Evaluation of Flow Forecasting Models for Adelaide Hills Catchments

T.J. Baker

A Thesis for the Degree of Master of Engineering Science

University of Adelaide
Department of Civil Engineering

November 1991

This Thesis embodies the results of supervised project work making up all of the work for the degree.

ERRATA

To qualify the comments made in Section [6.5.2] of the thesis, the following derivation is given for the solution of the $[\mathbf{A}]$ and $[\mathbf{B}]$ matrices for an AR(1) multisite model.

Chapter [3] sets out the development of the multisite generation equation, for both the annual case, and further, for the multiperiod, multisite case. The following derivation for the solution of the $[\mathbf{A}]$ and $[\mathbf{B}]$ matrices used in an AR(1) model is shown for the exact case to highlight the difference between the exact solution and the solution given in the thesis.

The general multisite model is given as –

$$[\mathbf{Z_t}] = [\mathbf{A}][\mathbf{Z_{t-1}}] + [\mathbf{B}][\epsilon_\mathbf{t}] \tag{1}$$

To derive the solution for the $[\mathbf{A}]$ matrix, both sides of Equation [1] are post-multiplied by $[\mathbf{Z_{t-1}}]^\mathbf{T}$ and expectations taken.

Thereby giving (using the notation given in Chapter 3) –

$$\mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_{t-1}}]^\mathbf{T}\} = [\mathbf{A}]\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^\mathbf{T}\} + \mathbf{E}\{[\mathbf{B}][\epsilon_\mathbf{t}][\mathbf{Z_{t-1}}]^\mathbf{T}\} \tag{2}$$

$$\Rightarrow [\mathbf{A}] = \mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_{t-1}}]^\mathbf{T}\}\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^\mathbf{T}\}^{-1} \tag{3}$$

$$[\mathbf{A}] = [\mathbf{M_1}][\mathbf{M_0^*}]^{-1}$$

$$\textit{where} \quad [\mathbf{M_0^*}] = \mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^\mathbf{T}\}$$

Similarly, for the solution of the $[\mathbf{B}]$ matrix, both sides of Equation [1] are postmultiplied by $[\mathbf{Z_t}]^\mathbf{T}$ and expectations taken, giving –

$$\mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_t}]^\mathbf{T}\} = [\mathbf{A}]\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_t}]^\mathbf{T}\} + \mathbf{E}\{[\mathbf{B}][\epsilon_\mathbf{t}][\mathbf{Z_t}]^\mathbf{T}\} \tag{4}$$

$$= [\mathbf{A}]\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_t}]^\mathbf{T}\} + [\mathbf{B}]\mathbf{E}\{[\epsilon_\mathbf{t}][\mathbf{Z_{t-1}}]^\mathbf{T}[\mathbf{A}]^\mathbf{T} + [\epsilon_\mathbf{t}][\epsilon_\mathbf{t}]^\mathbf{T}[\mathbf{B}]^\mathbf{T}\} \tag{5}$$

By substituting Equation [3] for [**A**], and rearranging,

$$\Rightarrow [\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = \mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_t}]^{\mathbf{T}}\} - \mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_{t-1}}]^{\mathbf{T}}\}\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^{\mathbf{T}}\}^{-1}\mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_t}]^{\mathbf{T}}\}$$

$$[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = [\mathbf{M_0}] - [\mathbf{M_1}][\mathbf{M_0^*}]^{-1}[\mathbf{M_1}]^{\mathbf{T}}$$

The general matrix solution given by Equations [3.19] & [3.22], and further developed for the multiperiod model given by Equations [3.43] & [3.44] have assumed that the underlying process describing the distribution is stationary. This implies that the covariance matrix calculated using $[\mathbf{Z_t}][\mathbf{Z_t}]^{\mathbf{T}}$ is equal to the covariance matrix calculated using $[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^{\mathbf{T}}$.

There is a subtle difference in using the latter form, and it has been noted by Kuzera [31] that for problems with a limited length of data, the above asymptotic solution may not hold true, and may lead to significantly different results.

By replacing $\quad [\mathbf{M_0^*}] = \mathbf{E}\{[\mathbf{Z_{t-1}}][\mathbf{Z_{t-1}}]^{\mathbf{T}}\}$

in lieu of $\quad [\mathbf{M_0}] = \mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_t}]^{\mathbf{T}}\}$

in Equations [3.19], [3.22],[3.43] & [3.44] the result will overcome matrix inconsistency for the solution of the [**A**] and [**B**] matrices.

The revised equations are thus –

For the general case,

$$[\mathbf{A}] = [\mathbf{M_1}][\mathbf{M_0^*}]^{-1} \tag{6}$$

$$[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = [\mathbf{M_0}] - [\mathbf{M_1}][\mathbf{M_0^*}]^{-1}[\mathbf{M_1}]^{\mathbf{T}} \tag{7}$$

and for the multiperiod case,

$$[\mathbf{A}_\tau] = [\mathbf{M}_{1,\tau}][\mathbf{M}^*_{0,\tau-1}]^{-1} \tag{8}$$

$$[\mathbf{B}_\tau][\mathbf{B}_\tau]^{\mathbf{T}} = [\mathbf{M}_{0,\tau}] - [\mathbf{M}_{1,\tau}][\mathbf{M}^*_{0,\tau-1}]^{-1}[\mathbf{M}_{1,\tau}]^{\mathbf{T}} \tag{9}$$

Kuczera [31] outlines a method to obtain consistent estimates of the [**A**] and [**B**] matrices when there is missing data in any of the records. This approach may have made better use of the streamflow records available in this study.

Furthermore, Crosby & Maddock [13] offer a solution technique to produce a consistent [**A**] and [**B**] matrix given a monotone sample (*i.e.* when continuous records have different sarting times).

REVISED TEXT

Page 3, Section 1.3, Paragraph 1: delete "at any point in time"

Page 40, Section 4.1, Paragraph 3: Replace "rain" with "precipitation".

Page 41, Replace "1700's (or 1800's)" with "1700s (or 1800s") respectively &,
Replace "world war one" (or two) with "World War One" (or Two).
(Also occurs on page 46)

Page 42, Figure [4.1]: Reference, South Australian Engineering &
Water Supply Department, publicity material (*Water Supply System)*.

Page 44, Figure [4.2]: Reference [12]
Crawley P.D. & Dandy G.C. (1989)
Optimal Operating Policies for Multiple Reservoir Systems
(University of Adelaide - Civil Engineering Department Report)

Page 58, Section 5.6.1.1, Paragraph 5:
Replace, "Although will not occur ...",
with "This will not occur ...".

Page 82, Figure [5.8], "Yields" measured in (Ml).

Page 106, Section 7.2.1, Paragraph 1: Remove "in toto".

Figures [5.3] to [5.5], The horizontal axis has the non dimensional units of
"Number of Standard Deviations from the Mean".

Tables [5.9] to [5.13], "Absolute Error" units are (Ml) for use in Tables.

Table [5.14], "Units of Yield" are in (Ml).

Chapter [2], The reference for the Air Passenger Data is -
Hyndman R.J. (1990) *PEST - User Manual*
(University of Melbourne)

Chapter [5], When referring to the "Warren" station, it has been incorrectly referred to as the "Warren River" station. The Warren station gauges the South Para River at the Warren Reservoir. (Occurs on pages 60 & 62).

Appendix [D], Units for all plots -
Horizontal axis - Number of Standard Deviations from the Mean
Vertical axis - Observed Yields, (Ml).
- Transformed Yields, (Non dimensional).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Throughout history many advances have been made in what at the time was considered purely theoretical mathematics. This is true to an extent with stochastic data generation. Much of the theoretical groundwork had been laid many years before fruitful applications were realized.

This has been mainly due to the fact that the methods require large computational effort, and that the techniques simply were not required until the late 20^{th} century when our engineering systems have become large, complex and expensive. In today's economic climate, the operations of such systems may affect millions of lives and impact harshly upon the environment.

Now that it has been recognised that human society can no longer exploit what were regarded historically as being infinite resources, such engineering systems have come under scrutiny and inevitably have to be operated more efficiently.

Stochastic data generation is one method which may be utilized to aid in system operation. The method may be used for any system or process that can be measured through time. In fact it may be defined as —

"The analysis of a time series that behaves in a probabilistic manner"

The areas of application include:

- economics —
  - for stock control
  - production; for determining how much of a commodity should be produced,
  - for forecasting stock market prices.
- traffic engineering
  - for forecasting future traffic demands on roads and highways.
  - in analysing traffic behaviour at intersections.

and, of course, hydrology on which this study centres.

Over the last century most water supply systems in Australia have been monitored, with special attention to the measurement of rainfall and gauging of flow volumes in streams.

In the operation of any reservoir system for urban supply or irrigation, decisions must be made regarding releases, pumpages, the imposition of restrictions and the declared allocations (in the case of irrigation). Such decisions are usually based on the current storage levels in the system, the likely future inflows and the demands placed on the system.

If more reliable forecasts of inflows and demands are available, less conservative operating decisions can be implemented. For example, there will be less chance of imposing unnecessary (and politically unpopular) water restrictions, or the undertaking of expensive pumping programs.

To illustrate the above, research reported by Dandy [12] indicated that up to a 20% saving in pumping costs could be achieved for the Adelaide Headworks system if perfect forecasts of future inflows were available. Obviously perfect forecasts cannot be achieved due to the natural variability and unpredictability of rainfall and catchment conditions. However, this figure does give an indication of the potential savings which could be achieved by improved forecasting techniques. Most other metropolitan supplies in Australian cities do not involve as much pumping as Adelaide. However, a crucial operating decision in all systems is the balance of storage maintained between reservoirs to maximize system reliability. Improved inflow forecasting can aid in decisions of this kind and hence result in increased reliability of supply.

In irrigation areas an allocation of water for irrigation is announced at the start of the growing season. The allocation is based on the current state of storage in the reservoirs and estimates of the probability of receiving various levels of inflow over the growing season. If the uncertainty in inflow forecasts can be reduced, irrigators will receive a better indication of the likely availability of water, thus having a better chance to optimize crop patterns.

It can be seen, therefore, that improvement of inflow forecasting is likely to result in benefits to the users of all major urban and rural supply systems in Australia.

1.2 Study Objectives

The objectives of this study are —

- To evaluate the use of unisite and multisite time series models as a technique for forecasting the runoff from water supply catchments.
- To apply the technique to a set of data for the Adelaide Hills catchments and identify any problems in the technique.
- To illustrate the use of the forecasting models developed, as input for a model used for determining optimum operating policies for the Adelaide Headworks system.

1.3 Methodology

When attempting at any point in time to forecast runoff from a catchment on a monthly basis there is certain background information which may be used. This includes the runoff from that and adjacent catchments in previous time periods, previous rainfall and the state of the catchment *e.g.* the soil moisture index and water table levels.

In this study use will be made of multisite time series models of inflows. These utilize the serial correlation of flows at a single site as well as the spatial correlation between sites to forecast future inflows. Other readily available data such as rainfall is also considered if it increases the forecasting ability of the model.

Forecasts of inflows for operational purposes are usually required for (1) to (24) months ahead. This being the case, a monthly multisite model such as that of Young & Pisano [54] is appropriate. Such models have been used by

Burton [8] to model the major tributary inflows of the Murray River. For the unisite case, a periodic Thomas & Fiering [51] model will be used. Although the approach has general application, this study will demonstrate its use by using data from a set of catchments in the Adelaide metropolitan water supply district.

It is envisaged that the model(s) developed will be used to assist in making operational decisions in the following manner:

At the start of any particular month, historical streamflow data for the previous months can be used to initialize the model. A large number of possible future inflow sequences will be generated by the time series model using Monte Carlo simulation. The generated flow data can then be used to estimate the flow at each river which will be exceeded with a specified probability over the next one to twelve months as needed. These forecast flows can then be used to make rational operating decisions.

Chapter 2

Data Analysis

2.1 Introduction

The practice of stochastic data generation has, at its core, the development of a data generation model which will produce replicates of time series data which are equally as likely to occur in the future as the historical series.

It is not the purpose of a model to exactly replicate the historical data values, since this would defeat the purpose of producing many sets of "feasible", although distinct data sets. The use of synthetic data allows the hydrologist to "test" proposed works over many feasible data sets, thus providing an insight to the risk behaviour of the works.

The purpose of this chapter, is to outline the broad type of data analysis required to undertake stochastic data generation, and some of the further model testing used to ensure that the models are adequate.

When considering the use of stochastic data generation, the overall method of analysis must be borne in mind. Any analysis will follow the same step by step procedure as a whole but may diverge at some point to overcome some difficulty and then return later to the main procedure of analysis.

The approach to stochastic data generation is, in general, well laid out in terms of overall requirements, although many methods may be used at each step to perform certain tasks.

The generalized procedure is shown in Figure [2.1]. As with any engineering project, the same three basic steps are followed to achieve an efficient and comprehensive solution. These are —

- Definition of the underlying problem.

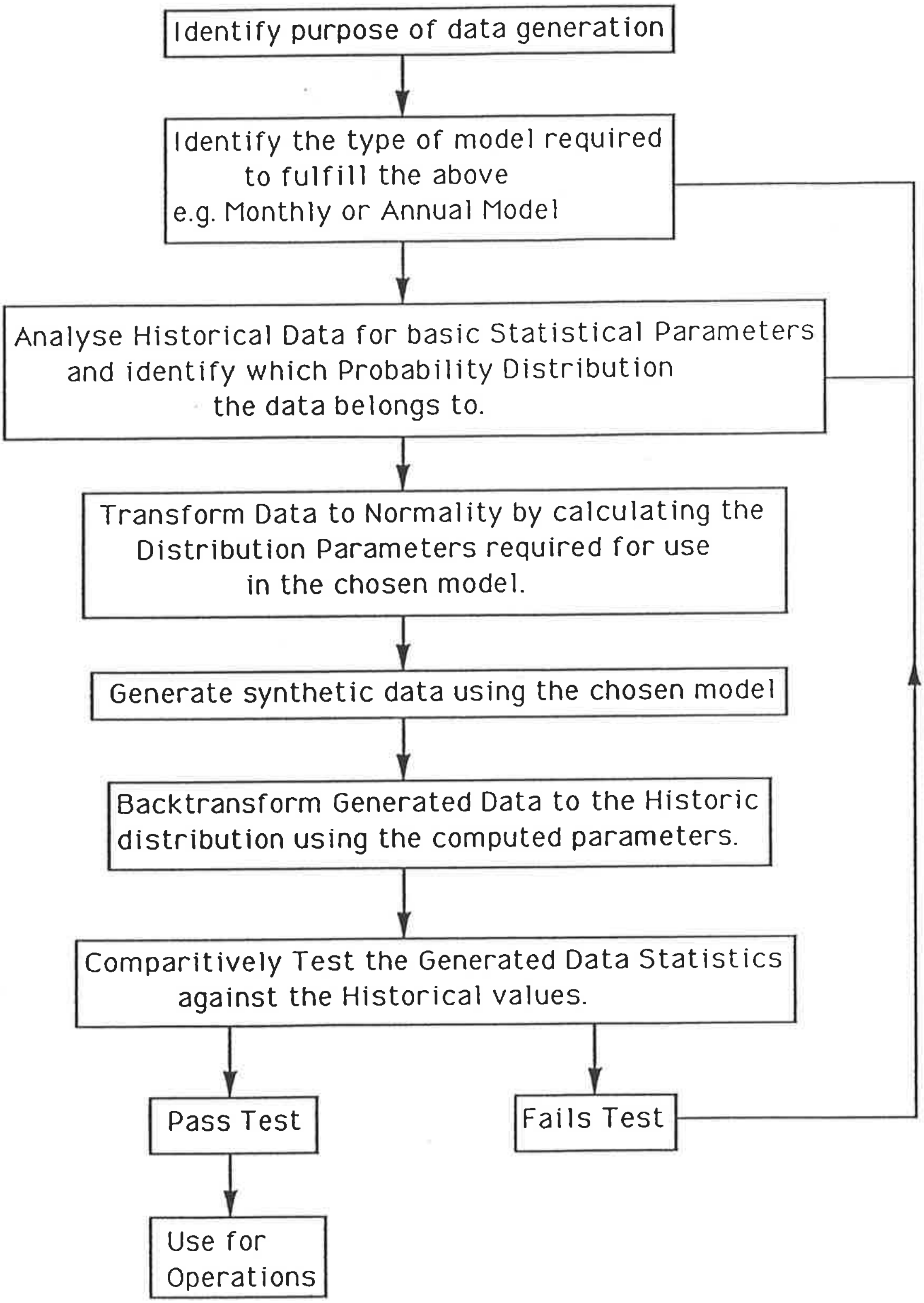

Figure 2.1: Stochastic Data Generation – Step Procedure

- Theoretical/Practical solution to the above.
- Implementation of the solution(s).

The following procedure outlines the broad analysis for stochastic data generation, but will be based around streamflow generation in particular.

For streamflow generation, the problem definition will be based, on the proposed use of the generated data *e.g.* for a proposed reservoir, theoretical testing of various sizes can be carried out to determine the cost/reliability/yield characteristics of each size. Or alternatively a periodic modelling process may be used to interface with given operating rules to analyse an existing system.

Given the statistical nature of the analysis, the relevant data is collated and analyzed on a preliminary basis. One of the most important steps in data generation is to identify the underlying probability distribution that the time series data belongs to. The reason for this being that the parameters used during data generation are usually based on a normal distribution. If the data is not "normal" then it will be transformed to normality by one of the methods outlined below.

Why the need for normality ? Statistical modelling involves the summation of terms in generation equations, and hence the summation of distributions. The normal distribution has the property that when a normal distribution is added to another normal distribution, then the result is also a normal distribution. This is not the case for most other distributions, although a special case of the gamma distribution also satisfies the property.

At the preliminary stage of the historical data analysis the same typical procedure is always followed. These steps are relatively straightforward and are not too time consuming.

Subsequent analysis will be dependent upon the outcome of results of the preliminary analysis and what is indicated with respect to the type of statistical distribution which best fits the data.

The steps of the data analysis phase of the research may be broadly outlined as follows –

- Determine the basic statistics and distribution of the raw data.
- On the basis of the above results, choose a theoretical distribution to fit the data.
- Calculate any parameters required, to transform the raw data to the chosen distribution.

- Calculate the statistics of the transformed data.
- Standardize the data if necessary given a chosen generation model.

2.2 General Analysis

In general a time series may be regarded as the combination of a set of distinct components.

For hydrologic data the components are usually identified as –

- a trend (t)
- a periodic or seasonal component (s)
- a deterministic or correlative component (d)
- a random component (ϵ)

Thus any value x in a data set will be the combination of the above –

$$x = t + s + d + \epsilon \tag{2.1}$$

To generate stochastic data it is necessary to identify each component and determine its relative significance.

The graph shown in Figure [2.2] represents a set of data known as the "Aircraft Passenger Numbers" for an airline. This data illustrates all of the above components, with a positive *trend*, a strong *seasonal* component, a regular *deterministic* value, and high frequency irregularities described by a *random* component.

Obviously if the observed values were used for analysis then the larger values due to trend or seasonality would bias the true shape of the underlying distribution and parameters, such as the mean. By identifying each component and its associated parameters, normally distributed generated values can be "moulded" to resemble the historical data set statistically.

With any stochastic data generation problem the first process after obtaining all useful data associated with the field of study under consideration, is to simply "eyeball" the data. This is best done graphically, by use of a commercially available package to show, for example, histograms or time series plots (such as Figure [2.2]).

Month to month serial scatter diagrams are also useful for the following purposes:

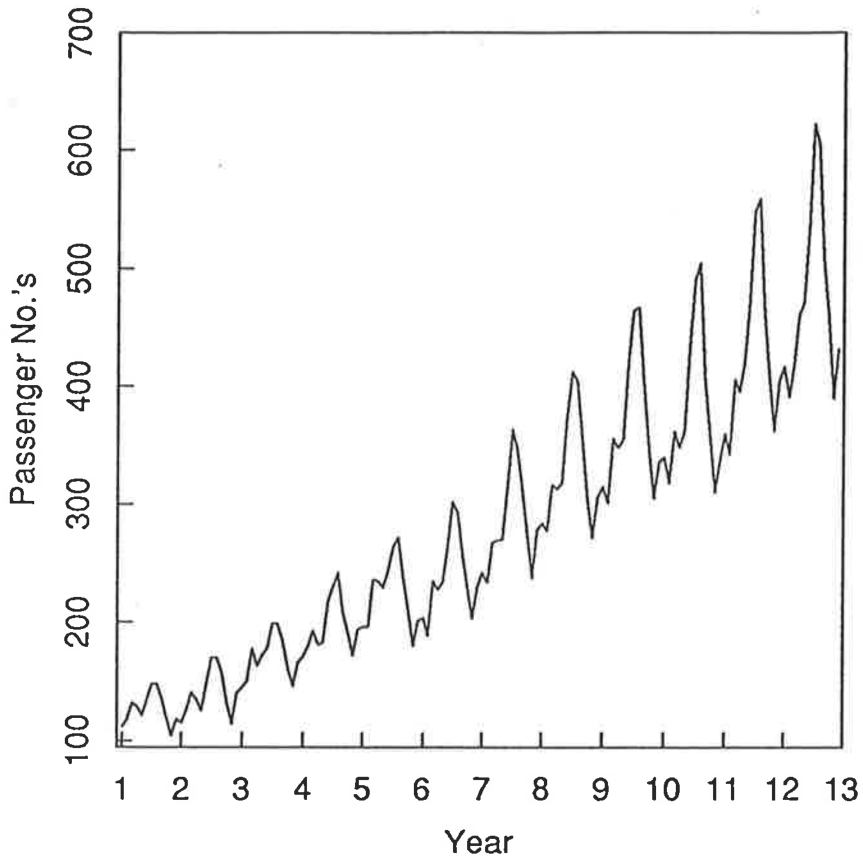

Figure 2.2: Aircraft Passenger Numbers – Raw Data

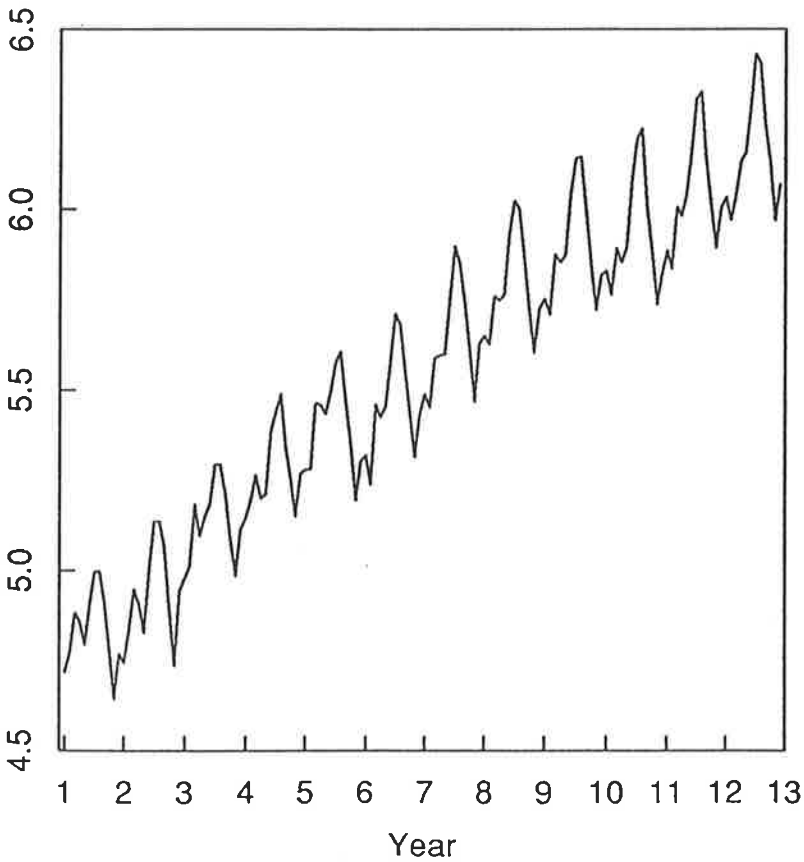

Figure 2.3: Aircraft Passenger Numbers – Transformed Data

To give a general idea of the lag one serial correlation of the series, and to indicate the presence of outliers.

From a computational standpoint it is useful to calculate the first three moments of a data set, namely –

- the mean $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
- the variance $\bar{V} = S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}$
- the coefficient of skewness $\bar{\gamma} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(nS^3)}$

where,

x_i	=	sample value.
n	=	number of samples.
$\bar{x}$	=	mean value of the sample.
S	=	unbiased sample, standard deviation.
$\bar{V}$	=	unbiased sample variance.
$\bar{\gamma}$	=	sample skewness.

The mean indicates the order of magnitude of the data, the standard deviation indicates the amount of relative spread about the mean, and the coefficient of skewness indicates the shape of the distribution. A positive skewness shows a longer tail of the distribution in the positive direction, and vice versa for a negative skewness. Zero skewness indicates a symmetrical distribution. The above graphical and computational results give the analyst a feel for the data as well as a preliminary insight to the type of distribution which will fit the data.

From the preliminary analysis it may be identified that the data is not described by a normal distribution. If this is the case, the data set is transformed from the distribution describing the raw process to the normal distribution by some technique, and is known as *normalizing* the data. Figure [2.3] shows the resultant data set based on Figure [2.2] after operating on the observed data by the *ln* function. This is one technique to transform log–normally distributed data to normally distributed data.

2.2.1 Trend

The analysis of any continuous random variate can only be attempted by using sampled data over a constant or variable time step. Either the instantaneous value of the variate at each time step is taken or the value is integrated over the time step. The latter is the case for hydrological data where the streamflow yield is the total volume of water yielded in the time step.

The process describing the variate may be changing with time, and as such, the distribution of the variate is said to be non–stationary. A time series is said to be stationary if its probabilistic behaviour is constant over time. *i.e.* the probability distribution of the series and its time structure does not vary through time. This is a very important point, since most of the analysis undertaken and derivation of models is based on the assumption of stationarity.

For some time series it may be apparent that some of the parameters of the distribution are changing through time. It is clear that the mean of the data shown in Figure [2.2] is increasing through time. For a hydrologic time series, this non stationarity may be due to changing land use functions, land management or global variations in the climate. One example is the widely publicized greenhouse effect, which refers to the global warming of the earth due to a build up of CO_2 and other gasses in the atmosphere.

Trend in a statistical parameter may be modelled in a number of ways, including a linear, exponential or power function of time.

2.2.2 Periodicity

Hydrologic or meteorologic data will usually possess a distinct cycle due to seasonal fluctuations in the climate.

This periodicity may be modelled in one of the following two ways –

- developing a periodic model in which the value of a variable is correlated with the corresponding value (p) time steps previously, where (p) is the period of the seasonal cycle.
- by removing the seasonal cycle using the following transformation.

$$y_{i,j} = \frac{x_{i,j} - \bar{x}_i}{s_i} \tag{2.2}$$

where,
$x_{i,j}$ = detrended data for season (i) and year (j).
$y_{i,j}$ = detrended & deseasonalized data for season (i) and year (j).
$\bar{x}_i$ = mean value for season (i).
s_i = standard deviation for season (i).

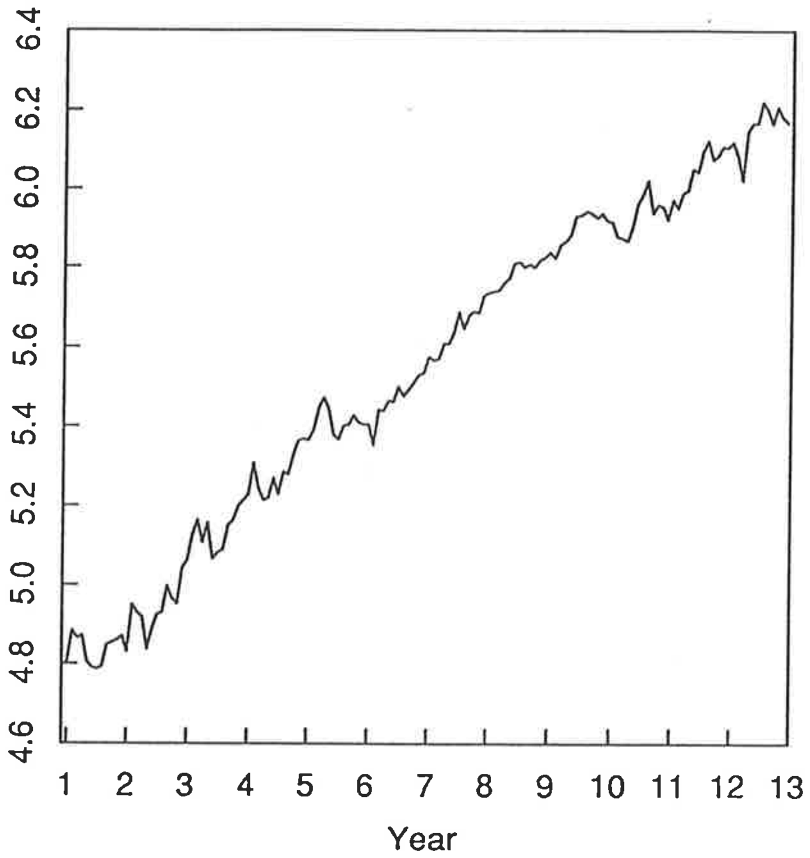

Figure 2.4: Air. Pass. – De-Seasonalized, Normalized Data

i.e.

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{i,j} \tag{2.3}$$

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{i,j} - \bar{x}_i)^2 \tag{2.4}$$

where n = the number of years in the sample.

Trend and periodicity is shown by example in Figure [2.3] which shows a clear seasonal component and trend embodied in the data, which will be removed once the process describing the trend and periodicity has been identified. Figures [2.4] and [2.5] show the deseasonalized and detrended data for the transformed aircraft passenger data shown in Figure [2.3].

2.3 Transformations

The process of transforming observed data from any given distribution to normality is an iterative one. Generally a distribution type (*e.g.* the gamma dis-

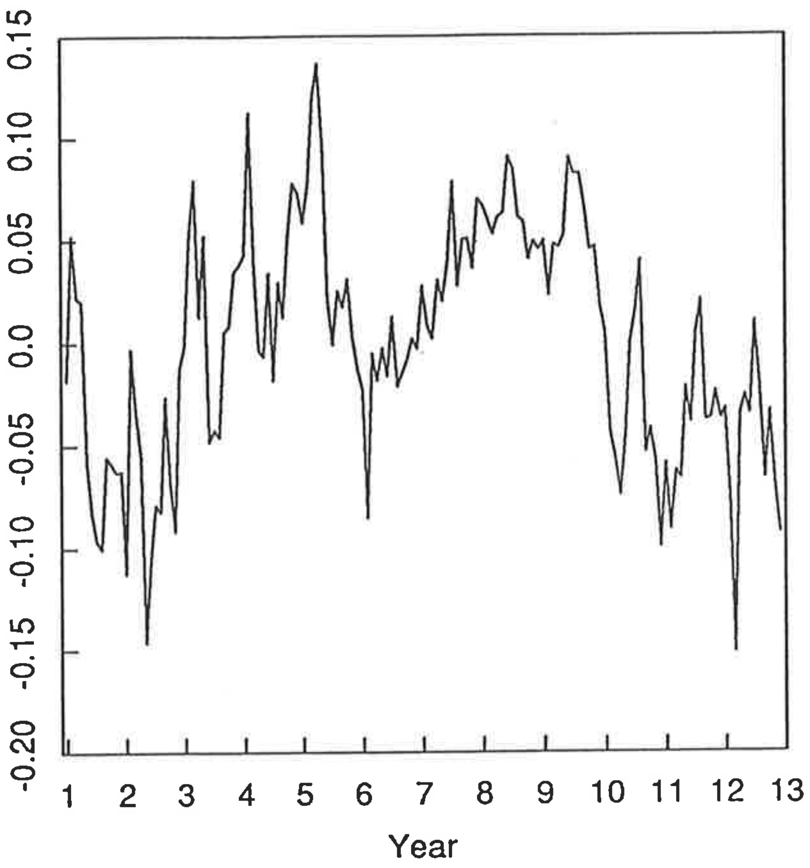

Figure 2.5: Air. Pass. – De–Trended, De–Seasonalized, Normalized Data

tribution), is selected and the assumption is checked subsequently. Therefore it is necessary to understand which distributions are most frequently encountered and the corresponding form of the transformations.

McMahon and Mein [36] list the following eight distributions that are frequently used with hydrologic data.

- Normal
- Log-Normal (3–parameter Log–Normal)
- Gamma
- Pearson type III
- Log-Pearson type III
- Kritzsky-Menkel
- Gumbel
- Weibull

The form of each distribution may be found in McMahon & Mein [36].

Each distribution has properties that describe a general shape. *e.g.* A log–normal distribution is defined only for positive values and the observed data is positively skewed.

By identifying the underlying distribution, not only is the form of transformation to normality known, but the probability of exceedance for a given value may also be calculated.

Studies of low flow hydrology frequently use a two-parameter gamma distribution, whereas for studies relating to continuous distributions of streamflow, a log-normal distribution is commonly found to be suitable.

The Pearson distributions follow from the gamma distribution and are in fact specialized cases of the gamma distribution. The gamma distribution involves both a shape and a scale parameter and by the addition of a location parameter the Pearson curves are derived.

Three commonly used methods for deriving the parameters of an assumed distribution to transform the data to normality, are outlined below.

These are —

- Parametric Transformations
- Moment Transformation Equations
- Maximum Likelihood

These are described in more detail below.

2.3.1 Parametric Transformations

This method applies when the parameters in the transformed domain are either known or assumed, and as such the observed data is transformed via these parameters and the resultant data set analysed by first principles.
e.g. If a variate is considered to be log normally distributed then all the data is transformed by taking the natural logarithm. Alternatively, the shape parameters of a gamma distribution may be known, and the observed data transformed to a new series given these parameters. The statistical properties of the new data set, such as mean and variance *etc* are derived in the usual manner using the equations given in section [2.2].

The following transformations are commonly used to produce normally distributed data. It is common to try several transformations, and use the one that produces the closest approximation to normality after transformation.

2.3.1.1 Log Transformation

The simplest transformation, and most frequently used for hydrologic data, is the log transformation. For this case it is hypothesised that by taking the natural logarithm of the observed data the resulting series will conform to a normal distribution.

$$i.e \qquad y_i = ln(x_i) \tag{2.5}$$

where, x_i $= i^{th}$ observed value
y_i $= i^{th}$ value after transformation

The series of y_i is then analysed as for the observed series x_i to determine if has been transformed to normality.

2.3.1.2 Shifted–Log Transformation

The natural extension of the above transformation, is to assume the data belongs to a three parameter log–normal distribution, and thus be transformed to normality using the shifted log transformation.

$$i.e \qquad y_i = ln(x_i - \tau) \tag{2.6}$$

where, x_i $= i^{th}$ observed value
y_i $= i^{th}$ value after transformation
τ = the location or shift parameter.

Clearly, τ must be less than the minimum value of x_i.

The value of τ may be determined so as to ensure zero skewness after transformation or using the method of maximum likelihood (refer Section 2.3.3).

2.3.1.3 Box–Cox Transformation

The Box–Cox transformation is of a power type defined as –

$$y_i = \frac{x_i^\lambda - 1}{\lambda} \quad for\ \lambda \neq 0 \tag{2.7}$$

$$y_i = ln(x_i) \quad for\ \lambda = 0 \tag{2.8}$$

The log–normal distribution is the result of a special case of the above, and it can be shown that –

$$as \quad \lambda \rightarrow 0 \quad \frac{x^\lambda - 1}{\lambda} \rightarrow ln(x) \tag{2.9}$$

If y_i has a normal distribution, x_i is said to have a "power normal" distribution for $\lambda \neq 0$. McMahon and Mein [36] calculated (λ) for seventeen Australian streams using annual streamflow data, and found the value of (λ) to range from (-0.26) to (0.70).

The value of (λ) has been determined explicitly by Chandler et al [10], but is most commonly found by choosing a value and iterating until the coefficient of skewness of the transformed data is as close as possible to zero.

Previous work by Burton [8] using data for the River Murray tributaries found that the Box–Cox transformation did not give superior results to using a shifted log transformation.

2.3.1.4 Wilson–Hilferty Transformation

The Wilson–Hilferty transformation is based on a "like–gamma" variate, and transforms skewed data to normality using Equation [2.10]

$$t_i = [\{(\frac{\gamma t_\gamma}{2} + 1)\}^{\frac{1}{3}} - 1 + \frac{\gamma^2}{36}]\frac{6}{\gamma} \tag{2.10}$$

where,
γ = coefficient of skewness of the raw data.
t_i = normal variate $N(0,1)$
t_γ = gamma–like variate with zero mean and unit variance.

The method is therefore –

- Standardize the observed values to produce (t_γ)
- Apply Equation [2.10] to produce normalized (t_i)

2.3.2 Moment Transformation Equations

Moment transformation equations relate the parameters used to describe a theoretical distribution to the parameters of the distribution in the transformed domain. Here the first (m) moments of a distribution are found and equated with the parameters in the transformed domain. Thus the parameters after transformation may be derived empirically from an analysis of the observed data rather than re–analysing a transformed data set.

In a paper by Matalas [37], the moment transformation equations for a three parameter log–normally distributed variate are reproduced from Aitchison & Brown [2] and are shown below.

$$\bar{x}_\tau = A_\tau + exp(0.5S_\tau^2 + \bar{X}_\tau) \tag{2.11}$$

$$s_\tau^2 = exp(2[S_\tau^2 + \bar{X}_\tau]) - exp(S_\tau^2 + 2\bar{X}_\tau) \tag{2.12}$$

$$\gamma_\tau = \frac{exp(3S_\tau^2) - 3exp(S_\tau^2) + 2}{[exp(S_\tau^2) - 1]^{\frac{3}{2}}} \tag{2.13}$$

$$r_\tau = \frac{exp(S_{\tau-1}S_\tau R_\tau) - 1}{\sqrt{exp(S_{\tau-1}^2) - 1}\sqrt{exp(S_\tau^2) - 1}} \tag{2.14}$$

where, $\bar{x}_\tau$ = sample mean in the raw domain.
s_τ = sample standard deviation in the raw domain.
γ_τ = sample coefficient of skewness in the raw domain.
r_τ = sample lag one serial correlation in the raw domain.

$\bar{X}_\tau$ = mean in the log domain.
S_τ = standard deviation in the log domain.
A_τ = location parameter for the 3–parameter log transformation.
R_τ = lag one serial correlation in the log domain.
τ = period under consideration.

Use of these equations ensures preservation of the first three moments of the data in the raw domain.

Further equations are used for multi–variate models to transform the lag zero and lag one cross correlations to the transformed domain.

These are given, for a time period τ as –

$$r_0^{p,q} = \frac{exp(S_p S_q R_0^{p,q}) - 1}{\sqrt{exp(S_p^2 - 1)}\sqrt{exp(S_q^2 - 1)}} \tag{2.15}$$

$$r_1^{p,q} = \frac{exp(S_p S_q R_1^{p,q}) - 1}{\sqrt{exp(S_p^2 - 1)}\sqrt{exp(S_q^2 - 1)}} \tag{2.16}$$

where, $r_0^{p,q}$ = the lag zero cross correlation between sites p & q, in the raw domain.

$r_1^{p,q}$ = the lag one cross correlation between sites p & q, in the raw domain.

$R_0^{p,q}$ = the lag zero cross correlation between sites p & q, in the transformed domain.

$R_1^{p,q}$ = the lag one cross correlation between sites p & q, in the transformed domain.

S_p = the standard deviation in the transformed domain at site p.

By inspection of the above equations it can be seen that S_τ may be found by solving Equation [2.13]. A_τ & $\bar{X}_\tau$ may only be found by solving Equations [2.11] and [2.12] iteratively.

Kite [26] rigorously analyzes the three parameter log–normal distribution and derives a set of independent equations which explicitly solve for the parameters in the log domain. Kite also produces equations to find the parameters by using the Method of Maximum Likelihood.

The "Kite"equations are given below for any time period τ.

Let (z_1&z_2) represent the coefficient of variation of the distributions [**X**] and [**X** − **A**] (respectively), then

$$z_1 = \frac{s}{\bar{x}} \tag{2.17}$$

$$z_2 = \frac{s}{\bar{x} - \mathbf{A}} \tag{2.18}$$

(**A**) is then given explicitly as –

$$\mathbf{A} = \bar{x}(1 - \frac{z_1}{z_2}) = \bar{x} - \frac{s}{z_2} \tag{2.19}$$

(z_1) can be computed from the raw data.
(z_2) is found by solving the following equation —

$$z_2 = \frac{1 - w^{\frac{2}{3}}}{w^{\frac{1}{3}}} \tag{2.20}$$

where

$$w = \frac{-\gamma + (\gamma^2 + 4)^{\frac{1}{2}}}{2} \tag{2.21}$$

It may also be shown that –

$$S = [ln(z_2^2 + 1)]^{\frac{1}{2}} \tag{2.22}$$

$$\bar{X} = ln(\frac{s}{z_2}) - \frac{1}{2}ln(z_2^2 + 1) \tag{2.23}$$

where, s = the standard deviation of observed values.
$\bar{x}$ = the mean of observed values.
γ = the coefficient of skewness of the observed values.
S = the standard deviation in the transformed domain.
$\bar{X}$ = the mean in the transformed domain.
A = the location parameter.

The solution technique, for any month (τ), is to first estimate $\bar{x}$, s and γ from the raw sample, then find w and z_2 using Equations [2.20] and [2.21]. Finally, solve for A, S & $\bar{X}$ using Equations [2.19], [2.22] & [2.23].

2.3.3 Maximum Likelihood

The method of Maximum Likelihood can be used to estimate the parameters of a distribution so as to give the best fit of that distribution to a set of observed data. The method produces asymptotically unbiased parameter estimates which have the smallest possible variance of any asymptotically unbiased estimator (Loucks et al [35]).

The drawback with the maximum likelihood approach is that it will not necessarily produce parameter estimates for all sets of data. The method may be described as follows –

Assume that a set of independent observations $(x_1, \cdots, x_n)$ have been made of a continuous random variable (X). The likelihood of making these observations given an assumed probability density function (pdf) for (X) is defined as follows –

$$L(x_1, \cdots, x_n|\underline{\Theta}) = f_x(x_1|\underline{\Theta}) \bullet f_x(x_2|\underline{\Theta}) \cdots \bullet f_x(x_n|\underline{\Theta}) \tag{2.24}$$

where $L(x_1, \cdots, x_n|\underline{\Theta})$ is the likelihood of making the observations $(x_1, \cdots, x_n)$ given the (pdf) of (X) has the parameter set $\underline{\Theta}$ and $f_x(x|\Theta)$ is the (pdf) of (X) for a given parameter set $\underline{\Theta}$.

As the observations are assumed to be independent, the probability of observing all of them is proportional to the product of the individual (pdf)'s. The maximum likelihood estimate of $\underline{\Theta}$ is the value that maximizes $L(x_1, \cdots, x_n|\underline{\Theta})$.

For example, consider a random variable (X) which is considered to have a three parameter log–normal distribution. Its (pdf) is given by —

$$f_x(x) = \frac{1}{x\sqrt{2\pi\sigma_y^2}} exp[\frac{-1}{2\sigma_y^2}(ln(x-\tau)-\mu_y)^2] \tag{2.25}$$

where (μ_y) and (σ_y^2) are the mean and variance of $ln(x)$ and (τ) is the location parameter. The likelihood function is given by –

$$L(x_1, \cdots, x_n|\mu_y, \sigma_y, \tau) = \Pi_{i=1}^n f_x(x_i|\mu_y, \sigma_y, \tau) \tag{2.26}$$

In this case it is easier to maximize the logarithm of the likelihood function.

$$i.e. \;\; ln(L) = ln\{\Pi_{i=1}^n f_x(x_i|\mu_y, \sigma_y, \tau)\} \tag{2.27}$$

$$ln(L) = \sum_{i=1}^{n} ln\{f_x(x_i|\mu_y, \sigma_y, \tau)\} \tag{2.28}$$

$$= -\sum_{i=1}^{n} ln(x_i\sqrt{2\pi}) - n\{ln(\sigma_y)\} - \frac{1}{2\sigma_y^2}\sum_{i=1}^{n}[ln(x_i-\tau)-\mu_y]^2 \tag{2.29}$$

to find the maximum of $ln(L)$, the partial derivatives of $ln(L)$ with respect to μ_y,σ_y and τ are found and set equal to zero. $i.e$

$$\frac{\partial lnL}{\partial \mu_y} = \frac{1}{\sigma_y^2}\sum_{i=1}^{n}[ln(x_i-\tau)-\mu_y] = 0 \tag{2.30}$$

$$\frac{\partial lnL}{\partial \sigma_y} = -\frac{n}{\sigma_y} + \frac{2}{\sigma_y^3}\sum_{i=1}^{n}[ln(x_i-\tau)-\mu_y]^2 = 0 \tag{2.31}$$

$$\frac{\partial lnL}{\partial \tau} = \frac{1}{\sigma_y^2}\sum_{i=1}^{n}[\frac{ln(x_i-\tau)-\mu_y}{(x_i-\tau)}] = 0 \tag{2.32}$$

From which

$$\mu_y = \frac{1}{n}\sum_{i=1}^{n} ln(x_i-\tau) \tag{2.33}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n}[ln(x_i-\tau)-\mu_y]^2 \tag{2.34}$$

τ may be found by substituting Equation [2.33] in Equation [2.32]. μ_y & σ_y are then given by Equations [2.33] & [2.34].

2.4 Tests for Normality

In Section [2.1] it was noted that it is usually convenient for modelling purposes for the generated data to be normally distributed with zero mean and unit variance (*i.e.* $N(0,1)$). This can be achieved by applying a suitable transformation to the raw data. Before transforming the data, the underlying distribution of the data needs to be determined.

No one method or test is available to determine the distribution of the sample data explicitly. One common process is that a distribution is assumed, the data is transformed and the assumption checked by a relevant test of normality on the transformed data. This raises two further questions; firstly, which distribution to try, and secondly, which test to use.

The question of distribution type can be found in previous literature or experience. For hydrologic data, the log–normal, gamma or log Pearson distributions have frequently been found to provide reasonable results.

The question of testing is more complex. One of the simplest and most widely used testing method is to determine if the coefficient of skewness of the transformed data is significantly different from zero. This is based on the symmetry property of the normal distribution.

Many authors in this field when publishing work based on actual data, present the work with the transformation type assumed or give details of already well known transformations applied to their data, with little or no rigorous testing. In fact little work has been compiled into the testing of distribution type.

Three methods were used in this study –

- Testing that the skewness is not significantly different from zero.
- The Shapiro–Wilk test for normality.
- Use of Quantile–Quantile plots for each data set.

Initially most emphasis was placed on the first method, with the remaining tests used to support the assumption of distribution type and subsequent transformation.

2.4.1 Skewness Test for Normality

In this method, the coefficient of skewness is calculated for the transformed data and then tested using the standard error of estimate ($S.E.E.$).

The equation given in Matalas [37] for the $(S.E.E.)$ on the coefficient of skewness is given as –

$$S.E.E. = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}} \qquad (2.35)$$

where (N) is the number of observations.

The significance level chosen in this study was 5%. Assuming the sampled coefficient of skewness is approximately normally distributed, it is not considered to be significantly different from zero if it lies within 1.96 standard errors from zero . The null hypothesis of normality can therefore not be rejected.

2.4.2 Shapiro–Wilk Test for Normality

The Shapiro–Wilk Test is shown by Pearson [41] to provide the best test for departure from normality.

This quantitative test attempts to weight the observed order statistics against the corresponding normal order statistics. Shapiro and Wilk [47] give the weighting factors on the basis of the best linear unbiased estimate of standard deviation, given as –

$$\hat{\sigma} = \sum_{i=1}^{h} b_{i,n}(x_{n-i+1} - x_i) \qquad (2.36)$$

where, σ = population standard deviation.
$\hat{\sigma}$ = best estimate of the sample standard deviation.
h = $\frac{1}{2}(n) or \frac{1}{2}(n-1)$ according to whether (n) is even or odd.
x_i = i^{th} sample value.
$b_{i,n}$ = weighting factor for the i_{th} normal order statistic given a sample of size (n).

The following description of the test is based on Pearson [41].

If the observed order statistics, x_i are plotted against the corresponding expected normal order statistics, $\zeta(\frac{i}{n})$ then the best linear unbiased estimate of the slope of this regression line is, apart from a normalizing constant, the estimate $(\hat{\sigma})$ of the population (σ) given in equation [2.36].

The test statistic (W) is proportional to the ratio of the square of an estimate based on this slope to the usual mean square estimate, given as –

$$W = \frac{(\sum_{i=1}^{n} a_{i,n} x_i)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{2.37}$$

the coefficients $a_{i,n}$ are given in Table [15] of Pearson [41] and are the normalized coefficients $b_{i,n}$ of equation [2.36].

i.e $$\sum a_{i,n}^2 = 1$$

Note that $(a_{n-i+1,n} = -a_{i,n})$, so that the numerator of the W–ratio can be written as –

$$A^2 = \{\sum_{i=1}^{h} a_{i,n}(x_{n-i+1} - x_i)\}^2 \tag{2.38}$$

Once the W–statistic has been computed for a data set then Table [16] of Pearson may be used to test the significance.

The drawback of this method is that exact values for $a_{i,n}$ are only available for $n = 20$, and approximate values for up to $n = 50$.

2.4.3 Quantile–Quantile Plots

The third test, is the use of quantile–quantile plots. This method is graphical by nature and thereby involves a physical judgement rather than an empirical test. Here, the order statistics are plotted against the corresponding normal order statistics on normal probability paper. The resulting line of best fit through the points should be straight if the assumption of normality for the transformed data proves correct. Sample plots are shown in Appendix [D].

2.5 Robust Methods

One approach that, in principle, lends itself to the estimation of statistical parameters in a more complete manner, is by the use of a field of statistics known as *robust statistics.* The field has been largely unrecognized by hydrologists with most literature being found in statistical texts with the case studies and data used being derived from the fields of economics or medicine.

Robust statistics tries to overcome contamination problems within data sets being due either to gross errors (outliers) or discontinuities. *i.e.* if 19 out of 20 points lie on a straight line and the 20th point is far from that line, then a linear regression will weight the outlier with as much importance as all other 19 values. A robust analysis will more heavily weight the correlative values.

Hampel et al [20] states that robust statistics is concerned with the fact that many assumptions commonly made in statistics (such as normality, linearity or independence) are, at best approximations to reality. Any method of dealing with the above form of problem such as subjective rejection (of outliers) or any other formal rejection rule belongs to the field of robust statistics in a broad sense.

Hampel et al [20] defines robust statistics as –

> *A body of knowledge partly formalised into "theories of robustness" relating to deviations from idealised assumptions in statistics*

and outlines the following areas that Robust Statistics may try to answer.

- Is the data unanimous in its message, or do different parts of the data give different impressions.?
  In this case, what does the bulk of the data infer ?
- Which minorities behave differently and how ?
- What is the influence of different parts of the data on the final result ?
- Which data are of crucial importance, either for model choice or for the final results, and which should be examined with special care ?
- How many gross errors can be tolerated by the design ?

Huber [23] is accredited with developing some of the modern techniques for robust analysis. Three of which are as follows, but not expounded upon here.

- Minimax approach
- Capacities approach
- Influence functions approach

The reader may further investigate such methods by reading Huber's text.

Chapter 3

Model Analysis

3.1 Introduction

The use of stochastic data generation models has evolved from the early years of this century when Hazen in (1914) attempted to produce a synthesized series of streamflow data by concatenating the annual yields for fourteen streams in the U.S.A.. Sudler in (1927) extended this theory by not only concatenating a series of a given length but choosing at random, (via the shuffling of cards) each sample within the series and repeating the exercise a number of times until a desired length of record had been formed. The purpose of such models was to produce a longer data sequence than the original one. This was required for use as an operational tool to test proposed works.

Whenever contemplating the use of data generation techniques, the end result must always be borne in mind. The statistical properties required to be reproduced by a particular model will greatly influence the type of model chosen and the degree of complexity required.

In the last quarter century the techniques of stochastic data generation have come to the fore in hydrologic analyses, with much documentation on the data analysis involved, and types of models that may be applied. In the nineteen fifties Hurst extensively studied the Nile river and postulated the now well known Hurst phenomenon of increasing ranges within data sets as the length of records increase. In the sixties Matalas was credited with progressing data generation into a new era with work involving regionalizing of parameters and multivariate techniques. In the seventies and eighties the advent of progressively superior digital computers has allowed these techniques to flourish and be used as a matter of course.

This chapter will outline the common types of models used for data generation. These models may be used for any form of continuous variable, but subsequent discussion will be illustrated using streamflow data.

Note also, that the models assume, except where noted, that the data is normally distributed.

Since, not all synthetic generation requirements are identical, a number of models have been developed, or more generally extended from the earliest mathematical models. These requirements may be based on the time period used or the number and type of parameters that the hydrologist wishes to preserve in the generation process.

3.2 Univariate Models

Univariate models are concerned with the temporal characteristics of a single time series. The models try and describe such characteristics based on some serial correlation with an event that has occurred previously. These models form the foundation of stochastic data generation and embody all the principles necessary to further develop the theory to higher order cases.

3.2.1 Autoregressive Models

The simplest and most commonly used model is a first order autoregressive model known as a Markov model. Markov (1856–1922) was a Russian mathematician who postulated that the outcome of a trial is somehow related (or dependent) upon the previous trial(s).

Hydrologically this seems intuitive as a high monthly streamflow is more likely to be followed by another month of high streamflow, or a dry month followed by another dry month rather than a very wet month. This process describes the persistence in hydrologic data, and it is this persistence that forms the basis of stochastic data generation. In fact it is the prime characteristic to be preserved in the generated data.

The Markov model in its simplest form is given as –

$$y_t = \Phi y_{t-1} + \epsilon_t \tag{3.1}$$

where,
- y_t = generated value for time period (t).
- Φ = autoregressive parameter, estimated from the sampled data.
- y_{t-1} = known value in time period (t-1).
- ϵ_t = random normal variate.

In order to preserve the mean and standard deviation , as well as the time structure of the series, the above equation takes the following form –

$$y_t = \bar{y} + \rho_1(y_{t-1} - \bar{y}) + \sigma_y\sqrt{1 - \rho_1^2}(\epsilon_t) \qquad (3.2)$$

where, $\bar{y}$ = Mean of normalized values.
σ_y = standard deviation of normalized values.
ρ_1 = lag one autocorrelation coefficient of the normalized values.
ϵ_t = random normal variate.
y_t = value of generated series in time period (t).

The above model is known as a lag one Markov model or a lag one autoregressive model. This type of model may be applied to a univariate case of, for example, annual streamflow yields.

The intuitive approach developed above may further be extended to an autoregressive model of order (p), and is denoted as an AR(p) model.
The model is written in general as –

$$y_t = \bar{y} + \Phi_1(y_{t-1} - \bar{y}) + \cdots + \Phi_p(y_{t-p} - \bar{y}) + \epsilon_t \qquad (3.3)$$

Or alternatively as –

$$y_t = \bar{y} + \sum_{j=1}^{p} \Phi_j(y_{t-j} - \bar{y}) + \epsilon_t \qquad (3.4)$$

where, Φ_j = the j^{th} autoregressive coefficient.
with other parameters defined as per equation [3.2]

The coefficients (Φ_j) may be found either by the method of moments or by maximum likelihood. Salas et al [43] gives a more complete description of solving for the autoregressive coefficients (Φ_p) as well as step by step procedures for AR(p) models, both annual and multi–period.

Thomas and Fiering [51] further extend the above approach to introduce a seasonal component. For this case, the model parameters are updated on a periodic basis and the persistence is described by the serial correlation coefficient between periods in lieu of a constant autocorrelation.

A periodic AR(1) model is given as –

$$y_t = \bar{y}_t + b_t(y_{t-1} - \bar{y}_{t-1}) + \sigma_t\sqrt{1 - \rho_t^2}(\epsilon_t) \qquad (3.5)$$

where,

$$b_t = \rho_t \frac{\sigma_t}{\sigma_{t-1}} \tag{3.6}$$

where, y_t = generated value in time period (t).
$\bar{y}_t$ = mean of normalized values in time period (t).
σ_t = standard deviation of the normalised values in time period (t).
ρ_t = lag one *serial* correlation coefficient between time periods (t) and (t-1).
ϵ_t = normal random variate in time period (t).
b_t = regression coefficient between time periods (t) & (t-1).

In the above type models it is possible to replace (ϵ_t) with a value that will introduce a skewness into the generated data similar to that of the historical data *i.e.* transform (ϵ) instead of (x_i).

For example, using the Wilson & Hilferty transformation, the procedure is as follows –

- Generate all (ϵ_t) values
- Apply the following equation to the above values to produce like-gamma values –

$$\epsilon_{\gamma,t} = \frac{2}{\gamma}(1 + \frac{\gamma\epsilon_t}{6} - \frac{\gamma^2}{36})^3 - \frac{2}{\gamma} \tag{3.7}$$

where, ϵ_t = normal random variate $N(0,1)$
$\epsilon_{\gamma,t}$ = like–gamma variate $G(0,1,\gamma)$
γ = the coefficient of skewness of $\epsilon_{\gamma,t}$

- Use the ($\epsilon_{\gamma,t}$) random variates in the model.

This can be used to produce a series $y_t(t = 1, \cdots, n)$ with a specified coefficient of skewness. McMahon & Mein [36] cite that the procedure breaks down for large values of skewness and autocorrelation.

McMahon & Mein [36] also suggests two methods of generating periodic data such that both the annual and monthly streamflow characteristics are preserved. These are —

- the two tier model — where monthly and annual data is generated with the monthly data proportioned to sum to the generated annual value.
- method of fragments — where only annual data is generated and then distributed to each month by choosing at random a "fragment" which is the fraction of monthly to annual yield for one of the observed yearly data sets.

Salas et al [43] and Box & Jenkins [6] extensively discuss properties and solutions of autoregressive models.

3.2.2 Autoregressive Moving Average Models – ARMA

The autoregressive models outlined above may be generalized to represent a wider range of time series by the inclusion of moving average terms.

A moving average model considers the magnitude of the stochastic component in the previous time step(s) in generating the next value.
If [**Y**] describes a normal variate and [**Z**] is defined as

$$z_i = \frac{y_i - \mu_y}{\sigma_y} \tag{3.8}$$

where, z_i = normalized and standardized sample values.
y_i = normalized sample values.
μ_y = mean of y_i.
σ_y = standard deviation of y_i

then [**Z**] may also be described as a series of weighted random variables –

$$z_t = \epsilon_t + \Theta_1 \epsilon_{t-1} + \Theta_2 \epsilon_{t-2} + \cdots \tag{3.9}$$

where, Θ_j = the j_{th} moving average parameter.

A moving average process of order (q) limits the above series to (q) weighted terms –

$$z_t = \epsilon_t - \Theta_1 \epsilon_{t-1} - \Theta_2 \epsilon_{t-2} - \cdots - \Theta_q \epsilon_{t-q} \tag{3.10}$$

$$z_t = \epsilon_t - \sum_{j=1}^{q} \Theta_j \epsilon_{t-j} \tag{3.11}$$

Combining equations [3.4] and [3.11] and using the standardized variate [**Z**] an ARMA model is fully described as –

$$z_t = \sum_{j=1}^{p} \Phi_j z_{t-j} - \sum_{j=1}^{q} \Theta_j \epsilon_{t-j} + \epsilon_t \tag{3.12}$$

for which $p + q + 2$ parameters must be evaluated from the observed data.

The reader is referred to Salas et al [43] and Box & Jenkins [6] for an extensive discussion of the estimation of parameters for ARMA(p,q) models as well as "goodness of fit" tests associated with these models.

3.2.3 ARIMA modelling

A model which is commonly referred to in the literature, and extensively noted by Box & Jenkins [6] is the autoregressive integrated moving average model. The ARIMA model is a more general case of the ARMA models outlined above. An ARIMA model is used when the observed data is found to be non–stationary.

A method of transforming a non–stationary series to a stationary series is by the use of *differencing*. This, simply stated is the transformation of the observed series by calculating the "difference" between adjacent observed values (d) times. Usually (d) only needs to be (1) or (2). Thus the series is said to be "integrated" and the transformed series is used for analysis in an ARMA model in the same way shown above.

3.3 Multivariate Models

The models outlined in section [3.2] are based on a single variable, and are applicable to systems that may adequately be described by a single process. In reality though, the design or ongoing operation of many real systems will be dependent on many components and will require a concurrent view of all components for decision making. For example, the Adelaide Metropolitan water supply system has ten reservoirs and associated catchments divided into two distinct distribution systems and augmented by three major pipelines from the Murray river. Such systems are geographically large and may involve several hydrologic and water use series.

For multivariate modelling, not only is the time dependent nature of a series preserved but also the spatial dependency between variates.

The principles and theory outlined for univariate analysis is directly applicable to multivariate analysis although an increased order of magnitude in effort is required to solve for the model parameters.

3.3.1 Multivariate Autoregressive Models

Two papers stand out in the literature as pioneering work in the field of multivariate stochastic data generation. Both papers were authored/co–authored by N.C. Matalas. In Matalas [5] a method is proposed where the statistical properties of a gauged streamflow site are used in conjunction with the generalised relationships of hydrological characteristics at an ungauged site. The theory for this multiple regression technique is not shown here as all sites for this study were gauged.

What is regarded as being the founding work for multivariate data generation is embodied in the second of the two papers by Matalas [37].

The technique outlined in this paper ensures that the means, standard deviations, lag one serial correlations and lag zero cross correlations of the historical series are reproduced in the synthetic series.

From section [3.2.1] it is recalled that the Markov process or lag one autoregressive model is defined as –

$$y_t = \bar{y} + \rho_1(y_{t-1} - \bar{y}) + \sigma_y\sqrt{1 - \rho_1^2}\epsilon_t \tag{3.13}$$

where the parameters are defined as for equation [3.2].

For a multivariate case the cross correlations between historic events needs to be considered with the estimates $(\bar{y})$, (σ_y) and (ρ_1).

The simplest method of generating multivariate data is based on a weakly stationary generating process defined as –

$$[\mathbf{Z_t}] = [\mathbf{A}][\mathbf{Z_{t-1}}] + [\mathbf{B}][\epsilon_\mathrm{t}] \tag{3.14}$$

where,

$[\mathbf{Z_t}]$	=	an $(n * 1)$ vector of generated values in time period (t).
$[\mathbf{A}]$	=	$(n * n)$ matrix to preserve time & spatial characteristics of the data.
$[\mathbf{B}]$	=	$(n * n)$ matrix similar to $[\mathbf{A}]$.
$[\epsilon_\mathrm{t}]$	=	$(n * 1)$ vector of random $N(0, 1)$ values.
n	=	the number of stations considered in the model.

Note that a typical element of the vector $[\mathbf{Z_t}]$ is generated from an equation of the form

$$Z_i(t) = \sum_{j=1}^{n} a_{i,j} Z_j(t-1) + \sum_{j=1}^{n} b_{i,j}\epsilon_j(t) \tag{3.15}$$

which is similar to a regression equation with correlated residuals, where the elements of the $[\mathbf{A_j}]$ matrix, $a_{i,j}$ are the generalised least–square regression coefficients.

For the case of two stations, expansion of equation [3.14] shows the dependence of each component in the generated $[\mathbf{Z}]$ matrix on the elements of the correlative matrices, $[\mathbf{A}]$ & $[\mathbf{B}]$.

$$z_t^1 = a_{1,1} z_{t-1}^1 + a_{1,2} z_{t-1}^2 + b_{1,1}\epsilon_t^1 + b_{1,2}\epsilon_t^2 \tag{3.16}$$

$$z_t^2 = a_{2,1} z_{t-1}^1 + a_{2,2} z_{t-1}^2 + b_{2,1}\epsilon_t^1 + b_{2,2}\epsilon_t^2 \tag{3.17}$$

where, z_t^i = the normal, standardized generated value in time period (t) at site (i).
$a_{p,q}$ = the p^{th} row and q^{th} column, element of the $[\mathbf{A}]$ matrix.
$b_{p,q}$ = the p^{th} row and q^{th} column, element of the $[\mathbf{B}]$ matrix.
ϵ_t^i = the random generated value in time period (t)

The $[\mathbf{A}]$ and $[\mathbf{B}]$ matrices are estimated in a manner similar to the (Θ) or (Φ) coefficients in a univariate model, such that the temporal and spatial characteristics of the historical records are preserved in the generated data.

Given that $[\mathbf{Z}]$ is in standardized and normalized format, by postmultiplying both sides of equation [3.14] by $[\mathbf{Z_{t-1}}]^{\mathbf{T}}$ and taking the expected value, a solution for the $[\mathbf{A}]$ matrix is given as –

$$[\mathbf{M_1}] = [\mathbf{A}][\mathbf{M_0}] \tag{3.18}$$

Rearranging gives

$$[\mathbf{A}] = [\mathbf{M_1}][\mathbf{M_0}]^{-1} \tag{3.19}$$

where

$$[\mathbf{M_0}] = E\{[\mathbf{Z_t}][\mathbf{Z_t}]^{\mathbf{T}}\} \tag{3.20}$$

$[\mathbf{M_0}]$ is an $(n * n)$ matrix whose elements are the lag zero cross correlations for a site (p) with site (q).
and

$$\mathbf{M_1} = \mathbf{E}\{[\mathbf{Z_t}][\mathbf{Z_{t-1}}]^{\mathbf{T}}\} \tag{3.21}$$

$[\mathbf{M_1}]$ is an $(n * n)$ matrix whose elements are the lag one cross correlations for site (p) with site (q).

The $[\mathbf{M_0}]$ matrix is symmetrical about the leading diagonal, with leading diagonal elements equal to (1).

The $[\mathbf{M_1}]$ matrix is not necessarily symmetrical and has leading diagonal elments equal to the lag one serial correlation at site (p), $(p = 1, .., n)$.

By post multiplying equation [3.14] by $[\mathbf{Z_t}]^{\mathbf{T}}$ and taking expected values the solution for the $[\mathbf{B}]$ matrix is obtained.

$$[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = [\mathbf{M_0}] - [\mathbf{M_1}][\mathbf{M_0}]^{-1}[\mathbf{M_1}]^{\mathbf{T}} \tag{3.22}$$

3.3.2 Solution for the B Matrix

The solution of $[\mathbf{B}]$ given that $[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = [\mathbf{C}]$ is a symmetrical matrix, does not possess a unique solution. Two methods are available to provide a satisfactory solution for $[\mathbf{B}]$.

The first method uses a technique of upper triangulation and subsequently solves for the lower triangular components by use of algebraic equations, and is suggested by Matalas [37], based on a method by Harman [21]. The second method is based on a principal components approach, deriving an explicit matrix solution. Both methods are outlined below —

3.3.2.1 Solution of the B–Matrix by Upper Triangulation

The method adopted by this study, which is outlined below is taken from Kottegoda [29].

The method assumes that the $[\mathbf{B}]$ matrix is lower triangular.
If $[\mathbf{B}][\mathbf{B}]^{\mathbf{T}}$ and $[\mathbf{C}]$ are written in full –

$$[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = \begin{vmatrix} b_{1,1} & 0 & 0 & \cdots & 0 \\ b_{2,1} & b_{2,2} & 0 & \cdots & 0 \\ b_{3,1} & b_{3,2} & b_{3,3} & \cdots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ b_{n,1} & b_{n,2} & \cdots & \cdots & b_{n,n} \end{vmatrix} \begin{vmatrix} b_{1,1} & b_{2,1} & \cdots & \cdots & b_{n,1} \\ 0 & b_{2,2} & \cdots & \cdots & b_{n,2} \\ 0 & 0 & b_{3,3} & \cdots & b_{n,3} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & \cdots & b_{n,n} \end{vmatrix}$$

$$[\mathbf{C}] = \begin{vmatrix} c_{1,1} & c_{1,2} & \cdots & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & \cdots & c_{2,n} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ c_{n,1} & c_{n,2} & \cdots & \cdots & c_{n,n} \end{vmatrix}$$

The diagonal elements are derived by –

$$b_{1,1} = (c_{1,1})^{\frac{1}{2}} \tag{3.23}$$

$$b_{2,2} = (c_{2,2} - b_{2,1}^2)^{\frac{1}{2}} \tag{3.24}$$

In general, for the leading diagonal

$$b_{k,k} = (c_{k,k} - b_{k,k-1}^2 - b_{k,k-2}^2 - \cdots - b_{k,1}^2)^{\frac{1}{2}} \tag{3.25}$$

Also, for the lower diagonal elements

$$b_{k,1} = \frac{c_{k,1}}{b_{1,1}} \tag{3.26}$$

$$b_{k,2} = \frac{(c_{k,2} - b_{2,1}b_{k,1})}{b_{2,2}} \tag{3.27}$$

and, in general for all remaining elements

$$b_{k,j} = \frac{(c_{k,j} - b_{j,1}b_{k,1} - b_{j,2}b_{k,2} - \cdots - b_{j,j-1}b_{k,j-1})}{b_{j,j}} \tag{3.28}$$

Kottegoda [29] also derives the general solution to the (p^{th}) order autoregressive multivariate model.
In general

$$[\mathbf{Z_t}] = \sum_{j=1}^{p} [\mathbf{A_j}][\mathbf{Z_{t-j}}] + [\mathbf{B}][\epsilon_t] \tag{3.29}$$

Note the similarity between the univariate and multivariate cases in the overall form of the equation.

In general $[\mathbf{M_j}]$ represents the covariance matrix with elements corresponding to the lag (j) cross correlation between two sites (p) and (q).

By post multiplying equation [3.29] by $[\mathbf{Z_{t-i}}]^{\mathbf{T}}$ and taking expectations.

$$M_i = \sum_{j=1}^{p} A_j M_{i-j} \quad (for \ i = 1, 2, 3, \cdots, p) \tag{3.30}$$

The general simultaneous solution being given as –

$$[A_1, A_2, \cdots, A_p] = [M_1, M_2, \cdots, M_p] \begin{vmatrix} M_0 & M_1 & \cdots & M_{p-1} \\ M_1^T & M_0 & \cdots & M_{p-2} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ M_{p-1}^T & M_{p-2}^T & \cdots & M_0 \end{vmatrix}$$

Similarly by post multiplying equation [3.29] by $[Z_t]^T$ and taking expectations, then –

$$[\mathbf{B}][\mathbf{B}]^{\mathbf{T}} = M_0 - \sum_{j=1}^{p} A_j M_j^T \tag{3.31}$$

and [**B**] is found from the method shown for the AR(1) case above, for each time period (t)

Salas et al [43] derive the [**A**] and [**B**] matrices for the multivariate AR(1) and AR(2) cases, and extend the theory to an ARMA(p,q) model.

3.3.2.2 Solution of the B–Matrix by Principal Components

This second method has been utilized by Rodriguez & Bras [7] together with associated adjustments to define the [**B**] matrix if a solution cannot be found directly.

Rodriguez and Bras outline the method as follows —

We know that there are an infinite number of solutions for [**B**], since equation [3.22] is satisfied by any matrix of the form [**B**] · [**D**] where [**D**] is orthogonal, implying $[\mathbf{D}][\mathbf{D}]^{\mathbf{T}} = [\mathbf{I}]$, for any such [**D**].

$$i.e. \qquad [\mathbf{C}] = [\mathbf{B}][\mathbf{D}][\mathbf{D}]^{\mathbf{T}}[\mathbf{B}]^{\mathbf{T}} = [\mathbf{B}][\mathbf{B}]^{\mathbf{T}} \tag{3.32}$$

Now define a further matrix [**P**] as follows -

$$[\mathbf{P}] = [\mathbf{P_1} \cdots \mathbf{P_n}] \tag{3.33}$$

where $\mathbf{P_i}$ is the i^{th} eigenvector of matrix [**C**]

The matrix [**P**] is also orthogonal *i.e.* $[\mathbf{P}][\mathbf{P}]^{\mathbf{T}} = [\mathbf{I}]$.

Define $(e_1 \cdots e_n)$ as the eigenvalues of matrix [**C**], and using the properties of eigenvalues and eigenvectors, it follows that —

$$[\mathbf{C}][\mathbf{P_i}] = \mathbf{P_i e_i} \tag{3.34}$$

Define $[\mathbf{E}]$ as a diagonal matrix of eigenvalues, as follows —

$$[\mathbf{E}] = \begin{vmatrix} e_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & e_n \end{vmatrix}$$

Therefore Equation [3.34] may be given in matrix form as

$$[\mathbf{C}][\mathbf{P}] = [\mathbf{P}][\mathbf{E}] \tag{3.35}$$

thus

$$[\mathbf{C}] = [\mathbf{P}][\mathbf{E}][\mathbf{P}]^{-1} = [\mathbf{B}][\mathbf{B}]^{\mathbf{T}} \tag{3.36}$$

therefore

$$[\mathbf{B}] = [\mathbf{P}][\mathbf{E}]^{1/2} \tag{3.37}$$

where

$$[\mathbf{E}]^{1/2} = \begin{vmatrix} e_1^{1/2} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & e_n^{1/2} \end{vmatrix}$$

Rodriguez & Bras state that the above procedure is limited by the algorithms used in finding the eigenvalues and eigenvectors and that for large matrices such procedures may result in errors or instabilities.

Even using the methods outlined above it is still possible in practice to produce a non positive definite covariance matrix $[\mathbf{B}][\mathbf{B}]^{\mathbf{T}}$.

Rodriguez & Bras attribute this to data transformations or to numerical anomalies, especially if the $z(t)'s$ are highly correlated.

To overcome this problem a method outlined in Rodriguez & Bras [7] (and attributed to Mejia & Millan, 1974) should produce positive definite matrices as required. This method was not used in this study, but is shown for completeness.

A new $[\mathbf{B}][\mathbf{B}]^{\mathbf{T}}$ matrix is defined as follows —

$$[\mathbf{B}'][\mathbf{B}']^{\mathbf{T}} = [\mathbf{B}][\mathbf{B}]^{\mathbf{T}} + \lambda_{\mathbf{j}} \tag{3.38}$$

where,

$$\lambda_j = \begin{vmatrix} |\lambda| & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & |\lambda| \end{vmatrix} \tag{3.39}$$

and λ is the most negative eigenvalue of the original $[\mathbf{B}][\mathbf{B}]^{\mathbf{T}}$ matrix. The above is repeated until $[\mathbf{B}'][\mathbf{B}']^{\mathbf{T}}$ become positive definite, but obtaining λ_j from the last $[\mathbf{B}'][\mathbf{B}']^{\mathbf{T}}$ matrix.

Once a $[\mathbf{B}']$ matrix has been found the model equation is modified as follows —

$$Z(t) = \frac{1}{\sqrt{1 + \sum_{j=1}^{m} \lambda_j}} [AZ(t-1) + B'\epsilon(t)] \tag{3.40}$$

This new model preserves the mean and variance of the historical data, affecting only the correlation coefficients by the following factor —

$$\frac{1}{\sqrt{1 + \sum_{j=1}^{m} \lambda_j}} \tag{3.41}$$

The degree of change may be calculated from the above and determined to be significant or not.

3.4 Multiperiod, Multivariate Models

In the analysis of a water resource system there is a need to consider the multiple components and demands of the system, but there may also be operational decisions made over relatively short time periods. These decisions are based on the cyclic nature of the inputs superimposed on the demands. This frequently means that within year decisions need to be made and multiperiod models are required to aid in such decision making.

The AR(1) model with periodic parameters is defined as –

$$Z_\tau = A_\tau Z_{\tau-1} + B_\tau \epsilon_\tau \tag{3.42}$$

where, τ = the period in question.
Z_τ = normalized and standardized generated values in time period τ.
A_τ = the $(n * n)$ matrix of coefficients to preserve the temporal and spatial characteristics between time periods τ and $\tau - 1$. (Similarly for B_τ).
ϵ_τ = an $(n * 1)$ vector of $N(0,1)$ random variates.

The periodic matrix parameters were derived by Salas & Pegram [44] as follows –

$$A_\tau = M_{1,\tau} M_{0,\tau-1}^{-1} \tag{3.43}$$

$$B_\tau B_\tau^T = M_{0,\tau} - M_{1,\tau} M_{0,\tau-1}^{-1} M_{1,\tau}^T \tag{3.44}$$

$$M_{k,\tau} = \begin{vmatrix} r_{k,\tau}^{1,1} & r_{k,\tau}^{1,2} & \cdots & r_{k,\tau}^{1,n} \\ r_{k,\tau}^{2,1} & r_{k,\tau}^{2,2} & \cdots & r_{k,\tau}^{2,n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ r_{k,\tau}^{n,1} & r_{k,\tau}^{n,2} & \cdots & r_{k,\tau}^{n,n} \end{vmatrix}$$

where $(r_{k,\tau}^{i,j})$ is found by correlating the values $(z_{v,\tau}^i)$ and $(z_{v,\tau-k}^j)$ for a period (τ), for $v = 1 \cdots, N - k$.

Salas et al [43] also derive the parameters for a multiperiod, multivariate AR(2) model.

3.5 Forecasting

The forecasting of data is a natural progression of data generation. For this case the value(s) at some lead time (l) are required to be known within a certain probability or confidence of the actual value.

The model that best fits the data is still used and a forecast found by conditional expectations. Box & Jenkins show that the forecast which has the minimum mean square error is given by –

$$\hat{z}_t(l) = E[z(t+l)|z(t), z(t-1), \cdots] \tag{3.45}$$

i.e the expected value of z_t given the preceding values through infinite history. Such forecasts are of great interest operationally and for this research will be utilized within an optimization program used to minimize pumping costs for the Adelaide metropolitan water supply system.

Chatfield [11] outlines and compares different forecasting models, but all quantitative models are based to some degree on the Box–Jenkins ARIMA modelling forecasts.

The three different approaches to forecasting are

- Subjective – Using judgement, intuition or practical knowledge.
- Univariate – Based on past observations; fitting a model and then extrapolating (projection methods).
- Multivariate – Based on taking observations on other variables into account. Regression methods are of this type. Also known as *causal* or *projection* methods.

Before choosing a forecasting procedure, it is essential to consider how the forecast is to be used, what accuracy is required, how many variables are to be forecast, how much data is available and how much "lead" time is necessary.

Chapter 4

System Background

4.1 Background to the Adelaide Water Supply System

The climate of South Australia is unique not only to Australia, but also with respect to land masses along similar latitudes in the Northern Hemisphere. This is due to the extensive ocean areas and the absence of a broad land mass connecting the Antarctic with the tropical regions. Australia, in general does not receive the same weather extremes characteristic of the Northern Hemisphere.

The South Australian climate is described as hot, dry summers with relatively mild nights, and cool but not severe winters with most rainfall occurring during the months of May to August.

South Australia is by far the driest of the Australian states and Territories with just over 80% of the state receiving an average of less than 250 millimetres of rain annually. Over the southern half of South Australia the main source of rain is from showers associated with unstable moist westerly airstreams occurring fairly regularly during the winter months of June to August. The wettest part of the state is in the Mount Lofty Ranges, immediately east of the capital, Adelaide. The average annual rainfall for this area is approximately 1200 mm. The Mount Lofty Ranges encompasses almost all of the catchment area available for metropolitan water supply. The topography of the area has a low flat plain from the sea to the ranges of approximately 20 *km*, with the undulating and hilly uplands of the ranges, generally running parallel to the coast.

4.2 The Headworks System

Since Australian settlement in the late 1700's and the steady growth of population for South Australia in the 1800's the general area has undergone many land use changes and significant human impact. During early colonization the populace had access to only rain water tanks or carting from rivers.

The increase in population in the late 1800's saw the need to augment the primitive water supply techniques, and in 1860 the state's first reservoir (Thorndon Park) was commissioned. Gradually through the years, and due to the climate and topography, Adelaide has been required to develop a complex reservoir and distribution network to maintain an acceptable supply of water to its consumers, (approximately one million in 1990). With expanding technology and industry in post world war two, and the need for a reliable water supply, the catchment areas could not cope with the consumer demand, and the water supply system required augmentation further by pumping from the Murray River.

Today, the Murray River supplies on average approximately 40% of Adelaide metropolitan supply, and in 1982–83 the value was as high as 80%, indicating a high dependence on this source.

The system now consists of nine metropolitan reservoirs, two major pipelines from the Murray River, and a further reservoir and pipeline to the north used to supplement the system. Figure [4.1] shows the area under study.

Together with approximately 8000 km of mains, 120 storage tanks and 48 pumping stations, the system requires a combination of experience and technological input to safely continue use without restrictions to the consumers. The Adelaide metropolitan system can be conveniently grouped into two main systems composed of a total of four main catchment areas.

The systems are described as —

- the *Northern* system
- the *Southern* system

The four catchment areas are —

- South Para system — 228 km^2
- Torrens system — 347 km^2
- Onkaparinga system — 451 km^2

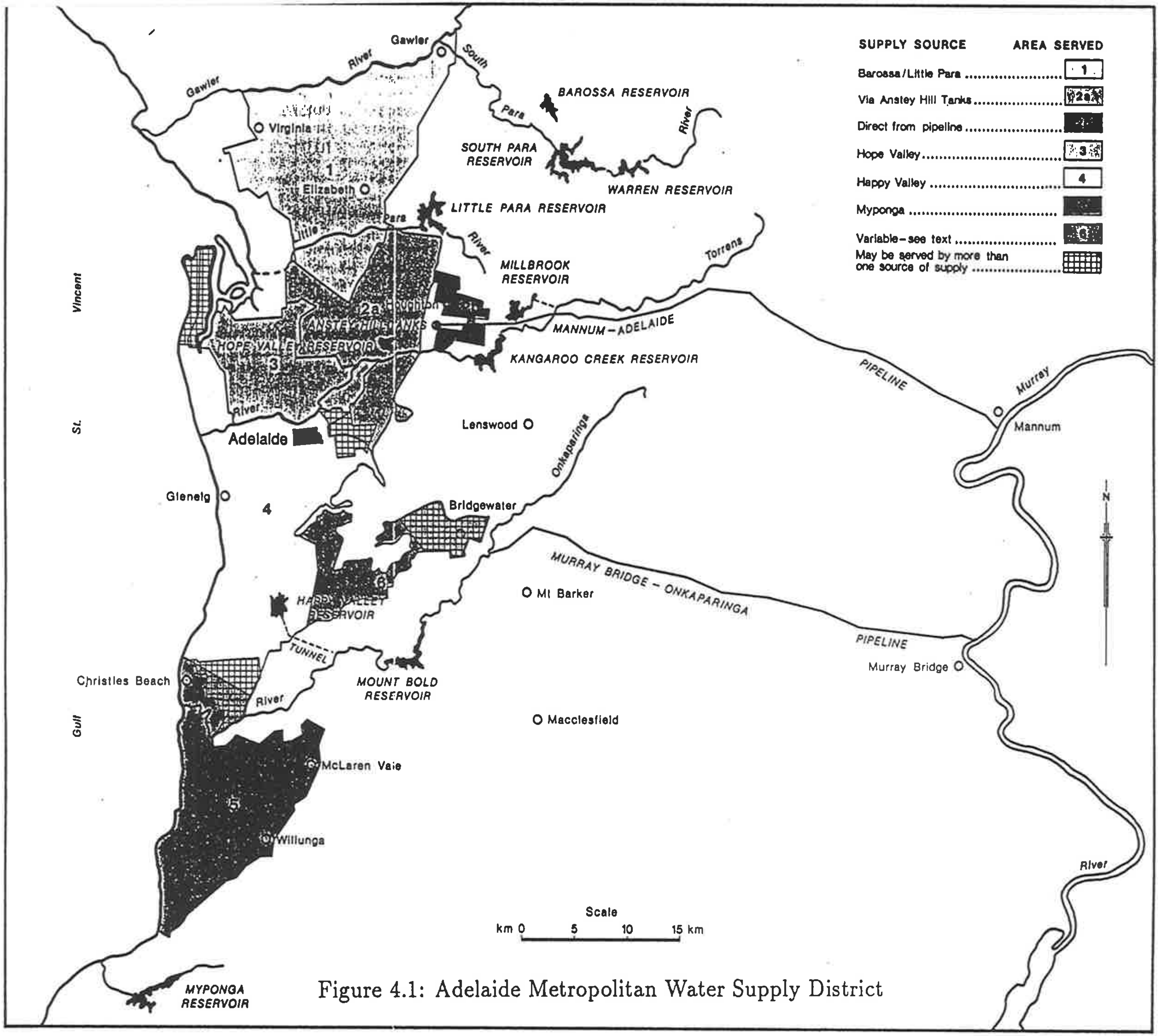

Figure 4.1: Adelaide Metropolitan Water Supply District

- Myponga system — 124 km^2

Figure [4.2] shows a schematic of the Adelaide Headworks System.

4.3 The Southern System

The Southern system consists of the Onkaparinga and Myponga catchment areas with three of the nine reservoirs and one pipeline, namely the Murray Bridge – Onkaparinga pipeline.

The reservoirs are —

- Mt. Bold Reservoir — This reservoir is an on–stream storage on the Onkaparinga River and has a catchment area of 388 km^2 and a capacity of 45.9 Gl.
- Myponga Reservoir — Situated on the Myponga River, it is another on–stream storage with a catchment size of 124 km^2 and a capacity of 26.8 Gl.
- Happy Valley Reservoir — This is an off–stream storage with no practical catchment associated with it and a capacity of 12.7 Gl.

Pipeline —

- Murray Bridge – Onkaparinga pipeline — The pipeline is 48 km in length and a (66″) MSCL pipe. The line transfers water from Murray Bridge to the southern system and discharges into the Onkaparinga River immediately south of the town of Verdun.

4.4 The Northern System

The Northern system consists of the South Para and Torrens catchment systems and contains the following reservoirs and pipelines. —

4.4.0.3 South Para system

- Warren Reservoir — The reservoir is constructed on the South Para River, has an associated catchment area of 119 km^2 and a capacity of

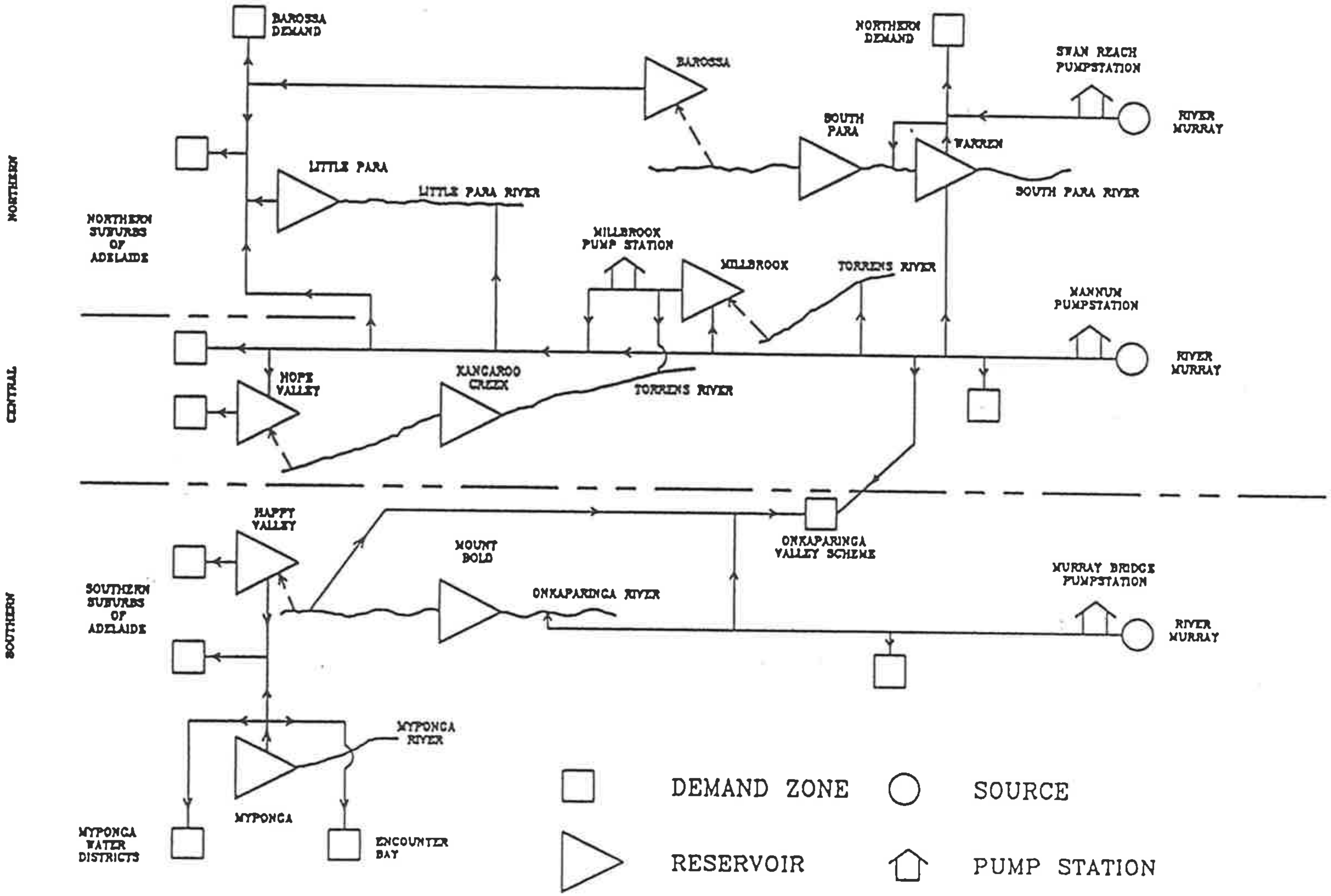

Figure 4.2: System Schematic

4.77 Gl. The reservoir is not strictly included in the metropolitan area and can be used to supplement the Northern system, although its major function is to supply the Barossa Valley region to the north of Adelaide and the Yorke Peninsula region.

- South Para Reservoir — The second of the on–stream storages built on the South Para River, it has an associated catchment of 109 km^2 and a capacity of 51.3 Gl.
- Barossa Reservoir — This reservoir is an off–stream storage, supplied by releases at the South Para reservoir. It to has no appreciable catchment area, but has a capacity of 4.51 Gl.

Pipeline —

- Swan–Reach Stockwell Pipeline — This pipeline is 54 km in length and transfers water from the Murray River at Swan Reach to the Warren Trunk Main. This is subsequently discharged to either the north of the state or to Warren reservoir.

4.4.1 Little Para Subsystem

- Little Para Reservoir — This reservoir is situated on the Little Para River and is the most recently constructed dam (1979) with a catchment area of 83 km^2 and a capacity of 20.8 Gl.

4.4.2 The Torrens System

- Millbrook Reservoir — This reservoir is an on–stream storage on the Torrens River with an associated catchment area of 233 km^2 and a capacity of 16.5 Gl.
- Kangaroo Creek Reservoir — This reservoir is downstream of Millbrook reservoir and is the second of the Torrens River on stream–storages. The catchment area is 55 km^2 and this figure does not include the Millbrook or upstream catchments. Capacity is surveyed as 19.0 Gl.
- Hope Valley Reservoir — Further downstream of Kangaroo Creek reservoir is the off–stream storage of Hope Valley reservoir. The catchment area associated with this reservoir alone is 57 km^2 and it has a capacity of 3.47 Gl. This reservoir is the site of Adelaide's first water filtration plant, commissioned in September 1977.

Pipeline —

- Mannum – Adelaide Pipeline — The pipeline was commissioned in 1954 and extends 60 *km* in length from the river town of Mannum to a terminal storage in the suburb of Modbury. The pipeline is not uniform in size over its length.

In South Australia streamflow gaugings generally began in systems where works were being constructed or feasibility studies made of proposed works. Thus the extent of streamflow records follows the gradual colonization of South Australia, starting with the Torrens catchment system in the late 1800's, Onkaparinga in the post world war one period, then the general South Para system in the post world war two period, and in recent history the Little Para subsystem in the late 1960's.

Chapter 5

Results

5.1 Introduction

In order to understand the difficulties associated with this form of stochastic analysis as well as the subtleties that may arise during a practical approach as opposed to a purely theoretical analysis, sets of hydrological data from Adelaide Hills catchments were used to apply the stochastic data generation models described previously.

This chapter outlines the methods adopted for data analysis and testing, as well as the models used for data generation.

5.2 The Raw Data

As cited in Chapter [1] this research uses the Adelaide Metropolitan Water Supply System as a case study. The study is based on hydrological data applicable to the region, namely streamflows and rainfall. As noted in the previous chapter, a monthly time step model has been adopted for operational considerations.

The raw data was supplied in two separate stages. Initially, seven streamflow data sets were supplied by the South Australian Engineering and Water Supply Department (E.&W.S.). These data sets consisted of the estimated natural monthly inflows into each of the respective catchments, expressed as a volume in (Ml).

Figure [5.1] shows the position of gauging stations.

These gauging sites are referred to as follows –

- Warren : G.S. 505 500
- South Para
  (at the Barossa diversion weir) : G.S. 505 501
- Little Para : G.S. 504 503
- Onkaparinga (At the Clarendon Weir) : G.S. 503 500
- Myponga : G.S. 502 501
- Torrens system (At the Gumeracha Weir) : G.S. 504 500
- Torrens system (At the Gorge Weir) : G.S. 504 501

Further to the above, seven rainfall data sets were supplied by the Bureau of Meteorology, Melbourne Office. These stations were chosen because they encompass all the metropolitan water catchment areas, as well as their proximity to the streamflow gauging stations.

In line with the above, these values represented the monthly total rainfall for a given station, expressed in units of *1/10th mm.*
Figure [5.1] shows the positions of the rainfall gauging stations.

These stations are –

- Thorndon Park : R.F. 023 027
- Clarendon P.O. : R.F. 023 710
- Millbrook Reservoir : R.F. 023 731
- Meadows : R.F. 023 730
- Myponga Reservoir : R.F. 023 738
- Mt. Bold Reservoir : R.F. 023 734
- Paracombe : R.F. 023 807

Table [5.1] gives the data set length, and period of record.

Fiering and Jackson [17] give some advice and quantitative guidance on the desirable length of historical records. In any case the hydrologist will always

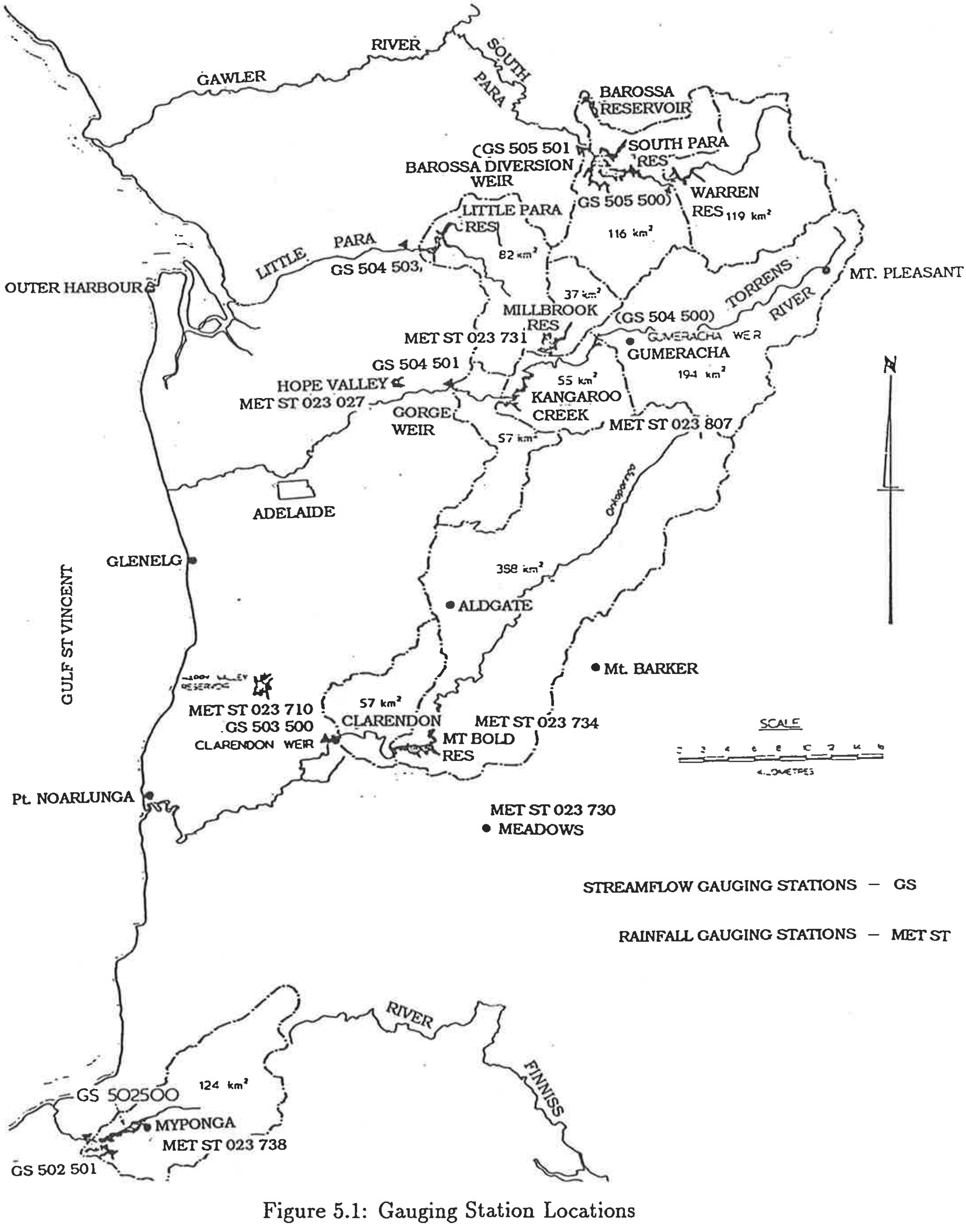

Figure 5.1: Gauging Station Locations

Streamflow Stations		
Station	Period	Length *yrs*
Warren	1939–1984	46
South Para	1939–1980	42
Myponga	1934–1984	51
Onkaparinga	1898–1984	87
Gorge	1884–1983	100
Gumeracha	1918–1983	66
Little Para	1969–1983	15
Rainfall Stations		
Thorndon	1879–1979	101
Clarendon PO	1875–1988	114
Mt. Bold Res	1939–1988	50
Millbrook Res	1914–1988	75
Meadows	1887–1988	102
Myponga Res.	1914–1983	70
Paracombe	1969–1988	20

Table 5.1: Station, Length & Period of Records

use the entire period of record available, and will only be able to use that period as a base. Thus the decision on length of record has, more generally than not, been set.

The length of record does however give an indication as to the expected confidence associated with the results.

5.3 Quality of Streamflow Data

It is commonly stated that stochastic data generation does not increase the amount of information available, nor increase the quality of historical data. This is not only a simple statement, but one that must always be remembered when reviewing results.

The streamflow yields supplied, although referred to as natural inflows to the catchments, are in reality reconstructed figures based on gauged values at the respective sites, and water balance equations composed of variables that ideally remove the human induced effects on the system. The water balance equations for some of the sites have up to twelve variables and include effects such as reservoir evaporation, changes in storage, volumes pumped from the Murray River *etc.*

As the individual components contain measurement and other errors, the yield data also contains errors which influence the results to an unknown extent. In many cases the water balance equations involve the difference of terms of similar magnitude. In such cases the errors are magnified as a percentage of the final estimated yield. The effect of the water balance equations is clear for some data, such as the South Para and Warren data sets where numerous problems resulted and these will be discussed below.

Two effects consistently lead to difficulties or infeasible solutions.–

- If, in the water balance equation the calculated inflow is negative, the value is truncated to zero. The logic being such, that a natural inflow cannot be negative. This has the effect of truncating the lower end of the distribution without giving due weighting to the magnitude of the calculated negative value.

- Secondly, it was apparent that potential outliers existed in the data sets. These unduly biased the historical statistics, especially the higher order moments of skewness and kurtosis.

  Although a correlation of single, large events (suspected outliers) with the rainfall data was attempted, no conclusive result could be drawn as whether these values should be deleted from the data set. This effect is more extensively discussed in Section [2.5] on robust statistics.

Almost all the streamflow data sets used contained some missing data, but generally less than five percent of values for any particular month were missing. These values were replaced by the mean values for the particular month determined from the remaining values.

To summarize the general quality of streamflow data, the Myponga data set was the least affected by errors in the water balance equation and as such is of good quality, the Onkaparinga and Torrens at Gorge and Gumeracha weirs are acceptable, although since the Gorge gauging includes the Gumeracha value, the two data sets are at times inconsistent. The South Para set includes the Warren values, which is to its detriment, as the Warren gauged data is clearly most affected by the truncation effect. Inclusion of Warren data with South Para data will only further contaminate the sample. These two data sets can only be described as poor and often lead to analysis problems due to large outliers and a grossly disproportionate number of zero yields throughout the year due to truncation.

The Little Para gauging station has only been recording since 1968, and was only compiled to 1984. This period of record is ideally too short for meaningful stochastic analysis but is required as an input into the optimization program developed for the Adelaide Metropolitan System for which the generated data will be used.

5.4 Quality of Rainfall Data

The rainfall records were received and analysed after preliminary analysis of the streamflow data. The analysis used was the same for both series. The seven data sets used were of high quality, consistently producing good results. No missing data existed in the files, nor did there seem to be any outliers.

Appendix [A] summarizes the statistical data only for the Clarendon P.O. rainfall gauging station since similar patterns were found throughout the other rainfall data sets.

5.5 Data Set Analysis

As cited in chapter [2] it is preferable to have normally distributed data for synthetic data generation. The raw data may belong to any one of a number of distributions. Using the raw data the parameters required to transform the data to a de-trended, de-seasonalized, zero mean, unit variance, normal distribution will be found.

Sections [2.3] & [2.4] described the types of distributions that may be encountered as well as the testing procedures considered. Figure [5.2] outlines the method of analysis adopted in this study. Initially the data sets were generally overviewed, missing data identified and basic statistics calculated. Throughout the study all parameters calculated refer to an individual month and site. Thus a total of $7(sites) * 12(months) = 84$ streamflow distributions

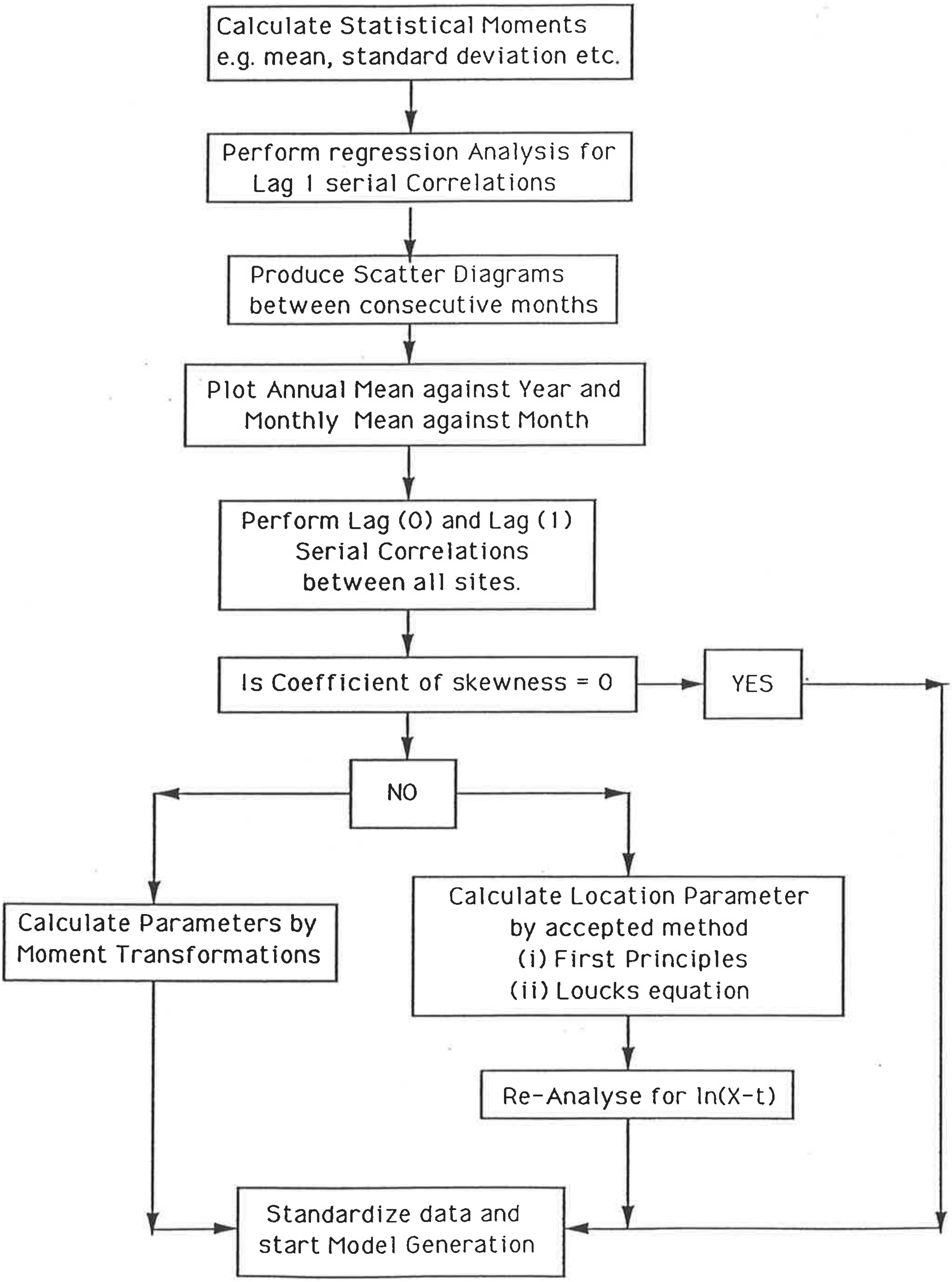

Figure 5.2: Step Procedure for Data Analysis

were considered rather than considering each gauged site as a single continuous data set.

To gain an appreciation for the type of data being used, the first four moments of the raw data series were computed, namely, the mean, standard deviation, coefficient of skewness and the coefficient of kurtosis. Initial data values were calculated on a personal computer with the use of the commercially packaged program "*Quattro*". For such analysis, as well as preliminary monthly serial scatter diagrams, this program was adequate, although limited.

In addition to the four statistical moments computed, the lag one serial correlation coefficients were calculated, giving an insight into the degree of persistence of the monthly data. The mean and standard deviation are required for use in a data generation model whereas the coefficient of skewness and coefficient of kurtosis were used to investigate if the data conforms to a normal distribution for which these values are zero and three respectively.

From the above results it was apparent that, in general, the data did not conform to a normal distribution for any monthly data set at a particular site. All gauged streamflow sites with the exception of Little Para produced highly skewed values and low monthly serial correlations. The Little Para site produced the largest serial correlations as was reflected in the monthly serial scatter diagrams, (an example of these are shown in Appendix [C] for the Myponga data set, together with the transformed values; see later). It is to be noted that the Little Para record is only fifteen complete years and as such has not been influenced by extreme events or cycles to the same extent as the other stations.

The statistics for the historical data at all sites are tabulated in Appendix [A].

In general the streams have high positive monthly skewness coefficients, in the order of (2.5) or more, with correspondingly high values of coefficient of variation, generally of the order of (1.0) to (1.5) but frequently these may rise to values in excess of (2.0). They thus have similar characterstics to other Australian streams. (McMahon & Mein [36])

Major streams in the Northern Hemisphere are characterised by low coefficients of variation, (generally less than 0.5) and very low coefficients of skewness by Australian standards, (mostly less than 0.5) with negative skewness common. The higher variability of streams under study (latitude $\sim 35°$ S) is also shown by McMahon to be reflected in world values where lower variation occurs in colder or tropical regions than in temperate climates.

Jacobs [25] extensively analysed the available records for the Adelaide Hills catchments, and obtained similar results to this study. The data sets used for each study originated from the same gauging stations yet differences occurred in basic statistics at some stations which shows the high variability of this regional data and the effect that specific values which have been included in one study but not the other can have on statistical moments.

In the analysis it was found that the use of higher moments to identify normality is prone to inconsistencies.

As

$$\text{the coefficient of skewness } \gamma_x = \frac{\sum_{i=1}^{n}(x_i-\mu)^3}{nS^3}$$

and

$$\text{the coefficient of kurtosis } \zeta_x = \frac{\sum_{i=1}^{n}(x_i-\mu)^4}{nS^4}$$

it can be seen that as $|x_i - \mu_x|$ becomes larger, it has an increased effect on the parameters. The parametric approach to statistics assumes that the data is free of outliers *i.e.* values which are not from the true population. If large outliers are a part of a series in question then they may have a disproportionate effect on the estimated parameters.

This, in fact, was the case for the data used in this study. Due to the large discontinuities and coefficients of variation (σ_x/μ_x) apparent in the data, it was suspected that some monthly data sets contained a small number of very large, highly suspect values. This can be seen from the lag one serial plots shown in Appendix [C] for the raw data, where, in some cases, the bulk of the data is concentrated in one corner of the plot with one or two values far from the centroid of the data.

An example of the effect on coefficients was found for the January monthly data for Little Para, where the summation of $(x-\mu)^3$ was almost entirely due to one of the fifteen values. Removal of the apparent outlier would result in a significantly lower coefficient of skewness.

Given the sizable effect of the above values on the skewness, the coefficient of kurtosis was affected by another order of magnitude. Due to this dominating effect on the kurtosis, it was subsequently deleted from the analysis. As this coefficient was not used beyond the initial stages it has not been tabulated with the other parameters.

Given that no trend has been reported in previous works published by the E.&W.S. Department of S.A. dealing with the data supplied, or in consultants' work undertaken using streamflow data from the Adelaide metropolitan area, no significant trend was expected within each monthly data set at any of the stations.

The method of trend analysis used was to produce a time series plot of mean annual values against time. Given the variability of the data and the length of records it was considered necessary only to identify if any trend in the mean was apparent and ignore potential changes in higher order moments.

The above plots were highly variable and did not indicate any definite trend. A *t-test* was performed on the annual mean values to test for statistical significance. The null hypothesis was that there is no significant trend.

The parameter tested was the *(m)* coefficient in the regression equation –

$$Y = mT + c \tag{5.1}$$

where, Y = annual yield.
T = time (years).
m, c = regression coefficients.

the t–value is defined as $m/Standard\,error$ which has $(N - 2)$ degrees of freedom. This t–value was tested at the 5 percent significance level for a two tailed t–test.

No significant trend was apparent in any of the data sets used.

5.6 Transformations

It was stated in Chapter [2] that hydrological data frequently possesses characteristics of a log–normal distribution. Physically this may be described by streamflow yields only taking on positive values, and the non–linear rainfall/runoff characteristics of a catchment.
Therefore the above analysis was repeated for the natural logarithms of all the data sets.

This transformation resulted in significantly lower skewness values, (in general below 0.5 in absolute value), and the monthly serial scatter diagrams (Appendix [C]) indicated a higher lag one serial correlation. Using the simple test for coefficient of skewness shown in section [2.4.1], a large number of monthly coefficient of skewness values, remained significantly different from zero.

Given the tendency to normality shown by the above transformation the analysis was extended to a three parameter log–transform.
The transformed values being defined as –

$$y_i = ln(x_i - \tau) \tag{5.2}$$

where, y_i = transformed value.
x_i = observed value.
τ = location parameter.

In the above equation, the location parameter (τ) is generally negative, resulting in a positive shift.

Two methods outlined in Sections [2.3.1] and [2.3.2] were used to identify the (τ) value for a given series.

- Parametric Transformation
- Moment Transformation

5.6.1 Parametric Transformation

The calculation of an appropriate shifting parameter (τ) by this method is an iterative one of choosing the distribution parameter values and subsequently testing the transformed sequence.
Two methods were used to choose an estimate of (τ) –

- an equation given by Loucks et al [35]
- systematic search.

5.6.1.1 Approximate Method using a Parametric Equation

The following equation for (τ) is given in Loucks et al [35]:

$$\tau = \frac{x_1 x_n - x_{0.5}^2}{x_1 + x_n - 2x_{0.5}} \tag{5.3}$$

where, x_1 = the minimum observed value.

$x_{0.5}$ = the median value.
x_n = the maximum observed value.

This method is very simple, requiring only three values to be sought from the observed series to define the shifting parameter.

Given the variability of the data used as mentioned above and the simplicity of Equation [5.3], the results were quite surprising. The method frequently produced transformed coefficient of skewness values less than (0.5) in absolute value, with small values (0.2–0.3) in the low flow months where small absolute changes in a shift can result in significant changes to the skewness.

This result is remarkable given the data used, since for some sites the value either side of the median, especially in the low flow months, can be significantly different.

The approach to the formulation of Equation [5.3] seems to be supported by a similar result derived by Sangal & Biswas [45], using only the mean, median and standard deviation of the observed data, given as –

$$\tau = x_{0.5} - \frac{\sigma_x^2}{2(\mu_x - x_{0.5})} \tag{5.4}$$

where, μ_x = mean of the sample
σ_x = standard deviation of sample.
$x_{0.5}$ = median of sample.
τ = location parameter.

By inspection of equation [5.3], it is evident that as $x_1 + x_n$ tends to $(2 * x_{0.5})$ the (τ) value tends to infinity. This will occur for data which has a symmetrical distribution *i.e.* a small coefficient of skewness. Although will not occur for a true log–normal distribution on which the derivation has been based, but was apparent for some of the actual monthly series used.

5.6.1.2 Systematic Search

A computer program "*TRANS*" was written to identify the (τ) value which produces zero skewness after transformation of a given series. The program simply uses a trial value of shifting parameter, starting at the minimum observed value of streamflow and transforms the data on the basis of this value using Equation [5.2]. The coefficient of skewness is then calculated and tested to determine whether it is within some predefined bounds of zero. In general the bounds used were (+/- 0.1), but where time allowed, bounds of (+/- 0.05) were used.

Note that these bounds are well within 1.96 standard errors of the estimate, which is generally around (0.6) for data sets of the size used. Of the seven streamflow sets, five behaved well and produced reasonable values of (τ) for all twelve monthly series. The Warren and South Para sites were troublesome, due to the high number of truncated values in many of the monthly series.

This individual method of parametric transformation is hereafter referred to as the method of *zero skew*.

5.6.2 Moment Transformation

The Moment Transformation Equations shown by Matalas [37] are given in Chapter [2] as Equations [2.11] to [2.14].

It is to be noted that these equations have no physical significance. The sole purpose of these equations is to produce parameters in the log domain which when used with generated normally distributed data preserve the statistics of the original data upon backtransformation.

The values obtained by this method for the shifting parameter (τ) were far greater than that required to produce zero skewness in the transformed data. Subsequently the resultant values of skewness after transformation are significantly different from zero and generally greater than the value calculated for the raw data.

By inspection of Equations [2.11] to [2.14] the transformed standard deviation is based only on the raw skewness of a series and is undefined for zero or negative skewness values.

5.7 Significance Testing

Significance testing was undertaken on the transformed data to test the transformed series for normality, when data sets were transformed by first principles. The tests used were –

- The Shapiro–Wilk Test
- Normal Probability Paper Quantile/Quantile plots (Q–Q plots)

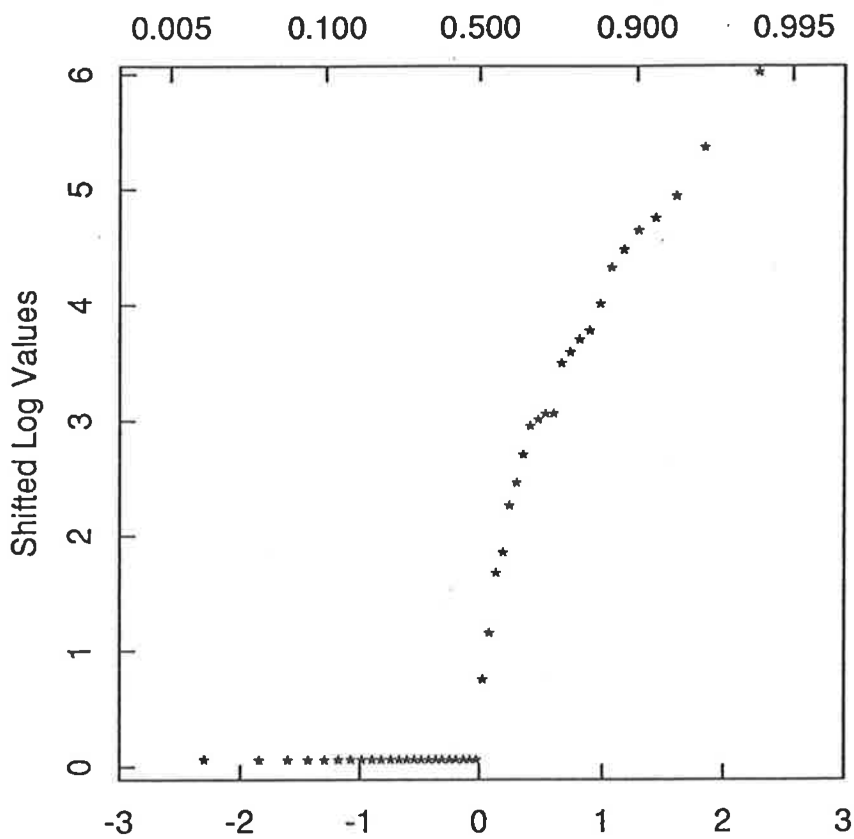

Figure 5.3: Q–Q plot, Transformed January values at Warren River

Although both of the above methods were used to test a transformed series, the Shapiro–Wilk test was only used to validate the result of the Quantile/Quantile plots, due to the restrictions of the Shapiro–Wilk test outlined in Section [2.4.2]. With the aid of the "S" statistically based computation/graphics package [42], the following quantile/quantile plots for each monthly series at each station were produced,

- Raw data
- Log Transformed data
- Shifted–log Transformed data
  (Based on a (τ) value found by systematic search)
- Shifted–log Transformed data
  (Based on a (τ) value derived from the moment transformation equations)

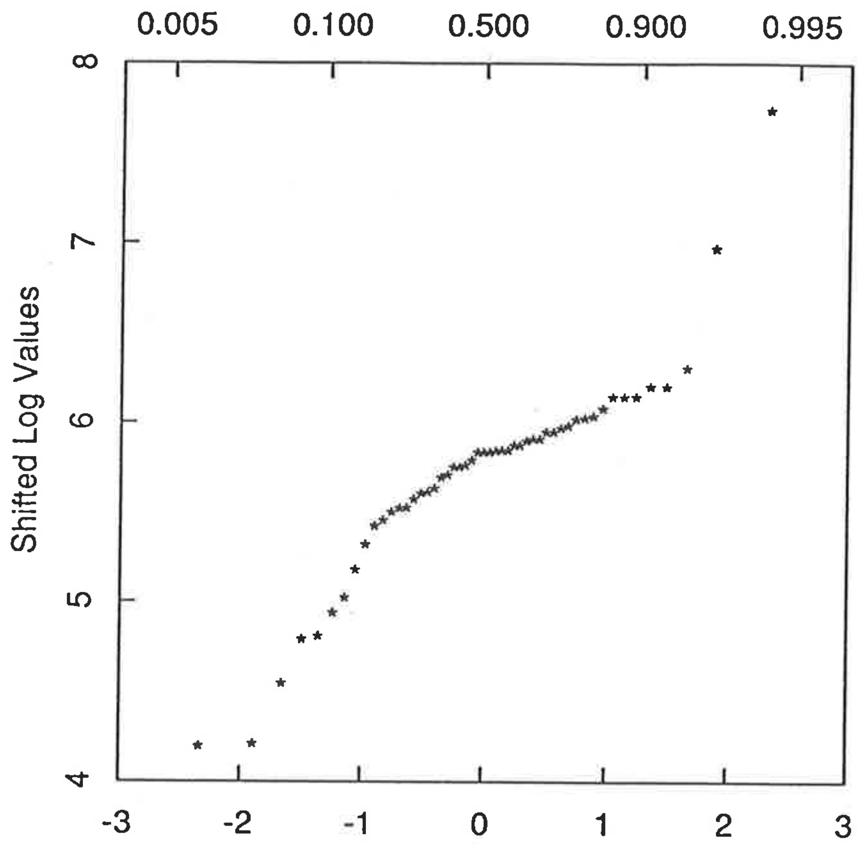

Figure 5.4: Q–Q plot, Transformed January values at Myponga River

A typical example is shown for two sites for the month of January in Figures [5.3] and [5.4].

For the plot based on the Warren station January series, it can be seen that the truncated values from the water balance equation have a significant effect on the tail of the distribution, since any undefined value for the transform $[(x_i - \tau) \leq 0]$ is set to zero. This result significantly affected the low flow months (November to April) at the Warren and South Para stations, and to a lesser extent, the Torrens at Gorge and Gumeracha sites, and the Onkaparinga at Clarendon site. These truncated values occurred too frequently to ignore for subsequent analysis.

In order to overcome this problem, the original records for all streamflow stations were reviewed and the truncated values replaced with the actual negative values calculated from the water balance equations. The reason for doing this was to identify the underlying distribution for each series on the assumption that the negative values were the result of consistent error in the data reconstruction process. The following stations required replacement of truncated values. –

- Warren

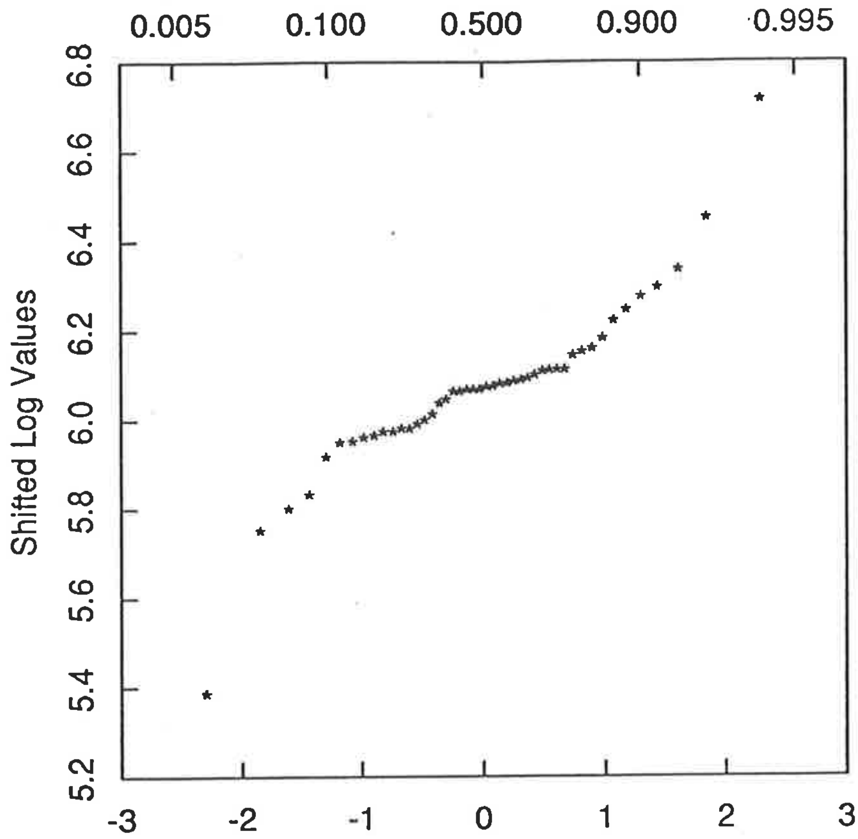

Figure 5.5: Q–Q plot, *Revised,* Transformed January values at Warren river

- South Para
- Torrens at Gorge
- Torrens at Gumeracha
- Onkaparinga at Clarendon

The summary statistics for the revised data sets are given in Appendix [B].

The shifting parameters (τ) were re–evaluated using the systematic search procedure and the quantile/quantile plots were recompiled. Figure [5.5] shows the revised series for the Warren station in January. Comparing figures [5.3] and [5.5] it can be seen that a vast improvement in the transformed series occurred. Subsequently, all series produced reasonable Q–Q plots.

For the Warren station, the monthly Shapiro–Wilk values are compared in Tables [5.2] & [5.3] for the two series for raw data and a shifted log transform. Values above 0.988 are significant at the 5% level *i.e.* normality may be assumed.

The above results indicate that the best fit to the underlying distribution

	Jan	Feb	Mar	Apr	May	Jun
Raw data – original record	0.5347	0.4790	0.6924	0.8921	0.4460	0.6256
Raw data – revised.	0.8776	0.7373	0.9142	1.0187	0.4564	0.6279
Trans. data – original record	0.8201	0.8805	0.8316	0.8789	0.9737	0.9698
Trans. data – revised.	0.9353	0.8568	0.8711	1.0239	1.0188	0.9044

Table 5.2: Monthly Shapiro–Wilk Values for Warren Data — January to June

	Jul	Aug	Sep	Oct	Nov	Dec
Raw data – original record	0.7778	0.8956	0.8680	0.6354	0.6840	0.7780
Raw data – revised.	0.7776	0.8956	0.8689	0.6376	0.7222	0.9575
Trans. data – original record	0.9540	0.9534	0.9481	0.9942	0.9586	0.8307
Trans. data – revised.	0.9548	0.9526	0.9527	1.0192	1.0090	1.0168

Table 5.3: Monthly Shapiro–Wilk Values for Warren Data — July to December

is obtained using a three parameter log normal distribution using the revised data sets. This indicates that the compilation of data at least reflects the pattern of temporal changes of streamflow, even if the absolute values are questionable.

5.8 Investigation of Fitting Methods

Extensive work on fitting two and three parameter log–normal distributions to hydrologic data has been carried out by Stedinger [49].
Stedinger writes — "the mean square error of estimation of selected quantities was used to evaluate the efficiency of alternative methods for fitting the two parameter and three parameter log normal distributions. Monte Carlo results show that use of a maximum likelihood parameter estimation dominates for fitting the 2–parameter log–normal distribution, for samples of 25 or more log–normal variates. For the 3–parameter, standard moment method performs best for log–normal distributions with low skew coefficients."

Stedinger goes on to say, that a good fitting procedure may be obtained by combining the moment or maximum likelihood methods already studied for the 2–parameter distribution with some technique which provides a reasonable estimate of (τ). Cohen (1957) essentially does this by combining the maximum likelihood estimates for the mean and standard deviation for a known (τ).

Stedinger gives a method for determining the location parameter by explicit solution, *i.e.* without the need to iterate.

It is also noted that the fitting technique may depend upon what the final result is required for, *e.g.* fitting the top end of the data or the bottom end.

The results of the above work qualify the findings and method used in this analysis, in that a known method was used to find the 3–parameter log–normal location parameter (τ), and then the other parameters found using Maximum Likelihood.

5.9 Summary of Tests Adopted

On the basis of the revised records, the transformation using a systematic search approach for each data set was based on the following –

- the computed τ value was not more than 3σ from μ

- The Coefficient of Skewness of the transformed series is within 1.96 standard errors of estimate.
- The straightness of the Q–Q plots.
- The shifting parameter τ does not truncate the observed series upon transformation.

Table [5.4] indicates the location parameters (τ) adopted.

A zero value indicates that the two–parameter log transformation sufficed and "*normal*" indicates that no transformation was required.

A complete tabulation for each of the data sets after transformation is given in Appendix [B].

5.10 Modelling

The above sections describe the methods and analysis required to extract all useful information from the data and to shape the data into a usable form for data generation.

The following stages of data generation are distinct from the data analysis phase, yet use the parameters identified above as input for a chosen model. The sections below outline the models chosen in this study, and use of the generated data with respect to the Adelaide Metropolitan Water Supply System.

In Chapter [3] it was shown that either a univariate or multivariate analysis could be used for data generation and that a multivariate analysis attempts to preserve the spatial correlation of the hydrological processes. Both univariate and multivariate modelling procedures were undertaken in this study, with each model analysis designed to produce the same performance parameters for comparison. These models are outlined below.

5.10.1 Multivariate Model

The sample $[\mathbf{M_0}]$ & $[\mathbf{M_1}]$ matrices are covariance matrices. Theoretically $[\mathbf{M_0}]$ & $[\mathbf{M_1}]$ should be positive semi–definite. Using practical streamflow data with the associated and inherent sampling errors together with mathematical manipulations that may lead to round–off error, it is possible for either or both of these matrices not to be positive semi–definite.

Final Shifting Parameters (τ) used												
	January	February	March	April	May	June	July	August	September	October	November	December
Warren	-425	-304	-10	normal	-67	-62	-184	-1094	-453	-77	-97	-140
South Para	-400	-342	normal	-26	-33	-53	-429	-712	-647	-44	-234	-325
Myponga	-64	-18	-265	-95	+166	-34	-323	-2503	+50	+23	-16	normal
Onkaparinga	0	-436	0	-84	-61	-322	-1367	-4815	-1095	-140	-292	-762
Gorge	-784	-956	0	-475	-377	+100	-588	-3068	-1570	-94	-242	-656
Gumeracha	-211	-166	-215	-1	+29	-102	-155	-710	-102	+21	-83	-69
Little Para	-78	-53	-14	-5	+4	+96	-482	-266	-327	-50	-66	-258
Clarendon	-42	-23	-66	-228	-650	-857	-1274	-510	-206	-338	-177	-131

Table 5.4: Location Parameters (τ) adopted

It has been noted by Kuczera [31], Crosby & Maddock [13] and Fiering [16] that if the records between variates are of differing lengths then problems may arise with the consistency of either the $[\mathbf{M_0}]$ matrix or the $[\mathbf{B}][\mathbf{B}]^\mathbf{T}$ matrix, thus not allowing a solution to either of the $[\mathbf{A}]$ or $[\mathbf{B}]$ matrices used in a multivariate generation equation.

If records of unequal length are used to estimate the lag zero and lag one covariance matrices $[\mathbf{M_0}]$ & $[\mathbf{M_1}]$, respectively, the covariance estimate $[\mathbf{C}]$ may not be positive definite, thereby preventing decomposition of $[\mathbf{C}]$. The problem may be overcome by truncating the larger records to the length of the shortest of the records at the expense of discarding useful information. If the missing data is due to the different records having different starting times then Crosby & Maddock [13] show how $[\mathbf{C}]$ can be made positive definite.

Crosby & Maddock [13] refer to the above entire sample as being monotone, *i.e* if we have (N) sets of continuous data, but they have differing start times then the sample is referred to as a *monotone sample.*

It may be that the existence of a monotone sample causes either (or both) of the $[\mathbf{M_0}]$ or $[\mathbf{B}][\mathbf{B}]^\mathbf{T}$ matrices not to be positive definite.

They further state that since the eigenvalues are variances in the principal component system, some of the variances are negative. A covariance matrix with negative eigenvalues is inconsistent. Even if $[\mathbf{M_0}]$ & $[\mathbf{M_1}]$ are consistent and are used to define the $[\mathbf{B}][\mathbf{B}]^\mathbf{T}$ matrix, the resultant may be inconsistent, therefore making it impossible to solve for a $[\mathbf{B}]$ matrix with all real values.

Fiering [16] & Beard [3] have both developed techniques for producing consistent estimates for the $[\mathbf{M_0}]$ matrix. In fact, both techniques can be used to produce a consistent $[\mathbf{M_0}]$ matrix when the data sets are not only monotone but have records missing in a non systematic way.

Crosby & Maddock show that neither Fiering's nor Beard's methods guarantee $[\mathbf{B}][\mathbf{B}]^\mathbf{T}$ to be consistent and go on to develop and apply their own method which not only produces a consistent $[\mathbf{M_0}]$ matrix, but a $[\mathbf{B}][\mathbf{B}]^\mathbf{T}$ matrix as well. Their method is based upon a maximum likelihood estimate developed by Anderson [1]. This method seems to be mathematically complex and difficult to apply.

The main thrust of this study was aimed at the production of a multivariate model for use as an operation tool. The model chosen is described in Chapter [3] as a multivariate, multiperiod autoregressive model of order (1), *i.e.* an AR(1) model.

The matrices associated with such a model were developed for an annual model by Matalas [37] and the derivation of matrices for the periodic case given by Salas et al [43].

This model is given by the equation [3.42]

$$[\mathbf{Z_t}] = [\mathbf{A_t}][\mathbf{Z_{t-1}}] + [\mathbf{B_t}][\epsilon_\mathbf{t}] \tag{5.5}$$

where $[\mathbf{A_t}]$ and $[\mathbf{B_t}]$ are given by equations [3.43] and [3.44], and for lag one are given by —

$$[A_t] = [M_{1,t}][M_{0,t-1}]^{-1} \tag{5.6}$$

and

$$[B_t][B_t]^T = [M_{0,t}] - [M_{1,t}][M_{0,t-1}]^{-1}[M_{1,t}]^T \tag{5.7}$$

$$\epsilon_t = \text{a vector of } N(0,1) \text{ random variates}$$

The computer program developed for this model is known as "*GENESIS*".

A typical analysis using this method involves taking the historical data sets of a given length, and firstly computing the historical statistics and then transforming the data to normality given a user–defined command.

The generation equation is given as equation [5.5] above, based on standardized, transformed values. These values must then be "shaped" to resemble the historical form by "backtransforming" the generated data, which is simply the reverse analysis of the data transformation sequence *i.e.* (for a three parameter log transformation)

$$y_t = \mu_t + \sigma_t z_t \tag{5.8}$$

$$x_t = e^{y_t} + \tau_t \tag{5.9}$$

where y_t = value in the log domain for time period (t).
x_t = value in the raw domain.
z_t = individual generated value.
τ_t = location parameter
μ_t = mean of y_t .
σ_t = standard deviation of y_t.

5.10.1.1 Solution of the "A" matrix

This section describes the problems encountered with solution of the $[\mathbf{A}]$ matrix when using data for six streamflow sites and one rainfall site. It was found that in this study a solution for either of the [**A**] or [**B**] matrices was not always defined.

The following two problems were encountered —

- the [$\mathbf{M_0}$] matrix is singular
- the [$\mathbf{M_0}$] matrix is close to singularity.

For a multiperiod, multisite model as adopted, there needs to be an [$\mathbf{A_t}$] and a [$\mathbf{B_t}$] matrix for each time period (t). This increases the probability of failure as the model will only require one of these matrices to be undefined to fail. Also, it can be shown mathematically, that as the number of stations increases, it becomes more likely that the [$\mathbf{M_{0,t-1}}$] matrix will be singular. This results in [$\mathbf{M_{0,t-1}}$]$^{-1}$ being undefined, thus no [$\mathbf{A_t}$] matrix can be found. This implies that the model has an upper bound on the number of stations for effective use.

The second problem is where an [$\mathbf{M_{0,t-1}}$] matrix is close to singularity. Here, the problem is of more concern as it can easily be overlooked during the model identification stage, although the final result will still be in error. When the [$\mathbf{M_{0,t-1}}$] matrix is close to singularity then one or more of its eigenvalues may be very small. (In the order of 1/100th or 1/1000th of the remaining eigenvalues.) This has the effect that some elements of the inverted matrix are of the order of 100 or 1000 times the remaining elements.

For the multisite model adopted, when these elements are used to define the [$\mathbf{A_t}$] matrix, the values in corresponding positions in the matrix are very large. Therefore, the [$\mathbf{A_t}$] matrix may well be defined, but when used to generate data it produces values 100 or 1000 times the order expected in the *log* domain. Subsequent backtransformation from the *log* domain to the *raw* domain requires the exponentiation of these already very high values, resulting in a computer overflow.

For the analysis undertaken in this study the above problem occurs with the simultaneous use of the Warren and South Para data sets or the Gorge and Gumeracha data sets within the same multivariate model, due to the high cross correlation between pairs of stations.

Ideally the elements of the [**A**] matrix should be bounded by (+/−1), for meaningful data generation. Thus the elements of the [**A**] matrix should be checked upon computation for the above effect.

5.10.1.2 Solution for the "B" matrix

Similarly to the previous section the description below pertains to an analysis using all streamflow sites.

Once any problems of defining the $[\mathbf{A_t}]$ matrix have been overcome, then attention can be turned to the inevitable problem of defining the $[\mathbf{B}]$ matrix. The method used to solve for the $[\mathbf{B_t}]$ matrix as shown in section [3.4.1] frequently leads to individual elements of the matrix being undefined. Again, the use of a multiperiod model magnifies the problem because of the increased number of matrices to define.

Lack of definition arises in this case since elements on the leading diagonal involve the square root function applied to the manipulation of various lag zero and lag one correlation components. This may lead to a negative value which has no real root. Subsequent calculations to define off diagonal values require the ill–defined diagonal value to be used, further complicating the problem.

For such cases offending elements were set to zero, so as to produce a solution. The $[\mathbf{B_t}]$ matrix only affects the stochastic component of the model, and this problem was not considered too significant for the end result. The number of occurrences of the problem needs to be checked by inspection of the $[\mathbf{B_t}]$ matrices to gain an idea as to the extent of the problem.

The above action seems reasonable, given that the $[\mathbf{B_t}]$ matrix is arbitrarily defined as lower triangular (in lieu of upper triangular, or the use of principal components to solve for the $[\mathbf{B_t}]$ matrix). In the generation of a new vector of flow values, the first value in the vector has only one component in the $[\mathbf{B_t}]$ matrix contributing to the solution, yet the (i^{th}) value has the summation of (i) stochastic components. Considering that some random component (ϵ_i) may occur in any position of the (ϵ) vector, then whether one value or n values are used should not be too significant as there is equal probability of a sum of these terms equalling zero. As long as each row of the $[\mathbf{B_t}]$ matrix has at least one non zero value, the desired result should be produced.

5.10.2 Comparison of Models in Generation & Forecasting

It has been shown that difficulties have occurred with using a large number of stations as well as data with questionable reliability. Does the result affect the viability of the particular model or not ? *i.e.* should more work be undertaken in this direction or should some other method be used ? To evaluate this we need to know how well the model can behave, and between what bounds and under what circumstances.

Given the above results and conclusions formed during the study it was decided to take the best of our data and models and to try to forecast better streamflows than those obtained previously.

As discussed, a multiperiod, multivariate model becomes unstable as its size increases. Also interdependence of data can result in mathematical inconsistencies.

To avoid these problems a five station model was adopted using only one streamflow site from each river system together with one rainfall station.

The following sites were chosen, for reasons of data quality and geographic position -

- South Para
- Myponga
- Torrens at the Gumeracha Weir
- Onkaparinga River at Clarendon Weir
- Millbrook Rainfall Station

The reason for choosing the Millbrook rainfall station rather than the previously used Clarendon P.O. station was that Millbrook is centrally located and its record length encompasses the concurrent record length of the streamflow sites. Also, all rainfall sites have good quality data and similar characteristics.

As noted previously, Crosby & Maddock [13] cite that the problem of matrix inconsistency is less likely to occur by using a concurrent data set. As such the records for the above stations were truncated to their concurrent data period, of 1939 to 1980. This is still a reasonable length to use for our purposes.

Using the above data records the same procedure to determine model parameters was completed *i.e.*

- Find the location parameter for each site and month for a 3–parameter log–normal distribution.
- Decide if the above parameters are reasonable and alter if required.
- Determine the coefficients for the multisite model using the above and then use the model to forecast data.

Table [5.5] shows the location parameters computed for each site and month given *(i)* the full length of record & *(ii)* the concurrent length of record (1939–1980). The results support the premise that the streamflow data does in fact conform to a stationary process, as the parameters are not too different between the two data records.

The data and subsequent parameters computed were in fact found to generate data well, with good correspondence between generated and historical statistics. Tables [5.6] & [5.7] show how well the model preserves the moments of each distribution during generation mode. The results are for the Onkaparinga streamflow station and are the summation statistics based on a five site multiperiod model where eighty replicates of eighty years of data were generated. Table [5.6] has the location parameter determined from a *Zero Skew* approach, whereas table [5.7] has the location parameter determined from *Matalas moment transformation* equations .

Comparing the generated *vs* historical values shown in tables [5.6] and [5.7] given two different procedures for transforming the historical data, it can be seen that the first three moments have been fitted by the moment transformation equations better than by the zero skew method, although the zero skew method still fits the data well. This is not surprising since the moment transformation method is designed to specifically reproduce the values for the first three moments of a distribution.

Comparison of both the median statistic and percentage zero value, indicates that the zero skew method has fitted the historical distribution better as it has replaced the lower end of the distribution more correctly than that of the moment transformation method.

Given that our most critical operating stages are during, or a result of, the low flow periods then fitting the lower end of the distribution is of more importance here, than particular values of distribution moments.

For the data used the location parameters could not just be set to zero when using moment transformation equations, as suggested by McMahon & Mein [36] if high percentage zero values are encountered, as this would result in a small or nil percentage zero value being generated. This simply is a result of the data being used for this study.

The reason for showing these tables when comparing forecasting models is to indicate that the ability of a model to reproduce sample statistics when used for generation is not a good indication of its ability to perform when forecasting. This places in perspective the difficulties that lie ahead for this form of hydrologic analysis.

Comparison of Location Parameters, Full Length *vs* Truncated Data Records													
Site	Record	January	February	March	April	May	June	July	August	September	October	November	December
South Para	Full	-400	-342	normal	-261	-33	-53	-429	-712	-647	-44	-234	-325
	Truncated	-400	-342	normal	-261	-33	-53	-429	-712	-647	-44	-234	-325
Myponga	Full	-64	-18	-265	-95	166	-34	-323	-2503	50	23	-16	normal
	Truncated	-59	-19	-284	-125	156	105	-293	-1937	98	14	-48	normal
Gumeracha	Full	-211	-166	-215	-1	29	-102	-155	-710	-102	21	-83	-69
	Truncated	-241	-165	-321	-5	42	7	-229	-569	-243	3	-96	-81
Onkaparinga	Full	0	-436	0	-84	-61	-322	-1367	-4815	-1095	-140	-292	-762
	Truncated	-3730	-461	-3481	-99	58	-4	-2089	-3137	-822	283	-222	-1160
Millbrook	Full	-55	-52	-35	-220	-436	-249	-944	-3951	-403	-536	-219	-82
	Truncated	-38	-53	-12	-422	-299	-387	-1570	-883	-225	-358	-330	-229

Table 5.5: Comparison of Location Parameters, Full Length *vs* Concurrent Data Records

Onkaparinga Generation Values (Ml) - Zero Skew Transformation													
Statistic	Type	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Minimum	Hist.	-1381	-405	-490	-78	114	97	779	1741	953	840	-50	-443
	Gen.	-955	-339	-389	-50	149	186	81	96	495	594	19	-361
Maximum	Hist.	2859	5927	946	16984	26617	62725	69879	66025	39687	31054	14061	2844
	Gen.	2206	4511	1043	8969	33355	*****	94448	95116	68228	76089	19636	3248
Mean	Hist.	367	515	270	1080	3544	8665	15858	18079	11997	7606	2795	827
	Gen.	362	495	273	964	3433	9292	16118	18437	12329	8107	2830	828
Median	Hist.	311	272	232	524	1830	3258	13393	15875	7965	4601	1366	708
	Gen.	321	238	264	521	1768	3879	11406	14152	8704	4536	1782	711
Std. Dev.	Hist.	664	1133	295	2595	5464	12919	14639	14388	10253	7979	3021	705
	Gen.	648	851	296	1418	5206	18577	16209	16688	12146	11554	3303	707
Skew	Hist.	1.368	3.500	0.205	5.473	2.968	2.557	1.652	1.148	1.066	1.524	1.850	1.019
	Gen.	0.407	2.283	0.207	3.281	3.309	3.996	2.355	2.072	2.204	3.400	2.661	0.974
% below zero	Hist.	14	21	14	2	0	0	0	0	0	0	2	5
	Gen.	31	30	18	5	0	0	1	1	0	0	1	8

(***** - Value inexcess of 100,000 Ml)

Table 5.6: Onkaparinga – Generation Statistics, Zero Skew Transformation

Onkaparinga Generation Values (Ml) - Moment Transformation													
Statistic	Type	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Minimum	Hist.	-1381	-405	-490	-78	114	97	779	1741	953	840	-50	-443
	Gen.	-673	-639	-545	-905	-2243	-6275	-4541	-5007	-5139	-3887	-1247	-520
Maximum	Hist.	2859	5927	946	16984	26617	62725	69879	66025	39687	31054	14061	2844
	Gen.	2863	10977	1295	13360	27936	59663	69818	65338	45208	36186	15080	4400
Mean	Hist.	367	515	270	1080	3544	8665	15858	18079	11997	7606	2795	827
	Gen.	383	799	283	1116	3578	8579	15534	17991	11915	7659	2833	893
Median	Hist.	311	272	232	524	1830	3258	13393	15875	7965	4601	1366	708
	Gen.	245	211	270	319	2019	5246	12585	15551	10443	6076	2119	713
Std. Dev.	Hist.	664	1133	295	2595	5464	12919	14639	14388	10253	7979	3021	705
	Gen.	704	1895	372	2453	5390	12416	14103	14242	10157	7825	3032	917
Skew	Hist.	1.368	3.500	0.205	5.473	2.968	2.557	1.652	1.148	1.066	1.524	1.850	1.019
	Gen.	1.158	2.885	0.247	2.686	2.082	1.695	1.362	0.963	0.906	1.189	1.563	1.274
% below zero	Hist.	14	21	14	2	0	0	0	0	0	0	2	5
	Gen.	33	41	23	39	24	25	7	6	9	13	13	13

Table 5.7: Onkaparinga – Generation Statistics, Moment Transformation

5.10.3 Key Station Approach

As noted in Sections [5.10.1.1] & [5.10.1.2] difficulties were encountered with the solution of the [$\mathbf{A_t}$] and [$\mathbf{B_t}$] matrices given inconsistency or near inconsistency.

The reasons for this are embodied in the time structure of the data itself and in particular for the data used for this study given the compilation procedure.

For the analysis undertaken in this study the above problem occurs with the simultaneous use of the Warren and South Para data sets or the Gorge and Gumeracha data sets within the same multivariate model, due to their interdependence.

The above case highlights that a "key–station" approach may be appropriate where correlations are high. *i.e.* remove one of the offending sites from the multisite model and subsequently correlate the transformed flows at this site to the transformed flows at another "*key site*".

5.10.4 White Noise Analysis

One of the major assumptions made when using the approach adopted in this study is that the processes describing the variants are stationary (or at least weakly stationary) and that the model parameters are calibrated on data belonging to a normal distribution. If either of these two assumptions are broken then data will not be generated in the raw domain to mirror that of the historic data.

The question may then be asked, how do we determine if the above two assumptions are held ? (and if not, what is the degree of difference in the answer ?)

In Section [5.5] it is described how the original data was tested for *trends* evident, and that no statistically significant trend was found in the lower order moments. This is one way of indicating that the data is at least weakly stationary, yet gives no indication that both assumptions hold together throughout the analysis.

One method at our disposal is to analyse the white noise component of the generation equation. Here, the calibrated model was taken, (*i.e.* given [$\mathbf{A_t}$] &[$\mathbf{B_t}$] $\forall(\mathbf{t})$ and the values for [$\mathbf{X_t}$]&[$\mathbf{X_{t-1}}$] over some period of the historic record are used in the following rearranged equation to compute the *white noise components* at each site for each time period.

$$\epsilon_t = [\mathbf{B_t}]^{-1}\{[\mathbf{Z_t}] - [\mathbf{A_t}][\mathbf{Z_{t-1}}]\} \tag{5.10}$$

The data used for this analysis was the period 1940 to 1980 for the sites below —

- South Para
- Myponga
- Gumeracha
- Millbrook R.F. station

The epsilon values (ϵ_i) were computed at each site, for each month and year over the above period. These (ϵ_i) values represent a set of random components required to produce perfect forecasts. If the analysis of these values is such that they may be able to be produced by randomly sampling an N(0,1) distribution, then we can confidently predict that the model will produce reliably generated data.

The following tests were used to analyse the ϵ_i values at each site —

- Quantile–Quantile Plots (Refer Figure [5.6])
- Q–Q Plots of ϵ_i values *vs* corresponding transformed data (Refer Figure [5.7])
- Histograms
- Time Series Plots
- Lag 0 & Lag 1 Covariance Matrices

The lag zero and one covariance matrices indicated that the epsilon values were in fact independent of serial and spatial correlation.

All other graphical plots were easily interpreted as the epsilon values being derived from an N(0,1) distribution.

Thus it may be concluded that the type of model and data used does conform to its base assumptions and that the calibration of the model has been achieved such that meaningful generation results will occur.

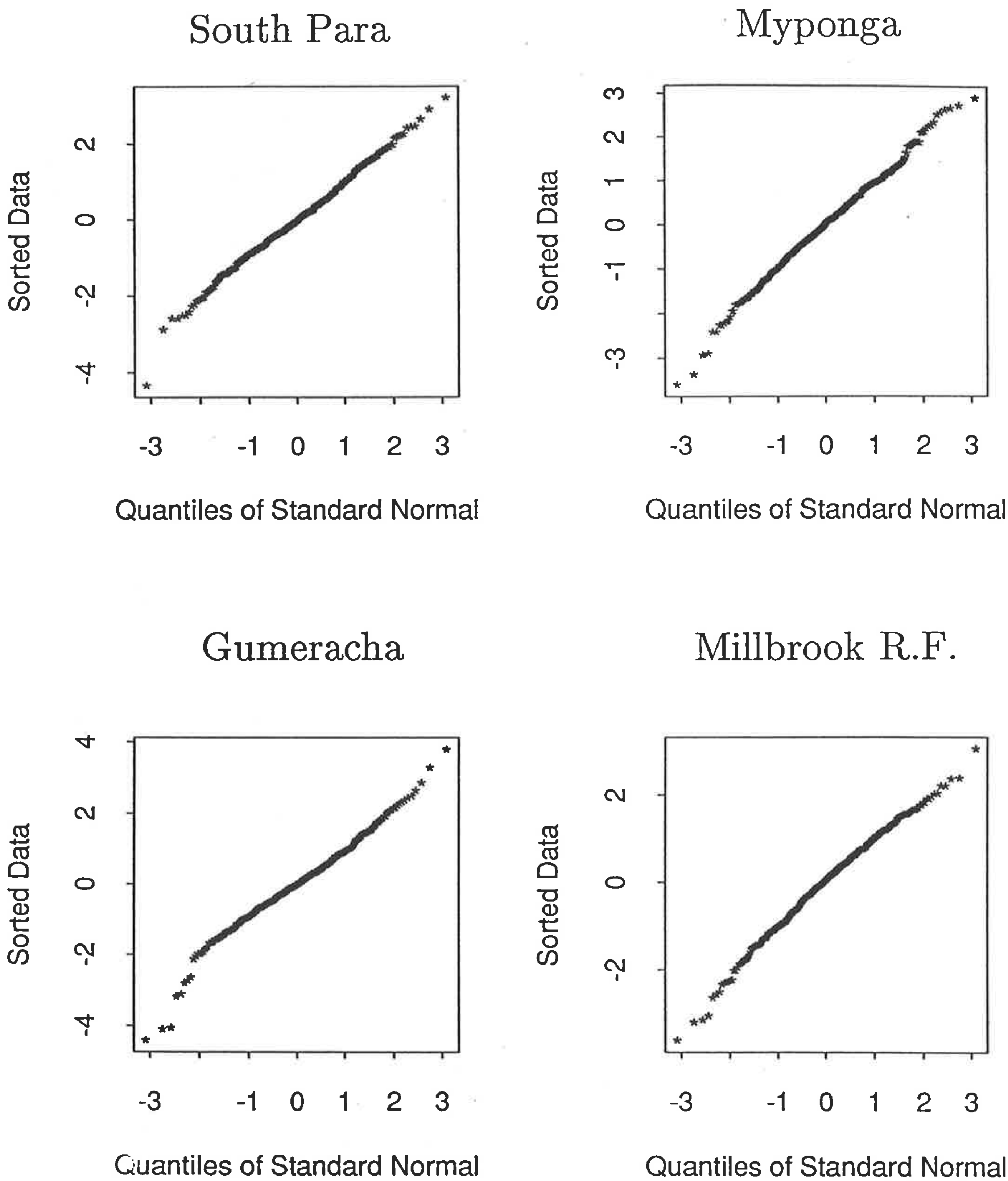

Figure 5.6: Quantile–Quantile Plot for White Noise

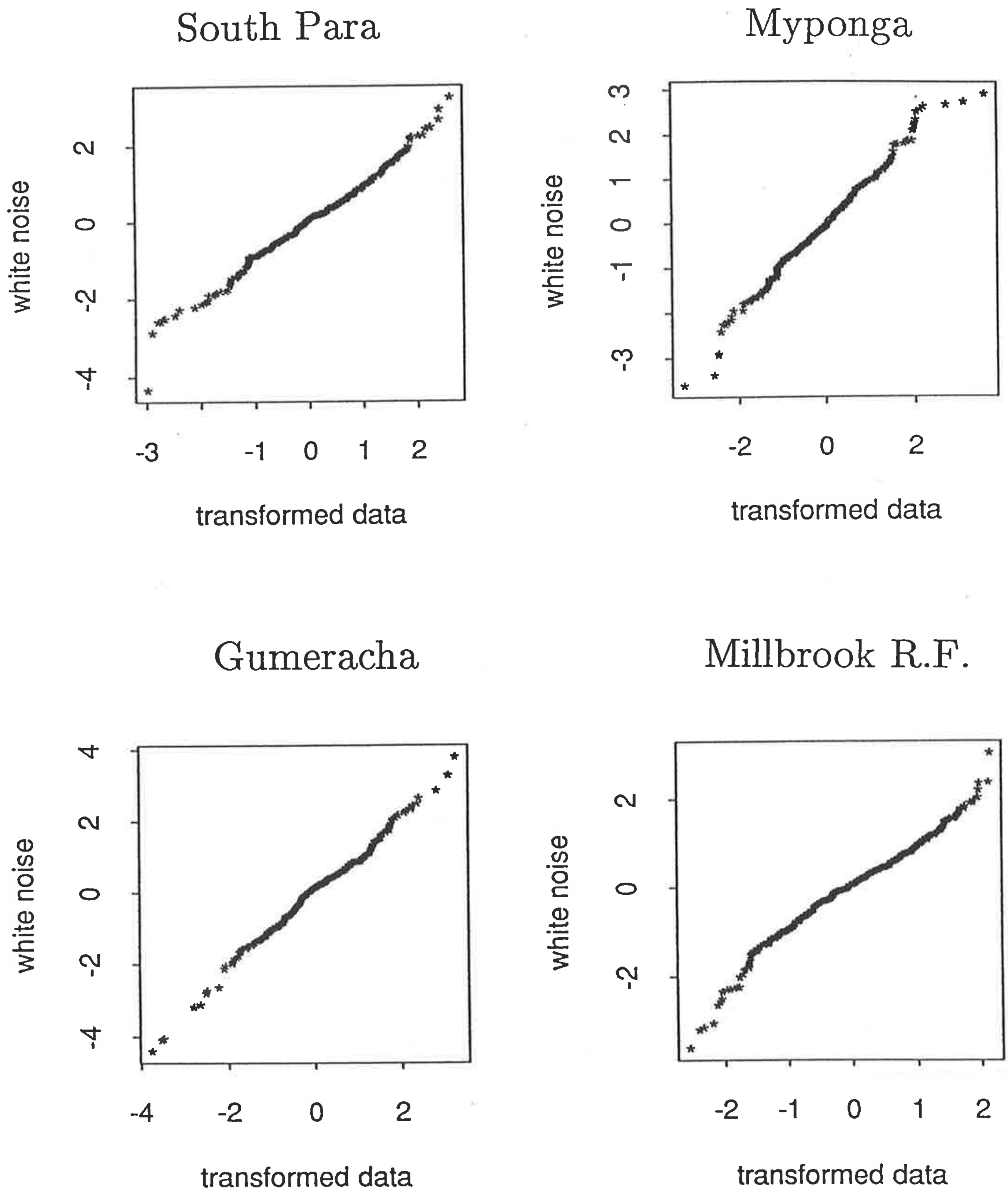

Figure 5.7: Q–Q Plot White Noise *vs* Transformed Data

5.10.5 Univariate Model

During the development of the multivariate model, it became evident that a multivariate model had definite shortcomings as well as being time consuming and possessing a high chance of failure for the operational component of this study.

A far simpler approach was initiated for comparison with the multivariate model to determine if the increased effort required was justified. Thus a univariate model was concurrently developed. The model adopted by this study was a periodic Thomas and Fiering model. See section [3.2]

The generation equation is reproduced here together with the necessary transformation.

$$y_t = \bar{y_t} + \rho_1 \frac{\sigma_t}{\sigma_{t-1}}(y_{t-1} - \bar{y_{t-1}}) + \sigma_t \sqrt{1 - \rho_t^2}(\epsilon_t) \tag{5.11}$$

where

$$y_t = ln(x_t - \tau_t) \tag{5.12}$$

where,

$\bar{y_t}$	=	Mean of normalized values for time period t.
σ_y	=	standard deviation of normalized values.
ρ_t	=	lag one autocorrelation coefficient of the normalized values.
ϵ_t	=	random normal variate.
y_t	=	value of generated series.
x_t	=	flow value in the raw domain.
τ_t	=	location parameter for a three parameter log transformation.

A periodic model was chosen to be in line with operational requirements, with forecasting of data a priority, rather than extensive lengths of continuously generated data. The computer program developed for this purpose is known as "*SINGEN*". This model is straightforward in development using the same principles, procedures and parameters from the data analysis phase as for the multivariate case. No major problems occurred in the development of this type of model and it has proven to be very successful in application.

5.10.6 Univariate Generation and Forecasting

Generation and forecasting was undertaken using both parametric and moment transformation approaches to parameter estimation.

It was generally found that the moment transformation equation approach, although physically irrelevant, produced superior results for the generation of monthly data sets of a given number of years, although the parameters used from the parametric transformation approach produced reasonable data.

Mejia [38] recommends that if the coefficient of variation of the data is greater than (0.75), then the third parameter (τ) of the log normal distribution should be set to zero and the remaining transformed parameters determined from the moment transformation equations. This has the effect of reducing the number of generated negative values. This was not attempted here.

During forecast mode, the Matalas moment transformations did not always backtransform to reasonable results for individual cases, thus relying on a parametric transformation approach to be used for forecasting.

The method of forecasting used is based upon running the generation equation approximately (100) times using the previous months value as an initialization. The forecast adopted was the mean or median (both were investigated) of the (100) values generated.

The generation equation, given above is of the form –

$$y_{t+1} = Coefficient * (y_t) + Stochastic\ Component$$

From this it can be seen that for (100) replicates (or ideally, an infinite series) that the expected value of the stochastic component will tend to zero for $N(0,1)$ randomly generated values. Thus the procedure for forecasting data is based solely on the deterministic component of the equation.

The above was proven graphically by producing (100) stochastic forecasts for each month of the year at the Myponga streamflow gauging station, and superimposing the deterministic forecasts, for various values of the yield for the previous month.

(Refer to figure [5.8] for comparison of forecasts)

The result indicated that the deterministic line coincided approximately with the average of the stochastic line. The stochastic line had large variations in forecasts for small variations in initializing values due to the random component.

Various performance statistics were used to evaluate the quality of the generation model. These were based upon running the model over a known historical record and comparing the percentage and absolute differences of the forecast values with the historical value.

The model performance was evaluated using updated monthly forecast values and compared with the use of a constant historical mean as the forecast value. The performance variables indicate a moderate increase in efficiency for

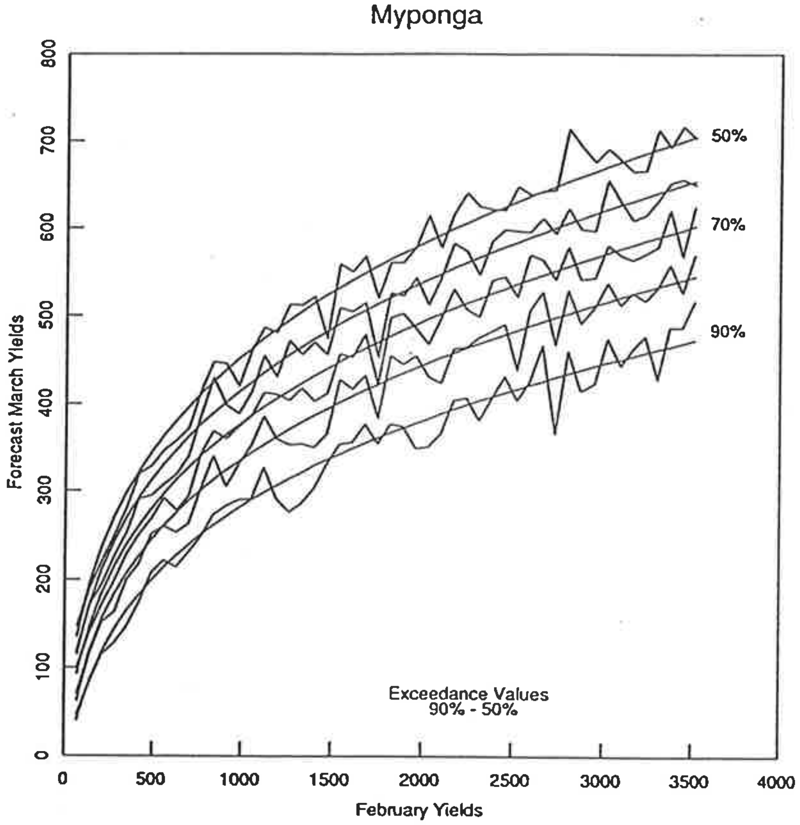

Figure 5.8: Comparison of Stochastic & Deterministic Forecasts

forecasts of one to two months ahead in the low flow months and up to three to four months in the higher flow months. The forecasting model gave no better results than the unconditional monthly mean for longer forecasting horizons.

Figure [5.9] illustrates the above result. It shows the average percentage of forecasting error for various lead times at the Myponga site for the month of August.

The forecasting model is better than the historical mean for lead times of up to 3 months but no better beyond that.

The operation of water supply systems often involves using a forecast inflow that has $p\%$ probability of exceedance. This study was required to forecast various values corresponding to many different exceedance probabilities.

In this case the usual testing of models for preservation of statistical moments, such as the mean, is less important, than the preservation of the tails of the generated distribution. This will be inherently more difficult as the variability in magnitude of forecasting values in the tails increases.

The use of the model for making forecasts with a specified probability of ex-

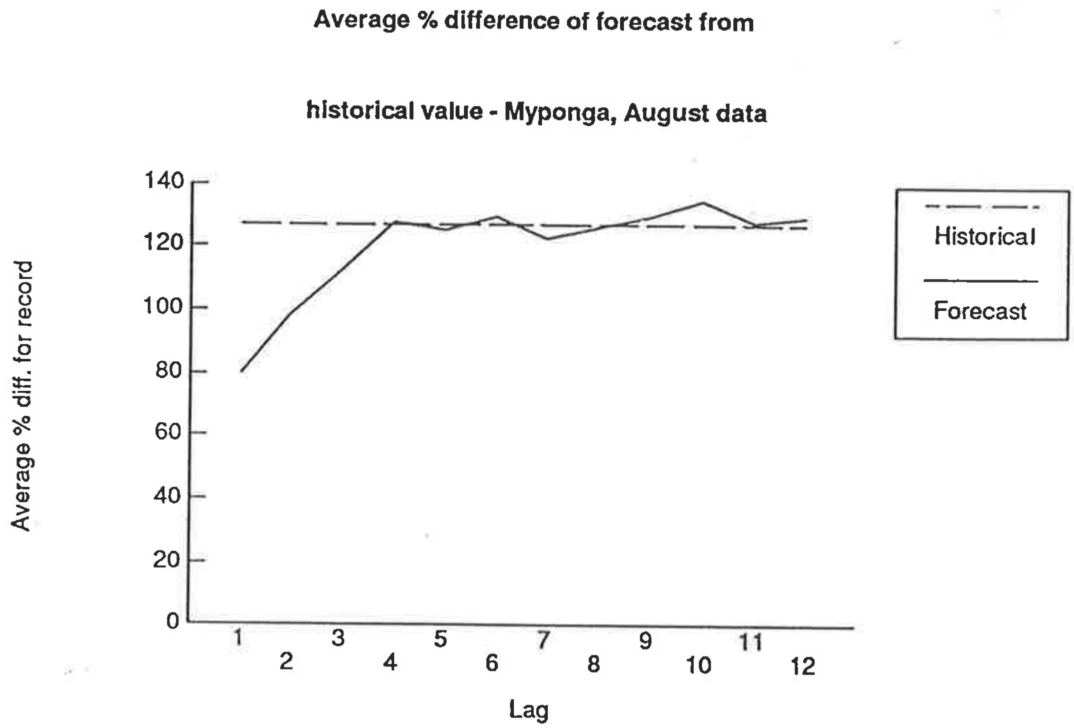

Figure 5.9: Ave % difference of forecast *vs* lag

ceedance is detailed in the following sections.

5.10.6.1 One Step Forecasts

For forecasting to assist in operations we need to estimate the value of a future streamflow which will be exceeded with a specified probability given the most recent streamflow information. The following univariate model was used in this study and is expressed in a slightly different terminology.

$$y_{i,t} = ln(x_{i,t} - \tau_t) \tag{5.13}$$

and

$$y_{i,t} = \bar{y}_t + \frac{\rho_t \sigma_t}{\sigma_{t-1}}(y_{i,t-1} - \bar{y}_{t-1}) + \sigma_t\sqrt{1-\rho_t^2}\epsilon_{i,t} \tag{5.14}$$

where, $x_{i,t}$ = streamflow in month (t) and year (i)
$y_{i,t}$ = transformed streamflow in month (t) and year (i).

p%	50	60	70	80	90
$\Phi(p)$	0.000	-0.253	-0.524	-0.842	-1.282

Table 5.8: Normal Distribution Values given Exceedance Levels

$\epsilon_{i,t}$ = standard normal variate for month (t) and year (i); N(0,1)
τ_t = location parameter for month (t)
$\bar{y}_t$ = mean of $y_{i,t}$ for month (t)
σ_t = standard deviation of $y_{i,t}$ for month (t)
ρ_t = correlation coefficient between $y_{i,t}$ and $y_{i,t-1}$

Now, let $y^p_{i,t}|\hat{y}_{i,t-1}$ be the value of $y_{i,t}$ which is exceeded with probability p given a known value $\hat{y}_{i,t-1}$. Substituting $\hat{y}_{i,t-1}$ for $y_{i,t-1}$ in Equation [5.14], it is apparent that the only random variable on the right hand side is $\epsilon_{i,t}$. As $\epsilon_{i,t}$ is normally distributed with zero mean and unit variance, therefore, $y_{i,t}|\hat{y}_{i,t-1}$ will also be normally distributed.

From Equation [5.14]:

$$\mathbf{E}[y_{i,t}|\hat{y}_{i,t-1}] = \bar{y}_t + \frac{\rho_t \sigma_t}{\sigma_{t-1}}(\hat{y}_{i,t-1} - \bar{y}_{i,t-1}) \tag{5.15}$$

and

$$\mathbf{Var}[y_{i,t}|\hat{y}_{i,t-1}] = \sigma_t^2(1 - \rho_t^2) \tag{5.16}$$

where $\mathbf{E}[X]$ denotes the expected value of X

and $\mathbf{Var}[X]$ denotes the variance of X.

Hence $y^p_{i,t}|\hat{y}_{i,t-1}$ can be found using tables of the standard normal distribution.

$$i.e. \quad y^p_{i,t}|\hat{y}_{i,t-1} = \bar{y}_t + \frac{\rho_t \sigma_t}{\sigma_{t-1}}(\hat{y}_{i,t-1} - \bar{y}_{t-1}) + \Phi(p)\sigma_t\sqrt{1 - \rho_t^2} \tag{5.17}$$

Values of $\Phi(p)$ are given in Table [5.8].

Equation [5.13] describes a deterministic, monotonic relationship between $(y_{i,t})$ and $(x_{i,t})$. It follows, therefore, that $x^p_{i,t}|\hat{x}_{i,t-1}$ can be found from $y^p_{i,t}|\hat{y}_{i,t-1}$ by using the inverse of Equation [5.13].

$$i.e. \quad x^p_{i,t}|\hat{x}_{i,t-1} = exp(y^p_{i,t}|\hat{y}_{i,t-1}) + \tau_t \tag{5.18}$$

Therefore to estimate a one–step forecast of $x^p_{i,t}$ given the previous value $\hat{x}_{i,t-1}$, Equations [5.17] and [5.18] should be used. The value $\hat{y}_{i,t-1}$ can be determined by substituting $\hat{x}_{i,t-1}$ in the right hand side of Equation [5.13].

5.10.6.2 Multiple Step Forecasts

Equations similar to [5.14] can be developed for multiple step forecasts of $y_{i,t}$. For the two step case we obtain –

$$y_{i,t+1} = \bar{y}_{t+1} + \frac{\rho_{t+1}\rho_t\sigma_{t+1}}{\sigma_{t-1}}(y_{i,t-1} - \bar{y}_{t-1}) + \sigma_{t+1}\{\rho_{t+1}\sqrt{1-\rho_t^2}\epsilon_{i,t} + \sqrt{1-\rho_{t+1}^2}\epsilon_{i,t+1}\} \quad (5.19)$$

and for the (m) step case –

$$y_{i,t+m} = \bar{y}_{t+1} + \frac{\rho_{t+m}\rho_{t+m-1} + \cdots + \rho_t\sigma_{t+m}}{\sigma_{t-1}}(y_{i,t-1} - \bar{y}_{t-1})$$
$$+\sigma_{t+m}\{\rho_{t+m}\rho_{t+m-1} + \cdots + \rho_{t+1}\sqrt{1-\rho_t^2}\epsilon_{i,t} + \cdots$$
$$+\rho_{t+m}\rho_{t+m-1}\sqrt{1-\rho_{t+m-2}^2}\epsilon_{i,t+m-2} + \rho_{t+m}\sqrt{1-\rho_{t+m-1}^2}\epsilon_{i,t+m-1} + \sqrt{1-\rho_{t+m}^2}\epsilon_{i,t+m}\} \quad (5.20)$$

From Equation [5.20]:

$$\mathbf{E}[y_{i,t+m}|\hat{y}_{i,t-1}] = \bar{y}_{t+m} + \frac{\rho_{t+m}\rho_{t+m-1} + \cdots + \rho_t\sigma_{t+m}}{\sigma_{t-1}}(\hat{y}_{i,t-1} - \bar{y}_{t-1}) \quad (5.21)$$

$$\mathbf{Var}[y_{i,t+m}|\hat{y}_{i,t-1}] = \sigma_{t+m}^2\{\rho_{t+m}^2\rho_{t+m-1}^2 \cdots \rho_{t+1}^2(1-\rho_t^2) + \cdots$$
$$+ \rho_{t+m}^2\rho_{t+m-1}^2(1-\rho_{t+m-2}^2) + \rho_{t+m}^2(1-\rho_{t+m-1}^2) + (1-\rho_{t+m}^2)\} \quad (5.22)$$

Therefore –

$$y^p_{i,t}|\hat{y}_{i,t-1} = \mathbf{E}[y_{i,t+m}|\hat{y}_{i,t-1}] + \Phi(p)\sqrt{\mathbf{Var}[y_{i,t+m}|\hat{y}_{i,t-1}]} \quad (5.23)$$

where values of $\Phi(p)$ are given in Table [5.8].
To find values of $x^p_{i,t+m}|\hat{x}_{i,t-1}$, use of the following is made –

- Find $(\hat{y}_{i,t-1})$ by substituting $(\hat{x}_{i,t-1})$ into the RHS of Equation [5.13]
- Find $(y^p_{i,t+m}|\hat{y}_{i,t-1})$ using Equations [5.21] & [5.22] and a specified value of (p)
- Find $(x^p_{i,t+m}|\hat{x}_{i,t-1})$ using the inverse transformation of equation [5.18]

From examination of Equations [5.21] and [5.22] it can be seen that as $\rho_{t+m} < 1$ and (m) becomes large –

$$\mathbf{E}[y^p_{i,t+m}|\hat{y}_{i,t-1}] \quad \rightarrow \quad \bar{y}_{t+m} = \mathbf{E}[y_{i,t+m}] \quad (5.24)$$

and the conditional distribution of $y_{i,t+m}$ approaches the unconditional distribution.

Values of the conditional forecasts $x^p_{i,t+m}|\hat{x}_{i,t-1}$ calculated using the above procedure compared well with those using Monte Carlo simulation of Equations [5.13] and [5.14] with (100) replicates, as shown in Figure [5.8].

5.11 Forecasting

Four models were used for comparison, these being —

- Multisite model with Zero Skew Transformation
- Multisite model with Moment Transformation
- Single site model with Zero Skew Transformation
- Single site model with Moment Transformation

The control used for model comparison was the unconditional median of an historical series. That is, for each site and month the unconditional median value of the series was used throughout as the forecast value and the performance parameters calculated accordingly. The historical median was found to perform better than the historical mean for this purpose.

A comparison of models used was undertaken to identify which model produces the highest performance and the differences between model performance.

A set of forecasts was made starting at some point in the historical record. The historical value was used to initiate the model and forecast the next twelve months of data. Two forecast values were chosen for comparison, namely the *mean* and *median* forecast of some (N) replicates. The above procedure was completed over a period of historical record, usually 1940 to 1979 (*i.e.* 40 *yrs*) as this is the concurrent data record for all sites that the program used can complete the forecasting procedure over.

It was found that the median forecast is superior to the mean forecast in almost all cases. Tables [5.9] to [5.13] show for each site the overall monthly performance parameters for a Lag 1 forecast. Only Lag 1 has been shown as this will obviously show the best forecast. As shown in Figure [5.9] the difference between using a forecasting model as opposed to using an unconditional value throughout rapidly converges at a lag of approximately three periods after initialization.

The two parameters chosen to best illustrate the general results of each forecasting model, are —

- Mean Percentage Difference between forecast & historical value for Lag 1.
- Mean Absolute Difference between forecast & historical value for Lag 1 in (Ml).

We can conclude from observation of Tables [5.9[to [5.13] that the multisite zero skew model generally provides the best forecast, followed by the single site, zero skew model. There is little to choose between the multisite and singlesite Matalas models but both are inferior to the zero skew models.

The tables in fact show that even the best model does not do any better than using the unconditional median for approximately 30% of the months, with these months generally being January to March.

The multisite model using Matalas moment transformation equations may be rejected as an operational tool as it is inferior to using the unconditional median for more than 50% of the months.

5.12 Application to Operational Hydrology

On the basis of the performance parameters used in making a comparison between the multivariate and univariate models developed, it was decided to adopt a univariate model for streamflow yield generation. The performance of the multivariate model was only slightly better than the univariate model and required more time and effort to produce results.

On the above basis, a univariate monthly Thomas & Fiering model was chosen to forecast streamflows at individual sites. The model is given as Equations [5.13] and [5.14]. The procedure outlined in Section [5.11.1.1] was used to forecast streamflows with specified probabilities of exceedance.

The resultant forecast streamflows were used as the streamflow input to an optimization model used for minimizing operational costs to the Adelaide metropolitan water supply system.

The optimization model (called HOMA) is described by Dandy & Crawley, [15]

HOMA is an optimisation model of the Adelaide Headworks system. It uses forecast streamflow yields for the metropolitan water supply catchments to aid

South Para Streamflow Station - Lag 1

Statistic	Model	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Average % Difference based on the Median Forecast	Unconditional Median	97	96	161	475	360	297	362	302	454	517	388	163
	Multisite Zero Skew	160	234	310	512	**587**	**210**	**113**	**71**	**114**	**200**	**201**	183
	Single Site Zero Skew	355	**202**	**237**	461	671	223	143	86	119	203	226	**158**
	Multisite Matalas	**155**	253	740	**432**	730	***	156	194	255	357	501	229
	Single Site Matalas	416	237	264	435	887	692	257	221	246	418	386	182
Average Absolute Error based on the Median Forecast	Unconditional Median	134	175	110	188	945	3239	4533	6166	4571	2495	631	162
	Multisite Zero Skew	143	185	155	194	987	**2851**	**3025**	**3561**	**2964**	2290	**464**	**134**
	Single Site Zero Skew	**134**	**167**	**98**	189	**985**	2857	3610	3653	3314	**2261**	501	160
	Multisite Matalas	139	193	255	209	1047	3227	3076	3739	3104	2434	539	144
	Single Site Matalas	**134**	186	100	**188**	998	3293	3721	3935	3372	2366	538	160

*** – Value is in excess of 1000%

Note: Minimum forecast error is shown in bold

Table 5.9: South Para – Forecasting Model Statistic Comparison

Myponga Streamflow Station - Lag 1													
Statistic	Model	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Average % Difference based on the Median Forecast	Unconditional Median	***	618	96	64	56	82	97	124	94	73	60	***
	Multisite Zero Skew	***	554	102	**44**	**53**	**68**	66	**63**	**59**	71	36	***
	Single Site Zero Skew	***	190	**62**	50	54	69	**65**	72	67	**70**	**35**	***
	Multisite Matalas	***	508	100	46	78	91	81	78	63	86	44	***
	Single Site Matalas	**842**	**48**	84	48	66	102	98	86	69	75	44	***
Average Absolute Error based on the Median Forecast	Unconditional Median	147	155	116	162	587	1824	2598	2558	1984	882	250	134
	Multisite Zero Skew	159	157	128	129	**560**	1608	**2132**	**1672**	**1420**	899	206	**92**
	Single Site Zero Skew	**138**	**152**	**85**	142	586	**1515**	2271	1983	1609	**864**	**194**	101
	Multisite Matalas	167	186	137	138	582	1665	2254	1784	**1420**	892	209	180
	Single Site Matalas	148	153	98	**123**	583	1611	2428	2065	1571	872	209	112

*** – Value is in excess of 1000%

Note: Minimum forecast error is shown in **bold**

Table 5.10: Myponga – Forecasting Model Statistic Comparison

Gumeracha Streamflow Station - Lag 1													
Statistic	Model	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Average % Difference based on the Median Forecast	Unconditional Median	79	69	77	255	99	225	174	257	183	163	146	118
	Multisite Zero Skew	112	124	**96**	185	**108**	**146**	**122**	**71**	**99**	**92**	**74**	**67**
	Single Site Zero Skew	111	138	106	266	113	178	141	101	120	98	79	**67**
	Multisite Matalas	91	**73**	129	**149**	216	446	185	196	148	169	102	78
	Single Site Matalas	**85**	110	104	243	190	541	290	180	208	142	129	71
Average Absolute Error based on the Median Forecast	Unconditional Median	92	137	71	104	545	2134	3466	4415	3028	1723	447	173
	Multisite Zero Skew	100	150	71	99	**532**	**1718**	**2705**	**2396**	**2386**	**1562**	**298**	122
	Single Site Zero Skew	98	156	73	**96**	544	1795	2843	3298	2897	1622	318	122
	Multisite Matalas	97	**142**	98	107	567	2072	2870	2675	2414	1790	299	**120**
	Single Site Matalas	**92**	161	**68**	102	565	2098	3003	3435	2846	1684	324	135

Note: Minimum forecast error is shown in bold

Table 5.11: Gumeracha – Forecasting Model Statistic Comparison

Onkaparinga Streamflow Station - Lag 1													
Statistic	Model	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Average % Difference based on the Median Forecast	Unconditional Median	269	156	389	203	130	172	174	126	106	106	155	205
	Multisite Zero Skew	248	146	442	151	104	**87**	**87**	**54**	**62**	**76**	98	**154**
	Single Site Zero Skew	230	138	442	225	**102**	96	92	72	76	78	**86**	169
	Multisite Matalas	255	**123**	444	**118**	207	164	105	79	79	122	107	227
	Single Site Matalas	**221**	148	**378**	216	157	201	145	83	98	103	164	178
Average Absolute Error based on the Median Forecast	Unconditional Median	393	493	224	837	2658	6954	9825	10733	7755	5269	1886	520
	Multisite Zero Skew	413	502	199	**811**	2553	**5901**	**7508**	**7445**	**5961**	5054	**1342**	**403**
	Single Site Zero Skew	**393**	**487**	199	841	2589	6078	8924	8325	7017	**4907**	1417	453
	Multisite Matalas	431	522	249	968	2742	6242	7849	7831	6189	5360	1407	**598**
	Single Site Matalas	401	689	**195**	937	**2510**	6002	9400	8489	7020	5110	1496	454

Note: Minimum forecast error is shown in bold

Table 5.12: Onkaparinga – Forecasting Model Statistic Comparison

Millbrook Rainfall Station - Lag 1													
Statistic	Model	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Average % Difference based on the Median Forecast	Unconditional Median	128	202	226	193	72	81	47	51	57	149	136	127
	Multisite Zero Skew	**159**	190	275	**159**	85	**69**	**44**	**36**	**50**	**148**	144	125
	Single Site Zero Skew	163	**155**	**260**	177	85	74	46	49	57	157	**122**	**118**
	Multisite Matalas	208	194	366	173	94	81	**44**	42	54	168	140	152
	Single Site Matalas	195	179	344	190	**82**	80	49	50	67	155	131	134
Average Absolute Error based on the Median Forecast	Unconditional Median	201	257	208	402	535	454	429	394	418	411	234	180
	Multisite Zero Skew	**212**	256	**209**	**382**	570	**425**	**409**	**308**	**349**	413	211	**173**
	Single Site Zero Skew	**212**	**250**	214	414	569	445	441	358	415	**411**	208	183
	Multisite Matalas	218	256	242	409	595	453	411	319	357	415	**206**	222
	Single Site Matalas	215	263	212	407	**558**	441	446	361	443	413	221	187

Note: Minimum forecast error is shown in bold

Table 5.13: Millbrook – Forecasting Model Statistic Comparison

in identifying optimum pumping decisions, which minimize costs and increase water quality to the city of Adelaide.

The forecast inflows currently used in the optimization model follow the method employed by the Pumping Engineer (E.&W.S.). This method uses a fixed exceedance value for a particular month for any given year. This value is derived from taking the 90% exceedance "year" and distributing this annual value to each month in proportion to the monthly yield in an average year. The method is only modified in very dry years, when the minimum recorded yield for the particular month is used.

Previous experience with the optimization model has shown that potential savings of up to 20% of total pumping costs can be achieved if perfect forecasts of monthly yields were available (Crawley & Dandy, [15]). The model uses a fixed twelve month window and bases its pumping policy on the inputs during that twelve month period (*i.e.* a water year), by stepping from one month to the next in the year and using forecast inputs for future months and actual values for past periods.

A more realistic approach is to continuously update the forecasts on a monthly basis using the most recent data as the optimization model steps through each period. In order to do this some of the source codes for the optimization program was revised to continuously update the forecasts using values from the univariate model developed in this study.

Chapter [4] outlines the two systems that metropolitan Adelaide is divided into. A separate optimization model is used for each system. Each model was run using the following cases for streamflow input forecasts.

- Monthly values based on a 90% exceedance year.
- Perfect values (*i.e.* the actual streamflow yields)
- Use of constant monthly 90% exceedance value
- Use of constant monthly 70% exceedance value.
- Continuously updated monthly 90% exceedance values.
- Continuously updated monthly 70% exceedance values.

The inputs for the constant exceedance value forecasts for the above runs are shown in Tables [5.15] to [5.17], and the results of the optimization runs are shown in Tables [5.18] & [5.19].

In Tables [5.18] & [5.19] "historical" refers to the estimated costs actually incurred by the E.&W.S. during the period. The results indicate that using

Northern System Annual 90% Exceedance Yields						
Month	Warren	South Para	Little Para	Millbrook	Kang. Ck.	Hope V.
Jul	210	700	590	1734	413	453
Aug	300	1000	890	2274	542	594
Sep	190	690	590	1721	410	449
Oct	70	260	230	754	180	196
Nov	20	80	90	233	56	61
Dec	0	30	40	107	25	28
Jan	0	20	20	53	13	14
Feb	0	20	30	47	11	12
Mar	0	10	10	40	10	10
Apr	0	60	60	153	37	40
May	30	230	230	334	80	86
Jun	120	340	340	1154	275	301

Table 5.14: Northern System Annual 90% Exceedance Yields

70% monthly exceedance values is very similar to using 90% annual exceedance values. Unfortunately the use of the updated forecasts does not show marked improvements compared with the use of constant values. The reason for this is difficult to determine. The forecasting model does give improved forecasts for one or two months lead times in low flow months (except for January to March) and three to four months in high flow months. The operational decisions, particularly pumping from the River Murray, depend more on the low flow periods than the high. Thus during the critical times (October to March) the forecasting model gives a limited improvement over the use of unconditional exceedance values.

This is undoubtedly due to the high variability and low monthly serial correlation during these periods.

On the positive side, the models developed can be used for the following purposes —

- to give monthly exceedance values at all sites for various probabilities of exceedance.
- to synthetically generate long streamflow sequences which can be used in the study of system reliability over a reasonable time horizon.

Southern System % Exceedance Yields (Ml)						
	Myponga			Mt. Bold		
Month	Annual 90%	Monthly 90%	Monthly 70%	Annual 90%	Monthly 90%	Monthly 70%
Jul	1550	682	1848	3600	2224	5870
Aug	1600	872	1861	5290	3526	8695
Sep	1090	568	1187	6450	2379	5243
Oct	450	290	621	4930	1252	2919
Nov	200	182	317	2100	592	1026
Dec	130	87	238	720	75	246
Jan	120	54	188	260	10	71
Feb	130	92	171	130	0	30
Mar	100	79	158	120	19	31
Apr	150	147	264	80	21	96
May	370	295	409	250	170	935
Jun	970	401	752	1070	994	2398

Table 5.15: Southern System Yields

Northern System Monthly 90% Exceedance Yields (Ml)						
Month	Warren	South Para	Little Para	Millbrook	Kang. Ck.	Hope V.
Jul	42	306	122	244	305	333
Aug	107	407	158	728	431	471
Sep	72	275	117	469	242	265
Oct	5	139	142	276	142	155
Nov	0*	17	69	110	102	112
Dec	0*	0*	27	21	22	24
Jan	0*	0*	18	27	0	0
Feb	0*	0*	5	16	0	0
Mar	0*	0*	18	18	0	0
Apr	0*	9	43	23	0	0
May	0*	46	78	52	74	81
Jun	0*	76	103	66	185	202

(* indicates a negative value was truncated)

Table 5.16: Northern System Monthly 90% Exceedance Yields

Northern System Monthly 70% Exceedance Yields (Ml)						
Month	Warren	South Para	Little Para	Millbrook	Kang. Ck.	Hope V.
Jul	335	1099	389	1146	601	658
Aug	1241	638	631	1628	860	941
Sep	482	362	480	1035	1075	1177
Oct	145	213	193	555	470	514
Nov	24	127	112	209	203	222
Dec	0*	31	75	64	116	127
Jan	0*	8	37	48	34	37
Feb	0*	0*	30	33	13	15
Mar	0*	0*	40	39	0	0
Apr	0	45	59	38	59	64
May	30	184	100	112	139	152
Jun	67	319	187	309	309	339

(* indicates a negative value was truncated)

Table 5.17: Northern System Monthly 70% Exceedance Yields

Southern System							
Year	Historical	Annual 90%	Perfect	Constant 90%	Constant 70%	Monthly Update 90%	Monthly Update 70%
75/76	0.041	0.272	0.232	0.559	0.268	0.558	0.316
76/77	4.104	2.991	2.844	3.061	3.005	3.062	3.006
77/78	4.442	3.896	3.863	3.827	3.879	3.850	3.900
78/79	1.024	0.914	0.881	0.955	0.863	0.936	0.819
79/80	0.557	0.975	0.322	1.047	0.936	1.063	0.976
80/81	1.775	1.914	1.929	1.958	1.931	1.950	1.951
81/82	0.247	0.456	0.483	0.606	0.496	0.508	0.656
82/83	5.340	3.871	3.780	3.848	3.904	3.822	3.870
83/84	0.606	0.992	0.541	0.992	0.971	0.989	0.830
84/85	0.453	1.190	0.697	1.323	1.170	1.256	1.186
85/86	1.786	1.988	2.041	1.926	1.996	1.910	1.975
86/87	0.015	0.613	0.342	0.697	0.599	0.672	0.590
Sub Total	20.390	20.072	17.928	20.799	20.018	20.576	20.075
End of period Storage	47.48	56.35	62.78	57.32	56.44	57.31	58.35
Adj.	0.842	0.354	—	0.300	0.349	0.301	0.244
Total	21.77	20.43	17.93	21.10	20.37	20.88	20.32

Table 5.18: Southern System – Annual Pumping Costs for Optimization Results \$m

Northern System							
Year	Historical	Annual 90%	Perfect	Constant 90%	Constant 70%	Monthly Update 90%	Monthly Update 70%
79/80	3.398	3.203	1.696	3.400	3.172	3.461	3.271
80/81	3.689	3.537	3.342	3.545	3.491	3.553	3.418
81/82	1.989	1.920	1.629	1.988	1.914	1.882	1.881
82/83	6.459	5.332	5.117	5.258	5.321	5.259	5.310
83/84	2.574	2.496	1.808	2.800	2.452	2.730	2.457
84/85	3.658	3.158	3.110	3.324	3.115	3.425	3.335
85/86	4.242	3.750	4.037	3.696	3.811	3.797	3.868
86/87	2.168	2.429	1.401	2.508	2.369	2.555	2.376
Sub Total	28.177	25.825	22.140	26.519	25.645	26.662	25.846
End of period Storage	70.32	69.10	64.36	71.74	69.10	69.65	67.71
Adj.	0.082	0.162	0.425	—	0.162	0.128	0.247
Total	28.26	25.99	22.59	26.52	25.81	26.79	26.09

Table 5.19: Northern System – Annual Pumping Costs for Optimization results $m

Chapter 6

Summary

6.1 Introduction

This study examined the use of single and multisite time series models for short term forecasting of streamflows. Such forecasts are useful to assist in the operations of water supply systems. Forecasting models were developed for the Adelaide metropolitan water supply system using data from seven Adelaide Hills catchments.

These models were then used, in conjunction with an existing optimization model of the Adelaide headworks system, to determine if improved operational efficiency could be achieved.

The conclusions reached in the report are summarized below –

6.2 Data

The quality of data to be used in such studies is most important, and cannot be over emphasized. For this study it was found that the streamflow data suffered from large errors as a result of the procedures used to compile the data leading to inconsistent results, in terms of both mathematical and physical properties.

The above problems are mainly the result of the natural streamflows having been estimated using a water balance procedure which makes adjustments for pumping and reservoir operations. The additive effect of errors in each component in the respective water balance equation used becomes excessive for two of the data sets (Warren and South Para) and the truncation of negative calculated values is too severe for the above two sites to provide meaningful results.

In addition, the Little Para River data is of limited use due to its short length. Data lengths should ideally be in excess of 30 years.

The streamflow data can only be described as generally poor, with only the Myponga River streamflow data behaving reasonably. In general, the data is highly variable between months and exhibits high skew throughout the monthly data sets.

No trend in any of the data series was found based on analysis of the annual moments of each series.

6.3 Transformations

In general, the data was found to conform to a 3–parameter log–normal distribution.

Two methods were employed to estimate the location parameter of the three parameter log–normal data. These were –

- Parametric Transformation —

  For this case the location parameter of the distribution was estimated so as to produce zero skewness in the transformed data. A test for normality was subsequently made. The technique used was a systematic search, although the approximate equation given by Loucks et al [35] produced comparable values for the location parameter.

- Moment Transformation Equations —

  For this method the moment transformation equations derived by Matalas [37] were used to estimate all three parameters of a three parameter log–normal distribution.

The parametric transformation provided a more realistic or physically relevant solution to the problem of data transformation, as the method reproduces the tails of the distribution more realistically than the moment transformation. Since the low flow months are of most concern operationally, it is the lower end of the distribution that requires to have the best fit.

6.4 Tests for Normality

The following tests were all found to be useful in testing for normality. They are listed in order of priority —

- Quantile–Quantile Plots
  This test is used if time and resources are available.
- Shapiro–Wilk Test
  This test can be used if the sample size is less than (100), or for more precise results with a sample of less than (50).
- Test of Skewness
  This test should be used as an adjunct to the above tests.

6.5 Modelling

Both univariate and multivariate models were evaluated for flow forecasting. In both cases the models were based on a periodic autoregressive model of the first order.

6.5.1 Univariate Model

For this study a monthly Thomas anf Fiering [51] model was developed. The model was found to be straightforward in application and use. Such models may be developed quickly and easily with considerably less effort than for the multivariate case.

It was found that the efficiency with regard to forecasting of streamflows using this model was at least directly comparable to that of the multivariate model.

6.5.2 Multivariate Model

The Autoregressive Multivariate, Multiperiod model is open to failure both by the nature of the model and by use of the poor data used by this study. The need to estimate a number of periodic matrices requires only one matrix to be undefined for model failure.

Two problems are frequently cited as developing problems with such a model and were found to occur for the model adopted by this study. These are —

- The covariance matrix [$\mathbf{M_0}$] used for the definition of the model [**A**] matrix may be singular or close to singularity, and thus not allowing the [**A**] matrix to be defined since the inverse of this covariance matrix is used. This is generally the result of high spatial correlation values within the [$\mathbf{M_0}$] matrix. This was encountered in the study for two catchment systems which each have two gauging stations used in the operational model. Obviously the gauging stations sharing the same catchment will be highly correlated.

  If the above problem arises, the model size will need to be reduced until matrix definition is attained. A key station approach may then be used to produce data at the sites excluded from the multivariate model.

- The method used to define the elements of the model [**B**] matrix does not always lead to real solutions. Since the [**B**] matrix is not unique, the method of Cholesky decomposition as adopted by this study does not guarantee a successful result for all cases. Another method, based on principal components has been outlined by Rodriguez & Bras [7] which may overcome difficulties with the definition of elements in the [**B**] matrix, and is further extended to an approximate solution if all the above fails using a method developed by Mejia & Millan and shown by Rodriguez & Bras [7].

  It was found for the complete seven site model, (six streamflow & one rainfall site), that matrix definition frequently did not occur, and was further exaggerated if moment transformation equations were used for parameter estimation. By using a five site model with concurrent data sets so as to reduce the effect of high station correlation and the likelihood of matrix singularity, the model was found to behave well, with all model matrices being defined, and subsequently producing generated data comparable in distribution and type to the historical series.

6.6 Generation

Summary statistics for the generation of data derived by the adopted models may be described as good. The univariate model statistics compare with the historical statistics on a site by site basis very well and are within two standard errors of estimate for each parameter evaluated.

The multivariate model produces good summary statistics, but benefits from a large averaging effect of high numbers of years or replicates used.

Moment transformation techniques used for model parameter estimation resulted in slightly superior summary statistics for the moments of mean, standard deviation *etc.* It was found although, that the percentage zero value produced was excessively high for the low flow periods of November to February and significantly higher in general than found in the historical data. By using a parametric transformation, the generated moments still compared well with the observed values, yet the occurrence of zero flows (or less, for the case study) fitted the observed data as well as for any moment statistic.

6.7 Forecasting

The ability of a model to forecast data given a good generation model does not always follow. A comparison of forecast data with historical values over some forty year period resulted in an overall increase in forecasting quality for only the first one to three months lead time during the high flow months reducing to only one to two months for the low flow periods of November to March.

Such time series models have a large attraction toward the average value of a series and thus find it difficult to follow historic sequences to any significant accuracy.

The models were evaluated in terms of performance by compiling the average percentage and average absolute differences between the forecast and associated historical value for a given site and month. The models may be ordered from best to worst as follows —

- Multisite model based upon a Parametric Transformation.
- Singlesite model based upon a Parametric Transformation
- Multisite model based upon a Moment Transformation.
- Singlesite model based upon a Moment Transformation

6.8 Application to Operational Hydrology

For the operational component of the study the periodic univariate model was chosen to forecast data, with the random component of the model set to a fixed exceedance value for an $N(0,1)$ distribution. This results in explicitly derived forecast values of a known exceedance level.

Forecast data was used as input to an optimization model of the Adelaide Headworks system. The use of a continuously updated forecast, based on the adopted monthly model, provided no significant increase in system efficiency compared to the results using a fixed forecast based on historical data. Similar results were gained for the Northern and Southern Headworks systems for inflows with various exceedance levels, thus indicating that the optimization model is relatively insensitive to streamflow forecasts.

Use of the forecasting model does not significantly improve the performance of the optimization model. This is most likely due to the high variability and low monthly serial correlation embodied in the data.

By virtue of the type of general model used (*i.e* an ARMA type model) any generated value will have a tendency to gravitate toward the mean of the series. This means that during forecasting the values will only be slightly influenced by the initialization value, and for lead periods of, at most, three periods.

Chapter 7

Recommendations for Further Work

7.1 General Recommendations

7.1.1 Estimation of Parameters

It is believed that superior estimation of distribution parameters will be provided by the use of *Robust statistical methods.*

This field is relatively embryonic for useful analysis and has not been extensively researched with respect to the field of hydrology. A suggested course of action is to review relevant literature and apply some of the methods to the data used for this study. Such methods may also be used to overcome the problems outlined above with respect to outliers evident within sampled data.

7.1.2 Modelling

Further investigation is required into the point at which a multi–site model, as developed by this study, breaks down. Quantitative methods are required to outline when a multisite model may effectively be used.

A further modelling procedure that may be investigated to overcome this problem is a two–tier type approach to analysis, where a full multisite model is used if the [**A**] and [**B**] matrices are well defined, or a smaller multisite model together with key station approach used if it is found that the problem of ill conditioning is present.

7.1.3 Model Verification

More work needs to be carried out into the verification of stochastic time series models in hydrology. In many studies using synthetic data generation a probability distribution and model form is assumed, with little or no model verification. In cases where a transformation to normality is used, standard tests of normality such as the Quantile–Quantile plots or the Shapiro–Wilk test should be carried out.

7.2 Specific Recommendations

7.2.1 Data

There is definite need to reconsider the procedure for data compilation and streamflow gauging. It is clear that severe errors/inconsistencies exist within all data sets used in this study. These will only be remedied by an extensive re–evaluation of all data in toto.

This work would involve estimating the errors associated with parameters in the water balance equations used as well as physical checking of the gauging stations.

7.2.2 Forecasting

Alternative univariate methods for flow forecasting could be applied to the Adelaide Hills data. For example, a general ARMA model (Box & Jenkins [6]) could be used instead of the first order autoregressive model used in this study.

Bibliography

[1] Anderson T. W. (1957) *Maximum Likelihood Estimates for a Multivariate Normal Distribution when Some Observations are Missing*
(Journal of American Statistical Assessment Vol. 52 pp. 200–203)

[2] Aitchison & Brown (1973) *The Log–Normal Distribution*
(University of Cambridge – Department of Applied Economics, Monograph 5)

[3] Beard L.R. (1967) *Monthly Streamflow Simulation*
(Engineering Center, U.S. Corps of Engineers, Washington, D.C.)

[4] Benjamin J.R & Cornell C.A (1970) *Probability, Statistics and Decisions for Civil Engineers*
(McGraw–Hill Book Company, New York)

[5] Benson M.A. & Matalas N.C. (1967) *Synthetic Hydrology Based on Regional Statistical Parameters*
(Water Resources Research Vol. 3 No.4 1967 pp. 931–935)

[6] Box G.E.P. & Jenkins G.M (1970) *Time Series Analysis - Forecasting and Control*
(Holden–Day, San Francisco)

[7] Bras R.L. & Rodriguez–Iturbe I. (1985) *Random Functions & Hydrology*
(Addison–Wesley Publishing Company)

[8] Burton C. (1988) *Management of the River Murray During Periods of Extended Drought*
(Submitted for the Degree of M.Eng.Sc. – University of Adelaide)

[9] Burton C. & Dandy G.C. (1988) *Low Flow Analysis of the River Murray using Synthetic Hydrology.*
(Hydrology & Water Resources Symposium 1988, – Canberra pp. 187–191)

[10] Chandler S, Spoila S, Kumar A. (1978) *Flood Frequency Analysis by Power Transformation*
(Journal of the Hydraulics Division ASCE, Vol 104, pp. 1495–1504)

[11] Chatfield C. (1975) *The Analysis of Time Series: Theory and Practice* (Chapman and Hall, London)

[12] Crawley P.D. & Dandy G.C. (1989) *Optimal Operating Policies for Multiple Reservoir Systems* (University of Adelaide - Civil Engineering Department Report)

[13] Crosby D.S. & Maddock T. (1970) *Estimating Coefficients of a Flow Generator for Monotone Samples of Data* (Water Resources Research, Vol 6, No. 4, pp. 1079–1086)

[14] Dandy G.C. (1988) *Developing Reliability–Cost Trade–offs for a Multiple Reservoir System* (Hydrology & Water Resources Symposium 1988, Canberra pp. 197–200)

[15] Dandy G.C. & Crawley P.D. (1991) *Application of Linear Programming to Planning and Operation of Reservoir Systems with Water Quality Considerations.* (International Hydrlogy & Water Resources Symposium Workshop — "Real World" Applications of Optimisation Techniques to Water Resources Management.)

[16] Fiering M.B. (1968) *Schemes for Handling Inconsistent Matrices* (Water Resources Research, Vol 4, No. 2, pp. 291–297)

[17] Fiering M.B. & Jackson B.B (1971) *Synthetic Streamflows* (American Geophysical Union — Water Resources Monograph 1)

[18] Franke J., Hardle W. & Martin D, (Editors), (1984) *Robust and Non Linear Time Series Analysis* (Lecture Notes in Statistics No.26 Springer–Verlag, New York)

[19] Gaundry M. & Laferrierre R. (1989) *The Box–Cox Transformation — Power Invariance & a New Interpretation* (Elsevier Scientific Publishing Company, Amsterdam)

[20] Hampel F.R, Ronchetti E.M, Rousseeuw P.J, Stahel W.A. (1986) *Robust Statistics — the approach based on influence functions* (John Wiley & Sons Inc, New York)

[21] Harman H.H. (1960) *Modern Factor Analysis* (University of Chicago Press)

[22] Harter H.L. (1961) *Expected Values of Normal Order Statistics* (Biometrica Vol 48, No. 1, pg. 151))

[23] Huber P.J. (1981) *Robust Statistics* (Wiley & Sons, New York)

[24] International Mathematical and Statistical Library (IMSL) *Problem solving software for use with Fortran programs.*

[25] Jacobs T.A. (1983) *Generation of Synthetic Streamflow and Rainfall in the Adelaide Hills*
(E&W.S. report, No.77/76)

[26] Kite G.W. (1977) *Frequency and Risk Analyses in Hydrology*
(Water Resources Publications, Fort Collins Colorado U.S.A.)

[27] Johnson & Kotz (1969) *Discrete Distributions*
(Houghton Mifflin Company, Boston)

[28] Kottegoda N.T. (1983) *Investigation of Outliers in Annual Maximum Flow Series*
(Elsevier Scientific Publishing Company, Amsterdam)

[29] Kottegoda N.T. (1980) *Stochastic Water Resources Technology*
(The McMillian Press Lt, London.)

[30] Kreyzig E. (1979) *Advanced Engineering Mathematics*
(John Wiley & Sons, New York)

[31] Kuczera G. (1987) *On Maximum Likelihood Estimators for the Multisite Lag One Streamflow Model: Complete & Incomplete Data Cases*
(Water Resources Research, Vol 23, No. 4, pp. 641–645)

[32] Kuczera G. (1988) *The Soil Dryness Index Streamflow Yield Model: An Overview of its Development and Capabilities*
(Hydrology & Water Resources Symposium 1988, Canberra, pp. 103–107)

[33] Lazaro R.C, Labadie J.W & Salas J.D. (1982) *Optimal Management of Multi–reservoir Systems using Streamflow Forecasts*
(*taken from* Time Series Methods in Hydrosciences 17, Elsevier Scientific Publishing Company, Amsterdam)

[34] Lindner M.A., Samad F.A. & Howell D.T. (1980) *The Use of Synthetic Hydrology in Decision Making in a Complex River Valley*
(Hydrology & Water Resources Symposium, Adelaide 1980, pp. 119–127)

[35] Loucks D.P, Stedinger J.R & Haith D.A (1981) *Water Resources Systems, Planning and Analysis*
(Prentice Hall, Englewood Cliffs, N.J.)

[36] McMahon T.A. & Mein R.G. (1986) *River and Reservoir Yield*
(Water Resources Publications, Colorado)

[37] Matalas N.C. (1967) *Mathematical Assessment of Synthetic Hydrology*
(Water Resources Research Vol 3, No 4, pp. 937–945)

[38] Mejia J.M. (1974) *Multivariate Generation of Mixtures of Normal & Log–normal Variables*
(Water Resources Research, Vol 10, pp. 691–693)

[39] Nazem S.M. (1988) *Applied Time Series Analysis for Business and Economic Forecasting*
(Marcel Dekker Inc, New York)

[40] Pankratz A. (1983) *Forecasting with Univariate Box–Jenkins Models*
(Wiley series in Probability & Mathematical Statistics, New York)

[41] Pearson E.S. & Hartley H.O. (1976) *Biometrika Tables for Statisticians Vol II*
(Cambridge University Press, London)

[42] Becker R.A & Chambers J.M (Editors)
S – An Interactive Environment for Data Analysis and Graphics
(Wadsworth Statistics/Probability Series, California)

[43] Salas J.D, Delleur J.W, Yevjevich V, Lane W. (1980) *Applied Modelling of Hydrologic Series*
(Water Resources Publications, Colorado)

[44] Salas J.D, & Pegram G.G.S. (1978) *A Seasonal Multivariate Multilag Autoregressive Model in Hydrology*
(Third Int. Hydrology Symposium, Colorado)

[45] Sangal B.P, & Biswas A.K. (1970) The Three Parameter Log–Normal Distribution & its Applications in Hydrology
(Water Resources Research, Vol 6, No. 2, pp 505–515)

[46] Shapiro S.S, & Francia R.S. (1972) *An Approximate Analysis of Variance Test for Normality*
(Journal of the American Statistical Association, Vol 67, No. 337)

[47] Shapiro S.S, Wilk M.B. (1965) *An Analysis of Variance Test for Normality*
(Biometrika, Vol 52, pp. 591–611)

[48] Srikanthan R, McMahon T.A, Codner G.P, Mein R.G. (1983) *Theory and Application of some Multisite Streamflow Generation Models*
(Civil Engineering Transactions – Institution of Engineers Aust. Vol 26, pp. 272–279)

[49] Stedinger J.R. (1980) *Fitting Log–Normal Distributions to Hydrologic Data*
(Water Resources Research, Vol 16, No. 3, pp. 481–490)

[50] Stedinger J, & Pei D. (1982) *An Annual Monthly Streamflow Model for Incorporating Parameter Uncertainty into Reservoir Simulation*
(In Time Series Methods in Hydrosciences, Elsevier Scientific Publishing Company, Amsterdam)

[51] Thomas H.A. & Fiering M.B. (1962) *Mathematical Synthesis of Streamflow Sequences of the Analysis of River Basins by Simulations.*
(In, Design of Water Resource Systems – A. Mass et al, Eds. Harvard University)

[52] Wilkinson J.H (1978) *Singular Value Decomposition – Basic Aspects* (In Numerical Software – Needs and Availability, Academic Press, London)

[53] Yevjevich V. (1972) *Stochastic Processes in Hydrology* (Water Resources Publications, Fort Collins Colorado, U.S.A.)

[54] Young G.Y & Pisano W.C (1968) *Operational Hydrology Using Residuals* (Proceedings of the ASCE, Journal of the Hydraulics Division, Vol 94, pp. 909–923)

Appendix A

Historical Statistics

This appendix contains the historical statistics of the original streamflow yield data files, as supplied by the E.&W.S. Dept.
The statistics for the Clarendon rainfall gauging station are also given here for a typical comparison with streamflow characteristics. The stations are —

- Warren River
- South Para River
- Myponga River
- Onkaparinga River at Clarendon Weir
- Torrens River at Gorge Weir
- Torrens River at Gumeracha Weir
- Little Para River
- Clarendon P.O. rainfall station.

Streamflow data is expressed in (Ml) and rainfall in (1/10th mm).

Catchment : Warren — Start Year : 1939 — Period : 46yrs

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	0	0	0	0	0	0	0	0	0
Max	387	504	273	218	5544	10872	15901	20708	11432	12916	1966	237
Mean	30.5	39.4	34.6	55.7	362.2	1597.7	3321.9	4759.9	3169.2	1467.6	254.3	48.0
Median	1.0	7.0	7.0	47.0	97.0	183.0	1963.5	3753.5	1983.0	478.0	70.0	13.5
St. Dev.	67.9	91.4	57.9	58.1	887.1	2943.5	4311.2	4848.4	3346.0	2516.6	418.5	66.4
Coef. Skew	3.573	3.784	2.240	0.799	4.535	1.948	1.635	1.129	0.949	2.752	2.190	1.308
Lag 1 S.C.	0.168	0.063	0.099	0.600	0.194	0.202	0.712	0.729	0.448	0.333	0.420	0.409
10% exc.	83.0	69.0	98.0	122.0	752.0	4983.0	9876.0	10656.0	8591.0	3617.0	911.0	147.0
90% exc.	0.0	0.0	0.0	0.0	0.0	0.0	42.0	107.0	72.0	5.0	0.0	0.0

Catchment : South Para — Start Year : 1939 — Period : 42yrs

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	0	0	0	0	0	0	0	0	0
Max	979	999	462	1244	11111	20407	29532	29312	27248	18694	5991	901
Mean	106.6	151.1	79.5	223.7	1078.2	3469.6	5883.9	7753.0	5504.4	2842.5	707.5	159.3
Median	22.0	21.0	26.5	132.0	348.5	649.0	3719.5	4578.0	2457.0	1211.0	250.0	61.0
St. Dev.	213.7	266.6	113.7	262.2	2020.4	5865.0	7190.4	8324.4	5939.3	3932.1	1135.4	218.7
Coef. Skew	2.531	1.185	1.809	1.775	3.312	1.759	1.889	1.145	1.386	2.085	2.800	1.771
Lag 1 S.C.	0.467	0.341	0.280	0.421	0.305	0.359	0.728	0.621	0.625	0.267	0.329	0.337
10% exc.	459.0	535.0	212.0	605.0	3437.0	13512.0	13299.0	18881.0	12574.0	8838.0	1715.0	494.0
90% exc.	0.0	0.0	0.0	0.0	30.0	71.0	310.0	402.0	302.0	137.0	10.0	0.0

Table A.1: Warren River & South Para River – Historical Statistics

Catchment : Myponga					Start Year : 1934				Period : 51yrs			
	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	1	27	59	195	85	360	462	290	119	119	0
Max	2138	3419	607	1425	4527	12169	14638	15509	14122	6032	1702	594
Mean	288.3	282.3	252.3	398.6	908.0	2387.3	4100.9	4421.7	2906.1	1361.8	541.4	308.7
Median	264.0	211.0	252.0	356.0	560.0	1191.0	2960.0	3920.0	2059.0	898.0	448.0	286.0
St. Dev.	303.2	461.7	140.4	242.8	968.7	2619.9	3470.8	3079.6	2868.4	1312.1	343.0	156.3
Coef. Skew	4.583	6.145	0.523	1.680	2.590	2.088	1.359	0.938	1.981	1.988	1.434	0.061
Lag 1 S.C.	0.124	0.417	0.466	0.402	0.290	0.436	0.442	0.567	0.555	0.499	0.496	0.563
10% exc.	384.0	330.0	409.0	700.0	1808.0	5346.0	10071.0	7556.0	7021.0	2844.0	844.0	528.0
90% exc.	54.0	92.0	79.0	147.0	295.0	401.0	682.0	872.0	568.0	290.0	182.0	87.0

Catchment : Onkaparinga at Clarendon						Start Year : 1898				Period : 87yrs		
	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	0	34	97	779	962	564	201	0	0
Max	2859	5927	2862	16984	26617	64207	81059	80446	66429	31054	14061	2844
Mean	338.7	351.3	240.3	804.9	3273.0	10950.2	16771.1	20181.7	15278.2	7029.1	2247.4	721.8
Median	205.0	74.0	118.0	392.0	1570.0	5289.0	11463.0	17035.5	9798.0	4198.0	1283.0	597.0
St. Dev.	469.0	830.4	366.3	1987.0	5109.8	13503.2	16162.2	15793.0	14279.0	7212.6	2471.1	655.8
Coef. Skew	3.185	4.903	3.817	6.597	3.045	2.004	1.591	1.201	1.469	1.734	2.389	1.181
Lag 1 S.C.	0.273	0.025	0.293	0.181	0.505	0.436	0.587	0.531	0.267	0.248	0.373	0.480
10% exc.	695.0	554.0	590.0	1465.0	8506.0	27724.0	37256.0	38824.0	36122.0	18610.0	5523.0	1810.0
90% exc.	10.0	0.0	19.0	21.0	170.0	994.0	2224.0	3526.0	2379.0	1252.0	592.0	75.0

Table A.2: Myponga River & Onkaparinga River – Historical Statistics

Catchment : Torrens at Gorge Weir   Start Year : 1884   Period : 100*yrs*

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	0	0	151	0	0	203	119	0	0
Max	1895	3735	1179	33031	18740	47532	51015	47181	35571	32588	8148	3842
Mean	342.7	263.3	188.1	848.4	1818.4	6237.3	9271.9	12385.8	9617.5	4578.3	1454.0	626.5
Median	248.5	128.0	128.5	307.0	719.0	2784.5	5750.0	10161.5	7099.5	2928.0	915.0	425.5
St. Dev.	382.3	540.7	235.9	3384.0	3079.9	8714.8	10074.3	10600.3	8754.6	5624.5	1462.7	601.6
Coef. Skew	1.978	4.799	1.942	8.702	3.242	2.162	1.782	0.904	1.126	2.590	2.279	2.264
Lag 1 S.C.	0.263	0.065	0.361	0.009	0.248	0.495	0.637	0.508	0.281	0.313	0.424	0.590
10% exc.	782.0	484.0	482.0	1104.0	4855.0	19056.0	20456.0	28692.0	22200.0	10621.0	3119.0	1432.0
90% exc.	0.0	0.0	0.0	0.0	208.0	454.0	882.0	1631.0	975.0	574.0	323.0	68.0

Catchment : Torrens at Gumeracha   Start Year : 1918   Period : 66*yrs*

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	0	32	0	39	128	110	69	0	0
Max	1007	1921	812	2898	10733	40896	24092	25203	29900	13739	3521	2128
Mean	150.4	150.8	115.0	176.2	870.7	2712.3	4232.4	6065.9	4326.8	2122.9	567.2	244.0
Median	71.5	48.0	52.0	65.0	193.5	651.5	1769.0	4658.5	2356.0	1090.0	354.5	137.0
St. Dev.	205.3	285.4	156.9	385.6	1959.5	5859.7	5132.5	5824.7	4908.7	2914.7	627.3	375.1
Coef. Skew	2.685	4.295	2.437	5.532	3.864	4.538	1.791	1.339	2.432	2.389	2.371	3.177
Lag 1 S.C.	0.272	0.070	0.300	0.146	0.202	0.569	0.706	0.542	0.220	0.369	0.402	0.602
10% exc.	355.0	380.0	288.0	346.0	1247.0	5417.0	11517.0	13055.0	10404.0	4899.0	1470.0	514.0
90% exc.	27.0	16.0	18.0	23.0	52.0	66.0	244.0	728.0	469.0	276.0	110.0	21.0

Table A.3: Gorge & Gumeracha Weirs – Historical Statistics

Catchment : Little Para	Start Year : 1969	Period : 15yrs										
	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	1	2	6	9	100	27	98	8	7	6	2
Max	300	310	232	748	3642	4863	5126	5995	4230	3434	863	367
Mean	131.2	108.6	84.7	146.0	378.5	903.1	1528.1	1650.1	1629.5	856.2	293.3	165.1
Median	148.0	100.0	59.0	83.0	140.0	288.0	1475.0	1134.0	1200.0	462.0	209.0	188.0
St. Dev.	106.1	93.0	74.4	180.6	905.9	1543.9	1412.1	1526.9	1511.7	100.5	243.1	116.2
Coef. Skew	0.183	0.631	0.814	2.417	3.097	1.877	0.957	1.403	0.610	1.495	0.744	0.135
Lag 1 S.C.	0.628	0.645	0.656	0.053	0.935	0.634	0.541	0.768	0.382	0.400	0.619	0.846
10% exc.	272.0	255.0	216.0	258.0	324.0	4462.0	3153.0	3144.0	4195.0	2800.0	515.0	318.0
90% exc.	18.0	5.0	18.0	43.0	78.0	103.0	122.0	158.0	117.0	142.0	69.0	27.0

Rainfall G.S. : Clarendon P.O.	Start Year : 1875	Period : 114yrs										
	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	15	69	147	199	64	160	38	11	0
Max	1586	1380	1750	2905	2652	2927	2900	3109	1899	1952	1382	1834
Mean	258.8	256.4	354.7	723.3	1024.8	1168.2	1125.3	1054.6	833.3	661.6	420.1	362.0
Median	198.5	147.0	262.5	624.5	947.0	1072.5	1074.0	980.5	755.5	622.0	377.5	314.5
St. Dev.	248.1	282.3	330.8	528.4	536.7	620.0	465.8	461.4	397.4	361.8	277.4	282.7
Coef. Skew	2.192	1.428	1.464	1.514	0.634	0.502	0.503	1.351	0.645	0.832	0.883	1.703
Lag 1 S.C.	0.020	-0.035	0.005	-0.070	0.019	0.169	0.114	0.230	0.035	-0.074	0.103	0.026
10% exc.	586.0	683.0	730.0	1427.0	1744.0	2161.0	1708.0	1559.0	1391.0	1160.0	780.0	676.0
90% exc.	28.0	5.0	33.0	150.0	376.0	398.0	506.0	550.0	366.0	223.0	99.0	49.0

Table A.4: Little Para & Clarendon rainfall – Historical Statistics

Appendix B

Revised & Transformed Statistics

This appendix tabulates the historical and transformed statistics of the final streamflow data sets adopted for analysis. The Clarendon rainfall station values are also given for comparison with streamflow data.
Streamflow data is expressed in (Ml), and rainfall in (1/10th mm). The stations are –

- Warren River
- South Para River
- Myponga River
- Onkaparinga River at the Clarendon Weir
- Torrens River at the Gorge Weir
- Torrens River at the Gumeracha Weir
- Little Para River
- Clarendon P.O. rainfall station.

Catchment : Warren Start Year : 1939 Period : $46yrs$

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	-210	-198	-327	-101	-67	-62	0	-2	-30	-63	-89	-87
Max	387	504	273	218	5544	10872	15901	20708	11432	12916	1966	237
Mean	8.2	21.9	14.0	46.6	357.7	1594.0	3320.4	4760.0	3168.0	1465.8	246.0	33.5
Median	1.0	7.0	7.0	42.5	97.0	183.0	1963.5	3753.0	1983.0	478.0	70.0	13.5
St. Dev.	87.0	104.8	86.7	69.3	889.0	2941.5	4311.2	4848.5	3347.2	2517.7	424.1	79.7
Skew	1.653	2.597	-0.483	0.262	4.519	1.948	1.636	1.129	0.948	2.749	2.138	0.804
Lag 1	0.338	0.206	0.265	0.307	0.212	0.205	0.713	0.729	0.448	0.334	0.420	0.439
10% exc.	83.0	69.0	98.0	122.0	752.0	4983.0	9876.0	10656.0	8591.0	3617.0	911.0	147.0
90% exc.	-59.0	-57.0	-50.0	-36.0	-14.0	-5.0	42.0	107.0	72.0	5.0	-44.0	-54.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	6.05	5.75	2.31	46.61	5.15	5.86	7.35	8.31	7.70	6.35	5.26	5.06
St. Dev.	0.19	0.28	1.97	69.25	1.37	1.91	1.40	0.90	1.06	1.49	1.09	0.46
Loc. Par	-425	-304	-10	normal	-67	-62	-184	-1094	-453	-77	-97	-140
Skew	-0.046	-0.038	-0.025	0.262	-0.583	-0.072	-0.049	-0.050	-0.050	-0.036	-0.047	-0.050
Lag 1	0.488	0.296	0.397	0.434	0.464	0.753	0.752	0.760	0.726	0.516	0.534	0.559

Table B.1: Warren River statistics

Catchment : South Para Start Year : 1939 Period : 42*yrs*

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	-278	-266	-447	-160	-19	-44	75	136	-37	15	-202	-216
Max	979	999	462	1244	11111	20407	29532	29312	27248	18694	5991	901
Mean	67.7	118.6	40.8	216.8	1110.4	3548.7	6024.0	7937.6	5633.9	2912.2	716.2	142.9
Median	2.5	10.5	23.5	159.5	365.5	678.0	4096.5	4987.5	3476.5	1621.5	258.5	70.0
St. Dev.	239.4	288.6	158.3	271.3	2012.6	5841.4	7130.0	8233.8	5876.8	3905.0	1136.6	238.9
Skew	1.946	1.538	0.014	1.563	3.307	1.745	1.893	1.137	1.386	2.086	2.761	1.339
Lag 1	0.446	0.398	0.393	0.209	0.299	0.352	0.722	0.613	0.617	0.254	0.322	0.349
10% exc.	451.0	535.0	212.0	605.0	3437.0	13512.0	13299.0	18881.0	12574.0	8838.0	1715.0	494.0
90% exc.	-116.0	-75.0	-119.0	-25.0	32.0	71.0	348.0	514.0	347.0	144.0	-27.0	-83.0

TRANSFORMED STATISTICS

Mean	6.05	5.97	40.81	4.62	6.08	6.90	8.24	8.57	8.27	7.13	6.38	6.03
St. Dev.	0.45	0.57	158.34	1.87	1.44	1.74	1.08	1.06	1.04	1.44	0.98	0.48
Loc. Par	-400	-342	normal	-26	-33	-53	-429	-712	-647	-44	-234	-325
Skew	-0.047	-0.042	-1.294	-0.029	-0.045	-0.050	-0.050	-0.050	-0.049	-0.040	-0.049	-0.049
Lag 1	0.470	0.452	0.487	0.057	0.183	0.710	0.681	0.775	0.764	0.531	0.597	0.505

Table B.2: South Para River statistics

Catchment : Myponga				Start Year : 1934					Period : 51*yrs*			
HISTORICAL STATISTICS												
	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	1	27	59	195	85	360	462	290	119	119	0
Max	2138	3419	607	1425	4527	12169	14638	15509	14122	6032	1702	594
Mean	208.3	282.3	252.3	398.6	908.0	2387.3	4100.9	4421.7	2906.1	1361.8	541.4	308.7
Median	264.0	211.0	252.0	356.0	560.0	1191.0	2960.0	3920.0	2059.0	898.0	448.0	286.0
St. Dev.	303.2	461.7	140.4	242.8	968.7	2619.9	3470.8	3079.6	2868.4	1312.1	343.0	156.3
Skew	4.583	6.145	0.523	1.680	2.590	2.088	1.359	0.938	1.981	1.988	1.434	0.061
Lag 1	0.124	0.417	0.466	0.402	0.290	0.436	0.442	0.567	0.555	0.499	0.496	0.563
10% conf.	384.0	330.0	409.0	700.0	1808.0	5346.0	10071.0	7556.0	7021.0	2844.0	844.0	528.0
90% conf.	54.0	92.0	79.0	147.0	295.0	401.0	682.0	872.0	568.0	290.0	182.0	87.0
TRANSFORMED STATISTICS												
Mean	5.68	5.40	6.21	6.10	6.07	7.32	8.12	8.75	7.61	6.83	6.15	308.7
St. Dev.	0.59	0.70	0.27	0.46	1.04	0.99	0.76	0.44	0.88	0.88	0.59	156.3
Loc. Par	-64	-18	-265	-95	166	-34	-323	-2503	-50	23	-16	normal
Skew	-0.006	0.029	-0.005	-0.003	-0.011	-0.006	-0.004	-0.005	0.115	-0.006	-0.004	0.061
Lag 1	0.458	0.743	0.619	0.413	0.333	0.521	0.601	0.641	0.601	0.506	0.700	0.638

Table B.3: Myponga River statistics

Catchment : Onkaparinga   Start Year : 1898   Period : 87*yrs*

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	-1381	-405	-490	-78	34	97	779	962	564	201	-180	-443
Max	2859	5927	2682	16984	26617	64207	81059	80446	66429	31054	14061	2844
Mean	300.4	333.0	230.1	804.0	3273.0	10950.2	16771.1	20181.7	15278.2	7029.1	2244.8	712.9
Median	205.0	74.0	118.0	392.0	1570.0	5289.0	11463.0	17035.0	9798.0	4198.0	1283.0	597.0
St. Dev.	526.7	840.5	377.3	1987.4	5109.8	13503.2	16162.2	15793.0	14279.0	7212.6	2473.6	667.7
Skew	1.897	4.771	3.472	6.594	3.045	2.004	1.591	1.201	1.469	1.734	2.383	1.088
Lag 1	0.248	-0.091	0.300	0.182	0.505	0.436	0.587	0.531	0.267	0.248	0.374	0.480
10% exc.	695.0	554.0	590.0	1465.0	8506.0	27724.0	37256.0	38824.0	36122.0	18610.0	5523.0	1810.0
90% exc.	10.0	-45.0	19.0	21.0	170.0	994.0	2224.0	3526.0	2379.0	1252.0	592.0	75.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	4.78	6.40	4.41	6.03	7.38	8.68	9.43	9.93	9.34	8.46	7.51	7.20
St. Dev.	1.94	0.64	1.80	1.17	1.22	1.21	0.90	0.64	0.88	0.94	0.80	0.44
Loc. Par	0	-436	0	-84	-61	-322	-1367	-4815	-1095	-140	-292	-762
Skew	-1.160	-0.019	-0.774	-0.004	-0.048	-0.050	-0.050	-0.050	-0.050	-0.050	-0.047	-0.049
Lag 1	0.375	0.244	0.458	0.411	0.401	0.516	0.645	0.652	0.478	0.489	0.616	0.507

Table B.4: Onkaparinga River statistics

Catchment : Torrens at Gorge Weir — Start Year : 1884 — Period : 100yrs

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	-461	-727	-856	-460	-357	151	0	0	203	119	-153	-378
Max	1895	3735	1179	33031	13782	47532	51015	47181	35571	32508	8148	3842
Mean	316.9	214.0	113.2	825.2	1745.0	6304.4	9271.9	12385.8	9617.5	4578.3	1452.5	615.3
Median	248.5	128.0	128.5	307.0	719.0	2784.5	5750.0	10161.5	7099.5	2928.0	915.0	425.5
St. Dev.	412.9	581.8	327.7	3390.7	2758.4	8721.2	10074.3	10600.3	8754.6	5624.5	1464.3	615.8
Skew	1.490	3.095	0.331	8.669	2.753	2.136	1.782	0.904	1.126	2.590	2.271	2.079
Lag 1	0.290	0.153	0.431	0.052	0.282	0.589	0.636	0.508	0.281	0.313	0.425	0.590
10% exc.	782.0	484.0	482.0	1104.0	4855.0	19056.0	20456.0	28692.0	22200.0	10621.0	3119.0	1432.0
90% exc.	-79.0	-181.0	-328.0	-91.0	208.0	454.0	882.0	1631.0	975.0	574.0	323.0	68.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	6.94	6.99	3.62	6.69	7.18	7.78	8.70	9.40	9.01	7.91	7.15	7.05
St. Dev.	0.35	0.39	2.56	0.80	0.94	1.49	1.05	0.72	0.82	1.06	0.75	0.44
Loc. Par	-784	-956	0	-475	-377	100	-588	-3068	-1570	-94	-242	-656
Skew	-0.049	0.048	-0.558	-0.013	-0.047	-0.048	-0.050	-0.050	-0.050	-0.050	-0.047	-0.049
Lag 1	0.415	0.430	0.521	0.420	0.389	0.611	0.618	0.617	0.508	0.575	0.734	0.674

Table B.5: Gorge Weir statistics

Catchment : Torrens at Gumeracha Weir	Start Year : 1918	Period : 66*yrs*

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	-167	-145	-160	-1	32	-95	39	128	110	69	-42	-57
Max	1007	1921	812	2898	10733	40896	24092	25023	29900	13739	3521	2128
Mean	147.9	146.4	110.9	176.2	870.7	2710.8	4234.2	6065.9	4326.8	2122.9	566.3	242.0
Median	71.5	48.0	52.0	65.0	193.5	651.5	1769.0	4658.5	2356.5	1090.0	354.5	137.0
St. Dev.	208.2	288.7	161.6	385.6	1959.5	5860.4	5132.5	5824.7	4908.7	2914.7	628.2	376.6
Skew	2.562	4.172	2.203	5.532	3.864	4.536	1.791	1.339	2.432	2.389	2.361	3.148
Lag 1	0.280	0.086	0.294	0.152	0.202	0.569	0.706	0.542	0.220	0.369	0.402	0.604
10% exc.	355.0	380.0	288.0	346.0	1247.0	5417.0	11517.0	13055.0	10404.0	4899.0	1470.0	514.0
90% exc.	27.0	16.0	18.0	23.0	52.0	66.0	244.0	728.0	469.0	276.0	110.0	21.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	5.77	5.54	5.69	4.36	5.29	6.86	7.74	8.45	7.84	6.93	6.13	5.32
St. Dev.	0.46	0.61	0.42	1.21	1.77	1.47	1.20	0.90	1.11	1.24	0.85	0.89
Loc. Par	-211	-166	-215	-1	29	-102	-155	-710	-102	21	-83	-69
Skew	-0.037	-0.021	-0.045	-0.027	-0.002	-0.008	-0.049	-0.050	-0.050	-0.045	-0.050	-0.023
Lag 1	0.446	0.311	0.376	0.457	0.419	0.526	0.631	0.679	0.483	0.458	0.655	0.633

Table B.6: Gumeracha Weir statistics

Catchment : Little Para	Start Year : 1969	Period : 15yrs

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	1	2	6	9	100	27	98	8	7	6	2
Max	300	310	232	748	3642	4863	5126	5995	4230	3434	863	367
Mean	131.2	108.6	84.7	146.0	378.5	903.1	1528.1	1650.1	1629.5	856.2	293.3	165.1
Median	148.0	100.0	59.0	83.0	140.0	288.0	1475.0	1134.0	1200.0	462.0	209.0	188.0
St. Dev.	106.1	93.0	74.4	180.6	905.9	1543.9	1412.1	1526.9	1511.7	100.5	243.1	116.2
Skew	0.183	0.631	0.814	2.417	3.097	1.877	0.957	1.403	0.610	1.495	0.744	0.135
Lag 1	0.628	0.645	0.656	0.053	0.935	0.634	0.541	0.768	0.382	0.400	0.619	0.846
10% exc.	272.0	255.0	216.0	258.0	324.0	4462.0	3153.0	3144.0	4195.0	2800.0	515.0	318.0
90% exc.	18.0	5.0	18.0	43.0	78.0	103.0	122.0	158.0	117.0	142.0	69.0	27.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	5.21	4.92	4.32	4.59	4.93	5.11	7.36	7.28	7.26	6.29	5.66	6.01
St. Dev.	0.55	0.61	0.79	0.94	1.31	2.04	0.74	0.80	0.87	1.09	0.72	0.28
Loc. Par	-78	-53	-14	-5	4	96	-482	-266	-327	-50	-66	-258
Skew	-0.100	-0.098	-0.091	-0.093	-0.044	-0.098	-0.100	-0.100	-0.101	-0.101	-0.100	-0.100
Lag 1	0.688	0.768	0.835	0.487	0.890	0.291	0.695	0.671	0.754	0.624	0.897	0.896

Table B.7: Little Para River statistics

Rainfall G.S. : Clarendon P.O. Start Year : 1875 Period : 114*yrs*

HISTORICAL STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Min	0	0	0	15	69	147	199	64	160	38	11	0
Max	1586	1380	1750	2905	2652	2927	2900	3109	1899	1952	1382	1834
Mean	258.8	256.4	354.7	723.3	1024.8	1168.2	1125.3	1054.6	833.3	661.6	420.1	362.0
Median	198.5	147.0	262.5	624.5	947.0	1072.5	1074.0	980.5	755.5	622.0	377.5	314.5
St. Dev.	248.1	282.3	330.8	528.4	536.7	620.0	465.8	461.4	397.4	361.8	277.4	282.7
Skew	2.192	1.428	1.464	1.514	0.634	0.502	0.503	1.351	0.645	0.832	0.883	1.703
Lag 1	0.020	-0.035	0.005	-0.070	0.019	0.169	0.114	0.230	0.035	-0.074	0.103	0.026
10% exc.	586.0	683.0	730.0	1427.0	1744.0	2161.0	1708.0	1559.0	1391.0	1160.0	780.0	676.0
90% exc.	28.0	5.0	33.0	150.0	376.0	398.0	506.0	550.0	366.0	223.0	99.0	49.0

TRANSFORMED STATISTICS

	January	February	March	April	May	June	July	August	September	October	November	December
Mean	5.42	5.07	5.74	6.72	7.37	7.57	7.76	7.32	6.87	6.84	6.29	6.05
St. Dev.	0.78	1.14	0.81	0.54	0.32	0.31	0.19	0.28	0.39	0.36	0.47	0.55
Loc. Par	-42	-23	-66	-228	-650	-857	-1274	-510	-206	-338	-177	-131
Skew	-0.091	-0.099	-0.096	-0.099	-0.100	-0.100	-0.100	-0.099	-0.099	-0.099	-0.099	-0.097
Lag 1	0.131	-0.012	-0.075	-0.031	-0.041	0.224	0.106	0.241	0.132	-0.015	0.073	0.101

Table B.8: Clarendon P.O. rainfall statistics

Appendix C

Myponga River, Serial Correlation Plots

This Appendix contains the Monthly Lag One serial correlation plots for –

- Raw Data
- Transformed Data (transformation of each month based on Table [5.4])

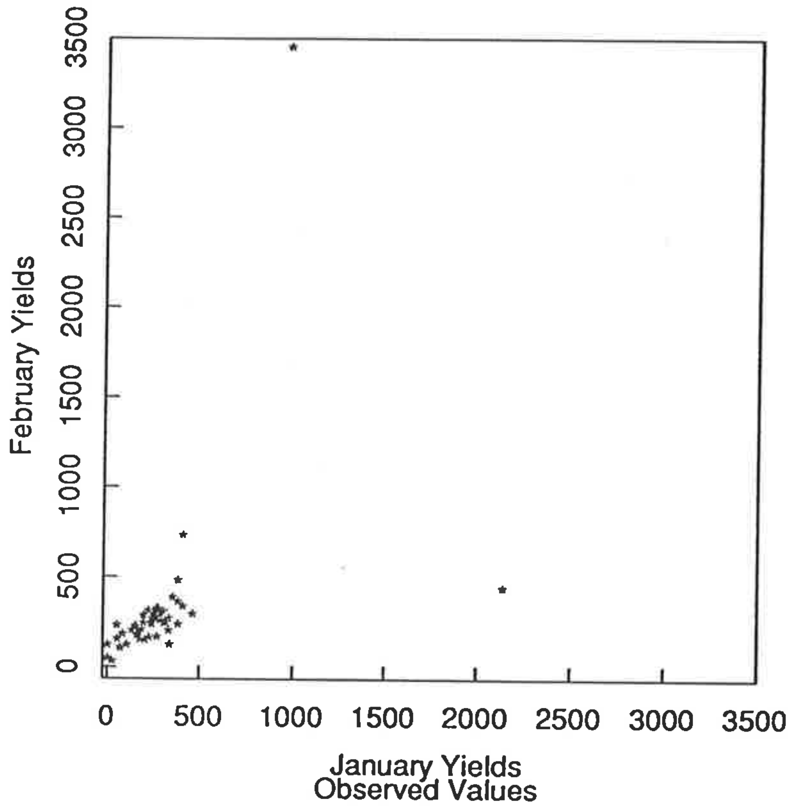

Figure C.1: Raw Data – January/February

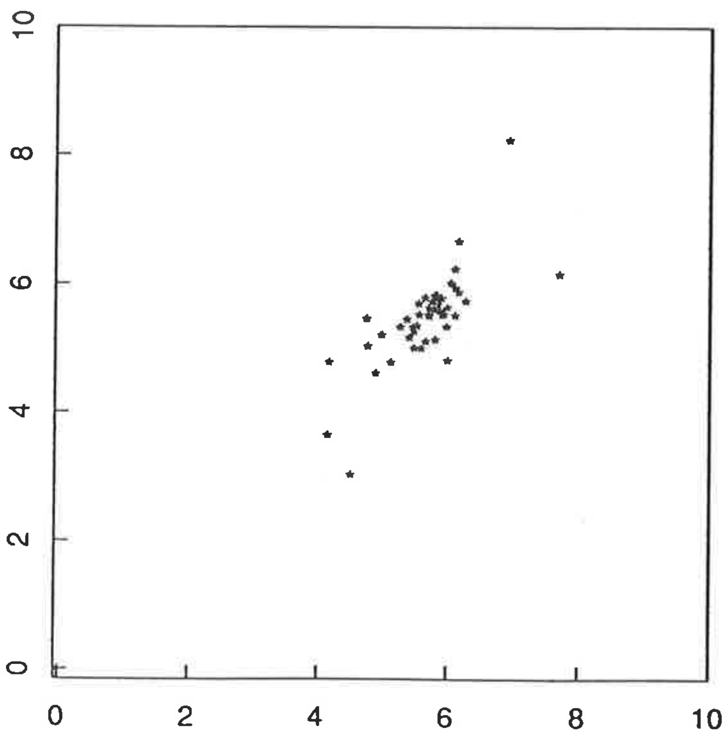

Figure C.2: Transformed Data – January/February

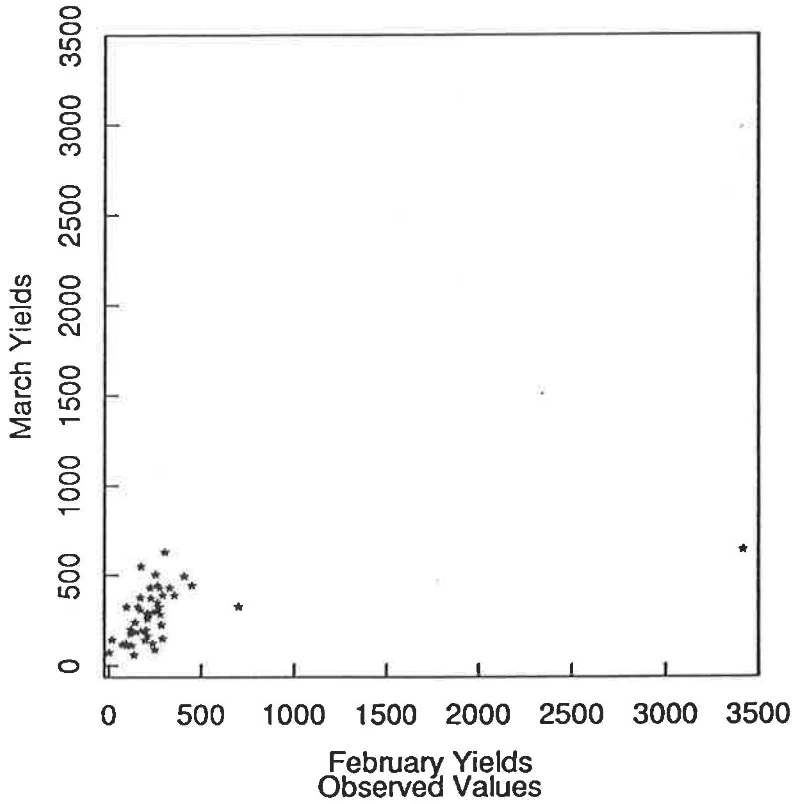

Figure C.3: Raw Data – February/March

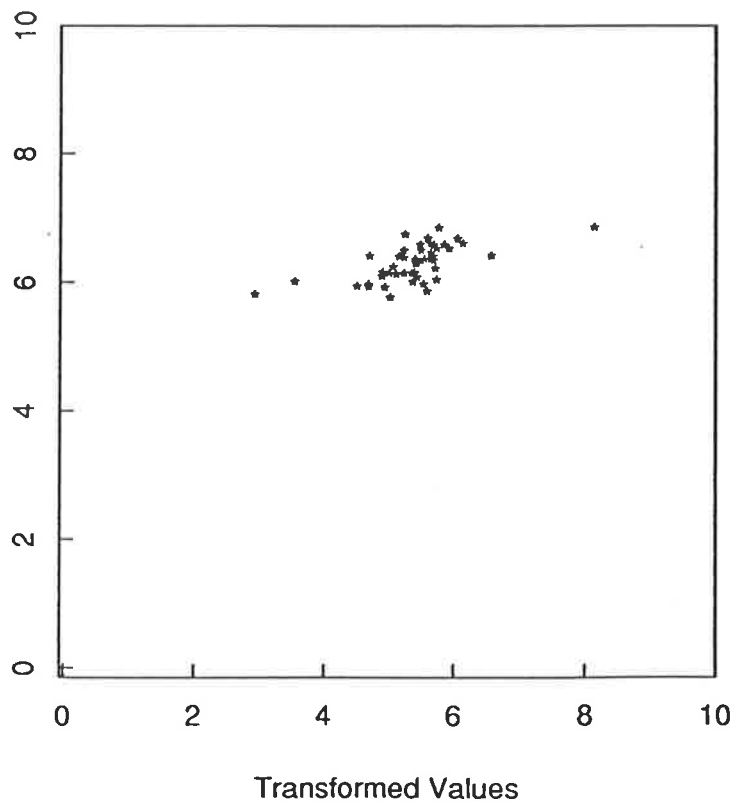

Figure C.4: Transformed Data – February/March

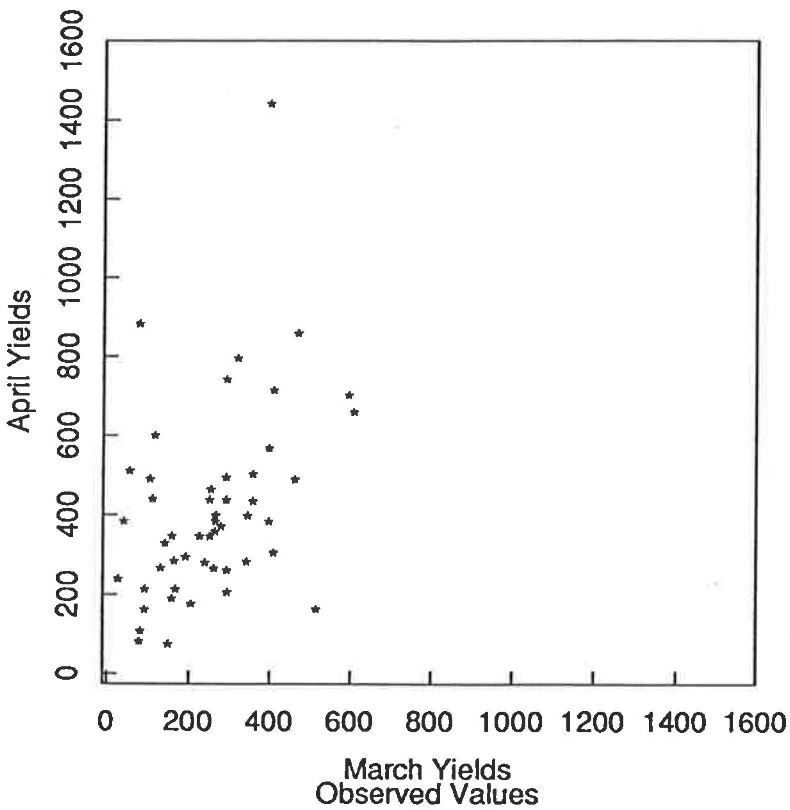

Figure C.5: Raw Data – March/April

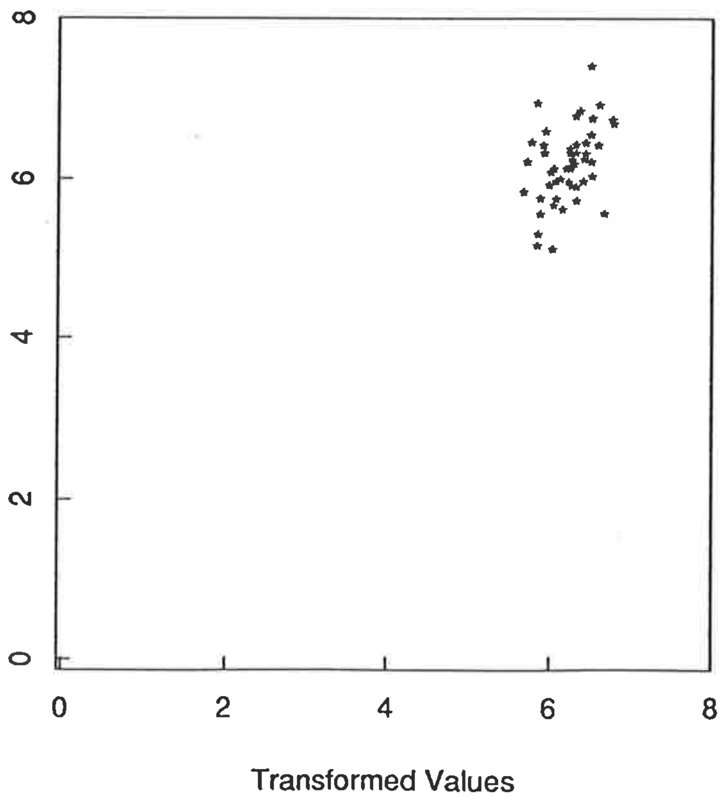

Figure C.6: Transformed Data – March/April

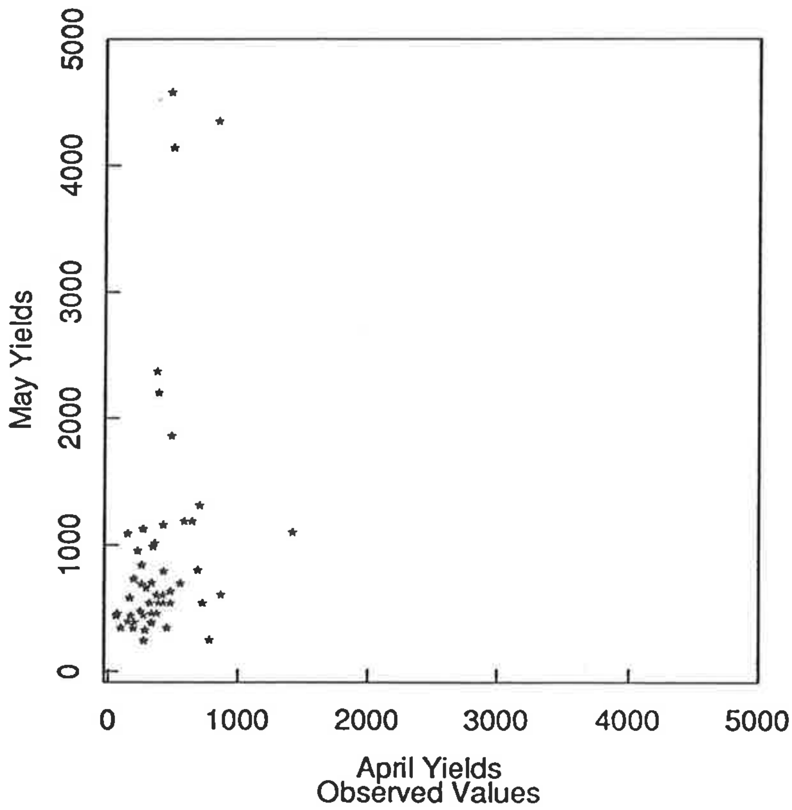

Figure C.7: Raw Data – April/May

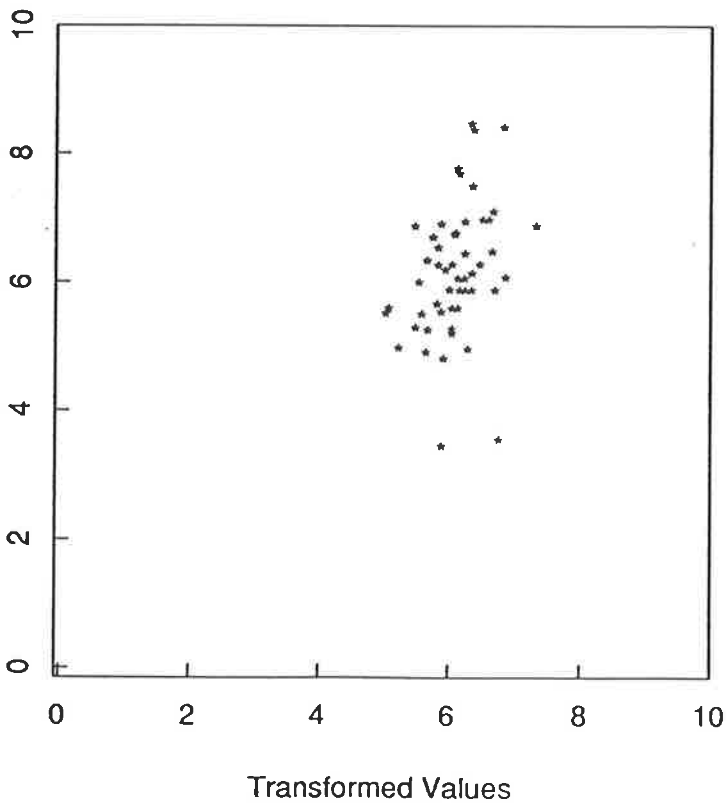

Figure C.8: Transformed Data – April/May

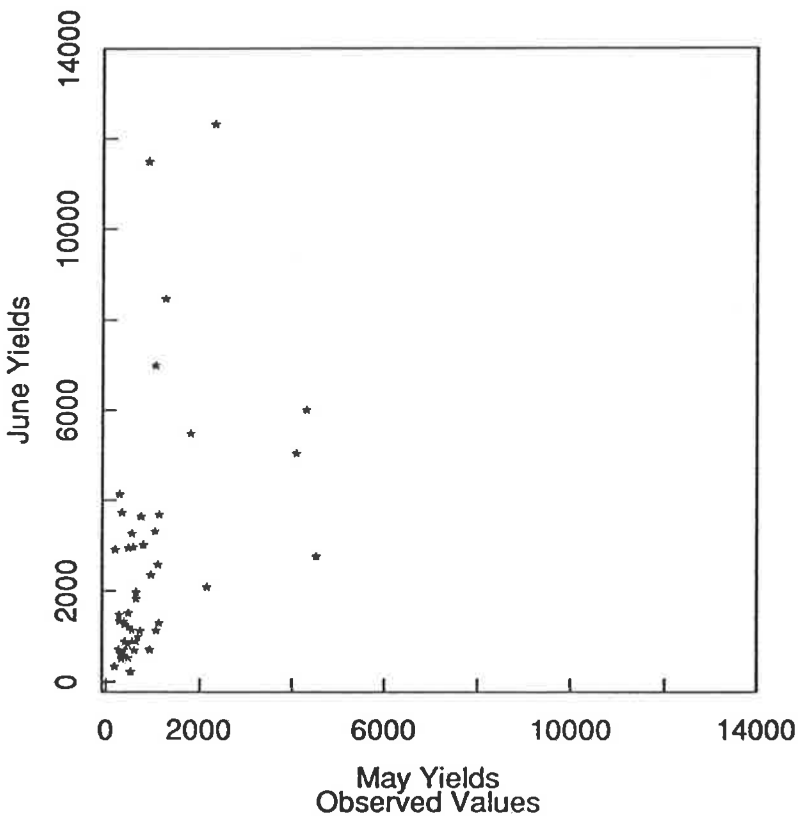

Figure C.9: Raw Data – May/June

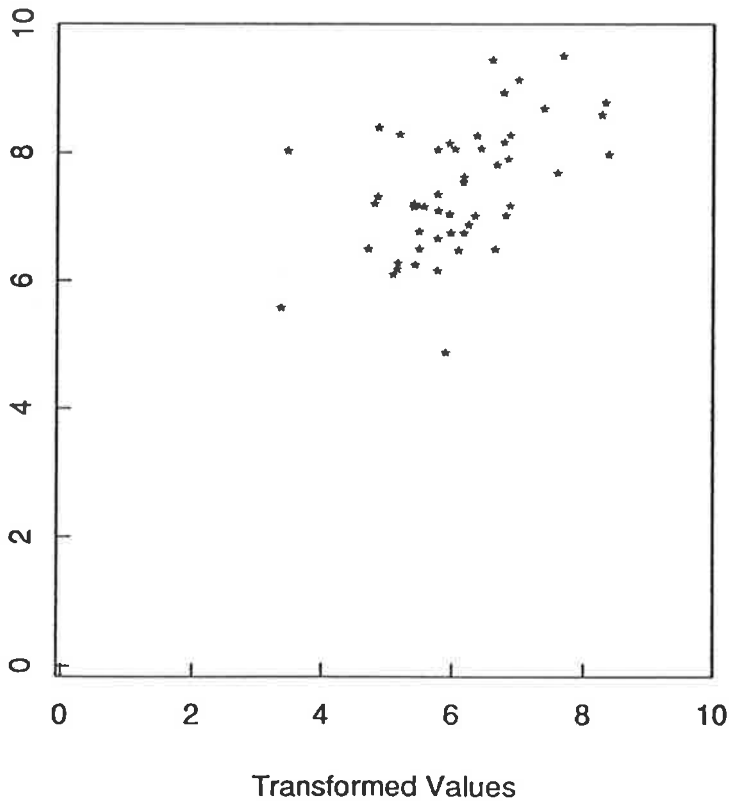

Figure C.10: Transformed Data – May/June

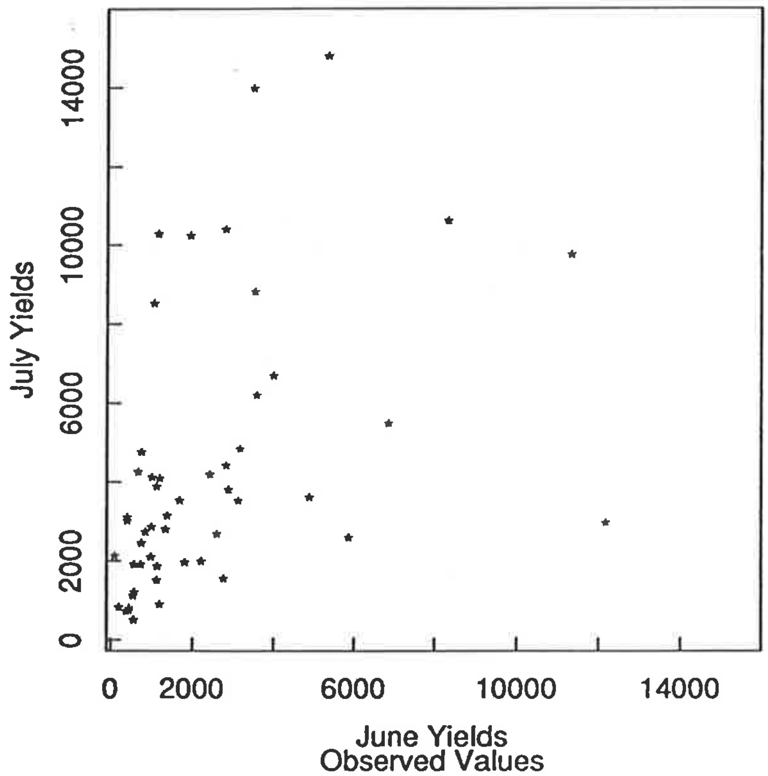

Figure C.11: Raw Data – June/July

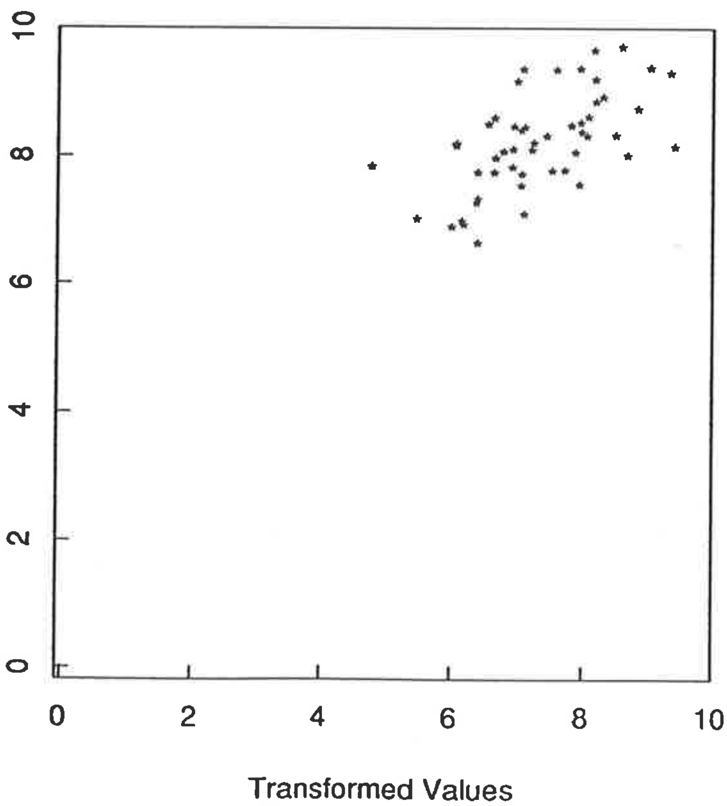

Figure C.12: Transformed Data – June/July

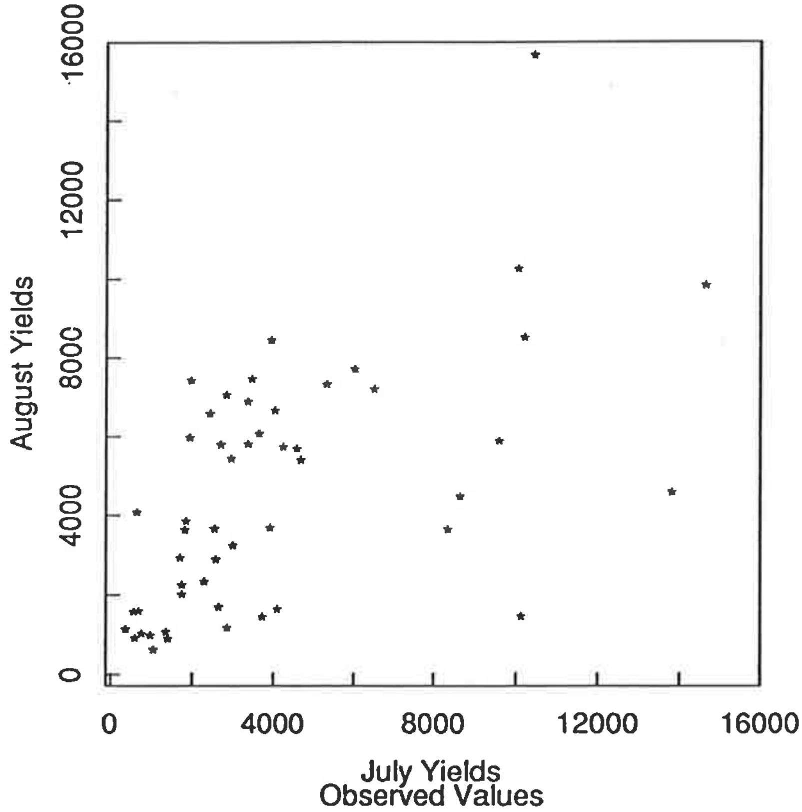

Figure C.13: Raw Data – July/August

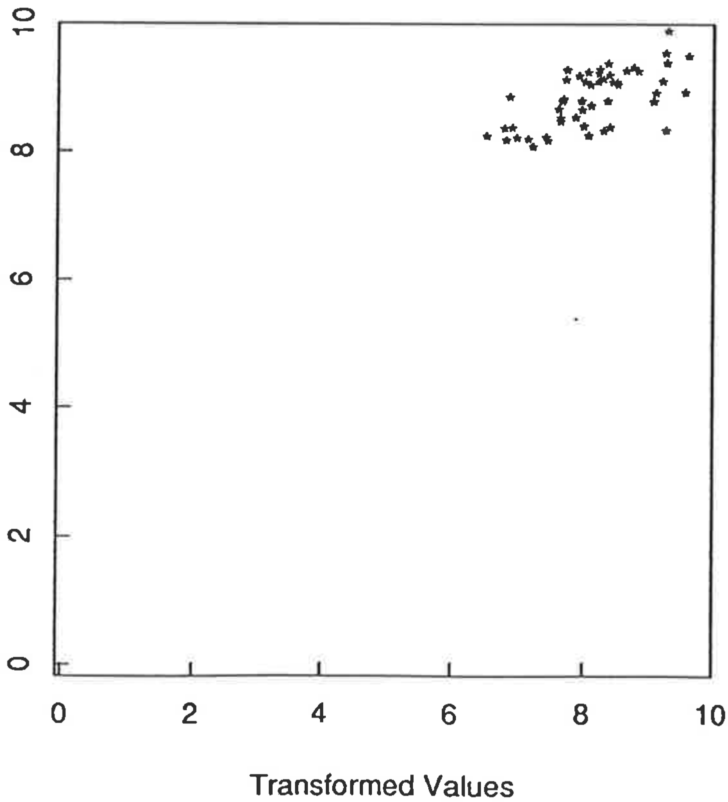

Figure C.14: Transformed Data – July/August

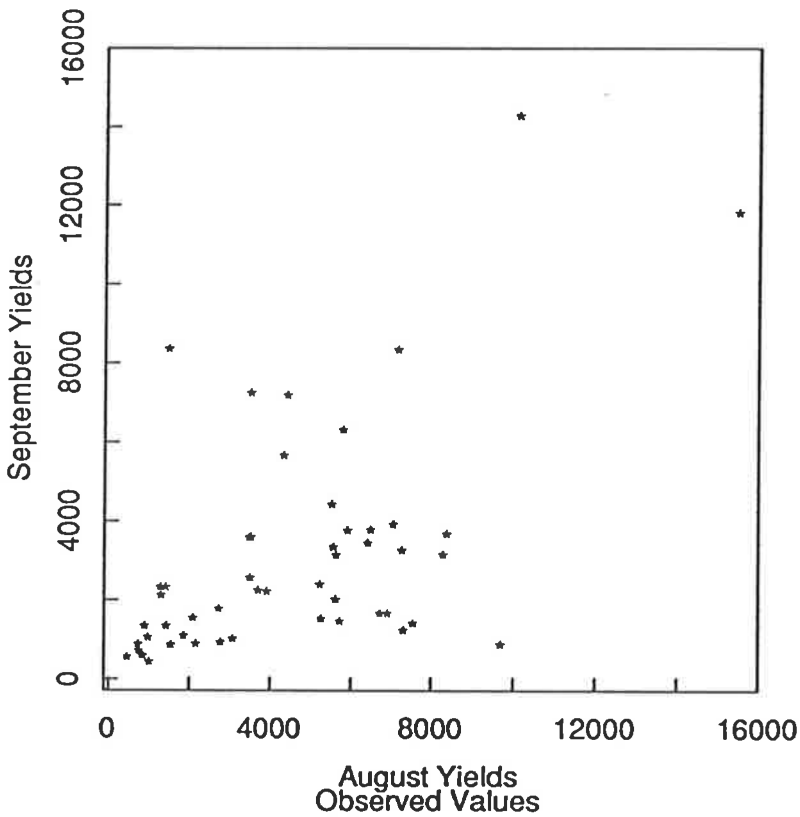

Figure C.15: Raw Data – August/September

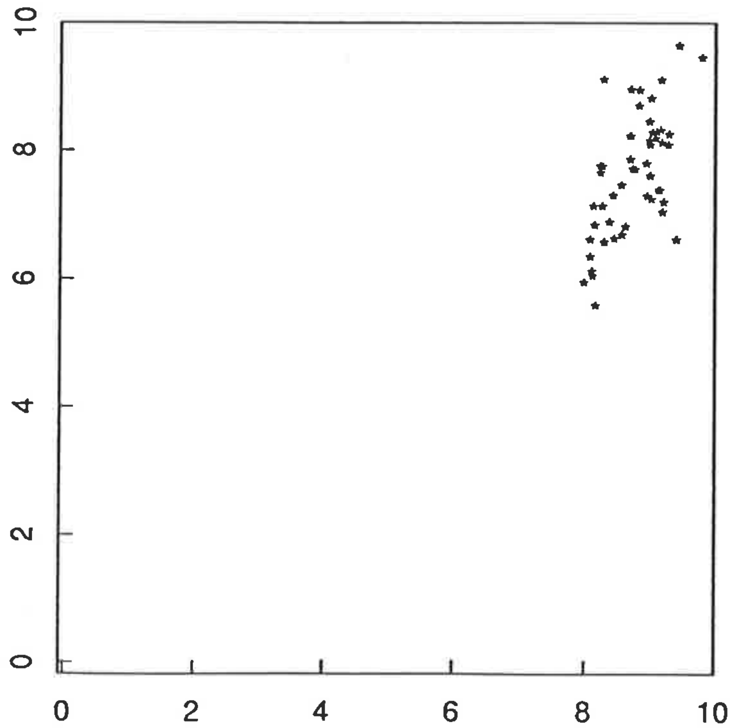

Figure C.16: Transformed Data – August/September

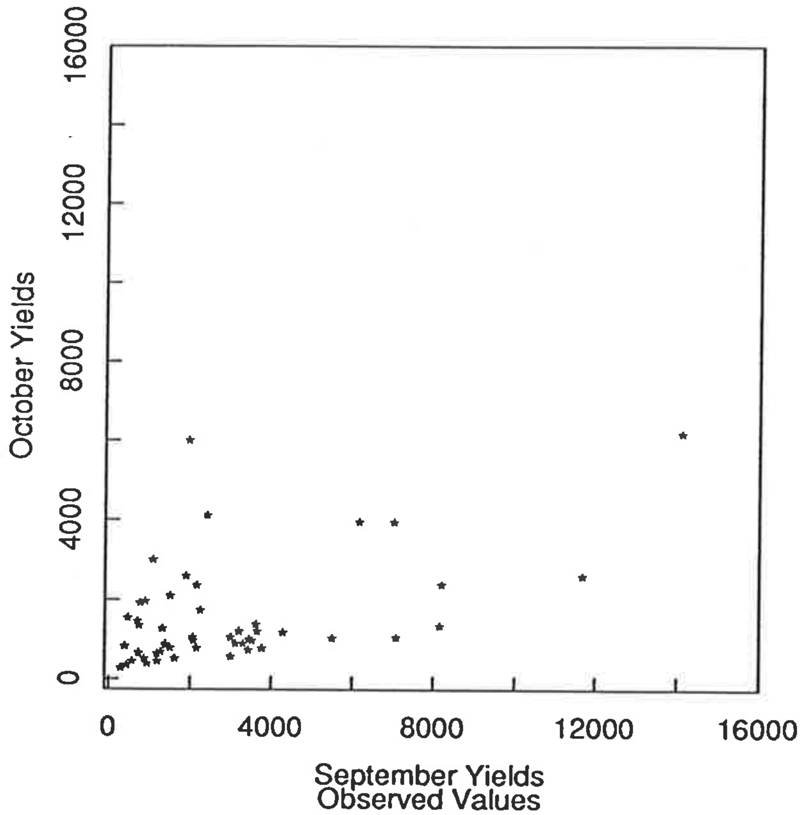

Figure C.17: Raw Data – September/October

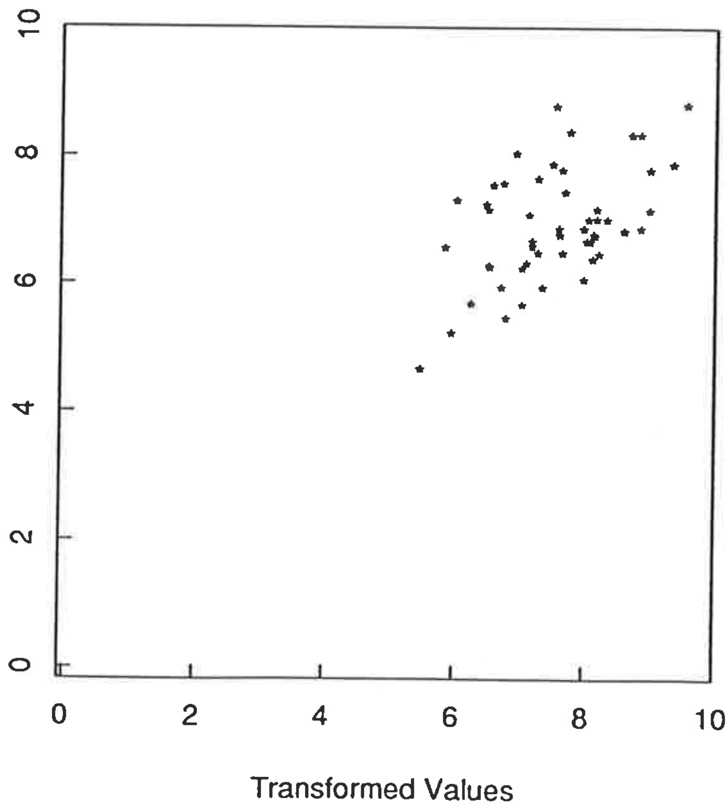

Figure C.18: Transformed Data – September/October

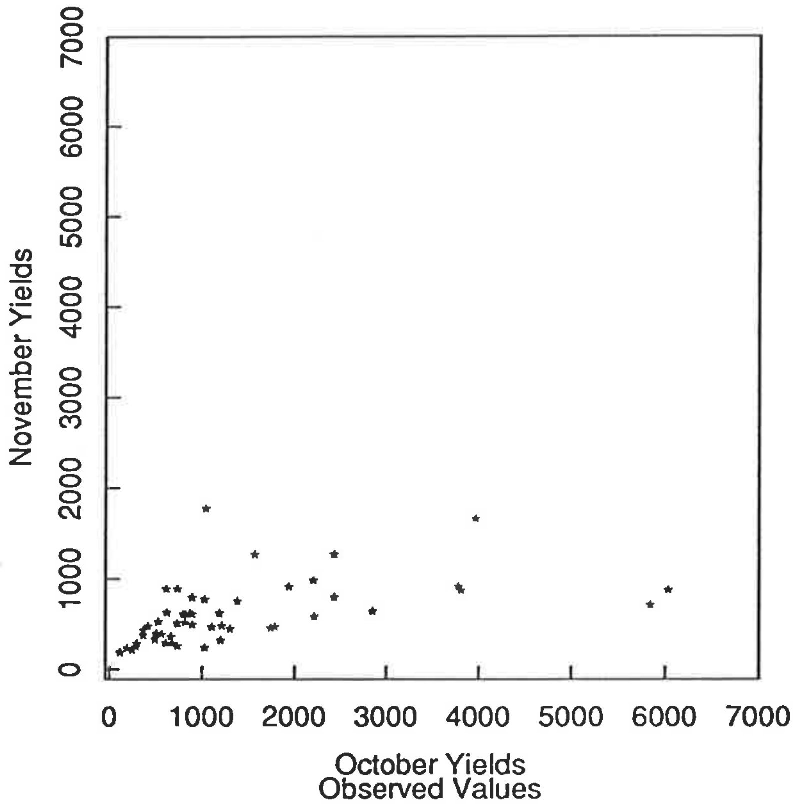

Figure C.19: Raw Data – October/November

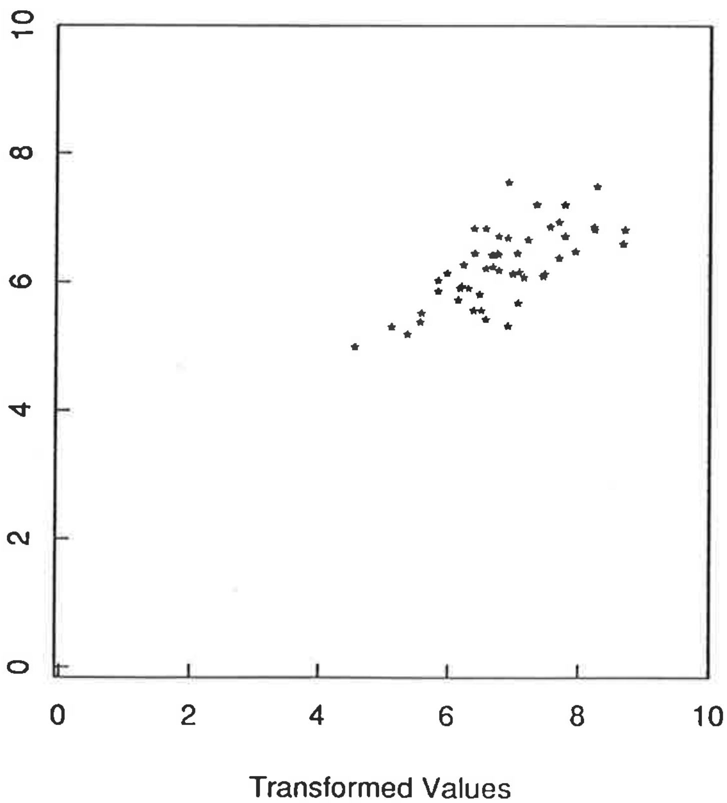

Figure C.20: Transformed Data – October/November

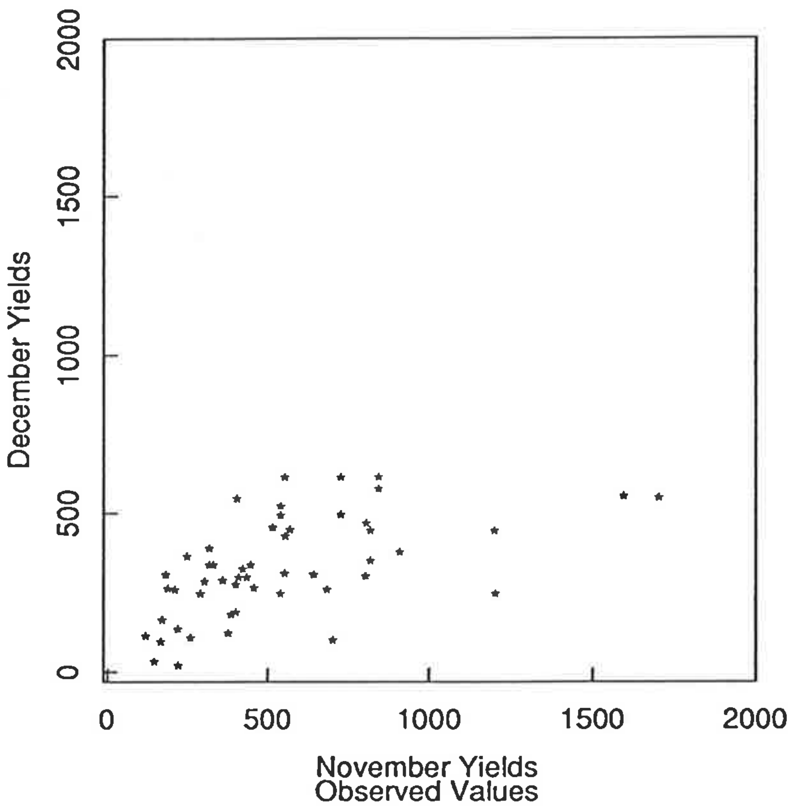

Figure C.21: Raw Data – November/December

TRANSFORMED DATA FOR ONE MONTH IN THE RAW DOMAIN AND, TRANSFORMED DATA FOR THE OTHER MONTH IN THE LOG DOMAIN

Figure C.22: Transformed Data – November/December

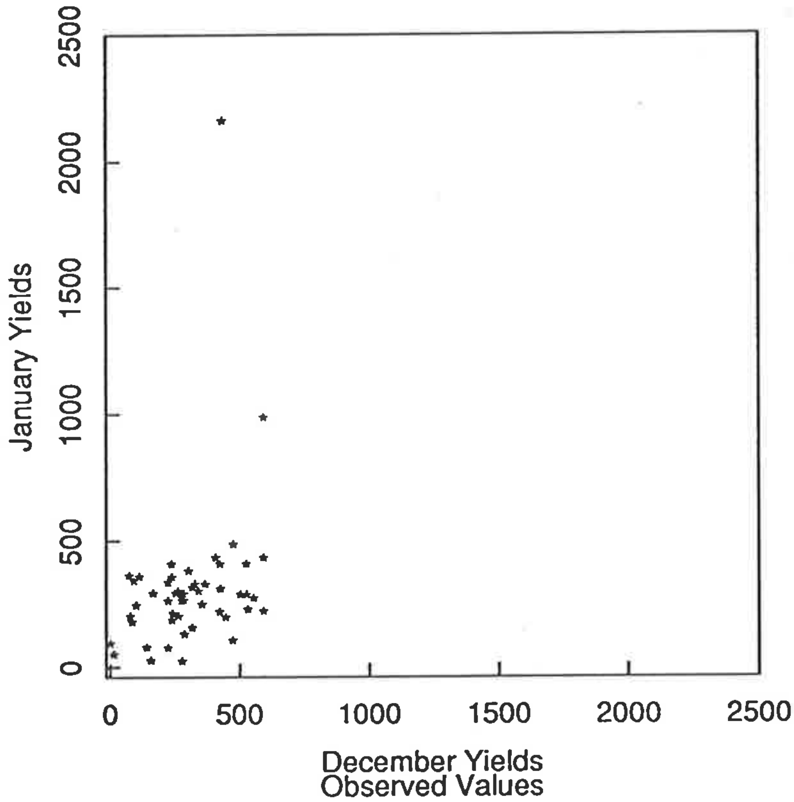

Figure C.23: Raw Data – December/January

TRANSFORMED DATA FOR ONE MONTH IN THE RAW DOMAIN AND,
TRANSFORMED DATA FOR THE OTHER MONTH IN THE LOG DOMAIN

Figure C.24: Transformed Data – December/January

Appendix D

Myponga River, Q–Q plots

This Appendix comprises a typical set of Q–Q plots, using the Myponga streamflow data as an example. For each monthly data set the following plots are produced.

- Raw Data
- Transformed Data (transformations based on data given in Table [5.4])

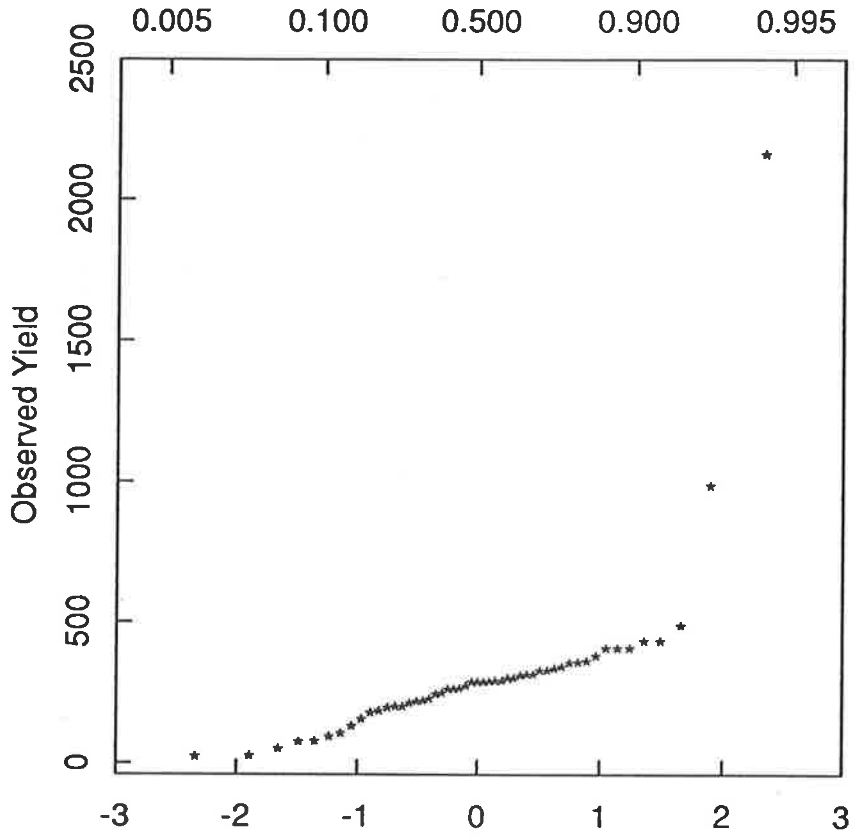

Figure D.1: Q–Q plot, January Raw data

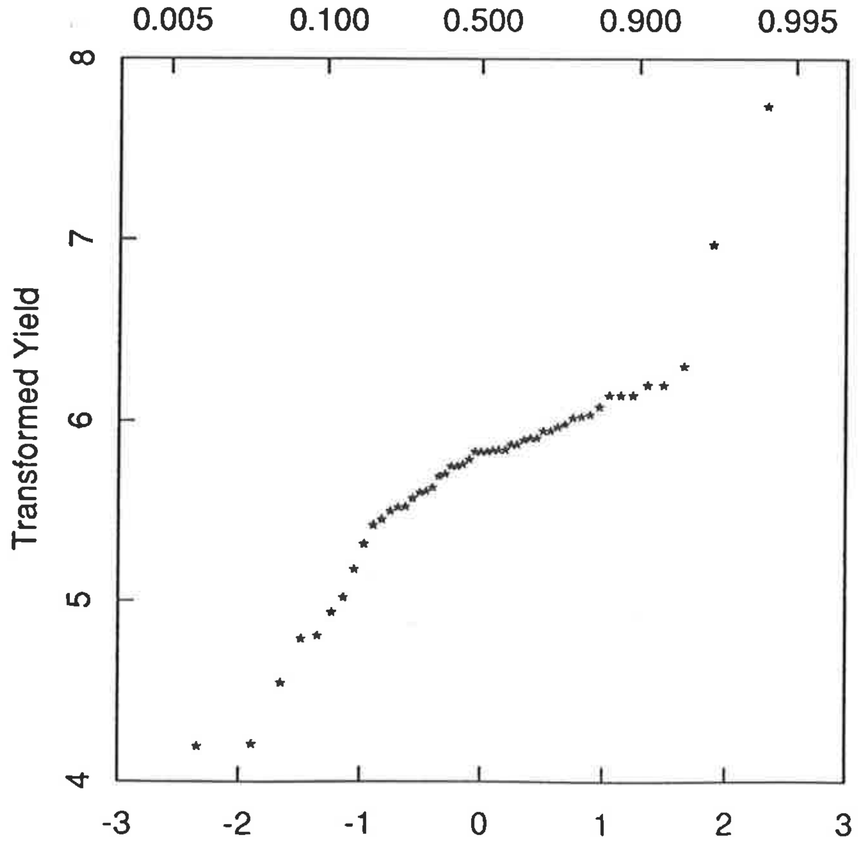

Figure D.2: Q–Q plot, January Transformed data

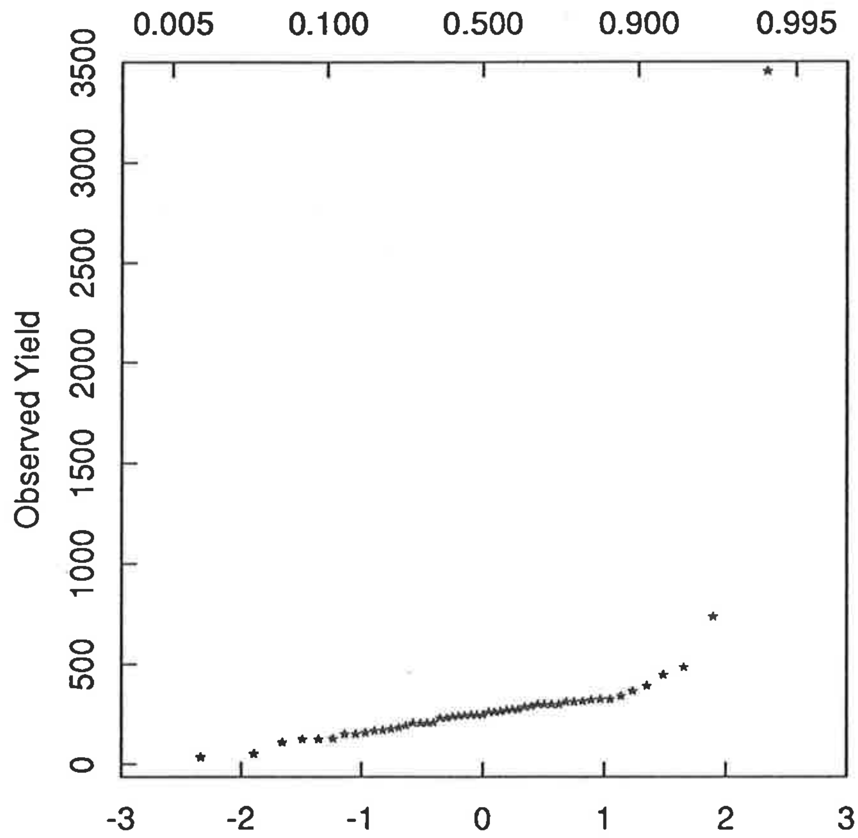

Figure D.3: Q–Q plot, February Raw data

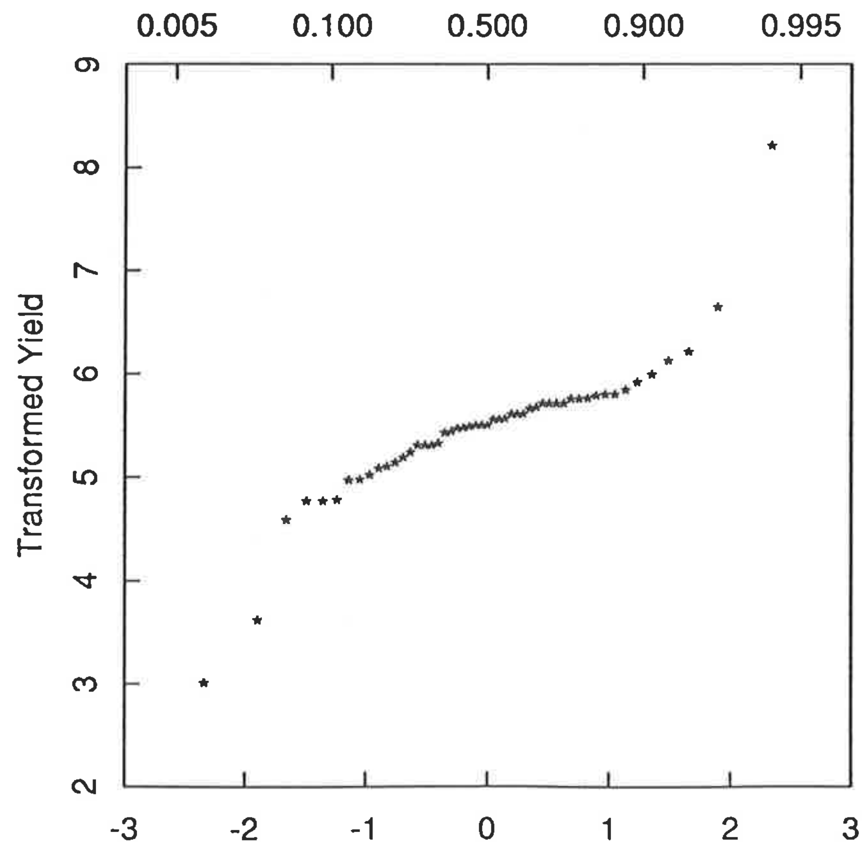

Figure D.4: Q–Q plot, February Transformed data

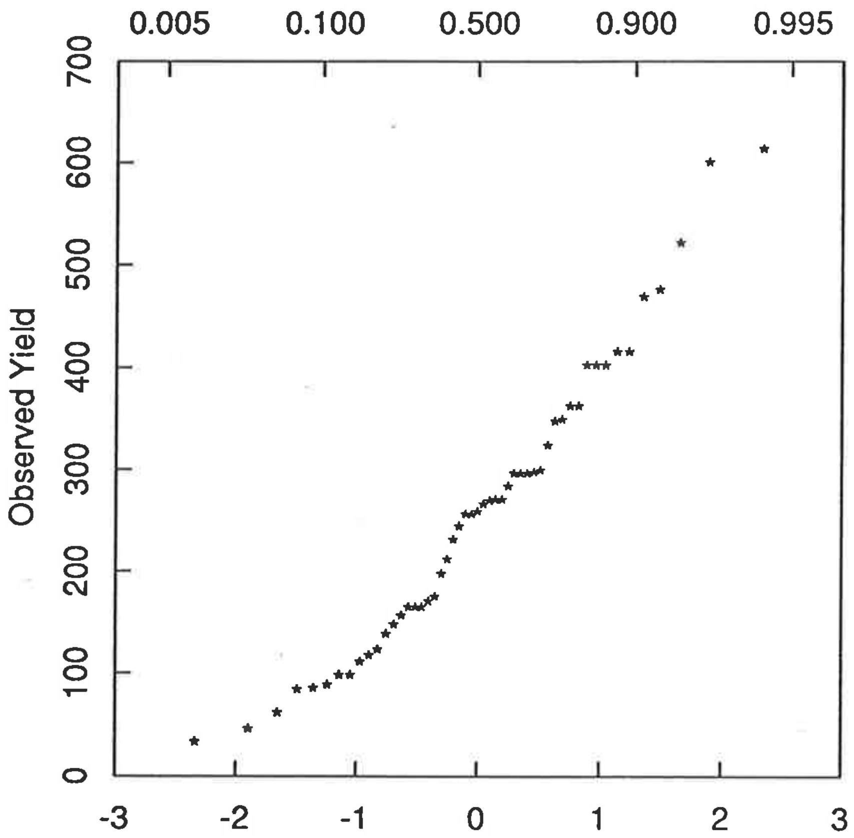

Figure D.5: Q–Q plot, March Raw data

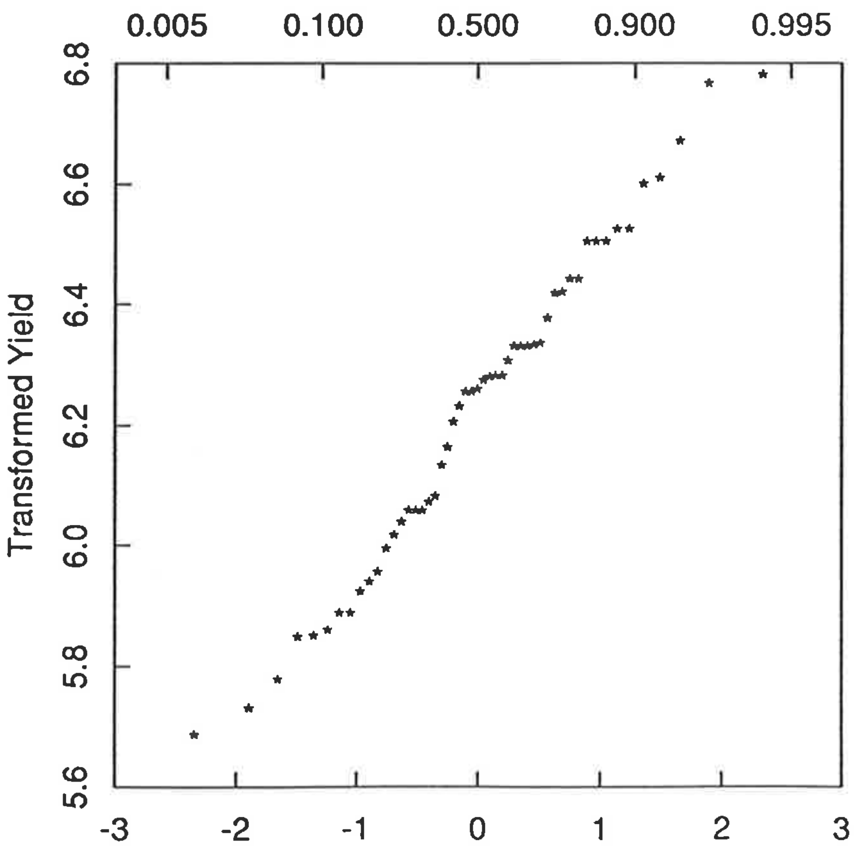

Figure D.6: Q–Q plot, March Transformed data

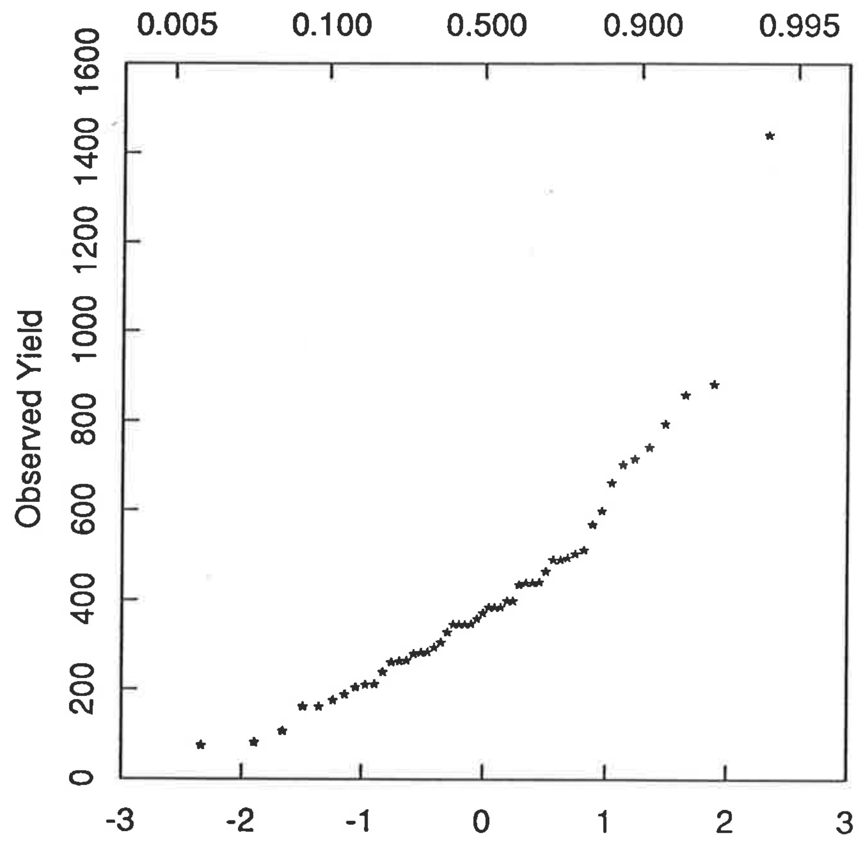

Figure D.7: Q–Q plot, April Raw data

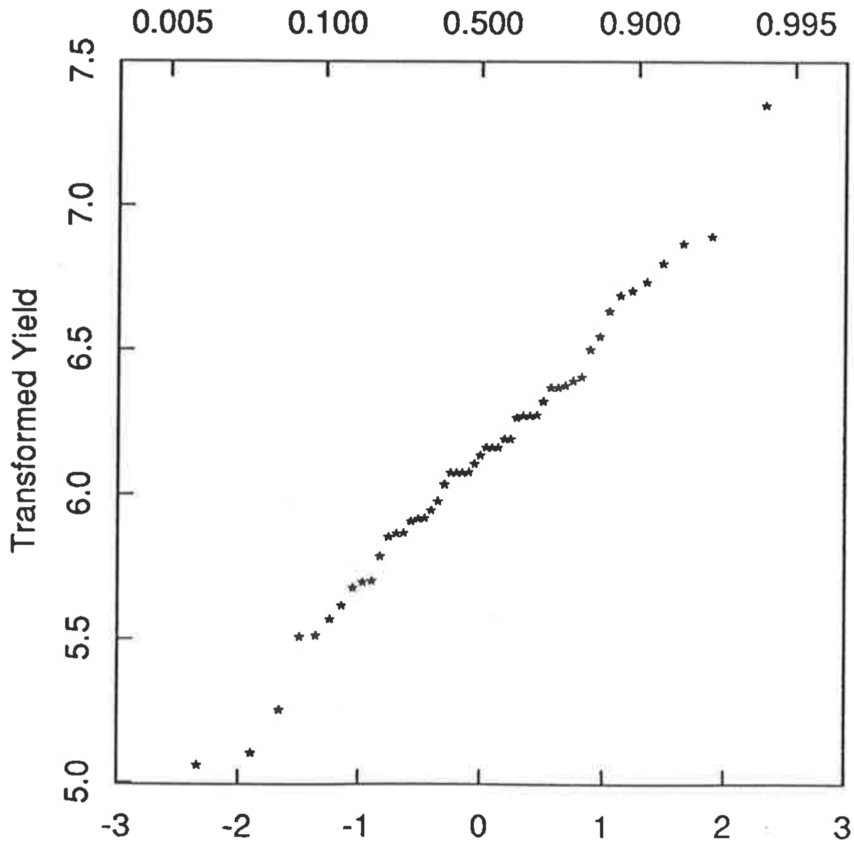

Figure D.8: Q–Q plot, April Transformed data

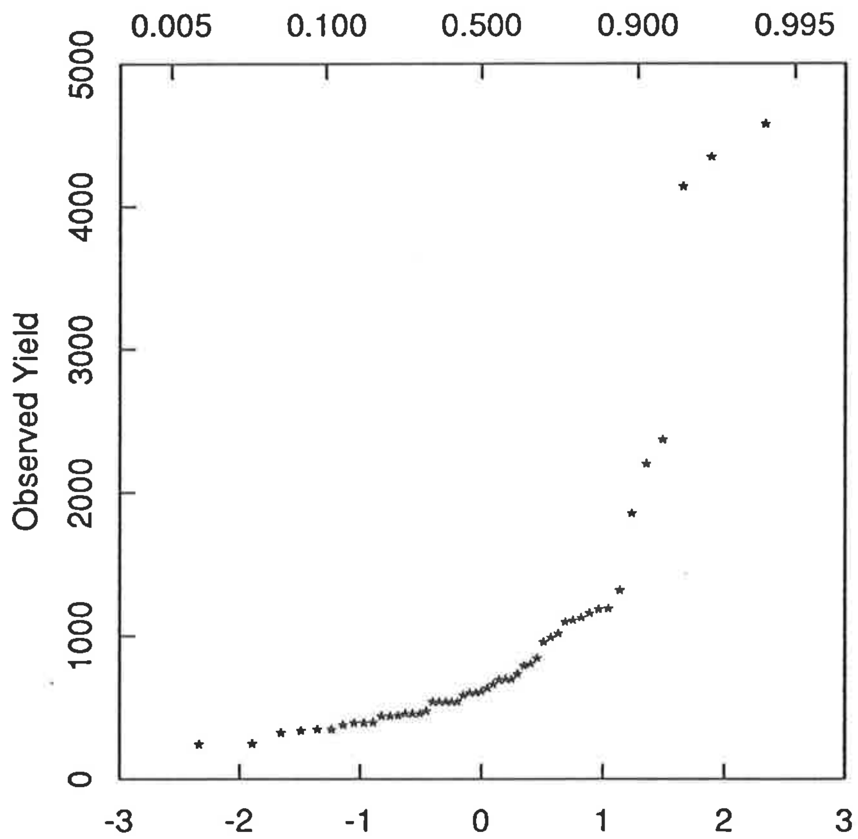

Figure D.9: Q–Q plot, May Raw data

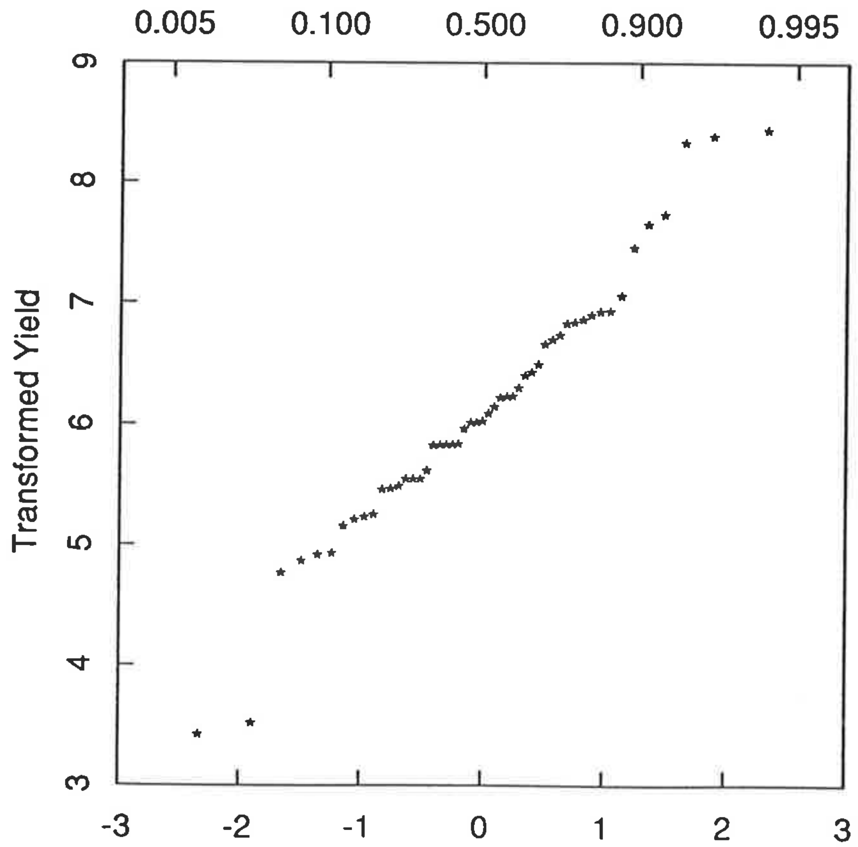

Figure D.10: Q–Q plot, May Transformed data

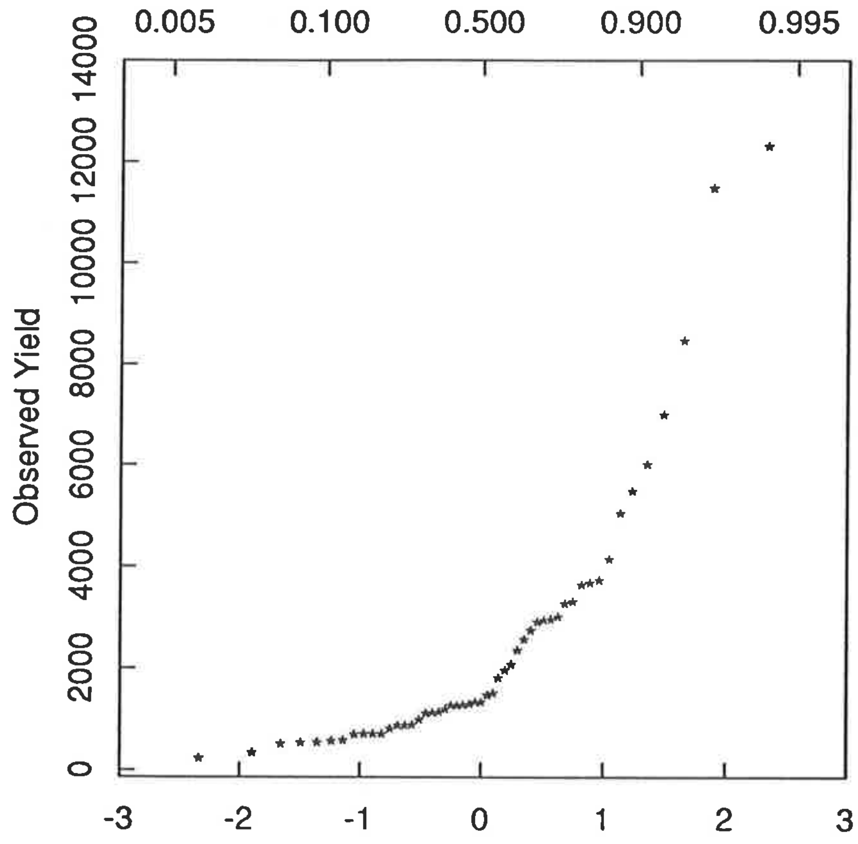

Figure D.11: Q–Q plot, June Raw data

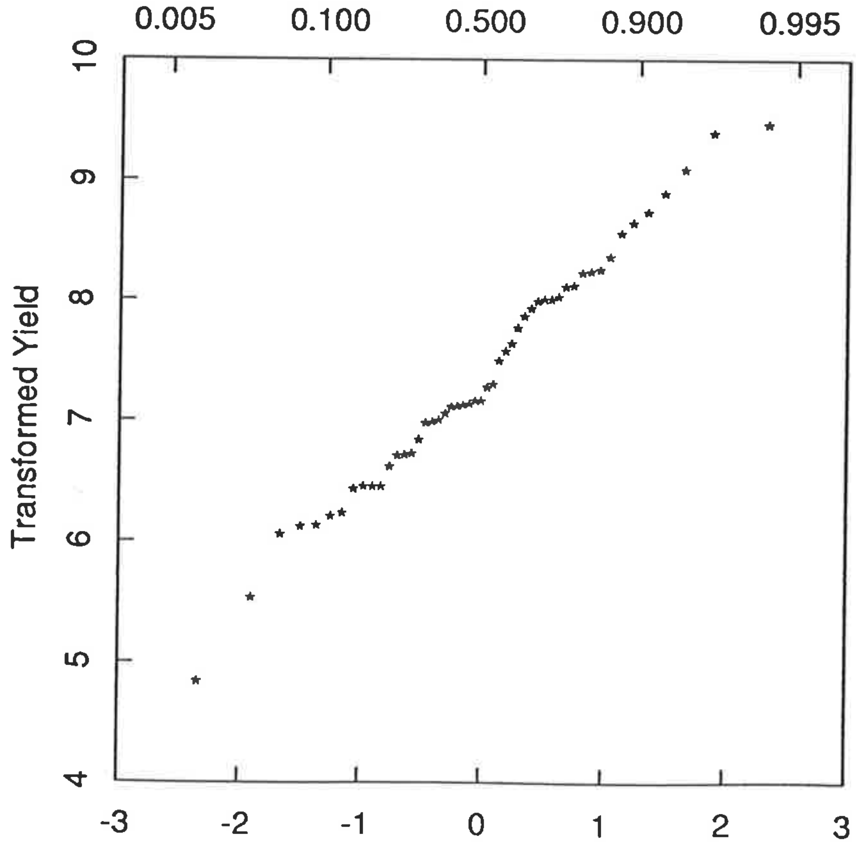

Figure D.12: Q–Q plot, June Transformed data

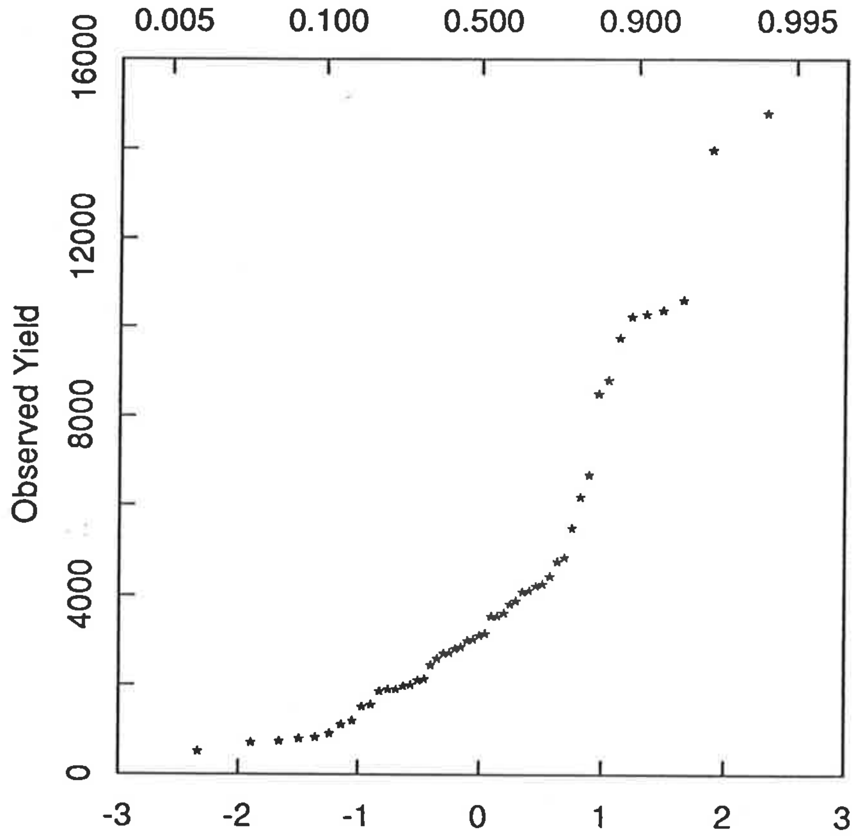

Figure D.13: Q–Q plot, July Raw data

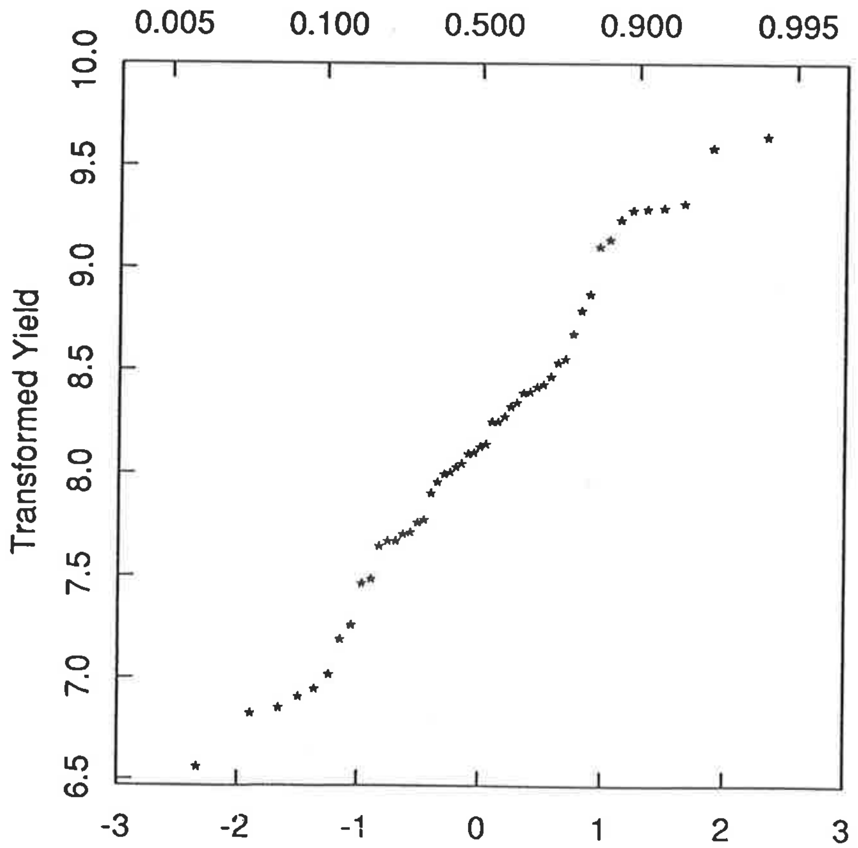

Figure D.14: Q–Q plot, July Transformed data

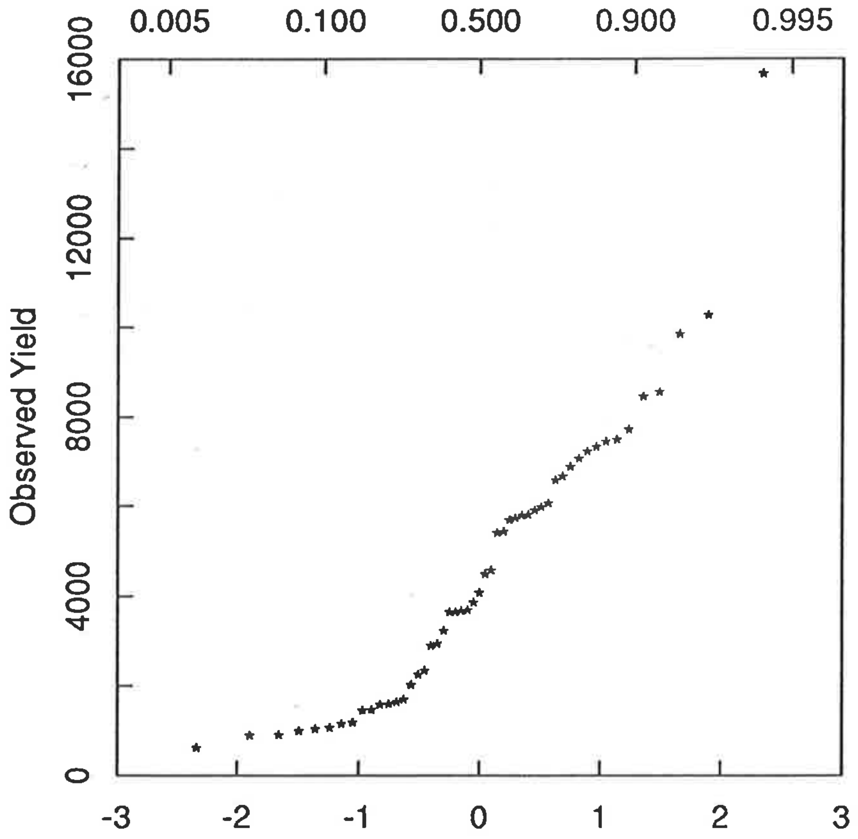

Figure D.15: Q–Q plot, August Raw data

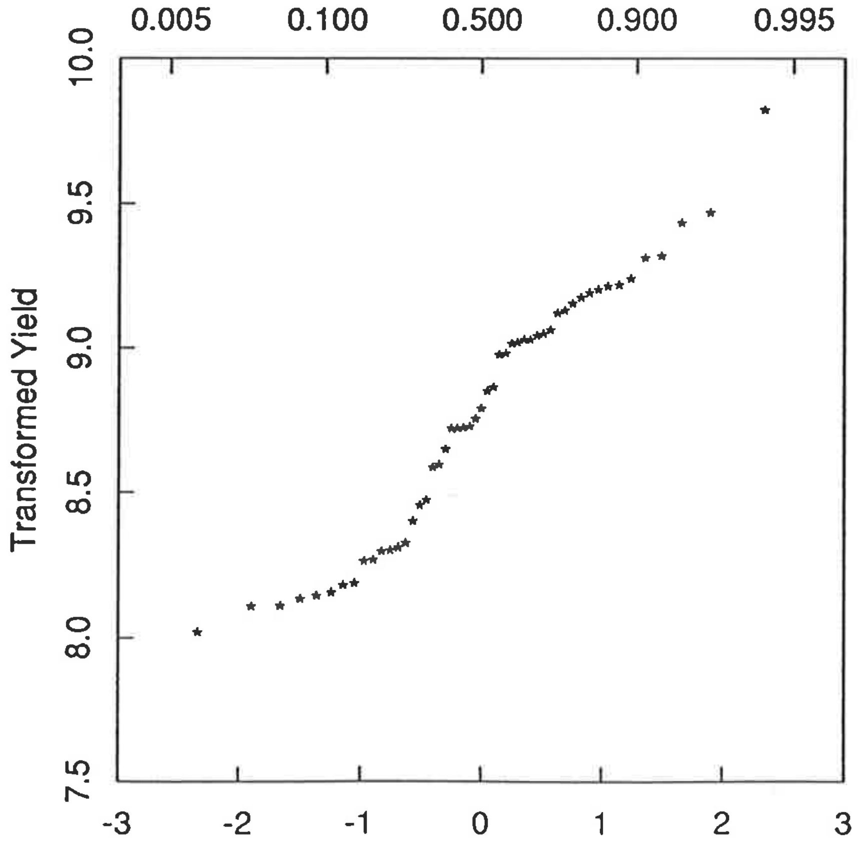

Figure D.16: Q–Q plot, August Transformed data

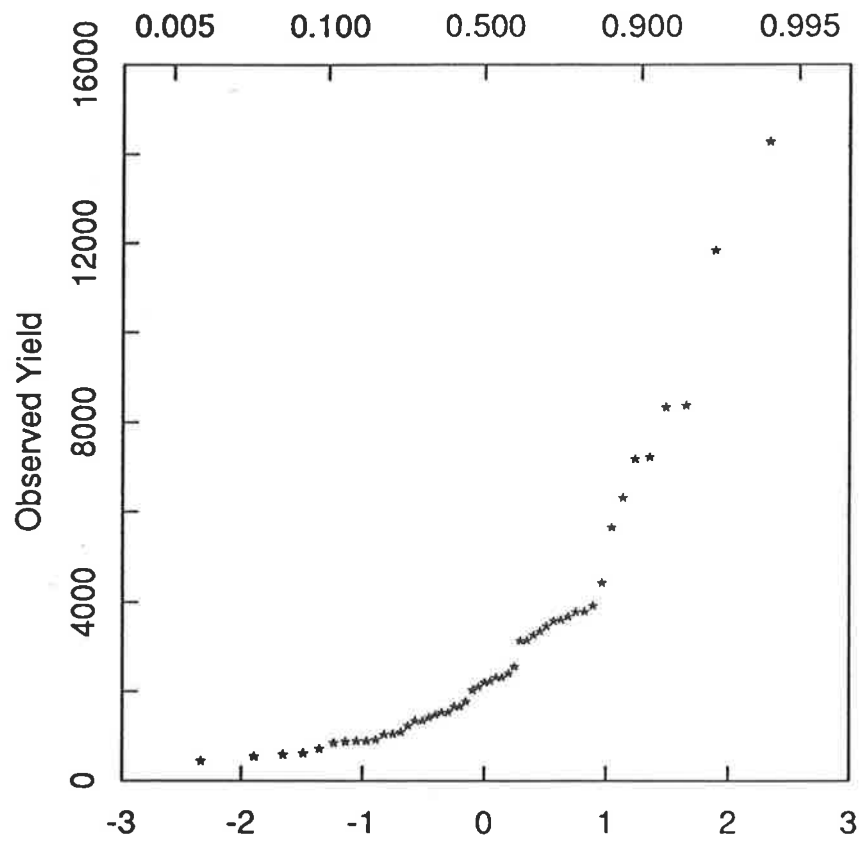

Figure D.17: Q–Q plot, September Raw data

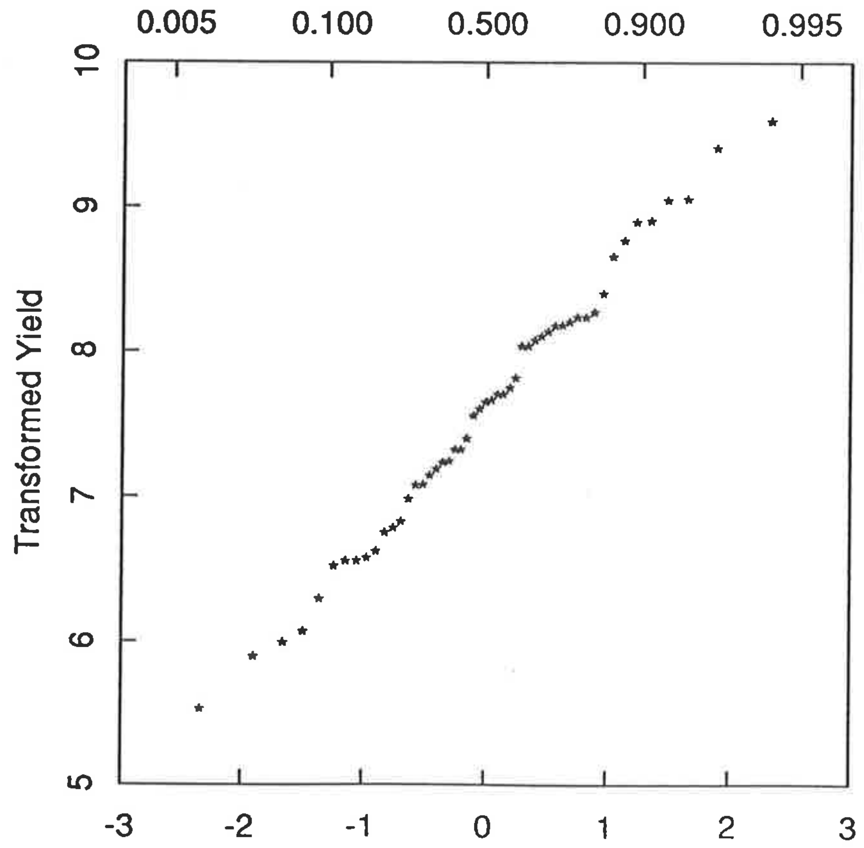

Figure D.18: Q–Q plot, September Transformed data

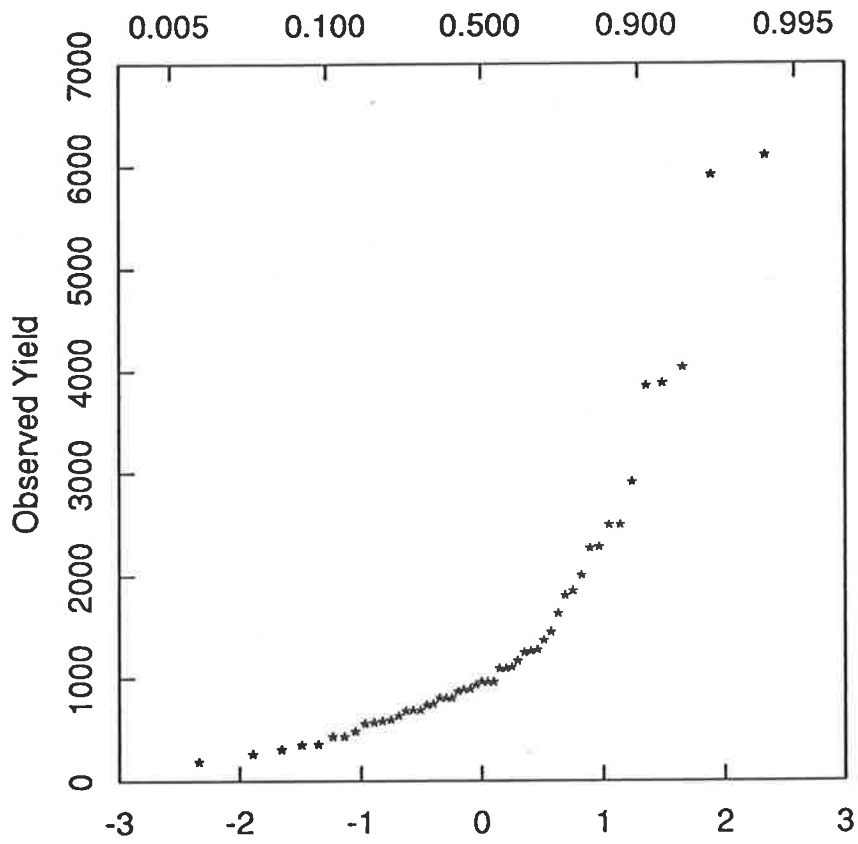

Figure D.19: Q–Q plot, October Raw data

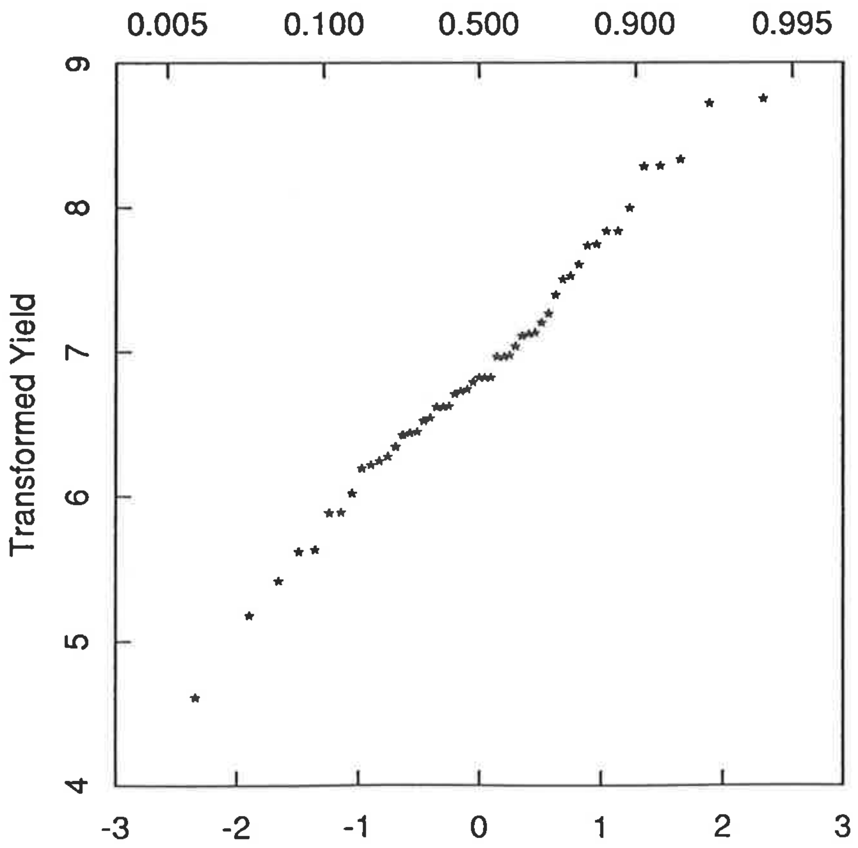

Figure D.20: Q–Q plot, October Transformed data

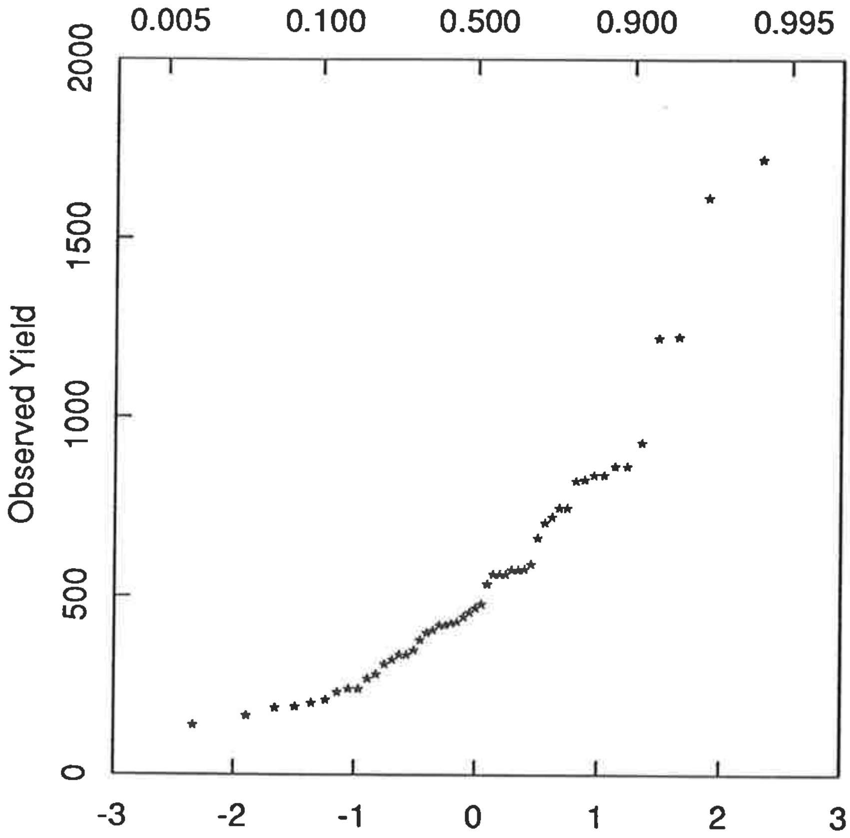

Figure D.21: Q–Q plot, November Raw data

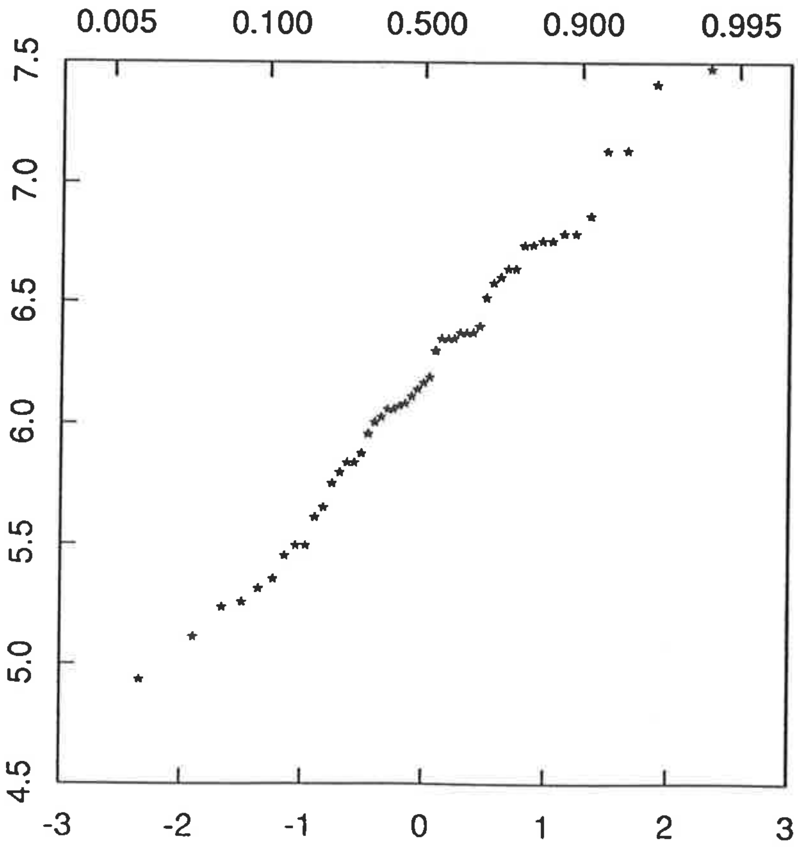

Figure D.22: Q–Q plot, November Transformed data

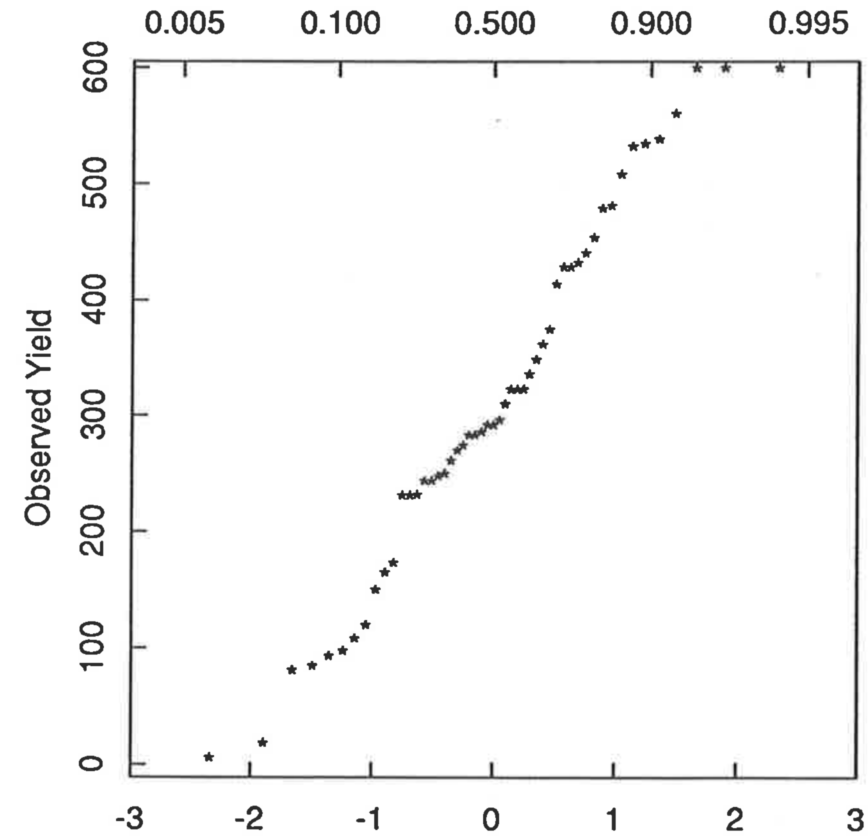

Figure D.23: Q–Q plot, December Raw data

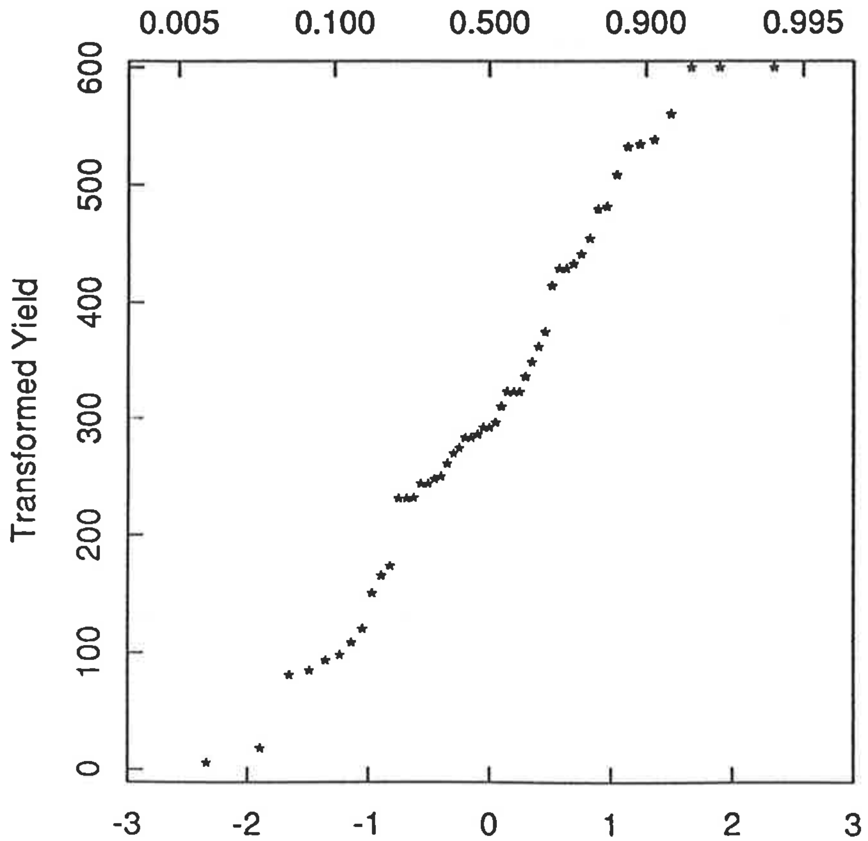

Figure D.24: Q–Q plot, December Transformed data

Appendix E

Streamflow Data Sets

This appendix contains the streamflow data used for parameter estimation for the data generation models. The values represent reconstructed data for each station, where truncated values have been replaced with the negative yields calculated from the water balance equations.
Thus, the data sets as supplied from the E.&W.S. department are as shown, except that the negative values were zero.

Values are given in (Ml's).

Warren River

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0	69	20	62	83	1652	3643	13319	2021	279	1017	23
14	−56	137	105	113	22	360	276	187	93	−44	−2
387	3	14	3	−8	70	934	1241	5306	1203	33	−28
−41	−3	−40	94	593	4983	9876	10212	10257	385	71	36
−28	69	−32	11	2	82	122	2414	1559	352	12	−69
−13	15	−29	−2	46	−38	33	−2	−30	5	24	−47
−43	5	−8	−16	30	47	22	107	242	145	205	−34
5	504	37	0	65	416	3907	1749	259	58	10	81
−38	46	49	55	0	229	4069	5401	1945	1050	430	43
−46	2	8	143	137	394	1427	5632	1026	753	1966	119
0	9	6	−36	49	−17	63	82	55	1886	548	−29
9	−57	10	−47	410	770	1018	1840	370	165	39	20
−2	−41	−50	22	450	4095	14159	9898	940	1507	62	36
19	−22	−36	16	1080	4539	3602	4380	3934	1413	1146	183
−9	2	−14	14	49	1516	7027	4838	4923	423	105	42
11	−5	−41	82	24	99	365	211	165	61	0	7
−2	33	−12	30	793	10265	2888	15025	2264	713	120	7
71	14	49	82	752	7548	13304	12612	8468	1313	105	33
−35	−44	−50	−15	−26	−5	3250	4659	3101	1434	241	33
8	21	14	46	350	1560	2440	8070	5185	5231	21	−15
−59	34	−11	−35	−67	−62	20	83	−26	−10	−56	−10
−38	96	3	93	5544	398	4426	6596	9081	508	92	−52
−35	−25	−15	99	15	90	643	846	1401	24	25	−48
3	−14	−8	−26	230	466	999	135	410	3412	−89	−55
−99	−72	−327	187	172	10872	15901	11020	4602	−7	−53	−87
−210	31	−105	−21	7	67	4615	2970	4557	5820	186	−29
−115	−198	−103	−61	−34	4	159	1585	564	−63	−61	−63
−89	−32	−6	−46	−14	56	2177	1832	3854	955	−7	237
−47	−81	−12	−101	−46	20	74	171	125	30	−25	−54

continued over

Warren River *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
2	27	31	85	1655	6855	5194	10379	1251	6175	911	124
98	357	89	113	205	98	2388	639	1234	12	−39	−28
41	−26	33	13	126	168	1236	6499	4668	585	39	12
−31	−58	−31	122	2053	4188	1758	10656	11432	1196	345	62
109	55	22	0	0	0	191	3161	482	111	27	15
20	154	53	108	120	391	2169	4345	8219	1289	69	147
200	78	29	80	412	217	10262	10239	8591	12916	1118	63
0	−52	0	0	0	0	0	4758	3168	1467	254	0
−23	28	16	39	67	−11	42	108	72	154	78	7
18	12	73	5	72	124	209	167	93	132	78	54
6	0	1	78	111	136	4645	6033	5115	312	36	187
38	37	46	99	65	40	82	1320	8805	8281	576	147
52	3	153	157	123	982	5069	1664	579	3617	1082	136
83	25	120	64	121	9472	15224	20708	3081	1238	273	129
131	13	98	73	51	100	58	44	0	1	28	0
20	0	273	218	254	228	2353	5520	9975	448	319	209
34	51	188	152	222	198	335	5517	2218	357	0	0

South Para River

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
30	119	103	166	303	2727	3902	29208	6956	1196	1658	52
34	−75	137	218	256	211	1148	901	514	208	16	77
702	47	161	43	94	244	2009	2150	11779	5364	183	−4
−30	3	−32	172	1399	14952	20501	29312	27248	2831	253	84
−5	81	−39	45	45	275	1120	5523	2269	567	12	−83
−35	15	−40	2	184	19	348	136	−37	15	24	−52
−46	5	−11	−18	30	223	75	402	663	403	330	−42
2	901	75	−25	214	982	6505	4360	830	182	4	104
−66	31	72	305	40	690	8394	13080	12055	5300	1715	42
−74	−16	−9	349	402	1150	3488	15000	4432	2049	5991	213
−18	747	−8	109	167	125	310	382	347	4077	856	−51
−19	−48	−11	−43	900	2039	2028	2015	555	310	54	16
−23	−54	−69	63	3437	10795	27414	15452	1595	2017	244	47
21	−42	−64	86	3609	11456	4843	4796	7282	3113	3384	901
75	9	−76	45	166	3254	12345	8463	9898	2071	237	225
3	−27	−53	143	243	615	1322	683	541	288	108	59
1	44	−2	23	1466	20407	8167	26168	5998	1226	357	56
58	−29	39	230	1505	17693	22720	16044	13227	2088	492	96
−60	−266	−28	8	−19	−44	5883	7753	5502	2844	698	139
64	116	40	210	1085	3466	4291	12671	8764	8838	107	17
−13	160	106	33	36	27	354	850	359	198	−65	310
18	424	55	442	11111	3408	6651	10952	15391	1061	505	−136
−48	−30	−28	769	−6	848	2018	1732	2521	132	151	−56
37	−48	6	−117	1046	1405	2051	635	819	4991	−69	−18
−71	−209	−447	594	740	19770	29532	22301	11013	130	−27	−145
−278	−31	−211	368	124	666	8039	5287	6865	9806	654	−26
−263	−207	−226	159	832	106	1237	3446	1289	263	433	333
−7	31	229	−160	325	523	5179	3482	6564	2252	156	690
−220	0	−119	−13	32	71	363	690	285	162	−202	−216

continued over

South Para River *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
−116	−75	−8	202	4440	13512	8906	17795	2210	9065	1487	592
188	999	149	319	647	484	5770	1523	2356	144	196	211
459	−49	108	322	407	752	3205	12014	6773	416	555	494
451	425	361	1244	5512	9567	3537	18881	16562	2866	1255	682
979	650	371	605	373	386	1434	6638	1204	358	27	15
54	498	58	408	377	1328	4811	5179	10295	2028	159	402
555	535	138	605	1009	618	13299	10573	8324	18694	1873	63
142	0	462	160	2367	476	4900	3490	2393	6777	2518	158
0	260	16	96	245	134	361	550	302	579	263	103
81	12	202	5	490	533	717	514	459	272	328	54
6	8	31	121	324	632	6625	7881	6801	656	213	187
224	37	64	171	358	179	277	2587	12574	12060	1482	271
52	30	212	640	323	2343	6927	1879	845	4415	1464	136

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
225	133	158	330	330	369	568	1425	2165	2204	911	356
238	211	132	250	792	2877	3656	5913	3629	1214	410	277
290	225	250	330	410	752	2296	2177	739	502	317	317
304	209	225	330	647	1821	1821	3485	3431	567	317	369
277	277	290	725	489	1373	2996	3075	871	369	304	264
304	277	396	369	554	3128	3379	6731	1504	621	819	330
317	211	250	423	489	396	2864	1017	290	119	119	92
2138	409	462	475	489	673	4105	1479	8209	2217	515	435
106	92	92	146	1043	6863	5346	7167	8169	1188	554	290
356	356	356	419	554	1003	3960	8288	2996	898	423	304
238	238	263	343	937	567	1056	462	396	673	290	225
198	211	238	264	396	435	646	3920	2059	898	727	594
963	3419	607	646	1135	3537	8658	4329	5508	871	554	594
409	304	594	687	752	3498	13819	4421	7021	3788	844	594
383	330	396	1425	1056	3167	4686	5240	2257	1571	1202	423
383	448	409	290	608	554	964	831	462	1386	686	238
132	171	277	356	964	2217	1848	3696	2085	819	448	317
277	290	356	488	1808	5346	14638	9688	727	1201	250	343
264	264	290	423	1109	2442	4052	6507	3642	1043	1702	528
264	238	264	383	488	2811	10229	8381	3524	805	541	502
290	264	317	779	198	2771	1400	739	725	515	330	317
250	264	264	369	2323	12169	2838	6916	1504	1940	844	554
462	264	409	700	1267	8328	10453	15509	11680	2429	727	475
264	225	396	554	646	739	1756	2098	1386	739	435	277
171	171	343	383	2152	1940	10071	10123	14122	6032	806	448
331	171	158	331	343	448	594	752	568	290	211	238
52	198	105	475	4527	2613	2548	3511	7075	898	541	225
27	1	40	369	409	567	1756	1861	937	238	146	13
1	92	79	66	409	1141	1705	2782	773	1746	386	159
268	136	27	223	905	11343	9601	5744	1321	1106	400	255

continued over

Myponga River *(cont.)*

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
204	117	141	313	491	1059	8360	3491	2409	3969	1595	532
177	118	168	196	682	850	2573	2737	1619	368	359	268
82	150	150	59	391	1191	3928	3528	3450	827	541	473
241	204	164	268	195	205	682	1437	1187	286	182	286
286	250	55	496	4087	4896	3464	7319	1082	2844	570	427
409	700	291	245	427	1132	3726	1289	2168	624	555	408
384	206	260	249	642	1682	3377	5647	2993	416	405	526
196	254	469	843	4297	5867	2439	6424	3306	738	819	423
334	237	92	197	340	401	2960	5275	1378	685	221	114
176	158	290	479	584	2828	4269	5581	3201	1033	703	79
266	290	117	584	1140	1155	10130	1312	1961	5844	644	286
312	227	341	267	1076	988	2703	5641	1881	2429	1205	226
268	286	191	278	277	566	360	989	897	1789	400	167
153	195	158	173	389	1126	1364	909	1188	492	261	87
70	74	82	92	297	4000	6536	7060	3769	613	221	0
0	17	111	424	738	978	1947	5816	6160	3811	804	280
221	279	252	449	295	1333	2649	1543	701	1304	378	102
188	175	515	147	347	3592	6037	7556	1262	531	459	244
338	94	293	190	289	1187	756	872	433	192	167	75
54	124	78	867	560	742	4597	5536	4280	1029	171	144
160	142	205	161	533	85	1980	7276	3113	740	190	242

Onkaparinga River at Clarendon Weir

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
74	67	74	72	8506	43946	30967	27840	3053	3330	1407	533
74	67	74	72	1007	9278	1783	8336	2016	1252	1123	185
163	67	74	750	2602	19435	5468	60691	10690	3388	682	144
74	67	74	72	74	16056	6397	6269	7965	5237	1664	154
74	58	45	55	123	5663	2224	4374	4107	2338	757	933
131	318	31	1650	1926	10274	21832	11719	36274	2972	1012	1035
1740	30	31	30	1212	5301	32032	14519	8411	3376	1133	31
254	28	31	1363	2158	20624	31413	14015	22103	20793	3832	400
136	28	31	30	813	12503	30695	38687	41593	12893	4051	984
319	128	31	30	935	3766	8590	27678	4156	1822	1265	149
245	30	31	30	4257	24708	8110	20150	29690	13831	1026	246
74	154	74	2035	15422	31289	42108	80446	20280	29286	4687	768
358	19	2682	729	4144	10974	36578	21384	20600	9574	2253	1813
335	67	74	301	2984	14660	8789	9742	9798	4194	643	282
417	69	74	72	108	1167	6245	4470	29571	3386	2320	1828
192	1937	551	799	505	363	997	4972	4959	4603	1169	236
120	17	19	240	1124	782	1763	962	564	201	132	89
31	28	21	21	1854	28744	19404	35211	41911	3828	999	174
37	27	24	18	105	24777	27705	33011	6868	6501	6478	1926
205	35	598	1340	25767	41209	81059	43448	64386	8206	2819	1381
332	90	73	350	1371	5345	16873	18195	6505	4198	1396	196
35	67	120	21	767	1499	3183	7502	16591	5177	592	92
19	18	19	18	1570	19475	24862	33097	17713	4432	1369	1081
1364	22	19	209	1345	4249	3779	9953	24475	7950	2039	1256
471	154	19	493	1703	7764	32516	26895	6048	6859	1353	769
529	25	19	18	9678	64207	54987	30012	50522	17889	2209	1955
845	1197	1204	1252	3840	15946	2579	14480	21305	10229	2621	513
56	734	135	204	2344	4240	4520	2869	30259	3878	1055	201
19	17	19	111	9070	6936	9713	39672	14409	23356	1218	421
158	107	38	55	2580	5732	14802	31367	5826	2922	2850	1025
22	352	160	205	2012	7966	14062	10876	1973	19925	545	24

continued over

Onkaparinga River *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
19	17	19	18	34	5002	8296	6304	8244	2777	605	2318
429	20	22	21	53	183	4910	20346	11863	5879	1239	125
31	20	22	21	492	16106	37256	22044	45763	3252	1437	668
40	34	36	1465	716	34809	33662	37587	11661	5658	1431	186
73	58	37	483	12209	8744	5332	17933	66429	5356	708	263
101	28	31	146	170	140	983	3794	4776	3761	9358	662
429	33	20	285	1368	8182	19088	37050	37279	2841	1086	75
53	49	201	66	106	1283	5500	7057	2609	2092	1143	749
420	118	58	36	1353	2470	2940	17035	23647	2666	301	911
191	526	143	7220	1917	4897	11237	19309	4737	1173	564	−54
272	351	577	1123	2624	11570	13711	38812	11237	2059	6321	792
789	263	636	1775	2676	1230	11463	5525	2786	1064	1063	112
2421	430	154	392	304	931	7805	3441	23668	8829	650	365
270	283	−94	−78	4527	27724	39865	42731	36122	3533	1283	457
−233	74	391	559	959	3135	10432	26378	12801	4400	1435	392
10	128	255	193	2281	994	7974	1741	953	1058	1126	280
107	356	4	211	776	669	779	8938	9740	4898	2970	626
−43	5927	911	489	2436	9033	27189	16145	8323	2279	909	2710
246	164	598	700	862	2992	25492	31064	17799	18610	4186	1119
506	165	125	1416	1967	3380	8499	23909	5092	7292	14061	1920
456	530	509	605	946	1385	2391	2834	2254	11267	7703	970
308	318	135	206	1920	4595	4512	13296	4287	3506	1170	475
313	550	118	107	6364	10710	69879	38824	4998	8845	1906	1057
355	−39	92	236	2480	9175	13690	12507	9664	3527	8975	2844
554	127	236	165	1073	19824	42947	21753	18294	5193	1395	988
468	429	793	2307	2254	4452	6790	5060	3907	2176	1337	779
292	56	232	622	5762	48939	14273	66025	5274	4592	4192	1334
648	257	283	2966	4985	62725	32342	17280	16943	6914	2254	1155
464	−45	4	800	1490	1632	5167	5117	5270	2050	1040	17

continued over

Onkaparinga River *(cont.)*

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
172	240	41	466	5394	1521	16353	27474	27964	22354	2557	1013
465	252	251	476	565	417	991	2683	961	857	408	967
238	554	−39	823	22958	18104	13305	13191	28863	4873	2094	114
−282	−405	−35	673	805	2957	9212	8053	5243	1281	682	597
17	−198	−135	96	2641	6416	4375	13918	3414	9359	1285	−443
582	460	178	143	1687	22999	45540	35304	15544	3965	968	478
628	871	43	64	305	1852	42812	16571	20004	27817	9938	2249
722	309	620	191	961	1808	3243	8695	4389	866	79	561
−793	−223	474	16	925	2389	16038	12277	11660	4610	1126	1810
71	133	136	72	114	97	1544	3403	2072	840	−50	−95
−65	−248	192	790	13424	22093	18874	46033	6817	25612	6052	1073
−1381	4315	455	1503	3925	3437	13764	4726	7606	1714	1035	647
940	47	231	1122	1740	4753	15494	26329	12996	2999	914	1206
380	395	590	16984	26617	20748	5870	37133	29311	7487	3347	1506
2859	893	516	800	1241	1420	10461	20924	7409	2747	895	247
365	760	230	1090	2248	8343	20154	20425	31419	7429	1936	1144
703	2079	513	1785	3652	2310	26431	18082	20159	31054	3077	914
695	520	946	1204	9090	3013	13481	17933	6974	16416	6596	730
52	552	171	418	634	1623	1590	3526	2379	9336	1208	259
646	280	336	438	1386	3676	4244	4110	3715	1516	1066	686
65	−85	−100	397	652	2265	15325	15824	12587	2039	891	194
17	−31	250	604	557	1295	6044	15925	39687	23263	1784	229
101	−171	−490	408	619	5289	15678	5412	3282	8914	5523	252
−259	−68	385	36	570	30216	55622	51089	8589	2919	1242	140
205	48	228	608	1285	4184	2669	2436	906	321	−180	−181
−273	−82	583	1013	2450	2398	13035	20993	22042	3022	1334	609
376	524	428	531	1293	1257	6498	40659	12171	2546	943	−2

Torrens at Gorge Weir

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
345	241	314	268	273	12047	882	13747	14052	3587	464	350
186	73	96	159	268	290	6533	29146	2491	873	518	496
295	150	178	218	259	291	1030	8696	18425	2668	987	1209
250	55	77	455	532	20257	20212	9001	12743	4473	1373	777
809	173	237	282	350	836	20135	13197	3550	672	323	147
1229	68	123	33031	8072	29493	1520	22258	7188	2705	542	349
208	64	128	172	240	5754	22804	34329	7931	8813	3599	731
402	229	221	303	292	368	2457	2179	1081	1137	634	326
146	91	100	161	251	816	1299	13146	7087	11428	1130	634
206	70	61	253	3978	26278	13688	648	23548	3243	1222	416
286	130	231	311	345	5799	11751	14391	9233	7726	675	413
303	107	138	588	415	2588	12529	14051	10002	710	380	254
171	95	104	1605	431	2273	3621	900	739	324	100	72
97	94	94	97	479	708	2132	10463	8769	1495	197	53
0	70	150	1156	6230	25098	12893	13307	2243	478	434	222
65	70	129	352	526	2985	659	1305	561	365	270	84
29	0	61	451	481	3849	1447	27494	6028	1659	427	175
150	51	42	259	203	7373	3164	1750	3631	3346	875	176
123	38	246	117	192	3740	877	1937	975	601	232	434
84	80	239	588	496	3169	8469	8590	13378	1411	706	352
1068	109	151	158	330	4472	20060	7355	4019	1234	441	92
18	4	10	585	1506	19056	31228	5932	15151	11421	3734	362
169	0	190	147	103	4330	14633	29368	22897	7379	2873	1511
238	163	213	554	363	1060	4414	12302	2055	1634	1244	390
106	0	182	225	5460	19959	13859	13265	25333	10125	1621	287
276	185	333	2116	13078	16913	25280	47181	12783	13174	2673	751
287	334	1023	60	1045	3991	20372	17125	13700	8299	3027	1615
29	0	0	0	0	4636	4333	3258	3287	1876	338	324
70	0	0	170	341	454	2263	4996	28498	1831	1441	1865
185	3318	465	313	213	362	551	1721	2585	4215	478	75

continued over

Torrens at Gorge Weir *(cont.)*

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
88	22	0	14	314	256	468	302	203	119	138	84
55	15	64	216	1828	13444	9600	20274	28336	4059	517	170
119	64	143	10	208	16464	13024	26625	5074	3865	5568	1244
336	256	184	832	1783	6115	9090	12263	9522	4533	1440	620
336	256	184	832	1783	6115	9090	13446	5563	3293	1316	193
339	706	252	211	1087	957	1119	3736	9511	2273	417	238
−292	−438	−263	210	1116	9085	9995	27130	17192	3636	1729	1375
593	−504	−359	−163	8	498	1301	4751	16001	3778	1893	802
1297	234	−34	−246	219	893	10461	13407	2884	1539	1102	718
711	327	−93	−191	13782	47532	51015	23225	35571	22116	2604	1765
775	648	536	531	2749	8415	1900	5395	15221	8425	1846	434
180	61	−152	−12	219	732	1431	2551	16365	3112	712	58
−192	−135	−327	150	5317	3265	5798	26689	7787	32588	1791	428
140	41	−433	−291	1460	1397	5886	28692	3108	1206	1034	636
37	484	−24	126	1508	5249	5086	1445	1617	10621	675	−378
55	−127	−212	105	335	2967	2653	2344	3451	914	709	2782
336	239	−210	−276	−357	151	1710	12400	5622	1804	780	298
77	160	219	72	528	6012	19737	15156	35470	2253	1165	677
−27	109	219	1487	646	16515	18020	21832	11284	3084	1587	543
1827	273	−357	718	7980	3612	1892	8235	35562	2856	1019	508
349	254	165	566	496	670	935	2556	2758	2056	3778	487
513	212	482	414	817	3396	8000	20671	11563	2215	950	349
264	252	102	414	474	937	2800	3132	1351	1035	448	397
754	250	−344	393	1063	1415	1483	10786	14856	1836	819	1078
378	264	160	8025	1357	2591	5702	8160	3924	1115	647	422
367	250	225	421	1345	5783	8155	33463	5737	1672	2403	725
521	272	46	1338	1288	535	4600	2578	1453	714	514	68
1895	126	375	239	364	754	4238	2501	15201	5317	856	302
173	170	−242	828	2127	15416	23018	30403	27915	3040	659	320
267	221	701	724	532	1368	4325	10165	4412	2509	679	369

continued over

Torrens at Gorge Weir *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
248	82	−163	185	858	601	2405	1004	532	626	596	223
−54	40	97	−124	285	329	561	2849	4043	1822	1052	347
174	3735	525	673	1270	3303	18606	4221	1491	1260	751	965
319	226	−549	886	1015	1935	16578	18241	9894	7122	2259	1535
552	241	−312	905	1053	2602	4329	17871	2924	5529	7712	3842
782	239	125	179	752	588	1615	1930	1266	7364	8148	639
97	142	89	105	1896	3502	3832	7421	3795	1754	821	280
231	125	78	432	2772	8057	43923	26138	5808	4676	1581	501
295	113	129	248	2288	9358	8268	8177	7997	3150	4273	1603
249	144	6	312	642	8908	24987	14990	14199	3332	1437	940
346	117	182	398	969	2045	3641	2354	2233	1370	763	298
417	279	531	406	3174	28343	9631	36522	7112	3728	1791	1105
235	484	159	1855	3125	33107	39396	26179	21245	5066	2218	1017
751	217	410	740	769	835	3150	2670	2862	1511	756	435
369	225	−119	−263	2162	920	9205	21033	20009	16935	2226	605
344	438	432	327	357	299	356	1542	417	496	57	307
−216	98	−328	73	11793	10449	8976	9906	22327	3359	1636	401
−241	−181	−215	591	172	1089	4286	3429	3355	287	175	−334
−266	−589	−69	29	2300	4179	3295	16454	1863	6944	491	−58
−19	−541	−856	413	1555	22183	34240	25536	15555	2009	566	−111
−327	−151	−110	146	594	1555	20456	8488	10943	16595	3080	806
−79	−429	−440	−183	744	734	2004	5875	3499	460	142	51
−461	−727	−429	−460	54	849	11272	7053	8450	3907	780	1616
−267	−153	−159	−91	−40	345	371	1687	579	156	−153	−219
−29	−327	−256	−23	7658	16439	13653	32882	4457	14019	3843	1432
267	1905	382	542	1358	1242	8031	3425	5048	1735	1525	1201
1421	−326	−95	212	563	1401	5875	18327	11572	3000	1084	591
165	266	259	4201	10953	13834	4107	29087	21461	6025	2350	1440
1346	492	376	590	409	694	4205	12709	3096	1391	880	457
226	917	171	311	1003	3112	8174	10789	19167	4366	1550	974

continued over

Torrens at Gorge Weir *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
896	1326	692	1134	2327	1183	19017	14012	15075	26168	3557	1107
839	522	762	735	4855	1617	0	0	4074	9794	5017	979
302	509	148	282	524	702	484	668	293	1661	627	381
462	47	405	294	696	1530	1129	1631	581	170	408	−20
−111	−304	−348	265	403	1164	9309	10158	12852	1257	665	383
496	162	226	364	742	462	1341	4971	22200	20578	1957	523
225	337	333	608	598	2987	9989	2772	1219	4323	3119	910
695	406	1179	534	655	23469	30517	35222	8587	3684	1719	904
552	10	538	734	951	1561	1203	1685	736	574	341	394
357	256	986	1104	1114	1140	7146	11325	20458	3141	696	423

Torrens at Gumeracha Weir

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
150	150	115	177	868	2698	4218	11864	1307	727	50	39
39	35	39	38	39	38	39	387	2564	680	46	39
150	150	115	177	868	2698	4218	6061	4323	2121	567	244
150	150	115	177	868	2698	4218	6061	4323	2121	567	244
150	150	115	177	868	2698	4139	12857	2082	1742	757	1725
731	426	472	456	10008	40896	24092	9227	29900	10205	1882	747
589	541	639	573	10733	1739	1027	3061	7703	3033	320	106
35	54	21	55	192	488	712	1757	6408	733	131	15
10	9	10	23	1373	944	2450	8361	1035	13739	438	21
86	35	39	38	174	556	1331	13055	469	454	269	106
39	380	40	38	189	753	1224	430	756	1780	366	39
39	35	39	38	44	559	891	1175	1311	276	133	905
51	35	39	38	39	62	482	3523	2083	1313	336	72
40	35	39	38	112	1630	6454	5427	12109	539	366	69
25	30	39	38	132	4323	6399	7796	5587	1180	428	21
1007	70	21	198	2225	1176	314	4403	10140	793	157	73
73	16	18	17	32	71	170	728	678	619	1506	94
75	47	52	73	213	782	3805	9580	2969	581	356	183
207	56	52	51	52	218	1305	1451	692	348	170	66
172	74	39	34	95	312	379	4626	4804	69	158	165
173	61	63	2898	527	904	2629	2026	1028	360	152	152
108	64	73	71	476	3114	3178	17329	2404	213	778	208
18	16	18	110	122	17	1481	801	452	157	55	64
991	47	123	51	73	199	1406	1037	6800	1538	110	70
63	57	55	553	1143	12771	13354	15661	12950	904	278	35
52	47	52	51	52	928	1066	5247	2641	1143	213	52
52	49	52	51	52	230	810	426	110	109	130	52
52	47	52	46	52	54	137	758	1502	466	273	78
52	1921	88	71	287	1058	9893	2298	364	280	116	145
66	113	75	56	333	620	10215	8361	4343	4135	687	210
83	49	52	162	100	677	1758	9583	808	1618	3521	2128
160	94	66	63	70	66	782	777	409	2857	1615	253

continued over

Torrens at Gumeracha Weir *(cont.)*

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
54	24	17	8	245	1089	1506	4815	865	608	337	56
35	32	35	34	583	2770	17441	14873	2308	2066	392	35
35	33	35	34	1159	4538	1780	7015	3724	3502	1658	273
85	32	35	34	49	3193	12896	7013	8759	1204	422	205
114	32	36	45	780	855	1146	773	827	486	245	28
39	24	36	134	1130	13587	3495	22871	3038	1136	555	150
218	118	246	532	823	13416	14796	12715	8062	1738	782	205
450	45	45	346	423	391	1378	1887	1269	450	277	59
27	27	27	27	652	464	5093	5815	6405	4922	784	196
33	229	288	114	96	96	302	716	319	195	100	142
35	40	222	14	6571	5417	6312	4546	10371	1196	545	132
28	250	115	121	132	287	1686	1424	1827	284	127	−31
74	−145	64	64	1022	1802	1350	7090	871	4213	334	−47
50	25	62	106	679	10547	11517	12678	10404	697	209	43
137	129	−160	12	79	340	12124	4418	6142	7640	959	260
70	29	54	66	158	278	531	2709	1369	555	−42	−57
−167	−102	−7	−1	188	309	5471	3056	5016	1565	331	992
12	24	36	32	104	53	244	786	250	85	−20	0
179	39	40	270	2451	9288	9037	20940	1762	6991	1470	355
56	1127	103	94	195	300	3161	1407	1861	312	353	172
130	−47	75	86	61	259	1546	6909	5311	815	172	70
11	14	17	953	3691	4416	1684	16294	12881	2949	699	253
209	32	24	52	185	171	1292	6374	919	416	228	182
279	508	111	20	122	626	2928	5059	10768	2939	995	131
34	442	410	266	983	415	10987	7864	8540	12477	1049	129
51	19	48	116	1247	491	3332	3381	2210	4899	2527	514
70	35	49	54	134	216	140	128	1277	1044	121	45
129	55	252	85	155	522	514	676	604	488	143	371
466	272	432	769	79	2698	4218	6029	5636	749	404	264
93	260	268	113	242	219	117	1628	11097	10100	739	417
355	117	327	115	219	1598	5235	1474	556	3136	1568	590
277	213	−103	31	33	11990	18510	25203	3630	1260	581	147
48	384	441	110	216	392	229	960	648	891	833	1027
357	301	812	136	172	−95	2763	4691	10990	1268	567	244

continued over

Little Para River

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
43	189	68	83	179	288	2895	597	767	193	99	75
55	18	21	53	100	293	1650	3144	1835	412	130	77
34	19	24	748	3642	4462	1337	1699	4195	1079	455	225
190	100	81	112	151	124	659	1000	480	183	89	34
21	60	40	59	124	761	1475	1586	2508	462	151	72
300	145	40	114	324	196	3153	2296	1200	2800	510	212
148	120	114	128	130	130	130	1134	544	1383	514	250
37	30	44	54	78	100	122	158	117	159	69	27
18	5	18	43	90	187	232	98	8	7	6	2
0	1	2	6	9	637	1856	2503	2258	398	209	102
184	84	59	136	89	103	27	981	3948	3434	515	367
261	255	188	246	221	645	2099	631	332	837	863	318
231	131	232	78	140	4863	5126	5995	1807	788	458	282
272	162	216	258	182	556	389	371	214	142	112	188
174	310	123	72	219	201	1771	2558	4230	566	219	245

Appendix F

Monthly Cross Correlation Matrices

This appendix documents the monthly, lag zero and one, cross correlation matrices for the *Revised* data sets, followed by the multisite [A] , [B] & [C] matrices for each month.
Each data set has the following computations for cross correlations undertaken —

- The raw data
- The transformed data, computed using the 3–parameter shifting parameters given in Table [5.4].

The values are represented by matrices, where each column (i) and row (j) of the matrix represents the correlation between stations (i) & (j).
The order of stations is as follows —

- Warren
- South Para
- Myponga
- Onkaparinga
- Gorge
- Gumeracha
- Little Para

Revised Data

Raw Data Lag 0 cross correlations.

January

1.000	0.745	0.629	0.382	0.721	0.586	0.695
0.745	1.000	0.401	0.613	0.744	0.353	0.405
0.629	0.401	1.000	0.477	0.583	0.615	0.064
0.382	0.613	0.477	1.000	0.590	0.339	0.144
0.721	0.744	0.583	0.590	1.000	0.603	0.349
0.586	0.353	0.615	0.339	0.603	1.000	-0.044
0.695	0.405	0.064	0.144	0.349	-0.044	1.000

February

1.000	0.763	0.763	0.849	0.863	0.877	0.211
0.763	1.000	0.535	0.747	0.736	0.677	0.324
0.763	0.535	1.000	0.832	0.832	0.826	0.207
0.849	0.747	0.832	1.000	0.824	0.869	0.158
0.863	0.736	0.832	0.824	1.000	0.849	0.339
0.877	0.677	0.826	0.869	0.849	1.000	0.371
0.211	0.324	0.207	0.158	0.339	0.371	1.000

March

1.000	0.661	0.142	0.136	0.565	0.512	0.565
0.661	1.000	0.233	0.320	0.546	0.229	0.402
0.142	0.233	1.000	0.385	0.283	-0.261	0.402
0.136	0.320	0.385	1.000	0.445	0.175	-0.101
0.565	0.546	0.283	0.445	1.000	0.302	0.666
0.512	0.229	-0.261	0.175	0.302	1.000	0.068
0.565	0.402	0.402	-0.101	0.666	0.068	1.000

April

1.000	0.636	0.396	0.252	0.445	0.318	0.259
0.636	1.000	0.320	0.641	0.646	0.415	0.858
0.396	0.320	1.000	0.428	0.419	0.256	0.539
0.252	0.641	0.428	1.000	0.725	0.575	0.914
0.445	0.646	0.419	0.725	1.000	0.880	0.926
0.318	0.415	0.256	0.575	0.880	1.000	0.622
0.259	0.858	0.539	0.914	0.926	0.622	1.000

May

1.000	0.957	0.827	0.829	0.873	0.960	0.987
0.957	1.000	0.863	0.881	0.922	0.956	0.928
0.827	0.863	1.000	0.919	0.933	0.864	0.968
0.829	0.881	0.919	1.000	0.785	0.561	0.948
0.873	0.922	0.933	0.785	1.000	0.733	0.919
0.960	0.956	0.864	0.561	0.733	1.000	0.940
0.987	0.928	0.968	0.948	0.919	0.940	1.000

June

1.000	0.983	0.863	0.833	0.934	0.921	0.933
0.983	1.000	0.909	0.855	0.940	0.940	0.989
0.863	0.909	1.000	0.827	0.862	0.846	0.799
0.833	0.855	0.827	1.000	0.927	0.857	0.974
0.934	0.940	0.862	0.927	1.000	0.926	0.967
0.921	0.940	0.846	0.857	0.926	1.000	0.893
0.933	0.989	0.799	0.974	0.967	0.893	1.000

July

1.000	0.976	0.712	0.879	0.941	0.905	0.914
0.976	1.000	0.734	0.879	0.941	0.881	0.832
0.712	0.734	1.000	0.811	0.831	0.811	0.711
0.879	0.879	0.811	1.000	0.829	0.911	0.884
0.941	0.941	0.831	0.829	1.000	0.955	0.951
0.905	0.881	0.811	0.911	0.955	1.000	0.888
0.914	0.832	0.711	0.884	0.951	0.888	1.000

August

1.000	0.932	0.672	0.844	0.902	0.922	0.916
0.932	1.000	0.609	0.878	0.927	0.906	0.733
0.672	0.609	1.000	0.669	0.710	0.655	0.656
0.844	0.878	0.669	1.000	0.918	0.887	0.859
0.902	0.927	0.710	0.918	1.000	0.915	0.827
0.922	0.906	0.655	0.887	0.915	1.000	0.848
0.916	0.733	0.656	0.859	0.827	0.848	1.000

September

1.000	0.891	0.581	0.887	0.950	0.956	0.903
0.891	1.000	0.652	0.867	0.926	0.906	0.961
0.581	0.652	1.000	0.664	0.720	0.566	0.864
0.887	0.867	0.664	1.000	0.862	0.798	0.862
0.950	0.926	0.720	0.862	1.000	0.886	0.944
0.956	0.906	0.566	0.798	0.886	1.000	0.923
0.903	0.961	0.864	0.862	0.944	0.923	1.000

October

1.000	0.959	0.816	0.875	0.948	0.951	0.894
0.959	1.000	0.877	0.928	0.969	0.965	0.922
0.816	0.877	1.000	0.906	0.895	0.824	0.868
0.875	0.928	0.906	1.000	0.890	0.878	0.904
0.948	0.969	0.895	0.890	1.000	0.966	0.958
0.951	0.965	0.824	0.878	0.966	1.000	0.952
0.894	0.922	0.868	0.904	0.958	0.952	1.000

November

1.000	0.899	0.522	0.793	0.752	0.792	0.841
0.899	1.000	0.653	0.830	0.773	0.880	0.806

0.522	0.653	1.000	0.799	0.673	0.660	0.459
0.793	0.830	0.799	1.000	0.864	0.827	0.772
0.752	0.773	0.673	0.864	1.000	0.879	0.762
0.792	0.880	0.660	0.827	0.879	1.000	0.678
0.841	0.806	0.459	0.772	0.762	0.678	1.000

December

1.000	0.683	0.006	0.359	0.502	0.425	0.426
0.683	1.000	0.252	0.528	0.485	0.288	0.174
0.006	0.252	1.000	0.646	0.468	0.053	0.119
0.359	0.528	0.646	1.000	0.724	0.352	-0.160
0.502	0.485	0.468	0.724	1.000	0.672	0.358
0.425	0.288	0.053	0.352	0.672	1.000	0.381
0.426	0.174	0.119	-0.160	0.358	0.381	1.000

Raw Data Lag 1 cross correlations.

January

0.338	0.229	-0.285	-0.132	0.102	0.029	0.196
0.197	0.446	-0.153	-0.009	0.121	-0.086	0.033
-0.044	-0.021	0.124	-0.005	0.042	-0.048	0.318
-0.031	0.103	0.079	0.248	0.193	0.030	-0.073
0.253	0.185	0.001	0.120	0.290	0.158	0.122
0.058	-0.082	-0.262	-0.082	0.022	0.280	0.049
0.779	0.456	-0.177	-0.085	0.347	0.167	0.628

February

0.206	0.135	0.259	-0.262	0.034	-0.016	-0.150
0.291	0.398	0.172	-0.072	0.172	-0.070	0.019
0.077	0.012	0.417	-0.087	0.029	-0.030	-0.166
0.156	0.149	0.322	-0.091	0.057	-0.065	-0.111
0.265	0.191	0.320	-0.085	0.153	0.032	0.096
0.171	0.032	0.252	-0.216	0.015	0.086	-0.010
0.381	0.041	0.021	-0.304	-0.044	0.212	0.645

March

0.265	0.332	0.059	0.052	0.288	0.238	0.816
0.198	0.393	0.084	0.122	0.274	0.085	0.219
0.299	0.228	0.466	0.302	0.448	0.284	-0.068
0.235	0.353	0.377	0.300	0.313	0.357	-0.001
0.292	0.383	0.188	0.225	0.431	0.291	0.624
0.068	0.143	-0.056	0.036	0.114	0.294	0.517
-0.095	-0.041	-0.086	-0.222	-0.018	0.040	0.656

April

0.307	0.022	-0.119	-0.043	0.038	0.455	0.219
-0.049	0.209	-0.073	-0.008	-0.065	0.129	0.142
0.129	0.045	0.402	0.176	0.071	0.166	-0.106
-0.008	0.379	0.262	0.182	0.128	-0.009	-0.263
0.151	0.368	0.202	0.065	0.052	0.031	-0.102
-0.037	0.199	0.040	0.016	0.032	0.152	-0.391
-0.237	0.488	0.500	0.123	0.009	-0.154	0.053

May

0.212	0.298	0.166	0.297	0.174	0.190	0.893
0.222	0.299	0.153	0.342	0.218	0.203	0.899
0.197	0.364	0.290	0.450	0.124	0.100	0.886
0.236	0.479	0.264	0.505	0.260	0.183	0.874
0.246	0.418	0.243	0.354	0.282	0.181	0.863
0.220	0.365	0.178	0.175	0.023	0.202	0.889
0.229	0.816	0.587	0.995	0.974	0.736	0.935

June

0.205	0.278	0.380	0.245	0.294	0.216	0.338
0.305	0.352	0.480	0.346	0.395	0.324	0.977
0.242	0.250	0.436	0.322	0.346	0.284	0.696
0.337	0.329	0.423	0.436	0.513	0.470	0.486
0.347	0.379	0.448	0.332	0.589	0.513	0.438
0.342	0.367	0.419	0.306	0.599	0.569	0.240
0.623	0.883	0.562	0.540	0.528	0.520	0.634

July

0.713	0.679	0.633	0.614	0.687	0.688	0.554
0.690	0.722	0.714	0.598	0.662	0.664	0.002
0.323	0.331	0.442	0.318	0.323	0.283	0.135
0.549	0.498	0.516	0.587	0.514	0.544	0.545
0.636	0.624	0.624	0.641	0.636	0.709	0.549
0.567	0.524	0.524	0.661	0.703	0.706	0.560
0.676	0.101	0.351	0.609	0.632	0.692	0.541

August

0.729	0.662	0.472	0.608	0.638	0.657	0.758
0.594	0.613	0.420	0.488	0.523	0.515	0.310
0.518	0.550	0.567	0.536	0.622	0.558	0.198
0.487	0.499	0.399	0.531	0.444	0.490	0.555
0.626	0.579	0.466	0.518	0.508	0.538	0.649
0.632	0.565	0.396	0.553	0.527	0.542	0.689
0.777	0.455	0.606	0.868	0.816	0.791	0.768

September

0.448	0.448	0.436	0.290	0.356	0.284	0.337
0.560	0.617	0.449	0.437	0.506	0.437	0.520
0.237	0.281	0.555	0.164	0.247	0.112	0.165
0.307	0.375	0.413	0.267	0.239	0.114	0.169
0.472	0.505	0.545	0.261	0.281	0.217	0.371
0.429	0.475	0.455	0.278	0.312	0.220	0.294
0.350	0.599	0.666	0.531	0.488	0.362	0.382

October

0.334	0.153	0.154	0.357	0.279	0.270	0.163
0.410	0.254	0.244	0.431	0.349	0.340	0.219
0.332	0.275	0.499	0.389	0.363	0.307	0.090
0.286	0.189	0.275	0.248	0.247	0.260	0.135
0.407	0.255	0.312	0.295	0.313	0.369	0.268
0.375	0.214	0.186	0.278	0.267	0.369	0.240
0.553	0.526	0.519	0.630	0.531	0.517	0.400

November

0.420	0.418	0.297	0.368	0.383	0.425	0.684
0.250	0.322	0.247	0.298	0.306	0.349	0.712
0.322	0.420	0.496	0.494	0.461	0.472	0.570
0.270	0.323	0.356	0.374	0.260	0.257	0.374
0.352	0.418	0.397	0.468	0.425	0.392	0.575

0.297	0.382	0.347	0.397	0.357	0.402	0.376
0.559	0.601	0.461	0.554	0.536	0.570	0.619

December

0.439	0.409	0.186	0.215	0.211	0.338	0.330
0.286	0.349	0.336	0.199	0.186	0.253	0.067
0.161	0.208	0.563	0.316	0.230	0.105	0.106
0.353	0.416	0.660	0.480	0.442	0.379	0.000
0.718	0.802	0.591	0.545	0.590	0.634	0.516
0.602	0.763	0.303	0.392	0.406	0.604	0.220
0.663	0.814	0.377	0.493	0.546	0.571	0.846

Transformed Data Lag 0 cross correlations.

January

1.000	0.800	0.341	0.052	0.672	0.415	0.645
0.800	1.000	0.202	0.177	0.652	0.193	0.456
0.341	0.202	1.000	0.288	0.545	0.273	0.041
0.052	0.177	0.288	1.000	0.431	0.257	-0.085
0.672	0.652	0.545	0.431	1.000	0.561	0.492
0.415	0.193	0.273	0.257	0.561	1.000	-0.063
0.645	0.456	0.041	-0.085	0.492	-0.063	1.000

February

1.000	0.706	0.466	0.448	0.646	0.574	0.284
0.706	1.000	0.414	0.516	0.594	0.467	0.361
0.466	0.414	1.000	0.694	0.549	0.260	0.153
0.448	0.516	0.694	1.000	0.612	0.430	0.055
0.646	0.594	0.549	0.612	1.000	0.655	0.509
0.574	0.467	0.260	0.430	0.655	1.000	0.469
0.284	0.361	0.153	0.055	0.509	0.469	1.000

March

1.000	0.485	0.147	0.251	0.329	0.383	0.527
0.485	1.000	0.248	0.180	0.463	0.269	0.431
0.147	0.248	1.000	0.399	0.358	-0.205	0.334
0.251	0.180	0.399	1.000	0.380	-0.029	0.170
0.329	0.463	0.358	0.380	1.000	0.217	0.686
0.383	0.269	-0.205	-0.029	0.217	1.000	-0.112
0.527	0.431	0.334	0.170	0.686	-0.112	1.000

April

1.000	0.651	0.348	0.326	0.479	0.435	0.229
0.651	1.000	0.353	0.402	0.502	0.440	0.634
0.348	0.353	1.000	0.585	0.540	0.309	0.560
0.326	0.402	0.585	1.000	0.512	0.394	0.530
0.479	0.502	0.540	0.512	1.000	0.605	0.736
0.435	0.440	0.309	0.394	0.605	1.000	0.133
0.229	0.634	0.560	0.530	0.736	0.133	1.000

May

1.000	0.766	0.645	0.721	0.761	0.570	0.629
0.766	1.000	0.640	0.776	0.823	0.710	0.695
0.645	0.640	1.000	0.743	0.796	0.549	0.650
0.721	0.776	0.743	1.000	0.788	0.776	0.673
0.761	0.823	0.796	0.788	1.000	0.684	0.708
0.570	0.710	0.549	0.776	0.684	1.000	0.565
0.629	0.695	0.650	0.673	0.708	0.565	1.000

June

1.000	0.911	0.775	0.850	0.897	0.766	0.862
0.911	1.000	0.830	0.855	0.889	0.849	0.909
0.775	0.830	1.000	0.859	0.884	0.827	0.830
0.850	0.855	0.859	1.000	0.937	0.849	0.867
0.897	0.889	0.884	0.937	1.000	0.829	0.869
0.766	0.849	0.827	0.849	0.829	1.000	0.583
0.862	0.909	0.830	0.867	0.869	0.583	1.000

July

1.000	0.937	0.802	0.837	0.928	0.897	0.954
0.937	1.000	0.830	0.884	0.866	0.939	0.834
0.802	0.830	1.000	0.900	0.867	0.871	0.765
0.837	0.884	0.900	1.000	0.912	0.901	0.795
0.928	0.866	0.867	0.912	1.000	0.896	0.946
0.897	0.939	0.871	0.901	0.896	1.000	0.850
0.954	0.834	0.765	0.795	0.946	0.850	1.000

August

1.000	0.973	0.789	0.867	0.847	0.910	0.932
0.973	1.000	0.768	0.879	0.869	0.904	0.912
0.789	0.768	1.000	0.843	0.779	0.815	0.759
0.867	0.879	0.843	1.000	0.916	0.895	0.883
0.847	0.869	0.779	0.916	1.000	0.925	0.820
0.910	0.904	0.815	0.895	0.925	1.000	0.898
0.932	0.912	0.759	0.883	0.820	0.898	1.000

September

1.000	0.957	0.739	0.900	0.948	0.944	0.929
0.957	1.000	0.761	0.893	0.933	0.924	0.969
0.739	0.761	1.000	0.840	0.781	0.724	0.840
0.900	0.893	0.840	1.000	0.933	0.887	0.895
0.948	0.933	0.781	0.933	1.000	0.942	0.965
0.944	0.924	0.724	0.887	0.942	1.000	0.931
0.929	0.969	0.840	0.895	0.965	0.931	1.000

October

1.000	0.924	0.755	0.849	0.879	0.844	0.879
0.924	1.000	0.786	0.864	0.879	0.895	0.902
0.755	0.786	1.000	0.900	0.843	0.754	0.702
0.849	0.864	0.900	1.000	0.914	0.845	0.769
0.879	0.879	0.843	0.914	1.000	0.864	0.965
0.844	0.895	0.754	0.845	0.864	1.000	0.871
0.879	0.902	0.702	0.769	0.965	0.871	1.000

November

1.000	0.829	0.526	0.680	0.724	0.714	0.803
0.829	1.000	0.673	0.792	0.830	0.751	0.813
0.526	0.673	1.000	0.754	0.760	0.645	0.508
0.680	0.792	0.754	1.000	0.850	0.726	0.586
0.724	0.830	0.760	0.850	1.000	0.846	0.811

0.714	0.751	0.645	0.726	0.846	1.000	0.783
0.803	0.813	0.508	0.586	0.811	0.783	1.000

December

1.000	0.708	0.049	0.358	0.607	0.603	0.422
0.708	1.000	0.213	0.491	0.594	0.362	0.252
0.049	0.213	1.000	0.619	0.532	0.139	0.136
0.358	0.491	0.619	1.000	0.696	0.446	-0.155
0.607	0.594	0.532	0.696	1.000	0.681	0.458
0.603	0.362	0.139	0.446	0.681	1.000	0.464
0.422	0.252	0.136	-0.155	0.458	0.464	1.000

Transformed Data Lag 1 cross correlations.

January

0.488	0.320	-0.251	-0.058	0.212	0.221	0.218
0.325	0.470	-0.196	-0.027	0.174	0.010	0.044
0.100	0.089	0.458	0.158	0.327	0.069	0.317
0.146	0.232	0.207	0.375	0.344	0.182	-0.318
0.446	0.255	0.067	0.168	0.415	0.364	0.252
0.205	-0.120	-0.179	-0.062	0.110	0.446	0.014
0.719	0.515	-0.127	-0.152	0.458	0.245	0.688

February

0.296	0.330	0.224	-0.345	0.146	0.077	-0.106
0.346	0.452	0.167	-0.180	0.159	-0.091	0.041
0.204	0.152	0.743	0.023	0.298	0.059	-0.123
0.124	0.177	0.495	0.244	0.245	0.047	-0.165
0.390	0.316	0.438	0.035	0.430	0.261	0.204
0.217	0.050	0.275	-0.142	0.224	0.311	0.102
0.512	0.159	0.076	-0.446	0.164	0.084	0.768

March

0.397	0.406	0.158	0.187	0.371	0.384	0.593
0.265	0.487	0.183	0.106	0.324	0.038	0.255
0.223	0.182	0.619	0.353	0.479	0.201	0.018
0.071	0.240	0.412	0.458	0.326	0.219	0.112
0.324	0.345	0.324	0.274	0.521	0.281	0.640
0.112	0.182	-0.099	-0.013	0.155	0.376	0.236
0.039	0.157	0.061	-0.170	0.309	0.178	0.835

April

0.434	0.022	-0.155	-0.060	0.088	0.394	0.239
0.169	0.057	-0.093	0.023	-0.010	0.111	0.330
0.076	0.020	0.413	0.232	0.222	0.145	0.090
0.313	0.432	0.296	0.411	0.330	0.250	-0.334
0.175	0.127	0.239	0.140	0.420	0.138	0.117
0.064	0.101	0.002	-0.071	0.131	0.457	-0.447
-0.220	0.585	0.478	0.320	0.556	-0.146	0.487

May

0.464	0.207	0.200	0.156	0.206	0.171	0.396
0.293	0.183	0.021	0.102	0.058	0.083	0.630
0.194	0.181	0.333	0.172	0.121	0.172	0.497
0.331	0.207	0.279	0.401	0.236	0.318	0.494
0.334	0.241	0.327	0.271	0.389	0.289	0.619
0.256	0.167	0.232	0.265	0.091	0.419	0.592
0.283	0.618	0.702	0.666	0.826	0.120	0.890

June

0.753	0.599	0.487	0.469	0.537	0.402	0.443
0.725	0.710	0.534	0.570	0.612	0.575	0.616
0.534	0.565	0.521	0.524	0.581	0.490	0.094
0.659	0.616	0.520	0.516	0.529	0.551	0.458
0.688	0.643	0.525	0.462	0.611	0.506	0.456
0.534	0.612	0.435	0.439	0.479	0.526	0.015
0.534	0.417	0.170	0.251	0.265	-0.163	0.291

July

0.752	0.676	0.719	0.695	0.733	0.620	0.668
0.713	0.681	0.764	0.719	0.755	0.724	0.649
0.568	0.537	0.601	0.508	0.514	0.414	0.425
0.638	0.616	0.632	0.645	0.623	0.563	0.485
0.712	0.663	0.703	0.652	0.618	0.609	0.616
0.647	0.643	0.674	0.653	0.641	0.631	0.587
0.665	0.496	0.480	0.518	0.544	0.276	0.695

August

0.760	0.802	0.702	0.712	0.693	0.737	0.631
0.738	0.775	0.703	0.677	0.678	0.683	0.475
0.520	0.568	0.641	0.590	0.567	0.557	0.175
0.578	0.635	0.627	0.652	0.619	0.606	0.425
0.723	0.655	0.648	0.643	0.617	0.610	0.671
0.726	0.735	0.674	0.646	0.662	0.679	0.633
0.671	0.594	0.798	0.802	0.675	0.697	0.671

September

0.726	0.691	0.611	0.563	0.550	0.577	0.756
0.753	0.764	0.603	0.602	0.623	0.629	0.794
0.487	0.513	0.606	0.424	0.401	0.389	0.510
0.606	0.611	0.615	0.478	0.481	0.430	0.606
0.704	0.685	0.663	0.498	0.508	0.514	0.760
0.672	0.655	0.627	0.493	0.487	0.483	0.682
0.664	0.721	0.761	0.743	0.694	0.651	0.754

October

0.516	0.435	0.320	0.463	0.444	0.438	0.447
0.560	0.531	0.393	0.511	0.490	0.512	0.379
0.449	0.425	0.510	0.472	0.433	0.430	0.189
0.449	0.402	0.394	0.489	0.466	0.421	0.280
0.581	0.527	0.438	0.549	0.575	0.508	0.563
0.524	0.481	0.328	0.369	0.397	0.458	0.373
0.737	0.660	0.472	0.632	0.655	0.642	0.624

November

0.534	0.530	0.417	0.455	0.461	0.445	0.745
0.519	0.597	0.461	0.514	0.557	0.597	0.771
0.546	0.633	0.700	0.652	0.681	0.586	0.623
0.634	0.607	0.635	0.616	0.567	0.498	0.557
0.684	0.665	0.647	0.677	0.734	0.640	0.789

0.691	0.692	0.581	0.557	0.625	0.655	0.669
0.803	0.851	0.525	0.621	0.809	0.767	0.897

December

0.559	0.491	0.187	0.300	0.381	0.478	0.408
0.275	0.505	0.311	0.245	0.349	0.303	0.234
0.222	0.238	0.638	0.361	0.342	0.143	0.183
0.436	0.444	0.564	0.507	0.486	0.346	0.049
0.628	0.687	0.658	0.524	0.674	0.577	0.655
0.614	0.554	0.334	0.323	0.479	0.633	0.246
0.747	0.854	0.322	0.303	0.584	0.754	0.896

Matrices for January.

[A] Matrix

	1	2	3	4	5	6	7
1	0.469	0.053	−0.225	−0.436	0.467	−0.076	−0.209
2	0.025	0.590	−0.168	−0.733	0.720	−0.140	−0.470
3	0.120	−0.157	0.415	−0.102	0.257	−0.238	0.226
4	0.015	0.212	0.381	−1.115	1.116	0.379	−1.289
5	0.408	−0.154	0.043	−0.408	0.592	0.064	−0.252
6	0.298	−0.242	0.269	−1.108	0.518	0.876	−0.903
7	0.636	0.042	−0.201	−0.567	0.804	−0.524	0.223

[C] Matrix

	1	2	3	4	5	6	7
1	0.6361	0.4953	0.3800	−0.0312	0.4244	0.2438	0.1340
2	0.4953	0.5583	0.2633	−0.0243	0.4967	0.1732	0.0297
3	0.3800	0.2633	0.6890	0.2861	0.4439	0.3719	−0.1425
4	−0.0312	−0.0243	0.2861	0.4261	0.2561	0.0045	−0.0434
5	0.4244	0.4967	0.4439	0.2561	0.7176	0.3512	0.1075
6	0.2438	0.1732	0.3719	0.0045	0.3512	0.4543	−0.1174
7	0.1340	0.0297	−0.1425	−0.0434	0.1075	−0.1174	0.0153

[B] Matrix

	1	2	3	4	5	6	7
1	0.7976	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.6210	0.4155	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.4764	−0.0784	0.6752	0.0000	0.0000	0.0000	0.0000
4	−0.0391	−0.0002	0.4513	0.4700	0.0000	0.0000	0.0000
5	0.5321	0.4000	0.3285	0.2738	0.3027	0.0000	0.0000
6	0.3056	−0.0398	0.3305	−0.2824	0.5723	0.0000	0.0000
7	0.1680	−0.1796	−0.3505	0.2582	0.4439	0.0000	0.0000

Matrices for February.

[A] Matrix

	1	2	3	4	5	6	7
1	1.557	−0.757	−0.433	−0.692	1.552	−1.101	−1.638
2	1.367	−0.454	−0.377	−0.392	1.167	−1.104	−1.294
3	1.085	−0.749	0.318	−0.322	1.105	−0.942	−1.125
4	0.891	−0.515	0.079	0.017	0.857	−0.792	−0.978
5	−0.094	0.218	0.392	−0.171	0.060	0.169	0.115
6	−0.331	0.176	0.401	−0.277	−0.233	0.529	0.343
7	−1.256	1.016	0.900	−0.207	−1.835	1.376	2.050

[C] Matrix

	1	2	3	4	5	6	7
1	0.3313	0.1777	0.0015	0.1742	0.4460	0.4578	0.4278
2	0.1777	0.4915	0.0505	0.2615	0.4329	0.4564	0.4824
3	0.0015	0.0505	0.2506	0.2014	0.2342	0.0896	0.3105
4	0.1742	0.2615	0.2014	0.6036	0.4291	0.3979	0.3586
5	0.4460	0.4329	0.2342	0.4291	0.7083	0.4545	0.3038
6	0.4578	0.4564	0.0896	0.3979	0.4545	0.7657	0.1865
7	0.4278	0.4824	0.3105	0.3586	0.3038	0.1865	−0.0678

[B] Matrix

	1	2	3	4	5	6	7
1	0.576	0.000	0.000	0.000	0.000	0.000	0.000
2	0.309	0.629	0.000	0.000	0.000	0.000	0.000
3	0.003	0.079	0.494	0.000	0.000	0.000	0.000
4	0.303	0.267	0.363	0.556	0.000	0.000	0.000
5	0.775	0.308	0.420	−0.073	0.000	0.000	0.000
6	0.795	0.335	0.123	0.041	0.000	0.063	0.000
7	0.743	0.402	0.560	−0.318	0.000	−9.498	0.000

Matrices for March.

[A] Matrix

	1	2	3	4	5	6	7
1	0.3008	0.0477	−0.1269	0.2264	−0.2227	−0.0205	0.6200
2	−0.1476	0.6263	0.0602	−0.2923	0.3955	−0.3348	0.0335
3	−0.3367	0.0396	0.7547	−0.4853	0.6672	0.1357	−0.3931
4	−0.4596	0.1914	0.2817	0.2318	0.1249	0.1451	−0.0139
5	0.1281	−0.0832	0.0120	0.2116	0.2104	−0.2895	0.6485
6	−0.1982	0.2320	−0.1016	−0.1714	−0.0213	0.4884	0.0150
7	−0.2024	0.0522	0.1585	−0.3045	0.1759	−0.1532	0.8483

[C] Matrix

	1	2	3	4	5	6	7
1	0.5620	0.2954	0.1698	0.1743	−0.0809	0.2266	0.1090
2	0.2954	0.6302	0.1100	0.0895	0.2216	0.2296	0.1947
3	0.1698	0.1100	0.4321	0.1221	0.2081	−0.1679	0.3103
4	0.1743	0.0895	0.1221	0.6934	0.2109	−0.0525	0.1269
5	−0.0809	0.2216	0.2081	0.2109	0.4823	0.1456	0.1750
6	0.2266	0.2296	−0.1679	−0.0525	0.1456	0.7836	−0.2562
7	0.1090	0.1947	0.3103	0.1269	0.1750	−0.2562	0.2030

[B] Matrix

	1	2	3	4	5	6	7
1	0.7496	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.3941	0.6891	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.2265	0.0300	0.6164	0.0000	0.0000	0.0000	0.0000
4	0.2325	−0.0031	0.1128	0.7916	0.0000	0.0000	0.0000
5	−0.1079	0.3833	0.3586	0.2485	0.3651	0.0000	0.0000
6	0.3023	0.1602	−0.3913	−0.0988	0.7716	0.0000	0.0000
7	0.1454	0.1994	0.4403	0.0556	−0.1574	0.0000	0.0000

Matrices for April.

[A] Matrix

	1	2	3	4	5	6	7
1	0.201	−0.286	−0.090	0.015	−0.306	0.504	0.550
2	−0.565	−0.023	−0.112	0.389	−1.031	0.701	1.395
3	−0.107	−0.200	0.480	0.098	−0.001	0.352	0.095
4	1.260	0.348	0.153	−0.275	1.640	−0.924	−2.382
5	0.487	−0.133	0.166	−0.265	0.963	−0.283	−0.786
6	0.761	0.023	0.159	−0.521	1.247	−0.285	−1.710
7	−1.227	0.695	0.198	0.394	−0.324	0.367	0.964

[C] Matrix

	1	2	3	4	5	6	7
1	0.603	0.385	0.318	0.566	0.495	0.507	0.455
2	0.385	0.529	0.355	1.094	0.749	0.945	0.450
3	0.318	0.355	0.732	0.467	0.397	0.224	0.398
4	0.566	1.094	0.467	−0.584	−0.033	−0.599	0.731
5	0.495	0.749	0.397	−0.033	0.656	0.219	0.732
6	0.507	0.945	0.224	−0.599	0.219	0.114	0.476
7	0.455	0.450	0.398	0.731	0.732	0.476	−0.133

[B] Matrix

	1	2	3	4	5	6	7
1	0.777	0.000	0.000	0.000	0.000	0.000	0.000
2	0.495	0.533	0.000	0.000	0.000	0.000	0.000
3	0.410	0.285	0.695	0.000	0.000	0.000	0.000
4	0.728	1.375	−0.321	0.000	0.000	0.000	0.000
5	0.637	0.814	−0.138	0.000	0.000	0.000	0.000
6	0.653	1.165	−0.541	0.000	0.000	0.000	0.000
7	0.586	0.299	0.105	0.000	0.000	0.000	0.000

Matrices for May.

[A] Matrix

	1	2	3	4	5	6	7
1	8.28	−11.74	−2.36	−2.34	−14.32	9.68	17.76
2	14.35	−20.99	−4.51	−4.17	−26.08	17.58	32.25
3	8.90	−13.16	−2.39	−2.74	−16.57	11.22	20.30
4	9.20	−13.51	−2.75	−2.39	−16.69	11.41	20.53
5	8.59	−12.64	−2.50	−2.53	−15.34	10.59	19.29
6	15.62	−23.22	−4.66	−4.56	−29.09	19.93	35.53
7	−4.27	6.39	1.45	1.45	8.16	−5.58	−9.03

[C] Matrix

	1	2	3	4	5	6	7
1	−5.30	−10.39	−6.39	−6.46	−6.00	−11.70	3.62
2	−10.39	−19.11	−11.99	−12.07	−11.22	−21.41	6.28
3	−6.39	−11.99	−7.07	−7.40	−6.85	−13.48	4.04
4	−6.46	−12.07	−7.40	−7.35	−7.00	−13.48	4.09
5	−6.00	−11.22	−6.85	−7.00	−6.36	−12.63	3.76
6	−11.70	−21.41	−13.48	−13.48	−12.63	−23.59	6.82
7	3.62	6.28	4.04	4.09	3.76	6.82	−1.76

[B] Matrix

	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0

Matrices for June.

[A] Matrix

	1	2	3	4	5	6	7
1	0.807	0.255	0.193	−0.305	−0.205	0.061	−0.051
2	0.497	0.367	0.138	−0.310	−0.247	0.251	0.201
3	0.207	0.415	0.344	0.030	0.290	0.153	−0.860
4	0.532	0.326	0.339	−0.341	−0.368	0.369	−0.042
5	0.542	0.339	0.261	−0.570	0.028	0.285	−0.088
6	0.334	0.837	0.470	−0.321	−0.169	0.397	−0.971
7	0.586	0.638	−0.298	0.456	−0.139	−1.091	0.079

[C] Matrix

	1	2	3	4	5	6	7
1	0.3968	0.3376	0.3535	0.3164	0.3348	0.2967	0.4476
2	0.3376	0.3655	0.4501	0.2902	0.3134	0.4181	0.5924
3	0.3535	0.4501	0.2976	0.4294	0.4181	0.0942	0.6790
4	0.3164	0.2902	0.4294	0.4585	0.3976	0.3500	0.6451
5	0.3348	0.3134	0.4181	0.3976	0.4136	0.3069	0.6006
6	0.2967	0.4181	0.0942	0.3500	0.3069	0.1331	0.4480
7	0.4476	0.5924	0.6790	0.6451	0.6006	0.4480	0.1926

[B] Matrix

	1	2	3	4	5	6	7
1	0.6299	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.5359	0.2799	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.5612	0.5336	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.5023	0.0749	0.0000	0.4479	0.0000	0.0000	0.0000
5	0.5316	0.1020	0.0000	0.2745	0.2126	0.0000	0.0000
6	0.4711	0.5917	0.0000	0.1542	−0.2174	0.0000	0.0000
7	0.7106	0.7559	0.0000	0.5168	0.0180	0.0000	0.0000

Matrices for July.

[A] Matrix

	1	2	3	4	5	6	7
1	0.759	−0.460	0.445	−0.089	0.035	0.066	0.070
2	0.422	−0.040	0.604	0.129	0.142	−0.076	−0.371
3	0.805	−1.709	0.211	−1.067	−0.335	1.364	1.530
4	0.570	−1.338	−0.095	−0.499	−0.162	1.232	1.144
5	1.017	−0.614	0.635	0.073	−1.039	0.408	0.370
6	0.439	−0.239	0.388	0.100	−0.352	0.296	0.150
7	1.194	−0.168	0.496	0.490	−0.400	−0.909	−0.142

[C] Matrix

	1	2	3	4	5	6	7
1	0.3682	0.3129	0.3194	0.3190	0.3320	0.3548	0.4230
2	0.3129	0.3605	0.2989	0.3087	0.2699	0.3744	0.4174
3	0.3194	0.2989	0.8333	0.6926	0.4079	0.4612	0.2714
4	0.3189	0.3087	0.6926	0.6955	0.4307	0.4385	0.3369
5	0.3320	0.2699	0.4079	0.4307	0.3532	0.3483	0.4259
6	0.3548	0.3744	0.4612	0.4385	0.3483	0.4942	0.4442
7	0.4230	0.4174	0.2714	0.3369	0.4259	0.4442	0.3634

[B] Matrix

	1	2	3	4	5	6	7
1	0.6068	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.5157	0.3075	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.5263	0.0893	0.7405	0.0000	0.0000	0.0000	0.0000
4	0.5256	0.1224	0.5470	0.3241	0.0000	0.0000	0.0000
5	0.5471	−0.0399	0.1667	0.1752	0.0000	0.0000	0.0000
6	0.5847	0.2370	0.1786	0.0139	0.0000	0.2532	0.0000
7	0.6971	0.1881	−0.1517	0.0938	0.0000	0.0706	0.0000

Matrices for August.

[A] Matrix

	1	2	3	4	5	6	7
1	1.38	0.14	0.16	−0.14	0.44	0.02	−1.24
2	5.05	−1.52	0.04	−1.09	3.26	0.48	−5.73
3	8.31	−3.91	−0.18	−1.73	6.09	1.66	−10.15
4	2.48	−0.72	−0.08	−0.50	2.58	0.33	−3.60
5	1.97	−1.03	0.53	1.25	−2.25	−0.31	0.65
6	1.08	0.16	0.36	−0.22	0.17	−0.12	−0.69
7	2.75	−3.06	0.67	2.51	−3.37	0.67	0.71

[C] Matrix

	1	2	3	4	5	6	7
1	0.294	0.108	−0.070	0.217	0.297	0.271	0.432
2	0.108	−0.655	−1.483	−0.262	0.432	0.174	0.577
3	−0.070	−1.483	−2.560	−0.702	0.603	0.154	0.572
4	0.217	−0.262	−0.702	0.135	0.592	0.349	0.557
5	0.297	0.432	0.603	0.592	0.252	0.380	−0.013
6	0.271	0.174	0.154	0.349	0.380	0.404	0.400
7	0.432	0.577	0.572	0.557	−0.013	0.400	−0.249

[B] Matrix

	1	2	3	4	5	6	7
1	0.5419	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.1999	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	−0.1290	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.3995	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.5477	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	0.5009	0.0000	0.0000	0.0000	0.0000	0.3912	0.0000
7	0.7963	0.0000	0.0000	0.0000	0.0000	0.0023	0.0000

Matrices for September.

[A] Matrix

	1	2	3	4	5	6	7
1	0.291	0.098	0.567	−1.093	0.736	−1.202	1.405
2	−0.571	0.935	0.504	−1.274	0.890	−1.170	1.536
3	−1.137	1.311	1.159	−1.195	0.639	−1.294	1.188
4	−0.337	0.899	1.025	−1.399	1.148	−1.939	1.357
5	−0.457	0.740	1.135	−2.018	1.247	−1.949	2.161
6	−0.048	0.633	0.948	−1.487	0.980	−1.747	1.509
7	−1.808	1.411	0.822	−0.498	0.636	−0.999	1.343

[C] Matrix

	1	2	3	4	5	6	7
1	0.2179	0.1604	0.1228	0.1485	0.0202	0.1301	0.2566
2	0.1604	0.1403	0.1092	0.1108	−0.0234	0.0912	0.2207
3	0.1228	0.1092	0.3267	0.1165	−0.0513	−0.0037	0.1575
4	0.1485	0.1108	0.1165	0.1523	−0.0455	0.0225	0.1707
5	0.0202	−0.0234	−0.0513	−0.0455	−0.2052	−0.0915	0.1441
6	0.1301	0.0912	−0.0037	0.0225	−0.0915	0.0939	0.2084
7	0.2566	0.2207	0.1575	0.1707	0.1441	0.2084	0.1213

[B] Matrix

	1	2	3	4	5	6	7
1	0.4668	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.3437	0.1491	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.2632	0.1256	0.4916	0.0000	0.0000	0.0000	0.0000
4	0.3182	0.0095	0.0643	0.2165	0.0000	0.0000	0.0000
5	0.0432	−0.2566	−0.0619	−0.2442	0.0000	0.0000	0.0000
6	0.2788	−0.0310	−0.1489	−0.2601	0.0000	0.0000	0.0000
7	0.5497	0.2131	−0.0283	−0.0201	0.0000	0.0000	0.0000

Matrices for October.

[A] Matrix

	1	2	3	4	5	6	7
1	−30.9	61.6	25.1	−33.4	62.7	11.7	−93.2
2	54.4	−103.1	−42.0	56.4	−105.3	−19.8	154.5
3	136.4	−262.0	−105.9	142.4	−266.0	−50.2	392.2
4	71.0	−136.4	−55.2	74.4	−137.7	−26.3	203.7
5	−19.6	39.1	16.0	−21.3	40.6	7.0	−59.2
6	27.6	−50.4	−20.4	27.1	−52.1	−9.6	75.6
7	−6.2	14.5	5.9	−7.9	14.8	2.5	−22.2

[C] Matrix

	1	2	3	4	5	6	7
1	6.3	−8.7	−23.1	−11.7	4.1	−4.0	1.8
2	−8.7	16.1	39.8	21.0	−5.4	8.1	−1.8
3	−23.1	39.8	100.3	52.4	−14.4	19.7	−5.3
4	−11.7	21.0	52.4	27.6	−7.2	10.6	−2.5
5	4.1	−5.4	−14.4	−7.2	2.9	−2.3	1.4
6	−4.0	8.1	19.7	10.6	−2.3	4.3	−0.6
7	1.8	−1.8	−5.3	−2.5	1.4	−0.6	0.7

[B] Matrix

	1	2	3	4	5	6	7
1	2.507	0.000	0.000	0.000	0.000	0.000	0.000
2	−3.478	2.004	0.000	0.000	0.000	0.000	0.000
3	−9.231	3.817	0.695	0.000	0.000	0.000	0.000
4	−4.667	2.371	0.333	0.255	0.000	0.000	0.000
5	1.644	0.181	0.065	0.135	0.349	0.000	0.000
6	−1.601	1.280	0.069	0.403	0.119	0.000	0.000
7	0.712	0.362	−0.089	−0.013	0.394	0.000	0.000

Matrices for November.

[A] Matrix

	1	2	3	4	5	6	7
1	0.314	2.246	−0.414	−3.725	6.075	0.403	−4.617
2	−0.346	1.759	−0.368	−2.552	4.304	0.664	−3.025
3	−0.521	0.453	0.503	−0.060	0.312	−0.140	0.187
4	0.529	0.586	0.361	−0.683	1.047	−0.197	−1.004
5	0.213	0.485	0.017	−1.193	2.365	0.145	−1.341
6	0.468	0.202	0.449	−0.619	−0.150	0.259	0.155
7	−0.006	0.514	−0.233	−0.009	−0.266	−0.121	0.971

[C] Matrix

	1	2	3	4	5	6	7
1	1.967	1.367	0.160	0.601	0.732	0.289	0.087
2	1.367	1.149	0.238	0.662	0.665	0.309	0.093
3	0.160	0.238	0.438	0.316	0.244	0.205	0.001
4	0.601	0.662	0.316	0.564	0.478	0.273	0.102
5	0.732	0.665	0.244	0.478	0.557	0.342	0.137
6	0.289	0.309	0.205	0.273	0.342	0.441	0.168
7	0.087	0.093	0.001	0.102	0.137	0.168	0.132

[B] Matrix

	1	2	3	4	5	6	7
1	1.403	0.000	0.000	0.000	0.000	0.000	0.000
2	0.975	0.446	0.000	0.000	0.000	0.000	0.000
3	0.114	0.284	0.587	0.000	0.000	0.000	0.000
4	0.428	0.547	0.190	0.213	0.000	0.000	0.000
5	0.522	0.351	0.144	0.164	0.338	0.000	0.000
6	0.206	0.241	0.192	0.076	0.324	0.439	0.000
7	0.062	0.072	−0.045	0.211	0.150	0.186	0.141

Matrices for December.

[A] Matrix

	1	2	3	4	5	6	7
1	0.617	0.567	−0.314	−0.576	0.375	0.499	−0.746
2	−0.088	1.504	−0.070	−1.016	0.797	0.195	−1.084
3	0.217	−0.531	1.024	0.115	0.000	−0.624	0.341
4	0.968	0.588	0.145	−0.915	1.620	0.063	−2.105
5	0.224	0.278	0.476	−0.455	0.327	−0.137	0.117
6	1.163	1.191	−0.341	−1.816	1.622	1.115	−2.605
7	−0.062	1.324	−0.123	−0.986	−0.270	0.662	0.212

[C] Matrix

	1	2	3	4	5	6	7
1	0.5310	0.3829	0.1214	−0.0129	0.2858	−0.1110	−0.1738
2	0.3829	0.4518	0.1838	0.0142	0.2550	−0.3037	−0.2760
3	0.1214	0.1838	0.4105	0.3232	0.1628	0.2323	0.2279
4	−0.0129	0.0142	0.3232	−0.0076	0.3200	−0.5240	−0.2551
5	0.2858	0.2550	0.1628	0.3200	0.3753	0.2786	−0.1545
6	−0.1110	−0.3037	0.2323	−0.5240	0.2786	−0.5144	−0.2142
7	−0.1738	−0.2760	0.2279	−0.2551	−0.1545	−0.2142	−0.2769

[B] Matrix

	1	2	3	4	5	6	7
1	0.7287	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.5255	0.4191	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.1666	0.2298	0.5744	0.0000	0.0000	0.0000	0.0000
4	−0.0177	0.0560	0.5453	0.0000	0.0000	0.0000	0.0000
5	0.3922	0.1167	0.1230	0.0000	0.4390	0.0000	0.0000
6	−0.1523	−0.5337	0.6620	0.0000	0.7269	0.0000	0.0000
7	−0.2385	−0.3595	0.6097	0.0000	−0.2141	0.0000	0.0000

Appendix G

Five Station Model Matrices

This appendix documents the monthly, $[\mathbf{M_0}]$ & $[\mathbf{M_1}]$ Covariance matrices, together with their associated $[\mathbf{A_t}]$ & $[\mathbf{B_t}]$ matrices for a five station multivariate model with parameters evaluated using a Parametric Transformation.

The order of stations is as follows —

- South Para
- Myponga
- Torrens at the Gumeracha Weir
- Onkaparinga River at the Clarendon Weir
- Millbrook Rainfall Station

Matrices for January.

$[\mathbf{M_{0,1}}]$ Matrix

	1	2	3	4	5
1	1.000	0.199	0.206	0.324	0.476
2	0.199	1.000	0.345	0.355	0.253
3	0.206	0.345	1.000	0.492	0.161
4	0.324	0.355	0.492	1.000	0.333
5	0.476	0.253	0.161	0.333	1.000

$[\mathbf{M_{1,1}}]$ Matrix

	1	2	3	4	5
1	0.4702	−0.1960	−0.0006	−0.0446	0.1630
2	0.0905	0.4466	0.1389	0.1529	0.2100
3	−0.1105	−0.1325	0.3526	−0.0853	−0.0835
4	0.0516	0.0215	0.1016	0.1087	−0.0185
5	0.0307	−0.3391	−0.1994	−0.3410	0.1488

$[\mathbf{A_1}]$ Matrix

	1	2	3	4	5
1	0.6493	−0.1853	−0.1110	−0.2019	0.0099
2	0.0265	0.5850	0.1511	−0.3397	0.1161
3	−0.2258	−0.0994	0.5570	−0.2017	0.0498
4	−0.0041	−0.0625	0.0539	0.1391	−0.0488
5	0.1543	−0.2332	−0.0773	−0.3183	0.2619

$[\mathbf{B_1}]$ Matrix

	1	2	3	4	5
1	0.8048	0.0000	0.0000	0.0000	0.0000
2	0.3314	0.7956	0.0000	0.0000	0.0000
3	0.3429	0.3006	0.7379	0.0000	0.0000
4	0.4071	0.2893	0.3324	0.7875	0.0000
5	0.3741	0.2816	−0.0599	0.2077	0.6981

Matrices for February.

$[\mathbf{M_{0,2}}]$ Matrix

	1	2	3	4	5
1	1.000	0.416	0.466	0.539	0.672
2	0.416	1.000	0.317	0.693	0.325
3	0.466	0.317	1.000	0.527	0.379
4	0.539	0.693	0.527	1.000	0.369
5	0.672	0.325	0.379	0.369	1.000

$[\mathbf{M_{1,2}}]$ Matrix

	1	2	3	4	5
1	0.4516	0.1658	−0.0987	−0.1352	0.2291
2	0.1529	0.7678	0.0699	0.0099	0.1631
3	0.0498	0.3206	0.2175	−0.1825	0.0370
4	0.1808	0.5040	0.0506	0.0892	0.2531
5	0.0845	0.0687	−0.1684	−0.3400	0.0841

$[\mathbf{A_2}]$ Matrix

	1	2	3	4	5
1	0.5143	0.2094	−0.1210	−0.3375	0.0629
2	0.0871	0.8883	−0.1190	−0.2781	0.0082
3	0.0577	0.3687	0.3190	−0.4995	0.0312
4	0.0676	0.5345	−0.1147	−0.1134	0.1416
5	0.1482	0.1929	−0.0534	−0.4739	0.1311

$[\mathbf{B_2}]$ Matrix

	1	2	3	4	5
1	0.8131	0.0000	0.0000	0.0000	0.0000
2	0.2191	0.5161	0.0000	0.0000	0.0000
3	0.4123	−0.1700	0.7199	0.0000	0.0000
4	0.4437	0.3130	0.3072	0.5561	0.0000
5	0.5832	0.0266	−0.0081	−0.0144	0.6723

Matrices for March.

$[\mathbf{M_{0,3}}]$ Matrix

	1	2	3	4	5
1	1.000	0.247	0.255	0.301	0.386
2	0.247	1.000	−0.144	0.393	0.321
3	0.255	−0.144	1.000	−0.160	0.155
4	0.301	0.393	−0.160	1.000	0.349
5	0.386	0.321	0.155	0.349	1.000

$[\mathbf{M_{1,3}}]$ Matrix

	1	2	3	4	5
1	0.4868	0.1836	0.0377	0.1080	0.0641
2	0.1820	0.6600	0.2103	0.3675	0.0184
3	0.1914	−0.0808	0.3156	−0.0408	0.0767
4	0.2867	0.4176	0.2202	0.4802	0.1323
5	0.0655	0.2136	0.0335	0.1246	0.0120

$[\mathbf{A_3}]$ Matrix

	1	2	3	4	5
1	0.9446	0.1844	−0.1299	−0.2848	−0.4768
2	0.0764	0.8224	0.1281	−0.2115	−0.2710
3	0.2873	−0.0668	0.4247	−0.3226	−0.1368
4	0.1208	0.1666	−0.0327	0.3627	−0.1244
5	0.0385	0.2505	−0.0178	−0.0323	−0.0768

$[\mathbf{B_3}]$ Matrix

	1	2	3	4	5
1	0.7567	0.0000	0.0000	0.0000	0.0000
2	0.1247	0.6953	0.0000	0.0000	0.0000
3	0.2042	−0.2106	0.8467	0.0000	0.0000
4	0.2407	0.1533	−0.1797	0.7947	0.0000
5	0.4359	0.1594	0.1446	0.2011	0.8202

Matrices for April.

$[\mathbf{M_{0,4}}]$ Matrix

	1	2	3	4	5
1	1.000	0.412	0.408	0.517	0.711
2	0.412	1.000	0.211	0.598	0.567
3	0.408	0.211	1.000	0.381	0.456
4	0.517	0.598	0.381	1.000	0.542
5	0.711	0.567	0.456	0.542	1.000

$[\mathbf{M_{1,4}}]$ Matrix

	1	2	3	4	5
1	0.1101	−0.0887	0.1362	−0.0150	−0.1808
2	0.0231	0.5738	0.0228	0.1817	0.0974
3	0.1348	−0.0202	0.4259	−0.0357	0.0509
4	0.4595	0.3874	0.2301	0.4279	0.2489
5	0.0575	0.0686	0.0781	0.1015	−0.1086

$[\mathbf{A_4}]$ Matrix

	1	2	3	4	5
1	0.1724	−0.0505	0.1397	0.0723	−0.2779
2	−0.1588	0.6547	0.1769	0.0322	−0.0898
3	0.0234	0.0421	0.4382	0.0281	−0.0493
4	0.2654	0.2642	0.2657	0.3183	−0.0904
5	0.0403	0.0914	0.1418	0.1566	−0.2300

$[\mathbf{B_4}]$ Matrix

	1	2	3	4	5
1	0.9531	0.0000	0.0000	0.0000	0.0000
2	0.4696	0.6375	0.0000	0.0000	0.0000
3	0.3577	0.0133	0.8285	0.0000	0.0000
4	0.4859	0.2460	0.0923	0.5439	0.0000
5	0.6881	0.2840	0.1928	0.0439	0.5884

Matrices for May.

$[\mathbf{M_{0,5}}]$ Matrix

	1	2	3	4	5
1	1.000	0.649	0.690	0.758	0.832
2	0.649	1.000	0.534	0.771	0.717
3	0.690	0.534	1.000	0.694	0.621
4	0.758	0.771	0.694	1.000	0.727
5	0.832	0.717	0.621	0.727	1.000

$[\mathbf{M_{1,5}}]$ Matrix

	1	2	3	4	5
1	0.2088	0.0243	0.0969	0.1348	0.2034
2	0.2456	0.3484	0.2032	0.2057	0.2626
3	0.2483	0.1805	0.2474	0.2261	0.2802
4	0.2646	0.2602	0.2773	0.4013	0.3605
5	−0.0352	−0.0215	0.0571	−0.0478	−0.0423

$[\mathbf{A_5}]$ Matrix

	1	2	3	4	5
1	0.1141	−0.1777	−0.0332	0.0929	0.1878
2	0.1170	0.3434	0.1299	−0.0979	−0.0215
3	0.0597	0.0137	0.1348	0.0629	0.1344
4	−0.0795	−0.0401	0.0971	0.3044	0.2304
5	−0.0110	0.0423	0.1124	−0.0719	−0.0708

$[\mathbf{B_5}]$ Matrix

	1	2	3	4	5
1	0.9659	0.0000	0.0000	0.0000	0.0000
2	0.6429	0.6616	0.0000	0.0000	0.0000
3	0.6501	0.0313	0.6884	0.0000	0.0000
4	0.7024	0.3177	0.1406	0.4305	0.0000
5	0.8760	0.2296	0.0696	0.1034	0.3899

Matrices for June.

$[\mathbf{M}_{0,6}]$ Matrix

	1	2	3	4	5
1	1.000	0.818	0.813	0.836	0.678
2	0.818	1.000	0.888	0.887	0.720
3	0.813	0.888	1.000	0.888	0.738
4	0.836	0.887	0.888	1.000	0.798
5	0.678	0.720	0.738	0.798	1.000

$[\mathbf{M}_{1,6}]$ Matrix

	1	2	3	4	5
1	0.7097	0.5411	0.5516	0.5626	0.6251
2	0.5693	0.5787	0.5214	0.5933	0.5648
3	0.5913	0.5221	0.5846	0.5502	0.5790
4	0.6158	0.5875	0.5482	0.6907	0.5817
5	0.1986	0.1813	0.2011	0.2320	0.1530

$[\mathbf{A}_6]$ Matrix

	1	2	3	4	5
1	0.5716	0.1613	0.1232	−0.1057	0.0345
2	0.1273	0.2455	0.1522	0.1376	0.0883
3	0.1685	0.1570	0.3042	−0.0204	0.1522
4	0.1564	0.1027	0.0790	0.4222	0.0218
5	0.1015	0.0380	0.0736	0.1654	−0.1247

$[\mathbf{B}_6]$ Matrix

	1	2	3	4	5
1	0.6906	0.0000	0.0000	0.0000	0.0000
2	0.5482	0.5235	0.0000	0.0000	0.0000
3	0.5162	0.3546	0.4140	0.0000	0.0000
4	0.5424	0.2624	0.1876	0.3122	0.0000
5	0.7674	0.2941	0.2323	0.3027	0.3434

Matrices for July.

$[\mathbf{M}_{0,7}]$ Matrix

	1	2	3	4	5
1	1.000	0.830	0.939	0.886	0.647
2	0.830	1.000	0.863	0.899	0.772
3	0.939	0.863	1.000	0.925	0.720
4	0.886	0.899	0.925	1.000	0.771
5	0.647	0.772	0.720	0.771	1.000

$[\mathbf{M}_{1,7}]$ Matrix

	1	2	3	4	5
1	0.6811	0.7640	0.7292	0.7199	0.6786
2	0.5377	0.6265	0.4786	0.4908	0.4542
3	0.6442	0.6851	0.6268	0.6486	0.6012
4	0.6171	0.6535	0.5645	0.6157	0.5870
5	0.1263	0.2307	0.1367	0.1431	0.2169

$[\mathbf{A}_7]$ Matrix

	1	2	3	4	5
1	0.1024	0.4756	0.1304	−0.1085	0.2568
2	0.2554	0.9667	−0.3625	−0.3871	0.1611
3	0.2338	0.4579	−0.0833	−0.0504	0.2145
4	0.2633	0.5311	−0.2910	−0.0259	0.2613
5	−0.0866	0.6249	−0.2612	−0.3351	0.2858

$[\mathbf{B}_7]$ Matrix

	1	2	3	4	5
1	0.6129	0.0000	0.0000	0.0000	0.0000
2	0.5736	0.4672	0.0000	0.0000	0.0000
3	0.6220	0.1300	0.2963	0.0000	0.0000
4	0.5778	0.2709	0.1329	0.3038	0.0000
5	0.7611	0.3221	0.1422	0.1838	0.3892

Matrices for August.

$[\mathbf{M_{0,8}}]$ Matrix

	1	2	3	4	5
1	1.000	0.770	0.904	0.878	0.735
2	0.770	1.000	0.804	0.839	0.704
3	0.904	0.804	1.000	0.929	0.790
4	0.878	0.839	0.929	1.000	0.848
5	0.735	0.704	0.790	0.848	1.000

$[\mathbf{M_{1,8}}]$ Matrix

	1	2	3	4	5
1	0.7753	0.7039	0.6830	0.6749	0.4103
2	0.5692	0.6252	0.5089	0.5630	0.2966
3	0.7369	0.6610	0.6617	0.6336	0.3025
4	0.6381	0.6339	0.5910	0.6234	0.2989
5	0.3317	0.3466	0.2550	0.2680	0.0304

$[\mathbf{A_8}]$ Matrix

	1	2	3	4	5
1	0.9680	0.5312	−0.4961	0.0153	−0.2808
2	0.4717	0.8239	−0.6299	0.3315	−0.4466
3	0.7487	0.5831	−0.1966	0.0155	−0.5026
4	0.4098	0.5799	−0.3399	0.4531	−0.5184
5	0.5153	0.7618	−0.5159	0.0112	−0.5283

$[\mathbf{B_8}]$ Matrix

	1	2	3	4	5
1	0.5650	0.0000	0.0000	0.0000	0.0000
2	0.3796	0.5820	0.0000	0.0000	0.0000
3	0.4310	0.1543	0.3545	0.0000	0.0000
4	0.5152	0.2045	0.2390	0.2828	0.0000
5	0.6386	0.1804	0.2751	0.2153	0.3838

Matrices for September.

$[\mathbf{M}_{0,9}]$ Matrix

	1	2	3	4	5
1	1.000	0.761	0.927	0.893	0.733
2	0.761	1.000	0.721	0.834	0.686
3	0.927	0.721	1.000	0.921	0.821
4	0.893	0.834	0.921	1.000	0.809
5	0.733	0.686	0.821	0.809	1.000

$[\mathbf{M}_{1,9}]$ Matrix

	1	2	3	4	5
1	0.7641	0.6073	0.6322	0.6083	0.4823
2	0.5165	0.6108	0.3952	0.4317	0.3359
3	0.6498	0.5991	0.5627	0.5852	0.4573
4	0.6149	0.6055	0.5076	0.5774	0.4565
5	0.2722	0.2940	0.1701	0.1794	0.1797

$[\mathbf{A}_9]$ Matrix

	1	2	3	4	5
1	1.059	0.224	−0.236	−0.225	−0.076
2	0.734	0.808	−0.663	−0.226	−0.057
3	0.681	0.314	−0.357	0.146	−0.106
4	0.692	0.387	−0.727	0.389	−0.080
5	0.616	0.454	−0.458	−0.432	0.134

$[\mathbf{B}_9]$ Matrix

	1	2	3	4	5
1	0.6148	0.0000	0.0000	0.0000	0.0000
2	0.4781	0.5269	0.0000	0.0000	0.0000
3	0.6573	−0.0442	0.3157	0.0000	0.0000
4	0.6362	0.1555	0.1485	0.2651	0.0000
5	0.7701	0.0831	0.3010	0.1057	0.3579

Matrices for October.

$[\mathbf{M}_{0,10}]$ Matrix

	1	2	3	4	5
1	1.000	0.788	0.897	0.859	0.688
2	0.788	1.000	0.827	0.906	0.738
3	0.897	0.827	1.000	0.922	0.700
4	0.859	0.906	0.922	1.000	0.757
5	0.688	0.738	0.700	0.757	1.000

$[\mathbf{M}_{1,10}]$ Matrix

	1	2	3	4	5
1	0.5306	0.3933	0.5034	0.5119	0.4433
2	0.4244	0.4733	0.4483	0.4727	0.3392
3	0.4766	0.3771	0.5505	0.5195	0.4325
4	0.4202	0.3698	0.4961	0.5113	0.3734
5	−0.0063	−0.0197	0.0008	0.0276	−0.0626

$[\mathbf{A}_{10}]$ Matrix

	1	2	3	4	5
1	0.5059	−0.1831	−0.2618	0.3470	0.1321
2	−0.2629	0.3604	0.4982	0.1414	−0.2389
3	−0.2973	−0.0506	0.7206	0.2547	−0.1130
4	−0.4595	−0.0612	0.5766	0.6263	−0.2283
5	−0.1381	−0.1215	0.0143	0.4347	−0.2417

$[\mathbf{B}_{10}]$ Matrix

	1	2	3	4	5
1	0.8362	0.0000	0.0000	0.0000	0.0000
2	0.6797	0.5195	0.0000	0.0000	0.0000
3	0.7556	0.1120	0.3117	0.0000	0.0000
4	0.7386	0.2897	0.0963	0.2366	0.0000
5	0.8211	0.3189	0.0848	0.0575	0.4280

Matrices for November.

$[\mathbf{M_{0,11}}]$ Matrix

	1	2	3	4	5
1	1.000	0.676	0.759	0.792	0.612
2	0.676	1.000	0.721	0.749	0.399
3	0.759	0.721	1.000	0.911	0.392
4	0.792	0.749	0.911	1.000	0.613
5	0.612	0.399	0.392	0.613	1.000

$[\mathbf{M_{1,11}}]$ Matrix

	1	2	3	4	5
1	0.5966	0.4610	0.5931	0.5224	0.4657
2	0.6323	0.6897	0.6385	0.6418	0.5124
3	0.6955	0.6271	0.7072	0.7100	0.6775
4	0.6055	0.5704	0.6164	0.6592	0.6586
5	0.1593	0.0170	0.1202	0.0916	0.1851

$[\mathbf{A_{11}}]$ Matrix

	1	2	3	4	5
1	0.3509	−0.1038	0.4923	−0.2482	0.1440
2	0.2109	0.6105	0.1998	−0.2525	−0.0318
3	0.2104	−0.1865	0.1878	0.2740	0.3315
4	0.1221	−0.2718	−0.0507	0.5471	0.3965
5	0.2753	−0.4688	0.0149	0.0361	0.3037

$[\mathbf{B_{11}}]$ Matrix

	1	2	3	4	5
1	0.7805	0.0000	0.0000	0.0000	0.0000
2	0.3807	0.5928	0.0000	0.0000	0.0000
3	0.3975	0.1678	0.4822	0.0000	0.0000
4	0.5177	0.2371	0.2680	0.3073	0.0000
5	0.6334	0.2010	−0.0680	0.4014	0.5436

Matrices for December.

$[\mathbf{M}_{0,12}]$ Matrix

	1	2	3	4	5
1	1.000	0.213	0.387	0.501	0.440
2	0.213	1.000	0.227	0.627	0.300
3	0.387	0.227	1.000	0.495	0.110
4	0.501	0.627	0.495	1.000	0.323
5	0.440	0.300	0.110	0.323	1.000

$[\mathbf{M}_{1,12}]$ Matrix

	1	2	3	4	5
1	0.5054	0.3114	0.3060	0.2534	0.3766
2	0.2379	0.6195	0.1853	0.3062	0.2097
3	0.5782	0.4981	0.7108	0.6355	0.3681
4	0.4388	0.5769	0.4008	0.4843	0.3471
5	−0.0840	−0.0920	−0.1266	−0.0969	0.1343

$[\mathbf{A}_{12}]$ Matrix

	1	2	3	4	5
1	0.555	0.171	0.717	−1.242	0.449
2	−0.131	0.976	−0.791	0.432	−0.055
3	0.031	−0.012	0.982	−0.408	0.219
4	0.059	0.510	−0.335	0.342	0.029
5	−0.182	0.028	0.200	−0.402	0.403

$[\mathbf{B}_{12}]$ Matrix

	1	2	3	4	5
1	0.7697	0.0000	0.0000	0.0000	0.0000
2	0.1665	0.6514	0.0000	0.0000	0.0000
3	0.1244	0.1593	0.6544	0.0000	0.0000
4	0.4120	0.2761	0.1855	0.5982	0.0000
5	0.5363	0.3657	0.0635	0.0468	0.7014