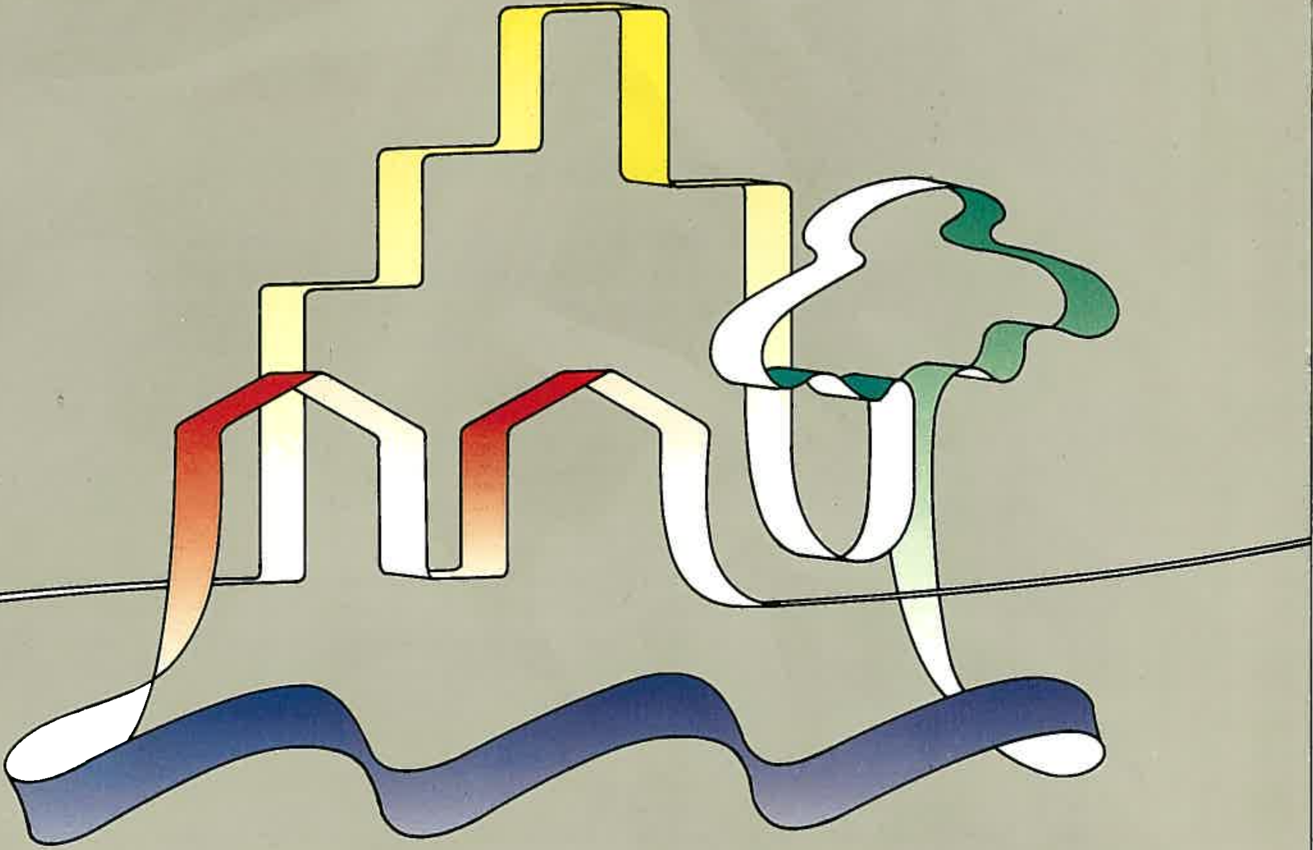
Urban Water Research Association of Australia

Forecasting Water Demand Using Weather Data



Research Report No. 30

URBAN WATER RESEARCH ASSOCIATION OF AUSTRALIA

The Association was formed in 1986 following initiatives by the Australian Water Research Advisory Council and the Major Urban Water Authorities of Australia. The Association's primary role is to foster and promote a comprehensive, co-ordinated and cost-effective approach to urban water research within Australia, for both metropolitan and non-metropolitan areas.

The Association invites proposals for research work through its member authorities and allocates funding to approved projects on an annual basis. The actual research is undertaken by water authorities, research organisations, universities, consultants and government agencies.

The UWRAA Research Report series presents information resulting from research projects supported by the Association and is published as a record of the work undertaken and as a means of disseminating the research findings. The Association also encourages the presentation of findings by the researchers in professional journals and at conferences. The Association's reports are indexed on STREAMLINE, the national water data base.

For further details contact:
Executive Officer
Urban Water Research Association of Australia
C/- Board of Works
GPO Box 4342 Telephone: (03) 615 5816
Melbourne 3001 Telex: AA34220
AUSTRALIA Fax: (03) 615 4408

Urban Water Research Association of Australia

Forecasting Water Demand Using Weather Data

M.N. Viswanathan

Research Report No.30
October 1991

Published for the Urban Water Research Association of Australia by the Melbourne and Metropolitan Board of Works.

ISBN 1 875298 30 4

Cover design by Gregory R Smith

FOREWORD

This report is based on UWRAA Research Project No WS-9: '*Forecasting water consumption*' which was undertaken during the period March 1987 to December 1989. Organisational responsibility for the project was as follows:

Sponsoring Authority : Hunter Water Board, NSW

Research Agency : Hunter Water Board

Project Officer : Mr R Wilson
Hunter Water Board

Principal Researcher : Dr M N Viswanathan
Hunter Water Board.

The project was funded by the Urban Water Research Association of Australia and by the Hunter Water Board.

SYNOPSIS

An outline is given of the purpose and nature of forecasting and of the various methods and models used for forecasting. A model is developed for forecasting water demand for an urban area using weather data (daily maximum temperature and daily rainfall) and base usage as inputs.

The model forecasts water demand for a given day given the anticipated weather parameters for that day and the weather parameters and water consumption for the preceding day. The model is linear and the development is based on the method of ‘Recursive Least Squares’. Coefficients of the model are assumed to be time dependent.

The model was used to forecast water demand for Newcastle with a lead time of one day. The model forecasted water demand within an accuracy of plus minus 5 percent.

Contents

Chapter 1

INTRODUCTION

1.1 General

Webster's dictionary defines forecasting as an activity " to calculate or predict some future event or condition, usually as a result of rational study or analysis of pertinent data". Forecasting is an important part of decision making, and many of our decisions are based on predictions of future unknown events. The ability to form good forecasts has been highly valued throughout history (Abraham and Ledolter 1983).

We all make forecasts, although we may not recognise them as forecasts. For example, a person waiting for a bus or parents expecting a telephone call from their children may not consider themselves forecasters. However, from past experience and from reading the bus schedule, the person waiting for the bus expects it to arrive at a certain time or within a certain time interval. Parents who have usually received calls from their children every weekend expect to receive one during the coming weekend also.

These people form expectations, and make forecasts. So does a bank manager who predicts the cash flow for the next quarter, or a control engineer who adjusts certain input variables to maintain the future value of some output variable as close as possible to a specified target. All make statements about future events, patterning the forecasts closely on previous occurrences and assuming that the future will be similar to the past.

Since the future events involve uncertainity, the forecasts are usually not perfect. The objective of forecasting is to reduce the forecast error: to produce forecasts that are seldom incorrect and that have small forecast errors. In business, industry and government, policymakers must anticipate the future behaviour of many critical variables before they make decisions. Their decisions depend on forecasts, and they expect these forecasts to be accurate; a forecast system is needed to make such predictions. Each situation that requires a forecast comes with its own unique set of problems, and the solutions to one are by no means the solutions in another situation. However, certain general principles are common to most forecasting problems and should be incorporated into any forcasting system.

1.2 Classification of Forecast Methods

Forecast methods may be broadly classified into:

i. Qualitative and

ii. Quantitative techniques.

Qualitative, or "subjective" methods are intuitive, largely educated guesses that may or may not depend on past data. Usually these forecasts cannot be reproduced by someone else, since the forecaster does not specify explicitly how the available information is incorporated into the forecast. Even though subjective forecasting is a nonrigorous approach, it may be quite appropriate and the only reasonable method in certain situations.

Forecasts that are based on mathematical or statistical models are called "quantitative". Once the underlying model or technique has been chosen, the corresponding forecasts are determined automatically; they are fully reproducible by any forecaster. Quantitative methods or models can be further classified as deterministic or probabilistic (also known as stochastic or statistical).

In deterministic models, the relationship between the variable of interest, Y, and the explanatory or predictor variables $X_1, ..., X_p$ is determined exactly:

$$Y = f(X_1 ..... X_p; B_1 ..... B_m)$$

The function f and the coefficients $B_1 ..... B_m$ are known with certainty. The traditional "laws" in the physical sciences are examples of such determininstic relationships.

In the social sciences, however, the relationships are usually stochastic. Measurement errors and variablity from other uncontrolled variables introduce random (stochastic) components. This leads to probabilistic or stochastic models of the form:

$$Y = f(X_1 ..... X_p; B_1 ..... B_m) + \text{noise}$$

where the noise or error component is a realization from a certain probability distribution.

Frequently, the function form f and the coefficients are not known and have to be determined from the past data. Usually the data occur in time ordered sequences referred to as "time series".

All quantitative forecasting methods make use of the following basic strategy. Past data are analysed in order to identify a pattern that can be used to describe them. Then this pattern is extrapolated, or extended , into the future in order to make forecasts. This strategy rests on the assumption that the pattern that has been identified will continue into the future. A forecasting technique cannot be expected to give good predictions unless this assumption is valid. This assumption is more likely to be valid in the short term than in the long term and so it is not surprising that, in general, short term forecasts are more accurate than long term forecasts.

In general, a quantitative forecast system consists of two major components, as illustrated in figure 1.1 . At the first stage, the model building stage, a forecasting model is constructed from pertinent data and available theory. At the second stage, the forecasting stage, the final model is used to obtain the forecasts. Since these forecasts depend on the specified model, one has to make sure that the model and its parameters stay constant during the forecast period. The stability of the forecast model can be assessed by checking the forecasts against the new observations. Forecast errors can be calculated, and possible changes in the model can be detected.

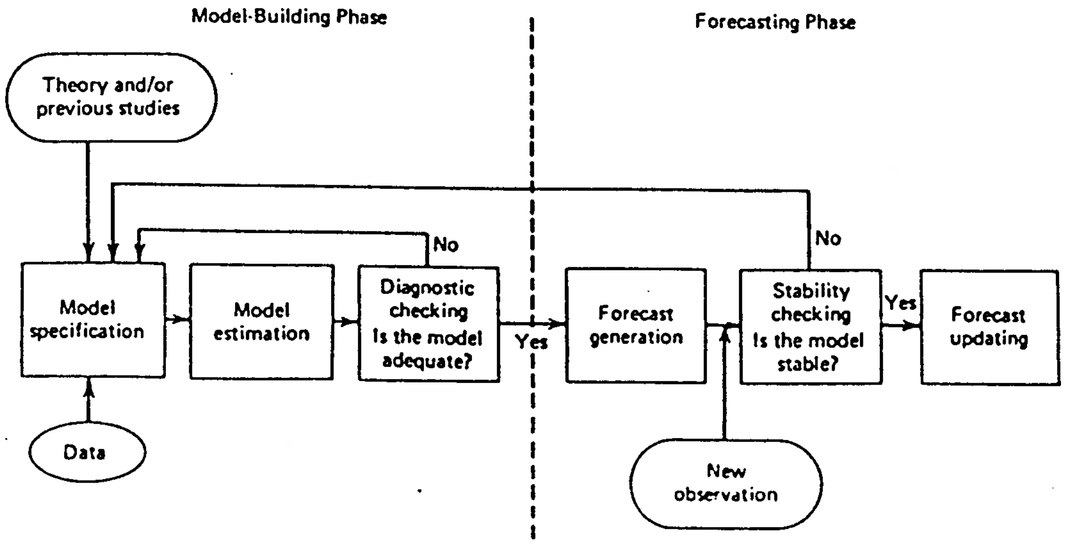


Figure 1.1: Conceptual framework of a forecasting system. (Abraham and Ledolter 1983)

1.3 Forecasting Methods

The time series models generate predictions that are based solely on the historical pattern of the variable to be forecast. Thus, any decisions management might implement will not alter the predictions generated by a time series model. Time series forecasting models are, therfore, most useful when conditions are expected to remain the same. They are not very useful in forecasting the impact of changes in management policies (O'Donovan 1983).

The most widely used types of time series models are outlined below:

i. Simple exponential smoothing:

A weighted average of the recently observed values of the variable under study is used as a forecast. An exponentially decreasing set of weights is assigned to these values so that the more recent values receive more weight than older values.

ii. Holt-Winters:

This is a more sophisticated version of exponential smoothing in which allowance is made for trend and seasonal patterns in the data.

iii. Decomposition:

A time series is regarded as having a number of components: trend, seasonality, cyclicity, and randomness. The first three of these components are estimated and used to forecast future values.

iv. Box-Jenkins:

A class of models known as Autoregressive Moving Average (ARMA) is studied. From the study of the data, an appropriate model is selected from this class and used to make forecasts.

v. Bayesian forecasting:

This method allows one to specify a range of models for the data, rather than a single model, with associated probabilities which are updated as more data become available. Estimates of the model parameters are also updated, using a technique known as Kalman Filtering. The method potentially includes both exponential smoothing and the ARMA model, though in practice, it is limited to a selection of small number of models which are in some ways equivalent to ARMA models and which have been found to cover many of the practically occurring situations.

Besides time series models, the other main type of quantitative type of forecasting method is causal models. These models involve the identification of other variables related to the variable to be predicted. Then a model is developed that describes the relationship between these variables and the variable to be predicted. This model is then used to forecast the variable of interest.

The main types of causal models are outlined below:

i. Multiple regression:

The relationship between the dependent and independent variables is known as the regression equation. Mulitple regression is the technique by which estimates are found for the coefficients of the independent variables in the regression equation. The regression equation is then used to forecast future values of the dependent variable.

ii. Econometrics:

Econometrics involves systems of interrelated regression equations. Regression analysis is used to estimate the coefficients of the variables in these equations.

iii. Multivariate Box-Jenkins:

The multivariate Box-Jenkins method is an extension of the univariate Box-Jenkins method which attempts to relate these independent variables to the dependent variable by means of the transfer function.

1.4 Choosing A Forecast Method

Main factors governing the choice of a forecasting method are:

1. Lead time:

The lead time is defined as the length of time into the future for which forecasts are required. The length of the lead time is usually categorized as follows:

Immediate: Less than one month.

Short term: One to three months.

Medium term: More than three months to less than two years.

Long term: Two years or more.

It has been found that because of the other factors considered below, certain forecasting methods are more suitable for a short lead time and other methods for a long lead time.

2. Time to prepare the forecast:

This means the total time needed to collect the necessary data, analyse them, and prepare the forecast. The various forecasting methods differ in the total time that this process takes. In any situation, forecasts must be ready before a certain time; otherwise they are useless for decision making. This makes some of the forecasting methods inappropriate when forecasts are urgently needed.

3. Pattern of data:

There are four basic subpatterns, some combination of which usually exists in any business or economic series data. These are (a) trend, (b) horizontal, (c) seasonal, and (d) cyclical. A trend exists when there is a pattern of growth or decline in the data over the time span being studied. A horizontal subpattern exists when the data are evenly distributed over time that is, when there is no apparent growth or decline over time. A seasonal subpattern is a periodic regular pattern with a constant period such as a day, a week, a month, or a year. Finally, a cyclic subpattern exists when the data are influenced by longer term economic fluctuations related to the general business cycle.

Knowledge of the types of subpatterns included in the data can be very useful in selecting the most appropriate forecasting method, since different methods vary in their ability to cope with different types of patterns.

4. Data requirements:

Since various forecasting methods require different amounts of historical data, the quantity of data available is important in determining the feasibilty of using alternative methods. If the needed historical data are not available, special data collecting procedures may have to be implemented.

5. Ease of understanding:

The ease with which a forecasting method is understood is very important. Managers are held responsible for the decisions they make and if they are to be expected to base their decisions on predictions generated by forecasting techniques, they must be able to understand these techniques.

6. Cost:

When choosing a forecasting technique, several costs are relevant. First, the cost developing the model must be considered. The complexity, and hence the cost, of this process varies from technique to technique. Secondly, the cost of storing the necessary data must be considered. Lastly, the cost of the actual operation of the forecasting technique is obviously very important.

7. Accuracy:

The accuracy of a method is defined as the extent to which forecast values generated by that method approximate to the actual future values that emerge in time. The level of accuracy needed in a given situation will depend on the importance of the decision being made and the role of the forecast in influencing the decision.

Chapter 2

WATER DEMAND FORECASTING

2.1 Nature of forecasting

Water demand forecasting is an important aspect of any water industry. Water demand forecasting can be broadly classified as:

i. Immediate forecasting

ii. Short term forecasting

iii. Medium term forecasting

iv. Long term forecasting

Immediate forecasting is forecasting water demand with a lead time of one or two days. This is used for optimum operation of the source and distribution system. For example, if a drop in demand is predicted for the following day because of weather conditions, the filling up of major reservoirs could be delayed and undertaken during off peak hours, thereby reducing energy costs.

Short term forecasting involves a lead time of 3 to 6 months. This forecasting is done to undertake major maintenance work on treatment plants, distribution system etc.,.

Medium term forecasting involves a lead time of upto 5 years. This is done to facilitate minor source amplifications and construction of reservoir systems.

Long term forecasting involves a lead time of about 20 years. This is done for major source amplifications.

2.2 Models

Several mathematical models were developed in the past to forecast water demands with various lead times. The following is a brief outline of some of the models.

Primeaux et al (1974) developed a model based on linear multiple regression analysis for forecasting annual water demands. The variables used were the number of persons per residence, the number of bathrooms per residence, annual rainfall, maximum temperatures and other economic data relating to the area. Clark (1982) developed a model for forecasting monthly demands based on total population and the spatial distribution of population levels. Both the above models are based on regression analysis. Yamuchi (1977) developed additive and multiplicative models to analyze the trend, cyclical, seasonal and irregular components of daily water consumption. Maidment (1985) developed transfer function models of daily water use. The method used by Maidment was based on a time series model of daily municipal water use as a function of rainfall and temperature. Kher (1986) proposed a water demand model from a set of noisy data using the noisy realisation theory. A mathematical programming based solution algorithm was developed to identify a first order lag dynamical model for the above noisy problem.

All the above models, forecasting either annual, monthly or daily water demands assume model and parameter invariance over a period of time. For forecasting models, with a lead time of more than few months, model invariance is not a valid assumption. However, for immediate and short term forecasting model invariance can still be assumed. Parameter invariance even for immediate and short term forecasting may not be a valid assumption. For example, variation of daily temperatures affect water demands more in summer months than during winter months. Similarly rainfall during winter months is likely to have very little effect on water demands than during summer months. Hence the parameters that represent rainfall and temperatures are likely to be different for summer and winter months.

2.3 Present study

The objective of the present study is to develop a forecasting method for forecasting water demands with a lead of time of a day or two. No attempt was made to develop a software package.

The input to the model is weather data. This includes daily maximum temperatures and rainfalls. The weather data is used as the input because this information is readily available and the cost of data collection is minimal. If other data, like economic, demographic variables are available, the method developed here can still be used with little or no modification.

In the present study model invariance is assumed. This is because the forecasting lead time is only a day or two. However, the model assumes that parameters vary as a function of time. One obvious question here is how one can use a model with varying parameters for forecasting water demand. Although the model assumes parameter variance, when it comes to forecasting, the model assumes parameter invariance for the period of the lead time. In other words, the model tracks parameter variance over a period of time and hence the parameters estimated for the month January will be different to that estimated for the month of June. Since forecasting is done only for a lead time of a day or two, the most recent parameters estimated are used for forecasting. This will give better forecasting results than using constant model parameters.

Chapter 3

THEORY

3.1 Recursive Parameter Estimation

Consider a linear regression problem (Young 1974) in which a variable x_o is known to be related to n other linearly independent variables x_j, j = 1,2,.....n by a linear relationship of the form:

$$x_o = a_1 x_1 + a_2 x_2 + ... + a_n x_n \; ...(1)$$

where a_j, j=1,2,...n are n unknown but constant parameters which characterize the relationship and are to be estimated in some manner. The variables x_j are assumed exactly known quantities but x_o, on the other hand, can be observed only in the presence of noise ϵ_y. If the observation of x_o is denoted by y, then clearly,

$$\mathrm{y} = x_o + \epsilon_y$$

or

$$\mathrm{y} = a_1 x_1 + a_2 x_2 + ... + a_n x_n + \epsilon_y$$

Suppose now that we have k such noisy observations $y_i, i = 1, 2, ...k$ and associated with the $i_t h$ observation is a set of the variables x_j denoted by x_{ji}, j = 1, 2, ... n, thus

$y_i = a_1x_{1i} + a_2x_{2i} + ... + a_{nx.ni} + \epsilon_{yi}$; ... (2)

i = 1, 2, ...k.

Where ϵ_{yi} is the error associated with the i_{th} observation. Suppose further that the sequence of errors ϵ_{yi}, i=1, 2, ... k has the following statiscal properties:

1. It is a zero mean sequence of random variables ie., E[ϵ_{yi}] = 0

2. The ϵ_{yi} are serially uncorrelated and have constant variance σ^2; ie.,

E[$\epsilon_{yi}.\epsilon_{yi}$] $= \sigma^2\delta_{ij}$

where $\delta_{ij} = 1; i = j$ and $\delta_{ij} = 0; i \neq j$, is the kronecker delta function

iii. The ϵ_{yi} are independent of the variables x_{ji}. ... (3)

The estimation problem posed in this manner is one of estimating the values of the unknown parameters a_j given the information $y_i, x_{1i}, x_{2i}, ...x_{ni}$ where i = 1, 2, ... k.

There are many ways of solving this kind of problem, but the best known and the one with which we are concerned here is the method of least squares in which the estimates $\hat{a}_j, j = 1, 2, ...n$ of the n unknown parameters are chosen to minimise the least squares cost function J, where:

J = $\sum_{i=1}^{n}[\sum_{j=1}^{n} x_{ji}\hat{a}_j - y_i]^2$... (4)

Since this cost function is unimodal in the cost function parameter space, minimisation with respect to the $\hat{a}_j$ simply requires that all the n partial derivatives of J with respect to the $\hat{a}_j$, j = 1, 2, ...n should be simultaneously set to zero. Such a procedure yields a set of n linear, simultaneous algebraic equations, the "normal equations" of linear regression analysis; equations that can be solved to obtain the "least squares" parameter estimates

$\hat{a}_{jk}$,j = 1, 2, ...n

where the subscript k denotes that the estimates are based upon the set of k observations.

A simple, concise statement of the least squares results in the multiparameter case can be obtained by using a vector matrix formulation: thus by writing (1) in the alternative vector form:

$x_o = x_a^T$ ... (5)

where
$x^T = [x_1 x_2 ... x_n]; a = [a_1, a_2, ... a_n]^T$,

we are able to define J as

$j = \sum_{i=1}^{k} [x_i^T \hat{a} - y_i]^2$ ... (6)

where

$y_i = x_i^T a + \epsilon_{yi}$; i = 1, 2, ...k. ...(7)

the normal equations then become

$\nabla_{\hat{a}(j_2)=[\sum_{i=1}^{k} x_i x_i^T]\hat{a}-\sum_{i=1}^{k} x_i y_i=0} \cdots$ (8)

Provided that the matrix $x_i x_i^T$ is non singular, the solution to the equation (8) takes the form

$\hat{a} = P_k b_k$ ... (9)

where

$P_k = [\sum_{i=1}^{k} x_i x_j^T]^{-1}$

and

$b_k = \sum_{i=1}^{k} x_i y_i$ ... (10)

Suppose now we wish to develop a recursive form of the solution given in (9) in which the estimate after k samples, $\hat{a}_k$, is a linear sum of the estimate obtained after k-1 samples $\hat{a}_{k-1}$, plus a corrective term based on the new information y_k and x_k received at the kth sampling instant. We are able to obtain such a solution by first noting from definitions in (10), that P_k and b_k can be related to their prior values P_{k-1} and b_{k-1}, by the following equations:

$P_k^{-1} = P_{k-1}^{-1} + x_k x_k^T$ ...(11)

$b_k = b_{k-1} + x_k y_k$ ... (12)

By straightforward matrix manipulation equation (11) can be transformed into the following recursive expression for P_k in terms of P_{k-1}

$P_k = P_{k-1} - P_{k-1} x_k [1 + x_k^T P_{K-1} x_k]^{-1} x_k^T P_{k-1}$ ... I(1)

which is often termed the "matrix inversion lemma". The recursive equation for $\hat{a}_k$ in terms of $\hat{a}_{k-1}$ can then be obtained by substituting from I(1) and (12) into (9) to yield,

$\hat{a}_k = \hat{a}_{k-1} - k_k [x_k^T \hat{a}_{k-1} - y_k]$ ... I(2)

in which k_k is a gain vector defined by

$k_k = P_{k-1} x_k [1 + x_k^T P_{k-1} x_k]^{-1}$

Finally, an alternative expression to I(2) can be found by multiplying k_k by $P_k P_{k-1}$ and substituting from equation (11) for P_k^{-1}, ie.,

$k_k = P_k [P_{k-1}^{-1} + x_k x_k^T] P_{k-1} x_k [1 + x_k^T P_{k-1} x_k]^{-1}$

$= P_k x_k$

Thus, equation I(2) can be written conveniently as,

$\hat{a}_k = \hat{a}_{k-1} - P_k[x_k x_k^T \hat{a}_{k-1} - x_k y_k]$ ... I(3)

Equation set I constitutes a recursive form of the least squares solution (9) and, although derived and presented differently, is equivalent to that suggested by Placket (Young 1974). In this form, however, the algorithm is completely deterministic, since it makes no use of the statistical assumptions (3) about the nature of the observation errors ϵ_{yi}. To incorporate such information into the algorithm, we must utilize some of the basic results of linear regression analysis. Since these are derived in most standard statistical texts, it suffices here to state that, provided the various assumptions made in the formulation of the linear regression problem are satisfied, the estimate $\hat{a}_k$ is statistically consistent and the estimation error vector

$\tilde{a} = \hat{a}_k - a$

has the following statistical properties:

1. Zero mean value $E[\hat{a}] = 0$

2. The variance covariance matrix $P_k^* = E[\tilde{a}_k \tilde{a}_k^T]$ is related to the matrix P_k in equation set I by the simple equation $P_k^* = \sigma^2 P_k$.

Noting the latter result, it is simple to obtain the recursive least squares regression algorithm by substituting P_k^*/σ^2 for P_k in the recursive least squares algorithm I to yield:

$\hat{a}_k = \hat{a}_{k-1} - P_{k-1}^* x_k[\sigma^2 + x_k^T P_{k-1}^* x_k]^{-1}[x_k^T \hat{a}_{k-1} - y_k]$

or

$\hat{a}_k = \hat{a}_{k-1} - P_k^*/\sigma^2[x_k x_k^T \hat{a}_{k-1} - x_k y_k]$ ... II(2)

while

$$P_k^* = P_{k-1}^* - P_{k-1}^* x_k[\sigma^2 + x_k^T P_{k-1}^* x_k]^{-1} x_k^T P_{k-1}^* \quad ... \quad \mathrm{II}(3)$$

This algorithm not only supplies the parameter estimates at each sampling instant, but also provides an indication of the accuracy of these estimates through the error covariance matrix P_k^*. It is, therefore, an extremely elegant algorithm which is particularly useful for processing data "on line", as may be required in several forecasting situations. But the algorithm is not only elegant it also provides some considerable advantages over the repeated or stagewise solution of (9) ; in addition to the new convenient recursive form, which provides additional information on the convergence of the estimates as well as the minimum computer storage, it will be noted that the term [1 + $x_k^T P_{k-1} x_k$] is simply a scalar quantity. As a result, there is no requirement for direct matrix inversion, even though the equivalent classical solution (9) applied repeatedly would entail the inversion of an nxn matrix at each step.

Before describing how the recursive regression algorithm II can be extended to accomodate problems in which the parameter vector may vary over the observation interval, it is worthwhile considering certain characteristics of the algorithm and how it performs in practice. For instance, since the algorithm is in recursive form, it is necessary to choose some starting values for $\hat{a}_o$ and P_o. One obvious approach is to first read in data for k=N, where N>n, to evaluate P_N and then solve the non recursive equation (9) to yield an initial estimate $\hat{a}_N$; the recursive algorithm can then be applied to all subsequent data for k>N.

A less obvious but more convenient approach becomes apparent if we consider the statistical importance of P_k^* as an estimation error covariance matrix; it then makes sense to choose P_o^* as consistent with the level of confidence one has in the initial estimate $\hat{a}_o$. Indeed with this interpretation, there is a direct link with Bayesian estimation, since $\hat{a}_o$ and the associated P_o^* represent the *apriori* statistics of the estimate which will be updated on receipt of the set of information y_1 and x_1 to yield the *posteriori* estimate $\hat{a}_1$ and the associated *aposteriori* covariance matrix P_1^*. These *aposteriori* statistics then become the *apriori* statistics for the next step in the recursion and the estimation proceeds in this manner.

In practice of course, it may well be that little is known *apriori* about the parameters in which case, it seems reasonable to choose $\hat{a}_o$ to be an arbitrary finite value ($\hat{a}_o$ = [0] seems most useful) and set P_0^* to be a diagonal matrix with large diagonal elements (of the order of 10^3 to 10^6), indicating little confidence in the initial estimate and no knowledge of the cross-variance properties of the estimates. This choice of $\hat{a}_o$ and P_o^* also makes sense in purely deterministic terms since it can be shown that the resulting performance of the algorithm is commensurate with the stagewise solution of the same problem obtained by repeated use of equation (9) for k>n. Indeed, not only are the results asymptotically equivalent as $k \rightarrow \infty$, but the convergence of the recursive values of $\hat{a}_k$ and P_k in algorithm I to the stagewise value is rapid, provided the diagonal elements of P_o are chosen to be large enough.

The interpretation of P_k^* as an estimation error covariance matrix is also interesting if we consider its function in equation II(2); reference to equations (6) and (8) shows that the correction term

$$P_k^*/\sigma^2(x_k x_k^T \hat{a}_{k-1} - x_k y_k)$$

can be considered as an instantaneous measure of the gradient of J at the kth instant, modulated or controlled by matrix P_k^*/σ^2; thus II(2) can be considered as a multi dimensional, discrete step, gradient procedure in the criterion function parameter space. Considering the consistency of the estimation procedure, we know P_k^* will be a strictly decreasing function of time, so as the estimation proceeds and confidence in the estimates increases, the weighting associated with the correction term is reduced, since it is more likely that the observed gradient is the result of observation error on y_k than the true estimation error; in effect the matrix P_k^*/σ^2 (or P_k in the deterministic case) acts to smooth or filter out the inaccuracy injected by the observational noise.

3.2 Parameter Variance

Basic least squares regression analysis, whether in its conventional block data processing form or in the recursive formulation described in the previous section, carries with it an implicit assumption that the parameters in the regression model are sensibly constant over the observational interval. If this is not the case, or if there is uncertainity on this point, it is clearly dangerous to use the analysis directly as biased and out of date estimates may be obtained.

One well known approach to this problem is to curtail the *memory* of the estimation procedure, for instance by the use of exponentially weighted past (EWP) averaging of the data. Recursive formulations of such EWP algorithms can be derived in a manner similar to that used in the previous section and useful in emphasising more recent data in comparison to the data received in the past; the exponential weighting introduces, as it were, a degree of *forgetting* into the algorithm.

The mechanics of this procedure rely on adjusting (II) to the form (Viswanathan 1982):

$$P_k^{-1} = \mu P_{k-1}^{-1} + x_k x_k^T \ \dots \ (13)$$

where μ is a *forgettingfactor* which usually takes the value $0.9<\mu<0.99$. Smaller μ correspond to more rapid forgetting of old data, while $\mu = 1$ corresponds to the normal arrangement with no forgetting.

EquationsI(1) and (3) are modified using (13). This gives:

$$P_k = 1/\mu[P_{k-1} - P_{k-1}x_k[\mu + x_k^T P_{k-1} x_k] x_k^T P_{k-1}] \ \dots \ (14)$$

Then substituting P_k^*/σ^2 for P_k one obtains

$$P_K^* = 1/\mu[P_{k-1}^* - P_{k-1}^* x_k[\sigma^2\mu + x_k^T P_{k-1}^* x_k] x_k^T P_{k-1} \ \dots \ (15)$$

Equation (15) along with II(1) constitute a recursive algorithm for the prediction of variable parameters in the regression model.

3.3 Transfer Function Models

Consider the estimation of parameters in a discrete time series or pulse(z) transform transfer function representation of a linear stochastic dynamic system. The basic deterministic part of the system, shown in figure 3.1 is block A, which is described by the following difference equation,

$$A[Z^{-1}]x_k = B[Z^{-1}]u_k \; ... \; (16)$$

where Z^{-1} is the backward shift operator ie.,

$$Z^{-1}x_k = x_{k-1}$$

while

$$A[Z^{-1}] = 1 + a_1 Z^{-1} + ... + a_n Z^{-n}$$

and

$$B[Z^{-1}] = b_o + b_1 Z^{-1} + ... + b_n Z^{-n}$$

are polynomials in Z^{-1} operating on the hypothetical deterministic output x_k and the deterministic input u_k respectively.

Utilizing the superposition property for linear systems, the various noise disturbances entering the system are lumped into an equivalent noise disturbance, which is added to the hypothetical noise free output, x_k to yield the observed output y_k as shown in figure 3.1. This noise term, which includes input and other disturbance effects as well as measurement noise, is assumed to have rational spectral density; ie., it is considered as the output of a transfer function as shown in block B of figure 3.1 whose input, e_k is a zero mean, serially uncorrelated sequence of random variables with variance σ^2 (discrete white noise) ie.,

$$E[e_k] = 0$$

$E[e_j e_k] = \sigma^2 \delta_{jk}$

where

$\delta_{jk} = 1 when j = k$and $\delta_{kj} = 0 when j \neq k$ ... (17)

Further e_k is independent of the deterministic input u_k ie.,

$E[e_j u_k] = 0$ for all j,k ... (18)

in this way ξ_k is generated from the discrete white noise e_k by the following difference equation

$C[Z^{-1}]\xi_k = D[Z^{-1}]e_k$ ... (19)

where

$C[Z^{-1}] = 1 + c_1 Z^{-1} + ... + C_n Z^{-1}$

$D[Z^{-1}] = d_o + d_1 Z^{-1} + ... d_n z^{-n}$

Equation (19) is reffered to as an Auto Regressive Moving Average process (ARMA). Equations (16) and (19) provide parametrically efficient models of the deterministic and stochastic aspects of the system. The model is completed by the introduction of the output equation ie.,

$y_k = x_k + \xi_k$ ... (20)

By substituting the observed equation (20) in (16), it is possible to describe the overall stochastic dynamic system by the following vector relationships

$y_k = z_k^T a + k$ ... (21)

$\xi_k = n_k^T c + d_o e_k$ ... (22)

where

$k = a_1\xi_{k-1} + ... + a_n\xi_{k-n} + \xi_k$

and

$z_k^T = [-y_{k-1} ... - y_{k-n}, u_k ... u_{k-n}]$

$a^T = [a_1 ... a_n, b_o ... b_n]$

$n_k^T = [-\xi_{k-1} ... - \xi_{k-n}, e_{k-1} ... e_{k-n}]$

$c^T = [c_1 ... c_n, d_1 ... d_n]$

The problem is then to utilise the sampled input data u_k and output data y_k to obtain statistically efficient estimates of the parameter vectors a and c that characterise the model.

Special case

A special degenerate case of figure 3.2 is where the noise model is characterised by a transfer function

$N[Z^{-1}] = 1/A[Z^{-1}]$

so that the model can be written

$y_k = B/A(u_k) + (1/A)e_k$ ... (23)

A particular example of this situation is the case shown in figure 3.3, where the system has no numerator dynamics ie.,

$B[Z^{-1}] = 1.0$

and the noise enters as serially uncorrelated white noise at the input to the system.

For the model of figure 3.2 equation 21 takes the reduced form

$$y_k = z_k^T a + e_k \; ... \; (24)$$

where e_k is a serially uncorrelated sequence of random variablers with variance σ^2 and z_k^T is given by

$$z_k^T = [-y_{k-1} \cdots - y_{k-n}, u_k \ldots u_{k-n}]$$

Since u_k is the deterministic input sequence, it is clear that this model is an extension of the simple Auto Regressive time series model. It is well known that an asymptotically unbiased and efficient estimate $\hat{a}_k$, of the parameter vector "a" can be obtained by simple least squares regression analysis. The recursive regression algorithm in this case can be obtained using equations (14) and (15).

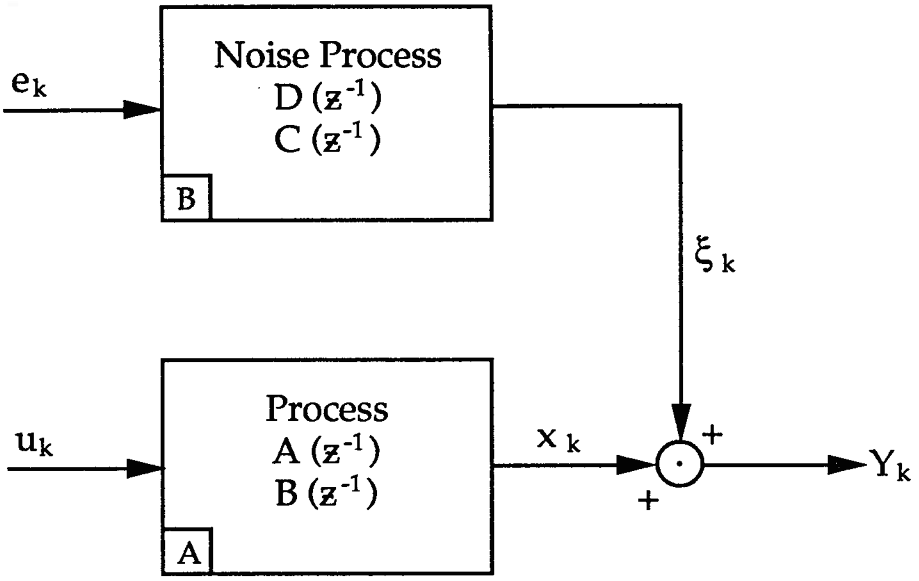


Figure 3.1: Transfer Function Model 1

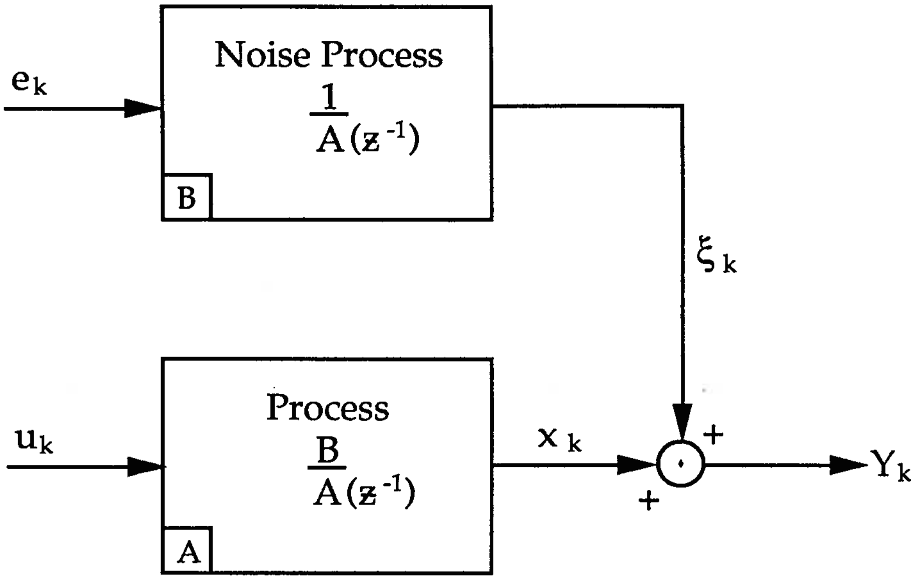


Figure 3.2: Transfer Function Model - case (a)

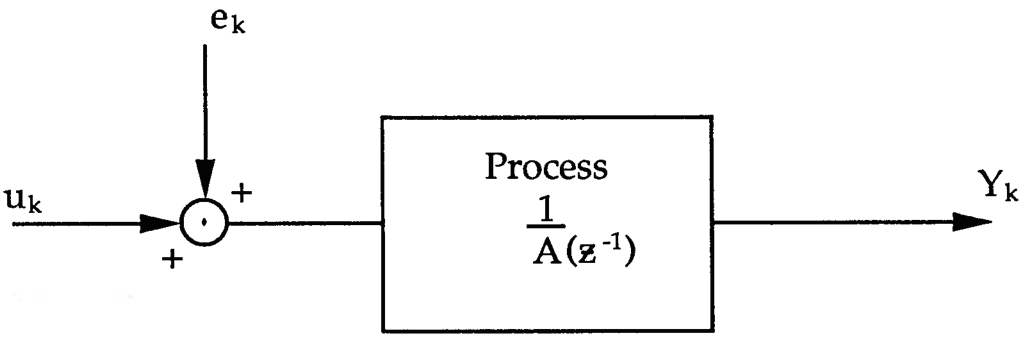


Figure 3.3: Transfer Function Model - case (b)

Chapter 4

PRESENT MODEL

Using methods outlined in the previous sections a model for Newcastle water demand is developed. The model is shown in figure 4.1 and has the same form as that described in section 3.0 except that an additional term, a constant in the first instance is inserted to account for the base usage.

The other inputs to the model are daily maximum temperatures and daily rainfall. The coefficients in the model are assumed to be time dependent. The model also assumes delayed effects of temperatures and rainfall. In other words a heavy rainfall on a particular day is likely to affect water demand on the following days. The constant term included represents other variables (other than temperature and rainfall) affecting water demand. In addition to the above deterministic variables, there could be a random component affecting the water demand and this is represented by the term e_k.

Equation 23 is written as:

$$y_k - a_1 y_{k-1} - a_2 y_{k-2} \ldots a_n y_{k-n} = [b_o T_k + b_1 T_{k-1} + \ldots b_n T_{k-n}]$$
$$+[c_o R_k + c_1 R_{k-1} + \ldots c_n R_{k-n}] + d_o + e_k \ldots \quad (25)$$

where

y_k=water demand on day'k'

T = daily maximum temperature

R = daily rainfall

e = error term
d_o = base usage

subscripts k,k-1 ... days of year.

Equation 25 is the general equation which express the water demand on the kth day in terms of temperatures, the rainfall on days k,k-1 ... , water consumptions on days k-1,k-2 ... and the error term.

The following equation is used for forecasting water demand on day k:

$$y_k = [a_1 y_{k-1} + \ldots a_n y_{k-n}] + [b_o T_k + \ldots b_n T_{k-n}] + [c_o R_k + \ldots c_n R_{k-n}] + d_o \ldots (26)$$

The number of coefficients a_n, b_n and c_n in the model were determined by trial and error. After several trials , it was found that the following equation represented the Newcastle water demand best:

$$y_k = a_1 y_{k-1} + b_o T_k + b_1 T_{k-1} + c_o R_k + c_1 R_{k-1} + d_o \ \ldots \ (27)$$

Using the recursive least squares algorithms (equations 15 and II(1)) the coefficients a_1, b_o, b_1, c_o, c_1 and d_o were estimated.

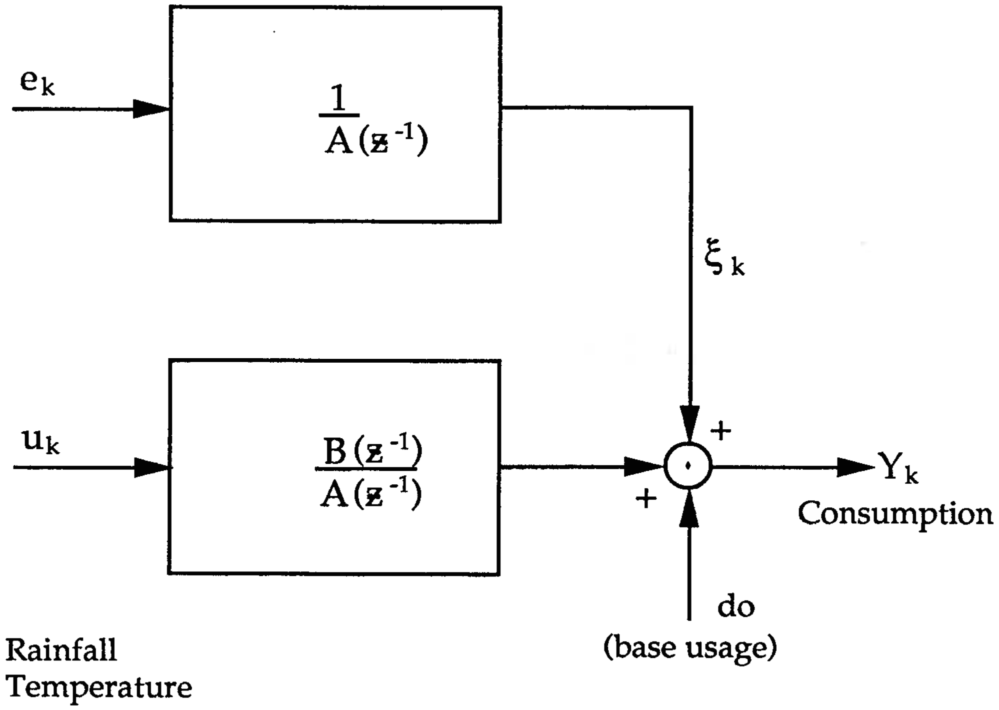


Figure 4.1: Model for Water Demand

Chapter 5

CASE STUDY

The model developed in the previous section is used to forecast water demand for Newcastle with a lead time of day 1. Newcastle has a population of about 500 000.

Variation of daily maximum temperatures and rainfalls for the year 1988 at Newcastle are shown in figures 5.1 and 5.2.

Figure 5.3 shows water demand forecasts and actual water consumptions.

Figure 5.4 shows the difference between the forecasts and actual consumption. The forecast error is within ±5 percent.

Figures 5.5 and 5.6 show the variation in parameters. The parameter variation is substantial between summer and winter months.

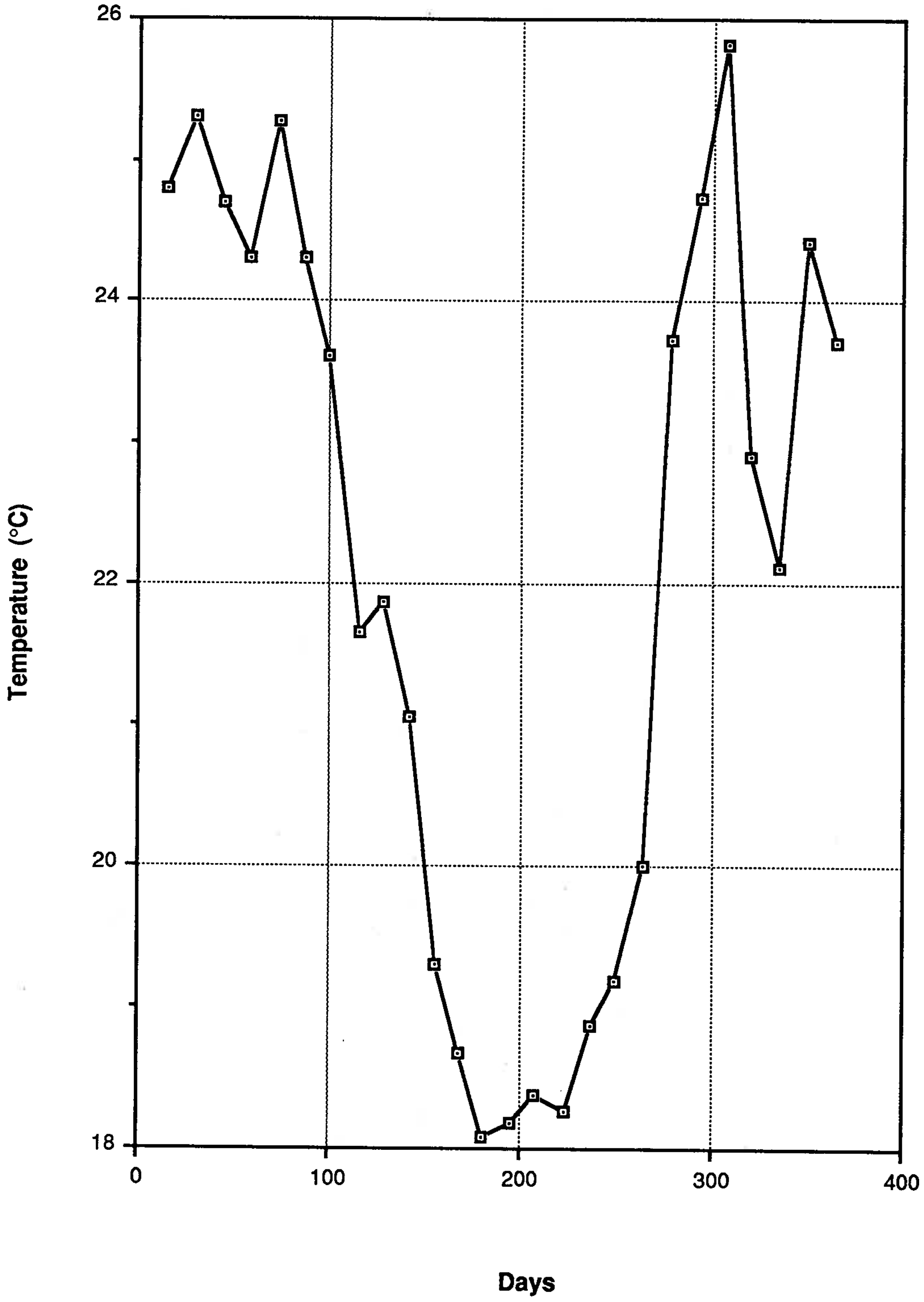


Figure 5.1: Temperature variations

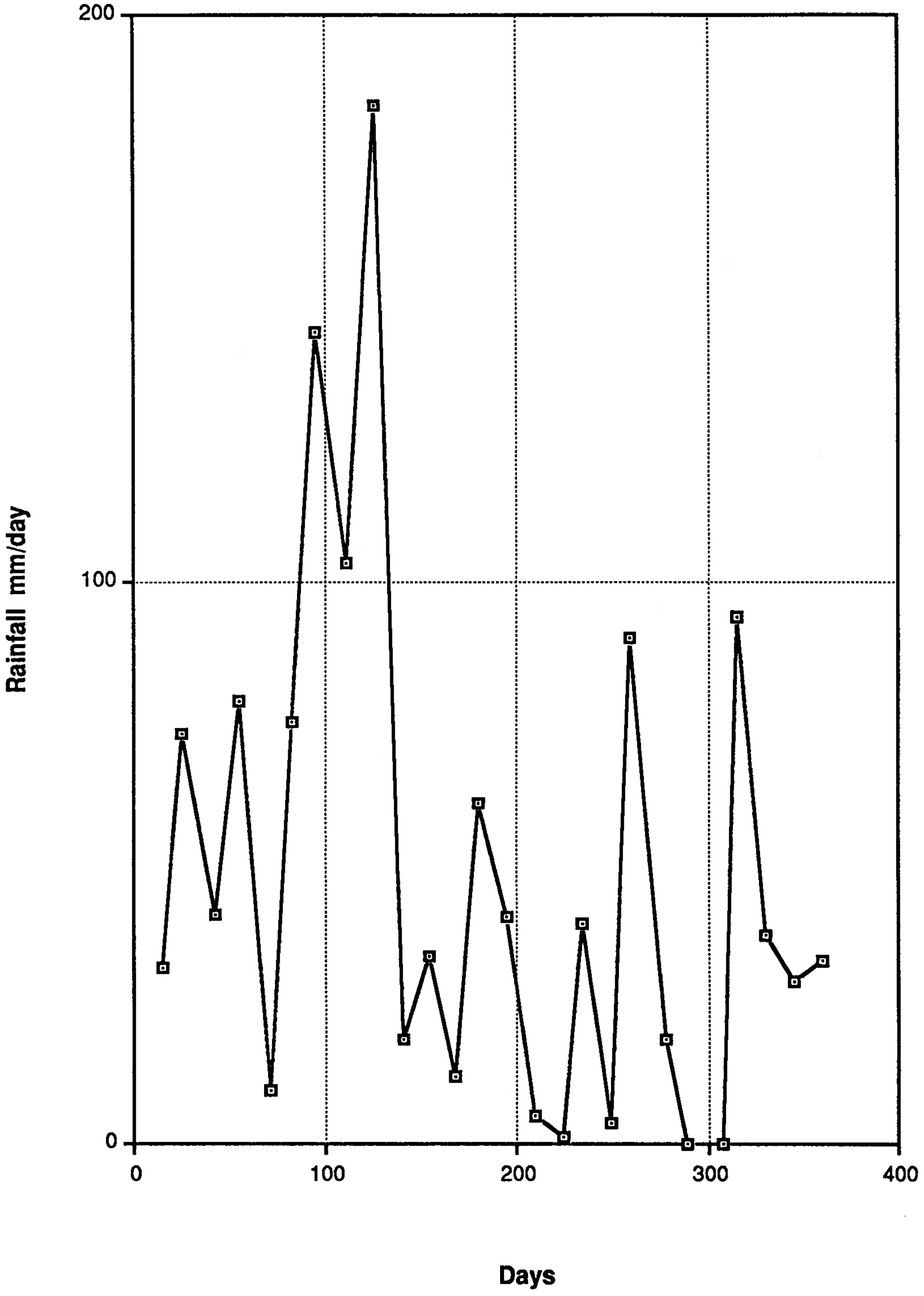


Figure 5.2: Rainfall

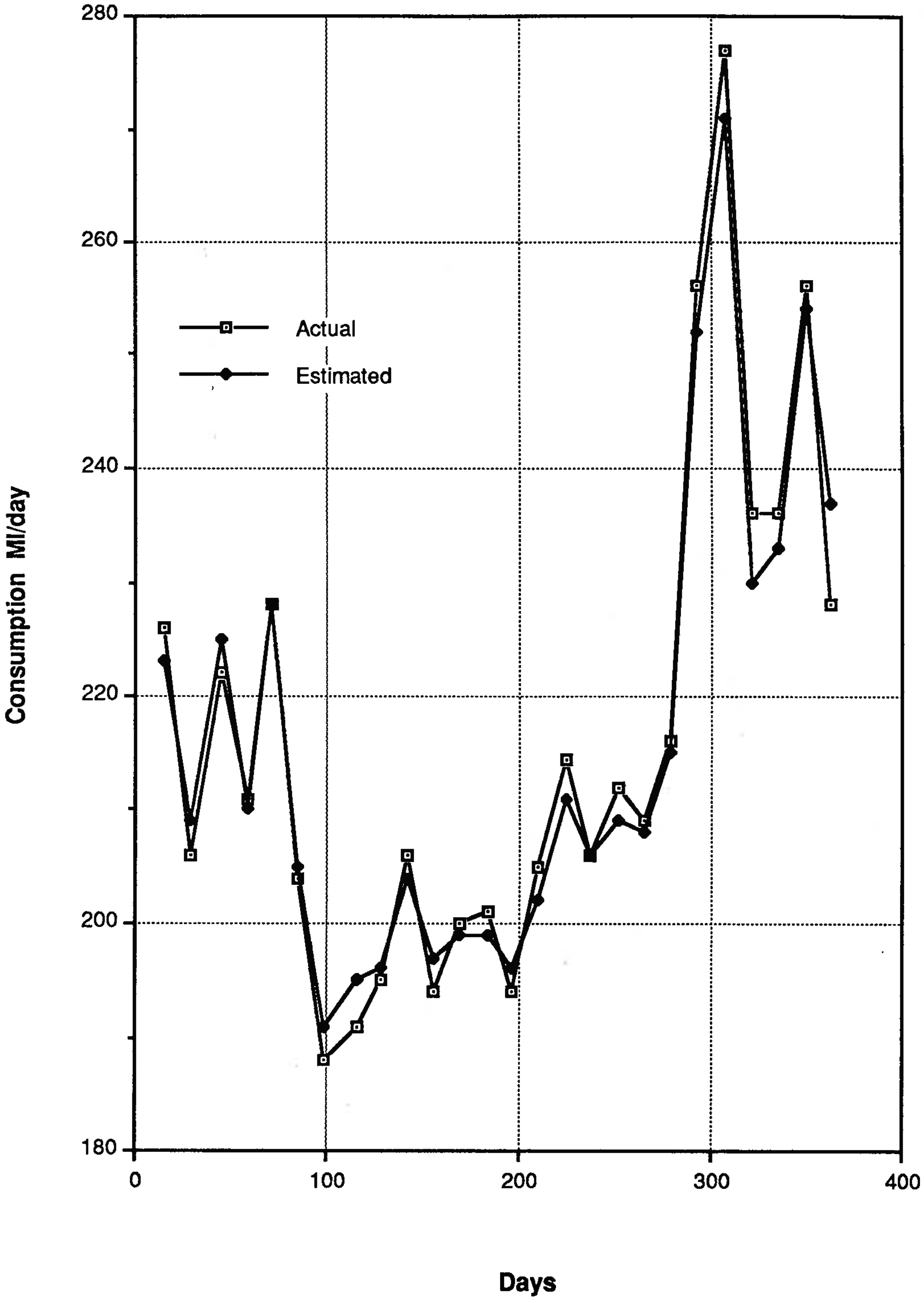


Figure 5.3: Water Demands - Actual and Estimated

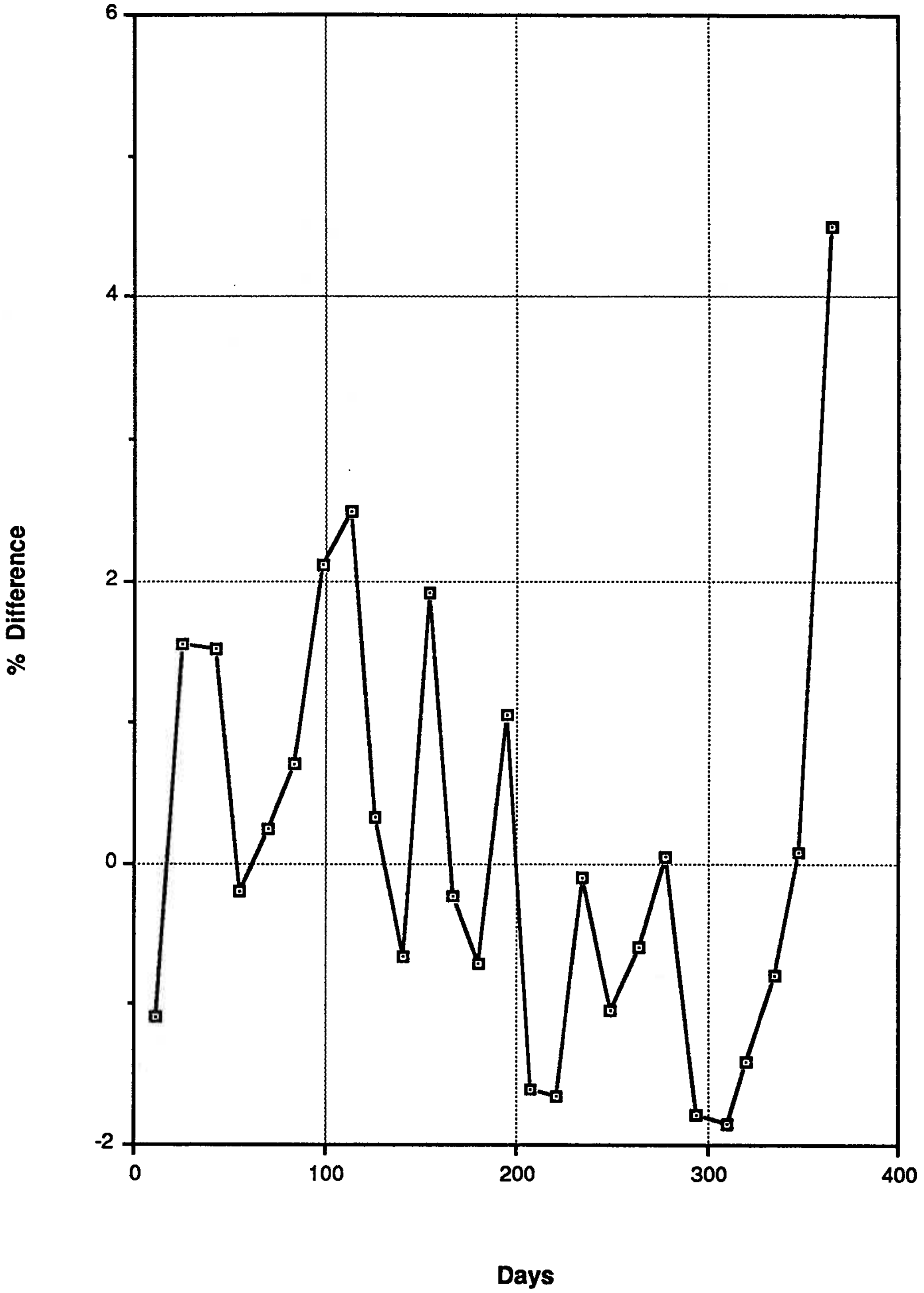


Figure 5.4: Differences in Estimated and Actual Water Demands

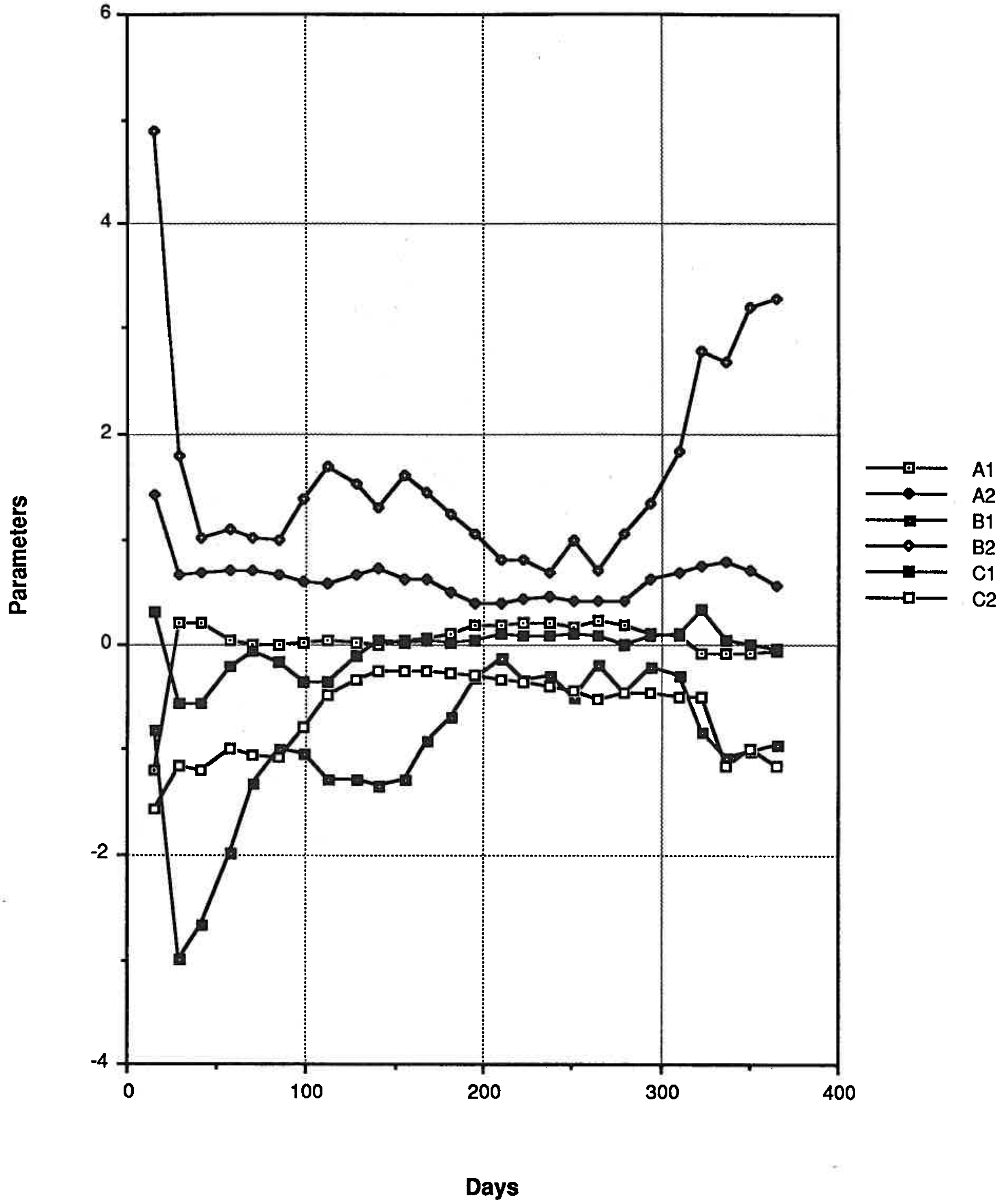


Figure 5.5: Variation of parameters

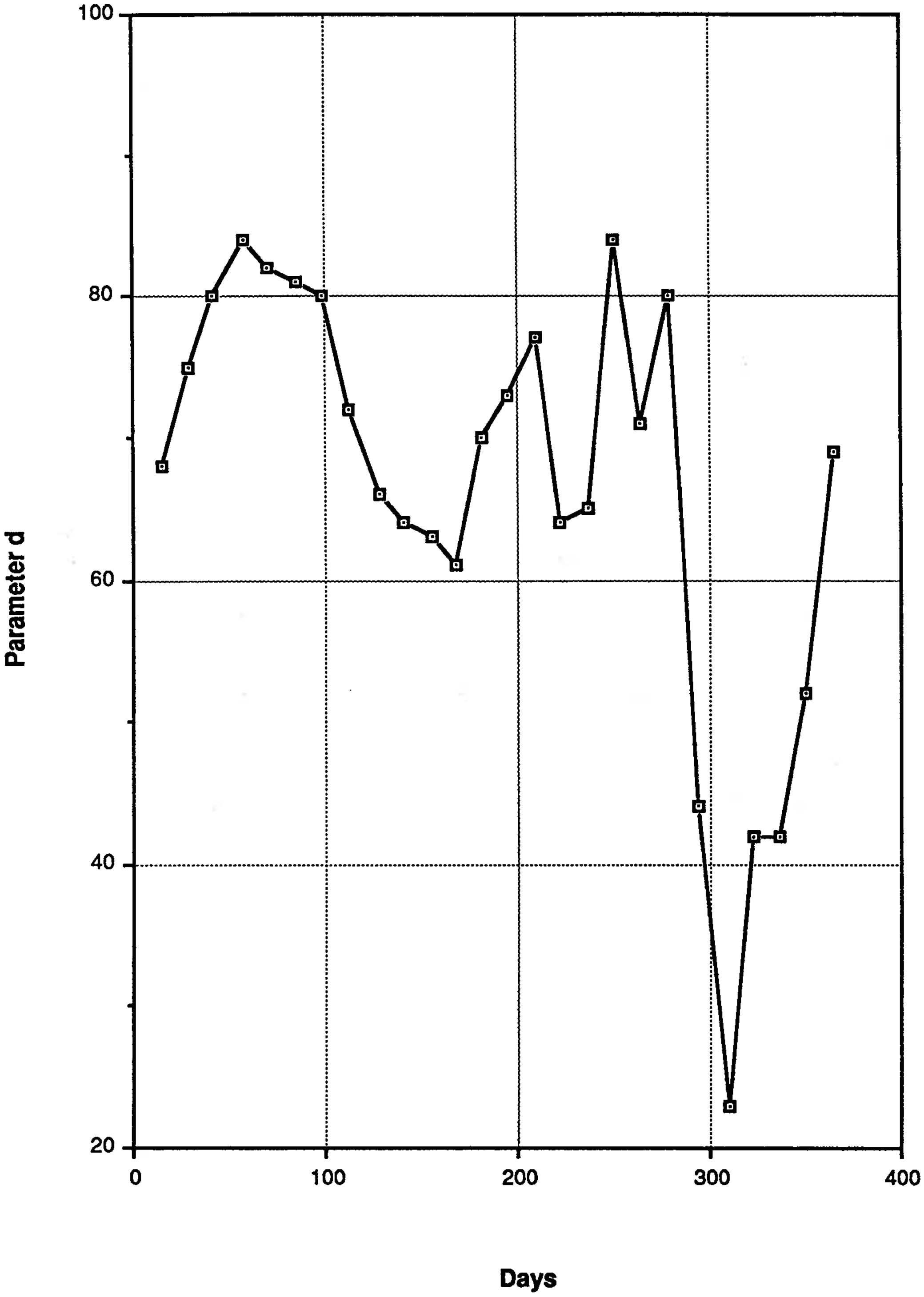


Figure 5.6: Variation of Parameter d

Chapter 6

CONCLUSIONS

1. A model was developed to forecast water demand for an urban area. The model forecasts water demand for day 'k' given the weather parameters for days 'k' and 'k-1' and water consumption for the day 'k-1'.

2. The model is linear and the development is based on the method 'Recursive Least Squares'.

3. Coefficients of the model are assumed to be time dependent.

4. The model was used to forecast water demand for Newcastle with a lead time of day 1. The model forecasts water demand within an accuracy of ±5 percent.

Chapter 7

REFERENCES

Abraham B and Ledolter J 1983
Statiscal methods for forecasting
John Wiley and Sons NewYork

Clark R M Gillean J I 1982
Predicting future water demand
Journal of the environmental division. Proceedings of the ASCE vol 108
EE6 December pp 1248-1264

Kher L K Soorooshian S 1986
Identification of water demand models from noisy data
Water resources research vol 22 no 3 March pp 322-330

Maidment D R Miau S P Crawford M M 1985
Transfer function models of daily urban water use
Water resources research vol 21 no 4 April pp 425-432

O'Donovan T M 1983
Short term forecasting
John Wiley and Sons NewYork

Primeaux W J Hollman K W 1974
Factors affecting residential water consumption - The managerial point of view.

Water and wastewater R138 - R144.

Viswanathan M N 1982
Determination of effectiveness of water consumption restrictions:
Methodology and application
M E Thesis University of Newcastle

Yamuchi H Huang W 1977
Alternative models for estimating the time series components of water consumption data
Water resources bulletin vol 13 no 3 June 599 - 610.

Young P 1974
Recurcive approaches to time series
The institute of mathematics and its applications May/June pp209 - 224

Chapter 8

ACKNOWLEDGEMENTS

The author thanks Hunter Water Board and Urban Water Research Association of Australia for funding this project.
The author also thanks Dr R Wilson for his comments on the first draft.

UWRAA RESEARCH REPORTS

Report No.	Title	Author
1	Trickling filter - solids contact process: Pilot plant studies.	M. Laginestra
2	A model of water pricing for Melbourne, Sydney and Perth	P.B. Dixon P.M. Norman
3	Taste generation associated with chloramination	M. Kerslake
4	Bacterial regrowth in water supplies	K. Power L.A. Nagy
5	Leakage management: Assessing the effect of pressure reduction on losses from water distribution systems	B. Horvath
6	Improving communication with the public on water industry policy issues	B.E. Nancarrow G.J. Syme
7	Water use efficiency of domestic appliances	I.J. Beith D.J. Horton
8	Pilot plant study of fermentation units to increase degradable COD fraction in sewage.	P.J. Bliss D. Barnes P.R. Evans I. Law
9	Artificial destratification of water storages in Australia.	T.F. McAuliffe R.S. Rosich
10	Taste thresholds of monochloramine and chlorine in water	R. O'Halloran C. Veres
11	Chromatographic analysis of chloramines using electro-chemical detection	R. O'Halloran Hai Lin Ge P. Spizziri
12	Glass reinforced plastic bore casing for large diameter and deep bores	R. Bowyer
13	A guide to improving communication with the public on water industry policy issues	B.E. Nancarrow G.J. Syme
14	Fouling and cleaning of fine bubble ceramic dome diffusers	K.J. Hartley
15	Chloramination of Water Supplies	P.M. Thomas (ed)
16	The 1988 Australian Winter Storms Experiment: Report on aircraft observations	J.B. Jensen
17	Pipeline assets: Life cycle management and economic life	R. Vass M. Anderson R. Lewis D. Samson
18	Development of empirical model for tradewaste discharges to small treatment plants	Camp Scott Furphy
19	PRELIM users guide (Amended): Australian Version	Camp Scott Furphy
20	Chemical regeneration of activated carbon: Preliminary studies	G. Newcombe
21	Management and display of dam surveillance data	D.M. Stirling G.L. Benwell A.B. Murnane
22	Evaluation and demonstration facilities for primary sensors	J.A. Lanaway M. Cavey
23	Modelling and design of reservoir aeration destratification systems	D.P. Lewis J.C. Patterson J. Imberger R.P. Wright S.G. Schadlow
24	Modelling optimum conditions for reservoir destratification using mechanical mixers	R.M.A. Velzeboer J.A. Cugley J.C. Patterson
25	Methods for detection of *Giardia* and *Crytosporidium* in water: A preliminary assessment	C.A. Bee P.E. Christy B.E. Robinson
26	Toxic cyanobacteria in water supplies: Analytical techniques	D.J. Flett B.C. Nicholson
27	Tracing toxic discharges to sewers by analysis of biofilms	D. Oliver T. Watson
28	Electronic meter reading: Link between water meter and house	P.J. Reid J.S. Renwick M.F. Prior
29	Identification of common noxious cyanobacteria: Part I – Nostocales	P. Baker
30	Forecasting water demand using weather data	M.N. Viswanathan
31	Effects of controls on water consumption	M.N. Viswanathan