



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 –

$$[\mathbf{Z}_t] = [\mathbf{A}][\mathbf{Z}_{t-1}] + [\mathbf{B}][\epsilon_t] \quad (1)$$

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

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

$$\mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_{t-1}]^T\} = [\mathbf{A}]\mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T\} + \mathbf{E}\{[\mathbf{B}][\epsilon_t][\mathbf{Z}_{t-1}]^T\} \quad (2)$$

$$\Rightarrow [\mathbf{A}] = \mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_{t-1}]^T\} \mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T\}^{-1} \quad (3)$$

$$[\mathbf{A}] = [\mathbf{M}_1][\mathbf{M}_0^*]^{-1}$$

$$\text{where } [\mathbf{M}_0^*] = \mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T\}$$

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

$$\mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_t]^T\} = [\mathbf{A}]\mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_t]^T\} + \mathbf{E}\{[\mathbf{B}][\epsilon_t][\mathbf{Z}_t]^T\} \quad (4)$$

$$= [\mathbf{A}]\mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_t]^T\} + [\mathbf{B}]\mathbf{E}\{[\epsilon_t][\mathbf{Z}_{t-1}]^T[\mathbf{A}]^T + [\epsilon_t][\epsilon_t]^T[\mathbf{B}]^T\} \quad (5)$$

By substituting Equation [3] for $[\mathbf{A}]$, and rearranging,

$$\begin{aligned} \Rightarrow [\mathbf{B}][\mathbf{B}]^T &= \mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_t]^T\} - \mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_{t-1}]^T\}\mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T\}^{-1}\mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_t]^T\} \\ &[\mathbf{B}][\mathbf{B}]^T = [\mathbf{M}_0] - [\mathbf{M}_1][\mathbf{M}_0^*]^{-1}[\mathbf{M}_1]^T \end{aligned}$$

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 $[\mathbf{Z}_t][\mathbf{Z}_t]^T$ is equal to the covariance matrix calculated using $[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T$.

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 $[\mathbf{M}_0^*] = \mathbf{E}\{[\mathbf{Z}_{t-1}][\mathbf{Z}_{t-1}]^T\}$

in lieu of $[\mathbf{M}_0] = \mathbf{E}\{[\mathbf{Z}_t][\mathbf{Z}_t]^T\}$

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

The revised equations are thus –

For the general case,

$$[\mathbf{A}] = [\mathbf{M}_1][\mathbf{M}_0^*]^{-1} \quad (6)$$

$$[\mathbf{B}][\mathbf{B}]^T = [\mathbf{M}_0] - [\mathbf{M}_1][\mathbf{M}_0^*]^{-1}[\mathbf{M}_1]^T \quad (7)$$

and for the multiperiod case,

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

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

Kuczera [31] outlines a method to obtain consistent estimates of the $[\mathbf{A}]$ and $[\mathbf{B}]$ 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 $[\mathbf{A}]$ and $[\mathbf{B}]$ matrix given a monotone sample (*i.e.* when continuous records have different starting 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

Synopsis	ii
List of Figures	viii
List of Tables	xi
1 Introduction	1
1.1 Introduction	1
1.2 Study Objectives	3
1.3 Methodology	3
2 Data Analysis	5
2.1 Introduction	5
2.2 General Analysis	8
2.2.1 Trend	10
2.2.2 Periodicity	11
2.3 Transformations	12
2.3.1 Parametric Transformations	14
2.3.2 Moment Transformation Equations	17
2.3.3 Maximum Likelihood	19

2.4	Tests for Normality	21
2.4.1	Skewness Test for Normality	21
2.4.2	Shapiro–Wilk Test for Normality	22
2.4.3	Quantile–Quantile Plots	23
2.5	Robust Methods	23
3	Model Analysis	25
3.1	Introduction	25
3.2	Univariate Models	26
3.2.1	Autoregressive Models	26
3.2.2	Autoregressive Moving Average Models – ARMA	29
3.2.3	ARIMA modelling	30
3.3	Multivariate Models	30
3.3.1	Multivariate Autoregressive Models	31
3.3.2	Solution for the B Matrix	33
3.4	Multiperiod, Multivariate Models	37
3.5	Forecasting	38
4	System Background	40
4.1	Background to the Adelaide Water Supply System	40
4.2	The Headworks System	41
4.3	The Southern System	43
4.4	The Northern System	43
4.4.1	Little Para Subsystem	45
4.4.2	The Torrens System	45

5 Results	47
5.1 Introduction	47
5.2 The Raw Data	47
5.3 Quality of Streamflow Data	50
5.4 Quality of Rainfall Data	52
5.5 Data Set Analysis	52
5.6 Transformations	56
5.6.1 Parametric Transformation	57
5.6.2 Moment Transformation	59
5.7 Significance Testing	59
5.8 Investigation of Fitting Methods	64
5.9 Summary of Tests Adopted	64
5.10 Modelling	65
5.10.1 Multivariate Model	65
5.10.2 Comparison of Models in Generation & Forecasting	70
5.10.3 Key Station Approach	76
5.10.4 White Noise Analysis	76
5.10.5 Univariate Model	80
5.10.6 Univariate Generation and Forecasting	80
5.11 Forecasting	86
5.12 Application to Operational Hydrology	87
6 Summary	99
6.1 Introduction	99
6.2 Data	99

6.3	Transformations	100
6.4	Tests for Normality	101
6.5	Modelling	101
6.5.1	Univariate Model	101
6.5.2	Multivariate Model	101
6.6	Generation	102
6.7	Forecasting	103
6.8	Application to Operational Hydrology	103
7	Recommendations for Further Work	105
7.1	General Recommendations	105
7.1.1	Estimation of Parameters	105
7.1.2	Modelling	105
7.1.3	Model Verification	106
7.2	Specific Recommendations	106
7.2.1	Data	106
7.2.2	Forecasting	106
	Bibliography	107
A	Historical Statistics	112
B	Revised & Transformed Statistics	117
C	Myponga River, Serial Correlation Plots	126
D	Myponga River, Q-Q plots	139

CONTENTS

vii

E Streamflow Data Sets	152
F Monthly Cross Correlation Matrices	169
G Five Station Model Matrices	198

List of Figures

2.1	Stochastic Data Generation – Step Procedure	6
2.2	Aircraft Passenger Numbers – Raw Data	9
2.3	Aircraft Passenger Numbers – Transformed Data	9
2.4	Air. Pass. – De-Seasonalized, Normalized Data	12
2.5	Air. Pass. – De-Trended, De-Seasonalized, Normalized Data . .	13
4.1	Adelaide Metropolitan Water Supply District	42
4.2	System Schematic	44
5.1	Gauging Station Locations	49
5.2	Step Procedure for Data Analysis	53
5.3	Q-Q plot, Transformed January values at Warren River	60
5.4	Q-Q plot, Transformed January values at Myponga River	61
5.5	Q-Q plot, <i>Revised</i> , Transformed January values at Warren river	62
5.6	Quantile-Quantile Plot for White Noise	78
5.7	Q-Q Plot White Noise <i>vs</i> Transformed Data	79
5.8	Comparison of Stochastic & Deterministic Forecasts	82
5.9	Ave % difference of forecast <i>vs</i> lag	83
C.1	Raw Data – January/February	127

C.2	Transformed Data – January/February	127
C.3	Raw Data – February/March	128
C.4	Transformed Data – February/March	128
C.5	Raw Data – March/April	129
C.6	Transformed Data – March/April	129
C.7	Raw Data – April/May	130
C.8	Transformed Data – April/May	130
C.9	Raw Data – May/June	131
C.10	Transformed Data – May/June	131
C.11	Raw Data – June/July	132
C.12	Transformed Data – June/July	132
C.13	Raw Data – July/August	133
C.14	Transformed Data – July/August	133
C.15	Raw Data – August/September	134
C.16	Transformed Data – August/September	134
C.17	Raw Data – September/October	135
C.18	Transformed Data – September/October	135
C.19	Raw Data – October/November	136
C.20	Transformed Data – October/November	136
C.21	Raw Data – November/December	137
C.22	Transformed Data – November/December	137
C.23	Raw Data – December/January	138
C.24	Transformed Data – December/January	138
D.1	Q-Q plot, January Raw data	140

D.2 Q-Q plot, January Transformed data	140
D.3 Q-Q plot, February Raw data	141
D.4 Q-Q plot, February Transformed data	141
D.5 Q-Q plot, March Raw data	142
D.6 Q-Q plot, March Transformed data	142
D.7 Q-Q plot, April Raw data	143
D.8 Q-Q plot, April Transformed data	143
D.9 Q-Q plot, May Raw data	144
D.10 Q-Q plot, May Transformed data	144
D.11 Q-Q plot, June Raw data	145
D.12 Q-Q plot, June Transformed data	145
D.13 Q-Q plot, July Raw data	146
D.14 Q-Q plot, July Transformed data	146
D.15 Q-Q plot, August Raw data	147
D.16 Q-Q plot, August Transformed data	147
D.17 Q-Q plot, September Raw data	148
D.18 Q-Q plot, September Transformed data	148
D.19 Q-Q plot, October Raw data	149
D.20 Q-Q plot, October Transformed data	149
D.21 Q-Q plot, November Raw data	150
D.22 Q-Q plot, November Transformed data	150
D.23 Q-Q plot, December Raw data	151
D.24 Q-Q plot, December Transformed data	151

List of Tables

5.1	Station, Length & Period of Records	50
5.2	Monthly Shapiro–Wilk Values for Warren Data — January to June	63
5.3	Monthly Shapiro–Wilk Values for Warren Data — July to December	63
5.4	Location Parameters (τ) adopted	66
5.5	Comparison of Location Parameters, Full Length <i>vs</i> Concurrent Data Records	73
5.6	Onkaparinga – Generation Statistics, Zero Skew Transformation	74
5.7	Onkaparinga – Generation Statistics, Moment Transformation .	75
5.8	Normal Distribution Values given Exceedance Levels	84
5.9	South Para – Forecasting Model Statistic Comparison	88
5.10	Myponga – Forecasting Model Statistic Comparison	89
5.11	Gumeracha – Forecasting Model Statistic Comparison	90
5.12	Onkaparinga – Forecasting Model Statistic Comparison	91
5.13	Millbrook – Forecasting Model Statistic Comparison	92
5.14	Northern System Annual 90% Exceedance Yields	94
5.15	Southern System Yields	95
5.16	Northern System Monthly 90% Exceedance Yields	96
5.17	Northern System Monthly 70% Exceedance Yields	96

5.18 Southern System – Annual Pumping Costs for Optimization Results \$m	97
5.19 Northern System – Annual Pumping Costs for Optimization results \$m	98
A.1 Warren River & South Para River – Historical Statistics	113
A.2 Myponga River & Onkaparinga River – Historical Statistics	114
A.3 Gorge & Gumeracha Weirs – Historical Statistics	115
A.4 Little Para & Clarendon rainfall – Historical Statistics	116
B.1 Warren River statistics	118
B.2 South Para River statistics	119
B.3 Myponga River statistics	120
B.4 Onkaparinga River statistics	121
B.5 Gorge Weir statistics	122
B.6 Gumeracha Weir statistics	123
B.7 Little Para River statistics	124
B.8 Clarendon P.O. rainfall statistics	125