ldots,p$.

To check the solution when $t=0$ for $g = 0,\ldots,p-1$

$$Y^{<g>}(0) = \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij}\lambda_j^g + \sum_{h=0}^{g-1} Z^{<h>}(0) \sum_{k=1}^{p} \kappa_k \lambda_k^{g-1-h} .$$

The first term is equal to $A^{\{g\}}$ by the definition of $\underset{\sim}{K}$.

Therefore

$$Y^{<g>}(0) = Y^{<g>}(0)$$

from the definition of $A^{\{g\}}$, and the first (p-1) derivatives of equation (3.3.2) make sense when $t=0$.

For $t > 0$

$$-\sum_{g=0}^{p} \xi_g Y^{<p-g>}(t) = -\sum_{g=0}^{p} \xi_g \sum_{i=1}^{p} A^{\{i-1\}} \sum_{j=1}^{p} K_{ij}\lambda_j^{p-g} e^{t\lambda_j}$$

$$-\sum_{g=0}^{p} \xi_g \sum_{h=0}^{p-g-1} Z^{<h>}(t) \sum_{k=1}^{p} \kappa_k \lambda_k^{p-g-1-h}$$

$$-\sum_{g=0}^{p} \xi_g \int_o^t Z(u) \sum_{k=1}^{p} \kappa_k \lambda_k^{p-g} e^{(t-u)\lambda_k} du ,$$

$$= -\sum_{i=1}^{p} \sum_{j=1}^{p} A^{\{i-1\}} K_{ij} e^{t\lambda_j} \{ \sum_{g=0}^{p} \xi_g \lambda_j^{p-g} \}$$

$$-\sum_{h=0}^{p-1} Z^{<h>}(t) \sum_{g=0}^{p-(h+1)} \xi_g \sum_{k=1}^{p} \kappa_k \lambda_k^{p-g-(h+1)}$$

$$-\int_o^t Z(u) \sum_{k=1}^{p} \kappa_k e^{(t-u)\lambda_k} \{ \sum_{g=0}^{p} \xi_g \lambda_k^{p-g} \} du ,$$

by reordering the expressions. The first and third terms are zero by the definition of $\underset{\sim}{\lambda}$, because the terms in curly brackets are no more than a re-expression of the polynomial of which $\lambda_1,\ldots,\lambda_p$ are the roots, and the second term, from the definition of $\underset{\sim}{\kappa}$, is

$$-\sum_{h=0}^{p-1} Z^{\langle h\rangle}(t)\ \psi_{q-h}\ .$$

What remains is a restatement of equation (3.3.1) for $t > 0$ and thus equation (3.3.2) has been demonstrated to be a solution.

3.3.3 Relation to cGARMA property

In this sub-section it will be proved that if the input variate, $Z(t)$, is a cGARMA(a,δ) process then the output variate, $Y(t)$, is a cGARMA$(p+a,\delta)$ process, provided that $Z(t)$ for $t > 0$ is distributed independently of $A^{\{0\}},\ldots,A^{\{p-1\}}$. This last restriction is rather arbitrary because, as $A^{\{0\}},\ldots,A^{\{p-1\}}$ are defined in terms of the derivatives of $Y(t)$ and $Z(t)$ at $t=0$, it is necessary for these derivatives to be correlated in order to ensure the independence condition. However, without the restriction the covariances of the derivatives of $Z(t)$ would be required to be known, not just the covariances of the process itself.

From equation (3.3.2), provided that $Z(t)$ for $t > 0$ is independent of $A^{\{0\}},\ldots,A^{\{p-1\}}$,

(3.3.3) $\operatorname{cov}(Y(s),Y(t)) =$

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\operatorname{cov}(A^{\{i-1\}},A^{\{j-1\}})\sum_{k=1}^{p}\sum_{\ell=1}^{p}K_{ik}K_{j\ell}\,e^{s\lambda_k}\,e^{t\lambda_\ell}$$

$$+\int_0^s\int_0^t \operatorname{cov}(Z(u),Z(v))\sum_{k=1}^{p}\sum_{\ell=1}^{p}\kappa_k\kappa_\ell\,e^{(s-u)\lambda_k}\,e^{(t-v)\lambda_\ell}\,dv\,du\ .$$

If $Z(t)$ is a cGARMA(a,δ) process, then

$$(3.3.4)\quad \mathrm{cov}(Z(u),Z(v)) = \sum_{g=1}^{a} \zeta^{\{g\}}(u)\ \eta^{\{g\}}(v) \qquad \text{for } u>v+\delta\ ,$$

for appropriately chosen functions $\zeta^{\{g\}}(t)$ and $\eta^{\{g\}}(t)$ specified over a continuous interval for $g=1,\ldots,p$. Therefore, restricting to the case when $s > t+\delta$,

$$\int_0^s\int_0^t \mathrm{cov}(Z(u),Z(v))e^{-u\lambda_k}\ e^{-v\lambda_\ell}\ dv\ du$$

$$= \int_0^{t+\delta}\int_0^t \mathrm{cov}(Z(u),Z(v))e^{-u\lambda_k}\ e^{-v\lambda_\ell}\ dv\ du$$

$$+ \int_{t+\delta}^s\int_0^t \sum_{g=1}^{a} \zeta^{\{g\}}(u)\ \eta^{\{g\}}(v)\ e^{-u\lambda_k}\ e^{-v\lambda_\ell}\ dv\ du\ ,$$

$$= \int_0^{t+\delta}\int_0^t \Big[\mathrm{cov}(Z(u),Z(v))- \sum_{g=1}^{a} \zeta^{\{g\}}(u)\ \eta^{\{g\}}(v)\Big]e^{-u\lambda_k}\ e^{-v\lambda_\ell}\ dv\ du$$

$$+ \sum_{g=1}^{a} \Big\{\int_0^s \zeta^{\{g\}}(u)e^{-u\lambda_k}\ du\Big\}\Big\{\int_0^t \eta^{\{g\}}(v)e^{-v\lambda_\ell}\ dv\Big\}\ ,$$

by substituting from equation (3.3.4) and reordering the expressions. Substituting this in equation (3.3.3) when $s > t+\delta$ and reordering gives

$$\mathrm{cov}(Y(s),Y(t)) = \sum_{k=1}^{p} \{e^{s\lambda_k}\}\Big\{\sum_{i=1}^{p}\sum_{j=1}^{p} \mathrm{cov}(A^{\{i-1\}},A^{\{j-1\}}) \sum_{\ell=1}^{p} K_{ik}K_{j\ell}\ e^{t\lambda_\ell}\Big\}$$

$$+ \sum_{k=1}^{p} \{e^{s\lambda_k}\}\Big\{\int_0^{t+\delta}\int_0^t \Big[\mathrm{cov}(Z(u),Z(v))- \sum_{g=1}^{a} \zeta^{\{g\}}(u)\ \eta^{\{g\}}(v)\Big] \sum_{\ell=1}^{p} \kappa_k\kappa_\ell\ e^{-u\lambda_k}\ e^{-(t-v)\lambda_\ell}\ dv\ du\Big\}$$

$$+ \sum_{g=1}^{a} \Big\{\int_0^s \zeta^{\{g\}}(u) \sum_{k=1}^{p} \kappa_k e^{(s-u)\lambda_k}\ du\Big\}\Big\{\int_0^t \eta^{\{g\}}(v) \sum_{\ell=1}^{p} \kappa_\ell e^{(t-v)\lambda_\ell}\ dv\Big\}\ .$$

This is the sum of $(p+a)$ products of a term dependent only on s and a term dependent only on t provided that $s > t+\delta$, therefore $Y(t)$ is a cGARMA$(p+a,\delta)$ process.

4. Stationary error processes: empirical autocorrelation parameterization

4.1 Introduction

If the departures of serially ordered observations from a fitted regression function are noticeably serially correlated, then either the regression function is inappropriate or the errors are not independent. In the latter case it is necessary to know the error covariances in order to estimate the regression parameters efficiently and obtain unbiased estimates of the variances of the parameter estimators. Except in exceptional circumstances where the mechanism causing the correlations is understood (see chapters 6, 7 and 8), these covariances will not be known and the data themselves must be used to identify an appropriate model. However, the range of potential models is too great to be considered unless some gross assumptions are made about the forms the correlations can take. A common choice in time series analysis and one that will be made in this and the following chapter, is that the error process is stationary. In this case the autocorrelation function fully describes the second-order properties of the errors.

In section 4.2 a particular class of models is presented in which the autocorrelations decay as a sum of exponentials after a certain time lag. Different ways of parameterizing this type of model are considered, including stationary solutions of the difference and differential equations in chapter 3. Model order identification is discussed for the special case of equally-spaced observations. In succeeding sections three data sets are used as

case studies of this approach to estimating regression parameters when errors are serially correlated. Maximum likelihood estimators and residual maximum likelihood estimators are compared. Finally, in section 4.6 the usefulness of the method is critically assessed.

The basic idea in this chapter, of identifying a stationary time series model from least-squares residuals and using it to re-estimate regression parameters, is not new, see [1.3.4]. The novelty lies in the choice of parametric model for the autocorrelation structure, the use of residual maximum likelihood estimators, and the application to particular data sets.

4.2 Theory

4.2.1 Model

The situation envisaged is where a sequence of observations $y_1,\ldots,y_n$ (denoted $\underset{\sim}{y}$) have been made at times $t_1,\ldots,t_n$ (denoted $\underset{\sim}{t}$) on a single experimental subject. It is assumed that the vector $\underset{\sim}{y}$ is a realisation from a multivariate normal distribution with mean $\underset{\sim}{f}$ and variance $\tau^2 \underset{\sim}{V}$, where

$$(\underset{\sim}{W}^{-1}\underset{\sim}{V}\,\underset{\sim}{W}^{-1})_{ij} = \rho(|t_i - t_j|) \qquad \text{for } i,j = 1,\ldots,n ,$$

$\underset{\sim}{W}$ is a diagonal matrix of weights used to standardize the variances of the errors and

$$\rho(u) = \sum_{\ell=1}^{p} \kappa_\ell e^{u\lambda_\ell} \qquad \text{for } u > \delta .$$

Therefore the autocorrelation function, $\rho(u)$, decays as the sum of p exponentials after a time lag of δ, with rates $\lambda_1,\ldots,\lambda_p$

(denoted $\underset{\sim}{\lambda}$) given weights $\kappa_1,\ldots,\kappa_p$ (denoted $\underset{\sim}{\kappa}$). At present, $\rho(u)$ will remain unspecified when $u<\delta$. One model order, p, must be a non-negative integer, but the other model order, δ, is unconstrained. However, if δ is negative then $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$ must be constrained to ensure that ρ remains symmetric about zero. It is immediately apparent that the definition in [2.2.1] is satisfied and $\underset{\sim}{y}$ is a GARMA(p,q) process with q dependent on the spacing of observation times as described in [2.2.7]. In particular, if $t_i = i$ for $i = 1,\ldots,n$, then $q = p+\delta-1$.

The parameters $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$ are constrained because $\underset{\sim}{V}$ is real, thus coefficients of $\underset{\sim}{\kappa}$ (and of $\underset{\sim}{\lambda}$) are either real or occur in complex-conjugate pairs. Also, $\underset{\sim}{V}$ must be positive-definite and this places a more complicated set of constraints on $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$. It is convenient to reparameterize $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$ in order to incorporate both these types of constraints. Alternative parameterizations also permit the addition of further constraints that may be deemed desirable. Four different choices will be considered in the following three subsections.

4.2.2 Parameterization by stationary solutions of stochastic difference equations

The parameterization by stationary difference equations is suitable only when the times between observations are equal or are all integer multiples of some basic time interval which can be taken to be unity without any loss of generality. The weight-corrected departures $(\underset{\sim}{a} = \underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}))$ are assumed to be the realisation at times $\underset{\sim}{t}$ of the discrete random vector $\underset{\sim}{Y}$ indexed by all integer times between t_1 and t_n, where $\underset{\sim}{Y}$ is generated

by the linear difference equation (3.2.1) with $\underset{\sim}{Z}$ a (q+1)-term moving average process of white noise denoted $\underset{\sim}{Z}^{\{w\}}$. Therefore,

$$(4.2.1) \qquad Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} - \sum_{j=0}^{q} \theta_j Z_{t-j}^{\{w\}} ,$$

where $\underset{\sim}{\phi}$ is the same vector of parameters as in equation (3.2.1) and $\underset{\sim}{\theta}$ is the vector of weights in the moving average process. This has been called an autogregressive-moving average process in the statistical literature.

It is possible to derive the autocorrelation coefficients of the stationary solution of equation (4.2.1) directly from equation (3.2.3) but is much simpler to use the method described by Box and Jenkins (1976, pp.74-75). From equation (4.2.1.)

$$(4.2.2) \qquad \mathrm{cov}(Y_s, Z_t^{\{w\}}) = \sum_{i=1}^{p} \phi_i \, \mathrm{cov}(Y_{s-i}, Z_t^{\{w\}}) - \sum_{j=0}^{q} \theta_j \, \mathrm{cov}(Z_{s-j}^{\{w\}}, Z_t^{\{w\}})$$

and

$$(4.2.3) \qquad \mathrm{cov}(Y_s, Y_t) = \sum_{i=1}^{p} \phi_i \, \mathrm{cov}(Y_{s-i}, Y_t) - \sum_{j=0}^{q} \theta_j \, \mathrm{cov}(Z_{s-j}^{\{w\}}, Y_t) .$$

Because $\underset{\sim}{Y}$ is stationary, $\mathrm{cov}(Y_s, Z_t^{\{w\}})$ and $\mathrm{cov}(Y_s, Y_t)$ are functions of $(s-t)$, but not of s or t separately. Equation (4.2.2) can be solved in turn for $t = s, s-1, s-2 \ldots$ to obtain $\mathrm{cov}(Y_s, Z_t^{\{w\}})$ because $\mathrm{cov}(Y_s, Z_t^{\{w\}}) = 0$ if $t>s$. Equation (4.2.3) can then be solved for $t = s, s-1, \ldots, s-p$ simultaneously to obtain $\mathrm{cov}(Y_s, Y_t)$ over this range of values of t, and then in turn for $t = s-p-1, s-p-2 \ldots$. In particular,

$$\mathrm{cov}(Y_s, Y_t) = \sum_{i=1}^{p} \phi_i \, \mathrm{cov}(Y_{s-i}, Y_t) \qquad \text{when} \quad s \geq t+q+1 .$$

Therefore the autocorrelation function of $\underset{\sim}{Y}$ is of the form

$$\rho(u) = \sum_{\ell=1}^{p} \kappa_\ell \, e^{u\lambda_\ell} \qquad \text{when} \quad u \geq q-p+1 ,$$

where $e^{\lambda_1},\ldots,e^{\lambda_p}$ are the roots of the pth order polynomial

$$x^p - \phi_1 x^{p-1} \ldots - \phi_p = 0 ,$$

which are all assumed to be distinct, and $\underset{\sim}{\kappa}$ is a vector of weights which can be expressed in terms of $\underset{\sim}{\phi}$ and $\underset{\sim}{\theta}$ by way of $\mathrm{cov}(Y_s, Y_t)$. The constraints on the parameters $\underset{\sim}{\phi}$ and $\underset{\sim}{\theta}$ necessary and sufficient to ensure $\underset{\sim}{V}$ is positive-definite, or equivalently that equation (4.2.1) has a stationary solution, are

$$|e^{\lambda_\ell}| < 1 \qquad \text{for} \quad \ell = 1,\ldots,p .$$

(See for example Box and Jenkins, 1976, pp 73-74.)

4.2.3 Parameterization by stationary solutions of stochastic differential equations

An alternative parametrization is in terms of stationary differential equations. There is no restriction to times between observations being integer multiples of some basic time interval. The random vector $Y(t)$ underlying the weight-corrected departures is assumed to be indexed over continuous time t and generated by the linear differential equation (3.3.1), where $Z(t)$ is the first derivative of a Weiner process. Strictly speaking, this definition of $Z(t)$ is nonsense because the Weiner process is nowhere differentiable, see for example Hoel, Port and Stone (1972, pp141-147). As already discussed in [3.3.1] this can be rectified by integrating both sides of equation (3.3.1) a certain number of times. Hoel, Port and Stone (1972, pp159-170) considered this equation in the restricted case when q is zero, and Phadke and Wu (1974) used the process with $p=2$ and $q=1$ to model annual sunspot numbers.

Examination of the general solution of equation (3.3.1) given in equation (3.3.2) reveals that the differential equation can have a stationary solution only when

$$|e^{\lambda_\ell}| < 1 \qquad \text{for} \quad \ell = 1,\ldots,p \ .$$

With this condition satisfied, the first double summation term in equation (3.3.2) decays to zero as t increases. Therefore, if the origin of the process is redefined to be in the infinite past, the solution is

$$(4.2.4) \qquad Y(t) = \int_{-\infty}^{t} Z(u) \sum_{k=1}^{p} \kappa_k^* e^{(t-u)\lambda_k} \, du \ ,$$

where $\underset{\sim}{\kappa}^*$ is used in place of $\underset{\sim}{\kappa}$ to distinguish it from another vector of parameters $\underset{\sim}{\kappa}$ which were introduced in [4.2.1] and will be used throughout this chapter. It follows that the autocovariance function of the process satisfies

$$\operatorname{cov}(Y(s),Y(t)) = \int_{-\infty}^{s}\int_{-\infty}^{t} \operatorname{cov}(Z(u),Z(v)) \sum_{k=1}^{p}\sum_{\ell=1}^{p} \kappa_k^*\kappa_\ell^* e^{(s-u)\lambda_k} e^{(t-v)\lambda_\ell} \, dv\,du,$$

$$= \sigma^2 \int_{-\infty}^{t} \sum_{k=1}^{p}\sum_{\ell=1}^{p} \kappa_k^*\kappa_\ell^* e^{(s-v)\lambda_k} e^{(t-v)\lambda_\ell} \, dv \qquad \text{for} \quad s \geq t \ ,$$

because $Z(u)$ and $Z(v)$ are uncorrelated when $u \neq v$,

$$= \sum_{k=1}^{p} \left(-\sigma^2 \sum_{\ell=1}^{p} \frac{\kappa_k^* \kappa_\ell^*}{\lambda_k+\lambda_\ell}\right) e^{(s-t)\lambda_k} \ .$$

Therefore the autocorrelation function of $Y(t)$ is

$$\rho(u) = \sum_{k=1}^{p} \kappa_k e^{u\lambda_k} \qquad \text{for} \quad u \geq 0 \ ,$$

where κ_k is proportional to $\sum_{\ell=1}^{p} \frac{\kappa_k^* \kappa_\ell^*}{\lambda_k+\lambda_\ell}$,

and $\underset{\sim}{\kappa}$ is normalized so that $\sum_{k=1}^{p} \kappa_k = 1$.

The constraints on the parameters $\underset{\sim}{\xi}$ and $\underset{\sim}{\psi}$ in equation (3.3.1) to ensure that a stationary solution exists and that $\underset{\sim}{V}$ is positive-definite are

$$|e^{\lambda_\ell}| < 1 \qquad \text{for} \quad \ell = 1,\ldots,p \quad .$$

The choice of parameterization by means of stationary solutions of stochastic differential equations is more restrictive on the range of possibilities for $\rho(u)$ than is the previous parameterization in [4.2.2]. In particular, δ is restricted to taking only the value zero, although this may be extended to the case where δ is an arbitrarily small positive number, by adding an independent white noise error to $Y(t)$. In this latter instance

$$\rho(u) = \sum_{k=1}^{p} \kappa_k e^{u\lambda_k} \qquad \text{for} \quad u > 0 \ ,$$

and

$$\sum_{k=1}^{p} \kappa_k < 1 \ .$$

An alternative model for $Y(t)$ is to define it by the weighted integral of a Weiner process given in equation (4.2.4). The autocorrelations remain as above but the reference to stochastic differential equations is avoided. This then serves as a third parameterization of the autocorrelation function.

4.2.4 Parameterization by a sum of positively-correlated Markov processes

A fourth parameterization is also applicable to any spacing of observation times. The weight-corrected departures can be considered to be the sum of p positively-correlated Markov processes if all elements in $\underset{\sim}{\kappa}$ are real and non-negative and all elements in $\underset{\sim}{\lambda}$ are real and non-positive. Because, then the

weight-corrected departures, $\underset{\sim}{a}$, can be expressed as

$$\underset{\sim}{a} = \sum_{\ell=1}^{p} \underset{\sim}{a}^{\{\ell\}} ,$$

where $\underset{\sim}{a}^{\{1\}},\ldots,\underset{\sim}{a}^{\{p\}}$ are positively-correlated Markov processes which are independent of one another and

$$\text{cov}(a_i^{\{\ell\}}, a_j^{\{\ell\}}) = \tau^2 \kappa_\ell e^{|i-j|\lambda_\ell} \qquad \text{for } i,j = 1,\ldots,n .$$

The above constraints on $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$ can be achieved by parameterizing with a (2p-1)-dimensional vector $\underset{\sim}{\alpha}$ where

$$\kappa_\ell = (\tfrac{1}{2} + \tfrac{1}{2} \sin \alpha_\ell) \prod_{k=1}^{\ell-1} (\tfrac{1}{2} - \tfrac{1}{2} \sin \alpha_k) \qquad \text{for } \ell = 1,\ldots,p-1,$$

$$\kappa_p = \prod_{k=1}^{p-1} (\tfrac{1}{2} - \tfrac{1}{2} \sin \alpha_k) ,$$

and $$e^{\lambda_\ell} = \tfrac{1}{2} + \tfrac{1}{2} \sin \alpha_{p+\ell-1} \qquad \text{for } \ell = 1,\ldots,p .$$

This ensures that all elements in $\underset{\sim}{\kappa}$ are non-negative, that all elements in $\underset{\sim}{\lambda}$ are non-positive, and that

$$\sum_{\ell=1}^{p} \kappa_\ell = 1 .$$

This parameterization is more restrictive than using either the difference or differential equations. The advantage of the sinusoid over the more commonly used logit transformation to constrain parameters to lie between 0 and 1 is that it allows the values 0 and 1 to be realised. For example, when p=1 and q=1 in equation (4.2.1) the stationary solution, denoted the ARMA(1,1) process by Box and Jenkins (1976), has an autocorrelation function:

$$\rho(u) = \kappa e^{u\lambda} \qquad \text{when } u>1 .$$

The same structure can be produced in this parameterization, provided $\kappa>0$, by setting p=2 and $\alpha_3 = -\pi/2$, because then $e^{\lambda_2} = 0$ and

$$\rho(u) = \kappa_1 + \kappa_2 = 1 \qquad \text{if } u = 0$$

$$= \kappa_1 e^{u\lambda_1} \qquad \text{if } u > 0 .$$

This is the parameterization that has been used in subsequent sections of this chapter. It has the advantage of restricting the parameters, $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$, to a sensible range for the data sets considered, and has less problems with iterative convergence in estimation when λ_ℓ is close to zero.

4.2.5 Model order identification

The orders, p and δ, in the autocorrelation function need to be known before $\underset{\sim}{\kappa}$ and $\underset{\sim}{\lambda}$ can be estimated. When observations are equally spaced this is equivalent to finding p and q, the orders in equation (4.2.1), because $\delta = q-p+1$. Autocorrelations $\hat{\underset{\sim}{\rho}}$ can be estimated from the weight corrected departures $(\hat{\underset{\sim}{a}})$ of the data from the regression function $(\hat{\underset{\sim}{f}})$ fitted, for example, by least-squares estimation. The sample autocorrelation at lag ℓ is

$$\rho_\ell = \Big(\sum_{i=1}^{n-\ell} \hat{a}_i \hat{a}_{i+\ell}\Big) \Big/ \Big(\sum_{j=1}^{n} \hat{a}_j^2\Big) \qquad \text{for } \ell = 0,1,2,\ldots ,$$

and $\hat{\rho}_\ell = \hat{\rho}_{-\ell}$. To help identify p and q from $\hat{\underset{\sim}{\rho}}$ I proposed (Glasbey, 1982) the use of generalized partial autocorrelation (GPA) coefficients, the (g,h)th of which is

$$\frac{\hat{\rho}_{g+h+1} - \sum_{i=1}^{g} \hat{\phi}_i^{\{gh\}} \hat{\rho}_{g+h-i+1}}{\sqrt{\sum_{j=-h}^{h} \Big(\sum_{k=0}^{\min(g,h-j)} \sum_{\ell=0}^{g} \hat{\phi}_k^{\{gh\}} \hat{\phi}_\ell^{\{gh\}} \hat{\rho}_{j+k-\ell}\Big)^2}} ,$$

where $\hat{\underset{\sim}{\phi}}^{\{gh\}}$ is the solution of the simultaneous linear equations

$$\hat{\rho}_\ell = \hat{\phi}_1^{\{gh\}} \hat{\rho}_{\ell-1} + \ldots + \hat{\phi}_g^{\{gh\}} \hat{\rho}_{\ell-g} \qquad \text{for } \ell = h+1,\ldots,h+g .$$

If $g \geq p$ and $h=q$ (or $g=p$ and $h \geq q$) then I proved that the

(g,h)th GPA coefficient is asymptotically normally distributed with zero mean and variance $1/n$, where n is the number of observations. Once the GPA coefficients have been calculated for a range of g and h, p and q can be selected as the smallest indices such that when $g>p$ and $h=q$, and when $g=p$ and $h>q$, the GPA coefficients are all within sampling range of zero.

When $h=0$, the GPA coefficients are simply the autocorrelation coefficients standardized to equal variances taking account of autocorrelations of lower lags (Bartlett, 1946) and assuming that autocorrelations at higher lags are zero. When $g=0$, the GPA coefficients are the partial autocorrelations proposed by Quenouille (1947).

The orders p and q identified from examination of an array of GPA coefficients are not necessarily those appropriate for parameterizations in [4.2.3] and [4.2.4] which require extra conditions to be satisfied. These can sometimes be ascertained by examining the values of $\hat{\underset{\sim}{\rho}}$. For example, for the weight-corrected departures to be represented as a sum of positively-correlated Markov processes all elements in $\underset{\sim}{\rho}$ should be non-negative. Of course, this is not a simple property to assess because of the sampling variability in $\hat{\underset{\sim}{\rho}}$. If there is doubt about the suitability of a model, then the options are either to use alternative parameterizations to test the assumptions, or to assess the fit by over-fitting using a higher order model.

Initial estimates of the error model parameters to provide starting values in the iterative optimization algorithm can be obtained by equating the first few coefficients in $\hat{\underset{\sim}{\rho}}$ and $\underset{\sim}{\rho}$.

4.2.6 Variances of parameter estimators

The distributions of maximum likelihood estimators have been discussed in [1.3.8]. The computer program, REGAME, which was described in section 2.4, has been used to fit models to the data sets in the following sections. Approximate variances of parameter estimators have been obtained by inverting the Hessian matrix of whichever optimization function is being used, whether the negative log-likelihood (L_M) or the negative residual log-likelihood (L_R) described in [2.3.2]. However, small simulations and marginal confidence intervals obtained directly from the likelihood surface have also been used.

4.3 Colquhoun's data: relaxation of drug-induced membrane currents

4.3.1 Introduction

One method of studying the mechanism of drug action has been by measuring the relaxation of drug-induced currents in the end-plate membrane of a muscle fibre after a voltage jump (see for example Colquhoun and Hawkes, 1977). The particular set of data that will be analysed in this thesis consists of 128 measurements made at intervals of 0.25 milliseconds after the voltage jump (Colquhoun, 1978). The first millisecond of data was discarded because of electrical transients leaving 124 observed currents which are plotted against time in figure 4.3.1. The theory of drug action predicts that the current decays as a sum of exponentials to a constant level. I used this data set as an

example in some earlier work (Glasbey 1980), the results of which will be referred to later.

4.3.2 Single exponential regression with independent errors

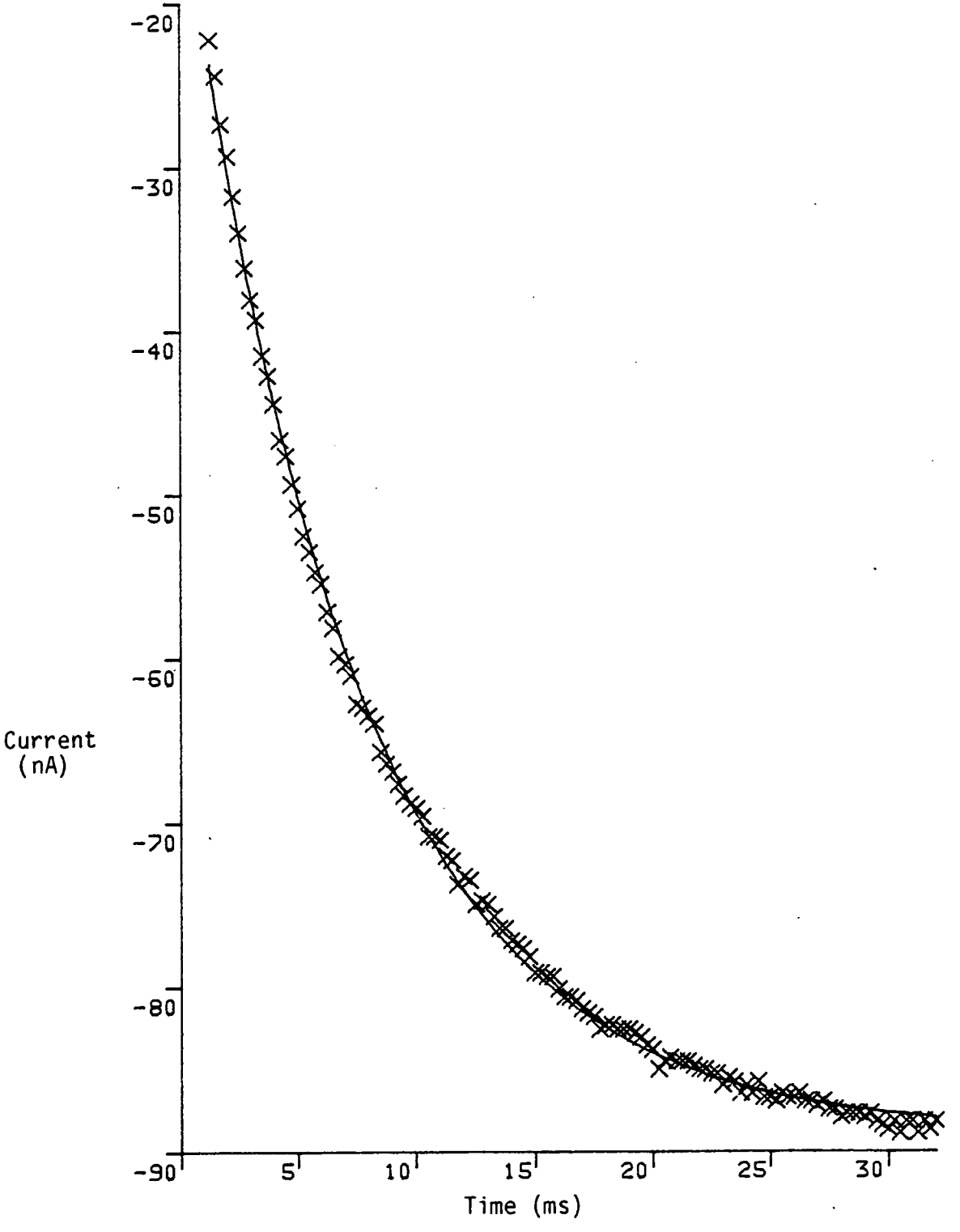
Initially the simplest regression model was considered (that is a single exponential response curve) and the simplest error model (independent errors with equal variances). Observations $\underset{\sim}{y}$ were therefore assumed to be independent normal deviates about a mean value of $\underset{\sim}{f}$ with variance matrix $\tau^2\underset{\sim}{I}$, where

$$f_i = \beta_1 + \beta_2 e^{-(1+i/4)/\beta_3} \qquad \text{for } i = 1, \ldots, 124,$$

with three unknown parameters, β_1, β_2 and β_3 (denoted $\underset{\sim}{\beta}$). From the definition in [2.2.1] $\underset{\sim}{y}$ is a GARMA(0,0) process, and the parameters $\underset{\sim}{\beta}$ and τ^2 can be estimated by minimizing L_M (defined in [2.3.2]) using the computer program REGAME described in section 2.4. At this stage the choice of L_M is of no importance because, for independent errors, the results would have been almost the same if L_R had been used as the optimizing function. The estimation proceeded iteratively on $\underset{\sim}{\beta}$ from a starting value which was obtained quite easily from the data which have only a small error component. The parameter estimates, standard errors and associated correlation coefficients for this simple model are given in table 4.3.1. As an example of the output from REGAME, the results to channel NOUT10 when this model was fitted, are given in Appendix B. The scatter plot produced by REGAME of departures of the data from the fitted regression show no indication that the error variances are unequal. However, the fitted curve in figure 4.3.1 exhibits systematic departures from the data, which implies that either a single exponential regression, or the assumption of

Figure 4.3.1

Colquhoun's data, single exponential regression with independent errors, fitted by least-squares estimation; observed currents (X) and fitted regression curve (——) plotted against time.



independent errors, is inappropriate. In most regression problems the more suspect assumption would be the single exponential model and this would then be modified, but in this example there is no a priori justification for assuming independent errors because the data were collected serially in time on a single subject. I will therefore consider extending the model to allow the error process to be correlated.

4.3.3 Generalized partial autocorrelation coefficients

To identify a suitable correlation model the autocorrelation and GPA coefficients of the departures of the data from the fitted regression curve were calculated. These are given in table 4.3.2. The estimated autocorrelation coefficients are significantly different from their zero expectations for independent errors. They appear to decline approximately exponentially after an initial drop from lag zero to lag one. This suggests an ARMA(1,1) process may be a suitable error model. The GPA coefficients indicate the same model because all terms below and to the right of the element where (p,q) is equal to (1,1) are small relative to the standard error of 9 (derived as $100/\sqrt{124}$). The estimated autocorrelation coefficients are also consistent with the errors being the sum of a positively-correlated Markov process and white noise, which was the parameterization considered in [4.2.4].

In some earlier work (Glasbey, 1980) using a similar regression function I restricted the choice of error models to autoregressive processes alone and found it necesary to use a third-order model, that is an ARMA(3,0) process in the present

Table 4.3.1

Colquhoun's data, single exponential regression with independent errors, fitted by least-squares estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients

	β_1	β_2	β_3	τ^2
estimate	-88.9	77.4	7.24	0.33
se	(0.1)	(0.3)	(0.06)	(0.04)
correlations				
β_2	0.18			
β_3	-0.82	-0.60		
τ^2	0.00	0.00	0.00	

Table 4.3.2

Colquhoun's data, single exponential regression with independent errors, fitted by least-squares estimation; 100×autocorrelation coefficients of residuals and 100×GPA coefficients.

Lag	1	2	3	4	5
autocorrelation	70	64	60	53	47
GPA					
p	0	1	2	3	4
q					
0	70	30	17	3	0
1	45	2	-5	-1	-6
2	36	-5	-2	-5	1
3	28	1	-4	7	5
4	24	-4	0	4	-2

Autocorrelation and GPA coefficients exceeding 18 (that is 2 standard errors) are underlined.

terminology. The estimated autocorrelation structure is very similar to that of an ARMA(1,1) process, but at the cost of one extra parameter.

4.3.4 Single exponential regression with ARMA(1,1) error model

The autocorrelation function was assumed to be of the form considered in [4.2.4] with $p=2$, e^{λ_2} held constant at zero and t_i taken to be i for $i = 1,\ldots,n$. The regression parameters ($\underset{\sim}{\beta}$) and error model parameters ($\underset{\sim}{\alpha}$) were jointly estimated by minimizing the negative log-likelihood L_M using REGAME. Starting values in the iterative process were taken as the least-squares estimated values of $\underset{\sim}{\beta}$ from [4.3.2] and $\underset{\sim}{\alpha}$ was estimated from the first few sample autocorrelation coefficients of the least-squares residuals. The results are given in table 4.3.3. The negative log-likelihood has decreased to -64.7 from the earlier value of -6.3 when errors were assumed to be independent. If the errors were independent then this difference would be distributed approximately as $\frac{1}{2}\chi^2_2$ because of the two extra parameters in the model. However, the 95 percentile for the $\frac{1}{2}\chi^2_2$ distribution is 3.0 which is greatly exceeded, so the improvement in fit is sufficient to reject the hypothesis of independent errors. The regression parameter estimates are approximately the same as those estimated assuming independent errors but the standard errors are much bigger. This illustrates the bias in the estimated variances of parameter estimators when an incorrect error model is assumed. The estimated autocorrelation function is

$$\hat{\rho}(u) = 0.85\ (0.976)^u \qquad \text{for } u > 0\ .$$

The output from fitting this model is given in the second half of Appendix B.

An alternative estimation method is to minimize L_R, the residual negative log-likelihood defined in [2.3.2]. The results are given in table 4.3.4. The estimated autocorrelation function is substantially different from above:

$$\hat{\rho}(u) = 0.99997\ (0.999995)^u \qquad \text{for } u > 0 .$$

Because the error process is very close to non-stationarity, problems were encountered with the difference equation parameterization of the autocorrelation function using ϕ_1 and θ_1, as described in [4.2.2]. The parameterization as a sum of positively-correlated Markov processes proved to be more robust. However, it was still necessary to experiment with different parameter scaling values in REGAME in order to achieve convergence. There are large correlations between the estimators of the error model parameters (α_1, α_2 and τ^2), but this would have been a problem with any parameterization. It is not of too great a concern because the main interest is in $\underset{\sim}{\beta}$. The fitted regression curve $(\hat{\underset{\sim}{f}})$ and the predicted fit using past observations (as described in [2.3.6]) are plotted against time in figure 4.3.2. The improvement in fit can be seen in the better agreement between predicted and observed values than in figure 4.3.1, although the fitted values in figure 4.3.2, which take no account of correlation in the errors, are worse than the fitted values when errors are assumed independent.

Further examination of the likelihood surface showed that it has no minimum and L_R converges to its lowest value of -68.8351 as α_1 converges to $\pi/2$ and

Table 4.3.3

Colquhoun's data, single exponential regression with ARMA(1,1) error model, fitted by maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	α_1	α_2	τ^2
estimate	-88.3	79.4	6.77	0.8	1.3	0.52
se	(0.7)	(1.0)	(0.23)	(0.3)	(0.1)	(0.37)
correlations						
β_2	-0.48					
β_3	-0.63	0.15				
α_1	0.22	-0.05	-0.53			
α_2	0.22	-0.10	-0.57	0.86		
τ^2	0.22	-0.07	-0.55	0.97	0.92	

Correlation coefficients exceeding 0.9 underlined.

Table 4.3.4

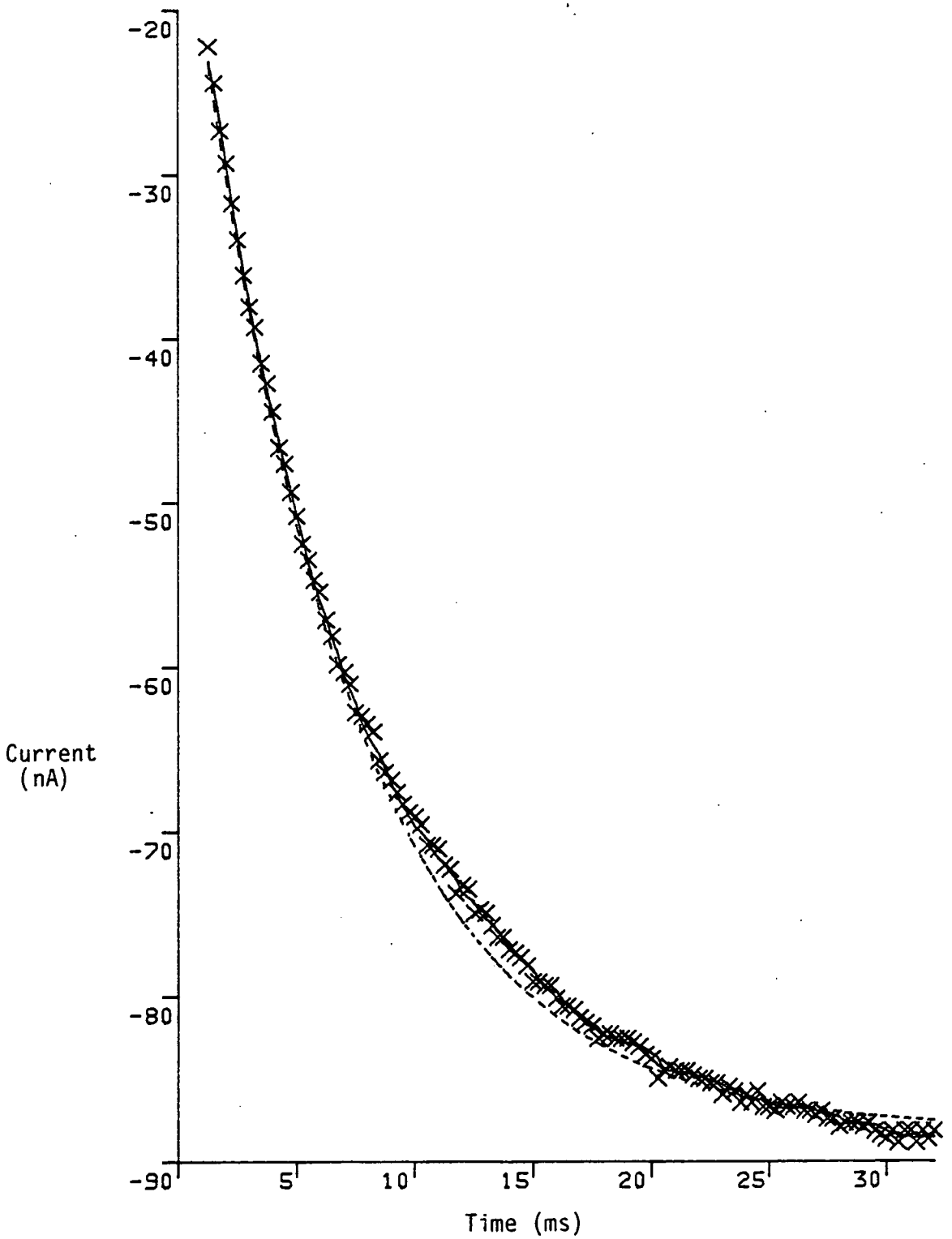
Colquhoun's data, single exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	α_1	α_2	τ^2
estimate	-88.2	78.6	6.59	1.559	1.566	2300.0
se	(47.7)	(1.6)	(0.24)	(0.109)	(0.042)	(43000.0)
correlations						
β_2	-0.02					
β_3	-0.01	0.48				
α_1	0.00	-0.01	-0.00			
α_2	0.00	-0.01	0.00	0.9997		
τ^2	0.00	-0.01	-0.00	0.99996	0.9998	

Correlation coefficients exceeding 0.9 underlined.

Figure 4.3.2

Colquhoun's data, single exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; observed currents (X), fitted regression curve (----) and predicted currents (——) conditional upon earlier observations, plotted against time.



$$\alpha_2 = \pi/2 - 0.379(\pi/2 - \alpha_1) .$$

As the limit is approached, all correlations between observations approach unity, and $\hat{\tau}^2$ becomes infinite, although the white noise component of the variance structure, that is $\kappa_2\tau^2$, converges to a finite value of 0.077. Therefore, the optimization algorithm gave the wrong answer for the minimum, probably because of numerical rounding errors as the limit was approached, although the value of L_R obtained (-68.8350) is close to its lower bound.

The high correlations in the error process imply that the regression function is not accounting for all the long-term trends in the data and as such are strong grounds for changing the regression function. However, I would like to persevere for a little while with the single exponential function because of its relationship with the simplest stochastic compartment models considered in section 7.3 as mechanistic models of Colquhoun's data.

The high correlations coud be eliminated by taking first differences of the series to obtain $y_2-y_1, y_3-y_2,\ldots, y_n-y_{n-1}$. The regression function becomes

$$\beta_2\{ e^{-(1+(i+1)/4)/\beta_3} - e^{-(1+i/4)/\beta_3}\} \qquad \text{for } i = 1,\ldots,n-1$$

with the loss of β_1, and the error series becomes a first-order moving average process with parameter θ defined in equation (4.2.1) and variance σ^2. The residual maximum likelihood estimates and standard errors are given in table 4.3.5. The error parameter estimates relate to the limiting relationships specified above between α_1, α_2 and τ^2 as follows:

$$\theta = 1 + (0.379)^2 - \sqrt{\{2(0.379)^2 + (0.379)^4\}} ,$$

$$\sigma^2 = 2(0.077)\ \{1 + (0.379)^2\}$$

Table 4.3.5

Colquhoun's data first differenced, exponential regression with ARMA(0,1) error model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_2	β_3	θ	σ^2
estimate	78.6	6.58	0.59	0.18
se	(1.6)	(0.24)	(0.07)	(0.02)
correlations				
β_3	0.48			
θ	-0.02	0.08		
σ^2	-0.01	0.03	0.39	

The estimates and standard errors of the parameters of primary interest, β_2 and β_3, are almost identical with those of table 4.3.4. Therefore it seems preferable to remain with the untransformed series to simplify comparisons between maximum likelihood estimation and residual maximum likelihood estimation.

A comparison of tables 4.3.3 and 4.3.4 shows that the regression parameter estimates are approximately the same but the standard errors of the residual maximum likelihood estimators are bigger, especially for $\hat{\beta}_1$. Both sets of parameter estimates satisfactorily account for the correlations in the data; examination of the residuals reveals no remaining autocorrelations.

4.3.5 Simulation to compare estimation methods

To compare the three estimation methods used so far, ten independent series, each of length 124, were generated from a model with a single exponential as regression function and stationary ARMA(1,1) errors with, as parameter values, the estimates from Colquhoun's data given in table 4.3.3. Pseudo-random standardized normal deviates were generated using the NAG routine G05DDF, starting from a seed set by the routine G05CBF. These were then transformed to have the appropriate mean and variance matrix by applying the technique described in [2.3.7], using subroutine REGGEN listed at the end of Appendix A. Separately, to each series of data, a single exponential regression model was fitted by three methods: (1) least-squares estimation, denoted LS (that is minimizing L_M on the assumption that the errors were independent); (2) maximum likelihood estimation, denoted ML (that is minimizing L_M on the assumption that the errors arose from an

ARMA(1,1) process); and (3) residual maximum likelihood estimation, denoted REML (that is minimizing L_R on the assumption that the errors arose from an ARMA(1,1) process). These are the three methods of estimation that have already been used with Colquhoun's data. Because of the heavy use of computer CPU time in fitting the models only ten series were simulated. It is recognised that sampling variability will remain a large component of any summary statistics derived from such a small sample, but large differences between methods of estimation should be identifiable.

For each parameter in the model and each method of estimation, three summary statistics were extracted from the ten independent estimates from the simulated series: bias, root-mean-square-error (abbreviated to r.m.s.e) and average standard error. To define what these statistics are, consider the regression parameter β_1 and denote its least-squares estimated value from the ith data set by $\hat{\beta}_1^{\{i\}}$. and its estimated standard error by $se(\hat{\beta}_1^{\{i\}})$, where i ranges between 1 and 10. The true value of β_1 is -88.3, given as the estimate of β_1 in table 4.3.3, and used to simulate the ten series. Therefore the estimated bias in the least-squares estimator of β_1 is

$$+88.3 + \sum_{i=1}^{10} \hat{\beta}_1^{\{i\}} / 10 ,$$

and the estimated r.m.s.e. is

$$\sqrt{[\sum_{i=1}^{10} (-88.3 - \hat{\beta}_1^{\{i\}})^2 / 10]} .$$

The mean standard error is defined simply as

$$\sum_{i=1}^{10} se(\hat{\beta}_1^{\{i\}}) / 10 \ .$$

Table 4.3.6 gives these summary statistics.

The whole procedure of simulation, estimation and summarization was repeated for a different set of model parameters, namely the estimates given in table 4.3.4. These summary statistics are given in table 4.3.7.

If it is assumed that the parameter estimators are normally distributed, then t-statistics can be derived in order to test whether the biases are significantly different from zero. The estimated variance of the bias can be calculated from the root-mean-square-error as

$$((r.m.s.e.)^2 - (bias)^2)/9 \ .$$

Because the 95 percentile for the t-distribution with 9 degrees of freedom is 2.26, biases on the threshold of significance at the 5% level satisfy:

$$\frac{bias}{\sqrt{(((r.m.s.e.)^2 - (bias)^2)/9)}} = \pm 2.26 \ .$$

This can be reexpressed as

$$bias = \pm 0.6017 \ (r.m.s.e.) \ .$$

Therefore biases exceeding 60% of the r.m.s.e.s are significantly different from zero at the 5% level and have been underlined in tables 4.3.6 and 4.3.7.

The most significant biases are for the maximum likelihood estimators of α_1 and α_2. The r.m.s.e.s are largest for the least-squares estimators as is to be expected because this is an inefficient estimation procedure when errors are correlated. The standard errors obtained by least-squares estimation are much smaller than the r.m.s.e.s and so are overestimating the precision

Table 4.3.6

Summary statistics derived from ten simulations of a single exponential regression with ARMA(1,1) error model, using as model parameters the estimates given in table 4.3.3, fitted by least-squares estimation (LS), maximum likelihood estimation (ML) and residual maximum likelihood estimation (REML).

	β_1	β_2	β_3	α_1	α_2	τ^2
bias						
LS	-0.1	-0.8	0.17			
ML	0.0	-0.4	0.11	-0.5	-0.4	-0.3
REML	0.1	-0.5	0.07	-0.1	-0.1	70.0
r.m.s.e.						
LS	0.7	1.3	0.23			
ML	0.6	0.9	0.15	0.6	0.5	0.3
REML	0.6	0.9	0.16	0.6	0.3	170.0
average se						
LS	0.1	0.2	0.04			
ML	0.3	0.5	0.11	0.3	0.2	0.1
REML	2.7	0.1	0.003	0.2	0.1	9.6

Biases significant at 5% level, i.e. exceeding 60% of r.m.s.e. underlined. Mean se's less than 80% of r.m.s.e. underlined. The three summary statistics are defined in [4.3.5].

Table 4.3.7

Summary statistics derived from ten simulations of a single exponential regression with ARMA(1,1) error model, using as model parameters the estimates given in table 4.3.4, fitted by least-squares estimation (LS), maximum likelihood estimation (ML) and residual maximum likelihood estimation (REML).

	β_1	β_2	β_3	α_1	α_2	τ^2
bias						
LS	-0.2	0.7	0.05			
ML	-0.2	0.8	0.05	-1.3	-0.7	-2300.0
REML	-0.3	0.7	0.05	-0.0	0.0	20.0
r.m.s.e.						
LS	24.7	1.2	0.2			
ML	24.8	1.3	0.1	1.4	0.7	2300.0
REML	24.9	1.2	0.1	0.001	0.0007	430.0
average se						
LS	0.1	0.2	0.04			
ML	0.2	0.5	0.01	0.3	0.2	0.06
REML	220.0	0.004	0.00002	0.003	0.0007	720.0

Biases significant at 5% level, i.e. exceeding 60% of r.m.s.e. underlined. Mean se's less than 80% of r.m.s.e. underlined. The three summary statistics are defined in [4.3.5].

of the parameter estimators. This is also as expected. However the ML, and especially the REML, standard errors are also much too small. Using ML, the true parameters are rejected on the basis of the negative log-likelihood, that is applying the likelihood ratio test. The average decreases are 4.8 and 8.8 in the two sets of simulations, but to be consistent with the true parameter values they should be distributed as $\frac{1}{20}\chi^2_{50}$. The 95 percentile for this distribution is 3.4, which is exceeded in both sets of simulations. For REML, the average decreases are 2.7 and 1.1 in the two sets of simulations, which are not significantly too large, so the true parameters are not rejected. This suggests that it is the quadratic approximation to the likelihood surface that has been used to obtain the REML standard errors which is causing them to be too small, possibly because of near singularity in the Hessian matrix, and likelihood-based confidence intervals should perform better.

4.3.6 Likelihood-based marginal confidence intervals

The quadratic approximation to the likelihood surface can be used to obtain 95% marginal confidence intervals for each regression parameter as

$$\hat{\beta}_j \pm se(\hat{\beta}_j)\ \sqrt{\{\chi^2_1\ (95\%)\}}$$

which is

$$\hat{\beta}_j \pm 1.96\ se(\hat{\beta}_j) \qquad \text{for } j = 1,2,3,$$

where $\chi^2_1(95\%)$ denotes the 95 percentile for the χ^2 distribution with 1 degree of freedom. To find 95% marginal confidence intervals based directly on the likelihood surface, the critical values of each parameter were found such that when that parameter

was held constant and all the other parameters estimated by minimizing L_R, the value of L_R was -66.9, that is $\frac{1}{2}\chi_1^2(95\%)$ greater than at its minimum value. The results are given in table 4.3.8. The only difference in marginal confidence intervals between the two methods of derivation is for β_1 which cannot apparently be estimated with any precision at all for this type of correlation amongst the errors. This is related to the point made in [4.3.4] that the likelihood surface has no minimum. As α_1 and α_2 approach $\pi/2$, the correlations all approach unity and β_1 cannot be estimated because it becomes confounded with the parameters in the error model. Therefore, no confidence limits can be placed on β_1.

4.3.7 Double exponential regression with independent errors

In [4.3.2] it was mentioned that the autocorrelations in the departures of the data from the fitted regression curve could have been caused by an inappropriate regression function. This was investigated further by considering a double exponential regression:

$$f_i = \beta_1 + \beta_2 e^{-(1+i/4)/\beta_3} + \beta_4 e^{-(1+i/4)/\beta_5}$$

for $i = 1,\ldots,124$.

The first fit obtained was by least-squares estimation and the parameter estimates are given in table 4.3.9. The autocorrelation and GPA coefficients of the departures are given in table 4.3.10. These are smaller than those obtained when fitting a single exponential regression, but some are sufficiently different from zero to reject the hypothesis of independent errors. The pattern is consistent with an ARMA(1,1) error model.

Table 4.3.8

Colquhoun's data, single exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; 95% marginal confidence intervals based on quadratic approximation and derived directly from the likelihood surface.

	Marginal confidence limits for		
	β_1	β_2	β_3
quadratic approximation			
lower	-181.6	75.5	6.11
upper	5.3	81.7	7.07
likelihood surface			
lower	$-\infty$	75.4	6.10
upper	∞	82.0	7.11

Table 4.3.9

Colquhoun's data, double exponential regression with independent errors, fitted by least-squares estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	β_5	τ^2
estimate	-90.4	25.3	3.28	58.6	9.22	0.12
se	(0.3)	(4.3)	(0.39)	(4.4)	(0.46)	(0.01)
correlations						
β_2	<u>-0.91</u>					
β_3	-0.80	<u>0.95</u>				
β_4	0.88	<u>-0.99</u>	<u>-0.98</u>			
β_5	<u>-0.95</u>	<u>0.99</u>	0.93	<u>-0.98</u>		
τ^2	-0.00	0.00	0.00	-0.00	0.00	

Correlation coefficients exceeding 0.9 underlined.

Table 4.3.10

Colquhoun's data, double exponential regression with independent errors, fitted by least-squares estimation; 100×autocorrelation coefficients of residuals and 100×GPA coefficients.

lag		1	2	3	4	5
autocorrelation		<u>22</u>	<u>22</u>	17	13	6
GPA						
	p	0	1	2	3	4
q						
0		<u>22</u>	18	10	5	-2
1		<u>21</u>	-2	-1	-3	-11
2		16	-1	-1	0	3
3		12	-2	0	3	0
4		6	-8	2	0	1

Autocorrelation and GPA coefficients exceeding 18 (that is 2 standard errors) are underlined.

4.3.8 Double exponential regression with ARMA(1,1) error model

Regression parameters $\underset{\sim}{\beta}$ and error parameters $\underset{\sim}{\alpha}$, as for the single exponential regression in [4.3.4], were jointly estimated: maximum likelihood estimates are given in table 4.3.11 and residual maximum likelihood estimates in table 4.3.12. The reduction in L_M, from -64.7 for a single exponential regression to -76.3 for a double exponential regression, is sufficiently large, for the addition of two extra parameters, to reject the hypothesis of a single exponential model using the asymptotic properties of the likelihood ratio test. Similarly, the reduction in L_R from -68.8 to -74.3 is sufficiently large to reject the same hypothesis. The regression parameter estimates in tables 4.3.11 and 4.3.12 are approximately the same as the least-squares estimates but the standard errors, especially the standard errors derived by residual maximum likelihood estimation, are larger. The maximum likelihood estimates of α_1 and α_2 are less than the residual maximum likelihood estimates. This is consistent with the results for the single exponential regression, which were shown by simulation to be due to the maximum likelihood estimates of the error parameters being biased downwards.

The experience with the single exponential regression indicates that marginal confidence intervals are more accurate when obtained directly from the likelihood surface. This also seems desirable because of the high correlations among the parameter estimates shown in table 4.3.12, which at the very least are likely to cause numerical problems in inverting the Hessian matrix. The 95% marginal confidence intervals are given in table 4.3.13. There are considerable discrepancies between the marginal

Table 4.3.11

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	β_5	α_1	α_2	τ^2
estimate	-90.4	26.1	3.41	57.6	9.29	-0.4	0.6	0.12
se	(0.4)	(7.6)	(0.64)	(7.7)	(0.80)	(0.3)	(0.3)	(0.02)
correlations								
β_2	-0.90							
β_3	-0.79	0.96						
β_4	0.87	-0.995	-0.98					
β_5	-0.94	0.99	0.93	-0.98				
α_1	0.04	-0.04	-0.03	0.03	-0.04			
α_2	-0.07	0.12	0.17	-0.14	0.11	-0.32		
τ^2	-0.00	0.03	-0.04	-0.03	0.02	0.35	0.20	

Correlation coefficients exceeding 0.9 underlined.

Table 4.3.12

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	β_5	α_1	α_2	τ^2
estimate	-90.6	31.5	3.86	51.8	9.83	0.1	1.1	0.19
se	(1.5)	(35.2)	(2.22)	(34.9)	(3.96)	(1.4)	(0.8)	(0.32)
correlations								
β_2	-0.94							
β_3	-0.91	0.99						
β_4	0.93	-0.9995	-0.995					
β_5	-0.96	0.996	0.98	-0.994				
α_1	-0.60	0.73	0.74	-0.73	0.70			
α_2	-0.62	0.74	0.75	-0.74	0.72	0.98		
τ^2	-0.61	0.73	0.74	-0.74	0.71	0.995	0.99	

Correlation coefficients exceeding 0.9 underlined.

Table 4.3.13

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; 95% marginal confidence intervals based on quadratic approximation and derived directly from the likelihood surface.

	Marginal confidence limits for				
	β_1	β_2	β_3	β_4	β_5
quadratic approximation					
lower	-93.6	-37.5	-0.48	-16.5	2.07
upper	-87.5	100.6	8.21	120.2	17.59
likelihood surface					
lower	$-\infty$	8.2	1.96	7.5	7.90
upper	∞	76.8	5.61	73.7	∞

confidence limits derived by quadratic approximation, that is based on the standard errors in table 4.3.12, and those derived directly from the likelihood surface. As for the single exponential regression in table 4.3.8, the likelihood-based limits for β_1 are infinite. For parameters β_2 , β_3 and β_4 the likelihood-based confidence intervals are shorter. For the fifth parameter the limits are asymmetric about $\hat{\beta}_5$, indicating that the estimator has a skew distribution.

The fitted curve and predicted fit from past observations based on residual maximum likelihood estimation are shown plotted against time in figure 4.3.3. The agreement between predicted and observed values is very good indeed, although it is difficult to see if it is any better than the agreement in figure 4.3.2.

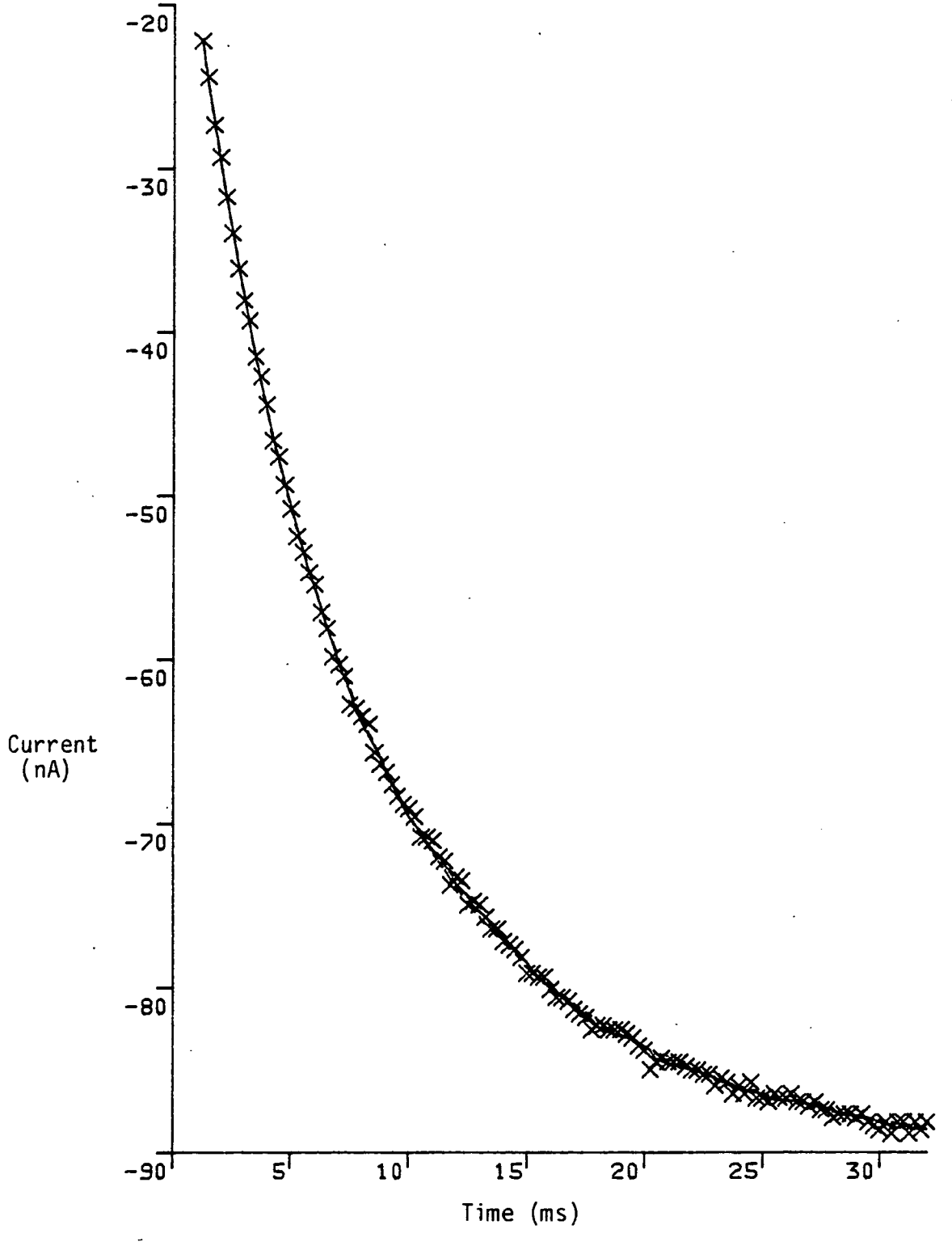
The adequacy of the ARMA(1,1) process to model the errors can be assessed either by examination of the residuals after fitting the model, or by fitting a model of higher order and testing the improvement in fit. Both these criteria support the existing model. The residuals, which are the estimates of $\underset{\sim}{z}$ in the notation of section 2.3, have very small autocorrelations at low lags. Alternatively, when an ARMA(2,2) model is fitted, the decrease in L_R is only 0.1 which is nowhere near approaching the magnitude required to reject the hypothesis of an ARMA(1,1) model using the likelihood ratio test.

4.3.9 Resumé

Regression parameters have been estimated and marginal confidence intervals derived for both single and double exponential functions. ARMA(1,1) error models estimated by residual maximum

Figure 4.3.3

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; observed currents (X), fitted regression curve (----) and predicted currents (____) conditional upon earlier observations, plotted against time.



likelihood with marginal confidence intervals based directly on the likelihood surface were found to be most appropriate for Colquhoun's data set. The improvement in fit of the double exponential model over the single exponential model was found to be sufficient to reject the simpler model.

This data set will be considered again in chapter 5 when a spectral parameterization of the error process will be considered. Then, in chapter 7 a mechanistic error model will be used in place of the empirical model considered in this chapter. Finally, in chapter 9 the parameter estimates obtained in this section will be re-used, but more generally valid estimates of the regression parameter estimator standard errors will be derived.

4.4 Dale's data: radioactive emission from a wheat leaf

4.4.1 Introduction

To study the effect of IAA, a growth regulating substance, on transport within a plant, a wheat leaf was fed a constant level of radioactive carbon dioxide for five minutes. The radioactive discharge was then measured every minute for 23 hours. The data, consisting of ten minute averages of Geiger counts, are plotted against time in figure 4.4.1. The theory of transport in the leaf predicts that the discharge decays as a sum of exponentials (Bauermeister, Dale, Williams, Scobie, 1980).

4.4.2 Regression with independent errors

Initially, the vector of observations, denoted $\underset{\sim}{y}$, were assumed independently normally distributed with mean $\underset{\sim}{f}$ where

$$f_i = \beta_1 e^{-\beta_2(10i-5)} + \beta_3 e^{-\beta_4(10i-5)}$$

for $i = 1,\ldots,138$,

and the ith observation has a variance of $\tau^2 f_i$. The choice of weighting for the variances was made because of the Poisson form of sampling variability inherent in radioactive emissions and for empirical reasons, arising from the observed scatter in the data revealed in some exploratory plots which are not included in this thesis. Estimation of the parameters by minimization of L_M, by the use of REGAME, gave the parameter estimates in table 4.4.1. The fitted curve is shown in figure 4.4.1. The autocorrelation and GPA coefficients of the weight-corrected departures of the data from the fitted regression are shown in table 4.4.2. These are sufficiently large to reject the hypothesis of independent errors and their pattern suggests an ARMA(1,1) process as a suitable error model. As for Colquhoun's data, the parameterization as a sum of positively-correlated Markov processes appears to be appropriate.

4.4.3 Regression with weighted ARMA(1,1) error model

With $t_i = i$ for $i = 1,\ldots,138$, and the same form of weights as in [4.4.2], the results of fitting the model by maximum likelihood estimation are shown in table 4.4.3 and by residual maximum likelihood in table 4.4.4. The regression parameter estimates in tables 4.4.3 and 4.4.4, especially $\hat{\beta}_2$, are quite different from the least-squares estimates. This can also be seen by comparing the curve fitted by residual maximum likelihood estimation which is plotted in figure 4.4.2 with the least-squares fit plotted in figure 4.4.1. As for Colquhoun's data, the maximum

Table 4.4.1

Dale's data, regression with weighted independent errors, fitted by least-squares estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	τ^2
estimate	818.0	0.0142	702.0	0.00067	0.61
se	(21.0)	(0.0006)	(7.0)	(0.00001)	(0.07)
correlations					
β_2	0.52				
β_3	-0.02	0.69			
β_4	-0.04	0.61	<u>0.93</u>		
τ^2	-0.00	0.00	-0.00	0.00	

Correlation coefficients execceding 0.9 underlined.

Table 4.4.2

Dale's data, regression with weighted independent errors, fitted by least-squares estimation; 100×autocorrelation coefficients of residuals and 100×GPA coefficients.

Lag		1	2	3	4	5
autocorrelation		<u>72</u>	<u>62</u>	<u>61</u>	<u>56</u>	<u>51</u>
GPA						
	p	0	1	2	3	4
q						
0		<u>72</u>	<u>22</u>	<u>22</u>	5	5
1		<u>44</u>	13	-7	3	-1
2		<u>36</u>	-7	-2	1	-6
3		<u>29</u>	1	1	-6	-0
4		<u>25</u>	1	-3	-1	3

Autocorrelation and GPA coefficients exceeding 17 (that is 2 standard errors) are underlined.

Figure 4.4.1

Dale's data, regression with weighted independent errors, fitted by least-squares estimation; observed radioactive discharges (X) and fitted regression curve(——) plotted against time.

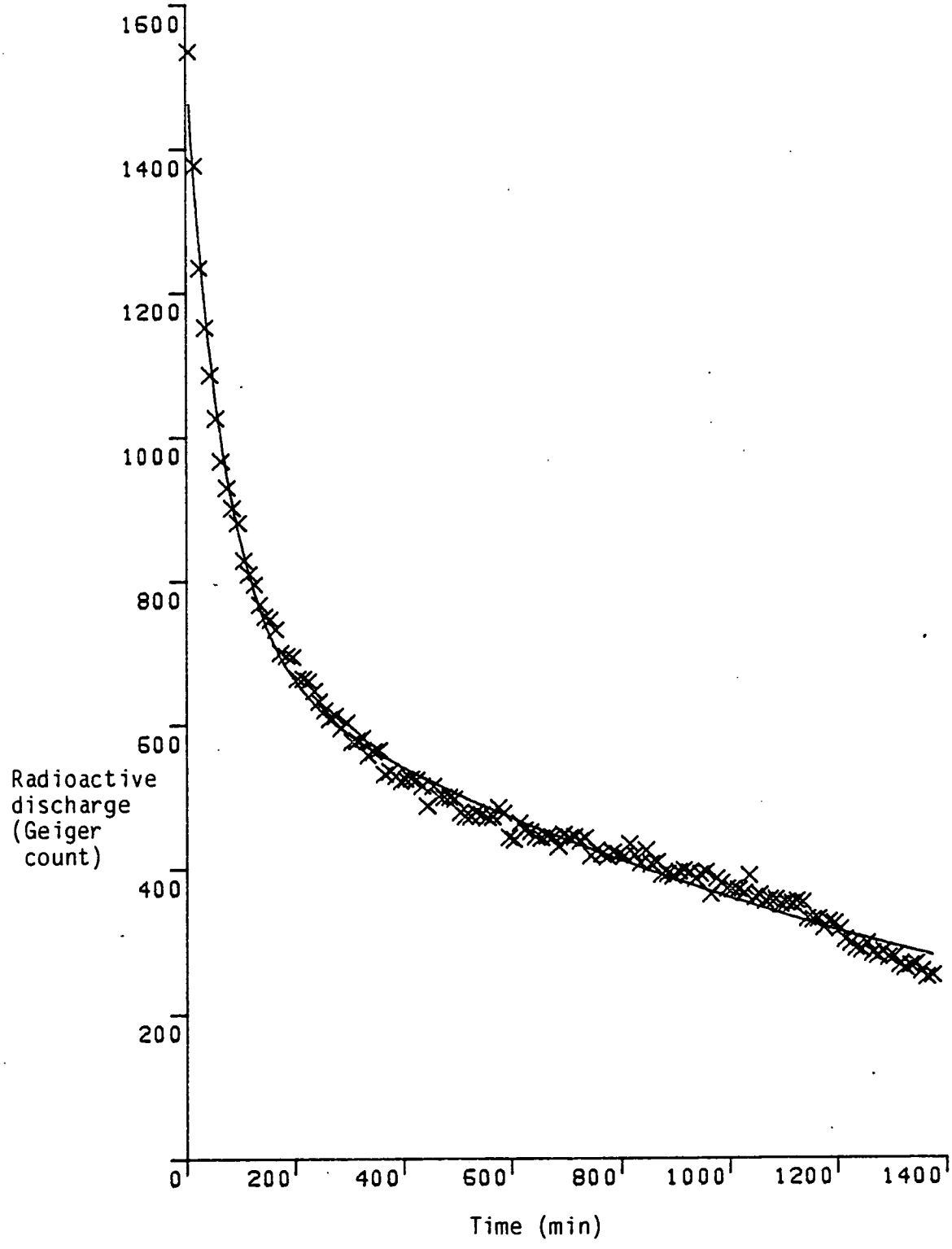


Table 4.4.3

Dale's data, regression with weighted ARMA(1,1) error model, fitted by maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	809.0	0.0205	785.0	0.00081	0.9	1.3	1.04
se	(38.0)	(0.0018)	(39.0)	(0.00006)	(0.2)	(0.1)	(0.64)
correlations							
β_2	-0.44						
β_3	-0.43	0.57					
β_4	-0.43	0.49	0.80				
α_1	-0.25	0.36	0.41	0.42			
α_2	-0.28	0.24	0.41	0.43	0.82		
τ^2	-0.27	0.32	0.40	0.44	<u>0.96</u>	0.89	

Correlation coefficients exceeding 0.9 underlined.

Table 4.4.4

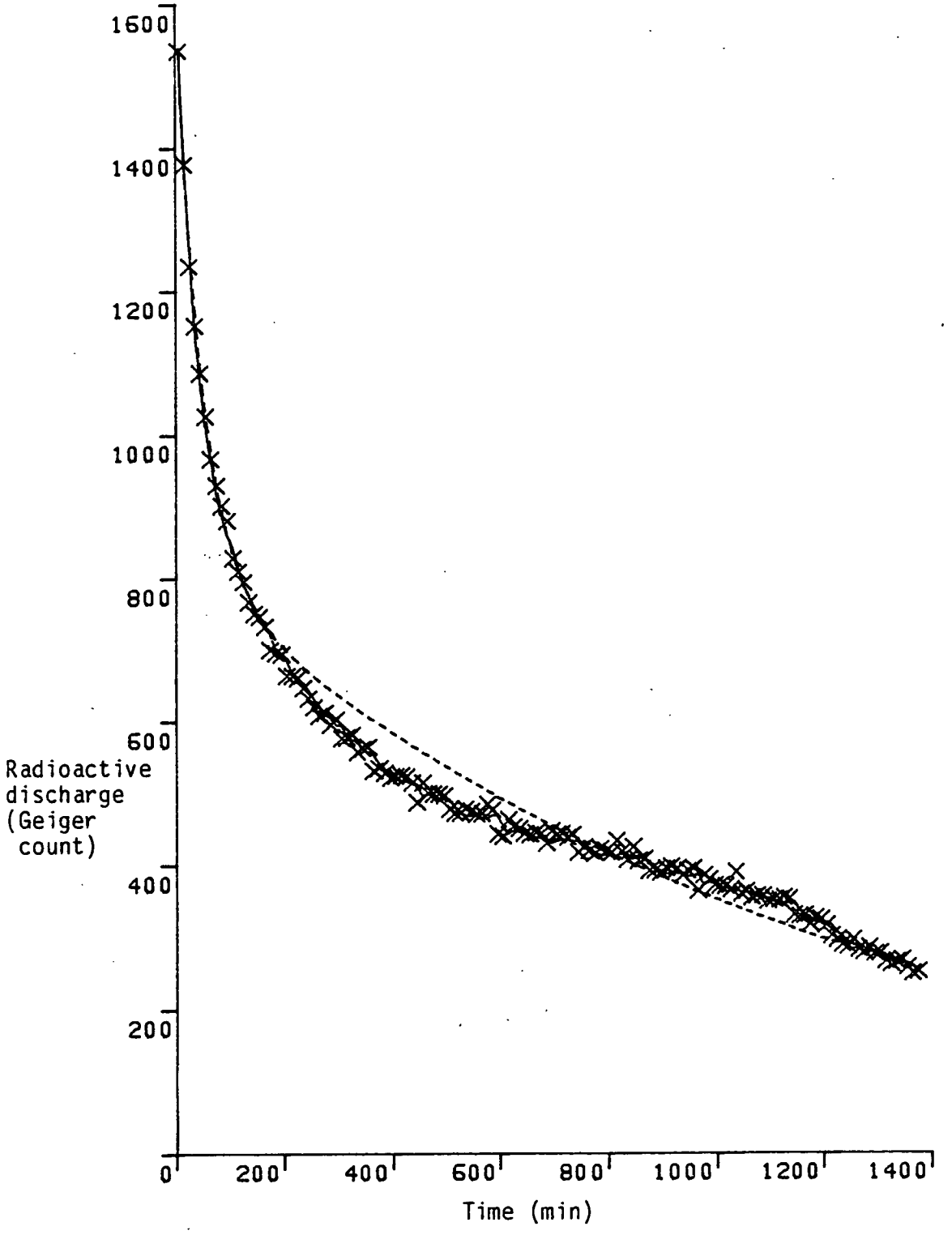
Dale's data, regression with weighted ARMA(1,1) errors model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	806.0	0.0210	816.0	0.00084	1.41	1.50	18.2
se	(74.0)	(0.0020)	(157.0)	(0.00013)	(0.77)	(0.36)	(175.0)
correlations							
β_2	-0.32						
β_3	0.72	0.12					
β_4	-0.64	0.34	-0.40				
α_1	0.55	-0.04	0.61	-0.41			
α_2	0.55	-0.06	0.61	-0.41	0.999		
τ^2	0.53	-0.05	0.60	-0.40	0.9996	0.9993	

Correlation coefficients exceeding 0.9 underlined.

Figure 4.4.2

Dale's data, regression with weighted ARMA(1,1) error model, fitted by residual maximum likelihood estimation; observed radioactive discharges (X), fitted regression curve (----) and predicted values (——) conditional upon earlier observations, plotted against time.



likelihood estimates of α_1 and α_2 are lower than the residual maximum likelihood estimates and this is probably the reason why regression parameter standard errors derived by maximum likelihood estimation are smaller. The improvement in fit over the model with independent errors, as measured by L_M decreasing from 458.7 to 385.3, is sufficient to reject the hypothesis of independent errors on the basis of the asymptotic properties of the likelihood ratio test. The residuals have very small autocorrelations at low lags, and so an ARMA(1,1) process appears to model adequately the correlation structure in the errors. The maximum likelihood estimate of the autocorrelation function is

$$\hat{\rho}(u) = 0.89\ (0.976)^u \qquad \text{for } u > 0 ,$$

and the residual maximum likelihood estimate is

$$\hat{\rho}(u) = 0.994\ (0.9987)^u \qquad \text{for } u > 0 .$$

The high correlations are a cause for concern as they were for Colquhoun's data in [4.3.4], although in this case L_R does have a minimum. The same reasons may be given here as in [4.3.4] for continuing with the present model.

4.4.4 Adjusted least-squares standard errors

The differences in the estimates of $\underset{\sim}{\beta}$ between table 4.4.1 and tables 4.4.3 and 4.4.4 may be because of the improvement in efficiency of maximum likelihood estimation compared with that of least-squares. On the assumption that the regression function is approximately linear in its parameters close to the best fitting value, the weighted least-squares estimate of $\underset{\sim}{\beta}$ is

$$\hat{\underset{\sim}{\beta}} \simeq \underset{\sim}{\beta}^* + (\underset{\sim}{X}^T \underset{\sim}{W}^{-1} \underset{\sim}{W}^{-1} \underset{\sim}{X})^{-1}\ \underset{\sim}{X}^T \underset{\sim}{W}^{-1} \underset{\sim}{W}^{-1} (\underset{\sim}{y} - \underset{\sim}{f}^*) , \tag{4.4.1}$$

where

$$\underset{\sim}{X} = \partial \underset{\sim}{f} / \partial \underset{\sim}{\beta}^T \Big|_{\underset{\sim}{\beta} = \underset{\sim}{\beta}^*} ,$$

$\underset{\sim}{W}$ is a diagonal matrix with $W_{ii} = \sqrt{f_i}$, $\underset{\sim}{\beta}^*$ is an arbitrary value of the vector of regression parameters close to the best fitting value, and $\underset{\sim}{f}^*$ is the corresponding value of the regression vector. Therefore, if $\underset{\sim}{y}$ has a variance $\tau^2\underset{\sim}{V}$, then

$$(4.4.2)\quad \mathrm{var}(\hat{\underset{\sim}{\beta}}) \simeq \tau^2(\underset{\sim}{X}^T\underset{\sim}{W}^{-1}\underset{\sim}{W}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{W}^{-1}\underset{\sim}{W}^{-1}\underset{\sim}{V}\,\underset{\sim}{W}^{-1}\underset{\sim}{W}^{-1}\underset{\sim}{X}\,(\underset{\sim}{X}^T\underset{\sim}{W}^{-1}\underset{\sim}{W}^{-1}\underset{\sim}{X}) .$$

On the assumption that $\underset{\sim}{V}$ has the same ARMA(1,1) structure as in [4.4.3], its parameters $(\underset{\sim}{\alpha})$ can be estimated using the departures of the data from the least-squares fitted regression. The maximum likelihood estimate of the error parameters and standard errors, and the least-squares estimated regression parameters with standard errors calculated using equation (4.4.2), are given in table 4.4.5. The corresponding residual maximum likelihood estimates are given in table 4.4.6. The standard errors are not large enough to account for the difference with tables 4.4.3 and 4.4.4.

The maximum likelihood and residual maximum likelihood estimators of the regression parameters, which were discussed in [4.4.3], can also be expressed approximately as linear functions of the observation vector $\underset{\sim}{y}$, analoguous to equation (4.4.1). Approximate correlations between these estimators and the least-squares regression estimators can then be calculated and the multivariate distance between different estimates of $\underset{\sim}{\beta}$ obtained. The squared Mahalanobis distance between the least-squares and maximum likelihood estimates is 22.0, which exceeds the 95 percentile of χ^2_4 (which is 9.5), and so is large enough to reject the hypothesis that the two estimators are both estimating $\underset{\sim}{\beta}$, on the assumption that the errors arose from an ARMA(1,1) process. A similar conclusion holds in a comparison between the least-squares and residual maximum likelihood estimates for which the Mahalanobis

Table 4.4.5

Dale's data, ARMA(1,1) error model fitted by maximum likelihood estimation to residuals after least-squares fit of regression, variances of regression parameter estimators recalculated; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	818.0	0.0142	702.0	0.00067	0.7	1.1	0.73
se	(42.0)	(0.0017)	(31.0)	(0.00005)	(0.2)	(0.2)	(0.32)
correlations							
β_2	-0.05						
β_3	-0.42	0.66					
β_4	-0.40	0.57	0.90				
α_1							
α_2					0.53		
τ^2					0.87	0.80	

Table 4.4.6

Dale's data, ARMA(1,1) error model fitted by residual maximum likelihood estimation to residuals after least-squares fit of regression, variances of parameter estimators recalculated; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	818.0	0.0142	702.0	0.00067	1.51	1.54	116.0
se	(151.0)	(0.0027)	(284.0)	(0.00023)	(0.24)	(0.13)	(14.0)
correlations							
β_2	-0.32						
β_3	0.86	-0.04					
β_4	-0.78	0.47	-0.59				
α_1	-						
α_2					0.997		
τ^2					0.9995	0.999	

Correlation coefficients exceeding 0.9 underlined.

squared distance is 16.1. Therefore, either asymptotic properties are not being closely approximated, or the errors are not from an ARMA(1,1) process, or equivalently (see [1.1.2]) the regression function is inappropriate.

4.4.5 Simulation to compare estimation methods

The validity of asymptotic approximations can be tested by simulation. To compare the estimation methods, ten independent series were generated in an analogous fashion to that described in [4.3.5], but this time the series were of length 138 and used the regression function in [4.4.2] with a weight-corrected stationary ARMA(1,1) model for the errors with, as parameter values, the estimates from Dale's data given in table 4.4.4. Separately, to each series of data, the regression model of [4.4.2] was fitted by four methods: (1) least-squares estimation, but with standard errors recalculated after an ARMA(1,1) process had been fitted by maximum likelihood estimation to the departures of the data from the fitted regression as described in [4.4.4], denoted LS/ML; (2) least-squares estimation as above, but with the ARMA(1,1) process fitted by residual maximum likelihood estimation, denoted LS/REML; (3) joint maximum likelihood estimation of both the regression and the ARMA(1,1) process, denoted ML; and (4) joint residual maximum likelihood estimation of both the regression and the ARMA(1,1) process, denoted REML. These are the four methods of estimation that have already been applied to Dale's data. Because of the heavy use of computer CPU time in fitting the models only ten series were simulated.

The results are summarized in table 4.4.7 using the summary statistics previously defined in [4.3.5]. None of the biases in

Table 4.4.7

Summary statistics derived from ten simulations of a regression with ARMA(1,1) error model, using as model parameters the estimates given in table 4.4.6, fitted by least-squares estimation and the error process by maximum likelihood estimation (LS/ML) or residual maximum likelihood estimation (LS/REML), full maximum likelihood estimation (ML) and residual maximum likelihood estimation (REML).

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
bias							
LS/ML	-4.0	-0.0005	12.0	-0.0000	-0.9	-0.6	-18.0
LS/REML	-4.0	-0.0005	12.0	-0.0000	-0.5	-0.3	- 3.0
ML	-2.0	-0.0003	14.0	0.0000	-0.9	-0.6	-18.0
REML	71.0	-0.0007	324.0	-0.0001	0.1	0.0	1100.0
r.m.s.e.							
LS/ML	26.0	0.0024	51.0	0.00006	0.9	0.6	19.0
LS/REML	26.0	0.0024	51.0	0.00006	0.6	0.4	21.0
ML	22.0	0.0018	51.0	0.00006	0.9	0.6	18.0
REML	207.0	0.0019	646.0	0.00027	0.2	0.1	1500.0
average se							
LS/ML	32.0	0.0018	20.0	0.00003	0.2	0.2	0.2
LS/REML	54.0	0.0027	85.0	0.00009	0.3	0.2	12.0
ML	30.0	0.0013	19.0	0.00003	0.2	0.2	0.2
REML	129.0	0.0011	465.0	0.00017	0.1	0.05	450.0

Biases significant at 5% level, i.e. exceeding 60% of r.m.s.e. underlined. Mean se's less than 80% of r.m.s.e. underlined. The three summary statistics are defined in [4.3.5].

the regression parameter estimators is significant at the 5% level, so the differences in estimates between table 4.4.1 and tables 4.4.3 and 4.4.4 remain unexplained. As with Colquhoun's data, α_1 and α_2 appear to be underestimated by maximum likelihood estimation. The average ses are smaller than the r.m.s.e.s with all the estimation methods and so are overestimating the precision of the parameter estimators. In ML, the average decrease in L_M from the true model to the estimated model is 6.0 which exceeds the 95 percentile of $\frac{1}{20}\chi^2_{60}$ which is 4.0, so the true model is rejected on the basis of the likelihood ratio test. In REML, the change in L_R of 2.3 is close to its expected value of 3.0 and the true model is not rejected. Therefore, as with the simulation in [4.3.5], the likelihood surface for L_R appears to provide a better measure of precision of estimation than do the standard errors based on a quadratic approximation.

4.4.6 Likelihood-based marginal confidence intervals

Marginal confidence intervals for the regression parameters were calculated using the standard errors in table 4.4.4, and also directly from the likelihood surface as described in [4.3.6]. The results are given in table 4.4.8 and show that the quadratic approximation is underestimating the standard errors. This is consistent with the simulation results.

4.4.7 Regression with non-stationary error model

Simulation has not shown the asymptotic properties to be far off the truth, so the inconsistencies between estimators must be due to the assumed model being inappropriate.

Examination of figure 4.4.1 and also of the line printer plots generated by REGAME (which have not been included in the thesis) indicate that the weight-corrected departures of the data from the least-squares fitted regression are not stationary: the departures show much greater changes in magnitude at the start of the series than at later times. When the errors were assumed to be stationary, the fitted regression function was adjusted so that the sample correlation coefficients among the departures at the start of the series increased. This can be seen in the departures of the data from the fitted curve in figure 4.4.2.

Therefore, as an experiment, the covariance model was modified by allowing the first three observations to be independent of the rest of the data which were once again modelled by an ARMA(1,1) process. This model corresponds to a GARMA(1,1) process. The parameters were estimated by residual maximum likelihood estimation and the results are given in table 4.4.9. The regression parameter estimates are a compromise between those given in tables 4.4.1 and 4.4.4 with the new estimate of β_2 in particular being much closer to the least-squares estimate. The fit is a slight improvement on the stationary model: L_R has decreased from 367.5 to 367.0 .

Once the possibility has been admitted that the error process is non-stationary, the number of alternative models from which to choose is enormous. The effects of different selections on the regression parameter estimators may be large, but the data are of little assistance in making the choice.

Table 4.4.8

Dale's data, regression with weighted ARMA(1,1) error model, fitted by residual maximum likelihood estimation; 95% marginal confidence intervals based on quadratic approximation and derived directly from the likelihood surface.

	Marginal confidence limits for			
	β_1	β_2	β_3	β_4
quadratic approximation				
lower	661.0	0.0171	509.0	0.00059
upper	950.0	0.0250	1123.0	0.00110
likelihood surface				
lower	523.0	0.0175	398.0	0.00039
upper	1174.0	0.0258	2034.0	0.00125

Table 4.4.9

Dale's data, regression with weighted ARMA(1,1) error model excluding the first three errors which are taken to be independent, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	749.0	0.0156	779.0	0.00078	1.1	1.4	2.2
se	(49.0)	(0.0025)	(43.0)	(0.00008)	(0.2)	(0.1)	(2.2)
correlations							
β_2	0.31						
β_3	-0.53	0.37					
β_4	-0.11	0.56	0.54				
α_1	-0.44	-0.48	0.07	-0.37			
α_2	-0.47	-0.57	0.06	-0.40	<u>0.92</u>		
τ^2	-0.44	-0.49	0.05	-0.35	<u>0.99</u>	<u>0.95</u>	

Correlation coefficients exceeding 0.9 underlined.

4.4.8 Resumé

Double exponential regression parameters have been estimated on the assumption that the errors arose from an ARMA(1,1) process. The least-squares and residual maximum likelihood estimates were not in agreement. An explanation for the discrepancy was sought in a small simulation experiment but none was found. Further examination of the data revealed a non-stationary error structure, but this opened up the possibility of so many alternative models that the analysis ground to a halt.

The effect on the regression parameter estimates of changing the error model reveals the sensitivity of the method described in this chapter to the assumptions made about $\underset{\sim}{V}$. This poses a problem because $\underset{\sim}{V}$ is unknown. This data set will be examined further in chapter 7 using a mechanistic error model rather than the empirical model considered here. Finally, in chapter 9 the parameter estimates obtained in this section will be re-used, but more generally valid estimates of the regression parameter estimator standard errors will be derived.

4.5 Bruce's data: energy demand of a mechanical model of a suckler cow

4.5.1 Introduction

A mechanical model of a suckler cow was built to measure the integrated energy demand of a real suckler cow (Burnett and Bruce, 1978; Bruce, 1980). Every hour for seven months between October and April the energy required to maintain the model at 39°C was

measured together with four climatic variables: temperature, radiation, wind speed, and rainfall. The objective was to relate the energy demand to the climatic variables and so be able to predict the energy demand of a suckler cow in any similar climatic environment.

The data were averaged over each 24 hours which reduced the number of observations to 200. The energy demand was standardized to unit surface area of model and log-transformed to standardize the variances because there was a greater scatter to the data at the higher energy levels. The log-transformed energy demand is plotted against time in figure 4.5.1. Consideration of the physics of energy balance and fluid flow led to the selection of a particular regression function with four unknown parameters:

$$\ell n\left[\frac{39 - \text{temperature} - \beta_1\ \text{radiation}/(5.3+7\ \text{wind}^{0.6})}{0.203 + \beta_2\{1-\min(\beta_3\ \text{rain},\ \beta_4)\} + 1/(5.3+7\ \text{wind}^{0.6})}\right] .$$

Because the optimization routine used in REGAME requires continuous first derivatives with respect to the parameters the coefficient of β_2 between curly brackets was approximated by:

$$1 - \beta_4 + 10^{-2}\ell n(1 + e^{(\beta_4-\beta_3 \text{rain})/10^{-2}}) .$$

4.5.2 Regression with independent errors

Initially, the observations $\underset{\sim}{y}$, that is the log-transformed energy demands per unit surface area of model, were assumed independently normally distributed with mean $\underset{\sim}{f}$ and variance $\tau^2\underset{\sim}{I}$. The least-squares estimated parameters are given in table 4.5.1 and the fitted curve is plotted in figure 4.5.1. The autocorrelation and GPA coefficients of the departures of the data from the regression function are given in table 4.5.2. These are

Figure 4.5.1

Bruce's data, regression with independent errors, fitted by least-squares estimation; observed log-energy demand (X) and fitted regression curve (——) plotted against time.

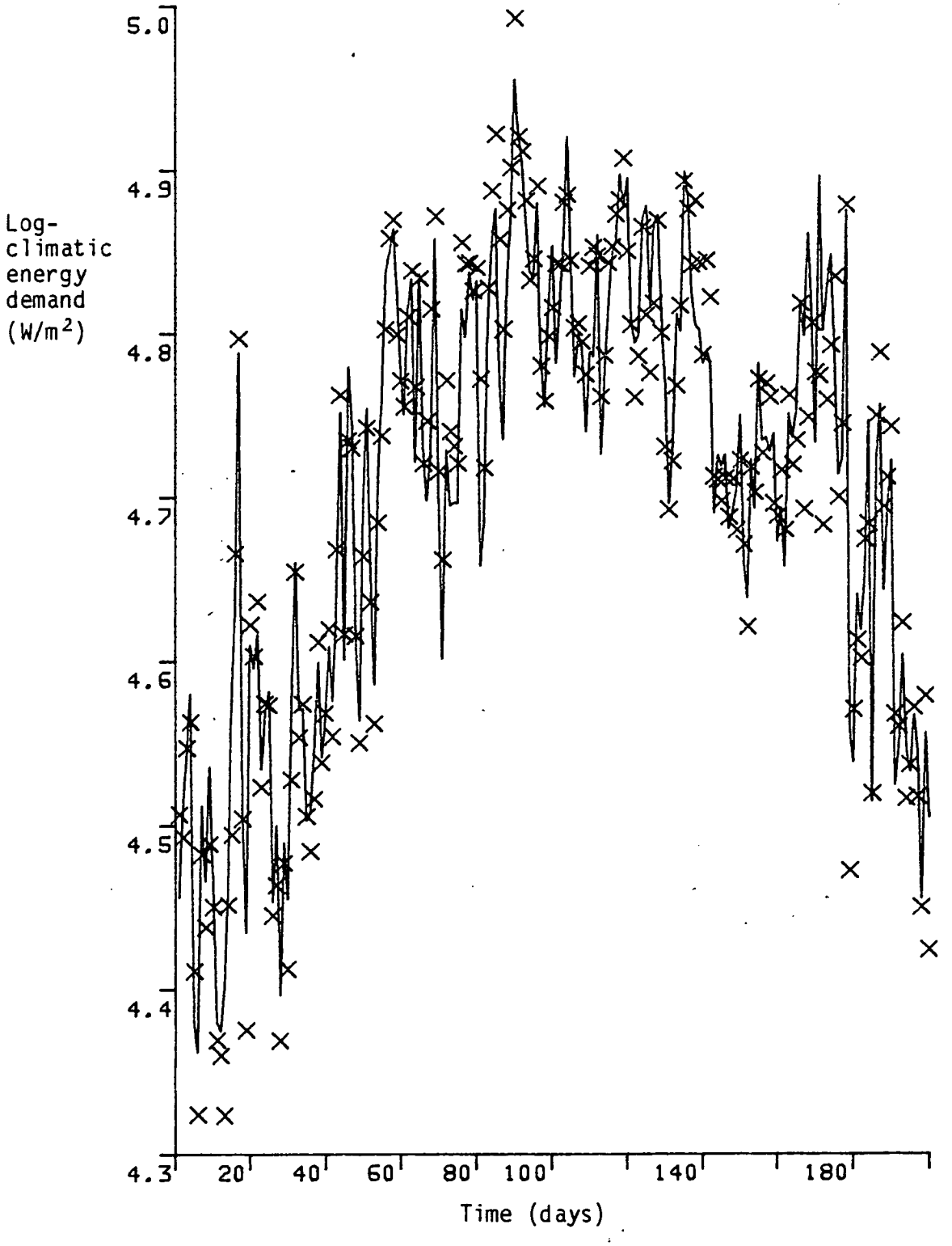


Table 4.5.1

Bruce's data, regression with independent errors, fitted by least-squares estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	τ^2
estimate	0.75	0.068	1.6	0.57	0.0013
se	(0.05)	(0.001)	(0.1)	(0.04)	(0.0001)
correlations					
β_2	0.05				
β_3	0.18	0.34			
β_4	0.07	0.20	0.30		
τ^2	0.00	0.00	0.00	0.00	

Table 4.5.2

Bruce's data, regression with independent errors, fitted by least-squares estimation; 100×autocorrelation coefficients of residuals and 100×GPA coefficients.

Lag	1	2	3	4	5
autocorrelation	35	25	28	32	31
GPA					
p	0	1	2	3	4
q					
0	35	15	18	19	15
1	23	8	-1	-2	-6
2	24	1	-3	0	1
3	26	-3	-1	1	0
4	24	-7	2	-1	-6

Autocorrelation and GPA coefficients exceeding 14 (that is 2 standard errors) are underlined.

sufficiently large to reject the hypothesis of independent errors and their pattern suggests an ARMA(1,1) model with the parameterization as a sum of positively-correlated Markov processes specified in [4.2.4].

4.5.3 Regression with ARMA(1,1) error model

The results of jointly estimating regression and error parameters by maximum likelihood are given in table 4.5.3 and by residual maximum likelihood in table 4.5.4. The regression function fitted by residual maximum likelihood estimation and the predicted fit using past observations (as described in [2.3.6]) are plotted in figure 4.5.2. The improvement in fit over the model with independent errors, as measured by L_M decreasing from -564.2 to -589.8, is sufficient to reject the hypothesis of independent errors on the basis of the asymptotic properties of the likelihood ratio test. The residuals have very small autocorrelations at low lags, and so an ARMA(1,1) process appears to model adequately the correlation structure in the errors. The maximum likelihood and residual maximum likelihood estimates are in good agreement on this occasion, possibly because α_2 is no longer close to $\pi/2$. However, the regression parameter estimates in tables 4.5.3 and 4.5.4 are substantially different from those in table 4.5.1.

To investigate whether this discrepancy is caused by the inefficiency of the least-squares estimators, an ARMA(1,1) model was fitted to the least-squares residuals. The standard errors of the least-squares regression parameter estimators were then adjusted, as described in [4.4.4]. The results of maximum likelihood estimation are given in table 4.5.5 and of residual

Table 4.5.3

Bruce's data, regression with ARMA(1,1) error model, fitted by maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	0.61	0.067	1.3	0.48	-0.0	0.9	0.0014
se	(0.08)	(0.003)	(0.1)	(0.04)	(0.2)	(0.2)	(0.0003)
correlations							
β_2	-0.01						
β_3	0.04	-0.15					
β_4	0.02	-0.23	0.54				
α_1	-0.09	-0.13	-0.14	-0.16			
α_2	-0.37	-0.06	0.08	0.10	0.26		
τ^2	-0.20	-0.11	-0.06	-0.07	0.75	0.55	

Table 4.5.4

Bruce's data, regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	0.60	0.067	1.3	0.48	0.0	1.0	0.0015
se	(0.08)	(0.003)	(0.1)	(0.04)	(0.2)	(0.2)	(0.0004)
correlations							
β_2	0.01						
β_3	0.02	-0.22					
β_4	0.01	-0.31	0.54				
α_1	-0.15	-0.17	-0.08	-0.09			
α_2	-0.38	-0.13	0.11	0.13	0.47		
τ^2	-0.24	-0.16	-0.01	-0.01	0.84	0.69	

Figure 4.5.2

Bruce's data, regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; observed log-energy demands (X), fitted regression curve (----) and predicted values (____) conditional upon earlier observations, plotted against time.

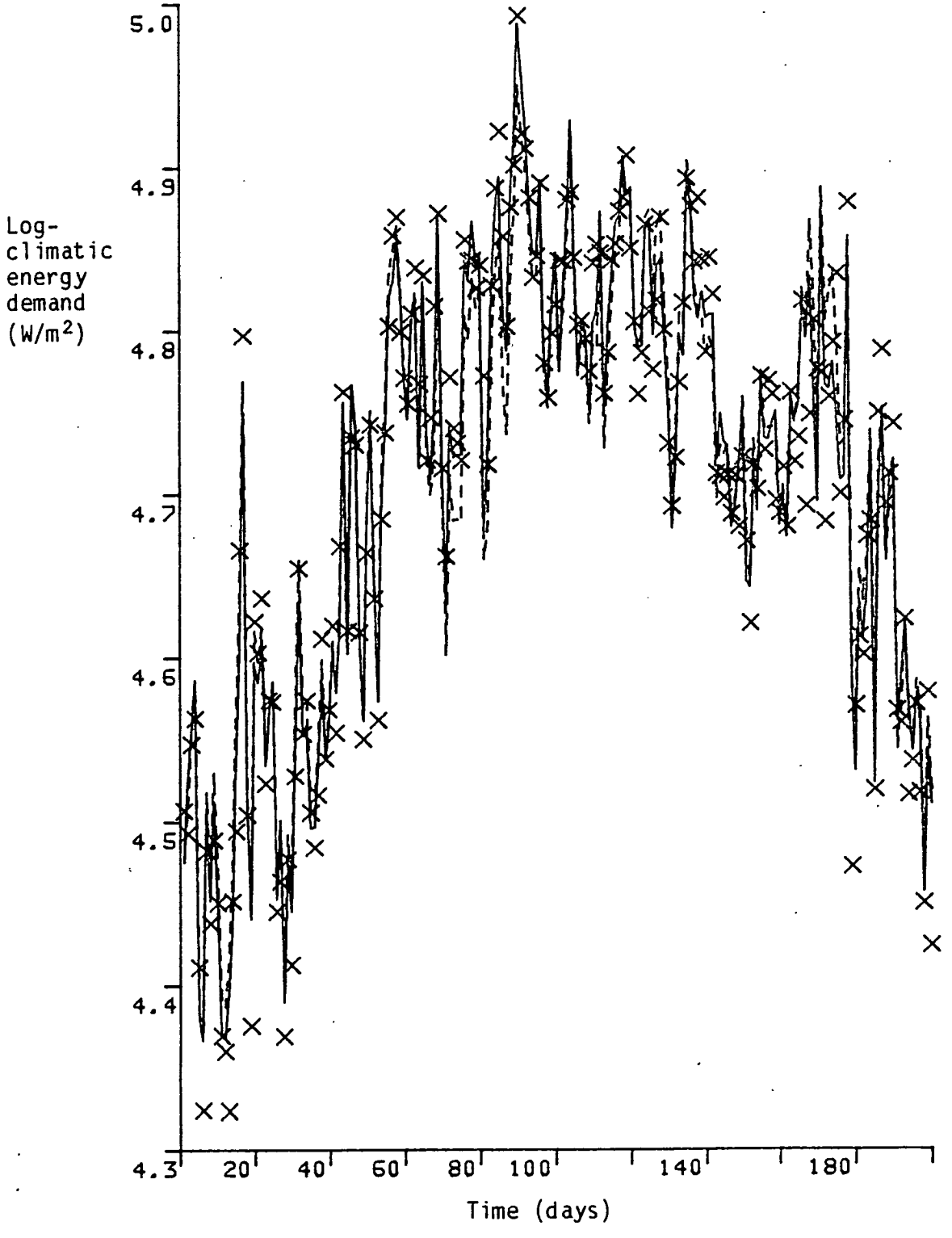


Table 4.5.5

Bruce's data, ARMA(1,1) error model fitted by maximum likelihood estimation to residuals after least-squares fit of regression, variances of regression parameter estimators recalculated; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	0.75	0.068	1.6	0.57	-0.2	0.9	0.0013
se	(0.10)	(0.002)	(0.2)	(0.05)	(0.2)	(0.2)	(0.0002)
correlations							
β_2	-0.05						
β_3	0.10	0.07					
β_4	0.13	-0.11	0.35				
α_1							
α_2					0.14		
τ^2					0.63	0.42	

Table 4.5.6

Bruce's data, ARMA(1,1) error model fitted by residual maximum likelihood to residuals after least-squares fit of regression, variances of regression parameter estimators recalculated; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
estimate	0.75	0.068	1.6	0.57	-0.2	0.9	0.0014
se	(0.11)	(0.002)	(0.2)	(0.05)	(0.2)	(0.2)	(0.0002)
correlations							
β_2	-0.06						
β_3	0.11	0.20					
β_4	0.15	-0.14	0.39				
α_1							
α_2					0.30		
τ^2					0.71	0.54	

maximum likelihood estimation in table 4.5.6. The squared Mahalanobis distance between the least-squares and residual maximum likelihood estimates was computed as in [4.4.4]. The value of 9.7 is close to the 95 percentile of χ^2_4 (which is 9.5) and so is large enough to cast doubt on the hypothesis that the two estimators are both estimating $\underset{\sim}{\beta}$, on the assumption that the errors arose from an ARMA(1,1) process. Therefore, as with the analysis of Dale's data, either asymptotic properties are not being closely approximated, or the errors are not from an ARMA(1,1) process.

4.5.4 Simulation to compare estimation methods

To compare the estimation methods, ten independent series were generated in an analogous fashion to that described in [4.3.5], but this time the series were of length 200 and used the regression function in [4.5.1] with an ARMA(1,1) model for the errors with, as parameter values, the estimates from Bruce's data given in table 4.5.4. Separately, to each series, the regression model of [4.5.1] was fitted by the same four methods described in [4.4.5]. Once again, because of the heavy use of computer CPU time in fitting the models only ten series were simulated. It is recognised that sampling variability will remain a large component of the summary statistics derived from such a small sample, but the earlier experiences in [4.3.5] and [4.4.5] suggest that valuable information can still be extracted.

The results are given in table 4.5.7 using the summary statistics previously defined in [4.3.5]. There are no biases in the regression parameter estimates significant at the 95% level so

Table 4.5.7

Summary statistics derived from ten simulations of a regression with ARMA(1,1) error model, using as model parameters the estimates given in table 4.5.6, fitted by least-squares estimation and the error process by maximum likelihood estimation (LS/ML) or residual maximum likelihood estimation (LS/REML), full maximum likelihood estimation (ML) and residual maximum likelihood estimation (REML).

	β_1	β_2	β_3	β_4	α_1	α_2	τ^2
bias							
LS/ML	0.02	0.000	0.1	0.02	-0.1	-0.1	0.0000
LS/REML	0.02	0.000	0.1	0.02	-0.0	-0.1	0.0001
ML	0.00	0.000	0.1	0.01	0.1	-0.0	0.0001
REML	0.00	0.000	0.1	0.01	0.1	0.0	0.0003
r.m.s.e.							
LS/ML	0.17	0.002	0.3	0.06	0.2	0.2	0.0003
LS/REML	0.17	0.002	0.3	0.06	0.2	0.2	0.0004
ML	0.09	0.002	0.2	0.03	0.2	0.2	0.0004
REML	0.09	0.003	0.2	0.03	0.2	0.2	0.0005
average se							
LS/ML	0.13	0.003	0.2	0.06	0.2	0.2	0.0003
LS/REML	0.15	0.003	0.3	0.06	0.2	0.2	0.0004
ML	0.07	0.003	0.2	0.05	0.2	0.2	0.0004
REML	0.07	0.004	0.2	0.05	0.3	0.2	0.0007

Biases significant at 5% level, i.e. exceeding 60% of r.m.s.e. underlined. Mean se's less than 80% of r.m.s.e. underlined.
The three summary statistics are defined in [4.3.5].

the differences in estimates between table 4.5.1 and tables 4.5.3 and 4.5.4 remain unresolved. The standard errors appear to be of the correct magnitude agreeing well with the r.m.s.e.s, so for this data set it would appear to be unnecessary to obtain marginal confidence intervals directly from the likelihood surface.

4.5.5 Regression with non-stationary error model

There is no a priori justification for the assumption that the error process is stationary. In part, the errors arise from the variability in other climatic variables that were not incorporated into the regression function, and climatic variables in general are known not to be stationary over as long a period as 200 days. The autocorrelation coefficients at lags one and two, calculated separately for each quarter of the series of departures of the data from the regression function fitted by residual maximum likelihood, also cast doubt on the assumption of stationarity. These are displayed in table 4.5.8.

To explore the possibility of the error process being non-stationary, the ARMA(1,1) model was generalized. In the stationary model, fitted in [4.5.3], elements in the variance matrix are of the form:

$$V_{ij} = \kappa\, e^{(i-j)\lambda} \qquad \text{if } i > j ,$$

which can be re-expressed as

$$V_{ij} = \kappa \prod_{\ell=j}^{i-1} e^{\lambda} \qquad \text{if } i > j .$$

This can be generalized to a non-stationary model by allowing κ and λ to be time dependent, to give

$$V_{ij} = \sqrt{(\kappa(i)\ \kappa(j))} \prod_{\ell=j}^{i-1} e^{\lambda(\ell)} \qquad \text{if } i > j$$

Table 4.5.8

Bruce's data, regression with ARMA(1,1) error model, fitted by residual maximum likelihood estimation; 100×autocorrelation coefficients of departures of data from fitted regression, separately for each quarter of data.

	100×autocorrelation coefficients			
quarter	1	2	3	4
lag				
1	18	88	28	53
2	20	76	21	21

where, for example,

$$\kappa(\ell) = \frac{1}{2} + \frac{1}{2} \sin (\alpha_1 + \alpha_1^* \ell) ,$$

$$e^{\lambda(\ell)} = \frac{1}{2} + \frac{1}{2} \sin (\alpha_2 + \alpha_2^* \ell) ,$$

and α_1^* and α_2^* are two additional parameters which determine the non-stationarity in the error process. In particular, if $\underset{\sim}{\alpha^*} = \underset{\sim}{0}$, then the model simplifies to its stationary form. This model serves as a simple non-stationary generalization of an ARMA(1,1) process of the form given in [4.2.4], whilst retaining the GARMA(1,1) property because

$$V_{ij} = (\sqrt{\kappa(i)} \prod_{\ell=1}^{i-1} e^{\lambda(\ell)})(\sqrt{\kappa(j)} \prod_{\ell=1}^{j-1} e^{-\lambda(\ell)}) \qquad \text{if } i > j .$$

The regression and error parameters were estimated by residual maximum likelihood and the estimates are given in table 4.5.9. The decrease in L_R from -577.7 to -583.2 is sufficiently large to reject the hypothesis of a stationary ARMA(1,1) error process, on the basis of a likelihood ratio test, because the 95 percentile of $\frac{1}{2} \chi_2^2$ (which is 3.0) is exceeded.

Once the possibility has been admitted that the error process is non-stationary, the choice of model becomes daunting. Unless further information is available about the error process the problem appears to be intractable.

4.5.6 Resumé

Regression parameters have been estimated on the assumption that the errors arose from an ARMA(1,1) process. The least-squares and residual maximum likelihood estimates were not in agreement. An explanation for the discrepancy was sought in a small simulation experiment but none was found. Further

Table 4.5.9

Bruce's data, regression with non-stationary ARMA(1,1) error model, fitted by residual maximum likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	α_1	α_1^*	α_2	α_2^*	τ^2
estimate	0.54	0.068	1.3	0.47	0.6	-0.003	1.4	-0.002	0.0019
se	(0.09)	(0.006)	(0.2)	(0.05)	(0.4)	(0.002)	(0.2)	(0.003)	(0.0010)
correlations									
β_2	0.30								
β_3	-0.18	-0.52							
β_4	-0.20	-0.63	0.56						
α_1	-0.52	-0.66	0.39	0.42					
α_1^*	0.08	0.01	0.01	0.00	-0.19				
α_2	-0.05	0.16	0.09	-0.00	-0.16	0.54			
α_2^*	-0.43	-0.65	0.34	0.45	0.81	-0.31	-0.53		
τ^2	-0.53	-0.66	0.43	0.46	0.91	0.15	0.10	0.72	

Correlation coefficients exceeding 0.9 underlined.

examination of the data revealed a non-stationary error structure, but this opened up the possibility of so many alternative models that the analysis ground to a halt.

In chapter 9, the parameter estimates obtained in this section will be re-used, but more generally valid estimates of the regression parameter estimator standard errors will be derived.

4.6 Discussion

The basic technique presented in this chapter is to fit a regression function to a series of observations by least-squares estimation, identify an ARMA model of suitable order by examination of the sample autocorrelations of the departures of the data from the fitted regression, and then jointly estimate the regression and error model parameters by maximum likelihood or residual maximum likelihood. The experience in applying the technique to three data sets shows that the method is not as simple as it seems. The only satisfactory analysis was of Colquhoun's data, for which a double exponential regression with ARMA(1,1) errors appeared to fit adequately. When a single exponential regression was fitted instead, problems were encountered in confounding between regression and error model parameters when estimating by maximizing the residual log-likelihood. For Dale's and Bruce's data inconsistencies between least-squares and maximum likelihood estimators of the regression parameters were encountered and were attributed to the lack of stationarity in the error processes. This highlights the major problem when estimating regression

parameters in the presence of unknown error covariances: the technique rests heavily upon gross assumptions that are made about the correlation structure of the observations. An impasse has been reached in this chapter and is the motivation for the novel approach taken in chapter 9.

Other points have arisen in the course of fitting models to the data sets. An ARMA(1,1) error model has been identified as the most appropriate ARMA process in every data set. Simulations have shown residual maximum likelihood estimation to be preferable to maximum likelihood estimation, and likelihood based marginal confidence intervals to be better than those derived from standard errors using the quadratic approximation to the likelihood surface. The comparison between maximum likelihood estimation and residual maximum likelihood estimation in this type of application warrants further study beyond the limited set of cases considered in this chapter or in the original paper by Cooper and Thompson, (1977).

5. Stationary error processes: empirical spectral parameterization

5.1 Introduction

This chapter is a natural sequel to the previous one. There are two dominant strands in the statistical time series literature: the time domain approach with its emphasis on the autocorrelation function; and the frequency domain approach with the spectral function at its heart. In chapter 4, the error process associated with a regression model was modelled empirically by specifying an autocorrelation function, and in this chapter it is modelled empirically by specifying a spectral function.

Robinson (1978) states that in certain situations it is easier to parameterize the spectrum than the autocorrelation function of a stationary process. He estimated the spectral parameters by maximizing a pseudo-likelihood, derived asymptotic variances of parameter estimators, and proved that the estimators are efficient when the data are multivariate normally distributed.

In section 5.2 the pseudo-likelihood is shown to be an approximation to the Gaussian likelihood, and it is demonstrated that the regression and spectral parameters can be estimated by minimizing L_p, the negative of the logarithm of the pseudo-likelihood, using REGAME. A technique analogous to that in chapter 4 is proposed; that is initially fit a regression function by least-squares estimation, identify an appropriate spectral function by examination of the periodogram of the departures of the data from the fitted regression, and then jointly estimate the regression and spectral parameters by minimizing L_p. The

technique is explored in section 5.3 by fitting a double exponential regression model to Colquhoun's data which were also used in section 4.3. The other two data sets in chapter 4 have not been re-used because the error processes were shown to be non-stationary, so the technique in this chapter is inappropriate.

Although the basic technique in chapter 4 is not new, the frequency domain approach in this chapter does appear to be original.

5.2 Theory

5.2.1 Model

The model to be considered is for a sequence of observations $y_1, \ldots, y_n$ (denoted $\underset{\sim}{y}$) which have been made at times 1, 2, ..., n on a single experimental subject. (This is rather more restrictive than [4.2.1] where observation times were allowed to be unequally spaced.) It is assumed that $\underset{\sim}{y}$ is a realisation from a multivariate normal distribution with mean $\underset{\sim}{f}$ and variance $\tau^2\underset{\sim}{V}$, where

$$(\underset{\sim}{W}^{-1}\underset{\sim}{V}\,\underset{\sim}{W}^{-1})_{jk} = \rho_{j-k} = \rho_{k-j} ,$$

and $\underset{\sim}{W}$ is a diagonal matrix of weights. Therefore, the vector of weight-corrected departures, denoted $\underset{\sim}{a}$ $(= \underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}))$, is a finite realisation of what is supposed to be a stationary process of infinite length with autocorrelation ρ_ℓ for any integer lag ℓ. The theoretical spectrum $(\underset{\sim}{d})$, denoted d_g at frequency $(g-1)/n$, is defined as

$$d_g = \tau^2 \sum_{\ell=-\infty}^{\infty} \rho_\ell \, e^{2\pi\imath\ell(g-1)/n} \quad , \tag{5.2.1}$$

for $g = 1, \ldots, n$,

where $\imath$ is defined to be $\sqrt{(-1)}$, and the periodogram $(\underset{\sim}{D})$, denoted D_g at frequency $(g-1)/n$, is defined as

$$D_g = \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} a_j \, e^{-2\pi\imath j(g-1)/n} \right|^2 ,$$

$$= \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} a_j \, a_k \, e^{2\pi\imath(k-j)(g-1)/n} \quad , \tag{5.2.2}$$

for $g = 1, \ldots, n$,

where a_j is the jth element in $\underset{\sim}{a}$. These are standard formulae, see for example Priestley (1981a, pp222-226 and 394-397), although different authors use different multiplying constants and frequency scales. In the above notation, g indexes the sinusoid component which completes (g-1) cycles in the series of length n.

5.2.2 Derivation of pseudo-likelihood

The negative log-likelihood of the observations is given in [2.3.2] as

$$L_M = \frac{n}{2} \ell n \, \tau^2 + \frac{1}{2} \ell n \left| \underset{\sim}{V} \right| + \frac{1}{2\tau^2} (\underset{\sim}{y} - \underset{\sim}{f})^T \underset{\sim}{V}^{-1} (\underset{\sim}{y} - \underset{\sim}{f})$$

$$= \frac{1}{2} \ell n \left| \tau^2 \underset{\sim}{V} \right| + \frac{1}{2\tau^2} \underset{\sim}{a}^T \underset{\sim}{W} \underset{\sim}{V}^{-1} \underset{\sim}{W} \underset{\sim}{a}$$

by substituting for $(\underset{\sim}{y} - \underset{\sim}{f})$. With $\underset{\sim}{H}$ defined as an $n \times n$ symmetric complex matrix with (j,k)th coefficient

$$H_{jk} = \frac{1}{\sqrt{n}} e^{2\pi\imath(j-1)(k-1)/n} \quad ,$$

$$L_M = \frac{1}{2} \ell n \left| \underset{\sim}{W} \bar{\underset{\sim}{H}} (\tau^2 \underset{\sim}{H} \underset{\sim}{W}^{-1} \underset{\sim}{V} \underset{\sim}{W}^{-1} \bar{\underset{\sim}{H}}) \underset{\sim}{H} \underset{\sim}{W} \right|$$

$$+ \frac{1}{2\tau^2} \underset{\sim}{a}^T \bar{\underset{\sim}{H}} (\tau^2 \underset{\sim}{H} \underset{\sim}{W}^{-1} \underset{\sim}{V} \underset{\sim}{W}^{-1} \bar{\underset{\sim}{H}})^{-1} \underset{\sim}{H} \underset{\sim}{a} \quad , \tag{5.2.3}$$

where $\bar{\underset{\sim}{H}}$ denotes the complex conjugate and also the inverse of $\underset{\sim}{H}$. It is the inverse because

$$(\bar{\underset{\sim}{H}}\,\underset{\sim}{H})_{jk} = \frac{1}{n}\sum_{\ell=1}^{n} e^{-2\pi\imath(j-1)(\ell-1)/n}\, e^{2\pi\imath(\ell-1)(k-1)/n}$$

$$= \frac{1}{n}\sum_{\ell=1}^{n} e^{2\pi\imath(k-j)(\ell-1)/n}$$

for $j,k = 1,\ldots,n$.

If $k=j$ this equals 1 , but otherwise it is a geometric series with a sum of

$$\frac{1}{n}\{\,1 - e^{2\pi\imath(k-j)}\}/\{\,1 - e^{2\pi\imath(k-j)/n}\} ,$$

which equals 0 because for integer values of j and k

$$e^{2\pi\imath(k-j)} = 1 .$$

Therefore $\bar{\underset{\sim}{H}}\,\underset{\sim}{H} = \underset{\sim}{I}$. It follows from this result that

$$|\bar{\underset{\sim}{H}}|\,|\underset{\sim}{H}| = |\underset{\sim}{I}| = 1 .$$

We now consider the $n\times n$ symmetric complex matrix $(\tau^2\underset{\sim\sim}{H}\underset{\sim\sim}{W}^{-1}\underset{\sim\sim}{V}\underset{\sim\sim}{W}^{-1}\bar{\underset{\sim}{H}})$, the (j,k)th element of which is

$$(\tau^2\underset{\sim\sim}{H}\underset{\sim\sim}{W}^{-1}\underset{\sim\sim}{V}\underset{\sim\sim}{W}^{-1}\bar{\underset{\sim}{H}})_{jk} = \frac{\tau^2}{n}\sum_{g=1}^{n}\sum_{h=1}^{n}\rho_{g-h}e^{2\pi\imath(j-1)(g-1)/n}\, e^{-2\pi\imath(k-1)(h-1)/n},$$

$$= \frac{\tau^2}{n}\sum_{\ell=1-n}^{n-1}\rho_{\ell}\sum_{v=\max(0,-\ell)}^{\min(n-1,n-1-\ell)} e^{2\pi\imath(j-1)(\ell+v)/n}\, e^{-2\pi\imath(k-1)v/n}$$

by transforming to a new scale with $v = h-1$ and $\ell = g-1-v$,

$$= \frac{\tau^2}{n}\sum_{\ell=1-n}^{n-1}\rho_{\ell}\, e^{2\pi\imath\ell(j-1)/n}\sum_{v=\max(0,-\ell)}^{\min(n-1,n-1-\ell)} e^{2\pi\imath(j-k)v/n} .$$

However,

$$\frac{1}{n}\sum_{v=0}^{n-1} e^{2\pi\imath(j-k)v/n} = I_{jk} ,$$

where I_{jk} denotes the (jk)th element in $\underset{\sim}{I}$, the identity matrix of size n, so provided ρ_ℓ decays rapidly to zero as ℓ increases,

$$(\tau^2 \underset{\sim\sim}{H} \underset{\sim\sim}{W}^{-1} \underset{\sim\sim}{V} \underset{\sim\sim}{W}^{-1} \underset{\sim}{\bar{H}})_{jk} \approx d_j \, I_{jk} \, .$$

Essentially, the assumption serves to omit from the likelihood the terms arising from end-effects in a finite realisation of an infinite process.

By substitution of the approximation for $(\tau^2 \underset{\sim\sim}{H} \underset{\sim\sim}{W}^{-1} \underset{\sim\sim}{V} \underset{\sim\sim}{W}^{-1} \underset{\sim}{\bar{H}})$, equation (5.2.3) becomes

$$L_M \approx \frac{1}{2} \ell n\left(|\underset{\sim}{W}|^2 |\underset{\sim}{\bar{H}}| \, |\underset{\sim}{H}| \prod_{j=1}^{n} d_j\right) + \frac{1}{2} \sum_{g=1}^{n} \frac{1}{d_g} \sum_{j=1}^{n} \sum_{k=1}^{n} (a_j \bar{H}_{jg})(H_{gk} a_k) \, ,$$

$$= \frac{1}{2} \sum_{j=1}^{n} \ell n(W_{jj}^2 d_j)$$

$$+ \frac{1}{2n} \sum_{g=1}^{n} \frac{1}{d_g} \sum_{j=1}^{n} \sum_{k=1}^{n} a_j a_k e^{-2\pi\imath(j-1)(g-1)/n} \, e^{2\pi\imath(g-1)(k-1)/n} \, ,$$

$$= \frac{1}{2} \sum_{j=1}^{n} \ell n(W_{jj}^2 d_j) + \frac{1}{2} \sum_{k=1}^{n} D_k/d_k \, , \qquad (5.2.4)$$

by the definition of $\underset{\sim}{D}$ in equation (5.2.2). This is the negative of the logarithm of the pseudo-likelihood, proposed by Robinson (1978), which will henceforth be denoted L_p. It has therefore been demonstrated that the pseudo-likelihood is an approximation to the Gaussian likelihood.

5.2.3 Estimation using REGAME

The computer program REGAME has not been designed to fit models by optimizing functions of the form given above for L_p. However, it can be tricked into doing so.

The negative log-likelihood of an n-vector of zero observations $\underset{\sim}{0}$, which are specified to be independently normally distributed with the jth observation having mean $W_{jj}\sqrt{D_j}$ and variance $W_{jj}^2 d_j$, has the same form as equation (5.2.4). REGAME can therefore be used to estimate regression parameters in $\underset{\sim}{f}$ and spectral parameters in $\underset{\sim}{d}$ by minimizing L_p. At each call, REGMOD has to set the jth pseudo-fitted value to $W_{jj}\sqrt{D_j}$, p and q to 0 and the jth diagonal element of the pseudo-variance matrix to $W_{jj}^2 d_j$.

Also of note on a computational theme, the periodogram ($\underset{\sim}{D}$) can be calculated very efficiently using the Fast Fourier Transform, for example using NAG library routine C06FAF (Numerical Algorithms Group, 1983).

5.2.4 Variances of parameter estimators

As the sample size (n) increases, L_p approximates more closely the negative log-likelihood, so maximum pseudo-likelihood estimators share the same asymptotic properties as maximum likelihood estimators already discussed in [1.3.8]. Robinson (1978) derived more general results for the spectral parameter estimators without resorting to the assumption of multivariate normality. In section 4.3, approximate variances of parameter estimators have been obtained by inverting the Hessian matrix of L_p using REGAME. The emphasis in this chapter is on a first attempt at modelling the spectrum of the error process, so the attention given in chapter 4 to checking asymptotic properties by small simulations and deriving likelihood-based marginal confidence intervals has not been repeated in this chapter.

5.3 Colquhoun's data

5.3.1 Choice of model

An important assumption made in [5.2.2], to derive L_p as an approximation to the Gaussian log-likelihood, was that the autocorrelation coefficients of the error process decayed rapidly to zero as the lags increased. However, it has already been shown in [4.3.4] that if the regression model is a single exponential function, then the error autocorrelations in Colquhoun's data decay very slowly. Therefore, the technique of estimation by minimizing L_p is likely to perform better when a double exponential function is fitted, for which the error autocorrelations decay much more quickly.

The first step in the analysis, as it was in chapter 4, is to fit the regression function by least-squares estimation, that is by minimizing the negative log-likelihood on the assumption that the errors are independent. It will first be shown that, for this particular correlation structure, L_p is identical to the negative log-likelihood.

5.3.2 Pseudo-likelihood when errors are independent

If errors are independent then the autocorrelation function is

$$\rho_\ell = \tau^2 \qquad \text{if } \ell = 0 ,$$
$$= 0 \qquad \text{otherwise,}$$

and from equation (5.2.1), the spectrum is flat, that is

$$d_g = \tau^2 \qquad \text{for } g = 1, \ldots, n .$$

Therefore

$$L_p = \frac{1}{2}\sum_{j=1}^{n} \ell n(W_{jj}^2\tau^2) + \frac{1}{2}\sum_{g=1}^{n} D_g/\tau^2$$

from equation (5.2.4),

$$= \frac{1}{2}\sum_{j=1}^{n} \ell n(W_{jj}^2\tau^2) + \frac{1}{2\tau^2}\sum_{g=1}^{n}\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_j\, a_k\, e^{2\pi\imath(k-j)(g-1)/n}$$

by substituting for $\underset{\sim}{D}$ from equation (5.2.2),

$$= \frac{1}{2}\sum_{j=1}^{n} \ell n(W_{jj}^2\tau^2) + \frac{1}{2\tau^2}\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_j\, a_k\, n\, I_{jk} \quad ,$$

$$= \frac{1}{2}\sum_{j=1}^{n} \ell n(W_{jj}^2\tau^2) + \frac{1}{2\tau^2}\sum_{j=1}^{n} a_j^2 \quad .$$

This is also the negative log-likelihood of $\underset{\sim}{y}$, which was stated in [5.2.2], if the errors are independent.

5.3.3 Double exponential regression with independent errors

The regression parameters $(\underset{\sim}{\beta})$ in the double exponential function given in [4.3.7], were estimated by minimizing L_p, using REGAME, with $d_g = \tau^2$ for $g = 1,\ldots,n$ and $\underset{\sim}{W} = \underset{\sim}{I}$. The results were almost identical to the maximum likelihood estimates with independent errors given in table 4.3.9. They would have been exactly the same, but the NAG library Fast Fourier Transform algorithm (C06FAF) required n to have no prime factor exceeding 19, so n was increased from 124 to 125 and a_{125} was set to zero.

5.3.4 Spectrum when the autocorrelation function is a sum of exponentials

To continue the analogy with section 4.3, the next model to be considered is the double exponential regression with an

ARMA(1,1) error process. Before this can be fitted it is necessary to derive the spectral function for this process.

When the autocorrelation function is of the form considered in chapter 4, as defined in [4.2.1], but with $\delta = 0$, then

$$\rho_\ell = \sum_{j=1}^{p} \kappa_j e^{\ell\lambda_j} \qquad \text{for} \quad \ell > 0 .$$

The spectrum, from equation (5.2.1), and making use of the fact that ρ_ℓ is symmetric about zero, has gth element

$$d_g = \tau^2 \sum_{\ell=-\infty}^{-1} \sum_{j=1}^{p} \kappa_j e^{-\ell\lambda_j} e^{2\pi\imath\ell(g-1)/n}$$

$$+ \tau^2 \sum_{j=1}^{p} \kappa_j$$

$$+ \tau^2 \sum_{\ell=1}^{\infty} \sum_{j=1}^{p} \kappa_j e^{\ell\lambda_j} e^{2\pi\imath\ell(g-1)/n} ,$$

$$= \tau^2 \sum_{j=1}^{p} \kappa_j \left\{ \frac{e^{\lambda_j - 2\pi\imath(g-1)/n}}{1-e^{\lambda_j - 2\pi\imath(g-1)/n}} + 1 + \frac{e^{\lambda_j + 2\pi\imath(g-1)/n}}{1-e^{\lambda_j + 2\pi\imath(g-1)/n}} \right\} ,$$

$$= \tau^2 \sum_{j=1}^{p} \kappa_j \frac{1-e^{2\lambda_j}}{(1-e^{\lambda_j - 2\pi\imath(g-1)/n})(1-e^{\lambda_j + 2\pi\imath(g-1)/n})} ,$$

$$= \tau^2 \sum_{j=1}^{p} \kappa_j \frac{1-e^{2\lambda_j}}{1-2e^{\lambda_j}\cos(2\pi(g-1)/n) + e^{2\lambda_j}} .$$

5.3.5 Double exponential regression with ARMA(1,1) error model

With the representation of the error process as a sum of positively-correlated Markov processes defined in [4.2.4], the regression parameters ($\underset{\sim}{\beta}$) and error model parameters ($\underset{\sim}{\alpha}$) for an ARMA(1,1) process were jointly estimated by minimizing L_p. The

parameter estimates are almost the same as the maximum likelihood estimates given in table 4.3.11. Figure 5.3.1 shows the first half of the periodogram $\underset{\sim}{D}$ and of the fitted spectrum $\underset{\sim}{\hat{d}}$ plotted against the number of harmonic cycles. Only the first half need be plotted because the second half is a mirror image of the first because

$$D_g = D_{n+2-g}$$

and

$$d_g = d_{n+2-g} \qquad \text{for } g = 2, \ldots, n .$$

The fitted spectral function is high at low numbers of cycles, or alternatively at low frequency, and plateaus to a low level at higher frequencies. The high value at low frequencies corresponds to the correlated component of the error process and is an attempt to fit the high values of the periodogram at 2,3 and 4 cycles. The plateau level of the spectrum corresponds to the white noise component of the error process.

Figure 5.3.2 shows the first half of the observed and fitted autocorrelation functions obtained as the Fourier transforms of $\underset{\sim}{D}$ and $\underset{\sim}{\hat{d}}$ respectively. In this case also, the second halves are mirror images of the first and so have not been plotted. The fitted autocorrelation function exhibits the appropriate form for an ARMA(1,1) process: an exponential decay to zero as the lag increases after an initial step change between lags zero and one. The observed autocorrelations agree well with the fit up to a lag of ten, but beyond that point the fitted autocorrelations no longer follow the trends in observed values. However, it should be borne in mind that the estimates of autocorrelation at adjacent lags can be highly correlated, so apparent trends may be no more than random fluctuations about an expectation of zero.

Figure 5.3.1

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by maximum pseudo-likelihood estimation; periodogram (X) of departures of the data from the fitted regression, and fitted spectrum (——), plotted against the number of cycles completed by each sinusoid component.

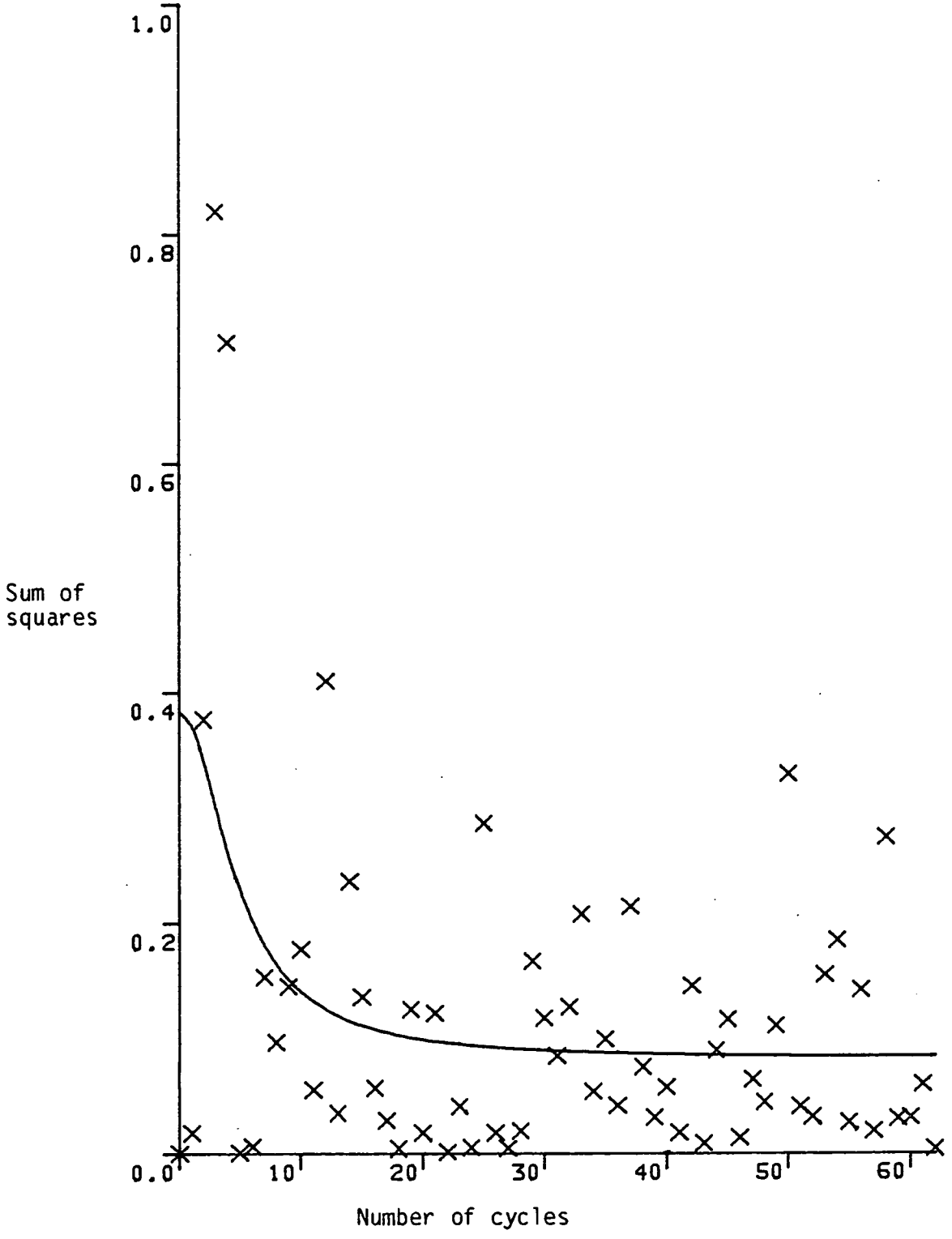
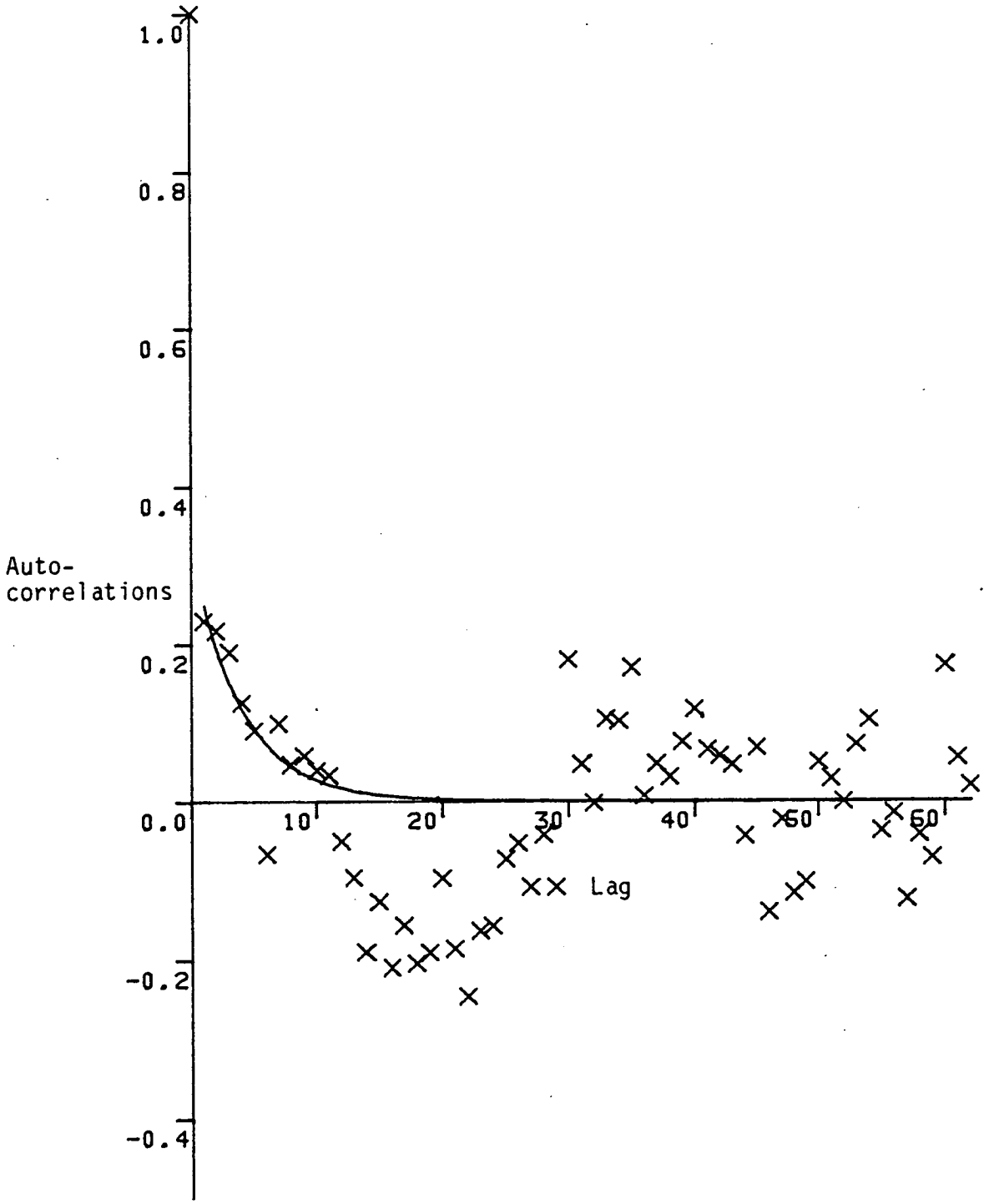


Figure 5.3.2

Colquhoun's data, double exponential regression with ARMA(1,1) error model, fitted by maximum pseudo-likelihood estimation; the autocorrelations (X) of the departures of the data from the fitted regression, and fitted autocorrelation function (——), plotted against the lag separation.



5.3.6 Double exponential regression with an empirical non-increasing error spectrum

The models considered so far have already been fitted in section 4.3. For the spectral approach to be of any use we have to consider spectral functions which are simpler than their corresponding autocorrelation functions. The first step is to examine the periodogram, plotted in figure 5.3.1, to identify a suitable spectral function. The main feature of the periodogram, as already commented on above, is that a few terms at low frequencies have high values. It is reasonable to suppose that the error terms are positively correlated with each other at low lags and therefore the spectrum is larger at low frequencies than at high frequencies.

On the assumption that $\underset{\sim}{d}$ is non-increasing at low frequencies one empirical parameterization is

$$d_g = d_{n+2-g} = \tau^2 \left(1 + \sum_{j=g}^{\ell} \mu_j^{*2}\right) \qquad \text{for } g = 1,\ldots,(n+2)/2 ,$$

for some small value of ℓ. The regression parameters $(\underset{\sim}{\beta})$ and error model parameters $(\underset{\sim}{\mu}^*)$ were jointly estimated by minimizing L_p for two values of ℓ (5 and 10). In both cases, all estimated coefficients in $\underset{\sim}{\mu}^*$ were zero except for μ_5^*. The results are given in table 5.3.1 with $\underset{\sim}{\mu}^*$ excluded except for μ_5^*. The fitted spectrum is shown in figure 5.3.3 and the fitted autocorrelation function is plotted in figure 5.3.4. The fitted spectrum is a step function approximation to the previous fitted spectrum in figure 5.3.1. The fitted autocorrelation function is also similar to that in figure 5.3.2, although it is no longer monotonically decreasing. L_p has been reduced from -77.4 for

Table 5.3.1

Colquhoun's data, double exponential regression with empirical non-increasing error spectrum, fitted by maximum pseudo-likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	β_5	μ_5^*	τ^2
estimate	-90.2	22.6	3.15	60.8	8.94	2.0	0.09
se	(0.4)	(6.9)	(0.67)	(7.1)	(0.70)	(0.6)	(0.01)
correlations							
β_2	-0.89						
β_3	-0.81	0.96					
β_4	0.87	-0.996	-0.98				
β_5	-0.94	0.99	0.94	-0.98			
μ_5^*	0.21	-0.19	-0.17	0.18	-0.19		
τ^2	-0.06	0.05	0.05	-0.05	0.05	-0.28	

Correlation coefficients exceeding 0.9 underlined.

Figure 5.3.3

Colquhoun's data, double exponential regression with empirical non-increasing error spectrum, fitted by maximum pseudo-likelihood estimation; periodogram (X) of departures of the data from the fitted regression, and fitted spectrum (——), plotted against the number of cycles completed by each sinusoid component.

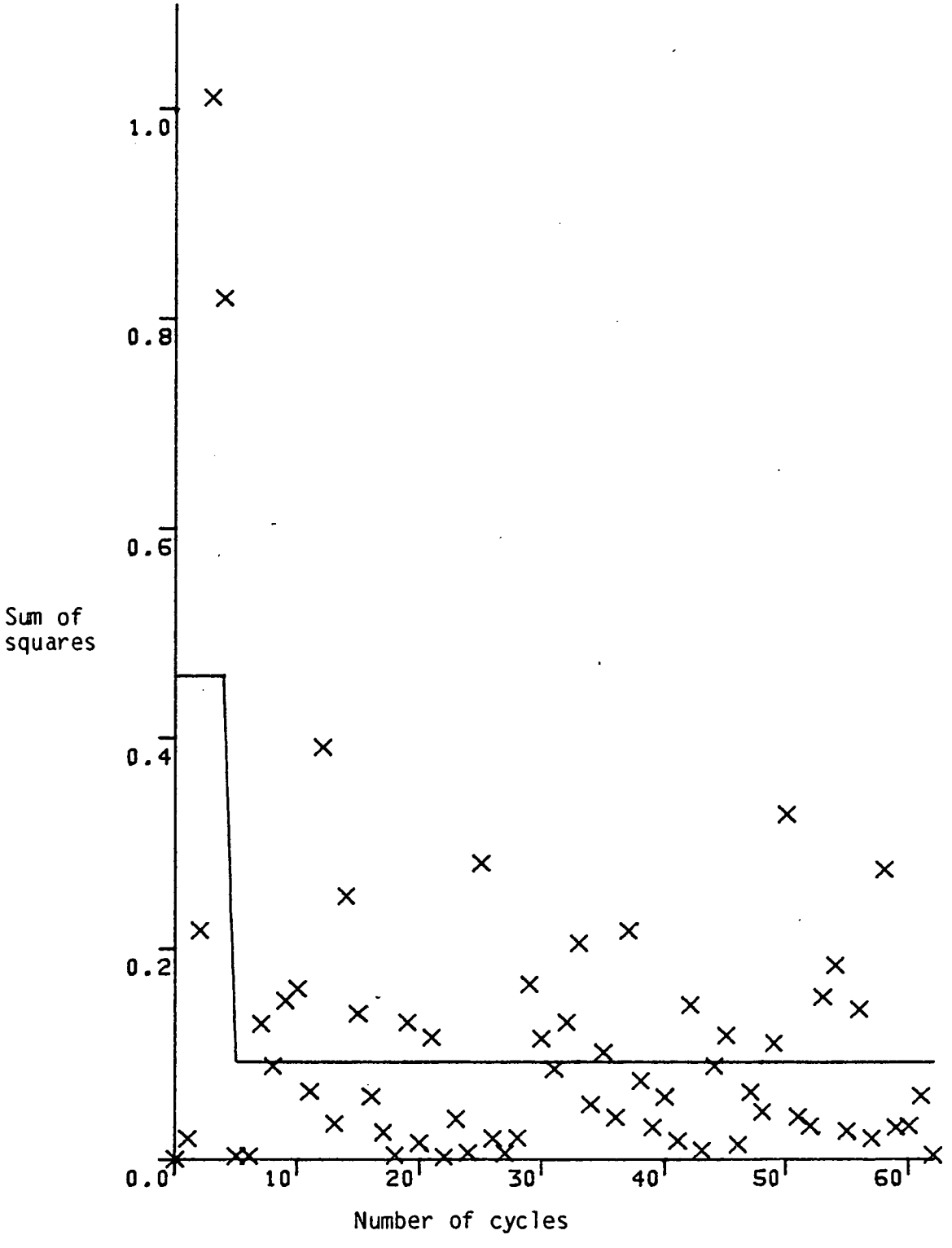
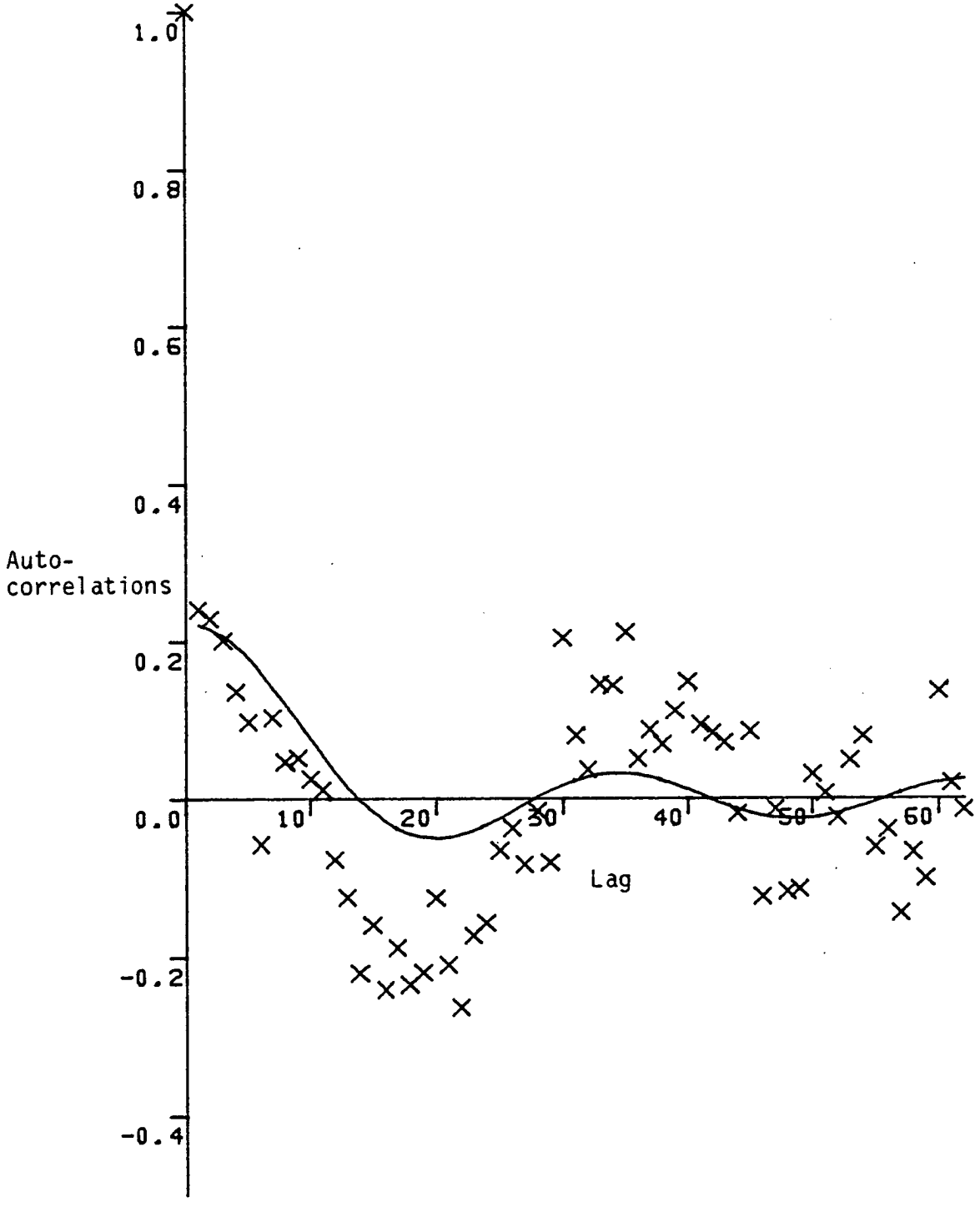


Figure 5.3.4

Colquhoun's data, double exponential regression with empirical non-increasing error spectrum, fitted by maximum pseudo-likelihood estimation; the autocorrelations (X) of the departures of the data from the fitted regression, and fitted autocorrelation function (——), plotted against the lag separation.



the ARMA(1,1) model to -79.2 with one less parameter being used (although some selection bias will have been introduced in choosing which parameter to use), so the new error model would appear to fit better. The regression parameter estimates and standard errors are in close agreement with those in table 4.3.11. At this stage, the spectral approach to modelling the error structure in order to estimate regression parameters has worked well.

5.3.7 Double exponential regression with an empirical error spectrum

To test the assumption that the error spectrum is non-increasing, $\underset{\sim}{d}$ was reparameterized as

$$d_g = d_{n+2-g} = \tau^2(1 + \mu_g^2) \qquad \text{for } g = 1,\ldots,\ell\ ,$$
$$= \tau^2 \qquad \text{for } g = (\ell+1),\ldots,(n+2)/2\ ,$$

for some small value of ℓ. The regression parameters $(\underset{\sim}{\beta})$ and error model parameters $(\underset{\sim}{\mu})$ were jointly estimated by minimizing L_p with ℓ equal to 5. Of the coefficients in $\underset{\sim}{\mu}$, only μ_4 and μ_5 were estimated to be non-zero. The parameter estimates are given in table 5.3.2 and the spectrum and autocorrelation function are plotted in figures 5.3.5 and 5.3.6 respectively. The observed and estimated autocorrelations are now in excellent agreement for all lags. L_p has been reduced to -83.4 for the addition of one extra parameter compared with the model in [5.3.6]. Therefore, on the basis of the asymptotic properties of the likelihood ratio test, the hypothesis of a non-increasing spectral function is rejected. However, the appropriateness of asymptotic results can be called into question for this model because the fitted autocorrelation function does not

Table 5.3.2

Colquhoun's data, double exponential regression with empirical error spectrum, fitted by maximum pseudo-likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_4	β_5	μ_4	μ_5	τ^2
estimate	-89.9	17.6	2.48	66.3	8.45	3.8	3.4	0.09
se	(0.2)	(2.1)	(0.36)	(2.5)	(0.24)	(2.1)	(1.9)	(0.01)
correlations								
β_2	<u>-0.90</u>							
β_3	-0.78	0.85						
β_4	0.87	<u>-0.95</u>	<u>-0.97</u>					
β_5	<u>-0.94</u>	<u>0.97</u>	<u>0.91</u>	<u>-0.98</u>				
μ_4	0.16	-0.18	-0.15	0.16	-0.16			
μ_5	0.13	-0.13	-0.15	0.15	-0.14	0.05		
τ^2	-0.04	0.04	0.04	-0.04	0.04	-0.13	-0.13	

Correlation coefficients exceeding 0.9 underlined.

Figure 5.3.5

Colquhoun's data, double exponential regression with empirical error spectrum, fitted by maximum pseudo-likelihood estimation; periodogram (X) of departures of the data from the fitted regression, and fitted spectrum (____), plotted against the number of cycles completed by each sinusoid component.

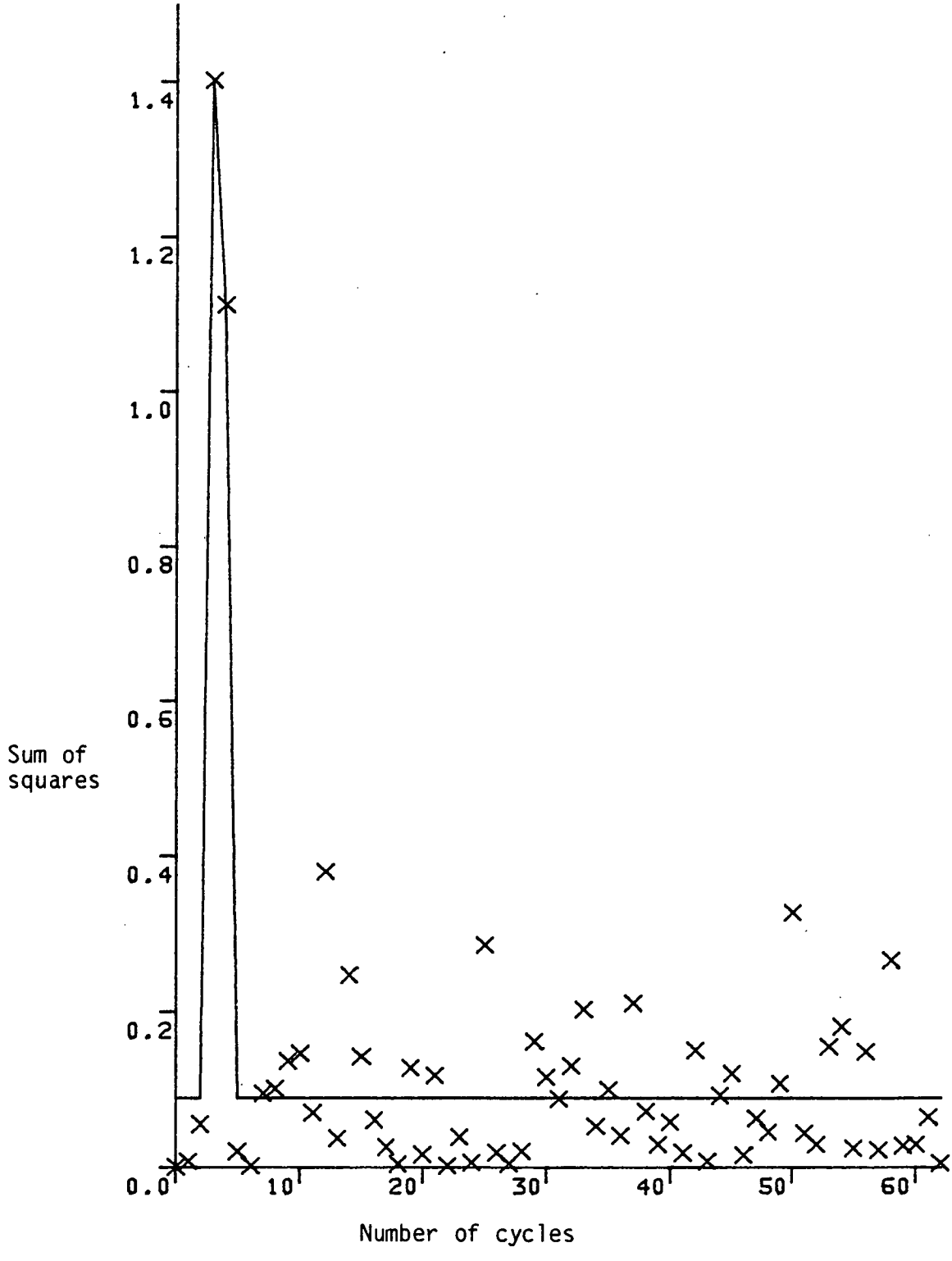
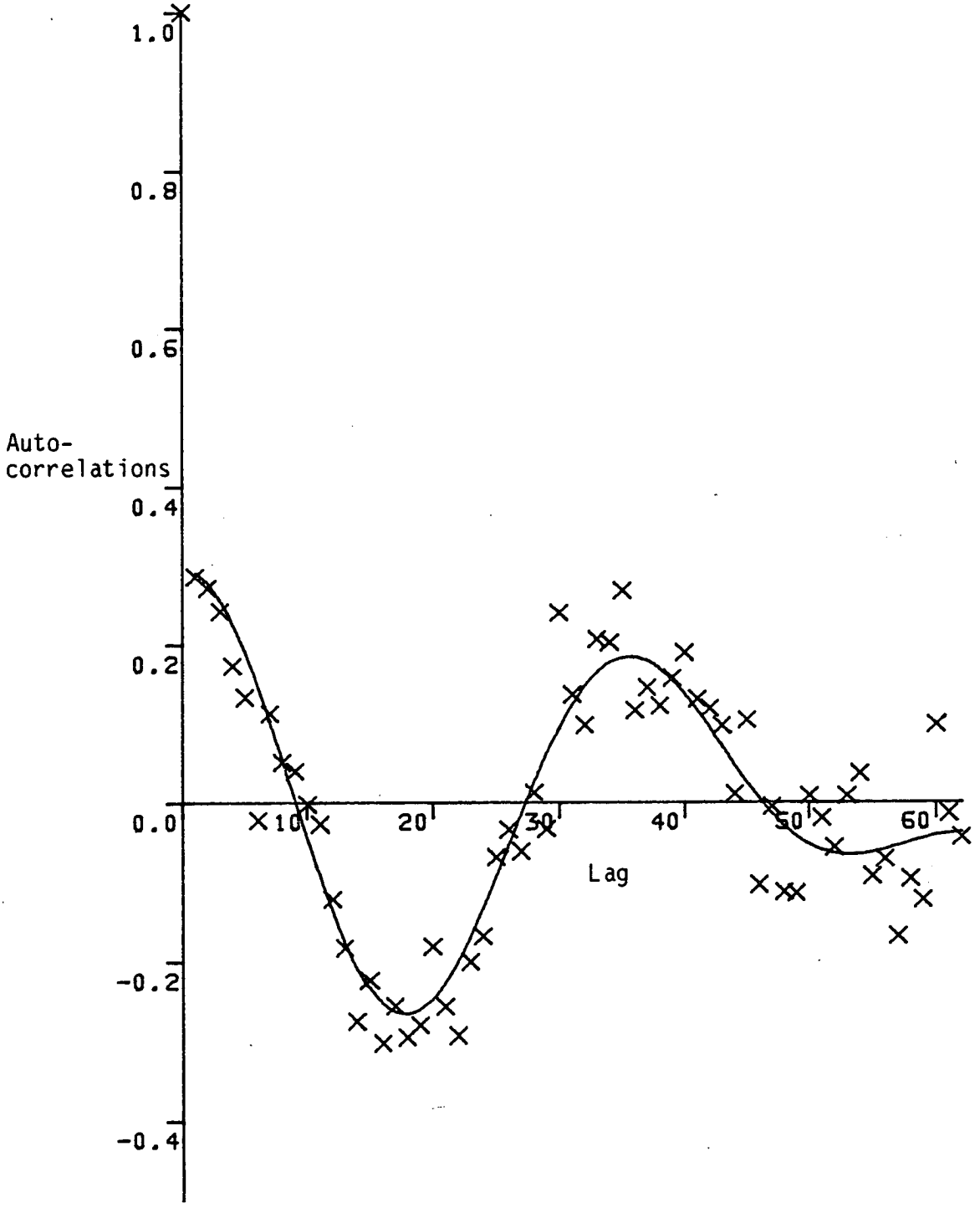


Figure 5.3.6

Colquhoun's data, double exponential regression with empirical error spectrum, fitted by maximum pseudo-likelihood estimation; the autocorrelations (X) of the departures of the data from the fitted regression, and autocorrelation function (——), plotted against the lag separation.



decay rapidly to zero, so the pseudo-likelihood fails to approximate the Gaussian likelihood. The regression parameter estimates have been changed substantially from those in tables 4.3.9, 4.3.11 and 5.3.1. The standard errors have decreased, probably because of the high negative autocorrelations in the error process for lags between 10 and 25 which is evident in figure 5.3.6.

5.3.8 Resumé

Double exponential regression parameters have been estimated for three assumed error spectral functions. The results when using the spectrum of a ARMA(1,1) process or an empirically chosen non-increasing function were in good agreement with the results obtained in section 4.3 when the autocorrelation of the error process was being modelled. However, when the constraint that the spectral function should be non-increasing was relaxed, the new estimates of the regression parameters were inconsistent with earlier results. The reasons for, and conclusions from, these results will be discussed in section 5.4. Colquhoun's data will be re-used as an example in chapters 7 and 9.

5.4 Discussion

The basic technique presented in this chapter is to fit a regression function to a series of observations by least-squares estimation, identify a spectral function by examination of the periodogram of the departures of the data from the fitted regression, and then jointly estimate the regression and error spectrum parameters by maximizing the pseudo-likelihood.

This method is more restrictive than that in chapter 4 because observations must be equally spaced in time, the autocorrelations of the error process must decay rapidly to zero, and the estimation method is an approximation to maximum likelihood, rather than to residual maximum likelihood estimation which was found to perform better in chapter 4. Also, if the error process is found to be non-stationary the technique in this chapter is unuseable, whereas the estimation methods in chapter 4 can still be used. The autocorrelations have to decay rapidly to zero for the pseudo-likelihood to be of use because end-effects in the error process are omitted. If the autocorrelations were to remain high these end-effects would not be of negligible magnitude. It is probably because this assumption is violated in the final model considered (in [5.3.7]) that the results were out of step with earlier regression parameter estimates.

Finally, the major problem, when estimating regression parameters in the presence of an unknown error spectral structure, is that the technique in this chapter rests heavily upon gross assumptions that are made about the spectral function of the error process. This is the same problem encountered in chapter 4 and is the motivation for the approach taken in chapter 9.

6. Models based on linear stochastic difference and differential equations

6.1 Introduction

Up to this point in the thesis, empirical models have been used to describe error processes in order to estimate regression parameters when errors are serially correlated. In other words, the data themselves have been used to identify the correlation structure among the errors. In this and the following two chapters a fundamentally different strategy will be adopted; instead of the data being used, the processes by which the data were generated will be used to identify the correlation structure. At least that is the theory, in practice the distinction is rather more blurred. This approach has been termed "mechanistic modelling" and was discussed in section 1.2.

Sometimes, regression functions fitted to data are selected because they are solutions of difference or differential equations. There is then an opportunity to add an error component directly into the original equation, rather than onto its deterministic solution, in order to model observations. This approach was used implicitly by Dhanoa and Le Du (1982) in modelling the milk yields of cows, and explicitly by Sandland and McGilchrist (1979) in modelling growth.

The simplest types of difference and differential equations are those which are linear with coefficients that are constant over time. General solutions of these have already been given in chapter 3. In sections 6.2, first-order equations are considered with input variates appropriate to generating lactation curves, and

the relationship between models in discrete and continuous time is explored. In an analogous fashion, in section 6.3 second-order equations are considered with input variates appropriate to generating animal growth curves. Neilson's data on the milk yields of cows are modelled by first-order differential equations in section 6.4 and the goodness-of-fit is tested. Similarly, in section 6.5 Thiessen's data on the weights of cattle are modelled by second-order differential equations. Finally, in section 6.6 the usefulness of stochastic difference and differential equations in estimating regression parameters in the presence of serially correlated errors is critically assessed.

The novelty in the modelling of lactation lies in the respecification of Dhanoa and Le Du's (1982) model, which makes clearer its relationship with other models and enables it to be extended to unequally spaced observation times, and in generalizing the model to test its goodness-of-fit. Regarding growth, the use of second-order stochastic differential equations is original, and may either be viewed as a generalization of Finney's (1958) and Mitchell's (1968) first-order models to permit a sigmoidal trend or a linearization of the "more biological" differential equations considered by Sandland and McGilchrist (1979).

6.2 A model for lactation using first-order equations

6.2.1 Model in discrete time

Dhanoa and Le Du (1982) proposed a model in which the observed milk yields $\underset{\sim}{y}$ from a single cow are a realisation of a

random vector $\underset{\sim}{Y}$ which is generated by the first-order difference equation

(6.2.1) $$Y_t = \phi Y_{t-1} + Z_t \qquad \text{for } t \geq 1 ,$$

where ϕ is a parameter, $Y_0 \equiv 0$, and Z_t is independently normally distributed for $t \geq 1$, with mean $(\alpha_1 + \alpha_2 t)$ and variance σ^2.

6.2.2 Solution for the model in discrete time

Equation (6.2.1) may be solved quite simply, but for illustrative purposes the general solution derived in [3.2.2] will be used. This gives

$$Y_t = Y_0 K^*_{11} e^{t\lambda^*_1} + \sum_{k=1}^{t} Z_k K^*_{11} e^{(t-k)\lambda^*_1} ,$$

where $e^{\lambda^*_1}$ is the root of

$$x - \phi = 0 ,$$

and K^*_{11} satisfies

$$K^*_{11} = 1 ,$$

with $\underset{\sim}{\lambda}^*$ and $\underset{\sim}{K}^*$ used in place of $\underset{\sim}{\lambda}$ and $\underset{\sim}{K}$ to distinguish these parameters from the ones used with the model in continuous time. Therefore, because Y_0 is identically zero,

$$Y_t = \sum_{k=1}^{t} Z_k \phi^{t-k} .$$

It follows that

$$E(Y_t) = \sum_{k=1}^{t} E(Z_k)\phi^{t-k}$$

$$= \sum_{k=1}^{t} (\alpha_1 + \alpha_2 k)\phi^{t-k}$$

$$= \alpha_1 \frac{(1-\phi^t)}{(1-\phi)} + \alpha_2 \frac{(\phi^{t+1} - (t+1)\phi + t)}{(1-\phi)^2}$$

$$= \{\frac{\alpha_1}{(1-\phi)} - \frac{\alpha_2\phi}{(1-\phi)^2}\}\{1-\phi^t\} + \frac{\alpha_2}{(1-\phi)} t \qquad (6.2.2)$$

by algebraic manipulation, and

$$\text{cov}\,(Y_s,Y_t) = \sum_{i=1}^{s}\sum_{j=1}^{t} \text{cov}(Z_i,Z_j)\,\phi^{s-i}\,\phi^{t-j}$$

$$= \sum_{i=1}^{t} \sigma^2\,\phi^{s-i}\,\phi^{t-i} \qquad \text{when } s \geq t$$

because Z_i and Z_j are uncorrelated when $i \neq j$,

$$= \sigma^2\,\phi^{s-t}\,\frac{(1-\phi^{2t})}{(1-\phi^2)} . \qquad (6.2.3)$$

As $\underset{\sim}{Z}$ is a GARMA(0,0) process it follows from [3.2.3] that $\underset{\sim}{Y}$ is a GARMA(1,0) process. This is also evident from equation (6.2.3). Equation (6.2.2) is an exponential function plus a linear trend and passes through the origin, and was proposed by Cobby and Le Du (1978) as an alternative regression model in place of Wood's (1967) curve:

$$E(Y_t) = \beta_1 t^{\beta_2} e^{-t\beta_3} ,$$

which has become the standard model for lactation data. For appropriate values of the parameters α_1, α_2 and ϕ equation (6.2.2) has the characteristic shape of a lactation curve: it rises rapidly to a maximum and then decreases more slowly towards zero.

6.2.3 Model in continuous time

The model may also be formulated in continuous time, which serves to shed light on the relationship between discrete and

continuous processes. It is assumed that the observations $\underset{\sim}{y}$ are a realisation in discrete time of a random variate $Y(t)$ indexed by t over a continuous interval, which is generated by the first-order differential equation

$$Y^{\langle 1\rangle}(t) - \xi\, Y(t) = Z(t) \qquad \text{for } t \geq 0 ,$$

where the term in angular brackets denotes the order of derivative with respect to t, ξ is a parameter, $Y(0) \equiv 0$, and $Z(t)$ is a random variate over a continuous interval with expectation $(\beta_1+\beta_2 t)$ and $(Z(t)-\beta_1-\beta_2 t)$ is the first derivative of the Weiner process with variance τ^2 for $t \geq 0$. Strictly speaking, this definition of $Z(t)$ is nonsense because the Weiner process is nowhere differentiable, but this does not invalidate the definition of $Y(t)$, which can be modified as discussed in [3.3.1].

6.2.4 Solution for the model in continuous time

Applying the results of [3.3.2] for illustrative purposes, rather than solving the differential equation directly, gives

$$Y(t) = A^{\{0\}} K_{11}\, e^{t\lambda_1} + \int_0^t Z(u)\, \kappa_1\, e^{(t-u)\lambda_1}\, du ,$$

where

$$A^{\{0\}} = Y(0) ,$$

λ_1 is the root of

$$x - \xi = 0 ,$$

K_{11} satisfies

$$K_{11} = 1 ,$$

and κ_1 satisfies

$$\xi_0\, \kappa_1 = \psi_0 ,$$

where ξ_0 and ψ_0 are defined to be -1. Therefore, because $Y(0)$ is identically zero,

$$Y(t) = \int_0^t Z(u)\, e^{(t-u)\xi}\, du \ .$$

It follows that

$$E(Y(t)) = \int_0^t (\beta_1+\beta_2 u)\, e^{(t-u)\xi}\, du \ ,$$

$$= \left(\frac{\beta_1}{\xi} + \frac{\beta_2}{\xi^2}\right) e^{t\xi} - \left(\frac{\beta_1}{\xi} + \frac{\beta_2}{\xi^2} + \frac{\beta_2 t}{\xi}\right) \ ,$$

$$= -\left(\frac{\beta_1}{\xi} + \frac{\beta_2}{\xi^2}\right)(1 - e^{t\xi}) - \frac{\beta_2}{\xi}\, t \qquad (6.2.4)$$

by algebraic manipulation, and

$$\mathrm{cov}(Y(s),Y(t)) = \int_0^s \int_0^t \mathrm{cov}(Z(u),Z(v))\, e^{(s-u)\xi}\, e^{(t-v)\xi}\, dv\, du \ ,$$

$$= \tau^2 \int_0^t e^{(s-u)\xi}\, e^{(t-u)\xi}\, du \qquad \text{when } s \geq t$$

because $Z(u)$ and $Z(v)$ are uncorrelated when $u \neq v$,

$$= \tau^2\, e^{(s-t)\xi}\, \frac{(e^{2t\xi} - 1)}{2\xi} \ . \qquad (6.2.5)$$

As $Z(t)$ is a cGARMA(0,ε) process, where ε is an arbitrarily small positive number, from [3.3.3] $Y(t)$ is a cGARMA(1,ε) process. However, it is evident from equation (6.2.5) which is valid when s equals t , and not just when s is greater than t , that $Y(t)$ is a cGARMA(1,0) process. Equation (6.2.4) is of the same functional form as equation (6.2.2) and the regression model proposed by Cobby and Le Du (1978).

6.2.5 Relation between models in discrete and continuous time

Provided ϕ is positive, the models in discrete and continuous time are equivalent because equation (6.2.2) is the same as equation (6.2.4) with

$$\phi = e^{\xi},$$

$$\frac{\alpha_1}{(1-\phi)} - \frac{\alpha_2\phi}{(1-\phi)^2} = -\left(\frac{\beta_1}{\xi} + \frac{\beta_2}{\xi^2}\right) \ ,$$

and $$\frac{\alpha_2}{(1-\phi)} = -\frac{\beta_2}{\xi} ,$$

and equation (6.2.3) is the same as equation (6.2.5) with

$$\frac{\sigma^2}{(1-\phi^2)} = -\frac{\tau^2}{2\xi} ,$$

where the relationship between ϕ and ξ ensures that σ^2 and τ^2 are of the same sign.

The model in continuous time will be used in section 6.4 because it ensures the desirable property that all correlations are positive and is more appropriate for the observation times in Neilson's data, which are unequally spaced.

6.3 A model for growth using second-order equations

6.3.1 Justification for choice of model

Many growth curves are derived as the solutions of deterministic difference or differential equations because the rate of growth is a fundamental variable in the process of growth. However, Sandland and McGilchrist (1979) comment that

> "Deterministic differential equations do not seem fully appropriate here. Growth is embedded in a stochastically fluctuating environment and it would be attractive to attempt to incorporate this environmental randomness into the class of growth models, rather than, as an afterthought, tacking a residual onto the expected value, which is either a polynomial in time or a solution to a differential equation."

It is difficult to work with stochastic versions of most biologically derived differential equations because they are non-linear. Linear equations are more mathematically tractable and may be justified either as approximations to non-linear equations or, more empirically and more specifically, the second-order equation is the simplest equation capable of yielding the

characteristic sigmoidal shape of a growth curve.

An alternative route to the second-order differential equation is via the work of Parks (1975a,b). He proposed that an animal's weight at time t, denoted $Y(t)$, satisfies

$$Y^{<1>}(t) - \mu_1 Y(t) = Z(t)$$

where the term in angular brackets denotes the order of derivative with respect to t, μ_1 is a constant and $Z(t)$ is proportional to the food intake of the animal at time t. A simple model of intake is

$$Z^{<1>}(t) - \mu_2 Z(t) = \mu_3 ,$$

with constants μ_2 and μ_3. After substitution for $Z(t)$ and $Z^{<1>}(t)$, this becomes

$$(Y^{<2>}(t) - \mu_1 Y^{<1>}(t)) - \mu_2 (Y^{<1>}(t) - \mu_1 Y(t)) = \mu_3 ,$$

which is a second-order differential equation. The analogous model in discrete time is a second-order difference equation.

6.3.2 Model in discrete time

Observations $\underset{\sim}{y}$ of an animal's weight are assumed to be a realisation of a random vector $\underset{\sim}{Y}$ whose elements are generated by the second-order difference equation

(6.3.1) $$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t \qquad \text{for } t>2 .$$

Here ϕ_1 and ϕ_2 are parameters, Z_t is a first-order moving average process with mean α for $t>2$, and Y_0 and Y_1 are constants.

6.3.3 Solution for the model in discrete time

Applying the results of [3.2.2] we get

(6.3.2) $$Y_t = \sum_{i=1}^{2} Y_{i-1} \sum_{j=1}^{2} K^*_{ij} e^{t\lambda^*_j} + \sum_{k=2}^{t} Z_k \sum_{\ell=1}^{2} K^*_{2\ell} e^{(t+1-k)\lambda^*_\ell} ,$$

where $e^{\lambda_1^*}$, $e^{\lambda_2^*}$ are the roots of

$$x^2 - \phi_1 x - \phi_2 = 0 ,$$

and the elements of $\underset{\sim}{K}^*$ satisfy

$$K_{11}^* e^{t\lambda_1^*} + K_{12}^* e^{t\lambda_2^*} = I_{1,t+1} \qquad \text{for } t = 0,1 ,$$

$$K_{21}^* e^{t\lambda_1^*} + K_{22} e^{t\lambda_2^*} = I_{2,t+1} \qquad \text{for } t = 0,1 ,$$

where $\underset{\sim}{I}$ is the identity matrix of size 2. (Parameters $\underset{\sim}{\lambda}^*$ and $\underset{\sim}{K}^*$ have been used in place of $\underset{\sim}{\lambda}$ and $\underset{\sim}{K}$ to distinguish them from the parameters used with the model in continuous time.) Therefore

$$K_{11}^* = \frac{e^{\lambda_2^*}}{(e^{\lambda_2^*} - e^{\lambda_1^*})} , \quad K_{12}^* = \frac{- e^{\lambda_1^*}}{(e^{\lambda_2^*} - e^{\lambda_1^*})} ,$$

$$K_{21}^* = \frac{- 1}{(e^{\lambda_2^*} - e^{\lambda_1^*})} , \quad K_{22}^* = \frac{1}{(e^{\lambda_2^*} - e^{\lambda_1^*})} .$$

It follows from equation (6.3.2) that

$$E(Y_t) = \sum_{i=1}^{2} Y_{i-1} \sum_{j=1}^{2} K_{ij}^* e^{t\lambda_j^*} + \sum_{k=2}^{t} \alpha \sum_{\ell=1}^{2} K_{2\ell}^* e^{(t+1-k)\lambda_\ell^*} ,$$

$$= \sum_{j=1}^{2} \{Y_0 K_{1j}^* e^{t\lambda_j^*} + Y_1 K_{2j}^* e^{t\lambda_j^*} + \alpha K_{2j}^* \frac{(e^{\lambda_j^*} - e^{t\lambda_j^*})}{(1 - e^{\lambda_j^*})}\} ,$$

$$= \sum_{j=1}^{2} \frac{\alpha K_{2j}^* e^{\lambda_j^*}}{(1 - e^{\lambda_j^*})} + \sum_{k=1}^{2} \{Y_0 K_{1k}^* + Y_1 K_{2k}^* - \frac{\alpha K_{2k}^*}{(1-e^{\lambda_k^*})}\} e^{t\lambda_k^*} \qquad (6.3.3)$$

by algebraic manipulation, and

$$\operatorname{cov}(Y_s, Y_t) = \sum_{i=2}^{s} \sum_{j=2}^{t} \operatorname{cov}(Z_i, Z_j) \sum_{k=1}^{2} \sum_{\ell=1}^{2} K_{2k}^* K_{2\ell}^* e^{(s+1-i)\lambda_k^*} e^{(t+1-j)\lambda_\ell^*} .$$

If $\underset{\sim}{Z}$ is a first-order moving average process with parameter θ, then $\operatorname{var}(Z_i) = \sigma^2(1+\theta^2)$, $\operatorname{cov}(Z_i, Z_{i-1}) = - \sigma^2\theta$, and Z_i and Z_j are uncorrelated if $|i-j| > 2$. Therefore, when $s \geq t+1$

$$\operatorname{cov}(Y_s,Y_t) = \sum_{j=3}^{t} \operatorname{cov}(Z_{j-1},Z_j) \sum_{k=1}^{2} \sum_{\ell=1}^{2} K^*_{2k}K^*_{2\ell}\, e^{(s+2-j)\lambda^*_k}\, e^{(t+1-j)\lambda^*_\ell}$$

$$+ \sum_{j=2}^{t} \operatorname{cov}(Z_j, Z_j) \sum_{k=1}^{2} \sum_{\ell=1}^{2} K^*_{2k}K^*_{2\ell}\, e^{(s+1-j)\lambda^*_k}\, e^{(t+1-j)\lambda^*_\ell}$$

$$+ \sum_{j=2}^{t} \operatorname{cov}(Z_{j+1},Z_j) \sum_{k=1}^{2} \sum_{\ell=1}^{2} K^*_{2k}K^*_{2\ell}\, e^{(s-j)\lambda^*_k}\, e^{(t+1-j)\lambda^*_\ell},$$

$$= \sum_{k=1}^{2} \sum_{\ell=1}^{2} \frac{\sigma^2 K^*_{2k}K^*_{2\ell}}{(1-e^{\lambda^*_k+\lambda^*_\ell})} \{-\theta(e^{(s+2-t)\lambda^*_k+\lambda^*_\ell} - e^{s\lambda^*_k+(t-1)\lambda^*_\ell})$$

$$+(1+\theta^2)(e^{(s+1-t)\lambda^*_k+\lambda^*_\ell} - e^{s\lambda^*_k+t\lambda^*_\ell}) - \theta(e^{(s-t)\lambda^*_k+\lambda^*_\ell} - e^{(s-1)\lambda^*_k+t\lambda^*_\ell})\},$$

$$= \sum_{k=1}^{2} \sum_{\ell=1}^{2} \frac{\sigma^2 K^*_{2k}K^*_{2\ell}}{(1-e^{\lambda^*_k+\lambda^*_\ell})} \{e^{s\lambda^*_k+t\lambda^*_\ell}(\theta e^{-\lambda^*_\ell} - (1+\theta^2) + \theta e^{-\lambda^*_k}) \tag{6.3.4}$$

$$- e^{(s-t)\lambda^*_k}(\theta e^{2\lambda^*_k+\lambda^*_\ell} - (1+\theta^2)e^{\lambda^*_k+\lambda^*_\ell} + \theta e^{\lambda^*_\ell})\}.$$

$\underset{\sim}{Y}$ is a GARMA(2,2) process by the definition in [2.2.1] because, when $s>t+1$, equation (6.3.4) is the sum of two terms, each of which is a product of a term $e^{s\lambda^*_k}$ dependent only on s and a term dependent only on t. However, $\underset{\sim}{Z}$ is a GARMA(0,1) process, so it follows from the last part of [3.2.3] that $\underset{\sim}{Y}$ has the more specific property of being a GARMA(2,1) process. Equation (6.3.3) consists of a constant term plus the weighted sum of two, possibly complex, exponential functions. Therefore, for appropriate values of the parameters it can take on the sigmoidal shape characteristic of growth curves.

6.3.4 Model in continuous time

Analogous with section 6.2, the growth model may also be formulated in continuous time in order to reveal some aspects of

the relationship between discrete and continuous processes. It is assumed that observations $\underset{\sim}{y}$ are the realisations in discrete time of a random variate $Y(t)$, indexed by t over a continuous interval, which is generated by the second-order differential equation

(6.3.5) $Y^{\langle 2\rangle}(t) - \xi_1 Y^{\langle 1\rangle}(t) - \xi_2 Y(t) = Z^{\langle 1\rangle}(t) - \psi_1 Z(t)$ for $t \geq 0$,

where the terms in angular brackets denote the orders of derivative with respect to t, ξ_1, ξ_2 and ψ_1 are parameters, $Y(0)$ and $(Y^{\langle 1\rangle}(0) - Z(0))$ are constants, and $Z(t)$ has an expectation of β and $(Z(t)-\beta)$ is the first derivative of a Weiner process with variance τ^2 for $t>0$. (A stationary version of this model was used by Phadke and Wu (1974) to model sunspot numbers.)

6.3.5 Solution for the model in continuous time

Applying the results of [3.3.2] we get

$$(6.3.6)\quad Y(t) = \sum_{i=1}^{2} A^{\{i-1\}} \sum_{j=1}^{2} K_{ij} e^{t\lambda_j} + \int_0^t Z(u) \sum_{k=1}^{2} \kappa_k e^{(t-u)\lambda_k} du ,$$

where

$$A^{\{i-1\}} = Y^{\langle i-1\rangle}(0) - \sum_{h=0}^{i-2} Z^{\langle h\rangle}(0) \sum_{k=1}^{2} \kappa_k \lambda_k^{i-2-h}$$

for $i=1,2$,

λ_1, λ_2 are the roots of

$$x^2 - \xi_1 x - \xi_2 = 0 ,$$

the elements of $\underset{\sim}{K}$ satisfy

$$K_{11}\lambda_1^t + K_{12}\lambda_2^t = I_{1,t+1} \qquad \text{for } t = 0,1,$$

$$K_{21}\lambda_1^t + K_{22}\lambda_2^t = I_{2,t+1} \qquad \text{for } t = 0,1,$$

where $\underset{\sim}{I}$ is the identity matrix of size 2, and the elements of $\underset{\sim}{\kappa}$ satisfy

$$\xi_0 \sum_{k=1}^{2} \kappa_k \lambda_k + \xi_1 \sum_{\ell=1}^{2} \kappa_\ell = \phi_1 ,$$

$$\xi_0 \sum_{k=1}^{2} \kappa_k = \phi_0 ,$$

where $\xi_0 = \phi_0 = -1$. Therefore

$$A^{\{0\}} = Y(0) , \qquad A^{\{1\}} = Y^{<1>}(0) - Z(0) ,$$

$$K_{11} = \frac{\lambda_2}{\lambda_2 - \lambda_1} , \qquad K_{12} = \frac{-\lambda_1}{\lambda_2 - \lambda_1} ,$$

$$K_{21} = \frac{-1}{\lambda_2 - \lambda_1} , \qquad K_{22} = \frac{1}{\lambda_2 - \lambda_1} ,$$

$$\kappa_1 = \frac{\lambda_2 - \xi_1 + \phi_1}{\lambda_2 - \lambda_1} \qquad \kappa_2 = \frac{-\lambda_1 + \xi_1 - \phi_1}{\lambda_2 - \lambda_1} .$$

It follows from equation (6.3.6) that

$$E(Y(t)) = \sum_{i=1}^{2} E(A^{\{i-1\}}) \sum_{j=1}^{2} K_{ij} e^{t\lambda_j} + \int_0^t \beta \sum_{k=1}^{2} \kappa_k e^{(t-u)\lambda_k} du ,$$

$$= \sum_{j=1}^{2} \{ Y(0) K_{1j} e^{t\lambda_j} + (E(Y^{<1>}(0)) - \beta) K_{2j} e^{t\lambda_j} + \beta\kappa_j \left(\frac{e^{t\lambda_j} - 1}{\lambda_j} \right) \} ,$$

$$(6.3.7) \qquad = -\sum_{j=1}^{2} \frac{\beta\kappa_j}{\lambda_j} + \sum_{k=1}^{2} \{ Y(0) K_{1k} + E(Y^{<1>}(0)) K_{2k} - \beta K_{2k} + \frac{\beta\kappa_k}{\lambda_k} \} e^{t\lambda_k}$$

by algebraic manipulation, and

$$\mathrm{cov}(Y(s), Y(t)) = \int_0^t \tau^2 \sum_{k=1}^{2} \sum_{\ell=1}^{2} \kappa_k \kappa_\ell e^{(s-u)\lambda_k} e^{(t-u)\lambda_\ell} du$$

for $s \geq t$,

because $Z(s)$ and $Z(t)$ are uncorrelated when $s \neq t$,

$$(6.3.8) \qquad = \tau^2 \sum_{k=1}^{2} \sum_{\ell=1}^{2} \kappa_k \kappa_\ell \left(\frac{e^{s\lambda_k + t\lambda_\ell} - e^{(s-t)\lambda_k}}{(\lambda_k + \lambda_\ell)} \right) .$$

As $Z(t)$ is a cGARMA$(0,\varepsilon)$ process, where ε is an arbitrarily small positive number, from [3.3.3] $Y(t)$ is a

cGARMA(2,ε) process. In fact it is evident from equation (6.3.8) above which holds when s is equal to t, and not only when s exceeds t, that $Y(t)$ has the more specific property of being a cGARMA(2,0) process. The functional form of equation (6.3.7) is the same as that for equation (6.3.3) which has already been discussed.

6.3.6 Relation between models in discrete and continuous time

The relation between models in discrete and continuous time is much more complicated in the case of second-order equations than in the case of first-order equations considered in [6.2.5]. What follows is a brief discussion of the similarities and differences, with no attempt made at rigorous proof.

In the model in discrete time, $e^{\lambda_1^*}$ and $e^{\lambda_2^*}$ are the roots of a real quadratic equation and are therefore either both real or a complex conjugate pair; whereas in the model in continuous time, λ_1 and λ_2 are the quadratic roots so that e^{λ_1} and e^{λ_2} are either both real and positive or a complex conjugate pair. The model in continuous time is evidently more restrictive than the model in discrete time. With the restriction that λ_1^* and λ_2^* are not real and negative, the expectations given by equations (6.3.3) and (6.3.7) are equivalent for all integer values of t provided that

$$\lambda_1^* = \lambda_1 ,$$

$$\lambda_2^* = \lambda_2 ,$$

$$\sum_{j=1}^{2} \alpha \frac{K_{2j}^* e^{\lambda_j^*}}{(1-e^{\lambda_j^*})} = \sum_{j=1}^{2} \frac{-\beta \kappa_j}{\lambda_j} ,$$

and $$Y_0K^*_{1k}+Y_1K^*_{2k} - \frac{\alpha K^*_{2k}}{(1-e^{\lambda^*_k})} = Y(0)K_{1k} + E(Y^{<1>}(0))K_{2k} - \beta K_{2k} + \beta\frac{\kappa_k}{\lambda_k}$$

for k=1,2.

For given values of $\underset{\sim}{\kappa}$ these five equations uniquely define ξ_1, ξ_2, β, $Y(0)$ and $E(Y^{<1>}(0))$ in terms of ϕ_1, ϕ_2, α, Y_0, Y_1, and conversely.

For the covariances given by equations (6.3.4) and (6.3.8) to be equivalent for integer values of s and t we require that

$$(6.3.9)\quad \sum_{\ell=1}^{2} \sigma^2 \frac{K^*_{2k}K^*_{2\ell}}{(1-e^{\lambda^*_k+\lambda^*_\ell})}\{\theta e^{2\lambda^*_k+\lambda^*_\ell} - (1+\theta^2)e^{\lambda^*_k+\lambda^*_\ell} + \theta e^{\lambda^*_\ell}\} = \sum_{\ell=1}^{2} \tau^2\frac{\kappa_k\kappa_\ell}{(\lambda_k+\lambda_\ell)} \qquad \text{for } k = 1,2$$

in order to equate the coefficients of $e^{(s-t)\lambda_k}$, and that

$$(6.3.10)\quad \sigma^2 \frac{K^*_{2k}K^*_{2\ell}}{(1-e^{\lambda^*_k+\lambda^*_\ell})}\{\theta e^{-\lambda^*_\ell} - (1+\theta^2) + \theta e^{-\lambda^*_k}\} = \tau^2\frac{\kappa_k\kappa_\ell}{(\lambda_k+\lambda_\ell)}$$

for $k = 1,2$, $\ell = 1,2$

in order to equate the coefficients of $e^{s\lambda_k+t\lambda_\ell}$. Equation (6.3.9) defines τ^2 and ψ_1 (included in $\underset{\sim}{\kappa}$) in terms of σ^2 and θ, and conversely, but these do not satisfy equation (6.3.10). Provided that $|e^{\lambda_1}|$ and $|e^{\lambda_2}|$ are less than unity, $e^{s\lambda_k+t\lambda_\ell}$ decreases to zero as s and t increase and the contribution to $\mathrm{cov}(Y(s), Y(t))$ from the components given in equation (6.3.10) declines to zero, so the models are equivalent asymptotically.

In order to achieve equivalence between the models for all times it is necessary to include a starting distribution other than the one so far considered in which initial variances are zero. This change has the effect of adding a term

$$\sum_{i=1}^{2}\sum_{j=1}^{2} \mathrm{cov}(Y_{i-1},Y_{j-1})\, K^{*}_{ik}\, K^{*}_{j\ell}$$

to the left side of equation (6.3.10) and a term

$$\sum_{i=1}^{2}\sum_{j=1}^{2} \mathrm{cov}(A^{\{i-1\}},A^{\{j-1\}})\, K_{ik}\, K_{j\ell}$$

to the right side. The variance of $(A^{\{0\}}, A^{\{1\}})$ is defined by, and defines, the variance of (Y_0,Y_1) and the models in discrete and continuous time are equivalent.

The model in continuous time will be used in section 6.5 to maintain similarity with section 6.4.

6.4 Neilson's data: milk yield of cows

6.4.1 Introduction

Daily milk yields were recorded approximately once a week from the date of calving for each of 23 British Friesian cows in a commercial herd at Langhill (Neilson, Whittemore, Lewis, Alliston, Roberts, Hodgson-Jones, Mills, Parkinson and Prescott, 1983). The number of measurements of milk yield per cow varied between 32 and 44 and were unequally spaced in time. The animals were part of a larger experiment to study the relationships between feed intake, milk output and body condition in high yielding dairy cows.

6.4.2 Models

A vector of observations $y_1, \ldots, y_n$, denoted $\underset{\sim}{y}$, of daily milk yields made on a single cow at times $t_1, \ldots, t_n$, denoted

$\underset{\sim}{t}$, measured in weeks are assumed to be multivariate normally distributed with mean $\underset{\sim}{f}$ where

$$f_i = -(\frac{\beta_1}{\xi} + \frac{\beta_2}{\xi^2})(1 - e^{\xi t_i}) - \frac{\beta_2}{\xi} t_i \qquad \text{for } i = 1,\ldots,n ,$$

as given in equation (6.2.4). This regression model was previously proposed by Cobby and Le Du (1978) because, for appropriate values of the parameters β_1, β_2 and ξ , it has the desirable features of a lactation curve of passing through the origin, rising rapidly to a maximum and subsequently decreasing much more slowly.

Several different models for the variance matrix $(\underset{\sim}{V})$ will be considered:

(0) Independent observations with equal variances τ^2 :

$$V_{ij} = \tau^2 I_{ij} ,$$

where $\underset{\sim}{I}$ is the identity matrix of size n .

(1) The standard variance structure appropriate to a first-order stochastic differential equation, for which

$$V_{ij} = \tau^2 \, e^{\xi(t_i - t_j)} \, \frac{(e^{2\xi t_j} - 1)}{2\xi} \qquad \text{for } i > j ,$$

as given in equation (6.2.5).

(2) The standard variance structure in (1) above but with a separate parameter ξ^*, independent of the parameter ξ in $\underset{\sim}{f}$, for which

$$V_{ij} = \tau^2 \, e^{\xi^*(t_i - t_j)} \, \frac{(e^{2\xi^* t_j} - 1)}{2\xi^*} \qquad \text{for } i > j .$$

(3) The standard variance structure in (1) plus an observation error, for which

$$V_{ij} = \tau^2[(\frac{1}{2} + \frac{1}{2}\sin\alpha^*)e^{\xi(t_i - t_j)} \, \frac{(e^{2\xi t_j} - 1)}{2\xi} + (\frac{1}{2} - \frac{1}{2}\sin\alpha^*)I_{ij}]$$

$$\text{for } i > j .$$

This uses the same type of parameterization as [4.2.4] to constrain terms to lies between 0 and 1.

(4) The variance structure with a separate parameter ξ^* plus an observation error, for which

$$V_{ij} = \tau^2[(\frac{1}{2} + \frac{1}{2}\sin\alpha^*)e^{\xi^*(t_i-t_j)}\frac{(e^{2\xi^* t_j} - 1)}{2\xi^*} + (\frac{1}{2} - \frac{1}{2}\sin\alpha^*)I_{ij}]$$

for $i \geq j$.

Models (2) to (4) enable the validity of Dhanoa and Le Du's (1982) model to be tested both for the connection between ξ in $\underset{\sim}{f}$ and ξ^* in $\underset{\sim}{V}$, and for the inclusion of an observation error. The particular choice of parameterization of the observation errors given by (3) and (4) above has been made to ensure that the parameters can fit the full range of possible models, whilst remaining bounded, without the use of more than one scaling parameter τ^2 . If α^* equals $\pi/2$ there is no observation error, and if α^* equals $-\pi/2$ observation error is the sole source of error. Model (0) corresponds to a GARMA(0,0) process, models (1) and (2) to GARMA(1,0) processes and models (3) and (4) to GARMA(1,1) processes.

It should be noted that these simple models do not take special account of any effects resulting from the onset of pregnancy or changes in environment such as leaving winter housing, rather they are subsumed within the general error model.

6.4.3 Method of estimation

The models were all fitted separately to the data from each cow by maximum likelihood estimation (that is by minimizing L_M), and by residual maximum likelihood estimation (that is by minimizing L_R), as defined in [2.3.2]. The results were in good

agreement, so only the maximum likelihood results have been presented because L_M is computationally simpler to evaluate, and the interpretation of L_R is not clear when parameters occur in both the regression function and in the variance matrix.

To obtain good initial estimates for the parameters in the iterative optimization routine the models were fitted sequentially. Parameter values were guessed from the data to provide starting values for model (0). The parameter estimates from model (0) were then used as starting values for model (1), those obtained for model (1) were used to start the iterations for models (2) and (3) and finally those for model (3) were used to start the iterations for model (4).

6.4.4 Results - likelihood comparisons

Table 6.4.1 shows the minimized values of L_M (the negative log-likelihood) for each model fitted to the data from each cow. For each cow the model with the lowest Akaike Information Criterion (AIC) (see Akaike, 1973), that is L_M plus the number of model parameters, is underlined. AIC has recently become popular in time-series analysis for choosing between models. On the basis of this criterion the choice of best model varies considerably among cows: on seven occasions model (0) is chosen, on eight occasions model (1) is chosen, on two occasions model (2) is chosen, once model (3) is chosen and on five occasions model (4) is chosen. When L_M is averaged over all cows AIC simply involves adding the number of model parameters as above, and the best fit is model (2).

Models (2), (3) and (4) are generalizations of model (1) so they can be compared on the basis of the asymptotic properties of the likelihood ratio test. Models (2) and (3) have one extra

Table 6.4.1

Neilson's data; minimum values of the negative log-likelihood (L_M) for different models.

		independent errors	correlation parameter		+ observation error	
			ξ	ξ^*	ξ	ξ^*
	Model	(0)	(1)	(2)	(3)	(4)
Cow	Number of Parameters	4	4	5	5	6
1		60.1	57.6	57.6	57.6	57.6
2		70.3	61.4	55.2*	61.4	55.2*
3		74.3	73.6	73.5	73.5	73.5
4		54.7	52.3	52.2	52.3	52.2
5		44.9	43.1	42.3	40.3*	38.1**
6		50.1	51.3	50.1	50.1	50.1
7		40.3	42.1	40.3	40.3	40.3
8		70.6	62.5	60.7	62.5	58.7**
9		34.8	40.2	34.8*	34.8*	34.8*
10		78.2	77.1	76.8	76.6	76.3
11		53.1	52.3	51.8	50.9	50.0
12		43.0	41.0	41.0	41.0	41.0
13		45.7	45.8	45.2	44.8	43.7
14		41.7	38.7	36.3*	38.7	34.3**
15		31.4	29.6	26.7*	29.6	26.6*
16		62.7	56.8	53.0*	56.8	50.6**
17		40.8	42.6	40.8	40.8	40.8
18		45.8	46.4	45.6	45.7	45.6
19		46.1	45.0	44.4	45.0	44.1
20		44.9	42.6	41.9	42.3	41.9
21		58.6	58.7	58.6	58.6	58.6
22		51.3	47.4	46.4	47.4	46.4
23		52.8	53.0	52.8	52.7	52.7
Average		52.0	50.5	49.0*	49.7	48.4*

For each cow the model with the lowest value of Akaike's Information Criterion has been underlined.

Models (2) to (4) with significantly lower L_M than model (1) at 5% level denoted by an asterisk.

Model (4) with significantly lower L_M than model (2) at 5% level denoted by a second asterisk.

parameter and so are a significant improvement over model (1) at the 5% level if L_M has been decreased by more than 1.9, that is the 95 percentile of $\frac{1}{2}\chi^2_1$. Model (4) has two extra parameters so a decrease in L_M of 3.0 (95 percentile of $\frac{1}{2}\chi^2_2$) is sufficient to reject model (1). Models which are a significant improvement are marked with an asterisk. In particular, model (4) is chosen in preference to model (1) on seven occasions.

Model (4) is also a generalization of model (2) with one extra parameter to specify the observation error, so a decrease of 1.9 is sufficient to reject model (2). Data sets for which model (2) is rejected are marked with a second asterisk in the table. Model (4) is prefered to model (2) on four occasions.

The same tests were also applied to the values of L_M for each model averaged over all cows. The significant decrease for a model with one extra parameter is 0.8 (the 95 percentile of $\frac{1}{46}\chi^2_{23}$) and with two extra parameters it is 1.4 (the 95 percentile of $\frac{1}{46}\chi^2_{46}$), once again appealing to the asymptotic properties of the likelihood ratio test. (Equivalent tests could have been performed using L_M summed over all cows.) On average, models (2) and (4) are better than model (1) but the improvement of (4) over (2) is not significant. However, the significant improvement from using model (4) for 4 cows is sufficient to justify its use because there is less than a 3 percent probability of model (2) being rejected on four or more occasions if it were in fact correct. The conclusion therefore is that Dhanoa and Le Du's model is found inadequate for the data set in two respects: the parameter ξ cannot be assumed to be common to both the regression function ($\underset{\sim}{f}$) and the variance matrix ($\underset{\sim}{V}$), and an observation error is necessary.

6.4.5 Results - parameter estimates

Parameter estimates and standard errors inevitably vary between cows for a particular model. In order to summarize the results three simple and widely used statistics have been derived: two statistics relate to the parameter estimates and are the mean and standard deviation of the between-cow distribution; the third statistic is the average standard error. To define these summary statistics more explicitly consider the parameter β_1 in a particular model: this was estimated as $\hat{\beta}_1^{\{j\}}$ with a standard error of s.e.$(\hat{\beta}_1^{\{j\}})$ using the data from the jth cow. The mean of the estimates is

$$\hat{\beta}_1 = \sum_{j=1}^{23} \hat{\beta}_1^{\{j\}}/23 ,$$

the standard deviation of the estimates is

$$\sqrt{\{ \sum_{j=1}^{23} (\hat{\beta}_1^{\{j\}} - \hat{\beta}_1)^2/22\}} ,$$

and the average standard error is

$$\sum_{j=1}^{23} \text{s.e.}(\hat{\beta}_1^{\{j\}})/23 .$$

Table 6.4.2 gives the results. These statistics are obviously discarding some information, but they are adequate for making some important comparisons between models and for assessing whether a parameter could take a common value for all cows.

In the table $e^{\hat{\xi}}$ and $e^{\hat{\xi}^*}$ are presented instead of $\hat{\xi}$ and $\hat{\xi}^*$ because on occasions the latter took large negative values. Also, $\hat{\tau}^2$ is not given because in models (1) to (4) it is too highly correlated with estimates of ξ and ξ^* to be interpretable when the latter took large negative values.

Table 6.4.2

Neilson's data; means and standard deviations of parameter estimates over all 23 cows for each model, and average errors (see [6.4.5]).

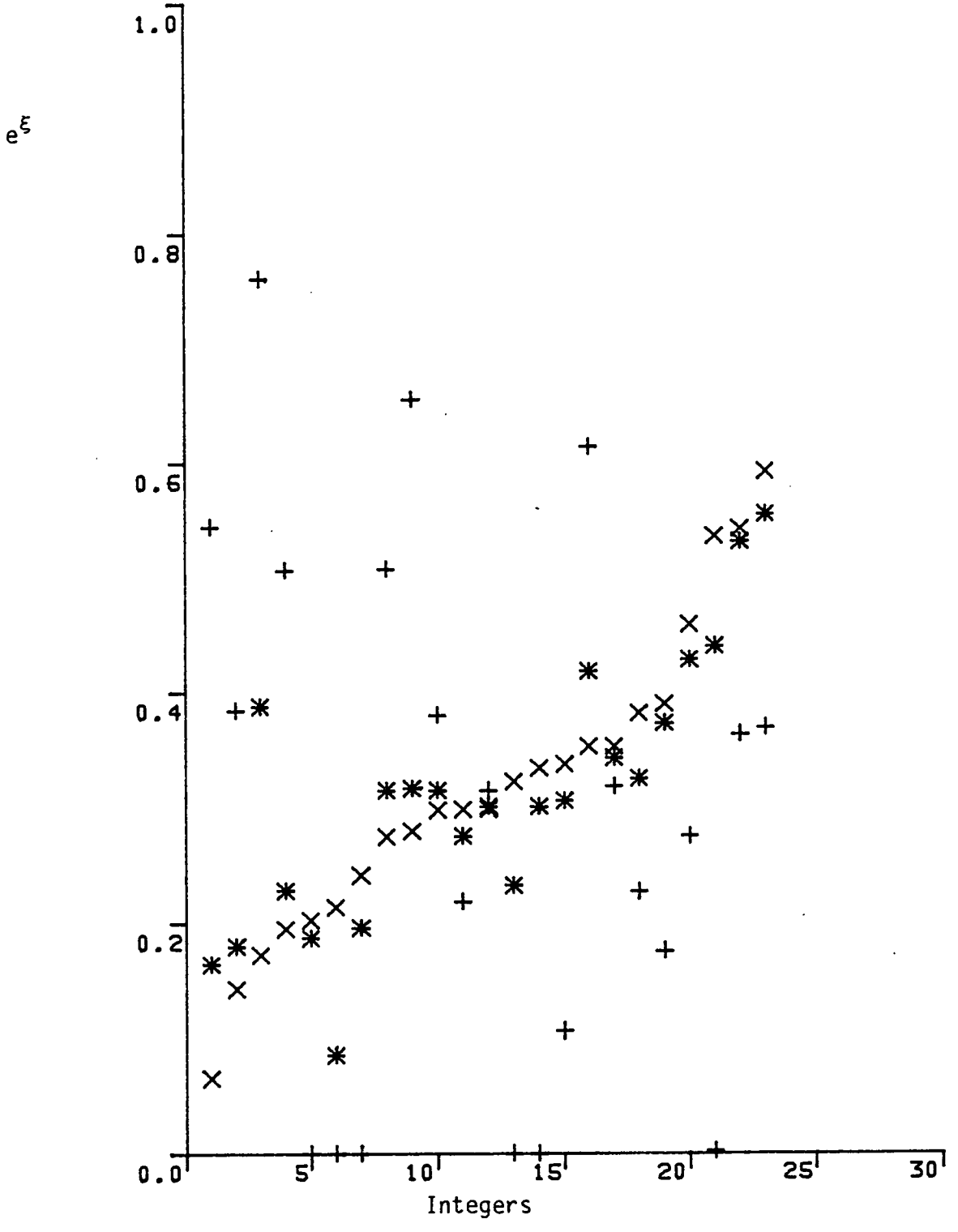
Model	β_1	β_2	e^{ξ}	e^{ξ^*}	α^*
		Means of estimates			
(0) indep. errors	46.	-0.82	0.33		
(1) cor. param. ξ	48.	-0.87	0.31		
(2) cor. param. ξ^*	47.	-0.85	0.32	0.29	
(3) ξ + obs. error	44.	-0.79	0.34		0.5
(4) ξ^* + obs.error	47.	-0.85	0.31	0.47	0.5
		standard deviations of estimates			
(0) indep. errors	17.	0.47	0.12		
(1) cor. param. ξ	24.	0.65	0.12		
(2) cor. param. ξ^*	18.	0.50	0.13	0.23	
(3) ξ + obs. error	17.	0.46	0.11		1.3
(4) ξ^* + obs.error	17.	0.49	0.12	0.34	1.0
		average standard errors			
(0) indep. errors	10.	0.20	0.07		
(1) cor. param. ξ	9.	0.16	0.06		
(2) cor. param. ξ^*	11.	0.21	0.07	0.11	
(3) ξ + obs. error	9.	0.18	0.07		1.3
(4) ξ^* + obs.error	11.	0.21	0.07	0.17	1.4

A feature of particular note in the table is the good agreement between the mean values of e^{ξ} (0.32) and e^{ξ^*} (0.29) in model (2). This seems to suggest model (1) is correct in assuming that ξ is equal to ξ^* and using a single parameter, and as a consequence reducing the average standard error of the estimator of e^{ξ} from 0.7 in models (0), (2), (3) and (4) to 0.6. However, although the average values of the estimates of e^{ξ} and e^{ξ^*} are in good agreement there is no evidence of a relationship between these two parameters for individual cows and the correlation coefficient between them over the 23 cows is -0.2 . Figure 6.4.1 shows the 23 estimates of e^{ξ} for model (2) ranked in ascending order, together with estimates of e^{ξ^*} for model (2) and estimates of e^{ξ} for model (1), arranged in the same order, plotted against the integers 1 to 23. From this it can be seen that using a single parameter biases the estimate of ξ . In fact the combined parameter consists of approximately 80% of the weight appropriate to the coefficient in the regression function and 20% to the coefficient appropriate to the correlation structure. Therefore, although the close agreement between the mean values of e^{ξ} and e^{ξ^*} in model (2) is of interest it is no justification for assuming that ξ and ξ^* are equal.

Although the likelihood ratio tests indicate significant differences between the models these do not seem to have any effect on average on the regression parameter estimates. The mean estimates, standard deviations and average standard errors are very similar for all models. This is probably because although the errors are significantly positively correlated this correlation coefficient is small in value so the assumption of independent errors does not reduce the efficiency by much. Also, standard

Figure 6.4.1

Neilson's data; the estimates of e^{ξ} (X) for model (2) ranked in ascending order, the estimates of e^{ξ^*} (+) for model (2) and the estimates of e^{ξ} (✳) for model (1), arranged in the same order, plotted against the integers 1 to 23.



deviations of estimates exceed the standard errors for all parameters except α^* , on average they are twice as big, so there is no evidence for parameters being constant over all cows.

Figures 6.4.2, 6.4.3, 6.4.4 and 6.4.5 show the data for cow number 8 together with the fitted regression function and predicted values using earlier observations (described in [2.3.6]) plotted against time for models (0), (1), (2) and (4) respectively. The eighth cow was selected as one of the four data sets for which model (4) gave the best fit. As such, it is not typical of the other cows, but rather serves to highlight the differences in fit between models. The fitted curve in model (1) reaches its linear asymptote less quickly than in the other three models because ξ is describing both the regression function and the correlation structure. The predicted values for model (4) give the best agreement with the data because this model can allow for more correlation between errors than model (1) whilst also allowing an observation error component which is absent from model (2).

6.4.6 Resumé

The evidence from Neilson's data is that the model advocated by Dhanoa and Le Du is inappropriate and if used will lead to biased estimates of ξ . An ARMA (1,1) process seems to be the simplest model capable of describing the error correlation structure, allowing for both persistent correlations and an observation error component. However, the difference this makes compared with least-squares estimation is small and does not appear to be sufficient to justify its use.

Figure 6.4.2

Neilson's data, cow number 8, model (0) fitted by maximum likelihood estimation; observed milk yields (X) and the fitted curve (——) plotted against time.

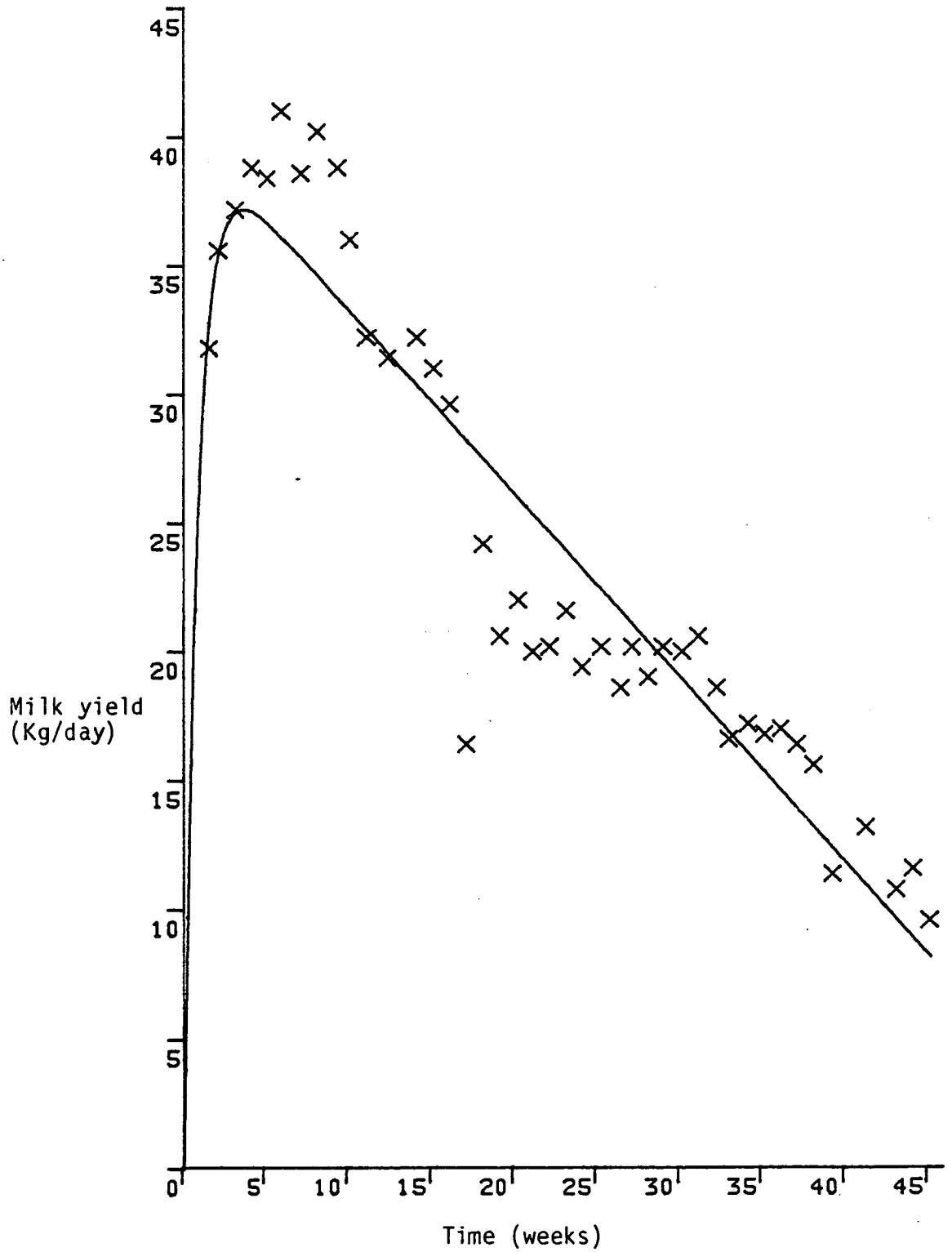


Figure 6.4.3

Neilson's data, cow number 8, model (1) fitted by maximum likelihood estimation; observed milk yields (X), the fitted curve (----) and the predicted values (——) conditional upon earlier observations plotted against time.

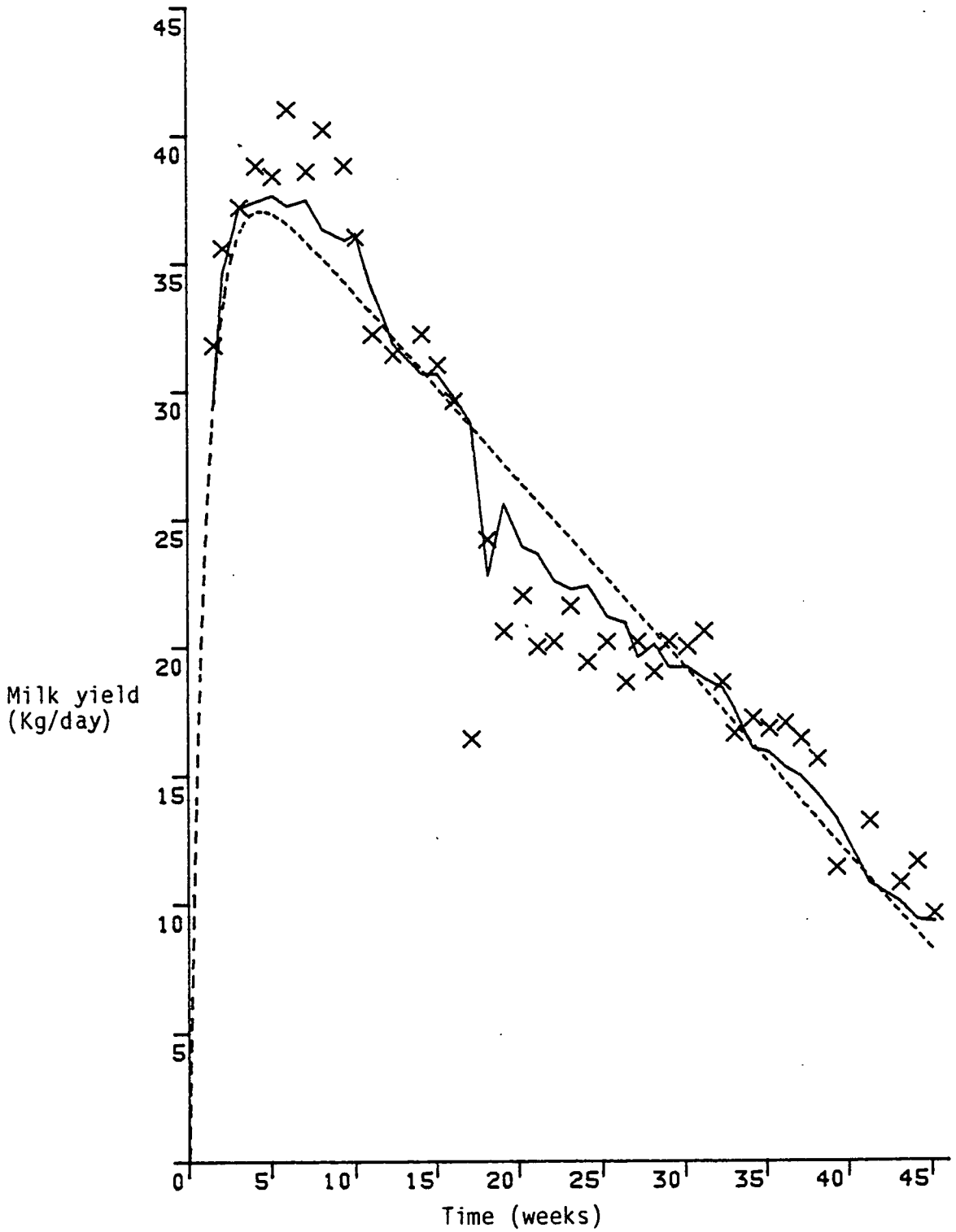


Figure 6.4.4

Neilson's data, cow number 8, model (2) fitted by maximum likelihood estimation; observed milk yields (X), the fitted curve (----) and the predicted values (____) conditional upon earlier observations plotted against time.

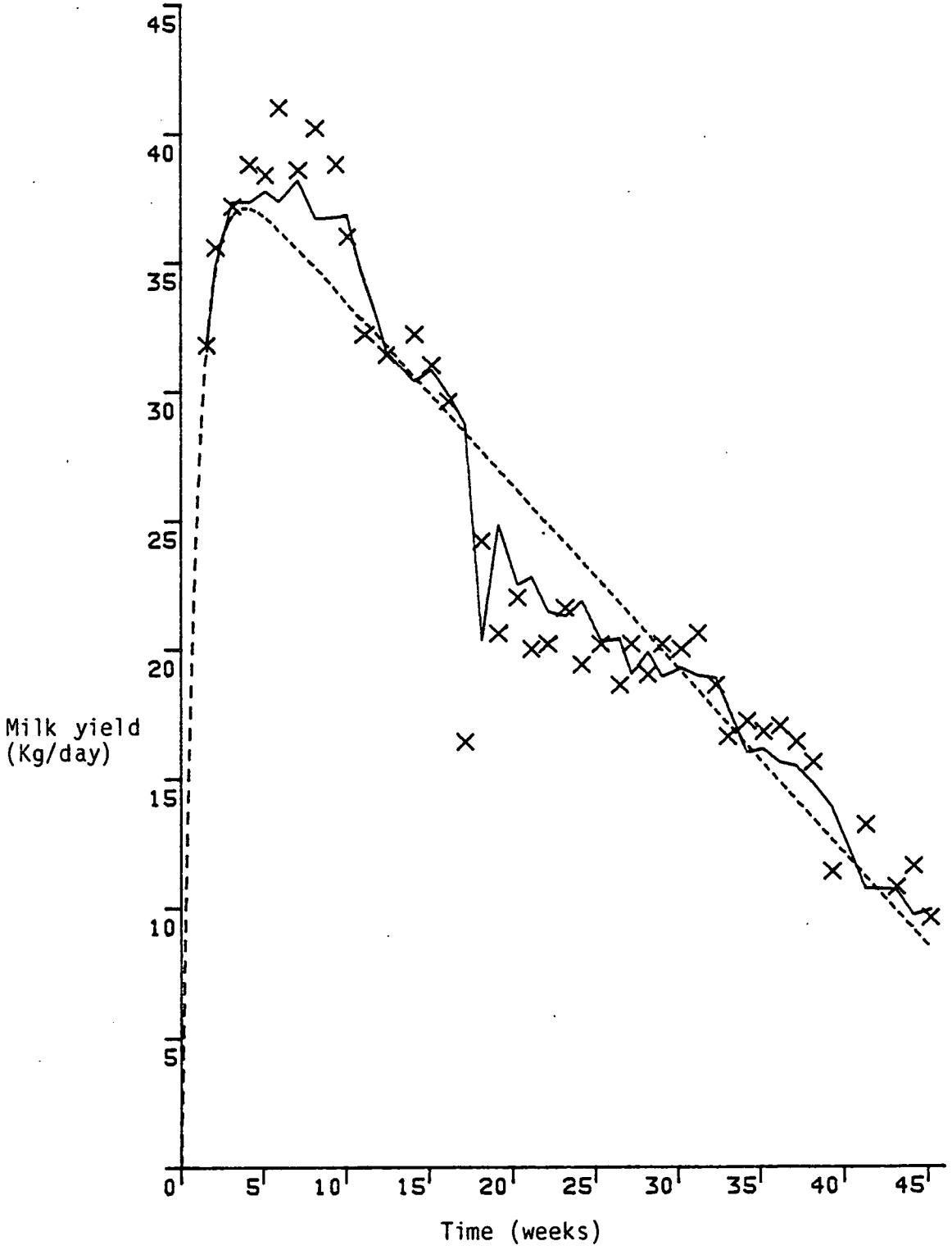
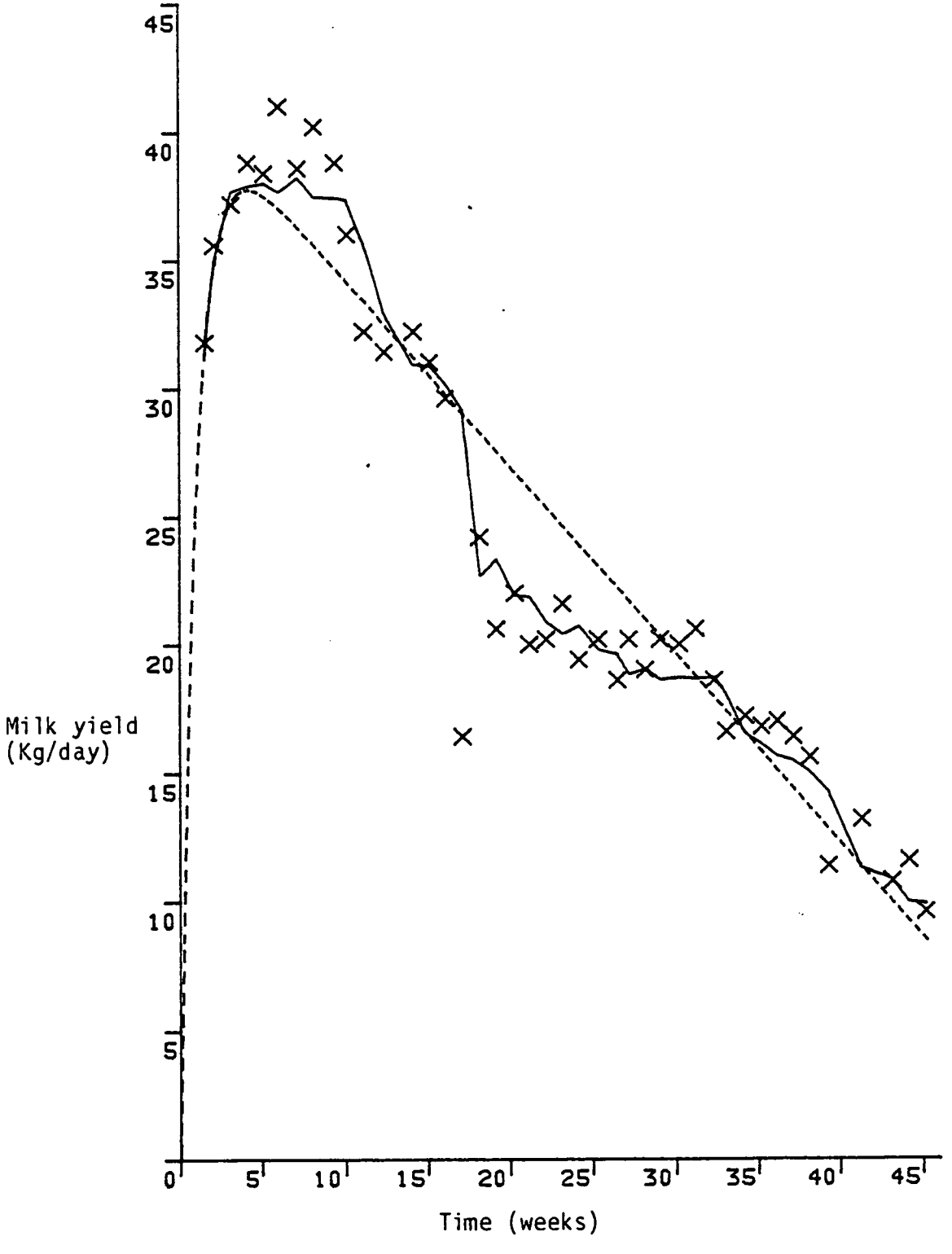


Figure 6.4.5

Neilson's data, cow number 8, model (4) fitted by maximum likelihood estimation; observed milk yields (X), the fitted curve (----) and the predicted values (——) conditional upon earlier observations plotted against time.



6.5 Thiessen's data: weights of cattle

6.5.1 Introduction

The weights of five non-pregnant female cattle fed ad libitum were recorded every second week from weaning at 12 weeks of age for periods of up to 5 years, resulting in between 82 and 133 observations per animal. These animals were part of a multibreed experiment, set up to study genetic variation between breeds and genetic inter-breed relationships for a wide spectrum of traits, in order to examine the problems of between-breed testing and selection (Thiessen, Hnizdo, Maxwell, Gibson and Taylor, 1984). The animals were of the following breed types: one Hereford, one Dairy Shorthorn, two British Whites and one British Friesian.

6.5.2 Models

A vector of observations $y_1, \ldots, y_n$, denoted $\underset{\sim}{y}$, of weights of a single animal at times $t_1, \ldots, t_n$, denoted $\underset{\sim}{t}$, measured in weeks are assumed to be multivariate normally distributed with mean $\underset{\sim}{f}$ where

$$f_i = f(\infty) + \sum_{k=1}^{2} \{f(0)\, K_{1k} + f^{<1>}(0)\, K_{2k} - f(\infty)\, \xi_2 \frac{K_{2k}}{\lambda_k}\}\, e^{t_i \lambda_k},$$

for $i = 1, \ldots, n$.

This is simply a restatement of equation (6.3.7) except for $Y(0)$ being replaced by $f(0)$, $E(Y^{<1>}(0))$ by $f^{<1>}(0)$ and some algebraic manipulation resulting from $(\beta\psi_1)$ having been replaced by $(\xi_2 f(\infty))$. This reparameterization reduces the number of parameters from six to five which is convenient because only five are estimable for a regression model described by a constant plus

the weighted sum of two exponential functions. The terminology $f(\infty)$ has been used because, provided $|e^{\underset{\sim}{\lambda}}|<1$, as t_i increases f_i tends to the limit $f(\infty)$.

Several different models for the variance matrix $(\underset{\sim}{V})$ will be considered:

(0) Independent observations with equal variances τ^2:

$$V_{ij} = \tau^2 I_{ij} ,$$

where $\underset{\sim}{I}$ is the identity matrix of size n.

(1) The standard variance structure appropriate to a second-order stochastic differential equation if the right side of equation (6.3.5) were simply $Z(t)$, plus an observation error, for which

$$V_{ij} = \tau^2 \left[\left(\frac{1}{2} + \frac{1}{2}\sin\alpha^*\right) \sum_{k=1}^{2}\sum_{\ell=1}^{2} K_{2k} K_{2\ell} \left(\frac{e^{t_i\lambda_k+t_j\lambda_\ell} - e^{(t_i-t_j)\lambda_k}}{(\lambda_k+\lambda_\ell)}\right) + \left(\frac{1}{2} - \frac{1}{2}\sin\alpha^*\right) I_{ij}\right] \quad \text{for } i > j .$$

This equation may be derived from equation (6.3.8) by noting that as ψ_1 increases, κ_1/ψ_1 converges to $-K_{21}$ and κ_2/ψ_1 converges to $-K_{22}$, and this is all that is required for equation (6.3.5) to approximate to the above specification in which the right side is dominated by the term in $Z(t)$. The parameterization of the observation error component using α^* is the same as in [6.4.2] to constrain terms to lie between 0 and 1.

(2) The standard variance structure as given by equation (6.3.8) plus an observation error, for which

$$V_{ij} = \tau^2 \left[\left(\frac{1}{2} + \frac{1}{2}\sin\alpha^*\right) \sum_{k=1}^{2}\sum_{\ell=1}^{2} \kappa_k \kappa_\ell \left(\frac{e^{t_i\lambda_k+t_j\lambda_\ell} - e^{(t_i-t_j)\lambda_k}}{(\lambda_k+\lambda_\ell)}\right) + \left(\frac{1}{2} - \frac{1}{2}\sin\alpha^*\right) I_{ij}\right] \quad \text{for } i > j .$$

(3) The variance structure but with separate parameter $\underset{\sim}{\xi}^*$, plus an observation error, for which

$$V_{ij} = \tau^2 \left[\left(\frac{1}{2} + \frac{1}{2}\sin\alpha^*\right) \sum_{k=1}^{2} \sum_{\ell=1}^{2} \kappa_k \kappa_\ell \frac{\left(e^{t_i\lambda_k^* + t_j\lambda_\ell^*} - e^{(t_i - t_j)\lambda_k^*}\right)}{\left(\lambda_k^* + \lambda_\ell^*\right)} \right.$$

$$\left. + \left(\frac{1}{2} - \frac{1}{2}\sin\alpha^*\right) I_{ij} \right] \qquad \text{for } i > j \ ,$$

where λ_1^* , λ_2^* are the roots of

$$x^2 - \xi_1^* x - \xi_2^* = 0 \ .$$

An observation error has been included in all the above models because there will inevitably be some errors of this sort when large animals are being weighed. As in section 6.4, the choice of parameterization of the observation errors has been so as to ensure that the parameters can fit the full range of possible models, whilst remaining bounded, without the use of more than one scaling parameter τ^2 . Model (1) enables model (2) to be tested against a simpler alternative and model (3) enables the assumption in model (2) that parameters are shared in common between the regression function $\underset{\sim}{f}$ and the variance matrix $\underset{\sim}{V}$ to be tested. Model (0) corresponds to a GARMA(0,0) process whilst models (1), (2) and (3) are all GARMA(2,2) processes.

6.5.3 Method of estimation

All the models were fitted separately to each data set by maximum likelihood estimation, that is by minimizing L_M as defined in [2.3.2]. For these models it would have been difficult to use residual maximum likelihood estimation because the first derivative of the regression function with respect to the model parameters would have been required.

Initial estimates of the parameter values for each model were obtained from the final estimates from the next "smaller" model and initial guesses were used for model (0). Convergence difficulties were encountered because of high correlations among the parameter estimators. This could possibly have been circumvented by reparameterizing, but for simplicity of programming, different initial estimates were tried instead and the number of iterations in the optimization algorithm increased. A consequence of this parameterization was that estimates of standard errors could not be obtained from a quadratic approximation to L_M at the minimum.

6.5.4 Results - likelihood comparisons

Table 6.5.1 shows the minimized values of L_M (the negative log-likelihood) for each model fitted to the data from each animal. For each animal the model with the lowest Akaike Information Criterion, that is L_M plus the number of model parameters, is underlined. On the basis of this criterion the best model is (3), except for animal 3 where model (2) is adequate.

Likelihood ratio tests are not necessary to show that the improvement of model (3) over models (0), (1) and (2) is sufficient to reject the simpler models for all except animal 3. The conclusion from this evidence is that the variance matrix needs to be parameterized separately from the regression function.

6.5.5 Results - parameter estimates

Table 6.5.2 shows the means and standard deviations of the parameter estimates over all animals for each model. The summary statistics were defined and discussed in [6.4.5]. As already

Table 6.5.2

Theissen's data; minimum values of the negative log-likelihood (L_M) for different models.

Animal	Model / Number of Parameters	independent errors (0) 6	model with ψ omitted (1) 7	standard model (2) 8	sep. cor. param. $\underset{\sim}{\xi}^*$ (3) 10
1		411.1	346.5	341.4	<u>330.7</u>
2		312.8	287.7	279.1	<u>274.7</u>
3		402.3	323.0	<u>306.0</u>	305.0
4		333.2	294.0	285.8	<u>279.0</u>
5		209.2	209.2	209.0	<u>194.5</u>
Average		333.7	292.1	284.3	<u>276.9</u>

For each animal the model with the lowest value of Akaike's Information Criterion has been underlined.

Table 6.5.2

Theissen's data; means and standard deviations of parameter estimates over all 5 animals for each model.

Model	$f(0)$	$f^{<1>}(0)$	$10^2\xi_1$	$10^4\xi_2$	$f(\infty)$	$10^2\psi$	$10^2\xi_1^*$	$10^4\xi_2^*$	α^*
				Means of estimates					
(0) indep. errors	96.	4.6	-6.8	-10.9	690.				
(1) model with ψ omitted	95.	4.6	-6.0	-9.2	710.				0.2
(2) standard model	95.	5.1	-5.8	-8.9	690.	2.5			1.4
(3) sep. cor. param. $\underset{\sim}{\xi}^*$	94.	4.8	-7.0	-10.8	690.	0.9	-16.	-200.	1.5
				standard deviations of estimates					
(0) indep. errors	7.	2.1	6.5	11.2	30.				
(1) model with ψ omitted	6.	1.5	4.8	7.9	30.				1.0
(2) standard model	5.	1.5	4.5	7.7	30.	2.3			0.1
(3) sep. cor. param. $\underset{\sim}{\xi}^*$	5.	1.2	6.1	9.8	30.	2.3	27.	330.	0.1

explained, standard errors could not be obtained for these models. It is recognised that some of the differences between animals are being ignored in forming these statistics, but this simple approach is adequate for making certain important comparisons between the different models.

The parameter estimates are in good agreement except that $\hat{\xi}_1$ and $\hat{\xi}_2$ in models (1) and (2) are greater than in models (0) and (3). This arises because $\underset{\sim}{\xi}$ parameterizes both the regression and error structure in models (1) and (2), whilst in models (0) and (3) $\underset{\sim}{\xi}$ only parameterizes the regression function, and in model (3) a separate parameter $\underset{\sim}{\xi}^*$ is used to model the errors and is estimated as being quite different from $\underset{\sim}{\xi}$.

The standard deviations are greatest for model (0) as is to be expected because the assumption of independent errors leads to inefficient estimators. The standard deviations for $\hat{\underset{\sim}{\xi}}$ are less in models (1) and (2) than in model (3) because the information in error structure is also being used in the estimation procedure. However, this appearance of increased precision is an illusion because models (1) and (2) have been rejected by the data so the estimators are meaningless anyway.

Figures 6.5.1, 6.4.2, 6.5.3 and 6.5.4 show the data for animal 4 together with the fitted regression function and predicted values using earlier observations (described in [2.3.6]) for models (0), (1), (2) and (3) respectively. The systematic departures of the fit in model (1) illustrates the effect of assuming a high correlation among observations, and was the reason for the choice of animal 4 for plotting.

Figure 6.5.1

Thiessen's data, animal number 4, model (0) fitted by maximum likelihood estimation; observed weights (X) and the fitted curve (——) plotted against time.

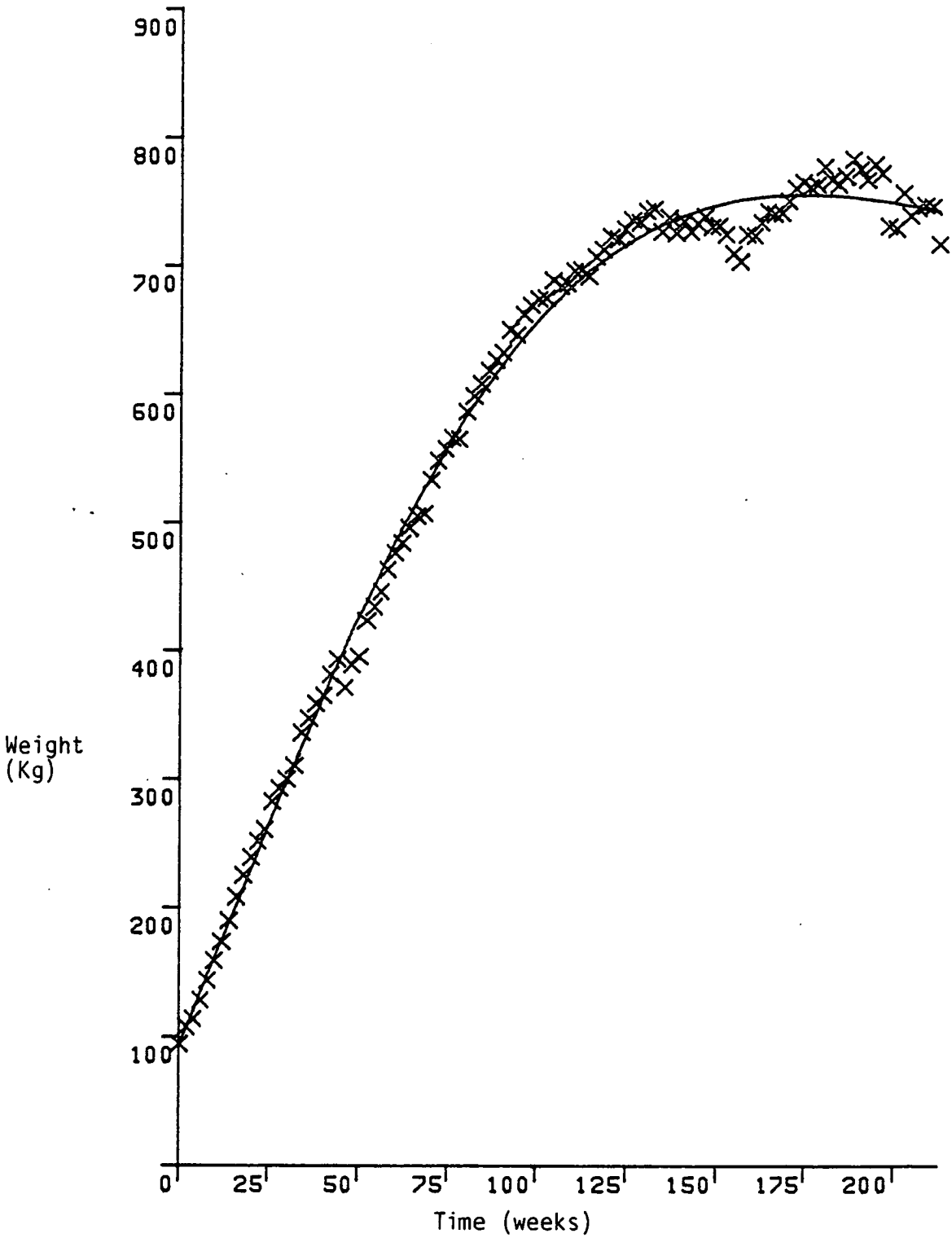


Figure 6.5.2

Thiessen's data, animal number 4, model (1) fitted by maximum likelihood estimation; observed weights (X), the fitted curve (----) and the predicted values (——) conditional upon earlier observations, plotted against time.

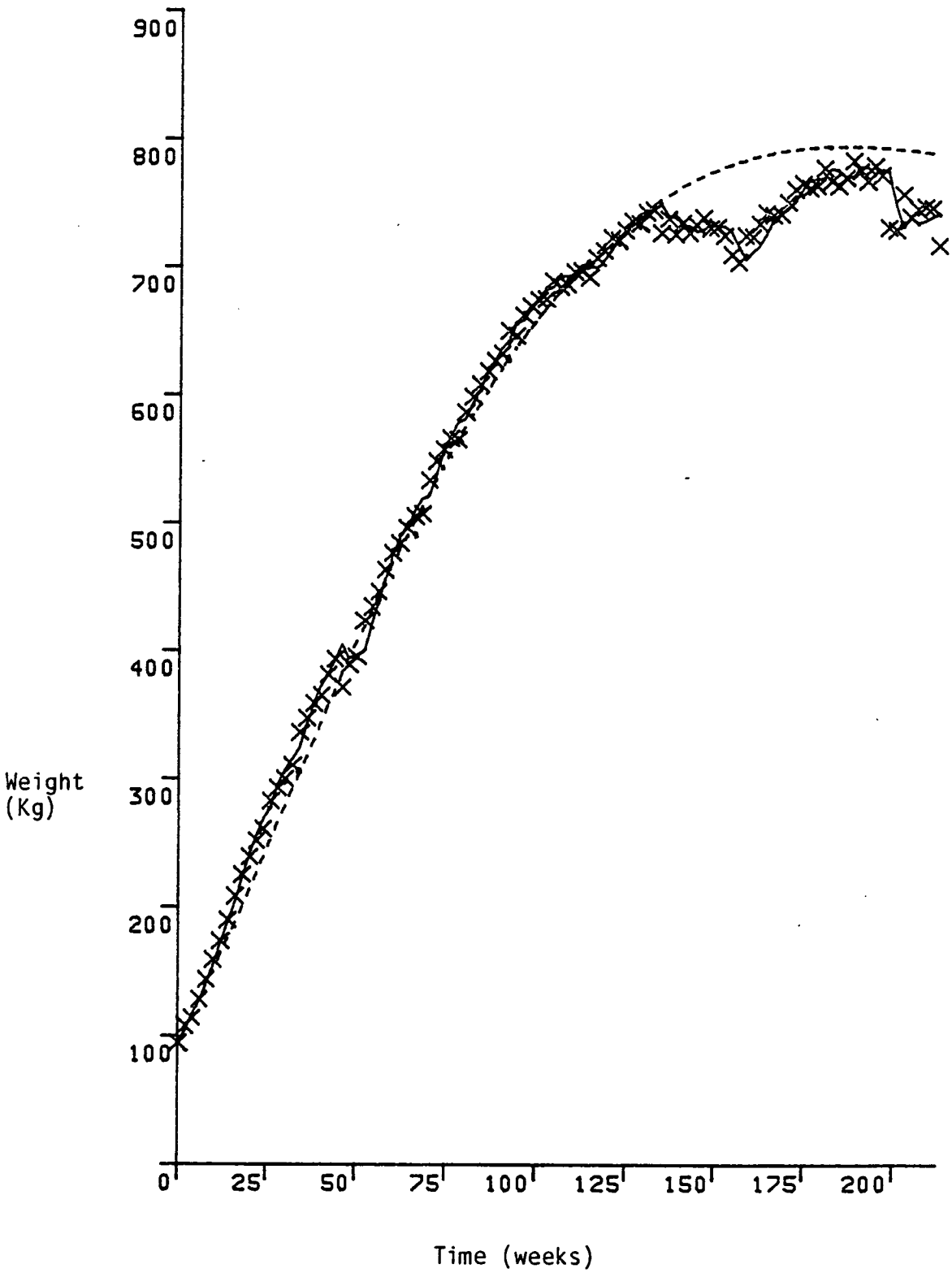


Figure 6.5.3

Thiessen's data, animal number 4, model (2) fitted by maximum likelihood estimation; observed weights (X), the fitted curve (----) and the predicted values (——) conditional upon earlier observations, plotted against time.

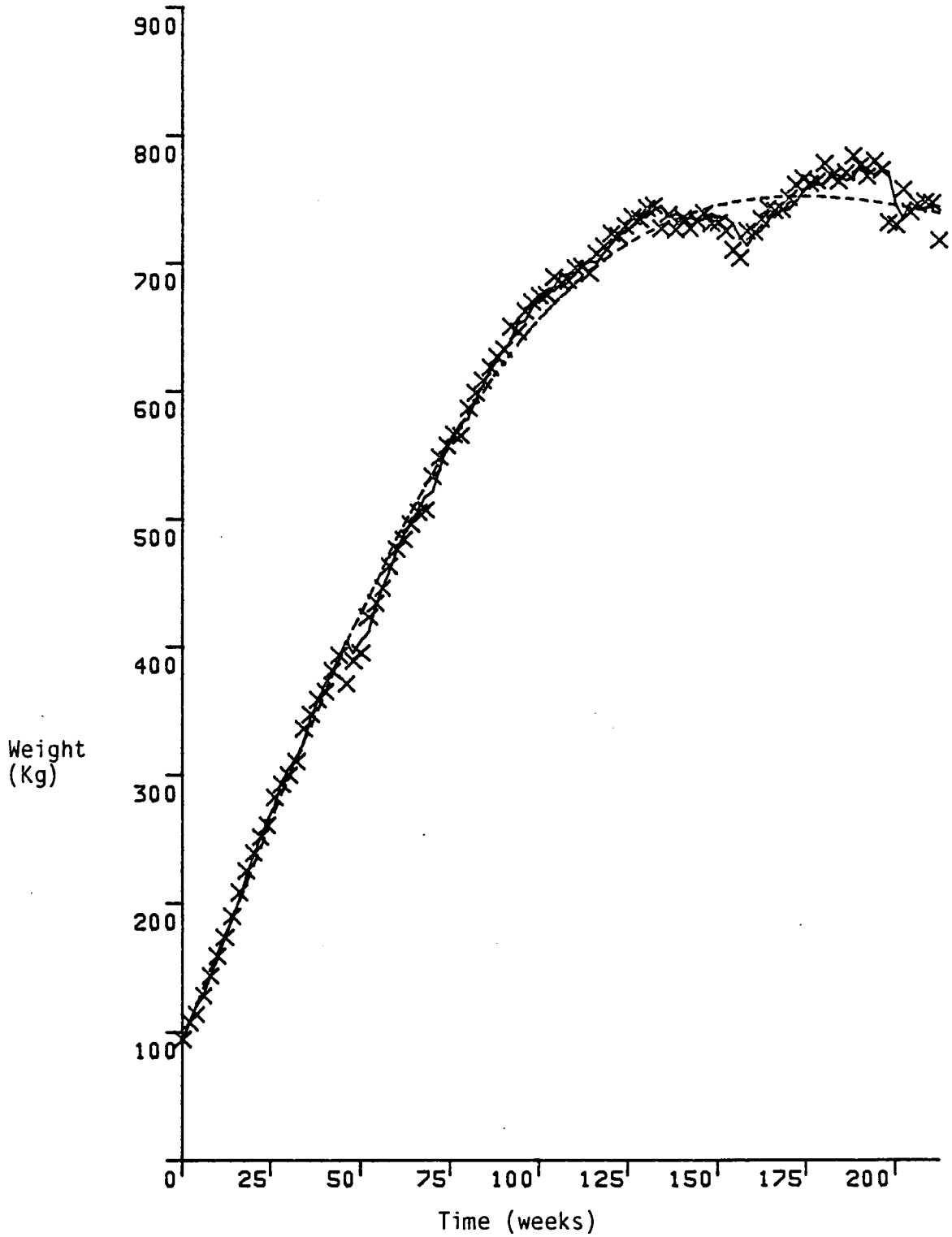
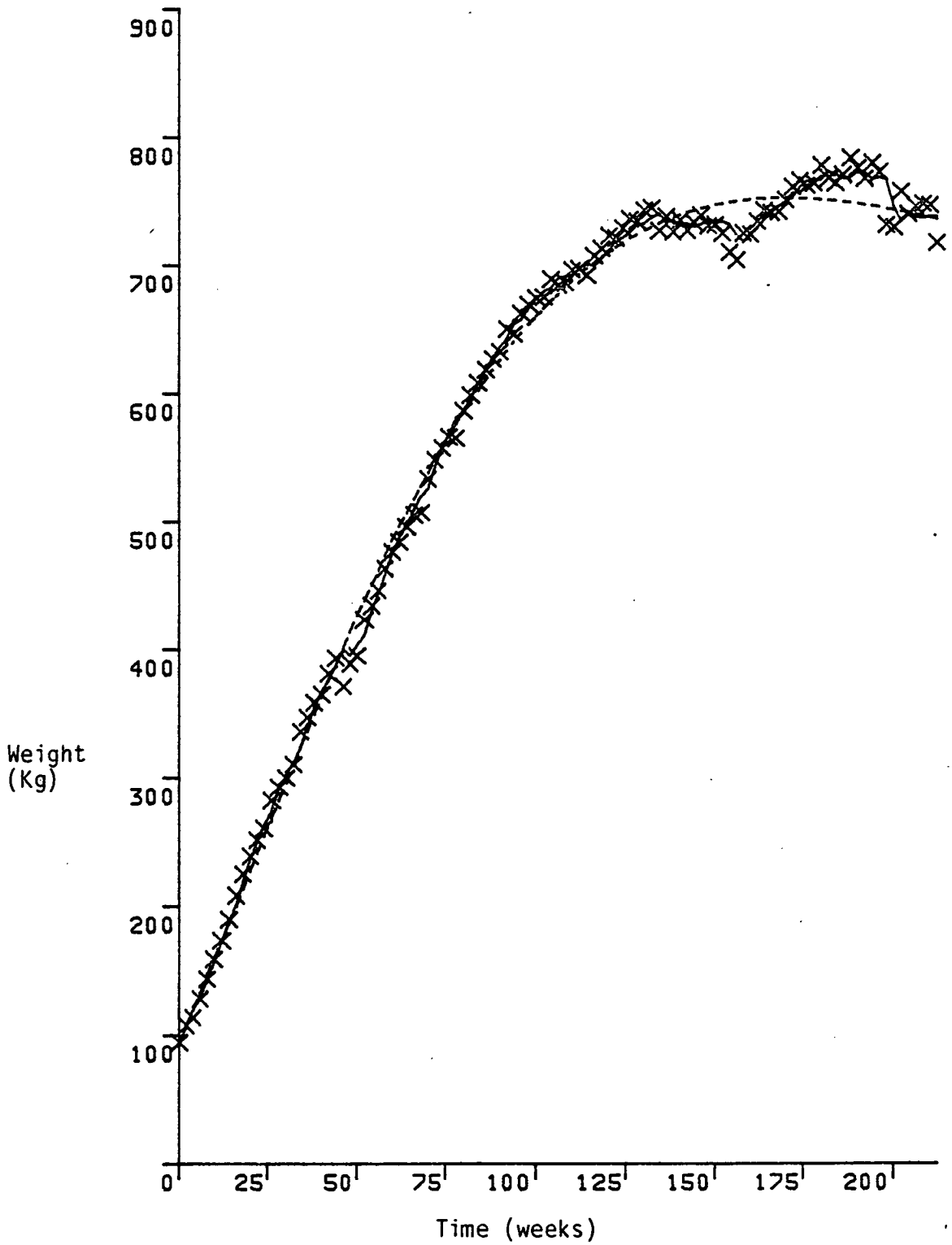


Figure 6.5.4

Thiessen's data, animal number 4, model (3) fitted by maximum likelihood estimation; observed weights (X), the fitted curve (----) and the predicted values (——) conditional upon earlier observations, plotted against time.



6.5.6 Resumé

The notion of using the same parameters to model both the regression function and the error structure in animal growth is an attractive one. However, in examples such as the above where the data do not support the model the approach cannot be recommended. It may give the appearance of improved precision but in fact the estimators are biased.

6.6 Discussion

The technique considered in this chapter is to model data by linear stochastic difference and differential equations with constant coefficients. The parameters in the model are estimated utilizing the information both in the main trend of the data and in the local stochastic fluctuations. The techniques were applied to two data sets, one of milk yields and the other of weights, for neither of which were there any theoretical justification for using such models. Therefore it was perhaps not surprising to find that in neither case did the models fit well, there being no evidence of a link between the parameters required to describe the regression function and those required to describe the correlation structure. If the parameters are linked together then the resulting estimators are biased, so the methods of chapters 4 and 5, where the error structure is modelled empirically on the basis of the observed data, would seem more appropriate to these types of application. In other words, if a differential equation is to be used then the deterministic form should be used with the error process "tacked on" at the end. This result is in direct contradiction to the arguments of Sandland and McGilchrist (1979), quoted in [6.3.1], and to the recommendations made implicitly by Dhanoa and Le Du.

7. Stochastic compartment models with constant coefficients

7.1 Introduction

This is the second of three chapters which consider the use of mechanistic models in order to estimate regression parameters when errors are serially correlated. Sometimes the theoretical justification for fitting a weighted sum of exponentials as a regression function is that a compartment system is a realistic model for the generation of the data. The fundamental assumption is that the system to be modelled may be divided into homogenous components called compartments, and that material flows between these compartments according to some specified kinetics. Such reasoning lay behind the choice of weighted sums of exponential functions to fit Colquhoun's and Dale's data in chapter 4, discussed for the respective cases by Colquhoun and Hawkes (1977) and Bauermeister, Dale, Williams and Scobie (1980). The opportunity arises, as it did for difference and differential equations in chapter 6, to incorporate stochastic variability into the basic model rather than to add error onto the deterministic solution in order to model the observations. Provided a stochastic compartment model is appropriate, this approach will improve the efficiency of parameter estimation and help identify the correct order of compartment system.

The stochastic variability considered in this chapter is that arising from having only a finite number of particles in the compartmental system, each behaving independently. This has been denoted type P1 by Matis and Wehrly (1979) who also considered other sources of variability. In section 7.2, the first and

second moments are derived for a type of variate specified by the model, and this variate is shown to be a cGARMA process. In section 7.3, Colquhoun's data is modelled using stochastic compartment models and the results compared with the empirical approach to modelling used in section 4.3. In section 7.4, Dale's data is modelled in a similar way. Finally, in section 7.5 the usefulness of stochastic compartment models in estimating regression parameters in the presence of serially correlated errors is critically assessed.

The mathematical development of stochastic compartment models given in section 7.2 is not new, see the literature review in [1.3.6]. At the theoretical level, what is original in this chapter is the link made with cGARMA processes which allows the Gaussian likelihood, or variants of it, to be evaluated simply. Therefore, parameters in a stochastic compartment model can be estimated using the computer program REGAME. The applications of the stochastic model to the two situations represented by the two data sets are also original, previous analyses not having utilized the link between the regression and error structures.

7.2. Theory

7.2.1 Model

The situation envisaged is where a sequence of observations $y_1,\ldots,y_n$ (denoted $\underset{\sim}{y}$) have been made at times $t_1,\ldots,t_n$ (denoted $\underset{\sim}{t}$) on a single experimental subject. It is assumed that the vector $\underset{\sim}{y}$ is a realisation at times $\underset{\sim}{t}$ of a random

variate $Y(t)$, indexed by t over a continuous finite or infinite interval, which is a weighted sum of the number of particles in each of p compartments of a stochastic compartment system. Therefore

$$(7.2.1) \qquad Y(t) = \sum_{i=1}^{p} \mu_i N_i(t)$$

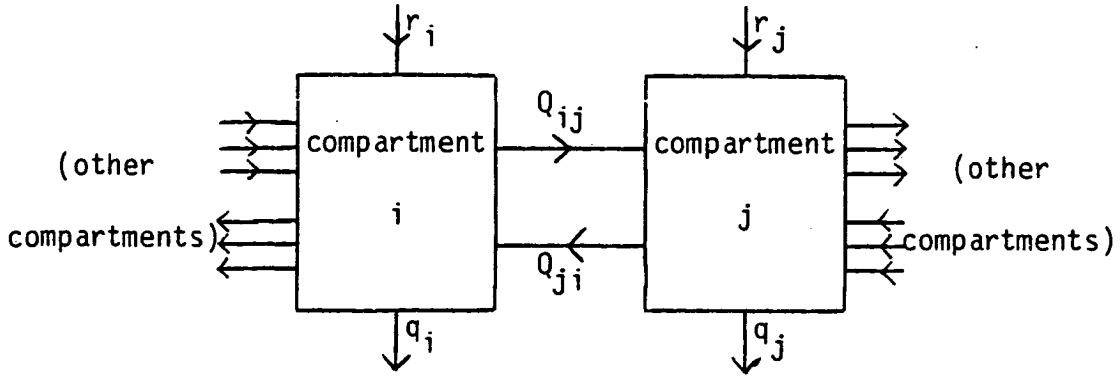
where $N_i(t)$ is the random number of particles in compartment i at time t and μ_i is the constant weight given to particles in compartment i.

The stochastic compartment system is defined as follows:

a) a particle in compartment i at time t is in compartment $j(\neq i)$ at time $t+\delta t$ with probability $Q_{ij}\delta t + o(\delta t)$, has left the system with probability $q_i\delta t + o(\delta t)$ and is otherwise still in compartment i , independently of what happens to all other particles and what happens at all earlier times, where $o(\delta t)$ denotes terms of order less than δt and $\delta t>0$;

b) the number of particles entering compartment i from outside the system between times t and $t+\delta t$ is Poisson distributed with expectation $r_i\delta t + o(\delta t)$.

Diagramatically this can be represented as:



The rate constant Q_{ij} denotes a coefficient in a square matrix

$\underset{\sim}{Q}$ of size p, all of whose off-diagonal elements must be non-negative for the definition above to make sense as δt approaches zero. Constants q_i and r_i, which denote coefficients in vectors $\underset{\sim}{q}$ and $\underset{\sim}{r}$, both of length p, must also be non-negative. For obvious reasons, $\underset{\sim}{r}$ is called the immigration rate and $\underset{\sim}{q}$ the emigration rate. If all coefficients in both $\underset{\sim}{q}$ and $\underset{\sim}{r}$ are zero the system is said to be closed, otherwise it is said to be open.

7.2.2 Solution

The solution that follows is an amalgam of the analyses given by Cox and Miller (1965, pp.178-186), Chiang (1980, pp.451-498) and Colquhoun and Hawkes (1977). The diagonal elements of $\underset{\sim}{Q}$, which were unspecified in [7.2.1], are defined by

$$Q_{ii} = -\sum_{j \neq i} Q_{ij} - q_i \qquad \text{for } i = 1,\ldots,n.$$

Therefore $(1 + Q_{ii}\delta t + o(\delta t))$ is the probability of a particle in compartment i at time t being in compartment i at time $t+\delta t$.

The probability that a particle in compartment i at time t will be in compartment j at time $s(\geq t)$ is dependent only upon the time separation $(s-t)$ and is denoted:

$$(7.2.2) \qquad P_{ij}(s-t) = \sum_{\ell=1}^{p} K_{ij}^{\{\ell\}} e^{(s-t)\lambda_\ell},$$

where the coefficients $\lambda_1,\ldots,\lambda_p$ in the p-vector $\underset{\sim}{\lambda}$ are the eigenvalues of $\underset{\sim}{Q}$, which are assumed to be all distinct, and $\underset{\sim}{K}^{\{1\}},\ldots,\underset{\sim}{K}^{\{p\}}$ are square matrices of size p, whose coefficients are specified by

$$K_{ij}^{\{\ell\}} = G_{i\ell}\,(\underset{\sim}{G}^{-1})_{\ell j} \qquad \text{for } i,j,\ell = 1,\ldots,p,$$

where $\underset{\sim}{G}$ is the matrix of right eigenvectors of $\underset{\sim}{Q}$.

The number of particles in compartment j at time s which were not in the system at an earlier time t is also dependent only upon the time separation $(s-t)$ and is Poisson distributed with expectation denoted by

$$(7.2.3) \qquad b_j(s-t) = \sum_{\ell=1}^{p} \kappa_j^{\{\ell\}}(e^{(s-t)\lambda_\ell} - 1) ,$$

where $\underset{\sim}{\kappa}^{\{1\}},\ldots,\underset{\sim}{\kappa}^{\{p\}}$ are vectors of length p whose coefficients are specified by

$$\kappa_j^{\{\ell\}} = \frac{1}{\lambda_\ell} \sum_{i=1}^{p} r_i \; K_{ij}^{\{\ell\}} \qquad \text{if } \lambda_\ell > 0,$$

and $$= 0 \qquad \text{if } \lambda_\ell = 0.$$

It is convenient to distinguish between the number of particles in the various compartments at time t that were in the system at time zero, denoted $\underset{\sim}{N}^{\{1\}}(t)$, and the number of particles that entered the system after time zero, denoted $\underset{\sim}{N}^{\{2\}}(t)$. From equation (7.2.1)

$$E(Y(t)) = \sum_{i=1}^{p} \mu_i \, E(N_i^{\{1\}}(t) + N_i^{\{2\}}(t)) ,$$

$$= \sum_{i=1}^{p} \mu_i [E\{E(N_i^{\{1\}}(t) \,|\, \underset{\sim}{N}(0))\} + E(N_i^{\{2\}}(t))]$$

by conditioning on $\underset{\sim}{N}(0)$, the number of particles in compartments at time zero,

$$(7.2.4) \qquad = \sum_{i=1}^{p} \mu_i [\sum_{j=1}^{p} n_j \, P_{ji}(t) + b_i(t)] ,$$

where $E(\underset{\sim}{N}(0))$ is denoted by $\underset{\sim}{n}$,

$$= \sum_{i=1}^{p} \mu_i [\sum_{j=1}^{p} n_j \sum_{k=1}^{p} K_{ji}^{\{k\}} e^{t\lambda_k} + \sum_{\ell=1}^{p} \kappa_i^{\{\ell\}}(e^{t\lambda_\ell} - 1)],$$

by substituting for $\underset{\sim}{P}(t)$ and $\underset{\sim}{b}(t)$ from equations (7.2.2) and (7.2.3),

$$(7.2.5) \qquad = -\sum_{i=1}^{p}\sum_{\ell=1}^{p} \mu_i \kappa_i^{\{\ell\}}$$

$$+\sum_{k=1}^{p}\{\sum_{i=1}^{p}\sum_{j=1}^{p} \mu_i n_j K_{ji}^{\{k\}} + \sum_{i=1}^{p} \mu_i \kappa_i^{\{k\}}\} e^{t\lambda_k} ,$$

by rearranging terms. From this it can be seen that the functional form of $E(Y)$ is a constant plus a weighted sum of p exponentials.

Also from equation (7.2.1), because $\underset{\sim}{N}^{\{1\}}(t)$ and $\underset{\sim}{N}^{\{2\}}(t)$ are independent

$$(7.2.6) \quad \mathrm{cov}(Y(s),Y(t)) = \sum_{i=1}^{p}\sum_{j=1}^{p} \mu_i\mu_j[\mathrm{cov}(N_i^{\{1\}}(s),N_j^{\{1\}}(t)) + \mathrm{cov}(N_i^{\{2\}}(s),N_j^{\{2\}}(t))] .$$

I shall consider $\underset{\sim}{N}^{\{1\}}$ and $\underset{\sim}{N}^{\{2\}}$ separately, and restrict to the case when $s>t$. By a standard probability formula, conditioning on $\underset{\sim}{N}(0)$,

$$\mathrm{cov}(N_i^{\{1\}}(s),N_j^{\{1\}}(t)) = E[\mathrm{cov}(N_i^{\{1\}}(s),N_j^{\{1\}}(t)|\underset{\sim}{N}(0))] + \mathrm{cov}[E(N_i^{\{1\}}(s)|\underset{\sim}{N}(0)), E(N_j^{\{1\}}(t)|\underset{\sim}{N}(0))] ,$$

$$= E[\sum_{k=1}^{p} N_k(0)(P_{kj}(t)P_{ji}(s-t)-P_{kj}(t)P_{ki}(s))]$$

$$+ \mathrm{cov}[\sum_{\ell=1}^{p} N_\ell(0)P_{\ell i}(s) , \sum_{k=1}^{p} N_k(0)P_{kj}(t)] ,$$

$$= \sum_{k=1}^{p} n_k(P_{kj}(t)P_{ji}(s-t)-P_{kj}(t)P_{ki}(s))$$

$$+ \sum_{k=1}^{p}\sum_{\ell=1}^{p} \mathrm{cov}(N_k(0),N_\ell(0))P_{kj}(t)P_{\ell i}(s) ,$$

where the derivation of the first term can be understood by considering a single particle: if it is in compartment k at time zero then it contributes 1 to $(N_i^{\{1\}}(s)\ N_j^{\{1\}}(t))$ only if it moves to compartment j by time t and then to compartment i by time s and otherwise contributes 0, contributes 1 to $N_i^{\{1\}}(s)$ by moving to compartment i by time s, and contributes 1 to $N_j^{\{1\}}(t)$ by moving to compartment j by time t. By applying the same standard formula to $\underset{\sim}{N}^{\{2\}}$, but this time conditioning on $\underset{\sim}{N}^{\{2\}}(t)$,

$$\operatorname{cov}(N_i^{\{2\}}(s),N_j^{\{2\}}(t)) = E[\operatorname{cov}(N_i^{\{2\}}(s),N_j^{\{2\}}(t)\,|\,\underset{\sim}{N}^{\{2\}}(t))] + \operatorname{cov}[E(N_i^{\{2\}}(s)\,|\,\underset{\sim}{N}^{\{2\}}(t)),\ E(N_j^{\{2\}}(t)\,|\,\underset{\sim}{N}^{\{2\}}(t))] .$$

Because

$$\operatorname{var}(N_j^{\{2\}}(t)\,|\,\underset{\sim}{N}^{\{2\}}(t)) = 0 ,$$

$$\operatorname{cov}(N_i^{\{2\}}(s),\ N_j^{\{2\}}(t)\,|\,\underset{\sim}{N}^{\{2\}}(t)) = 0 ,$$

and the first term above is zero. Therefore

$$\operatorname{cov}(N_i^{\{2\}}(s),N_j^{\{2\}}(t)) = \operatorname{cov}\Big[\sum_{k=1}^{p} N_k^{\{2\}}(t)P_{ki}(s-t)+b_i(s-t),\ N_j^{\{2\}}(t)\Big] ,$$

because particles that are in compartment i at time s, but were not in the system at time zero, either were in some compartment k at time t and moved to compartment i before time s, or entered the system between times t and s,

$$= \sum_{k=1}^{p} \operatorname{cov}(N_k^{\{2\}}(t),N_j^{\{2\}}(t))P_{ki}(s-t) ,$$

$$= b_j(t)P_{ji}(s-t) ,$$

because $N_k^{\{2\}}(t)$ and $N_j^{\{2\}}(t)$ are uncorrelated unless k equals j, in which case $\operatorname{var}(N_j^{\{2\}}(t))$ equals $E(N_j^{\{2\}}(t))$ because the variate is Poisson distributed.

When these results are incorporated into equation (7.2.6) they give

$$(7.2.7)\quad \operatorname{cov}(Y(s),Y(t)) =$$

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\mu_i\mu_j\Big[\sum_{k=1}^{p} n_k P_{kj}(t)\left(P_{ji}(s-t) - P_{ki}(s)\right)$$

$$+ \sum_{k=1}^{p}\sum_{\ell=1}^{p}\operatorname{cov}(N_k(0),N_\ell(0))P_{kj}(t)P_{\ell i}(s) + b_j(t)P_{ji}(s-t)\Big] .$$

7.2.3 Distribution of $\underset{\sim}{N}(0)$

Up to this point no mention has been made of the distribution of $\underset{\sim}{N}(0)$, that is the number of particles in the system at time zero, beyond denoting its expectation by $\underset{\sim}{n}$. Three particular distributions which have appeared in the literature and tie in well with the other components of a stochastic compartmental model, are:

a) The null distribution, that is $\underset{\sim}{N}(0)$ is a vector of constants and in particular the covariance coefficients required in equation (7.2.7) are specified by
$\operatorname{cov}(N_k(0),N_\ell(0)) = 0$.

b) The multinomial distribution, so that in particular
$\operatorname{cov}(N_k(0),N_\ell(0)) = n_k \{I_{k\ell} - n_\ell / \sum_{i=1}^{p} n_i \}$,
where $I_{k\ell}$ is the $(k\ell)$th coefficient in the identity matrix $\underset{\sim}{I}$ of size p .

c) Each coefficient of $\underset{\sim}{N}(0)$ is an independent Poisson random variate, so that in particular
$\operatorname{cov}(N_k(0),N_\ell(0)) = n_k I_{k\ell}$ for $k,\ell = 1, \ldots, p$.

The uses of different starting distributions are discussed as they arise practically in sections 7.3 and 7.4.

7.2.4 Relation to cGARMA property

By substitution of $\underset{\sim}{P}(t)$ and $\underset{\sim}{b}(t)$ from equations (7.2.2) and (7.2.3), equation (7.2.7) becomes

$$\begin{aligned}\operatorname{cov}(Y(s),Y(t)) = &\sum_{i=1}^{p}\sum_{j=1}^{p}\mu_i\mu_j\Big[\sum_{k=1}^{p} n_k \sum_{g=1}^{p}\sum_{h=1}^{p} K_{kj}^{\{g\}} e^{t\lambda_g} \\ &\qquad (K_{ji}^{\{h\}} e^{(s-t)\lambda_h} - K_{ki}^{\{h\}} e^{s\lambda_h}) \\ &+ \sum_{k=1}^{p}\sum_{\ell=1}^{p}\operatorname{cov}(N_k(0),N_\ell(0)) \sum_{g=1}^{p}\sum_{h=1}^{p} K_{kj}^{\{g\}} e^{t\lambda_g} K_{\ell i}^{\{h\}} e^{s\lambda_h} \\ &+ \sum_{g=1}^{p}\sum_{h=1}^{p}\kappa_j^{\{g\}}(e^{t\lambda_g}-1) K_{ji}^{\{h\}} e^{(s-t)\lambda_h}\Big],\end{aligned}$$

$$\begin{aligned}= \sum_{h=1}^{p} e^{s\lambda_h}\Big\{&\sum_{i=1}^{p}\sum_{j=1}^{p}\mu_i\mu_j\Big[\sum_{k=1}^{p}\sum_{g=1}^{p} n_k K_{kj}^{\{g\}} e^{t\lambda_g} (K_{ji}^{\{h\}} e^{-t\lambda_h} - K_{ki}^{\{h\}}) \\ &+ \sum_{k=1}^{p}\sum_{\ell=1}^{p}\sum_{g=1}^{p}\operatorname{cov}(N_k(0),N_\ell(0)) K_{kj}^{\{g\}} e^{t\lambda_g} K_{\ell i}^{\{h\}} \\ &+ \sum_{g=1}^{p}\kappa_j^{\{g\}}(e^{t\lambda_g}-1) K_{ji}^{\{h\}} e^{-t\lambda_h}\Big]\Big\} \qquad \text{when } s \geq t\end{aligned}$$

by rearranging terms. This is the sum of p products of a term dependent only upon s and a term dependent only upon t. Therefore $\underset{\sim}{Y}$ is a cGARMA(p,0) process by the definition in [2.2.7] and, of more importance, the discrete realisation $\underset{\sim}{y}$ is a GARMA(p,p-1) process. Therefore, the computer program REGAME can be used to fit a stochastic compartment model to a vector $\underset{\sim}{y}$ of observations.

7.2.5 Parameter estimators

Unlike the models considered in previous chapters, the realisation $\underset{\sim}{y}$ of a stochastic compartment model is not normally

distributed, in fact a single element is distributed as a weighted sum of Poisson variates and terms in a multinomial random variate, and joint distributions between elements are even more complicated. Because the likelihood function is so complicated, Matis and Wehrly (1979) fitted the models by generalized least-squares estimation. In the examples that follow, other sources of error have been included and Gaussian likelihood estimation (that is using L_M) has been used instead of generalized least-squares estimation (that is using L_S), because, as discussed in [2.3.2], L_S cannot be used when the variance matrix has parameters in it which are not in the regression function. The third optimization function considered in the thesis, the residual Gaussian likelihood, has not been employed in this chapter because of the complications involved in deriving the design matrix $\underset{\sim}{X}$, defined in [2.3.2]. As discussed in [1.3.8], asymptotic properties do apply to regression parameter estimators, but this is because of the replication in particles, rather than the replication in observation times. The computer program, REGAME which was described in section 2.4, has been used to fit the models in the following sections, and has derived approximate variances of parameter estimators by inverting the Hessian matrix.

7.3 Colquhoun's data

7.3.1 Introduction

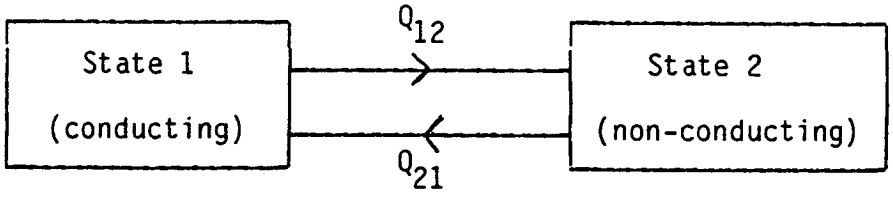
Compartment models have been proposed to explain the relaxation of drug-induced currents, see for example Colquhoun and

Hawkes (1977). Colquhoun's data, first considered in section 4.3, is an example of this type of data. The theory derives from the observed current being the sum of the currents flowing through each of the ion-channels in a muscle end-plate membrane. Each channel is assumed to be in one of several states, some of which are conducting with a particular constant conductance and some of which are non-conducting. It is assumed that each channel changes state independently of all other channels, according to the probabilistic laws of a stochastic compartment model. With these assumptions, the observed current is a linear combination of the number of channels in each state and corresponds with the variate $Y(t)$ considered in section 7.2.

The compartments actually correspond to different possible molecular states of each channel, the transitions between which are known, from chemical physics, to be well approximated by a Markov process. It is therefore slightly misleading to use the term "compartment model", which originated in order to describe much more empirical situations. However, I do not consider the distinction between empirical and mechanistic models to be sufficient to warrant the use of a different name to describe mathematically equivalent stochastic processes.

7.3.2 Closed two-compartment model

The simplest model of drug action is one in which each channel has only a single conducting state and a single non-conducting state. Every channel is in one or other of these states, there being neither immigration nor emigration, so the system is closed. It may be represented diagramatically as:



The current at time t is

$$Y(t) = \mu_1 N_1(t) ,$$

where μ_1 is the current flowing through a single ion-channel in state one and $N_1(t)$ denotes the number of channels in this state at time t. Because of the direction in which the current was measured μ_1 is actually the negative of the current (denoted τ^2 because, for reasons that will become apparent later, it may be viewed as a scaling parameter).

The distribution of channels between states at time zero, $\underset{\sim}{N}(0)$, is assumed to be multinomial (binomial in this case) because this ensures the desirable property that $\underset{\sim}{N}(t)$ is then multinomially distributed for all later times t. The electrical circuits by which $\underset{\sim}{y}$ was recorded introduce measurement error which, for lack of any other information, is assumed to be white noise. This serves to add an extra term $\sigma^2 I_{st}$ into the expression for $\mathrm{cov}(Y(s),Y(t))$. Therefore, from equation (7.2.4)

$$E(Y(t)) = - \sum_{j=1}^{2} \tau^2 n_j P_{j1}(t) ,$$

and from the modification to equation (7.2.7)

$$\begin{aligned} \mathrm{cov}(Y(s),Y(t)) = \tau^2 \{ & \sum_{k=1}^{2} \tau^2 n_k P_{k1}(t) \left(P_{j1}(s-t) - P_{k1}(s)\right) \\ & + \sum_{k=1}^{2} \sum_{\ell=1}^{2} \tau^2 \, \mathrm{cov}(N_k(0),N_\ell(0)) \, P_{k1}(t) \, P_{\ell 1}(s) \\ & + \sigma^2/\tau^2 \, I_{st} \} , \end{aligned}$$

with terms in $\underset{\sim}{b}(t)$ omitted because there is no immigration. The

current, τ^2 , can be treated as a scaling parameter, and therefore estimated separately, provided that $\tau^2 \underset{\sim}{n}$ is estimated in place of $\underset{\sim}{n}$ and σ^2/τ^2 in place of σ^2 . From the above covariance structure $\underset{\sim}{y}$, the realisation of $\underset{\sim}{Y}$ at discrete times $\underset{\sim}{t}$, is a GARMA(2,2) process. However, because the system is closed, one of the eigenvalues of the transition matrix $\underset{\sim}{Q}$ is zero and as a consequence $\underset{\sim}{y}$ is a GARMA(1,1) process.

The model was fitted to Colquhoun's data by minimizing the negative Gaussian log-likelihood (L_M) using transformations of the parameter estimates obtained in chapter 4 as starting values in the iterative process. The parameter estimates, standard errors and correlation coefficients are given in table 7.3.1. The minimum value of L_M was -65.4 .

7.3.3 $\tau^2 n_2$ held constant

The iterative convergence was slow, from which, upon examination, it became apparent that $\tau^2 n_2$ is known with little precision despite its estimated standard error (based on the quadratic approximation to L_M) being small. With $\tau^2 n_2$ held constant at either 200 or 1000, the model was refitted and the parameter estimates that were obtained are given in tables 7.3.2 and 7.3.3 respectively. The increase in L_M to -65.0 is small and, in particular, is not sufficiently large to reject the hypothesis that $\tau^2 n_2$ is equal to 1000, on the basis of the asymptotic properties of the likelihood ratio test. This uses the property that the Gaussian likelihood converges to the true likelihood function as the number of particles in the compartment system increases. Therefore, purely on the basis of the data, it

Table 7.3.1

Colquhoun's data, closed two-compartment model fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 n_1$	$\tau^2 n_2$	σ^2/τ^2	τ^2
estimate	0.0057	0.135	9.4	83.3	2.6	0.030
se	0.0118	0.009	1.0	7.8	1.8	0.021
correlations						
Q_{21}	-0.88					
$\tau^2 n_1$	-0.30	0.07				
$\tau^2 n_2$	0.99	-0.87	-0.41			
σ^2/τ^2	0.78	-0.70	-0.20	0.77		
τ^2	-0.86	0.77	0.24	-0.85	-0.97	

Correlation coefficients exceeding 0.9 underlined.

Table 7.3.2

Colquhoun's data, closed two-compartment model with $\tau^2 n_2$ fixed at 200, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	0.085	0.062	9.0	11.4	0.0067
se	0.002	0.001	0.6	4.8	0.0022
correlations					
Q_{21}	0.92				
$\tau^2 n_1$	-0.50	-0.68			
σ^2/τ^2	-0.06	-0.08	0.13		
τ^2	0.04	0.06	-0.11	-0.92	

Correlation coefficients exceeding 0.9 underlined.

is not possible to distinguish whether $\tau^2 n_2$ is equal to 83 or 1000. This is possibly because direct observations are made on compartment 1 but not on compartment 2.

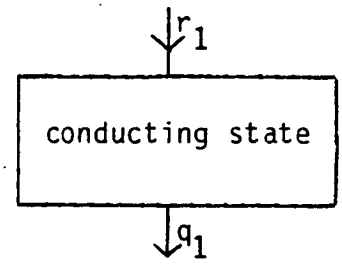
However, consideration of the experimental conditions under which the data were collected indicates that the current τ^2 is closer to 0.0045 nAmps, as estimated when $\tau^2 n_2 = 1000$, than to 0.0300 nAmps, as estimated when $\tau^2 n_2 = 83$. This is because the conductance of a single ion-channel is somewhere between 20 and 40 pSiemens and the potential across the membrane was 160 mVolts, which implies a current of between 0.0032 and 0.0064 nAmps.

An understanding of why problems arose when estimating parameters in a two-compartment model may be gained by examining the functional form of the expectations. This is a single exponential function plus a mean term and therefore only requires three parameters to describe it. However, there are four parameters in the compartment model: n_1, n_2, Q_{12}, Q_{21} ; so information on the fourth parameter can only be obtained from the covariance structure. We are led to conclude that there is very little extra information coming from the second-order moments.

7.3.4 Open one-compartment model

The problem of overparameterization can be circumvented if it is assumed that the expected number of channels in a non-conducting state at time zero (n_2) is sufficiently large to be considered infinite. There is some justification in making this assumption because it has already been noted that when $\tau^2 n_2$ is large the estimated current $\hat{\tau}^2$ takes a more reasonable value than when $\tau^2 n_2$ is small. The closed two-compartment model becomes an open

one-compartment model, which can be represented diagramatically as:



In practical terms, this means that at any particular time nearly every ion-channel is in a non-conducting state, and these provide a large reservoir of potentially conducting channels.

If it assumed that $N_1(0)$ is Poisson distributed, then this ensures that $N_1(t)$ is Poisson distributed for all later times t, and in a sense this may be considered to be the "natural" starting distribution. The conductance τ^2 can once again be treated as a scaling parameter provided that $\tau^2 r_1$ is estimated in place of r_1. The observations $\underset{\sim}{y}$ once again constitute a GARMA(1,1) process. From equations (7.2.4) and (7.2.7), with the Poisson starting distribution incorporated,

$$E(Y(t)) = -\tau^2\{n_1 P_{11}(t) + b_1(t)\},$$

$$= -\frac{\tau^2 r_1}{q_1} + \left(\frac{\tau^2 r_1}{q_1} - \tau^2 n_1\right) e^{-tq_1},$$

and $$\operatorname{cov}(Y(s),Y(t)) = \tau^4\{n_1 P_{11}(t)\, P_{11}(s-t) + b_1(t)\, P_{11}(s-t) + \sigma^2/\tau^4\, I_{st}\},$$

$$= -\tau^2\{E(Y(t))\, e^{-(s-t)q_1} - \sigma^2/\tau^2\, I_{st}\}.$$

The model was fitted by minimizing L_M and the estimates are given in table 7.3.4. Figure 7.3.1 shows the data, fitted curve of expectations and predicted values based on past observations (as described in [2.3.7]) plotted against time. L_M is minimized as -64.9, which is close to the value of -65.0 for the two-compartment model with $\tau^2 n_2 = 1000$. This is not surprising because the two

Table 7.3.3

Colquhoun's data, closed two-compartment model with $\tau^2 n_2$ fixed at 1000, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	0.134	0.0129	9.0	16.9	0.0045
se	0.003	0.0003	0.6	7.2	0.0015
correlations					
Q_{21}	0.97				
$\tau^2 n_1$	-0.61	-0.66			
σ^2/τ^2	-0.08	-0.09	0.15		
τ^2	0.06	0.07	-0.13	-0.92	

Correlation coefficients exceeding 0.9 underlined.

Table 7.3.4

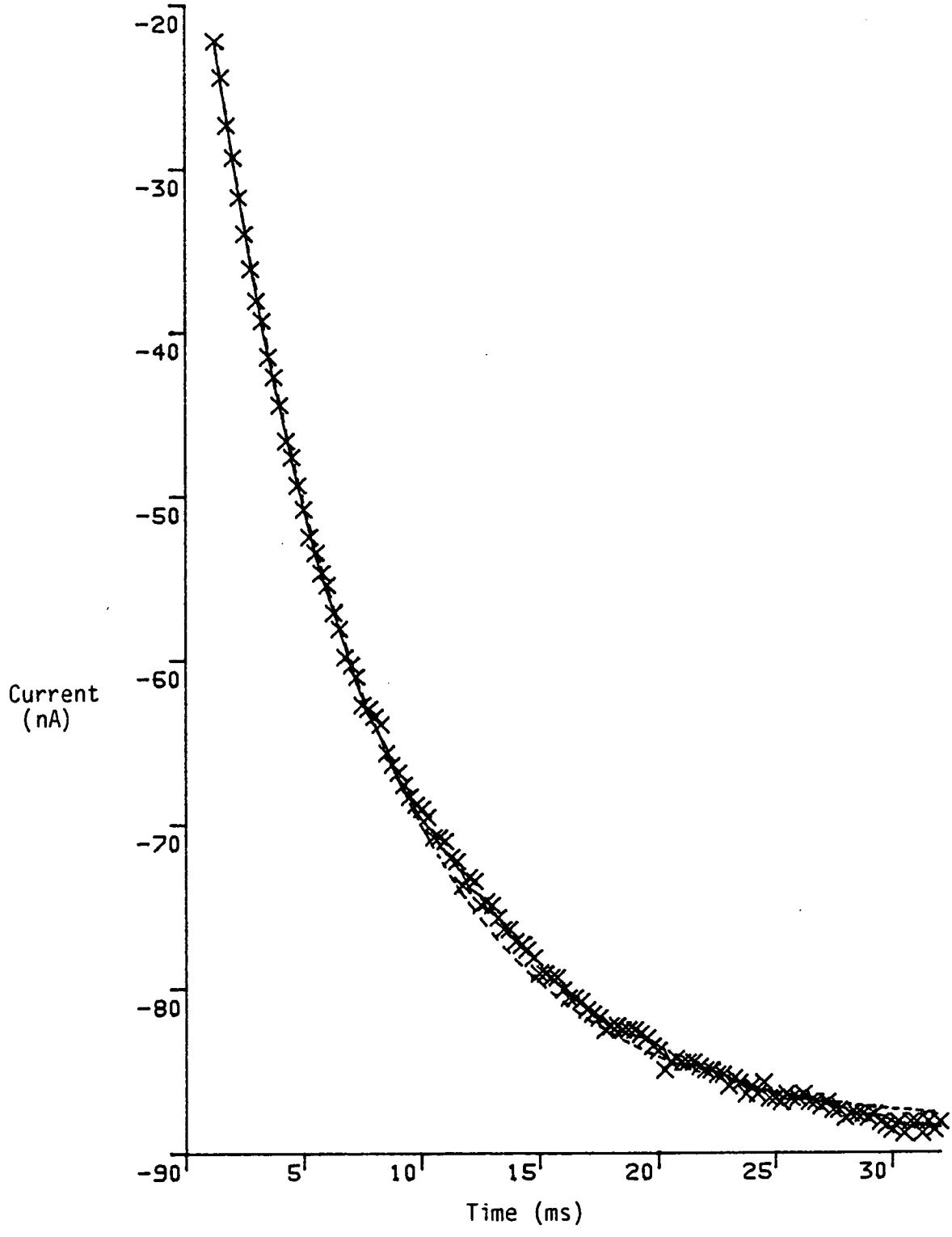
Colquhoun's data, open one-compartment model fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	$\tau^2 r_1$	q_1	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	13.0	0.147	9.0	18.3	0.0042
se	0.3	0.003	0.6	7.8	0.0014
correlations					
q_1	0.97				
$\tau^2 n_1$	-0.65	-0.63			
σ^2/τ^2	-0.09	-0.09	0.15		
τ^2	0.07	0.07	-0.13	-0.92	

Correlation coefficients exceeding 0.9 underlined.

Figure 7.3.1

Colquhoun's data, open one-compartment model fitted by Gaussian likelihood estimation; observed currents (X), expected currents (----) and predicted currents (____) conditional upon earlier observations, plotted against time.



models are very similar: as $\tau^2 n_2$ increases, the two-compartment approximates to a one-compartment model, with $(\tau^2 n_2\ Q_{21})$ converging to $\tau^2 r_1$ and Q_{12} converging to q_1 .

7.3.5 Separate covariance parameterization

To test the adequacy of the stochastic compartment model a different emigration parameter, denoted q_1^* , was introduced into the variance matrix. Therefore

$$\text{cov}(Y(s), Y(t)) = -\tau^2\{E(Y(t))\ e^{-(s-t)q_1^*} - \sigma^2/\tau^2\ I_{st}\} ,$$

and the covariances decay exponentially at rate q_1^* rather than at rate q_1 as in [7.3.4]. The reparameterization facilitates the assumption to be tested that covariances between observations decay at the same rate q_1 as the expectation function approaches its asymptote.

The parameters were estimated by minimizing L_M . The decrease in L_M to -65.3 is not sufficiently large for the addition of one extra parameter to reject the hypothesis that q_1^* is equal to q_1 , on the basis of the asymptotic properties of the likelihood ratio test. The parameter estimates are given in table 7.3.5. It can be seen that q_1 is estimated much more precisely than q_1^* , the standard errors being 0.005 and 0.106 respectively. Therefore, if the two parameters are combined into one, as they were in [7.3.4], the pooled estimate is dominated by the contribution from the component which controls the rate at which the function of expected values approaches its asymptote. This is borne out by $\hat{q}_1$ in table 7.3.4 (0.147) being much closer to $\hat{q}_1$ (0.150) in table 7.3.5 than it is to $\hat{q}_1^*$ (0.059).

Table 7.3.5

Colquhoun's data, open one-compartment model with separate covariance parameter, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	$\tau^2 r_1$	q_1	$\tau^2 n_1$	q_1^*	σ^2/τ^2	τ^2
estimate	13.2	0.150	8.8	0.059	19.7	0.0040
se	0.4	0.005	0.6	0.106	8.8	0.0015
correlations						
q_1	<u>0.96</u>					
$\tau^2 n_1$	-0.68	-0.65				
q_1^*	-0.56	-0.62	0.31			
σ^2/τ^2	-0.21	-0.22	0.21	0.26		
τ^2	0.26	0.28	-0.23	-0.39	<u>-0.93</u>	

Correlation coefficients exceeding 0.9 underlined.

Table 7.3.6

Colquhoun's data, open one-compartment model parameterized without the use of a scaling parameter, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	r_1	q_1	n_1	σ^2	τ^2
estimate	3140.	0.147	2160.0	0.076	0.0042
se	1060.	0.003	770.0	0.013	0.0014
correlations					
q_1	-0.01				
n_1	<u>0.98</u>	-0.17			
σ^2	0.35	-0.09	0.36		
τ^2	<u>0.998</u>	-0.06	<u>0.98</u>	0.35	

Correlation coefficients exceeding 0.9 underlined.

7.3.6 Estimation without use of scaling parameter

In order to obtain direct estimates of r_1, n_1 and σ^2, the conductance τ^2 cannot be treated as a scaling parameter as previously. When this reparameterized model was fitted using REGAME the convergence was very slow, which illustrates the usefulness of the earlier parameterization that enabled τ^2 to be omitted from the iterative procedure (see [2.3.4]). The final results are summarized in table 7.3.6.

7.3.7 Reparameterization using $\underset{\sim}{\beta}$

It is illuminating to relate the results of this section to those given in section 4.3, where a single exponential regression function (parameterized by $\underset{\sim}{\beta}$) with an ARMA(1,1) error process was fitted to the same set of data. The one-compartment model differs from this earlier model in only two respects: the variances change through time; and the covariances decay at the same rate as the expectation (or regression) function approaches its asymptote.

We may reparameterize the compartment model using $\underset{\sim}{\beta}$, where

$$E(Y(t)) = \beta_1 + \beta_2 e^{-t/\beta_3} ,$$

by letting

$$\tau^2 n_1 = -\beta_1 - \beta_2 ,$$

$$\tau^2 r_1 = -\beta_1/\beta_3 ,$$

and

$$q_1 = 1/\beta_3 .$$

The parameters, $\underset{\sim}{\beta}$ together with σ^2/τ^2 and τ^2 as in [7.3.4], can be estimated by minimizing L_M. The estimates, which are simply a transformation of those given in table 7.3.4, are given in table 7.3.7.

Table 7.3.7

Colquhoun's data, open one-compartment model reparameterized using regression parameters $\underset{\sim}{\beta}$, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	σ^2/τ^2	τ^2
estimate	-88.3	79.3	6.78	18.3	0.0042
se	0.5	0.7	0.16	7.8	0.0014
correlations					
β_2	-0.57				
β_3	-0.69	0.02			
σ^2/τ^2	-0.05	-0.09	0.09		
τ^2	0.03	0.09	-0.07	<u>-0.92</u>	

Correlation coefficients exceeding 0.9 underlined.

Table 7.3.8

Colquhoun's data, open one-compartment model reparameterized using regression parameters $\underset{\sim}{\beta}$, with separate covariance parameter, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	β_1	β_2	β_3	β_3^*	σ^2/τ^2	τ^2
estimate	-88.0	79.2	6.67	17.0	19.7	0.0040
se	1.0	1.0	0.23	30.6	8.8	0.0015
correlations						
β_2	-0.82					
β_3	-0.78	0.43				
β_3^*	0.57	-0.41	-0.62			
σ^2/τ^2	-0.18	0.06	0.22	-0.26		
τ^2	0.24	-0.12	-0.28	0.39	<u>-0.93</u>	

Correlation coefficients exceeding 0.9 underlined.

If a separate parameter β_3^* is used in the covariances, as in [7.3.5], where

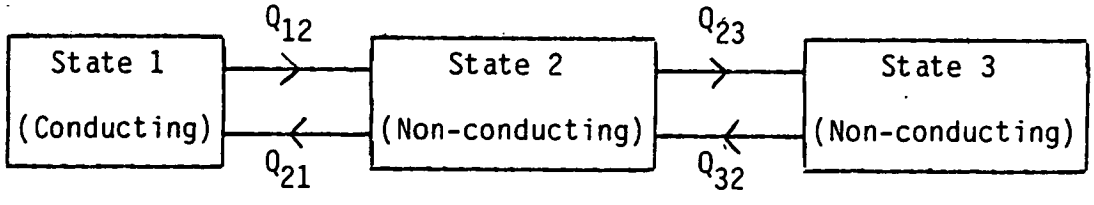
$$q_1^* = 1/\beta_3^* ,$$

and the model is fitted by minimizing L_M, then the estimates, which are simply a transformation of those given in table 7.3.5, are given in table 7.3.8.

The estimate of $\underset{\sim}{\beta}$ in table 7.3.8 is very close to that in table 4.3.3 which is not surprising because the models differ only in whether or not the variances are constant. The estimate of $\underset{\sim}{\beta}$ is also very close to that in table 7.3.7, but the standard errors are larger, because the covariances among the observations are not being used to give extra information about β_3. This illustrates the gain in efficiency of estimation, when appropriate, through using the same parameters to describe both the function of expectations and the covariance structure of data.

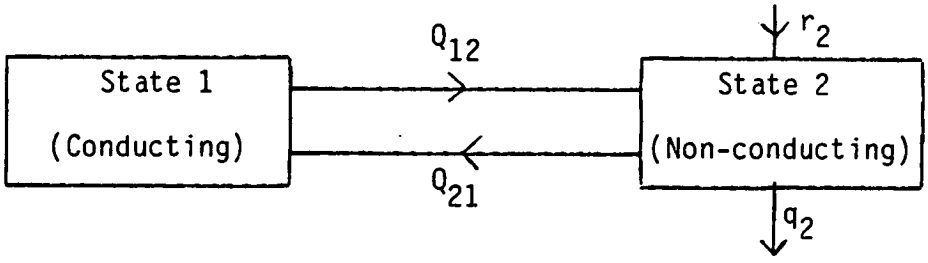
7.3.8 Open two-compartment model

Colquhoun proposed a model in which each ion-channel can be in one of three states, only the first of which is conducting, to account for the double exponential form of the regression previously fitted in [4.3.7]. The model may be represented as follows:



which requires seven parameters once n_1, n_2 and n_3 have been included. If the number of channels in the third state is considered to be large then, analogous with the preceding

simplification of a two-compartment model, the model reduces to an open two-compartment model as follows:



There are six parameters, including n_1 and n_2, to be estimated, although the expectation function can be parameterized by five. Mindful of the difficulties encountered in fitting a closed two-compartment model, I held $\tau^2 n_2$ constant in order to obtain good starting values for the optimization algorithm; values for the remaining five parameters were found so that they were consistent with:

$$E(Y(t)) = -90.4 + 25.0\, e^{-t/3.3} + 58.6\, e^{-t/9.2},$$

which was the least-squares fit of the double exponential regression function to Colquhoun's data, given in table 4.3.9. For a range of values of $\tau^2 n_2$, the values for the other parameters were found and are given in table 7.3.9. When $\tau^2 n_2$ was less than 200, values could not be found for the other parameters.

Because the system is open, the distribution of $\underset{\sim}{N}(0)$ was assumed to be Poisson, as in [7.3.4]. The inclusion of observation errors means that $\underset{\sim}{y}$ is a GARMA(2,2) process. With $\tau^2 n_2$ held constant at either 500 or 8000 the remaining parameters were estimated by minimizing L_M, and the resulting estimates are given in tables 7.3.10 and 7.3.11 respectively. The corresponding values of L_M are -74.5 and -74.9. There is little difference in L_M between the two models although the parameter estimates are

Table 7.3.9

Alternative sets of parameter values in open two-compartment model (specified in [7.3.8]) consistent with

$$E(Y(t)) = -90.4 + 25.0\, e^{-t/3.3} + 58.6\, e^{-t/9.2} .$$

Set No.	Q_{12}	Q_{21}	$\tau^2 r_2$	q_2	$\tau^2 n_1$	$\tau^2 n_2$
1	0.18	0.082	56.0	0.28	5.4	200.0
2	0.15	0.028	61.0	0.12	7.5	500.0
3	0.19	0.015	107.0	0.09	7.4	1000.0
4	0.22	0.007	248.0	0.09	7.1	2000.0
5	0.23	0.004	537.0	0.10	7.1	4000.0
6	0.25	0.002	1143.0	0.10	7.0	8000.0

Table 7.3.10

Colquhoun's data, open two-compartment model with $\tau^2 n_2$ fixed at 500, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 r_2$	q_2	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	0.24	0.031	94.0	0.135	6.9	114.0	0.00077
se	0.07	0.002	26.0	0.008	0.7	40.0	0.00025
correlations							
Q_{21}	0.97						
$\tau^2 r_2$	0.98	0.90					
q_2	0.53	0.36	0.70				
$\tau^2 n_1$	-0.69	-0.82	-0.58	-0.04			
σ^2/τ^2	-0.17	-0.22	-0.11	0.14	0.26		
τ^2	0.08	0.14	0.03	-0.17	-0.19	-0.92	

Correlation coefficients exceeding 0.9 underlined.

Table 7.3.11

Colquhoun's data, open two-compartment model with $\tau^2 n_2$ fixed at 8000, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 r_2$	q_2	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	0.28	0.0020	1420.0	0.109	6.9	131.0	0.00067
se	0.08	0.0001	490.0	0.016	0.8	90.0	0.00041
correlations							
Q_{21}	<u>0.98</u>						
$\tau^2 r_2$	<u>0.99</u>	<u>0.94</u>					
q_2	<u>0.95</u>	0.87	<u>0.99</u>				
$\tau^2 n_1$	-0.79	-0.87	-0.71	-0.62			
σ^2/τ^2	0.05	0.09	0.03	-0.00	-0.15		
τ^2	-0.09	-0.12	-0.06	-0.03	0.17	<u>-0.98</u>	

Correlation coefficients exceeding 0.9 underlined.

quite different. It would appear that once again we are trying to estimate more parameters than the data can identify. However, the agreement between the data and the fitted models is excellent. For example, figure 7.3.2 shows the fit when $\tau^2 n_2$ is held constant at 8000.

One solution would be to assume that $\tau^2 n_2$ is large. The system would reduce to one-compartment but the rate of immigration would change with time according to an exponential relationship and this would move us outside the theory of section 7.2 in which all rate parameters are constant over time.

7.3.9 Resumé

The open one-compartment system appears to be successful in modelling the data. This is in accord with existing belief in the scientific community that ion-channels behave approximately like particles in stochastic compartment models. However, the estimation procedure does not seem to extend readily to larger systems because of lack of identifiability of larger numbers of parameters.

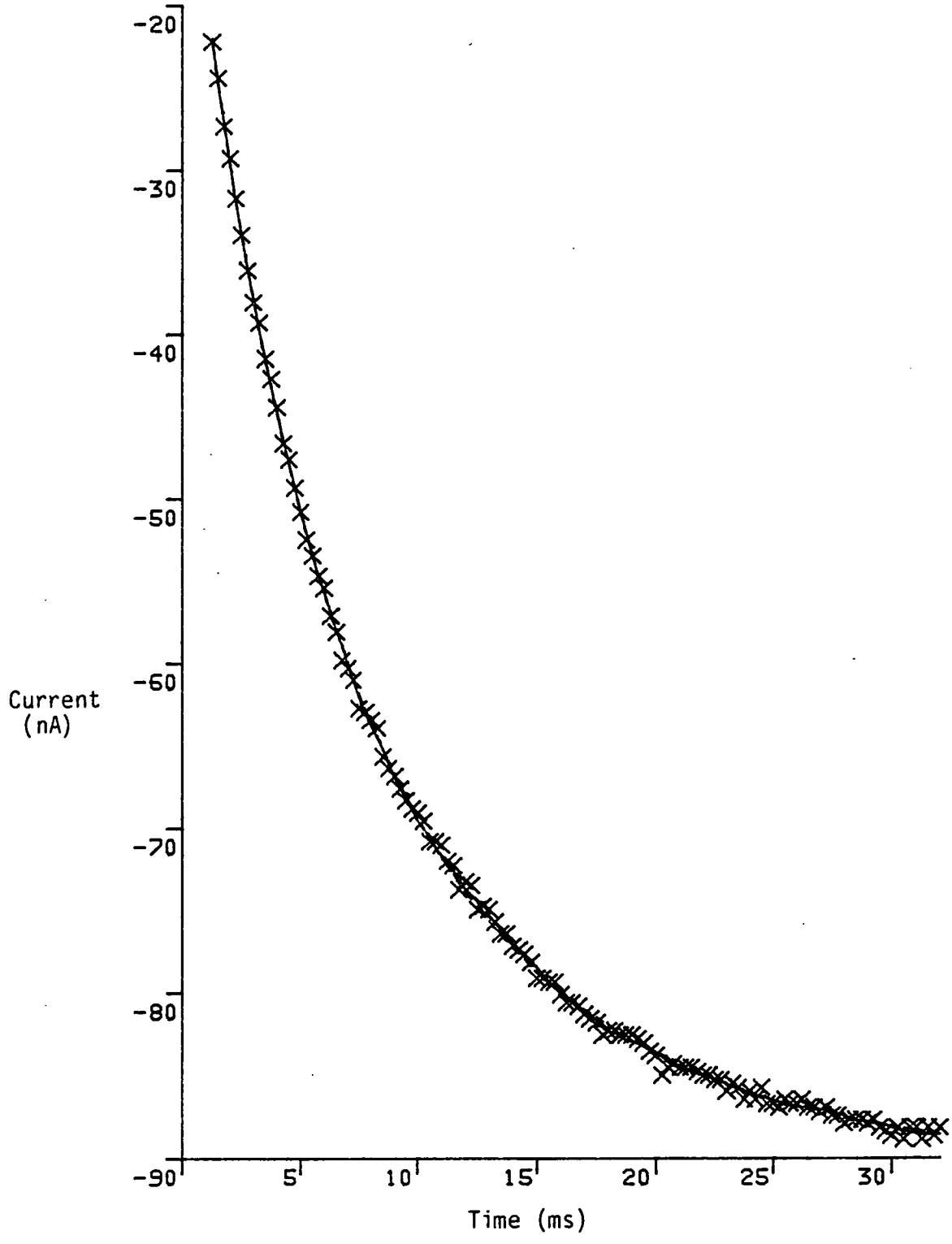
7.4 Dale's data

7.4.1 Introduction

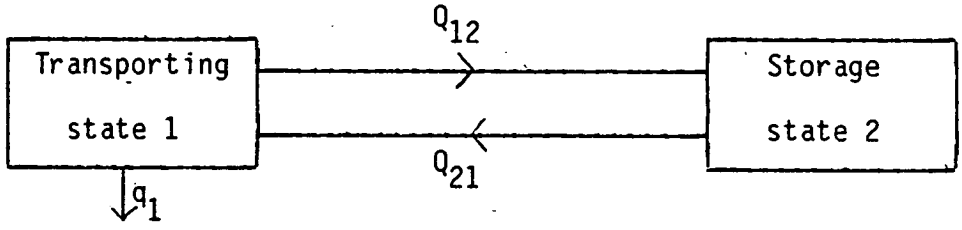
Compartment models have been proposed to explain leaf transport systems, see for example Bauermeister, Dale, Williams and Scobie, 1980. Dale's data, first considered in section 4.4, is an example of this type of data. It is postulated that ^{14}C atoms in

Figure 7.3.2

Colquhoun's data, open two-compartment model with $\tau^2 n_2$ fixed at 8000, fitted by Gaussian likelihood estimation; observed currents (X), expected currents (----) and predicted currents (____) conditional upon earlier observations, plotted against time.



the leaf may be either in molecules which are available for transportation out of the leaf or in molecules which are stored within the leaf. On the assumption that molecules switch between states or leave the leaf independently of one another, at probabilistic rates constant over time, the process may be modelled using an open two-compartment system, thus:



The observations $\underset{\sim}{y}$ are a constant proportion of the ^{14}C atoms in the two compartments, with the constant (denoted τ^2) being the rate of emission of β-particles from a single ^{14}C atom multiplied by the rate of detection by a Geiger counter. As in section 7.3, τ^2 may be treated as a global scaling parameter by reparameterizing.

7.4.2 Estimation

The distribution of $\underset{\sim}{N}(0)$ was assumed to be Poisson. Observation errors were included in the model with variance $\sigma^2 E(Y(t))$ at time t, as in section 4.4, to take account of the Poisson nature of the sampling variability associated with radioactive emission. With these assumptions, the observed counts $\underset{\sim}{y}$ form a GARMA(1,1) process.

There are five parameters in the model, including n_1 and n_2, but the expectation which is a weighted sum of two exponentials has only four, so we may expect problems of indeterminancy. To obtain good starting values for the

optimization algorithm, $\tau^2 n_2$ was held constant and values for the remaining four parameters found consistent with:

$$E(Y(t)) = 818\ e^{-0.0142t} + 702\ e^{-0.00067t},$$

which was the least-squares fit of the double exponential regression function to Dale's data, given in table 4.4.1. For a range of values of $\tau^2 n_2$, the values for the other parameters were found and are given in table 7.4.1.

The parameters were estimated by minimizing L_M for the two extreme cases in table 4.4.1, that is with $\tau^2 n_2 = 0$ in one case and with $Q_{12} = 0$ in the other case, and the estimates of the remaining parameters are given in tables 7.4.2 and 7.4.3 respectively. In both cases, and also in fact in all intermediate cases between these extremes, L_M was minimized as 387.4. Therefore, purely on the basis of the data, it is impossible to distinguish among the alternative sets of parameter values.

7.4.3 Temporary immigration to generate the initial distribution

It is necessary to introduce further information about the experiment which generated the data in order to reduce the number of parameters in the model. ^{14}C was fed to the leaf as radioactive carbon dioxide at a constant rate for five minutes prior to time zero. This can be represented as temporary immigration at rate r_1 into the first compartment between times -5 and 0. After time zero the immigration rate is zero. This single parameter replaces both n_1 and n_2.

The new parameter set was estimated by minimizing L_M and the estimates are given in table 7.4.4. Once again L_M is equal to 387.4. Temporary immigration in this case is equivalent to

Table 7.4.1

Alternative sets of parameter values in two-compartment model consistent with $E(Y(t)) = 818\ e^{-0.0142t} + 702\ e^{-0.00067t}$.

Set number	Q_{12}	Q_{21}	q_1	$\tau^2 n_1$	$\tau^2 n_2$
1	0.00572	0.00120	0.0080	1520.0	0.0
2	0.00467	0.00104	0.0092	1320.0	200.0
3	0.00320	0.00088	0.0108	1120.0	400.0
4	0.00165	0.00076	0.0125	970.0	550.0
5	0.00	0.00067	0.0142	851.0	669.0

Table 7.4.2

Dale's data, two-compartment model with $\tau^2 n_2$ fixed at 0, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	q_1	$\tau^2 n_1$	σ^2/τ^2	τ^2
estimate	0.0080	0.0015	0.0101	1606.0	0.034	3.6
se	0.0014	0.0002	0.0011	80.0	0.013	1.1
correlations						
Q_{21}	0.65					
q_1	0.68	0.23				
$\tau^2 n_1$	0.10	0.03	0.16			
σ^2/τ^2	-0.15	0.02	-0.09	-0.04		
τ^2	0.07	-0.05	0.12	-0.12	-0.89	

Table 7.4.3

Dale's data, two-compartment model with Q_{12} fixed at 0, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{21}	q_1	$\tau^2 n_1$	$\tau^2 n_2$	σ^2/τ^2	τ^2
estimate	0.0008	0.0188	860.0	745.0	0.034	3.6
se	0.0001	0.0023	67.0	63.0	0.013	1.1
correlations						
q_1	0.45					
$\tau^2 n_1$	-0.27	-0.29				
$\tau^2 n_2$	0.34	0.48	-0.25			
σ^2/τ^2	0.10	-0.13	0.07	-0.12		
τ^2	-0.05	0.09	-0.06	-0.09	-0.89	

Table 7.4.4

Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 r_1$	q_1	σ^2/τ^2	τ^2
estimate	0.0079	0.0015	329.0	0.0103	0.034	3.6
se	0.0013	0.0002	17.0	0.0011	0.013	1.1.
correlations						
Q_{21}	0.65					
$\tau^2 r_1$	0.14	0.04				
q_1	0.68	0.23	0.21			
σ^2/τ^2	-0.15	0.03	-0.04	-0.09		
τ^2	0.06	-0.05	-0.11	0.12	-0.89	

assuming that the coefficients in $\underset{\sim}{N}(0)$ are both Poisson distributed, with $\tau^2 n_1 = 1574$ and $\tau^2 n_2 = 31$. The data, fitted curve of expected values and predicted values from past observations (as described in [2.3.7]) are plotted against time in figure 7.4.1.

7.4.4 Separate covariance parameterization

The models fitted so far in this section assume that the covariances between observations decay to zero with increasing time separation at a rate which is the sum of two exponentials, which are the same two exponentials that describe the rate of decay of the function of expectations to zero. To test this assumption, a different set of parameters corresponding to Q_{12}, Q_{21} and q_1 were used in the covariances, denoted Q^*_{12}, Q^*_{21} and q^*_1. These are the three parameters which jointly specify the two exponential terms. This is analogous to the approach taken in [7.3.5] when the assumption of a common single exponential term was tested in Colquhoun's data.

The parameters were estimated by minimizing L_M and the estimates are given in table 7.4.5. Figure 7.4.2 shows the data, fitted curve and predicted values plotted against time. The value of L_M has been reduced to 384.4. This is a decrease of 3.0 compared with the results in [7.4.3] for the addition of 3 parameters which, on the basis of asymptotic likelihood theory, is not a sufficient change to reject the model with shared parameters. This uses the property that the Guassian likelihood converges to the true likelihood function as the number of particles in the compartment system increases.

Figure 7.4.1

Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), fitted by Gaussian likelihood estimation; observed radioactive discharges (X), fitted values (----) and predicted values (____) conditional upon earlier observations, plotted against time.

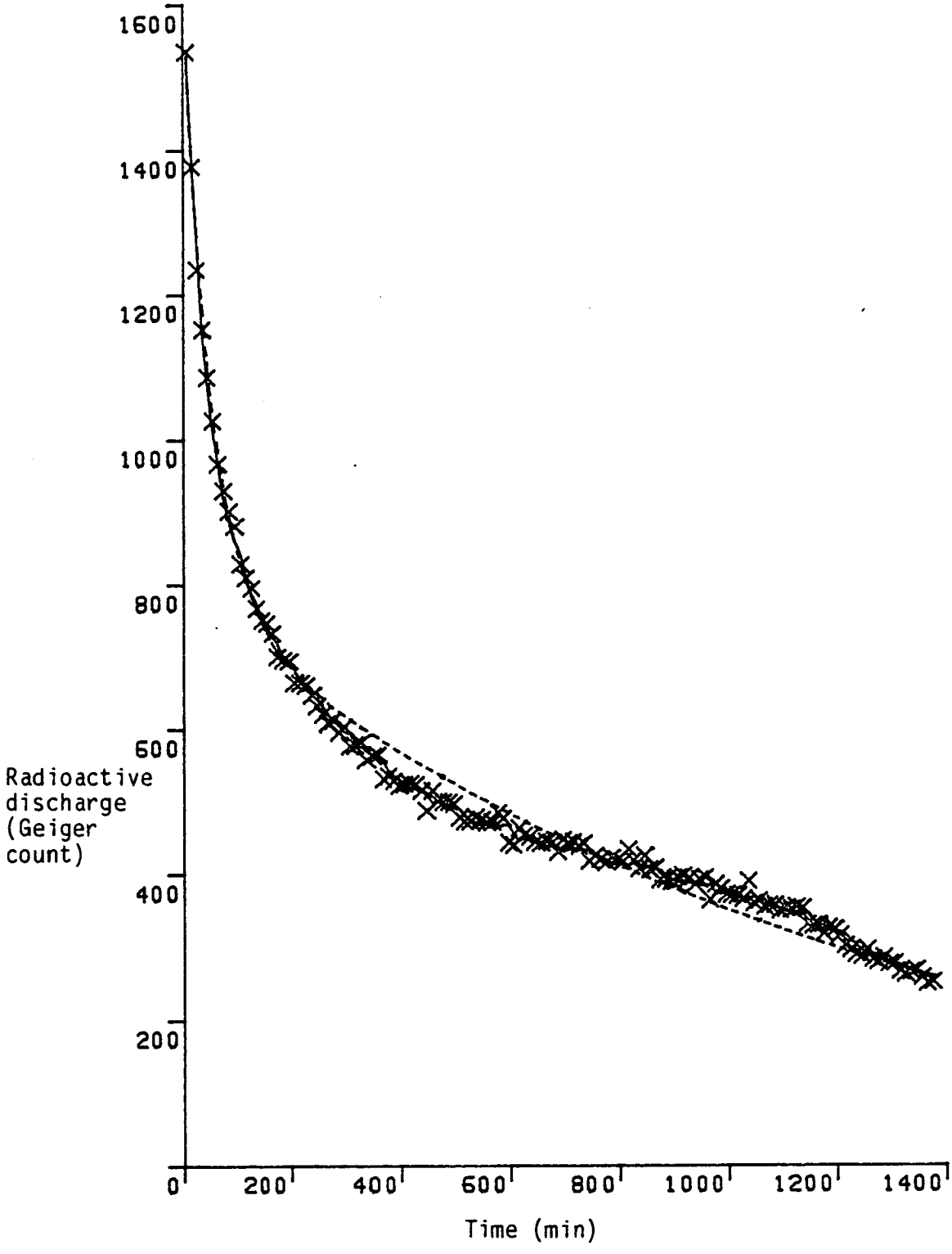


Table 7.4.5

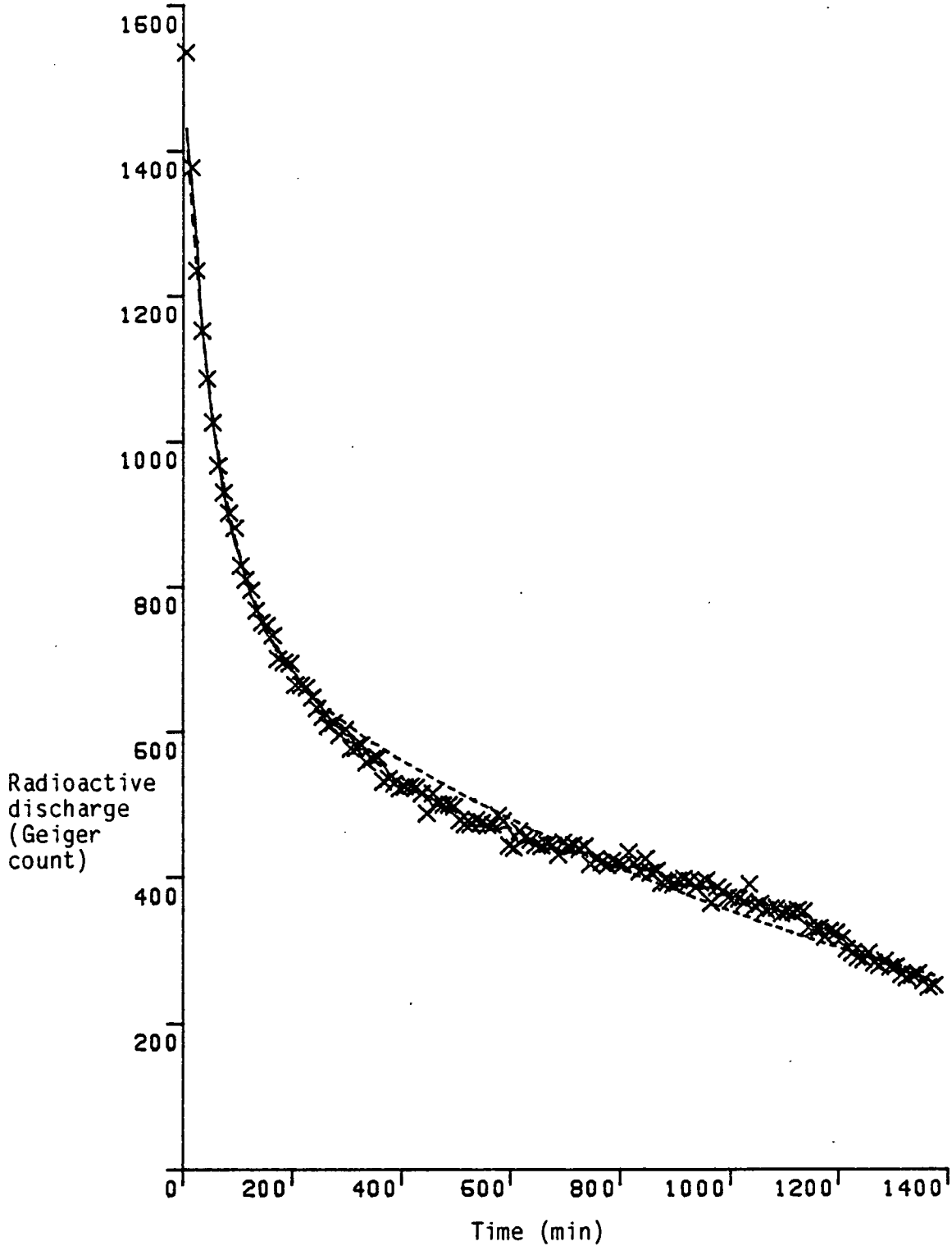
Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]) and separate covariance parameterization, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	$\tau^2 r_1$	q_1	Q^*_{12}	Q^*_{21}	q^*_1	σ^2/τ^2	τ^2
estimate	0.0067	0.0014	304.0	0.0082	0.0197	0.0006	0.18	0.0056	21.3
se	0.0011	0.0002	16.0	0.0012	0.0352	0.0006	0.15	0.0105	39.6
correlation									
Q_{21}	0.54								
$\tau^2 r_1$	0.70	0.24							
q_1	0.77	0.35	0.42						
Q^*_{12}	0.25	0.06	0.23	0.25					
Q^*_{21}	0.25	0.21	0.10	0.38	-0.03				
q^*_1	0.02	-0.01	0.05	0.05	-0.76	-0.13			
σ^2/τ^2	0.20	0.12	0.12	0.25	0.86	0.35	-0.88		
τ^2	-0.22	-0.11	-0.17	-0.25	-0.87	-0.35	0.88	<u>-0.995</u>	

Correlation coefficients exceeding 0.9 underlined.

Figure 7.4.2

Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), and separate covariance parameterization, fitted by Gaussian likelihood estimation; observed radioactive discharges (X), fitted values (----) and predicted values (____) conditional upon earlier observations, plotted against time.



7.4.5 Estimation without use of scaling parameter and with σ^2 fixed at 0.1

To obtain direct estimates of r_1, q_1 and σ^2, the rate of emission τ^2 cannot be treated as a scaling parameter. The estimates resulting from minimizing L_M are given in table 7.4.6.

Particularly worthy of note is the estimate of σ^2 as 0.12. The observations are ten-minute averages of Geiger counts, so if in any single minute the distribution is Poisson, then the variance of the average will be one tenth of the mean. To test this, σ^2 was fixed at 0.1 and the other parameters were re-estimated. The results are summarized in table 7.4.7 and figure 7.4.3 shows the fitted values plotted against time. L_M is minimized at 388.0 which is only 0.6 more than when $\sigma^2 = 0.12$. Therefore the data is consistent with the hypothesis that σ^2 is equal to 0.1, on the basis of the asymptotic properties of the likelihood ratio test.

7.4.6 Resumé

Although the preceding analysis has appeared to produce a model which fits the data adequately, one particular feature nullifies this conclusion: the estimate of τ^2 is 3.6. If molecules in the leaf behave independently of one another then τ^2 should represent the rate of emission of a single ^{14}C atom multiplied by the rate of detection of a Geiger counter. However, the product of these rates is miniscule and the value 3.6 is utter nonsense. Therefore the stochastic compartment model does not describe the transport mechanism, maybe because molecules do not act independently but in groups.

There is certainly not the same degree of belief in the stochastic compartment model describing the transport mechanism as

Table 7.4.6

Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), parameterized without the use of a scaling parameter, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	r_1	q_1	σ^2	τ^2
estimate	0.0079	0.0015	92.0	0.0103	0.12	3.6
se	0.0013	0.0002	28.0	0.0011	0.02	1.1
correlations						
Q_{21}	0.65					
r_1	-0.04	0.06				
q_1	0.68	0.23	-0.08			
σ^2	-0.22	-0.03	0.15	0.01		
τ^2	0.06	-0.05	<u>-0.99</u>	0.12	-0.20	

Correlation coefficients exceeding 0.9 underlined.

Table 7.4.7

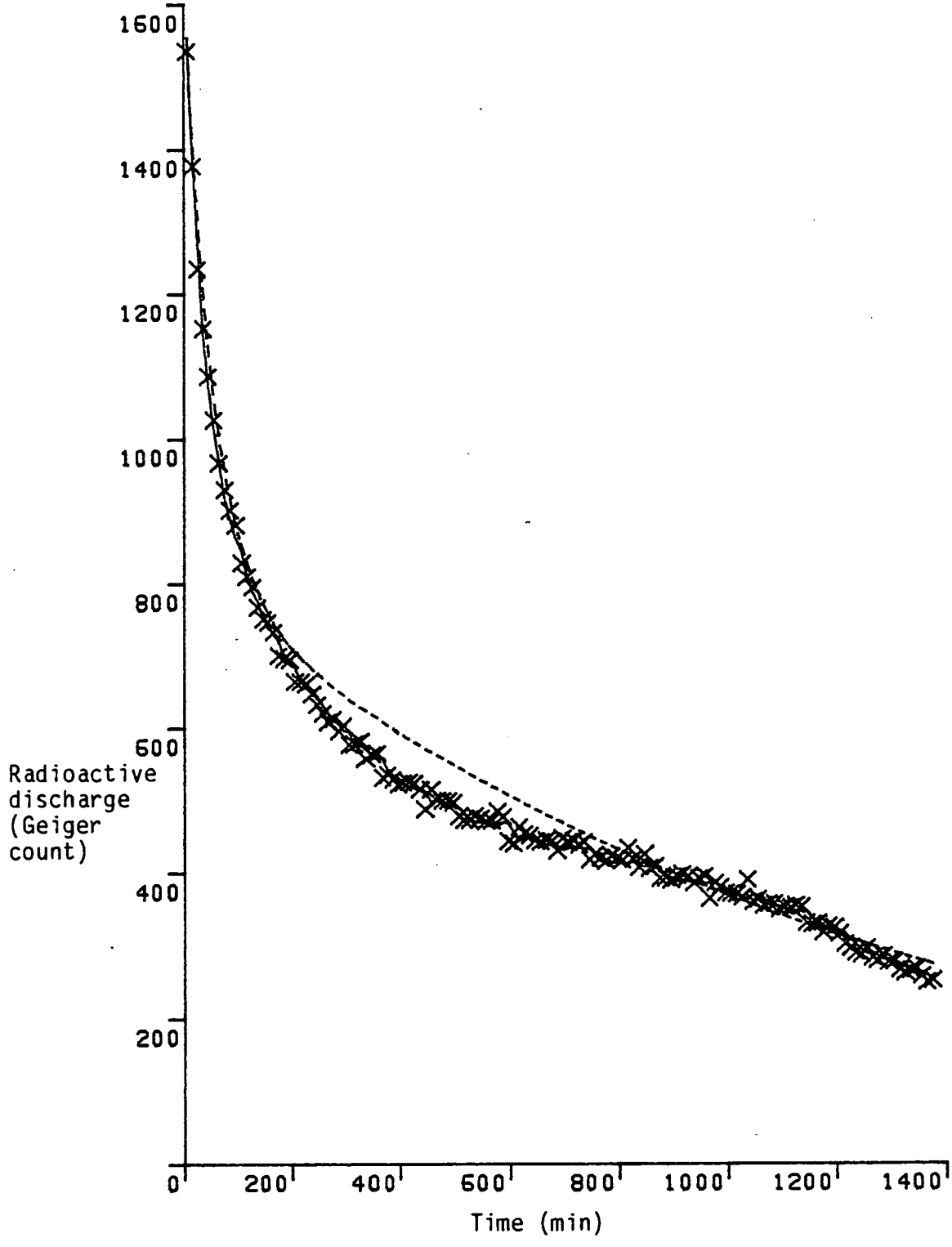
Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), σ^2 fixed at 0.1, parameterized without the use of a scaling parameter, fitted by Gaussian likelihood estimation; parameter estimates with associated estimated standard errors and estimated correlation coefficients.

	Q_{12}	Q_{21}	r_1	q_1	τ^2
estimate	0.0082	0.0015	85.0	0.0103	3.9
se	0.0014	0.0002	26.0	0.0012	1.1
correlations					
Q_{21}	0.65				
r_1	-0.03	0.06			
q_1	0.70	0.22	-0.07		
τ^2	0.04	-0.06	<u>-0.99</u>	0.11	

Correlation coefficients exceeding 0.9 underlined.

Figure 7.4.3

Dale's data, two-compartment model with temporary immigration (specified in [7.4.3]), σ^2 fixed at 0.1, fitted by Gaussian likelihood estimation; observed radioactive discharges (X), fitted values (----) and predicted values (____) conditional upon earlier observations, plotted against time.



there is in it describing the behaviour of ion-channels, considered in section 7.3. At best it is an approximation to reality. Therefore, although it may be reasonable to expect the long term trend in the data to be adequately described by a compartment model, it may be taking the model too far to expect it also to explain the covariance structure of the observations. In that case the empirical approach of chapter 4 would seem to be preferable to the analysis in the present section which may be compared to building a tower on sand. However, the method of chapter 4 also has its deficiencies, a partial resolution of which will be presented in chapter 9.

7.5 Discussion

The technique considered in this chapter is to model data, which are postulated to have arisen from a compartment system, by a simple type of stochastic compartment model. The parameters in the model are estimated utilizing the information both in the main trend of the data and in the local stochastic fluctuations. There are theoretical reasons for proposing compartment models as generating mechanisms for both Colquhoun's and Dale's data. Judged on the basis of goodness-of-fit, stochastic compartment models appear to perform well in both applications. However, at this point the similarities end. For Colquhoun's data, the parameter estimates were consistent with other information about values they should take. But, for Dale's data this was not the case. I think the difference between the two problems is that in one the stochastic model comes very close to describing the true

mechanism, whereas in the other, although a deterministic model may be a reasonable approximation to reality, the sources of stochastic variability are much more complex than they are assumed to be in a stochastic compartment model. Misleading results can be produced in fitting an inappropriate stochastic model because parameters do not have the same interpretation as in a deterministic model, and may have no physical meaning at all. Therefore, if there is any doubt about the validity of a stochastic compartment model, it would be better to fit a deterministic model with an empirical choice of error model, as described in chapters 4 and 5.

8. Cumulative count models

8.1 Introduction

This is the last of three chapters describing the use of mechanistic models in order to estimate regression parameters when errors are serially correlated. In the approach adopted here the correlations arise through cumulative numbers of events up to particular times being used as the dependent variate. The situation envisaged is where a fixed number of events, N say, occur in a particular interval of time, and the times of the events are independent identically distributed random variables.

In section 8.2, the variate indexed by continuous time of the cumulative number of events is shown to be a cGARMA(1,0) process and the generalized sum of squares, S, (defined in [2.3.2]) is shown to be Berkson's χ^2 statistic. This helps to illuminate the relation between the correlated structure of the cumulative numbers of events and the approximate independence of events in non-overlapping time intervals. However, the formulation as a GARMA process permits other types of correlation structure to be used intermediate between that predicted by the mechanistic model and the simple-minded assumption of independence. In section 8.3, the technique is applied to seed germination data previously analysed by Hunter, Glasbey and Naylor (1984). Finally, in section 8.4 the technique is evaluated critically.

8.2 Theory

8.2.1 Model

The model to be considered is for a sequence of observations $y_1, \ldots, y_n$ (denoted $\underset{\sim}{y}$) which have been made at monotonically increasing times $t_1, \ldots, t_n$ (denoted $\underset{\sim}{t}$), where y_i is the number of events that have occurred up to time t_i. It is assumed that $\underset{\sim}{y}$ is a realisation at discrete times $\underset{\sim}{t}$ of a random variate $Y(t)$ indexed by continuous time, which is itself the sum of N independent random variates $Y^{\{1\}}(t), \ldots, Y^{\{N\}}(t)$, where $Y^{\{k\}}(t)$ is an indicator random variable that denotes whether or not the event indexed by k occurred before time t. If the event labelled k occurs at time $t^{\{k\}}$, then

$$Y^{\{k\}}(t) = 0 \qquad \text{if } t < t^{\{k\}}$$
$$1 \qquad \text{if } t \geq t^{\{k\}} .$$

It is further assumed that $t^{\{k\}}$ is a positive random variate with cumulative probability function $p(t)$, therefore

$$P(t^{\{k\}} \leq t) = p(t) \qquad \text{for all } t \geq 0 .$$

8.2.2 Solution

From [8.2.1] it follows that

$$E(Y(t)) = \sum_{k=1}^{N} E(Y^{\{k\}}(t)) ,$$
$$= N\, E(Y^{\{1\}}(t))$$

because the N events occur independently of one another and are identically distributed,

$$= N\, P(t \geq t^{\{1\}})$$

by the definition of $Y^{\{1\}}(t)$ above,

$$= Np(t) .$$

Also,

$$\begin{aligned}
\mathrm{cov}(Y(s),Y(t)) &= \sum_{k=1}^{N} \mathrm{cov}(Y^{\{k\}}(s),\, Y^{\{k\}}(t)) \,, \\
&= N\, \mathrm{cov}(Y^{\{1\}}(s),\, Y^{\{1\}}(t)) \,, \\
&= N\,[E(Y^{\{1\}}(s)Y^{\{1\}}(t)) - E(Y^{\{1\}}(s))E(Y^{\{1\}}(t))] \,, \\
&= N[P(s,t > t^{\{1\}}) - P(s > t^{\{1\}})P(t > t^{\{1\}})] \,, \\
&= N(p(t)-p(s)p(t)) \qquad\qquad \text{for } s > t \,, \\
&= N(1-p(s))p(t) \,.
\end{aligned}$$

This is the product of a term dependent only upon s and a term dependent only upon t, so $Y(t)$ satisfies the definition in [2.2.7] and is a cGARMA(1,0) process. As a consequence, $\underset{\sim}{y}$ is a GARMA(1,0) process.

At this point it is worth noting that another way of viewing the above model is as an open one-compartment system with N particles in the system at time zero, no immigration and a time-varying rate of emigration,

$$q(t) = \frac{1}{1-p(t)} \frac{\partial p(t)}{\partial t} \,,$$

provided $p(t)$ is differentiable with respect to t and not equal to unity. This is closely related to the model applied by Matis and Wehrly (1979) to fish survival data in their review of stochastic compartment models.

8.2.3 Relation between the generalized sum of squares and Berkson's χ^2 statistic

In what follows I forge a link between the approach to correlated errors taken in this thesis and the more conventional way of tackling the particular type of covariance structure associated with cumulative counts data. At the same time, this

serves as another illustration of the decomposition of the variance of a GARMA process in terms of lower-triangular band matrices, as specified in [2.2.2].

The vector of observations $(\underset{\sim}{y})$ is a realisation of a multivariate distribution with mean $\underset{\sim}{f}$ and variance matrix $\underset{\sim}{V}$, where

$$f_i = Np_i \qquad \text{for } i = 1,\ldots,n ,$$

$$V_{ij} = N(1-p_i)p_j \qquad \text{for } i \geq j ,$$

and $p(t_i)$ is denoted p_i. Because $\underset{\sim}{y}$ is a GARMA(1,0) process, from [2.2.2] $(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi})^T$ is diagonal, where $\underset{\sim}{\Phi}$ is a lower triangular matrix with bandwidth 2 defined by

$$\Phi_{ii} = 1 \qquad \text{for } i = 1,\ldots,n$$

and

$$\Phi_{ii}V_{i,i-1} + \Phi_{i,i-1}V_{i-1,i-1} = 0 \qquad \text{for } i = 2,\ldots,n .$$

Therefore

$$\Phi_{i,i-1} = \frac{-(1-p_i)}{(1-p_{i-1})} \qquad \text{for } i = 2,\ldots,n .$$

It follows that

$$(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)_{ii} = \Phi_{ii}^2V_{ii} + 2\Phi_{ii}\Phi_{i,i-1}V_{i,i-1} + \Phi_{i,i-1}^2V_{i-1,i-1}$$

$$= N(1-p_i)p_i - 2\frac{(1-p_i)}{(1-p_{i-1})}N(1-p_i)p_{i-1} + \frac{(1-p_i)^2}{(1-p_{i-1})^2}N(1-p_{i-1})p_{i-1}$$

$$\text{for } i = 1,\ldots,n ,$$

where p_0 is taken to be zero,

$$= \frac{N(1-p_i)(p_i-p_{i-1})}{(1-p_{i-1})} .$$

From [2.2.2], $\underset{\sim}{\Theta}$ is a diagonal matrix with

$$\Theta_{ii} = \sqrt{(\underset{\sim}{\Phi}\,\underset{\sim}{V}\,\underset{\sim}{\Phi}^T)_{ii}} \qquad \text{for } i = 1,\ldots,n .$$

In [2.3.3], the generalized sum of squares, S , is defined by

$$S = \underset{\sim}{z}^T \underset{\sim}{z} \quad ,$$

where

$$\underset{\sim}{z} = \underset{\sim}{\Theta}^{-1} \underset{\sim}{\Phi}\, \underset{\sim}{e} \quad ,$$

and $\underset{\sim}{e} = \underset{\sim}{y} - \underset{\sim}{f}$.

Therefore,

$$z_i = (\Phi_{ii} e_i + \Phi_{i,i-1} e_{i-1})/\Theta_{ii} \qquad \text{for } i = 1,\ldots,n \ ,$$

where e_o is defined to be zero,

$$= \frac{e_i - (1-p_i)e_{i-1}/(1-p_{i-1})}{\sqrt{\{N(1-p_i)(p_i-p_{i-1})/(1-p_{i-1})\}}} \ ,$$

and

$$S = \sum_{i=1}^{n} \frac{e_i^2 - 2(1-p_i)e_i e_{i-1}/(1-p_{i-1}) + (1-p_i)^2 e_{i-1}^2/(1-p_{i-1})^2}{N(1-p_i)(p_i-p_{i-1})/(1-p_{i-1})} \ ,$$

$$= \sum_{i=1}^{n} e_i^2 \left[\frac{(1-p_{i-1})}{N(1-p_i)(p_i-p_{i-1})} + \frac{(1-p_{i+1})}{N(1-p_i)(p_{i+1}-p_i)} \right]$$

$$- \sum_{i=2}^{n} \frac{2e_i e_{i-1}}{N(p_i-p_{i-1})}$$

with p_{n+1} taken as 1 . Rearrangement of terms leads to

$$S = \sum_{i=1}^{n} e_i^2 \left[\frac{1}{N(p_i-p_{i-1})} + \frac{1}{N(p_{i+1}-p_i)} \right] - \sum_{i=2}^{n} \frac{2e_i e_{i-1}}{N(p_i-p_{i-1})}$$

$$= \sum_{i=1}^{n+1} \frac{(e_i-e_{i-1})^2}{N(p_i-p_{i-1})}$$

with e_{n+1} defined to be zero. This is Berkson's χ^2 statistic: the ith term has a numerator: $(e_i - e_{i-1})^2$ which is the squared difference between the observed and the expected number of events between times t_{i-1} and t_i , and a denominator equal to the

expected number of events. Thus, there are two routes to the same optimization function S: either take account of the correlation structure between cumulative counts and form a generalized sum of squares; or form a new variate, the number of events between observation times, and construct Berkson's χ^2 statistic.

8.3 Naylor's data: seed germination tests

8.3.1 Introduction

Approximately one hundred seeds were placed in each of 33 petri-dishes inside a germination cabinet. Every few hours for the following week the seeds which had germinated (that is had a radicle at least 4mm in length) were counted and removed. The numbers of seeds remaining ungerminated at the end of the test were also counted. The number of observations per dish varied between 7 and 12. These data have been used by Hunter, Glasbey and Naylor (1984) to illustrate an efficient way to fit a curve to cumulative counts taking account of the correlations among the observations. Previous methods presented in the seed germination literature (see for example Nicholas and Heydecker, 1968) had inappropriately assumed the observations to be independent. In this section the flexibility of REGAME in handling different variance matrices will be used to illustrate the use of the model considered in section 8.2 and to explore the use of related models.

8.3.2 Models

The germination times of seeds are assumed to be log-normally distributed with a time origin 48 hours after the start of the experiment, except for a probability of non-germination (denoted β_3). Therefore the probability of a seed having germinated by observation time $t_i (> 48)$ is

$$p_i = (1-\beta_3) \int_{-\infty}^{(\ell n(t_i-48)-\beta_1)/\sqrt{\beta_2}} \frac{e^{-x^2/2}}{\sqrt{(2\pi)}}\, dx$$

for $i = 1,\ldots,n$,

where β_1 and β_2 are the mean and variance of the distribution of log-times to germination of seeds which germinate. This is the distribution that was found to be appropriate by Hunter, Glasbey and Naylor (1984).

Several different models for the variance matrix ($\underset{\sim}{V}$) of the observations will be considered:

(0) Independent observations with binomial type variances:

$$V_{ij} = \tau^2 N(1-p_i)p_i I_{ij} ,$$

where τ^2 is a variance parameter and I_{ij} denotes the (ij)th coefficient of the identity matrix ($\underset{\sim}{I}$) of size n. This is a type of variance matrix commonly assumed in the germination literature, although it should be inappropriate for cumulative data.

(1) The standard variance structure, for which

$$V_{ij} = N(1-p_i)p_j \qquad \text{for } i \geq j ;$$

shown in [8.2.2] to be appropriate to cumulative data, where N is the number of seeds in a petri-dish.

(2) The standard variance structure in (1) together with an independent observation error with a standard deviation

proportional to the rate of germination, denoted error type (a), this leads to variance matrix coefficients:

$$V_{ij} = N(1-p_i)p_j + \left(\sigma^{\{a\}} \frac{\partial p_i}{\partial t_i}\right)^2 I_{ij} \qquad \text{for } i > j ,$$

where $\sigma^{\{a\}}$ is a proportionality constant.

(3) The standard variance structure in (1) together with an independent observation error with constant standard deviation $\sigma^{\{b\}}$, denoted error type (b), this leads to variance matrix coefficients:

$$V_{ij} = N(1-p_i)p_j + (\sigma^{\{b\}})^2 I_{ij} \qquad \text{for } i > j .$$

(4) A scaled version of the standard variance structure in (1), for which

$$V_{ij} = \tau^2 N(1-p_i)p_j \qquad \text{for } i > j ,$$

with τ a proportionality constant. This may for example result from the responses of seeds being correlated rather than independent.

(5) A scaled version of the variance structure in (1) plus observation error of the type denoted (a), which leads to variance matrix coefficients:

$$V_{ij} = \tau^2 N(1-p_i)p_j + \left(\sigma^{\{a\}} \frac{\partial p_i}{\partial t_i}\right)^2 I_{ij} \qquad \text{for } i > j .$$

(6) A scaled version of the variance structure in (1) plus an observation error of the type denoted (b), which leads to variance matrix coefficients:

$$V_{ij} = \tau^2 N(1-p_i)p_j + (\sigma^{\{b\}})^2 I_{ij} \qquad \text{for } i > j .$$

If observation errors arise through misclassifying seeds that are on the point of germinating then error type (a) will be appropriate, because the chance of misclassifying is proportional to the rate of germination at that time. Observation error type

(b) has been included to provide an alternative form for the variances. Inclusion of a scaled version of the variance structure enables the magnitude of the variances assumed in model (1) to be tested.

Model (0) corresponds to a GARMA(0,0) process, models (1) and (4) to GARMA(1,0) processes and the remaining models to GARMA(1,1) processes.

8.3.3 Method of estimation

Among the various models only (1) could be fitted by minimizing the generalized sum of squares S (which was shown in [8.2.3] to be equivalent to Berkson's χ^2 statistic) because the other models have a scaling term in the variance matrix. Models (0) and (4) have single global scaling parameters and so could be fitted by minimizing L_S, defined in [2.3.2]. The other models require the term $\ell n|\underset{\sim}{V}|$ to be in the function to be minimized, so the choice when using the computer program REGAME is between L_M and L_R, as defined in [2.3.2]. In this application L_M is the negative Gaussian log-likelihood (Whittle, 1961) and not the true negative log-likelihood, because the observations are not normally distributed. When minimizing L_M, the scaling parameters are estimated without taking account of the number of other parameters to be estimated. This can introduce severe biases when the average number of observations per petri-dish is only ten and three other parameters are being estimated. Therefore the negative residual Gaussian log-likelihood L_R was chosen as the optimizing function, and for consistency was used to fit all seven models.

To obtain good initial estimates for the parameters in the iterative optimization routine the models were fitted

sequentially. Parameter values obtained by Hunter, Glasbey and Naylor (1984) were used to provide starting values for model (1). The parameter estimates from model (1) were then used as starting values for models (0), (2), (3) and (4), and those finally obtained from model (4) were used as starting values for models (5) and (6). With this scheme no convergence problems were encountered. I could have eliminated one parameter from the iterative procedure by reparameterizing $\underset{\sim}{V}$ to have a global scaling parameter, as I did in chapter 7. But, because there were no convergence difficulties I preferred to retain the original parameters.

8.3.4 Results - likelihood comparisons

Table 8.3.1 shows the minimized values of L_R (the negative residual Gaussian log-likelihood) for each model fitted to the data from each petri-dish. For each dish the model with the lowest value for an approximation to the Akaike Information Criterion (AIC) using L_R, instead of L_M, plus the number of model parameters, is underlined. AIC (Akaike, 1973) has already been used in sections 6.4 and 6.5. Model (1), the one considered in section 8.2, is preferred on this basis for 60% of dishes.

Models (2) to (6) are all generalizations of model (1) and so can be compared on the basis of the asymptotic properties of the likelihood ratio test. This uses the fact that the residual Gaussian likelihood converges to the true likelihood function as the number of seeds increases. Models (2), (3) and (4) each have one extra parameter and so are a significant improvement over model (1) at the 5% level if L_R has been decreased by more than 1.9, that is the 95 percentile of $\frac{1}{2}\ \chi^2_1$. Models (5) and (6) have two

Table 8.3.1

Naylor's data; minimum values of the negative residual Gaussian log-likelihood (L_R) for different models

		Independent errors	standard variances			scaled variances		
				+ error (a)	+ error (b)		+ error (a)	+ error (b)
	Model	(0)	(1)	(2)	(3)	(4)	(5)	(6)
Dish	Number of Parameters	4	3	4	4	4	5	5
1		5.6	5.4	4.7	5.4	4.8	4.1	4.8
2		8.1	7.6	6.2	5.0*	5.9	5.9	4.8
3		5.9	4.2	4.0	4.1	4.1	3.9	4.1
4		10.1	8.8	7.7	8.6	8.1	7.7	8.1
5		7.9	6.8	6.8	6.8	6.7	6.7	6.7
6		8.4	6.1	6.1	6.1	6.1	6.1	6.1
7		7.4	5.3	5.3	5.3	4.6	3.7	4.6
8		7.5	5.8	5.8	5.8	5.8	5.8	5.8
9		6.0	5.1	5.0	5.1	5.1	4.4	5.1
10		9.8	8.0	8.0	8.0	8.0	8.0	8.0
11		7.4	7.1	7.1	7.1	7.1	7.1	7.1
12		7.6	7.9	7.5	7.6	7.5	7.4	7.5
13		9.2	8.7	7.8	7.5	7.9	7.8	7.5
14		10.6	8.1	6.7	8.0	7.3	6.6	7.3
15		7.2	4.9	4.8	4.9	4.8	4.4	4.8
16		13.9	12.4	11.9	12.1	11.1	11.1	11.1
17		6.1	5.0	5.0	5.0	4.2	4.2	4.2
18		9.6	8.0	8.0	7.8	7.8	7.8	7.8
19		0.1	4.1	4.1	4.1	2.2*	2.2	0.2*
20		8.5	10.6	8.2*	8.4*	8.1*	8.0	8.1
21		8.5	9.2	8.9	7.8	8.1	8.1	7.8
22		4.2	4.4	4.4	4.4	3.4	3.4	3.4
23		8.4	5.8	5.7	5.8	5.5	1.9*	5.5
24		9.0	8.8	7.9	8.8	7.9	7.8	7.9
25		3.6	4.2	4.2	4.2	2.6	2.6	2.6
26		7.3	5.6	5.6	5.6	5.6	5.6	5.6
27		6.3	7.5	7.5	6.2	7.2	7.2	6.1
28		8.0	8.2	8.2	7.7	7.7	7.7	7.6
29		6.9	5.9	5.9	5.7	5.8	5.8	5.3
30		8.9	9.9	7.7*	9.3	7.8*	7.6	7.8
31		10.7	9.5	9.5	8.8	9.3	9.3	8.7
32		9.2	11.5	11.4	9.0*	10.0	10.0	8.4*
33		10.4	11.5	11.2	10.0	10.0	10.0	9.7
Average		7.8	7.3	6.9	6.8	6.6	6.4	6.4

For each dish the model with the lowest value of an approximation to Akaike's Information Criterion has been underlined. Models (2) to (6) with significantly lower L_R than model (1) at 5% level denoted by an asterisk.

extra parameters so a decrease in L_R of 3.0 (the 95 percentile of $\frac{1}{2}\chi^2_2$) is sufficient to reject model (1). Models which are significant improvements are marked with an asterisk. Only 11 of the 165 tests are significant which is close to the expected number of 8 arising by chance if model (1) is correct. However, because the 5 tests for each petri-dish are correlated with each other it is not possible to assess how substantial the difference is between the observed number of 11 and the expected number of 8. If model (1) is correct then the 33 differences between L_R in model (1) and model (2) should be independent deviates from $\frac{1}{2}\chi^2_1$. To assess this, the 33 differences were ranked and plotted against the corresponding quantiles of $\frac{1}{2}\chi^2_1$ in figure 8.3.1. Except for a preponderance of observations at zero the agreement with the predicted straight line is very good. Figure 8.3.2 shows a similar plot of the ranked differences between L_R in models (1) and (6) against the quantiles of $\frac{1}{2}\chi^2_2$ with even better agreement.

The values of L_R for each model averaged over all dishes are given at the foot of table 8.3.1. The significant decrease for a model with one extra parameter is 0.7 (the 95 percentile of $\frac{1}{66}\chi^2_{33}$), and with two extra parameters 1.3 (the 95 percentile of $\frac{1}{66}\chi^2_{66}$), once again appealing to the asymptotic properties of the likelihood ratio test. Model (1) is adequate as it is not rejected in favour of any of the other models on the basis of this test. The improvement in fit obtained by using a scaled version of the variance matrix (model (4)) is close to the 5% level and, because the mean estimate of τ is 1.2, this result suggests that the covariances are slightly larger than those predicted in the standard model for cumulative counts.

Figure 8.3.1

Naylor's data; ranked differences in minimum values of the negative residual Gaussian log-likelihood between models (1) and (2) plotted against quantiles of $\frac{1}{2}\chi^2_1$ distribution.

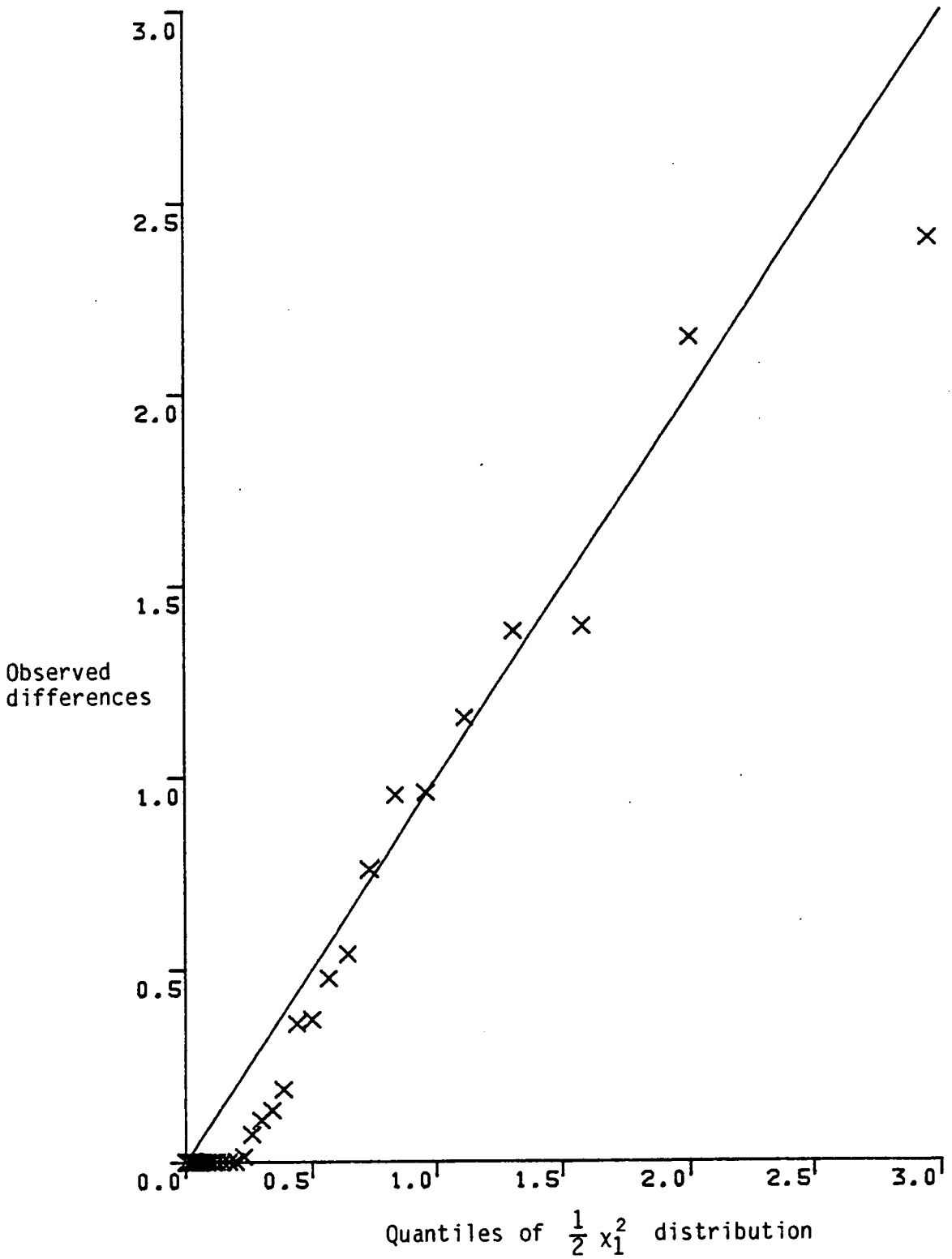
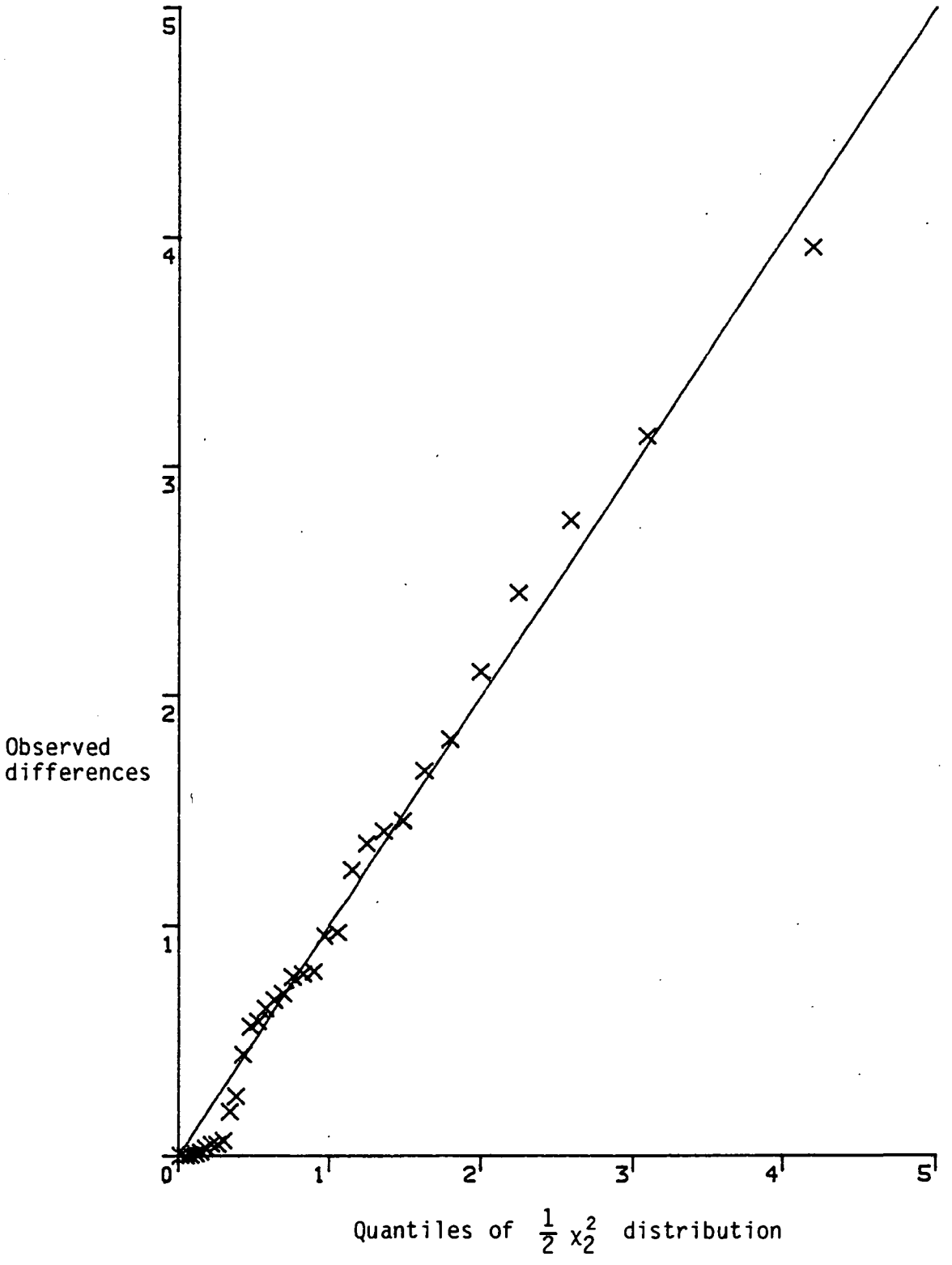


Figure 8.3.2

Naylor's data; ranked differences in minimum values of the negative residual Gaussian log-likelihood between models (1) and (6) plotted against quantiles of $\frac{1}{2}\chi^2_2$ distribution.



8.3.5 Results - parameter estimates

Table 8.3.2 shows the means and the standard deviations of the parameter estimates over all dishes for each model together with the average standard errors supplied by REGAME. To define what these summary statistics are (they were previously discussed and used in [6.4.5] and [6.5.5]) consider the parameter β_1 in a particular model; this was estimated as $\hat{\beta}_1^{\{j\}}$ with a standard error of $se(\hat{\beta}_1^{\{j\}})$ using the data from the jth petri-dish. The mean of the estimates is

$$\hat{\beta}_1 = \sum_{j=1}^{33} \hat{\beta}_1^{\{j\}} / 33 ,$$

the standard deviation of the estimates is

$$\sqrt{\{ \sum_{j=1}^{33} (\hat{\beta}_1^{\{j\}} - \hat{\beta}_1)^2 / 32\}} ,$$

and the average standard error is

$$\sum_{j=1}^{33} se(\hat{\beta}_1^{\{j\}}) / 33 .$$

Some of the differences between petri-dishes is being ignored in forming these statistics, but this simple approach is adequate for making some important comparisions between models and for assessing whether a parameter could take a common value for all dishes. The mean parameter estimates are in close agreement over all models, as are the standard deviations, but the standard errors are on average only half the size in model (0) that they are in the other models. The agreement among models (1) to (6) is not surprising since models (2) to (6) are generalizations of model (1) which were found in [8.3.4] not to be significant improvements. The estimated precision in model (0) is too great because the assumption that the

Table 8.3.2

Naylor's data; means and standard deviations of parameter estimates over all 33 dishes for each model, and average standard errors (see [8.3.5]).

Model		β_1	β_2	β_3	τ	$\sigma^{\{a\}}$	$\sigma^{\{b\}}$
				means of estimates			
(0)	indep. errors	3.24	0.39	0.085	0.6		
(1)	standard variances	3.25	0.39	0.082			
(2)	+ obs error(a)	3.25	0.39	0.084		0.7	
(3)	+ obs error(b)	3.24	0.37	0.086			0.5
(4)	scaled variances	3.25	0.38	0.084	1.2		
(5)	+ obs error(a)	3.25	0.39	0.084	0.9	0.6	
(6)	+ obs error(b)	3.24	0.37	0.087	1.0		0.3
				standard deviations of estimates			
(0)	indep. errors	0.19	0.20	0.046	0.2		
(1)	standard variances	0.19	0.18	0.045			
(2)	+ obs error(a)	0.19	0.18	0.045		0.7	
(3)	+ obs error(b)	0.19	0.18	0.044			0.6
(4)	scaled variances	0.19	0.18	0.044	0.3		
(5)	+ obs error(a)	0.19	0.17	0.045	0.4	0.8	
(6)	+ obs error(b)	0.19	0.19	0.044	0.4		0.5
				average standard errors			
(0)	indep. errors	0.03	0.04	0.014	0.2		
(1)	standard variances	0.07	0.07	0.029			
(2)	+ obs error(a)	0.07	0.08	0.029		1.3	
(3)	+ obs error(b)	0.07	0.07	0.031			1.1
(4)	scaled variances	0.07	0.08	0.033	0.3		
(5)	+ obs error(a)	0.07	0.07	0.027	0.4	1.6	
(6)	+ obs error(b)	0.07	0.07	0.031	0.4		1.0

observations are independent assumes that the data contain more information than they actually do. With the exception of model (0) the results agree very well with those of Hunter, Glasbey and Naylor (1984) which were estimated by full maximum likelihood and therefore took account of distributional properties of the observations and not simply first and second moments.

Figures 8.3.3 and 8.3.4 show the data for dish 2 together with the fitted curve and predicted fit using earlier observations (described in [2.3.6]) plotted against time for models (1) and (3) respectively. Although model (3) gave a significant improvement over model (1) for this particular dish, that being the reason for its selection for plotting, the differences are difficult to detect in the plots because there are so few data points. The figures are shown in the same form as in other chapters for completeness.

8.3.6 Resumé

For Naylor's data the standard variance model is adequate and is superior to the model in which errors are assumed to be independent. This confirms the assumptions made by Hunter, Glasbey and Naylor (1984). It has been demonstrated that REGAME is capable of fitting models which combine both types of error, although the full generality was found unnecessary for this data set.

8.4 Discussion

Data which are cumulative counts are serially correlated, but in a way that can be modelled very easily. So this is a rather

Figure 8.3.3

Naylor's data, dish number 2, model (1) fitted by residual Gaussian maximum likelihood estimation; observed germination counts (X), the fitted curve (----) and predicted values (——) conditional upon earlier observations, plotted against time.

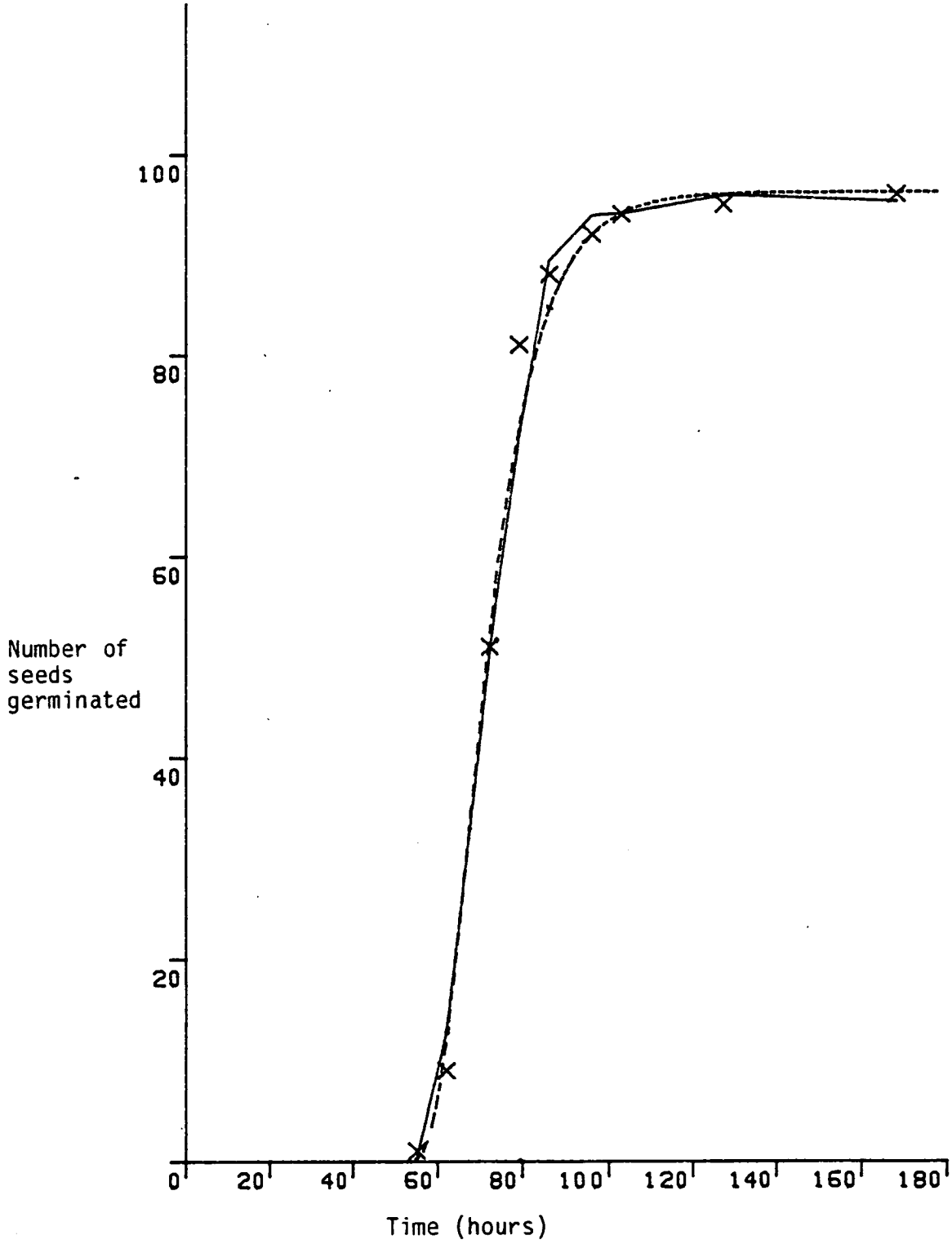
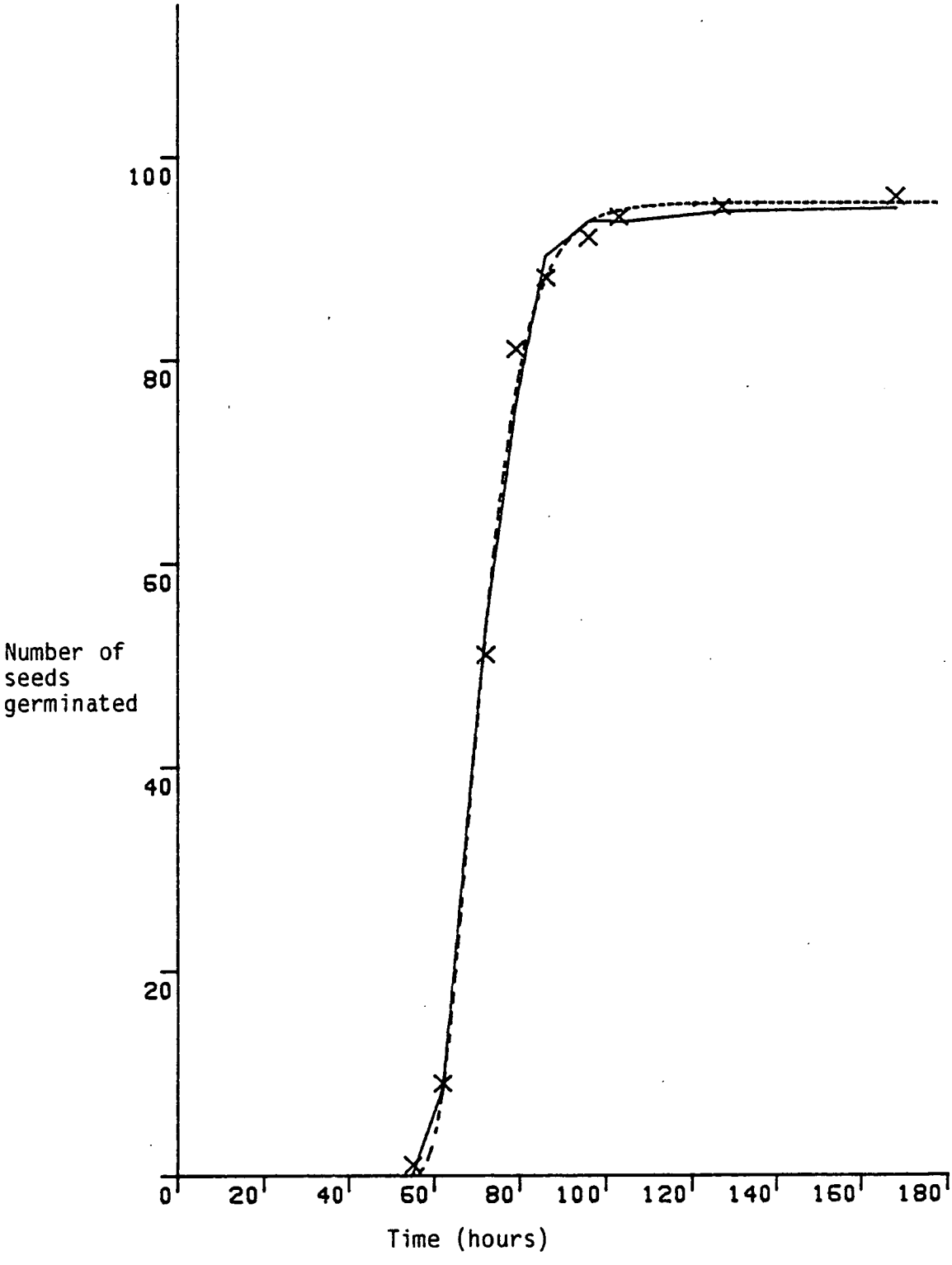


Figure 8.3.4

Naylor's data, dish number 2, model (3) fitted by residual Gaussian maximum likelihood estimation; observed germination counts (X), the fitted curve (----) and predicted values (____) conditional upon earlier observations, plotted against time.



trivial example of a mechanistic approach to specifying the error covariance structure. However, by setting this within the context of the preceding chapters in the thesis, we acquire the capability to include and test for other sources of error. In the particular illustration, using Naylor's data, these extensions proved to be unnecessary. The general modelling technique would be of particular use if cumulative counts were observed subject to unavoidable sampling errors.

Of the three chapters considering mechanistic modelling, the type considered in this chapter is the least suspect. For data such as Naylor's, there is no doubt that the mechanistic approach outperforms by far an empirical approach to modelling the error covariance structure.

9. Conservative estimates of the variances of regression parameter estimators

9.1 Introduction

This chapter is quite distinct from all the other chapters in the thesis, which were concerned with the estimation of regression parameters conditional upon the error variance matrix being known except for a few parameters. Here the emphasis is on obtaining valid estimates of the variances of regression parameter estimators without making gross assumptions about the error covariance structure. The comments of Efron (1982, p3) in introducing jackknifing and bootstrapping are equally valid here:

> "An important theme of what follows is the substitution of computational power for theoretical analysis. The payoff, of course, is freedom from the constraints of traditional parametric theory, with its overreliance on a small set of standard models for which theoretical solutions are available."

Previous chapters may be divided into those in which the choice of variance matrix is empirical (4 and 5) and those in which the variances are an integral part of the overall model (6, 7 and 8). In the empirical case, the choice of a variance matrix "close" to the true matrix will improve the efficiency of estimation of regression parameters relative to ordinary least-squares estimation. However, if this approximating variance matrix is assumed true, the estimated variances of the parameter estimators will be biased. The discussions in sections 4.6 and 5.4 highlight this problem. Essentially, there is a difficulty in estimating regression parameters from observations which are correlated in an unknown manner.

For a particular fitted regression it is desirable to estimate variances of the parameter estimators such that they will be valid for a wide range of possible error variance matrices. A general method of approaching the problem is presented in section 9.2. A computer program, CEVOPE, which can be used to implement this technique on real data sets, is described in section 9.3. In section 9.4 the method is illustrated by a small example. Then in sections 9.5, 9.6 and 9.7 the method is applied to Colquhoun's, Dale's and Bruce's data which were originally analysed in chapter 4. Finally, the usefulness of the technique is evaluated in section 9.8.

9.2 Method

9.2.1 Regression parameter estimators

It is assumed that the m-vector of unknown regression parameters, $\underset{\sim}{\beta}$, is estimated by generalized least-squares by minimizing the optimization function

$$\underset{\sim}{e}^T \underset{\sim}{R}^{-1} \underset{\sim}{e} ,$$

$$\text{where } \underset{\sim}{e} = \underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}) .$$

Here $\underset{\sim}{y}$ is the n-vector of observations, $\underset{\sim}{f}$ is the n-vector of regression values (which are functions of $\underset{\sim}{\beta}$), $\underset{\sim}{W}$ is a known n by n diagonal matrix of weights, $\underset{\sim}{e}$ is the n-vector of weight-corrected departures of the observations from the regression and $\underset{\sim}{R}$ is a known n by n positive-definite symmetric matrix. It is further assumed that $\underset{\sim}{f}$ is linear in $\underset{\sim}{\beta}$ and so can be expressed as

$$\underset{\sim}{f} = \underset{\sim}{W}\,\underset{\sim}{X}\,\underset{\sim}{\beta} ,$$

where $\underset{\sim}{X}$ is the n by n weight-corrected regression design matrix.

It is a standard result that the estimator of $\underset{\sim}{\beta}$, denoted $\hat{\underset{\sim}{\beta}}$, is given by

$$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-1}\ \underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{W}^{-1}\underset{\sim}{y}\ ,$$

$$= \underset{\sim}{U}^T\underset{\sim}{W}^{-1}\underset{\sim}{y}\ , \qquad (9.2.1)$$

if an n by m matrix $\underset{\sim}{U}$ is defined by

$$\underset{\sim}{U} = \underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-1}\ .$$

It is convenient at this stage to note that $\underset{\sim}{X}^T\underset{\sim}{U} = \underset{\sim}{I}$.

At this point some comments my help to clarify the situation. The use of a matrix of weights may appear rather arbitrary, but it is convenient in what follows to be able to handle heteroscedasticity separately from other features of the variance matrix of $\underset{\sim}{y}$. The choice of $\underset{\sim}{R}$ does not affect the mathematical development in this section and so, for the sake of generality, has been left deliberately vague. Finally, the assumption that the regression function is linear in $\underset{\sim}{\beta}$ is not too restrictive in practice because, provided $\underset{\sim}{f}$ is differentiable with respect to $\underset{\sim}{\beta}$, $\underset{\sim}{f}$ may be approximated by a linear function in the region of $\underset{\sim}{\beta}$ close to $\hat{\underset{\sim}{\beta}}$ with

$$\underset{\sim}{X} = \underset{\sim}{W}^{-1}\left.\frac{\partial \underset{\sim}{f}}{\partial \underset{\sim}{\beta}^T}\right|_{\underset{\sim}{\beta} = \hat{\underset{\sim}{\beta}}}\ .$$

Provided the approximation, which is precisely that used in non-linear regression to obtain standard errors of parameter estimators, is good what follows in this section still applies in the non-linear case.

If it is assumed that $E(\underset{\sim}{y}) = \underset{\sim}{f}$, then $\hat{\underset{\sim}{\beta}}$ is an unbiased estimator of $\underset{\sim}{\beta}$. Further, if $\underset{\sim}{V}$ is the n by n symmetric

positive-definite variance matrix of the weight-corrected observations, that is $\text{var}(\underset{\sim}{W}^{-1}\underset{\sim}{y}) = \underset{\sim}{V}$, then $\text{var}(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{U}^T\underset{\sim}{V}\,\underset{\sim}{U}$, and in particular

$$\text{var}(\hat{\beta}_s) = \underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u} \quad ,$$

where $\underset{\sim}{u}$ denotes the sth column of $\underset{\sim}{U}$. Strictly, the notation should include an indication of s in $\underset{\sim}{u}$, but this has been omitted in order to simplify what follows. All subsequent references in this section will be to a single parameter β_s and must be repeated for each value of s from 1 to m .

9.2.2 Form of estimator of $\text{var}(\hat{\beta}_s)$

The variance of $\hat{\beta}_s$ may be rewritten as

$$\text{tr}(\underset{\sim}{u}\,\underset{\sim}{u}^T\underset{\sim}{V})$$

where tr denotes the trace of a square matrix. In order to estimate this variance it is necessary to estimate $\underset{\sim}{V}$ using the only information available, the vector of weight-corrected departures, $\hat{\underset{\sim}{e}}$, where

$$\hat{\underset{\sim}{e}} = \underset{\sim}{W}^{-1}(\underset{\sim}{y} - \hat{\underset{\sim}{f}}) \quad , \tag{9.2.2}$$

and $\hat{\underset{\sim}{f}}$ is the estimated regression function. The ideal would be to find an estimator of $\underset{\sim}{V}$ which was unbiased for a wide range of possible variance matrices, say all matrices belonging to some set Ω . Unless very restrictive assumptions are made about Ω , such as

$$\Omega = \{\underset{\sim}{V} : \underset{\sim}{V} = \tau^2\underset{\sim}{I} \ , \ \tau^2 > 0\}$$

(using the set notation defined in the conventions section at the front of the thesis), this does not seem to be possible.

The next best alternative is for the estimator of $\text{var}(\hat{\beta}_s)$ to be conservative, that is upward biased (or more precisely, not downward biased) for all $\underset{\sim}{V}$ belonging to Ω because marginal

confidence intervals placed on β_s will then never have higher error probability levels than they should have.

I considered this problem for some time but failed to find an estimator of $\underset{\sim}{V}$ which ensured that $\text{var}(\hat{\beta}_s)$ was conservative for any reasonable set Ω. So instead, I decided to estimate $\underset{\sim}{V}$ simply using $\hat{\underset{\sim}{e}}\,\hat{\underset{\sim}{e}}^T$, but sought to substitute a symmetric n by n matrix denoted $\underset{\sim}{C}$ in place of $\underset{\sim\sim}{uu}^T$ in the estimator of $\text{var}(\hat{\beta}_s)$. The problem is then to find $\underset{\sim}{C}$ such that

$$E(\text{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{e}}\,\hat{\underset{\sim}{e}}^T)) \geq \text{tr}(\underset{\sim}{u}\,\underset{\sim}{u}^T\underset{\sim}{V}) \quad \text{for all} \quad \underset{\sim}{V} \in \Omega .$$

The estimator may be rewritten as

$$\hat{\underset{\sim}{e}}^T\,\underset{\sim}{C}\,\hat{\underset{\sim}{e}} ,$$

and it becomes apparent that it is a quadratic form of the weight-corrected departures $\hat{\underset{\sim}{e}}$. It is desirable for $\underset{\sim}{C}$ to be positive-semidefinite because then the variance estimator cannot take negative values. For the sake of generality Ω will remain unspecified at present, but will be discussed in detail in subsections [9.2.11] to [9.2.17].

9.2.3 Expectation of $\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\,\hat{\underset{\sim}{e}}$

It follows from equation (9.2.2) that

$$\begin{aligned}\hat{\underset{\sim}{e}} &= \underset{\sim}{W}^{-1}(\underset{\sim}{y}-\underset{\sim}{W}\,\underset{\sim}{X}\,\hat{\underset{\sim}{\beta}}) \quad , \\ &= (\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{U}^T)\underset{\sim}{W}^{-1}\underset{\sim}{y} \quad ,\end{aligned}$$

after substitution for $\hat{\underset{\sim}{\beta}}$ from equation (9.2.1). Therefore $\hat{\underset{\sim}{e}}$ is orthogonal to $\underset{\sim}{U}$ because

$$\underset{\sim}{U}^T\hat{\underset{\sim}{e}} = (\underset{\sim}{U}^T-\underset{\sim}{U}^T)\,\underset{\sim}{W}^{-1}\underset{\sim}{y} = \underset{\sim}{0} .$$

Also,

$$E(\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}) = \mathrm{tr}(\underset{\sim}{C}\, E(\hat{\underset{\sim}{e}}\, \hat{\underset{\sim}{e}}^T)) \quad ,$$

$$= \mathrm{tr}(\underset{\sim}{C}(\underset{\sim}{I}-\underset{\sim}{X}\,\underset{\sim}{U}^T)\underset{\sim}{V}(\underset{\sim}{I}-\underset{\sim}{U}\,\underset{\sim}{X}^T)) \ ,$$

$$= \mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}) \ ,$$

provided that $\underset{\sim}{C}\,\underset{\sim}{X} = \underset{\sim}{0}$.

Without loss of generality we can restrict our choice to $\underset{\sim}{C}$ orthogonal to $\underset{\sim}{X}$ because, as $\hat{\underset{\sim}{e}}$ is orthogonal to $\underset{\sim}{U}$, only the component of $\underset{\sim}{C}$ orthogonal to $\underset{\sim}{X}$ contributes to $\hat{\underset{\sim}{e}}^T \underset{\sim}{C}\, \hat{\underset{\sim}{e}}$. Discussion of other distributional properties of $\hat{\underset{\sim}{e}}^T \underset{\sim}{C}\, \hat{\underset{\sim}{e}}$ is deferred to [9.2.18].

9.2.4 Specification of relative bias

If a symmetric matrix $\underset{\sim}{C}$ can be found such that

$$\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}) \geq \underset{\sim}{u}^T \underset{\sim}{V}\,\underset{\sim}{u} \quad \text{for all } \underset{\sim}{V} \in \Omega \ ,$$

and $\underset{\sim}{C}$ is orthogonal to $\underset{\sim}{X}$, then $\hat{\underset{\sim}{e}}^T \underset{\sim}{C}\, \hat{\underset{\sim}{e}}$ can be used as a conservative estimator of $\mathrm{var}(\hat{\beta}_s)$. However, if a matrix $\underset{\sim}{C}$ exists that meets the specification, then it is not unique; for example $2\underset{\sim}{C}$ is also valid. Therefore, some means of choosing between competing matrices is needed.

A sensible approach would seem to be to choose the estimator with smallest bias. The relative upward bias of $\hat{\underset{\sim}{e}}^T \underset{\sim}{C}\, \hat{\underset{\sim}{e}}$ as an estimator of $\mathrm{var}(\hat{\beta}_s)$ is

$$\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V})/(\underset{\sim}{u}^T \underset{\sim}{V}\,\underset{\sim}{u}) - 1$$

when the true error variance matrix is $\underset{\sim}{V}$. There are various ways in which this bias can be specified for all $\underset{\sim}{V} \in \Omega$, such as

(i) $\underset{\underset{\sim}{V} \in \Omega}{\text{average}} \ (\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V})/(\underset{\sim}{u}^T \underset{\sim}{V}\,\underset{\sim}{u}) - 1)$

where the averaging is with respect to some well behaved measure on Ω ,

or (ii) $\max_{\underset{\sim}{V} \in \Omega} (\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V})/(\underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u}) - 1)$,

but if Ω is large, in some unspecified sense, then many unrealistic variance matrices will be included. Therefore, and also to simplify the problem, the bias I have chosen to minimize is that for a specific choice of $\underset{\sim}{V}$, denoted $\hat{\underset{\sim}{V}}$. The way in which $\hat{\underset{\sim}{V}}$ is estimated or otherwise chosen is irrelevant to what follows in this section. It will be considered in subsequent sections when real data are encountered.

The search for an estimator of $\mathrm{var}(\hat{\beta}_s)$ can now be expressed formally as an optimization problem:

minimize $\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}})$ with respect to $\underset{\sim}{C}$,

subject to constraints $\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}) \geq \underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u}$ for all $\underset{\sim}{V} \in \Omega$,

for $\underset{\sim}{C}$ symmetric and $\underset{\sim}{C}\,\underset{\sim}{X} = \underset{\sim}{0}$.

To illustrate what the constraints mean, consider the case when

$\Omega = \{\underset{\sim}{V} : V_{ij} > 0 \text{ if } |i-j| < \upsilon \text{ and } V_{ij} = 0 \text{ if } |i-j| > \upsilon\}$

for some non-negative integer υ . The constraints are then equivalent to

$$C_{ij} \geq u_i u_j \quad \text{for } |i-j| < \upsilon .$$

9.2.5 Relation to ordinary least-squares

It may be enlightening at this stage to relate the above remarks to ordinary least-squares when $\underset{\sim}{V} = \tau^2\underset{\sim}{I}$. For this subsection alone let

$$\underset{\sim}{R} = \underset{\sim}{W} = \underset{\sim}{I} ,$$

and $\Omega = \{\tau^2\underset{\sim}{I} : \tau^2 \geq 0\}$,

then $\underset{\sim}{U} = \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}$,

and if $$\underset{\sim}{C} = (\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{U}^T)\,\underset{\sim}{u}^T\underset{\sim}{u}/(n-m) \ ,$$

the constraints in the optimization problem are exactly satisfied because

$$\begin{aligned} \mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}) &= \mathrm{tr}(\tau^2\,\underset{\sim}{C}) \ , \\ &= \tau^2\,\mathrm{tr}(\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{U}^T)\,\underset{\sim}{u}^T\underset{\sim}{u}/(n-m) \ , \\ &= \tau^2\,\underset{\sim}{u}^T\underset{\sim}{u} \ , \\ &= \underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u} \qquad \text{for all } \underset{\sim}{V} \in \Omega \ , \end{aligned}$$

and $$\underset{\sim}{C}\,\underset{\sim}{X} = \underset{\sim}{0} \ .$$

Therefore, $\underset{\sim}{C}$ minimizes $\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}})$ provided that $\hat{\underset{\sim}{V}} \in \Omega$, and the estimator of the variance of $\hat{\beta}_s$ is

$$\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\,\hat{\underset{\sim}{e}} = \hat{\tau}^2\,\underset{\sim}{u}^T\underset{\sim}{u} \ ,$$

where

$$\hat{\tau}^2 = \hat{\underset{\sim}{e}}^T\hat{\underset{\sim}{e}}/(n-m) \ .$$

Therefore, the conservative estimator of variance chosen by the optimization problem specified in [9.2.4] with the restrictive assumption about Ω, is the same as the standard estimator.

9.2.6 Optimization problem

I now return to the general theory. The optimization problem in [9.2.4] has $n(n+1)/2$ variables corresponding to the elements in $\underset{\sim}{C}$, nm constraints arising from the requirement that $\underset{\sim}{C}$ be orthogonal to $\underset{\sim}{X}$, and a potentially large number of other constraints dependent on the choice of Ω. Therefore this is a very large problem. For example, if n is equal to 200 and m is equal to 4, then there are 20100 variables and at least 800 constraints. Also, the solution of the problem may not result in a choice of $\underset{\sim}{C}$ which is positive-semidefinite. Therefore $\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\,\hat{\underset{\sim}{e}}$

may take negative values with non-zero probability, which is undesirable for a variance estimator.

Any symmetric matrix $\underset{\sim}{C}$ may be decomposed as

$$\underset{\sim}{C} = \sum_{k=1}^{t} x_k \, \underset{\sim}{c}^{\{k\}} \, \underset{\sim}{c}^{\{k\}T} \quad ,$$

for some value of t not exceeding n, with coefficients $x_1,\ldots,x_t$, denoted $\underset{\sim}{x}$, and vectors $\underset{\sim}{c}^{\{1\}},\ldots,\underset{\sim}{c}^{\{t\}}$, for example by using the eigenvalues and eigenvectors of $\underset{\sim}{C}$. Constraining all the coefficients x_k to be non-negative ensures that $\underset{\sim}{C}$ is positive-semidefinite. The decomposition also enables the optimization problem to be reformulated in a computationally more tractable form.

The optimization problem may be restated as:

$$\text{minimize} \quad \sum_{k=1}^{t} x_k \, (\underset{\sim}{c}^{\{k\}T} \, \hat{\underset{\sim}{V}} \, \underset{\sim}{c}^{\{k\}}) \quad ,$$

with respect to x_k and $\underset{\sim}{c}^{\{k\}}$ for $k = 1,\ldots,t$,

$$\text{subject to} \quad \sum_{k=1}^{t} x_k \, (\underset{\sim}{c}^{\{k\}T} \, \underset{\sim}{V} \, \underset{\sim}{c}^{\{k\}}) \geq \underset{\sim}{u}^T \underset{\sim}{V} \, \underset{\sim}{u} \quad \text{for all} \quad \underset{\sim}{V} \in \Omega \quad ,$$

$$\text{for} \qquad x_k \geq 0 \quad \text{and} \quad \underset{\sim}{c}^{\{k\}T} \underset{\sim}{X} = \underset{\sim}{0} \qquad k = 1,\ldots,t \quad .$$

The constraints on $\underset{\sim}{c}^{\{k\}}$ are sufficient, but not necessary, to ensure that $\underset{\sim}{C}$ is orthogonal to $\underset{\sim}{X}$. This problem is linear in $\underset{\sim}{x}$ and quadratic in $\underset{\sim}{c}^{\{k\}}$ for $k = 1,\ldots,t$.

9.2.7 Dual optimization problem

Provided Ω is a convex set, the optimization problem with respect to $\underset{\sim}{x}$ with $\underset{\sim}{c}^{\{1\}},\ldots,\underset{\sim}{c}^{\{t\}}$ held fixed (orthogonal to $\underset{\sim}{X}$) has a dual:

$$\text{maximize} \quad \underset{\sim}{u}^T \underset{\sim}{V} \, \underset{\sim}{u} \quad \text{with respect to} \quad \underset{\sim}{V} \quad ,$$

subject to $\underset{\sim}{c}^{\{k\}T}\underset{\sim}{V}\,\underset{\sim}{c}^{\{k\}} \leq \underset{\sim}{c}^{\{k\}T}\hat{\underset{\sim}{V}}\,\underset{\sim}{c}^{\{k\}}$ for $k = 1,\ldots,t$,

for $\underset{\sim}{V} \in \Omega$.

The solution to this problem is also the solution to the primal problem in [9.2.6]: the maximum and minimum are equal and the Lagrangian multipliers are the optimal coefficients in $\underset{\sim}{x}$ in the primal problem. The optimal $\underset{\sim}{V}$ is the matrix within Ω for which the relative bias of $\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\,\hat{\underset{\sim}{e}}$ as an estimator of $\mathrm{var}(\hat{\beta}_s)$ is most severe.

For a particular set of vectors, $\underset{\sim}{c}^{\{1\}},\ldots,\underset{\sim}{c}^{\{t\}}$, the linear optimization problem in [9.2.6] can be solved to find the coefficients $\underset{\sim}{x}$ to associate with these vectors in order to obtain the estimator of $\mathrm{var}(\hat{\beta}_s)$ with smallest relative bias if the error variance matrix is $\hat{\underset{\sim}{V}}$. In order to improve upon this solution, in other words to find an estimator with an even smaller relative bias, it is necessary to change the set of c-vectors. Examination of the dual optimization problem suggests a way in which this can be done, and this will be considered in the following subsections.

9.2.8 Selection of $\underset{\sim}{c}^{\{t+1\}}$

Denote the optimal choice $\underset{\sim}{V}$ in the dual problem by $\overset{\prime\prime}{\underset{\sim}{V}}$. If a vector $\underset{\sim}{c}$ can be found such that

$$\underset{\sim}{c}^T\,\overset{\prime\prime}{\underset{\sim}{V}}\,\underset{\sim}{c} > \underset{\sim}{c}^T\,\hat{\underset{\sim}{V}}\,\underset{\sim}{c} \quad ,$$

and $\underset{\sim}{c}$ is added to the constraints as $\underset{\sim}{c}^{\{t+1\}}$, then $\overset{\prime\prime}{\underset{\sim}{V}}$ no longer satisfies all the constraints and so becomes infeasible. Therefore, the maximum in the new problem, and hence the bias in the estimator of $\mathrm{var}(\hat{\beta}_s)$, will be reduced.

To ensure that $\underset{\sim}{c}$ is orthogonal to $\underset{\sim}{X}$, it can be found as an eigenvector of

$$(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)(\hat{\underset{\sim}{V}} - \overset{''}{\underset{\sim}{V}})(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T) \tag{9.2.3}$$

with a negative eigenvalue. Because this last matrix is real and symmetric all eigenvalues and eigenvectors are real. The columns of $\underset{\sim}{X}$ are eigenvectors with zero eigenvalues, so provided the eigenvalue corresponding to $\underset{\sim}{c}$ is non-zero, $\underset{\sim}{c}$ will be orthogonal to $\underset{\sim}{X}$. Also, although $\underset{\sim}{c}$ will not in general be an eigenvector of $(\hat{\underset{\sim}{V}}-\overset{''}{\underset{\sim}{V}})$, $\underset{\sim}{c}^T(\hat{\underset{\sim}{V}}-\overset{''}{\underset{\sim}{V}})\underset{\sim}{c}$ will be negative. (As an aside, it may be noted that this would not have been the case if $\underset{\sim}{c}$ had been found as an eigenvector of

$$(\underset{\sim}{I} - \underset{\sim}{U}\,\underset{\sim}{X}^T)(\hat{\underset{\sim}{V}}-\overset{''}{\underset{\sim}{V}})(\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{U}^T) \quad .$$

This transformation would have ensured that $\underset{\sim}{c}$ was orthogonal to $\underset{\sim}{X}$, but because in general

$$(\underset{\sim}{I} - \underset{\sim}{X}\,\underset{\sim}{U}^T)\underset{\sim}{c} \neq \underset{\sim}{c} \ ,$$

$\underset{\sim}{c}^T(\hat{\underset{\sim}{V}}-\overset{''}{\underset{\sim}{V}})c$ would not necessarily have been negative.)

If all the eigenvalues of equation (9.2.3) are non-negative, then the existing solution to the dual optimization problem cannot be made infeasible. Therefore, the positive-semidefinite matrix $\underset{\sim}{C}$ has been found which gives a conservative estimator of $\text{var}(\hat{\beta}_s)$ with smallest possible relative bias if the error variance matrix is $\hat{\underset{\sim}{V}}$.

9.2.9 Interpretation of $\underset{\sim}{c}^{\{t+1\}}$

The constraint $\underset{\sim}{c}^T\,\underset{\sim}{V}\,\underset{\sim}{c} \leq \underset{\sim}{c}^T\,\hat{\underset{\sim}{V}}\,\underset{\sim}{c}$ is linear in elements of $\underset{\sim}{V}$, so every vector $\underset{\sim}{c}$ orthogonal to $\underset{\sim}{X}$ defines a constraining hyperplane in the space of $\underset{\sim}{V}$. Figure 9.2.1 gives a crude

Figure 9.2.1

An illustration in the space V_{11}, V_{22} of the optimal value of $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ subject to the boundary imposed by the global set of constraints $\underset{\sim}{c}^T \underset{\sim}{V} \underset{\sim}{c} \leq \underset{\sim}{c}^T \hat{\underset{\sim}{V}} \underset{\sim}{c}$ for all vectors $\underset{\sim}{c}$ orthogonal to $\underset{\sim}{X}$.

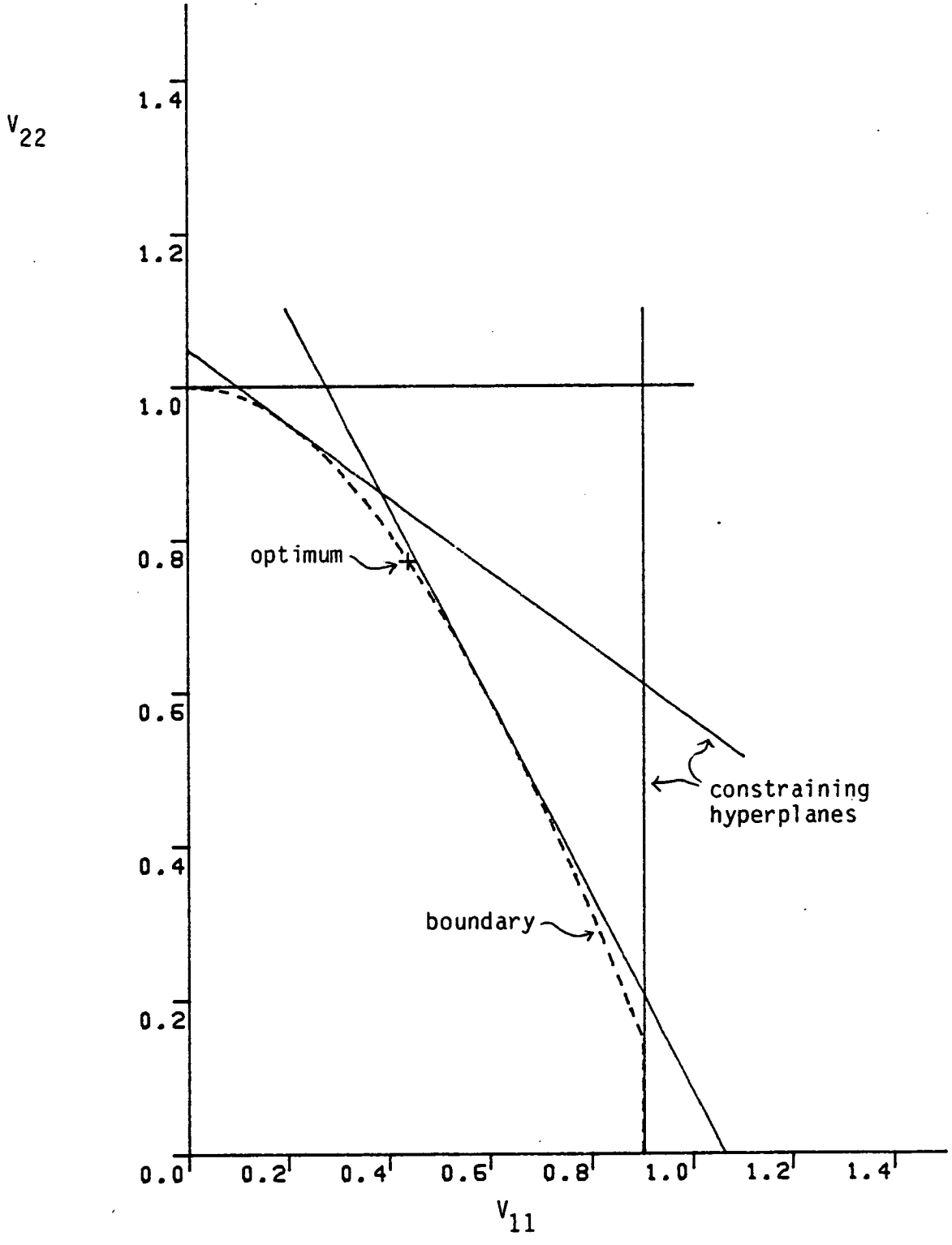
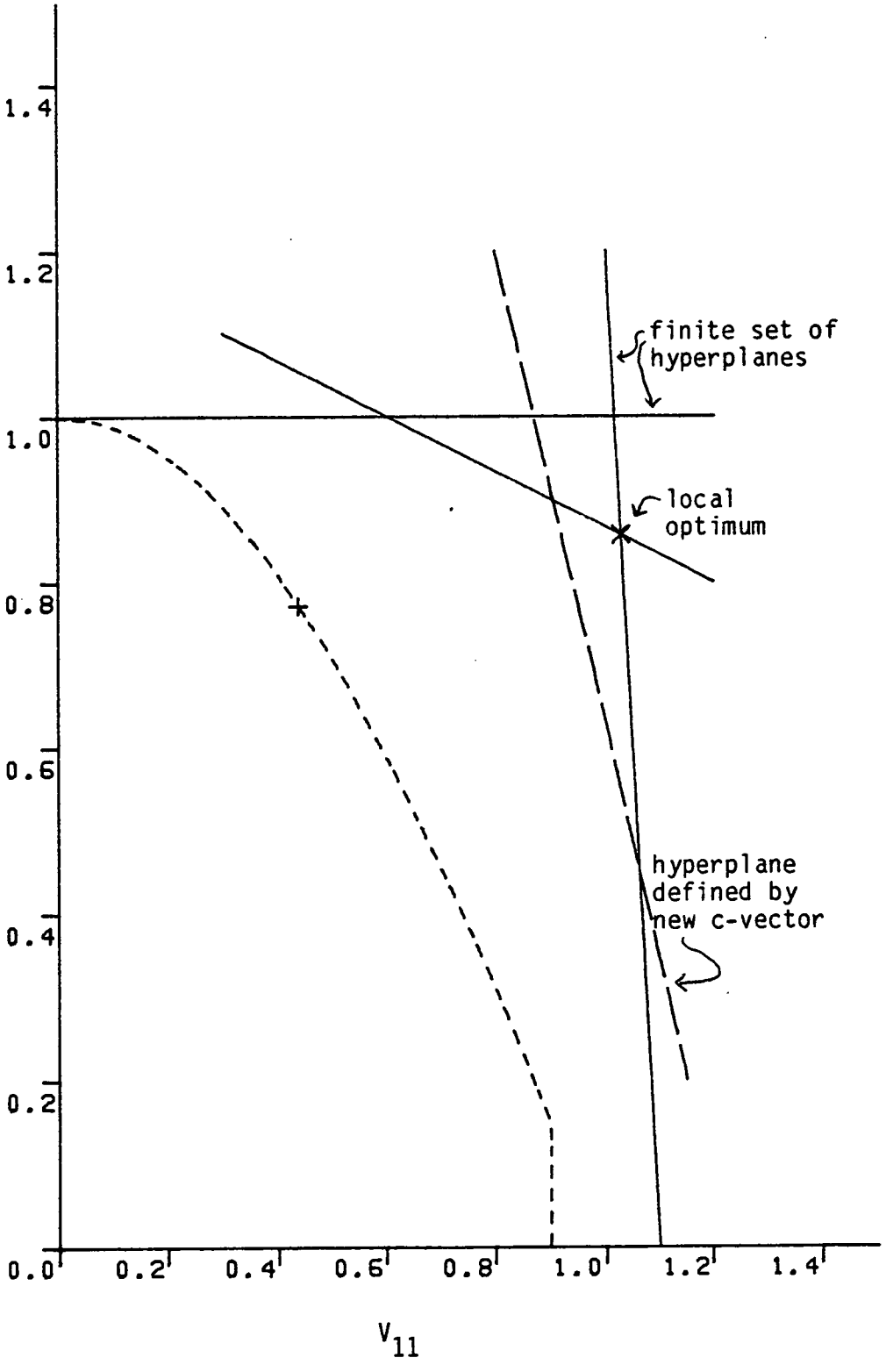


Figure 9.2.2

An illustration in the space V_{11}, V_{22} of the optimal value of $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ subject to a finite set of constraints of the form $\underset{\sim}{c}^T \underset{\sim}{V} \underset{\sim}{c} \leq \underset{\sim}{c}^T \hat{\underset{\sim}{V}} \underset{\sim}{c}$, and the selection of a new c-vector so that the optimum value is no longer feasible.



diagrammatic representation, simply in terms of two elements in $\underset{\sim}{V}$. There is a global boundary defined by all possible c-vectors. The point on this boundary at which $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ is maximized corresponds to the best possible positive-semidefinite matrix $\underset{\sim}{C}$. A finite set of constraints corresponding to $\underset{\sim}{c}^{\{1\}},\ldots, \underset{\sim}{c}^{\{t\}}$ is less restrictive and the value $\underset{\sim}{\overset{''}{V}}$ maximizing $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ will, in general, lie outside the region defined by all $\underset{\sim}{c}$ (see figure 9.2.2). The new vector $\underset{\sim}{c}^{\{t+1\}}$ is so chosen that $\underset{\sim}{\overset{''}{V}}$ is outside the boundary of the new constrained region. The vector $\underset{\sim}{c}$ which would maximize the separation of $\underset{\sim}{\overset{''}{V}}$ from the new boundary would be the one that lay on the global constraining surface in the direction of $\underset{\sim}{\overset{''}{V}}$, that is the vector $\underset{\sim}{c}$ will maximize

$$\frac{\underset{\sim}{c}^T \underset{\sim}{\overset{''}{V}} \underset{\sim}{c}}{\underset{\sim}{c}^T \underset{\sim}{\hat{V}} \underset{\sim}{c}} \quad .$$

This can be found, after incorporation of the constraint $\underset{\sim}{c}^T \underset{\sim}{X} = \underset{\sim}{0}$, as the eigenvector with largest negative generalized eigenvalue (μ) which satisfies

$$(\underset{\sim}{I}-\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{\hat{V}}(\underset{\sim}{I}-\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\ \underset{\sim}{c} = \mu(\underset{\sim}{I}-\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{\overset{''}{V}}(\underset{\sim}{I}-\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\ \underset{\sim}{c}$$

9.2.10 Algorithm for obtaining $\underset{\sim}{c}^{\{t+1\}}$

A simple algorithm exists for finding the eigenvector of a matrix with the eigenvalue of greatest absolute magnitude, originally due to Aitken (1937) and refined by Wilkinson (1954). This can be modified to enable $\underset{\sim}{c}$ to be found as specified in [9.2.8]. Starting with any initial vector $\underset{\sim}{w}^{\{0\}}$, we may transform this to

$$\underset{\sim}{w}^{\{1\}} = (\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)(\underset{\sim}{\hat{V}} - \underset{\sim}{\overset{''}{V}})(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\ \underset{\sim}{w}^{\{0\}}$$

and standardize the resultant vector to have a largest element of

unity. After repeated transformations, $\underset{\sim}{w}^{\{i\}}$ converges to the eigenvector with eigenvalue of largest absolute magnitude (say μ) provided this is unique. If the eigenvalue of largest absolute magnitude happens to be positive then a new transformation can be applied:

$$\underset{\sim}{w}^{\{i+1\}} = (\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)(\hat{\underset{\sim}{V}} - \underset{\sim}{\overset{''}{V}})(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T) - \mu\underset{\sim}{I})\ \underset{\sim}{w}^{\{i\}} ,$$

which after repeated application ensures $\underset{\sim}{c}$ converges to the eigenvector with the largest negative eigenvalue. If the largest negative eigenvalue is not unique the algorithm will not converge. However, in the course of iteration $\underset{\sim}{w}^{\{i\}}$ may be found such that

$$\underset{\sim}{w}^{\{i\}T}\ \underset{\sim}{\overset{''}{V}}\ \underset{\sim}{w}^{\{i\}} > \underset{\sim}{w}^{\{i\}T}\ \hat{\underset{\sim}{V}}\ \underset{\sim}{w}^{\{i\}}$$

and this is all that is required since every $\underset{\sim}{w}^{\{i\}}$ is automatically orthogonal to $\underset{\sim}{X}$.

The alternative approach for finding $\underset{\sim}{c}^{\{t+1\}}$ as the solution of the generalized eigenvalue problem, which was discussed in [9.2.9], has not been pursued. No algorithm equivalent to the above exists and more general algorithms for finding all the eigenvectors proved to be prohibitively costly in computer time.

9.2.11 The choice of Ω

At this stage the choice of Ω needs to be considered. The first point that should be made is that Ω cannot be made so large as to contain all positive-definite symmetric matrices. If it did then $\tau^2\underset{\sim}{I} + \sigma^2\ \underset{\sim}{X}\ \underset{\sim}{X}^T$ would belong to Ω provided that τ^2 and σ^2 were positive. Applying the results of [9.2.1], for variance matrices of this form,

$$\begin{aligned} \text{var}(\hat{\underset{\sim}{\beta}}) &= \underset{\sim}{U}^T(\tau^2\underset{\sim}{I} + \sigma^2\ \underset{\sim}{X}\ \underset{\sim}{X}^T)\underset{\sim}{U} \ , \\ &= \tau^2\underset{\sim}{U}^T\ \underset{\sim}{U} + \sigma^2\underset{\sim}{I} \ , \end{aligned}$$

because $\underset{\sim}{X}^T \underset{\sim}{U} = \underset{\sim}{I}$, and from [9.2.3]

$$E(\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}) = tr(\underset{\sim}{C}(\tau^2 \underset{\sim}{I} + \sigma^2 \underset{\sim}{X} \underset{\sim}{X}^T)) ,$$

$$= tr(\tau^2 \underset{\sim}{C}) ,$$

because $\underset{\sim}{C}$ is orthogonal to $\underset{\sim}{X}$. From the above we notice that the variance of $\hat{\underset{\sim}{\beta}}$ depends on σ^2 , but the quadratic form of departures is independent of σ^2 . Therefore, no matrix $\underset{\sim}{C}$ exists such that

$$E(\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}) \geq var(\hat{\beta}_s) \qquad \text{for all } \sigma^2 > 0 .$$

One explanation of this problem is that when the variance matrix is of the above form σ^2 is confounded with the regression parameters. The situation will be encountered again in [9.4.9].

In order that the upward bias in the variance estimator of $\hat{\beta}_s$ should be small, the set Ω should be as small as possible consistent with containing all "reasonable" error variance matrices and leading to a mathematically tractable optimization problem. The dual problem in [9.2.7] has $n(n+1)/2$ variables (corresponding to the elements in the symmetric matrix $\underset{\sim}{V}$) and t constraints, and so is already very big. Two particular limitations can be placed on Ω without making the optimization problem any larger by adding more constraints. These can be represented as:

$$\underset{\sim}{V} \in \Omega(\omega,\upsilon) \text{ if and only if}$$

(i) $\underset{\sim}{V}$ is symmetric

(ii) $(\Delta^{\omega} \underset{\sim}{V})_{ij} \geq 0 \qquad \text{for } i \geq j$

and (iii) $V_{ij} = 0 \qquad \text{if } i > j+\upsilon ,$

where ω and υ are non-negative integers, Δ is the matrix operator

$$(\Delta \underset{\sim}{V})_{ij} = V_{ij} + V_{i+1,j-1} - V_{i+1,j} - V_{i,j-1} , \quad \text{for } i \geq j ,$$

with $V_{n+1,j} = V_{i,0} = 0 ,$

and $(\Delta^0 \underset{\sim}{V}) = \underset{\sim}{V} .$

First of all I will demonstrate how the choice of Ω can be incorporated into the optimization problem when ω is equal to 0, 1 or 2. Then I will go on to discuss how Ω can be interpreted and what other choices are possible.

9.2.12 Dual Problem for $\Omega(0,\upsilon)$

If $\underset{\sim}{V} \in \Omega(0,\upsilon)$, then $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ can be re-expressed as

$$\sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} V_{ij}(2u_i u_j/(1 + I_{ij})) ,$$

where I_{ij} is an element in the identity matrix, $\underset{\sim}{I}$, of size n. Therefore, the dual optimization problem of [9.2.7] can be restated as:

$$\text{maximize} \quad \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} V_{ij}(2u_i u_j/(1+I_{ij}))$$

with respect to coefficients V_{ij} ,

$$\text{subject to} \quad \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} V_{ij}(2c_i^{\{k\}} c_j^{\{k\}}/(1+I_{ij})) \leq \underset{\sim}{c}^{\{k\}T} \hat{\underset{\sim}{V}} \underset{\sim}{c}^{\{k\}}$$

$$\text{for} \quad k = 1,\ldots,t ,$$

$$\text{for} \quad V_{ij} \geq 0 \qquad j = 1,\ldots,n,\ i = j,\ldots,\min(j+\upsilon,n) .$$

This is a linear programming problem. The number of variables V_{ij} have been reduced to $(2n-\upsilon)\cdot(\upsilon+1)/2$, and no constraints have been added because the boundary constraints $V_{ij} \geq 0$ are incorporated into the standard linear program.

9.2.13 Dual problem for $\Omega(1,\upsilon)$

Lower-triangular elements in an $n \times n$ matrix $\underset{\sim}{\Psi}$ are specified by

$$\Psi_{ij} = (\Delta \underset{\sim}{V})_{ij} \qquad \text{for} \quad i > j .$$

If $\underset{\sim}{V} \in \Omega(1,\upsilon)$, then

$$\Psi_{ij} = 0 \qquad \text{for } i > j+\upsilon ,$$

and

$$V_{gh} = \sum_{i=g}^{n} \sum_{j=1}^{h} \Psi_{ij} \qquad \text{for } g > h$$

because

$$\sum_{i=g}^{n} \sum_{j=1}^{h} \Psi_{ij} = \sum_{i=g}^{n} \sum_{j=1}^{h} V_{ij} + \sum_{i=g+1}^{n} \sum_{j=1}^{h-1} V_{ij}$$

$$- \sum_{i=g+1}^{n} \sum_{j=1}^{h} V_{ij} - \sum_{i=g}^{n} \sum_{j=1}^{h-1} V_{ij}$$

$$= \sum_{j=1}^{h} V_{gj} - \sum_{j=1}^{h-1} V_{gj}$$

$$= V_{gh} .$$

Therefore

$$\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u} = \sum_{h=1}^{n} \sum_{g=h}^{n} 2u_g u_h \left(\sum_{j=1}^{h} \sum_{i=g}^{n} \Psi_{ij} \right)/(1+I_{gh}),$$

$$= \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} \Psi_{ij} \left(\sum_{h=j}^{i} \sum_{g=h}^{i} 2u_g u_h /(1+I_{gh}) \right) .$$

It is convenient at this point to note that $\underset{\sim}{u}^T \underset{\sim}{V} \underset{\sim}{u}$ is always non-negative because each coefficient Ψ_{ij} is non-negative by definition and it is multiplied by a term which can be re-expressed as a sum of squares:

$$\left(\sum_{h=j}^{i} u_h \right)^2 ,$$

and so is also non-negative. Therefore, all matrices in $\Omega(1,\upsilon)$ are positive-semidefinite.

The optimization problem of [9.2.7] can be restated as:

$$\text{maximize} \quad \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} \Psi_{ij} \left(\sum_{h=j}^{i} \sum_{g=h}^{i} 2u_g u_h/(1+I_{gh}) \right)$$

with respect to coefficients Ψ_{ij} ,

$$\text{subject to } \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} \Psi_{ij} \left(\sum_{h=j}^{i} \sum_{g=h}^{i} 2c_i^{\{k\}} c_j^{\{k\}} / (1+I_{ij}) \right) \leq \underset{\sim}{c}^{\{k\}T} \hat{\underset{\sim}{V}} \underset{\sim}{c}^{\{k\}}$$

$$\text{for } k = 1,\ldots,t ,$$

$$\text{for } \Psi_{ij} > 0 \qquad j = 1,\ldots,n,\ i = j,\ldots,\min(j+\upsilon,n) .$$

Once again this is a linear programming problem, and it has the same numbers of variables and constraints as the problem in [9.2.12].

9.2.14 Dual problem for $\Omega(2,\upsilon)$

Lower-triangular elements in an $n \times n$ matrix $\underset{\sim}{T}$ are specified by

$$T_{ij} = (\Delta^2 \underset{\sim}{V})_{ij} \qquad \text{for } i > j .$$

If $\underset{\sim}{V} \in \Omega(2,\upsilon)$, then as in [9.2.13],

$$T_{ij} = 0 \qquad \text{for } i > j+\upsilon ,$$

$$\text{and} \qquad \Psi_{gh} = \sum_{i=g}^{n} \sum_{j=1}^{h} T_{ij} \qquad \text{for } g > h .$$

The optimization problem of [9.2.7] can be restated as:

$$\text{maximize} \quad \sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} T_{ij} \left(\sum_{h=j}^{i} \sum_{g=h}^{i} 2u_g u_h (i+1-g)(h+1-j)/(1+I_{gh}) \right)$$

with respect to coefficients T_{ij} , subject to

$$\sum_{j=1}^{n} \sum_{i=j}^{\min(j+\upsilon,n)} T_{ij} \left(\sum_{h=j}^{i} \sum_{g=h}^{i} 2c_g^{\{k\}} c_h^{\{k\}} (i+1-g)(h+1-j)/(1+I_{gh}) \right)$$

$$\leq \underset{\sim}{c}^{\{k\}T} \hat{\underset{\sim}{V}} \underset{\sim}{c}^{\{k\}} \qquad \text{for } k = 1,\ldots,t ,$$

$$\text{for } T_{ij} > 0 \qquad j = 1,\ldots,n,\ i = j,\ldots,\min(j+\upsilon,n) .$$

Once again this is a linear programming problem with the same number of variables and constraints as the problem in [9.2.12].

9.2.15 Interpretation of $\Omega(\omega,n-1)$ for a Markov process

If $\underset{\sim}{V}$ is a symmetric matrix with a Markov structure, that is

$$V_{ij} = \begin{cases} 1 & \text{when } i=j \\ \prod_{\ell=j}^{i-1} \rho_\ell & \text{when } i>j \end{cases}$$

for a set of coefficients $\rho_1, \ldots, \rho_n$, then:

(i) $\underset{\sim}{V} \in \Omega(0,n-1)$ if and only if

$$\rho_\ell > 0 \qquad \text{for } \ell=1,\ldots,(n-1);$$

(ii) $\underset{\sim}{V} \in \Omega(1,n-1)$ if in addition

$$\Psi_{ij} = \left(\prod_{\ell=j}^{i-1} \rho_\ell\right)(1-\rho_i)(1-\rho_{j-1}) > 0 \qquad \text{for } i \geq j,$$

where $\rho_0 = 0$, so in this instance

$$0 < \rho_\ell < 1 \qquad \text{for } \ell=1,\ldots,(n-1);$$

(iii) $\underset{\sim}{V} \in \Omega(2,n-1)$ if in addition

$$T_{ij} = \left(\prod_{\ell=j}^{i-1} \rho_\ell\right)(1-2\rho_i+\rho_i\rho_{i+1})(1-2\rho_{j-1}+\rho_{j-1}\rho_{j-2}) > 0 \qquad \text{for } i \geq j,$$

where $\rho_{-1} = \rho_0 = 0$ and $\rho_n = 1$, so in this instance

$$0 < \rho_\ell < 1,$$

$$\rho_\ell > (2\rho_{\ell-1}-1)/\rho_{\ell-1}$$

and $\rho_\ell < 1/(2-\rho_{\ell-1})$ for $\ell=1,\ldots,n$, with $\rho_0=\rho_n=0$.

9.2.16 Discussion about $\Omega(\omega,\upsilon)$

The set $\Omega(\omega,\upsilon)$ has the desirable properties that if

$$\underset{\sim}{V}^{\{1\}}, \underset{\sim}{V}^{\{2\}} \in \Omega(\omega,\upsilon),$$

then $$V^{\{1\}} + V^{\{2\}} \in \Omega(\omega,\upsilon),$$

and $$k\,\underset{\sim}{V}^{\{1\}} \in \Omega(\omega,\upsilon) \qquad \text{provided } k > 0.$$

Therefore, in particular, sums of Markov processes satisfying the restrictions of [9.2.15] also belong in $\Omega(\omega,n-1)$.

The choice of ω equal to 1 appears to have much to commend it. It ensures that $\underset{\sim}{V}$ is positive-semidefinite and that its elements decrease monotonically away from the diagonal in any direction. It is also necessary that the rate of decay should decrease, so the matrix

$$\begin{bmatrix} 6 & 4 & 1 & 0 & 0 \\ 4 & 6 & 4 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \\ 0 & 1 & 4 & 6 & 4 \\ 0 & 0 & 1 & 4 & 6 \end{bmatrix}$$

is not a member of $\Omega(1,3)$ because the lower-triangular elements of $\underset{\sim}{\Psi}$ are

$$\begin{bmatrix} 2 & & & & \\ 3 & -1 & & & \\ 1 & 2 & -1 & & \\ 0 & 1 & 2 & -1 & \\ 0 & 0 & 1 & 3 & 2 \end{bmatrix}$$

which are not all non-negative. I consider this to be an undesirable feature of the set. It may be noted that

$$\Psi_{ij} = -\operatorname{cov}(e_{i+1} - e_i \, , \, e_j - e_{j-1}) \quad \text{for} \quad i > j \, ,$$

so the condition $\omega=1$ is the same as the assumption that the covariances among the first differences of the error process are all non-positive.

Specifying $\omega>2$ imposes additional, possibly unrealistic, constraints on $\underset{\sim}{V}$. In particular, the constraint on ρ_ℓ in [9.2.15] may be unreasonable if the observation times are unequally spaced. Also ρ_1 and ρ_{n-1} must each be less than 0.5.

One choice of general interest, although not of interest in a thesis on correlated errors, is $\Omega(0,0)$, which is the set of all diagonal matrices with non-negative elements. Thus, it would be appropriate for a situation where errors are independent but have heterogeneous variances of an unknown form.

9.2.17 Other choices for Ω

Other possible restrictions that could be placed on n by n matrices $\underset{\sim}{V}$ that belong to Ω are:

(i) $V_{ii} = \tau^2$ for $i = 1,\ldots,n$;

(ii) $V_{ij} > V_{i+1,j}$ and $V_{i,j-1}$ for $i>j$.

The second set of constraints is more restrictive than those above when $\omega = 0$ but less restrictive than when $\omega = 1$. These alternative sets of constraints, although realistic in some applications, cannot be incorporated into the optimization problem without greatly increasing its complexity.

9.2.18 Distribution of $\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}$

Up to this point attention has focused on the expectation of $\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}$ and its distribution has been ignored, except to note that the variate is always non-negative because $\underset{\sim}{C}$ is constrained to be positive-semidefinite by the way it is constructed.

If $\underset{\sim}{y}$ is normally distributed then

$$E((\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}})^2) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n} C_{ij}\, C_{k\ell}\, (E(\hat{e}_i\hat{e}_j)E(\hat{e}_k\hat{e}_\ell) + E(\hat{e}_i\hat{e}_k)E(\hat{e}_j\hat{e}_\ell) + E(\hat{e}_i\hat{e}_\ell)E(\hat{e}_j\hat{e}_k)) ,$$

$$= (\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}))^2 + 2\,\mathrm{tr}(\underset{\sim}{C}\,\underset{\sim}{V}\,\underset{\sim}{C}\,\underset{\sim}{V}) .$$

Therefore

$$\text{var}(\hat{\underset{\sim}{e}}\ \underset{\sim}{C}\ \hat{\underset{\sim}{e}}) = 2\ \text{tr}(\underset{\sim}{C}\ \underset{\sim}{V}\ \underset{\sim}{C}\ \underset{\sim}{V})\ .$$

A quadratic form of normal deviates is not, in general, χ^2 distributed. However, it may be approximated by a χ^2 distribution by equating the first and second moments (Satterthwaite, 1946). In this example it means assuming that $\hat{\underset{\sim}{e}}^T\ \underset{\sim}{C}\ \hat{\underset{\sim}{e}}$ is distributed as $\frac{1}{r}\text{tr}(\underset{\sim}{C}\ \underset{\sim}{V})\chi^2_r$, where the non-integer degrees of freedom r is defined by

$$r = \frac{(\text{tr}\ (\underset{\sim}{C}\ \underset{\sim}{V}))^2}{\text{tr}(\underset{\sim}{C}\ \underset{\sim}{V}\ \underset{\sim}{C}\ \underset{\sim}{V})}\ .$$

The degrees of freedom can be used, after substituting $\hat{\underset{\sim}{V}}$ for $\underset{\sim}{V}$, to obtain t-statistics in order to construct approximate marginal confidence intervals for β_s.

It would be desirable to incorporate the variance of $\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\ \hat{\underset{\sim}{e}}$ in the optimization problem rather than just concentrating on the expectation of the estimator. However, this cannot be achieved without greatly complicating the problem and, in particular, forfeiting the linearity in the optimization problem.

It is important to note that these degrees of freedom are conditional upon $\hat{\underset{\sim}{V}}$ being the true error variance matrix which may seem rather inconsistent since the whole justification for this chapter is that the error variance matrix cannot be assumed to be $\hat{\underset{\sim}{V}}$. It will always be conservative, and therefore safe, to assume that r is equal to unity, although the marginal confidence intervals may be much wider than otherwise.

9.3 Computer program CEVOPE

9.3.1 Programming philosophy

The computer program CEVOPE, Conservative Estimates of the Variances Of regression Parameter Estimators, implements the method described in section 9.2. It is the program used in subsequent sections to apply the techniques to real data sets. In fact, without the existence of powerful computers the method in section 9.2 would only be of academic interest. The program is written in FortranIV. All arrays are declared in the main program and transferred to subroutines where all operations are performed except for input and output. Because of the amount of computer time required by the algorithm, efforts have been made to reduce the CPU and disc times to as low values as possible. Under these restrictions clarity of code has been a priority. There is strict checking of input parameters, clear error messages and well annotated output. An indicator variable can be set to generate the printing of all intermediate results in the program.

A brief description of the program structure, input and output follows. Then, some specific features of the program are considered, which serves to bridge the gap between the mathematical discussion in section 9.2 and the problems associated with computer programming. Elsewhere, I have discussed the computational problem in more detail (Glasbey, 1984).

9.3.2 Program structure

The main program calls a sequence of subroutines:

CVINPT, to input the arrays associated with the model (output from REGAME);

CVVARS, to calculate the model dependent parameter estimator variances;

CVSTTC, to select the first set of c-vectors;

CVINIT, to set-up the input arrays for the linear program;

CVLINP, to find the best combination of c-vectors using NAG library routine H01ADF;

CVEIGC, to find a new c-vector by solving an eigenvalue problem;

CVMODI, to modify input arrays for the linear program;

CVINCR, to increase the first dimension of array XLHS if necessary.

Then CVLINP is called again, followed by CVEIGC, and so the algorithm loops round. Once a stopping condition has been reached CVRESU is called to calculate the results of using the conservative estimator.

Auxiliary subroutines which are used are:

REGTRA, already described in chapter 2;

CVIXXT, to apply the transformation $(\underset{\sim}{I}-\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)$;

CVOMEG, to calculate the parameters in the linear program corresponding to a choice of set $\Omega(\omega,\upsilon)$;

CVVOPT, to apply the transformation $\underset{\sim}{V}''$.

General output routines /NINOUT/,MATPN0,MATPN1,MATPR0,MATPR1 and MATPR2 are also used. The program is listed in Appendix C.

9.3.3 Input

Control is from channel NIN5. Ten parameters are required to initiate each run of the program, in 10I7 format:

ND, the number of data sets to advance from channels NIN3, NIN4, (<0 to stop the program);

NVARS, 1 to calculate model dependent results in CVVARS;

NS, number of parameter (s) for which conservative estimate sought;

NSEED, seed for NAG random number generator (if 0 seed chosen at random, if < 0 c-vectors restored from channel NIN2);

NT, initial number (t) of c-vectors (unless NSEED<0)

NOMEGA, difference operator in Ω (i.e. ω in $\Omega(\omega,\upsilon)$);

NLAG, lag beyond which covariances are zero (i.e. υ in $\Omega(\omega,\upsilon)$);

NITER, maximum number of iterations;

NCPU, maximum CPU time;

NPRI, 1 to output intermediate calculations.

Arrays associated with a model are input from channels NIN3 and NIN4. These correspond to the output from REGAME to channels NOUT8 and NOUT9.

If c-vectors found by a previous run of CEVOPE are to be re-used, these are restored from channel NIN2.

9.3.4 Output

Output is to channel NOUT6. Initially, if NVARS=1, the following is output from CVVARS:

Parameter estimates ($\hat{\underset{\sim}{\beta}}$) ;

standard errors $\surd(\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{u})$;

parameter estimator variances assuming an error variance of $\hat{\underset{\sim}{V}}$, $(\underset{\sim}{U}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{U})$;

lower bounds on variances, $(\underset{\sim}{X}^T\hat{\underset{\sim}{V}}^{-1}\underset{\sim}{X})^{-1}$;

efficiency of estimators, $\surd(\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{u})/\surd((\underset{\sim}{X}^T\hat{\underset{\sim}{V}}^{-1}\underset{\sim}{X})^{-1})_{ss}$ for s=1,...,m .

Then there follows a heading giving s , $\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{u}$, $\surd(\underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u})$ and NSEED and the iterative history:

ITER, number of iterations;

ICPUT, total CPU time used in seconds;

NT, number of c-vectors (t) in current solution;

ITEIG, number of iterations required in CVEIGC to find new c-vector;

ICPU, CPU time used by CVEIGC on last call;

IDIS, disc time used by CVEIGC on last call;

QUAD, the value of $\underset{\sim}{c}^{\{t+1\}^T}(\hat{\underset{\sim}{V}}-\underset{\sim}{V}'')\,\underset{\sim}{c}^{\{t+1\}}$;

ITLP, number of iterations required by linear program;

ICPU, CPU time used by linear program on last call;

IDIS, disc time used by linear program on last call;

XOPT, lowest estimate of standard error so far found for $\surd tr(\underset{\sim}{C}\,\hat{\underset{\sim}{V}}) \equiv \surd(\underset{\sim}{u}^T\underset{\sim}{V}''\,\underset{\sim}{u})$;

BIAS, relative bias of conservative estimate of standard error $\surd(\underset{\sim}{u}^T\underset{\sim}{V}''\,\underset{\sim}{u})/\surd(\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{u}) - 1$.

Once iterations have ceased the following are output from CVRESU:

Number of data sets advanced, ND;

Number of parameter, NS;

Estimate of parameter value, $\hat{\beta}_s$;

Model dependent results

Variance of estimator assuming error variance $\hat{\underset{\sim}{V}}$, $\underset{\sim}{u}^T\underset{\sim}{V}\,\underset{\sim}{u}$;

S.e. of estimator assuming $\hat{\underset{\sim}{V}}$, $\surd(\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\,\underset{\sim}{u})$;

Expected results from conservative estimator

Difference operator applied to Ω , NOMEGA;

Lag beyond which covariances are zero, NLAG;

Expected value of conservative estimator of variance assuming $\hat{\underset{\sim}{V}}$, $tr(\underset{\sim}{C}\,\hat{\underset{\sim}{V}})$;

Expected conservative estimator of s.e. assuming $\hat{\underset{\sim}{V}}$,

$\surd tr(\underset{\sim}{C}\ \hat{\underset{\sim}{V}})$;

Relative bias in s.e. assuming $\hat{\underset{\sim}{V}}$, $\surd(tr(\underset{\sim}{C}\ \hat{\underset{\sim}{V}}))/\surd(\underset{\sim}{u}^T\hat{\underset{\sim}{V}}\ \underset{\sim}{u}) - 1$;

Variance of conservative estimator of variance,

$2tr(\underset{\sim}{C}\ \hat{\underset{\sim}{V}}\ \underset{\sim}{C}\ \hat{\underset{\sim}{V}})$;

Approximate degrees of freedom (r) of conservative estimator

$tr(\underset{\sim}{C}\ \hat{\underset{\sim}{V}})^2/tr(\underset{\sim}{C}\ \hat{\underset{\sim}{V}}\ \underset{\sim}{C}\ \hat{\underset{\sim}{V}})$;

Approximate 95% t-statistic ;

Actual results from conservative estimator

Estimate of variance using conservative estimator, $\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\ \hat{\underset{\sim}{e}}$;

Estimate of s.é. using conservative estimator, $\surd(\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\ \hat{\underset{\sim}{e}})$;

Approximate 95% lower confidence bound for parameter,

$\hat{\beta}_s - t_r(95\%)\surd(\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\ \hat{\underset{\sim}{e}})$;

Approximate 95% upper confidence bound for parameter,

$\hat{\beta}_s + t_r(95\%)\surd(\hat{\underset{\sim}{e}}^T\underset{\sim}{C}\ \hat{\underset{\sim}{e}})$.

An example of output is given in appendix D, and is discussed further in section 9.4.

The final c-vectors are output to channel NOUTIO for possible future restoration via channel NIN2.

9.3.5 Stopping conditions

Iterations cease when one of four conditions is reached:

a) there are no negative eigenvalues, so the optimal matrix $\underset{\sim}{C}$ has been found;

b) XOPT has not decreased in the last iteration;

c) the maximum number of iterations has been reached;

d) the CPU time is about to exceed its maximum allocation.

Only in case (a) can the user be sure of having found the best matrix $\underset{\sim}{C}$. In case (b) a better solution may be found starting from a different seed. In cases (c) and (d) the existing solution can be restored from channel NIN2 and further improved.

If any of these conditions is met then CVRESU is called. CVRESU is also called in the event of two error conditions being encountered:

e) NT has exceeded its maximum value (30);

f) CVEIGC fails to converge after 1000 iterations.

In case (e) all that can be done is to restart from a different random number generator seed in the hope of finding a better solution, because the program cannot be restarted from the final set of c-vectors as they exceed the capacity of the program. Condition (f) may mean that the current solution is close to being optimal.

If any other failure condition is encountered then meaningful output cannot be produced by CVRESU, so the run ends and new input is requested.

9.3.6 Algorithm for linear program

The revised simplex algorithm (see for example Garvin, chapter 13, 1960) has been used to solve the dual optimization problem in order to avoid having to use the enormous square matrix corresponding to the $(2n-\upsilon)(\upsilon+1)/2$ auxiliary variables required by the inequality constraints given in [9.2.12] for example. Because elements in the large matrix of constraints are accessed by the algorithm row by row whereas FortranIV stores array elements by columns, it is more efficient to store the transpose of the

matrix. Also, if the first dimension of the array XLHS holding the constraints matrix is of fixed size 30, that is the maximum value of t, and there are 10,000 constraints, then with double precision variables the array is of size 2.4 Mbytes. However, the central processor unit (CPU) space available on the Prime 550 computer is only 1.7 Mbytes. Therefore, in paged virtual memory parts of the array have to be moved in and out of the CPU as required. This incurs disc I/O time and dramatically increases the total run time for the program. If instead, the first dimension of the array is set close to t then, when t is less then 30, elements in the constraints matrix may be placed in closer proximity in computer store and if no other job is running concurrently on the computer, disc I/O may be unnecessary. In order to achieve this, the first dimension of the array has to be variable so that it can be changed as t is changed between linear programs. Another technique that is being used is to place all variables used in the linear programming algorithm into a single common block. This increases the chances of variables required consecutively being on the same page of virtual memory, which reduces access time.

A linear programming routine in the NAG library, H01ADF (Numerical Algorithms Group, 1983), has been used in which any degeneracy encountered is resolved by the perturbed problem technique described for example by Garvin (chapter 14, 1960). It is almost certainly possible to improve upon the general routine for this particular problem, for example by using as a starting value the solution of the preceding linear program rather than allowing H01ADF to set all the variables to zero initially. Also,

until the final few iterations, the time-consuming search performed by the simplex algorithm could be restricted to a subset only of the large number of constraints.

9.3.7 Algorithm for calculating XLHS

Because of the size of the array XLHS the algorithm used to calculate its elements is of critical importance. In order to calculate the element in a row of XLHS, denoted $\underset{\sim}{w}$, corresponding to a c-vector, denoted $\underset{\sim}{c}$, for $\Omega(\omega,\upsilon)$ a doubly-recursive algorithm has been used:

First, k = 0

```
For  i = 1,...,n
     For  j = i,i-1,..,max(1,i-υ)
     k = k+1
     w_k = 2c_i c_j/(1+I_ij)
     End
End.
```

This gives the elements appropriate for $\Omega(0,\upsilon)$, previously specified in [9.2.12].

Then if $\omega > 0$

```
For  ℓ = 1,...,ω
     k = 0
     h = 0
     For  i = 1,...,n
          For  j = i,i-1,...,max(1,i-υ)
               k = k+1
               If  j < i   w_k = w_k+w_k-1+v_i-j
               If  j < i-1  w_k = w_k-v_i-j-1
          End
```

```
                For  j = i,i-1,...,max(1,i-υ)
                     h = h+1
                     v_{i-j+1} = w_h
                End
            End
        End.
```

This gives the elements appropriate for $\Omega(\omega,\upsilon)$, with the same results for ω equal to 1 or 2 as previously given in [9.2.13] and [9.2.14]. A listing of this algorithm implemented in FortranIV is given in subroutine CVOMEG in Appendix C.

9.3.8 Algorithm for obtaining eigenvectors

From a random initial vector $\underset{\sim}{w}^{\{0\}}$ orthogonal to $\underset{\sim}{X}$, $\underset{\sim}{w}^{\{1\}}$ is derived by the transformation

$$\underset{\sim}{w}^{\{1\}} = (\underset{\sim}{I} - (\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)[(\hat{\underset{\sim}{V}}-\underset{\sim}{\overset{''}{V}}) - \mu^{\{0\}}\underset{\sim}{I}]\,\underset{\sim}{w}^{\{0\}} ,$$

which is obtained by transforming by each component matrix in turn. The transformation $\underset{\sim}{\overset{''}{V}}$ is given in [9.3.9]. Then $\underset{\sim}{w}^{\{1\}}$ is standardized to have a largest element of unity, by dividing by its element of largest absolute magnitude, denoted $\lambda^{\{1\}}$. The difference between $\underset{\sim}{w}^{\{1\}}$ and $\underset{\sim}{w}^{\{0\}}$ is calculated as

$$x^{\{1\}} = \sum_{i=1}^{n} |w_i^{\{1\}} - w_i^{\{0\}}| .$$

This is then repeated for $\underset{\sim}{w}^{\{2\}}, \underset{\sim}{w}^{\{3\}}, \ldots$.

At the start $\mu^{\{0\}} = 0$, but every 5 iterations if $\lambda^{\{i\}} > 0$ then

$$\mu^{\{i\}} = \mu^{\{i-1\}} + \lambda^{\{i\}} .$$

After every 250 iterations, or when $x^{\{i\}} < 10^{-10}$, the ratio of quadratic forms

$$y^{\{i\}} = \underset{\sim}{w}^{\{i\}T}(\underset{\sim}{\hat{V}}-\underset{\sim}{\overset{\shortparallel}{V}})\underset{\sim}{w}^{\{i\}}/\underset{\sim}{w}^{\{i\}T}\underset{\sim}{w}^{\{i\}}$$

is calculated as

$$(\lambda^{\{i+1\}}\underset{\sim}{w}^{\{i\}T}\underset{\sim}{w}^{\{i+1\}} + \mu^{\{i\}}\underset{\sim}{w}^{\{i\}T}\underset{\sim}{w}^{\{i\}})/\underset{\sim}{w}^{\{i\}T}\underset{\sim}{w}^{\{i\}} .$$

If $y^{\{i\}}$ is negative, or if $x^{\{i\}} < 10^{-10}$ and $\lambda^{\{i+1\}} < 0$, then the algorithm can stop with $\underset{\sim}{c}^{\{t+1\}}$ set to $\underset{\sim}{w}^{\{i\}}$. Otherwise, iteration continues with

$$\mu^{\{i+1\}} = \mu^{\{i\}} + y^{\{i\}}$$

until either the above conditions are satisfied or 1000 iterations are completed, in which case an error state is entered. A listing of this algorithm implemented in FortranIV is given in subroutine CVEIGC in Appendix C.

9.3.9 Algorithm for transforming with respect to $\underset{\sim}{\overset{\shortparallel}{V}}$

The first time the transformation is to be used, a vector $\underset{\sim}{w}$ is created using the output from the linear program H01ADF (MOPT(1,...,t) and VOPT(1,...,t)). For $\Omega(\omega,\upsilon)$ a doubly-recursive algorithm has been used which in some senses is the inverse of that given in [9.3.7]:

```
w = 0
For  ℓ = 1,...,t
     k = 0
     For i = 1,...,n
         For j = i,i-1,...,max(1,i-υ)
         k = k+1
         if  MOPT(ℓ) = k    w_k = VOPT(ℓ)
         End
     End
End.
```

If $\omega > 0$

For $\ell = 1,\ldots,\omega$

$\underset{\sim}{v} = \underset{\sim}{0}$

$k = (2n-\upsilon)(\upsilon+1)/2+1$

$h = k$

For $i = n,n-1,\ldots,1$

For $j = \max(1,i-\upsilon),\ldots,i-1,i$

$k = k - 1$

$w_k = w_k + v_j$

If $j > \max(1,i-\upsilon)$ $w_k = w_k + w_{k+1} - v_{j-1}$

End

For $j = \max(1,i-\upsilon),\ldots,i-1,i$

$h = h-1$

$v_j = w_h$

End

End

End.

Then, and on subsequent occasions, the transformation

$$\underset{\sim}{v}^{\{2\}} = \underset{\sim}{\overset{\cdot\cdot}{V}}\ \underset{\sim}{v}^{\{1\}}$$

is obtained by

$\underset{\sim}{v}^{\{2\}} = \underset{\sim}{0}$

$k = 0$

For $i = 1,\ldots,n$

For $j = i,i-1,\ldots,\max(1,i-\upsilon)$

$k = k+1$

$v_i^{\{2\}} = v_i^{\{2\}} + w_k v_j^{\{1\}}$

If $i \neq j$ $v_j^{\{2\}} = v_j^{\{2\}} + w_k v_i^{\{1\}}$

End

End.

A listing of this algorithm implemented in FortranIV is given in subroutine CVVOPT in Appendix C.

9.3.10 Algorithm for t-statistics

The degrees of freedom of the conservative estimator of variance are

$$r = (\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}}))^2/\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}}\,\underset{\sim}{C}\,\hat{\underset{\sim}{V}}) \ ,$$

as given in [9.2.18]. This can be calculated using the decomposition of $\underset{\sim}{C}$ in [9.2.6] and

$$A_{k\ell} = \sqrt{x_k x_\ell}\ \underset{\sim}{c}^{\{k\}T}\ \hat{\underset{\sim}{V}}\ \underset{\sim}{c}^{\{\ell\}} \qquad \text{for } k,\ell = 1,\ldots,t \ ,$$

because then

$$\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}}) = \sum_{k=1}^{t} A_{kk} \ ,$$

and

$$\mathrm{tr}(\underset{\sim}{C}\,\hat{\underset{\sim}{V}}\,\underset{\sim}{C}\,\hat{\underset{\sim}{V}}) = \sum_{k=1}^{t}\sum_{\ell=1}^{t} A_{k\ell}\,A_{\ell k} \ .$$

The appropriate t-statistics are then obtained by interpolation between the t-statistics at integer degrees of freedom using a geometric scale.

9.4 An illustrative example

9.4.1 Model

To illustrate the algebra of section 9.2, a particularly simple numerical example will be analysed. A vector $\underset{\sim}{y}$ of 5 observations is defined to be a realisation from the model:

$$\underset{\sim}{y} \sim N(\underset{\sim}{f}\ ,\ \tau^2\ \underset{\sim}{I}) \ ,$$

where $f_i = \beta_1 + \beta_2 i$ for $i = 1,\ldots,5$,

$\tau^2 = 1$

and $\underset{\sim}{\beta}^T = (0,0)$.

Therefore, the design matrix $\underset{\sim}{X}$ is

$$\begin{bmatrix} 1, & 1 \\ 1, & 2 \\ 1, & 3 \\ 1, & 4 \\ 1, & 5 \end{bmatrix} .$$

If $\underset{\sim}{\beta}$ is estimated by ordinary least-squares then, in the notation of [9.2.1],

$$\underset{\sim}{W} = \underset{\sim}{R} = \underset{\sim}{I}$$

and $$\underset{\sim}{U} = \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} = \begin{bmatrix} 0.8, & -0.2 \\ 0.5, & -0.1 \\ 0.2, & 0.0 \\ -0.1, & 0.1 \\ -0.4, & 0.2 \end{bmatrix} .$$

In what follows the conservative estimators of $\mathrm{var}(\hat{\beta}_s)$ (for s=1 or 2) with the minimum biases will be found for a range of choices of Ω , on the assumption that $\hat{\underset{\sim}{V}} = \underset{\sim}{I}$. One particular case, that is when s is equal to 2 and the error variance matrix is restricted to belong to the set $\Omega(1,4)$, is considered in detail in subsections [9.4.2] to [9.4.8]. Other cases with s equal to 2 are considered in [9.4.9] to [9.4.11], and cases with s equal to 1 are considered in [9.4.12].

The effect of the choice of $\hat{\underset{\sim}{V}}$ on the best conservative estimator is illustrated in [9.4.13]. Finally, in [9.4.14] the potential of the technique to aid in the selection of the estimators of the regression parameters themselves is demonstrated.

9.4.2 Dual optimization problem for s=2, Ω (1, 4)

If $\hat{\underset{\sim}{V}}$ is assumed to be diagonal with constant elements, then for simplicity $\hat{\underset{\sim}{V}}$ can be specified as $\underset{\sim}{I}$. Further, if it is assumed that $\underset{\sim}{V} \in \Omega(1,4)$, then transforming to a new set of variates $\underset{\sim}{\Psi}$ in accord with [9.2.13], the optimization problem becomes, in the case when s=2:

$$\begin{aligned}
\text{maximize} \quad & 0.04\,\Psi_{11} \\
& + 0.09\,\Psi_{21} + 0.01\,\Psi_{22} \\
& + 0.09\,\Psi_{31} + 0.01\,\Psi_{32} + 0.00\,\Psi_{33} \\
& + 0.04\,\Psi_{41} + 0.00\,\Psi_{42} + 0.01\,\Psi_{43} + 0.01\,\Psi_{44} \\
& + 0.00\,\Psi_{51} + 0.04\,\Psi_{52} + 0.09\,\Psi_{53} + 0.09\,\Psi_{54} + 0.04\,\Psi_{55}
\end{aligned}$$

$$\begin{aligned}
\text{Subject to} \quad & c_1^{\{k\}2}\Psi_{11} \\
& + (c_1^{\{k\}} + c_2^{\{k\}})^2\Psi_{21} + c_2^{\{k\}2}\Psi_{22} \\
& + (c_1^{\{k\}} + c_2^{\{k\}} + c_3^{\{k\}})^2\Psi_{31} + (c_2^{\{k\}} + c_3^{\{k\}})^2\Psi_{32} + c_3^{\{k\}2}\Psi_{33} \\
& + (c_1^{\{k\}} + c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}})^2\Psi_{41} + (c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}})^2\Psi_{42} \\
& \quad + (c_3^{\{k\}} + c_4^{\{k\}})^2\Psi_{43} + c_4^{\{k\}2}\Psi_{44} \\
& + (c_1^{\{k\}} + c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}} + c_5^{\{k\}})^2\Psi_{51} \\
& \quad + (c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}} + c_5^{\{k\}})^2\Psi_{52} \\
& \quad + (c_3^{\{k\}} + c_4^{\{k\}} + c_5^{\{k\}})^2\Psi_{53} \\
& \quad + (c_4^{\{k\}} + c_5^{\{k\}})^2\Psi_{54} + c_5^{\{k\}2}\Psi_{55} \\
& \leq \underset{\sim}{c}^{\{k\}T}\underset{\sim}{c}^{\{k\}} \qquad \text{for } k = 1,\ldots,t
\end{aligned}$$

$$\text{for} \quad \underset{\sim}{\Psi} \geq \underset{\sim}{0} \ .$$

9.4.3 Solution when t=1

The above problem requires the c-vectors to be specified. These vectors, or in fact a single vector, can be generated by a

random process. For simplicity consider a single vector, that is $t=1$. A vector orthogonal to $\underset{\sim}{X}$ is chosen at random and standardized to have a largest element of unity:

$$\underset{\sim}{c}^{\{1\}^T} = (-0.1275,\ -0.4467,\ 1.0000,\ -0.1498,\ -0.2760) .$$

The above optimization problem is solved by

$$\Psi_{ij} = \begin{cases} 80.81 & \text{if } (i,j)=(5,2) \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, back-transforming, the optimal choice of $\underset{\sim}{V}$ in $\Omega(1,4)$ is

$$\underset{\sim}{\overset{\text{"}}{V}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 80.81 & 80.81 & 80.81 & 80.81 \\ 0 & 80.81 & 80.81 & 80.81 & 80.81 \\ 0 & 80.81 & 80.81 & 80.81 & 80.81 \\ 0 & 80.81 & 80.81 & 80.81 & 80.81 \end{bmatrix} ,$$

and $\underset{\sim}{u}^T\underset{\sim}{\overset{\text{"}}{V}}\,\underset{\sim}{u} = 3.233$, where $\underset{\sim}{u}$ is the second column of $\underset{\sim}{U}$. The Lagrangian multiplier gives the solution to the primal problem as

$$x_1 = 2.459$$

and $$tr(\underset{\sim}{C}) = x_1\ \underset{\sim}{c}^{\{1\}^T}\ \underset{\sim}{c}^{\{1\}} = 3.233 .$$

Therefore, the solutions of the primal and dual problems agree as they should do. The relative bias in the standard error is

$$\sqrt{(\underset{\sim}{u}^T\underset{\sim}{\overset{\text{"}}{V}}\,\underset{\sim}{u})}/\sqrt{(\underset{\sim}{u}^T\underset{\sim}{u})}-1 = 4.685 .$$

9.4.4 The search for $\underset{\sim}{c}^{\{2\}}$

Once the optimization problem has been solved the next step (as described in [9.2.8]) is to find $\underset{\sim}{c}^{\{2\}}$ as the eigenvector with largest negative eigenvalue in

$$(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)(\underset{\sim}{I} - \underset{\sim}{\overset{\text{"}}{V}})(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T) .$$

Starting with a vector orthogonal to $\underset{\sim}{X}$ chosen at random,

$$(0.5771, -0.9315, -0.1105, 0.7069, -0.2420)^T ,$$

this is transformed, using the algorithm described in [9.3.8], on the first iteration to

$$(1.0 , -0.9804, -0.5099, -0.0391, 0.5294)^T$$

with $\lambda^{\{1\}} = -18.08$, on the second iteration to

$$(-0.9994, 1.0 , 0.4994, -0.0012, -0.4988)^T$$

with $\lambda^{\{2\}} = 31.34$, thence to

$$(1.0, -1.0, -0.5, 0.0, 0.5) , \lambda^{\{3\}} = 31.30 ,$$

$$(-1.0, 1.0, 0.5, 0.0, -0.5) , \lambda^{\{4\}} = 31.32 ,$$

and $$(1.0, -1.0, -0.5, 0.0, 0.5) , \lambda^{\{5\}} = 31.32 .$$

It appears that the matrix has at least two eigenvalues of equal largest magnitude because the algorithm is not converging. At this point, after 5 iterations $\lambda^{\{5\}} > 0$ so $\mu^{\{5\}}$ is set to 31.32. On the next iteration the vector is transformed to

$$(1.0, -1.0, -0.5, 0.0, 0.5) , \lambda^{\{6\}} = -62.65 ,$$

to which it converges exactly in the course of the next ten iterations. Therefore the matrix has a single eigenvalue of -31.3 and at least one eigenvalue of 31.3, and introducing the shift of 31.32 has ensured the eigenvalue of largest magnitude is unique and negative.

9.4.5 Solution when t=2

The new optimization problem with $\underset{\sim}{c}^{\{2\}}$ incorporated is solved by

$$\Psi_{ij} = \begin{cases} 7.188 & \text{if } (i,j)=(3,1) \\ 0.703 & \text{if } (i,j)=(5,2) \\ 0 & \text{otherwise .} \end{cases}$$

Therefore

$$\underset{\sim}{V}'' = \begin{bmatrix} 7.188 & 7.188 & 7.188 & 0.0 & 0.0 \\ 7.188 & 7.891 & 7.891 & 0.703 & 0.703 \\ 7.188 & 7.891 & 7.891 & 0.703 & 0.703 \\ 0.0 & 0.703 & 0.703 & 0.703 & 0.703 \\ 0.0 & 0.703 & 0.703 & 0.703 & 0.703 \end{bmatrix}$$

and $\underset{\sim}{u}^T \underset{\sim}{V}'' \underset{\sim}{u} = 0.6750$.

The primal solution is

$$\underset{\sim}{x}^T = (0.4514,\ 0.0327) ,$$

and the relative bias equals 1.598.

9.4.6 New c-vectors and bias

From this solution, a new c-vector can be found

$$(-0.5515,\ 0.3101,\ 1.0000,\ -0.7242,\ -0.0343)^T$$

which reduces the relative bias to 1.491 when incorporated in the optimization problem. On succeeding iterations the c-vectors and relative biases are:

$$(\ 0.4924,\ 0.0341,\ -0.5452,\ -0.9812,\ 1.0000)^T ,\quad 1.491 ,$$

$$(-0.6156,\ 0.1218,\ 1.0000,\ 0.0970,\ -0.6032)^T ,\quad 1.314 ,$$

$$(-0.3735,\ -0.0968,\ 1.0000,\ -0.2154,\ -0.3142)^T ,\quad 1.263 ,$$

$$(-0.3184,\ -0.3794,\ 1.0000,\ 0.4117,\ -0.1393)^T ,\quad 1.182 ,$$

$$(-0.2607,\ -0.3326,\ 1.0000,\ 0.0405,\ -0.4472)^T ,\quad 1.170 ,$$

$$(-0.2509,\ -0.2776,\ 1.0000,\ -0.1633,\ -0.3081)^T ,\quad 1.160 .$$

9.4.7 Final solution

After 18 iterations the relative bias stops decreasing and 4 c-vectors remain. However, they are all very similar and the solution has 1.00008 degrees of freedom (r). Examination of the solution reveals that the optimum corresponds to $\underset{\sim}{C} = x\ \underset{\sim}{c}\ \underset{\sim}{c}^T$, for

a single vector with x equal to 0.36 and

$$\underset{\sim}{c}^T = (-1/3, -1/6, 1, -1/6, -1/3)$$

$$= (-0.3333, -0.1667, 1.0000, -0.1667, -0.3333) .$$

It is because of the effects of rounding error that the final solution does not give this single vector exactly.

The solution of the optimization problem has

$$\underset{\sim}{\overset{\text{"}}{V}} = \begin{bmatrix} 2.5556 & 2.5556 & 0.0 & 0.0 & 0.0 \\ 2.5556 & 2.5556 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 2.5556 & 2.5556 \\ 0.0 & 0.0 & 0.0 & 2.5556 & 2.5556 \end{bmatrix} ,$$

$$\underset{\sim}{u}^T \underset{\sim}{\overset{\text{"}}{V}} \underset{\sim}{u} = 0.46 ,$$

$$\underset{\sim}{C} = \begin{bmatrix} 0.04 & 0.02 & -0.12 & 0.02 & 0.04 \\ 0.02 & 0.01 & -0.06 & 0.01 & 0.02 \\ -0.12 & -0.06 & 0.36 & -0.06 & -0.12 \\ 0.02 & 0.01 & -0.06 & 0.01 & 0.02 \\ 0.04 & 0.02 & -0.12 & 0.04 & 0.04 \end{bmatrix} ,$$

$$\text{tr}(\underset{\sim}{C}) = 0.46 ,$$

relative bias = 1.145 ,

and d.f. (r) = 1.0 .

9.4.8 Output from CEVOPE

Appendix D shows the output from CEVOPE for the preceding problem, including the numerical results when

$$\underset{\sim}{y}^T = (0.2247, 0.4873, -1.081, -0.8700, -0.5833) .$$

In this case

$$\underset{\sim}{\hat{\beta}}^T = (0.5274, -0.2973) ,$$

$$\hat{\tau}^2 = 0.3354,$$

and $$\underset{\sim}{\hat{e}}^T = (-0.0054, 0.5545, -0.7169, -0.2081, 0.3759) .$$

The model dependent standard error of $\hat{\beta}_2$ is

$$\sqrt{(\hat{\tau}^2 \, \underset{\sim}{u}^T \underset{\sim}{u})} = 0.1831$$

which has 3 degrees of freedom. Therefore 95% marginal confidence bounds are

$$-0.2973 \pm 0.1831 \; t_3(95\%)$$
$$= -0.88 \quad \text{and} \quad 0.29 \; ,$$

which encompass the true value of zero as we would expect. The 95% bounds based on the conservative estimator are

$$-7.14 \quad \text{and} \quad 6.55 \; .$$

These are much wider and are the price that has to be paid in order to have marginal confidence limits which are valid for all variance matrices in the set $\Omega(1,4)$.

9.4.9 Dual optimization problem for s=2, $\Omega(0,1)$

To illustrate the effect of Ω on the problem, consider the situation as above but with the weaker assumption that $\underset{\sim}{V} \in \Omega(0,1)$. The optimization problem is, in accord with [9.2.12]:

$$\begin{aligned}
\text{maximize} \quad & 0.04 \, V_{11} \\
& + 0.04 \, V_{21} + 0.01 \, V_{22} \\
& \quad + 0.00 \, V_{32} + 0.00 \, V_{33} \\
& \quad\quad + 0.00 \, V_{43} + 0.01 \, V_{44} \\
& \quad\quad\quad + 0.04 \, V_{54} + 0.04 \, V_{55}
\end{aligned}$$

$$\begin{aligned}
\text{subject to} \quad & c_1^{\{k\}2} V_{11} \\
& + 2c_1^{\{k\}} c_2^{\{k\}} V_{12} + c_2^{\{k\}2} V_{22} \\
& \quad + 2c_2^{\{k\}} c_3^{\{k\}} V_{32} + c_3^{\{k\}2} V_{33} \\
& \quad\quad + 2c_3^{\{k\}} c_4^{\{k\}} V_{43} + c_4^{\{k\}2} V_{44} \\
& \quad\quad\quad + 2c_4^{\{k\}} c_5^{\{k\}} V_{54} + c_5^{\{k\}2} V_{55} \\
& \quad\quad\quad\quad \text{for} \quad k = 1, \ldots, t \; ,
\end{aligned}$$

for $\underset{\sim}{V} \geq \underset{\sim}{0}$.

This has a dual (which is in fact the original primal problem):

$$\text{minimize} \quad \sum_{k=1}^{t} x_k \, \underset{\sim}{c}^{\{k\}T} \, \underset{\sim}{c}^{\{k\}}$$

$$\text{subject to} \quad \sum_{k=1}^{t} x_k \, c_1^{\{k\}2} \geq 0.04$$

$$\sum_{k=1}^{t} x_k \, 2c_1^{\{k\}} \, c_2^{\{k\}} \geq 0.04$$

$$\sum_{k=1}^{t} x_k \, c_2^{\{k\}2} \geq 0.01$$

$$\sum_{k=1}^{t} x_k \, 2c_2^{\{k\}} \, c_3^{\{k\}} \geq 0.00$$

$$\sum_{k=1}^{t} x_k \, c_3^{\{k\}2} \geq 0.00$$

$$\sum_{k=1}^{t} x_k \, 2c_3^{\{k\}} \, c_4^{\{k\}} \geq 0.00$$

$$\sum_{k=1}^{t} x_k \, c_4^{\{k\}2} \geq 0.01$$

$$\sum_{k=1}^{t} x_k \, 2c_4^{\{k\}} \, c_5^{\{k\}} \geq 0.04$$

$$\sum_{k=1}^{t} x_k \, c_5^{\{k\}2} \geq 0.04$$

for $\underset{\sim}{x} \geq \underset{\sim}{0}$.

For this problem to be solvable it is necessary, in particular, for the expression

$$\sum_{k=1}^{t} x_k (2c_1^{\{k\}} \, c_2^{\{k\}} + c_2^{\{k\}} \, c_3^{\{k\}} + c_3^{\{k\}} \, c_4^{\{k\}} + 2c_4^{\{k\}} \, c_5^{\{k\}})$$

to be greater than zero. However, each $\underset{\sim}{c}$ is orthogonal to $\underset{\sim}{X}$ so

$$c_1 + c_2 + c_3 + c_4 + c_5 = 0$$

and $$c_1 + 2c_2 + 3c_3 + 4c_4 + 5c_5 = 0 \quad .$$

Therefore $$c_4 = \frac{1}{2}(c_1 - c_3 - 3c_5) \ ,$$

$$c_2 = \frac{1}{2}(-3c_1 - c_3 + c_5) \ ,$$

and $$2c_1c_2 + c_2c_3 + c_3c_4 + 2c_4c_5$$

$$= \frac{1}{2}(-3c_1-c_3+c_5)(2c_1+c_3) + \frac{1}{2}(c_1-c_3-3c_5)(c_3+2c_5)$$

$$= -3c_1^2 - c_3^2 - 3c_5^2 - 2c_1c_3 + 2c_1c_5 - 2c_3c_5$$

$$= -(c_1-c_5)^2 - 2(c_1+c_3/2)^2 - 2(c_5+c_3/2)^2$$

$$< 0 \ .$$

Therefore no c-vectors exist satisfying all the constraints. So, if no greater constraint can be placed on $\underset{\sim}{V}$ than that it belongs to $\Omega(\omega,\upsilon)$ with $\omega = 0$ and $\upsilon \geq 1$, then no bound can be placed on the variance of the estimator $\hat{\beta}_2$.

9.4.10 Solutions when s=2

Table 9.4.1 gives the relative biases of the minimum biased estimators for a range of values of ω and υ in $\Omega(\omega,\upsilon)$ when s=2. As ω increases, or υ decreases, the relative bias decreases. This is as it should be because

$$\Omega(\omega,\upsilon) \subset \Omega(\omega-1,\upsilon) \ ,$$

and $$\Omega(\omega,\upsilon) \subset \Omega(\omega,\upsilon+1) \ .$$

Also, when $\upsilon=0$ the differencing operator has no effect so

$$\Omega(\omega,o) = \Omega(0,0) \quad \text{for all} \quad \omega \ ,$$

and this is consistent with the relative biases all being equal in the table when $\upsilon=0$.

When $\upsilon=0$, the optimal choice of $\underset{\sim}{C}$ is $x_1 \underset{\sim}{c}\,\underset{\sim}{c}^T$ where

$$\sqrt{x_1}\,\underset{\sim}{c}^T = (1/5, -2/15, -2/15, -2/15, 1/5)$$

$$= (0.2, -0.1333, -0.1333, -0.1333, 0.2) \ .$$

This example will be used further in [9.4.11]. When $\omega=1$ and $\upsilon=1,2,3$ or 4, once again a single vector is optimal:

$$\sqrt{x_1}\,\underset{\sim}{c}^T = (-0.2,\ -0.1,\ 0.6,\ -0.1,\ -0.2)$$

which was given in [9.4.7]. When $\omega=2$ and $\upsilon=1,2,3$ or 4 the combination of two vectors is optimal, with the second being the reverse of the first:

$$\sqrt{x_1}\,\underset{\sim}{c}^{\{1\}T} = (-0.1993,\ 0.2728,\ 0.0351,\ -0.0915,\ -0.0171)$$

$$\sqrt{x_2}\,\underset{\sim}{c}^{\{2\}T} = (-0.0171,\ -0.0915,\ 0.0351,\ 0.2728,\ -0.1993).$$

Similarly when $\omega=3$ and $\upsilon=1,2,3$ or 4, two vectors are optimal:

$$\sqrt{x_1}\,\underset{\sim}{c}^{\{1\}T} = (-0.1846,\ 0.2277,\ 0.0214,\ 0.0126,\ -0.0770)$$

$$\sqrt{x_2}\,\underset{\sim}{c}^{\{2\}T} = (-0.0770,\ 0.0126,\ 0.0214,\ 0.2277,\ -0.1846).$$

9.4.11 Comparison with standard estimator of variance

The standard least-squares estimator of the variance of $\hat{\beta}_2$ is

$$\hat{var}(\hat{\beta}_2) = \hat{\tau}^2\ \underset{\sim}{u}^T\underset{\sim}{u}$$

where

$$\hat{\tau}^2 = \underset{\sim}{\hat{e}}^T\underset{\sim}{\hat{e}}/3\ .$$

Therefore

$$E(\hat{var}(\hat{\beta}_2)) = \frac{0.1}{3}\ tr((\underset{\sim}{I} - \underset{\sim}{X}\underset{\sim}{U}^T)\underset{\sim}{V}(\underset{\sim}{I} - \underset{\sim}{U}\ \underset{\sim}{X}^T))\ .$$

If $\underset{\sim}{V} \in \Omega(0,0)$, that is the variance matrix is diagonal, then

$$E(\hat{var}(\hat{\beta}_2)) = 0.013V_{11} + 0.023V_{22} + 0.027V_{33} + 0.023V_{44} + 0.013V_{55}\ .$$

This may be compared with

$$var(\hat{\beta}_2) = \underset{\sim}{u}^T\underset{\sim}{V}\ \underset{\sim}{u}$$

$$= 0.040V_{11} + 0.010V_{22} + 0.000V_{33} + 0.010V_{44} + 0.040V_{55}\ ,$$

and the expected value of the conservative estimator given in [9.4.10]:

$$E(\hat{\underset{\sim}{e}}^T \underset{\sim}{C} \hat{\underset{\sim}{e}}) = 0.040V_{11} + 0.018V_{22} + 0.018V_{33} + 0.018V_{44} + 0.040V_{55}.$$

Whatever the positive values of diagonal elements in $\underset{\sim}{V}$

$$E(\hat{e}^T C \hat{e}) \geq \text{var}(\hat{\beta}_2) \quad ,$$

but $\hat{\text{var}}(\hat{\beta}_2)$ is sometimes downward biased. Figure 9.4.1 illustrates this for the particular case when $V_{11} = V_{33} = V_{55} = 1$ and $V_{22} = V_{44}$. When $V_{22} = 1$ all the error variances are equal (this is the assumption made in the standard analysis) and $\hat{\text{var}}(\hat{\beta}_2)$ is unbiased. However, when $V_{22} < 1$ $\hat{\text{var}}(\hat{\beta}_2)$ on average underestimates the variance of $\hat{\beta}_2$.

9.4.12 Solutions when s=1

Table 9.4.2 gives the relative biases of the minimum biased estimators for a range of values of ω and υ in $\Omega(\omega,\upsilon)$ when s=1. When ω=0 and υ>1 no solution is possible for the same reason as in [9.4.9] when s=2. Also, when ω=1 and υ=4 the primal problem includes a constraint

$$\sum_{k=1}^{t} x_k \, (c_1^{\{k\}} + c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}} + c_5^{\{k\}})^2 \geq 1.0 \quad ,$$

but

$$c_1^{\{k\}} + c_2^{\{k\}} + c_3^{\{k\}} + c_4^{\{k\}} + c_5^{\{k\}} = 0 \quad ,$$

so once again no solution is possible.

It has not been possible to find in all cases the optimal c-vectors in simple form as it was when s=2. However it has been possible in three cases. When υ=0 the optimal $\underset{\sim}{C}$ is the combination of two vectors:

$$\surd x_1 \, \underset{\sim}{c}^{\{1\}T} = (0.8, \; -0.8, \; -0.4, \; 0.0, \; 0.4)$$

$$\surd x_2 \, \underset{\sim}{c}^{\{2\}T} = (0, \; -1/70, \; -2/70, \; 1/10, \; -4/70)$$

$$= (0.0, \; -0.0143, \; -0.0286, \; 0.1, \; -0.0571) \quad .$$

Table 9.4.1

Relative biases (degrees of freedom in brackets) of best conservative estimators of standard error of $\hat{\beta}_2$ for a range of values of the parameters in $\Omega(\omega,\upsilon)$ and the model specified in [9.4.1], with $\hat{\underset{\sim}{V}} = \underset{\sim}{I}$.

ω =	0	1	2	3
υ				
0	0.153 (1.0)	0.153 (1.0)	0.153 (1.0)	0.153 (1.0)
1	-	1.145 (1.0)	0.575 (1.8)	0.360 (1.8)
2	-	1.145 (1.0)	0.575 (1.8)	0.360 (1.8)
3	-	1.145 (1.0)	0.575 (1.8)	0.360 (1.8)
4	-	1.145 (1.0)	0.575 (1.8)	0.360 (1.8)

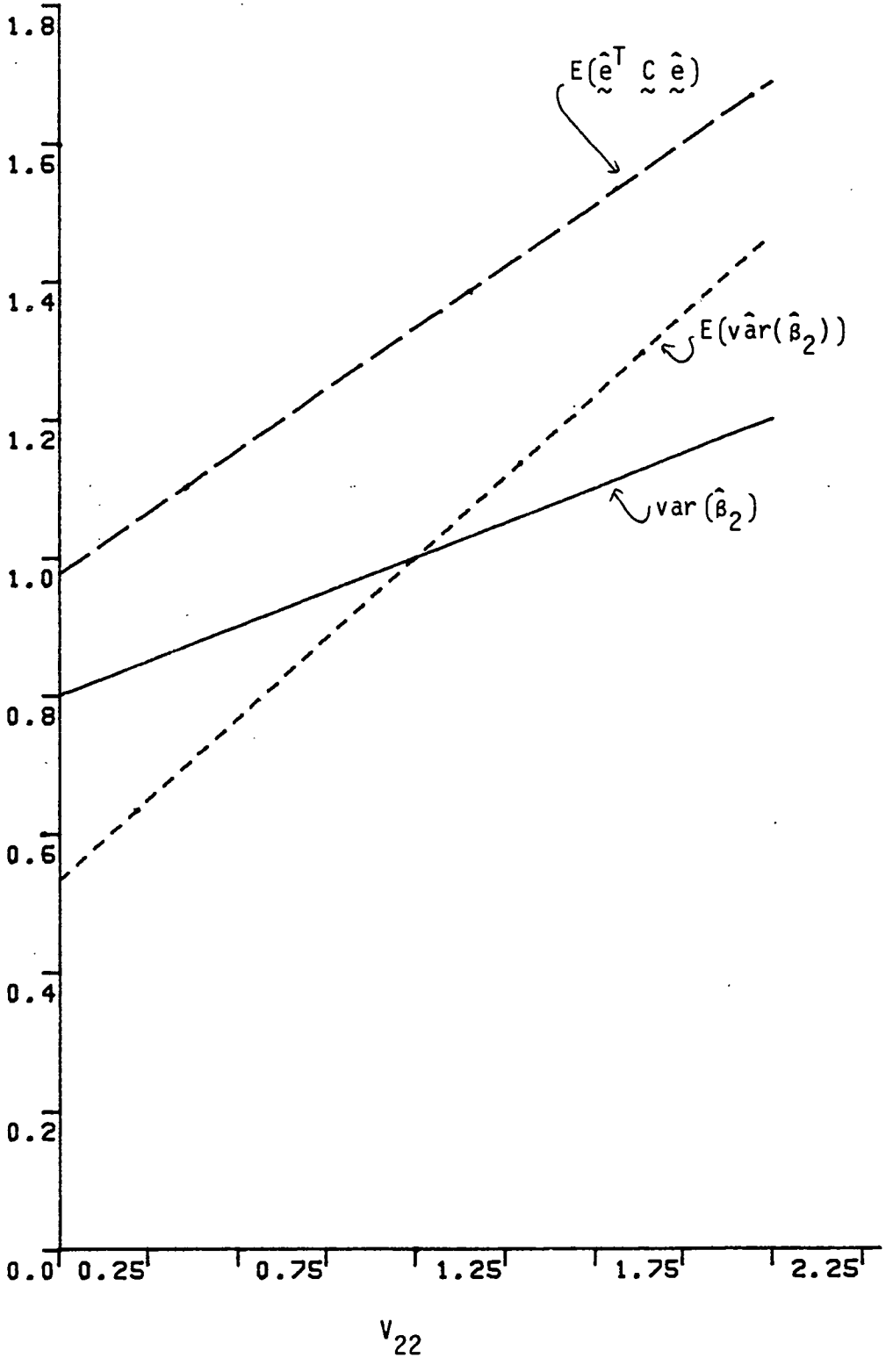
Table 9.4.2

Relative biases (degrees of freedom in brackets) of best conservative estimators of standard error of $\hat{\beta}_1$ for a range of values of the parameters in $\Omega(\omega,\upsilon)$ and the model specified in [9.4.1], with $\hat{\underset{\sim}{V}} = \underset{\sim}{I}$.

ω =	0	1	2	3
υ				
0	0.211 (1.0)	0.211 (1.0)	0.211 (1.0)	0.211 (1.0)
1	-	1.564 (1.4)	0.581 (1.3)	0.342 (1.3)
2	-	2.000 (1.0)	0.609 (1.0)	0.352 (1.1)
3	-	2.387 (1.1)	0.609 (1.0)	0.352 (1.1)
4	-	-	0.609 (1.0)	0.352 (1.1)

Figure 9.4.1

The variance of $\hat{\beta}_2$ and expected values of the conventional estimator and a conservative estimator (for $\Omega(0,0)$) of var $(\hat{\beta}_2)$ for the model specified in [9.4.1], except that $V_{11}=V_{33}=V_{55}=1$ and $V_{22}=V_{44}$, plotted against V_{22}.



When $\omega=1$ and $\upsilon=2$ a single vector is optimal:

$$\surd x_1 \underset{\sim}{c}^T = (0.8,\ 0.5,\ -2.8,\ 0.9,\ 0.6).$$

When $\omega=2$ and $\upsilon=2,3$ or 4, a single vector is optimal:

$$\surd x_1 \underset{\sim}{c}^T = (-(9+\surd 145)/20,\ (\surd 145)/10,\ (19-\surd 145)/20,\ -1/10,\ -4/10)$$
$$= (-1.0521,\ 1.2042,\ 0.3479,\ -0.1,\ -0.4).$$

9.4.13 Dependence on $\hat{\underset{\sim}{V}}$

All the preceding results in this section have been dependent on the assumption that

$$\hat{\underset{\sim}{V}} = \underset{\sim}{I} .$$

For other choices of $\hat{\underset{\sim}{V}}$ the expected relative biases will be different and the best estimators may be different. For example, if $\hat{\underset{\sim}{V}}$ has a AR(1) structure so that

$$\hat{V}_{ij} = \rho^{i-j} , \quad \text{when} \quad i > j$$

for some value of ρ between 0 and 1, $\hat{\underset{\sim}{V}} \in \Omega(1,4)$. In the case when $s=2$ and it assumed that $\underset{\sim}{V} \in \Omega(1,4)$ the single c-vector given in [9.4.7] remains optimal but the relative bias changes with ρ , as illustrated in table 9.4.3. The relative bias is never less than zero because of the constraints in the optimization problem.

Although the c-vector in [9.4.7] was optimal for the above choices of $\hat{\underset{\sim}{V}}$ it is not optimal for all variance matrices in $\Omega(1,4)$. For example, if

$$\hat{\underset{\sim}{V}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} ,$$

Table 9.4.3

Relative biases of best conservative estimator of standard errors of $\hat{\beta}_2$ for $\Omega(1,4)$ and the model specified in [9.4.1], when $\hat{V}$ has an AR(1) structure as specified in [9.4.13].

ρ	Relative bias
0.0	1.145
0.2	0.894
0.4	0.646
0.6	0.413
0.8	0.199
0.9	0.098
0.99	0.010

then the preceding c-vector has a relative bias of 18.0. The best choice of vector is

$$\sqrt{x_1}\,\underset{\sim}{c}^T = (0.3,\ -0.6,\ 0.0,\ 0.6,\ -0.3)$$

with a relative bias of 2.0 .

9.4.14 Optimal choice of $\hat{\underset{\sim}{\beta}}$

An interesting but unresolved question raised by the approach to conservative estimators developed in this chapter can be illustrated by example. If the true variance of $\underset{\sim}{y}$ is proportional to $\underset{\sim}{W}\,\underset{\sim}{R}\,\underset{\sim}{W}$ ($\underset{\sim}{I}$ in this example), then estimation of $\underset{\sim}{\beta}$ by the use of $\underset{\sim}{U}$ given in [9.2.1] provides the unbiased estimator with the minimum variance. However, it does not necessarily give the linear unbiased estimator with the minimum expected conservative estimator of variance for a particular choice of $\Omega\,(\omega,\upsilon)$. The question is: what choice of $\underset{\sim}{U}$ does minimize the expected conservative estimator of variance?

In the above example, when $s=2$ and $\underset{\sim}{V}=\underset{\sim}{I}$, $\hat{\beta}_2 = \underset{\sim}{e}^T\underset{\sim}{u}$ (with $\underset{\sim}{u}^T = (-0.2,\ -0.1,\ 0.0,\ 0.1,\ 0.2)$) is the best estimator of β_2 , with a variance of 0.1 . But if we assume only that $\underset{\sim}{V} \in \Omega\,(1,4)$, the expected value of the conservative estimator of variance is 0.46 if $\underset{\sim}{V}$ is equal to $\underset{\sim}{I}$. When β_2 is instead estimated using

$$\underset{\sim}{u}^T = (-2/7,\ \ 1/14,\ \ 0,\ \ -1/14,\ \ 2/7)$$
$$= (-0.2857,\ 0.0714,\ 0.0,\ -0.0714,\ 0.2857)$$

its variance is 0.1735 when $\underset{\sim}{V}$ is equal to $\underset{\sim}{I}$, but the best conservative estimator on the assumption that $\underset{\sim}{V} \in \Omega\,(1,4)$ is derived from a single vector:

$$\sqrt{x_1}\,\underset{\sim}{c}^T = (-2/7,\ 1/14,\ 3/7,\ 1/14,\ -2/7)$$
$$= (0.2857,\ 0.0714,\ 0.4286,\ 0.0714,\ -0.2857)$$

with an expected value of 0.3571 if $\underset{\sim}{V}$ is equal to $\underset{\sim}{I}$. Therefore, if no confidence can be placed in the conjucture that $\underset{\sim}{V}$ is equal to $\underset{\sim}{I}$, and a conservative estimator of the variance of $\hat{\beta}_2$ is to be used by assuming only that $\underset{\sim}{V} \in \Omega(1,4)$, then the second estimator above is preferable to the conventional first estimator.

9.5 Colquhoun's data

9.5.1 Choice of regression parameter estimators and $\hat{\underset{\sim}{V}}$

Both a single exponential regression and a double exponential regression model were fitted to Colquhoun's data by three methods described in section 4.3: least-squares estimation (LS); maximum likelihood estimation with an ARMA(1,1) error model (ML); and residual maximum likelihood estimation with an ARMA(1,1) error model (REML). The least-squares residuals were found to be incompatible with the assumed error variance matrix $\tau^2\underset{\sim}{I}$, so it does not seem appropriate to assume that $\hat{\underset{\sim}{V}}$ is equal to $\tau^2\underset{\sim}{I}$ when obtaining conservative estimators. Instead, ARMA(1,1) models were fitted to the least-squares residuals by both maximum likelihood (LS/ML) and residual maximum likelihood (LS/REML) estimation, thus giving four choices of estimator and $\hat{\underset{\sim}{V}}$. Strictly, the method in section 9.2 is not applicable when regression and variance parameters have been jointly estimated from the same data because dimensionality in the residuals will have been lost to the variance model as well as to the regression model. However, this has been ignored in what follows.

9.5.2 Choice of $\Omega(\omega,\upsilon)$

The smallest set of matrices to which $\hat{\underset{\sim}{V}}$ belongs in the Ω class is $\Omega(1,123)$. The fact that ω is equal to 1 follows directly from [9.2.15] by observing that $\hat{\underset{\sim}{V}}$ is the sum of two matrices with Markov structure and ω cannot exceed 1 because the correlations exceed 0.5. The value of υ is 123 because

$$\hat{V}_{124,1} > 0 ,$$

and in particular when the single exponential regression model is fitted by REML the estimated correlation between y_1 and y_{124} exceeds 0.999.

9.5.3 Single exponential regression results

Considering first the single exponential regression model, the variance of $\hat{\beta}_1$ is unbounded when $\underset{\sim}{V}$ belongs to $\Omega(1,123)$ for all three methods of estimation (LS, ML and REML) This is a repeat of the situation encountered when estimating a global mean in [9.4.12]. Conservative estimators of variances were found for both of the other regression parameters for all four combinations of estimators and $\hat{\underset{\sim}{V}}$. In all cases, the maximum CPU time of 30,000 seconds was reached before the best estimator could be found. Table 9.5.1 summarises the results.

The four pairs of marginal confidence intervals are equally valid, each being at least at the 95% level for all variance matrices in $\Omega(1,123)$. The expected standard errors derived from the conservative estimators of variance, and hence the expected widths of marginal confidence intervals, are very similar for each parameter over all four methods of estimation. It does not necessarily follow that because ML and REML estimators are more

Table 9.5.1

Colquhoun's data, single exponential regression with ARMA(1,1) error model estimated by four methods; summary results of using conservative estimators of the variances of the regression parameter estimators on the assumption that $\underset{\sim}{V} \in \Omega(1,123)$.

Estimation method	Parameter number	Parameter estimate	Model based s.e.	Expected s.e.	Relative bias	Aprox. d.f.	Estimated s.e.	Approximate 95% confidence limits	
LS/ML									
	2	77.4	1.0	2.2	1.16	3.3	3.9	65.5	89.4
	3	7.24	0.22	0.52	1.33	3.2	0.34	6.19	8.28
LS/REML									
	2	77.4	2.1	2.4	0.16	3.3	3.0	68.3	86.5
	3	7.24	0.43	0.60	0.38	3.5	0.23	6.56	7.91
ML									
	2	79.4	1.0	1.7	0.74	3.0	3.2	69.1	89.6
	3	6.77	0.18	0.47	1.61	3.7	0.63	4.94	8.60
REML									
	2	78.6	1.6	1.7	0.12	3.0	2.6	70.4	86.8
	3	6.59	0.23	0.53	1.29	3.6	0.82	4.21	8.97

LS/ML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by maximum likelihood from residuals

LS/REML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by residual maximum likelihood from residuals

ML regression parameters and ARMA(1,1) error variance matrix jointly estimated by maximum likelihood

REML regression parameters and ARMA(1,1) error variance matrix jointly estimated by residual maximum likelihood

efficient than LS estimators that they will give smaller conservative estimates of variances. (This relates to the optimal choice of $\hat{\underset{\sim}{\beta}}$ considered in [9.4.14].) The estimated standard errors show less consistency between methods of estimation; in part this is because a different value of $\hat{\underset{\sim}{V}}$ is being used in each case, and in part it results from sampling variability.

It is not possible to choose from among the estimators the one with the smallest estimated standard error without invalidating the 95% interpretation of the marginal confidence interval. If a choice is to be made between estimators then it should be on the basis of the expected standard error, the approximate degrees of freedom and the appropriateness of $\hat{\underset{\sim}{V}}$. The estimation methods ML and REML have slightly lower expected standard errors and about the same degrees of freedom as LS/ML and LS/REML, but REML is preferable in being based on a probably more realistic estimate of $\underset{\sim}{V}$. It has already been observed in chapter 4 that ML does seem to give downward biased estimates of $\underset{\sim}{V}$. This may account for the estimated standard errors of $\hat{\beta}_2$ for both LS/ML and ML exceeding what would have been their 95% limits if the assumed value of $\hat{\underset{\sim}{V}}$ had been correct.

9.5.4 Double exponential regression results

The variance of $\hat{\beta}_1$ is unbounded as in the single exponential regression case. Conservative estimators of variances were found for the other four parameters, although not ones with minimum relative bias because the maximum CPU time was reached during the iterations. Table 9.5.2 gives the best results found.

The expected standard errors for each parameter are similar for LS/ML, LS/REML and ML but larger values were obtained for

Table 9.5.2

Colquhoun's data, double exponential regression with ARMA(1,1) error model estimated by four methods; summary results of using conservative estimators of the variances of the regression parameter estimators on the assumption that $\underset{\sim}{V} e \Omega(1,123)$.

Estimation method	Parameter number	Parameter estimate	Model based s.e.	Expected s.e.	Relative bias	Aprox. d.f.	Estimated s.e.	Approximate 95% confidence limits	
LS/ML	2	25.3	7.7	26.3	2.43	6.3	36.8	-63.9	114.5
	3	3.28	0.68	2.27	2.32	7.1	2.77	-3.25	9.82
	4	58.6	7.9	27.0	2.43	6.9	33.0	-19.8	137.0
	5	9.22	0.82	2.83	2.45	6.2	3.50	0.70	17.73
LS/REML	2	25.3	12.7	33.0	1.60	5.8	35.0	-61.1	111.7
	3	3.28	1.07	2.92	1.71	6.6	3.12	-4.17	10.74
	4	58.6	12.9	34.4	1.67	6.0	36.3	-30.3	147.5
	5	9.22	1.36	3.57	1.62	5.4	3.30	0.92	17.51
ML	2	26.1	8.5	28.3	2.32	6.3	37.9	-65.7	117.8
	3	3.41	0.71	2.33	2.27	7.6	2.91	-3.38	10.19
	4	57.6	8.7	28.9	2.34	6.8	36.4	-29.0	144.2
	5	9.29	0.90	3.02	2.35	5.9	3.94	-0.40	18.98
REML	2	31.5	21.0	51.8	1.47	6.0	49.1	-88.8	151.8
	3	3.86	1.29	3.48	1.70	5.9	2.90	-3.28	11.00
	4	51.8	20.5	51.8	1.52	6.5	46.6	-60.2	163.8
	5	9.83	2.45	6.01	1.45	6.1	5.33	-3.20	22.85

LS/ML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by maximum likelihood from residuals

LS/REML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by residual maximum likelihood from residuals

ML regression parameters and ARMA(1,1) error variance matrix jointly estimated by maximum likelihood

REML regression parameters and ARMA(1,1) error variance matrix jointly estimated by residual maximum likelihood

REML. This may be due, in part, to elements in $\hat{\underset{\sim}{V}}$ estimated by REML being larger than for the other estimation methods. The estimated standard errors are all greater than expectation except for the REML ones. However, interpretation is hindered by the potentially large correlations between the conservative estimators.

As already stated, all four sets of marginal confidence intervals are valid. If a choice has to be made then, given the closeness of expected standard errors and degrees of freedom for LS/ML, LS/REML and ML, the LS/REML estimates seem to be preferable because of the ease of interpretation of least-squares estimators and the approximate unbiasedness of residual maximum likelihood estimates of $\underset{\sim}{V}$ from the least-squares residuals.

9.6 Dale's data

9.6.1 Choice of regression parameter estimators, $\hat{\underset{\sim}{V}}$ and $\Omega(\omega,\upsilon)$

As for Colquhoun's data in [9.5.1], there are four choices of estimator and $\hat{\underset{\sim}{V}}$ available from the models fitted in section 4.4. The smallest set of matrices to which $\hat{\underset{\sim}{V}}$ belongs is $\Omega(1,137)$.

9.6.2 Results

Conservative estimators of variances were found for all four parameters for all combinations of estimators and $\hat{\underset{\sim}{V}}$. The maximum CPU times were reached before minimum biased estimators were found. Table 9.6.1 summarises the best results found.

The expected relative biases are larger than those previously encountered and most of the final estimators have only one degree

Table 9.6.1

Dale's data, regression with ARMA(1,1) error model estimated by four methods; summary results of using conservative estimators of the variances of the regression parameter estimators on the assumption that $\underset{\sim}{V} \in \Omega(1,137)$.

Estimation method	Parameter number	Parameter estimate	Model based s.e.	Expected s.e.	Relative bias	Aprox. d.f.	Estimated s.e.	Approximate 95% confidence limits	
LS/ML	1	818.	42.	765.	17.21	1.0	1123.	-13204.	14841.
	2	0.0142	0.0017	0.0048	1.88	2.3	0.0064	-0.0101	0.0385
	3	702.	31.	1516.	47.30	1.0	1772.	-21750.	23154.
	4	0.00067	0.00005	0.00097	17.19	1.0	0.00125	-0.01486	0.01619
LS/REML	1	818.	151.	1202.	6.94	1.0	1326.	-15837.	17474.
	2	0.0142	0.0027	0.0066	1.49	1.6	0.0054	-0.0199	0.0483
	3	702.	284.	2290.	7.07	1.0	2647.	-32810.	34214.
	4	0.00067	0.00023	0.00152	5.71	1.0	0.00162	-0.01962	0.02096
ML	1	809.	36.	442.	11.44	1.0	108.	-506.	2125.
	2	0.0205	0.0015	0.0044	1.99	3.4	0.0079	-0.0028	0.0439
	3	785.	35.	940.	25.54	1.0	130.	-864.	2434.
	4	0.00081	0.00006	0.00069	11.12	1.0	0.00015	-0.00103	0.00265
REML	1	806.	63.	494.	6.85	1.0	68.	-26.	1637.
	2	0.0210	0.0016	0.0049	1.96	3.5	0.0083	-0.0034	0.0454
	3	816.	125.	1064.	7.55	1.0	254.	-2361.	3994.
	4	0.00084	0.00013	0.00078	5.19	1.0	0.00016	-0.00111	0.00279

LS/ML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by maximum likelihood from residuals

LS/REML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by residual maximum likelihood from residuals

ML regression parameters and ARMA(1,1) error variance matrix jointly estimated by maximum likelihood

REML regression parameters and ARMA(1,1) error variance matrix jointly estimated by residual maximum likelihood

of freedom. The unexplained exception is found for the second parameter. As a consequence, the marginal confidence intervals are much wider than those obtained in chapter 4. If a choice is to be made between estimators, then ML and REML have much smaller expected standard errors, and REML is preferable as it is based on a more reliable matrix $\hat{\underset{\sim}{V}}$.

9.7 Bruce's data

9.7.1 Choice of regression parameter estimators, $\hat{\underset{\sim}{V}}$ and $\Omega(\omega,\upsilon)$

As for Colquhoun's and Dale's data there are four choices of estimator and $\hat{\underset{\sim}{V}}$ from models fitted in section 4.5. The smallest set of matrices to which $\hat{\underset{\sim}{V}}$ belongs is $\Omega(1,199)$. However, this means that the dual optimization problem has 20100 variables whereas CEVOPE has an upper limit of 10000. The correlation coefficients between the errors in the models fitted to Bruce's data decay more rapidly than in either Colquhoun's or Dale's data, so that for example the REML estimated correlation between y_i and y_{i+30} is 0.05. It seems reasonable to assume that after 30 days have elapsed there should be no carry-over of errors. Therefore, it is assumed that the true error variance matrix belongs to $\Omega(1,30)$.

9.7.2 Results

Conservative estimators of variances were found for all four methods of estimation. Before minimum biased estimators could be found the maximum number of c-vectors (30) was reached in half the

cases and the maximum CPU time reached in the rest. It is possible that the large number of c-vectors and the large number of degrees of freedom in the final solution could be because the parameter υ is less than its maximum possible value of n-1. Table 9.7.1 summarises the best results found.

As already seen in section 4.5, there is little to choose between ML and REML in this example. In general, both have lower expected standard errors and larger degrees of freedom than either LS/ML or LS/REML so both appear to be acceptable.

9.8 Discussion

An impasse was reached in chapters 4 and 5. Empirical models of the error variances were there assumed and used to estimate regression parameters, but the final results were sensitive to the choice of model. Because no information was available for choosing $\underset{\sim}{V}$ except for that contained in the data, the problem appeared to be intractable. The method presented in this chapter has gone a long way towards overcoming the problem. When applied to real data the new technique provides marginal confidence intervals which have very general validity, although they are very wide in certain circumstances.

The method should be applicable to any situation where regression parameters and standard errors are to be estimated from data with serial correlation of unknown form except for some very general specification. Partly because of the novelty of the approach it has several unresolved problems. These are discussed in the category of future work in section 10.2.

Table 9.7.1

Bruce's data, regression with ARMA(1,1) error model estimated by four methods; summary results of using conservative estimators of the variances of the regression parameter estimators on the assumption that $\underset{\sim}{V} \in \Omega(1,30)$.

Estimation method	Parameter number	Parameter estimate	Model based s.e.	Expected s.e.	Relative bias	Aprox. d.f.	Estimated s.e.	Approximate 95% confidence limits	
LS/ML	1	0.75	0.10	0.27	1.64	3.1	0.23	0.04	1.46
	2	0.068	0.002	0.004	1.00	5.7	0.005	0.057	0.080
	3	1.6	0.2	0.5	1.36	1.4	0.2	-0.1	3.33
	4	0.57	0.05	0.21	3.33	2.1	0.18	-0.15	1.29
LS/REML	1	0.75	0.11	0.19	0.66	4.4	0.13	0.40	1.10
	2	0.068	0.002	0.006	1.26	5.4	0.004	0.059	0.078
	3	1.6	0.2	0.7	2.20	2.8	0.5	-0.1	3.3
	4	0.57	0.05	0.09	0.79	5.0	0.09	0.34	0.81
ML	1	0.61	0.07	0.18	1.55	4.7	0.15	0.23	0.99
	2	0.067	0.003	0.003	0.31	9.0	0.003	0.060	0.074
	3	1.3	0.2	0.4	1.42	7.0	0.5	0.1	2.5
	4	0.48	0.04	0.14	2.21	3.1	0.19	-0.12	1.08
REML	1	0.60	0.07	0.21	1.89	3.3	0.20	0.00	1.20
	2	0.067	0.003	0.004	0.31	9.3	0.004	0.058	0.076
	3	1.3	0.2	0.5	1.84	2.8	0.5	-0.3	3.0
	4	0.48	0.04	0.10	1.22	6.7	0.12	0.19	0.77

LS/ML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by maximum likelihood from residuals

LS/REML regression parameters estimated by least-squares, ARMA(1,1) error variance matrix estimated by residual maximum likelihood from residuals

ML regression parameters and ARMA(1,1) error variance matrix jointly estimated by maximum likelihood

REML regression parameters and ARMA(1,1) error variance matrix jointly estimated by residual maximum likelihood

10. Conclusions/Future work

10.1 Conclusions

10.1.1 General

Ordinary least-squares regression parameter estimators are inefficient and estimated standard errors are biased when errors are correlated. If an incorrect variance structure is assumed and used in the estimation procedure, then the new estimators may be even less efficient and the estimated standard errors more biased. Moreover, regression parameters change their meaning in the presence of a correlated errors model. This is because it is the conjunction of a regression model and an error model which describe a data set, so if the error model is changed the regression model has also to change in order to compensate. For example, if errors are assumed to be highly correlated, then the fitted regression model will exhibit much larger systematic departures from the data than if the errors had been assumed to be independent. Therefore, although within the last few years it has become computationally easy to estimate regression parameters with almost any choice of variance matrix, it is statistically difficult and fraught with dangers. In a word, contrary to most advice in the literature, beware!

10.1.2 Empirical models

It was found in the research reported in this thesis to be easier to model empirically the correlation coefficients of an error process than to model the spectral function. The recommended procedure is:

a) Estimate the regression parameters by (possibly weighted) least-squares.

b) Examine the residuals for autocorrelation, and if these are inconsistent with the assumption of independent errors identify a low-order ARMA model.

c) Jointly estimate the regression and error parameters by residual maximum likelihood using the variance matrix of the assumed ARMA model.

d) Compare the results with the least-squares estimates, if they are inconsistent return to the least-squares results and obtain variances of the estimators by some other means, such as using the method described in chapter 9 to derive conservative estimates of the variances valid over wide classes of possible error variance structures.

10.1.3 Mechanistic model

If the assumed model is (almost) correct, then the use of a mechanistic model makes for the most efficient use of the data, but if not then the results can be very misleading. The recommended procedure is always to include an observation error, always to test the goodness-of-fit by refitting the model with separate sets of parameters for the regression and error components, and never to use a model without a sound scientific basis.

10.1.4 Conservative estimates of the variances of regression parameter estimators

The technique described in chapter 9 makes heavy use of computer CPU time but light use of explicit statistical models.

Estimates of the variances of regression parameter estimators are obtained which are valid when errors are serially correlated in an unknown manner. The price that sometimes has to be paid for this robustness is the large upward biases in the estimates of the variances of regression parameter estimators. The present method may still be some way from being of routine use but it has potential for further development, as discussed in [10.2.3].

10.2 Future work

10.2.1 Empirical models

The comparison between maximum likelihood and residual maximum likelihood estimates of regression parameters when errors are correlated warrants further study beyond the small-scale simulations reported in chapter 4, particularly when the process is close to non-stationarity.

Empirically modelling the spectral function of the error process is a more recent technique than modelling the correlation coefficients in order to estimate regression parameters, and seems to offer more potential for future work. It ought to extend more easily to handle large numbers of regression parameters such as treatment effects in designed experiments, and to higher-dimensional correlation structures such as spatial processes. Another possibility is to explore the relationship between the parameterized spectrum approach in this thesis and the non-parametric approach of Duncan and Jones (1966).

10.2.2 Mechanistic models

I expect many more models in future literature to be mechanistic rather than empirical, although how mechanistic most of them will be is open to question. One particular gap of which I am aware is the lack of any successful attempt to introduce stochasticity into non-linear growth equations.

It is possible to generalize the difference and differential equations considered in chapter 6, and the compartment models considered in chapter 7, by allowing the parameters to vary with time. In fact, the model considered in chapter 8 may be thought of as a single compartment with a time-varying rate of emigration. Whether or not the solutions are GARMA and cGARMA processes has yet to be resolved, although it would seem that the solution of a p-compartment model is a cGARMA(p^2,0) process, rather than a cGARMA(p,0) process as it is with time-constant parameters.

10.2.3 Conservative estimates of the variances of regression parameter estimators

The area of conservative estimates of the variances of regression parameter estimators is the one in which I am most conscious of the potential for further work. The computer CPU time used by the program CEVOPE should be reducible through better linear programming and eigenvector algorithms and through special algorithms for specific choices of Ω such as $\Omega(0,0)$ and $\Omega(1,\upsilon)$. Also, improved algorithms are required to cope with rounding errors, especially as the optimum solution is approached.

Further work could be put into other choices of Ω, for example more restrictive assumptions to reduce the variance bias in

some applications, and to other types of correlation structure such as might arise in spatial data or two-way tables.

More fundamentally, methods are required to incorporate the variance of the variance estimator into the optimization problem (see [9.2.18]) and to find the regression parameter estimator which minimizes the expected value of the conservative estimator of variance (see [9.4.14]).

Appendix A

REGAME computer program listing (see section 2.4).

```
0001    C*************************************************************************C
0002    C*************************************************************************C
0003    C**                                                                     **C
0004    C**     RRRRRR    EEEEEEE    GGGGG      AAA     M     M   EEEEEEE    **C
0005    C**     R     R   E         G     G    A   A    MM   MM   E          **C
0006    C**     R     R   E         G         A     A   M M M M   E          **C
0007    C**     RRRRRR    EEEE      G  GGGG   AAAAAAA   M  M  M   EEEE       **C
0008    C**     R   R     E         G     G   A     A   M     M   E          **C
0009    C**     R    R    E         G     G   A     A   M     M   E          **C
0010    C**     R     R   EEEEEEE    GGGGG    A     A   M     M   EEEEEEE    **C
0011    C**                                                                     **C
0012    C**                                                                     **C
0013    C**           REGRESSION ESTIMATION WITH GENERALISED                    **C
0014    C**            AUTOREGRESSIVE-MOVING AVERAGE ERRORS                     **C
0015    C**                                                                     **C
0016    C**                        C.A.GLASBEY                                  **C
0017    C**                  A.F.R.C. UNIT OF STATISTICS                        **C
0018    C**                    UNIVERSITY OF EDINBURGH                          **C
0019    C**                                                                     **C
0020    C*************************************************************************C
0021    C*************************************************************************C
0022    C
0023          SUBROUTINE REGAME(YT,NLT,ACCTOL,NCALMX,NLIKT,NSCALT,NMT,NDERT,
0024         1                  NVAR,NPARVR,NOUTT,NCVI,PARAM,NPART,SE,XLIK,
0025         1                  ZT,PARVAR,IFAIL)
0026    C
0027    C
0028    C-------------------------------------------------------------------------
0029    C                        REGAME                                         !
0030    C                        ======                                         !
0031    C                                                                       !
0032    C         MAIN CONTROLLING SUBROUTINE                                   !
0033    C                                                                       !
0034    C  INPUT                                                                !
0035    C   YT(NLT)   =Y(NL)       DATA VECTOR OF OBSERVATIONS                  !
0036    C   NLT       =NL          NUMBER OF OBSERVATIONS (BETWEEN 1 AND 200)   !
0037    C   ACCTOL                 ACCURACY TO WHICH PARAMETERS REQUIRED (>0.0) !
0038    C   NCALMX                 MAXIMUM NUMBER OF FUNCTION EVALS. IN E04JBF  !
0039    C   NLIKT     =NLIK        CHOICE OF OPTIMIZATION CRITERION             !
0040    C                          1  MAXIMUM LIKELIHOOD (GAUSSIAN LIKELIHOOD)  !
0041    C                          2  RESTRICTED MAXIMUM LIKELIHOOD             !
0042    C                          3  SUM OF SQUARES (OR CHI-SQUARED)           !
0043    C   NSCALT    =NSCALE      1  IF SCALE PARAMETER TO BE ESTIMATED         !
0044    C   NMT       =NM          NUMBER OF REGRESSION PARAMETERS, USED WHEN   !
0045    C                             NLIK=2 OR NCVI=1 (BETWEEN 0 AND 5)        !
0046    C   NDERT     =NDER        1  IF THE DERIVATIVE SUBROUTINE REGDER       !
0047    C                             SUPPLIED, CALLED WHEN NLIK=2 OR NCVI=1    !
0048    C   NVAR                   1  IF PARAMETER VARIANCES TO BE ESTIMATED BY  !
0049    C                             DIFFERENCE METHODS, OTHERWISE THE         !
0050    C                             APPROX. IN E04JBF WILL BE USED            !
0051    C   NPARVR                  DIMENSIONS OF PARVAR                         !
0052    C                             ( >= NPAR, NPAR+1 IF NSCALE=1)            !
0053    C   NOUTT     =NOUT         CONTROLS OUTPUT                              !
0054    C                          0  ITERATIONS TO NOUT6 (WIDTH 80 CHAR.)      !
0055    C                          1          AND NOUT10 (WIDTH 120 CHAR.)      !
0056    C                          2  FINAL RESULTS TO NOUT10                   !
0057    C                          3       AND PLOTS                            !
0058    C                          4  INTERMEDIATE RESULTS IN REGAME, REGVAR     !
0059    C                          5                AND IN REGLIK, REGAMD       !
0060    C   NCVI                   1  CREATES OUTPUT TO NOUT8 AND NOUT9 FOR     !
0061    C                             INPUT TO PROGRAM CEVOPE                   !
0062    C                                                                       !
0063    C  INPUT/OUTPUT                                                         !
0064    C   PARAM(16)              PARAMETER STARTING VALUES / ESTIMATES        !
```

```
0065     C   NPART    =NPAR        NUMBER OF PARAMETERS (BETWEEN 1 AND 15)
0066     C                             / INCREASED BY 1 IF NSCALE=1
0067     C   SE(16)                 APPROX. STANDARD ERRORS (>=0.0), IF 0.0 THEN
0068     C                             THAT PARAMETER HELD CONSTANT
0069     C                             / PARAMETER STANDARD ERRORS
0070     C
0071     C  OUTPUT
0072     C   XLIK                   FINAL VALUE OF OPTIMIZATION FUNCTION
0073     C   ZT(NLT)  =Z(NL)        VECTOR OF RESIDUALS FROM MODEL
0074     C   PARVAR(NPARVR,NPARVR) PARAMETER VARIANCES, CORRELS. BELOW DIAG.
0075     C   IFAIL                  FAILURE INDICATOR
0076     C                            0  MODEL FITTING SUCCESSFUL
0077     C                            1  ONE OF PARAMETERS OUT OF RANGE
0078     C                            2  TOO MUCH ROUNDING ERROR FOUND BY E04HBF
0079     C                                 FOR STARTING PARAM TO BE USEABLE
0080     C                            3  NCALMX EXCEEDED
0081     C                            4  NOT ALL CONDITIONS FOR OPTIMUM SATISFIED
0082     C                                BUT NO BETTER POINT FOUND
0083     C                            5  OPTIMUM NOT FOUND IN LOCAL SEARCH,
0084     C                                PROBLEM ILL-CONDITIONED
0085     C                           +10  VARIANCES ONLY FOUND APPROXIMATELY
0086     C
0087     C-------------------------------------------------------------------
0088     C
0089     C              MODEL DEPENDENT SUBROUTINES
0090     C
0091     C-------------------------------------------------------------------
0092     C                          REGMOD
0093     C                          ======
0094     C        SUBROUTINE SPECIFYING F AND VAR FOR PARTICULAR
0095     C        PARAMETER VALUES.
0096     C
0097     C  INPUT
0098     C   PARAM(NPAR)           VECTOR OF MODEL PARAMETERS
0099     C   NPAR                  NUMBER OF MODEL PARAMETERS
0100     C   NLMAX                 FIRST DIMENSION OF VAR
0101     C   NL                    NUMBER OF OBSERVATIONS
0102     C   NFDIF                 1  IF F IS REQUIRED, BUT NOT VAR
0103     C
0104     C  OUTPUT
0105     C   F(NL)                 VECTOR OF FITTED VALUES
0106     C   VAR(NLMAX,11)         BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
0107     C   NPQA                  NP+NQ+1
0108     C   NP                    GENERALIZED AUTOREGRESSIVE ORDER (<=5)
0109     C   NQ                    GENERALIZED MOVING AVERAGE ORDER (<=5)
0110     C   NTEQ                  1  IF ALL ROWS OF PHI ARE EQUAL IN
0111     C                             DECOMPOSITION OF VAR
0112     C
0113     C  CONTROL
0114     C   NOUT                  CONTROLS OUTPUT
0115     C   IFAIL                 1  IF F OR VAR CANNOT BE EVALUATED AT PARAM
0116     C
0117     C-------------------------------------------------------------------
0118     C
0119     C                          REGDER
0120     C                          ======
0121     C        SUBROUTINE SUPPLYING DERIV FOR PARTICULAR PARAMETER VALUES
0122     C        IF NDER=1 AND EITHER NLIK=2 OR NCVI=1
0123     C
0124     C  INPUT
0125     C   PARAM(NPAR)           VECTOR OF MODEL PARAMETERS
0126     C   NPAR                  NUMBER OF MODEL PARAMETERS
0127     C   NLMAX                 FIRST DIMENSION OF DERIV
0128     C   NL                    NUMBER OF OBSERVATIONS
```

```
0129     C   NM                    SECOND DIMENSION OF DERIV                    !
0130     C                                                                      !
0131     C  OUTPUT                                                              !
0132     C   DERIV(NLMAX,NM)       MATRIX OF FIRST DERIVATIVES OF F W.R.T. PARAM !
0133     C                                                                      !
0134     C  CONTROL                                                             !
0135     C   NOUT                  CONTROLS OUTPUT                              !
0136     C   IFAIL                 1  IF DERIV CANNOT BE EVALUATED AT PARAM     !
0137     C                                                                      !
0138     C______________________________________________________________________!
0139     C
0140     C                        NAG ROUTINES
0141     C
0142     C______________________________________________________________________!
0143     C                          E04HBF                                      !
0144     C                          E04JBF                                      !
0145     C                          E04JBQ                                      !
0146     C                          F01ABF                                      !
0147     C                          F03ABF                                      !
0148     C                          F03AGF                                      !
0149     C                          F04ATF                                      !
0150     C                          F04JGF                                      !
0151     C                          G05CBF                                      !
0152     C                          G05DDF                                      !
0153     C                          X02AAF                                      !
0154     C______________________________________________________________________!
0155     C
0156     C                  AUXILIARY NAG ROUTINES
0157     C
0158     C______________________________________________________________________!
0159     C        E04JBM  E04KBV  E04KBY  E04JBN  E04JBP                        !
0160     C        E04JBR  E04JBS  E04JBT  E04JBU  E04JBV                        !
0161     C        E04JBW  E04JBX  E04JBY  E04JBZ  E04KBQ                        !
0162     C        E04KBR  E04KBS  E04KBT  E04KBU  E04KBW                        !
0163     C        E04KBX  E04KBZ  E04LBT  E04LBU  E04LBV                        !
0164     C        E04LBW  E04LBX  E04LBY  E04LBZ  E04ABZ                        !
0165     C        F04AHF  F04AQF  F03AEF  F03AFF  F04AJF                        !
0166     C        F01ACF  F01ADF  F01DEF  F04JGZ  F02WDF                        !
0167     C        F02WDZ  F01QAF  F01QAY  F02WAZ  F04JAY                        !
0168     C        F04JGR  F04JGW  F01LZF  F01QAZ  F02SZF                        !
0169     C        F02WAY  F02WCW  F02WCY  F02WDY  F04JGS                        !
0170     C        F04JGV  F04JGY  F01LZW  F01LZX  F01LZY                        !
0171     C        F01LZZ  F01QAW  F01QAX  F02SZZ  F04JGT                        !
0172     C        F04JGU  F04JGX  X02ACF  X02AGF  X02BCF                        !
0173     C        X02BDF  X02DAF  X03AAF  P01AAF  X04AAF                        !
0174     C        Y13AAF  Y13ACF  Y13AEF  Y13AFF  Y13AGF                        !
0175     C        Y13ADF  Y13ABF  G05CAF  G05CCF  G05CGZ                        !
0176     C        G05CFZ  BG05CC                                                !
0177     C______________________________________________________________________!
0178     C
0179     C                 AUXILIARY OUTPUT ROUTINES
0180     C
0181     C______________________________________________________________________!
0182     C                         /NINOUT/                                     !
0183     C                          MATPN1                                      !
0184     C                          MATPRO                                      !
0185     C                          MATPR1                                      !
0186     C                          MATPR2                                      !
0187     C                          PLTACR                                      !
0188     C                          PLTLNG                                      !
0189     C______________________________________________________________________!
0190     C
0191     C        LIST OF SUBROUTINES AND ROUTINES CALLED BY THEM,
0192     C        ARRANGED HIERARCHICALLY.
```

```
0193     C
0194     C                                                    MODEL      NAG       OTHER
0195     C                                                    ======     ======    ======
0196     C  1  REGAME:    E04HBF
0197     C     =======    E04JBF
0198     C                            REGLIK
0199     C                 REGVAR
0200     C                 REGOUT
0201     C                 REGCVI
0202     C                                                                         MATPN1
0203     C                                                                         MATPRO
0204     C                                                                         MATPR1
0205     C                 .          .          .          .          .          .
0206     C
0207     C                 E04HBF:    REGLIK
0208     C                 =======
0209     C
0210     C                 E04JBF:    REGLIK
0211     C                 =======    REGMON
0212     C                                                              E04JBQ
0213     C                 .          .          .          .          .          .
0214     C
0215     C  2                         REGLIK:    REGAMD
0216     C                            =======    REGTRA
0217     C                                                    REGMOD
0218     C                                                    REGDER
0219     C                                                              F03ABF
0220     C                                                                         MATPRO
0221     C                                                                         MATPR1
0222     C                                                                         MATPR2
0223     C                 .          .          .          .          .          .
0224     C
0225     C  3                         REGMON:
0226     C                            =======
0227     C                 .          .          .          .          .          .
0228     C
0229     C  4              REGVAR:    REGLIK
0230     C                 =======               REGTRA
0231     C                (REGVRF)                                      F01ABF
0232     C                                                                         MATPR2
0233     C                 .          .          .          .          .          .
0234     C
0235     C  5              REGOUT:                                                 PLTACR
0236     C                 =======                                                 PLTLNG
0237     C                 .          .          .          .          .          .
0238     C
0239     C  6              REGCVI:               REGAMD
0240     C                 =======               REGTRA
0241     C                                                    REGMOD
0242     C                                                    REGDER
0243     C                                                              F01ABF
0244     C                                                                         MATPR2
0245     C                 .          .          .          .          .          .
0246     C
0247     C  7                                    REGAMD:                F03AGF
0248     C                                       =======                F04ATF
0249     C                                                              F04JGF
0250     C                                                              X02AAF
0251     C                                                                         MATPRO
0252     C                                                                         MATPR1
0253     C                                                                         MATPR2
0254     C                 .          .          .          .          .          .
0255     C
0256     C  8                                    REGTRA:                           MATPR1
```

```
0257     C                                        =======
0258     C                .         .         .         .         .         .
0259     C
0260     C  9  REGGEN:                          REGAMD
0261     C     =======                          REGTRA
0262     C                                                 REGMOD
0263     C                                                          GO5CBF
0264     C                                                          GO5DDF
0265     C
0266     C
0267     C-----------------------------------------------------------------------
0268     C        THE TYPES OF ALL VARIABLES ARE DECLARED.
0269     C        VARIABLES BEGINNING WITH A-H OR P-Z ARE DOUBLE PRECISION REALS
0270     C        THE REST ARE INTEGERS EXCEPT FOR OCCASIONAL LOGICAL VARIABLES
0271     C        BEGINNING WITH THE LETTER L.
0272     C        THE ORDER OF DECLARATION IS:
0273     C                        VARIABLES TRANSFERRED IN PARAMETER LISTS,
0274     C                        VARIABLES IN COMMON,
0275     C                        LOCAL VARIABLES IN ALPHABETICAL ORDER.
0276     C
0277           DOUBLE PRECISION
0278          1    YT(NLT),ACCTOL,PARAM(16),SE(16),XLIK,ZT(NLT),
0279          1 PARVAR(NPARVR,NPARVR),
0280          1    PARAMM(16),PARAMS(16),XLIKM,XLIKS,STEPMX,SCALE,
0281          1 Y(200),F(200),Z(200),VAR(200,11),THETA(200,6),
0282          1    BOUNDL(15),BOUNDU(15),EPS(15),ETA,HESD(15),HESL(105),
0283          1 PARAMT(16),RSCALE,WK(135),XLIKD(16),XLIKT,XLKMIN,XX
0284           INTEGER
0285          1    NLT,NCALMX,NLIKT,NSCALT,NMT,NDERT,NVAR,NPARVR,NOUTT,NCVI,
0286          1 NPART,IFAIL,
0287          1    MLEST(16),NPAR,NOUT,NSCALE,NSTAGE,
0288          1 NL,NPQA,NPA,NQA,NP,NQ,NLIK,NM,NDER,NLMAX,NPMAX,NQMAX,
0289          1    I,IFAILT,J,JZ,K,KU,MCONST(15),MWK(2),NBOUND,NCALL,
0290          1 NCALMT,NFLAG,NHESL,NMMAX,NMONIT,NMWK,NN,NPARMX,NPEST,
0291          1 NPRI,NSTYPE,NWK
0292           LOGICAL
0293          1    LOGJBF
0294           COMMON
0295          1 /REGCM1/ PARAMM,PARAMS,XLIKM,XLIKS,STEPMX,MLEST,NPAR,NOUT
0296          1 /REGCM2/ SCALE,NSCALE,NSTAGE
0297          1 /REGCM3/ Y,F,Z,VAR,THETA,NL,NPQA,NPA,NQA,NP,NQ,
0298          1          NLIK,NM,NDER,NLMAX,NPMAX,NQMAX
0299          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0300     C
0301     C-----------------------------------------------------------------------
0302     C
0303     C  /REGCM1/ SHARED BY REGAME, REGLIK, REGMON, REGVAR
0304     C
0305     C   PARAMM(16)           LOCATION SHIFT TRANSFORMING PARAM TO PARAMT
0306     C   PARAMS(16)           SCALE SHIFT TRANSFORMING PARAM TO PARAMT
0307     C   XLIKM                LOCATION SHIFT TRANSFORMING XLIK TO XLIKT
0308     C   XLIKS                SCALE SHIFT TRANSFORMING XLIK TO XLIKT
0309     C   STEPMX               MAX. CHANGE OF PARAMT IN ONE STEP BY E04JBF
0310     C   MLEST(16)            LOCATION OF PARAMETERS TO BE ESTIMATED
0311     C   NPAR                 NUMBER OF MODEL PARAMETERS
0312     C   NOUT                 CONTROLS OUTPUT
0313     C
0314     C  /REGCM2/ SHARED BY REGAME, REGLIK, REGVAR
0315     C
0316     C   SCALE                SCALE PARAMETER
0317     C   NSCALE               1  IF SCALE PARAMETER TO BE ESTIMATED
0318     C   NSTAGE               1  FOR ESTIMATION STAGE
0319     C                        2  FOR CALCULATION OF PARVAR STAGE
0320     C
```

```
C  /REGCM3/ SHARED BY REGAME, REGLIK
C
C   Y(200)              DATA VECTOR OF OBSERVATIONS
C   F(200)              VECTOR OF FITTED VALUES
C   Z(200)              VECTOR OF RESIDUALS FROM MODEL
C   VAR(200,11)         BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
C   THETA(200,6)        SECOND MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
C   NL                  NUMBER OF OBSERVATIONS
C   NPQA                NUMBER OF COLUMNS IN VAR (NP+NQ+1)
C   NPA                 NUMBER OF COLUMNS IN PHI (NP+1)
C   NQA                 NUMBER OF COLUMNS IN THETA (NQ+1)
C   NP                  GENERALIZED AUTOREGRESSIVE ORDER
C   NQ                  GENERALIZED MOVING AVERAGE ORDER
C   NLIK                OPTIMIZATION CRITERION
C   NM                  NUMBER OF REGRESSION PARAMETERS, AND SECOND
C                          DIMENSION OF DERIV (WHEN NLIK=2 OR NCVI=1)
C   NDER                1  IF DERIVATIVE SUBROUTINE REGDER AVAILABLE
C   NLMAX               MAXIMUM SIZE ALLOWED FOR NL (=200)
C   NPMAX               MAXIMUM SIZE ALLOWED FOR NP (=5)
C   NQMAX               MAXIMUM SIZE ALLOWED FOR NQ (=5)
C
C
C-----------------------------------------------------------------------
C  LOCAL
C   BOUNDL(15)          LOWER BOUNDS ON PARAMT IN E04JBF
C   BOUNDU(15)          UPPER BOUNDS ON PARAMT IN E04JBF
C   EPS(15)             STEP LENGTH TO CALC. DERIVS. IN E04JBF
C   ETA                 ACCURACY OF 1-D SEARCH IN E04JBF
C   HESD(15)            DIAG. OF DECOMPOSITION OF SECOND DERIVS.
C   HESL(105)           LOWER TRIANGLE OF DECOMP. OF SECOND DERIVS.
C   I                   LENGTH INDEX
C   IFAILT              FAILURE INDICATOR IN REGVAR
C   J                   INDEX
C   JZ                  NPAR+1-J
C   K                   INDEX
C   KU                  UPPER BOUND ON K DO-LOOP
C   LOGJBF              DUMMY VARIABLE (LOGICAL)
C   MCONST(15)          DENOTES CONSTRAINED PARAMETERS IN E04JBF
C   MWK(2)              WORK SPACE IN E04HBF, E04JBF
C   NBOUND              SPECIFIES TYPE OF BOUNDARIES IN E04JBF
C   NCALL               NUMBER OF CALLS TO REGLIK BY E04HBF
C   NCALMT              MAX. CALLS TO REGLIK FROM E04JBF,
C                          HALVED IF IFAIL RETURNED SET TO 4
C   NFLAG               DUMMY USED IN CALLING REGLIK
C   NHESL               DIMENSION OF HESL
C   NMMAX               MAXIMUM OF NM
C   NMONIT              SPECIFIES FREQUENCY OF CALLS TO REGMON
C   NMWK                DIMENSION OF MWK
C   NN                  NUMBER OF ELEMENTS USED IN HESL
C   NPARMX              MAXIMUM OF NPAR
C   NPEST               NUMBER OF PARAMETERS TO BE ESTIMATED
C   NPRI                1  IF INTERMEDIATE RESULTS OUTPUT
C   NSTYPE              SPECIFIES START CONDITIONS IN E04JBF
C   NWK                 DIMENSION OF WK
C   PARAMT(16)          RESCALED PARAM
C   RSCALE              SQUARE-ROOT OF SCALE
C   WK(135)             WORK SPACE IN E04HBF, E04JBF
C   XLIKD(16)           FIRST DERIVATIVE OF XLIKT W.R.T. PARAMT
C   XLIKT               RESCALED OPTIMIZATION FUNCTION
C   XLKMIN              LOWER BOUND ON XLIKT USED BY E04JBF
C   XX                  WORK SPACE
C
C
C-----------------------------------------------------------------------
```

```
0385            EXTERNAL REGLIK,REGMON,E04JBQ
0386            DATA NMMAX,NPARMX,NHESL,NMWK,NWK,LOGJBF
0387           1    /    5,    15,  105,    2, 135, .TRUE./
0388            NOUT=NOUTT
0389            WRITE (NOUT6,938)
0390            IF (NOUT.GE.1) WRITE (NOUT10,942)
0391            NLMAX=200
0392            NPMAX=5
0393            NQMAX=5
0394      C
0395      C        CHECK PARAMETERS ARE IN RANGE
0396      C
0397            NL=NLT
0398            NLIK=NLIKT
0399            NSCALE=NSCALT
0400            NM=NMT
0401            NDER=NDERT
0402            NPAR=NPART
0403            IF ((NL.LT.1).OR.(NL.GT.NLMAX)) GO TO 017
0404            IF (ACCTOL.LE.0.0D0) GO TO 019
0405            IF ((NLIK.LT.1).OR.(NLIK.GT.3)) GO TO 021
0406            IF ((NM.LT.0).OR.(NM.GT.NMMAX)) GO TO 023
0407            K=NPAR
0408            IF (NSCALE.EQ.1) K=NPAR+1
0409            IF (NPARVR.LT.K) GO TO 025
0410            IF ((NPAR.LT.1).OR.(NPAR.GT.NPARMX)) GO TO 027
0411            DO 001 I=1,NL
0412.01   001    Y(I)=YT(I)
0413.01   C
0414.01   C        SET PARAMETERS TO BE ESTIMATED
0415.01   C
0416            K=0
0417            DO 002 J=1,NPAR
0418.01         IF (SE(J).LT.-1.0D-12) GO TO 029
0419.01         PARAMM(J)=PARAM(J)
0420.01         PARAMS(J)=SE(J)
0421.01         IF (SE(J).LT.1.0D-12) GO TO 002
0422.01         K=K+1
0423.01         MLEST(K)=J
0424.01         PARAMT(K)=0.0D0
0425.01   002   CONTINUE
0426            NPEST=K
0427            K=K+1
0428            MLEST(K)=0
0429            IF (NPEST.LT.1) GO TO 032
0430      C
0431      C        OUTPUT HEADINGS
0432      C
0433            WRITE (NOUT6,939)
0434            IF (NOUT.LT.1) GO TO 003
0435            WRITE (NOUT10,943)
0436            WRITE (NOUT10,944)
0437            WRITE (NOUT10,945) NCALMX,NLIK,NSCALE,NM,NDER,ACCTOL,NVAR,NPEST,
0438           1 (PARAMS(I),I=1,NPAR)
0439            WRITE (NOUT10,946)
0440      C
0441      C        INITIALISE PARAMETERS FOR E04JBF
0442      C
0443      003   NPRI=0
0444            IF (NOUT.GE.4) NPRI=1
0445            NSTAGE=1
0446            XLIKM=0.0D0
0447            XLIKS=FLOAT(NPEST)
0448            CALL REGLIK(NFLAG,NPEST,PARAMT,XLIKT,XLIKD,MWK,NMWK,WK,NWK)
```

```
0449           IF (XLIKT.GT.9.9D19) GO TO 031
0450           XLIKM=XLIKT*XLIKS-XLIKS
0451     C
0452           IFAIL=1
0453           CALL E04HBF(NPEST,REGLIK,PARAMT,NCALL,EPS,HESL,NHESL,HESD,
0454          1            XLIKT,XLIKD,MWK,NMWK,WK,NWK,IFAIL)
0455           IF (NPRI.NE.1) GO TO 004
0456           CALL MATPR1(EPS,NPEST,1)
0457           CALL MATPR1(HESD,NPEST,1)
0458           CALL MATPRO(XLIKT,1)
0459           CALL MATPR1(XLIKD,NPEST,1)
0460     004   IF (IFAIL-1) 007,034,005
0461     005   WRITE (NOUT6,906)
0462           IF (NOUT.GE.1) WRITE (NOUT10,906)
0463     906   FORMAT (' *** REGAME: E04JBF CANNOT START FROM THIS POINT ***')
0464           GO TO 012
0465     C
0466     007   IF (NCALMX.GT.0) GO TO 008
0467           IFAIL=3
0468           GO TO 012
0469     008   NCALMT=NCALMX
0470           NMONIT=1
0471           NSTYPE=1
0472           ETA=0.5D0
0473           IF (NPEST.EQ.1) ETA=0.0D0
0474           STEPMX=1.0D0
0475           XLKMIN=-1.0D0
0476           NBOUND=1
0477           IFAIL=1
0478     C
0479     C        OPTIMIZATION USING E04JBF
0480     C
0481     009   CALL E04JBF(NPEST,REGLIK,REGMON,NMONIT,LOGJBF,NSTYPE,E04JBQ,
0482          1            NCALMT,ETA,ACCTOL,STEPMX,XLKMIN,EPS,NBOUND,BOUNDL,
0483          1            BOUNDU,PARAMT,HESL,NHESL,HESD,MCONST,XLIKT,
0484          1            XLIKD,MWK,NMWK,WK,NWK,IFAIL)
0485           IF (IFAIL.EQ.1) GO TO 036
0486           IF (IFAIL.NE.4) GO TO 011
0487           WRITE (NOUT6,910)
0488           IF (NOUT.GE.1) WRITE (NOUT10,910)
0489     910   FORMAT (' *** REGAME: E04JBF FAILED BUT BEING RESTARTED ***')
0490           NSTYPE=0
0491           NCALMT=NCALMT/2
0492           GO TO 009
0493     011   CALL REGLIK(NFLAG,NPEST,PARAMT,XLIKT,XLIKD,MWK,NMWK,WK,NWK)
0494           IF (IFAIL.EQ.3) IFAIL=4
0495           IF (IFAIL.EQ.2) IFAIL=3
0496           IF (NPRI.NE.1) GO TO 012
0497           NN=((NPEST*(NPEST-1)))/2
0498           CALL MATPR1(HESL,NN,1)
0499           CALL MATPR1(HESD,NPEST,1)
0500           CALL MATPN1(MCONST,NPEST,1)
0501           CALL MATPR1(XLIKD,NPEST,1)
0502     C
0503     C        CALCULATE PARAMETER VARIANCES AND RESULTS FOR OUTPUT
0504     C
0505     012   CALL REGVAR(NVAR,HESL,NHESL,HESD,ACCTOL,NPEST,XLIKT,NL,NPARVR,
0506          1            PARAMT,PARVAR,SE,MWK,NMWK,WK,NWK,NPRI,IFAILT)
0507           IF (IFAILT.NE.0) IFAIL=IFAIL+10
0508           IF (NSCALE.EQ.1) XLIKD(NPAR)=0.0D0
0509           K=NPEST
0510           DO 013 JZ=1,NPAR
0511.01        XX=0.0D0
0512.01        J=NPAR+1-JZ
```

```
0513.01          IF (K.EQ.0) GO TO 013
0514.01          IF (MLEST(K).NE.J) GO TO 013
0515.01          XX=XLIKD(K)
0516.01          K=K-1
0517.01   013    XLIKD(J)=XX
0518             NPART=NPAR
0519      C
0520             XLIK=XLIKT*XLIKS+XLIKM
0521             K=1
0522             DO 014 J=1,NPAR
0523.01          XX=0.0D0
0524.01          IF (MLEST(K).NE.J) GO TO 014
0525.01          XX=PARAMT(K)
0526.01          K=K+1
0527.01   014    PARAM(J)=XX*PARAMS(J)+PARAMM(J)
0528.01   C
0529             RSCALE=DSQRT(SCALE)
0530             DO 016 I=1,NL
0531.01          Z(I)=Z(I)/RSCALE
0532.01          ZT(I)=Z(I)
0533.01          DO 015 J=1,NQA
0534.02   015    THETA(I,J)=THETA(I,J)*RSCALE
0535.01          DO 016 J=1,NPQA
0536.02   016    VAR(I,J)=VAR(I,J)*SCALE
0537.02   C
0538.02   C        OUTPUT OF RESULTS
0539.02   C
0540             WRITE (NOUT6,940) IFAIL,XLIK,(PARAM(J),J=1,NPAR)
0541             WRITE (NOUT6,941) (SE(J),J=1,NPAR)
0542             IF (NPAR.GE.2) WRITE (NOUT6,950) PARVAR(2,1)
0543             IF (NPAR.LT.3) GO TO 51
0544             DO 50 J=3,NPAR
0545.01          KU=J-1
0546.01   50     WRITE (NOUT6,951) (PARVAR(J,K),K=1,KU)
0547      51     IF (NOUT.GE.1) CALL REGOUT(IFAIL,XLIK,XLIKD,PARAM,NPAR,SE,
0548             1                           PARVAR,NPARVR,Y,F,Z,THETA,NLMAX,
0549             1                           NL,NQA,VAR,NPQA,NP,NQ,NOUT)
0550              IF (NCVI.EQ.1) CALL REGCVI(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,
0551             1                           Y,F,PARAM,PARAMM,PARAMS,PARAMT,
0552             1                           MLEST,NPAR,NM,NDER,NOUT,IFAIL)
0553              RETURN
0554      C
0555      C        ERROR MESSAGES
0556      C
0557      017    WRITE (NOUT6,918) NL
0558             IF (NOUT.GE.1) WRITE (NOUT10,918) NL
0559      918    FORMAT (' *** REGAME: NL =',I5,' OUT OF RANGE ***')
0560             IFAIL=1
0561             RETURN
0562      019    WRITE (NOUT6,920) ACCTOL
0563             IF (NOUT.GE.1) WRITE (NOUT10,920) ACCTOL
0564      920    FORMAT (' *** REGAME: ACCTOL =',1PG12.4,' OUT OF RANGE ***')
0565             IFAIL=1
0566             RETURN
0567      021    WRITE (NOUT6,922) NLIK
0568             IF (NOUT.GE.1) WRITE (NOUT10,922) NLIK
0569      922    FORMAT (' *** REGAME: NLIK =',I5,' OUT OF RANGE ***')
0570             IFAIL=1
0571             RETURN
0572      023    WRITE (NOUT6,924) NM
0573             IF (NOUT.GE.1) WRITE (NOUT10,924) NM
0574      924    FORMAT (' *** REGAME: NM =',I5,' OUT OF RANGE ***')
0575             IFAIL=1
0576             RETURN
```

```
0577     025   WRITE (NOUT6,926) NPARVR
0578           IF (NOUT.GE.1) WRITE (NOUT10,926) NPARVR
0579     926   FORMAT (' *** REGAME: NPARVR =',I5,' OUT OF RANGE ***')
0580           IFAIL=1
0581           RETURN
0582     027   WRITE (NOUT6,928) NPAR
0583           IF (NOUT.GE.1) WRITE (NOUT10,928) NPAR
0584     928   FORMAT (' *** REGAME: NPAR =',I5,' OUT OF RANGE ***')
0585           IFAIL=1
0586           RETURN
0587     029   WRITE (NOUT6,930) SE(J)
0588           IF (NOUT.GE.1) WRITE (NOUT10,930) SE(J)
0589     930   FORMAT (' *** REGAME: SE =',1PG12.4,' OUT OF RANGE ***')
0590           IFAIL=1
0591           RETURN
0592     032   WRITE (NOUT6,933)
0593           IF (NOUT.GE.1) WRITE (NOUT10,933)
0594     933   FORMAT (' *** REGAME: ALL PARAMETERS CONSTRAINED ***')
0595           IFAIL=1
0596           RETURN
0597     031   WRITE (NOUT6,934)
0598           IF (NOUT.GE.1) WRITE (NOUT10,934)
0599     934   FORMAT (' *** REGAME: XLIK CANNOT BE EVAL. AT PARAM START ***')
0600           IFAIL=1
0601           RETURN
0602     034   WRITE (NOUT6,935)
0603           IF (NOUT.GE.1) WRITE (NOUT10,935)
0604     935   FORMAT (' *** REGAME: E04HBF PARAMETER OUT OF RANGE ***')
0605           RETURN
0606     036   WRITE (NOUT6,937)
0607           IF (NOUT.GE.1) WRITE (NOUT10,937)
0608     937   FORMAT (' *** REGAME: E04JBF PARAMETER OUT OF RANGE ***')
0609           RETURN
0610     C
0611     C        FORMAT STATEMENTS
0612     C
0613     938   FORMAT (//9X,'R E S U L T S   F R O M   P R O G R A M   ',
0614          1 'R E G A M E')
0615     939   FORMAT (//' NITER NCALL  XLIK       XLIKCH   SLOPE  ',
0616          1 'PARACH   COND'/21X,'PARAM.....')
0617     940   FORMAT (/' IFAIL=',I3,1PG14.6/19X,5G12.4/
0618          1 19X,5G12.4/19X,5G12.4/19X,5G12.4)
0619     941   FORMAT (13X,'S.E.= ',1P5G12.4/
0620          1 19X,5G12.4/19X,5G12.4/19X,5G12.4)
0621     950   FORMAT (/12X,'CORREL ',5(F8.4,4X)/
0622          1 19X,5(F8.4,4X)/19X,5(F8.4,4X)/19X,5(F8.4,4X))
0623     951   FORMAT (19X,5(F8.4,4X)/
0624          1 19X,5(F8.4,4X)/19X,5(F8.4,4X)/19X,5(F8.4,4X))
0625     942   FORMAT (1H1////22X,'R E S U L T S   F R O M   P R O G R A M',
0626          1 '   R E G A M E  -  C . A . G L A S B E Y'/)
0627     943   FORMAT (/1X,119('*')
0628          1  /5X,'R',9X,'D',39X,'D',9X,'R',39X,'R',9X,'D'
0629          1  /3X,'F',3X,'U',5X,'U',3X,'O',35X,'U',3X,'O',5X,'F',3X,'U'
0630          1  ,35X,'F',3X,'U',5X,'U',3X,'O'
0631          1  /1X,'A',3X,'*',3X,'S',1X,'E',3X,'*',3X,'S'
0632          1  ,31X,'E',3X,'*',3X,'S',1X,'A',3X,'*',3X,'S'
0633          1  ,31X,'A',3X,'*',3X,'S',1X,'E',3X,'*',3X,'S'
0634          1  /1X,'E',3X,'+',3X,'S',1X,'A',3X,'+',3X,'S'
0635          1  ,31X,'A',3X,'+',3X,'S',1X,'E',3X,'+',3X,'S'
0636          1  ,31X,'E',3X,'+',3X,'S',1X,'A',3X,'+',3X,'S'
0637          1  /3X,'U',3X,'O',5X,'F',3X,'U',35X,'F',3X,'U',5X,'U',3X,'O'
0638          1  ,35X,'U',3X,'O',5X,'F',3X,'U'
0639          1  /5X,'D',9X,'R',39X,'R',9X,'D',39X,'D',9X,'R'
0640          1  /1X,119('*')/)
```

```
0641     944   FORMAT (///'  I T E R A T I O N S')
0642     945   FORMAT (//'      NCALMX  NLIK NSCALE  NM   NDER     ACCTOL',
0643          1 '     NVAR  NPEST    PARAMS...../
0644          1 4X,I6,1X,4I6,5X,1PD8.1,1X,2I6,3X,4G14.6/62X,4G14.6/
0645          1 60X,4G14.6/58X,4G14.6)
0646     946   FORMAT (//' NITER NCALL  XLIK       XLIKCH   SLOPE  ',
0647          1 'PARACH   COND           PARAM.....')
0648           END
0649
0650           SUBROUTINE REGLIK(NFLAG,NPEST,PARAMT,XLIKT,XLIKD,MWK,NMWK,WK,NWK)
0651     C
0652     C
0653     C-----------------------------------------------------------------------
0654     C                          REGLIK
0655     C                          ======
0656     C
0657     C         EVALUATES OPTIMIZATION FUNCTION AT PARAMT CALLING REGMOD,
0658     C         AND REGDER IF NLIK=2, NDER=1
0659     C
0660     C  INPUT
0661     C   NPEST                 NUMBER OF PARAMETERS TO BE ESTIMATED
0662     C   PARAMT(NPEST)         RESCALED PARAMETERS TO BE ESTIMATED
0663     C
0664     C  OUTPUT
0665     C   XLIKT                 RESCALED OPTIMIZATION FUNCTION
0666     C                            (SET TO 1.0D20 IF IT CANNOT BE EVALUATED)
0667     C
0668     C  DUMMY
0669     C   NFLAG
0670     C   XLIKD(NPEST)
0671     C   MWK(NMWK)
0672     C   NMWK
0673     C   WK(NWK)
0674     C   NWK
0675     C
0676     C-----------------------------------------------------------------------
0677     C
0678           DOUBLE PRECISION
0679          1    PARAMT(NPEST),XLIKT,XLIKD(NPEST),WK(NWK),
0680          1    PARAMM(16),PARAMS(16),XLIKM,XLIKS,STEPMX,SCALE,
0681          1 Y(200),F(200),Z(200),VAR(200,11),THETA(200,6),
0682          1    DD(5,5),DDDET,DDTDET,DERIV(200,5),E(200),EPSL,PARAM(16),
0683          1 PHI(200,6),SS,W1(200),W2(200),XLDET,XX,XY
0684           INTEGER
0685          1    NFLAG,NPEST,MWK(NMWK),NMWK,NWK,
0686          1    MLEST(16),NPAR,NOUT,NSCALE,NSTAGE,
0687          1 NL,NPQA,NPA,NQA,NP,NQ,NLIK,NM,NDER,NLMAX,NPMAX,NQMAX,
0688          1    I,IFAIL,J,K,NFDIF,NPRI,NTEQ
0689           COMMON
0690          1 /REGCM1/ PARAMM,PARAMS,XLIKM,XLIKS,STEPMX,MLEST,NPAR,NOUT
0691          1 /REGCM2/ SCALE,NSCALE,NSTAGE
0692          1 /REGCM3/ Y,F,Z,VAR,THETA,NL,NPQA,NPA,NQA,NP,NQ,
0693          1          NLIK,NM,NDER,NLMAX,NPMAX,NQMAX
0694          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0695     C
0696     C-----------------------------------------------------------------------
0697     C
0698     C  LOCAL
0699     C   DD(5,5)               SUM OF SQARES OF DERIV, AND TRANSFORMED DERIV
0700     C   DDDET                 DETERMINANT OF DD
0701     C   DDTDET                DETERMINANT OF DD AFTER TRANSFORMATION
0702     C   DERIV(200,5)          FIRST DERIVATIVE OF F(NL) W.R.T. PARAM(NM)
0703     C   E(200)                VECTOR OF DEPARTURES
0704     C   EPSL                  STEPLENGTH USED TO CALC DERIV IF NDER NOT 1
```

```
0705     C   I                   LENGTH INDEX
0706     C   IFAIL               FAILURE INDICATOR
0707     C   J                   INDEX
0708     C   K                   INDEX
0709     C   NFDIF               1  IF ONLY F REQUIRED FROM REGMOD
0710     C   NPRI                1  IF INTERMEDIATE RESULTS OUTPUT
0711     C   NTEQ                1  IF ALL ROWS OF PHI EQUAL
0712     C   PARAM(16)           VECTOR OF MODEL PARAMETERS
0713     C   PHI(200,6)          FIRST MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
0714     C   SS                  SUM OF SQUARES OF RESIDUALS
0715     C   W1(200)             WORK SPACE
0716     C   W2(200)             WORK SPACE
0717     C   XLDET               LOG DETERMINANT OF THETA
0718     C   XX                  WORK SPACE
0719     C   XY                  WORK SPACE
0720     C
0721     C
0722     C-----------------------------------------------------------------------
0723           EPSL=1.0D-4
0724           NPRI=0
0725           IF (NOUT.GE.5) NPRI=1
0726           K=1
0727           DO 001 J=1,NPAR
0728.01        XX=0.0D0
0729.01        IF (MLEST(K).NE.J) GO TO 001
0730.01        XX=PARAMT(K)
0731.01        K=K+1
0732.01  001   PARAM(J)=XX*PARAMS(J)+PARAMM(J)
0733.01  C
0734           NFDIF=0
0735           IFAIL=0
0736           CALL REGMOD(PARAM,NPAR,NLMAX,NL,NFDIF,F,VAR,NPQA,NP,NQ,NTEQ,
0737          1            NOUT,IFAIL)
0738           IF (IFAIL.NE.0) GO TO 022
0739           IF ((NP.LT.0).OR.(NP.GT.NPMAX)) GO TO 024
0740           IF ((NQ.LT.0).OR.(NQ.GT.NQMAX)) GO TO 026
0741           IF (NPQA.NE.NP+NQ+1) GO TO 028
0742     C
0743           NPA=NP+1
0744           NQA=NQ+1
0745           CALL REGAMD(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,NTEQ,PHI,THETA,XLDET,
0746          1            NOUT,NPRI,IFAIL)
0747           IF (IFAIL.NE.0) GO TO 030
0748     C
0749           DO 002 I=1,NL
0750.01  002   E(I)=Y(I)-F(I)
0751           CALL REGTRA(E,NL,PHI,NLMAX,NPA,0,0,W1,NOUT,NPRI,IFAIL)
0752           CALL REGTRA(W1,NL,THETA,NLMAX,NQA,1,0,Z,NOUT,NPRI,IFAIL)
0753           IF (IFAIL.NE.0) GO TO 030
0754           SS=0.0D0
0755           DO 003 I=1,NL
0756.01  003   SS=SS+Z(I)*Z(I)
0757           GO TO (004,006,020),NLIK
0758     C
0759     C        NLIK=1  MAXIMUM LIKELIHOOD
0760     C        ======  ==================
0761     C
0762     C        XLIK=0.5*NL*LOG(SCALE)+0.5*LOG(DET(VAR))+0.5*SS/SCALE
0763     C
0764     C
0765     004   IF (NSTAGE.EQ.2) GO TO 005
0766           SCALE=1.0D0
0767           IF (NSCALE.EQ.1) SCALE=SS/FLOAT(NL)
0768     005   XLIKT=((FLOAT(NL)*DLOG(SCALE)+SS/SCALE)/2.0D0+XLDET-XLIKM)/XLIKS
```

```
0769            RETURN
0770      C
0771      C        NLIK=2  RESTRICTED MAXIMUM LIKELIHOOD
0772      C        ======  ============================
0773      C
0774      C        XLIK=0.5*(NL-NM)*LOG(SCALE)+0.5*LOG(DET(VAR))+0.5*SS/SCALE
0775      C            +0.5*LOG(DET(DERIV*VAR(-1)*DERIV)/DET(DERIV*DERIV))
0776      C
0777      C
0778      006   IF (NM.EQ.0) GO TO 004
0779            IF (NDER.NE.1) GO TO 007
0780            IFAIL=0
0781            CALL REGDER(PARAM,NPAR,NLMAX,NL,NM,DERIV,NOUT,IFAIL)
0782            IF (IFAIL.NE.0) GO TO 031
0783            GO TO 010
0784      C
0785      007   NFDIF=1
0786            IFAIL=1
0787            K=1
0788            DO 009 J=1,NM
0789.01         XX=EPSL
0790.01         IF (MLEST(K).EQ.J) XX=EPSL*PARAMS(J)
0791.01         PARAM(J)=PARAM(J)+XX
0792.01         CALL REGMOD(PARAM,NPAR,NLMAX,NL,NFDIF,W1,VAR,NPQA,NP,NQ,NTEQ,
0793.01        1            NOUT,IFAIL)
0794.01         XY=0.0D0
0795.01         IF (MLEST(K).NE.J) GO TO 008
0796.01         XY=PARAMT(K)
0797.01         K=K+1
0798.01   008   PARAM(J)=XY*PARAMS(J)+PARAMM(J)
0799.01         IF (IFAIL.NE.0) GO TO 033
0800.01         DO 009 I=1,NL
0801.02   009   DERIV(I,J)=(W1(I)-F(I))/XX
0802.02   C
0803      010   DO 012 J=1,NM
0804.01         DO 012 K=1,J
0805.02         XX=0.0D0
0806.02         DO 011 I=1,NL
0807.03   011   XX=XX+DERIV(I,J)*DERIV(I,K)
0808.02   012   DD(K,J)=XX
0809            IFAIL=1
0810            CALL F03ABF(DD,5,NM,DDDET,W1,IFAIL)
0811            IF (IFAIL.NE.0) GO TO 035
0812            IF (NPRI.NE.1) GO TO 013
0813            CALL MATPR2(DERIV,NLMAX,NL,NM,1)
0814            CALL MATPR2(DD,5,NM,NM,1)
0815            CALL MATPRO(DDDET,1)
0816      C
0817      013   DO 017 J=1,NM
0818.01         DO 014 I=1,NL
0819.02   014   W1(I)=DERIV(I,J)
0820.01         CALL REGTRA(W1,NL,PHI,NLMAX,NPA,0,0,W2,NOUT,NPRI,IFAIL)
0821.01         CALL REGTRA(W2,NL,THETA,NLMAX,NQA,1,0,W1,NOUT,NPRI,IFAIL)
0822.01         IF (IFAIL.NE.0) GO TO 030
0823.01         DO 015 I=1,NL
0824.02   015   DERIV(I,J)=W1(I)
0825.01         DO 017 K=1,J
0826.02         XX=0.0D0
0827.02         DO 016 I=1,NL
0828.03   016   XX=XX+DERIV(I,J)*DERIV(I,K)
0829.02   017   DD(K,J)=XX
0830            IFAIL=1
0831            CALL F03ABF(DD,5,NM,DDTDET,W1,IFAIL)
0832            IF (IFAIL.NE.0) GO TO 035
```

```
      IF (NPRI.NE.1) GO TO 018
      CALL MATPR2(DERIV,NLMAX,NL,NM,1)
      CALL MATPR2(DD,5,NM,NM,1)
      CALL MATPRO(DDTDET,1)
C
018   IF (NSTAGE.EQ.2) GO TO 019
      SCALE=1.0D0
      IF (NSCALE.EQ.1) SCALE=SS/FLOAT(NL-NM)
019   XLIKT=((FLOAT(NL-NM)*DLOG(SCALE)+SS/SCALE)/2.0D0+XLDET
     1      +DLOG(DDTDET/DDDET)/2.0D0-XLIKM)/XLIKS
      RETURN
C
C        NLIK=3  SUM OF SQUARES
C        ======  ==============
C
C        XLIK=0.5*NL*LOG(SCALE)+0.5*SS/SCALE
C
C
020   IF (NSTAGE.EQ.2) GO TO 021
      SCALE=1.0D0
      IF (NSCALE.EQ.1) SCALE=SS/FLOAT(NL)
021   XLIKT=((FLOAT(NL)*DLOG(SCALE)+SS/SCALE)/2.0D0-XLIKM)/XLIKS
      RETURN
C
C        ERROR MESSAGES
C
022   WRITE (NOUT6,923)
      IF (NOUT.GE.1) WRITE (NOUT10,923)
923   FORMAT (' *** REGLIK: REGMOD CANNOT EVALUATE F OR VAR ***')
      XLIKT=1.0D20
      RETURN
024   WRITE (NOUT6,925) NP
      IF (NOUT.GE.1) WRITE (NOUT10,925) NP
925   FORMAT (' *** REGLIK: NP =',I5,' OUT OF RANGE ***')
      XLIKT=1.0D20
      RETURN
026   WRITE (NOUT6,927) NQ
      IF (NOUT.GE.1) WRITE (NOUT10,927) NQ
927   FORMAT (' *** REGLIK: NQ =',I5,' OUT OF RANGE ***')
      XLIKT=1.0D20
      RETURN
028   WRITE (NOUT6,929) NPQA
      IF (NOUT.GE.1) WRITE (NOUT10,929) NPQA
929   FORMAT (' *** REGLIK: NPQA =',I5,' NOT EQUAL TO NP+NQ+1 ***')
      XLIKT=1.0D20
      RETURN
030   XLIKT=1.0D20
      RETURN
031   WRITE (NOUT,932)
      IF (NOUT.GE.1) WRITE (NOUT10,932)
932   FORMAT (' *** REGLIK: REGDER CANNOT EVALUATE DERIV ***')
      XLIKT=1.0D20
      RETURN
033   WRITE (NOUT6,934)
      IF (NOUT.GE.1) WRITE (NOUT10,934)
934   FORMAT (' *** REGLIK: REGMOD CANNOT EVALUATE F FOR NLIK=2 ***')
      XLIKT=1.0D20
      RETURN
035   WRITE (NOUT6,936)
      IF (NOUT.GE.1) WRITE (NOUT10,936)
936   FORMAT (' *** REGLIK: DETERMINANT FOR NLIK=2 NOT POS. DEF. ***')
      XLIKT=1.0D20
      RETURN
      END
```

```
0897
0898           SUBROUTINE REGMON(NPEST,PARAMT,XLIKT,XLIKD,MCONST,XLNORM,COND,
0899          1                  LOGJBF,NITER,NCALL,MWK,NMWK,WK,NWK)
0900      C
0901      C
0902      C-----------------------------------------------------------------------
0903      C                          REGMON
0904      C                          ======
0905      C
0906      C         MONITORS ITERATIONS BY E04JBF
0907      C
0908      C  INPUT
0909      C   NPEST                 NUMBER OF PARAMETERS TO BE ESTIMATED
0910      C   PARAMT(NPEST)         RESCALED PARAMETERS TO BE ESTIMATED
0911      C   XLIKT                 RESCALED OPTIMIZATION FUNCTION
0912      C   XLIKD(NPEST)          VECTOR OF FIRST DERIVATIVES OF XLIKT
0913      C   MCONST(NPEST)         DENOTES CONSTRAINED PARAMETERS IN E04JBF
0914      C   XLNORM                NORM OF XLIKD
0915      C   COND                  RATIO OF LARGEST TO SMALLEST VALUES IN HESD
0916      C   NITER                 NUMBER OF ITERATIONS
0917      C   NCALL                 NUMBER OF CALLS BY E04JBF TO REGLIK
0918      C
0919      C  DUMMY
0920      C   LOGJBF
0921      C   MWK(NMWK)
0922      C   NMWK
0923      C   WK(NWK)
0924      C   NWK
0925      C
0926      C
0927      C-----------------------------------------------------------------------
0928            DOUBLE PRECISION
0929           1    PARAMT(NPEST),XLIKT,XLIKD(NPEST),XLNORM,COND,WK(NWK),
0930           1    PARAMM(16),PARAMS(16),XLIKM,XLIKS,STEPMX,
0931           1    BLA,CEE,PARACH,PARAM(16),PARATX(16),SIGN,SLOPE,
0932           1 XLIK,XLIKCH,XLIKTX,XX,XY
0933            INTEGER
0934           1    NPEST,MCONST(NPEST),NITER,NCALL,MWK(NMWK),NMWK,NWK,
0935           1    MLEST(16),NPAR,NOUT,
0936           1    J,K
0937            LOGICAL
0938           1    LOGJBF
0939            COMMON
0940           1 /REGCM1/ PARAMM,PARAMS,XLIKM,XLIKS,STEPMX,MLEST,NPAR,NOUT
0941           1 /MONCM1/ XLIKTX,PARATX
0942           1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0943      C
0944      C
0945      C-----------------------------------------------------------------------
0946      C  LOCAL
0947      C   BLA                   BLANK CHARACTER
0948      C   CEE                   'C' CHARACTER, DENOTES CONSTRAINED STEP
0949      C   J                     INDEX
0950      C   K                     INDEX
0951      C   PARACH                CHANGE IN PARAMT
0952      C   PARAM(16)             VECTOR OF MODEL PARAMETERS
0953      C   PARATX(16)            PARAMT AT LAST ITERATION
0954      C   SIGN                  SET TO BLA OR CEE
0955      C   SLOPE                 SLOPE OF XLIKT W.R.T. PARAMT
0956      C   XLIK                  OPTIMIZATION FUNCTION
0957      C   XLIKCH                CHANGE IN XLIKT
0958      C   XLIKTX                XLIKT AT LAST ITERATION
0959      C   XX                    WORK SPACE
0960      C   XY                    WORK SPACE
```

```
0961     C
0962     C
0963     C-----------------------------------------------------------------------
0964           DATA CEE,BLA/'C',' '/
0965     C
0966     C        CALCULATE CONVERGENCE CRITERIA
0967     C
0968           XLIK=XLIKT*XLIKS+XLIKM
0969           K=1
0970           DO 001 J=1,NPAR
0971.01        XX=0.0D0
0972.01        IF (MLEST(K).NE.J) GO TO 001
0973.01        XX=PARAMT(K)
0974.01        K=K+1
0975.01  001   PARAM(J)=XX*PARAMS(J)+PARAMM(J)
0976.01  C
0977           XLIKCH=1.0D-10
0978           PARACH=0.0D0
0979           XX=0.0D0
0980           IF (NITER.EQ.0) GO TO 003
0981     C
0982           XLIKCH=DSIGN(DSQRT(DABS(XLIKTX-XLIKT)/(1.0D0+DABS(XLIKT))),
0983          1                          XLIKTX-XLIKT)+1.0D-10
0984           DO 002 J=1,NPEST
0985.01        XX=XX+PARAMT(J)*PARAMT(J)
0986.01        XY=PARATX(J)-PARAMT(J)
0987.01  002   PARACH=PARACH+XY*XY
0988           PARACH=DSQRT(PARACH)
0989     003   SIGN=BLA
0990           IF (PARACH.GE.(1.0D0-1.0D-6)*STEPMX) SIGN=CEE
0991           PARACH=PARACH/(1.0D0+DSQRT(XX))+1.0D-10
0992     C
0993           DO 004 J=1,NPEST
0994.01  004   PARATX(J)=PARAMT(J)
0995           XLIKTX=XLIKT
0996           SLOPE=XLNORM/(1.0D0+DABS(XLIKT))
0997     C
0998     C        OUTPUT
0999     C
1000           WRITE (NOUT6,905) NITER,NCALL,XLIK,XLIKCH,SLOPE,PARACH,COND,
1001          1 SIGN,(PARAM(J),J=1,NPAR)
1002           IF (NOUT.LT.1) RETURN
1003           WRITE (NOUT10,906) NITER,NCALL,XLIK,XLIKCH,SLOPE,PARACH,COND,
1004          1 SIGN,(PARAM(J),J=1,NPAR)
1005     905   FORMAT (I4,I6,1PG14.6,3D8.1,D8.1,2X,A1/
1006          1 19X,5G12.4/19X,5G12.4/19X,5G12.4/19X,5G12.4)
1007     906   FORMAT (I4,I6,1PG14.6,3D8.1,D8.1,2X,A1,
1008          1 5X,4G14.6/62X,4G14.6/60X,4G14.6/58X,4G14.6)
1009           RETURN
1010           END
1011
1012           SUBROUTINE REGVAR(NVAR,HESL,NHESL,HESD,ACCTOL,NPEST,XLIKT,NL,
1013          1                  NPARVR,PARAMT,PARVAR,SE,MWK,NMWK,WK,NWK,
1014          1                  NPRI,IFAILT)
1015     C
1016     C-----------------------------------------------------------------------
1017     C
1018     C                        REGVAR
1019     C                        ======
1020     C
1021     C        CALCULATES APPROXIMATE VARIANCES FOR PARAM BY
1022     C        INVERTING THE MATRIX OF SECOND DERIVATIVES OF
1023     C        XLIK OBTAINED BY DIFFERENCING.
1024     C
```

```
1025     C  INPUT
1026     C   NVAR                  1  IF PARVAR TO BE EST. BY DIFFERENCE METHOD
1027     C   HESL(NHESL)           LOWER TRIANGLE OF APPROX. TO SECOND DERIVS.
1028     C   NHESL                 DIMENSION OF HESL
1029     C   HESD(NPEST)           DIAGONAL OF APPROX. TO SECOND DERIVATIVES
1030     C   ACCTOL                ACCURACY TO WHICH PARAMETERS EST. BY E04JBF
1031     C   NPEST                 NUMBER OF PARAMETERS ESTIMATED
1032     C   XLIKT                 RESCALED OPTIMIZATION FUNCTION AT MAX.
1033     C   NL                    NUMBER OF OBSERVATIONS
1034     C   NPARVR                DIMENSIONS OF PARVAR
1035     C
1036     C  INPUT/OUTPUT
1037     C   PARAMT(16)             RESCALED ESTIMATED PARAMETERS AT MAX.
1038     C
1039     C  OUTPUT
1040     C   PARVAR(NPARVR,NPARVR) ESTIMATED VARIANCES OF PARAMETERS,
1041     C                            CORRELATIONS BELOW DIAGONAL
1042     C   SE(16)                STANDARD ERRORS OF PARAMETERS
1043     C
1044     C  DUMMY
1045     C   MWK(NMWK)
1046     C   NMWK
1047     C   WK(NWK)
1048     C   NWK
1049     C
1050     C  CONTROL
1051     C   NPRI                  1  INTERMEDIATE RESULTS OUTPUT
1052     C   IFAILT                1  IF PARVAR CANNOT BE EVALUATED
1053     C
1054     C
1055     C-----------------------------------------------------------------------
1056           DOUBLE PRECISION
1057          1    HESL(NHESL),HESD(NPEST),ACCTOL,XLIKT,PARAMT(16),
1058          1 PARVAR(NPARVR,NPARVR),SE(16),WK(NWK),
1059          1    PARAMM(16),PARAMS(16),XLIKM,XLIKS,STEPMX,SCALE,
1060          1    EPSV,REGVRF,TRAN(16,16),W1(16),W2(16),XLDER(17,16),
1061          1 XMM,XMP,XPM,XPP
1062           INTEGER
1063          1    NVAR,NHESL,NPEST,NL,NPARVR,MWK(NMWK),NMWK,NWK,NPRI,IFAILT,
1064          1    MLEST(16),NPAR,NOUT,NSCALE,NSTAGE,
1065          1    I,IFAILV,J,JF,JS,K,L,NFLAG,NPARN,NPESTA,NPESTN,NPESTZ,
1066          1 NTRAN,NXLDER
1067           COMMON
1068          1 /REGCM1/ PARAMM,PARAMS,XLIKM,XLIKS,STEPMX,MLEST,NPAR,NOUT
1069          1 /REGCM2/ SCALE,NSCALE,NSTAGE
1070          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
1071     C
1072     C-----------------------------------------------------------------------
1073     C
1074     C  LOCAL
1075     C   EPSV                  STEP USED TO CALC SECOND DERIVS OF XLIKT
1076     C   I                     INDEX
1077     C   IFAILV                INTERNAL FAILURE INDICATOR
1078     C   J                     INDEX
1079     C   JF                    UPPER BOUND ON J DO-LOOP
1080     C   JS                    LOWER BOUND ON J DO-LOOP
1081     C   K                     INDEX
1082     C   L                     INDEX
1083     C   NFLAG                 DUMMY VARIABLE IN REGLIK
1084     C   NPARN                 NPAR (+1 IF NSCALE=1)
1085     C   NPESTA                NPESTN+1
1086     C   NPESTN                NPEST (+1 IF NSCALE=1)
1087     C   NPESTZ                NPESTN-1
1088     C   NTRAN                 DIMENSION OF TRAN
```

```
1089     C   NXLDER              DIMENSION OF XLDER
1090     C   REGVRF              FUNCTION TO EVAL. XLIKT NEAR PARAMT
1091     C   TRAN(16,16)         LOWER TRIANGULAR MATRIX FORM HESD, HESL
1092     C   W1(16)              WORK SPACE
1093     C   W2(16)              WORK SPACE
1094     C   XLDER(17,16)        SECOND DERIVS. OF XLIKT W.R.T. PARAMT
1095     C   XMM                 XLIKT AT PARAMT -EPSV -EPSV
1096     C   XMP                 XLIKT AT PARAMT -EPSV +EPSV
1097     C   XPM                 XLIKT AT PARAMT +EPSV -EPSV
1098     C   XPP                 XLIKT AT PARAMT +EPSV +EPSV
1099     C
1100     C
1101     C-------------------------------------------------------------------
1102           DATA NTRAN,NXLDER/16,17/
1103           IFAILT=0
1104           EPSV=10.0D0*ACCTOL
1105           NSTAGE=2
1106     C
1107     C        LOWER TRIANGULAR TRANSFORMATION MATRIX  TRAN
1108     C
1109           NPARN=NPAR
1110           NPESTN=NPEST
1111           IF (NSCALE.NE.1) GO TO 001
1112           NPARN=NPAR+1
1113           NPESTN=NPEST+1
1114     001   NPESTZ=NPESTN-1
1115           DO 003 I=1,NPEST
1116.01        DO 002 J=1,NPESTZ
1117.02  002   TRAN(I,J)=0.0D0
1118.01  003   TRAN(I,NPESTN)=DSQRT(HESD(I))
1119.01  C
1120           IF (NSCALE.NE.1) GO TO 005
1121           DO 004 J=1,NPESTZ
1122.01  004   TRAN(NPESTN,J)=0.0D0
1123           TRAN(NPESTN,NPESTN)=DSQRT(FLOAT(NL)/(2.0D0*XLIKS))/SCALE
1124           PARAMS(NPARN)=1.0D0
1125           PARAMM(NPARN)=0.0D0
1126           MLEST(NPESTN)=NPARN
1127           PARAMT(NPESTN)=SCALE
1128     C
1129     005   K=0
1130           DO 006 I=2,NPEST
1131.01        JS=NPESTN+1-I
1132.01        L=0
1133.01        DO 006 J=JS,NPESTZ
1134.02        K=K+1
1135.02        L=L+1
1136.02  006   TRAN(I,J)=HESL(K)*TRAN(L,NPESTN)
1137           IF (NPRI.EQ.1) CALL MATPR2(TRAN,NTRAN,NPESTN,NPESTN,1)
1138           DO 007 I=1,NPESTN
1139.01  007   W1(I)=0.0D0
1140           IFAILV=0
1141     C
1142     C        MATRIX OF SECOND DERIVATIVES ON TRANSFORMED SCALE  XLDER
1143     C        (SHOULD APPROXIMATE TO IDENTITY MATRIX)
1144     C
1145           IF (NVAR.EQ.1) GO TO 033
1146           WRITE (NOUT6,935)
1147           IF (NOUT.GE.1) WRITE (NOUT10,935)
1148     935   FORMAT (' *** REGVAR: NVAR NOT 1, VARIANCES AN APPROX. ***')
1149           GO TO 034
1150     033   DO 009 I=1,NPESTN
1151.01        DO 009 J=1,I
1152.02        XPP=REGVRF(I,J, 1, 1,W1,W2,EPSV,NPESTN.TRAN,NTRAN.PARAMT,NPEST,
```

```
1153.02          1           MWK,NMWK,WK,NWK,NOUT,NPRI,IFAILV)
1154.02           IF (IFAILV.NE.0) GO TO 010
1155.02           XMM=REGVRF(I,J,-1,-1,W1,W2,EPSV,NPESTN,TRAN,NTRAN,PARAMT,NPEST,
1156.02          1           MWK,NMWK,WK,NWK,NOUT,NPRI,IFAILV)
1157.02           IF (IFAILV.NE.0) GO TO 010
1158.02           IF (J.EQ.I) GO TO 008
1159.02           XMP=REGVRF(I,J,-1, 1,W1,W2,EPSV,NPESTN,TRAN,NTRAN,PARAMT,NPEST,
1160.02          1           MWK,NMWK,WK,NWK,NOUT,NPRI,IFAILV)
1161.02           IF (IFAILV.NE.0) GO TO 010
1162.02           XPM=REGVRF(I,J, 1,-1,W1,W2,EPSV,NPESTN,TRAN,NTRAN,PARAMT,NPEST,
1163.02          1           MWK,NMWK,WK,NWK,NOUT,NPRI,IFAILV)
1164.02           IF (IFAILV.NE.0) GO TO 010
1165.02           XLDER(I,J)=(XPP-XMP-XPM+XMM)/(4.0D0*EPSV*EPSV)
1166.02           XLDER(J,I)=XLDER(I,J)
1167.02           GO TO 009
1168.02  008      XLDER(I,J)=(XPP-2.0D0*XLIKT+XMM)/(EPSV*EPSV)
1169.02  009      CONTINUE
1170              GO TO 014
1171        C
1172        010   IF (IFAILT.EQ.1) GO TO 025
1173              WRITE (NOUT6,911)
1174              IF (NOUT.GE.1) WRITE (NOUT10,911)
1175        911   FORMAT (' *** REGVAR: FAILURE, VARIANCES AN APPROX. ***')
1176        034   IFAILT=1
1177              DO 013 I=1,NPESTN
1178.01           DO 012 J=1,NPESTN
1179.02  012      XLDER(I,J)=0.0D0
1180.01  013      XLDER(I,I)=1.0D0
1181              GO TO 014
1182        014   IF (NPRI.EQ.1) CALL MATPR2(XLDER,NXLDER,NPESTN,NPESTN,1)
1183        C
1184        C        MATRIX OF SECOND DERIVATIVES ON  PARAMT  SCALE  XLDER
1185        C
1186              DO 016 I=1,NPESTN
1187.01           DO 015 J=1,NPESTN
1188.02  015      W1(J)=XLDER(J,I)
1189.01           CALL REGTRA(W1,NPESTN,TRAN,NTRAN,NPESTN,0,0,W2,NOUT,NPRI,IFAILV)
1190.01           DO 016 J=1,NPESTN
1191.02  016      XLDER(J,I)=W2(J)
1192.02  C
1193              DO 018 I=1,NPESTN
1194.01           DO 017 J=1,NPESTN
1195.02  017      W1(J)=XLDER(I,J)
1196.01           CALL REGTRA(W1,NPESTN,TRAN,NTRAN,NPESTN,0,0,W2,NOUT,NPRI,IFAILV)
1197.01           DO 018 J=1,NPESTN
1198.02  018      XLDER(I,J)=W2(J)
1199.02  C
1200.02  C        VARIANCE MATRIX  PARVAR , STANDARD ERRORS  SE
1201.02  C        (OBTAINED BY INVERTING RESCALED  XLDER)
1202.02  C
1203              IFAILV=1
1204              CALL F01ABF(XLDER,NXLDER,NPESTN,PARVAR,NPARVR,W1,IFAILV)
1205              NPESTA=NPESTN+1
1206              IF (NPRI.EQ.1) CALL MATPR2(XLDER,NXLDER,NPESTA,NPESTN,1)
1207              IF (IFAILV.NE.0) GO TO 010
1208        C
1209              K=1
1210              DO 024 I=1,NPARN
1211.01           IF (MLEST(K).EQ.I) GO TO 021
1212.01           DO 020 J=I,NPARN
1213.02           PARVAR(J,I)=0.0D0
1214.02  020      PARVAR(I,J)=0.0D0
1215.01           GO TO 024
1216.01  021      L=K
```

```
1217.01          DO 023 J=I,NPARN
1218.02          IF (MLEST(L).EQ.J) GO TO 022
1219.02          PARVAR(J,I)=0.0D0
1220.02          PARVAR(I,J)=0.0D0
1221.02          GO TO 023
1222.02  022     L=L+1
1223.02          PARVAR(J,I)=0.0D0
1224.02          PARVAR(I,J)=XLDER(L,K)*PARAMS(I)*PARAMS(J)/XLIKS
1225.02  023     CONTINUE
1226.01          K=K+1
1227.01  024     CONTINUE
1228             GO TO 029
1229       C
1230       025   WRITE (NOUT6,926)
1231             IF (NOUT.GE.1) WRITE (NOUT10,926)
1232       926   FORMAT (' *** REGVAR: APPROX. VARIANCES ALSO FAIL ***')
1233             DO 028 I=1,NPARN
1234.01          DO 027 J=1,NPARN
1235.02  027     PARVAR(I,J)=0.0D0
1236.01  028     PARVAR(I,I)=PARAMS(I)
1237       029   DO 030 I=1,NPARN
1238.01  030     SE(I)=DSQRT(PARVAR(I,I))
1239.01  C
1240             IF (NPARN.EQ.1) GO TO 032
1241             DO 031 I=2,NPARN
1242.01          JF=I-1
1243.01          DO 031 J=1,JF
1244.02          IF ((SE(I).LT.1.0D-12).OR.(SE(J).LT.1.0D-12)) GO TO 031
1245.02          PARVAR(I,J)=PARVAR(J,I)/(SE(I)*SE(J))
1246.02  031     CONTINUE
1247       032   CONTINUE
1248             IF (NPRI.EQ.1) CALL MATPR2(PARVAR,NPARVR,NPARN,NPARN,1)
1249       C
1250             IF (NSCALE.EQ.1) SCALE=PARAMT(NPESTN)
1251             CALL REGLIK(NFLAG,NPEST,PARAMT,XLIKT,W1,MWK,NMWK,WK,NWK)
1252             NPAR=NPARN
1253             RETURN
1254             END
1255
1256            DOUBLE PRECISION FUNCTION
1257            I REGVRF(I,J,IS,JS,W1,W2,EPSV,NPESTN,TRAN,NTRAN,PARAMT,
1258            1       NPEST,MWK,NMWK,WK,NWK,NOUT,NPRI,IFAILV)
1259             DOUBLE PRECISION
1260            1    W1(NTRAN),W2(NTRAN),EPSV,TRAN(NTRAN,NTRAN),PARAMT(NTRAN),
1261            1 WK(NWK),
1262            1    SCALE,
1263            1    XLIKD(15)
1264             INTEGER
1265            1    I,J,IS,JS,NPESTN,NTRAN,NPEST,MWK(NMWK),NMWK,NWK,NOUT,
1266            1 NPRI,IFAILV,
1267            1    NSCALE,NSTAGE,
1268            1    K,NFLAG
1269             COMMON
1270            1 /REGCM2/ SCALE,NSCALE,NSTAGE
1271       C
1272       C-------------------------------------------------------------------
1273       C
1274       C  LOCAL
1275       C   K                    INDEX
1276       C   NFLAG                DUMMY VARIABLE IN REGLIK
1277       C   XLIKD(15)            DUMMY VARIABLE USED IN REGLIK
1278       C
1279       C-------------------------------------------------------------------
1280       C
```

```
1281           W1(I)=FLOAT(IS)*EPSV
1282           W1(J)=FLOAT(JS)*EPSV
1283           CALL REGTRA(W1,NPESTN,TRAN,NTRAN,NPESTN,1,1,W2,NOUT,NPRI,IFAILV)
1284           IF (IFAILV.NE.0) RETURN
1285           DO 001 K=1,NPESTN
1286.01  001    W2(K)=W2(K)+PARAMT(K)
1287           IF (NSCALE.EQ.1) SCALE=W2(NPESTN)
1288           CALL REGLIK(NFLAG,NPEST,W2,REGVRF,XLIKD,MWK,NMWK,WK,NWK)
1289           IF (REGVRF.GT.9.9D19) IFAILV=1
1290           W1(I)=0.0D0
1291           W1(J)=0.0D0
1292           RETURN
1293           END
1294
1295           SUBROUTINE REGOUT(IFAIL,XLIK,XLIKD,PARAM,NPAR,SE,PARVAR,NPARVR,
1296          1                  Y,F,Z,THETA,NLMAX,NL,NQA,VAR,NPQA,NP,NQ,NOUT)
1297      C
1298      C
1299      C-----------------------------------------------------------------------
1300      C                        REGOUT
1301      C                        ======
1302      C
1303      C         OUTPUT COMPREHENSIVE SET OF RESULTS TO NOUT10
1304      C
1305      C  INPUT
1306      C   IFAIL                FAILURE INDICATOR FOR REGAME
1307      C   XLIK                 FINAL VALUE OF OPTIMIZATION FUNCTION
1308      C   XLIKD(NPAR)          FIRST DERIVATIVES OF XLIKT W.R.T. PARAMT
1309      C   PARAM(NPAR)          PARAMETER ESTIMATES
1310      C   NPAR                 NUMBER OF MODEL PARAMETERS
1311      C   SE(NPAR)             STANDARD ERRORS OF PARAMETERS
1312      C   PARVAR(NPARVR,NPARVR)  PARAMETER VARIANCES, CORRELATIONS BELOW DIAG.
1313      C   NPARVR                DIMENSIONS OF PARVAR
1314      C   Y(NL)                 DATA VECTOR OF OBSERVATIONS
1315      C   F(NL)                 VECTOR OF FITTED VALUES
1316      C   Z(NL)                 VECTOR OF STANDARDIZED RESIDUALS FROM MODEL
1317      C   THETA(NLMAX,NQA)      SECOND MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
1318      C   NLMAX                 FIRST DIMENSION OF THETA, VAR
1319      C   NL                    NUMBER OF OBSERVATIONS
1320      C   NQA                   BAND WIDTH OF THETA
1321      C   VAR(200,NPQA)         BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
1322      C   NPQA                  NUMBER OF COLUMNS IN VAR
1323      C   NP                    GENERALIZED AUTOREGRESSIVE ORDER
1324      C   NQ                    GENERALIZED MOVING AVERAGE ORDER
1325      C
1326      C  CONTROL
1327      C   NOUT                  CONTROLS OUTPUT
1328      C                          1  IFAIL,XLIK,XLIKD,PARAM,SE,PARVAR OUTPUT
1329      C                          2    AND Y,F,Z,VAR,NP,NQ
1330      C                          >2   AND PLOTS OF Y,F,Z,VAR
1331      C
1332      C-----------------------------------------------------------------------
1333
1334           DOUBLE PRECISION
1335          1    XLIK,XLIKD(NPAR),PARAM(NPAR),SE(NPAR),PARVAR(NPARVR,NPARVR),
1336          1 Y(NL),F(NL),Z(NL),THETA(NLMAX,NQA),VAR(NLMAX,NPQA),
1337          1    AST,BLA,SGN,VARCHR(11),VARMAX(11),VARMIN(11),VARNUM(11),
1338          1 W(200,4),WCHR(4),WSGN(16)
1339           INTEGER
1340          1    IFAIL,NPAR,NPARVR,NLMAX,NL,NQA,NPQA,NP,NQ,NOUT,
1341          1    I,IZ,J,JF,JFF,L,MAXIS(200),MCHN(11),MPL(2,1)
1342           COMMON
1343          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
1344      C
```

```
1345     C
1346     C-------------------------------------------------------------------
1347     C  LOCAL
1348     C   AST                  '*' CARACTER
1349     C   BLA                  BLANK CHARACTER
1350     C   I                    INDEX
1351     C   IZ                   I-1
1352     C   J                    INDEX
1353     C   JF                   MIN(NPQA,4)
1354     C   JFF                  MIN(NPQA,8)
1355     C   L                    INDEX
1356     C   MAXIS(200)           NUMBERS 1-200 USED BY PLTNLG
1357     C   MCHN(11)             CHANNEL NUMBERS USED BY PLTLNG
1358     C   MPL(2,1)             PLOT NUMBERS USED BY PLTACR
1359     C   SGN                  SET TO AST TO DENOTE RESIDUALS > 2.5
1360     C   VARCHR(11)           CHARACTERS USED TO PLOT VAR
1361     C   VARMAX(11)           UPPER BOUNDS USED BY PLTLNG
1362     C   VARMIN(11)           LOWER BOUNDS USED BY PLTLNG
1363     C   VARNUM(11)           NUMBERS STORED AS CHARACTERS FOR VARCHR
1364     C   W(200,4)             WORK SPACE
1365     C   WCHR(4)              CHARCTERS USED TO PLOT Y,F,Z
1366     C   WSGN(16)             SET TO AST TO DENOTE CORRELS.  > 2.5
1367     C
1368     C-------------------------------------------------------------------
1369     C
1370           DATA AST,BLA,WCHR/'**','  ','Y','F','P','Z'/, VARNUM/'0',
1371          1 '1','2','3','4','5','6','7','8','9','A'/
1372     C
1373     C        OUTPUT IFAIL,XLIK,XLIKD,PARAM,SE,PARVAR
1374     C
1375           WRITE (NOUT10,911)
1376           WRITE (NOUT10,912) IFAIL
1377           WRITE (NOUT10,913) XLIK
1378           WRITE (NOUT10,914) (XLIKD(J),J=1,NPAR)
1379           WRITE (NOUT10,917) (J,J=1,NPAR)
1380           WRITE (NOUT10,915) (PARAM(J),J=1,NPAR)
1381           WRITE (NOUT10,916) (SE(J),J=1,NPAR)
1382     C
1383           WRITE (NOUT10,918)
1384           DO 001 I=1,NPAR
1385.01  001    WRITE (NOUT10,919) I,(PARVAR(J,I),J=1,I)
1386.01  C
1387           IF (NPAR.EQ.1) GO TO 004
1388           WRITE (NOUT10,925)
1389           DO 003 I=2,NPAR
1390.01        IZ=I-1
1391.01        DO 002 J=1,IZ
1392.02        WSGN(J)=BLA
1393.02        IF (DABS(PARVAR(I,J)).GT.0.9D0) WSGN(J)=AST
1394.02  002   CONTINUE
1395.01  003   WRITE (NOUT10,926) I,(PARVAR(I,J),WSGN(J),J=1,IZ)
1396           WRITE (NOUT10,920) AST
1397     004   IF (NOUT.LT.2) RETURN
1398     C
1399     C        NOUT>1  OUTPUT Y,F,Z,VAR,NP,NQ
1400     C
1401           WRITE (NOUT10,927) NP,NQ
1402           JF=MINO(NPQA,4)
1403           JFF=MINO(NPQA,8)
1404           DO 005 I=1,NL
1405.01        W(I,3)=Y(I)-THETA(I,NQA)*Z(I)
1406.01        SGN=BLA
1407.01        IF (DABS(Z(I)).GT.2.5D0) SGN=AST
1408.01        WRITE (NOUT10,928) I,Y(I),F(I),W(I,3),Z(I),SGN,(VAR(I,J),J=1,JF)
```

```
1409.01          IF (NPQA.GT.4) WRITE (NOUT10,929) (VAR(I,J),J=5,JFF)
1410.01          IF (NPQA.GT.8) WRITE (NOUT10,930) (VAR(I,J),J=9,NPQA)
1411.01    005   CONTINUE
1412             WRITE (NOUT10,921) AST
1413             IF (NOUT.LT.3) RETURN
1414       C
1415       C        NOUT>2  OUTPUT PLOTS OF Y,F,Z,VAR
1416       C
1417             WRITE (NOUT10,931)
1418             MCHN(1)=1
1419             MCHN(2)=1
1420             MCHN(3)=1
1421             DO 006 I=1,NL
1422.01           MAXIS(I)=I
1423.01           W(I,1)=Y(I)
1424.01    006    W(I,2)=F(I)
1425             CALL PLTLNG(W,NLMAX,NL,3,VARMAX,VARMIN,0,MCHN,WCHR,1,MAXIS,1,0)
1426       C
1427             MCHN(2)=0
1428             MCHN(4)=2
1429             DO 007 I=1,NL
1430.01           W(I,1)=W(I,1)-F(I)
1431.01           W(I,3)=W(I,3)-F(I)
1432.01    007    W(I,4)=Z(I)
1433             CALL PLTLNG(W,NLMAX,NL,4,VARMAX,VARMIN,0,MCHN,WCHR,1,MAXIS,1,1)
1434       C
1435             MPL(1,1)=4
1436             MPL(2,1)=2
1437             CALL PLTACR(W,NLMAX,NL,4,VARMAX,VARMIN,0,MPL,1,MAXIS,0)
1438       C
1439             VARMAX(1)=0.0D0
1440             VARMIN(1)=0.0D0
1441             DO 008 I=1,NL
1442.01           DO 008 J=1,NPQA
1443.02           VARMAX(1)=DMAX1(VARMAX(1),VAR(I,J))
1444.02    008    VARMIN(1)=DMIN1(VARMIN(1),VAR(I,J))
1445             DO 009 J=1,NPQA
1446.01           VARMAX(J)=VARMAX(1)
1447.01    009    VARMIN(J)=VARMIN(1)
1448             DO 010 J=1,NPQA
1449.01           MCHN(J)=1
1450.01           L=NPQA+1-J
1451.01    010    VARCHR(J)=VARNUM(L)
1452             CALL PLTLNG(VAR,NLMAX,NL,NPQA,VARMAX,VARMIN,1,MCHN,VARCHR,1,
1453            1          MAXIS,1,1)
1454             RETURN
1455       C
1456       911   FORMAT (////'  R E S U L T S')
1457       912   FORMAT (//16X,'IFAIL',I4)
1458       913   FORMAT (/'  NEG. LOG-LIKELIHOOD ',1PG14.6)
1459       914   FORMAT (/'    FIRST DERIVATIVES ',1P7G14.6/22X,7G14.6/22X,7G14.6)
1460       917   FORMAT (/12X,7I14/12X,7I14/12X,7I14)
1461       915   FORMAT ('  PARAMETER ESTIMATES ',1P7G14.6/22X,7G14.6/22X,7G14.6)
1462       916   FORMAT ('      STANDARD ERRORS ',1P7G14.6/22X,7G14.6/22X,7G14.6)
1463       918   FORMAT (/'        (CO)VARIANCES ')
1464       919   FORMAT (15X,I6,1X,1P7G14.6/22X,7G14.6/22X,7G14.6)
1465       925   FORMAT (/'         CORRELATIONS ')
1466       926   FORMAT (15X,I6,1X,7(F10.6,A2,2X)/22X,7(F10.6,A2,2X)/
1467            1  22X,7(F10.6,A2,2X))
1468       920   FORMAT (/22X,'(CORRELATIONS EXCEEDING 0.9 IN MAGNITUDE',
1469            1 ' INDICATED BY ',A2,')')
1470       927   FORMAT ((///7X,'DATA(Y)',8X,'FIT(F)',5X,'PREDICTIONS',4X,
1471            1 'RESIDUALS(Z)',6X,'COVARIANCES(VAR) ( NP =',I2,' , NQ =',I2,' )')
1472       928   FORMAT (I4,1P3G14.6,0PF12.6,A2,4X,1P4G14.6)
```

```
1473     929   FORMAT (62X,1P4G14.6)
1474     930   FORMAT (60X,1P4G14.6)
1475     921   FORMAT (/22X,'(RESIDUALS Z EXCEEDING 2.5 IN MAGNITUDE',
1476          1 ' INDICATED BY ',A2,')')
1477     931   FORMAT (////'  THERE FOLLOW:'
1478          1 /' (A)  A LONGITUDINAL PLOT OF DATA(Y),FIT(F),PREDICTION     ',
1479          1 '         (REVEALS LACK OF FIT)'
1480          1 /' (B)  A LONGITUDINAL PLOT OF DATA-FIT,PREDICTION-FIT,RESIDU',
1481          1 'ALS(Z)   (REVEALS CORRELATION OF ERRORS)'
1482          1 /' (C)  A CROSS-SECTIONAL PLOT OF RESIDUALS(Z) AGAINST FIT(F)',
1483          1 '         (REVEALS HETEROGENEITY OF VARIANCE AND OUTLIERS)'
1484          1 /' (D)  A LONGITUDINAL PLOT OF COVARIANCES OF ERRORS         ',
1485          1 '         (DISPLAYS FITTED COVARIANCE STRUCTURE)')
1486           END
1487
1488           SUBROUTINE REGCVI(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,Y,F,PARAM,
1489          1              PARAMM,PARAMS,PARAMT,MLEST,NPAR,NM,NDER,NOUT,IFAIL)
1490     C
1491     C
1492     C-----------------------------------------------------------------------
1493     C                        REGCVI
1494     C                        ======
1495     C
1496     C        CREATES INPUT FILES FOR CEVOPE.
1497     C        RESULTS OUTPUT TO NOUT9:  NL,IFAIL          (2I5)
1498     C                                  E(1...NL)        (5G24.16)
1499     C                                  NM
1500     C                                  PARAM(1...NM)
1501     C                                  DERIV(1...NL,J)
1502     C                                                   J=1...NM
1503     C                      TRANSFORMED  DERIV(1...NL,J)
1504     C                                                   J=1...NM
1505     C
1506     C                           NOUT8:  NPA,NQA
1507     C                                   PHI(1...NL,J)
1508     C                                                J=1...NPA
1509     C                                   THETA(1...NL,J)
1510     C                                                J=1...NQA
1511     C
1512     C  INPUT
1513     C   VAR(NLMAX,11)         BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
1514     C   NLMAX                 FIRST DIMENSION OF VAR
1515     C   NL                    NUMBER OF OBSERVATIONS
1516     C   NPQA                  NP+NQ+1
1517     C   NPA                   NP+1
1518     C   NQA                   NQ+1
1519     C   NP                    GENERALIZED AUTOREGRESSIVE ORDER
1520     C   NQ                    GENERALIZED MOVING AVERAGE ORDER
1521     C   Y(NL)                 DATA VECTOR OF OBSERVATIONS
1522     C   F(NL)                 VECTOR OF FITTED VALUES
1523     C   PARAM(NPAR)           PARAMETER ESTIMATES
1524     C   PARAMM(NPAR)          LOCATION SHIFT TRANSFORMING PARAM TO PARAMT
1525     C   PARAMS(NPAR)          SCALE SHIFT TRANSFORMING PARAM TO PARAMT
1526     C   PARAMT(NPAR)          RESCALED ESTIMATED PARAM
1527     C   MLEST(16)             LOCATION OF ESTIMATED PARAMETERS
1528     C   NPAR                  NUMBER OF MODEL PARAMETERS
1529     C   NM                    NUMBER OF REGRESSION PARAMETERS
1530     C   NDER                  1  IF REGDER AVAILABLE
1531     C   NOUT                  >0 OUTPUT TO NOUT10
1532     C                         >3 INTERMEDIATE RESULTS OUTPUT
1533     C   IFAIL                 FAILURE INDICATOR IN REGAME
1534     C
1535     C-----------------------------------------------------------------------
1536     C
```

```
1537           DOUBLE PRECISION
1538          1    VAR(NLMAX,11),Y(NL),F(NL),PARAM(NPAR),PARAMM(NPAR),
1539          1 PARAMS(NPAR),PARAMT(NPAR),
1540          1    DD(6,5),DDI(5,5),DERIV(200,5),DERIVT(200,5),E(200),
1541          1 EPSL,PHI(200,6),THETA(200,6),W1(200),W2(200),WT(200),
1542          1 XLDET,XX,XY
1543           INTEGER
1544          1    NLMAX,NL,NPQA,NPA,NQA,NP,NQ,MLEST(16),NPAR,NM,NDER,NOUT,IFAIL,
1545          1    I,IFAILP,J,K,NDD,NDDI,NFDIF,NPRI,NTEQ,NWRITE
1546           COMMON
1547          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
1548     C
1549     C
1550     C-----------------------------------------------------------------------
1551     C  LOCAL
1552     C   DD(6,5)              SUM OF SQUARES OF DERIV
1553     C   DDI(5,5)             INVERSE OF DD
1554     C   DERIV(200,5)         FIRST DERIVATIVE OF F(NL) W.R.T. PARAM(NM)
1555     C   DERIVT(200,5)        TRANSFORM OF DERIV
1556     C   E(200)               VECTOR OF WEIGHT CORRECTED DEPARTURES
1557     C   EPSL                 STEPLENGTH USED TO CALC. DERIV IF NDER NOT 1
1558     C   I                    INDEX
1559     C   IFAILP               FAILURE INDICATOR
1560     C   J                    INDEX
1561     C   K                    INDEX
1562     C   NDD                  DIMENSION OF DD
1563     C   NDDI                 DIMENSION OF DDI
1564     C   NFDIF                SET TO 1 WHEN CALLING REGMOD IF NDER NOT 1
1565     C   NPRI                 1  IF INTERMEDIATE RESULTS OUTPUT
1566     C   NTEQ                 1  IF ALL ROWS OF PHI EQUAL
1567     C   NWRITE               SET TO LOWER OF NM,NPAR
1568     C   PHI(200,6)           FIRST MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
1569     C   THETA(200,6)         SECOND MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
1570     C   W1(200)              WORK SPACE
1571     C   W2(200)              WORK SPACE
1572     C   WT(200)              VECTOR OF WEIGHTS
1573     C   XLDET                LOG DETERMINANT OF THETA
1574     C   XX                   WORK SPACE
1575     C   XY                   WORK SPACE
1576     C
1577     C
1578     C-----------------------------------------------------------------------
1579           DATA NDD,NDDI/6,5/
1580           NPRI=0
1581           IF (NOUT.GE.4) NPRI=1
1582     C
1583     C        WEIGHT CORRECTED DEPARTURES,STANDARDIZED TO SUM
1584     C        OF SQUARES OF UNITY
1585     C
1586           XX=0.0D0
1587           DO 001 I=1,NL
1588.01         WT(I)=DSQRT(VAR(I,NPQA))
1589.01         E(I)=(Y(I)-F(I))/WT(I)
1590.01  001    XX=XX+E(I)*E(I)
1591           XX=DSQRT(XX/FLOAT(NL))
1592           DO 002 I=1,NL
1593.01         WT(I)=WT(I)*XX
1594.01  002    E(I)=E(I)/XX
1595           WRITE (NOUT9,903) NL,IFAIL
1596     903   FORMAT (2I5)
1597           WRITE (NOUT9,904) (E(I),I=1,NL)
1598     904   FORMAT (1P5G24.16)
1599     C
1600     C        PARAMETER ESTIMATES
```

```
1601     C
1602           NWRITE=MINO(NM,NPAR)
1603           WRITE (NOUT9,903) NM,NWRITE
1604           IF (NWRITE.GT.0) WRITE (NOUT9,904) (PARAM(J),J=1,NWRITE)
1605     C
1606     C        WEIGHT CORRECTED PHI, THETA DECOMPOSITION OF VAR
1607     C
1608           DO 005 I=1,NL
1609.01         DO 005 J=1,NPQA
1610.02         K=I+J-NPQA
1611.02         XX=1.0D0
1612.02         IF (K.GT.0) XX=WT(K)
1613.02  005    VAR(I,J)=VAR(I,J)/(WT(I)*XX)
1614           NTEQ=0
1615           IFAILP=0
1616           CALL REGAMD(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,NTEQ,PHI,THETA,XLDET,
1617          1            NOUT,NPRI,IFAILP)
1618     C
1619           IF (IFAILP.NE.0) GO TO 021
1620           WRITE (NOUT8,903) NPA,NQA
1621           DO 006 J=1,NPA
1622.01  006    WRITE (NOUT8,904) (PHI(I,J),I=1,NL)
1623           DO 007 J=1,NQA
1624.01  007    WRITE (NOUT8,904) (THETA(I,J),I=1,NL)
1625.01  C
1626.01  C        WEIGHT CORRECTED FIRST DERIVATIVES
1627.01  C
1628           IF (NM.LE.0) RETURN
1629           IF (NDER.NE.1) GO TO 008
1630           IFAILP=0
1631           CALL REGDER(PARAM,NPAR,NLMAX,NL,NM,DERIV,NOUT,IFAILP)
1632           IF (IFAILP.NE.0) GO TO 023
1633           GO TO 011
1634     C
1635     008   EPSL=1.0D-4
1636           NFDIF=1
1637           IFAILP=0
1638           K=1
1639           DO 010 J=1,NM
1640.01         XX=EPSL
1641.01         IF (MLEST(K).EQ.J) XX=EPSL*PARAMS(J)
1642.01         PARAM(J)=PARAM(J)+XX
1643.01         CALL REGMOD(PARAM,NPAR,NLMAX,NL,NFDIF,W1,VAR,NPQA,NP,NQ,NTEQ,
1644.01        1            NOUT,IFAILP)
1645.01         XY=0.0D0
1646.01         IF (MLEST(K).NE.J) GO TO 009
1647.01         XY=PARAMT(K)
1648.01         K=K+1
1649.01  009    PARAM(J)=XY*PARAMS(J)+PARAMM(J)
1650.01         IF (IFAILP.NE.0) GO TO 025
1651.01         DO 010 I=1,NL
1652.02  010    DERIV(I,J)=(W1(I)-F(I))/XX
1653.02  C
1654     011   IF (NPRI.EQ.1) CALL MATPR2(DERIV,NLMAX,NL,NM,1)
1655           DO 014 J=1,NM
1656.01         DO 012 I=1,NL
1657.02         W1(I)=(DERIV(I,J))/WT(I)
1658.02  012    DERIV(I,J)=W1(I)
1659.01         WRITE (NOUT9,904) (DERIV(I,J),I=1,NL)
1660.01  C
1661.01  C        TRANSFORMED DERIVATIVES
1662.01  C
1663.01         CALL REGTRA(W1,NL,PHI,NLMAX,NPA,0,0,W2,NOUT,NPRI,IFAILP)
1664.01         CALL REGTRA(W2,NL,THETA,NLMAX,NQA,1,0,W1,NOUT,NPRI,IFAILP)
```

```
1665.01          IF (IFAILP.NE.0) GO TO 027
1666.01          CALL REGTRA(W1,NL,THETA,NLMAX,NQA,1,1,W2,NOUT,NPRI,IFAILP)
1667.01          IF (IFAILP.NE.0) GO TO 027
1668.01          CALL REGTRA(W2,NL,PHI,NLMAX,NPA,0,1,W1,NOUT,NPRI,IFAILP)
1669.01          DO 013 I=1,NL
1670.02  013     DERIVT(I,J)=W1(I)
1671.01  014     CONTINUE
1672.01  C
1673             DO 016 J=1,NM
1674.01          DO 016 K=J,NM
1675.02          XX=0.0D0
1676.02          DO 015 I=1,NL
1677.03  015     XX=XX+DERIVT(I,J)*DERIV(I,K)
1678.02          DD(K,J)=XX
1679.02  016     DD(J,K)=XX
1680             IF (NPRI.EQ.1) CALL MATPR2(DD,NDD,NM,NM,1)
1681             IFAILP=1
1682             CALL F01ABF(DD,NDD,NM,DDI,NDDI,W1,IFAILP)
1683             IF (IFAILP.NE.0) GO TO 029
1684             DO 017 J=1,NM
1685.01          DO 017 K=J,NM
1686.02  017     DDI(J,K)=DDI(K,J)
1687             IF (NPRI.EQ.1) CALL MATPR2(DDI,NDDI,NM,NM,1)
1688     C
1689             DO 020 J=1,NM
1690.01          DO 019 I=1,NL
1691.02          XX=0.0D0
1692.02          DO 018 K=1,NM
1693.03  018     XX=XX+DERIVT(I,K)*DDI(K,J)
1694.02  019     DERIV(I,J)=XX
1695.01  020     WRITE (NOUT9,904) (DERIV(I,J),I=1,NL)
1696             IF (NPRI.EQ.1) CALL MATPR2(DERIV,NLMAX,NL,NM,1)
1697             RETURN
1698     C
1699     C        ERROR MESSAGES
1700     C
1701     021     WRITE (NOUT6,922)
1702             IF (NOUT.GE.1) WRITE (NOUT10,922)
1703     922     FORMAT (' *** REGCVI, REGAMD FAILS ***')
1704             RETURN
1705     023     WRITE (NOUT6,924)
1706             IF (NOUT.GE.1) WRITE (NOUT10,924)
1707     924     FORMAT (' *** REGCVI, REGDER FAILS ***')
1708             RETURN
1709     025     WRITE (NOUT6,926)
1710             IF (NOUT.GE.1) WRITE (NOUT10,926)
1711     926     FORMAT (' *** REGCVI, REGMOD FAILS ***')
1712             RETURN
1713     027     WRITE (NOUT6,928)
1714             IF (NOUT.GE.1) WRITE (NOUT10,928)
1715     928     FORMAT (' *** REGCVI, REGTRA FAILS ***')
1716             RETURN
1717     029     WRITE (NOUT6,930)
1718             IF (NOUT.GE.1) WRITE (NOUT10,930)
1719     930     FORMAT (' *** REGCVI, F01ABF FAILS, MATRIX NOT POS. DEF. ***')
1720             RETURN
1721             END
1722
1723             SUBROUTINE REGAMD(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,NTEQ,PHI,
1724            1                  THETA,XLDET,NOUT,NPRI,IFAIL)
1725     C
1726     C-----------------------------------------------------------------------
1727     C
1728     C                          REGAMD
```

```
C                       ======
C
C       GENERALIZED AUTOREGRESSIVE-MOVING AVERAGE DECOMPOSITION
C       DECOMPOSES VAR INTO LOWER TRIANGULAR BAND MATRICES
C       PHI AND THETA.
C
C  INPUT
C   VAR(NLMAX,NPQA)       BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
C   NLMAX                 FIRST DIMENSION OF VAR,PHI,THETA
C   NL                    NUMBER OF OBSERVATIONS
C   NPQA                  NP+NQ+1
C   NPA                   NP+1
C   NQA                   NQ+1
C   NP                    GENERALISED AUTOREGRESSIVE ORDER
C   NQ                    GENERALISED MOVING AVERAGE ORDER
C   NTEQ                  1  IF ALL ROWS OF PHI ARE EQUAL
C
C  OUTPUT
C   PHI(NLMAX,NPA)        FIRST MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
C   THETA(NLMAX,NQA)      SECOND MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
C   XLDET                 LOG DETERMINANT OF THETA
C
C  CONTROL
C   NOUT                  >0 ERROR MESSAGES OUTPUT TO CHANNEL NOUT10
C   NPRI                  1  OUTPUT OF INTERMEDIATE RESULTS
C   IFAIL                 1  IF VAR NOT POSITIVE DEFINITE
C
C
C-----------------------------------------------------------------------
      DOUBLE PRECISION
     1    VAR(NLMAX,NPQA),PHI(NLMAX,NPA),THETA(NLMAX,NQA),XLDET,
     1    PHIVAR(11),TOL,W(5,5),W1(5),W2(5),WA(5,5),WB(5),WC(5),
     1 X02AAF,XX
      INTEGER
     1    NLMAX,NL,NPQA,NPA,NQA,NP,NQ,NTEQ,NOUT,NPRI,IFAIL,
     1    I,IJ,IL,ILK,IS,J,JA,JAF,JF,JP,K,KA,KAF,KP,L,LA,LAF,
     1 LAQ,LKA,LKAF,LQ,NCJGF,NDET,NJGF,NPMAX,NV1,NV2,NWMAX
      LOGICAL
     1    LOGJGF
      COMMON
     1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
C
C
C-----------------------------------------------------------------------
C  LOCAL
C   I                     LENGTH INDEX
C   IJ                    I-J
C   IL                    I-L
C   ILK                   I-L-K
C   IS                    LOWER BOUND ON I DO-LOOP
C   J                     INDEX
C   JA                    J+1
C   JAF                   UPPER BOUND ON JA DO-LOOP
C   JF                    UPPER BOUND ON J DO-LOOP
C   JP                    NPA-J
C   K                     INDEX
C   KA                    K+1
C   KAF                   UPPER BOUND ON KA DO-LOOP
C   KP                    NPA-K
C   L                     INDEX
C   LA                    L+1
C   LAF                   UPPER BOUND ON LA DO-LOOP
C   LAQ                   L+1-NQA
C   LKA                   L+K+1
```

```
1793     C   LKAF                UPPER BOUND ON LKA DO-LOOP
1794     C   LOGJGF              USED BY F04JGF (LOGICAL)
1795     C   LQ                  NQA-L
1796     C   NCJGF               NUMBER OF CALLS TO F04JGF
1797     C   NDET                POWER IN DETERMINANT OF THETA
1798     C   NJGF                1 IF F04JGF CALLED
1799     C   NPMAX               MAXIMUM OF NP
1800     C   NV1                 FIRST TERM IN VAR
1801     C   NV2                 SECOND TERM IN VAR
1802     C   NWMAX               NUMBER OF ELEMENTS IN W
1803     C   PHIVAR(11)          WORK SPACE
1804     C   TOL                 ACCURACY USED BY F04JGF
1805     C   W(5,5)              WORK SPACE
1806     C   W1(5)               WORK SPACE
1807     C   W2(5)               WORK SPACE
1808     C   WA(5,5)             WORK SPACE
1809     C   WB(5)               WORK SPACE
1810     C   WC(5)               WORK SPACE
1811     C   X02AAF              SUPPLIES ACC. TO WHICH SOLN. OF F04JGF REQUIRED
1812     C   XX                  WORK SPACE
1813     C
1814     C
1815     C-----------------------------------------------------------------------
1816     C
1817     C                    PHI
1818     C                    ===
1819     C
1820     C        CALCULATE RECTANGULAR MATRIX PHI(NL,NPA) FROM
1821     C        VAR(NL,NPQA) BY SOLVING SIMULTANEOUS EQUATIONS
1822     C
1823     C        FOR L FROM NQA TO NQA-1+MIN(NP,I-NQA)
1824     C
1825     C        SUMJ PHI(I,NPA-J)*COV(I-J,I-L)= -COV(I,I-L)
1826     C
1827     C        SUMJ FROM 1 TO MIN(NP,I-NQA)
1828     C
1829     C        WHERE COV(NC1,NC2)=VAR(NV1,NV2)
1830     C                      NV1 = MAX(NC1,NC2)
1831     C                      NV2 = NPQA-DIF(NC1,NC2)
1832     C
1833     C        FOR I FROM NQA+1 TO NL
1834     C
1835     C
1836           DATA NPMAX,NWMAX/5,25/
1837           DO 002 I=1,NPQA
1838.01        DO 001 J=1,NP
1839.02  001   PHI(I,J)=0.0D0
1840.01  002   PHI(I,NPA)=1.0D0
1841.01  C
1842           IS=NQA+1
1843           NJGF=0
1844           NCJGF=0
1845           IF (IS.GT.NL) GO TO 011
1846           DO 008 I=IS,NL
1847.01        PHI(I,NPA)=1.0D0
1848.01        IF (NP.EQ.0) GO TO 008
1849.01        IF ((NTEQ.EQ.1).AND.(I.GT.NPQA).AND.(NJGF.EQ.0)) GO TO 006
1850.01  C
1851.01        JF=MINO(NP,I-NQA)
1852.01        DO 004 LAQ=1,JF
1853.02        L=NQA+LAQ-1
1854.02        IL=I-L
1855.02        DO 003 J=1,JF
1856.03        IJ=I-J
```

```
1857.03         NV1=MAXO(IJ,IL)
1858.03         NV2=NPQA-IABS(IJ-IL)
1859.03  003    WA(LAQ,J)=VAR(NV1,NV2)
1860.02         NV2=NPQA-L
1861.02  004    WB(LAQ)=-VAR(I,NV2)
1862.02  C
1863.01         NJGF=0
1864.01         IFAIL=1
1865.01         CALL F04ATF(WA,NPMAX,WB,JF,WC,W,NPMAX,W1,W2,IFAIL)
1866.01         IF (IFAIL.EQ.0) GO TO 006
1867.01  C
1868.01  C        IF SIMULTANEOUS EQNS. NOT OF FULL RANK USE F04JGF
1869.01  C
1870.01         NJGF=1
1871.01         NCJGF=NCJGF+1
1872.01         IFAIL=1
1873.01         TOL=X02AAF(XX)
1874.01         DO 005 J=1,JF
1875.02  005    WC(J)=WB(J)
1876.01         CALL F04JGF(JF,JF,WA,NPMAX,WC,TOL,LOGJGF,XX,J,W,NWMAX,IFAIL)
1877.01         IF (IFAIL.NE.0) GO TO 021
1878.01  C
1879.01  006    DO 007 J=1,JF
1880.02         JP=NPA-J
1881.02  007    PHI(I,JP)=WC(J)
1882.01  008    CONTINUE
1883.01  C
1884            IF (NCJGF.EQ.0) GO TO 011
1885            WRITE (NOUT6,910) NCJGF
1886            IF (NOUT.GE.1) WRITE (NOUT10,910) NCJGF
1887     910    FORMAT (' *** REGAMD: AUTOREGRESSIVE ORDER<NP IN',I5,' ROWS ***')
1888     011    IF (NPRI.NE.1) GO TO 012
1889     C
1890            CALL MATPR2(VAR,NLMAX,NL,NPQA,1)
1891            CALL MATPR2(PHI,NLMAX,NL,NPA,1)
1892     C
1893     C                      THETA
1894     C                      =====
1895     C
1896     C        CALCULATE RECTANGULAR MATRIX THETA(NL,NQA) FROM
1897     C        MATRICES PHI(NL,NPA) AND VAR(NL,NPQA) USING
1898     C
1899     C        THETA(I,NQA-L) = SUMK PHI(I-L,NPA-K)*PHIVAR(L+K+1)
1900     C
1901     C                         SUMK FROM 0 TO MIN(NPA,I-L,NQA-L)-1
1902     C        FOR L FROM 0 TO MIN(NQA,I)-1
1903     C
1904     C        WHERE
1905     C        PHIVAR(L+K+1) = SUMJ PHI(I,NPA-J)*COV(I-J,I-L-K)
1906     C
1907     C                        SUMJ FROM 0 TO MIN(NPA,I)-1
1908     C        FOR L+K+1 FROM 1 TO MIN(NPQA,I)
1909     C
1910     C
1911     C        FOR I FROM 1 TO NL
1912     C
1913     C
1914     012    DO 013 I=1,NQA
1915.01          DO 013 J=1,NQA
1916.02  013    THETA(I,J)=0.0D0
1917.02  C
1918            DO 017 I=1,NL
1919.01         LKAF=MIN0(NPQA,I)
1920.01         DO 015 LKA=1,LKAF
```

```
1921.02          ILK=I+1-LKA
1922.02          XX=0.0D0
1923.02          JAF=MIN0(NPA,I)
1924.02          DO 014 JA=1,JAF
1925.03          J=JA-1
1926.03          IJ=I-J
1927.03          NV1=MAX0(IJ,ILK)
1928.03          NV2=NPQA-IABS(IJ-ILK)
1929.03          JP=NPA-J
1930.03  014     XX=XX+PHI(I,JP)*VAR(NV1,NV2)
1931.02  015     PHIVAR(LKA)=XX
1932.02  C
1933.01          LAF=MIN0(NQA,I)
1934.01          DO 017 LA=1,LAF
1935.02          L=LA-1
1936.02          IL=I-L
1937.02          XX=0.0D0
1938.02          KAF=MIN0(NPA,IL,NQA-L)
1939.02          DO 016 KA=1,KAF
1940.03          K=KA-1
1941.03          KP=NPA-K
1942.03          LKA=LA+K
1943.03  016     XX=XX+PHI(IL,KP)*PHIVAR(LKA)
1944.02          LQ=NQA-L
1945.02  017     THETA(I,LQ)=XX
1946.02  C
1947             IF (NPRI.EQ.1) CALL MATPR2(THETA,NLMAX,NL,NQA,1)
1948     C
1949     C           CHOLESKY DECOMPOSITION OF BAND MATRIX THETA
1950     C
1951             IFAIL=1
1952             CALL F03AGF(NL,NQ,THETA,NLMAX,THETA,NLMAX,NQA,XLDET,NDET,IFAIL)
1953             IF (IFAIL.NE.0) GO TO 019
1954     C
1955             DO 018 I=1,NL
1956.01  018     THETA(I,NQA)=1.0D0/THETA(I,NQA)
1957             XLDET=(DLOG(XLDET)+FLOAT(NDET)*DLOG(2.0D0))/2.0D0
1958             IF (NPRI.NE.1) RETURN
1959     C
1960             CALL MATPRO(XLDET,1)
1961             CALL MATPR2(THETA,NLMAX,NL,NQA,1)
1962             RETURN
1963     C
1964     C           ERROR MESSAGES
1965     C
1966     021     WRITE (NOUT6,922)
1967             IF (NOUT.GE.1) WRITE (NOUT10,922)
1968     922     FORMAT (' *** REGAMD: F04JGF FAILS, PHI CANNOT BE FOUND ***')
1969             IFAIL=1
1970             RETURN
1971     019     WRITE (NOUT6,920)
1972             IF (NOUT.GE.1) WRITE (NOUT10,920)
1973     920     FORMAT (' *** REGAMD: VARIANCE MATRIX NOT POSITIVE DEFINITE ***')
1974             RETURN
1975             END
1976
1977             SUBROUTINE REGTRA(X,NL,BAND,NLMAX,NBAND,NINV,NTRA,Y,NOUT,NPRI,
1978            1                  IFAIL)
1979     C
1980     C-----------------------------------------------------------------------
1981     C
1982     C                          REGTRA
1983     C                          ======
1984     C
```

```
1985  C        TRANSFORM VECTOR  X  TO VECTOR  Y  BY PRE-MULTIPLYING BY
1986  C        A LOWER TRIANGULAR MATRIX  BAND  WITH BANDWIDTH NBAND WHICH
1987  C        CAN BE INVERTED (NINV=1) AND/OR TRANSPOSED (NTRA=1).
1988  C
1989  C  INPUT
1990  C   X(NL)                 VECTOR TO BE TRANSFORMED
1991  C   NL                    DIMENSION OF X,Y (BETWEEN 1 AND NLMAX)
1992  C   BAND(NLMAX,NBAND)     LOWER TRIANGULAR BAND MATRIX
1993  C   NLMAX                 FIRST DIMENSION OF BAND (> 0)
1994  C   NBAND                 BAND WIDTH OF BAND (> 0)
1995  C   NINV                  1  IF INVERSE OF BAND TO BE USED
1996  C   NTRA                  1 IF TRANSPOSE OF BAND TO BE USED
1997  C
1998  C  OUTPUT
1999  C   Y(NL)                 TRANSFORMED VECTOR
2000  C
2001  C  CONTROL
2002  C   NOUT                  >0 OUTPUT TO NOUT10
2003  C   NPRI                  1  OUTPUT OF INTERMEDIATE RESULTS
2004  C   IFAIL                 1  IF NINV=1 BUT RANK(BAND)<NL
2005  C
2006  C
2007  C-------------------------------------------------------------------
2008        DOUBLE PRECISION
2009       1    X(NL),BAND(NLMAX,NBAND),Y(NL),
2010       1    XX
2011        INTEGER
2012       1    NL,NLMAX,NBAND,NINV,NTRA,NOUT,NPRI,IFAIL,
2013       1    I,IA,IZ,J,JF,K,KS,L,NBANDA,NBANDZ,NLZ
2014        COMMON
2015       1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
2016  C
2017  C-------------------------------------------------------------------
2018  C
2019  C  LOCAL
2020  C   I                     INDEX
2021  C   IA                    I+1
2022  C   IZ                    I-1
2023  C   J                     INDEX
2024  C   JF                    UPPER BOUND ON J DO-LOOPS
2025  C   K                     INDEX
2026  C   KS                    LOWER BOUND ON K DO-LOOPS
2027  C   L                     INVERSE/ TRANSPOSE OPTION INDEX
2028  C   NBANDA                NBAND+1
2029  C   NBANDZ                NBAND-1
2030  C   NLZ                   NL-1
2031  C   XX                    WORK SPACE
2032  C
2033  C-------------------------------------------------------------------
2034  C
2035        IF ((NL.LT.1).OR.(NL.GT.NLMAX)) GO TO 017
2036        IF (NBAND.LT.1) GO TO 019
2037        IF (NLMAX.LT.1) GO TO 021
2038        IFAIL=0
2039        NBANDZ=NBAND-1
2040        NBANDA=NBAND+1
2041        L=1
2042        IF (NTRA.EQ.1) L=L+1
2043        IF (NINV.EQ.1) L=L+2
2044        GO TO (001,004,007,012),L
2045  C
2046  C     NINV = 0   NTRA = 0
2047  C     ========   ========
2048  C
```

```
2049      C     Y(I) = SUMK BAND(I,K) * X(I+K-NBAND)
2050      C
2051      C           SUMK FROM MAX(NBAND+1-I,1) TO NBAND
2052      C
2053      C
2054      001   DO 003 I=1,NL
2055.01         XX=0.0D0
2056.01         KS=MAXO(NBANDA-I,1)
2057.01         DO 002 K=KS,NBAND
2058.02         J=I+K-NBAND
2059.02   002   XX=XX+BAND(I,K)*X(J)
2060.01   003   Y(I)=XX
2061            IF (NPRI.EQ.1) CALL MATPR1(Y,NL,1)
2062            RETURN
2063      C
2064      C     NINV = 0   NTRA = 1
2065      C     ========   ========
2066      C
2067      C     Y(I) = SUMJ BAND(J,NBAND+I-J) * X(J)
2068      C
2069      C           SUMJ FROM I TO MIN(I+NBAND-1,NL)
2070      C
2071      C
2072      004   DO 006 I=1,NL
2073.01         XX=0.0D0
2074.01         JF=MINO(I+NBANDZ,NL)
2075.01         DO 005 J=I,JF
2076.02         K=NBAND+I-J
2077.02   005   XX=XX+BAND(J,K)*X(J)
2078.01   006   Y(I)=XX
2079            IF (NPRI.EQ.1) CALL MATPR1(Y,NL,1)
2080            RETURN
2081      C
2082      C     NINV = 1   NTRA = 0
2083      C     ========   ========
2084      C
2085      C     Y(I) = ( X(I) - SUMK BAND(I,K) * Y(I+K-NBAND) ) / BAND(I,NBAND)
2086      C
2087      C                 SUMK FROM MAX(NBAND+1-I,1) TO NBAND-1
2088      C
2089      C
2090      007   IF (BAND(1,NBAND).EQ.0.0D0) GO TO 023
2091            Y(1)=X(1)/BAND(1,NBAND)
2092            IF (NL.EQ.1) GO TO 011
2093            DO 010 I=2,NL
2094.01         XX=X(I)
2095.01         IF (NBAND.EQ.1) GO TO 009
2096.01         KS=MAXO(NBANDA-I,1)
2097.01         DO 008 K=KS,NBANDZ
2098.02         J=I+K-NBAND
2099.02   008   XX=XX-BAND(I,K)*Y(J)
2100.01   009   IF (BAND(I,NBAND).EQ.0.0D0) GO TO 023
2101.01   010   Y(I)=XX/BAND(I,NBAND)
2102      011   IF (NPRI.EQ.1) CALL MATPR1(Y,NL,1)
2103            RETURN
2104      C
2105      C     NINV = 1   NTRA = 1
2106      C     ========   ========
2107      C
2108      C     Y(I) = ( X(I) - SUMJ BAND(J,NBAND+I-J) * Y(J) ) / BAND(I,NBAND)
2109      C
2110      C                 SUMJ FROM I+1 TO MIN(I+NBAND-1,NL)
2111      C
2112      C
```

```
      012   NLZ=NL-1
            IF (BAND(NL,NBAND).EQ.0.0D0) GO TO 023
            Y(NL)=X(NL)/BAND(NL,NBAND)
            IF (NL.EQ.1) GO TO 016
            DO 015 IZ=1,NLZ
            I=NL-IZ
            XX=X(I)
            IF (NBAND.EQ.1) GO TO 014
            IA=I+1
            JF=MINO(I+NBANDZ,NL)
            DO 013 J=IA,JF
            K=NBAND+I-J
      013   XX=XX-BAND(J,K)*Y(J)
      014   IF (BAND(I,NBAND).EQ.0.0D0) GO TO 023
      015   Y(I)=XX/BAND(I,NBAND)
      016   IF (NPRI.EQ.1) CALL MATPR1(Y,NL,1)
            RETURN
      C
      C        ERROR MESSAGES
      C
      017   WRITE (NOUT6,918) NL
            IF (NOUT.GE.1) WRITE (NOUT10,918) NL
      918   FORMAT (' *** REGTRA: NL =',I5,' OUT OF RANGE ***')
            IFAIL=1
            RETURN
      019   WRITE (NOUT6,920) NBAND
            IF (NOUT.GE.1) WRITE (NOUT10,920) NBAND
      920   FORMAT (' *** REGTRA: NBAND =',I5,' OUT OF RANGE ***')
            IFAIL=1
            RETURN
      021   WRITE (NOUT6,922) NLMAX
            IF (NOUT.GE.1) WRITE (NOUT10,922) NLMAX
      922   FORMAT (' *** REGTRA: NLMAX =',I5,' OUT OF RANGE ***')
            IFAIL=1
            RETURN
      023   WRITE (NOUT6,924)
            IF (NOUT.GE.1) WRITE (NOUT10,924)
      924   FORMAT (' *** REGTRA: BAND MATRIX NOT OF FULL RANK ***')
            IFAIL=1
            RETURN
            END

            SUBROUTINE REGGEN(PARAM,NPAR,SCALE,NSEED,NL,Y,NOUT,IFAIL)
      C
      C
      C-----------------------------------------------------------------------
      C                           REGGEN
      C                           ======
      C
      C        TO SIMULATE DATA WITH GENERALIZED AUTOREGRESSIVE-MOVING
      C        AVERAGE ERRORS, CALLING REGMOD.
      C        RESULTS OUTPUT TO NOUT7:  NL,NSEED  (2I5)
      C                                  Y(1...NL) (5G24.16)
      C
      C  INPUT
      C   PARAM(NPAR)          VECTOR OF MODEL PARAMETERS
      C   NPAR                 NUMBER OF MODEL PARAMETERS
      C   SCALE                SCALE PARAMETER
      C   NSEED                SEED TO SET RANDOM NUMBER GENERATOR
      C   NL                   LENGTH OF VECTOR TO BE SIMULATED
      C                                       (BETWEEN 1 AND 200)
      C
      C  OUTPUT
      C   Y(NL)                VECTOR OF SIMULATED OBSERVATIONS
```

```
2177     C
2178     C  CONTROL
2179     C   NOUT                >0 ERROR MESSAGES OUTPUT TO NOUT10
2180     C                       5  INTERMEDIATE RESULTS OUTPUT
2181     C   IFAIL               1  IF REGGEN FAILS
2182     C
2183     C
2184     C-----------------------------------------------------------------------
2185           DOUBLE PRECISION
2186          1    PARAM(NPAR),SCALE,Y(NL),
2187          1    E(200),F(200),G05DDF,PHI(200,6),THETA(200,6),VAR(200,11),
2188          1 W1(200),XLDET,Z(200)
2189           INTEGER
2190          1    NPAR,NSEED,NL,NOUT,IFAIL,
2191          1    I,J,NFDIF,NLMAX,NP,NPA,NPMAX,NPQA,NPRI,NQ,NQA,NQMAX,NTEQ
2192           COMMON
2193          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
2194     C
2195     C
2196     C-----------------------------------------------------------------------
2197     C  LOCAL
2198     C   E(200)              VECTOR OF DEPARTURES
2199     C   F(200)              VECTOR OF FITTED VALUES
2200     C   G05DDF              SUPPLIES SIMULATED STANDARD NORMAL DEVIATES
2201     C   I                   INDEX
2202     C   J                   INDEX
2203     C   NFDIF               0  BECAUSE BOTH F AND VAR REQUIRED
2204     C   NLMAX               MAXIMUM SIZE ALLOWED FOR NL (=200)
2205     C   NP                  GENERALIZED AUTOREGRESSIVE ORDER (MAX=5)
2206     C   NPA                 NP+1
2207     C   NPMAX               MAXIMUM SIZE ALLOWED FOR NP (=5)
2208     C   NPQA                NP+NQ+1
2209     C   NPRI                1  FOR INTERMEDIATE RESULTS OUTPUT
2210     C   NQ                  GENERALIZED MOVING AVERAGE ORDER (MAX=5)
2211     C   NQA                 NQ+1
2212     C   NQMAX               MAXIMUM SIZE ALLOWED FOR NQ (=5)
2213     C   NTEQ                1  IF ALL ROWS OF PHI EQUAL
2214     C   PHI(200,6)          FIRST MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
2215     C   THETA(200,6)        SECOND MATRIX IN TRIANGULAR BAND DECOMP. OF VAR
2216     C   VAR(200,11)         BELOW DIAGONAL ELEMENTS OF VARIANCE MATRIX
2217     C   W1(200)             WORK SPACE
2218     C   XLDET               LOG DETERMINANT OF THETA
2219     C   Z(200)              VECTOR OF SIMULATED RESIDUALS FROM MODEL
2220     C
2221     C
2222     C-----------------------------------------------------------------------
2223           DATA NLMAX,NPMAX,NQMAX/200,5,5/
2224           IF ((NL.LT.1).OR.(NL.GT.NLMAX)) GO TO 006
2225           NPRI=0
2226           IF (NOUT.GE.5) NPRI=1
2227     C
2228     C        USE REGMOD TO OBTAIN F AND VAR
2229     C
2230           NFDIF=0
2231           IFAIL=0
2232           CALL REGMOD(PARAM,NPAR,NLMAX,NL,NFDIF,F,VAR,NPQA,NP,NQ,NTEQ,
2233          1            NOUT,IFAIL)
2234           IF (IFAIL.NE.0) GO TO 008
2235           IF ((NP.LT.0).OR.(NP.GT.NPMAX)) GO TO 010
2236           IF ((NQ.LT.0).OR.(NQ.GT.NQMAX)) GO TO 012
2237           IF (NPQA.NE.NP+NQ+1) GO TO 014
2238           NPA=NP+1
2239           NQA=NQ+1
2240           DO 001 I=1,NL
```

```
2241.01          DO 001 J=1,NPQA
2242.02    001   VAR(I,J)=VAR(I,J)*SCALE
2243.02    C
2244.02    C        DECOMPOSE VAR INTO PHI AND THETA
2245.02    C
2246             CALL REGAMD(VAR,NLMAX,NL,NPQA,NPA,NQA,NP,NQ,NTEQ,PHI,THETA,XLDET,
2247            1           NOUT,NPRI,IFAIL)
2248             IF (IFAIL.NE.0) GO TO 016
2249      C
2250      C        SIMULATE RESIDUALS Z
2251      C
2252             IF (NSEED.NE.0) CALL G05CBF(NSEED)
2253             DO 002 I=1,NL
2254.01     002   Z(I)=G05DDF(0.0D0,1.0D0)
2255.01    C
2256.01    C        TRANSFORM TO DEPARTURES E
2257.01    C
2258             CALL REGTRA(Z,NL,THETA,NLMAX,NQA,0,0,W1,NOUT,NPRI,IFAIL)
2259             CALL REGTRA(W1,NL,PHI,NLMAX,NPA,1,0,E,NOUT,NPRI,IFAIL)
2260      C
2261      C        ADD TO F TO OBTAIN SIMULATED OBSERVATIONS Y
2262      C
2263             DO 003 I=1,NL
2264.01     003   Y(I)=F(I)+E(I)
2265.01    C
2266.01    C        OUTPUT RESULT
2267.01    C
2268             WRITE (NOUT7,904) NL,NSEED
2269      904    FORMAT (2I5)
2270             WRITE (NOUT7,905) Y
2271      905    FORMAT (1P5G24.16)
2272             RETURN
2273      C
2274      C        ERROR MESSAGES
2275      C
2276      006    WRITE (NOUT6,907) NL
2277             IF (NOUT.GE.1) WRITE (NOUT10,907) NL
2278      907    FORMAT (' *** REGGEN: NL =',I5,' OUT OF RANGE ***')
2279             IFAIL=1
2280             RETURN
2281      008    WRITE (NOUT6,909)
2282             IF (NOUT.GE.1) WRITE (NOUT10,909)
2283      909    FORMAT (' *** REGGEN: REGMOD CANNOT EVALUATE F OR VAR ***')
2284             IFAIL=1
2285             RETURN
2286      010    WRITE (NOUT6,911) NP
2287             IF (NOUT.GE.1) WRITE (NOUT10,911) NP
2288      911    FORMAT (' *** REGGEN: NP =',I5,' OUT OF RANGE ***')
2289             IFAIL=1
2290             RETURN
2291      012    WRITE (NOUT6,913) NQ
2292             IF (NOUT.GE.1) WRITE (NOUT10,913) NQ
2293      913    FORMAT (' *** REGGEN: NQ =',I5,' OUT OF RANGE ***')
2294             IFAIL=1
2295             RETURN
2296      014    WRITE (NOUT6,915) NPQA
2297             IF (NOUT.GE.1) WRITE (NOUT10,915) NPQA
2298      915    FORMAT (' *** REGGEN: NPQA =',I5,' NOT EQUAL TO NP+NQ+1 ***')
2299             IFAIL=1
2300             RETURN
2301      016    IFAIL=1
2302             RETURN
2303             END
```

Appendix B

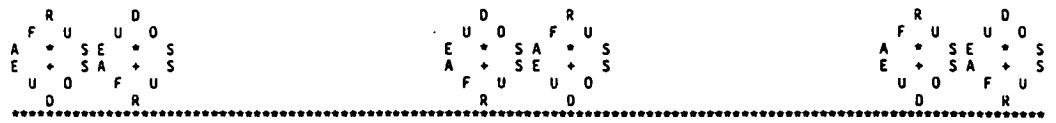
Example of output from REGAME to channel NOUT10 (see [2.4.7]).

(1) Colquhoun's data, single exponential regression with independent errors, fitted by least-squares estimation (see [4.3.2]);

(i) history of iterations,
(ii) parameter estimates and standard errors,
(iii) fitted model,
(iv) line printer plots.

(2) Colquhoun's data, single exponential regression with ARMA(1,1) errors, fitted by maximum likelihood estimation (see [4.3.4]).

RESULTS FROM PROGRAM REGAME - C.A.GLASBEY



ITERATIONS

NCALMX	NLIK	NSCALE	NM	NDER	ACCTOL	NVAR	NPEST	PARAMS.....			
100	1	1	3	1	1.00-04	1	3	1.00000	1.00000	0.200000	0.0000000+00

NITER	NCALL	XLIK	XLIKCH	SLOPE	PARACH	COND		PARAM.....			
0	0	103.692	1.00-10	1.40+01	1.00-10	1.60+01		-90.0000	80.0000	7.00000	-1.57080
1	6	59.6059	1.00+00	1.80+00	5.00-01	7.60+00	C	-89.8687	80.1165	7.19689	-1.57080
2	13	34.0910	6.10-01	3.90-01	3.40-01	1.80+01		-89.5777	79.6061	7.34952	-1.57080
3	18	15.3577	4.60-01	4.50-01	3.00-01	7.10+01	C	-89.3397	78.6399	7.36940	-1.57080
4	23	-1.70056	4.00-01	3.00-01	2.50-01	3.00+01	C	-89.0894	77.6844	7.33817	-1.57080
5	27	-4.19857	1.50-01	1.70-01	2.40-01	2.80+01	C	-88.7165	77.1758	7.18296	-1.57080
6	31	-6.33802	1.40-01	7.70-03	1.10-01	2.70+01		-88.8766	77.4458	7.24023	-1.57080
7	35	-6.34258	6.40-03	1.00-02	1.10-02	3.00+01		-88.8567	77.4558	7.23293	-1.57080
8	39	-6.34460	4.30-03	1.10-03	3.50-03	3.50+01		-88.8626	77.4476	7.23488	-1.57080
9	43	-6.34463	4.80-04	9.80-05	2.80-04	3.30+01		-88.8629	77.4483	7.23504	-1.57080
10	50	-6.34463	6.80-05	2.70-06	6.30-05	3.10+01		-88.8628	77.4483	7.23500	-1.57080
10	57	-6.34463	1.00-10	2.70-06	1.00-10	3.10+01		-88.8628	77.4483	7.23500	-1.57080

RESULTS

IFAIL 0

NEG. LOG-LIKELIHOOD -6.34463

FIRST DERIVATIVES -6.0737170-05 2.6092640-05 -7.1728820-05 0.0000000+00 0.0000000+00

	1	2	3	4	5
PARAMETER ESTIMATES	-88.8628	77.4483	7.23500	-1.57080	0.332096
STANDARD ERRORS	0.121258	0.299208	5.8111930-02	0.0000000+00	4.2176150-02

(CO)VARIANCES

1	1.4703400-02				
2	6.4304520-03	8.9525260-02			
3	-5.7519350-03	-1.0415150-02	3.3769970-03		
4	0.0000000+00	0.0000000+00	0.0000000+00	0.0000000+00	
5	2.0546160-08	9.1697910-08	-1.7864890-08	0.0000000+00	1.7788270-03

CORRELATIONS

2	0.177239			
3	-0.816281	-0.599001		
4	0.000000	0.000000	0.000000	
5	0.000004	0.000007	-0.000007	0.000000

(CORRELATIONS EXCEEDING 0.9 IN MAGNITUDE INDICATED BY **)

	DATA(Y)	FIT(F)	PREDICTIONS	RESIDUALS(Z)	COVARIANCES(VAR) (NP = 0 , NQ = 0)
1	-22.2170	-23.7032	-23.7032	2.578899**	0.332096
2	-24.4140	-25.9162	-25.9162	2.606813**	0.332096
3	-27.3440	-28.0542	-28.0542	1.232338	0.332096
4	-29.2970	-30.1195	-30.1195	1.427225	0.332096
5	-31.7380	-32.1146	-32.1146	0.653574	0.332096
6	-33.9350	-34.0420	-34.0420	0.185741	0.332096
7	-36.1330	-35.9040	-35.9040	-0.397423	0.332096
8	-38.0860	-37.7027	-37.7027	-0.665180	0.332096
9	-39.3070	-39.4403	-39.4403	0.231274	0.332096
10	-41.5040	-41.1189	-41.1189	-0.668311	0.332096
11	-42.7240	-42.7404	-42.7404	0.028538	0.332096
12	-44.4330	-44.3069	-44.3069	-0.218734	0.332096
13	-46.6310	-45.8202	-45.8202	-1.406881	0.332096
14	-47.6070	-47.2821	-47.2821	-0.563710	0.332096
15	-49.3160	-48.6944	-48.6944	-1.078657	0.332096
16	-50.7810	-50.0587	-50.0587	-1.253430	0.332096
17	-52.4900	-51.3766	-51.3766	-1.932017	0.332096
18	-53.4670	-52.6498	-52.6498	-1.418058	0.332096
19	-54.6870	-53.8797	-53.8797	-1.400808	0.332096
20	-55.4200	-55.0679	-55.0679	-0.610969	0.332096
21	-57.1290	-56.2157	-56.2157	-1.584785	0.332096
22	-58.1050	-57.3246	-57.3246	-1.354294	0.332096
23	-59.8140	-58.3957	-58.3957	-2.461110	0.332096
24	-60.3030	-59.4305	-59.4305	-1.514021	0.332096
25	-61.0350	-60.4301	-60.4301	-1.049592	0.332096
26	-62.7440	-61.3958	-61.3958	-2.339442	0.332096
27	-62.9880	-62.3287	-62.3287	-1.144029	0.332096
28	-63.4760	-63.2299	-63.2299	-0.427004	0.332096
29	-63.9650	-64.1005	-64.1005	0.235170	0.332096
30	-65.6740	-64.9416	-64.9416	-1.271002	0.332096
31	-66.4060	-65.7540	-65.7540	-1.131377	0.332096
32	-66.8940	-66.5389	-66.5389	-0.616229	0.332096
33	-67.6270	-67.2971	-67.2971	-0.572482	0.332096
34	-68.3590	-68.0296	-68.0296	-0.571686	0.332096
35	-68.8470	-68.7371	-68.7371	-0.190651	0.332096
36	-69.0920	-69.4207	-69.4207	0.570352	0.332096
37	-69.5800	-70.0810	-70.0810	0.869397	0.332096
38	-70.8010	-70.7189	-70.7189	-0.142433	0.332096
39	-70.8010	-71.3352	-71.3352	0.926913	0.332096
40	-71.0450	-71.9305	-71.9305	1.536532	0.332096
41	-72.0210	-72.5056	-72.5056	0.840843	0.332096
42	-72.2650	-73.0611	-73.0611	1.381482	0.332096
43	-73.7300	-73.5978	-73.5978	-0.229394	0.332096
44	-73.2420	-74.1163	-74.1163	1.517093	0.332096
45	-73.4860	-74.6171	-74.6171	1.962801	0.332096
46	-74.9510	-75.1010	-75.1010	0.260220	0.332096
47	-74.7070	-75.5684	-75.5684	1.494708	0.332096
48	-74.9510	-76.0199	-76.0199	1.854834	0.332096
49	-75.6830	-76.4561	-76.4561	1.341533	0.332096
50	-76.4160	-76.8775	-76.8775	0.800789	0.332096
51	-76.4160	-77.2845	-77.2845	1.507168	0.332096
52	-77.1480	-77.6778	-77.6778	0.919333	0.332096
53	-77.3920	-78.0577	-78.0577	1.155135	0.332096
54	-77.6370	-78.4247	-78.4247	1.366813	0.332096
55	-78.1250	-78.7792	-78.7792	1.135190	0.332096
56	-79.1010	-79.1217	-79.1217	0.035858	0.332096
57	-79.1010	-79.4525	-79.4525	0.609971	0.332096
58	-79.3450	-79.7721	-79.7721	0.741177	0.332096
59	-79.3450	-80.0809	-80.0809	1.276953	0.332096
60	-80.0780	-80.3791	-80.3791	0.522575	0.332096
61	-80.5660	-80.6673	-80.6673	0.175761	0.332096
62	-80.5660	-80.9456	-80.9456	0.658779	0.332096
63	-80.8100	-81.2145	-81.2145	0.701984	0.332096
64	-81.2990	-81.4743	-81.4743	0.304199	0.332096
65	-81.5430	-81.7252	-81.7252	0.316246	0.332096
66	-81.7870	-81.9677	-81.9677	0.313504	0.332096
67	-82.5190	-82.2019	-82.2019	-0.550340	0.332096
68	-82.2750	-82.4281	-82.4281	0.265642	0.332096
69	-82.2750	-82.6466	-82.6466	0.644884	0.332096
70	-82.5190	-82.8578	-82.8578	0.587838	0.332096
71	-82.5190	-83.0617	-83.0617	0.941756	0.332096
72	-82.5190	-83.2587	-83.2587	1.283654	0.332096
73	-82.7630	-83.4491	-83.4491	1.190532	0.332096
74	-83.0080	-83.6329	-83.6329	1.084457	0.332096
75	-83.4960	-83.8106	-83.8106	0.545873	0.332096
76	-83.7400	-83.9822	-83.9822	0.420227	0.332096
77	-84.9610	-84.1479	-84.1479	-1.410896	0.332096
78	-84.4720	-84.3081	-84.3081	-0.284467	0.332096
79	-84.2280	-84.4628	-84.4628	0.407381	0.332096
80	-84.4720	-84.6122	-84.6122	0.243298	0.332096
81	-84.4720	-84.7566	-84.7566	0.493814	0.332096
82	-84.4720	-84.8960	-84.8960	0.735821	0.332096
83	-84.7170	-85.0308	-85.0308	0.544467	0.332096
84	-84.9610	-85.1609	-85.1609	0.346907	0.332096
85	-84.9610	-85.2866	-85.2866	0.565084	0.332096
86	-85.2050	-85.4081	-85.4081	0.352443	0.332096
87	-85.2050	-85.5254	-85.5254	0.556051	0.332096
88	-85.9370	-85.6388	-85.6388	-0.517478	0.332096
89	-85.4490	-85.7483	-85.7483	0.519349	0.332096
90	-85.6930	-85.8541	-85.8541	0.279500	0.332096
91	-86.4260	-85.9563	-85.9563	-0.815132	0.332096
92	-85.9370	-86.0550	-86.0550	0.204719	0.332096
93	-86.4260	-86.1503	-86.1503	-0.478347	0.332096
94	-85.6930	-86.2425	-86.2425	0.953473	0.332096
95	-86.6700	-86.3315	-86.3315	-0.587458	0.332096

96	-86.6700	-86.4174	-86.4174	-0.438270	0.332096
97	-86.9140	-86.5005	-86.5005	-0.717556	0.332096
98	-86.4260	-86.5807	-86.5807	0.268485	0.332096
99	-86.6700	-86.6582	-86.6582	-0.020425	0.332096
100	-86.6700	-86.7331	-86.7331	0.109505	0.332096
101	-86.4260	-86.8054	-86.8054	0.658428	0.332096
102	-86.9140	-86.8753	-86.8753	-0.067133	0.332096
103	-86.9140	-86.9428	-86.9428	0.050003	0.332096
104	-87.1580	-87.0080	-87.0080	-0.260248	0.332096
105	-86.9140	-87.0710	-87.0710	0.272473	0.332096
106	-87.4020	-87.1319	-87.1319	-0.468741	0.332096
107	-87.4020	-87.1907	-87.1907	-0.366727	0.332096
108	-87.8900	-87.2475	-87.2475	-1.114992	0.332096
109	-87.6460	-87.3023	-87.3023	-0.596382	0.332096
110	-87.6460	-87.3553	-87.3553	-0.504414	0.332096
111	-87.6460	-87.4065	-87.4065	-0.415568	0.332096
112	-87.8900	-87.4560	-87.4560	-0.753148	0.332096
113	-87.6460	-87.5038	-87.5038	-0.246828	0.332096
114	-88.1350	-87.5499	-87.5499	-1.015282	0.332096
115	-88.3790	-87.5945	-87.5945	-1.361313	0.332096
116	-88.6230	-87.6376	-87.6376	-1.709972	0.332096
117	-88.1350	-87.6792	-87.6792	-0.790948	0.332096
118	-88.8670	-87.7194	-87.7194	-1.991413	0.332096
119	-88.1350	-87.7582	-87.7582	-0.653804	0.332096
120	-88.1350	-87.7957	-87.7957	-0.588705	0.332096
121	-88.8670	-87.8320	-87.8320	-1.796039	0.332096
122	-88.1350	-87.8670	-87.8670	-0.465065	0.332096
123	-88.6230	-87.9008	-87.9008	-1.253191	0.332096
124	-88.1350	-87.9335	-87.9335	-0.349682	0.332096

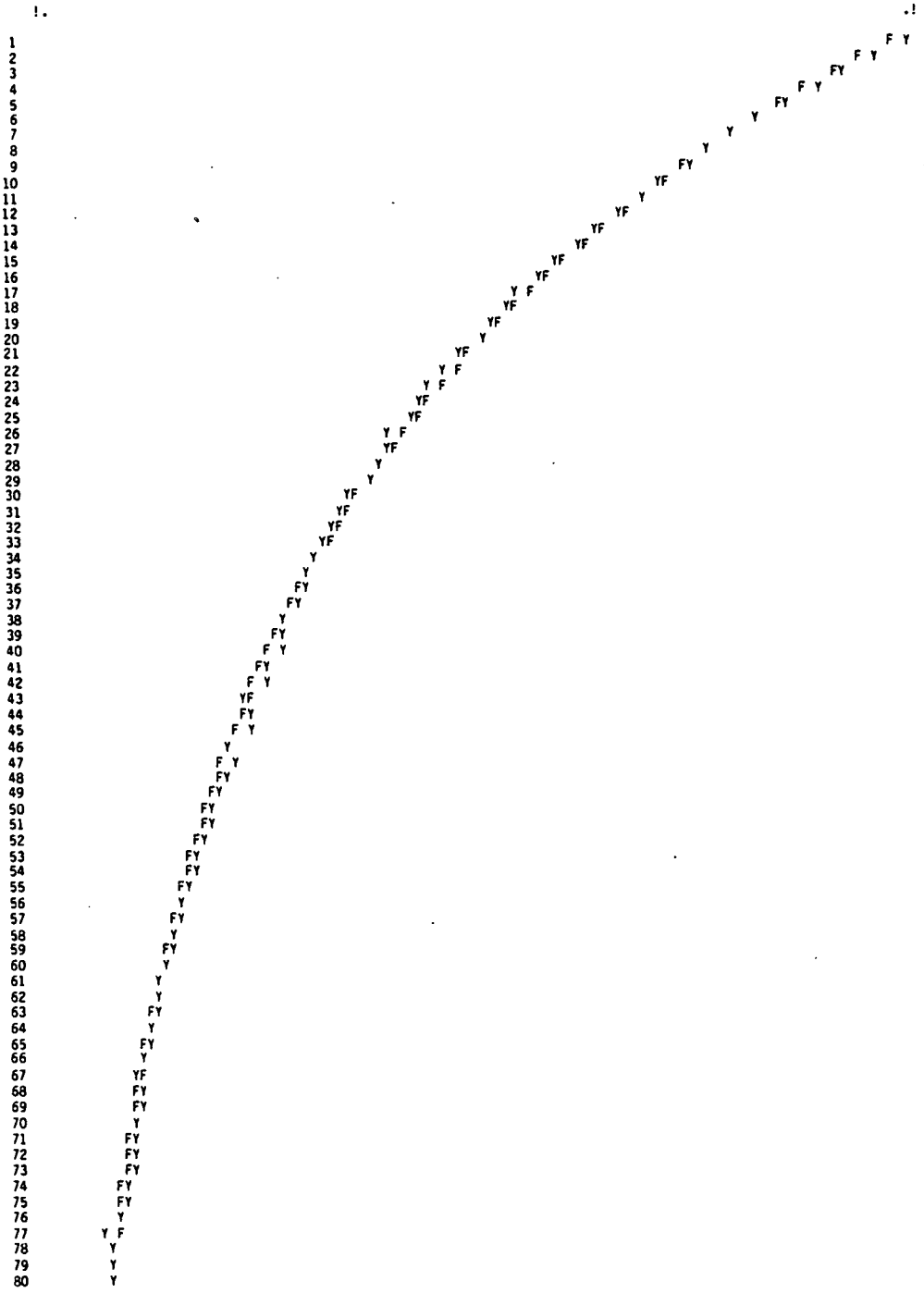
(RESIDUALS Z EXCEEDING 2.5 IN MAGNITUDE INDICATED BY **)

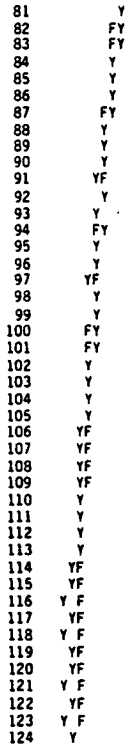
THERE FOLLOW:

(A)	A LONGITUDINAL PLOT OF DATA(Y),FIT(F),PREDICTION	(REVEALS LACK OF FIT)
(B)	A LONGITUDINAL PLOT OF DATA-FIT,PREDICTION-FIT,RESIDUALS(Z)	(REVEALS CORRELATION OF ERRORS)
(C)	A CROSS-SECTIONAL PLOT OF RESIDUALS(Z) AGAINST FIT(F)	(REVEALS HETEROGENEITY OF VARIANCE AND OUTLIERS)
(D)	A LONGITUDINAL PLOT OF COVARIANCES OF ERRORS	(DISPLAYS FITTED COVARIANCE STRUCTURE)

OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

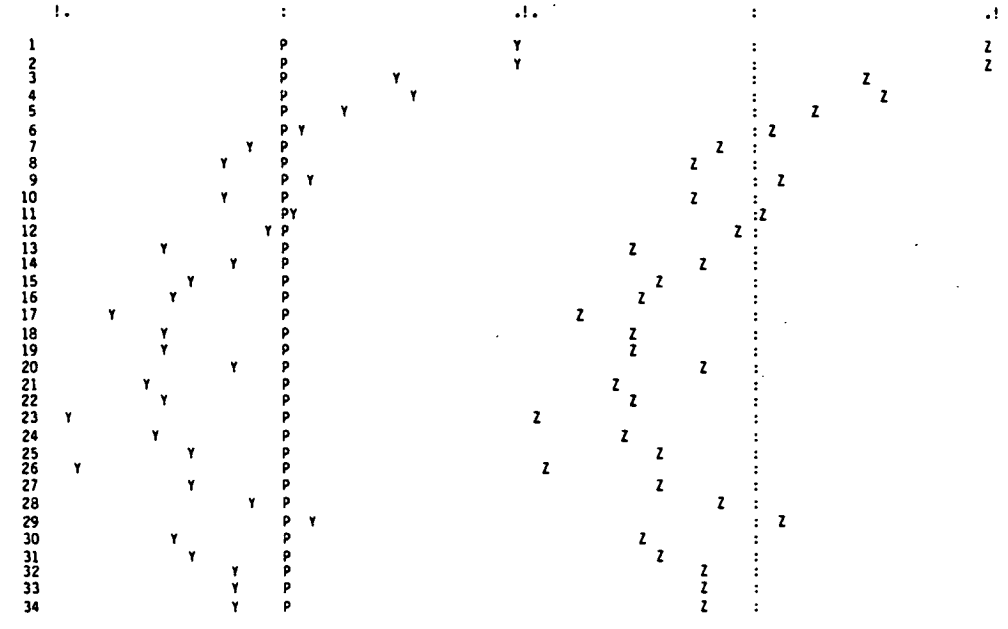
VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	-88.867000	-22.217000	1.6054014	145.16720	Y
2	1	-88.867000	-22.217000	1.6054014	145.16720	F
3	1	-88.867000	-22.217000	1.6054014	145.16720	P

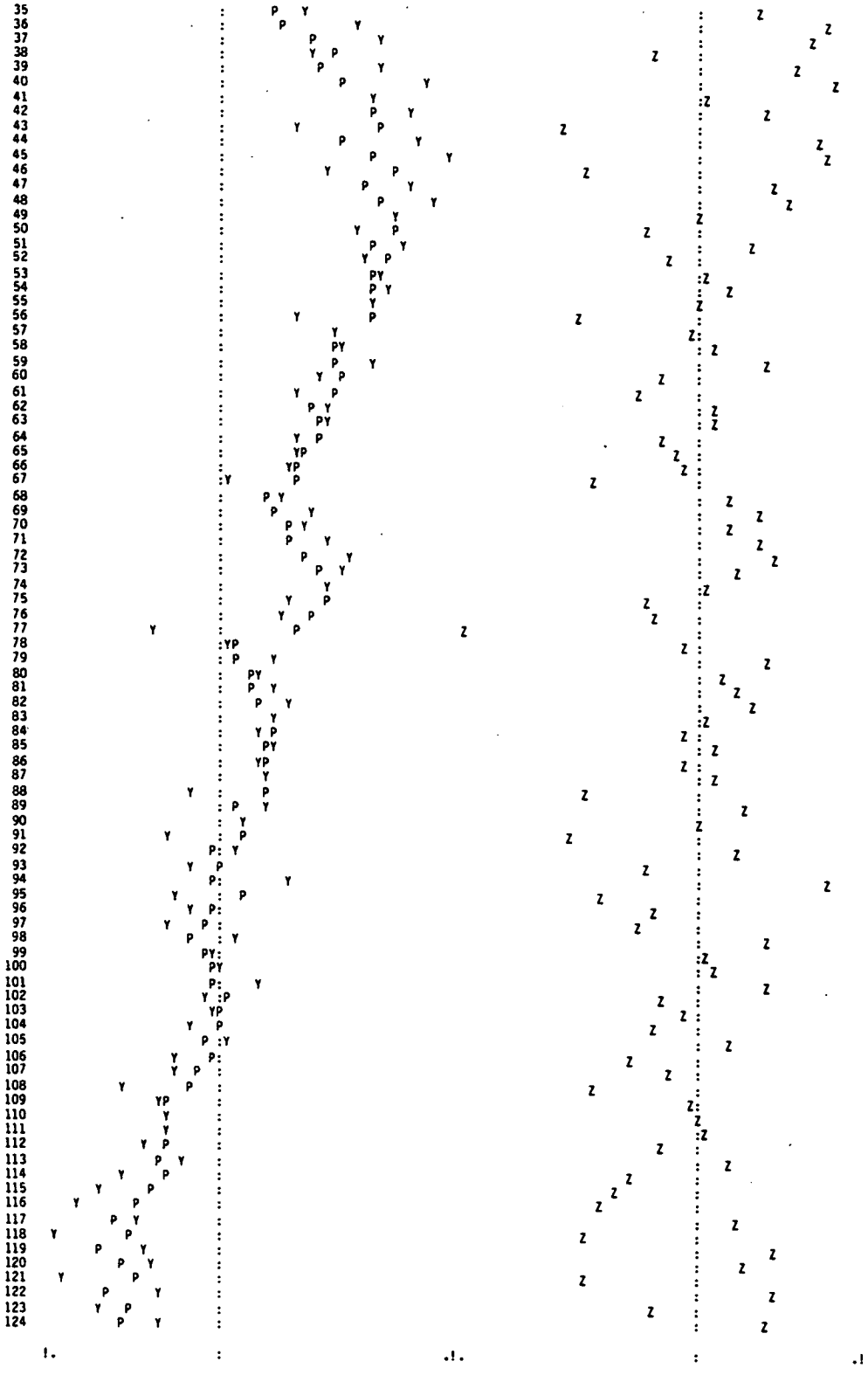




OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	-1.4182817	1.5022474	17.804993	27.752495	Y
3	1	-1.4182817	1.5022474	17.804993	27.752495	P
4	2	-2.4611098	2.6068135	10.260613	81.752495	Z





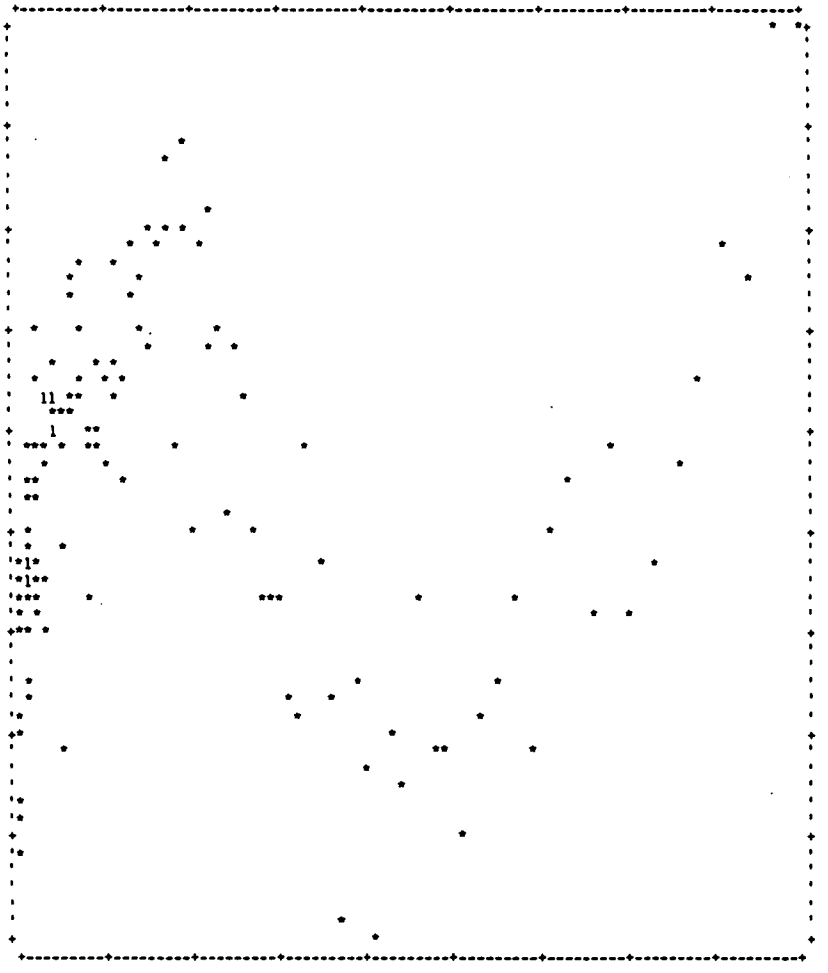
OUTPUT FROM SUBROUTINE PLTACR - C.A.GLASBEY

VARIABLE	MINIMUM	MAXIMUM	*VTV(1)	+VTV(2)	*VTA(1)	+VTA(2)
2	-87.933487	-23.703159	0.84072429	76.427819	1.4012072	125.71303
4	-2.4611096	2.6068134	10.655252	28.723742	17.758753	46.206237

COUNTS UP TO	0	1	2	4	8	16	32	64	128	256	512	OVER
CHARACTERS		*	1	2	3	4	5	6	7	8	9	0

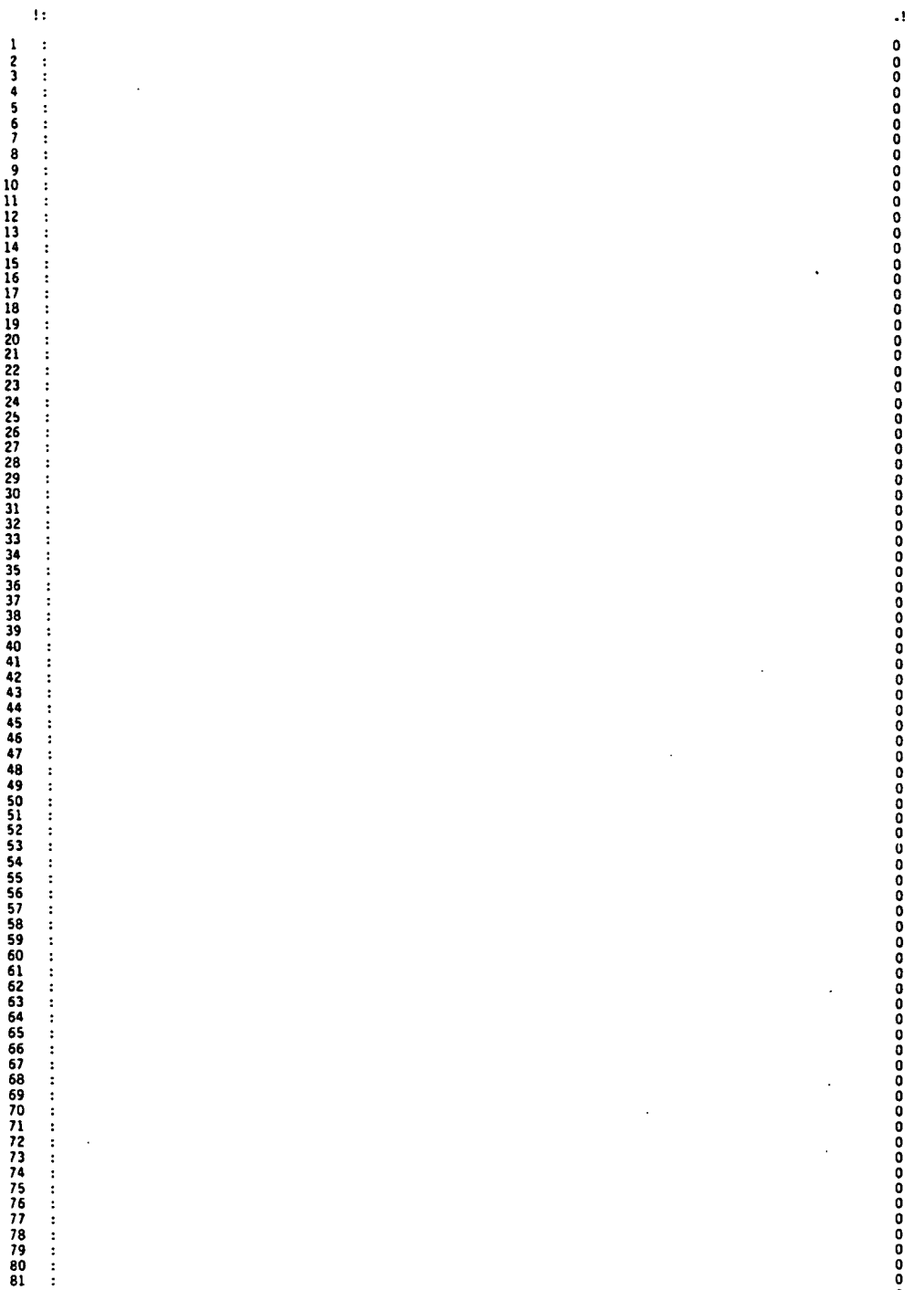
PLOTS 4 VERSUS 2

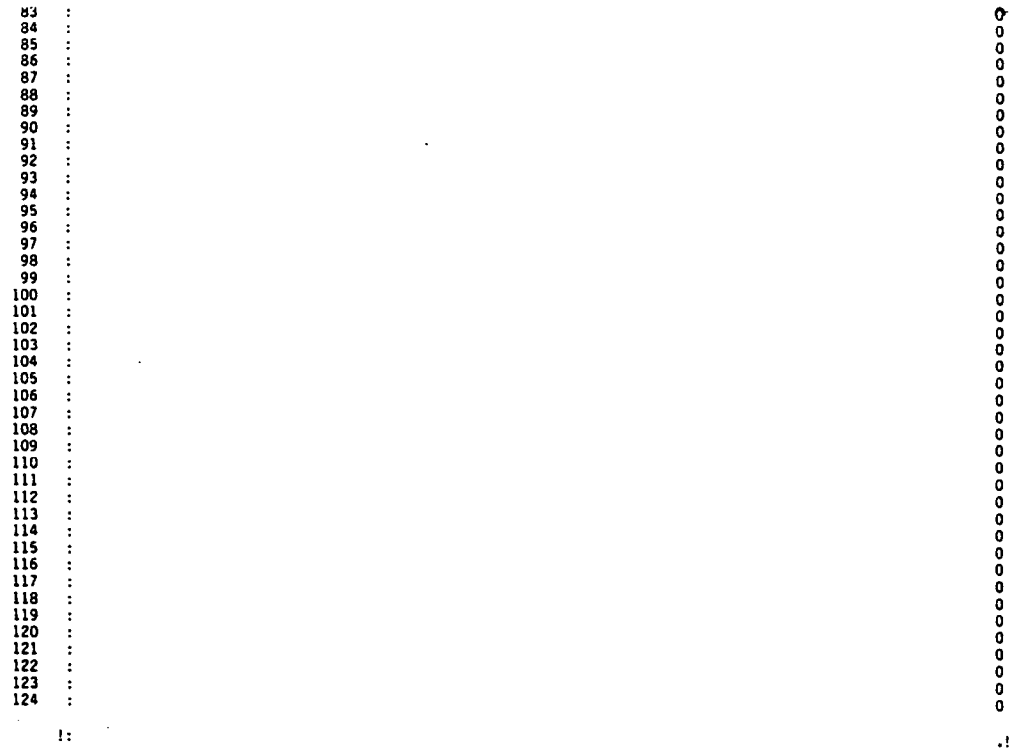
VERTICAL VARIABLE 4 (-2.461 , 2.607) HORIZONTAL VARIABLE 2 (-87.93 , -23.70)



OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	0.000000000+00	0.33209551	322.19647	2.5000000	0





AUTOCORRELATIONS OF RESIDUALS

0.6953660 0.6379363 0.5984393 0.5295662 0.4735971 0.4007144 0.4184132 0.3884956 0.3422894 0.2864343

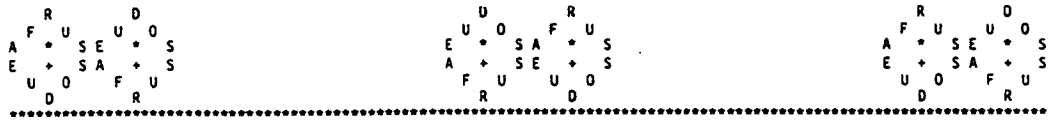
GENERALIZED PARTIAL AUTOCORRELATIONS

NQV	NPV=	0		1		2		3	4	5	6	7	8	9
0		69.54	X	29.90	X	16.92	+	2.70	-0.13	-5.86	12.68	4.45	-2.25	-8.70
1		45.48	X	2.04		-4.57		-0.54	-5.84	0.14	4.94	-3.10	-6.81	
2		35.89	X	-4.82		-2.43		-5.10	0.62	1.21	1.94	-5.62		
3		28.32	X	0.79		-4.10		6.78	4.50	2.24	-1.17			
4		23.51	X	-3.60		-0.22		3.83	-1.57	3.51				
5		18.88	X	13.06		3.96		0.30	3.52					
6		19.04	X	-6.47		-7.67		-8.95						
7		17.07	+	-2.75		-0.65								
8		14.62	+	-2.37										
9		11.98												

STANDARD ERROR = 8.98

* VALUES EXCEEDING 1.5 SE DENOTED BY +
VALUES EXCEEDING 2.0 SE DENOTED BY X

RESULTS FROM PROGRAM REGAME - C.A.GLASBEY



ITERATIONS

NCALMX	NLIK	NSCALE	NM	NDER	ACCTOL	NVAR	NPEST	PARAMS.....			
200	1	1	3	1	1.0D-04	1	5	1.00000	1.00000	0.500000	0.500000
								0.500000	0.0000000+00		

NITER	NCALL	XLIK	XLIKCH	SLOPE	PARACH	COND		PARAM.....			
0	0	-60.1152	1.0D-10	4.3D-01	1.0D-10	1.2D+01		-88.8600	77.4500	7.23500	0.600000
								1.00000	-1.57080		
1	7	-62.0261	4.9D-01	1.0D+00	5.0D-01	1.1D+01	C	-88.7900	78.4331	7.19978	0.611338
								1.07611	-1.57080		
2	14	-63.7521	5.2D-01	7.5D-01	3.4D-01	1.1D+01	C	-88.8552	79.3587	7.05543	0.631651
								1.19220	-1.57080		
3	20	-63.8697	1.4D-01	8.5D-01	1.9D-01	1.1D+01		-88.8661	79.9548	6.92813	0.729627
								1.14741	-1.57080		
4	28	-64.3515	2.9D-01	2.6D-01	5.1D-02	1.2D+01		-88.8137	79.8802	6.91838	0.698945
								1.21835	-1.57080		
5	35	-64.5542	1.9D-01	2.0D-01	1.4D-01	2.0D+01		-88.6942	79.4715	6.88506	0.762024
								1.24405	-1.57080		
6	42	-64.6742	1.5D-01	2.5D-01	8.7D-02	1.7D+01		-88.5440	79.5282	6.78034	0.814035
								1.28393	-1.57080		
7	48	-64.6994	6.8D-02	1.7D-01	2.7D-02	1.7D+01		-88.4958	79.4580	6.77065	0.820533
								1.27147	-1.57080		
8	55	-64.7337	8.0D-02	9.7D-02	6.9D-02	1.6D+01		-88.3650	79.2862	6.76365	0.797585
								1.27131	-1.57080		
9	62	-64.7466	4.9D-02	1.4D-02	2.9D-02	1.5D+01		-88.3369	79.3653	6.76929	0.779641
								1.25701	-1.57080		
10	69	-64.7468	6.4D-03	3.1D-03	5.1D-03	1.5D+01		-88.3493	79.3753	6.77164	0.778737
								1.25721	-1.57080		
11	76	-64.7469	2.0D-03	2.4D-03	1.3D-03	1.1D+01		-88.3491	79.3741	6.77203	0.776897
								1.25656	-1.57080		
12	83	-64.7469	6.4D-04	2.5D-04	1.5D-04	1.1D+01		-88.3494	79.3742	6.77187	0.776859
								1.25648	-1.57080		
13	90	-64.7469	8.1D-05	1.5D-05	4.0D-05	1.2D+01		-88.3495	79.3742	6.77192	0.776839
								1.25646	-1.57080		
13	95	-64.7469	1.0D-10	1.5D-05	1.0D-10	1.2D+01		-88.3495	79.3742	6.77192	0.776839
								1.25646	-1.57080		

RESULTS

IFAIL 0

NEG. LOG-LIKELIHOOD -64.7469

FIRST DERIVATIVES 1.204753D-06 7.385462D-06 1.125540D-05 -4.425280D-06 6.883670D-06 0.0000000+00 0.0000000+00

	1	2	3	4	5	6	7
PARAMETER ESTIMATES	-88.3495	79.3742	6.77192	0.776839	1.25646	-1.57080	0.517483
STANDARD ERRORS	0.651737	0.978568	0.226243	0.302573	0.149382	0.0000000+00	0.367237

(CO)VARIANCES

1	0.424762						
2	-0.305306	0.957595					
3	-9.324698D-02	3.228359D-02	5.118609D-02				
4	4.291613D-02	-1.611335D-02	-3.640921D-02	9.155037D-02			
5	2.174327D-02	-1.430660D-02	-1.924066D-02	3.890992D-02	2.231512D-02		
6	0.0000000+00	0.0000000+00	0.0000000+00	0.0000000+00	0.0000000+00	0.0000000+00	
7	5.324483D-02	-2.464051D-02	-4.577291D-02	0.107927	5.024349D-02	0.0000000+00	0.134863

CORRELATIONS

2	-0.478709					
3	-0.632392	0.145819				
4	0.217630	-0.054421	-0.531870			
5	0.223333	-0.097869	-0.569304	0.860856		
6	0.000000	0.000000	0.000000	0.000000	0.000000	
7	0.222463	-0.068567	-0.550917	0.971304**	0.915869**	0.000000

(CORRELATIONS EXCEEDING 0.9 IN MAGNITUDE INDICATED BY **)

	DATA(Y)	FIT(F)	PREDICTIONS	RESIDUALS(Z)	COVARIANCES(VAR) (NP = 1 ,		NQ = 1)
1	-22.2170	-22.3539	-22.3539	0.190295	0.0000000+00	0.0000000+00	0.517483
2	-24.4140	-24.7458	-24.6323	0.543507	0.0000000+00	0.429344	0.517483
3	-27.3440	-27.0511	-26.8295	-1.390202	0.418825	0.429344	0.517483
4	-29.2970	-29.2728	-29.2751	-0.060700	0.418825	0.429344	0.517483
5	-31.7380	-31.4140	-31.4249	-0.872721	0.418825	0.429344	0.517483
6	-33.9350	-33.4775	-33.6101	-0.908146	0.418825	0.429344	0.517483
7	-36.1330	-35.4663	-35.7210	-1.152388	0.418825	0.429344	0.517483
8	-38.0860	-37.3830	-37.7901	-0.827988	0.418825	0.429344	0.517483
9	-39.3070	-39.2302	-39.7412	1.215051	0.418825	0.429344	0.517483
10	-41.5040	-41.0105	-41.3420	-0.453397	0.418825	0.429344	0.517483
11	-42.7240	-42.7263	-43.1119	1.085650	0.418825	0.429344	0.517483
12	-44.4330	-44.3798	-44.6069	0.486716	0.418825	0.429344	0.517483
13	-46.6310	-45.9735	-46.1281	-1.407407	0.418825	0.429344	0.517483
14	-47.6070	-47.5094	-47.8535	0.689972	0.418825	0.429344	0.517483
15	-49.3160	-48.9896	-49.2305	-0.239189	0.418825	0.429344	0.517483
16	-50.7810	-50.4161	-50.6840	-0.271341	0.418825	0.429344	0.517483
17	-52.4900	-51.7910	-52.0896	-1.120577	0.418825	0.429344	0.517483
18	-53.4670	-53.1160	-53.5612	0.263755	0.418825	0.429344	0.517483
19	-54.6870	-54.3930	-54.7911	0.291368	0.418825	0.429344	0.517483
20	-55.4200	-55.6237	-55.9721	1.545043	0.418825	0.429344	0.517483
21	-57.1290	-56.8098	-56.9374	-0.536210	0.418825	0.429344	0.517483
22	-58.1050	-57.9530	-58.1511	0.128907	0.418825	0.429344	0.517483
23	-59.8140	-59.0547	-59.2302	-1.633922	0.418825	0.429344	0.517483
24	-60.3030	-60.1164	-60.5121	0.585164	0.418825	0.429344	0.517483
25	-61.0350	-61.1397	-61.4453	1.148297	0.418825	0.429344	0.517483
26	-62.7440	-62.1259	-62.2663	-1.337033	0.418825	0.429344	0.517483
27	-62.9880	-63.0763	-63.3969	1.144484	0.418825	0.429344	0.517483
28	-63.4760	-63.9923	-64.1479	1.880392	0.418825	0.429344	0.517483
29	-63.9650	-64.8751	-64.7686	2.249003	0.418825	0.429344	0.517483
30	-65.6740	-65.7259	-65.3131	-1.010115	0.418825	0.429344	0.517483
31	-66.4060	-66.5459	-66.2819	-0.347297	0.418825	0.429344	0.517483
32	-66.8940	-67.3362	-67.1263	0.650226	0.418825	0.429344	0.517483
33	-67.6270	-68.0978	-67.8038	0.494716	0.418825	0.429344	0.517483
34	-68.3590	-68.8318	-68.4770	0.330302	0.418825	0.429344	0.517483
35	-68.8470	-69.5392	-69.1477	0.841692	0.418825	0.429344	0.517483
36	-69.0920	-70.2209	-69.7235	1.767342	0.418825	0.429344	0.517483
37	-69.5800	-70.8780	-70.1500	1.595173	0.418825	0.429344	0.517483
38	-70.8010	-71.5112	-70.5819	-0.613118	0.418825	0.429344	0.517483
39	-70.8010	-72.1215	-71.2992	1.394323	0.418825	0.429344	0.517483
40	-71.0450	-72.7097	-71.7160	1.877927	0.418825	0.429344	0.517483
41	-72.0210	-73.2765	-72.0492	0.079042	0.418825	0.429344	0.517483
42	-72.2650	-73.8228	-72.6148	0.978876	0.418825	0.429344	0.517483
43	-73.7300	-74.3493	-73.0364	-1.941187	0.418825	0.429344	0.517483
44	-73.2420	-74.8568	-73.8426	1.680989	0.418825	0.429344	0.517483
45	-73.4860	-75.3458	-74.1256	1.790089	0.418825	0.429344	0.517483
46	-74.9510	-75.8171	-74.3809	-1.595477	0.418825	0.429344	0.517483
47	-74.7070	-76.2713	-75.0895	1.070475	0.418825	0.429344	0.517483
48	-74.9510	-76.7091	-75.4092	1.282280	0.418825	0.429344	0.517483
49	-75.6830	-77.1310	-75.6868	0.010574	0.418825	0.429344	0.517483
50	-76.4160	-77.5376	-76.1273	-0.807960	0.418825	0.429344	0.517483
51	-76.4160	-77.9294	-76.6647	0.696065	0.418825	0.429344	0.517483
52	-77.1480	-78.3071	-76.9777	-0.476489	0.418825	0.429344	0.517483
53	-77.3920	-78.6711	-77.4397	0.133615	0.418825	0.429344	0.517483
54	-77.6370	-79.0219	-77.8023	0.462741	0.418825	0.429344	0.517483
55	-78.1250	-79.3599	-78.1067	-0.051140	0.418825	0.429344	0.517483
56	-79.1010	-79.6858	-78.4703	-1.765237	0.418825	0.429344	0.517483
57	-79.1010	-79.9998	-79.0565	-0.124473	0.418825	0.429344	0.517483
58	-79.3450	-80.3024	-79.3994	0.152134	0.418825	0.429344	0.517483
59	-79.3450	-80.5940	-79.6922	0.971843	0.418825	0.429344	0.517483
60	-80.0780	-80.8751	-79.8619	-0.604700	0.418825	0.429344	0.517483
61	-80.5660	-81.1460	-80.2407	-0.910371	0.418825	0.429344	0.517483
62	-80.5660	-81.4071	-80.6490	0.232368	0.418825	0.429344	0.517483
63	-80.8100	-81.6587	-80.8873	0.216345	0.418825	0.429344	0.517483
64	-81.2990	-81.9012	-81.1190	-0.503816	0.418825	0.429344	0.517483
65	-81.5430	-82.1349	-81.4411	-0.285299	0.418825	0.429344	0.517483
66	-81.7870	-82.3602	-81.7225	-0.180551	0.418825	0.429344	0.517483
67	-82.5190	-82.5773	-81.9800	-1.508551	0.418825	0.429344	0.517483
68	-82.2750	-82.7865	-82.4110	0.380731	0.418825	0.429344	0.517483
69	-82.2750	-82.9881	-82.5696	0.824407	0.418825	0.429344	0.517483
70	-82.5190	-83.1824	-82.6609	0.397138	0.418825	0.429344	0.517483
71	-82.5190	-83.3697	-82.8064	0.804360	0.418825	0.429344	0.517483
72	-82.5190	-83.5502	-82.8902	1.038904	0.418825	0.429344	0.517483
73	-82.7630	-83.7241	-82.9376	0.488716	0.418825	0.429344	0.517483
74	-83.0080	-83.8918	-83.0574	0.138265	0.418825	0.429344	0.517483
75	-83.4960	-84.0533	-83.2204	-0.771279	0.418825	0.429344	0.517483
76	-83.7400	-84.2090	-83.5025	-0.664772	0.418825	0.429344	0.517483
77	-84.9610	-84.3591	-83.7612	-3.358024**	0.418825	0.429344	0.517483
78	-84.4720	-84.5037	-84.3817	-0.252774	0.418825	0.429344	0.517483
79	-84.2280	-84.6431	-84.5588	0.925758	0.418825	0.429344	0.517483
80	-84.4720	-84.7774	-84.5680	0.268731	0.418825	0.429344	0.517483
81	-84.4720	-84.9069	-84.6657	0.542121	0.418825	0.429344	0.517483
82	-84.4720	-85.0317	-84.7219	0.699460	0.418825	0.429344	0.517483
83	-84.7170	-85.1519	-84.7537	0.102670	0.418825	0.429344	0.517483
84	-84.9610	-85.2678	-84.8652	-0.268032	0.418825	0.429344	0.517483
85	-84.9610	-85.3795	-85.0236	0.175204	0.418825	0.429344	0.517483
86	-85.2050	-85.4872	-85.1159	-0.249370	0.418825	0.429344	0.517483
87	-85.2050	-85.5909	-85.2630	0.162295	0.418825	0.429344	0.517483
88	-85.9370	-85.6909	-85.3487	-1.646458	0.418825	0.429344	0.517483
89	-85.4490	-85.7872	-85.6796	0.645399	0.418825	0.429344	0.517483
90	-85.6930	-85.8801	-85.6865	-0.018311	0.418825	0.429344	0.517483
91	-86.4260	-85.9696	-85.7832	-1.798976	0.418825	0.429344	0.517483
92	-85.9370	-86.0559	-86.1211	0.515356	0.418825	0.429344	0.517483
93	-86.4260	-86.1390	-86.1319	-0.823155	0.418825	0.429344	0.517483
94	-85.6930	-86.2191	-86.3252	1.769465	0.418825	0.429344	0.517483
95	-86.6700	-86.2963	-86.1568	-1.436294	0.418825	0.429344	0.517483
96	-86.6700	-86.3707	-86.4319	-0.666318	0.418825	0.429344	0.517483

97	-86.9140	-86.4425	-86.5937	-0.896538	0.418825	0.429344	0.517483
98	-86.4260	-86.5116	-86.7822	0.996966	0.418825	0.429344	0.517483
99	-86.6700	-86.5782	-86.7053	0.098693	0.418825	0.429344	0.517483
100	-86.6700	-86.6424	-86.7528	0.231714	0.418825	0.429344	0.517483
101	-86.4260	-86.7043	-86.7801	0.991120	0.418825	0.429344	0.517483
102	-86.9140	-86.7639	-86.7018	-0.593985	0.418825	0.429344	0.517483
103	-86.9140	-86.8214	-86.8423	-0.200547	0.418825	0.429344	0.517483
104	-87.1580	-86.8767	-86.9248	-0.652773	0.418825	0.429344	0.517483
105	-86.9140	-86.9301	-87.0666	0.427155	0.418825	0.429344	0.517483
106	-87.4020	-86.9816	-87.0561	-0.968217	0.418825	0.429344	0.517483
107	-87.4020	-87.0311	-87.2368	-0.462363	0.418825	0.429344	0.517483
108	-87.8900	-87.0789	-87.3430	-1.530773	0.418825	0.429344	0.517483
109	-87.6460	-87.1250	-87.5929	-0.148649	0.418825	0.429344	0.517483
110	-87.6460	-87.1694	-87.6462	0.000621	0.418825	0.429344	0.517483
111	-87.6460	-87.2121	-87.6772	0.087393	0.418825	0.429344	0.517483
112	-87.8900	-87.2533	-87.6950	-0.545614	0.418825	0.429344	0.517483
113	-87.6460	-87.2931	-87.7989	0.427923	0.418825	0.429344	0.517483
114	-88.1350	-87.3314	-87.7660	-1.032684	0.418825	0.429344	0.517483
115	-88.3790	-87.3683	-87.9341	-1.245115	0.418825	0.429344	0.517483
116	-88.6230	-87.4038	-88.1268	-1.388623	0.418825	0.429344	0.517483
117	-88.1350	-87.4381	-88.3341	0.557323	0.418825	0.429344	0.517483
118	-88.8670	-87.4711	-88.2687	-1.674586	0.418825	0.429344	0.517483
119	-88.1350	-87.5030	-88.5110	1.052244	0.418825	0.429344	0.517483
120	-88.1350	-87.5336	-88.3724	0.664488	0.418825	0.429344	0.517483
121	-88.8670	-87.5632	-88.2902	-1.614384	0.418825	0.429344	0.517483
122	-88.1350	-87.5917	-88.5226	1.084801	0.418825	0.429344	0.517483
123	-88.6230	-87.6192	-88.3783	-0.684965	0.418825	0.429344	0.517483
124	-88.1350	-87.6456	-88.4802	0.966161	0.418825	0.429344	0.517483

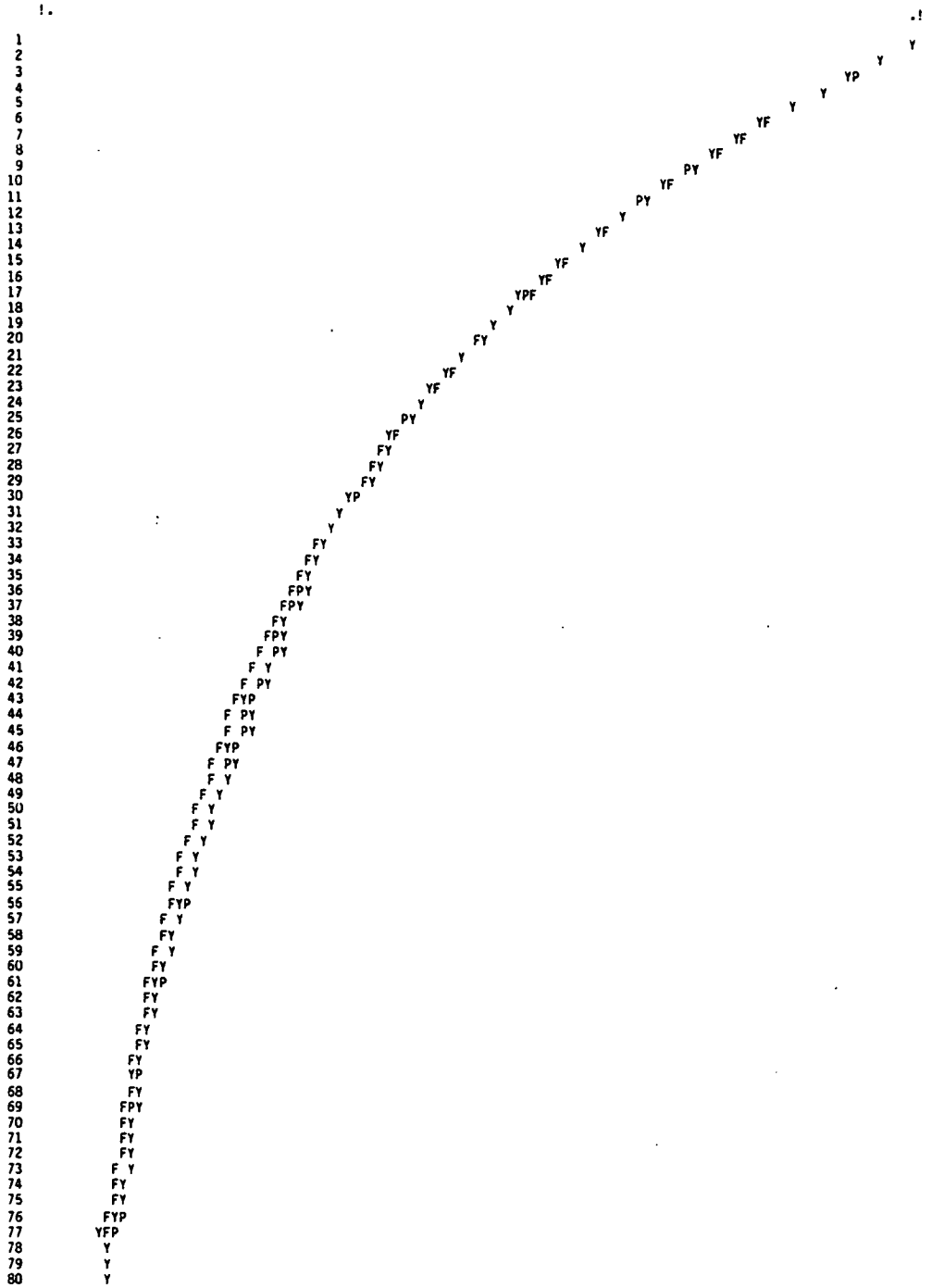
(RESIDUALS Z EXCEEDING 2.5 IN MAGNITUDE INDICATED BY **)

THERE FOLLOW:

(A)	A LONGITUDINAL PLOT OF DATA(Y),FIT(F),PREDICTION	(REVEALS LACK OF FIT)
(B)	A LONGITUDINAL PLOT OF DATA-FIT,PREDICTION-FIT,RESIDUALS(Z)	(REVEALS CORRELATION OF ERRORS)
(C)	A CROSS-SECTIONAL PLOT OF RESIDUALS(Z) AGAINST FIT(F)	(REVEALS HETEROGENEITY OF VARIANCE AND OUTLIERS)
(D)	A LONGITUDINAL PLOT OF COVARIANCES OF ERRORS	(DISPLAYS FITTED COVARIANCE STRUCTURE)

OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

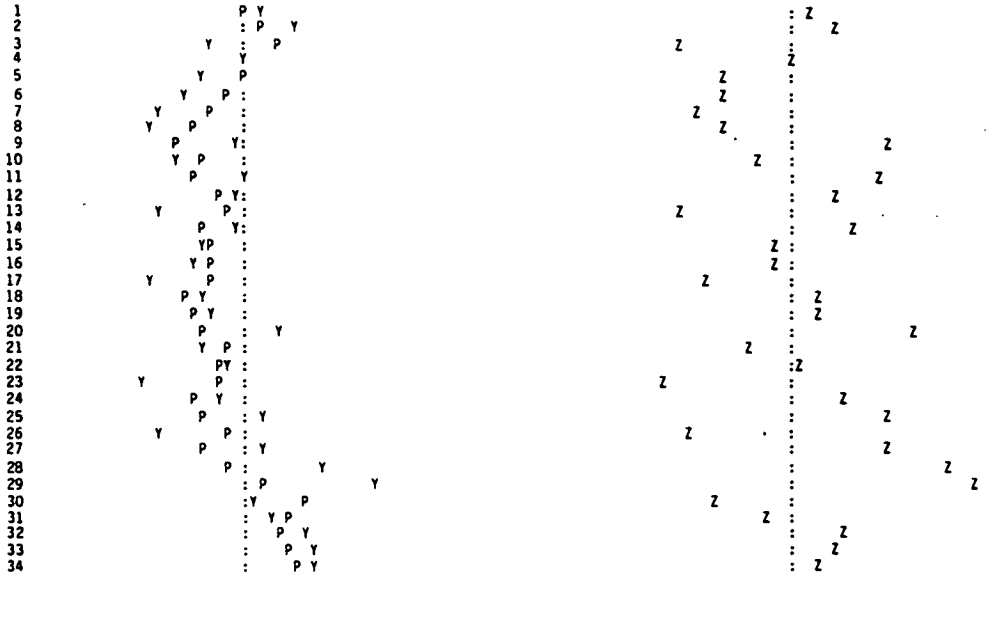
VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	-88.867000	-22.217000	1.6054014	145.16720	Y
2	1	-88.867000	-22.217000	1.6054014	145.16720	F
3	1	-88.867000	-22.217000	1.6054014	145.16720	P

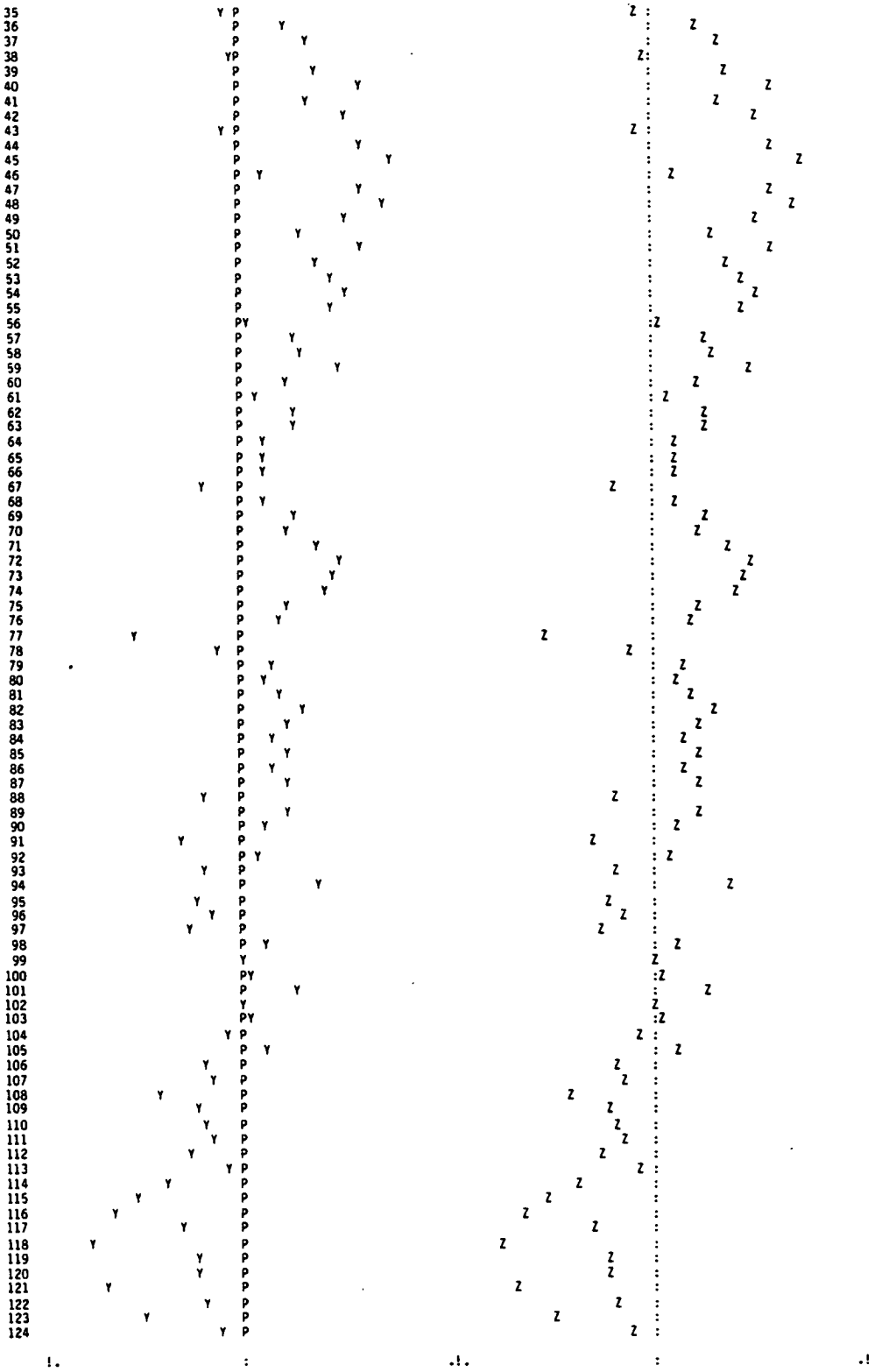


```
 81        FY
 82        FY
 83        FY
 84        Y
 85        Y
 86       FY
 87       FY
 88       YP
 89        Y
 90        Y
 91      YF
 92      PY
 93       Y
 94      FY
 95       Y
 96       Y
 97     YF
 98     PY
 99     PY
100     PY
101     FY
102      Y
103      Y
104      Y
105      Y
106    YF
107    YF
108    YF
109    YF
110    YF
111    YF
112    YF
113    YF
114   YF
115   YF
116  YPF
117   YF
118  YPF
119   YF
120   YF
121  YPF
122   YF
123  YPF
124   YF
```

OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	-1.3958692	1.8597870	15.972203	24.795105	Y
3	1	-1.3958692	1.8597870	15.972203	24.795105	P
4	2	-3.3580244	2.2490033	9.2740758	87.642573	Z





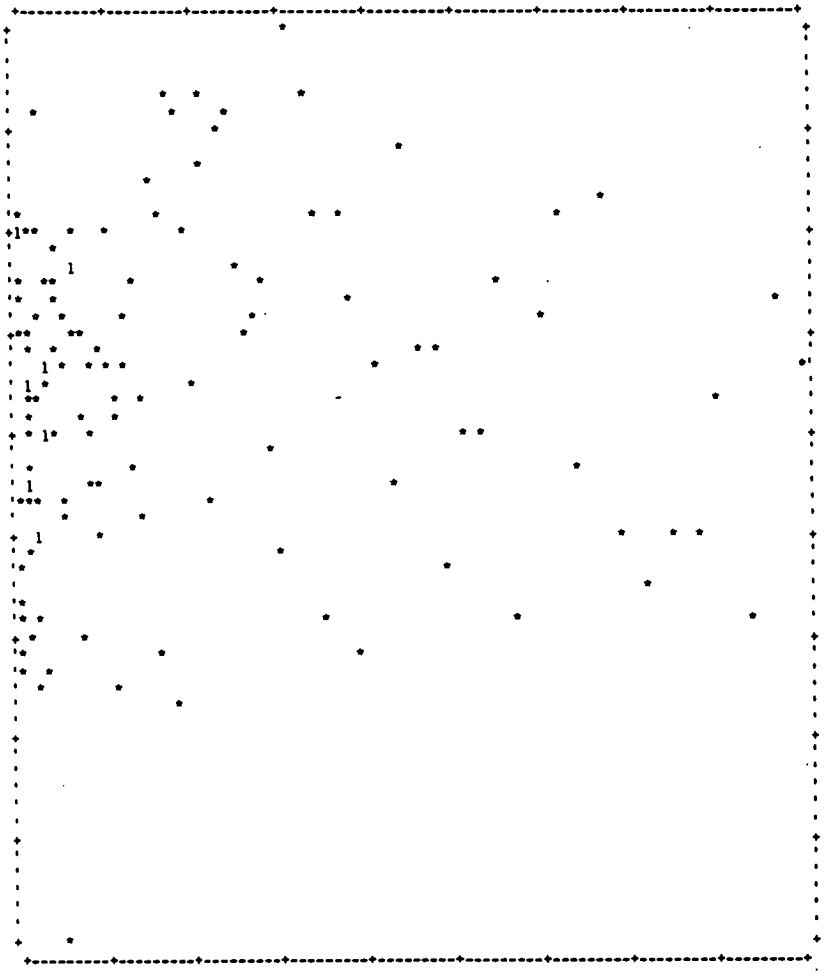
OUTPUT FROM SUBROUTINE PLTACR - C.A.GLASBEY

VARIABLE	MINIMUM	MAXIMUM	*VTV(1)	+VTV(2)	*VTA(1)	+VTA(2)
2	-87.645645	-22.353889	0.82705685	74.987931	1.3784281	123.31322
4	-3.3580246	2.2490034	9.6307705	34.840364	16.051284	56.400607

COUNTS UP TO	0	1	2	4	8	16	32	64	128	256	512	OVER
CHARACTERS		*	1	2	3	4	5	6	7	8	9	0

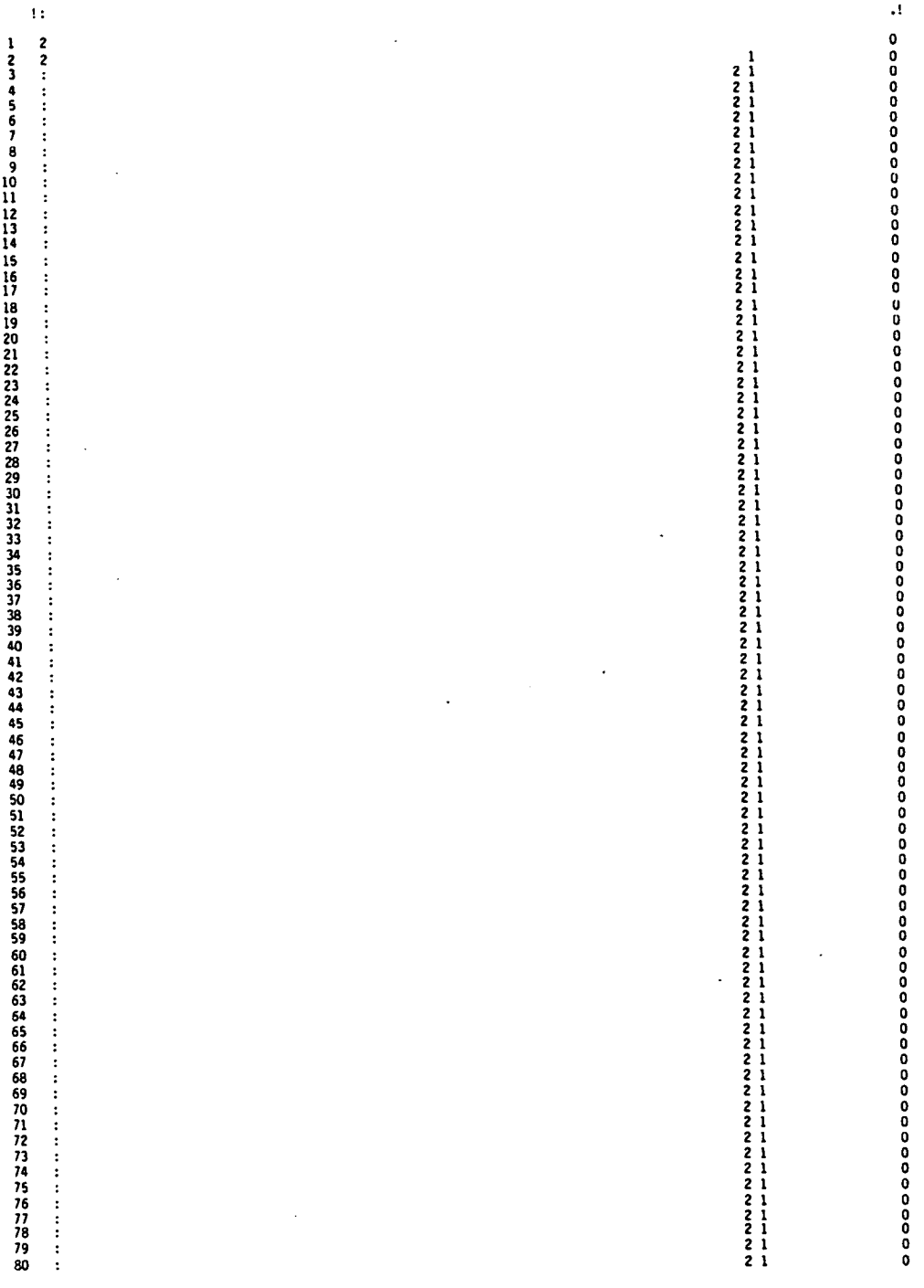
PLOTS
4 VERSUS 2

VERTICAL VARIABLE 4 (-3.358 , 2.249) HORIZONTAL VARIABLE 2 (-87.65 , -22.35)



OUTPUT FROM SUBROUTINE PLTLNG - C.A.GLASBEY

VARIABLE	CHANNEL	MINIMUM	MAXIMUM	*VT(1)	+VT(2)	CHARACTER
1	1	0.000000000+00	0.51748311	206.77003	2.5000000	2
2	1	0.000000000+00	0.51748311	206.77003	2.5000000	1
3	1	0.000000000+00	0.51748311	206.77003	2.5000000	0



```
 81  :                                                       2 1          0
 82  :                                                       2 1          0
 83  :                                                       2 1          0
 84  :                                                       2 1          0
 85  :                                                       2 1          0
 86  :                                                       2 1          0
 87  :                                                       2 1          0
 88  :                                                       2 1          0
 89  :                                                       2 1          0
 90  :                                                       2 1          0
 91  :                                                       2 1          0
 92  :                                                       2 1          0
 93  :                                                       2 1          0
 94  :                                                       2 1          0
 95  :                                                       2 1          0
 96  :                                                       2 1          0
 97  :                                                       2 1          0
 98  :                                                       2 1          0
 99  :                                                       2 1          0
100  :                                                       2 1          0
101  :                                                       2 1          0
102  :                                                       2 1          0
103  :                                                       2 1          0
104  :                                                       2 1          0
105  :                                                       2 1          0
106  :                                                       2 1          0
107  :                                                       2 1          0
108  :                                                       2 1          0
109  :                                                       2 1          0
110  :                                                       2 1          0
111  :                                                       2 1          0
112  :                                                       2 1          0
113  :                                                       2 1          0
114  :                                                       2 1          0
115  :                                                       2 1          0
116  :                                                       2 1          0
117  :                                                       2 1          0
118  :                                                       2 1          0
119  :                                                       2 1          0
120  :                                                       2 1          0
121  :                                                       2 1          0
122  :                                                       2 1          0
123  :                                                       2 1          0
124  :                                                       2 1          0

    !:                                                                   .!
```

A U T O C O R R E L A T I O N S   O F   R E S I D U A L S

-0.0365390 0.0289507 0.0600407 0.0474054 0.0186853 -0.1104928 0.0999713 0.0820289 0.1208935 0.0676121

G E N E R A L I Z E D   P A R T I A L   A U T O C O R R E L A T I O N S

NQV NPV=	0	1	2	3	4	5	6	7	8	9
0	-3.65	2.77	6.22	5.14	1.91	-11.67	8.56	9.34	13.89 +	7.54
1	2.89	4.37	-1.30	-1.14	-9.87	-1.60	8.41	1.34	-3.47	
2	5.99	-1.21	0.19	-2.95	1.22	-0.54	6.09	-3.97		
3	4.71	-0.90	-3.02	-0.16	1.47	4.90	0.57			
4	1.85	-8.76	0.97	1.40	0.90	6.22				
5	-10.96	-1.52	-0.01	4.95	5.66					
6	9.80	7.95	4.97	0.01						
7	7.96	2.42	-3.25							
8	11.66	-2.68								
9	6.44									

STANDARD ERROR = 8.98

VALUES EXCEEDING 1.5 SE DENOTED BY +
VALUES EXCEEDING 2.0 SE DENOTED BY X

Appendix C

CEVOPE computer program listing (see section 9.3).

```
0001  C*************************************************************************C
0002  C*************************************************************************C
0003  C**                                                                       **C
0004  C**      CCCCC    EEEEEEE   V      V    00000    PPPPPP    EEEEEEE     **C
0005  C**    C     C   E         V      V   0     0   P     P   E           **C
0006  C**    C         E         V      V   0     0   P     P   E           **C
0007  C**    C         EEEE      V      V   0     0   PPPPPP    EEEE        **C
0008  C**    C         E          V    V    0     0   P         E           **C
0009  C**    C     C   E           V  V     0     0   P         E           **C
0010  C**     CCCCC    EEEEEEE      V       00000    P         EEEEEEE     **C
0011  C**                                                                       **C
0012  C**                                                                       **C
0013  C**                        PROGRAM TO CALCULATE                           **C
0014  C**              CONSERVATIVE (UPWARD BIASED) ESTIMATES OF                 **C
0015  C**            THE VARIANCES OF REGRESSION PARAMETER ESTIMATORS            **C
0016  C**               FOR CLASSES OF SERIALLY CORRELATED ERRORS                **C
0017  C**                                                                       **C
0018  C**                           C.A.GLASBEY                                 **C
0019  C**                     A.F.R.C. UNIT OF STATISTICS                       **C
0020  C**                       UNIVERSITY OF EDINBURGH                         **C
0021  C**                                                                       **C
0022  C*************************************************************************C
0023  C*************************************************************************C
0024  C
0025  C
0026  C                          SUBROUTINES
0027  C
0028  C ┌──────────────────────────────────────────────────────────────────────┐
0029  C │                          CVINPT                                      │
0030  C │                          CVVARS                                      │
0031  C │                          CVSTTC                                      │
0032  C │                          CVIXXT                                      │
0033  C │                          CVINIT                                      │
0034  C │                          CVOMEG                                      │
0035  C │                          CVEIGC                                      │
0036  C │                          CVVOPT                                      │
0037  C │                          CVMODI                                      │
0038  C │                          CVINCR                                      │
0039  C │                          CVLINP                                      │
0040  C │                          CVRESU                                      │
0041  C └──────────────────────────────────────────────────────────────────────┘
0042  C
0043  C                        REGAME ROUTINE
0044  C
0045  C ┌──────────────────────────────────────────────────────────────────────┐
0046  C │                          REGTRA                                      │
0047  C └──────────────────────────────────────────────────────────────────────┘
0048  C
0049  C                         NAG ROUTINES
0050  C
0051  C ┌──────────────────────────────────────────────────────────────────────┐
0052  C │                          F01AAF                                      │
0053  C │                          G05CBF                                      │
0054  C │                          G05CCF                                      │
0055  C │                          G05DDF                                      │
0056  C │                          G05DYF                                      │
0057  C │                          H01ADF                                      │
0058  C │                          X02AAF                                      │
0059  C │                          X02ACF                                      │
0060  C └──────────────────────────────────────────────────────────────────────┘
0061  C
0062  C                    AUXILIARY NAG ROUTINES
0063  C
0064  C ┌──────────────────────────────────────────────────────────────────────┐
```

```
0065     C        F03AFF  F04AJF  X03AAF  P01AAF  X04AAF                    !
0066     C        Y13AAF  Y13ACF  Y13AEF  Y13AFF  Y13AGF                    !
0067     C        Y13ADF  Y13ABF  G05CAF  G05CGZ  G05CFZ                    !
0068     C        BG05CC                                                    !
0069     C_____________________________________________________________________!
0070     C
0071     C               AUXILIARY OUTPUT ROUTINES
0072     C
0073     C_____________________________________________________________________!
0074     C                       /NINOUT/                                      !
0075     C                        MATPN0                                       !
0076     C                        MATPN1                                       !
0077     C                        MATPR0                                       !
0078     C                        MATPR1                                       !
0079     C                        MATPR2                                       !
0080     C_____________________________________________________________________!
0081     C
0082     C          IMPLEMENTATION DEPENDENT ROUTINES
0083     C
0084     C_____________________________________________________________________!
0085     C                        CLOCK                                        !
0086     C                        DTIM$A                                       !
0087     C_____________________________________________________________________!
0088     C
0089     C
0090     C        ARRAY DIMENSIONS
0091     C
0092     C       DOUBLE PRECISION
0093     C      1 PARAM(NPARMX),RESID(NLMAX),
0094     C      1 U(NLMAX,NPARMX),X(NLMAX,NPARMX),XT(NLMAX,NPARMX),
0095     C      1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0096     C      1 UVU(NPARMX,NPARMX),XVIX(NPARMX,NPARMX),
0097     C      1 XVIXI(NPARMX,NPARMX),SE(NPARMX),
0098     C      1 C(NLMAX,NTMAX),CWT(NTMAX),VOPT(NTBMAX),CNEW(NLMAX),
0099     C      1 XLHS(NTAMAX,NVARMX),XRHS(NTMAX),W1(NLMAX),W2(NLMAX),
0100     C      1 W3(NVARMX),W4(NTBMAX,NTAMAX),W5(NLMAX),
0101     C      1 ACCTOL,XOPT,XOPTLD,QUAD,RBIAS,DTIM$A
0102     C       INTEGER
0103     C      1 MINEQ(NTMAX),MOPT(NTMAX),MW1(NTMAX),MW2(NTVRMX),
0104     C      1 MW3(NTTMAX)
0105     C
0106           DOUBLE PRECISION
0107          1 PARAM(5),RESID(200),
0108          1 U(200,5),X(200,5),XT(200,5),
0109          1 PHI(200,6),THETA(200,6),
0110          1 UVU(5,5),XVIX(5,5),
0111          1 XVIXI(5,5),SE(5),
0112          1 C(200,30),CWT(30),VOPT(32),CNEW(200),
0113          1 XLHS(31,10000),XRHS(30),W1(200),W2(200),
0114          1 W3(10000),W4(32,31),W5(200),
0115          1 ACCTOL,XOPT,XOPTLD,QUAD,RBIAS,DTIM$A
0116           INTEGER
0117          1 MINEQ(30),MOPT(30),MW1(30),MW2(10030),
0118          1 MW3(60)
0119           INTEGER*4 INT4
0120           COMMON
0121          1 /CVCOM1/ U,X,XT,PHI,THETA,C,W3
0122          1 /CVCOM2/ VOPT,XRHS,W4,XOPT,XLHS
0123          1 /CVCOM3/ MINEQ,MOPT,MW1,MW3,NLHS,NT,NVAR,NTBMAX,ITLP,IFAIL,MW2
0124          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0125     C
0126     C        INITIALISE VARIABLES
0127     C
0128           DATA
```

```
0129          1 NLMAX,NPARMX,NPAMAX,NQAMAX,NTMAX,NVARMX,NLHSSP,ACCTOL
0130          1/  200,     5,     6,     6,   30, 10000,     5,1.0D-10/
0131     C
0132           NTAMAX=NTMAX+1
0133           NTBMAX=NTMAX+2
0134           NTVRMX=NTMAX+NVARMX
0135           NTTMAX=2*NTMAX
0136     C
0137     C        INPUT AND CHECK PARAMETERS
0138     C
0139     001    WRITE (NOUT6,902)
0140     902    FORMAT (/' INPUT (10I7):'/'     ND  NVARS',
0141          1 '     NS  NSEED     NT NOMEGA   NLAG  NITER   NCPU   NPRI')
0142           READ (NIN5,903) ND,NVARS,NS,NSEED,NT,NOMEGA,NLAG,NITER,NCPU,NPRI
0143     903   FORMAT (10I7)
0144     C
0145     C
0146     C  ND      NUMBER OF DATA SETS TO ADVANCE FROM CHANNELS NIN3,NIN4
0147     C          (NEGATIVE TO STOP PROGRAM)
0148     C  NVARS   1 TO CALCULATE MODEL DEPENDENT RESULTS IN CVVARS
0149     C  NS      NO. OF PARAMETER FOR WHICH CONSERVATIVE ESTIMATE BEING SOUGHT
0150     C  NSEED   SEED FOR NAG RANDOM NUMBER GENERATOR USED IN CVINIT, CVEIG
0151     C          (NEGATIVE TO START PROGRAM FROM CHANNEL NIN2)
0152     C  NT      INITIAL NUMBER OF VECTORS IN C-ARRAY (UNLESS NSEED NEGATIVE)
0153     C  NOMEGA  DIFFERENCE OPERATOR APPLIED TO CLASS OF VARIANCE MATRICES
0154     C  NLAG    LAG BEYOND WHICH VARIANCE TERMS ARE ZERO
0155     C  NITER   MAXIMUM NUMBER OF ITERATIONS
0156     C  NCPU    MAXIMUM CPU TIME
0157     C  NPRI    1 TO OUTPUT INTERMEDIATE CALCULATIONS
0158     C
0159     C
0160           IF (ND.LT.0) STOP
0161           IF ((NT.LT.1).OR.(NT.GT.NTMAX)) GO TO 023
0162           IF (NOMEGA.LT.0) GO TO 025
0163           IF (NLAG.LT.0) GO TO 027
0164           IF (NITER.LE.0) GO TO 029
0165           IFAIL=0
0166           CALL CLOCK(TCPUST)
0167           TCPULD=TCPUST
0168           TCPULM=TCPUST+FLOAT(NCPU)
0169           TDISLD=SNGL(DTIM$A(INT4))
0170     C
0171     C        INPUT AND CHECK MATRICES, USING CVINPT
0172     C
0173           IF (ND.EQ.0) GO TO 005
0174           DO 004 ID=1,ND
0175.01        CALL CVINPT(PARAM,NPARMX,NPAR,RESID,NLMAX,NL,U,X,XT,
0176.01       1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0177.01       1            XVIX,XVIXI,W1,ACCTOL,NPRI,IFAIL)
0178.01        IF (IFAIL.NE.0) GO TO 001
0179.01  004   CONTINUE
0180     005   IF (NLAG.GT.NL-1) GO TO 027
0181           NVAR=INTS((INTL(2*NL-NLAG)*INTL(NLAG+1))/2)
0182           IF (NVAR.GT.NVARMX) GO TO 031
0183     C
0184     C        CALCULATE PARAMETER ESTIMATOR VARIANCES
0185     C        ASSUMING ERROR VARIANCE MATRIX IS V, USING CVVARS
0186     C
0187           CALL CVVARS(PARAM,NPARMX,NPAR,U,NLMAX,NL,X,
0188          1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0189          1            UVU,XVIX,XVIXI,SE,W1,W2,
0190          1            NVARS,NPRI,IFAIL)
0191           IF (IFAIL.NE.0) GO TO 001
0192           IF ((NS.LE.0).OR.(NS.GT.NPAR)) GO TO 033
```

```
0193     C
0194     C       SELECT FIRST SET OF C-VECTORS, USING CVSTTC
0195     C
0196           CALL CVSTTC(XT,NLMAX,NL,NPARMX,NPAR,X,NSEED,C,NTMAX,NT,
0197          1            W1,W5,ACCTOL,NPRI,IFAIL)
0198           IF (IFAIL.NE.0) GO TO 001
0199           WRITE (NOUT6,906) NS,UVU(NS,NS),SE(NS),NSEED
0200     906   FORMAT (//18X,'*** O U T P U T   F R O M   C E V O P E ***'//
0201          1 11X,'MODEL DEPENDENT ESTIMATE OF THE VARIANCE OF THE ESTIMATOR'/
0202          1 6X,'OF PARAMETER NUMBER',I3,'  IS',1PG12.4,
0203          1 '  AND THE S.E. IS',G12.4//18X,
0204          1 'SEARCH FOR ROBUST ESTIMATOR WITH MINIMUM BIAS (NSEED=',I5,')'//
0205          1 '                    ADD NEW C-VECTOR    ',
0206          1 '   FIND BEST COMBINATION OF C-VECTORS'/
0207          1 '              ========================',
0208          1 '  ====================================='/
0209          1 ' ITER ICPUT NT ITEIG ICPU IDIS   QUAD  ',
0210          1 '  ITLP ICPU IDIS    XOPT      REL BIAS'/
0211          1 ' ==== ===== == ===== ==== ====  =======',
0212          1 '  ==== ==== ====  ==========  ==========')
0213     C
0214     C       SET INPUT FOR LINEAR PROGRAM, USING CVINIT
0215     C
0216           NLHS=NTAMAX-((NTMAX-NT)/NLHSSP)*NLHSSP
0217           CALL CVINIT(U,NLMAX,NL,NPARMX,NPAR,PHI,NPAMAX,NPA,
0218          1            THETA,NQAMAX,NQA,C,NTMAX,NT,NOMEGA,NLAG,NS,
0219          1            XLHS,NLHS,NTA,NVARMX,NVAR,XRHS,MINEQ,
0220          1            W1,W2,W3,ACCTOL,NPRI,IFAIL)
0221           IF (IFAIL.NE.0) GO TO 001
0222           ITER=0
0223           XOPTLD=1.0D20
0224           GO TO 008
0225     C
0226     C       FIND NEW C-VECTOR, USING CVEIGC
0227     C
0228     007   ITER=ITER+1
0229           CALL CVEIGC(XT,NLMAX,NL,NPARMX,NPAR,X,PHI,NPAMAX,NPA,
0230          1            THETA,NQAMAX,NQA,VOPT,NTBMAX,MOPT,NTMAX,NT,NOMEGA,
0231          1            NLAG,NVAR,CNEW,QUAD,W1,W2,W3,NVARMX,
0232          1            W5,ITEIG,ACCTOL,NPRI,IFAIL)
0233           IF (IFAIL.NE.0) GO TO 021
0234           IF (QUAD.GE.-ACCTOL) GO TO 013
0235     C
0236     C       MODIFY INPUT FOR LINEAR PROGRAM, USING CVMODI
0237     C
0238           CALL CVMODI(PHI,NLMAX,NL,NPAMAX,NPA,THETA,NQAMAX,NQA,
0239          1            C,NTMAX,NT,CWT,CNEW,NOMEGA,NLAG,NVAR,
0240          1            XLHS,NLHS,NVARMX,XRHS,
0241          1            W1,W2,W3,MW1,MW3,NTTMAX,ACCTOL,NPRI,IFAIL)
0242           IF (IFAIL.NE.0) GO TO 021
0243     C
0244     C       INCREASE FIRST DIMENSION OF XLHS IF NECESSARY, USING CVINCR
0245     C
0246           IF (NT+1.LE.NLHS) GO TO 008
0247           NLHS=NLHS+NLHSSP
0248           CALL CVINCR(XLHS,NLHS,NVARMX,NVAR,W1,NLMAX,W3,NT,NPRI)
0249     C
0250     C       OPTIMIZE COMBINATION OF C-VECTORS BY LINEAR PROGRAM,
0251     C       USING CVLINP
0252     C
0253     008   CALL CLOCK(TCPUA)
0254           TDIS=SNGL(DTIM$A(INT4))
0255           CALL CVLINP(XLHS,NLHS,NVARMX,NVAR,XRHS,NTMAX,NT,
0256          1            MINEQ,XOPT,CWT,VOPT,NTBMAX,MOPT,NL,
```

```
     1           W4,NTAMAX,MW1,MW2,NTVRMX,MW3,NTTMAX,ITLP,NPRI,IFAIL)
      IF (IFAIL.NE.0) GO TO 001
C
C        CALCULATE CPU AND DISC TIMES
C
      CALL CLOCK(TCPU)
      ICPUT=INT(TCPU-TCPUST+0.5)
      ICPU1=INT(TCPUA-TCPULD+0.5)
      ICPU2=INT(TCPU-TCPUA+0.5)
      IDIS1=INT(TDIS-TDISLD+0.5)
      TDISLD=TDIS
      TDIS=SNGL(DTIM$A(INT4))
      IDIS2=INT(TDIS-TDISLD+0.5)
      TDISLD=TDIS
      RBIAS=XOPT/SE(NS)-1.0D0
      IF (ITER.GT.0) GO TO 010
      WRITE (NOUT6,909) ITER,ICPUT,NT,ICPU1,IDIS1,
     1 ITLP,ICPU2,IDIS2,XOPT,RBIAS
909   FORMAT (I5,I6,I3,6X,2I5,9X,I6,2I5,1PG12.4,G12.4)
      GO TO 012
010   WRITE (NOUT6,911) ITER,ICPUT,NT,ITEIG,ICPU1,IDIS1,QUAD,
     1 ITLP,ICPU2,IDIS2,XOPT,RBIAS
911   FORMAT (I5,I6,I3,I6,2I5,1PD9.0,I6,2I5,G12.4,G12.4)
C
C        CHECK STOPPING CONDITIONS
C
012   IF (XOPT.GT.XOPTLD*(1.0D0-ACCTOL)) GO TO 015
      IF (ITER.GE.NITER) GO TO 017
      IF (2.0*(TCPU-TCPULD).GT.(TCPULM-TCPU)) GO TO 019
      XOPTLD=XOPT
      TCPULD=TCPU
      GO TO 007
C
C        OUTPUT STOPPING CRITERION
C
013   WRITE (NOUT6,914)
914   FORMAT (/' ITERATIONS CEASED, OPTIMUM REACHED')
      GO TO 021
015   WRITE (NOUT6,916)
916   FORMAT (/' ITERATIONS CEASED, NO DECREASE IN XOPT')
      GO TO 021
017   WRITE (NOUT6,918)
918   FORMAT (/' ITERATIONS CEASED, NITER REACHED')
      GO TO 021
019   WRITE (NOUT6,920)
920   FORMAT (/' ITERATIONS CEASED, CPU LIMIT MAY BE EXCEEDED ON NEXT ',
     1 'ITERATION')
C
C        CALCULATE FINAL RESULTS, USING CVRESU
C
021   CONTINUE
      CALL CVRESU(PARAM,NPARMX,NPAR,RESID,NLMAX,NL,
     1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,UVU,
     1            C,NTMAX,NT,CWT,ND,NS,NOMEGA,NLAG,
     1            W1,W2,W4,NTBMAX,NTAMAX,NPRI)
      GO TO 001
C
C        ERROR MESSAGES
C
023   WRITE (NOUT6,924)
924   FORMAT (' *** CEVOPE, NT OUT OF RANGE ***')
      GO TO 001
025   WRITE (NOUT6,926)
926   FORMAT (' *** CEVOPE, NOMEGA OUT OF RANGE ***')
```

```
0321            GO TO 001
0322      027   WRITE (NOUT6,928)
0323      928   FORMAT (' *** CEVOPE, NLAG OUT OF RANGE ***')
0324            GO TO 001
0325      029   WRITE (NOUT6,930)
0326      930   FORMAT (' *** CEVOPE, NITER OUT OF RANGE ***')
0327            GO TO 001
0328      031   WRITE (NOUT6,932)
0329      932   FORMAT (' *** CEVOPE, NL AND NVAR COMBINED OUT OF RANGE ***')
0330            GO TO 001
0331      033   WRITE (NOUT6,934)
0332      934   FORMAT (' *** CEVOPE, NS OUT OF RANGE ***')
0333            GO TO 001
0334            END
0335
0336            SUBROUTINE CVINPT(PARAM,NPARMX,NPAR,RESID,NLMAX,NL,U,X,XT,
0337           1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0338           1            XVIX,XVIXI,W1,ACCTOL,NPRI,IFAIL)
0339      C
0340      C       INPUT AND CHECK MATRICES
0341      C
0342            DOUBLE PRECISION
0343           1 PARAM(NPARMX),RESID(NLMAX),
0344           1 U(NLMAX,NPARMX),X(NLMAX,NPARMX),XT(NLMAX,NPARMX),
0345           1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0346           1 XVIX(NPARMX,NPARMX),XVIXI(NPARMX,NPARMX),W1(NLMAX),ACCTOL,XX
0347            COMMON
0348           1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0349      C
0350      C
0351      C       READS OUTPUT FILE CREATED BY REGAME.
0352      C              INPUT FROM NIN4:  NL                (2I5)
0353      C                                RESID(1...NL)     (5G24.16)
0354      C                                NPAR
0355      C                                PARAM(1...NPAR)
0356      C                                X(1...NL,J)
0357      C                                        J=1...NPAR
0358      C                                U(1...NL,J)
0359      C                                        J=1...NPAR
0360      C
0361            READ (NIN4,901) NL
0362      901   FORMAT (2I5)
0363            IF ((NL.LT.1).OR.(NL.GT.NLMAX)) GO TO 017
0364            READ (NIN4,902) (RESID(I),I=1,NL)
0365      902   FORMAT (5G24.16)
0366            IF (NPRI.EQ.1) CALL MATPR1(RESID,NL,1)
0367            READ (NIN4,901) NPAR
0368            IF ((NPAR.LT.1).OR.(NPAR.GT.NPARMX)) GO TO 019
0369            READ (NIN4,902) (PARAM(J),J=1,NPAR)
0370            IF (NPRI.EQ.1) CALL MATPR1(PARAM,NPAR,1)
0371            DO 003 J=1,NPAR
0372.01   003   READ (NIN4,902) (X(I,J),I=1,NL)
0373            IF (NPRI.EQ.1) CALL MATPR2(X,NLMAX,NL,NPAR,1)
0374            DO 004 J=1,NPAR
0375.01   004   READ (NIN4,902) (U(I,J),I=1,NL)
0376            IF (NPRI.EQ.1) CALL MATPR2(U,NLMAX,NL,NPAR,1)
0377      C
0378      C       CALCULATES XT=X(XTX)-
0379      C
0380            DO 010 I=1,NPAR
0381.01         DO 010 J=1,NPAR
0382.02         XX=0.0D0
0383.02         DO 009 K=1,NL
0384.03   009   XX=XX+X(K,I)*X(K,J)
```

```
0385.02  010   XVIX(I,J)=XX
0386           IF (NPRI.EQ.1) CALL MATPR2(XVIX,NPARMX,NPAR,NPAR,1)
0387     C
0388           IFAIL=1
0389           CALL F01AAF(XVIX,NPARMX,NPAR,XVIXI,NPARMX,W1,IFAIL)
0390           IF (IFAIL.NE.0) GO TO 021
0391           IF (NPRI.EQ.1) CALL MATPR2(XVIXI,NPARMX,NPAR,NPAR,1)
0392     C
0393           DO 012 I=1,NL
0394.01        DO 012 K=1,NPAR
0395.02        XX=0.0D0
0396.02        DO 011 J=1,NPAR
0397.03  011   XX=XX+X(I,J)*XVIXI(J,K)
0398.02  012   XT(I,K)=XX
0399           IF (NPRI.EQ.1) CALL MATPR2(XT,NLMAX,NL,NPAR,1)
0400     C
0401     C
0402     C                INPUT FROM NIN3:  NPA,NQA
0403     C                                  PHI(1...NL,J)
0404     C                                                J=1...NPA
0405     C                                  THETA(1...NL,J)
0406     C                                                  J=1...NQA
0407     C
0408     C
0409           READ (NIN3,901) NPA,NQA
0410           IF ((NPA.LT.1).OR.(NPA.GT.NPAMAX)) GO TO 023
0411           IF ((NQA.LT.1).OR.(NQA.GT.NQAMAX)) GO TO 025
0412           DO 013 J=1,NPA
0413.01  013   READ (NIN3,902) (PHI(I,J),I=1,NL)
0414           IF (NPRI.EQ.1) CALL MATPR2(PHI,NLMAX,NL,NPA,1)
0415           DO 014 I=1,NL
0416.01        IF (PHI(I,NPA).EQ.0.0D0) GO TO 027
0417.01  014   CONTINUE
0418.01  C
0419           DO 015 J=1,NQA
0420.01  015   READ (NIN3,902) (THETA(I,J),I=1,NL)
0421           IF (NPRI.EQ.1) CALL MATPR2(THETA,NLMAX,NL,NQA,1)
0422           DO 016 I=1,NL
0423.01        IF (THETA(I,NQA).EQ.0.0D0) GO TO 029
0424.01  016   CONTINUE
0425           RETURN
0426     C
0427     C        ERROR MESSAGES
0428     C
0429     017   WRITE (NOUT6,918)
0430     918   FORMAT (' *** CVINPT, NL OUT OF RANGE ***')
0431           IFAIL=1
0432           RETURN
0433     019   WRITE (NOUT6,920)
0434     920   FORMAT (' *** CVINPT, NPAR OUT OF RANGE ***')
0435           IFAIL=1
0436           RETURN
0437     021   WRITE (NOUT6,922)
0438     922   FORMAT (' *** CVINPT, X NOT OF FULL RANK ***')
0439           RETURN
0440     023   WRITE (NOUT6,924)
0441     924   FORMAT (' *** CVINPT, NPA OUT OF RANGE ***')
0442           IFAIL=1
0443           RETURN
0444     025   WRITE (NOUT6,926)
0445     926   FORMAT (' *** CVINPT, NQA OUT OF RANGE ***')
0446           IFAIL=1
0447           RETURN
0448     027   WRITE (NOUT6,928)
```

```
0449     928   FORMAT (' *** CVINPT, PHI SINGULAR ***')
0450           IFAIL=1
0451           RETURN
0452     029   WRITE (NOUT6,930)
0453     930   FORMAT (' *** CVINPT, THETA SINGULAR ***')
0454           IFAIL=1
0455           RETURN
0456           END
0457
0458           SUBROUTINE CVVARS(PARAM,NPARMX,NPAR,U,NLMAX,NL,X,
0459          1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0460          1            UVU,XVIX,XVIXI,SE,W1,W2,
0461          1            NVARS,NPRI,IFAIL)
0462     C
0463     C        CALCULATE PARAMETER ESTIMATOR VARIANCES
0464     C        ASSUMING ERROR VARIANCE MATRIX IS V
0465     C
0466           DOUBLE PRECISION
0467          1 PARAM(NPARMX),U(NLMAX,NPARMX),X(NLMAX,NPARMX),
0468          1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0469          1 UVU(NPARMX,NPARMX),XVIX(NPARMX,NPARMX),
0470          1 XVIXI(NPARMX,NPARMX),SE(NPARMX),
0471          1 W1(NLMAX),W2(NLMAX),XX
0472           COMMON
0473          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0474     C
0475           DO 004 J=1,NPAR
0476.01        DO 001 I=1,NL
0477.02  001   W1(I)=U(I,J)
0478.01        CALL REGTRA(W1,NL,PHI,NLMAX,NPA,1,1,W2,0,NPRI,IFAIL)
0479.01        CALL REGTRA(W2,NL,THETA,NLMAX,NQA,0,1,W1,0,NPRI,IFAIL)
0480.01        CALL REGTRA(W1,NL,THETA,NLMAX,NQA,0,0,W2,0,NPRI,IFAIL)
0481.01        CALL REGTRA(W2,NL,PHI,NLMAX,NPA,1,0,W1,0,NPRI,IFAIL)
0482.01        DO 003 K=1,J
0483.02        XX=0.0D0
0484.02        DO 002 I=1,NL
0485.03  002   XX=XX+W1(I)*U(I,K)
0486.02        UVU(K,J)=XX
0487.02  003   UVU(J,K)=XX
0488.01  004   SE(J)=DSQRT(UVU(J,J))
0489           IF (NVARS.NE.1) RETURN
0490     C
0491     C        IF NVARS=1 OUTPUT RESULTS
0492     C
0493           WRITE (NOUT6,905)
0494     905   FORMAT (/' PARAMETER ESTIMATES')
0495           CALL MATPR1(PARAM,NPAR,0)
0496           WRITE (NOUT6,906)
0497     906   FORMAT (/' STANDARD ERRORS')
0498           CALL MATPR1(SE,NPAR,0)
0499           WRITE (NOUT6,907)
0500     907   FORMAT (/' PARAMETER ESTIMATOR VARIANCES ASSUMING V')
0501           CALL MATPR2(UVU,NPARMX,NPAR,NPAR,0)
0502     C
0503     C        CALCULATE EFFICIENCY OF ESTIMATORS
0504     C
0505           DO 010 J=1,NPAR
0506.01        DO 008 I=1,NL
0507.02  008   W1(I)=X(I,J)
0508.01        CALL REGTRA(W1,NL,PHI,NLMAX,NPA,0,0,W2,0,NPRI,IFAIL)
0509.01        CALL REGTRA(W2,NL,THETA,NLMAX,NQA,1,0,W1,0,NPRI,IFAIL)
0510.01        CALL REGTRA(W1,NL,THETA,NLMAX,NQA,1,1,W2,0,NPRI,IFAIL)
0511.01        CALL REGTRA(W2,NL,PHI,NLMAX,NPA,0,1,W1,0,NPRI,IFAIL)
0512.01        DO 010 K=1,J
```

```
0513.02          XX=0.0D0
0514.02          DO 009 I=1,NL
0515.03  009     XX=XX+W1(I)*X(I,K)
0516.02          XVIX(K,J)=XX
0517.02  010     XVIX(J,K)=XX
0518             IF (NPRI.EQ.1) CALL MATPR2(XVIX,NPARMX,NPAR,NPAR,1)
0519       C
0520             IFAIL=1
0521             CALL F01AAF(XVIX,NPARMX,NPAR,XVIXI,NPARMX,W1,IFAIL)
0522             IF (IFAIL.NE.0) GO TO 014
0523             WRITE (NOUT6,911)
0524       911   FORMAT (/' LOWER BOUNDS ON VARIANCES')
0525             CALL MATPR2(XVIXI,NPARMX,NPAR,NPAR,0)
0526       C
0527             DO 012 J=1,NPAR
0528.01  012     W1(J)=XVIXI(J,J)/UVU(J,J)
0529             WRITE (NOUT6,913)
0530       913   FORMAT (/' EFFICIENCY OF ESTIMATORS')
0531             CALL MATPR1(W1,NPAR,0)
0532             RETURN
0533       C
0534       014   WRITE (NOUT6,915)
0535       915   FORMAT (' *** CEVOPE, XVIXI SINGULAR ***')
0536             RETURN
0537             END
0538
0539             SUBROUTINE CVSTTC(XT,NLMAX,NL,NPARMX,NPAR,X,NSEED,C,NTMAX,NT,
0540            1            W1,W5,ACCTOL,NPRI,IFAIL)
0541       C
0542       C       SELECT FIRST SET OF C-VECTORS
0543       C
0544             DOUBLE PRECISION
0545            1 XT(NLMAX,NPARMX),X(NLMAX,NPARMX),
0546            1 C(NLMAX,NTMAX),
0547            1 W1(NLMAX),W5(NLMAX),ACCTOL,G05DDF,XX
0548             INTEGER
0549            1 G05DYF
0550             COMMON
0551            1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0552       C
0553             IF (NSEED) 006,001,002
0554       C
0555       C        INITIAL C-VECTORS CHOSEN AT RANDOM
0556       C
0557       001   CALL G05CCF
0558             NSEED=G05DYF(1,9999)
0559       002   CALL G05CBF(NSEED)
0560             DO 005 J=1,NT
0561.01          DO 003 I=1,NL
0562.02  003     W1(I)=G05DDF(0.0D0,1.0D0)
0563.01          IF (NPRI.EQ.1) CALL MATPR1(W1,NL,1)
0564.01          CALL CVIXXT(W1,NLMAX,NL,XT,NPARMX,NPAR,X,W5,ACCTOL,
0565.01         1            NPRI,IFAIL)
0566.01          IF (IFAIL.NE.0) RETURN
0567.01          XX=0.0D0
0568.01          DO 004 I=1,NL
0569.02          IF (DABS(W1(I)).GT.DABS(XX)) XX=W1(I)
0570.02  004     CONTINUE
0571.01          DO 005 I=1,NL
0572.02  005     C(I,J)=W1(I)/XX
0573             IF (NPRI.EQ.1) CALL MATPR2(C,NLMAX,NL,NT,1)
0574             RETURN
0575       C
0576       C        C-VECTORS RESTORED FROM CHANNEL NIN2
```

```
0577      C
0578      006   NSEED=-NSEED
0579            CALL G05CBF(NSEED)
0580            READ (NIN2,907) NT
0581      907   FORMAT (I5)
0582            DO 008 J=1,NT
0583.01   008   READ (NIN2,909) (C(I,J),I=1,NL)
0584      909   FORMAT (5D24.16)
0585            IF (NPRI.EQ.1) CALL MATPR2(C,NLMAX,NL,NT,1)
0586            RETURN
0587            END
0588
0589            SUBROUTINE CVIXXT(W1,NLMAX,NL,XT,NPARMX,NPAR,X,W5,ACCTOL,
0590           1            NPRI,IFAIL)
0591      C
0592      C        PERFORM TRANSFORMATION W1=(I-(XT)X)W1
0593      C
0594            DOUBLE PRECISION
0595           1 XT(NLMAX,NPARMX),X(NLMAX,NPARMX),
0596           1 W1(NLMAX),W5(NLMAX),ACCTOL,XX,XY,XZ
0597            COMMON
0598           1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0599            DO 005 ITER=1,3
0600.01         DO 002 I=1,NPAR
0601.02         XX=0.0D0
0602.02         DO 001 J=1,NL
0603.03   001   XX=XX+X(J,I)*W1(J)
0604.02   002   W5(I)=XX
0605.01         IF (NPRI.EQ.1) CALL MATPR1(W5,NPAR,1)
0606.01   C
0607.01         XY=0.0D0
0608.01         XZ=0.0D0
0609.01         DO 004 I=1,NL
0610.02         XY=XY+DABS(W1(I))
0611.02         XX=0.0D0
0612.02         DO 003 J=1,NPAR
0613.03   003   XX=XX+XT(I,J)*W5(J)
0614.02         W1(I)=W1(I)-XX
0615.02   004   XZ=XZ+DABS(XX)
0616.01         IF (NPRI.EQ.1) CALL MATPR1(W1,NL,1)
0617.01         IF (XZ/XY.LT.ACCTOL) RETURN
0618.01   005   CONTINUE
0619            WRITE (NOUT6,906)
0620      906   FORMAT (' *** CVIXXT, FAILURE TO CONVERGE ***')
0621            IFAIL=1
0622            RETURN
0623            END
0624
0625            SUBROUTINE CVINIT(U,NLMAX,NL,NPARMX,NPAR,PHI,NPAMAX,NPA,
0626           1            THETA,NQAMAX,NQA,C,NTMAX,NT,NOMEGA,NLAG,NS,
0627           1            XLHS,NLHS,NTA,NVARMX,NVAR,XRHS,MINEQ,
0628           1            W1,W2,W3,ACCTOL,NPRI,IFAIL)
0629      C
0630      C       SET INPUT FOR LINEAR PROGRAM
0631      C
0632            DOUBLE PRECISION
0633           1 U(NLMAX,NPARMX),
0634           1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0635           1 C(NLMAX,NTMAX),
0636           1 XLHS(NLHS,NVARMX),XRHS(NTMAX),W1(NLMAX),W2(NLMAX),
0637           1 W3(NVARMX),ACCTOL,CVC
0638            INTEGER
0639           1 MINEQ(NTMAX)
0640            COMMON
```

```
0641          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0642     C
0643           DO 001 I=1,NTMAX
0644.01  001   MINEQ(I)=-1
0645.01  C
0646           DO 004 J=1,NT
0647.01        DO 002 I=1,NL
0648.02  002   W1(I)=C(I,J)
0649.01        CALL CVOMEG(W1,NLMAX,NL,PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0650.01       1            NOMEGA,NLAG,CVC,W3,NVARMX,NVAR,
0651.01       1            W2,ACCTOL,NPRI,IFAIL)
0652.01        IF (IFAIL.NE.0) RETURN
0653.01        DO 003 K=1,NVAR
0654.02  003   XLHS(J,K)=W3(K)
0655.01        XRHS(J)=CVC
0656.01  004   CONTINUE
0657           IF (NPRI.EQ.1) CALL MATPR1(XRHS,NT,1)
0658     C
0659           DO 005 I=1,NL
0660.01  005   W1(I)=U(I,NS)
0661           CALL CVOMEG(W1,NLMAX,NL,PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0662          1            NOMEGA,NLAG,CVC,W3,NVARMX,NVAR,
0663          1            W2,ACCTOL,NPRI,IFAIL)
0664           IF (IFAIL.NE.0) RETURN
0665           NTA=NT+1
0666           DO 006 K=1,NVAR
0667.01  006   XLHS(NTA,K)=-W3(K)
0668           RETURN
0669           END
0670
0671           SUBROUTINE CVOMEG(W1,NLMAX,NL,PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0672          1            NOMEGA,NLAG,CVC,W3,NVARMX,NVAR,
0673          1            W2,ACCTOL,NPRI,IFAIL)
0674     C
0675     C        CALCULATES CVC=(W1)V(W1)
0676     C        AND W3 IS L.H.S. IN LINEAR PROGRAM (DEPENDS ON NOMEGA,NLAG)
0677     C
0678           DOUBLE PRECISION
0679          1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0680          1 W1(NLMAX),W2(NLMAX),
0681          1 W3(NVARMX),ACCTOL,CVC,XX
0682           COMMON
0683          1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0684     C
0685     C        CALCULATE CVC
0686     C
0687           CALL REGTRA(W1,NL,PHI,NLMAX,NPA,1,1,W2,0,NPRI,IFAIL)
0688           CALL REGTRA(W2,NL,THETA,NLMAX,NQA,0,1,W3,0,NPRI,IFAIL)
0689           CVC=0.0D0
0690           DO 001 I=1,NL
0691.01  001   CVC=CVC+W3(I)*W1(I)
0692           IF (NPRI.EQ.1) CALL MATPRO(CVC,1)
0693           IF (CVC.LE.0.0D0) GO TO 007
0694     C
0695     C        CALCULATES W3 IF NOMEGA=0
0696     C
0697           NLAGA=NLAG+1
0698           K=0
0699           DO 002 I=1,NL
0700.01        JZF=MINO(NLAGA,I)
0701.01        DO 002 JZ=1,JZF
0702.02        J=I+1-JZ
0703.02        K=K+1
0704.02        W3(K)=W1(I)*W1(J)
```

```
0705.02          IF (I.NE.J) W3(K)=2.0D0*W3(K)
0706.02  002     CONTINUE
0707             IF (NPRI.EQ.1) CALL MATPR1(W3,NVAR,1)
0708             IF (NOMEGA.EQ.0) RETURN
0709       C
0710       C        CALCULATES W3 IF NOMEGA>0
0711       C
0712             DO 005 IOMEGA=1,NOMEGA
0713.01          K=0
0714.01          KK=0
0715.01          DO 004 I=1,NL
0716.02          JZF=MIN0(NLAGA,I)
0717.02          DO 003 JZ=1,JZF
0718.03          J=I+1-JZ
0719.03          JZZ=JZ-1
0720.03          JZY=JZ-2
0721.03          XX=0.0D0
0722.03          IF (JZ.GE.2) XX=XX+W3(K)+W2(JZZ)
0723.03          IF (JZ.GE.3) XX=XX-W2(JZY)
0724.03          K=K+1
0725.03  003     W3(K)=W3(K)+XX
0726.02          DO 004 JZ=1,JZF
0727.03          KK=KK+1
0728.03  004     W2(JZ)=W3(KK)
0729.01          IF (NPRI.EQ.1) CALL MATPR1(W3,NVAR,1)
0730.01  005     CONTINUE
0731.01  C
0732             DO 006 K=1,NVAR
0733.01          IF (W3(K).LE.ACCTOL) W3(K)=0.0D0
0734.01  006     CONTINUE
0735             RETURN
0736       C
0737       007   WRITE (NOUT6,908)
0738       908   FORMAT (' *** CVOMEG, CVC NON-POSITIVE ***')
0739             IFAIL=1
0740             RETURN
0741             END
0742
0743             SUBROUTINE CVEIGC(XT,NLMAX,NL,NPARMX,NPAR,X,PHI,NPAMAX,NPA,
0744            1            THETA,NQAMAX,NQA,VOPT,NTBMAX,MOPT,NTMAX,NT,NOMEGA,
0745            1            NLAG,NVAR,CNEW,QUAD,W1,W2,W3,NVARMX,
0746            1            W5,ITEIG,ACCTOL,NPRI,IFAIL)
0747       C
0748       C        FIND NEW C-VECTOR
0749       C
0750             DOUBLE PRECISION
0751            1 XT(NLMAX,NPARMX),X(NLMAX,NPARMX),
0752            1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0753            1 VOPT(NTBMAX),CNEW(NLMAX),
0754            1 W1(NLMAX),W2(NLMAX),
0755            1 W3(NVARMX),W5(NLMAX),QUAD,CNEIG,ACCTOL,SHIFT,XDIF,XX,XY,G05DDF
0756             INTEGER
0757            1 MOPT(NTMAX)
0758             COMMON
0759            1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0760       C
0761       C
0762       C        FROM A RANDOM STARTING VECTOR REPEATEDLY APPLY TRANSFORMATION:
0763       C
0764       C        CNEW = (I-(XT)X) ((V-VOPT) CNEW - (SHIFT) CNEW)
0765       C
0766       C        IN THE FOLLOWING STEPS:
0767       C        W1=CNEW
0768       C        W5=(VOPT)W1
```

```
0769     C        W1=(V)W1    BY:   W2=((PHI)T-)W1
0770     C                          W1=((THETA)T)W2
0771     C                          W2=(PHI)W1
0772     C                          W1=((THETA)-)W2
0773     C        W1=W1-W5-(SHIFT)CNEW
0774     C        W1=(I-(XT)X)W1
0775     C        CNEW=STANDARDISED(W1)
0776     C
0777     C
0778           SHIFT=0.0D0
0779           DO 001 I=1,NL
0780.01  001   W1(I)=G05DDF(0.0D0,1.0D0)
0781           CALL CVIXXT(W1,NLMAX,NL,XT,NPARMX,NPAR,X,W5,ACCTOL,
0782          1            NPRI,IFAIL)
0783     C
0784           DO 008 ITEIG=1,1000
0785.01        DO 002 I=1,NL
0786.02  002   CNEW(I)=W1(I)
0787.01        IF (IFAIL.NE.0) RETURN
0788.01        CALL CVVOPT(W1,NLMAX,NL,VOPT,NTBMAX,NT,MOPT,NTMAX,
0789.01       1            NOMEGA,NLAG,NVAR,ITEIG,W5,W2,
0790.01       1            W3,NVARMX,NPRI)
0791.01        CALL REGTRA(W1,NL,PHI,NLMAX,NPA,1,1,W2,0,NPRI,IFAIL)
0792.01        CALL REGTRA(W2,NL,THETA,NLMAX,NQA,0,1,W1,0,NPRI,IFAIL)
0793.01        CALL REGTRA(W1,NL,THETA,NLMAX,NQA,0,0,W2,0,NPRI,IFAIL)
0794.01        CALL REGTRA(W2,NL,PHI,NLMAX,NPA,1,0,W1,0,NPRI,IFAIL)
0795.01        DO 003 I=1,NL
0796.02  003   W1(I)=W1(I)-W5(I)-SHIFT*CNEW(I)
0797.01        IF (NPRI.EQ.1) CALL MATPR1(W1,NL,1)
0798.01        CALL CVIXXT(W1,NLMAX,NL,XT,NPARMX,NPAR,X,W5,ACCTOL,
0799.01       1            NPRI,IFAIL)
0800.01        IF (IFAIL.NE.0) RETURN
0801.01        CNEIG=0.0D0
0802.01        DO 004 I=1,NL
0803.02        IF (DABS(W1(I)).GT.DABS(CNEIG)) CNEIG=W1(I)
0804.02  004   CONTINUE
0805.01        XDIF=0.0D0
0806.01        DO 005 I=1,NL
0807.02        W1(I)=W1(I)/CNEIG
0808.02        XDIF=XDIF+DABS(W1(I)-CNEW(I))
0809.02  005   CONTINUE
0810.01        IF (NPRI.EQ.1) CALL MATPRO(XDIF,1)
0811.01        IF (NPRI.EQ.1) CALL MATPRO(CNEIG,1)
0812.01        IF (NPRI.EQ.1) CALL MATPR1(W1,NL,1)
0813.01        IF ((MOD(ITEIG,250).NE.0).AND.(XDIF.GT.ACCTOL)) GO TO 007
0814.01  C
0815.01  C        CHECK STOPPING CONDITIONS BY CALCULATING:
0816.01  C
0817.01  C        QUAD = CNEW((I-(XT)X) (V-VOPT) (I-(XT)X) CNEW - (SHIFT) CNEW)
0818.01  C
0819.01        XX=0.0D0
0820.01        XY=0.0D0
0821.01        DO 006 I=1,NL
0822.02        XX=XX+CNEW(I)*W1(I)
0823.02        XY=XY+CNEW(I)*CNEW(I)
0824.02  006   CONTINUE
0825.01        QUAD=(CNEIG*XX+SHIFT*XY)/XY
0826.01        IF (NPRI.EQ.1) CALL MATPRO(QUAD,1)
0827.01        IF (QUAD.LT.0.0D0) RETURN
0828.01        IF ((XDIF.LT.ACCTOL).AND.(CNEIG.LT.0.0D0)) RETURN
0829.01        SHIFT=QUAD+SHIFT
0830.01        GO TO 008
0831.01  C
0832.01  007   IF ((MOD(ITEIG,5).EQ.0).AND.(CNEIG.GT.0.0D0))
```

```
0833.01       1 SHIFT=CNEIG+SHIFT
0834.01         IF (NPRI.EQ.1) CALL MATPRO(SHIFT,1)
0835.01  008    CONTINUE
0836.01  C
0837            WRITE (NOUT6,909)
0838     909    FORMAT (' *** CVEIGC, FAILURE TO CONVERGE ***')
0839            IFAIL=1
0840            ITEIG=1001
0841            QUAD=0.0D0
0842            RETURN
0843            END
0844
0845            SUBROUTINE CVVOPT(W1,NLMAX,NL,VOPT,NTBMAX,NT,MOPT,NTMAX,
0846           1             NOMEGA,NLAG,NVAR,ITEIG,W5,W2,
0847           1             W3,NVARMX,NPRI)
0848     C
0849     C        CALCULATES W5=(VOPT)(W1)
0850     C
0851            DOUBLE PRECISION
0852           1 VOPT(NTBMAX),
0853           1 W1(NLMAX),W2(NLMAX),
0854           1 W3(NVARMX),W5(NLMAX),XX
0855            INTEGER
0856           1 MOPT(NTMAX)
0857            IF (ITEIG.NE.1) GO TO 008
0858     C
0859     C        CALCULATE W3 WHEN ITEIG=1 IF NOMEGA=0
0860     C
0861            DO 001 I=1,NVAR
0862.01  001    W3(I)=0.0D0
0863            NLAGA=NLAG+1
0864            DO 003 L=1,NT
0865.01         IF (MOPT(L).GT.NVAR) GO TO 003
0866.01         K=0
0867.01         DO 002 I=1,NL
0868.02         JZF=MINO(NLAGA,I)
0869.02         DO 002 JZ=1,JZF
0870.03         J=I+1-JZ
0871.03         K=K+1
0872.03         IF (MOPT(L).NE.K) GO TO 002
0873.03         W3(K)=VOPT(L)
0874.03  002    CONTINUE
0875.01  003    CONTINUE
0876            IF (NPRI.EQ.1) CALL MATPR1(W3,NVAR,1)
0877            IF (NOMEGA.EQ.0) GO TO 008
0878     C
0879     C        CALCULATE W3 WHEN ITEIG=1 IF NOMEGA>0
0880     C
0881            NLA=NL+1
0882            DO 007 IOMEGA=1,NOMEGA
0883.01         DO 004 J=1,NL
0884.02  004    W2(J)=0.0D0
0885.01         K=NVAR+1
0886.01         KK=NVAR+1
0887.01         DO 006 IZ=1,NL
0888.02         I=NLA-IZ
0889.02         JS=MAXO(I-NLAG,1)
0890.02         DO 005 J=JS,I
0891.03         XX=W2(J)
0892.03         JZ=J-1
0893.03         IF (J.GT.JS) XX=XX+W3(K)-W2(JZ)
0894.03         K=K-1
0895.03  005    W3(K)=W3(K)+XX
0896.02         DO 006 J=JS,I
```

```
0897.03          KK=KK-1
0898.03    006   W2(J)=W3(KK)
0899.01          IF (NPRI.EQ.1) CALL MATPR1(W3,NVAR,1)
0900.01    007   CONTINUE
0901.01    C
0902.01    C        CALCULATE W5
0903.01    C
0904       008   NLAGA=NLAG+1
0905             DO 009 I=1,NL
0906.01    009   W5(I)=0.0D0
0907             K=0
0908             DO 010 I=1,NL
0909.01          JZF=MINO(NLAGA,I)
0910.01          DO 010 JZ=1,JZF
0911.02          J=I+1-JZ
0912.02          K=K+1
0913.02          W5(I)=W5(I)+W3(K)*W1(J)
0914.02          IF (I.NE.J) W5(J)=W5(J)+W3(K)*W1(I)
0915.02    010   CONTINUE
0916             IF (NPRI.EQ.1) CALL MATPR1(W5,NL,1)
0917             RETURN
0918             END
0919
0920             SUBROUTINE CVMODI(PHI,NLMAX,NL,NPAMAX,NPA,THETA,NQAMAX,NQA,
0921            1            C,NTMAX,NT,CWT,CNEW,NOMEGA,NLAG,NVAR,
0922            1            XLHS,NLHS,NVARMX,XRHS,
0923            1            W1,W2,W3,MW1,MW3,NTTMAX,ACCTOL,NPRI,IFAIL)
0924       C
0925       C        MODIFY INPUT FOR LINEAR PROGRAM
0926       C
0927             DOUBLE PRECISION
0928            1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
0929            1 C(NLMAX,NTMAX),CWT(NTMAX),CNEW(NLMAX),
0930            1 XLHS(NLHS,NVARMX),XRHS(NTMAX),W1(NLMAX),W2(NLMAX),
0931            1 W3(NVARMX),ACCTOL,CVC
0932             INTEGER
0933            1 MW1(NTMAX),MW3(NTTMAX)
0934             COMMON
0935            1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
0936             NTA=NT+1
0937       C
0938       C        REMOVE C-VECTORS FOR WHICH CWT=0
0939       C
0940             NEWLOC=NTA
0941             NSHIFT=0
0942             K=0
0943             L=NTA
0944       001   K=K+1
0945             IF (K.EQ.L) GO TO 004
0946             IF (CWT(K).GE.ACCTOL) GO TO 001
0947             IF (NEWLOC.NE.NTA) GO TO 002
0948             NEWLOC=K
0949             GO TO 001
0950       002   L=L-1
0951             IF (K.EQ.L) GO TO 004
0952             IF (CWT(L).LT.ACCTOL) GO TO 002
0953             NSHIFT=NSHIFT+1
0954             MW1(NSHIFT)=K
0955             MW3(NSHIFT)=L
0956             DO 003 I=1,NL
0957.01    003   C(I,K)=C(I,L)
0958             CWT(K)=CWT(L)
0959             XRHS(K)=XRHS(L)
0960             GO TO 001
```

```
0961      C
0962      004   IF (NEWLOC.EQ.NTA) K=K+1
0963            IF (K.GT.NTMAX) GO TO 013
0964            NT=K-1
0965            IF (K.GT.NLHS) GO TO 005
0966            IF (K.EQ.NTA) GO TO 005
0967            NSHIFT=NSHIFT+1
0968            MW1(NSHIFT)=K
0969            MW3(NSHIFT)=NTA
0970      005   IF (NPRI.NE.1) GO TO 006
0971            CALL MATPN0(NSHIFT,1)
0972            IF (NSHIFT.EQ.0) GO TO 006
0973            CALL MATPN1(MW1,NSHIFT,1)
0974            CALL MATPN1(MW3,NSHIFT,1)
0975      C
0976      C        ADD NEW C-VECTOR CNEW
0977      C
0978      006   DO 007 I=1,NL
0979.01   007   C(I,NEWLOC)=CNEW(I)
0980            CWT(NEWLOC)=0.0D0
0981            CALL CVOMEG(CNEW,NLMAX,NL,PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,
0982           1            NOMEGA,NLAG,CVC,W3,NVARMX,NVAR,
0983           1            W2,ACCTOL,NPRI,IFAIL)
0984            IF (IFAIL.NE.0) RETURN
0985            XRHS(NEWLOC)=CVC
0986            IF (NT+1.GT.NLHS) GO TO 012
0987            NTA=NT+1
0988            IF (NSHIFT.EQ.0) GO TO 010
0989            DO 009 I=1,NVAR
0990.01         DO 008 J=1,NSHIFT
0991.02         K=MW1(J)
0992.02         L=MW3(J)
0993.02   008   XLHS(K,I)=XLHS(L,I)
0994.01   009   XLHS(NEWLOC,I)=W3(I)
0995            GO TO 012
0996      010   DO 011 I=1,NVAR
0997.01   011   XLHS(NEWLOC,I)=W3(I)
0998.01   C
0999      012   IF (NPRI.NE.1) RETURN
1000            CALL MATPR2(C,NLMAX,NL,NT,1)
1001            CALL MATPR1(CWT,NT,1)
1002            CALL MATPR2(XLHS,NLHS,NTA,NVAR,1)
1003            CALL MATPR1(XRHS,NT,1)
1004            RETURN
1005      C
1006      013   WRITE (NOUT6,914)
1007      914   FORMAT (' *** CVMODI, NEW NT OUT OF RANGE ***')
1008            IFAIL=1
1009            RETURN
1010            END
1011
1012            SUBROUTINE CVINCR(XLHS,NLHS,NVARMX,NVAR,W1,NLMAX,W3,NT,NPRI)
1013      C
1014      C        INCREASE FIRST DIMENSION OF XLHS
1015      C
1016            DOUBLE PRECISION
1017           1 XLHS(NLHS,NVARMX),W1(NLMAX),W3(NVARMX)
1018      C
1019            NTA=NT+1
1020            JOLD=INTS((INTL(NVAR)*INTL(NT)-1)/INTL(NLHS)+1)
1021            IOLD=INTS(INTL(NVAR)*INTL(NT)-INTL(JOLD-1)*INTL(NLHS)+1)
1022            J=NVAR+1
1023            DO 002 JZ=1,NVAR
1024.01         J=J-1
```

```
1025.01          I=NTA
1026.01          DO 001 IZ=1,NT
1027.02          I=I-1
1028.02          IOLD=IOLD-1
1029.02          IF (IOLD.GE.1) GO TO 001
1030.02          IOLD=NLHS
1031.02          JOLD=JOLD-1
1032.02  001     W1(I)=XLHS(IOLD,JOLD)
1033.01          W1(NTA)=W1(NT)
1034.01          W1(NT)=W3(J)
1035.01          DO 002 I=1,NTA
1036.02  002     XLHS(I,J)=W1(I)
1037             IF (NPRI.EQ.1) CALL MATPR2(XLHS,NLHS,NTA,NVAR,1)
1038             RETURN
1039             END
1040
1041             SUBROUTINE CVLINP(XLHS,NLHS,NVARMX,NVAR,XRHS,NTMAX,NT,
1042            1            MINEQ,XOPT,CWT,VOPT,NTBMAX,MOPT,NL,
1043            1            W4,NTAMAX,MW1,MW2,NTVRMX,MW3,NTTMAX,ITLP,NPRI,IFAIL)
1044     C
1045     C        OPTIMIZE COMBINATION OF C-VECTORS BY LINEAR PROGRAM
1046     C
1047             DOUBLE PRECISION
1048            1 CWT(NTMAX),VOPT(NTBMAX),
1049            1 XLHS(NLHS,NVARMX),XRHS(NTMAX),
1050            1 W4(NTBMAX,NTAMAX),XMIN,X02AAF,XOPT
1051             INTEGER
1052            1 MINEQ(NTMAX),MOPT(NTMAX),MW1(NTMAX),MW2(NTVRMX),
1053            1 MW3(NTTMAX)
1054             COMMON
1055            1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
1056     C
1057             XMIN=X02AAF(XOPT)
1058             NITLP=NT*NL
1059             NTA=NT+1
1060             NTB=NT+2
1061             NTVAR=NT+NVAR
1062             NTT=2*NT
1063             IFAIL=1
1064             CALL H01ADF(XLHS,NLHS,NT,NVAR,MINEQ,XRHS,XMIN,NITLP,
1065            1            NTVAR,NTT,NTA,NTB,MW1,W4,NTBMAX,MW2,
1066            1            MW3,MOPT,VOPT,XOPT,ITLP,IPAR,IFAIL)
1067             IF (IFAIL.NE.0) GO TO 003
1068             IF (IPAR.NE.0) GO TO 005
1069     C
1070             XOPT=DSQRT(-XOPT)
1071             DO 002 I=1,NT
1072.01 002     CWT(I)=DSQRT(DABS(W4(NTA,I)))
1073.01  C
1074             IF (NPRI.NE.1) RETURN
1075             CALL MATPR2(W4,NTBMAX,NTA,NT,1)
1076             CALL MATPR1(CWT,NT,1)
1077             CALL MATPN1(MOPT,NT,1)
1078             CALL MATPR1(VOPT,NT,1)
1079             RETURN
1080     C
1081     003     WRITE (NOUT6,904) IFAIL
1082     904     FORMAT (' *** CVLIN, H01ADF FAILS, IFAIL =',I3,' ***')
1083             RETURN
1084     005     WRITE (NOUT6,906) IPAR
1085     906     FORMAT (' *** CVLIN, H01ADF FAILS, IPAR =',I3,' ***')
1086             IFAIL=1
1087             RETURN
1088             END
```

```
1089
1090          SUBROUTINE CVRESU(PARAM,NPARMX,NPAR,RESID,NLMAX,NL,
1091         1            PHI,NPAMAX,NPA,THETA,NQAMAX,NQA,UVU,
1092         1            C,NTMAX,NT,CWT,ND,NS,NOMEGA,NLAG,
1093         1            W1,W2,W4,NTBMAX,NTAMAX,NPRI)
1094     C
1095     C        CALCULATE FINAL RESULTS
1096     C
1097          DOUBLE PRECISION
1098         1 PARAM(NPARMX),RESID(NLMAX),
1099         1 PHI(NLMAX,NPAMAX),THETA(NLMAX,NQAMAX),
1100         1 UVU(NPARMX,NPARMX),
1101         1 C(NLMAX,NTMAX),CWT(NTMAX),
1102         1 W1(NLMAX),W2(NLMAX),
1103         1 W4(NTBMAX,NTAMAX),TSTAT(34),TDF(34),
1104         1 XX,SUVU,TRCV,STRCV,RBIAS,VARECE,DF,T,ECE,SECE,CILOW,CIUPP
1105          COMMON
1106         1 /NINOUT/ NIN1,NIN2,NIN3,NIN4,NIN5,NOUT6,NOUT7,NOUT8,NOUT9,NOUT10
1107          DATA
1108         1 TSTAT
1109         1/ 12.706D0, 4.303D0, 3.182D0, 2.776D0, 2.571D0,
1110         1   2.447D0, 2.365D0, 2.306D0, 2.262D0, 2.228D0,
1111         1   2.201D0, 2.179D0, 2.160D0, 2.145D0, 2.131D0,
1112         1   2.120D0, 2.110D0, 2.101D0, 2.093D0, 2.086D0,
1113         1   2.080D0, 2.074D0, 2.069D0, 2.064D0, 2.060D0,
1114         1   2.056D0, 2.052D0, 2.048D0, 2.045D0, 2.042D0,
1115         1   2.021D0, 2.000D0, 1.980D0, 1.960D0/,
1116         1 TDF
1117         1/  1.0D0,   2.0D0,   3.0D0,   4.0D0,   5.0D0,
1118         1   6.0D0,   7.0D0,   8.0D0,   9.0D0,  10.0D0,
1119         1  11.0D0,  12.0D0,  13.0D0,  14.0D0,  15.0D0,
1120         1  16.0D0,  17.0D0,  18.0D0,  19.0D0,  20.0D0,
1121         1  21.0D0,  22.0D0,  23.0D0,  24.0D0,  25.0D0,
1122         1  26.0D0,  27.0D0,  28.0D0,  29.0D0,  30.0D0,
1123         1  40.0D0,  60.0D0, 120.0D0,   1.0D10/
1124     C
1125          WRITE (NOUT6,901)
1126     901  FORMAT (///)
1127     C
1128          WRITE (NOUT6,902) ND
1129     902  FORMAT (/' NUMBER OF DATA SETS ADVANCED (ND)
1130         1                          ',7X,I5)
1131          WRITE (NOUT6,903) NS
1132     903  FORMAT (/' NUMBER OF PARAMETER (NS)
1133         1                          ',7X,I5)
1134          WRITE (NOUT6,904) PARAM(NS)
1135     904  FORMAT (/' ESTIMATE OF PARAMETER VALUE (PARAM(NS))
1136         1                          ',1PG16.8/)
1137     C
1138     C        MODEL DEPENDENT RESULTS
1139     C
1140          WRITE (NOUT6,905)
1141     905  FORMAT (/' MODEL DEPENDENT RESULTS',
1142         1        /' ======================')
1143          WRITE (NOUT6,906) UVU(NS,NS)
1144     906  FORMAT (/' VARIANCE OF ESTIMATOR ASSUMING V IS ERROR VARIANCE MATR
1145         1IX (UVU(NS,NS))           ',12X,1PG16.8)
1146          SUVU=DSQRT(UVU(NS,NS))
1147          WRITE (NOUT6,907) SUVU
1148     907  FORMAT (/' S.E. OF ESTIMATOR ASSUMING V IS ERROR VARIANCE MATRIX (
1149         1SQRT(UVU(NS,NS)))         ',1PG16.8/)
1150     C
1151     C        EXPECTED RESULTS FROM CONSERVATIVE ESTIMATOR
1152     C
```

```
1153           WRITE (NOUT6,908)
1154      908  FORMAT (/' EXPECTED RESULTS FROM CONSERVATIVE ESTIMATOR',
1155          1       /' ========================================')
1156           DO 012 K=1,NT
1157.01        DO 009 I=1,NL
1158.02   009  W1(I)=CWT(K)*C(I,K)
1159.01        CALL REGTRA(W1,NL,PHI,NLMAX,NPA,1,1,W2,0,NPRI,IFAIL)
1160.01        CALL REGTRA(W2,NL,THETA,NLMAX,NQA,0,1,W1,0,NPRI,IFAIL)
1161.01        CALL REGTRA(W1,NL,THETA,NLMAX,NQA,0,0,W2,0,NPRI,IFAIL)
1162.01        CALL REGTRA(W2,NL,PHI,NLMAX,NPA,1,0,W1,0,NPRI,IFAIL)
1163.01        DO 011 L=1,NT
1164.02        XX=0.0D0
1165.02        DO 010 I=1,NL
1166.03   010  XX=XX+C(I,L)*W1(I)
1167.02   011  W4(L,K)=XX*CWT(L)
1168.01   012  CONTINUE
1169           IF (NPRI.EQ.1) CALL MATPR2(W4,NTBMAX,NT,NT,1)
1170      C
1171           WRITE (NOUT6,913) NOMEGA
1172      913  FORMAT (/' DIFFERENCE OPERATOR APPLIED TO CLASS OF VARIANCE MATRIC
1173          1ES (NOMEGA)              ',19X,I5)
1174           WRITE (NOUT6,914) NLAG
1175      914  FORMAT (/' LAG BEYOND WHICH VARIANCE TERMS ARE ZERO (NLAG)
1176          1                         ',19X,I5)
1177      C
1178           TRCV=0.0D0
1179           DO 015 K=1,NT
1180.01   015  TRCV=TRCV+W4(K,K)
1181           WRITE (NOUT6,916) TRCV
1182      916  FORMAT (/' EXPECTED VALUE OF CONSERVATIVE ESTIMATOR OF VARIANCE AS
1183          1SUMING V (TR(CV))        ',12X,1PG16.8)
1184           STRCV=DSQRT(TRCV)
1185           WRITE (NOUT6,917) STRCV
1186      917  FORMAT (/' EXPECTED VALUE OF CONSERVATIVE ESTIMATOR OF S.E. ASSUMI
1187          1NG V (SQRT(TR(CV)))      ',1PG16.8)
1188           RBIAS=STRCV/SUVU-1.0D0
1189           WRITE (NOUT6,918) RBIAS
1190      918  FORMAT (/' EXP. REL. BIAS OF CONSERVATIVE EST. OF S.E. (SQRT(TR(CV
1191          1))/SQRT(UVU(NS,NS))-1)  ',24X,1PG16.8/)
1192      C
1193           VARECE=0.0D0
1194           DO 019 K=1,NT
1195.01        DO 019 L=1,NT
1196.02   019  VARECE=VARECE+W4(K,L)*W4(L,K)
1197           VARECE=2.0D0*VARECE
1198           WRITE (NOUT6,920) VARECE
1199      920  FORMAT (/' VARIANCE OF CONSERVATIVE ESTIMATOR OF VARIANCE (2*TR(CV
1200          1CV))                     ',24X,1PG16.8)
1201           DF=2.0D0*TRCV*TRCV/VARECE
1202           WRITE (NOUT6,921) DF
1203      921  FORMAT (/' APPROX. D.F. OF CONSERVATIVE ESTIMATOR OF VARIANCE (TR(
1204          1CV)**2/TR(CVCV))         ',24X,1PG16.8)
1205      C
1206           IF (DF.LT.1.0D0) DF=1.0D0
1207           DO 022 I=1,33
1208.01        IF (DF.LT.TDF(I)) GO TO 023
1209.01   022  CONTINUE
1210      023  IZ=I-1
1211           IF (IZ.LT.1) IZ=1
1212           T=TSTAT(I)+(TSTAT(IZ)-TSTAT(I))*(1.0D0/DF-1.0D0/TDF(I))
1213          1                              /(1.0D0/TDF(IZ)-1.0D0/TDF(I))
1214           WRITE (NOUT6,924) T
1215      924  FORMAT (/' APPROXIMATE 95% T-STATISTIC
1216          1                         ',24X,1PG16.8/)
```

```
1217     C
1218     C        ACTUAL RESULTS FROM CONSERVATIVE ESTIMATOR
1219     C
1220           WRITE (NOUT6,925)
1221     925   FORMAT (/' ACTUAL RESULTS FROM CONSERVATIVE ESTIMATOR',
1222          1        /' ========================================')
1223           ECE=0.0D0
1224           DO 027 K=1,NT
1225.01        XX=0.0D0
1226.01        DO 026 I=1,NL
1227.02  026   XX=XX+RESID(I)*C(I,K)
1228.01        XX=XX*CWT(K)
1229.01  027   ECE=ECE+XX*XX
1230           WRITE (NOUT6,928) ECE
1231     928   FORMAT (/' ESTIMATE OF VARIANCE USING CONSERVATIVE ESTIMATOR (ECE)
1232          1                          ',12X,1PG16.8)
1233           SECE=DSQRT(ECE)
1234           WRITE (NOUT6,929) SECE
1235     929   FORMAT (/' ESTIMATE OF S.E. USING CONSERVATIVE ESTIMATOR (SQRT(ECE
1236          1))                        ',1PG16.8/82X,'==========='/)
1237     C
1238           CILOW=PARAM(NS)-T*SECE
1239           WRITE (NOUT6,930) CILOW
1240     930   FORMAT (/' APPROXIMATE 95% LOWER CONFIDENCE BOUND FOR PARAMETER VA
1241          1LUE                       ',1PG16.8)
1242           CIUPP=PARAM(NS)+T*SECE
1243           WRITE (NOUT6,931) CIUPP
1244     931   FORMAT (/' APPROXIMATE 95% UPPER CONFIDENCE BOUND FOR PARAMETER VA
1245          1LUE                       ',1PG16.8)
1246     C
1247     C        DUMP OF FINAL C-VECTOR TO CHANNEL NOUT10
1248     C
1249           WRITE (NOUT10,932) NT
1250     932   FORMAT (I5)
1251           DO 033 J=1,NT
1252.01   033  WRITE (NOUT10,934) (C(I,J),I=1,NL)
1253     934   FORMAT (1P5D24.16)
1254           RETURN
1255           END
```

Appendix D

Example of output from CEVOPE (see [9.3.4] and [9.4.8]).

```
INPUT (10I7):
     ND  NVARS     NS  NSEED     NT NUMEGA   NLAG  NITER   NCPU   NPRI
      8      1      2    214      1      1      4     50   1000      0

PARAMETER ESTIMATES
 0.5274     -0.2973

STANDARD ERRORS
 0.6074      0.1831

PARAMETER ESTIMATOR VARIANCES ASSUMING V
 0.3689     -0.1006
-0.1006      3.3535D-02

LOWER BOUNDS ON VARIANCES
 0.3689     -0.1006
-0.1006      3.3535D-02

EFFICIENCY OF ESTIMATORS
  1.000      1.000

                *** O U T P U T   F R O M   C E V O P E ***

        MODEL DEPENDENT ESTIMATE OF THE VARIANCE OF THE ESTIMATOR
     OF PARAMETER NUMBER  2  IS  3.3535D-02  AND THE S.E. IS  0.1831

                SEARCH FOR RUBUST ESTIMATOR WITH MINIMUM BIAS (NSEED=  214)

                    ADD NEW C-VECTOR                FIND BEST CUMBINATION OF C-VECTORS
                    ----------------                ----------------------------------
ITER  ICPUT  NT  ITEIG  ICPU  IDIS   QUAD      ITLP  ICPU  IDIS     XOPT       REL BIAS
****  *****  **  *****  ****  ****  ********  ****  ****  ****  ----------  ----------
   0      2   1              2     1                3     0     0     1.041       4.685
   1      3   2     15       0     0  -1.D+01       3     0     0    0.4758       1.598
   2      3   3     16       0     0  -4.D-01       3     0     0    0.4562       1.491
   3      4   4     36       1     0  -6.D-01       4     0     0    0.4237       1.314
   4      5   5     26       1     0  -1.D-01       3     0     0    0.4143       1.263
   5      6   4     53       1     0  -4.D-02       3     0     0    0.3997       1.182
   6      8   4     67       2     0  -6.D-02       4     0     0    0.3973       1.170
   7      9   4     51       1     0  -2.D-02       4     0     0    0.3956       1.160
   8     10   4     58       1     0  -9.D-03       5     0     0    0.3945       1.154
   9     12   4     75       2     0  -4.D-03       4     0     0    0.3938       1.150
  10     14   4     72       2     0  -2.D-03       4     0     0    0.3933       1.148
  11     17   4    128       3     0  -1.D-03       4     0     0    0.3929       1.146
  12     19   4     92       2     0  -4.D-04       4     0     0    0.3929       1.146
  13     21   4     88       2     0  -3.D-04       4     0     0    0.3928       1.145
  14     25   4    130       3     0  -1.D-04       4     0     0    0.3928       1.145
  15     27   4     94       2     0  -1.D-04       4     0     0    0.3928       1.145
  16     29   4     93       2     0  -5.D-05       4     0     0    0.3928       1.145
  17     31   4     91       2     0  -9.D-05       4     0     0    0.3928       1.145
  18     33   4     92       2     0  -3.D-05       5     0     0    0.3928       1.145

ITERATIONS CEASED, UPTIMUM REACHED

NUMBER OF DATA SETS ADVANCED (ND)                                          8

NUMBER OF PARAMETER (NS)                                                   2

ESTIMATE OF PARAMETER VALUE (PARAM(NS))                          -0.29730000

MODEL DEPENDENT RESULTS
***********************

VARIANCE OF ESTIMATOR ASSUMING V IS ERROR VARIANCE MATRIX (UVU(NS,NS))                3.353504790-02

S.E. OF ESTIMATOR ASSUMING V IS ERROR VARIANCE MATRIX (SQRT(UVU(NS,NS)))    0.18312577

EXPECTED RESULTS FROM CONSERVATIVE ESTIMATOR
********************************************

DIFFERENCE OPERATOR APPLIED TO CLASS OF VARIANCE MATRICES (NOMEGA)                             1

LAG BEYOND WHICH VARIANCE TERMS ARE ZERO (NLAG)                                               4

EXPECTED VALUE OF CONSERVATIVE ESTIMATOR OF VARIANCE ASSUMING V (TR(CV))              0.15426567

EXPECTED VALUE OF CONSERVATIVE ESTIMATOR OF S.E. ASSUMING V (SQRT(TR(CV)))   0.39276669

EXP. REL. BIAS OF CONSERVATIVE EST. OF S.E. (SQRT(TR(CV))/SQRT(UVU(NS,NS))-1)                  1.1447920

VARIANCE OF CONSERVATIVE ESTIMATOR OF VARIANCE (2*TR(CVCV))                                    4.759187990-02

APPROX. D.F. OF CONSERVATIVE ESTIMATOR OF VARIANCE (TR(CV)**2/TR(CVCV))                        1.0000823

APPROXIMATE 95% T-STATISTIC                                                                    12.704617

ACTUAL RESULTS FROM CONSERVATIVE ESTIMATOR
******************************************

ESTIMATE OF VARIANCE USING CONSERVATIVE ESTIMATOR (ECE)                0.29039159

ESTIMATE OF S.E. USING CONSERVATIVE ESTIMATOR (SQRT(ECE))    0.53887994
                                                             **********

APPROXIMATE 95% LOWER CONFIDENCE BOUND FOR PARAMETER VALUE   -7.1435635

APPROXIMATE 95% UPPER CONFIDENCE BOUND FOR PARAMETER VALUE    6.5489635
```

REFERENCES

AITKEN, A.C. (1934), On least-squares and linear combinations of observations, Proceedings of the Royal Society of Edinburgh, 55, 42-48.

AITKEN, A.C. (1937), The evaluation of the latest roots and vectors of a matrix, Proceedings of the Royal Society of Edinburgh, 57, 269-304.

AKAIKE, H. (1973), Information theory and an extension of the maximum likelihood principle, 2nd International Symposium on Information Theory, Budapest: Publishing House of the Hungarian Academy of Sciences, 267-281.

ANDERSON, R.L. (1954), The problem of autocorrelation in regression analysis, Journal of the American Statistical Association, 49, 113-129.

ANDERSON, T.W. (1971), The Statistical Analysis of Time Series, New York: Wiley.

ANSLEY, C.F. (1979), An algorithm for the exact likelihood of a mixed autoregressive-moving average process, Biometrika, 66, 59-65.

ANSLEY, C.F. and NEWBOLD, P. (1979), On the finite sample distribution of residual autocorrelations in autoregressive-moving average models, Biometrika, 66, 547-553.

ATKINSON, A.C. (1969), The use of residuals as a concomitant variable, Biometrika, 56, 33-41.

BARD, Y. (1974), Nonlinear Parameter Estimation, New York: Academic Press.

BARTLETT, M.S. (1946), On the theoretical specification and sampling properties of autocorrelated time-series, Supplement to Journal of Royal Statistical Society, 8, 27-41.

BARTLETT, M.S. (1978), An Introduction to Stochastic Processes, with special reference to Methods and Applications, (3rd edition), Cambridge: University Press.

BAUERMEISTER, A., DALE, J.E., WILLIAMS, E.J. and SCOBIE, J. (1980), Movement of ^{14}C- and ^{11}C-labelled assimilate in wheat leaves: the effect of IAA, Journal of Experimental Botany, 31, 1199-1209.

BLIGHT, B.J.N. and OTT, L. (1975), A Bayesian approach to model inadequacy for polynomial regression, Biometrika, 62, 79-88.

BLOOMFIELD, P. and WATSON, G.S. (1975), The inefficiency of least squares, Biometrika, 62, 121-128.

BOX, G.E.P. and JENKINS, G.M. (1976), Time Series Analysis; Forcasting and Control (Revised Edition), San Francisco: Holden-Day.

BOX, G.E.P. and PIERCE, D.A. (1970), Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, Journal of the American Statistical Association, 65, 1509-1526.

BRAND, L. (1966), Differential and Difference Equations, New York: Wiley.

BRUCE, J.M. (1980), Modelling the climatic energy demand of suckler cows, Animal Production, 30, 449-450.

BURNETT, G.A. and BRUCE, J.M. (1978), Thermal simulation of a suckler cow, Farm Building Progress, 54, 11-13.

CAMPBELL, M.J. and WALKER, A.M. (1977), A survey of statistical work on the Mackenzie river series of annual Canadian lynx trappings for the years 1821-1934 and a new analysis, Journal of the Royal Statistical Society, Series A, 140, 411-431.

CHATTERJEE, S. and PRICE, B. (1977), Regression Analysis by Example, New York: Wiley.

CHIANG, C.L. (1956), On regular best asymptotically normal estimates, Annals of Mathematical Statistics, 27, 336-351.

CHIANG, C.L. (1980), An Introduction to Stochastic Processes and Their Applications, New York: Krieger.

CHIPMAN, J.S. (1979), Efficiency of least-squares estimation of linear trend when residuals are autocorrelated, Econometrica, 47, 115-128.

CLIFF, A.D. and ORD, J.K. (1981), Spatial Processes; Models and Applications, London: Pion.

COBBY, J.M. and LE DU, Y.L.P. (1978), On fitting curves to lactation data, Animal Production, 26, 127-133.

COCHRANE, D. and ORCUTT, G.H. (1949), Application of least squares regression to relationships containing autocorrelated error terms, Journal of the American Statistical Association, 44, 32-61.

COLQUHOUN, D. (1978), Critical Analysis of numerical biological data, The 6th International CODATA Conference, CODATA Bulletin, 29, 26-33.

COLQUHOUN, D. and HAWKES, A.G. (1977), Relaxation and fluctuations of membrane currents that flow through drug-operated channels, Proceedings of the Royal Society of London B, 199, 231-262.

COOK, D.G. and POCOCK, S.J. (1983), Multiple regression in geographical mortality studies, with allowance for spatially correlated errors, Biometrics, 39, 361-371.

COOPER, D.M. and THOMPSON, R. (1977), A note on the estimation of the parameters of the autoregressive-moving average process, Biometrika, 64, 625-628.

COX, D.R. and MILLER, H.D. (1965), The Theory of Stochastic Processes, London: Methuen.

CRAMÉR, H. (1946), Mathematical Methods of Statistics, Princeton: University Press.

CROWDER, M.J. (1976), Maximum likelihood estimation for dependent observations, Journal of the Royal Statistical Society, Series B, 38, 45-53.

CROWDER, M.J. and TREDGER, J.A. (1981), The use of exponentially damped polynomials for biological recovery data, Applied Statistics, 30, 147-152.

DHANOA, M.S. and LE DU, Y.L.P. (1982), A partial adjustment model to describe the lactation curve of a dairy cow, Animal Production, 34, 243-247.

DUNCAN, D.B. and JONES, R.H. (1966), Multiple regression with stationary errors, Journal of the American Statistical Association, 61, 917-928.

DURBIN, J. (1960), Estimation of parameters in time-series regression models, Journal of the Royal Statistical Society, Series B, 22, 139-153.

DURBIN, J. and WATSON, G.S. (1950), Testing for serial correlation in least squares regression I, Biometrika, 37, 409-428.

DURBIN, J. and WATSON, G.S. (1951), Testing for serial correlation in least squares regression II, Biometrika, 38, 159-178.

DURBIN, J. and WATSON, G.S. (1971), Testing for serial correlation in least-squares regression III, Biometrika, 58, 1-19.

EFRON, B. (1982), The Jackknife, The Bootstrap and Other Resampling Plans, Philadelphia: Society for Industrial and Applied Mathematics.

ENGLE, R.F. (1974), Specification of the disturbance for efficient estimation, Econometrica, 42, 135-146.

FINNEY, D.J. (1958), The efficiencies of alternative estimators for an asymptotic regression equation, Biometrika, 45, 370-388.

FINNEY, D.J. (1978), Growth curves: their nature, uses, and estimation, in Patterns of Growth & Development in Cattle, (eds. H. de Boer and J. Martin) The Hague: Martinus Hijhoff, 658-672.

FISK, P.R. (1967), Models of the second kind in regression analysis, Journal of the Royal Statistical Society, Series B, 29, 266-281.

GALLANT, A.R., and GOEBEL, J.J. (1976), Nonlinear regression with autocorrelated errors, Journal of the American Statistical Association, 71, 961-967.

GARVIN, N.W. (1960), Introduction to linear programming, New York: McGraw-Hill.

GASTER, M. and ROBERTS, J.B. (1977), The spectral analysis of randomly sampled records by a direct transform, Proceedings of the Royal Society of London, Series A, 354, 27-58.

GILL, P.E. and MURRAY, W. (1972), Quasi-Newton methods for unconstrained optimization, Journal of the Institute of Mathematics and its Applications, 9, 91-108.

GLASBEY, C.A. (1979), Correlated residuals in non-linear regression applied to growth data, Applied Statistics, 28, 251-259.

GLASBEY, C.A. (1980), Nonlinear regression with autoregressive time series errors, Biometrics, 36, 135-140.

GLASBEY, C.A. (1982), A generalization of partial autocorrelations useful in identifying ARMA models, Technometrics, 24, 223-228.

GLASBEY, C.A. (1984), CEVOPE - a program for finding conservative estimates of the variances of regression parameter estimators when errors are serially correlated, Compstat 1984, Vienna: Physica-Verlag, 55-60.

GRIZZLE, J.E. and ALLEN, D.M. (1969), Analysis of growth and dose response curves, Biometrics, 25, 357-381.

HARVEY, A.C. and PHILLIPS, G.D.A. (1979), Maximum likelihood estimation of regression models with autoregressive-moving average disturbances, Biometrika, 66, 49-58.

HOEL, P.G., PORT, S.C. and STONE, C.J. (1972), Introduction to Stochastic Processes, Boston: Houghton Mifflin.

HUNTER, E.A., GLASBEY, C.A. and NAYLOR, R.E.L. (1984), The analysis of data from germination tests, Journal of Agricultural Science, 102, 207-213.

JENNRICH, R.I. (1969), Asymptotic properties of non-linear least squares estimators, Annals of Mathematical Statistics, 40, 633-643.

JOHNSTON, J. (1972), Econometric Methods (2nd edition), New York: McGraw-Hill.

JONES, R.H. (1981), Fitting a continuous time autoregression to discrete data, in Applied Time Series Analysis II, (ed. D.F. FINDLEY) New York: Academic Press.

JUDGE, G.G., GRIFFITHS, W.E., HILL, R.C., and LEE, T-C. (1980), The Theory and Practice of Econometrics, New York: Wiley.

KODELL, R.L. and MATIS, J.H. (1976), Estimating the rate constants in a two-compartment stochastic model, Biometrics, 32, 377-390.

KOWALSKI, C.J. and GUIRE, K.E. (1974), Longitudinal Data Analysis, Growth, 38, 131-169.

KRÄMER, W. (1980), Finite sample efficiency of ordinary least squares in the linear regression model with autocorrelated errors, Journal of the American Statistical Association, 75, 1005-1009.

LEHOCZKY, J.P. and GAVER, D.P. (1977), A diffusion-approximation analysis of a general n-compartment system, Mathematical Biosciences, 36, 127-148.

LJUNG, G.M. (1982), The likelihood function for a stationary Gaussian autoregressive-moving average process with missing observations, Biometrika, 69, 265-268.

LJUNG, G.M. and BOX, G.E.P. (1978), On a measure of lack of fit in time series models, Biometrika, 65, 297-303.

MALINVAUD, E. (1980), Statistical Methods of Econometrics, (3rd edition), Amsterdam: North Holland.

MANN, H. and WALD, A. (1943), On the statistical treatment of linear stochastic difference equations, Econometrica, 11, 173-220.

MATIS, J.H. and HARTLEY, H.O. (1971), Stochastic compartmental analysis: Model and least squares estimation from time series data, Biometrics, 27, 77-102.

MATIS, J.H. and WEHRLY, T.E. (1979), Stochastic models of compartmental systems, Biometrics, 35, 199-220.

MITCHELL, A.F.S. (1968), Exponential regression with correlated observations, Biometrika, 55, 149-162

NEILSON, D.R., WHITTEMORE, C.T., LEWIS, M., ALLISTON, J.C., ROBERTS, D.J., HODGSON-JONES, L.S., MILLS, J., PARKINSON, H. and PRESCOTT, J.H.D. (1983), Production characteristics of high yielding cows, Animal Production, 36, 321-335.

NEWBOLD, P. (1981), Some recent developments in time series analysis, International Statistical Review, 49, 53-66.

NICHOLS, M.A. and HEYDECKER, W. (1968), Two approaches to the study of germination data, Proceedings of International Seed Testing Association, 33, 531-540.

NUMERICAL ALGORITHMS GROUP (1983), Library Manual Mk 10, NAG Central Office, 256 Banbury Road, Oxford, OX2 7DE.

PANDIT, S.M. and WU, S.M. (1975), Unique estimates of the parameters of a continuous stationary stochastic process, Biometrika, 62, 497-501.

PARKS, J.R. (1975a), The animal growth phase plane, Journal of Theoretical Biology, 55, 371-380.

PARKS, J.R. (1975b), A theory of animal weight response to controlled feeding, Journal of Theoretical Biology, 55, 381-391.

PATTERSON, H.D. and SILVEY, V. (1980), Statutory and recommended list trials of crop varieties in the United Kingdom, Journal of the Royal Statistical Society, Series A, 143, 219-252.

PHADKE, M.S. and WU, S.M. (1974), Modeling of continuous stochastic processes from discrete observations with application to sunspots data, Journal of the American Statistical Association, 69, 325-329.

PIERCE, D.A. (1971a), Distribution of residual autocorrelations in the regression model with autoregressive-moving average errors, Journal of the Royal Statistical Society, Series B, 33, 140-146.

PIERCE, D.A. (1971b), Least squares estimation in the regression model with autoregressive-moving average errors, Biometrika, 58, 299-312.

PIERCE, D.A. (1972), Least squares estimation in dynamic-disturbance time series models, Biometrika, 59, 73-78.

PRIESTLEY, M.B. (1981a), Spectral Analysis and Time Series, Volume 1: Univariate Series, London: Academic Press.

PRIESTLEY, M.B. (1981b), Spectral Analysis and Time Series, Volume 2: Multivariate Series, Prediction and Control, London: Academic Press.

PURDUE, P. (1979), Stochastic compartmental models: a review of the mathematical theory with ecological applications, in Compartmental Analysis of Ecosystem Models (eds. J.H. Matis, B.C. Patten and G.C. White), 223-260, Maryland: International Co-operative Publishing House.

QUENOUILLE, M.H. (1947), A large-sample test for the goodness of fit of autoregressive schemes, Journal of the Royal Statistical Society, 110, 123-129.

QUENOUILLE, M.H. (1958), Discrete autoregressive schemes with varying time-intervals, Metrika, 1, 21-27.

RAO, P and GRILICHES, Z. (1969), Small-sample properties of several two-stage regression methods in the context of auto-correlated errors, Journal of the American Statistical Association, 64, 253-272.

RIPLEY, B.D. (1981), Spatial Statistics, New York: Wiley.

ROBINSON, P.M. (1977), Estimation of a time series model from unequally spaced data, Stochastic Processes and their Applications, 6, 9-24.

ROBINSON, P.M. (1978), Alternative models for stationary stochastic processes, Stochastic Processes and their Applications, 8, 141-152.

ROSS, G.J.S. (1978), Exact and approximate confidence regions for functions of parameters in non-linear models, Compstat 1978, Vienna: Physica-Verlag, 110-116.

SANDLAND, R.L. and McGILCHRIST, C.A. (1979), Stochastic growth curve analysis, Biometrics, 35, 255-271.

SATTERTHWAITE, F.E. (1946), An approximate distribution of estimates of variance components, Biometrics, 2, 110-114.

SPITZER, J.J. (1979), Small-sample properties of nonlinear least squares and maximum likelihood estimators in the context of autocorrelated errors, Journal of the American Statistical Association, 74, 41-47.

THIESSEN, R.B., HNIZDO, E., MAXWELL, D.A.G., GIBSON D. and TAYLOR, St.C.S. (1984), Multibreed comparisons of British Cattle. Variation in body weight, growth rate and food intake, Animal Production, 38, 323-340.

THORNLEY, J.H.M. (1976), Mathematical Models in Plant Physiology: a Quantitative Approach to Problems in Plant and Crop Physiology, London: Academic Press.

WATSON, G.S. (1955), Serial correlation in regression analysis I, Biometrika, 42, 327-341.

WATSON, G.S. and HANNAN, E.J. (1956), Serial correlation in regression analysis II, Biometrika, 43, 436-448.

WEBER, J.E. and MONARCHI, D.E. (1982), Performance of the Durbin-Watson test and WLS estimation when the disturbance term includes serial dependence in addition to first-order autocorrelation, Journal of the American Statistical Association, 77, 117-128.

WEDDERBURN, R.W.M. (1974), Quasi-likelihood functions, generalized linear models and the Gauss-Newton method, Biometrika, 61, 439-447.

WHITE, J.S. (1958), The limiting distribution of the serial correlation coefficient in the explosive case, Annals of Mathematical Statistics, 29, 1188-1197.

WHITTLE, P. (1961), Gaussian estimation in stationary time series, Bulletin of the International Statistical Institute, 33, 105-129.

WILKINSON, G.N., ECKERT, S.R., HANCOCK, T.W. and MAYO, O. (1983), Nearest neighbour (NN) analysis of field experiments, Journal of the Royal Statistical Society, Series B, 45, 151-211.

WILKINSON, J.H. (1954), The calculation of the latent roots and vectors of matrices on the pilot model of the A.C.E., Proceedings of the Cambridge Philosophical Society, 50, 536-566.

WILKINSON, J.H. and REINSCH, C. (1971), Handbook for Automatic Computation. Volume II Linear Algebra, Berlin: Springer-Verlag.

WOOD, P.D.P. (1967), Algebraic model of the lactation curve in cattle, Nature, 216, 164-165.

WOODWARD, W.A. and GRAY, H.L. (1983), A comparison of autoregressive and harmonic component models for the lynx data, Journal of the Royal Statistical Society, Series A, 146, 71-73.

ZYSKIND, G. (1967), On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models, Annals of Mathematical Statistics, 38, 1092-1109.