

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/282122426>

# Reconstruction of bore hydrograph trends in fractured rock aquifers using data mining techniques

Thesis · September 2014

CITATIONS  
0

READS  
78

## 1 author:



Aleksandra Rančić

15 PUBLICATIONS 104 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Dryland salinity in the Murray–Darling Basin in south-eastern Australia [View project](#)



CATplus [View project](#)

**PLEASE TYPE**

**THE UNIVERSITY OF NEW SOUTH WALES**  
**Thesis/Dissertation Sheet**

Surname or Family name: **Rancic**

First name: **Aleksandra**

Other name/s:

Abbreviation for degree as given in the University calendar:

School: **Civil and Environment Engineering**

Faculty:  
**UNSW**

Title: Reconstruction of bore hydrograph trends in fractured rock aquifers using data mining techniques

**Abstract**

The lack of recorded groundwater hydrographs presents a problem for many groundwater studies. This is especially the case in the areas of fractured rock basement. In the New South Wales (NSW) portion of the Murray-Darling Basin (MDB), Australia, long-term hydrograph trends were required to separate groundwater components of climate forcing from the effects of deforestation on groundwater rise. To overcome the problem of missing long-term hydrographs, the data-mining technique was developed to enable reconstruction of aquifer behaviour from independent point data acquired at different times.

The multi-decadal descriptive data-mining methodology described here uses independent Standing Water Level data (SWL - depth to water table), recorded in the groundwater database after borehole construction. Data are restricted to the Palaeozoic and early Mesozoic fractured rock aquifers closest to the land-surface. Catchments with pronounced rainfall gradients are divided to preserve homogeneity and avoid spatial-temporal confounding. Annual SWL time-series are created using a median function. The noise is filtered by applying 21-year moving average. The methodology is demonstrated on the Namoi Catchment. Results are validated using existing short-term hydrographs and long-term proxy datasets: rainfall trends. Evaluation is assessed using Root Mean Square Error method. It was validated on the Lachlan and Murrumbidgee catchments, where historic hydrographs exist.

The approach described here has been successful in recreating past multi-decadal groundwater level trends in the fractured rock areas in eastern NSW. Aquifers rapidly responded to wetter climatic conditions and equilibrated within two to seventeen years. The influences of changes in land-use were detected in the Lachlan West section, where wheat-growing replaced pastoralism, and was separated from climatic drivers. Everywhere else the climate was the main driver of groundwater rise and was directly implicated in triggering dryland salinity. This observation helps clarify the mechanisms of dryland salinity occurrence.

The methodology for reconstruction of aquifer behaviour has obvious applicability in many other parts of the world where only basic bore completion data exist, but long-term monitoring has not been carried out. In particular it can be applied to the basement areas in Africa, Australia and India.

**Declaration relating to disposition of project thesis/dissertation**

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

*Aleksandra Rancic*

Signature

*R. C. Aarwout*

Witness

*5/9/2014*

Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

**FOR OFFICE USE ONLY**

Date of completion of requirements for Award:

**THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS**

#### COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.'

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed ..... *Aleksandrs Rasevics* .....

Date ..... *05/09/2014* .....

#### AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed ..... *Aleksandrs Rasevics* .....

Date ..... *05/09/2014* .....

#### **ORIGINALITY STATEMENT**

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed ..... *Aleksandar Radojević* .....

Date ..... *05/09/2014* .....

# Reconstruction of bore hydrograph trends in fractured rock aquifers using data mining techniques

Aleksandra Rančić

Submitted in total fulfillment of the requirements of the degree of  
Doctor of Philosophy.

*UNIVERSITY OF NEW SOUTH WALES*  
*School of Civil and Environmental Engineering*  
September 5, 2014



# Abstract

The lack of recorded groundwater hydrographs presents a problem for many groundwater studies. This is especially the case in the areas of fractured rock basement. In the New South Wales (NSW) portion of the Murray-Darling Basin (MDB), Australia, long-term hydrograph trends were required to separate groundwater components of climate forcing from the effects of deforestation on groundwater rise. To overcome the problem of missing long-term hydrographs, the data-mining technique was developed to enable reconstruction of aquifer behaviour from independent point data acquired at different times.

The multi-decadal descriptive data-mining methodology described here uses independent Standing Water Level data (SWL – depth to water table), recorded in the groundwater database after borehole construction. Data are restricted to the Palaeozoic and early Mesozoic fractured rock aquifers closest to the land-surface. Catchments with pronounced rainfall gradients are divided to preserve homogeneity and avoid spatial-temporal confounding. Annual SWL time-series are created using a median function. The noise is filtered by applying 21-year moving average. The methodology is demonstrated on the Namoi Catchment. Results are validated using existing short-term hydrographs and long-term proxy datasets: rainfall trends. Evaluation is assessed using Root Mean Square Error method. It was validated on the Lachlan and Murrumbidgee catchments, where historic hydrographs exist.

The approach described here has been successful in recreating past multi-decadal groundwater level trends in the fractured rock areas in eastern NSW. Aquifers rapidly responded to wetter climatic conditions and equilibrated within two to seventeen years. The influences of changes in land-use were detected in the Lachlan West section, where wheat-growing replaced pastoralism, and was separated from climatic drivers. Everywhere else the climate was the main driver of groundwater rise and was directly implicated in triggering dryland salinity. This observation helps clarify the mechanisms of dryland salinity occurrence.

The methodology for reconstruction of aquifer behaviour has obvious applicabil-

ity in many other parts of the world where only basic bore completion data exist, but long-term monitoring has not been carried out. In particular it can be applied to the basement areas in Africa, Australia and India.

# Acknowledgments

I want to express my thankfulness to the contagiously inspirational hydrogeologist, Gabriel Salas. He had a patience to tutor, teach, guide and encourage me and pass to me his supporting network, including a historian, Kathryne Welsh, and hydrogeologist, Ann Smithson, who also helped me on this journey.

I am sincerely grateful to Dr David Read, who tirelessly encouraged me and diligently proofread and re-read the text of this thesis and made sure they were finished and written in English, not in Serbglish. My gratitude to Dr Dugald Black, who, I believe, trusted my abilities and appointed me to work on the NSW Murray-Darling Basin Salinity Audit, from which this research originated. Dr Martin Sillance encouraged me to turn the research into the thesis. I would like to thank Dr Greg Summerell for being the most supportive boss in the world; and all others from the Wagga Wagga Research Centre (Dr Peter Barker, Sandy Grant, Frank Harvey, Sarrah McGouch, Geoffrey Beale, Michelle Miller and Lisa Pike) and the Office of Environment and Heritage and their predecessors for supporting my studies. My appreciation to Dr Amrit Kathuria, Dr Hue Jones and Dr Terry Coen for statistical help; to my supervisors, Dr Ian Acworth and Dr Bill Johnston for the advice, constructive criticism and insightful clarifications; to Rex Wagner who loaned me his thesis; to my colleagues from the CATplus team: Brendan Christy, Dr Kym Lowell and Terry McLean for their support; to my doctors, Michael Kefaloukas and Rod Burgess and surgeon George Kourtesis for making sure I am alive, well and healthy; to my daughter Milica, for her love and emotional support and to my son Milos for being always there for me, with no exception.

And many thanks to all others, friends and colleagues who helped and supported me on the way.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Water table fluctuations . . . . .	1
1.2	The role of recharge in water cycle . . . . .	2
1.3	Hypothesis . . . . .	6
1.4	Aim . . . . .	6
1.5	Study area . . . . .	8
1.5.1	Topography . . . . .	8
1.5.2	Climate . . . . .	8
1.5.3	Geology . . . . .	12
1.5.4	Hydrogeology . . . . .	15
1.5.5	History of land-use change in the study area . . . . .	18
1.6	Structure of the Thesis . . . . .	21
<b>2</b>	<b>Literature Review</b>	<b>23</b>
2.1	Data mining . . . . .	23
2.1.1	Data mining tasks . . . . .	24
2.1.2	Components of data mining algorithms . . . . .	26
2.1.3	Data mining in hydrology . . . . .	27
2.1.4	Data mining in hydrogeology . . . . .	29
2.2	Methods for reconstructing aquifer behaviour from point data . . . . .	31
2.3	Summary . . . . .	33
<b>3</b>	<b>Methods</b>	<b>34</b>
3.1	Data sources for data mining . . . . .	34
3.2	Bore data mining method . . . . .	36
3.2.1	Data selection . . . . .	40

3.2.2	Creation of composite SWL time-series . . . . .	47
3.2.3	Noise filtration . . . . .	57
3.3	Summary . . . . .	62
<b>4</b>	<b>Results</b>	<b>63</b>
4.1	Data sources for data mining . . . . .	63
4.2	Bore data mining . . . . .	64
4.2.1	Bore selection . . . . .	65
4.2.2	Creation of composite SWL time-series . . . . .	66
4.2.3	Noise filtration . . . . .	78
4.3	Summary . . . . .	86
<b>5</b>	<b>Validation</b>	<b>87</b>
5.1	Validation of aquifer behaviour based on recorded hydrographs . . . . .	87
5.2	Validation of aquifer behaviour using rainfall as the surrogate test data	96
5.3	Summary . . . . .	101
<b>6</b>	<b>Discussion</b>	<b>102</b>
6.1	Bore data mining technique . . . . .	102
6.2	New knowledge gained from the hydrograph trend reconstruction . .	104
6.3	Summary . . . . .	109
<b>7</b>	<b>Conclusions</b>	<b>110</b>
<b>Bibliography</b>		<b>113</b>
<b>A Step-change analysis of individual rainfall stations</b>		<b>127</b>
<b>B Conference Publications</b>		<b>131</b>

# List of Figures

1.1	Conceptual model of a soil water balance. . . . .	5
1.2	Study area. . . . .	7
1.3	Topography. . . . .	9
1.4	Average annual rainfall in the study area. . . . .	10
1.5	Geology of eastern NSW. . . . .	13
1.6	Legend for Figure 1.5 - continuation . . . . .	14
1.7	Historic growth of a) sheep numbers in NSW and b) NSW population and area under cereal crops. . . . .	20
3.1	Standing water level. . . . .	35
3.2	Bores projected on Geology map of Namoi catchment. . . . .	41
3.3	Namoi section: choice of the central tendency model for aggregation of SWL points into annual data: Annual median (Med Ann SWL), arithmetic average (Av Ann SWL) and spatially weighted average (SpAv Ann SWL). . . . .	48
3.4	Normalised, 21-year moving average filter, for rainfall and SWL in Namoi section. . . . .	50
3.5	Splines filtering technique - Namoi section. . . . .	52
3.6	Exploratory analysis for Namoi Section: Timeseries of bore altitude and SWL and spatial distribution in relation to time of construction. Points are coloured in red for bores constructed in 1902–1949 period, in green for 1950–1969, in magenta for 1970–1989 and in black for 1990–2004 periods. . . . .	55
3.7	Residual mass curves of rainfall and SWL in Namoi catchment. . . . .	59

4.1 Spatial distribution of bores in Lachlan catchment. Top: Bores in Laclan West section coloured in magenta, in Mid section in green and in Eastern section in red. Bottom: Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	70
4.2 Altitude of bores in Lachlan catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	71
4.3 SWL of bores in Lachlan catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	72
4.4 Spatial distribution of bores in Murrumbidgee catchment. Top: Bores in Murrumbidgee West section coloured in black, in North (Inner West) section in green, in Northeast section in magenta and in Southeast section in red. Bottom: Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	73
4.5 Elevation of bores in Murrumbidgee catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	74
4.6 SWL of bores in Murrumbidgee catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black. . . . .	75
4.7 Water table trend in Lachlan: a) West, b) Mid and c) East section, in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL). . . . .	79
4.8 Water table trend in Murrumbidgee West and Inner West sections in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL). . . . .	80
4.9 Water table trend in Murrumbidgee East a) combined and b) North and c) South portion in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL). . . . .	81
4.10 Residual mass curve of median annual SWL (Med Ann SWL) in Lachlan: a) West, b) Mid and c) East section. . . . .	82

4.11	Residual mass curve of median annual SWL (Med Ann SWL) in Murrumbidgee West section and its Inner West portion. . . . .	83
4.12	Residual mass curve of median annual SWL (Med Ann SWL) in Murrumbidgee East section and its northern and southern portions. . . . .	84
5.1	Location of bores used for model validation. . . . .	88
5.2	Monitoring bores in Wattle Retreat and Yass catchments. . . . .	89
5.3	Modelled SWL trend (21 yr Med Ann SWL) in Lachlan Mid section matches a) recorded hydrograph of bore GW036478-2 and b) its annual average. c) Wattle Retreat hydrographs exhibit falling tend. Med Ann SWL represents Median Annual SWL. . . . .	90
5.4	Modelled SWL trend (21yr Med Ann SWL) in North portion of Murrumbidgee East section matches a) recorded hydrograph of bore GW036792-1 and b) its annual average. c) Hydrographs in Yass catchment exhibit falling tend. Med Ann SWL represents Median Annual SWL. . . . .	91
5.5	Modelled SWL trend (21 yr Med Ann SWL) in Murrumbidgee West section matches recorded hydrograph of bore GW030386 from Tarcutta Creek alluvial, which replicates trends captured by GW061119 and GW062952 from mid and upper Tarcutta Catchment. Med Ann SWL represents Median Annual SWL. . . . .	93
5.6	Trends in rainfall (21yr Ann Rain) shifted forward by a) 5 b) 11 and c) 16 years, compared to trends in SWL (21yr Med Ann SWL) and corresponding confidence intervals in the East, Mid and West sections of the Lachlan River catchment. . . . .	98
5.7	Residual mass curves of rainfall (Ann Rainfall) shifted forward by a) 5 b) 11 and c) 16 years, compared to residual mass curves of median annual SWL (Med Ann SWL) in the East, Mid and West sections of the Lachlan River catchment. . . . .	99
6.1	Water level record for Lake George in southern NSW. . . . .	108

# List of Tables

3.1	Data mining algorithmic components. . . . .	37
3.2	Namoi section – Choice of length (years) for the moving average filter. . . . .	51
3.3	Number of data points by year in Namoi section. . . . .	53
3.4	Change-point analysis of SWL data in Namoi section. . . . .	60
3.5	Change-point analysis of individual rainfall series for Namoi section. . . . .	60
4.1	Available SWL data points. . . . .	63
4.2	Selection of PPD* rainfall stations. . . . .	64
4.3	Bore Selection. . . . .	65
4.4	Number of data points by year in each catchment section. . . . .	66
4.5	Post-hoc segmentation. . . . .	77
4.6	Change-point analysis of SWL data. . . . .	85
4.7	Change-point analysis of composite rainfall series. . . . .	86
5.1	Root Mean Square Error (RMSE) between normalised SWL trends and annual timeseries of monitored hydrographs. . . . .	95
5.2	Abrupt changes and lags in groundwater response to rainfall. . . . .	100
A.1	Change-point analysis of individual rainfall series. . . . .	127
A.2	Change-point analysis of rainfall variability for individual rainfall sta- tions. . . . .	129



# Chapter 1

## Introduction

There is a growing concern about the impacts of current climate change on water resource availability. However, this concern can only be allayed if information on the way in which groundwater systems have behaved and reacted to past climate change is understood. Without this information it is not possible to calibrate, with any degree of accuracy, a model on which to base reliable predictions. In many parts of the world, information describing the way in which groundwater levels respond to climate variation does not exist, because the monitoring records of bores are rare and short. This is particularly the case with the low-yielding aquifers in fractured rock, where exploitation is restricted mainly to stock and domestic use, or small municipal exploitation, such as in large areas of Africa, Australia and India [Acworth, 1981]. Studies of fractured rock aquifers and the determination of aquifer parameters in these areas has often been neglected because of the lack of investment and interest that accompanies huge commercial exploitation.

However, databases often contain a huge amount of independent point data of depth to water table when bores were drilled, that reach back decades or even centuries. They are the remnants of the records left during production bore drilling, necessary for adequate pump positioning. This thesis demonstrates a data mining method for reconstruction of groundwater trends in fractured rock aquifers using depth to water table records, collected when the bores were constructed.

### 1.1 Water table fluctuations

Water tables can be characterised by their spatial and temporal variability.

The depth to groundwater varies within a catchment, depending on the topographical position and proximity to recharge or discharge areas. Upper-catchment recharge areas are usually characterised by greater water table depths, larger vertical hydraulic gradients, and unconfined groundwater conditions. In discharge areas and in the lower parts of catchments, groundwater is shallower, has an upward hydraulic gradient, and is often confined by materials of lower permeability. Between these two positional extremes, depth to water table has a smaller range of values and vertical hydraulic gradients are less pronounced.

The hinge-point between recharge and discharge areas does not stay constant with time, but can move. Hence, the same area towards the middle of the system can act as a discharge area under wet conditions or as a recharge area under dry conditions.

Aquifers can respond to recharge by hydraulic transmission of pressure changes, which can be rapid, and as physical flow of water from areas of high to low hydraulic head. The time needed by a groundwater system to adapt to a new hydrologic equilibrium depends primarily on its storage size, length and hydraulic conductivity.

## 1.2 The role of recharge in water cycle

Water balance studies are concerned with the quantification of fluxes associated with the myriad of pathways that water molecules follow within the water cycle.

Aquifers are primarily recharged by rainfall or leakage from rivers. There are two main drivers for recharge fluctuation, and hence, water table fluctuation:

- climatic fluctuations and
- broad-scale land-use changes.

Water table fluctuations can also be caused by a number of secondary factors, including:

- groundwater abstractions with localised or widespread impacts, or by
- river regulation by altering flooding and flow regimes.

Groundwater levels respond to climate change because that change produces variation in recharge. Recharge variations can be the result of superimposed climatic

cycles, with annual cycles superimposed upon longer cycles with various periodicities and amplitudes, ranging from sub-decadal, decadal or to changes over more than hundreds or thousands of years, such as glacial periods.

An increase in recharge, irrespective of its cause, results in a rise in the water table. The more significant the recharge increase, the more the water table becomes shallower and in more locations the water table approaches the ground surface to produce seeps or springs.

Water discharges from aquifers naturally by evapotranspiration or through seepage to streams, lakes or the ocean.

During a rainfall event a portion of water that reaches the ground surface can assume a lateral path (as the surface runoff) and the remaining flux infiltrates vertically into the soil. A large proportion of the infiltrated water later leaves the soil through evapotranspiration. The replenished soil profile increases its moisture content, and if its holding capacity is exceeded, the surplus drains from the profile. In layered soils with sharp contrasting hydraulic properties a portion of this surplus can be evacuated from the profile laterally, as a shallow subsurface flux. The remaining flux moves vertically down, becoming deep drainage and recharging aquifers. Groundwater storages are replenished through this recharge and through the seepage from river channels. Water leaves aquifers through discharge to streams, lakes or oceans, through evapotranspiration, where the water table is shallow enough for the plant roots to tap into the capillary fringe and also within the discharge zones, where the water table reaches the ground surface, forming seeps and springs.

When recharge increases the groundwater storage starts to fill, and the water table rises, expanding the discharge area. Gradually the discharge increases to equilibrate with recharge. Recharge component is generally the smallest component of the waterbalance with evapotranspiration being the largest across the slopes of the Murray-Darling Basin in NSW [Tuteja et al., 2003]. Paradoxically, the volume of groundwater exceeds by far that of surface water, due to its slow travel and long residence time.

Thus, there are two major factors that can substantially increase recharge and fill groundwater storages, (i) increase in rainfall and (ii) decrease in evapotranspiration. Abrupt increase (or decrease) in rainfall frequently occurs as a manifestation of natural climate variability. Decrease in transpiration is caused by reduced plant demand or its ability to extract water. In nature it accompanies a decrease in solar

radiation. However, it can also be imposed by human intervention and arise as a consequence of the land clearing. Removal of deep-rooted, native trees, adapted to the local climate and their replacement with shallow-rooted plant species reduces the depth of root zone from which water can be efficiently evapotranspirated. Increase in recharge, irrespective of its cause, results in water table rise, unless counteracted by equal and simultaneous increase in output, such as through extractions. The more significant the recharge increase, the more water tables become shallower and in more points they come close to or intersect the ground surface, increasing the number and extent of discharge areas, seeps or springs. Where discharge is slowed down by the low hydraulic conductivity of the clay-rich soils, seeps occur and waterlogged areas form. In climates where the potential evapotranspiration substantially exceeds rainfall, water evaporates, leaving and accumulating salts near and on the land surface. If rainfall is insufficiently powerful to effectively remove accumulated salt land degradation can occur.

A good approximation to the calculation of recharge to a shallow aquifer is possible using the water balance approach [Lerner et al., 1990, Rushton, 2003]. This is shown diagrammatically in Fig. 1.1.

The soil is considered to provide a storage capacity that is determined by the soil porosity and rooting depth. If a soil is wetted above field capacity, it can no longer hold moisture in tension and water drains from the base of the soil. This is represented in Fig. 1.1 as the soil reservoir filling and overflowing down the central drain pipe to become aquifer recharge. The amount of moisture held in the soil reservoir is determined by rainfall or irrigation additions and by evapotranspiration. Lateral flow is ignored in this simple conceptual model.

Evapotranspiration reduces the available soil moisture and creates a soil moisture deficit. Rainfall is required in excess of evapotranspiration to return the soil to field capacity (eliminating the soil moisture deficit) and to become freely draining again. The most successful recharge modelling requires the water balance to be run daily so that recharge is removed from the soil each day [Rushton, 2003]. This is a reasonable approximation and is confirmed by the fact that soil profiles dry out at the surface fairly quickly after rainfall in the study area.

The significance of the rooting depth of the crops is that the magnitude of the soil reservoir is the product of the available water in the soil and the depth of the roots. A soil that has no plant cover can not develop a soil moisture deficit as there

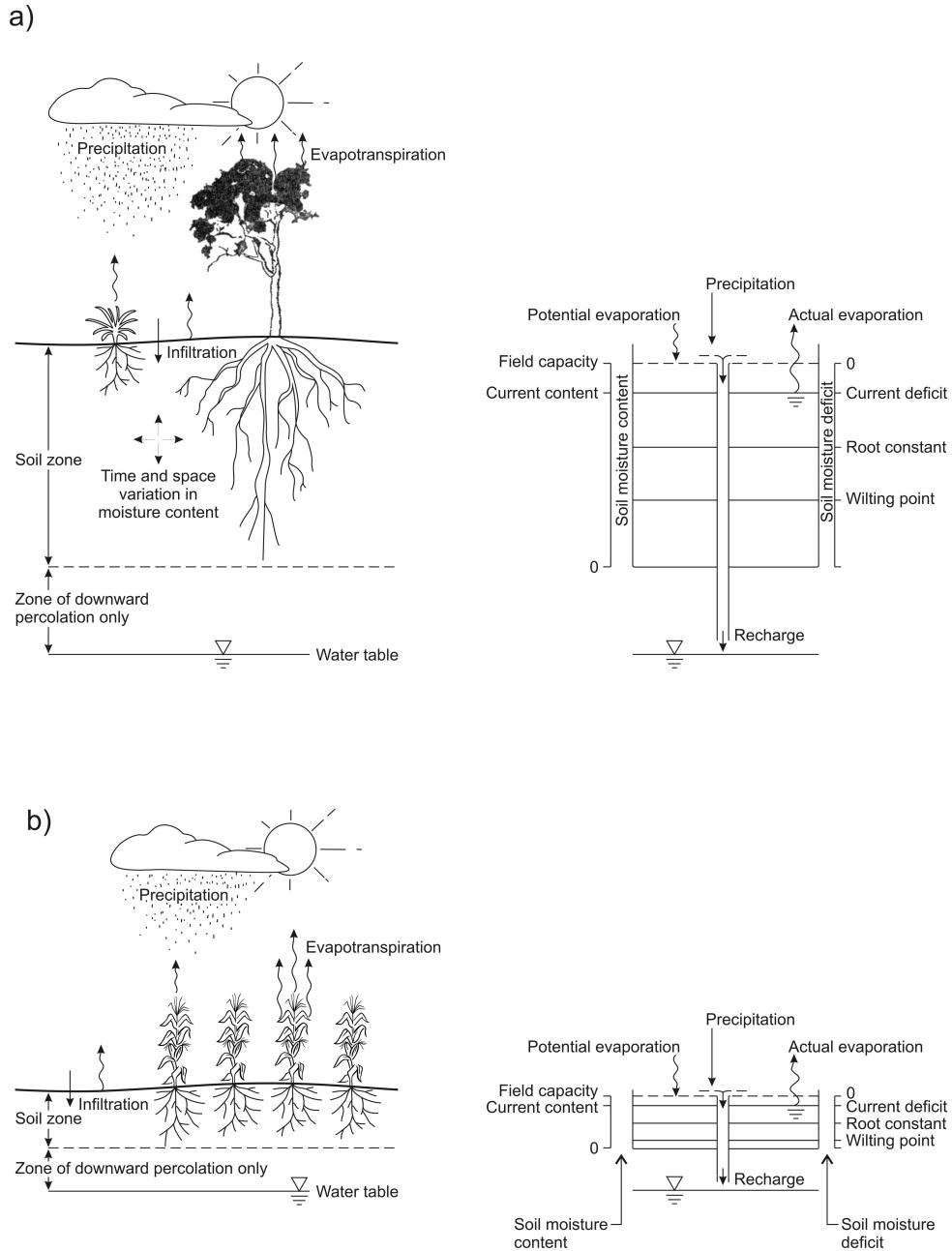


Figure 1.1: Conceptual model of a soil water balance.

is effectively no transpiration loss (accepting that minor evaporation loss of moisture direct from the soil to the atmosphere can occur [Allen et al. (1998)]). Rainfall on a bare soil therefore rapidly extinguishes the soil moisture deficit, returning the soil to a freely draining condition and allowing groundwater recharge. By contrast, a natural vegetation cover (Fig. 1.1a), with tree roots extending to several metres,

can develop a soil moisture deficit of more than 250mm. Such a large deficit requires a long period of rainfall with lower evapotranspiration before recharge can occur. Actual evapotranspiration (ET<sub>a</sub>) is considered to continue at the maximum rate possible (ET<sub>0</sub>) until the availability of soil moisture limits the growth of the crop [Rushton, 2003]. Actual transpiration (ET<sub>a</sub>) then reduces to zero at the point at which the crop dies (wilting point). A soil planted with shallow rooted crops (Fig. 1.b) reacts more closely to a bare soil than a fully vegetated soil. For this reason, recharge occurs more frequently in a soil where the natural deep rooted vegetation has been replaced by shallow rooted crops.

### 1.3 Hypothesis

Major temporal trends and the change points in the position of the water table in an area of fractured rock can be reconstructed by an agglomeration of independent observations of the water table depth made in bores at the time they were constructed.

### 1.4 Aim

The aim of this thesis is to describe and validate a technique for reconstruction of groundwater trends and major shifts in position of water table from independent point data. This data-mining technique draws on scattered point-data of standing water level (SWL) of production bores, recorded while the bores were drilled. SWL represents depth from the ground surface to the water level in the bore (see Figure 3.1). These scattered point-data can be aggregated across the catchment areas over time to describe the changes in the system as a whole. Median annual SWL time-series can be constructed and further processed to describe trends and shifts in hydrologic regime in the system as a whole.

The resulting trends can be validated by comparison with monitoring data and rainfall trends. Shifts in water table can be compared with the corresponding rainfall shifts.

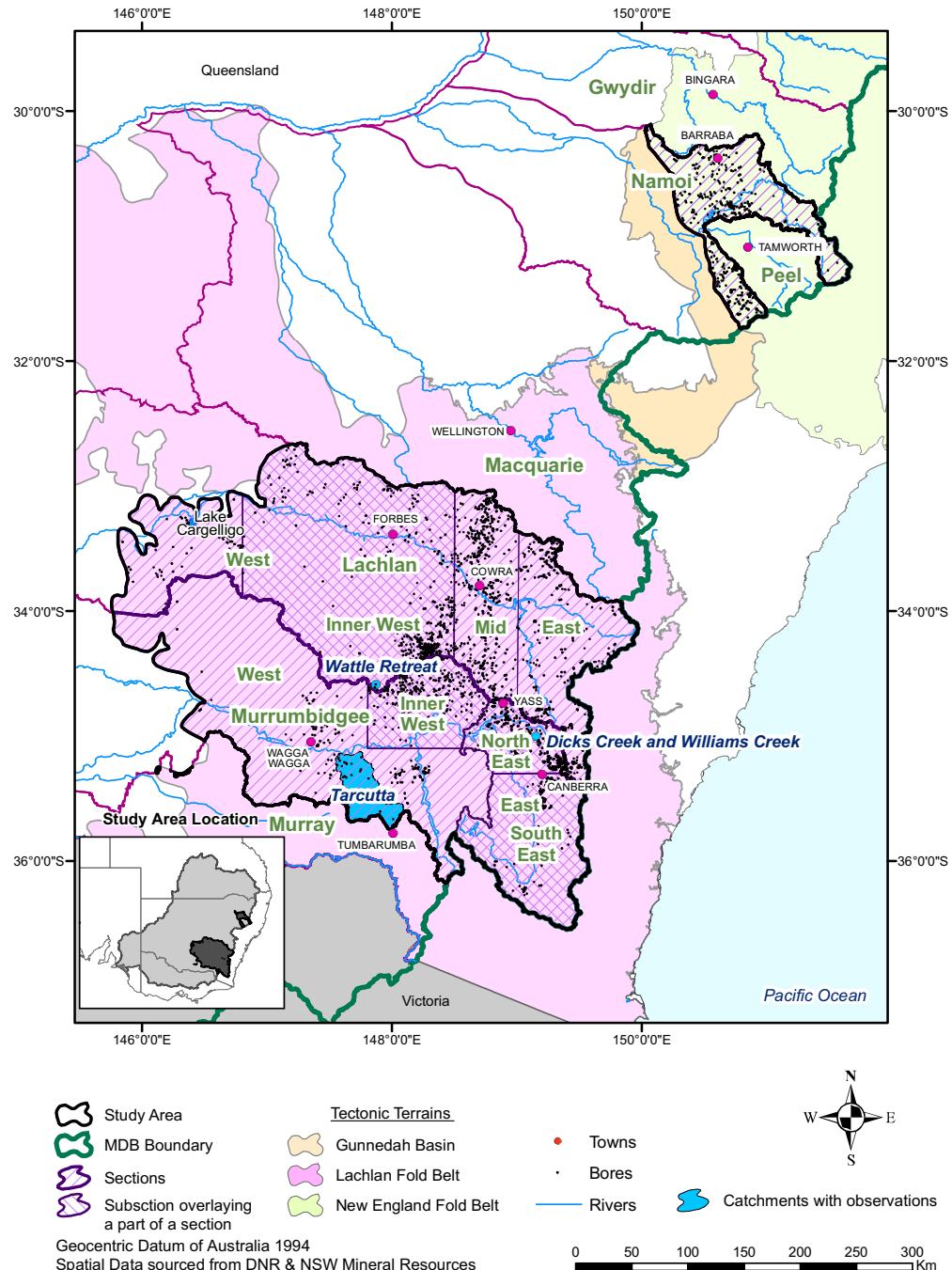


Figure 1.2: Study area.

## 1.5 Study area

The study area in which this hypothesis was developed and tested covers three catchments of the western slopes of the Great Dividing Range in NSW, Australia (Figure 1.2 and Table 4.5):

- a portion of the Namoi catchment in the New England Fold Belt, from the summer rainfall zone;
- a portion of the Lachlan catchment in the Lachlan Fold Belt, from the uniform rainfall zone and
- a portion of the Murrumbidgee catchment in the Lachlan Fold Belt, covering uniform and winter rainfall zones.

### 1.5.1 Topography

The western watershed of the Great Dividing Range includes mountainous to hilly erosive landscapes in the east, with some peaks exceeding 2000 m above sea level, grading westwards into depositional plains with low to gentle slopes and some protruding relict mountain ranges (Fig. 1.3).

### 1.5.2 Climate

#### Rainfall

Rainfall decreases with the distance from the coast and increases with rising altitude. Hence, the elevated areas along the Great Dividing Range receive the highest average annual rainfall, over 800 mm/year (Figure 1.4). Rainfall declines to less than 400 mm/year in the west. However, from north to south it is possible to define three seasonal rainfall zones:

Summer rainfall zone. The southern catchments of the Queensland and the northern catchments of the NSW, including the Namoi catchment, experience a dominance of summer over winter rainfall. Summer rains originate from troughs between the travelling anticyclones (high-pressure cells) that draw in moist, maritime airflows from tropical regions to the north [Gentilli, 1971], as well as the remnants of tropical cyclones that come down from the north.

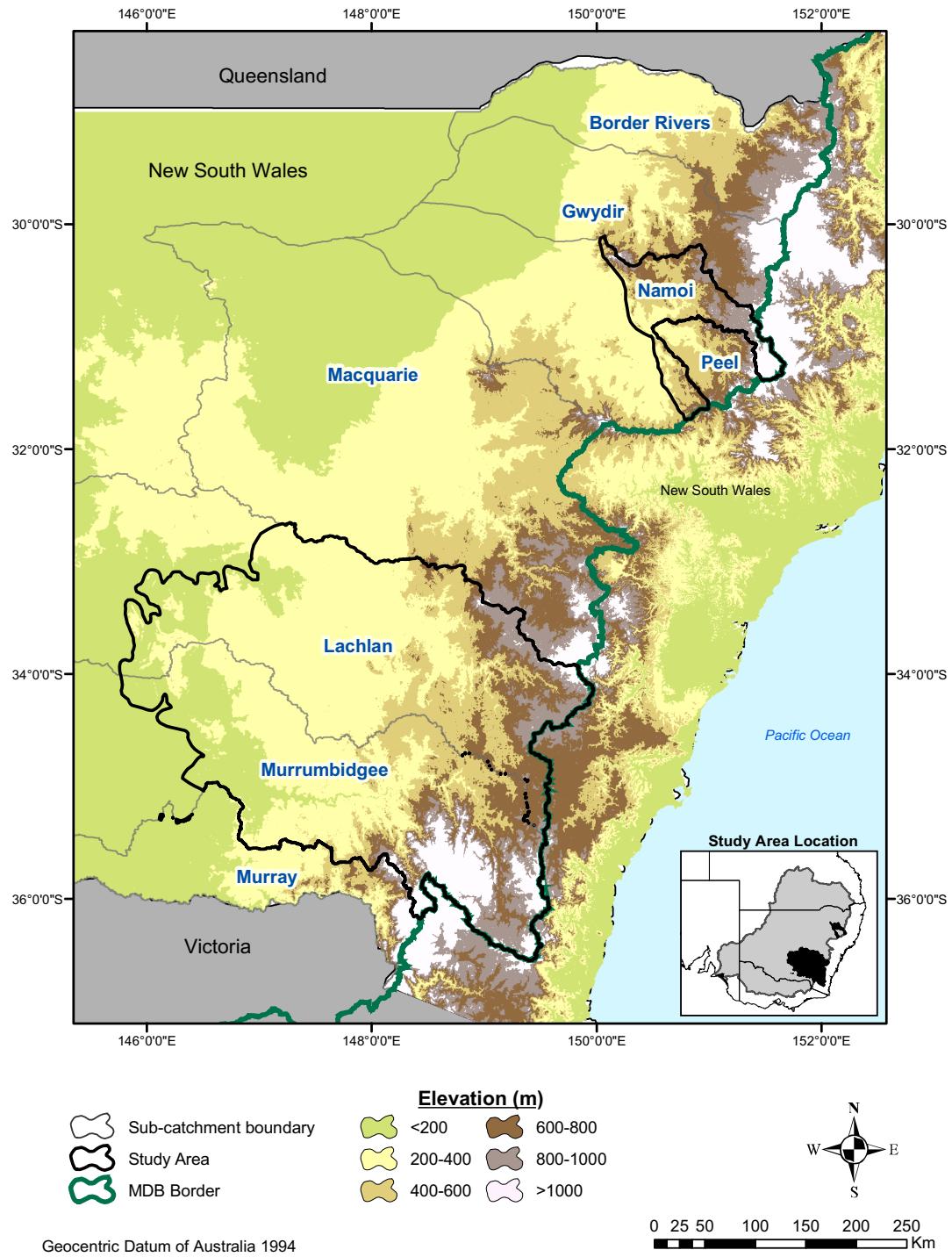


Figure 1.3: Topography.

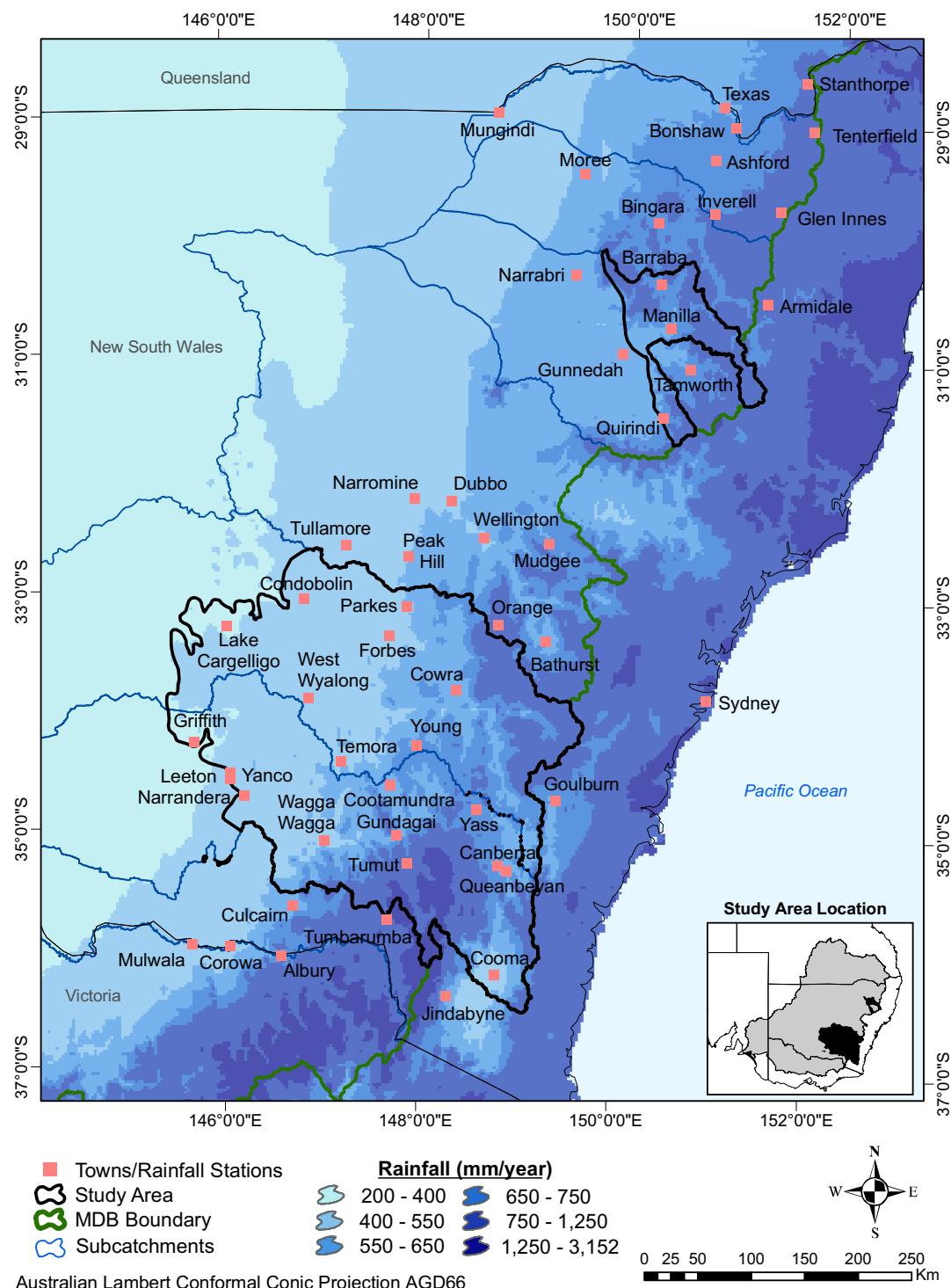


Figure 1.4: Average annual rainfall in the study area.

Uniform rainfall zone. The central catchments of the NSW, including the Lachlan catchment, and the upper portion of the Murrumbidgee catchment above Burinjuck Dam, experience a non-seasonal rainfall pattern, with no more than 30% variation between summer and winter. The dominance of the winter rain-bearing westerlies over the summer rain-bearing troughs diminishes northwards.

Winter rainfall zone. The southern catchments the NSW (south-western Murrumbidgee and Murray), Victoria and Tasmania experience a dominance of winter over summer rainfall. The weather in this region is dominated by eastward-moving anticyclones in summer, and rain-bearing westerly winds, low-pressure cells and associated troughs in winter [Dodson, 1998].

Differences in seasonal rainfall distribution between the north and south of the state are the result of the change in the average latitudinal position of the *subtropical high-pressure ridge*, which follows the Sun in its annual cycle. In summer, the Sun is over the Southern Hemisphere and the ridge passes to the south of mainland Australia [Deacon, 1953, Pittock, 1973]. This exposes southern NSW to easterly winds, and northern areas to intrusions of warm, moisture-laden tropical cyclones. The easterly winds bring moisture onto the southern NSW coast and ranges, but the inland is generally dry. In winter, as the Sun moves to the Northern Hemisphere and the subtropical high-pressure systems also move north, southern Australia comes under the influence of mid-latitude cyclones, and predominantly westerly to south-westerly winds. This brings rain-bearing low-pressure cells and associated cold frontal systems across the inland of NSW, while areas such as the southern coastal strip and ranges receive little rain [Deacon, 1953].

## Evapotranspiration

Throughout the study area, potential evapotranspiration is highest in summer and lowest in winter, and the average evapotranspiration declines from north to south and with increasing altitude. Although in southern Australia there is usually a surplus of rainfall over evapotranspiration during winter, it is only in high-altitude, high-rainfall zones (generally where rainfall exceeds 850 mm) that there is an annual surplus.

## Temperature

Daytime temperatures are generally mild to hot in summer, and mild to cool in winter. Throughout the area, frost is common in winter. Snowfalls are common in winter only at elevations above 1500 m.

### 1.5.3 Geology

Four major geological provinces occur in the study area [Branagan and Packham, 2000], as shown in Figure 1.5: the Lachlan Fold Belt, the New England Fold Belt, the Late Cretaceous to Tertiary volcanics and the unconsolidated sediments in upland tributaries of Tertiary and Quaternary age.

The Lachlan Fold Belt (LFB) and the New England Fold Belt (NEFB) are composite orogenic belts consisting of: (a) meta-sediments of marine origin, which dominate the area; (b) granites, which occupy approximately one fifth of the area; and (c) minor extent of meta-volcanics. The Lachlan Fold Belt is of Mid-Cambrian to Early Carboniferous age, while the New England Fold Belt formation lasted from Cambrian to Triassic. They were formed during subduction of the Pacific oceanic plate under Antarctic continental plate, with the subduction margin gradually moving to the east of what are today Australia and Antarctica. The compression has resulted in strongly folded rocks, most of which became slightly metamorphosed. Meta-sediments comprise mainly of low-grade flysch sediments of deep marine environment such as siltstone, shale, sandstone and some greywacke, while shallow sea limestone rocks are rare. Slate, formed by higher grade metamorphism is associated with areas that were subjected to intense pressure and deformation, such as in proximity of regional faults, or subjected to heat, originating from molten lava. Much of the volcanic rock extruded during the subduction process has been eroded, so its extent within the fold belts is minor. As granitic rocks weather slower than meta-sedimentary, they usually occupy the catchment uplands and form catchment divides. The exposed granitic batholiths reflect their origin in composition. Low temperature S-type granites of supracrustal origin are lighter in colour than I-type low- and high-temperature infracrustal granites [Chappell and White, 2001]. Eroded material was transported to the west and accumulated within the Great Artesian Basin (GAB) during the Mesozoic and Cainozoic. The weight of sediments caused subduction of the rocks under GAB, while erosion of the eastern highlands resulted

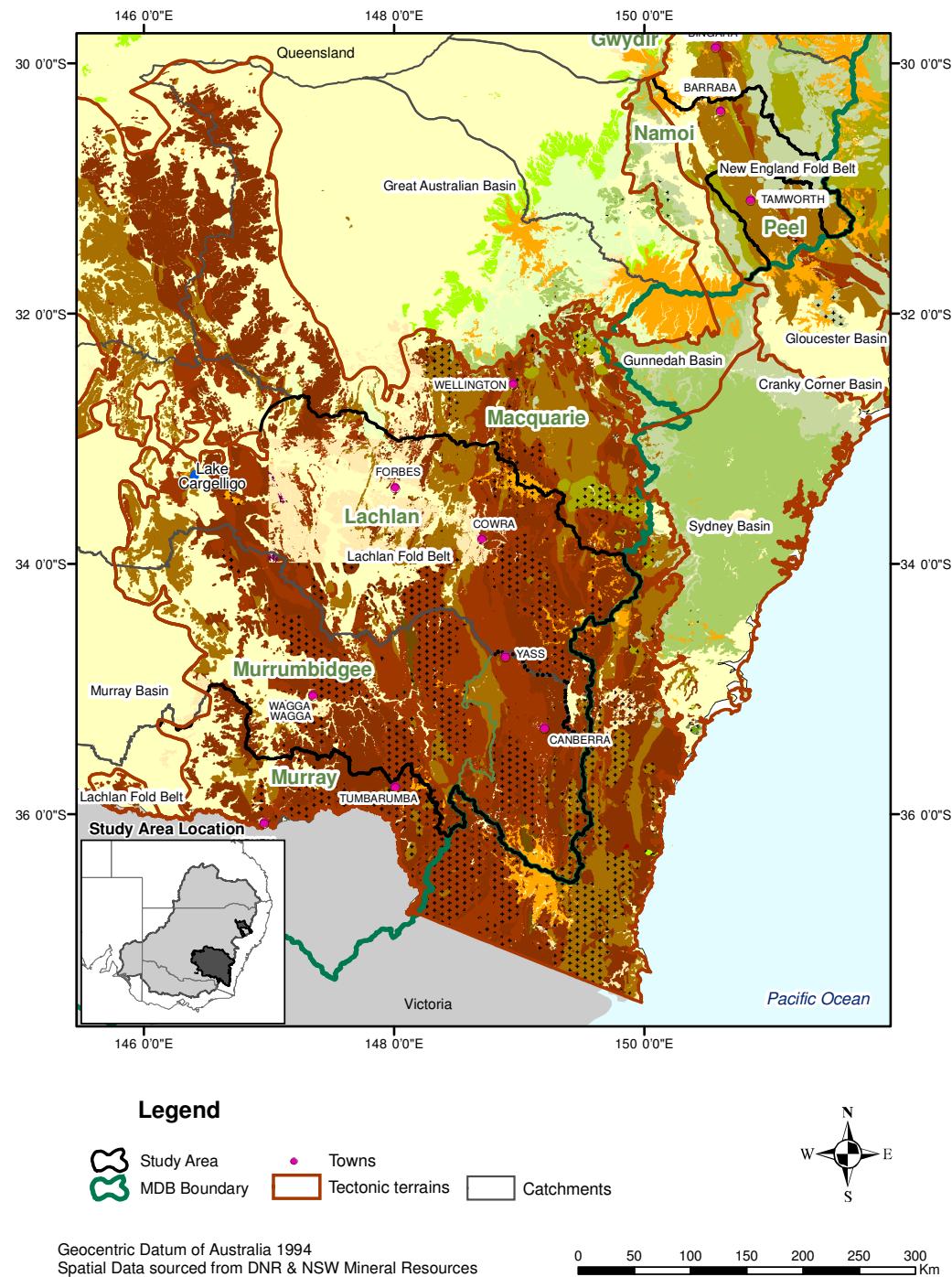


Figure 1.5: Geology of eastern NSW.

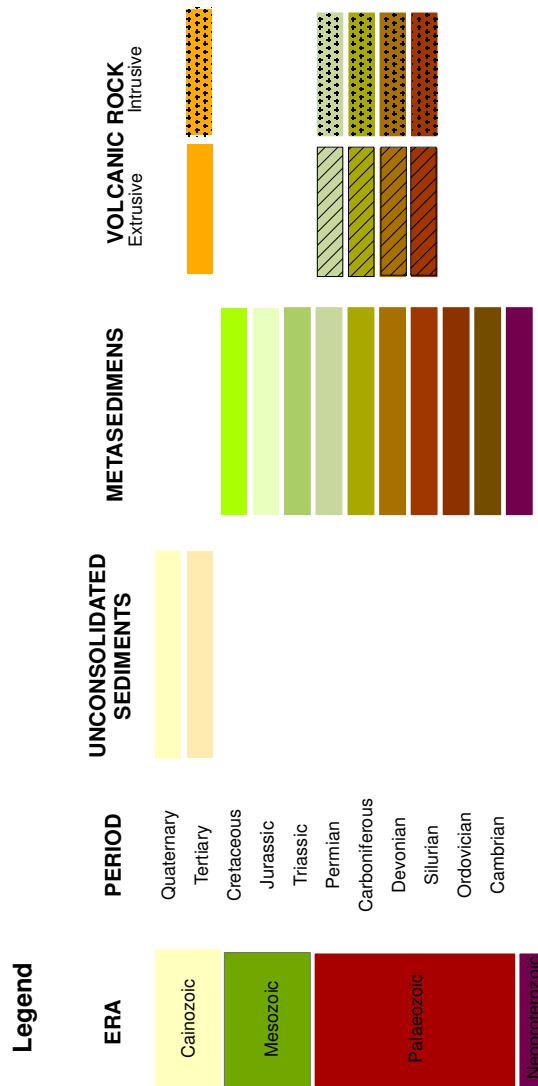


Figure 1.6: Legend for Figure 1.5 - continuation

in their further uplift, which together produced the tilting of the fold belt surfaces and dipping to the west. The break-up of Gondwana during the Mesozoic resulted in two rift zones followed by sea-floor spreading around Australia: opening of the Tasman sea to the east and separation of Australia from Antarctica to the south. The first event lasted between 90 and 64 Ma, subjecting a large part of the eastern Australian margin (from central New South Wales to the Marion Plateau) to strike-slip motion [Gaina et al., 1998]. This extension, and strike-slip deformation have resulted in a strongly faulted north-north-west structural grain. Bores associated with this geological province were the only one selected for this study.

The Late Cretaceous to Tertiary volcanics are formed by widespread volcanism associated with mantle plume activity, which produced mainly basaltic, volcanic complexes, thick volcanic piles, lava fields, minor sills, dykes and laccoliths. No bores associated with this geological province were selected for this study.

Unconsolidated sediments in upland tributaries are of Tertiary and Quaternary age. No bores associated with this geological province were selected for this study.

The study area encompasses a large region where many fracturing styles exist. The most prominent are near vertical faults, often trending in NNW, N, or NNE directions such as, among many others, the Humula Fault in Tarcutta River catchmentthe, the Lake George fault, (Lachlan Fold belt) which truncated the Yass River in the Murrumbidgee catchment and caused formation of the Lake George terminal basin and the Peel regional fault in the Namoi catchment (New England Fault belt). Fractures that the bores in the study area encounter are usually near horizontal. They originated from unloading due to weathering.

#### **1.5.4 Hydrogeology**

Three broad groundwater systems occur within the study area: aquifers in fractured metamorphosed fold belt terrains, in fractured late-Cretaceous–Tertiary volcanics and in unconsolidated Cainozoic sediments.

##### **Aquifers in the fold belt terrains**

The fold belt terrains consist of Palaeozoic to Triassic metamorphosed sedimentary, volcanic and intrusive fractured rock. The hydrogeological characteristics of the terrain most pertinent to this study are greatly reduced primary porosity and per-

meability. Structural deformation, igneous intrusion, weathering and solution have created secondary fracture porosity and permeability. At the regional scale, bulk permeability is generally very low, but at the local scale it can be extremely variable depending on fracture frequency, aperture, asperity, connectivity and degree of mineral infill or dissolution. Saturated fold belt systems are generally in good hydraulic connection unless very poorly fractured. Large permeable fractures may act as primary conduits between deeper and shallower aquifers.

The effects of weathering are usually limited to depths of 100 m. Below these depths, fractures tend to be closed by the weight of the overlying rock. Therefore, the permeability of these aquifers generally decreases with depth [Price, 1985]. However, water-yielding zones have been found even at depths of several hundred meters. The existence of production bores deeper than a hundred meters, and even more than two hundred meters in the study area confirms that the deeper water-yielding zones exist. However, these bores are rare.

The weathering thickness is variable and generally increases down the catchment. For example, metasedimentary bedrock often protrudes on the ridges. On these ridges, the highly permeable and thin horizon, with a stony content forms above the fractured bedrock. Upper slopes, adjacent to the ridges, are often characterised by a thicker weathered material with increase in loamy/clayey content with depth and corresponding reduction in vertical hydraulic conductivity. If this change in the particle size is abrupt, it can inhibit vertical infiltration and direct laterally a portion of infiltrated water [Rassam and Littleboy, 2003]. The depth of the regolith further increases in the mid-slopes. Vertical saturated hydraulic conductivity ( $K_{sat}$ ) decreases with depth due to the increase in clay content and compaction. Its range along the profile can, in places, exceed even three orders of magnitude. For example,  $K_{sat}$  values obtained by well piezometer measurements for recharge modelling at experimental paired catchments at Wagga Wagga ranged from over 100 cm/day at the surface to around 1 cm/day or even less at 1.3-1.35 m depth, in soil pits 1e and 2c (red chromosol) and 2d (sodosol) [Johnston et al., 2003].

Within the alluvial valleys the regolith is the deepest and alluvial depth can reach tens of meters in the creeks and more than hundred meters around the lowland rivers.

Quartz, as a product of granitic weathering, contributes to the formation of more permeable soils, with significant sand content. However, westerly winds, especially prominent during the ice ages brought fine dust and salt (Parna) from the salt pans

of the arid Australian interior characterised by the terminal drainage, and deposited them on the western flanks of the Great Dividing Range [Bowler, 1976, McTainsh, 1989, Melis and Acworth, 2001, Smithson, 2002, Summerell, 2004]. So even the regolith formed on the granites has the fine aeolian component mixed with the parent material, which reduces the soil permeability compared to the pure granitic soils east of the divide.

On the regional scale, the regolith thickness increases westward, in the lower, flatter parts of the study area, where partially eroded peneplane dips and is covered by the thicker unconsolidated sediments.

The level of saturation at the surface of the regolith is highly variable: from fully saturated, after some rainfall events, containing 30-40% water, to very dry, containing only 4-6% of water after the prolonged warm and dry conditions. These extremes are more pronounced in the southern parts of the study area, in the winter rainfall zone. This is because summer drought coincides in the south with the high evapotranspiration produced by the intense solar input. The variability in soil moisture diminishes with the depth of soil profile. While capable of being fully saturated, the lower margin of water content at depth is generally higher, due to the decreasing root density which limits water evacuation. Even though hydraulic conductivity of the regolith has strong tendency to decrease with depth, which can act as impervious barrier and as impediment to recharge, the spatial variability is very high and places exist where recharge occurs through these deeper regolith layers. Due to this variability, recharge tends to favour preferential pathways, such as through paleo-channels [Summerell, 2004] and/or through cavities formed by roots from withered vegetation, burrowing animals and through fractures. The hydraulic conductivity of the regolith is also characterised by seasonal and temporal variability, especially if the regolith contains cracking soil, which dries up in summer, and is capable of forming sufficiently wide and deep fractures, which allow water to penetrate down to the bedrock, until the fractures close up [Greve et al., 2001].

Owing to their much lower porosity, fractured-rock aquifers have a much smaller storativity than sedimentary aquifers, such as the Great Artesian Basin or the Murray Basin. Thus, they respond faster and with greater magnitude to changes in the water balance than large sedimentary basins do.

The bores used in this study are located in fold belt terrains, where they are rarely influenced by irrigation and other human activities.

### Aquifers in the Cainozoic volcanics

The late-Cretaceous to Tertiary volcanics within the study area are generally well fractured and moderately permeable owing to the formation of shrinkage cracks during lava cooling. These systems are characterised by relatively fresh groundwater, high fracture permeability (orders of magnitude greater than fold belt fractured rocks) and moderate storage capacity. Cainozoic volcanic fractured rocks often form productive groundwater supply systems, and in NSW many are managed as Groundwater Management Areas (GMAs). However, these systems were excluded from the present study because of their vastly different hydrogeological characteristics from those of the older, more widespread fold belt fractured-rock terrains and because of the potential impact of abstractions.

### Aquifers in the unconsolidated Cainozoic sediments

The hydraulic and chemical characteristics of the unconsolidated sedimentary aquifers are highly variable, depending on the environment of deposition and their location in the landscape. Some thick riverine plain sequences composed of sandy gravels have very high primary porosity, permeability and yield characteristics. Dryland salinity is more closely associated with the thinner, recent and more variable upland valley-fill sediments that are inhomogeneous and often have much lower permeability due to increased clay content.

In valleys and riverine plains, the fold belt and Tertiary volcanic fractured-rock systems may be overlaid by unconsolidated sediments of variable thicknesses. At these locations, the unconsolidated sediments generally control shallow groundwater characteristics and responses. Tertiary basalts may also overlie older, unconsolidated sediments.

#### **1.5.5 History of land-use change in the study area**

The history of tree clearing and land-use in the study area is presented because they have had impact on the depth to groundwater, through increased recharge, due to the replacement of deep rooted woodlands with pastures and cropping.

After the construction of a road over the Blue Mountains, colonisation of inland NSW began in Bathurst in 1815 and spread from there north, west and south [Salas and Smithson, 2002]. The naturally occurring woody vegetation was progressively

cleared to allow agriculture to develop, as well as for other purposes [Rančić et al., 2009]. As the distribution of native vegetation and the extent and density of tree cover that existed before the European settlement had not been recorded systematically, several independent attempts to reconstruct native vegetation cover resulted in a wide range of estimates [Brett, 1993, Graetz et al., 1995, AUSLIG, 1997, Geoscience Australia, 2009], and is still a topic of much debate [Butzer and Helgren, 2005]. However, Walker et al. [1993] estimated that 60 percent of trees have been cleared in the Murray-Darling Basin.

By 1860, squatter-pastoralists had established a monopoly in land-holding in the colony [Docker, 2005]. The period between 1860 and 1900 was characterised with the most intense, large-scale clearing [historian, Kathryn Wells, personal communication, 4th December 2004], due to:

1. Crown Land Acts of 1861 [Legislative Assembly of NSW, 1861] and 1875, which encouraged land development through the influx of large number of small farmers, and imposed major clearing within three years of occupation, or loss of land;
2. Introduction of wire fencing, which removed the need for shepherds [Harrison, 2004, Docker, 2005], but encouraged clearing to maximise the area under pasture;
3. Ringbarking, mostly using inexpensive Chinese labour, which efficiently removed the trees and opened the areas for grazing [Docker, 2005]; and
4. Influx of former gold-diggers from Victorian fields to NSW who further fuelled population expansion [Docker, 2005].

Approximately 90 percent of the tree clearing in the slopes west of the Great Dividing Range was finished by or at the beginning of the twentieth century [Kathryn Wells, personal communication, 4th December 2004]. This claim agrees with research conducted for:

- Mumbil catchment, in Macquarie catchment, [Anderson, 1992], where most of the clearing was completed before 1912;
- Wattle Retreat catchment in Lachlan catchment [Bradd, 1988], where clearing was carried out between 1874 and 1899;

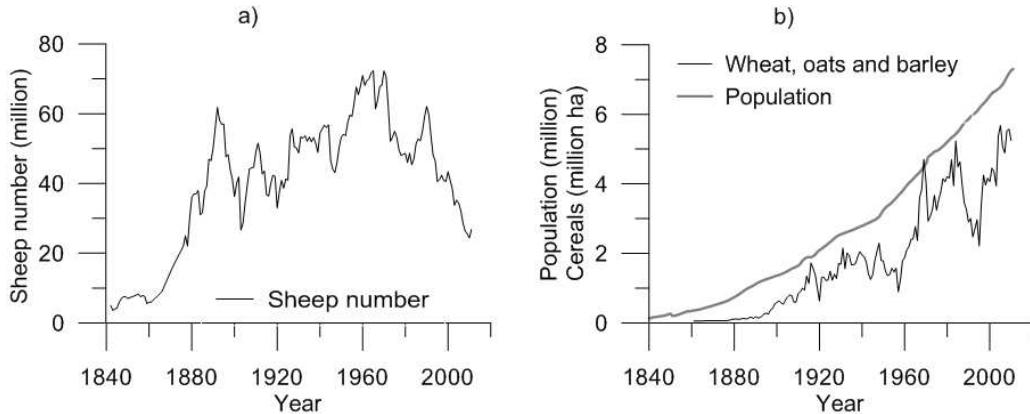


Figure 1.7: Historic growth of a) sheep numbers in NSW and b) NSW population and area under cereal crops.

*Data compiled from [Wadham and Wood, 1950] and Australian Bureau of Statistics (ABS): Historical selected Agriculture Commodities, by State (1861 to Present) (7124.0), Agricultural Commodities, Australia, 2010-11 (7121.0), Australian Historical Population Statistics, 2008 (3105.0.65.001), Australian Demographic Statistics, Jun 2011 (3101.0) and State and Regional Indicators, Dec 2010 (1338.1).*

- Tarcutta Creek catchment, in Murrumbidgee catchment [Rančić et al., 2011], where clearing commenced in 1830s and most of the catchment was ring-barked between 1860s and 1890s; and
- Yass catchment, in Murrumbidgee catchment, where according to Wagner [2001], pg 44:

*“...the significant dryland salinity started developing in 1940s, 50-60 years after the clearing.”*

Progress of pastoralist settlement is correlated and vividly reflected in sheep numbers (Figure 1.7a), which increased in NSW from 5 to over 60 million between 1842 and 1892. Sheep numbers declined from 1892, initiated by soil degradation and followed by the dryer spells of the first half of the twentieth century, which started with the Federation drought (1895-1903). Unfavorable soil and climate conditions reduced pasture production, preventing any further development and halting stock-related land clearing. Sheep numbers exceeded the 1892 record only between 1956 and 1972, during an extremely wet period, favorable for pasture production, which temporarily increased stock carrying capacity (Figure 1.7a).

Grain production progressively took hold during the twentieth century (Figure

1.7b) to accommodate increasing population. Cropping initially intensified in the southern parts of the sheep-wheat belt, then spread north and west [Bedward et al., 2007]. Bedward et al. [2007] demonstrated that cropping accelerated in the early 1900s before becoming more stable between 1920 and 1960, when it then doubled (Figure 1.7b).

Even though stock-related clearing was over after the 1892 crash, the cropping intensification in the wheatbelt area between 1900 and 1920 and associated tree clearing were significant in the south-western flatter portion of the study area of the Lachlan and Murrubidgee catchments.

## 1.6 Structure of the Thesis

Following the Introduction in Chapter 1, data mining techniques are reviewed in Chapter 2 of the thesis, together with papers that deal with other methods of using point data to gain knowledge of a system behaviour.

The methods used for this study is presented in Chapter 3, starting with the methods used to select representative bores, followed by the methods used for aggregation of point-data into data series and ending by the methods for reconstruction of aquifer behaviour: plotting of trends and residual mass curves and change-point analysis. The method is illustrated using data from the Namoi catchment. This approach is taken because the Namoi has the most comprehensive dataset available. Using this comprehensive data set, it is demonstrated that it is possible to reconstitute groundwater level trends back to the beginning of the twentieth century.

Chapter 4 summarises the results of the study.

Chapter 5 is concerned with validation. To validate the method, it is necessary to make use of the limited groundwater hydrograph data that is available. Direct validation is carried out using historic hydrographs of limited duration. Unfortunately, the Namoi catchment does not have groundwater hydrograph data. For this reason it was necessary to turn to other catchments in NSW fractured-rock terrains to provide the necessary validation. Those with the longest available hydrograph data from the study area (Figure 1.2) were chosen :

- Wattle Retreat in the Lachlan Catchment,
- Dicks Creek and Williams Creek catchments in the eastern section of Mur-

- rumbidgee and situated in the uniform rainfall rainfall zone and
- the Tarcutta Creek catchment from the western portion of Murrumbidgee catchment, situated mainly in the winter rainfall zone.

In addition to direct validation that spans only a limited duration, the model is also validated over a longer time-span indirectly. This is achieved by use of the long-term rainfall data that are widely available. The model is validated by comparison of derived groundwater trends, residual mass curves and change points with the respective rainfall trends, residual mass curves and change points using data from Namoi, Lachlan and Murrumbidgee River catchments.

Chapter 6 discusses the domain of the model applicability, user-friendliness, robustness and the target group. It also discusses the wealth of knowledge gained by its application to South-East Australian fractured rock aquifers.

Chapter 7 summarises the conclusions and provides recommendations for further research.

# Chapter 2

## Literature Review

This chapter provides an overview of literature relevant to data mining techniques, their application in hydrology and hydrogeology and methods for reconstructing aquifer behaviour from point data.

### 2.1 Data mining

According to Fayyad et al. [1996], knowledge discovery from databases (KDD) refers to the overall process of discovering useful knowledge from data and making sense of data. Data mining represents a step in this process: the application of specific algorithms for extracting patterns from data [Fayyad et al., 1996]. The KDD process involves data:

- (a) selection,
- (b) preprocessing,
- (c) transformation,
- (d) mining to extract patterns and relationships, and
- (e) interpretation and assessment of the discovered structures [Hand et al., 2001].

Roots of data mining date back to 1970s. Lovell [1980] published the first paper with the title Data Mining, referring to the related work in the 1970s, which describes new techniques for analysis of data extracted from growing databases.

Replacement of desk calculators by electronic computers and development of electronic data repositories allowed extraction of useful information from these data [Lovell, 1980]. Increased computer power and data storage availability fueled increases in digital data collection and exponential growth of databases [Gopalan and Sivaselvan, 2009]. Increased computer speed permitted the extraction of useful information from ever increasing bodies of data. Improvement of data mining algorithms further contributed to advances in data mining and KDD techniques [Hand et al., 2001].

Data mining is an interdisciplinary exercise. It emerged from statistics, machine learning, artificial intelligence (AI), data management, pattern recognition and visualisation techniques and therefore encompasses concepts from all these disciplines. Statistics uses data collected for the specific analysis. Whereas data mining uses data that were previously collected for some other purpose. Data sets examined in data mining are often large. When only small data sets are involved, the problem is reduced to classical exploratory data analysis as practiced by statisticians. Databases often contain some missing values and the measurement noise can represent an additional problem [Hand et al., 2001].

A data set of  $n$  objects, for which the same  $p$  measurements are taken, can be represented by the  $n \times p$  matrix. The  $n$  rows represent the  $n$  objects and may be referred to as individuals, entities, cases or records. The  $p$  columns represent  $p$  measurements and may be referred to as variables, features, attributes or fields. Variables can be quantitative and categorical. Quantitative variables can be measured in numerical scale, while categorical take only certain values. Categorical variables may be ordinal, if they can be placed in order, or nominal, if the categories are defined by name. The  $n \times p$  data matrix can be referred to as the data set, training data, sample or database [Hand et al., 2001].

### **2.1.1 Data mining tasks**

Data mining tasks can be categorised into types of tasks, corresponding to the objectives of the analysis [Hand et al., 2001]:

- (a) exploratory data analysis,
- (b) descriptive modelling,

- (c) predictive modelling,
- (d) discovering patterns and rules,
- (e) retrieval by content, etc.

Exploratory data analysis (EDA) is a typically interactive and visual way of exploring the data, through bar and column charts, scatter plots, pie charts, radar graphs (coxoncomb plots) etc. While graphical display is suitable for small dimensional datasets ( $p = 1, 2, 3$ ), for  $p$  higher than 3 or 4, projection techniques, such as principal component analysis, are applied to produce informative low-dimensional projection of data [Hand et al., 2001].

Modelling can be *descriptive* or *predictive*. The purpose of *descriptive* modelling is to describe all of the data or the process generating the data [Hand et al., 2001]. Associations are analysed across the data set with no distinction between independent and dependent variables [Wedel and Kamakura, 1999]. Some examples of *descriptive* modelling are:

- (a) probability distribution modelling, and in particular, density estimation;
- (b) partitioning of the  $p$ -dimensional space into groups: segmentation, and cluster analysis (forming of homogeneous groups based on patterns and data); and
- (c) dependency modelling (describing the relationship between variables) [Hand et al., 2001].

Spate et al. [2002] point out that there is a trade-off between the amount of information stored in the clustering regime and understandability or usefulness of that information: enough clusters are required so that internal objects are close, but not so many that there are similar clusters.

The purpose of *predictive* modelling is to predict the value of a variable. For example, classification is used for modelling of categorical *predictive* variables and regression for modelling of quantitative variables [Hand et al., 2001]. Predictive methods analyse associations between two sets of variables: a set of dependent variables, which need to be explained or predicted and a set of independent variables [Wedel and Kamakura, 1999].

A *segmentation* approach is called *a priori* when the type and number of segments are defined in advance by the researcher. Examples of *a priori, descriptive*

*segmentation* analysis are contingency tables and log-linear models, and of *a priori*, *predictive segmentation* methods are cross-tabulation, regression, logit and discriminant analysis. A segmentation approach is called *post-hoc* when the type and number of segments are determined on the basis of data analysis [Wedel and Kamakura, 1999]. Cluster analysis belongs to the *descriptive, post-hoc* segmentation approach [Wedel and Kamakura, 1999] and its aim is to discover natural groups in data [Hand et al., 2001]. Other *post-hoc, predictive* methods are AID, CART, Clusterwise regression, Artificial Neural Networks (ANN) and mixture models [Wedel and Kamakura, 1999].

*Discovering patterns and rules* is a task concerned with data points that significantly differ from the rest (outliers) or combination of items that frequently or rarely occur together. This problem has been addressed using algorithmic techniques based on association rules [Hand et al., 2001]

*Retrieval by content* is a task in which the user has a pattern of interest and wishes to find similar patterns in the text or image dataset. Examples of this type of tasks are web search for the text data sets and fingertip identification for the image data set [Hand et al., 2001].

Although each task is different, they share many common components, such as similarity or distance between any two data vectors and score functions used to assess how well a model or pattern fits the data [Hand et al., 2001].

### 2.1.2 Components of data mining algorithms

According to Hand et al. [2001], the process of seeking relationships within a data set involves the choice of:

- (a) the model or the pattern structure representation,
- (b) the score function, to evaluate data representations,
- (c) the optimisation and search method algorithms and
- (d) the data management strategy:
  - *Structures* represent the general functional forms of the models or patterns, with unspecified parameter values. A fitted model or pattern has specific values for its parameters. Model structures are a global summary

of the dataset, valid at any point in the full measurement space. Pattern structures are restricted to certain regions of the space.

- *Score functions* evaluate how well a model or parameter structure fits a given data set. Several score functions, such as likelihood, sum of squared errors, and misclassification rate are widely used for this purpose.
- The goal of *optimization* and search is to determine the structure and the parameter values that achieve the best fit by finding a minimum or maximum value of the score function. Optimisation techniques deal with estimation of model parameters. Combinatorial search techniques, often accomplished using heuristic search techniques, deal with finding interesting patterns (such as rules) from a large family of potential patterns.
- *Data management strategy* refers to the ways in which the data are stored, indexed, and accessed. In case of large datasets, data might have to be accessed not only from RAM memory but also from secondary (disk) or tertiary (tape) storage, so the strategies for efficient data access are necessary.

### 2.1.3 Data mining in hydrology

Spate et al. [2002] demonstrate the relevance of data mining techniques and tools to the science of rainfall-runoff hydrological modelling. They cover tools available for extracting trends, characteristics or rules from hydrologic data, such as:

1. *Clustering*, which could be used for grouping hydrograph peaks according to shape using the k-means algorithm or the k-medoids algorithm. The k-medoids algorithm should be used preferentially to k-means when the data contain outliers, because an extreme value would not skew the medoid as much as the mean value.
2. *Classification*, which could be used for detection of the high intensity/short duration events from the low intensity events:
  - (a) Bayesian methods (decision trees used to apply markers, called classifiers such as J48 algorithm in the Weka data mining package and the "divide-and-conquer method".

- (b) Neural networks and
  - (c) Genetic algorithms.
3. *Association Rule Finding* ( $A \Rightarrow^T B$ : where A and B are events with the rule *if A occurs then B occurs within time T*. A could be a rainfall event type and B a shape of flow peak) and
4. *Dominant Mode Analysis / Series Similarity Measures*, which approximate a series by decomposition with basis functions and evaluate agreement of two time series via agreement of their power spectrum [Ehret and Zehe, 2011]:
- (a) Fourier analysis, for separation of the time-series of streamflow into a set of frequencies with amplitudes [Spate et al., 2002] and
  - (b) Wavelet transformations, to regionalise catchments by taking the wavelet spectra of streamflow records as a signature of catchment response characteristics and clustering on these [Spate et al., 2002].

The *Artificial Neural Networks* (classification technique) and the *Fourier Analysis* (dominant mode analysis technique) have the the most frequent application in hydrologic studies, with *Wavelet Transformations* recently gaining popularity.

Maier and Dandy [2000] present a comprehensive review of 43 papers published by 1998 in which neural networks were used for the prediction and forecasting of water resources variables: Algal concentration (3), Cyanobacteria concentration (3), Flow (18), Rainfall (13), salinity (3), pH (1), water table level (2) and water level (1). Feedforward networks were used in 41 out of 43 papers. The majority of the networks were trained using the backpropagation algorithm.

Liang and Liang [2001] used a neural network algorithm with the multi-resolution learning paradigm (NNMLP) to investigate the potential applicability and limitations of the streamflow forecasting approach. The predicted long-term range streamflows using the NNMLP requires only historical streamflow information. The model is applicable in regions with limited available information, and for being combined with other approaches to improve long-term range streamflow forecasts.

Hewett [2003] applied an induction technique called the second-order table compression learning system (SORCER), to generate predictive models of future water inflows of Lake Okeechobee, in south Florida, by creating a predictive relationship

which integrates solar and ocean-atmospheric conditions into forecasts of regional water flow. The SORCER method uses a process that transforms a table consisting of training data into a second order table (which has sets of atomic values as entries) with fewer rows, by merging rows in consistency preserving ways. The model performed favourably when tested against neural networks, decision tree learning and associational rule mining techniques.

Montanari and Toth [2007] applied Fourier analysis, using the Whittle estimator [Whittle, 1953] in the rainfall-runoff modelling; Schaeifi and Zehe [2009] proposed a method for rainfall-runoff model calibration and performance analysis in the wavelet-domain by fitting the estimated wavelet-power spectrum.

Recent years saw the explosion in the ANN hydrological applications, such as in:

- streamflow discharge, flood and runoff forecast [Bhatia et al., 2013, Pan et al., 2013, Ghalkhani et al., 2013, Lalozaee et al., 2013, Awchi, 2014, Prakash et al., 2014],
- sediment yield, sediment concentration [Singh et al., 2012, Cao et al., 2013], and
- water quality forecasting [Wang et al., 2013, Banejad et al., 2013].

#### **2.1.4 Data mining in hydrogeology**

Neural networks are the most frequent data-mining technique used in hydrogeology. It is often applied to analyse complex hydrochemical datasets and for short-term prediction of groundwater levels.

The unsupervised artificial neural network and self-organizing map (SOM) algorithm in hydrogeological research was used by: Hong and Rosen [2001] to characterise and diagnose the groundwater quality in a fractured-rock aquifer; Sanchez-Martos et al. [2002] to classify a hydrochemical data set, from a detritic aquifer in a semi-arid region, into classes of different chemical composition; and Lischeid [2003] to investigate spatial and temporal trends in water quality. Peeters et al. [2007] proposed a variant of Kohonen's self-organising map, the GEO3DSOM, capable of explicitly incorporating three-dimensional spatial knowledge into the algorithm. They used it to perform exploratory data analysis and clustering of multivariate

spatial hydrogeological data. GEO3DSOM evolved from GEO-SOM [Bacao et al., 2005].

Applications of data-mining techniques in predicting water table fluctuations are confined to the Artificial Neural Network (ANN) technique. This research originated in Canada in the mid 1990s. Yang et al. [1996] developed a model, based on the ANN technique for simulation of water table fluctuations under subsurface drainage system and applied it to the corn field in Quebec, Canada. Backpropagation neural network was selected to deal with non-linear problems. Daily rainfall and evapotranspiration data from 1960 to 1983 and from 1991-1992 were available and used to estimate daily water table (WTD) data from 01 May until 31 October, producing 4784 daily records, using DRAINMOD [Skaggs, 1980]. The first 13 years of rainfall, ET and WDT data were used to train the ANN model and the other half to test it. Model predictions were done for each of the 13 years separately, using initial WDT, rainfall and ET data and achieved a Root Mean Square (RMS) error of 0.12.

Shukla et al. [1996] investigated the use of ANNs in transient drainage design to predict water-table fluctuations, as an alternative to use of the Boussinesq equation. The numerical model based on a Boussinesq equation was used to produce 52,280 solutions. Half of these were used to train 15 fully connected, feed-forward ANNs with slightly different architectures. The best performing model was further subjected to 21 training regimes. The best ANN had mean relative errors of -2.04% and 0.50% for the 12 and 24-h delay. For both versions, the absolute mean of the differences between the training and recall outputs was less than 9 cm. The ANN model generated solutions for the 52,280 cases in 140 seconds, while the numerical model based on Boussinesq equation required circa 100 hours.

Application of ANNs for short-term forecasts for up to 18 months, but mostly for only 1-3 months, gained exponential momentum during the past decade. [Coppola et al., 2003, Daliakopoulos et al., 2005, Michael et al., 2005, Nayak et al., 2006, Asefa et al., 2007, Feng et al., 2007, Uddameri, 2007, Krishna et al., 2008, Nourani et al., 2008, Chu and Chang, 2009, Dash et al., 2010, Ghose et al., 2010, Mohanty et al., 2010, Chen et al., 2011, Jalalkamali and Jalalkamali, 2011, Jalalkamali et al., 2011, Nourani et al., 2011, Sreekanth et al., 2011, Yoon et al., 2011, Zhu et al., 2011, Li et al., 2012, Rakhshandehroo et al., 2012, Taormina et al., 2012, Sahoo and Jha, 2013, Moosavi et al., 2013a,b, Shiri et al., 2013, Shirmohammadi et al., 2013, Chitsazan et al., 2013, Chattopadhyay and Rangarajan, 2014, Jha and Sahoo, 2014]

and others. The popularity arose because of:

- (a) the need for accurate short-term forecasts in groundwater management, planning and well operation.
- (b) the ANN ability to handle highly non-linear relationships,
- (c) the superiority of ANN in accuracy over physically-based groundwater models for short-term predictions, and
- (d) the modest data requirements: short length of training data and no need for aquifer parametrisation.

Some other applications of ANN in hydrogeology include use of ANN models for: determination of aquifer parameters [Lin and Chen, 2006, Samani et al., 2007] and development of boundary conditions [Huo et al., 2011].

Most of the applications of data-mining techniques in predicting water table fluctuations have been restricted to short-term predictions. All applications use true timeseries with groundwater and other data sampled at equal time intervals. No application was found in the literature that deals with prediction of water-table fluctuations based on:

- (a) data sampled irregularly,
- (b) individual readings from independent bores and
- (c) agglomeration of water-table point data scattered in time and space.

The applications of ANN in water-table modeling are mostly concerned with the short-term forecasts for management purposes, rather than with the reconstruction of the past hydrographs based on the long term but patchy records. No data-mining application was found that recreates long-term (multidecadal) trends of water table fluctuations.

## **2.2 Methods for reconstructing aquifer behaviour from point data**

Pioneering work by Salas and Smithson [2002] in the Macquarie River catchment in NSW paved the way for reconstruction of aquifer behaviour from scattered point

data.

Salas and Smithson [2002] selected 255 bores drilled in fractured rock, far away from rivers and watercourses, that could temporarily influence the water table. Change the in water table,  $\Delta SWL$ , between (a) the value measured when the bore was constructed, available from the NSW groundwater database GDS, and (b) the value obtained through the 1988 survey [Salas and Garland, 1989], was calculated for each bore. Bores drilled during the same year were grouped together and the median value  $MED_y$  was calculated for year  $y$ . The median was chosen as a measure of central tendency, rather than the mean, because the number of points in each year was small. The resulting series  $MED_y$  was used to construct the residual mass curve (Section 3.2.3). The constructed residual mass curve was plotted against the residual mass curve of rainfall. The authors concluded that there was a close association between the shapes of the two graphs, which suggested a lag of 4 to 13 years in the response of the groundwater system to changes in the rainfall pattern.

The method that Salas and Smithson [2002] employed to reconstruct aquifer behaviour from point data followed the principles used during reconstruction of surfaces from scattered point data, reviewed by Schall and Samozino [2005]. Instead of approximating scattered data with the three-dimensional geometric function, Salas and Smithson [2002] operate within the two-dimensional space, in which one dimension represents time and the other, the position of the water table relative to the 1988 level. Therefore, the problem defaults to the reconstruction of a time-series function from incomplete scattered point data. The time-series  $MED_y$  represents a very noisy and incomplete dataset, due to the small annual sample size. Schall and Samozino [2005] demonstrated that surface reconstruction can be drastically improved if the appropriate preliminary filtering method was applied to the original scattered data. To overcome the noise and introduce the smoothing effect, Salas and Smithson [2002] used residual mass curves as the filtering technique, reconstructing in fact, the first derivative of the original  $MED_y$  function.

The method of Salas and Smithson [2002] touches the boarder-line of the data-mining methods, as:

- (a) one of the points used to generate  $\Delta SWL$  was drawn from the database and was collected for a purpose which was unrelated to their analysis;
- (b) database information was used for discovering useful knowledge from data;

and

- (c) this was implemented by applying a specific algorithm for extracting patterns from data [Fayyad et al., 1996], (Section 2.1).

However, the second point Salas and Smithson [2002] used was surveyed and the number of points used for the study was relatively small. Nevertheless, the methodology Salas and Smithson [2002] employed is highly significant as it introduces the idea of aquifer behaviour reconstruction from scattered point data which could be mined from groundwater databases.

## 2.3 Summary

Data mining represents the assembly of new techniques for discovering patterns and useful knowledge from data extracted from databases. The ANN technique is the most common application of the data mining techniques in hydrology and hydrogeology. Application of ANNs for modelling of water table fluctuations are restricted to *predictive* models with short-term forecasts for up to 18 months. ANNs require true time-series of recorded hydrographs and other input data sampled regularly with equal time intervals.

The method of Salas and Smithson [2002], based on preliminary filtering of scattered water table data, is highly significant as it opens the door to multi-decadal *descriptive* modeling by mining the scattered point data from available groundwater databases. It is this technique that is further developed in this thesis.

# Chapter 3

## Methods

This chapter provides a description of the data-mining technique used for reconstruction of aquifer behaviour from independent point data. The multi-decadal, *descriptive* modeling is carried out after the filtering of scattered water table data found in the groundwater database. Data were restricted to the fractured rock aquifers closest to the land-surface. The multi-decadal descriptive modeling technique is illustrated using the Namoi catchment dataset as an example.

### 3.1 Data sources for data mining

The primary data used in this study were standing water level (SWL) data (Figure 3.1). SWL data were drawn from the NSW Groundwater Database System (GDS). They originated from drillers' records, which were collected during bore construction and pump installation. These data were not collected with the purpose of recreating historic water table fluctuations or trends.

Thus, tapping into the GDS and gathering useful information from it represents use of a data mining method. The GDS [DIPNR, 2004a] held data for a total of 15 467 bores across the study area (Table 4.1 and Table 4.3). As some bores had several water-bearing zones, each with SWL data, the database provided a total of 18 250 SWL data points (Table 4.3). In addition to their date of completion and the SWL on that date, the database also contained information relating to the rock types and other features of the site in the form of drillers' reports.

As noted in Chapter 2, the SWL records can be referred to as objects. Their associated attributes can be referred to as variables or fields and they can be presented

**STANDING WATER LEVEL (SWL)**  
Depth from the ground surface to water table

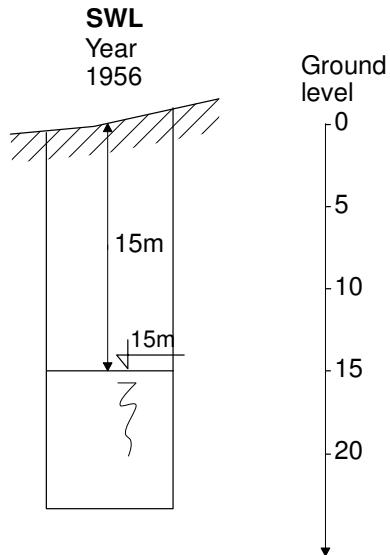


Figure 3.1: Standing water level.

by an  $n \times p$  matrix, referred to as the data set or sample [Hand et al., 2001]. The set of variables used in data-mining of the GDS consists of a combination of quantitative variables (depth of water-bearing zones, depth of each geologic horizon, longitude and latitude of the bore, year of construction), and qualitative, categorical, nominal variables (type of water-bearing zone: fractured, unconsolidated, unknown); type of horizons (sand, clay, soil, loam, rock, granite, sandstone, limestone, siltstone, mudstone, basalt, etc...)

Secondary datasets consisted of:

- spatial digital data - Geographic Information System (GIS), used for bore selection : (1) geology map (Figure 1.5); (2) Groundwater Management Area (GMA) map; and (3) topographic map (DEM, Figure 1.3) with catchment boundaries and drainage network;
- existing short-term groundwater hydrographs, used for model validation over the limited time-frame; and
- log-term rainfall monthly Patched Point Data (PPD), used as an additional

resource for model validation over the entire modeled time-frame.

Monthly PPD rainfall data (January 1880–July 2004) were obtained from the Bureau of Meteorology [NRME, 2004]. The PPD dataset is a semi-synthetic dataset that combines the original record with gaps in data in-filled by using a trivariate thin-plate smoothing spline Wahba and Wendelberger [1980], and latitude, longitude and elevation as independent variables. To overcome potential bias resulting from data filling, the rainfall stations with the most reliable data and the fewest missing data were selected for inclusion in the final dataset. This procedure identified 54 long-term stations having fewer than 5% missing records (Table 3.4).

Three additional rainfall stations with a shorter length of record (>50 years) and a low proportion of missing data were included in the dataset to improve the spatial coverage. Where rainfall stations were closely clustered, or when their data were inconsistent with the datasets of adjoining stations, stations were removed from the primary dataset. Methods employed to detect the inconsistencies in the rainfall time-series were: cross correlation of monthly time-series and visual inspection of: a) annual time series with 5-year, 7-year and 21-year moving average filters plotted in the central year and b) residual mass curves. The shorter, 5-year and 7-year moving average filters accentuated the effect of ENSO - Walker oscillation (east-west climatic sea-saw). The longer, 21-year moving average filter, filtered these shorter frequencies out and illustrated the much slower cycle, of near century frequency, associated with the movement of the high-pressure ridge in north-south direction.

For each subdivision of the major basins (Figure 1.2), data from the verified rainfall stations were used to calculate a single, area-weighted dataset. This procedure was implemented after the bore selection was finished. The weighting coefficients were calculated based on Thiessen Polygons: as the ratio between number of bores located closest to each contributing station, within the corresponding Thiessen Polygon, and the total number of bores in the subdivision.

## 3.2 Bore data mining method

The approach adopted in this thesis consists of developing a composite SWL time series from the SWL that was recorded when each bore was constructed. Annual

Table 3.1: Data mining algorithmic components.

<b>Components</b>	<b>Description</b>
TASKS	Classification; discovering patterns and rules to deal with outliers; segmentation; model building: trend, residual mass curve and change-point analysis.
STRUCTURE	Decision tree.
SCORE FUNCTION	Root Mean Square Error (RMSE).
SEARCH METHOD	Unspecified.
DATA MANAGE- MENT TECHNIQUE	Unspecified.

samples are derived for each year for each catchment or sub-section. The task is similar to that of time-series reconstruction from scattered point data.

The main difference between this approach and a similar approach used in the Macquarie Valley study [Salas and Smithson, 2002], described in Section 2.2, is that their method is based on the difference in SWL at one bore location between two points in time (date of construction and 1988;  $\Delta$  SWL), whereas the method described here uses only the SWL data recorded by the driller when the bore was constructed. As this approach does not require a second measurement, it is more time-efficient, and it potentially provides a larger statistical sample.

The data mining algorithmic components applied in this method are presented in Table 3.1.

The method consists of three phases:

1. Data selection;
2. Creation of composite SWL time-series; and
3. Noise filtration.

Data for the fractured bores had to be selected from the complete dataset. Bores in alluvial aquifers had to be eliminated from the study for several reasons: influence of streams on water table, groundwater extractions due to development, river regulation (dams) with altered flow and recharge regime, and difference in: aquifer

size, hydraulic properties, system response to hydrologic changes and time needed for the system to equilibrate after a hydrologic disturbance.

Cainozoic volcanic rocks are present in the Eastern Australia mainly as Tertiary basalts, formed while Australia was drifting northwards during the Cainozoic and passing over hot-spots. These rocks vastly differ in their hydraulic properties and response times from the Paleozoic fractured rocks and are subject to groundwater exploitation due to their high specific yield, thus they had to be eliminated.

The data selection procedure is designed to minimise the possibility of including non-fractured-rock bores, or bores whose SWL data may be influenced by streams. The GDS database contains information about bores in different aquifers, but it does not have a single attribute field that allows bores in fractured rock to be simply selected on that basis. However, this information can be logically deduced based on the position of the bore relative to mapped geology and a set of attributes from the GDS, such as type of aquifer geology the water bearing zone is situated in, its depth etc...

This involves classification of the bores into groups, an example of the supervised discriminative predictive modeling data-mining technique (Section 3.2.1). The value of one categorical variable (type of aquifer,  $c_j$ ) is predicted from the known values of other input variables  $x_i$ , ( $i = 1, p$ ) represented as the  $p$ -dimensional vector. The classification data-mining technique is used to derive a classification model  $f(x; \Theta)$ , by defining a domain over the  $p$ -dimensional space of the vector  $x_i$  that defines the decision boundary for class  $c_j$  to be predicted [Hand et al., 2001]. The set of logical rules can be presented as a decision tree. Some of the rules were derived using simple mathematical logic, and some were derived using algorithmic technique: discovering patterns and rules to eliminate points that significantly differ from others (outliers).

The GDS database includes the year of construction, the SWL, and the latitude and longitude, which allows individual bores to be allocated and the SWL analysed separately for each of the catchments in the study area (Figure 1.2). A bore selection method that involves two levels of classification is thus devised. It allows identification of fractured-rock bores on the basis of geographic information system (GIS) and the GDS database information.

Due to a small sample size, especially during the early years, the annual data were characterised by high variability. To eliminate the noise, introduced by small annual sample size and years with missing data, which are often encountered in the data

mining from databases [Hand et al., 2001], the annual time series were subjected to filtering techniques: moving average, splines and residual mass curves. These techniques represent low-pass filters, which smooth the data, attenuate short-term variability and accentuate underlying trends. As pointed out in Section 2.2, Schall and Samozino [2005] demonstrated that surface reconstruction can be drastically improved if the appropriate preliminary filtering method was applied to original scattered data.

In summary, the three phases of the adopted data-mining method involve:

1. Data selection:

- (a) Predictive modeling; **Spatial classification of bores** based on aquifers they were drilled in, using 2 digital maps and GIS software: (1) geological (G) and (2) GMA maps, into 3 *a priori* defined *Classes*: (I) Unconsolidated sediments, (II) Cainozoic volcanic rock – Volcanics, and (III) Palaeozoc fractured rock – Fractured Rock (Table 4.3);
- (b) Predicative modeling; Second level of **classification** to form *Groups* within the three previously defined *Classes* **based on the groundwater database attributes**. Seven groups are formed in total (Table 4.3);
- (c) Descriptive modeling; **Discovering** patterns and rules to eliminate points that significantly differ from others (outliers) and developing **selection rules based on the database attributes** that are applied on each *Group* of bores;
- (d) Predicative modeling; Supervised discriminative modeling; Applying the derived **rules for** each *Group*, with the purpose of **selecting the bores** that were in shallow fractured rock, based on designed algorithm (Table 4.3);
- (e) Predicative modeling; **Screening** and elimination of those fractured rock **bores** that may have been **influenced by alluvial aquifers**, based on Root Mean Square Error criteria (Table 4.3);
- (f) Agglomeration of all selected points into the **final dataset** (Table 4.3);

2. Creation of composite SWL time-series:

- (a) Descriptive modeling; Model building; **Choice of central tendency model for annual time-series** based on Root Mean Square Error criteria;
- (b) Descriptive modeling; Model building; Testing of **alternative filtering techniques**: choice of the length of moving average filter, and of degrees of freedom in splines;
- (c) Exploratory data analysis; **Exploratory analysis of bore data**
- (d) Descriptive modeling; **Post-hoc segmentation** to define geographic boundaries with optimum amount of data for trend and change-point analysis: sufficient amount of data to extract meaningful information but keeping the geographic domain compact enough to insure homogeneity and avoid confounding effects;

3. Noise filtration:

- (a) Residual mass curves with change-point analysis;
- (b) Historic SWL trends;

Data mining method is described in detail and illustrated using the Namoi catchment.

### 3.2.1 Data selection

#### Spatial classification of bores

Spatial classification of bores was performed by projecting the latitude and longitude of each bore onto the geological map (Figure 1.5), (G) and Groundwater Management Map (GMA). The predictive modeling was carried out using the attributes of GIS maps. Bores were placed into three *a priori* defined *Classes*: (I) (*Unconsolidated sediments*); (II) Cainozoic volcanic rock (*Volcanics*); and (III) Fractured rock of the New England Fold Belt and Lachlan Fold Belt (*Fractured Rock*). Bores and associated water bearing zones were placed in *Class* (III) when both maps indicated so. If one map indicated *Class* (III) and the other *Class* (I) (or *Class* (II)), the bores were placed in the *Class* (I) (or *Class* (II)):

$$\text{Fractured Rock} = \text{Fractured Rock}_G \cap \text{Fractured Rock}_{GMA} \quad (3.1)$$

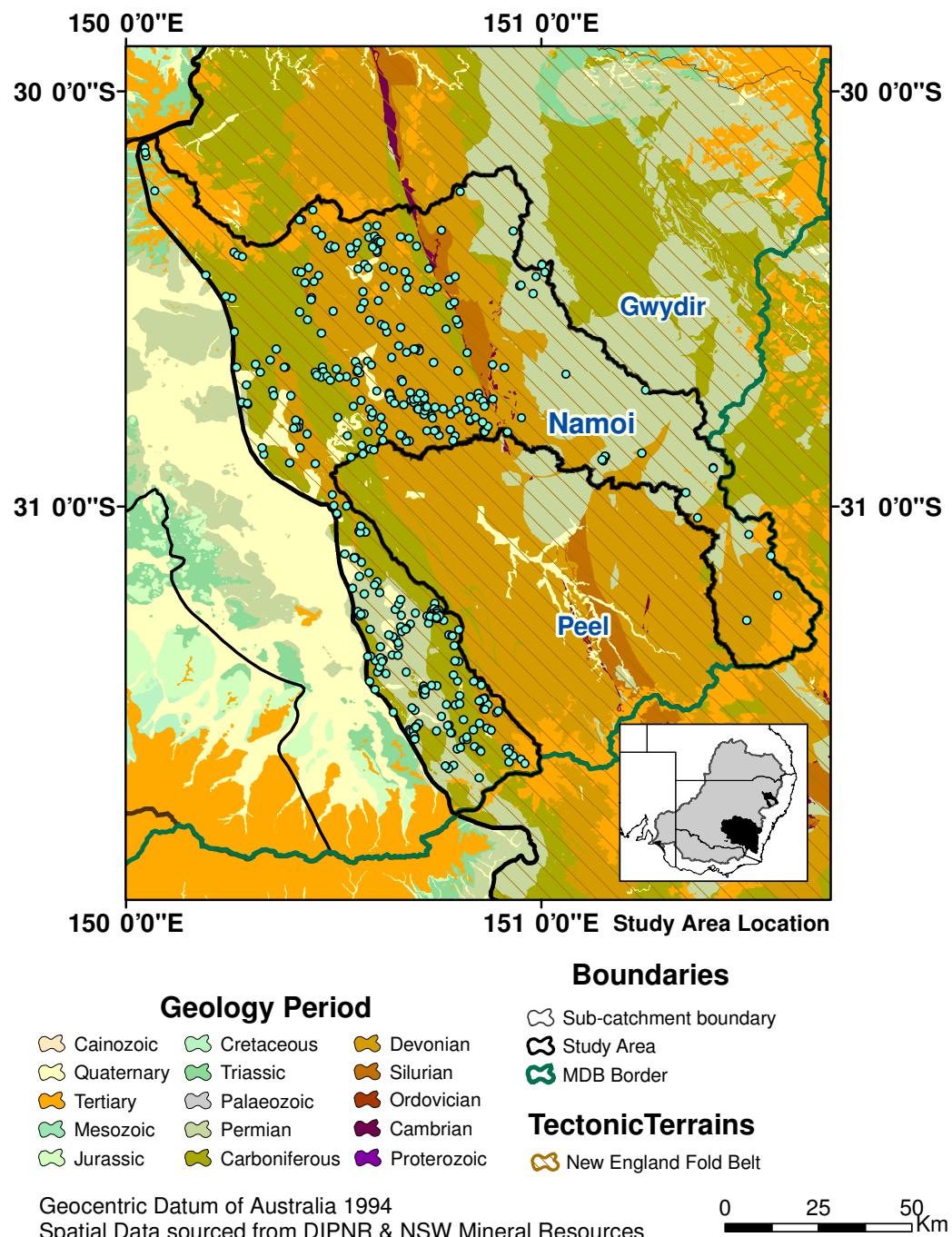


Figure 3.2: Bores projected on Geology map of Namoi catchment.

The GDS database contained 1817 bores in the Namoi catchment with SWL information. As some bores had multiple water bearing zones, the total of 2197 WBZ with SWL were found. Bores for each *Class* were selected using the ArcGIS software. In Figure 3.2 Cainozoic and Quaternary geology is associated with *Class* (I)- (*Unconsolidated sediments*), Tertiary geology with *Class* (II)- (*Volcanics*), and the remaining geology with *Class* (III)- (*Fractured Rock*). The spatial extent of each *Class* was defined using selection by attributes of the geology shapefile. For example, Cainozoic and Quaternary unconsolidated sediments defined *Class* I (Fig. 3.2). If alluvials or Cainozoic volcanics were mapped over the wider extent in groundwater management map, then these additional areas would have been incorporated into the *Classes* I and II.

The bores from the shapefile containing Bore ID, Longitude and Latitude were selected from *Class* I: (*Select, by location, features from bore.shp, that are contained within, selected features from geology.shp*). Selected records from the Bore shapefile were exported from GIS and imported into a Microsoft Excel worksheet. This was repeated in turn for each *Class*. Approximately two thirds of bores projected on the fractured bore geology (1219 bores/1359 WBZ), one third on the alluvials (577 bores/797 WBZ) and a negligible number (21 bores/31 WBZ) on the *Cainozoic volcanics* (Fig. 3.2). WBZ and associated attributes from the GDS database were loaded from the GDS into each corresponding worksheet for further processing, using the Microsoft Excel function VLOOKUP.

### **Classification based on the database attributes**

Spatial classification of bores based solely on geology could be insufficient for selection of fractured rock bores because of the lack of accuracy in bore coordinates and imprecision of geologic mapping, which could place the bore which is close to the border of geologic material in the wrong material. To deal with these problems, a second level of classification was performed, based on the GDS attributes, after which the derived set of rules allowed for the final fractured rock bore selection.

The attribute used from the GDS database for the second level of classification was the water-bearing zone material. The bores were categorised as: fractured rock (Fr), if the attribute indicated *fractured* or *consolidated*, unconsolidated (Unc) or unknown (N/a) (Table 4.3). Due to the small sample size of the Cainozoic volcanic rocks, the second level of classification was not performed for *Class* II (Volcanics).

This resulted in seven *Groups* of bores: three categories for areas defined by the geological map as unconsolidated sediments (*Class I*), three for fractured-rock areas (*Class III*), and a single remaining category for Cainozoic volcanic areas (*Class II*) (Table 4.3). Information from WBZ material in Namoi section revealed that 131 WBZ from Class I (*Unconsolidated sediments*) were drilled in fractured rock - *Group 1* (Fr), 334 in *Group 2* (Unc) and 332 in *Group 3* (N/a). Out of 1369 WBZ in Class III (*Fractured rock*), 670 WBZ belonged to *Group 5* (Fr), 181 to *Group 6* (Unc) and 518 to *Group 7* (N/a), (Table 4.3). Bores from each *Group* were placed in a separate Microsoft Excel worksheet.

After the bores were classified based on the GDS attributes, the set of rules were derived to allow for the final fractured rock bore selection.

### **Discovering selection rules based on the database attributes**

Adopted bores classification in *Classes*, and in *Groups* was insufficient for bore selection, because: (1) shallow bores could be sunk in perched aquifers in unconsolidated sediments above fractured rock; and (2) very deep water bearing zones could belong to a deeper aquifer, typically producing outliers that should be filtered out. To deal with these problems, descriptive modelling was applied to discover the rules for bore selection and eliminate outliers.

Two depth rules were imposed. The minimum depth rule aimed to eliminate shallow bores whose water-bearing zones may occur in perched (unconsolidated) aquifers overlying the fractured rock. The maximum depth rule aimed to minimise SWL outliers (extreme values), which generally coincided with deep water-bearing zones and high vertical hydraulic gradients, typical of recharge zones near the top of ridges and discharge zones in catchment lowlands. Deep fractured-rock water-bearing zones are found in catchment lowlands on account of the thick deposits of unconsolidated materials there. The SWL is either a very small positive value, or more often a negative value representing artesian conditions. Eliminating deeper water-bearing zones also provided confidence that only the local and locally recharged groundwater systems were sampled.

The thresholds in the rules were derived using iterative procedure. Thresholds for *Groups* from *Class III* were derived in a pilot study, based on the dataset of bores that were contained between 148°E and 149°30' E and 32°30' S and 34°S, by minimizing Root Mean Square Error (RMSE) between trends in normalised rainfall

and SWL. The trends were in a form of a 21-year moving average. Thresholds for remaining *Classes* were derived by minimising RMSE between *Class III* and the *Class* in question.

The selection rules within each of the seven *Groups* were derived to predict if the water-bearing zone is representative of fractured rock or not. Rules were based on the groundwater database attributes that were easily accessible in Excel:

- the features of the water bearing zone: (1) depth to the upper margin of the water bearing zone, (2) depth to the deeper margin of the water bearing zone,
- the features of the deepest logged geologic horizon: (1) depth to the upper margin of the deepest logged geologic horizon(2) depth to the lower margin of the deepest logged geologic horizon and (3) geological material, if available, and
- the number of water-bearing zones.

Only bores that had a single water-bearing zone were included for automated bore selection in Excel. Those with multiple water-bearing zones were processed individually.

### **Rules for selecting the bores**

The following set of rules was applied during the predictive, supervised discriminative modelling phase of bore selection:

Water-bearing zones were considered shallow if they were located within the top:

- 10 m, for bores located in fractured-rock areas: fractured-rock zone (*Group 5* in Table 4.3)
- 16 m, for bores in fractured-rock areas: unknown and unconsolidated zone (*Groups 6 and 7*)
- 20 m, for bores located in unconsolidated sediments (*Groups 1, 2 and 3*).

Bores within these categories were eliminated unless a driller's report indicated that the entire thickness of the water-bearing zone was fully contained within fractured rock.

Water-bearing zones were defined as deep if the top of the water-bearing zone occurred below 60 m, and eliminated. Bores remained in the dataset if their water-bearing zones were fully contained within the top 45 m.

Remaining bores (i.e. water-bearing zones reaching deeper than 45 m, with the upper margin not deeper than 60 m) were not eliminated, provided drillers' reports indicated that they were fully contained within a single fractured rock horizon starting at a depth shallower than 30 m, because:

- For highland fractured-rock recharge areas, it was assumed that 30 m was the depth of considerable weathering. Water-bearing zones that were fully contained within a single rock material, which includes the weathering zone, would most likely be recharged locally, and thus be representative of the shallow fractured-rock groundwater system. Where the profile consisted of series of layered rock materials, some materials could limit the water-bearing zone capacity to being recharged locally.
- For lowland areas, the 30 m depth constraint ensured less than 30 m of unconsolidated overburden, which means that the selected bore was not deep in the alluvial flats, but outside or just at the alluvial-fractured-rock boundary, reducing the possibility of the water-bearing zone being influenced by water in the alluvial systems.

As indicated by drillers' logs, the water-bearing zones of some bores in *Groups 2* and *6* (unconsolidated material; Table 4.3) were likely to have represented weathered fractured-rock zones. Provided the other criteria were met, these bores were included in the final dataset.

Bores in Cainozoic volcanic areas (*Class II, Group 4*) were only included if the driller's report showed that the associated rock material was not related to the Cainozoic volcanic rocks. In addition, the bore had to satisfy the same criteria as the other fractured-rock bores from *Group 5*.

These procedures resulted in three possible outcomes: data were eliminated, selected or nominated for further investigation. Some of the bore records nominated for further investigations were checked individually, taking into account the full driller's record, proximity to streams [DIPNR, 2004b] and records of nearby bores in addition to already mentioned sources. This was done for the records which generated small annual sample size, and especially for the early records.

Bores with multiple water-bearing zones had to be processed individually. In the first half of the twentieth century, bores were drilled using a cable tool with simultaneous progressive sinking of the screen (drill and drive method). Drillers using this technique were able to record SWL in each water-bearing zone. Where this was the case, the SWL data for the shallowest fractured-rock water-bearing zone was used. With the introduction of rotary mud drilling, pumping tests were generally performed after the bore was drilled, and the resulting SWL was not representative of a particular zone. These data were not included unless the topmost water-bearing zone in the fractured rock was the only zone that was screened.

While stock and domestic bores generally have a single pipe running down the borehole, monitoring bores may have multiple pipes. If the information was available, construction details and drillers' reports were used to select appropriate water-bearing zones.

After applying selection rules to Namoi section, a total of 422 bores were selected (Table 4.3):

- 41 out of 797 WBZ from *Class I: Group 1 – 29; Group 2 – 2; Group 3 – 10;*
- 3 out of 31 WBZ from *Class II, Group 4;*
- 378 out of 1369 WBZ from *Class III: Group 5 – 272; Group 6 – 7; Group 7 – 99;*

### **Screening of bores influenced by alluvial aquifers**

A small number of fractured-rock bores were selected based on the GDS attributes from the areas mapped as unconsolidated sediments - Group 1 in Table 4.3. These bores were close to the border between fractured rock (with water-bearing zones) and alluvial flats. They either had a shallow alluvial layer or were entirely in the fractured rock, but owing to the imprecise mapping or to scale problems, they appeared as mapped in unconsolidated sediments. The final step in bore selection was to compare trends in SWL between these bores and adjacent bores located in mapped fractured-rock areas (*Group 5*). This precautionary step ensured that the behaviour of fractured-rock bores mapped within unconsolidated geology was not influenced by nearby alluvial groundwater systems. Root Mean Square Error was used as the criteria.

## Final dataset

Selected bores from all *Groups* were joined to form the final dataset (Table 4.3).

### 3.2.2 Creation of composite SWL time-series

#### Choice of central tendency model for annual time-series

The final dataset consisted of a variable number of SWL points for each year. Three possible approaches were investigated for transforming this data into the composite annual time series: arithmetic means, spatially weighted averages and medians. The spatially weighted average was explored in order to reduce the bias that could be introduced by the multiple bores in close proximity drilled during the same year. This was done by splitting the section into 30 seconds squares. Average SWL was calculated for each square for a particular year. Then the average was calculated from these values.

It was found that annual medians give the most convenient and robust measure of central tendency. This is clearly demonstrated by the 95% Confidence intervals (CF) plotted for the Namoi section (Figure 3.3).

$$CF = \frac{StDev * Tinv(p, m)}{\sqrt{n}} \quad (3.2)$$

where:

$n$  is the sample size (21: for a 21 year moving average);  $StDev$  is the standard deviation;  $p$  is the probability associated with the two-tailed Student's t-distribution (0.05 for the 95% confidence interval);  $m$  is number of degrees of freedom ( $m = 21$  for the 21-year moving average) and  $Tinv(p, m)$  is the inverse of the t-value of the Student's t-distribution.

Confidence intervals are positioned closer to the trendline for the Median central tendency model than for the other two models.

Root Mean Square error between lagged 21-year moving average of SWL and rainfall was used as the testing criteria for the model choice. The spatially weighted average technique showed a marginally better accuracy only in one of the tested datasets outside of the study area (in the Peel catchment, not presented here), where bore density was high (1 point per  $5 \text{ km}^2$  between 1912 and 2004), [Rančić et al., 2009] and bores were geographically clustered in some years. As the benefit

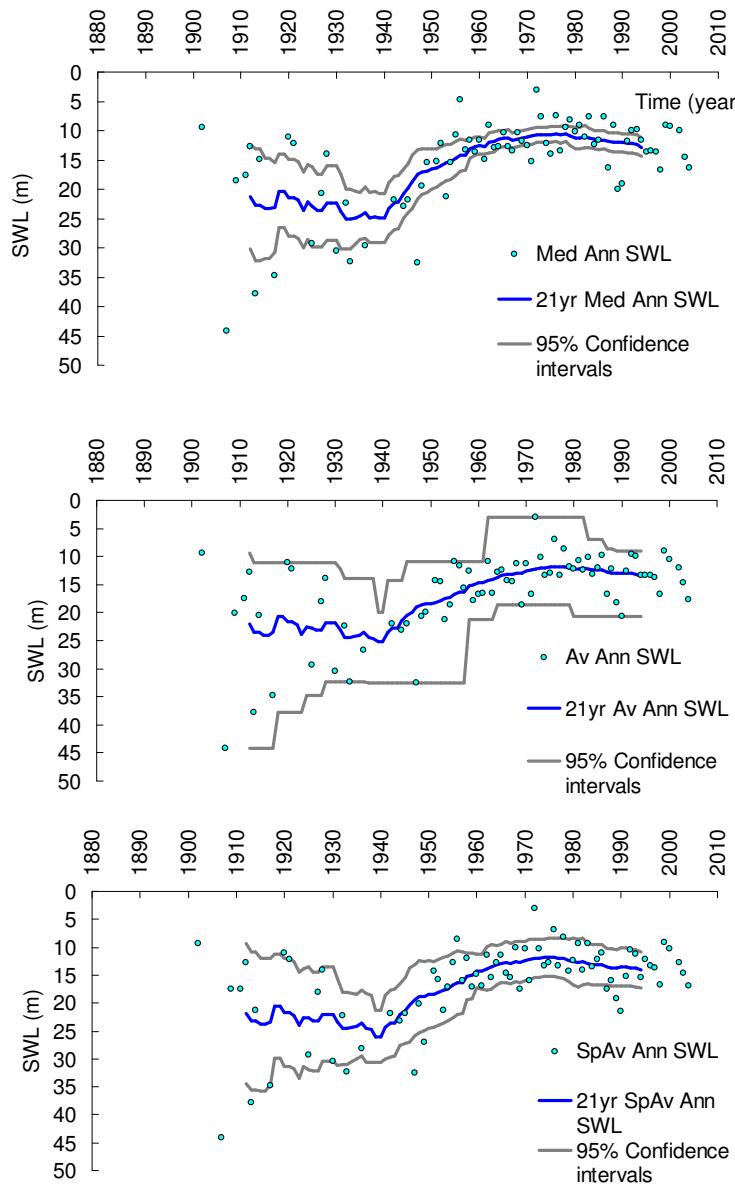


Figure 3.3: Namoi section: choice of the central tendency model for aggregation of SWL points into annual data: Annual median (Med Ann SWL), arithmetic average (Av Ann SWL) and spatially weighted average (SpAv Ann SWL).

was marginal and restricted to only one case study, and the median model was much simpler than the spatial average, for each section (Figure 1.2) and subsection, composite SWL time series were thus constructed using the median annual SWL values.

### Alternative filtering techniques

Two alternative filtering techniques for trend representation were tested: a simple moving-average filter and splines.

Moving average filters of odd length, ranging from 5 to 23 years, were tested on the dataset from the Macquarie valley used by Salas and Smithson [2002]. The odd length was chosen to be able to plot the filter on the central year ( $Y$ ). Only SWLs at a time of bore construction were used. Rainfall data from Salas and Smithson [2002] were used as a surrogate training data-set for the choice of filter length. Moving average filters of standing water level and rainfall were plotted, compared, lagged, normalised and evaluated with the Root Mean Square Error (RMSE). To evaluate the goodness of fit using root mean square error, it is necessary to perform a linear transformation of each series by normalising the series. The convention of SWL axes pointing downwards (Figure 3.1) demanded to inverse the sign of the SWL data when normalised:

$$Nswl(i) = -\frac{SWL(i) - Av}{StDev} \quad (3.3)$$

where  $Nswl$  is normalised  $SWL$ ,  $i$  is the year,  $Av$  represents arithmetic average for the  $SWL$ , and  $StDev$  the standard deviation.

The RMSE followed a monotonically decreasing function, until it reached the minimum at 21 years. The presented optimum was unique. Therefore, the 21-year moving average filter was adopted. The analysis was performed using Microsoft Office Excel.

Validation on the Namoi section (Figure 3.4) confirmed that the 21 year is the optimum filter length. Composite rainfall, which included weighted values from five rainfall stations in the Namoi section were used for this exercise. The results of the validation on the Namoi catchment (Table 3.2) demonstrate that the RMSE steeply declines when approaching the 21 year optimum. The table presents RMSE calculated over 83 years (spanning the full SWL record available in Namoi section for the 21 year filter: Rainfall 1908–1990; SWL 1912-1994) and over 81 years (cor-

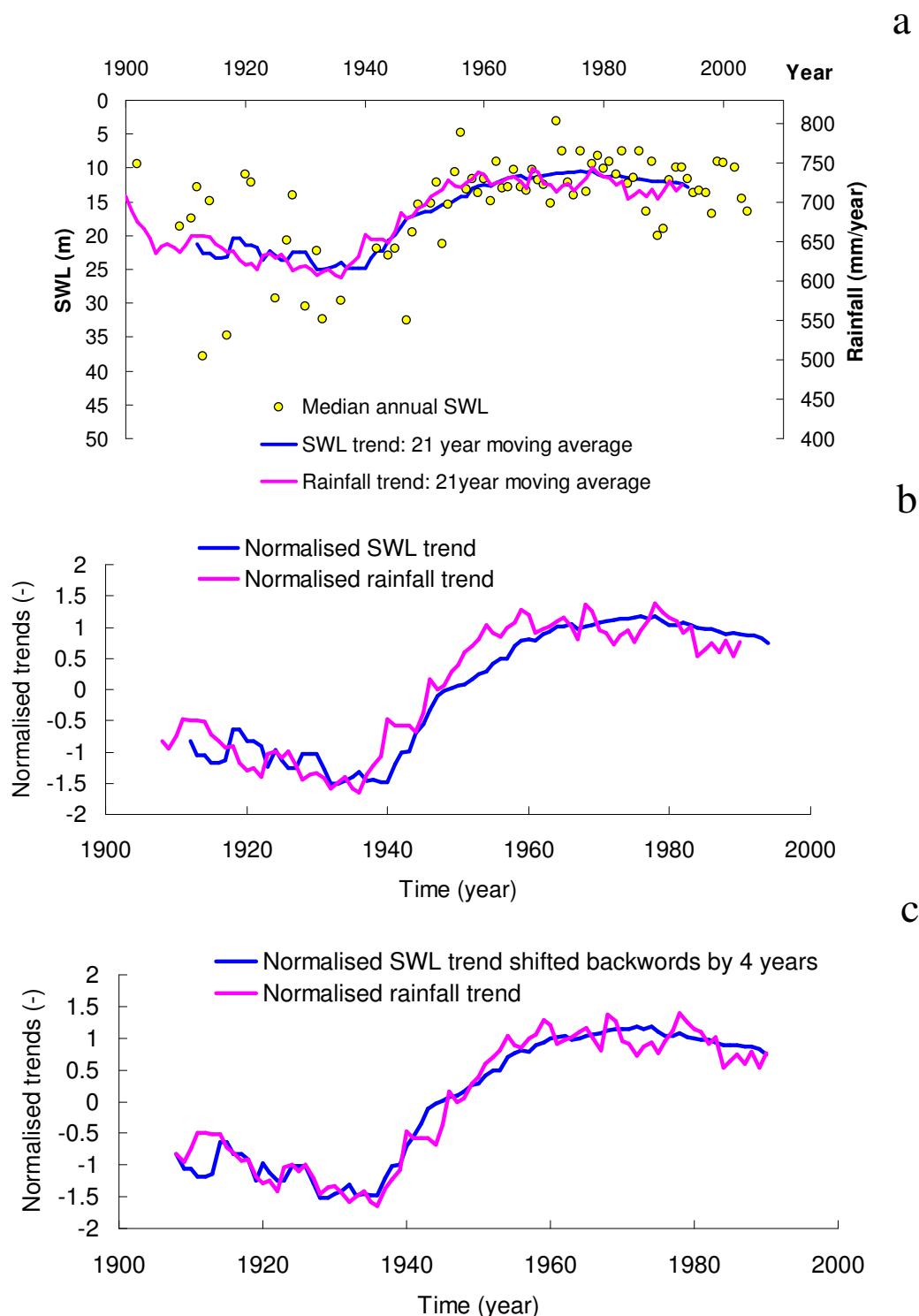


Figure 3.4: Normalised, 21-year moving average filter, for rainfall and SWL in Namoi section.

Table 3.2: Namoi section – Choice of length (years) for the moving average filter.

<b>Length (years)</b>	<b>RMSE*</b>	
	<b>n=81</b>	<b>n=83</b>
15	0.4322	0.4357
17	0.4030	0.4027
19	0.3369	0.3393
21	0.2430	0.2401
23		0.2498

RMSE\*–Root Mean Square Error;  
n–Sample size.

responding to the span of the 23 year moving average filter, starting a year later and finishing a year earlier). As the SWL trend lagged by 4 years after the rainfall trend (Figure 3.4 a and b), it was shifted back prior to RMSE calculation (Figure 3.4 c). Although obvious from the visual inspection, the optimum lag of four years for the Namoi section was confirmed by the RMSE.

Splines were tested on the Namoi catchment using the R software package. A weighted smoothing spline was fitted to the median SWL data. The weights were equal to the number of SWL data available per year. The degree of smoothing from 2 to 9 was investigated (Figure 3.5). Weighted spline order 7 was chosen to represent the trend.

### Exploratory analysis of bore data

The exploratory data analysis aimed to maximise insights into the data, to uncover their underlying structure, to extract important variables, to detect outliers and anomalies, and to test underlying assumptions. Histograms, scatter plots, and three-dimensional surface splines were used. The summary statistics of the number of data points available for analysis each year were also calculated.

Using the latitude and longitude of individual bores, the altitude was estimated based on the DEM network and relationship between altitude and SWL was investigated. Regression between SWL and elevation indicated a weak inverse relationship: the water table was generally deeper in the dryer and less recharged lowlands in the

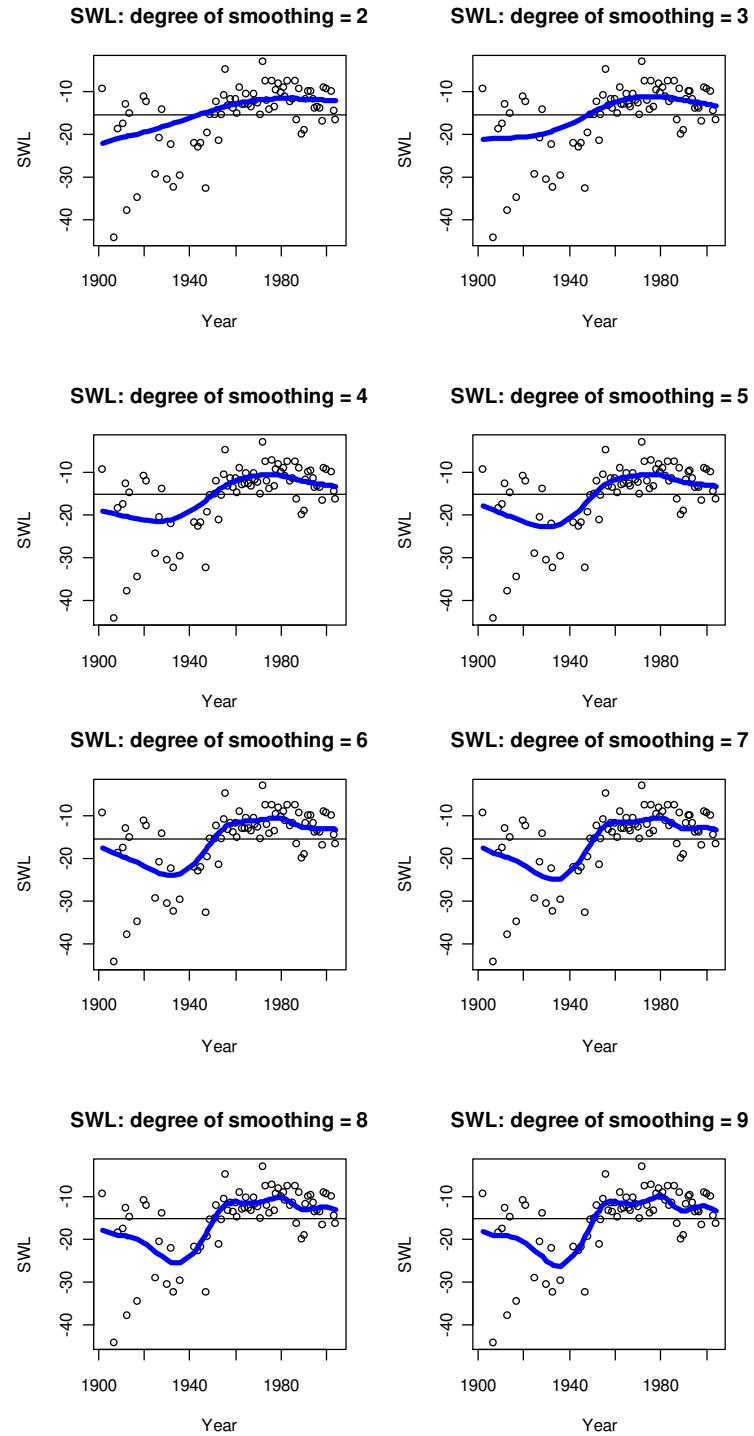


Figure 3.5: Splines filtering technique - Namoi section.

western part of the study area.

Bore data from each section were categorised into time periods (pre-1950, 1951–1970, 1971–1990 and post-1990), as this allowed for the examination of temporal and spatial distribution. Scatter plots of bore elevation and SWL against time were drawn to evaluate uniformity: Non-uniform spatial bore spread can prevent us from knowing whether a change in SWL over time is due only to processes that operated over the time period, or whether the difference in bore position within the landscape also played a role in the calculated SWL values. This inability to distinguish between the influence of two variables (such as space and time) on the third variable (SWL) is called spatial/temporal confounding. Figure 3.6 demonstrates graphs produced for the exploratory analysis of the Namoi section.

Although the rainfall data represented a true time series, the bore data consisted of a variable number of SWL points for each year up to August 2004 (Tables 3.3 and 4.4). SWL data were not available for every year. Data were sparse in the early records, and had to be checked for potential geographic clustering, regional and altitude-related change in depth to water table, to avoid spatial-temporal confounding. As the geographic clustering proved not to be the problem (Section 3.2.2), more of the concern was the regional and altitude-related change in depth to water table in some of the sections.

The specific details sought in the exploratory phase of data analysis were:

- the number of data points available for analysis;
- the relationship between SWL and elevation, longitude and latitude;
- issues such as confounding of SWL by elevation, longitude or latitude over time; and
- the spatial confounding of SWL data by land use, based on the land-use capability map.

Table 3.3: Number of data points by year in Namoi section.

<b>Namoi</b>												
Year	1902	1907	1909	1911	1912	1913	1914	1917	1920	1921	1925	1927
SWL no	1	1	8	2	2	1	5	1	1	1	2	3
Year	1928	1930	1932	1933	1936	1942	1944	1945	1947	1948	1949	1951
SWL no	1	1	1	1	4	1	3	1	1	3	6	3

Table 3.3: (continued)

Year	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963
SWL no	3	1	6	15	10	8	10	3	14	8	4	4
Year	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975
SWL no	1	12	5	5	6	7	5	5	1	5	9	8
Year	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987
SWL no	4	3	6	5	15	27	21	11	4	4	3	10
Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
SWL no	6	9	8	4	11	6	15	11	1	1	2	2
Year	2000	2002	2003	2004								
SWL no	5	7	7	5								

The exploratory analysis was performed using routines in S-Plus 7 software [Incorporated, 1995].

### Post-hoc segmentation

Study area segmentation is carried out iteratively. The aim of segmentation is to form homogeneous groups based on patterns and data. As Spate et al. [2002] point out, there is a trade-off between the amount of information stored in the groups and usefulness of information: enough groups are required so that the internal objects are close, but not so many that there are similar groups. Sections had to be small enough to preserve hydrologic homogeneity and avoid spatial-temporal confounding, but large enough to contain sufficient number of bores from which the aquifer behaviour can be reconstructed. This was particularly applicable for the long, and therefore pronounced rainfall gradients found in the elongated Lachlan and Murrumbidgee catchments.

The initial segmentation of the study area was done based on river catchment boundaries, with underlying assumption that each watershed is underlined by a separate aquifer that feeds a corresponding stream. Exploratory analysis was then performed to ensure data homogeneity. Out of all the variables that were tested, only elevation showed some relationship with SWL. The regression analysis indicated a very weak relationship between SWL and elevation, which was, an implicit reflection of the rainfall gradient. The relationship was less pronounced in the compact Namoi section and more pronounced in the two southern, elongated catchments, that stretched from well watered highlands in the east to the low-lying, dryer plains in the

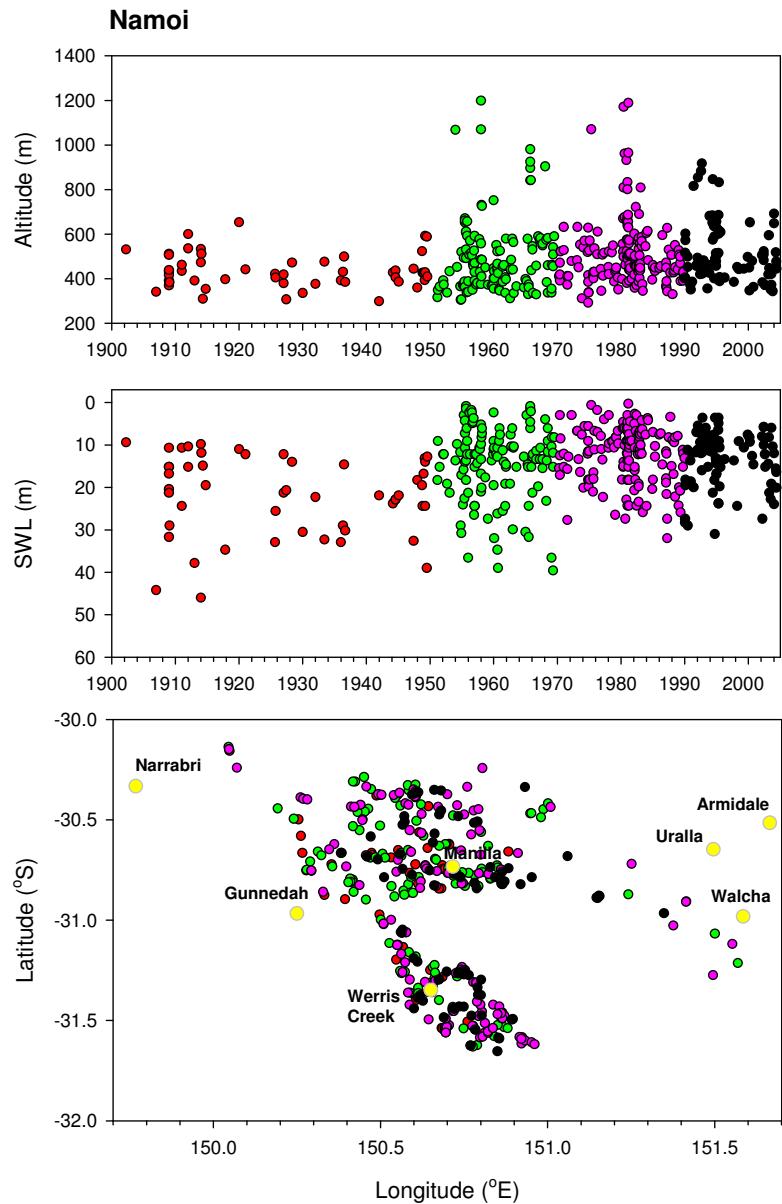


Figure 3.6: Exploratory analysis for Namoi Section: Timeseries of bore altitude and SWL and spatial distribution in relation to time of construction. Points are coloured in red for bores constructed in 1902–1949 period, in green for 1950–1969, in magenta for 1970–1989 and in black for 1990–2004 periods.

west where water table was deeper (Figure 1.3 and Figure 1.4). This highlighted the need to subdivide them into smaller, less elongated, and more hydrologically homogeneous units.

Segmentation resulted in a total of 6 sections:

- Namoi catchment (1 section);
- Lachlan catchment (3 sections): East, Mid, West; and
- Murrumbidgee catchment (2 sections): Murrumbidgee East, also known as the Upper Murrumbidgee, the entire catchment area upstream of Burrinjuck Dam; and Murrumbidgee West, the catchment downstream of Burrinjuck Dam within the Lachlan Fold Belt.

As the rivers in the study area generally flow in westerly direction, so that Lachlan and Murrumbidgee sections were elongated in east-west direction following the direction of the rainfall gradient. Therefore the dissection was performed in north-south direction, delineating eastern, (mid) and western portion of the Murrumbidgee and Lachlan catchments. Three criteria were respected:

- number of bores in each segment had to be sufficient to derive trends;
- areas had to be split in more hydrologically homogeneous sections, minimising rainfall range, with similar spatial coverage; and
- dissection of uplands from lowlands was performed following some natural geographic features:
  1. Burrinjuck Dam, which separates the Snowy Mountains highlands from the lower, flatter landscape in the west of the Murrumbidgee Catchment, and coincides with the split between uniform and winter rainfall zones;
  2. Confluence of Lachlan River and its north-flowing tributary, which separates the steeper, hilly area in the east from the rolling hills with the alluvial valleys in the mid section of the Lachlan catchment;
  3. Change from the rolling hills landscape into much flatter lowlands with deeper and wider unconsolidated sediment layers in the west.

The newly formed sections were again examined using exploratory analysis. This indicated three potentially non-homogeneous sections. One was Murrumbidgee East, with the potential to be further subdivided into North-East and South-East sections. However, Murrumbidgee West and Lachlan West could not be subdivided, due to the insufficient number of bores in the west portions of these sections. Therefore analysis was performed only on the eastern part of these sections: Lachlan Inner West and Murrumbidgee Inner West (Figure 1.2) and compared with the results for the un-dissected Lachlan West and Murrumbidgee West sections. This resulted in a total of 9 subsections (Table 4.5).

Trends in the Lachlan Inner West and Murrumbidgee Inner West showed similar patterns to those in the entire western subsections, indicating that spatial/temporal confounding was not a major issue, and eliminating the need to perform separate statistical analysis for the Inner West subsets (Table 4.5).

### 3.2.3 Noise filtration

Due to a small sample size, especially during the early years (Tables 3.3 and 4.4), the annual SWL data were characterised by high variability (Figures 3.3 and 3.5). To eliminate the noise, introduced by small annual sample size and years with missing data, which are often encountered in the data mining from databases [Hand et al., 2001], the annual time series were subjected to filtering techniques: residual mass curves, linear moving average and splines. These techniques represent low-pass filters, which smooth the data, attenuate short-term variability and accentuate underlying trends. As pointed out in Section 2.2, Schall and Samozino [2005] demonstrated that surface reconstruction can be drastically improved if the appropriate preliminary filtering method was applied to original scattered data.

The creation of trendlines for Namoi catchment as a product of noise filtration has already been presented in detail in Section 3.2.2:

- **the 21 year moving average filter** of median annual SWL in Figure 3.3, 21 year moving average filter of normalised median annual SWL shifted back 4 years and the 21 year moving average of normalised annual rainfall in Figure 3.4; and
- **the weighted spline order 7** of median annual SWL in Figure 3.5.

The results of the filtering for the rest of the study area is given in the Chapter 4.

The third form of the low-pass filter applied in this study are the residual mass curves. They are used to detect abrupt hydrologic changes. This technique complements the trend analysis, which by its nature smoothes the function it is applied to, and therefore distorts and masks the abruptness of the change.

### Residual mass curves

A residual mass curve (Figure 3.7) represents the cumulative sum of deviations from the long-term average:

$$\sum_{i=1}^n (X_i - m) \cdot \Delta t, \quad (3.4)$$

where  $\Delta t$  is the time step or time interval (year, month, day, hour, minute, second ...);  $X_i$  is the  $i$ th element of the series, representing the average value over time interval  $\Delta t$  (mm/year, mm/month ...); and  $m$  is the long-term average (mean), expressed in the same units as  $X_i$ .

The residual mass curve is an approximation of an integral of deviations from the average (the sum becomes an integral when the time-step approaches zero):

$$\int [X(t) - m] dt \quad (3.5)$$

The extremes of the residual mass curves are often used to detect and illustrate abrupt changes in the mean. Slopes of the segments between the extremes (first derivative of the integral) correspond to the values of the mean preserved during each segment. A positive slope (rising trend) indicates that the period is characterised by the mean being greater than the long-term mean, and a negative slope (falling trend) indicates the opposite. If the residual mass curve shows a pointy 'V' shape (discontinuity of the first derivative), such as in Namoi section, (Figure 3.7), the system abruptly changed from one regime (mean) to another. A smooth 'U'-shaped curve (continuity of the first derivative is preserved) indicates a smooth transition.

Extremes in residual mass curves underpin the Change-point analysis methodology. In addition to relying on residual mass curves for detection of change points, by identifying the extremes in residual mass curves, the approach uses bootstrap analysis for calculation of confidence levels. Bootstrapping is a distribution-free approach. The only assumption for a bootstrap is that the re-sampled data are independent

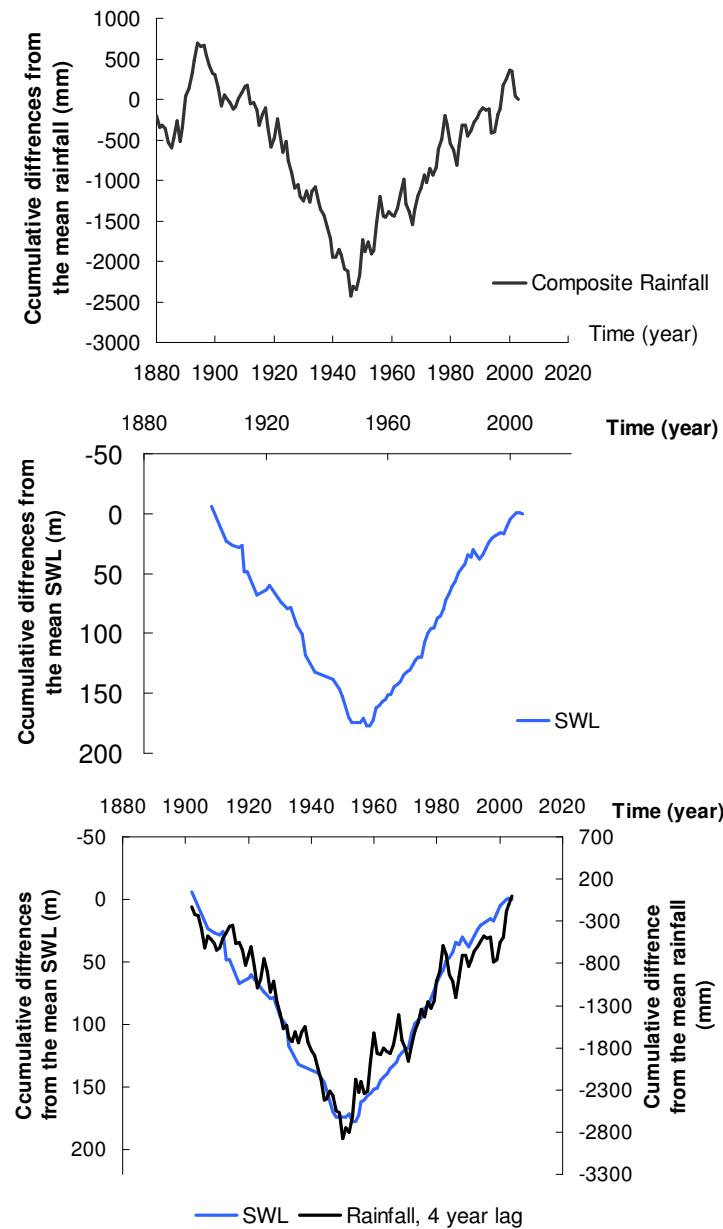


Figure 3.7: Residual mass curves of rainfall and SWL in Namoi catchment.

and identically distributed. Once a change has been detected, the data are segmented each side of the change-point, and the analysis is repeated for each segment. The procedure is iteratively repeated until all abrupt changes with a confidence level above the set limit (minimum 90%) have been accounted for. Change-point analysis was undertaken using the software package Change-Point Analyser v. 2.3 [Taylor, 2000].

As the residual mass curves plots indicated abrupt hydrologic change in rainfall data around the middle of the twentieth century and the rise of water table following this change, the change-point analysis was performed to evaluate the precise timing and the statistical significance of this event.

Table 3.4: Change-point analysis of SWL data in Namoi section.

Section/subsection	Year change occurred	Confidence level	From (m)	To (m)
Namoi	1949	>99.5%	23.2	15.9
	1955	95%	15.9	11.6

Confidence level >95%; bootstraps = 1000.

The results of the change point analysis for SWL and rainfall for Namoi catchment are presented in Tables 3.4 and 3.5.

Table 3.5: Change-point analysis of individual rainfall series for Namoi section.

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
<b>Namoi</b>				
Gunnedah Pool	1897	94%	675.3	536.7
	1947	>99.5%	536.7	661.2
Goonoo Goonoo Station	no change			
Barraba Post Office	1947	97%	650.4	730.1

Confidence level of >90% was used to indicate a significant change.

Table 3.5: (continued)

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
Wallabadah (Woodton)	1889	95%	653.4	980.0
	1895	98%	980.0	694.0
	1949	>99.5%	694.0	834.0
Manilla Post Office	no change			
<b>Weighted-average rainfall</b>	<b>1947</b>	<b>97%</b>	<b>647.8</b>	<b>726.7</b>

Confidence level of >90% was used to indicate a significant change; bootstraps = 1000.

The results indicate the two-step rise in water table in Namoi section: in 1949 and 1955 (Table 3.5), following abrupt change from dry to wet hydrologic regime in 1947. The lags are therefore two and eight years, the average being five, close to the lag of four years captured by RMSE between the moving average of rainfall and the moving average of SWL.

### 3.3 Summary

This chapter describes a data mining method for reconstruction of aquifer behaviour in fractured rock, which uses depth to water table records, collected when the bores were constructed. Bores in shallow fractured Palaeozoic rock were selected, and those in deeper rocks, in unconsolidated sediments and in Cainozoic volcanics eliminated. Annual standing water level time-series were created using median function. The noise in the series was filtered out using residual mass curves and trends.

A worked example has been given to demonstrate the use of the methodology. The data for a part of the Namoi catchment has been used in this worked example.

In the following Chapter, the detailed application is given for the sections of Lachlan and Murrumbidgee catchments.

# Chapter 4

## Results

This chapter presents the results of the data-mining technique for reconstruction of aquifer behaviour from independent point data across the study area (Figure 1.2): sections of the Lachlan and Murrumbidgee catchments. Results for the Namoi section have already been presented in Chapter 3, to illustrate the method. Some of these results are repeated to facilitate comparison with the other sections.

### 4.1 Data sources for data mining

Bore information is the primary data source used for this study. The spatial distribution of the selected bores is presented in the Figure 1.2.

Table 4.1 provides the summary of all the SWL points drawn from the database, from which the bores in fractured rock were selected. Some bores had multiple water-bearing zones (WBZ) so the total number of associated SWL points was greater than the number of bores.

Table 4.1: Available SWL data points.

	WBZ*	Bores
Namoi, New England FB**	2197	1817
Lachlan, Lachlan FB**	8837	7459
Murrumbidgee, Lachlan FB**	7216	6191
Total	18253	15467

\*WBZ – Water Bearing Zones; \*\*FB – Fold Belt.

In general, before 1950, and even in some years after 1950, data were sparse.

Secondary sources of data used in this study consist of: a) spatial digital maps, required for bore selection and/or exploratory data analysis, b) recorded ground-water hydrographs of limited duration, used for partial model validation and c) the long-term rainfall data, used as an additional hydrologic resource for the model validation.

Spatial digital maps used as the secondary data source were the Topographic and Geology maps. Topographic map (Figure 1.3 in Section 1.5.1), was required for a) catchment boundary and stream network definition, and b) exploratory analysis of bore data. Geology map (Figure 1.5 in Section 1.5.3) was required for selection of bores in fractured rock.

Recorded groundwater hydrographs of limited duration, used for model validation are presented in Section 5.1.

Table 4.2: Selection of PPD\* rainfall stations.

Catchment	Available PPD rainfall stations	< 95% recorded data since 1900	Selected PPD stations
Namoi	24	10	5
Lachlan	87	24	20
Murrumbidgee	108	20	15

\*PPD=Patch Point Data

Out of more than 200 rainfall stations across the study area only 40 were reliable. Those were the stations with a long term record, with less than 5% of missing data and with the record consistent with the stations around them (Table 4.2). Poorly managed rainfall stations in the south of NSW resulted in only a small portion of stations with relatively complete and reliable record. The 21-year moving average filters of composite rainfall time-series are presented in Section 5.2

## 4.2 Bore data mining

This section describes the results of the three phases of the bore data mining method: a) bore selection, b) creation of composite SWL time series and c) noise filtering.

### 4.2.1 Bore selection

Table 4.3 illustrates bore selection process:

- spatial classification of bores into 3 *a priori* defined *Classes*,
- second level of classification to form *Groups*,
- selection of bores, based on rules described in Section 3.2.1, and
- final selection of bores, after screening and elimination of bores influenced by alluvial aquifers (presented in italic).

Table 4.3: Bore Selection.

Class	Class	Group	Group	Namoi New		Lachlan		Murrum-		Total		
				no	England FB	Avail-able	Selec-ted	Avail-able	Selec-ted	Avail-able	Selec-ted	
Uncon-soli-dated	I	Fr	1	131	29	795	178	460	100	1386	207	
		Unc	2	334	2	2977	53	1269	0	3580	55	
		N/a	3	332	10	1153	60	865	5	2350	60	
		To-tal	WBZ	797	41	4925	291	2594	105	7315	332	
		Bores		577	41	4167	291	2269	105	7013	332	
		Volca-nics	WBZ	4	31	3	397	19	56	0	487	22
			Bores		21	3	282	19	5	0	303	22
Fra-ctu-red	III	Fr	5	670	272	2011	789	2210	608	4891	1669	
		Unc	6	181	7	273	10	225	7	679	24	
		N/a	7	518	99	1231	267	2130	409	3879	775	
		To-tal	WBZ	1369	378	3515	1066	4566	1024	7216	1024	
		Bores		1219	379	3010	1066	3877	1024	6191	1024	
		To-tal	WBZ	2197	422	8837	1376	7216	1024	18250	2822	
		Bores		1817	422	7459	1376	6191	1024	15467	2822	
Additional SWL points						0	26			0	26	
<b>Total selected SWL points</b>				<b>422</b>		<b>1402</b>		<b>1024</b>		<b>2848</b>		

Selected bores in Italic; FB = Fold belt.

A small number of the additional points (26) used in Lachlan catchment were drawn from subsequent measurements of the selected bores. The two readings from the same bore were several decades apart, and therefore could be considered independent. This was implemented to increase the number of points in the years where the total number of points was small.

### 4.2.2 Creation of composite SWL time-series

The data availability for developing composite SWL time series was variable between catchments (Table 4.3) and over time (Table 4.4). In general, before 1950, data were sparse, i.e. containing gaps and with few observations per year, or, except for the Namoi section, missing entirely the first decades of the twentieth century. Bore drilling commenced much earlier in the flatter, western parts of the Lachlan and Murrumbidgee catchments, than in the hillier east (Table 4.4), allowing longer hydrograph reconstruction.

Table 4.4: Number of data points by year in each catchment section.

<b>Lachlan East</b>													
Year	1941	1945	1946	1947	1949	1950	1952	1953	1954	1955	1956	1957	
SWL no	1	1	2	1	1	1	6	5	5	3	1	10	
Year	1960	1961	1962	1964	1965	1966	1967	1968	1969	1970	1972	1973	
SWL no	3	6	4	3	4	9	5	9	1	1	2	11	
Year	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	
SWL no	10	8	6	8	6	10	24	14	7	12	6	8	
Year	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998	
SWL no	14	3	9	3	11	5	4	7	6	2	2	8	
Year	1999	2000	2001	2002	2003								
SWL no	6	7	13	13	5								
<b>Lachlan Mid</b>													
Year	1930	1939	1941	1942	1946	1950	1951	1952	1953	1954	1955	1956	
SWL no	1	1	1	2	2	2	7	3	8	3	7	3	
Year	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	
SWL no	5	10	14	11	11	16	5	11	15	24	17	26	
Year	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	
SWL no	11	3	2	3	9	4	11	1	4	4	10	24	
Year	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	
SWL no	28	16	11	5	8	5	9	9	1	3	5	9	
Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003		
SWL no	3	6	8	3	15	24	10	4	13	4	7		
<b>Lachlan West</b>													
Year	1911	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	
SWL no	1	1	1	1	3	4	2	14	7	8	3	2	
Year	1933	1937	1938	1939	1942	1944	1945	1946	1947	1948	1949	1950	
SWL no	1	4	7	2	1	6	5	6	6	6	6	5	
Year	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	
SWL no	1	3	8	11	8	3	10	8	5	11	9	9	
Year	1963	1964	1965	1966	1967	1968	1969	1970	1971	1973	1974	1975	
SWL no	12	9	32	15	23	25	7	2	2	14	5	5	
Year	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	
SWL no	2	8	8	9	27	15	26	7	5	6	9	11	

Table 4.4: (continued)

Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
SWL no	9	1	3	5	9	3	6	8	3	15	24	10
Year	2000	2001	2002	2003								
SWL no	4	13	4	7								
<b>Murrumbidgee East</b>												
Year	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1964
SWL no	13	6	1	11	7	8	5	1	3	6	1	5
Year	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1977
SWL no	2	4	4	9	3	3	3	5	17	6	3	5
Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
SWL no	13	6	16	12	9	18	12	43	20	11	3	3
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
SWL no	4	5	5	5	24	6	3	16	34	31	29	34
Year	2002	2003	2004									
SWL no	46	15	9									
<b>Murrumbidgee West</b>												
Year	1920	1921	1922	1923	1924	1928	1929	1930	1932	1935	1936	1937
SWL no	1	2	13	3	2	4	1	2	1	1	2	1
Year	1940	1942	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953
SWL no	2	1	1	8	6	2	7	8	2	6	4	4
Year	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965
SWL no	8	8	2	2	6	1	6	13	5	4	23	33
Year	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977
SWL no	15	30	26	7	3	2	5	10	3	2	1	1
Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
SWL no	9	2	12	13	1	3	1	10	3	7	3	2
Year	1990	1991	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
SWL no	1	2	4	2	8	2	7	14	12	10	3	13
Year	2003	2004										
SWL no	11	9										

Median annual SWL timeseries are constructed for each section of the Study area. The resulting hydrographs contain gaps in the years for which there are no data. Due to the years with small sample size, the noise in the data is significant, so that this resulting series does not represent the true system behaviour. Plots of these composite SWL time-series are presented together with their filtering in the Section 4.2.3.

### Alternative filtering techniques

Use of three low pass filtering techniques: moving average, splines and residual mass curves is illustrated using Namoi section of the Study area in: Sections 3.2.2

(Table 3.2 and Figures 3.4 and 3.5); and Section 3.2.3 (Figure 3.7). The moving average filter of the median annual SWL timeseries for sections of Lachlan and Murrumbidgee catchments is presented in Section 4.2.3 with the 95% confidence intervals. The residual mass curves of the median annual SWL for the Lachlan and Murrumbidgee catchments is also be presented in the Section 4.2.3 to illustrate the timing of abrupt rise in water tables. The significance of this abrupt rise is be evaluated using change-point analysis.

### **Exploratory analysis of bore data**

The aim of exploratory data analysis is to maximise the insights into the data, uncover underlying structure, extract important variables and detect outliers and anomalies. In data-mining it is also used as a vehicle for post-hoc segmentation. It aids in defining the geographic boundaries with optimum amount of data for trend and change-point analysis: sufficient amount of data to extract meaningful information, but keeping the geographic domain compact, to insure homogeneity and avoid confounding effects.

Data for developing composite SWL time series were sparse before 1950 (Table 4.4). Therefore, sections of the study area were examined in relation to relative uniformity of altitude, latitude, longitude and land-use of bores between the first and the second halves of the century. Years when there were few data points caused considerable apparent variability within the time series, as the small sample size leads to an estimate of a median that diverges substantially from the population median, introducing artificial variability in the time series.

The graphs for the Namoi, Lachlan and Murrumbidgee cathments show:

- (a) Spatial bore distribution in Figures 4.1 and 4.4
- (b) Timeseries of bore elevation in Figures, 4.2 and 4.5, and
- (c) SWL versus time, for each bore used for creation of median annual time-series in Figures 4.3 and 4.6.

Graphs were plotted and analysed to detect non-homogeneity, non-uniformity and possible confounding effects. The relationship between SWL and elevation, longitude and latitude was examined. The exploratory analysis of SWL data found that of all the variables examined, only elevation was weakly related to SWL.

At the regional scale, bores in the western, arid lowlands tend to have slightly deeper groundwater than the bores from the well watered highlands. Correlation between SWL and elevation confirmed the existence of this weak, inverse relationships in most catchments. This non-homogeneity of data would have introduced bias if the bores from short spectrum of elevations dominated certain temporal periods. This type of problem is called spatial-temporal confounding. To avoid this problem, Lachlan catchment was dissected into three (east, mid and west) and Murrumbidgee catchments into two (east and west) smaller units, and the exploratory analysis was repeated: the three types of graphs were constructed for each of these smaller units (Figures 4.1 to 4.6), median annual SWL calculated and trends estimated.

Analysis of Figure 3.6 shows only a marginal degree of non-uniformity in the Namoi catchment, as before 1950 slightly more bores were drilled in the lower elevations than in the higher, whereas later in the century the spread is equal and uniform. However, the extent of this non-uniformity is mild enough not to require dissection into smaller units.

The inspection of the Lachlan West catchment indicated that the earliest drilled bores from the very west part of the section could be possibly introducing bias to the beginning of the SWL reconstruction, due to their naturally deeper water table. To examine this, the Eastern portion of the Lachlan West section, which contained the majority of bores was defined as the Inner Lachlan West for further analysis.

In Murrumbidgee East section, the area north of Canberra, mainly containing the Yass River catchment, was at lower elevation and had a higher bore density than the Murrumbidgee headwaters. To examine the effects of that potential non-uniformity Murrumbidgee East was split into its Northern and Southern sections.

In Murrumbidgee West its north-eastern segment had a higher density than the remainder of the section, so it was defined as the Murrumbidgee Inner West to be subjected to further examination.

The procedure was repeated for the third time, iteratively: graphs for each section were produced (Figures 4.1 to 4.6), annual median time-series constructed and filtering then applied.

In the northern and southern portions of Murrumbidgee East section (upstream of Burrinjuck Dam), only a few bores were drilled before 1950, so that the relative uniformity between the first and second halves of the century was not possible to judge (Figure 4.5).

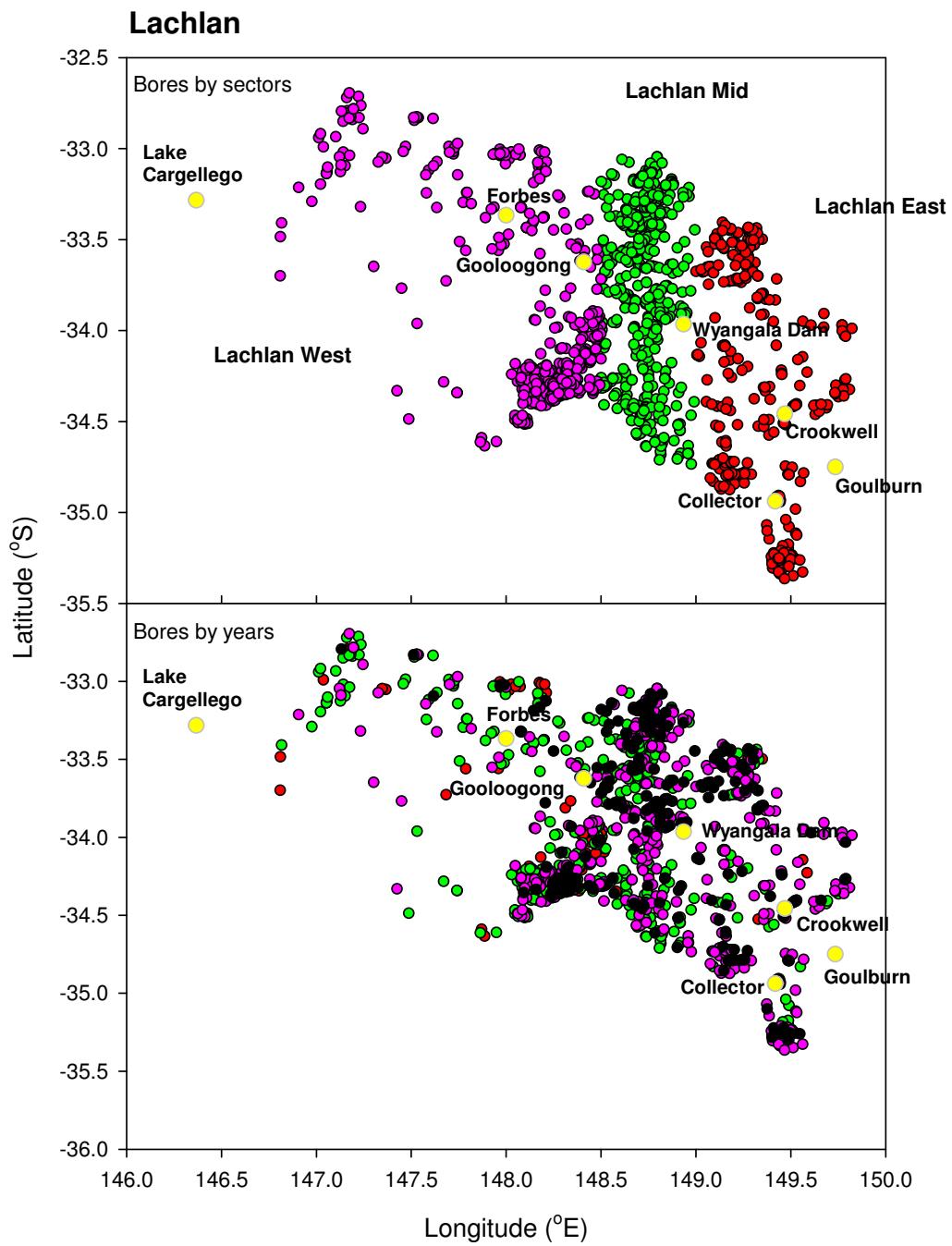


Figure 4.1: Spatial distribution of bores in Lachlan catchment. Top: Bores in Laclan West section coloured in magenta, in Mid section in green and in Eastern section in red. Bottom: Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

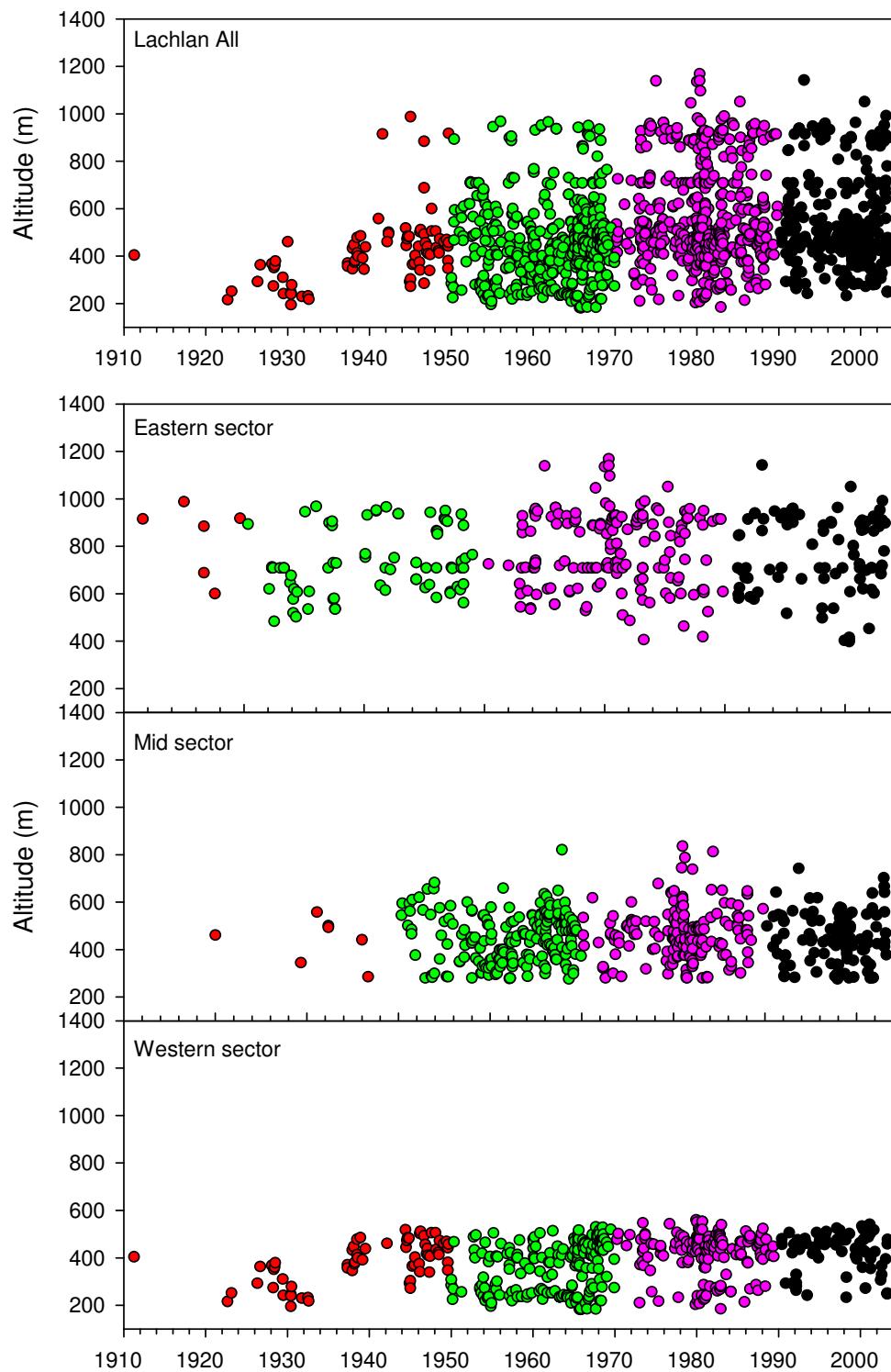


Figure 4.2: Altitude of bores in Lachlan catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

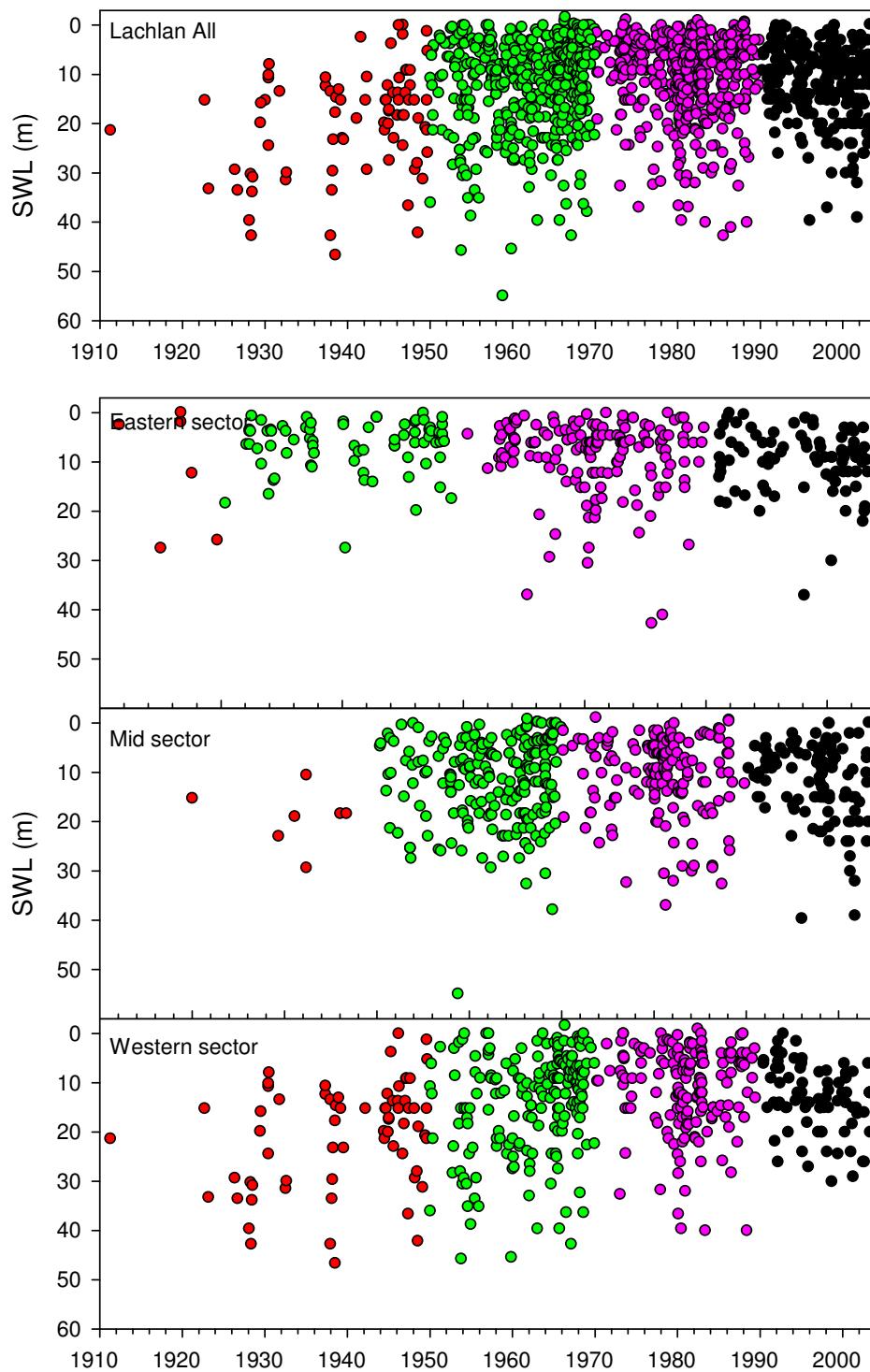


Figure 4.3: SWL of bores in Lachlan catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

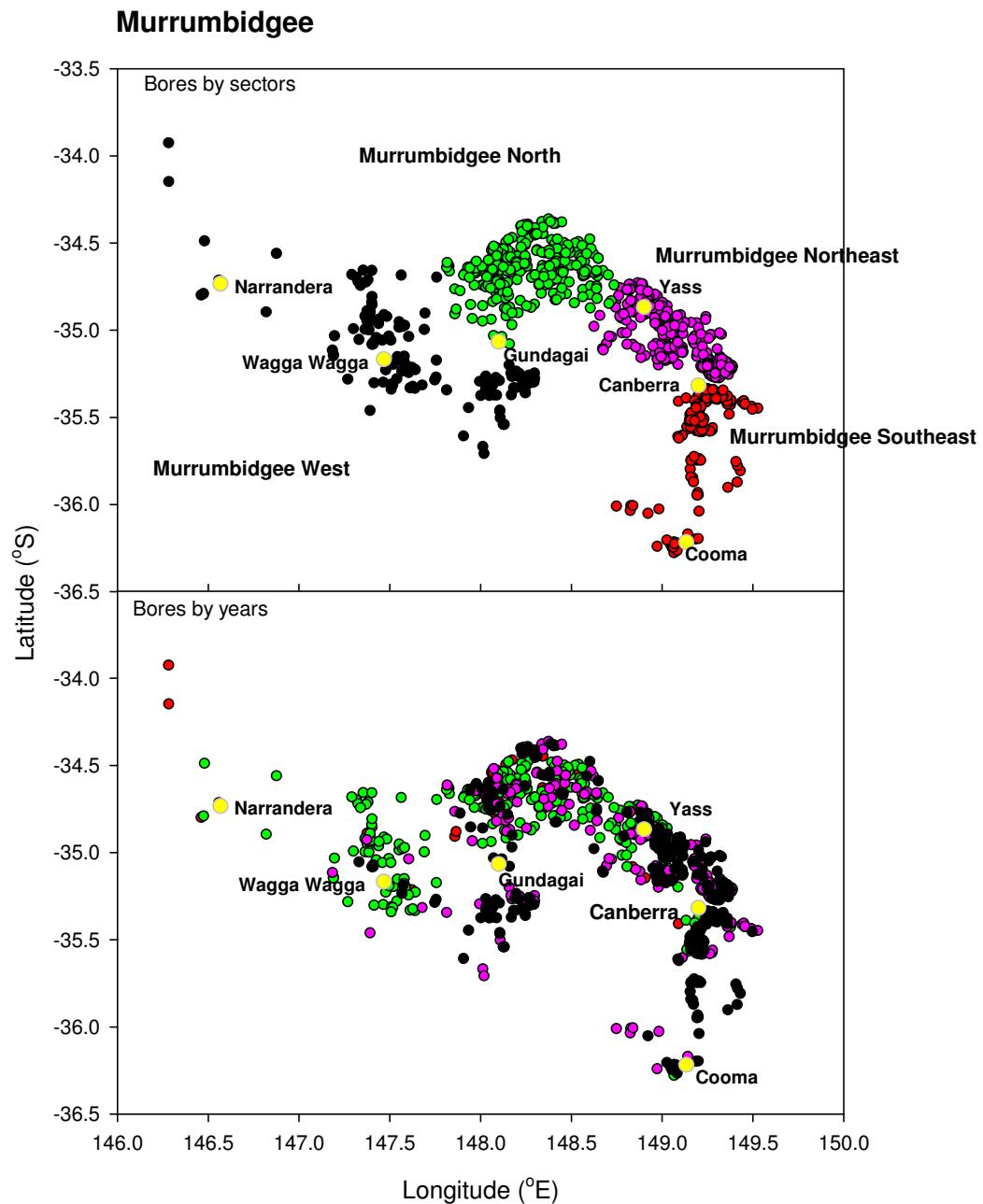


Figure 4.4: Spatial distribution of bores in Murrumbidgee catchment. Top: Bores in Murrumbidgee West section coloured in black, in North (Inner West) section in green, in Northeast section in magenta and in Southeast section in red. Bottom: Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

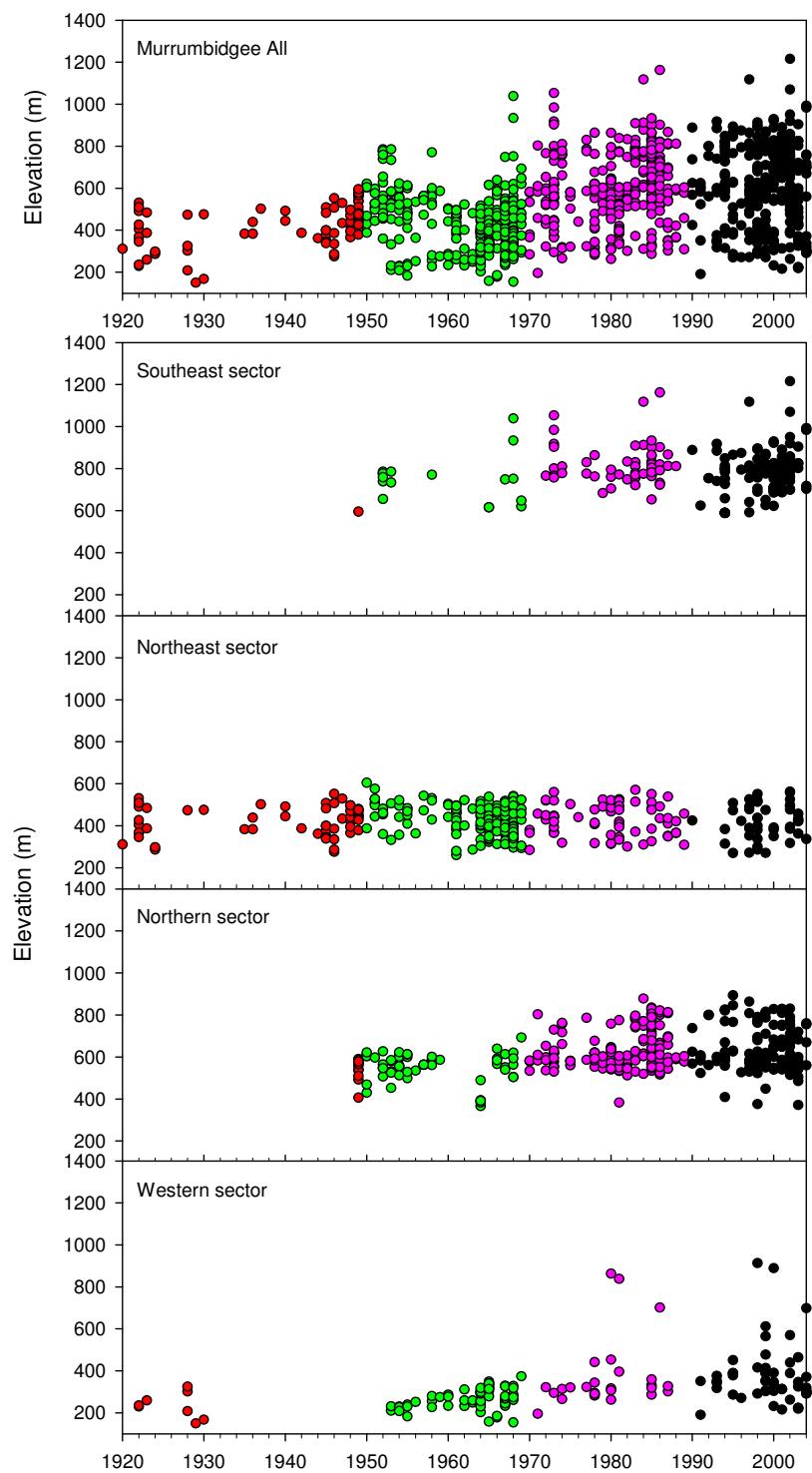


Figure 4.5: Elevation of bores in Murrumbidgee catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

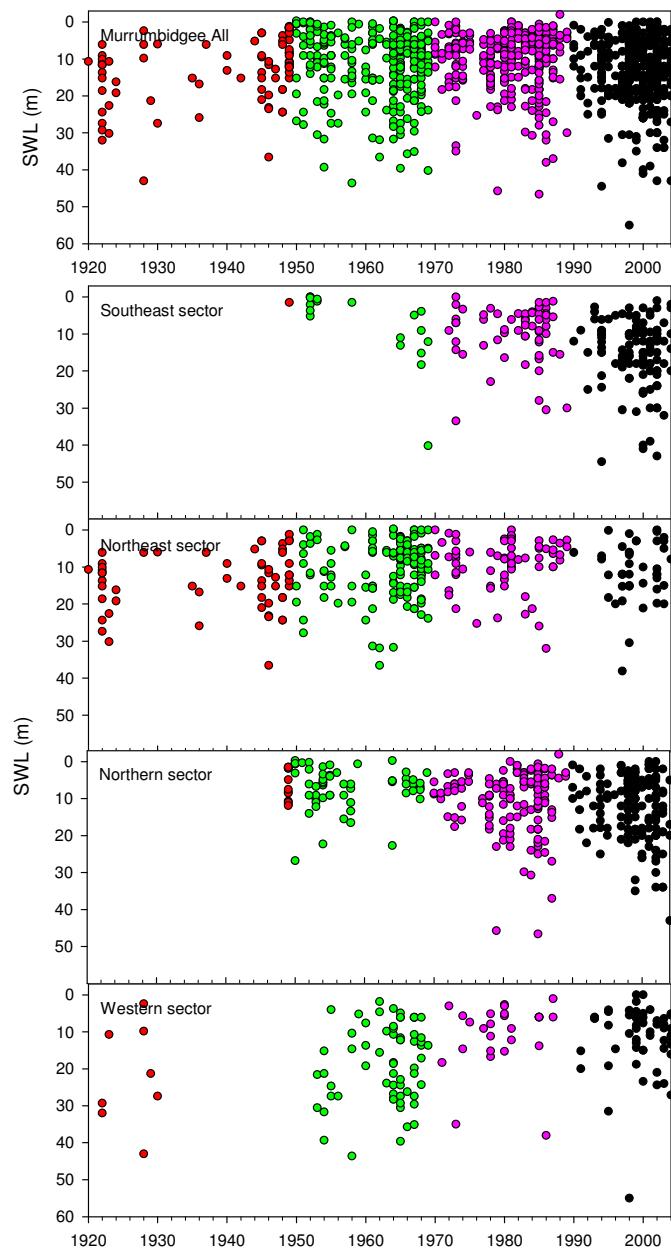


Figure 4.6: SWL of bores in Murrumbidgee catchment. Bores constructed before 1950 coloured in red, 1950-1969 in green, 1970-1989 in magenta and 1990-2004 in black.

In all of the Lachlan and Murrumbidgee West sections, there were only a few bores at higher altitudes, outside the range captured by the early bores, so mild non-uniformity was not an alarming issue.

In other sections the bores providing SWL over time were fairly uniformly distributed across all elevations.

A further dissection of any of the sections would have resulted in insufficient bore numbers and would have prevented further bore analysis. This is because the sample size would have been substantially reduced, the accuracy of the calculated annual median and twenty one year moving average would have declined and the resulting trend estimate would have diverged from the reality.

## Post-hoc segmentation

The results of study area segmentation is presented in Table 4.5.

Table 4.5: Post-hoc segmentation.

Section	Approximate geographic limits
<b>Summer rainfall zone</b>	
NE portion of Namoi	East of the boundary with the Gunnedah Basin to the crest of the Great Dividing Range
<b>Uniform rainfall zone</b>	
Lachlan East	East of 149 °E (~40 km west of Cowra) to the crest of the GDR
Lachlan Mid	Between 148.5 °E and 149 °E (from ~20 km west to ~40 km east of Cowra)
Lachlan Inner West*	Portion of Lachlan West between 148.5 °E and 146.8 °E (near Kacatoo)
Murrumbidgee East	Murrumbidgee catchment upstream of Burrinjuck Dam
Murrumbidgee South-East	Portion of Murrumbidgee East south of 35.3 °S, including areas around Canberra and Cooma
Murrumbidgee North-East	portion of Murrumbidgee east, north of 35.3 °S (Yass River Catchment and surrounding area)
<b>Winter rainfall zone</b>	
Murrumbidgee West	Entire contributing catchment within the study area downstream of Burrinjuck Dam
Murrumbidgee Inner West* (Northern)	Subset of Murrumbidgee west between Burrinjuck Dam and 147.8 °E (5 km east of Tarcutta) and north of (35.1 °S near South Gundagai)

\*additional subsection in which only a part of already formed section is analysed: Lachlan Inner West is part of Lachlan West; Murrumbidgee East is split into South East and North East.

### 4.2.3 Noise filtration

The description of the methodology can be found in Chapter 3. The development of the method and its application are illustrated using Namoi section example in Sections 3.2.2 and 3.2.3. The results of the moving average filtering technique across the sections of the Lachlan and the Murrumbidgee catchments are presented in Figures 4.7, 4.8 and 4.9.

Reconstructed SWL trends demonstrate the rise in water table from deeper to shallower depth after the mid of the twentieth century, persistence of quasi-equilibrium condition at this shallower depth and the mild deepening of the water table towards the end of the century. The equilibrium conditions, that followed the water table rise, prevailed during the 1970s and 1980s in the mid and western portions of the Lachlan and Murrumbidgee sections (Figures 4.7, 4.8 and 4.9). The water table rise started and terminated earlier in the Namoi catchment (Section 3.2.2 Figures 3.3 and 3.4), and eastern sections of the Lachlan and Murrumbidgee catchments (Figures 4.7 and 4.9), which were closer to the topographic high of the boundary imposed by the Great Divide.

In Chapter 5, trends in SWL are plotted against the corresponding rainfall trends.

The residual mass curves of annual SWL time-series in Lachlan and Murrumbidgee catchments are presented in Figure 4.10, Figure 4.11 and Figure 4.12, respectively.

The plots demonstrate that the minimum SWL occur shortly after the middle of the twentieth century, indicating an abrupt rise in the water table. The precise timing and the significance of this event was further investigated using change-point analysis (Table 4.6).

Reducing the level of significance to a confidence level of 90% detected change-points in SWL in all of the catchments and most of the sections. The same analysis performed on the rainfall data (Table 4.7 and Appendix, Table A), detected the climate shift in 1947 from dry to wet hydrologic regime in the summer and uniform rainfall zones of the study area. The rise in water table across the study area affected by the rainfall shift, lags behind the rainfall signal by: two and eight years in the Namoi section, five years in the Lachlan East section and 16 years in the Lachlan West section (Table 4.6). The bore dataset from the Murrumbidgee East section does not start early enough to capture the impact of climate shift on water

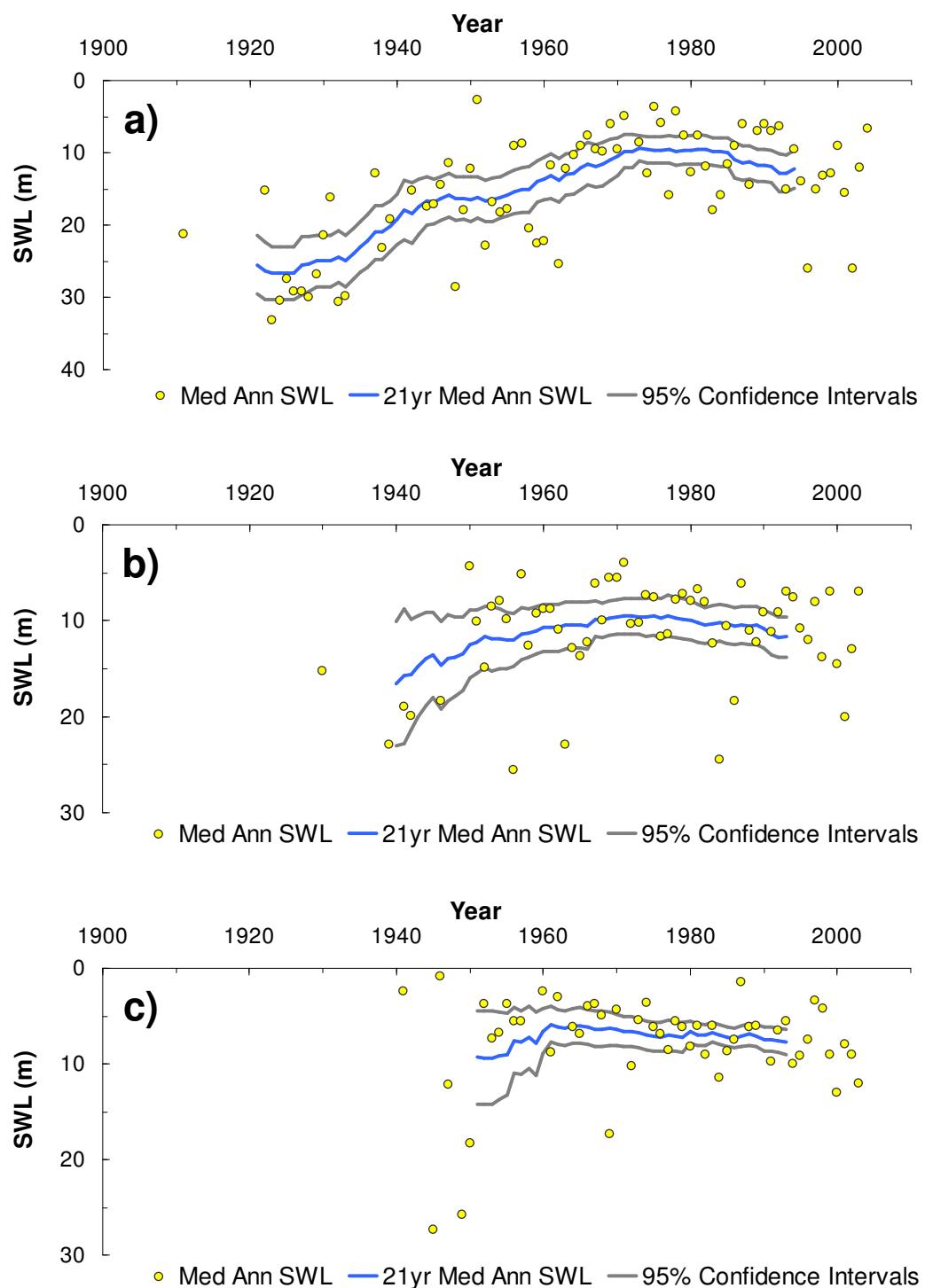
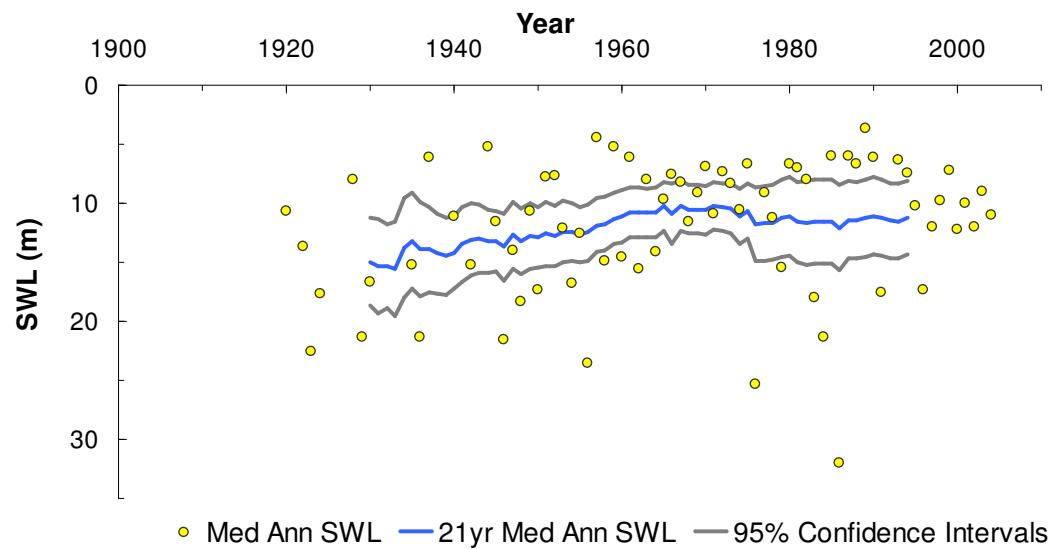


Figure 4.7: Water table trend in Lachlan: a) West, b) Mid and c) East section, in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL).

### a) Murrumbidgee West



### b) Murrumbidgee Inner West

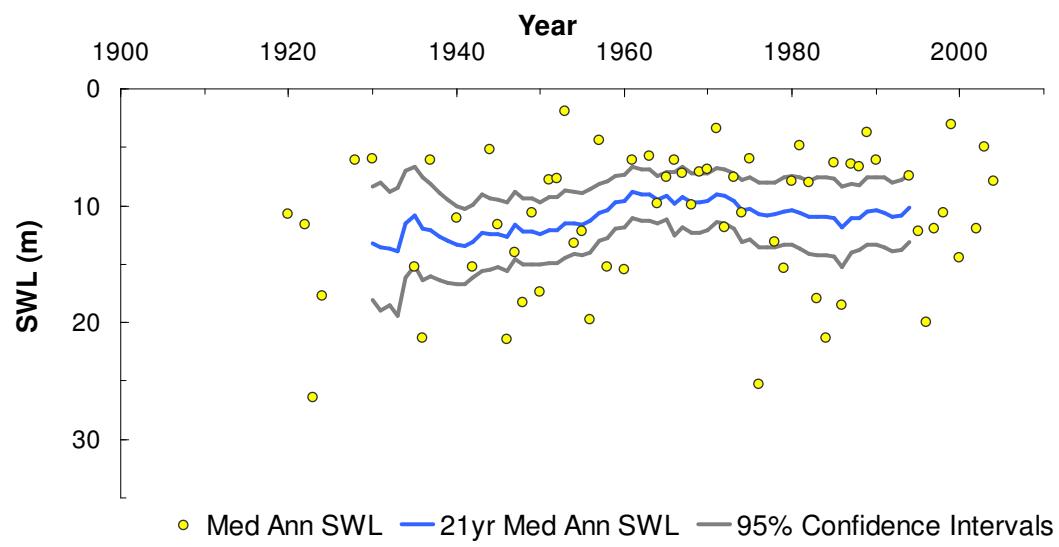


Figure 4.8: Water table trend in Murrumbidgee West and Inner West sections in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL).

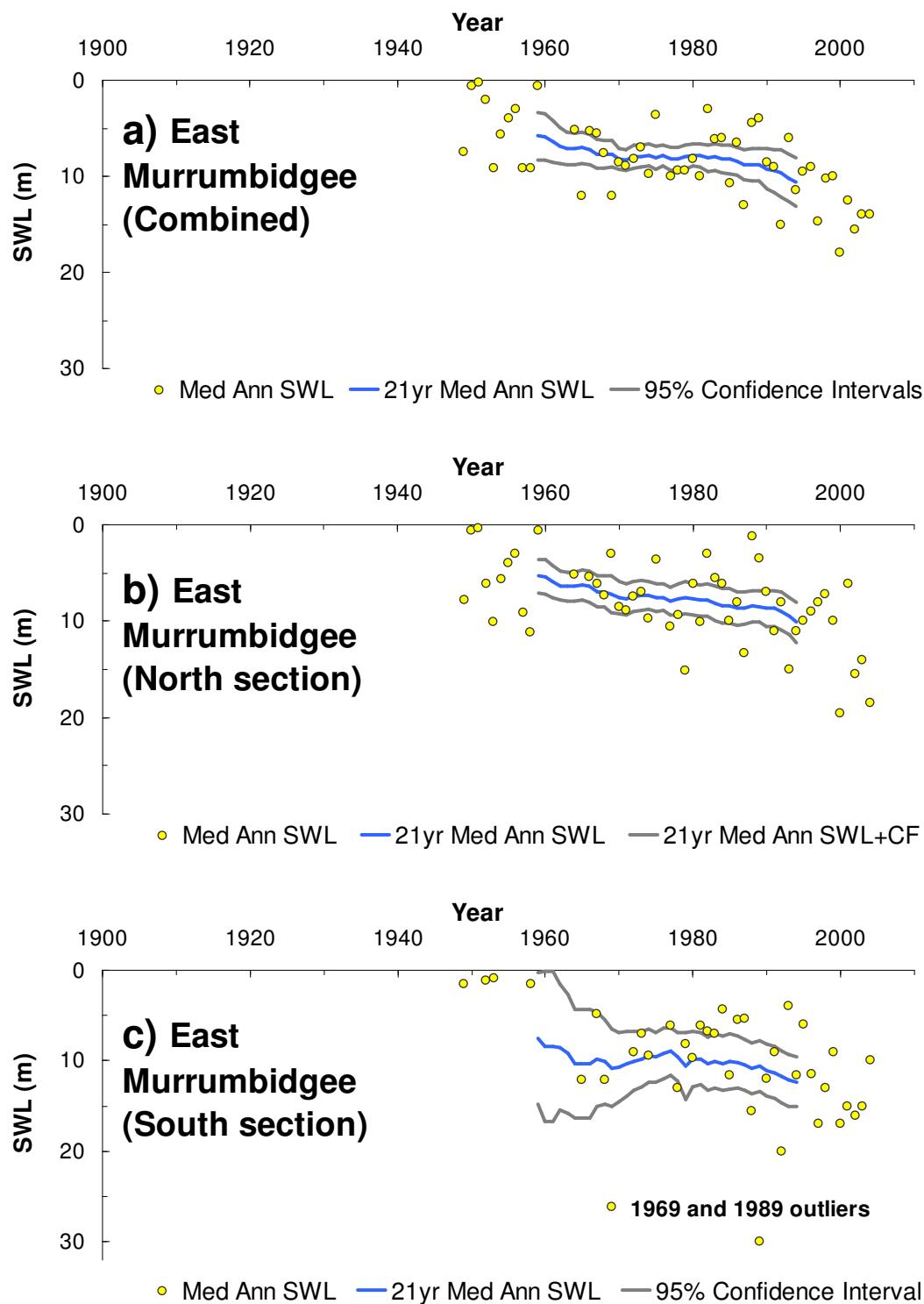


Figure 4.9: Water table trend in Murrumbidgee East a) combined and b) North and c) South portion in a form of the 21 year moving average filter (21yr Med Ann SWL) of median annual SWL (Med Ann SWL).

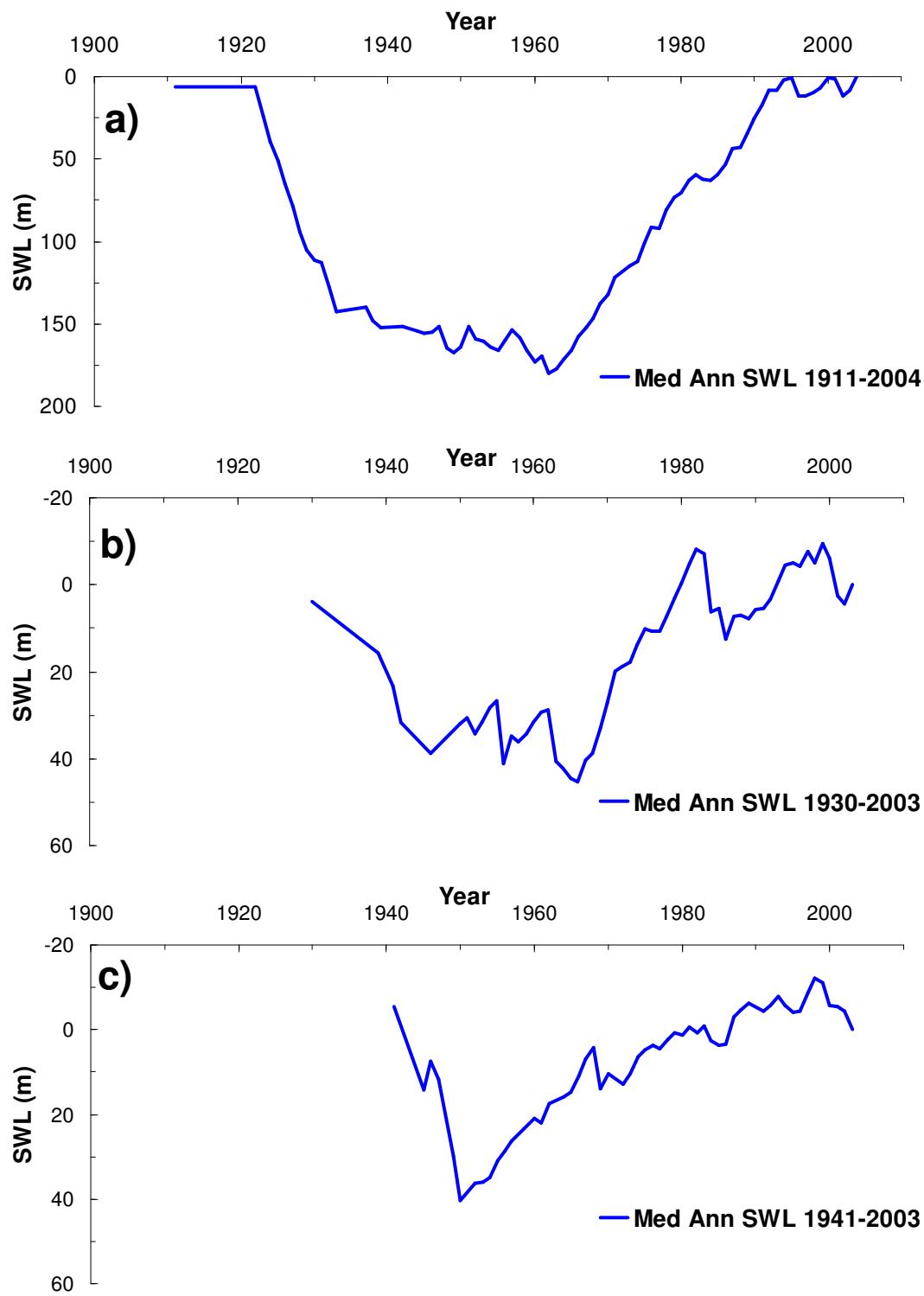


Figure 4.10: Residual mass curve of median annual SWL (Med Ann SWL) in Lachlan: a) West, b) Mid and c) East section.

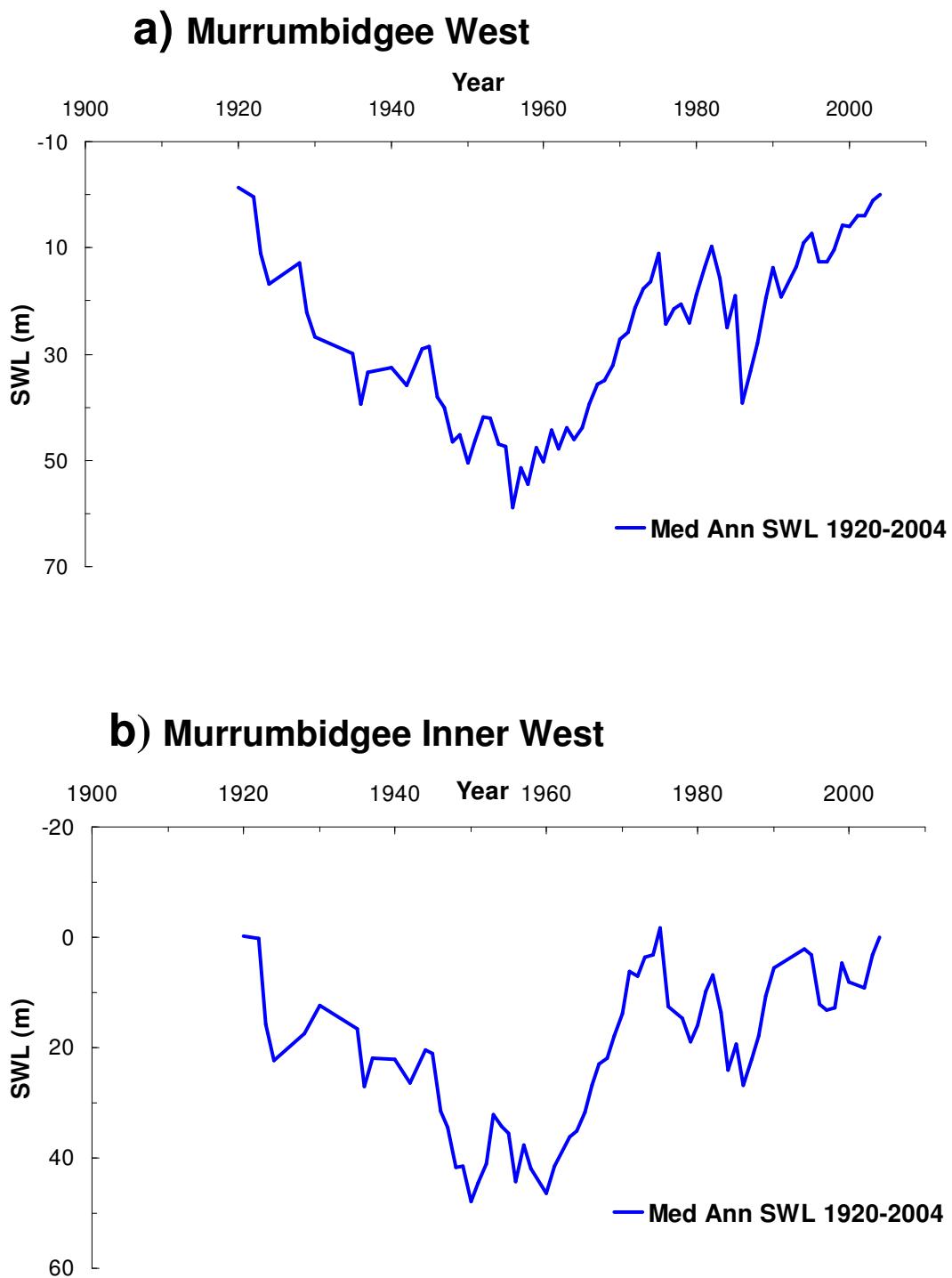


Figure 4.11: Residual mass curve of median annual SWL (Med Ann SWL) in Murrumbidgee West section and its Inner West portion.

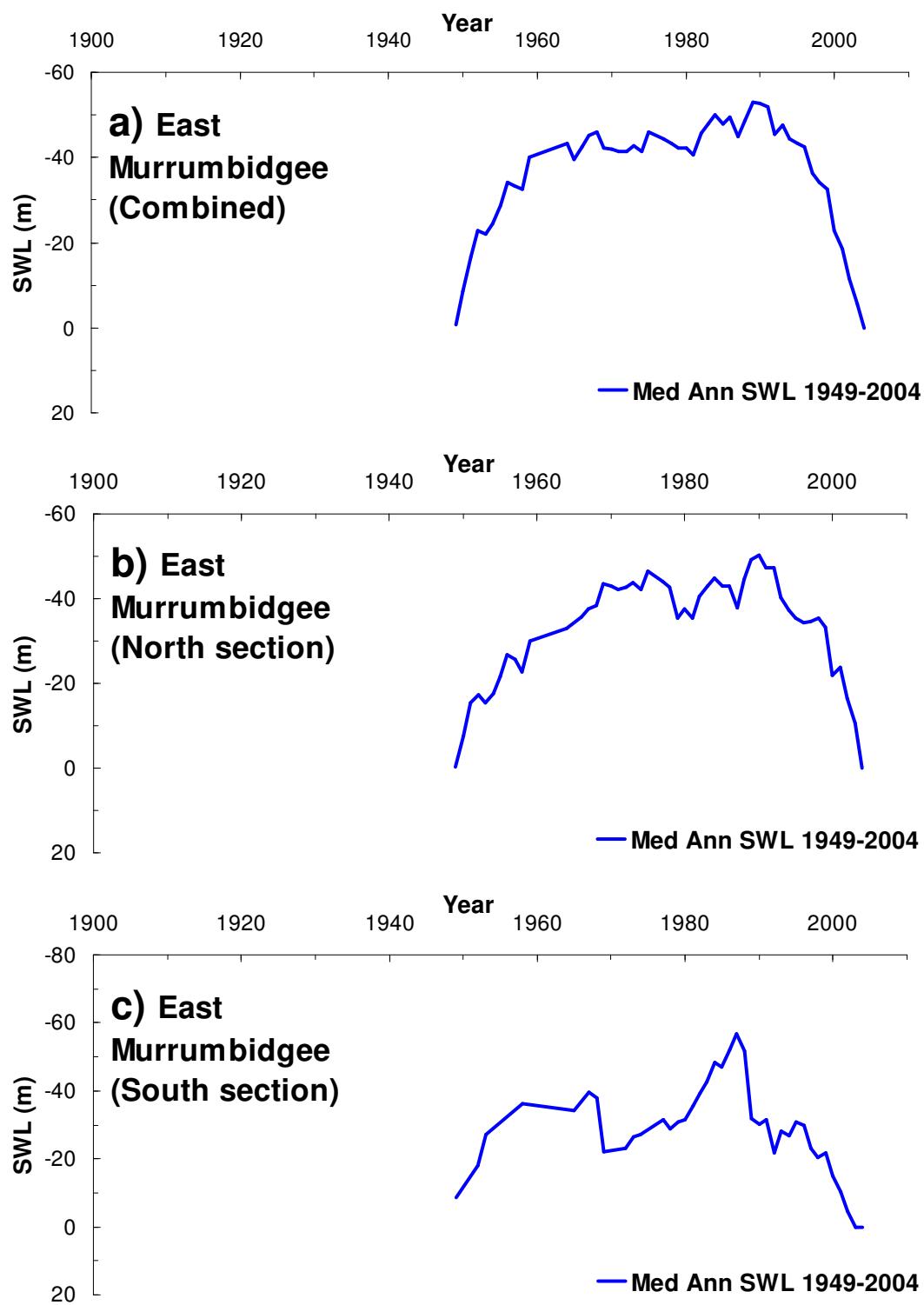


Figure 4.12: Residual mass curve of median annual SWL (Med Ann SWL) in Murrumbidgee East section and its northern and southern portions.

Table 4.6: Change-point analysis of SWL data.

Section/subsection	Year change occurred	Confidence level	From (m)	To (m)
Namoi	1949	>99.5%	23.2	15.9
	1955	95%	15.9	11.6
Lachlan East	1952	92%	14.5	6.9*
Lachlan Mid	no change			
Lachlan West	1937	>99.5%	26.2	16.9
	1963	>99.5%	16.9	9.2
	1993	97%	9.2	14.6
Murrumbidgee East	1965	93%	4.7	8.3
	1997	>99.5%	8.3	13.6
Murrumbidgee West	1937	94%	21.1	11.4

\*Analysis of ranked data showed no significant change.

Confidence level >90%; bootstraps = 1000.

table (Figures 4.9 and 4.12). The Lachlan Mid section demonstrates the rise in water table, (Figure 4.7), but the minimum of the residual mass curve in 1966 (Figure 4.10) fails to reach the threshold of 90% for confidence levels. The rainfall shift is not pronounced in the winter rainfall zone, (Table 4.7 and Appendix Table A) and its absence is mirrored by the same absence of SWL shift in the Murrumbidgee West section. However, the increase in rainfall over the twentieth century is present in the Murrumbidgee West section, and it is consistent with the rise in water table (Figure 4.8). The abrupt rise in SWL in the Lachlan West section in 1937 is interpreted and discussed in Chapter 5.

Table 4.7: Change-point analysis of composite rainfall series.

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
Namoi	1947	97%	647.8	726.7
Lachlan East	1947	99%	668.6	775.4
Lachlan Mid	1895	99%	650.9	568.5
	1947	98%	568.5	668.7
Lachlan West	1947	97%	569.9	648.3
Murrumbidgee East (North)	1947	96%	638.0	720.9
Murrumbidgee East (South)	1947	>99.5%	562.7	635.5
Murrumbidgee	1895	96%	625.7	508.7
West	1916	94%	508.7	632.4

Confidence level of >90% was used to indicate a significant change.

### 4.3 Summary

The application of the bore data mining method across the sections of the study area was presented in this Chapter. The bores in fractured rock with the water bearing zones closest to the ground surface were selected from the database. Median value of SWL was calculated for each year to create the composite SWL hydrograph. Filtering techniques were applied to eliminate the noise in the series. These results are discussed and validated in Chapter 5.

# **Chapter 5**

## **Validation**

In this chapter the model results are validated. Model validation generally relies on comparison of modelled and observed data. However, as the recorded hydrographs in the study area are of a very short duration, they are not sufficient for validation of the modelling results. To overcome this constraint, smoothed rainfall data are used as a surrogate dataset for validation in conjunction with the existing recorded SWL hydrographs.

### **5.1 Validation of aquifer behaviour based on recorded hydrographs**

The results are validated using the longest available hydrograph data from the fractured-rock geology in the NSW (Figure 5.1 and Figure 5.2):

- the Lachlan catchment (Wattle Retreat) from the mid section of the Lachlan catchment (Figure 5.3) and
- the Murrumbidgee catchment: Yass catchment in the north-eastern section of Murrumbidgee (Figure 5.4) and the Tarcutta Creek catchment (Figure 5.5) from the western portion of Murrumbidgee catchment.

Figures 5.3a and 5.3b demonstrate that the hydrograph of the bore GW036748-2, recorded at Wattle Retreat closely follows the modelled trendline. Other hydrographs from the site exhibit similar patterns with a downward trend (Figure 5.3c), caused by diminishing rainfall. Hydrographs of the piezometers GW036753

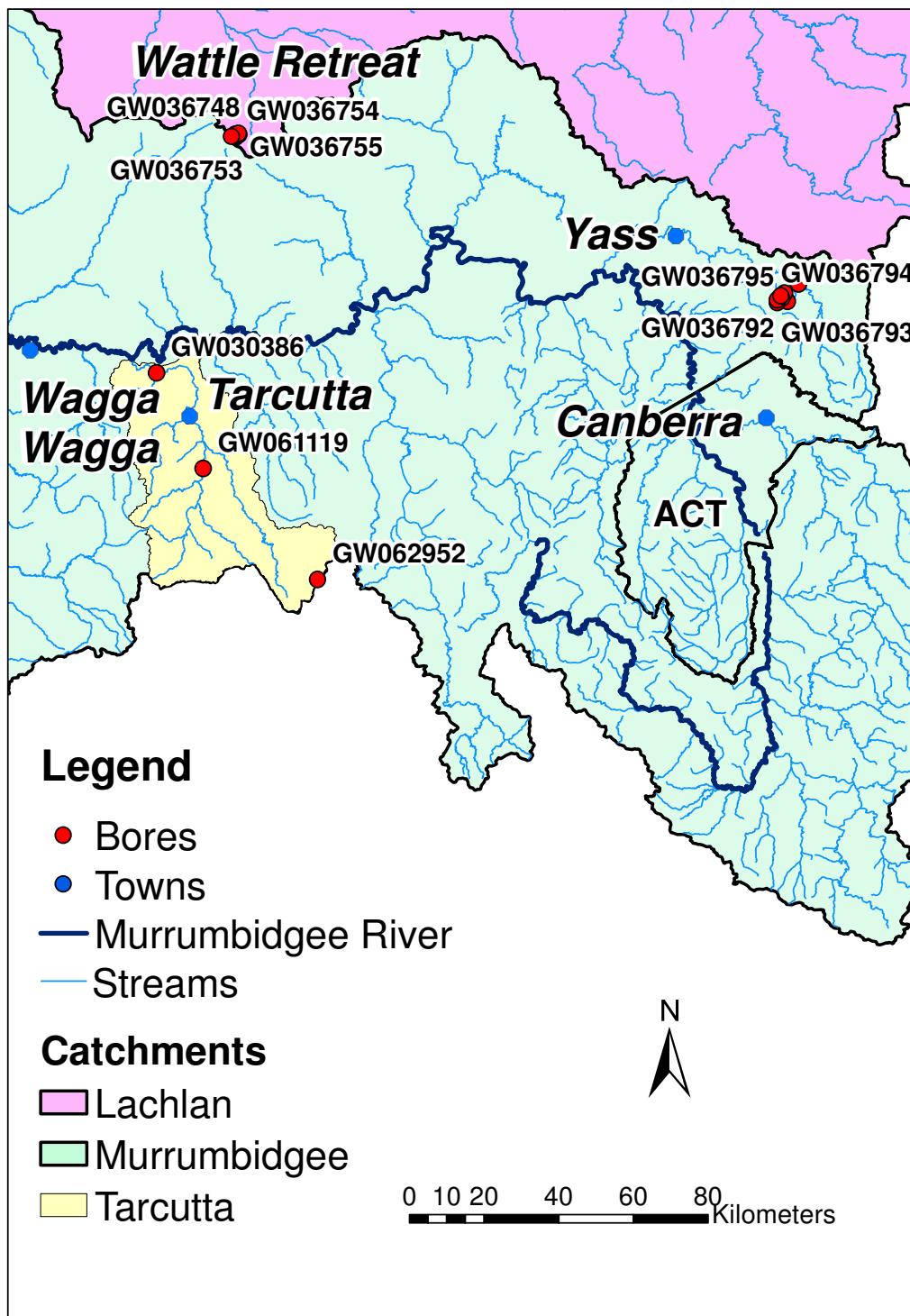


Figure 5.1: Location of bores used for model validation.

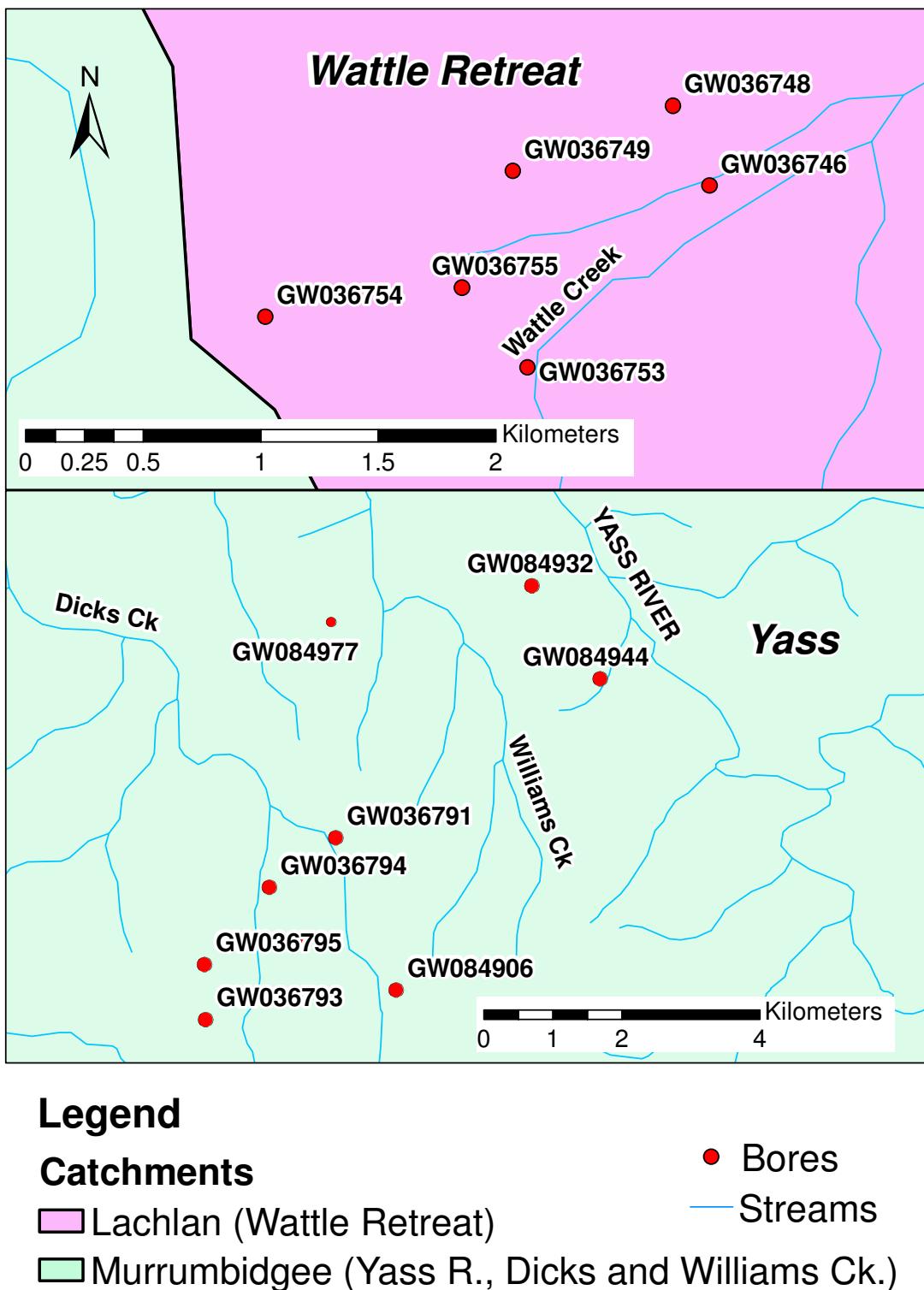


Figure 5.2: Monitoring bores in Wattle Retreat and Yass catchments.

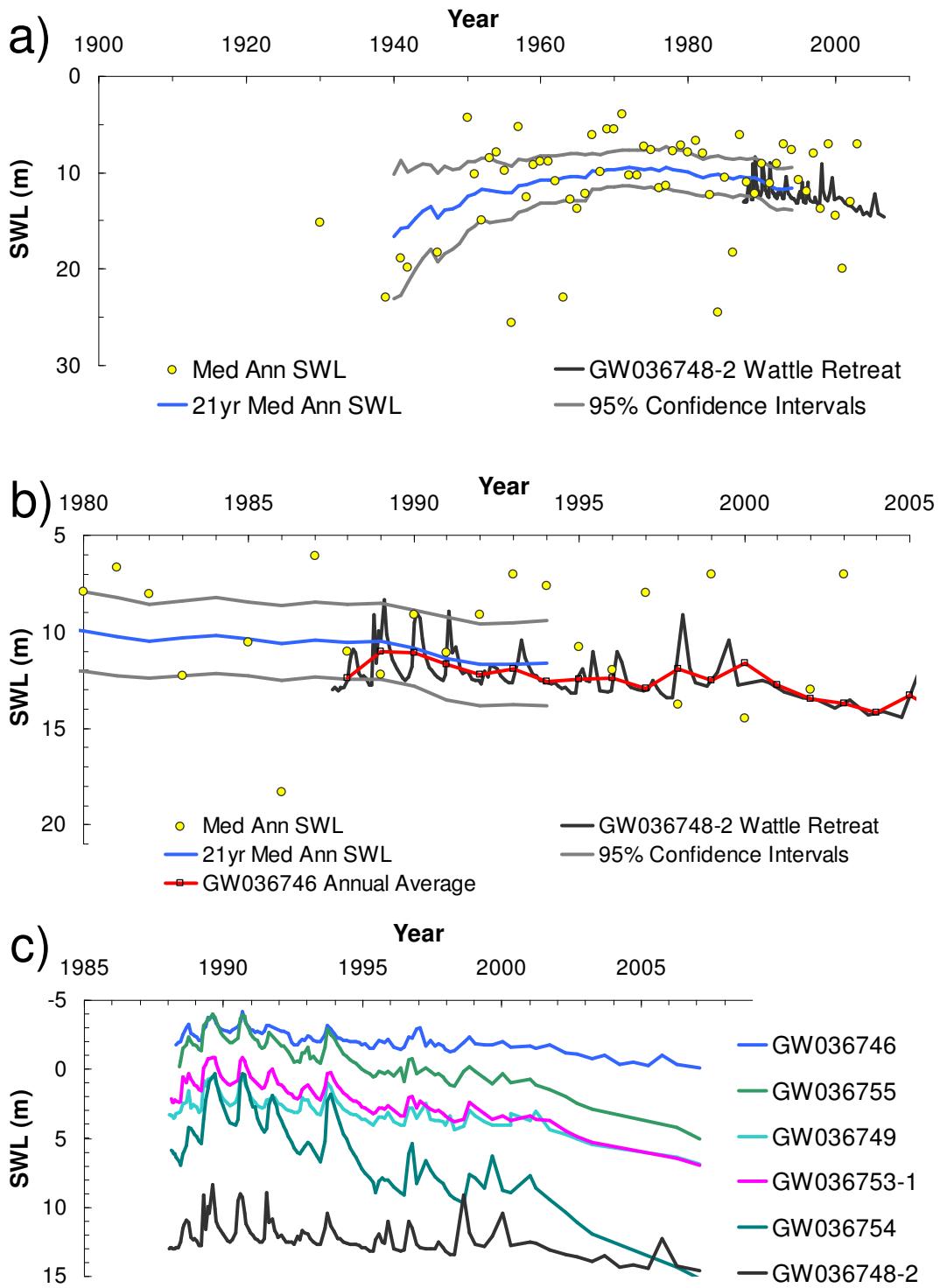


Figure 5.3: Modelled SWL trend (21 yr Med Ann SWL) in Lachlan Mid section matches a) recorded hydrograph of bore GW036478-2 and b) its annual average. c) Wattle Retreat hydrographs exhibit falling tend. Med Ann SWL represents Median Annual SWL.

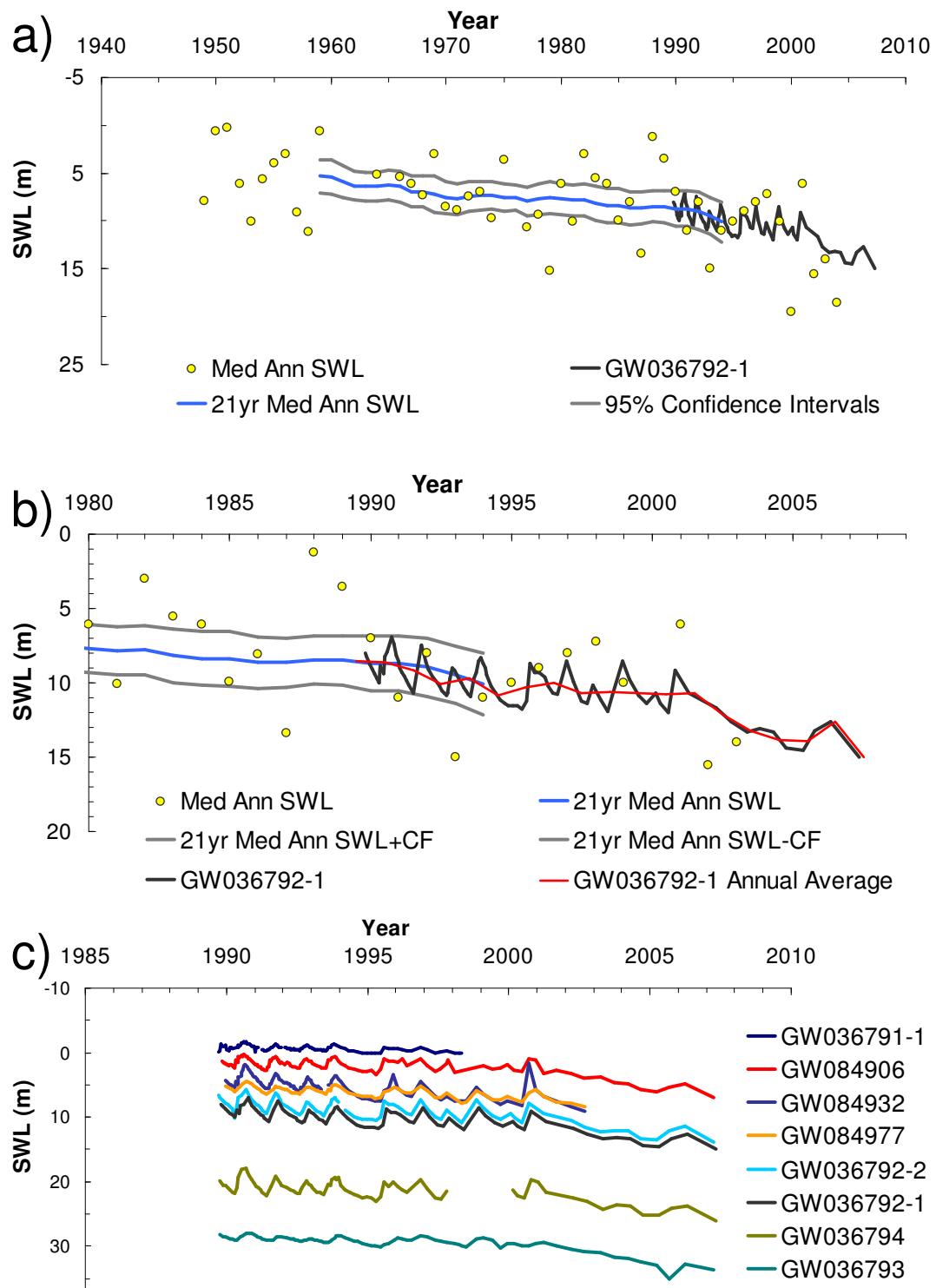


Figure 5.4: Modelled SWL trend (21yr Med Ann SWL) in North portion of Murrumbidgee East section matches a) recorded hydrograph of bore GW036792-1 and b) its annual average. c) Hydrographs in Yass catchment exhibit falling tend. Med Ann SWL represents Median Annual SWL.

and GW036749 are closer to the land surface because these bores are positioned closer to the discharge area than the bore GW036748. Bores at the discharge areas are under artesian conditions during the wettest period (GW036746, GW036755, GW036753-1). Variability of all hydrographs diminishes towards the end of the record. This is due to the compounding effect of: a) dryer climatic conditions and b) gradual reduction in frequency of observations from monthly (1988-1994) to annual (2006-2007).

Similar conclusions can be drawn for the bore hydrographs in the Yass catchment from the Figure 5.4. Figures 5.4a and 5.4b again demonstrate that the recorded hydrograph closely follows the modelled trendline, and the other hydrographs from the site (Figure 5.4c) also exhibit similar patterns with a downward trend, caused by diminishing rainfall. Hydrographs differ in their annual amplitude. This is caused by the signal being attenuated through energy dissipation along its pathway. For example, the annual amplitude of deeper bores, such as bore GW036793 close to the Quarry Lake on the Picarree Hill and bores GW084906 and GW036791-1 at the discharge areas are smaller than the amplitude at the remaining bores. Even though hydrographs differ in the amplitude of annual variation and base level, the imprint of drying climate is consistently reflected in the downward trend across the entire range of presented hydrographs. Deep piezometers GW084906, GW085932 and GW084977 were drilled to the contact with the rock. Even though their intake is partially in unconsolidated sediments above the fractured rock, they are not in an independent, perched aquifer, but under influence of the fractured rock aquifer, so their behaviour is consistent with the behaviour of the bores with an intake in fractured rock.

Piezometers and bores drilled in fractured rock in the Yass area and Wattle Retreat catchment were part of the early salinity investigations in NSW, prompted by concern that the water table would rise, causing the expansion of water-logged areas and salinity. Monitoring of most of these bores stopped in 2007, when it was clear that the water table was falling, apart from a few deeper bores, which were re-surveyed in 2012.

Some private bores in fractured rock were monitored with irregular intervals between the 1990s and 2006, such as GW062952 and GW061119 in the upper and mid Tarbutta Creek Catchment. Depth to water table at the time of their construction was appended to the monitoring record of these two bores (Figure 5.5). Apart from

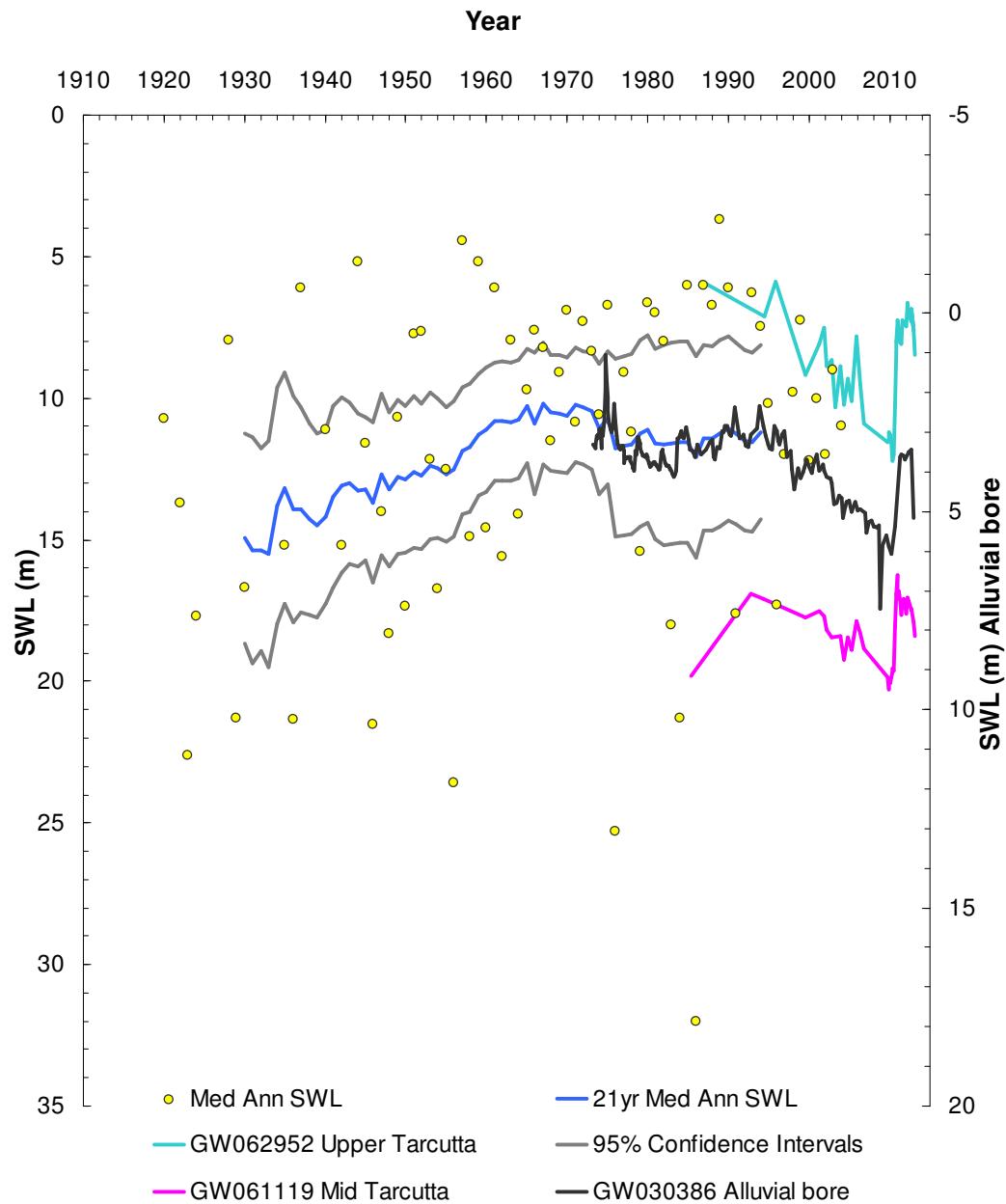


Figure 5.5: Modelled SWL trend (21 yr Med Ann SWL) in Murrumbidgee West section matches recorded hydrograph of bore GW030386 from Tarcutta Creek alluvial, which replicates trends captured by GW061119 and GW062952 from mid and upper Tarcutta Catchment. Med Ann SWL represents Median Annual SWL.

the beginning of hydrographs' records, which, due to the infrequent observations exhibit variation, the two hydrographs share common behavior (Figure 5.5). Monitoring of these two bores together with ca. 60 bores and piezometers in Tarcutta Creek Catchment re-commenced in 2009 through Future Farm Industries Cooperative Research Centre (FFI CRC) [Rančić et al., 2011]. It captured the end of the Millennium Drought in 2010 and abrupt shift to wet conditions, triggered by a strong La Niña event. The third hydrograph presented in Figure 5.5 belongs to the alluvial bore GW030386-1, which had been monitored initially on a monthly basis and lately less frequently. Even though this is not a fractured rock bore, Figure 5.5 clearly demonstrates that it behaves in a similar manner during the overlapping period of observations as the presented fractured rock bores. The hydrograph of this alluvial bore demonstrates:

- quasi-equilibrium conditions that lasted through 1970s, to mid 1990s;
- consistent deepening of water table since the mid 1990s as a consequence of the Millennium drought;
- sudden rise in water table caused by high rainfall, brought about by La Niña in 2009; and
- the recession in 2012, after La Niña ceased.

As this alluvial bore demonstrates common behaviour with the fractured rock bores in the area, we can infer commonality in response to climate of fractured rock and this particular bore from the alluvial aquifer in the Tarcutta Creek Catchment. If we assume that this commonality of aquifer response existed since the beginning of the alluvial bore record in 1973, the hydrograph of the bore GW030386 would allow validation over two decades.

Root mean square error is calculated between normalised modeled trends (21 year moving average) and corresponding bore hydrographs for all three sites: Wattle Retreat, Yass and Tarcutta (Table 5.1). Before the SWLs were normalised, annual SWLs were calculated for each bore using the average of all observations during a year. The table shows the closest fit between the modeled trend and annual hydrographs derived from recorded data for Yass bores (RMSE between 0.89 and 1.04), followed by Tarcutta bore (RMSE 1.02) and Wattle Retreat bores (RMSE between 1.06 and 1.27). Tighter fit in Yass area correlates with higher density

Table 5.1: Root Mean Square Error (RMSE) between normalised SWL trends and annual timeseries of monitored hydrographs.

	RMSE	n	Site
GW036746	1.14	7	Wattle Retreat
GW036748-2	1.06	7	
GW036749	1.18	7	
GW036753-1	1.21	7	
GW036754	1.27	7	
GW036755	1.17	7	
GW036792-2	0.89	6	Yass
GW036793	1.04	6	
GW036794-2	0.89	6	
GW084906	1.04	6	
GW084932	0.89	6	
GW084977	1.03	6	
GW030386-1	1.02	22	Tarcutta

n=sample size

of bores which were used to derive long-term groundwater trend for North-East Murrumbidgee.

The annual monitoring data were compared to the modeled trend, instead of trend-to-trend comparison, because of the short monitoring record. Short term variability captured by the annual data inflates RMSE, i.e. RMSE in the trend-to-trend evaluation would have been lower.

The provided examples of bore hydrographs are in excellent agreement with the modeled SWL trends in all three regions where these rare records are available. The evidence from monitoring sites suggests a very strong climate driver, which forces the aquifers of each area to produce hydrographs with consistent patterns. This climate forcing of the aquifers is exploited to further validate results, by comparing SWL trends with trends in rainfall over the entire modelling period.

## 5.2 Validation of aquifer behaviour using rainfall as the surrogate test data

Comparison of monitoring hydrographs with modeled SWL trends validated the model prediction and suggested a strong climate forcing of water table fluctuations. This climate forcing was exploited to further validate the model, for the entire prediction time-span.

The results are validated using composite standing water level and composite rainfall data from Namoi and Lachlan catchments and the corresponding sections (Figure 1.2). The graphical representation of the results for the Murrumbidgee catchment are omitted, for brevity. They can be found in Rančić et al. [2009], together with the graphical representation of groundwater and rainfall trends of other fractured rock areas of NSW part of the Murray-Darling Basin.

As with the recorded hydrographs, trends in form of the 21-year moving average of standing water level and rainfall are used for model validation. Use of splines as an alternative filter is tested on the Namoi catchment.

Long validation period captured sudden increase in rainfall reported to be around the middle of the century. This allowed for additional validation through change point analysis: timing of the shift from dry to wet hydrologic regime of standing water level and rainfall and graphical representation of this event through residual mass curves.

Graphical comparison of trends is accompanied with Root Mean Square Error (RMSE) evaluation. Groundwater and rainfall trends are normalised, so that the RMSE test can be used for evaluation. Rainfall and SWL are negatively correlated (more rainfall produces shallower, smaller SWL). Therefore, the sign of the normalised SWL series is inverted before calculation of RMSE. Confidence of abrupt shift in the regime is evaluated using bootstrapping, as a part of change-point analysis.

Most of the validation of SWL trend in Namoi section was presented in Chapter 3, where the methodology was described:

- Graphs of 21-year moving average filters of rainfall and SWL and corresponding normalised trends are presented in Figure 3.4. RMSE of 0.24 is presented in Table 3.2;

- Weighted splines derived for Namoi catchment are presented in Figure 3.5; RMSE between 21 year moving average of rainfall and weighted spline with 7 degrees of freedom is 0.25; For equivalent unweighted spline RMSE is 0.27; Count=70; This demonstrates that simple and robust moving average technique is not inferior to splines; Following the principle of parsimony it is therefore recommended;
- Residual mass curves are presented in Figure 3.7, and
- Change-point analysis shows climate shift in 1947 (Table 3.5) and detects two-fold change in SWL in Namoi: in 1949 and 1955 (Table 3.4).

The results for the Lachlan catchment are presented in Figure 5.6 and Figure 5.7. They demonstrate that the trend in SWL is following the rainfall trend in Eastern and Mid sections of the Lachlan River catchment with 5 and 11 years of delay. Rainfall trends are therefore shifted forward by 5 and 11 years respectively, so the fit can be evaluated. In these two sections the shifted rainfall trends are fully contained within the confidence intervals around SWL trendline. The RMSE between normalised SWL and rainfall trends is 0.57 (n=43) for the Lachlan East section and 0.53 (n=54) in the Lachlan Mid section.

In Lachlan West section the groundwater trend delays by 16 years behind the trend in rainfall for most of the period. The exception is divergence between the trendlines in the early record, suggesting the presence of a driver unrelated to climate, which caused additional recharge and a steeper rise in water table. This divergence is consistent with the increased recharge, due to advancement of cropping and associated clearing in the Murray-Darling Basin wheatbelt area, documented and analysed by Bedward et al. [2007]. The match from 1944 between delayed rainfall and SWL trends is excellent, producing a very low RMSE of 0.28 (n=51).

Figure 5.7 confirms delayed SWL response to rainfall in the Lachlan catchment. The delay increases from East to West, with the distance from the Great Dividing Range. The minima of the residual mass curves determine the timings when rainfall abruptly increased and SWL rose and underpins change-point analysis. The results of the SWL time-series change-point analysis are given in Table 4.6. The results of the change-point analysis of the weighted-average annual rainfall time series are given in Table 4.7 for each catchment section. The results of the analyses of indi-

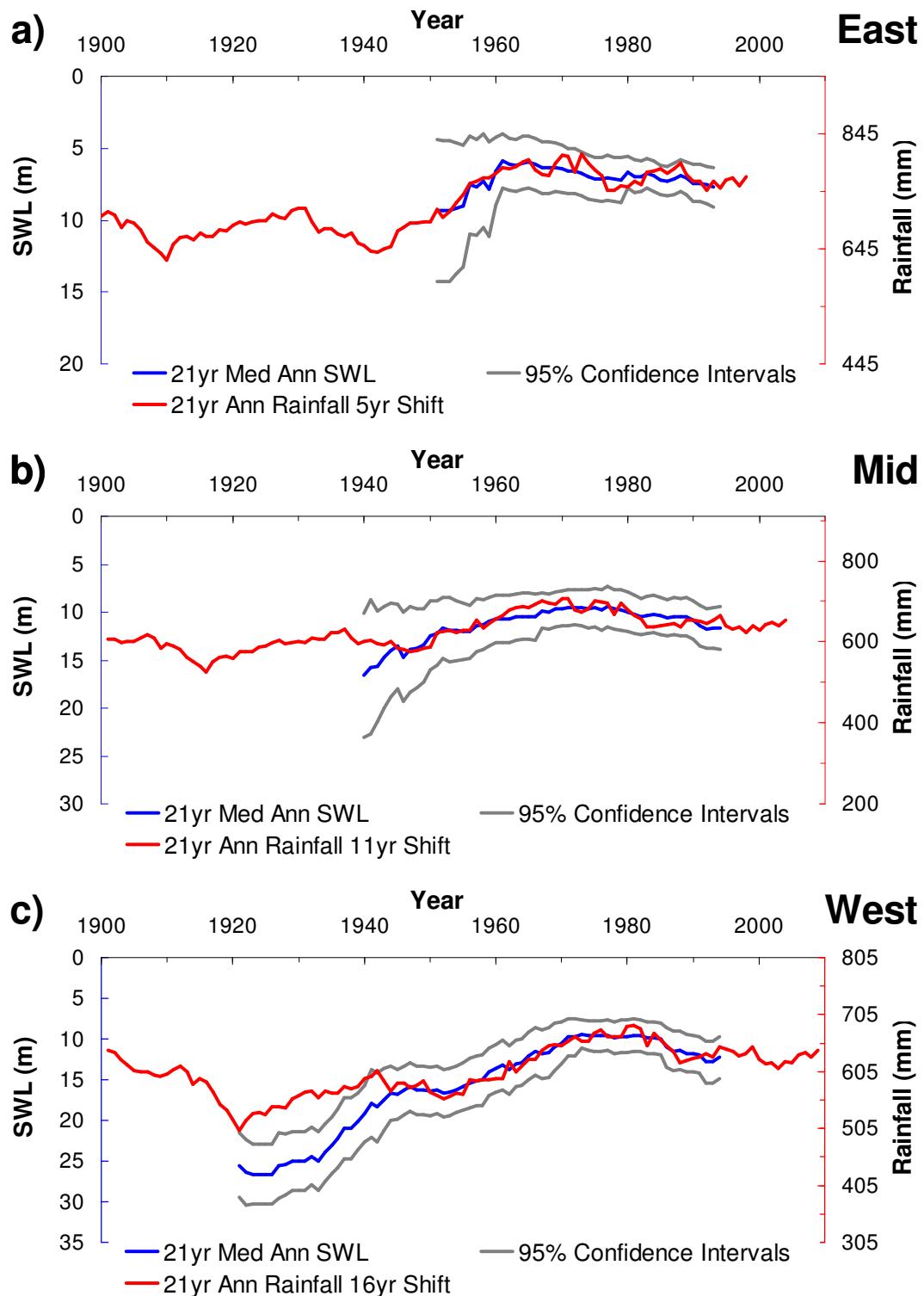


Figure 5.6: Trends in rainfall (21yr Ann Rain) shifted forward by a) 5 b) 11 and c) 16 years, compared to trends in SWL (21yr Med Ann SWL) and corresponding confidence intervals in the East, Mid and West sections of the Lachlan River catchment.

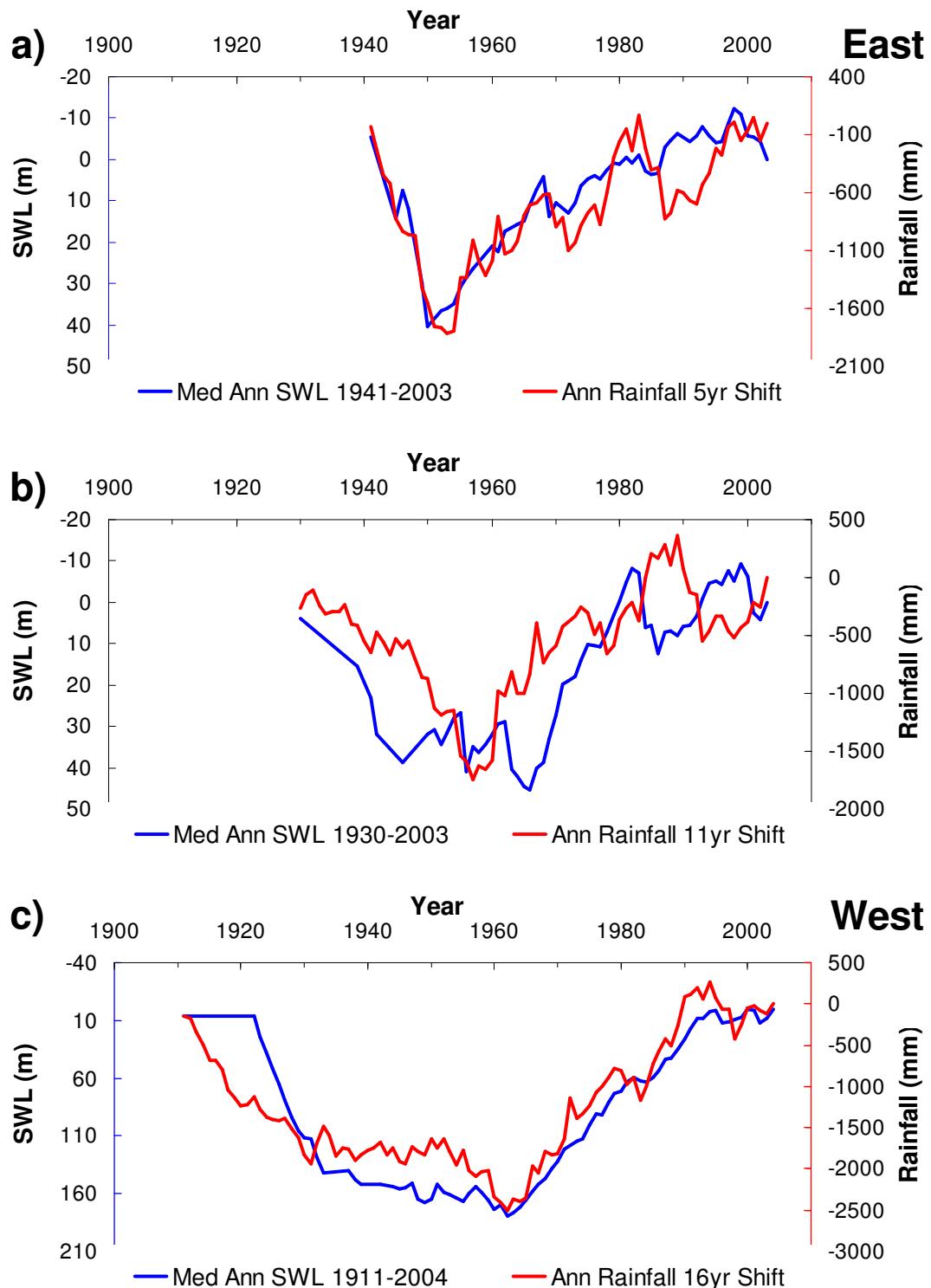


Figure 5.7: Residual mass curves of rainfall (Ann Rainfall) shifted forward by a) 5 b) 11 and c) 16 years, compared to residual mass curves of median annual SWL (Med Ann SWL) in the East, Mid and West sections of the Lachlan River catchment.

vidual rainfall stations that were used in the study can be found in the Appendix in Table A.1 and Table A.2.

A statistically significant (>90%) step-point change was detected in all the sections except Murrumbidgee West (Table 4.7). The change-point, which occurred in 1947, was associated with an abrupt increase in the average rainfall of between 73 and 107 mm (Murrumbidgee East and Lachlan East, respectively). The rainfall shift was most pronounced in the uniform rainfall zone and at higher altitudes. Its influence diminished in the winter rainfall zone, at lower altitudes and with increasing aridity.

The summary of change-points and lags is presented in Table 5.2. The summary shows that the trend analysis can provide information about the delayed response of the aquifers in more instances than the change-point analysis: when the SWL record does not reach back to capture the shift, as in case of Murrumbidgee East section, including its North East and South East divisions, or when the abrupt change in SWL is mild and close to the start of the record, and therefore statistically not significant, such as in case of Lachlan Mid section. Where change-point analysis captures statistically significant shifts in rainfall and SWL, the estimated delay is either identical (Lachlan East and Lachlan West section) or close, such as in Namoi catchment, where SWL abruptly changes after two and eight years, averaging five years, which is close to four years estimated by trend-point analysis. This provides additional validation of the accuracy and robustness of the 21-year moving average methodology.

Table 5.2: Abrupt changes and lags in groundwater response to rainfall.

Section/ subsection	Year of shift in Rainfall	Year of shift in SWL	Delay (years)	Moving average fit
Namoi	1947	1949, 1955	2, 8	4–5
Lachlan East	1947	1952	5	5
Lachlan Mid	1947	(NS)	(-)	11
Lachlan West	1947	1963	16	16
Murrumbidgee NE	1947	1991	#	N/A
Murrumbidgee SE	1947	1988	#	N/A
Murrumbidgee East	1947	1965, 1997	#	N/A
Murrumbidgee West	(NS)	1937	(-)	11

# The delay between shift in rainfall and shift in SWL was not calculated for the Murrumbidgee, as the shifts indicated a decrease in SWL, which was unrelated to the rainfall behaviour.

NS=Not statistically significant. N/A=Not applicable.

Lags that are presented in this analysis should be seen as approximate and informative, not as accurate values.

### **5.3 Summary**

This chapter validates the methodology for reconstruction of bore hydrograph trends. The method was validated using recorded bore hydrographs in fractured rock areas and rainfall timeseries, as a surrogate hydrologic dataset. The criterium was the Root Mean Square Error and the model was validated successfully in both instances.

# **Chapter 6**

## **Discussion**

This chapter discusses the model applicability in relation to the age and the condition of fracturing in the rocks; the target user group; the user-friendliness of its application in respect to software requirements and model transparency; the degree of model robustness and the causes of spatial variability in the dataset. The application of this methodology led to a much improved understanding of the recent changes in groundwater levels in the upland areas of NSW.

### **6.1 Bore data mining technique**

The presented methodology for hydrograph trend reconstruction has been derived and validated using bores drilled in the Palaeozoic fractured rocks. Caution should be exercised if the methodology is extrapolated outside this time boundary. The sedimentary and volcanic rocks in the study area were fractured, but not weathered. Secondary porosity caused by fracturing was dominant in these rocks, irrespective of their origin. That allowed for bores in granitic and meta-sedimentary rocks to be analysed together. If the model is applied for younger rocks, in which fracturing of rocks differs between sedimentary, volcanic intrusive and volcanic extrusive rocks, the study area should be segmented to reflect the geology and each geologic formation should be modeled separately.

The method uses GIS software and Microsoft Excel, typically available and familiar to hydrogeologists around the world. The target user group comprises of hydrogeologists, hydrologists and water engineers. Application of the methodology is inexpensive, as it does not require purchase of specialised statistical or data-mining

software and the training associated with their use.

The methodology uses the annual median of standing water level as a measure of central tendency to compose the SWL points in annual time-series. When tested against the alternative average and spatially weighted average (Figure 3.3) it was superior, clearly demonstrated by tighter confidence intervals. However, if the bore density is high, it is also recommended to investigate use of the spatially weighted average.

The model uses a very simple and robust 21-year moving average filter. Although other filter lengths were investigated, the 21-year moving average was chosen based on the minimal Root Mean Square Error. When tested against the more sophisticated splines, it demonstrated robustness and superiority. Confidence intervals around splines proved to be wider, and the RMSE was marginally higher. The only advantage of using splines would be a longer time-span: capturing the results for the entire sample interval. This would, however, come with very low reliability at the edges of the interval due to widening of the confidence intervals. The extension of the filter could also be achieved by allowing the moving average filter to capture less than 21 year span at the start and the end of the sample interval. The optimal length of the filter could be a function of bore density. It is therefore recommended to further investigate the optimal filter size.

Tightness of the confidence intervals depends on the sample size:

- (a) the number of years in which fractured rock bores were drilled within the 21 year period, and
- (b) the scatter of the annual sample data, as shown by the standard deviation.

As the early data are generally more sparse than the recent, the confidence intervals are wider at the start of the record.

There are two causes of spatial variability in SWL records, both linked to the scale. Figures 5.3 c and 5.4 c illustrate variability at the local scale; for example where water table is deeper at the top of the hill than at the bottom. The methodology deals with this variability through the central tendency model (median annual SWL). However, variability at the larger, regional scale, is primarily due to the spatial rainfall variability, and can be observed by comparing depth to groundwater in three sections of the Lachlan catchment: deepest water table in the more arid Lachlan West section and shallowest in the high rainfall Lachlan East section (Figure 5.6).

As the rainfall and altitude are positively correlated, the water table is deeper in lowland areas. The methodology deals with this large-scale spatial variability by sub-dividing the study area. This prevents spatial-temporal confounding.

It was possible to use rainfall as a training dataset while investigating the choice of filtering technique because the changes in land-use, pumping and abstractions from the aquifer were negligible in the Namoi section during the entire data interval evaluated.

## 6.2 New knowledge gained from the hydrograph trend reconstruction

The application of the methodology to the NSW portion of the Murray-Darling Basin made possible a reconstruction of the long-term groundwater trends. Understanding of the hydrogeologic history in turn contributed to a number of discoveries and new knowledge, gained through comparison of groundwater and rainfall trends. It explained the origin of shallow groundwater levels and seeps observed during the second portion of the twentieth century and linked it directly to the climate shift in 1947. This had major implications in understanding of the origins and drivers of dryland salinity and has led to significant presentations and discussions at regional conferences [Rancic et al., 2006, Rančić and Acworth, 2008] and an on-line publishing of a peer reviewed report [Rančić et al., 2009]. As the conference papers are not widely available, they are included as Appendix B.

Analysis showed that the reconstructed water table trends were very similar to the trends in rainfall across the study area, with the exception of the initial period in the Lachlan West section. The water table trends lag variably behind the rainfall trends. Lags estimated by trend analysis match well with lags deducted from the minima in residual mass curves. Observed lags are short close to the Great Dividing Range and increase down the catchment in a westerly direction. This is an interesting finding and requires further research to clarify its causes. These observed lags reduced the perception of the response time of the aquifers by more than an order of magnitude, compared with previous assumption introduced by Coram et al. [2000].

Comparison of rainfall and water table trends allowed separation of climatic

from non-climatic drivers in the study area. Following the same logic as Stoll et al. [2011], if certain patterns of changes are observable in the groundwater level data, but not in the rainfall data, the observed patterns are likely to be related to land use changes.

For all the sections in the study area, except the Lachlan West section, the similarity of rainfall and water table patterns indicated that the recharge was not substantially altered by the non-climatic effects during the period of observation. Climate was therefore the major driver of water table fluctuations during the period of observation everywhere in the study area except in the early period in the Lachlan West section. Coarse groundwater data coverage prevented detection of possible minor, temporary or local effects of human influence on recharge, such as localised changes in vegetation cover or land management practices, minor groundwater pumping, dam construction and river regulation, which occurred throughout the twentieth century.

Diverging patterns in the Lachlan West section indicated additional recharge in the initial period due to a cause unrelated to climate. It was consistent with the rise in water table from the change in land-use. Indeed, the expansion of the wheatbelt analysed and documented by Bedward et al. [2007] coincides spatially and temporally with the rise in water table in Lachlan West section. At the start of the twentieth century this area was covered by pastures and mostly cleared of trees. Land-use change from grassland to cropping and elimination of any remaining trees triggered additional recharge and caused groundwater rise unrelated to climate. The match between delayed rainfall and SWL trends is excellent after 1944, indicating the climate as the primary driver since then.

It is unlikely that the change in storativity, due to the change in groundwater level is significant for the water level trends in the study area. Only a small portion of the aquifer is under unconfined condition: at the ridges, where only the highly permeable A horizon forms. Existence of the less permeable loamy and clayey B horizon elsewhere imposes confined conditions, so change in storativity in these areas due to the water-table variation can not be significant. Confined aquifer conditions are confirmed by the changes in barometric pressure in observation bores wherever the water level has been monitored with high frequency (less than hourly), such as in the bore GW367921 in the Yass valley [Nicoll and Scown, 1993]. Insignificant change in storativity as a function of water table is also reflected in the fact that the

water level trends follow the rainfall trends with the delay both during deep (first half of the twentieth century) and shallow (second half of the twentieth century) episodes of the groundwater level.

There is no strong evidence to link the spatial variability in response times to variability in styles of fracturing. However, it is possible that fracture density is higher in the east than in the west part of the study area, due to the stronger effects from the extensional forces during opening of Tasman Sea closer to the cost. Unfortunately, the fault mapping in NSW is very poor and inconsistent, to confirm this hypothesis. It is very likely that this variability is correlated to: (a) the thickness of the regolith, which effects recharge time, and (b) the flatness of the slopes, which drive the hydraulic gradient, as the response times increases in westward direction, away from the divide, where regolith is thicker and slopes are flatter. However, this hypothesis is yet to be confirmed.

The groundwater abstractions in fractured rock have not been metered yet. Therefore it is not possible to quantitatively prove that the effect of diversions on watertable trends was negligible. However, most of the early pumps, until 1970s have been equipped with the wind turbines only, and many still are. They have very limited pump capacity and the bore density is extremely low, which points out to very low abstraction levels and explains the lack of metering. The study area contains more bores that draw water from alluvial aquifers than from the fractured rock aquifers. In the alluvial aquifers, in so-called alluvial groundwater management areas: the yielding capacity of these aquifers is much higher; the bores are generally equipped with electronic pumps, which allow more intense abstractions compared to bores equipped with windmill pumps; and the usage has been metered recently. It can be therefore concluded, with certainty, that the abstractions from the alluvial aquifers are much higher in total and per unit area than the abstractions from the fractured rock aquifers in the study area. In the Tarcutta catchment bore density, ratio of fractured to alluvial bores and the pumping styles are typical of the study area, and the abstractions were metered since 1999. Evaluation of the abstractions from alluvial aquifers during the hydrologic study in the Tarcutta catchment [Rančić et al., 2011, 2014], showed that the maximum abstractions in the 1999-2004 period occurred in 1999: 202 ML/year, which equates to 0.17 percent of streamflow and less than 0.4 percent of baseflow. On the basis of this, it can be concluded that the abstractions from fractured rock aquifers are insignificant in the overall water

balance in the study area and during the period the study covers. Therefore, even though these abstractions act in the same direction as evapotranspiration in water removal, their effect on the fractured rock aquifers is negligible.

The exclusion of deeper wells from the analysis dramatically reduced the noise in the data and the RMSE when the trends in SWL and rainfall were compared. The deeper wells, which access deeper water bearing zones are situated mainly in three locations: (a) very top of the catchments and ridge lines, corresponding to intense recharge areas with downwards hydraulic gradient; (b) deeper water bearing zones associated with longer groundwater pathways, often not discharging locally; and (c) in fractured rock underlying alluvial systems, in discharge areas, associated with upward hydraulic gradient. It would be possible to look at these three categories of wells if data were available. Only the third type of deep bores exist in sufficiently large numbers in the study area to perform separate analysis. However, the groundwater levels in these bores could be influenced by alluvial aquifers, stream heights and abstractions, which might compromise and invalidate conclusions.

Filling of the aquifers and the response to increased recharge. As the aquifers generally fill faster than empty, the time needed to reach equilibrium during the drying phase could take longer.

The watertable rose across the study area during the second half of the twentieth century. In the summer and uniform rainfall zones this rise was abrupt, and consistent with the rainfall shift, while in the winter zone, which dominates the Murrumbidgee West section, the transition from the dryer to wetter hydrologic regime exist, but it is not abrupt, which is consistent with the rainfall patterns. In the winter rainfall zone the increase in rainfall has been detected, but below the 90 percent significance level. In the Lachlan East data due to the short bore record prior to the rise, the change-point does not appear significant when evaluated with the method employed by the Change-point analyser software. However, the residual mass curve constructed from these data illustrates the rise (Figure 5.7a). The bore record in the Murrumbidgee East section can not depict this rise as the record begins after the rainfall shift.

The wetter climate forced groundwater levels in fractured rock to rise during the second half of the twentieth century across the study area. The higher position of the groundwater table allowed the groundwater systems to reach the local depressions in the terrain and to discharge at a larger number of points. Where evaporation

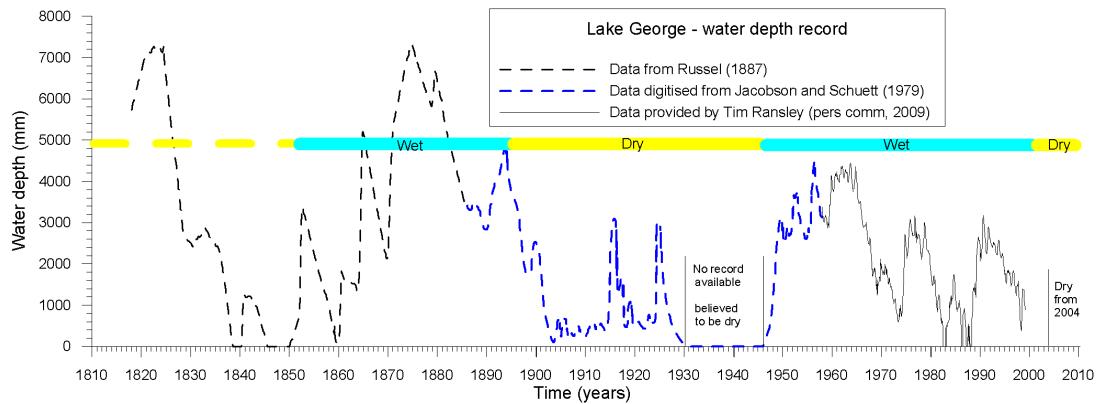


Figure 6.1: Water level record for Lake George in southern NSW.

exceeded discharge salt became concentrated, forming saline patches on the surface of the land. This increased both the number of salinised areas and their total size.

When looking at the relative importance of land use change and climate on the current position of a water table, it is important to look at the problem from a wide historical perspective, and acknowledge that climate changes, and has always been changing, and that stationary climatic mode, that can allow development of hydrologic equilibrium is exceptional and short-lived. The record of levels of Lake George [Jacobson et al., 1991] is one of the longest Australian hydrologic records, and it clearly illustrates (Figure 6.1) that the climatic variation that caused so many salinity outbreaks over the last half of the twentieth century was in fact very mild. Lake George levels fluctuated in the nineteenth century almost twice as much (from 0 to 7.5 m) as in the twentieth (from 0 to 4.5 m). This range is again negligible when compared with relatively recent history. Between 27,000 to 21,000 years ago the level reached 37 m and then water spilled over Geary's Gap into the Yass River Valley catchment. The high of 37 m represents about nine times the height reached in the twentieth century. This excess of rainfall over evaporation is thought to be as much due to reduced evaporation, a result of lower temperatures, as it was to increased rainfall. In either case, land clearing was not a cause at this time!

Short response times of groundwater systems and completion of most clearing by the end of the nineteenth century are together responsible for the absence of detectable clearing-induced water table rise across the study area in the early part of the twentieth century, except in Lachlan West section. This rise must have preceded the available bore record. Even though the climate was the primary driver

of water table fluctuations during the period of observation, if there was no clearing these fluctuations would have occurred at the deeper base level. The clearing that took place more than a century ago changed the hydrologic properties in the study area, allowing higher recharge than if the previous land use had remained intact. The dry climatic phase between 1895 and 1947 and the cleared terrain had opposing effects on the recharge rate and on the hydrologic balance. After 1947, the effects of clearing and the wetter climate on recharge were superimposed. If the clearing had not taken place, the water table would have been deeper, the overall effect of the rainfall shift would have been smaller, and dryland salinity would have been less prominent.

The better understanding of groundwater hydrograph changes, as being primarily caused by climatic factors overriding the secondary importance of changes to the water balance caused by clearing during the period of observation, has led to a much better understanding of processes in these fractured rock systems. Previous studies Coram et al. [2000] had assumed implicitly that rainfall was constant and that the only hydrologic change in the system was due to land clearing. This led to the improbable conclusion that the groundwater system response lags clearing by as much as 50 years. The new understanding developed in this Thesis better explains the perceived lags a result of an abrupt change in rainfall and the position of water table as a combination of factors.

### 6.3 Summary

This chapter discusses the developed methodology for reconstruction of bore hydrograph trends and the new knowledge gained by its application to South-Eastern Australian fractured rock aquifers. It recommends to use this robust and user-friendly method for reconstruction of groundwater trends in shallow fractured rock aquifers. Reconstruction of the hydrograph trends dramatically reduced the perception of the time needed for the aquifers to equilibrate after the hydrologic change: by one to two orders of magnitude. The comparison with rainfall linked the climatic shift in 1947 from dryer to wetter hydrologic regime to groundwater rise, waterlogging, saline seeps and increase in dryland salinity. This, with no doubt, added the climate to the clearing as the important factor in the dryland salinity expression.

# Chapter 7

## Conclusions

A major outcome from this research has been the realisation that historic aquifer behaviour can be reconstructed from depth to water table data recorded at the time of bore completion.

These data offer the opportunity to:

- (a) reconstruct long-term trends of standing water level in representative areas of fractured basement;
- (b) detect change points in groundwater hydrographs and evaluate statistically their confidence intervals;
- (c) estimate the response time of fractured-rock aquifers to climatic variations; and
- (d) detect the influence of changes in land-use and be able to separate them from climatic drivers.

This work also demonstrates that the major shifts in climate (rainfall) are the significant drivers of recharge change.

The approach described here has been successful in recreating past multi-decadal groundwater level hydrograph trends in the fractured rock areas in eastern New South Wales in Australia. This has allowed important conclusions to be drawn concerning the groundwater level response to climate and land-use change. Using this approach, it has also been possible to provide a different interpretation for the mechanisms that lead to the development of dryland salinity in this area.

Application of the methodology demonstrated that aquifers in fractured rock basement in the study area respond to hydrologic changes rapidly and reach a state of dynamic equilibrium within a half decade in the Namoi section and eastern highlands of the Lachlan and Murrumbidgee River catchments while the process takes up to 16 years in the western lowlands of the Lachlan River catchment. The available borehole records reached back to the first half of the twentieth century, with the earliest borehole sank in the Namoi catchment in 1902. In most of the sections of the study area, including the Namoi section, the hydrograph trends in the form of the 21-year moving average have been the delayed replica of the equivalent rainfall trends, indicating that climate forcing has been the main driver during the period of observation, with the land-use change effects undetectable on the section-scale. Both findings came as a surprise, as the appearance of waterlogged areas and saline seeps in the second half of the twentieth century had been assumed to be a direct consequence of a massive deforestation for agriculture development. As the peak of deforestation had happened a century earlier before the expansion of salinity, much longer response to hydrologic changes had been deduced incorrectly. While the initial purpose of deforestation was for pastoralism, with major deforestation occurring between 1860s and 1900s, the second wave of land-use change occurred in the so-called wheat-belt, in the first half of the twentieth century. The remnant tree vegetation was eliminated and pastures were replaced by grain production on the western, flatter fringes of the study area. The Lachlan West section clearly demonstrates the ability of the methodology to detect the changes in recharge triggered by the switch in land-use, as the early hydrograph trend rose steeply, deviating from the rainfall trend in this section. In conclusion, climate was the major driver of groundwater fluctuations during the twentieth century, but clearing had previously contributed to the elevated water table. Both drivers are therefore implicated in the expression of dryland salinity in this part of the Australian continent.

As water shortages becomes pressing in many parts of the world, there is mounting pressure to develop the resources in shallow fractured basement rocks. This is certainly the case in Eastern Australia, where the current negotiations concerning surface water and groundwater within the Murray-Darling Basin make the assumption that large volumes of groundwater in fractured rocks are available in the Basin.

Studies of fractured rock aquifers and the determination of aquifer parameters in these areas have often been neglected. This is probably the case because each

individual borehole resource has only limited commercial value.

In addition to this, the growing concern about the impact of current climate change on water resource availability is a significant factor for management of the resource. Without the information on past groundwater behaviour, that can be derived from an application of the methodology described in this thesis, it is not possible to calibrate groundwater models to predict the impacts of climate change with any degree of accuracy.

Consequently, the methodology has obvious applicability in many other parts of the world where basic bore completion data exists, but long-term monitoring data have not been collected, as the application of the down-hole hammer technique for bore drilling in 1960s marked the very beginning of hydro-geological investigations [Lachassagne et al., 2014]. This is certainly the case in large areas of Africa, South America, Australia and India where due to water shortages the exploitation of water from the fractured rocks is crucial [Krásný, J and Sharp, J.M., 2007, Lachassagne et al., 2014].

# Bibliography

- R. I. Acworth. *Investigation of Groundwater Resources in the Fractured Basement of Northern Nigeria*. PhD thesis, Department of Geological Sciences, The University of Birmingham, UK, 1981.
- R. G. Allen, L. S. Periera, D. Raes, and M. Smith. Crop evapotranspiration - guidelines for computing crop water requirements. Irrigation and Drainage Paper 56, Food and Agriculture Organisation of the United Nations (FAO), FAO, Rome, 1998.
- J. Anderson. Hydrological assessment of the dryland salinisation of the Snake Creek catchment. Technical report, Department of Water Resources, Technical Services Division, February 1992. Report TS 92.057.
- T. Asefa, N. Wanakule, and A. Adams. Field-scale application of three types of neural networks to predict ground-water levels. *JAWRA Journal of the American Water Resources Association*, 43(5):1245–1256, 2007. ISSN 1752-1688. doi: 10.1111/j.1752-1688.2007.00107.x. URL <http://dx.doi.org/10.1111/j.1752-1688.2007.00107.x>.
- AUSLIG. Australia: Natural vegetation map (1:5,000,000), 1997.
- T. A. Awchi. River discharges forecasting in northern Iraq using different ANN techniques. *Water Resources Management*, 28(3):801–814, 2014. ISSN 0920-4741. doi: 10.1007/s11269-014-0516-3. URL <http://dx.doi.org/10.1007/s11269-014-0516-3>.
- F. Bacao, V. Lobo, and M. Painho. The self-organizing map, the Geo-SOM, and relevant variants for geosciences. *Computers and Geosciences*, 31:155–163, 2005.

- H. Banejad, M. Kamali, K. Amirmoradi, and E. Olyaei. Forecasting Some of the Qualitative Parameters of Rivers Using Wavelet Artificial Neural Network Hybrid (W-ANN) Model (Case of study: Jajroud River of Tehran and Gharaso River of Kermanshah). *Iranian Journal of Health and Environment*, 6(3):18, 2013.
- M. Bedward, C. Simpson, M. Ellis, and L. M. Metcalfe. Patterns and Determinants of Historical Woodland Clearing in Central-Western New South Wales, Australia. *Geographical Research*, pages 348–357, 2007. doi: 10.1111/j.1745-5871.2007.00474.x.
- N. Bhatia, L. Sharma, S. Srivastava, N. Katyal, and R. Srivastav. Streamflow Decomposition Based Integrated ANN Model. *Open Journal of Modern Hydrology*, 3:15, 2013.
- J. M. Bowler. Aridity in Australia: age, origins and expression in aeolian landforms and sediments. *Earth-Science Reviews*, 12:279–312, 1976.
- J. M. Bradd. Hydrological controls on Dryland salinity at Wattle Retreat, Cootamundra, New South Wales, 1988. BSc Thesis. Department of Applied Geology, University of New South Wales.
- D. F. Branagan and G. H. Packham. *Field Geology of New South Wales*. New South Wales Department of Mineral Resources, Sydney, Australia, third edition, 2000.
- D. Brett. In: *Land Care: science in action*, chapter Restoring the Murray-Darling Basin, pages 30–35. CSIRO Publications, Canberra, ACT, Australia, 1993.
- K. W. Butzer and D. M. Helgren. Livestock, Land Cover and Environmental History: The Tablelands of New South Wales, Australia 1820-1920. *Annals of Association of American Geographers*, 95(1):80–111, 2005.
- Y. Cao, Z. Liang, Q. Huang, S. Huo, K. Xu, and W. Chang. Study on application of BMA method to forecasting of sediment concentration hydrograph at Longmen Hydrological Station on Yellow River. *Water Resources and Hydropower Engineering*, 6:4, 2013.
- B. W. Chappell and A. J. R. White. Two contrasting granite types: 25 years later. *Australian Journal of Earth Sciences*, 48:489–499, 2001.

- P. B. Chattopadhyay and R. Rangarajan. Application of ANN in sketching spatial nonlinearity of unconfined aquifer in agricultural basin. *Agricultural Water Management*, 133:81–91, 2014.
- L. Chen, C. Chen, and D. Lin. Application of Integrated Back-Propagation Network and Self-Organizing Map for Groundwater Level Forecasting. *Journal of Water Resources Planning and Management*, 137(4):352–365, 2011. doi: 10.1061/(ASCE)WR.1943-5452.0000121.
- M. Chitsazan, G. Rahmani, and A. Neyamadpour. Groundwater level simulation using artificial neural network: a case study from Aghili plain, urban area of Gotvand, south-west Iran. *Geopersia*, 3(1):35–46, 2013.
- H. Chu and L. Chang. Optimal control algorithm and neural network for dynamic groundwater management. *Hydrological Processes*, 23(19): 2765–2773, 2009. ISSN 1099-1085. doi: 10.1002/hyp.7374. URL <http://dx.doi.org/10.1002/hyp.7374>.
- E. Coppola, F. Szidarovszky, M. Poulton, and E. Charles. Artificial Neural Network Approach for Predicting Transient Water Levels in a Multilayered Groundwater System under Variable State, Pumping, and Climate Conditions. *Journal of Hydrologic Engineering*, 8:348–360, 2003. doi: 10.1061/(ASCE)1084-0699(2003)8:6(348).
- J. E. Coram, P. R. Dyson, P. A. Houlder, and W. R. Evans. Australian groundwater flow systems contributing to dryland salinity. Technical report, Bureau of Rural Science, Canberra, 2000. Report for the Dryland Salinity Theme of the National Land and Water Resources Audit, Canberra (CD contains report, maps, and spatial data).
- I. N. Daliakopoulos, P. Coulibaly, and I. K. Tsanis. Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309:229–240, 2005. doi: 10.1016/j.jhydrol.2004.12.001.
- N. Dash, S. Panda, R. Remesan, and N. Sahoo. Hybrid neural modeling for groundwater level prediction. *Neural Computing and Applications*, 19:1251–1263, 2010. doi: 10.1007/s00521-010-0360-1.

E. L. Deacon. Climatic change in Australia since 1880. *Australian Journal of Physics*, 6(2):209–218, 1953.

DIPNR. Groundwater Database System (GDS). Digital data set, 2004a. NSW Department of Infrastructure, Planning and Natural Resources. Parramatta, NSW, Australia.

DIPNR. NSW Stream Network. Digital data set, 2004b. NSW Department of Infrastructure, Planning and Natural Resources. Parramatta, NSW, Australia.

E. W. Docker. *The Bardwells of Bardwell Park*. Mini-publishing, 2005.

J. R. Dodson. Timing and response of vegetation change to Milankovitch forcing in temperate Australia and New Zealand. *Global and Planetary Change*, 18:161–174, 1998.

U. Ehret and E. Zehe. Series distance – an intuitive metric to quantify hydrograph similarity in terms of occurrence, amplitude and timing of hydrological events. *Hydrology and Earth System Sciences*, 15(3):877–896, 2011. doi: 10.5194/hess-15-877-2011. URL <http://www.hydrol-earth-syst-sci.net/15/877/2011/>.

U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI magazine*, 17:37–54, 1996.

S. Feng, S. Kang, Z. Huo, S. Chen, and X. Mao. Neural Networks to Simulate Regional Ground Water Levels Affected by Human Activities. *Ground Water*, 46: 80–90, 2007.

C. Gaina, R. D. Roest, W. R. and Mller, and P. Symonds. The opening of the Tasman Sea: a gravity anomaly animation. *Earth Interactions*, 2:1–23, 1998.

J. Gentilli. Word survey of climatology. In Joseph Gentilli, editor, *Climates of Australia and New Zealand*, volume 13, page 405. Elsevier, 1971. ISBN 0444408274.

Geoscience Australia. Vegetation - Pre-European Settlement (1788), 1:5 000 000, 2009. URL <http://www.ga.gov.au/meta/ANZCW0703005425.html#citeinfo>. Cited on 30/06/2010.

- H. Ghalkhani, S. Golian, B. Saghafian, A. Farokhnia, and A. Shamseldin. Application of surrogate artificial intelligent models for real-time flood routing. *Water and Environment Journal*, 27(4):535–548, 2013.
- D. K. Ghose, S. S. Panda, and P. C. Swain. Prediction of water table depth in western region, Orissa using BPNN and RBFN neural networks. *Journal of Hydrology*, 394 (3–4):296 – 304, 2010. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2010.09.003. URL <http://www.sciencedirect.com/science/article/pii/S0022169410005664>.
- N. P. Gopalan and B. Sivaselvan. *Data Mining: Techniques and Trends*. PHI Learning Private Limited, New Delhi, India, 2009. ISBN 978-81-203-3812-8.
- R. D. Graetz, M. A. Wilson, and S. K. Campbell. Land cover disturbance over the Australian continent: A contemporary assessment. In *Biodiversity Series*, volume 7. Department of Environment, Sport and Territories, 1995.
- A. K. Greve, R. I. Acworth, and B. F. J. Kelly. 3D Cross-hole resistivity tomography to monitor water percolation during irrigation on cracking soil. *Soil Research*, 49 (8):661–669, 2001. URL <http://dx.doi.org/10.1071/SR11270>.
- David J. Hand, Heikki Mannila, and Padhraic Smyth. *Principles of data mining*. The MIT Press, Cambridge, Massachusetts, 2001. ISBN 10:0-262-08290-X.
- R. Harrison. *Shared Landscapes: Archaeologies of Attachment and the Pastoral Industry in New South Wales (Studies in the Cultural Construction of Open Space)*. UNSW Press, 2004. ISBN-10: 0868405590.
- R. Hewett. Data mining for generating predictive models of local hydrology. *Applied Intelligence*, 19:157–170, 2003. ISSN 0924-669X. URL <http://dx.doi.org/10.1023/A:1026005922241>. 10.1023/A:1026005922241.
- Y. S. Hong and M. R. Rosen. Intelligent characterisation and diagnosis of the groundwater quality in an urban fractured-rock aquifer using an artificial neural network. *Urban Water*, 3(3):193–204, 2001. doi: [http://dx.doi.org/10.1016/S1462-0758\(01\)00045-0](http://dx.doi.org/10.1016/S1462-0758(01)00045-0).
- Z. Huo, S. Feng, S. Kang, X. Mao, and F. Wang. Numerically modelling groundwater in an arid area with ANN-generated dynamic boundary conditions. *Hydrological*

*Processes*, 25(5):705–713, 2011. ISSN 1099-1085. doi: 10.1002/hyp.7858. URL <http://dx.doi.org/10.1002/hyp.7858>.

Insightful Incorporated. S-Plus 7 software, 1995.

G. Jacobson, J. Jankowski, and R. S. Abell. Groundwater and surface water interaction at Lake George, New South Wales. *BMR Journal of Australian Geology & Geophysics*, 11:177–188, 1991. 10.1139/e74-015.

A. Jalalkamali and N Jalalkamali. Groundwater modeling using hybrid of artificial neural network with genetic algorithm. *African Journal of Agricultural Research*, 6:5775–5784, 2011.

A. Jalalkamali, H. Sedghi, and M. Manshouri. Monthly groundwater level prediction using ANN and neuro-fuzzy models: a case study on Kerman plain, Iran. *Journal of Hydroinformatics*, 13:867–876, 2011. doi: 10.2166/hydro.2010.034.

M. K. Jha and S. Sahoo. Efficacy of neural network and genetic algorithm techniques in simulating spatio-temporal fluctuations of groundwater. *Hydrological Processes*, 2014.

W. H. Johnston, D. L. Garden, A. Rancic, T. B. Koen, K. B. Dassanayake, C. M. Langford, N. J. S. Ellis, M. A. Rab, N. K. Tuteja, M. Mitchell, J. Wadsworth, D. Dight, K. Holbrook, R. LeLievre, and S. M. McGeoch. The impact of pasture development and grazing on water-yielding catchments in the Murray-Darling basin in south-eastern Australia. *Australian Journal of Experimental Agriculture*, 43:817–841, 2003.

Krásný, J and Sharp, J.M. Hydrology of fractured rocks from particular fractures to regional approaches: State-of-the-art and future challenges. In *Groundwater in Fractured Rocks: IAH Selected Paper Series*, chapter Introductory paper, pages 1–30. Taylor and Francis/Balkema, Leiden, The Netherlands, 2007. ISBN ISBN13:978-0-415-41442-5 (Hbk).

B. Krishna, Y. R. Satyaji Rao, and T. Vijaya. Modelling groundwater levels in an urban coastal aquifer using artificial neural networks. *Hydrologic Processes*, 22(1):1180–1188, 2008. ISSN 1099-1085. doi: 10.1002/hyp.6686. URL <http://dx.doi.org/10.1002/hyp.6686>.

- Patrick Lachassagne, Benoit Dewandel, and Roberts Wyns. The conceptual model of weathered hard rock aquifers and its practical applications. In J Krásný and J.M. Sharp, editors, *Groundwater in Fractured Rocks: IAH Selected Paper Series*, chapter 2, pages 13–45. 2014.
- A. Lalozaee, A. Pahlavanravi, F. Bahreini, H. Ebrahimi, and H. Iezadi. Efficiency Comparison of IHACRES Model and Artificial Neural Networks (ANN) in Rainfall-Runoff Process Simulation in Kameh Watershed (A Case Study in Khorsan Province, NE Iran). *International Journal of Agriculture: Research and Review*, 3:900–907, 2013. ISSN 2228–7973. URL <http://ecisi.com/wp-content/uploads/2013/08/900-907.pdf>.
- Legislative Assembly of NSW. The Act for Regulating the Alienation of Crown Lands (No. 1 of 1861), 1861. URL <http://www.foundingdocs.gov.au/item.asp?sdID=80>.
- D. N. Lerner, A. S. Issar, and I. Simmers. *Groundwater Recharge - A Guide to Understanding and estimating Natural Recharge*. Verlag Heinz Heise, 1990. Volume 8 of the International Contributions to Hydrogeology.
- X. Li, L. Shu, L. Liu, D. Yin, and J. Wen. Sensitivity analysis of groundwater level in Jinci Spring Basin (China) based on artificial neural network modeling. *Hydrogeology Journal*, 20:727–738, 2012. ISSN 1431-2174. URL <http://dx.doi.org/10.1007/s10040-012-0843-5>. 10.1007/s10040-012-0843-5.
- X. Liang and Y. Liang. Applications of Data Mining in Hydrology. In *Proceedings of the 2001 IEEE International Conference on Data Mining, ICDM '01*, pages 617–620, Washington, DC, USA, 2001. IEEE Computer Society. ISBN 0-7695-1119-8. URL <http://dl.acm.org/citation.cfm?id=645496.658025>.
- G. F. Lin and G. R. Chen. An improved neural network approach to the determination of aquifer parameters. *Journal of Hydrology*, 316:281–289, 2006.
- G. Lischeid. Taming awfully large data sets: using self-organizing maps for analyzing spatial and temporal trends of water quality data. *Geophysical Research Abstracts*, 22, 2003.
- M. C. Lovell. Data mining. *The Review of Economics and Statistics*, 65:1–12, 1980.

- H. R. Maier and G. C. Dandy. Neural networks for the prediction and forecasting of water resources variables : a review of modelling issues and applications. *Environmental Modelling & Software*, 15(1):101–124, 2000. URL <https://cours.etsmtl.ca/sys843/pdf/Maier2000.pdf>.
- G. H. McTainsh. Quaternary aeolian dust processes and sediments in the australian region. *Quaternary Science Reviews*, 8:235–253, 1989.
- M. I. Melis and R. I. Acworth. A possible aeolian component in Pleistocene and Holocene valley aggradation, Dicks Creek Catchment, New South Wales. *Australian Journal of Soil Science*, 39(1):13–38, 2001.
- W. J. Michael, B. S. Minsker, D. Tcheng, A. J. Valocchi, and J. J. Quinn. Integrating data sources to improve hydraulic head predictions: A hierarchical machine learning approach. *Water Resources Research*, 41:15, 2005. doi: 10.1029/2003WR002802. W03020.
- S. Mohanty, M. Jha, A. Kumar, and K. Sudheer. Artificial Neural Network Modeling for Groundwater Level Forecasting in a River Island of Eastern India. *Water Resources Management*, 24:1845–1865, 2010. ISSN 0920-4741. URL <http://dx.doi.org/10.1007/s11269-009-9527-x>. 10.1007/s11269-009-9527-x.
- A. Montanari and E. Toth. Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins? *Water Resources Research*, 43, 2007. doi: 10.1029/2006wr005184. W05434.
- V. Moosavi, M. Vafakhah, B. Shirmohammadi, and N. Behnia. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water resources management*, 27(5):1301–1321, 2013a.
- V. Moosavi, M. Vafakhah, B. Shirmohammadi, and M. Ranjbar. Optimization of Wavelet-ANFIS and Wavelet-ANN Hybrid Models by Taguchi Method for Groundwater Level Forecasting. *Arabian Journal for Science and Engineering*, pages 1–12, 2013b.
- P. C. Nayak, R. Y. R. Satyaji, and K. P. Sudheer. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, 20:77–90, 2006. doi: 10.1007/s11269-006-4007-z.

- C. Nicoll and J. Scown. Dryland salinity in the Yass valley: processes and management. Technical report, Departement of Conservation and Land Management, NSW, 1993. Technical Report.
- V. Nourani, A. A. Mogaddam, and A. O. Nadiri. An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrological Processes*, 22(26):5054–5066, 2008. ISSN 1099–1085. doi: 10.1002/hyp.7129. URL <http://dx.doi.org/10.1002/hyp.7129>.
- V. Nourani, R. G. Ejlali, and M. T. Alami. Spatiotemporal Groundwater Level Forecasting in Coastal Aquifers by Hybrid Artificial Neural Network-Geostatistics Model: A Case Study. *Environmental Engineering Science*, 28:217–228, 2011. doi: 10.1089/ees.2010.0174.
- NRME. Climate data. NSW Monthly Rainfall, 2004. Department of Natural Resources, Mines and Energy. Data CD prepared from data provided by Commonwealth Beaurau of Meteorology.
- T. Pan, Y. Yang, H. Kuo, Y. Tan, J.g Lai, T. Chang, C. Lee, and K. Hsu. Improvement of watershed flood forecasting by typhoon rainfall climate model with an ANN-based southwest monsoon rainfall enhancement. *Journal of Hydrology*, 506:90–100, 2013.
- L. Peeters, F. Bacao, V. Lobo, and A. Dassargues. Exploratory data analysis and clustering of multivariate spatial hydrogeological data by means of GEO3DSOM, a variant of Kohonen’s Self-Organizing Map. *Hydrology and Earth System Sciences*, 11:1309–1321, 2007.
- A. B. Pittock. Global meridional interactions in stratosphere and troposphere. *Quarterly Journal of the Meterological Society*, 99(421):424–437, 1973.
- O. Prakash, K. Sudheer, and K. Srinivasan. Improved higher lead time river flow forecasts using sequential neural network with error updating. *Journal of Hydrology and Hydromechanics*, 62(1):60–74, 2014.
- M. Price. *Introducing Groundwater*. Allen and Unwin, London, UK, 1985.

- G. Rakhshandehroo, M. Vaghefi, and M. Aghbolaghi. Forecasting Groundwater Level in Shiraz Plain Using Artificial Neural Networks. *Arabian Journal for Science and Engineering*, pages 1–13, 2012. ISSN 1319-8025. doi: 10.1007/s13369-012-0291-5.
- A. Rancic, I. Acworth, A. Kathuria, G. Salas, and B. Johnston. Effect of rainfall on groundwater trends over the past century in fractured rocks of the New England Fold Belt in the Namoi catchment, NSW. In *Proceedings of the 10th Murray-Darling Basin Groundwater Workshop: Does the science hold water?*, pages 1–1, Canberra, 2006.
- A. Rančić and I. Acworth. The relationship between the 1947 shift in climate and the expansion of dryland salinity on the western slopes of the Great Dividing Range in NSW. In *Proceedings of the 2nd International Salinity Forum: Salinity, water and society-global issues, local action*, pages 1–5, Adealide, 2008.
- A. Rančić, G. Salas, A. Kathuria, I. Acworth, W. Johnston, A. Smithson, and G. Beal. Climatic influences on shallow fractured-rock groundwater system in the Murray-Darling Basin. Reserch report, Department of Environment and Climate Change, NSW, 59-61 Goulburn Street, PO Box A290, Sydney NSW 2000, Australia, February 2009. URL <http://www.environment.nsw.gov.au/resources/salinity/09108GroundwaterMDB.pdf>. DECC2009/108.
- A. Rančić, B. Christy, T. McLean, and G. Summerell. In *CATPlus Final Technical Report CMI 102529*, chapter Tarcutta Creek Catchment, pages 70–151. Department of Primary Industries, Victoria, 2011.
- A. Rančić, B. Christy D. Read and, T. Mc Lean, I. Hume, and G. Summerell. CAT-Plus modelling in the Tarcutta River Catchment final report. Reserch report, Office of Environment and Heritage, New South Wales Government, 59-61 Goulburn Street, PO Box A290, Sydney NSW 2000, Australia, 2014. Report OEH 2014/0421.
- D. Rassam and M. Littleboy. Identifying vertical and lateral components of drainage flux in hill slopes. In D. A. Post, editor, in *Proceedings of the Inter-*

- national Congress on Modelling and Simulation, MODSIM 2003*, pages 183–188), Townsville, Australia, 2003.
- K. R. Rushton. *Groundwater Hydrology - Conceptual and Computational Models*. Wiley, John Wiley and Sons Ltd, Chichester, West Sussex, PO19 8SQ, England, 2003.
- S. Sahoo and M. K. Jha. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. *Hydrogeology Journal*, 21(8):1865–1887, 2013.
- G. Salas and N. Garland. A Survey of Standing Water Level Changes in Bores in the Macquarie Region. Unpublished report. Groundwater Unit, Department of Water Resources. Dubbo, NSW, Australia, 1989.
- G. Salas and A. Smithson. Rainfall controls on standing water levels in a fractured rock area of central NSW, Australia. In A. Puhalovich, editor, *Proceedings of the IAH International Groundwater Conference - Balancing the Groundwater Budget*, pages 1–8, 2002. Darwin, 12 - 17 May, 2002.
- N. Samani, M. Gohari-Moghadam, and A. A. Safavi. A simple neural network model for the determination of aquifer parameters. *Journal of Hydrology*, 340:1–11, 2007. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2007.03.017. URL <http://www.sciencedirect.com/science/article/pii/S0022169407001862>.
- F. Sanchez-Martos, P. A. Aguilera, A. Garrido-Frenich, J. A. Torres, and A. Pulido-Bosch. Assessment of groundwater quality by means of self-organizing maps: application in a semiarid area. *Environmental Management*, 30:716–726, 2002.
- B. Schaeffli and E. Zehe. Hydrological model performance and parameter estimation in the wavelet-domain. *Hydrology and Earth System Sciences*, 13(10):1921–1936, 2009. doi: 10.5194/hess-13-1921-2009. URL <http://www.hydrol-earth-syst-sci.net/13/1921/2009/>.
- O. Schall and M. Samozino. Surface from Scattered Points: A Brief Survey of Recent Developments. In Bianca; Falcidieno and Nadia Magnenat-Thalmann, editors, *Proceedings of the 1st International Workshop towards Semantic Virtual Environment*, pages 138–147, Villars, Switzerland, 2005. MIRALab.

- J. Shiri, O. Kisi, H. Yoon, K. Lee, and H. A. Nazemi. Predicting groundwater level fluctuations with meteorological effect implications-A comparative study among soft computing techniques. *Computers & Geosciences*, 56:32–44, 2013.
- B. Shirmohammadi, M. Vafakhah, V. Moosavi, and A. Moghaddamnia. Application of several data-driven techniques for predicting groundwater level. *Water Resources Management*, 27(2):419–432, 2013.
- M. B. Shukla, R. Kok, S. O. Prasher, G. Clark, and R. Lacroix. Use of artificial neural networks in transient drainage design. *Transactions of the ASAE*, 39:119–124, 1996.
- A. Singh, M. Imtiyaz, R. K. Isaac, and D. M. Denis. Comparison of soil and water assessment tool (SWAT) and multilayer perceptron (MLP) artificial neural network for predicting sediment yield in the Nagwa agricultural watershed in Jharkhand, India. *Agricultural Water Management*, 104:113–120, 2012.
- R. W. Skaggs. DRAINMOD: Reference Report, 1980. Raleigh: North Carolina State University.
- A. Smithson. An Investigation of unconsolidated sedimentary units and their role in the development of salinity in the Snake Gully Catchment, Central New South Wales, 2002.
- J. M. Spate, B. F. W. Croke, and .A J. Jakeman. Data mining in hydrology. *Focus*, 19(7):1511–1515, 2002.
- P. D. Sreekanth, P. D. Sreedevi, S. Ahmed, and N. Geethanjali. Comparison of FFNN and ANFIS models for estimating groundwater level. *Environmental Earth Sciences*, 62(6):1301–1310, 2011. doi: 10.1007/s12665-010-0617-0.
- S. Stoll, H. J. Hendricks Franssen, R. Barthel, and W. Kinzelbach. What can we learn from long-term groundwater data to improve climate change impact studies? *Hydrology and Earth System Sciences*, 15(12):3861–3875, 2011.
- G. Summerell. *Understanding the processes of salt movement from the landscape to the stream in dryland catchments*. PhD thesis, Department of Civil and Environmental Engineering University of Melbourne, 2004.

- R. Taormina, K. Chau, and R. Sethi. Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Engineering Applications of Artificial Intelligence*, 2012. ISSN 0952-1976. doi: 10.1016/j.engappai.2012.02.009. URL <http://www.sciencedirect.com/science/article/pii/S0952197612000462>.
- W. Taylor. *Change-Point Analyzer 2.0 Shareware Program*. Taylor Enterprises, Libertyville, IL, USA, 2000.
- N. K. Tuteja, G. Beale, W. Dawes, J. Vaze, B. Murphy, P. Barnett, A. Rancic, R. Evans, G. Geeves, D. W. Rassam, and M. H. Miller. Predicting the effects of landuse change on water and salt balance - a case study of a catchment affected by dryland salinity in NSW, Australia. *Journal of Hydrology*, 283:67–90, 2003. doi: 10.1016/S0022-1694(03)00236-1,2002.
- V. Uddameri. Using statistical and artificial neural network models to forecast potentiometric levels at a deep well in South Texas. *Environmental Geology*, 51: 885–895, 2007. doi: 10.1007/s00254-006-0452-5.
- S. M. Wadham and G. L. Wood. *Land utilization in Australia*. Melbourne University Press, Melbourne, Vic, second edition, 1950.
- R. Wagner. Dryland salinity in the south-east region NSW, 2001. M. Sc. Thesis, Australian National University, Canberra, ACT.
- G. Wahba and J. Wendelberger. Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review*, 108:1122–1143, 1980.
- J. Walker, F. Bullen, and J. Williams. Ecohydrological changes in the Murray-Darling Basin. I. The number of trees cleared over two centuries. *Journal of Applied Ecology*, 30:265–73, 1993.
- Y. Wang, T. Zheng, Y. Zhao, J. Jiang, Y. Wang, L. Guo, and P. Wang. Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China. *Environmental Science and Pollution Research*, 20(12):8909–8923, 2013.

M. Wedel and W. A. Kamakura. *Market Segmentation: Conceptual and Methodological Foundations*. Kluwer Academic Publishers, Boston, 2nd edition, 1999.

P. Whittle. Estimation and information in stationary time series. *Arkiv fr Matematik*, 2:423–434, 1953. ISSN 0004-2080. URL <http://dx.doi.org/10.1007/BF02590998>. 10.1007/BF02590998.

C. C. Yang, S. O. Prasher, and R. Lacroix. Applications of artificial neural networks to land drainage engineering. *Transactions of the ASAE*, 39(2):525–533, 1996.

H. Yoon, S. Jun, Y. Hyun, G. Bae, and K. Lee. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1–2):128–138, 2011. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2010.11.002.

C. Zhu, Q. Luan, Z. Hao, and Q. Ju. Integration of Grey with Neural Network Model and its Application in Data Mining. *Journal of Software*, 6:716–723, 2011.

# Appendix A

## Step-change analysis of individual rainfall stations

Table A.1: Change-point analysis of individual rainfall series.

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
<b>Namoi</b>				
Gunnedah Pool	1897	94%	675.3	536.7
	1947	>99.5%	536.7	661.2
Goonoo Goonoo Station	no change			
Barraba Post Office	1947	97%	650.4	730.1
Wallabadah (Woodton)	1889	95%	653.4	980.0
	1895	98%	980.0	694.0
	1949	>99.5%	694.0	834.0
Manilla Post Office	no change			
<b>Weighted-average rainfall</b>	<b>1947</b>	97%	647.8	726.7
<b>Lachlan East</b>				
Golspie (Ayrston)	1949	92%	694.2	774.9
Bungendore (Gidleigh)	1948	94%	583.0	665.8
Bigga (Woolbrook)	1890	95%	592.6	958.6
	1895	97%	958.6	537.4
	1900	90%	537.4	723.6
	1947	99%	723.6	864.9
Millthorpe (Inala)	1947	99%	729.3	861.7
<b>Weighted-average rainfall</b>	<b>1947</b>	99%	668.6	775.4

Confidence level of >90% was used to indicate a significant change.

Table A.1: (continued)

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
<b>Lachlan Mid</b>				
Manildra (Hazeldale)	1947	93%	618.3	714.5
Yass (Derringullen)	1947	91%	662.2	758.8
Cowra Ag Research Station	1886	91%	476.3	769.6
	1895	>99.5%	469.7	550.8
	1947	98%	550.8	652.8
Murringo (Windermere)	1895	>99.5%	679.3	509.6
	1915	93%	509.6	696.1
	1977	96%	696.1	601.5
Canowindra (Canowindra Street)	1947	97%	552.3	644.6
Cudal Post Office	no change			
Boorowa Post Office	1895	94%	629.7	421.0
	1903	95%	421.0	625.8
<b>Weighted-average rainfall</b>	1895	99%	650.9	568.5
	<b>1947</b>	98%	568.5	668.7
<b>Lachlan West</b>				
Grenfell (Quondong Rd)	1886	92%	455.1	820.2
	1895	97%	820.2	508.4
	1916	97%	508.4	635.8
Wombat (Tumbleton)	1947	93%	653.1	738.7
Trundle (Murrumbogie)	1895	98%	580.2	399.2
	1947	>99.5%	399.2	524.3
Condobolin Retirement Village	1895	95%	509.2	391.5
	1947	>99.5%	391.5	482.5
Lake Cargelligo Airport	1947	97%	392.0	461.6
Goonumbla (Coradgery)	1895	96%	594.6	465.0
	1947	98%	465.0	565.5
Manildra (Hazeldale)	1947	93%	618.3	714.5
Forbes (Camp Street)	1886	96%	408.0	678.9
	1895	>99.5%	678.9	469.0
	1950	99%	469.0	568.2
Warroo (Geeron)	1895	97%	511.0	388.1
	1948	>99.5%	388.1	474.9
Barmedman Post Office	no change			
<b>Weighted-average rainfall</b>	<b>1947</b>	97%	569.9	648.3
<b>Murrumbidgee East (North)</b>				
Canberra Airport	1895	97%	729.0	541.3
	1947	97%	541.3	639.5
Fairlight Station	1895	99%	829.6	621.3
	1915	>99.5%	621.3	844.2
Hall (Lochleigh)	1895	>99.5%	729.4	560.2
	1915	94%	560.2	711.0
Yass (Derringullen)	1947	91%	662.2	758.8
<b>Weighted-average rainfall</b>	<b>1947</b>	96%	638.0	720.9

Confidence level of >90% was used to indicate a significant change.

Table A.1: (continued)

Section/subsection	Year change occurred	Confidence level	Pre-change average rainfall (mm/year)	Post-change average rainfall (mm/year)
<b>Murrumbidgee East (South)</b>				
Canberra Airport	1895	97%	729.0	541.3
	1947	97%	541.3	639.5
Fairlight Station	1895	99%	829.6	621.3
	1915	>99.5%	621.3	844.2
Hall (Lochleigh)	1895	>99.5%	729.4	560.2
	1915	94%	560.2	711.0
Michelago (Soglio)	1895	98%	626.0	448.9
	1913	>99.5%	448.9	633.2
Cooma (Kiaora)	1948	95%	484.0	558.9
Yass (Derringullen)	1947	91%	662.2	758.8
<b>Weighted-average rainfall</b>	<b>1947</b>	<b>&gt;99.5%</b>	<b>562.7</b>	<b>635.5</b>
<b>Murrumbidgee West</b>				
Tarcutta Post Office	1895	98%	687.3	548.6
	1915	93%	548.6	682.6
Mundarlo (Yabtree)	no change			
Old Junee (Millbank)	1950	94%	479.0	546.4
Grong Grong (Berembed)	no change			
Leeton Caravan Park	1947	92%	402.6	456.3
Adelong (Gundagai Street)	1947	90%	662.2	758.8
Henty Post Office	no change			
Bethungra (Retreat)	1895	97%	539.7	440.4
	1916	94%	440.4	565.4
Boorowa Post Office	1895	93%	629.7	421.0
	1903	96%	421.0	625.8
<b>Weighted-average rainfall</b>	1895	96%	625.7	508.7
	1916	94%	508.7	632.4

Confidence level = 90%, bootstraps = 1000. Some rainfall stations were used in two adjoining catchment sections (e.g. Glenn Innes PO in Border Rivers and Gwydir) to determine average rainfall for the section.

Table A.2: Change-point analysis of rainfall variability for individual rainfall stations.

Section/subsection	Year change occurred	Confidence level	Pre-change rainfall standard deviation (mm/year)	Post-change rainfall standard deviation (mm/year)
<b>Namoi</b>				
Barraba Post Office	1924	94%	113.9	196.3
Tingha (Crystal Hill)	1978	98%	193.1	455.0
	1988	98%	455.0	128.8

Bootstraps = 1000.

Data are shown only for stations where change was detected with a confidence level >90%.

Table A.2: (continued)

Section/subsection	Year change occurred	Confidence level	Pre-change rainfall standard deviation (mm/year)	Post-change rainfall standard deviation (mm/year)
<b>Lachlan East</b>				
Golspie (Ayrston)	1904	>99.5%	220.3	77.5
	1944	99%	77.5	242.2
Bungendore (Gidleigh)	1950	91%	129.1	344.2
	1984	98%	344.2	121.5
Bigga (Woolbrook)	1944	97%	134.4	263.0
Millthorpe (Inala)	1950	93%	165.0	298.6
<b>Average</b>	1944	97%	95.3	229.1
<b>Lachlan Mid</b>				
Manildra (Hazeldale)	1944	>99.5%	107.1	227.3
Yass (Derringullen)	1908	99%	160.2	85.1
	1930	92%	85.1	221.6
Cowra Ag Research Station	1944	99%	122.7	225.9
Canowindra (Canowindra St)	1944	93%	84.7	216.0
Cudal Post Office	1944	97%	134.3	204.4
<b>Average</b>	1944	98%	130.4	182.2
<b>Lachlan West</b>				
Grenfell (Quondong Road)	1926	93%	90.8	221.4
Condobolin Retirement Village	1950	91%	120.9	156.4
Lake Cargelligo Airport	1896	91%	179.6	50.6
	1926	99%	50.6	162.4
Goonumbla (Coradgery)	1972	93%	121.1	424.2
	1980	98%	424.2	136.1
Manildra (Hazeldale)	1944	>99.5%	107.1	227.3
Forbes (Camp Street)	1934	98%	87.9	223.4
Warroo (Geeron)	1944	98%	116.7	191.4
Barmedman Post Office	1928	99%	80.1	196.4
Murrumbidgee East (North)	see Peel			
<b>Average</b>	1950	93%	130.4	213.8
<b>Murrumbidgee East (South)</b>				
Michelago (Soglio)	1904	95%	192.6	61.3
	1950	>99.5%	61.3	212.7
Yass (Derringullen)	1908	99%	160.2	85.1
	1930	92%	85.1	221.6
<b>Average</b>	1950	98%	104.7	209.5
<b>Murrumbidgee West</b>				
Bethungra (Retreat)	1934	92%	96.4	190.8

Bootstraps = 1000.

Data are shown only for stations where change was detected with a confidence level >90%.

## **Appendix B**

### **Conference Publications**

# **Effect of rainfall on groundwater trends over the past century in fractured rocks of the New England Fold Belt in the Namoi catchment, NSW**

**Aleksandra Rancic**, NSW Department of Natural Resources, **Ian Acworth**, School of Civil and Environmental Engineering, The University of New South Wales, **Amrit Kathuria**, NSW Department of Primary Industries, **Gabriel Salas**, **Bill Johnston**, NSW Department of Natural Resources

Before 1987, the groundwater levels in fractured rocks of the Namoi sector of the New England Fold Belt were monitored infrequently. Therefore, this study exploited data collected for the standing water level (SWL) at the time of bore construction. Although patchy, such records extend back to the beginning of the twentieth century.

Based on these data, the median SWL was calculated for each year. Residual mass curves and a 21-year moving average low-pass filter were applied both to the annual SWL data and annual rainfall data for the Barraba Post Office rainfall station (Commonwealth Bureau of Meteorology no. 54003).

Change point analysis was used to detect abrupt changes in the mean. The analysis iteratively uses a combination of residual mass curve charts and bootstrapping. Rainfall change-point analysis for 1900 to 2003 showed two distinct periods, separated by an abrupt climate shift in 1947 that was characterised by an increase in average rainfall from 627 mm/yr to 730 mm/yr ( $P=0.03$ ). A sudden rise in groundwater began in 1949, with the median groundwater level rising from 23.2 m to 15.9 m below ground ( $P=0.01$ ). This finished in 1955, when the level stabilised at 11.7 m below ground ( $P=0.05$ ). Thus, there was a lag in the groundwater response to climatic change of two to eight years, and total rise in the watertable of 11.5 m.

An alternative methodology, in the form of a simple long-term moving average produced similar result, but in a more transparent form. This showed two states of dynamic equilibrium of rainfall and groundwater, separated by a mid-century hydrological shift. A time lag of around 5 years between changes in Barraba rainfall and groundwater was observed. The abrupt hydrologic shift in the middle of the century was replicated, as well as the deepening of the groundwater levels associated with a gradual drying of the climate in the past few decades.

The major clearing and associated change in land use in this area had finished by the end of the nineteenth century, and the short groundwater response suggests that any rise in groundwater levels triggered by clearing must have been completed before groundwater recording began. Groundwater levels in this system are currently showing a trend towards a fall, following a similar trend towards a decrease in rainfall.

# The relationship between the 1947 shift in climate and the expansion of dryland salinity on the western slopes of the Great Dividing Range in NSW

Aleksandra Rančić<sup>1,2</sup>, Ian Acworth<sup>2</sup>

<sup>1</sup>Department of Environment and Climate Change (DECC), NSW

<sup>2</sup>School of Civil and Environmental Engineering, The University of New South Wales (UNSW)

## Abstract

This paper summarises some findings from research undertaken for the latest Murray-Darling Basin Salinity Audit. An increase in dryland salinity outbreaks west of the Great Dividing Range in NSW occurred after massive land-clearing at the end of the nineteenth century. The effect of the hydrological imbalance produced by the land-clearing was not immediate, as the climate also shifted around 1894 and brought a substantial reduction in rainfall, so that the overall change in recharge was small. However, when the rainfall increased again after 1947, recharge increased significantly and the elevation of the water tables also increased in response. It took between five and seventeen years for the groundwater in fractured rock of New England and Lachlan Fold Belts to adapt to the new hydrologic regime and to achieve equilibrium. The response times were faster near the Divide and increased further west.

The last two decades of the twentieth century have seen a slight decrease in rainfall, and a corresponding reduction in the water table elevation. Drought conditions are present again since the turn of the century and we can only speculate at this time if this marks a significant change back to the longer-term drier climatic regime that existed in the first half of the 20<sup>th</sup> century. This drier climatic scenario should see a significant reduction in water table elevation and a consequent reduction in the area of salinised land.

The key findings of this paper are that: (1) the response time of the major groundwater systems to change in hydrologic regime appears to be shorter than currently conceptualised and increases with the distance from the Great Dividing Range; (2) that major shifts in climate have had the most important impact on land salinisation; and (3) the 1894 climate shift masked the true impact of land use changes on catchment hydrology during the early part of the twentieth century.

## Introduction

This paper summarises the finding of one of three studies undertaken to provide new, evidence-based science to underpin the NSW MDB Salinity Audit update in 2007. The aim of this study was to discover and examine the historic trends in groundwater levels in fractured rocks in the NSW portion of the Murray-Darling Basin over the period 1900-2003, and re-examine the role of climate and land-use changes as the major drivers for fluctuation in groundwater levels.

## Method

Historical land-use data and water level monitoring data for the region are unfortunately not available. The first bores drilled to monitor water levels in NSW fractured rock were installed in 1987 as part of the MDB salinity and drainage strategy. However, more than 33,000 production bores have been drilled within the study area, and information about the standing water level (SWL) at the time of drilling was recorded and stored in the state groundwater database. This information provides a valuable indication of water level change. This study uses the drillers' SWL records to estimate the trends in depth to groundwater. It is noted that these levels do not represent groundwater elevations as there are only very few locations where the surface elevation of the bore has been established. However, the SWL can still be used as an important indicator of groundwater response. The system behaviour was composed by agglomerating the individual records (point-data) over each section of

the study area from all 5,473 records held in the Departmental Groundwater Data Systems (GDS) database that are genuinely located in shallow fractured rock groundwater systems.

The study area covers sections of the Lachlan Fold Belt and the New England Fold Belt. Bores within these areas were subdivided into the major river catchments. Data from eight river catchments were analysed (**Figure 1**). The three southern catchments (Lachlan, Murrumbidgee and Murray) were further sub-divided as shown in **Table 1**.

The analysis is based on a comparison of groundwater and rainfall trends derived from the long term record of rainfall held by the Bureau of Meteorology. In particular, the similarities and deviations between the two have been studied, followed by a statistical analysis of the available annual datasets, to bring statistical rigour to the findings. Detailed statistical analysis has included: change-point, spectral, cross correlation and multiple regression analysis and the preparation of statistical models. The precise timing of the climate and standing water level (SWL) shifts were determined using change-point analysis. This statistical method uses residual mass curves to detect change points, by identifying the extremes in residual mass curves, and evaluates their statistical significance based on boot-strapping.

## Results

The most significant outcome of this study has been that trends in SWL have been following trends in rainfall across the study area for the whole period of available SWL data. Significantly, the response in groundwater level indicated by the initial SWL data appears to lag the rainfall change with the lag increasing with distance westward from the Great Dividing Range. The only noticeable deviation from this pattern of direct response to rainfall variation, that is consistent with an alternative model based upon water level response caused solely by land clearing, was found in the early record of the West Lachlan section, prior to 1937 (see **Figure 2**).

Inter-decadal climatic variation of the NSW region, probably linked to north-south energy transfer and movement of the dominant high-pressure ridge, changed around 1894 to a drier phase; back to a wetter phase in 1947; and probably back to a drier phase again in 2000. This approximately 50 year climatic signal is the most pronounced in the uniform rainfall zone of the Macquarie, Lachlan, and East Murrumbidgee areas. In every river catchment the groundwater table trend exhibits a delayed replica of the smoothed rainfall signal (**Figure 2**).

Groundwater has been in dynamic equilibrium with rainfall inputs during at least the last four decades. Change-point analysis shows that there was an average annual rainfall increase after 1947. In the last decade it has been decreasing in most areas. Groundwater levels in fractured rock have followed these trends. They rose, stabilised and currently are falling in most areas.

## Discussion

Groundwater flow systems in fractured rock show a rapid response to changed recharge conditions. The shortest lags, of up to half a decade have been observed close to the Great Dividing Range. They increase down the catchments, away from the Great Divide, and reach 16-17 years in the West Lachlan section (see

Table 1). The increase in lag time is interesting and requires further work. Three factors may impact on the lag time and explain the increase away from the higher parts of the landscape.

The first factor is the time needed for recharge to infiltrate and reach fractured rock groundwater. This will be a function of the depth to groundwater beneath the surface. A second factor that will determine the rate of rise is the amount of unconfined storage that requires filling (or emptying) before the equilibrium is achieved. A third factor may relate to the time taken for a pressure pulse to propagate through the system. Each of the three components can contribute to the highly variable lags within the study area.

During the wetter phase of the second half of the twentieth century, the increased recharge due to land-clearing was superimposed on the wetter climate phase. Together, they caused the watertable to rise significantly. Where the rising water table saturated surface deposits containing salts, many seeps and scalds were reported (Acworth and Jankowski, 2001). This research has revealed that climatic factors currently dominate watertable and land-salinity trends, with land-use change impacts masked by the overriding influence of climate.

The West Lachlan record (Fig. 2d) indicates a more extensive response to clearing during the early part of the twentieth century than occurred in the other systems studied.

### **Conclusions**

Perhaps the most significant conclusion that this work reveals is the absolute necessity to implement and maintain a regional monitoring network throughout the area. Much time and expense could have been avoided over the past 30 years if actual groundwater level data were available to inform debate.

Based upon the analysis of standing water levels during drilling it has been possible to demonstrate that groundwater flow systems in fractured rock show a much faster response to changed recharge conditions than previously appreciated. This is significant when compared with the former slower estimates included (implicitly) in the previous Salinity Audit. Current estimates as conceptualised within the Groundwater Flow System (GFS) Framework (Coram 1998, Coram et al. 2000) are also based on these slower estimates. It had been assumed that the groundwater rise could extend for several centuries after the hydrologic change, with current estimates assuming 30-50 years for a local system and 50-100 years for an intermediate groundwater system. The analysis reported above makes this slow response unlikely.

Recent and Paleo-climate records (Jacobson et al, 1991) demonstrate that the climatic variation, which brought to life so many salinity outbreaks over the last half of the century, was in fact very mild compared to climatic variability over longer time-scales. This suggests that the salt-build up in the landscape is not a mere consequence of a single, wet, longer-term climatic episode superimposed with the land-use change, but has its origin in the long sequence of preceding hydrologic perturbations.

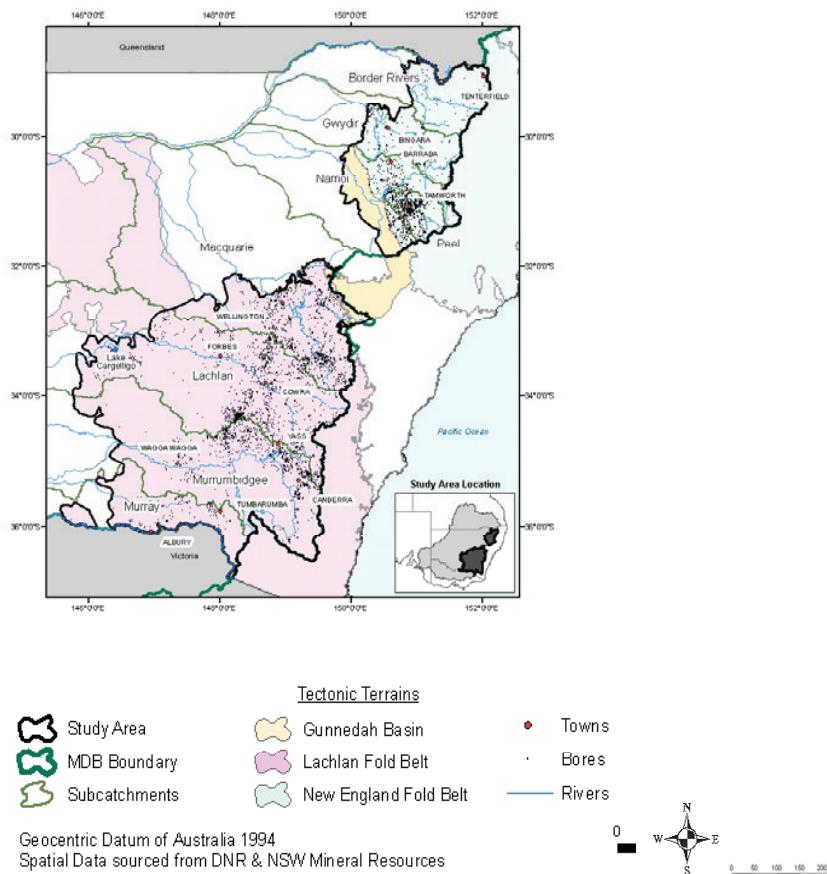
### **Acknowledgments**

Gratitude is expressed to Gabriel Salas, for his vision to initiate this work and generosity to see it all the way through, Amrit Kathuria, Geoffrey Bealle, Ann Smithson and William Johnston who are the co-authors of the Audit Report (Rančić et al. 2007), as well as Dugald Black, Jirka Šimunek, Gregory Summerell and John Cooper for their contribution. The New South Wales Salinity Strategy (2001-2005) and the Murray Darling Basin Commission contributed funding for this research.

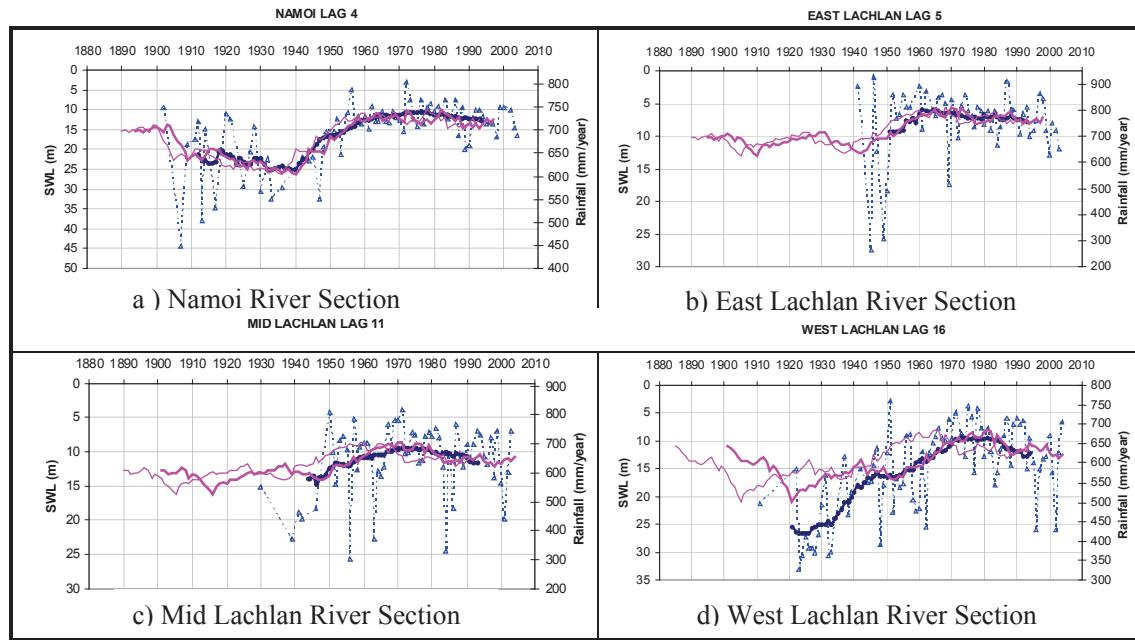
### **References**

1. Acworth, R.I. and Jankowski, J. 2001, *Salt source for dryland salinity – evidence from an upland catchment on the Southern Tablelands of New South Wales*. Australian Journal of Soil Research, **39.1**, 39 – 59.
2. Coram, J.E. (ed.) 1998, *National Classification of Catchments for land and river salinity control*. RIRDC Publication No. **98/78**, Rural Industries Research and Development Corporation, Canberra, ISBN 0 6425 7884 2.
3. Coram, J.E., Dyson, P.R., Houlder, P.A. & Evans, W.R. 2000, *Australian groundwater flow systems contributing to dryland salinity*. Report by Bureau of Rural Science for the Dryland Salinity Theme of the National Land and Water Resources Audit, Canberra (CD contains report, maps, and spatial data).

4. Jacobson, G., Jankowski, J. & Abell, R. 1991, 'Groundwater and surface water interaction at Lake George, New South Wales', *BMR Journal of Australian Geology and Geophysics*, **12**, 161-189.
5. Priestley, M. 1981, *Spectral analysis and time series*, Academic Press, New York, NY, USA.
6. Rančić, A., Salas, G., Kathuria, A., Johnston, W., Smithson, A. & Beale, G. 2007, *MDB Salinity Audit: Climatic influence on shallow fractured rock groundwater systems in the Murray-Darling Basin, NSW*, DECC 2007, ISBN 978 1 74122 589 1, DECC 2007/458.
7. Salas, G. & Smithson, A. 2002, 'Rainfall controls on standing water levels in a fractured rock area of central NSW, Australia', in *Balancing the Groundwater Budget, Proceedings IAH International Groundwater Conference*, Darwin, 12-17 May 2002, International Association of Hydrogeologists, Melbourne, Victoria.
8. Wagner, R. 2001, *Dryland salinity in the south-east region NSW*, M.Sc. Thesis, Australian National University, Canberra, ACT.



**Figure 1** Study area



**Figure 2** Trends in Rainfall and Standing Water Levels in form of the 21 year moving average filter. Median annual SWL indicated by blue triangles and dashed line, SWL trend in bold blue, rainfall trend in pale red and time-lagged rainfall trend in bold red. Rainfall is spatially averaged.

**Table 1** Lags based on Change-point analysis, Trend analysis and Cross Correlation. NSS=Not Statistically Significant ( $P < 90\%$ ). \* Lags statistically significant in the multiple regression analysis. # un-related shifts.

Division	Year of shift in		Delay (years)	Moving average fit (years)	Cross correlation
	Rainfall	SWL			
Border Rivers	1947	N/A	(-)	0-1	(-)
Gwydir	NSS	NSS	(-)	8	NSS
Namoi	1947	1949, 1955	2, 8	4-5	1*, 7, 10*
Peel	1947	1952	5	3	3*
Macquarie	1947	1953	6	5-6	1*, 2*, 7*
East Lachlan	1947	1952	5	5	NSS
Mid Lachlan	1947	NSS	(-)	11	19*
West Lachlan	1947	1963	16	16	1, 16*, 17*
Murrumbidgee NE	1947	1991	#	N/A	0*, 1*
Murrumbidgee SE	1947	1988	#	N/A	3*
Murrumbidgee E	1947	1965, 1997	#	N/A	0*, 1*
Murrumbidgee W	NSS	1937	(-)	11	19*
East Murray	NSS	NSS	(-)	14	(-)
West Murray	NSS	NSS	(-)	27-30	(-)