FLOW AND SALINITY IN THE MURRAY RIVER

Study of flow and salt loads in the Murray River of the Murray-Darling Basin

Ingrid van den Brink

Bachelor Thesis 17-08-2015

UNIVERSITEIT TWENTE.

Flow and Salinity in the Murray River

Study of flow and salt loads in the Murray River of the Murray-Darling Basin

Australian National University, Canberra

17-08-2015

UNIVERSITEIT
TWENTE.

Colophon

Date
12-08-2015

Student
Ingrid van den Brink
i.m.vandenbrink@student.utwente.nl
+31649734865

The Netherlands:
Witbreuksweg 401-209
7522 ZA Enschede

Supervisor University of Twente
M. Pahlow
m.pahlow@utwente.nl
+31 53 489 4705

Supervisor Australian National University
B. Croke
Barry.Croke@anu.edu.au
+61 2 6125 0666

iCAM, Bldg 48a, Linnaeus Way
The Australian National University
Canberra ACT 0200 Australia

Information about the Fenner school
The Fenner School is unique in Australia. There are very few places in the world where economists and hydrologists, historians and ecologists, foresters, geographers and climatologists work together. The Fenner School consists of several research groups (DirectorFennerSchool, 2015). This research will be part of the iCAM (Integrated Catchment Assessment and Management Centre) (iCAM, 2015)

Preface

This report is my Bachelor Thesis for the study Civil Engineering at the University of Twente. After finishing my Bachelor, I intend to start the master 'Water engineering and Management'. For this reason I was interested in doing research on flow modelling, to learn more about modelling and water processes. I have done my bachelor thesis at the Fenner School of Environment & Society, College of Medicine, Biology & Environment at the Australian National University in Canberra (ANU). More information on the Fenner school is given in the Colophon. I am glad the ANU gave me the opportunity to do research on flow modelling using a top-down approach.

The Murray-Darling Basin, where the Murray River is located, suffers from enormous salinity problems which affect the flora and fauna. Also there are consequences for the quality and use of for example irrigation water. The river processes are very complex due to, for example, tributaries, lakes, dams, and anabranches. This research is the first step towards the development of a salt model which will include the complexity of the river. For this thesis a top-down approach is used for modelling the flow in a specific reach in the Murray River. The top-down approach means that research starts by data analysis of flow and salinity data and that the model is built based on these analysis. Further, there is a lack of knowledge on uncertainties in these data used for model development and since the river processes are very complex, it is important to also take these uncertainties into account.

I would like to thank the people who helped me to accomplish this thesis. First of all, I would like to thank Mister Croke from the ANU for teaching me many things about the top-down approach and all the different aspects which are important during the development of a model (including uncertainties). Further, I would like to thank Mister Pahlow from the University of Twente for giving me very good feedback and advice about the research itself, but especially about the process of doing research. At last, Martijn Booij for arranging this project.

Ingrid van den Brink
August 2015, Canberra

Summary

The Murray-Darling Basin is Australia's largest river system. It covers 1,059,000 square kilometres (MDBA, sd) and includes a series of interlinked sedimentary aquifers. The Murray River is the main river in this basin. Much of the groundwater underlying the basin contains of significant amounts of salt. The salinity problems are enormous and are affecting the flora, fauna, irrigation and drink water.

To reduce the salinity, many regulations are employed, such as salt interception schemes, injection of fresh water, artificial flooding and dams. Salinity management requires an understanding of catchment data and processes in the Murray-Darling Basin (Fitzpatrick et al., 2007). The routing model used to investigate the effect of salinity regulations is MSM-BIGMOD. It calculates the salt loads using salinity and flow data from sources such as tributaries, anabranches, salt interception schemes etc. However, there is a lack of information about three aspects. First, there are unquantified sources which are not taken into account since the river processes are too complex. The model refers to this as 'unaccounted salt loads'. Second, there is a lack of information about uncertainties in the input flow and salinity data. Third, the uncertainties in the parameters and model structure are unknown since the model has too many parameters for a proper uncertainty analysis. To get a better insight in the uncertainties and river processes, it is important to develop a simplified conceptual model using a top-down approach. The long term objective is to look at the flow and salinity data including uncertainties, to understand the signals and see if the signals support the processes that are included in MSM-BIGMOD. This research focusses on the first two stages:

Conceptualising and testing of a flow model for a particular reach of the Murray River based on data analysis and quantification of uncertainty of the input flow and salinity from nearest upstream sites.

Before conceptualizing the flow model, a flow and salinity data analysis is needed to obtain a better understanding of river processes. This analysis shows the complexity of the river processes due to tributaries, anabranches, groundwater recharge and discharge and floodplains. The differences between sites are varying from 24% to 79% caused by the river processes and implemented regulations. Second, this research is about identifying and giving advice how to reduce the different uncertainties in the salinity and flow data. The flow data is obtained using a rating curve instead of direct measurements. The salinity data includes several assumptions or 'rule of thumbs'. To convert the salinity data to salt load, a conversion factor K is used. In literature, a range for factor K is found, varying from 0.45 to 0.9. Changing the factor with 10%, the salt load will change with 18%. Most researchers and decision-makers are making the assumption that the input data is error free, but only the uncertainty of the factor K on the output can already be 180%. Further, this assumption cannot be made since the influence of the other uncertainties on the modelled output is unknown. To get a better insight in this and to reduce the uncertainties, more information about the river cross section, ionic composition of the river, amount of rating curves, parameters in rating curves and measurement equipment is needed. At last, the data and uncertainty analysis are used to choose the specific reach for conceptualizing the flow model. The reach from Lock 9 to Lock 5 is chosen since this reach had less anabranches and tributaries than other reaches. This is important since a simple model structure is needed to understand uncertainties and river processes. The model structure which gives the highest objective function values (NSE > 0.9, RVE < 0.15) contains of a single store with only one flow path. However, when comparing the modelled output with the flow output, there are at least three additional modules needed to cover the river processes of the reach. The first is a floodplain module which covers the flow peak when the water level reaches a specific height where it overflows the floodplain, the second is an additional groundwater module and the third are the additional tributaries and anabranches. Further research is very important to make sure the most reliable model is used to study which management strategies are most effective to reduce the salinity in the Murray River.

Table of context

Glossary

Accounted salt load inflows
A MSM-BIGMOD term used for salt inflows to the Murray River from tributaries and drains which are quantified using flow and salinity data (Telfer et al., 2012).

Anabranches
Branches of river that leave the main stream and re-join it downstream (Telfer et al., 2012).

Discharge
Water entering the river system.

Floodplain
Land adjacent to a stream or river that stretches from the banks of its channel to the base of the enclosing valley walls and experiences flooding during periods of high discharge. It includes the floodway, which consists of the stream channel and adjacent areas that carry flood flows, and the flood fringe, which are areas covered by the flood, but which do not experience a strong current (Telfer et al., 2012).

Gaining floodplain
Reaches where the regional groundwater system is discharging into the floodplain alluvium (Telfer et al., 2012).

Gaining stream
Reaches of river where groundwater is discharging from the floodplain alluvial sediments into the river (Telfer et al., 2012).

Losing floodplain
Reaches where the groundwater flow is from the floodplain sediments to the regional groundwater system (Telfer et al., 2012).

Losing streams
Reaches of river where the river is losing water to the floodplain alluvia sediments (Telfer et al., 2012).

Reach
Part of the river which can be distinguished by specific river processes.

Recharge
The process of aquifer replenishment, usually from rainfall, irrigation accessions and losses from surface water bodies such as rivers and lakes; water entering the groundwater system (Telfer et al., 2012).

Salinity data
From this point the salinity data will be referred to as the salinity (EC) in the Murray River. The salinity indicates the accumulated amount of salt in the water.

Salt load
To measure the salt load first the salinity (EC [µS/cm]) is converted using a factor k = 0.55 which results in a salt concentration (mg/L). The salt load [kg/d] is measured by the product of flow [ML/d] and salt

concentration [mg/L]. Salt load is the amount of dissolved salts in water carried past a designed point over a specified period of time and is usually expressed as tonnes per day.

Slow flow and quick flow

The slow flow interacts with the bedding of the river. It is used to infer the groundwater contributions to slow flow (Ivkovic et al., 2014). However, this division in quick and slow pathway may be an artefact of representing the transport mechanism as a combination of exponentially decaying stores, rather than physical processes. The quick flow component might have a shorter time constant than the slow flow component (Ivkovic et al., 2014)

Through flow floodplain

Reaches where the regional groundwater flow lines show that groundwater flows beneath or through the floodplain. In these reaches, the floodplain alluvium is potentially gaining water from the upgradient side, but is losing water to the regional groundwater system on the downgradient side (Telfer et al., 2012).

Unaccounted salt load inflows

A MSM-BIGMOD term used for salt inflows to the Murray River from all groundwater inflows and unaccounted surface water discharges. Many discharges to the river are either un-regulated or not measured, such as discharges from evaporation basis and outflow from anabranches and lagoons (Telfer et al., 2012).

Tributaries

A stream that flows to another stream, the Murray River in this research.

List of figures

List of tables

1 Introduction

1.1 Background

The Murray-Darling Basin is Australia's largest river system. It covers 1,059,000 square kilometres (MDBA, sd) and includes a series of interlinked sedimentary aquifers. The basin's waterways sustain over two million people and hundreds of species of native fauna and flora; its rivers are the primary source of water for irrigation, municipal water supply and recreation (Burnell et al., 2013). The Murray River is the main river in the Lower Murray-Darling basin which in turn is part of the Murray-Darling Basin (Bekesi et al., 2014), as shown in Figure 1. The focus of this research is on the Murray River.

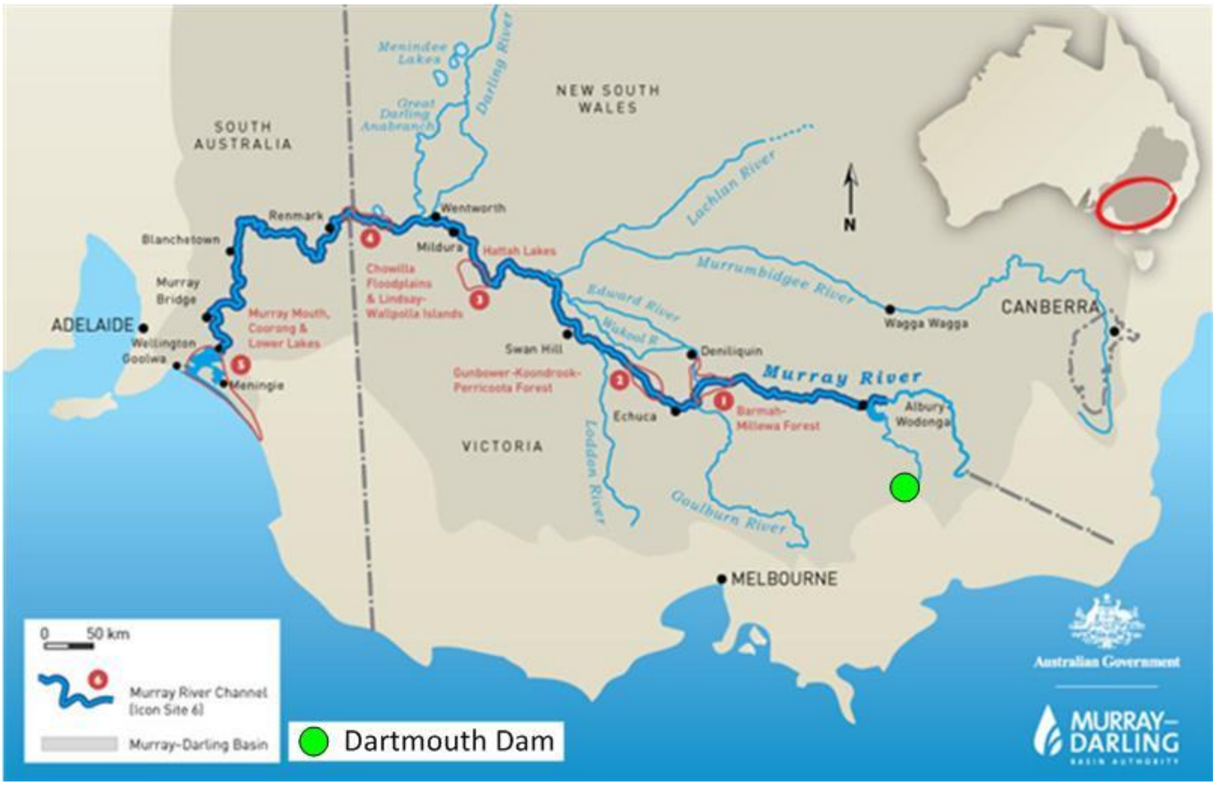

Figure 1 Murray River in the Murray-Darling Basin (Authority, 2011)

Much of the groundwater underlying the Murray Basin contains significant amounts of salt. About 30% of the basin contains more than $1.4x10^4$ mg/L total dissolved solids (TDS), and 2% has salinities above that of sea water. The salinity affects most water uses, such as irrigation and drinking water, and the environment. It also represents a threat to the environmental circumstances of floodplains, wetlands and irrigated crops (Burnell et al., 2013). There are three different types of salinity which are important to understand the salinity problems (NSW, 2013):

- *Dryland salinity*

  Native vegetation is effective at using most of the water entering the soil profile from rainfall, allowing only a small proportion of rainfall to reach the groundwater system (recharge). Since European settlement, the native vegetation is replaced with crops and pastures which have shallower roots and different seasonal growth patterns. These plants use less water, resulting in more water percolating from beneath the root zone into the groundwater. This extra groundwater results in a rising groundwater table which moves dissolved salts to the surface. In some cases this results in white salt on the soil surface, particularly in low-lying areas such as rivers, streams and wetlands (Audit, 2000).

- *Irrigation salinity*
  This occurs when there is a localised rise in the groundwater level caused by the application of large volumes of irrigation water (NSW, 2013).

- *River salinity*
  River salinity is the concentration of dissolved salts in a stream, river or lake (NSW, 2013). In the lower part of the basin, most groundwater discharges to the floodplain of the Murray River and transfers significant salt loads into the river. (Bekesi et al., 2014). To understand the river salinity in the Murray River, it is important to know something about the instream processes in the Lower Murray-Darling basin. Annual rainfall averages approximately 300 mm/year over the Lower Murray-Darling Basin and is relatively evenly distributed throughout the year (Bekesi et al., 2014). Average evapotranspiration at approximately 2000 mm/year, greatly exceeds rainfall in most months. This suggests that groundwater recharge from local rainfall may be small and only occurs during wet periods. Groundwater recharge is not uniformly distributed in time. Gaining and losing conditions change frequently along the Murray River, depending on current and past river levels, lateral and vertical groundwater flow into the floodplain sediments, and seasonal changes in evapotranspiration in the floodplain. It is difficult to recognise when the river is gaining or losing without careful analysis (Bekesi et al., 2014). The focus of this research lies on *river salinity* in the Murray River.

In order to manage the problem and to protect the ecology and biodiversity along the river, a range of management strategies are being employed including the development of salt interception schemes (SIS), injection of fresh water and artificial flooding (Fitzpatrick et al., 2007). Salt Interception Schemes (SIS) are the most viable solution to instream salinity problems in the Murray Basin as it can be implemented in a short time frame and can operate for decades (Telfer et al., 2013). A typical SIS is made up of a line of relatively shallow bores that intercept saline groundwater flow adjacent to or within the river floodplain before it gets a chance to enter the river (Bekesi et al., 2014). For both farmers and government not only the reduction of river salinity is relevant, but the reduction of dryland salinity as well. For both it is important to make the most effective choice for implementing a strategy. Nigel Hall et al. (2004) describe the use of spreadsheet models to help farmers and their advisors to make decisions on land and water use to manage dryland salinity (Hall et al., 2004). Sadoddin et al. (2005) have developed a new tool for integrated management of dryland salinity. A Bayesian decision network was used to demonstrate the impacts of various management scenarios on terrestrial and riparian ecology taking into account the economic, ecological, social and biophysical system components. The Stage Two Report outlines the general principles for managing and preventing dryland salinity, like increasing water use in discharge areas (Sadoddin et al., 2005).

Salinity management requires an understanding of catchment data and processes in the Murray-Darling Basin. Methods to monitor the temporal state of river and particularly river-groundwater interactions, have been in place for many years now. There are several methods to collect the data needed for routing flow and salinity models. These models are existing as well, but since there is a lack of information about the river processes in the Murray-Darling Basin and since there is a lack of information about uncertainties in the input data and models, the models and methods to monitor the river still not have the capacity to define variability at a resolution appropriate for developing effective management strategies (Fitzpatrick et al., 2007).

1.2 State of the art

Before decision-makers can develop effective strategies to reduce the salinity in the Murray River, the used models need to be more reliable. There are different aspects which are important to improve the management strategies. In the State of the art information is given about research which has already been done. The first aspect is the possible methods to collect data. The second contains information about the available routing models and the third aspect is about uncertainties in the input data and the model structures. This information leads to the research gap where this research is about.

1.2.1 Collecting data of the Murray River

Monitoring the Murray River involves collecting flow and salinity data. Flow is measured in ML/d and salinity is measured in electrical conductivity [EC]. The flow data is measured using a rating curve. This rating curve converts the observable quantity stage height into the discharge rate (NationalWeatherService). The salinity data can be measured in several ways. One method is Run-of-River (RoR) surveys which involves the electrical conductivity [EC] measurement of river water (Fitzpatrick et al., 2007). The EC is measured during RoR surveys with one km intervals using a boat equipped with a pump, continuous flow-through cell and an EC meter logging data on a laptop computer. EC and river flow are normally inversely related. For this reason RoR surveys are completed at low (less than 4000 ML/day) and steady river flows when groundwater discharge to the river may be considered constant (Burnell et al., 2013). Another way of collecting salinity [EC] data is by using monitoring sites along the Murray River. The frequency of salinity recording varies from continuous to daily, weekly and monthly. Continuous monitoring of in-stream salinity covers more than 50 monitoring sites. Generally, these sites are spaced between 20 and 30 kilometres apart at the start and end of river reaches (Jin & Close, 2012). More detailed information about collecting flow and salinity data is given in section 5.1.1.

1.2.2 MSM-BIGMOD model of Murray River

The MSM-BIGMOD model of the Murray River is a comprehensive flow and salinity routing model, used to assess the impacts of potential changes in river management on river flow and salinity levels. This model begins with the inflow from Dartmouth Dam (Figure 1) and incorporates tributaries, storages, weirs, irrigation and urban diversions, salt interception schemes, drainage diversions and wetlands. The model operates through a process of hydrological routing, which involves dividing the river into reaches, each with different flow parameters and variation due to the different inputs (Ravalico et al., 2011). The MSM-BIGMOD incorporates the salt inflow from most sources, “accounted salt load data”, but there are unquantified sources which are not taken into account. This model refers to this as “unaccounted salt inflows”. Those unaccounted salt loads need to be added to balance the salt budget in the model (Telfer et al., 2012).

Floodplain Salt Conceptual Model

The 2007-08 IAG (Independent Audit Group) report includes different recommendations for The Murray Darling Basin Authority (MDBA). The IAG-Salinity is very concerned about the potential for a significant rise in salinity levels which are expected to follow the next flood after a drought. This happened at the end of the Millennium Drought. This drought was broken by flood events in late 2010, with floods scattered through 2011 into early 2012 across large parts of the Murray Darling Basin. The salt mobilised during the flood recession (the period after a flood peak when river flow continues to decrease). They recommend developing a conceptual model of flood recession salt mobilisation in the floodplains in preparation for the next high flow events (Shepherd et al., 2009).

In 2012 the MDBA presented a report which includes a literature review of previous studies of Murray River floodplain processes and the development of a Floodplain Salt Conceptual Model. The conceptual model consists of three elements: regional, floodplain and river. The regional elements include sources of salt for the floodplain landscape and measures that reduce the salt inputs. The floodplain elements address the storage and mobilisation of salt within the floodplain and the surface waters. The river elements address the salt inputs, river flow and river salinity. This model is not a floodplain salt predictive model, but encapsulates the existing understanding of flood recession salt mobilisation. The MSM-BIGMOD data is used in this model to provide new insights into floodplain salt delivery processes.

1.2.3 Uncertainty in input data and model

In order for a model to be reliable and credible, the modeler must be conscious of the uncertainties involved. In particular, the modeler must address the uncertainty that model assumptions are accurate, and hence to what extent model results will match reality. Doing so will in turn help to minimize the risk that decisions based on the model may lead to adverse impacts because of what the modeler did not or could not know. In general, Jin et al. (2010) refer to three principal sources contributing to model uncertainty in conceptual models: errors associated with input and calibration data, improper model structure, and uncertainty in the model parameters (Jin et al., 2010). These three *sources* correspond with those determined by Guillaume et al., (2010) (Guillaume et al., 2010). Errors in the input may result in errors in estimated parameters and hence errors in simulated discharge (Tillaart, 2010). Jake et al., describes the necessity of a systematic approach to minimize the risk of ignoring uncertainties, for example by checking through each potential source of uncertainty or using a decision tree (Guillaume et al., 2012). The influence of the three sources on the modelled output need to be determined. This can be done by the use of performance indicators. The role of performance indicators is to give an accurate indication of the fit between a model and the system being modelled (Croke B. et al., 2012).

<u>Input data MSM-BIGMOD</u>

The input data is the flow and salinity data. Commonly the uncertainty of flow records is not quantitatively assessed, so the data are used with the implicit assumption of being error free (Chiew et al., 2008). This assumption might be incorrect because the errors in streamflow data are possibly quite large because flow itself is usually not directly measured but rather derived from a proxy of stream height (stage) (Herschy, 2009). As described in the section 'Collecting data of the Murray River' this method is used in the Murray River. Tomkins studied 36 gauges in the Namoi River shown in Figure 2 to provide information on the uncertainty of streamflow data used in rainfall-runoff and river models. The Namoi River is located in the Murray-Darling basin, but is located further North East from the Murray River. However, this is the only research found about the effect of rating curves in the Murray-Darling basin. The uncertainty in rating curve and the reliability of flows are highly variable over time and stage within each gauge and between gauges. Tomkins looked at the proportion of gauges which exceeds a deviation of 10% when comparing empirical data. Of the analysed gauges, 39% had a deviation exceeding between 21% and 50% (Tomkins, 2014).

Figure 2 Namoi Catchment (Restoring the Balance in the Murray-Darling Basin, sd)

<u>Uncertainties parameters and model structure MSM-BIGMOD</u>

Ravalico et al., (2011) state that sensitivity analysis of MSM-BIGMOD is important, given that decisions are made about management of the Murray River based on outputs from the model. The large number of model inputs and parameters arising from the inclusion of the many tributaries, storages, drains, and diversions pose a challenge for traditional sensitivity analysis methods. Ravalico et al., (2011) used the Management Option Rank Equivalence (MORE) method developed especially for use with complex models used for decision-making. This method is based on the premise that potential management options are ranked based on model output. At the end Ravalico et al., (2011) concludes that in order to gain a better understanding of the different contributions of each parameter it would be beneficial to perform further sensitivity analysis on the model (Ravalico et al., 2011).

1.3 Research gap

Models can structure and evaluate our knowledge to help anticipate future consequences. However, models are fallible and predictive uncertainty needs to be addressed systematically for modelled outputs to reliably support decision making (Guillaume et al., 2010).

There are different management strategies which are implemented to decrease the effect of salinity. There are also different methods to obtain data about the Murray-Darling basin. In addition, there is a model (MSM-BIGMOD) which calculates the salt loads by using the salinity and flow data. The MDBA is also responsible for the development of a Floodplain Conceptual Salt Model which can assist with improving the current understanding of the sources of salt, the storage locations in the floodplain landscape, the mobilisation processes, the transport pathways to the river, and the river salinity impacts (Telfer et al., 2012).

There is not much information known about the uncertainties in the input flow and salinity data. Further, due to the curse of dimensionality (MSM-BIGMOD has many dimensions, in this case parameters (Ravalico et al., 2011)) it is very difficult to do a proper uncertainty analysis. It is necessary to have a good insight in these uncertainties to make reliable management decisions to reduce the salinity problems. Apart from the uncertainties, the processes in the Murray River are very complex. Due to the complexity of the MSM-BIGMOD model it is difficult to discover which river processes are included and which are not. Obtaining a good insight in these river processes is important to reduce the unaccounted salt load which still needs to be added to balance the total salt load. At least, there is no model which can predict the floodplain salt. Developing such a model is too premature at this stage, because significant major data sources have not been evaluated (Telfer et al., 2012).

2 Research objectives and Research questions

Based on the research gap (section 1.3) the research objective is formulated. The research objective leads to several sub-questions.

2.1 Research objective and limitations

To get a better insight in the uncertainties and river processes, it is important to develop a simplified, conceptual model. The long term objective is to look at the flow and salinity data including uncertainties, to understand the signals and see if the signals support the processes that are included in MSM-BIGMOD and therefor the Floodplain Salt Conceptual Model. This research focusses on the first two stages, and will build from this to the third step in further researches.

The research objective for this research is:

Conceptualising and testing of a flow model for a particular reach of the Murray River based on data analysis and quantification of uncertainty of the input flow and salinity from nearest upstream sites.

<u>Limitations</u>

In this research a conceptual flow model will be developed by looking at the signals obtained from the flow data and salinity data. Although developing the salt model is for later research, it is important to keep this development in mind during data analysis (including uncertainties) as this will ease its development in the future. Further, because of the same reason, choosing the 'particular reach' is based on both flow and salinity data analysis.

2.2 Research questions

To reach the purpose of this research it is needed to find an answer on the following main question:

What is the structure of a conceptual model suitable for use in modelling flow at a particular reach along the Murray river in the Murray-Darling Basin?

The question can be split up into five sub questions:

1. What for information can be obtained from the connection between flow and EC data at the different sites along the Murray River over time?
2. What are the contributors to uncertainty in flow and EC measurements at observation sites along the Murray River?
3. How do the uncertainties identified in and other uncertainties propagate to the uncertainty in the salt load L_s?
4. Based on the data analysis, which reach with accompanying data is most suitable for conceptualising a flow model?
5. What would a top-down model for estimating the flow at the selected particular site look like?

2.3 Research diagram

Figure 3 shows the research diagram of this study. It explains the relation between the different sub-questions. Developing a conceptual flow model involves calibration to obtain the parameter values and validation to test the reliability of the model. It also involves an uncertainty analysis about the parameters and model structure. In order to develop a top-down flow model for a specific reach (sub-question 5), the most suitable reach need to be selected (sub-question 4). The reach is chosen according to the results of the data analyses which determines the connections between the raw EC and flow data over time (sub-question 1). Further, the uncertainties in the data input (EC and flow data) (sub-question 2) and therefore in the salt loads (sub-question 3) are determined. The salt loads are the accounted and unaccounted salt loads. The unaccounted salt loads obtained in this research are compared with the unaccounted salt loads from MSM-BIGMOD. The salt load is obtained by the product of the concentration [S] and the associated flow [Q]. The concentration can be calculated by converting the EC data by using the parameter K.

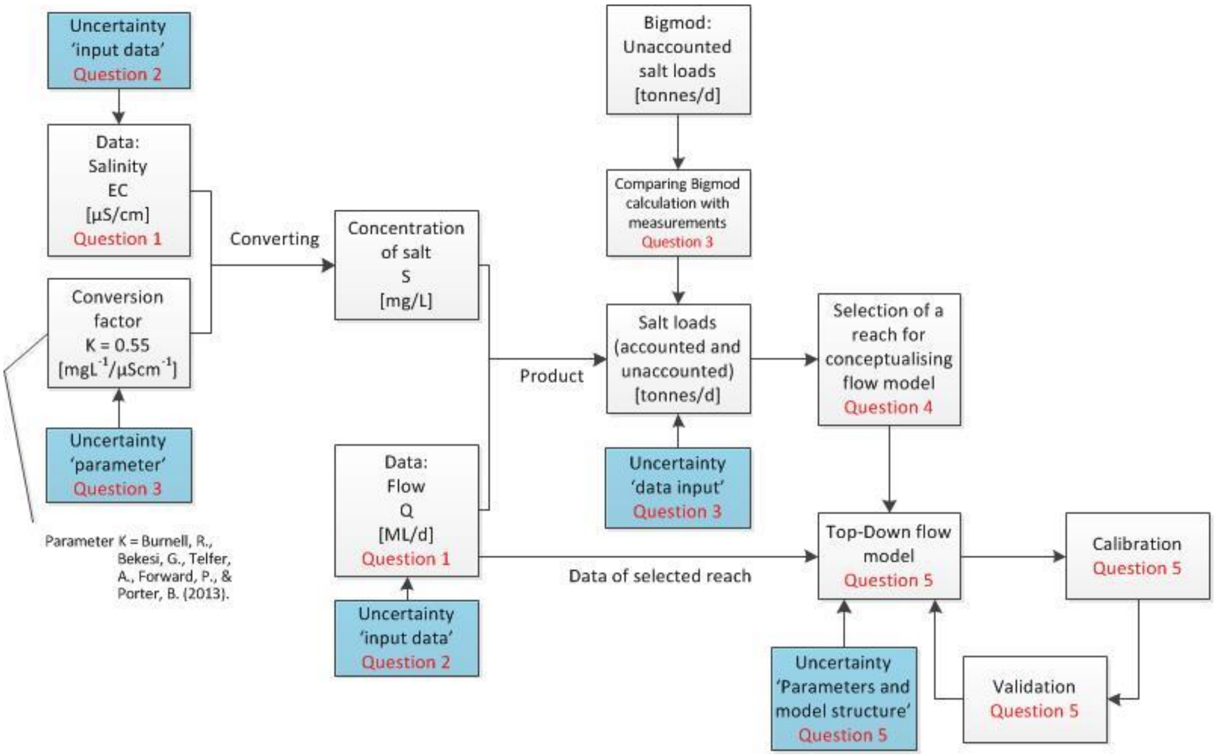

Figure 3 Research diagram which explains the link between different sub-questions

2.4 Report structure

The research diagram explains the relations between the five sub-questions. The report structure gives the structure which is used to present results of the sub-questions. The five sub-questions are divided into three different chapters as shown in Figure 4; Chapter 4 "Data analysis", Chapter 5 "Uncertainties" and Chapter 6 "Model structure".

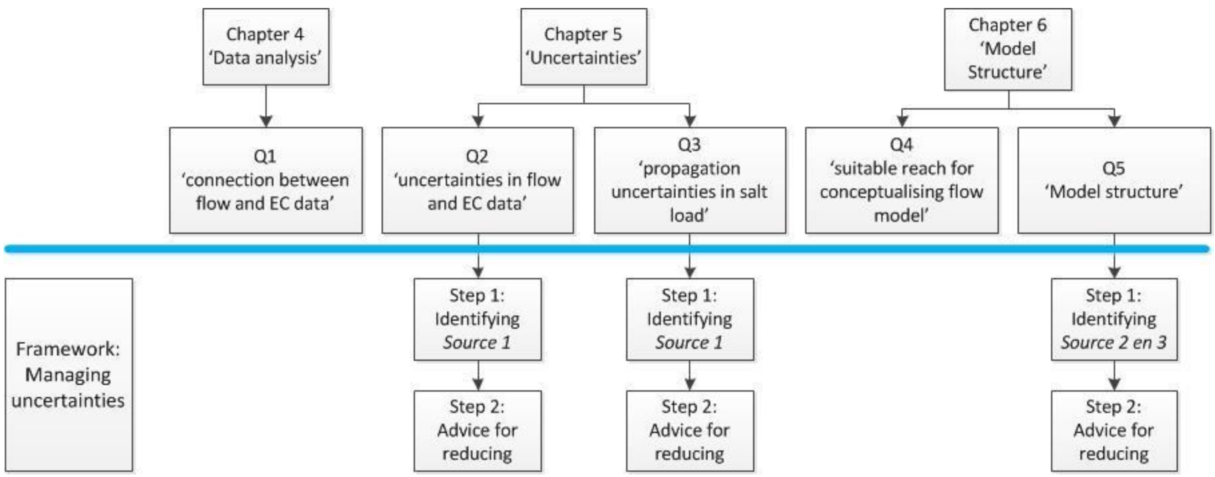

Figure 4 Framework managing uncertainties

As shown in section 2.3 and Figure 4, the uncertainties are an overall subject interrelated with the sub-questions 2, 3 and 4. This results in a general ‘Framework to manage uncertainties’.

2.4.1 Framework to manage uncertainties

Guillaume et al., (2010) identifies tasks required to manage uncertainty related to the consequences of decisions. Tasks are organized within a framework to guide the selection of methods, which can help ensure that uncertainty is treated systematically, coherently and transparently during analysis and decision making. This research uses the two steps from that framework; Step one ‘Identifying the uncertainties’ and step two ‘reducing the uncertainties’. Figure 4 depicts the framework which is adapted in both Chapter 5 ‘Uncertainties’ and Chapter 6 ‘Model Structure’. In Chapter 3 ‘Methodology’ the framework is described in more detail per sub-question.

3 Methodology

Chapter 3 describes the methodology to obtain the results for the five sub-questions. The same structure is used in the results. Section 3.1 describes the methodology for obtaining the results described in Chapter 4 'Data Analysis', section 3.2 for Chapter 5 'Uncertainties' and section 3.3 for Chapter 6 'Model Structure'.

3.1 Methodology: Data analysis

To obtain the results from sub-question 1 'Connection between flow and EC data', several steps were needed. First the available data were investigated to configure if it is usable for conceptualizing a flow model for a reach in the Murray River. Second, the flow and EC data were compared. To get a better understanding of the relationship between both data, it is important to compare the salt load as well. The salt load is determined through observations of flow [mL/d] and salinity [EC] (Telfer et al., 2012). To impute the salt load, the salt concentration is measured. The salinity [EC] is converted to the unit salt concentration [mg/L]. This may be approximated by: $mgL^{-1}/\mu Scm^{-1}$

$$S\ [mg/L] = K * EC\ \ [\mu S/cm]$$

In this research the value $K = 0.55mg\,L^{-1}/\mu Scm^{-1}$ is used (Burnell et al., 2013). In Chapter 5 the uncertainties in using this factor K are obtained. After converting the salinity from EC to concentration, the accounted salt load are determined. The salt loads are the product of flow and salinity from tributaries, anabranches and the river. The salt load consist of the accounted and unaccounted salt load. The differences between the flow, EC and salt load at different sites over time are explained on base of literature review. This literature review gives insight in the study area and the different catchments of the Murray River with all the anabranches and tributaries. It also reflects the losing and gaining areas around the Murray River. At the end the data analysis is used to decide on an appropriate reach for developing the flow model.

3.2 Methodology: Uncertainties

As described in section 1.1 'Background', in general, Jin et al. (2010) refer to three principal sources contributing to model uncertainty in conceptual models; errors associated with input and calibration data, improper model structure, and uncertainty in the model parameters (Jin et al., 2010). These three *sources* correspond with those determined by Guillaume et al., (2010) in the first step 'Identifying' of the framework to manage uncertainties (Guillaume et al., 2010). The uncertainties in Chapter 5 represent the uncertainties associated with the input data, source 1.

3.2.1 Q2: Uncertainties in flow and EC data

The second sub-question is about the uncertainties in the data input (salinity and flow measurements). The potential sources and possible values of the uncertainties will be identified by doing a literature review.

3.2.2 Q3: Uncertainties in salt load

The uncertainties in the salt load consist of the uncertainties in the conversion factor $K\ [mg\,L^{-1}/\mu Scm^{-1}]$ to obtain the salt load and other uncertainties during calculation of the salt load. Section 3.1 explains the calculation. These uncertainties are obtained and qualified by doing a literature review.

Another important aspect are the unaccounted salt load which are used in MSM-BIGMOD to balance the salt load. This unaccounted salt load explains there is another uncertainty in the salt load between different sites. The routed salt load at a site has to be the same as the measured salt load, but in many situations this is not.

accounted salt load
The accounted salt load is determined from tributaries and drains, and the extraction for consumptive use (irrigation, stock and domestic uses) (Telfer et al., 2013).

unaccounted salt loads
Unaccounted salt loads refer to inflows from unquantified sources including groundwater flow into the river and surface water inputs from unmonitored tributaries and anabranches (Telfer et al., 2013).

Data is available of the adjusted monthly averages of daily unaccounted salt inflow in MSM-BIGMOD. This data set is referred to as 'unaccounted salt load MSM-BIGMOD'. The flow, EC and therefore salt load data used during this research will be referred to as 'salt load observed data'. The two datasets are compared by comparing the 'unaccounted salt load MSM-BIGMOD' added to the model between site X upstream and site Y downstream (including Y) with the difference between the 'salt load observed data' at site X upstream and site Y downstream.

3.3 Methodology: Model Structure

Section 3.3 contains the methodology to obtain the results for the fourth and fifth sub-question. It also involves the uncertainties for source 2 and 3 'improper model structure and uncertainty in the model parameters' for the first step 'identifying' of the framework as mentioned in section 3.2 (Jin et al., 2010).

3.3.1 Q4: The most suitable reach for conceptualizing a flow model

Depending on the results from the data analysis, a reach will be chosen which is most suitable for conceptualizing a flow model. The most appropriate reach is where there are as less as possible anabranches and tributaries or groundwater recharge or discharge. When the reach is as less complex as possible, the focus lies on the simple river processes which gives a better insight in which processes are modelled and which are not. In the graphs this is shown when the difference between the flow between two sites is as small as possible. Also the difference in salinity needs to be small.

3.3.2 Q5: Top-down conceptual flow model

To find an answer on this sub-question, an iterative approach was used. The top-down approach means the process starts with analyzing the available data. Depending on the conclusion drawn on base of the data analysis, the model structure is decided (Step 1 'Model Structure'). After that the parameters in the conceptual model are calibrated (Step 2 'Calibration'). The third step contains model performance analysis using objective functions which might result in the fact that the Model Structure needs to be modified (Step 3 'Model Performance'). The fourth step is determining the parameter uncertainties on the model output by using Monte Carlo (Step 4 'Parameter Uncertainties'). The last step is testing the predictive ability of the model by testing it against an independent data set (Step 4 'Validating'). Before starting the iterative steps, the calibration and validation method is chosen.

Calibration and Validation
For the calibration and validation several techniques which will be listed in this section, could be used. First, the Split-sample test means that the available record should be split into two segments one of which should be used for calibration and the other for validation. Another test is the Proxy-basin test

which can be used when the flow from an ungauged basin C needs to be simulated using two gauged basins A and B which are available in the same region. Another test is the differential split-sample test which is required whenever a model from for example gauge X, needs to be used to simulate flows in gauge basin Y under conditions different from the conditions corresponding to the available data from gauge X (Klemes, 1986). Another test which looks like the split-sample test is K-fold partitioning. The data is split into K sets, one set is used for calibrating and the remaining K-1 sets are used for validating. The hold out method can then be repeated K times allowing all results to be averaged (Bennett et al., 2013).

The proxy-basin test and the differential split-sample test are not usable in this research since there is no question of an ungauged basin and there is the sample of the data is big enough to split the sample. In this research the split-sample test in combination with the idea of the K-fold partitioning test are used.

Step 1: Model Structure

The first step is deciding which model structure is used. This step 1 contains of different sub-steps. First the autocorrelation and cross-correlation of the input and output data from the chosen reach are obtained. Correlation functions are useful time series analysis tools and yield physical information such as the time delay between two related processes. These results are combined to obtain the Unit Hydrograph of the reach. The shape of the Unit Hydrograph gives a lot information about the shape of the model structure.

Autocorrelation of the input data

The autocorrelation is a measure of how closely a quantity observed at a given time is related to the same quantity at another time. It measures the degree of resemblance ρ of the signal with itself as time passes. When the time lag is zero, ρ is by definition one since this means the flow series of the input is compared with itself at the same time. Equation (1) is the autocorrelation equation where $\bar{X}$ is the mean of the values X in the series (Scargle, 1989). For hydrology the autocorrelation is useful for exploring the seasonality of the input flow, as well as the persistence of the flow between time steps.

$$\rho_x(k) = (\tfrac{1}{N}) \sum_{n=1}^{N-k} [X(t_n) - \bar{X}][X(t_{n+k}) - \bar{X}] \qquad (1)$$

Cross-correlation of the input and output data

The cross-correlation function measures how closely two different observables are related to each other at the same or different times (Scargle, 1989). Equation (2) is the cross-correlation function obtained where the two series are not symmetrical; that is: $r_{+k} \neq r_{-k}$ (Padilla & Pulido-Bosch, 1994). The coefficient r is a measurement of the size and direction of the relationship between x and y. The sample non-normalized cross-correlation of two inputs signals requires that r be computed by a sample shift (time-shifting) along one of the input signals (Lyon, 2010). The cross-correlation graph shows the hydrograph (discharge of flow) of the actual flow data in a specific period (depending on the period of the data). It is useful for exploring the average response of the catchment across the data period (Croke & Shin, 2015). The peak of the correlation coefficient, shows the degree to which the input flow represents the output flow (Croke & Littlewood, 2005). A negative correlation coefficient means there is an anti-correlation between the two shapes. Further, the cross-correlation graph also shows the seasonality of the relationship variations.

$$r_{+k} = r_{xy}(k) = \frac{C_{xy}(k)}{\sqrt{C_x^2(0) C_y^2(0)}} \qquad (2)$$

$$r_{-k} = r_{yx}(k) = \frac{C_{yx}(k)}{\sqrt{C_x^2(0)C_y^2(0)}} \quad (2)$$

Where

$$C_{xy}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(x_t - \bar{x})(y_{t+k} - \bar{y}) \quad (3)$$

$$C_{yx}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \bar{y})(x_{t+k} - \bar{x}) \quad (3)$$

$$C_x(0) = \frac{1}{n}\sum_{t=1}^{n}(x_t - \bar{x})^2 \quad (4)$$

$$C_y(0) = \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})^2 \quad (4)$$

Shape Unit Hydrograph obtained from deconvolution

In this research the cross-correlation and the auto-correlation are obtained. The cross-correlation shows the hydrograph (discharge of flow) of the actual flow data over the available time period. Using deconvolution of the autocorrelation in combination with the cross-correlation, gives the shape of the Unit Hydrograph for the input and output data (Croke B. , 2005). More information about the definition of the Unit Hydrograph can be found in Appendix I. The shape of this Unit Hydrograph is a non-parametric empirical estimate.

Formulation of equation Unit Hydrograph

After the shape of the Unit Hydrograph is obtained, a possible formulation of the equation of the Unit Hydrograph that has the potential to reproduce the shape of the UH of this research is formulated. Jakeman et al., (1990) have written the general discrete convolution equation of the Unit Hydrograph shown in equation (5) into an autoregressive formulation shown in equation (6) which is the formulation used to reproduce the shape of the Unit Hydrograph obtained from the research data. Since in this research the effective rainfall will not be used as the input, the u_k is replaced for the flow input $I(\zeta)$. The application of autoregressive models has been attractive mainly because the autoregressive form has an intuitive type of time dependence (the value of a variable at the present time depends on the values at previous time) and they are the most simple models to use (Salas et al., 1980). It is an efficient formulation in terms of writing it down and in terms of decreasing the calculation time of the flow. This formulation also offers a powerfull tool to estimate the parameters (Jakeman et al., 1990). More information about the Unit Hydrograph and this translation is defined in Appendix II (Jakeman et al., 1990).

$$Q_k = h_0u_k + h_1u_{k-1} + h_2u_{k-2} + \cdots + h_{k-1}u_1 + \zeta_k \quad (5)$$

$$Q(\zeta) = \left(\frac{\beta}{1+\alpha\zeta^{-1}}\right)^n * I(\zeta) \quad (6)$$

Parameter α

This equation represents a model of a linear reservoir with n storages all connected in series. The parameter α is related to the time constant τ for a linear reservoir (Jakeman et al., 1990).

$$-\alpha = e^{-\frac{\Delta t}{\tau}} = e^{-\frac{1}{\tau}} \quad (7)$$

Parameter β

The parameter β is related to the fractional throughput Steady State Gain (Jakeman et al., 1990). The Steady State Gain (SSG) is approximately the ratio of the temporal sum of the output of a system (streamflow output) to the temporal sum of the input (streamflow input). When the Steady State Gain is one, there is no water mass which is conserved between the input and output of the reach, for example, there are no losses (through anabranches, evaporation or infiltration), or gains (through tributaries, exfiltration from aquifer to the river). The parameter β govern the height of the unit hydrograph peaks (Ivkovic et al., 2014) . This can be written down in the following equation:

$$\frac{Q(\zeta)}{I(\zeta)} = \frac{\beta}{1+\alpha} = G_{ss} \quad (8)$$

Parameter n

The parameter n is the number of stores (Nash cascade) which need to be used in the model structure. The Nash cascade connects identical linear reservoirs in series. The output from the first reservoir is the input for the second reservoir etc. (Nash J. , 1958).

Step 2: Calibration

The least square method will be used for optimizing the parameters (Albritton et al., 1976). The parameter values are changed till the minimum sum of squared residuals (observed flow minus modelled flow) are obtained. The sum of squared residuals is the numerator in the Nash-Sutcliffe objective function, equation (9) (Nash & Sutcliffe, 1970). In other words, when minimizing in the least square, the NSE (Nash-Sutcliffe efficiency) will be maximised. In the NSE equation, the *n* is the number of time steps, o_i is observed flow at time step i (daily here), $\bar{o}_i$ is the mean of the observed flow and m_i is the modelled flow. NSE exists in the interval (- ∞ to 1.0]. The closer the value of NSE is to 1, the more accurate the model performs. It assesses the quality of the shape of the hydrograph (Tillaart, 2010). This means the parameter values which are giving the NSE the closest to 1, are the best parameters. When the NSE ≤ 0, the model is not better than using the observed mean as a predictor.

$$NSE = 1 - \frac{\sum_i^n (o_i - m_i)^2}{\sum_i^n (o_i - \bar{o}_i)^2} \quad (9)$$

Step 3: Model performance using observations and objective functions

There are different kinds of objective functions which can determine the model performance given a certain parameter set (see review by Bennett et al., 2013 for an extensive discussion regarding the approaches for model performance). These objective functions are given an indication of the fit between a model and the system being modelled.

First the NSE, used to optimize the parameters, gives an indication of the model performance. A problem with using NSE is its oversensitivity for higher flows. The second objective function is a proposal to ease this. It is the logarithmic from the NSE, thus the log from the observed and the log from the modelled flow (Muleta). The third objective function is the RVE (Relative Volume Error) which is shown in eq (10). RVE is aimed at the relative volume difference between the observed and modelled flow output and has an optimum value at zero (Tillaart, 2010).

$$RVE = \frac{\sum_i^n (o_i - m_i)}{\sum_i^n (o_i)} \quad (10)$$

The combined objective function of NSE and RVE used in this research is called y (Tillaart, 2010) and is defined as follows:

$$Y = \frac{NSE}{1+|RVE|} \quad (11)$$

Before using the objective functions to give information about the model performance, observations are used. It is not possible to get an idea about which river processes are taken into account in the model and which are not by only looking at the objective functions. By analysing the plots from the modelled output and observed output against the time and the plots from the residuals (modelled output minus observed output), a general idea about which processes are captured by the model structure is given.

Step 4: Parameter uncertainties using Monte Carlo
During step 4 the uncertainty in the model structure due to the parameter values are obtained by using a Monte Carlo simulation (Croke, 2009). The Monte Carlo simulation is commonly adopted for uncertainty analysis of deterministic models. It requires the random generation of many realisations of the inputs that are run through the model in order to derive confidence limits for a given flow output (Loveridge et al., 2013). As such, a Monte Carlo simulation can give the impact of uncertainty in the model parameter or the impact of uncertainty in the model input on the model output. In this research the impact of the uncertainty in the model parameters is obtained. To determine the uncertainties of the model input, information about these uncertainties is needed which is not currently available.

In this research the Monte Carlo simulation is done for the different parameters by repeatedly adding random noise to one parameter using the mean and the variance of the parameter, while the other parameters are keeping their constant value (J.C.Clarke). Another approach can be that the random noise is added to all the parameters at the same time. Information about this approach is given in Appendix III. Since the optimization process gives one optimized value for the parameter, this value is the mean μ of the parameter values used in the Monte Carlo simulation. Further, the variance of the parameters is needed which can be determined by obtaining the variance-covariance matrix.

Variance-Covariance Matrix
Equation (12) shows the equation to obtain the variance-covariance matrix, where σ_y^2 is the variance of the residuals (modelled output minus observed output) and J^TJ is the Jacobian matrix of the parameters (tue). The Jacobi matrix gives the partial derivatives of the least square function F with respect to the parameters, $Jij = \delta Fi/\ \delta\beta j$ (Tellinghuisen, 2001). The Jacobian matrix represents the local sensitivity of the objective function F to variation in the parameters p (Gavin, 2013).

$$V = \sigma_y^2 (J^TJ)^{-1} \qquad (12)$$

The variance-covariance matrix V is a symmetric matrix, shown in matrix (13) where the diagonal represents the variance of the parameters $(p_1,\ p_2 \dots p_n)$ (Albritton et al., 1976). The non-diagonal values are representing the covariance between two parameters where c is the correlation coefficient. In this situation the variance σ^2 (the diagonal of matrix V) of the specific parameter is used to obtain the standard deviation σ.

$$V = \begin{matrix} \sigma_{p1}^2 & \sigma_{p1}\sigma_{p2}c_{p1,p2} & \sigma_{p1}\sigma_{p3}c_{p1,p3} \\ \sigma_{p2}\sigma_{p1}c_{p2,p1} & \sigma_{p2}^2 & \sigma_{p2}\sigma_{p3}c_{p2,p3} \\ \sigma_{p3}\sigma_{p1}c_{p3,p1} & \sigma_{p3}\sigma_{p2}c_{p3,p2} & \sigma_{p3}^2 \end{matrix} \qquad (13)$$

When both the σ and the μ are obtained, equation (14) shows the implementation in Monte Carlo to repeatedly add random noise to one parameter. $N(0,1)$ means the 'random noise' which is added to the mean μ of the parameter and which is normal distributed with mean 0 and standard deviation σ obtained from the variance-covariance matrix.

$$Parameter = \mu + N(0, \sigma) \qquad (14)$$

Amount of runs and confidence interval

For the amount of runs a balance was found between the running time in MatLab in combination with the available time for this research project and the accuracy of the uncertainty (Owen, 2009-2013). For the output a 95% confidence interval is used. This means there is 95% confidence that a random sample of the flow output lies within plus or minus 1.96 standard deviation of the mean. The amount of runs in this research is 1000 which should give 25 samples below and 25 samples above the 95% confidence interval.

Step 5: Validation

For the validation the other dataset (according to split-sample test) is used as the input for the flow model. Objective functions are used to determine the model performance given a certain data set. The objective functions are NSE, log NSE, RVE and Y as described in step 1: 'Calibration'.

Managing uncertainties

At the of Chapter 6 Modelled Structure, information is given about the uncertainties in using the NSE as objective function and the least square as optimization tool.

4 Results: Data analysis

In this this chapter the first sub-question of the connection between flow and EC data is answered. First the available data is determined (4.1) and second the relationship between flow, salinity (EC) and salt load is obtained (4.2).

4.1 Available data and location of sites

For this project EC data at 54 sites is available and flow data at 46 sites from 01-07-1990 to 30-04-2012. There are 25 sites measuring flow data and 33 sites measuring EC whose locations are unknown which makes these sites unusable. Further, the flow model is developed for a reach in the Murray River. Therefore, the sites which are located in anabranches and tributaries are not important for the data analysis. At last, for the data analysis only the flow and EC data from the same site can be compared. At the end there are still 9 sites available which have a known location, are located in the Murray River and have both EC and flow data available. The 9 sites are shown in Figure 5.

Figure 5 The nine sites available for data analysis

4.2 Relationship flow, salinity and salt load

In section 4.2 the important results are shown according to the relationship between the flow, salinity and salt load, but also according to the difference between several sites.

Flow data

Figure 6, Figure 7, Figure 8 are showing the flows for the total available data period of forty years.

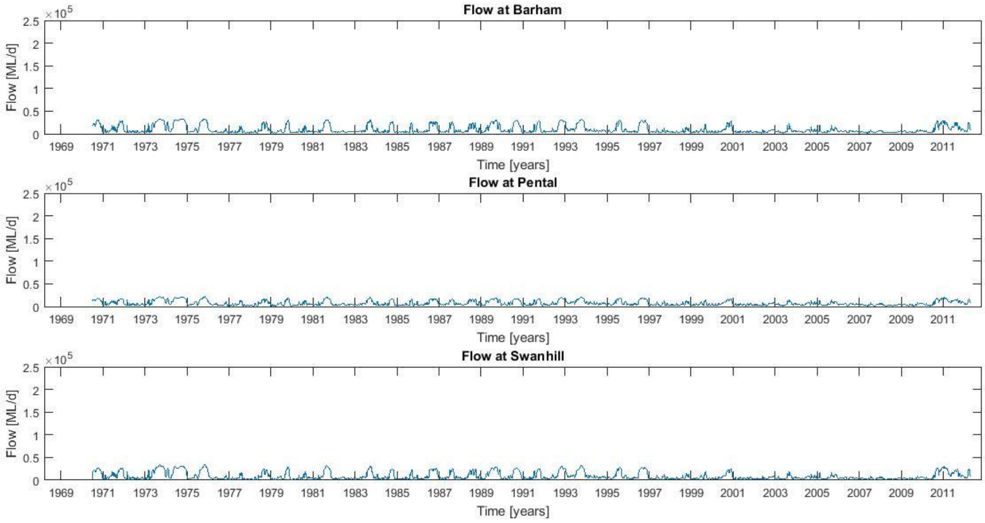

Figure 6 Flow data Barham, Pental and Swan Hill

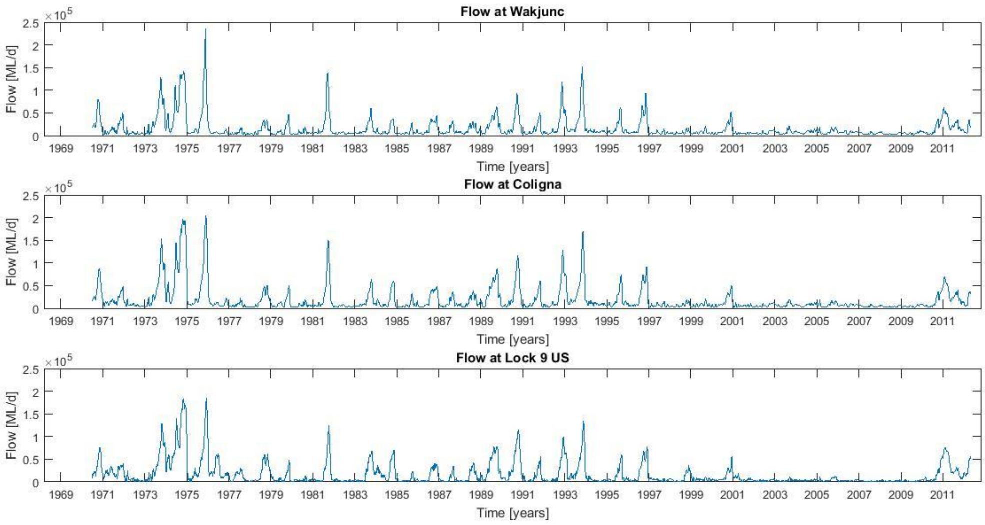

Figure 7 Flow data at Wakool junction, Coligna and Lock 9

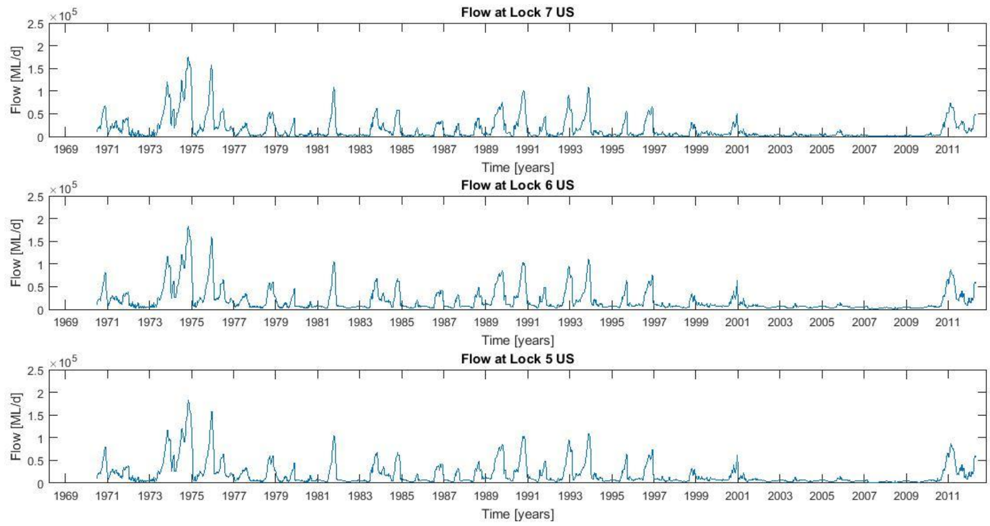

Figure 8 Flow data at Lock 7, Lock 6 and Lock 5

The period 1997 un till 2009 was during the Millennium Drought in Australia which causes the lower flow. The nine different sites are divided in three reaches:

Reach 1 - Upstream

The first reach is Barham, Pental and Swan Hill. The flow at these sites is lower than at the sites downstream, for example the difference in average percentage between Barham and Wakool Junction is 64%.

Reach 2 - Middle

The second reach is Wakool junction and Coligna. Due to the Edward river downstream from Swan Hill, the flow rises with an average percentage of 19%.

Reach 3 - Downstream

The third reach exist of Lock 9 to Lock 5. The locks in general are used to regulate the amount of water (Telfer et al., 2012). The lock closest to Lock 9 upstream is Lock 11 and then Lock 15 (Coligna is in between both locks). The distance between Lock 15 and Lock 11 is much bigger than the distances between Lock 9, 7, 6 and 5 (Telfer et al., 2012). Since the distance between Lock 9, 7, 6 and 5 is smaller, the peak flows are very similar. The average difference in percentage between Lock 9 and Lock 5 is 14% where the difference between Barham and Wakool Junction is 64%.

Difference between Reach 1 (Pental, Barham, Swan Hill) and Reach 2 (Wakool Junction, Coligna)

Figure 9 shows the Murray River between Barham, Pental, Swan Hill (reach 1) and Wakool junction (reach 2). Between Barham and Pental, Reedy Creek and the anabranch the Little Murray River flows into the Murray River. More information about the Little Murray can be found in "**Barham, Pental and Swan Hill (reach 1)**". Between Swan Hill and Wakool Junction, Edward River (the biggest tributary), Speewa Creek and Bingera Creek flow into the Murray. There are also two small, unknown streams.

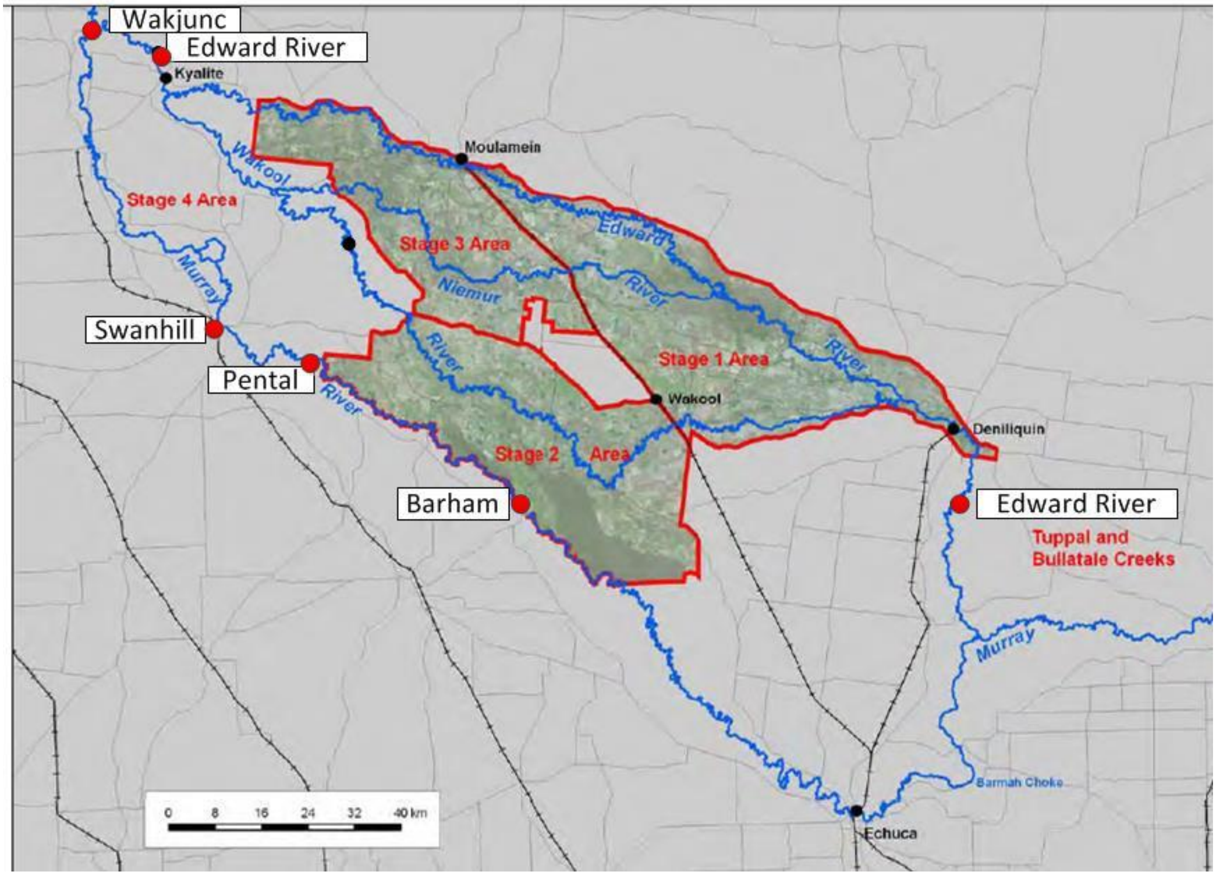

Figure 9 Area Reach 1-Reach 2

Figure 10 shows the differences in the peak flow at Swan Hill and Wakool junction. Between the flow at Swan Hill and Wakool Junction, the average, percentage change is 69%. This percentage change fluctuates during the years which shows that the input from groundwater and anabranches fluctuates as well. The difference between both locations is caused by the anabranch the Edward River. The Edward River is part of the Edward and Wakool Rivers floodplain which is shown in Figure 9. Flooding in this floodplain is independent of local rainfall. The average annual rainfall is less than 400 mm. Figure 9 shows the complexity of the floodplain. More information about the flooding's in the Edward River and in this floodplain is necessary to predict the effect of the flooding's in the Edward River on the Murray River at Wakool junction (Environment-Climate-Change-&-Water-NSW, 2011).

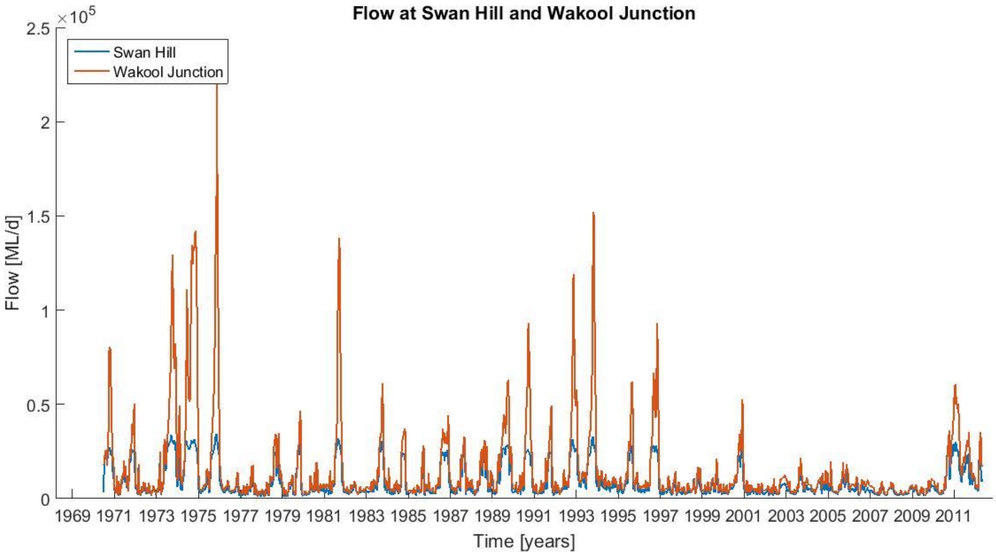

Figure 10 Flow Swan Hill and Wakool junction

Difference sites Wakool junction (reach 2) – Coligna (reach 2)

Figure 11 shows there is no big difference between the flow at Wakool junction and Coligna. The average percentage change is 19% which is much lower in comparison with the difference between Swan Hill and Wakool Junction (69%). However, Figure 12 shows there are many tributaries (Murrumbidgee is the biggest), flowing into the Murray River which might clarify the 19% difference. There is relatively little information available regarding water use and management downstream of Balranald. The channel capacity downstream of Balranald is estimated to be 11000 to 13000 ML/d. The volume of flow decreases from Balranald, with only a small volume of water making it to the junction with the Murray River (Murray-Darling-Basin-Authority, Environmental Water Delivery - Murrumbidgee Valley, 2012). According to the MDBA (Murray-Darling Basin Authority), the contribution of the Murrumbidgee in amounts of flow to the Murray River is small which is in agreement with Figure 11. Further, Wee Wee Creek, Bridge Creek and Tualka Creek are tributaries before Murrumbidgee Valley and Taila Creek and Chalka Creek are tributaries between Wakool junction and Coligna behind the Murrumbidgee Valley. There is little information available about these tributaries, but the contribution of these tributaries including the contribution from for example groundwater, is around 19% which makes the contribution much lower than the contribution from for example the Edward River.

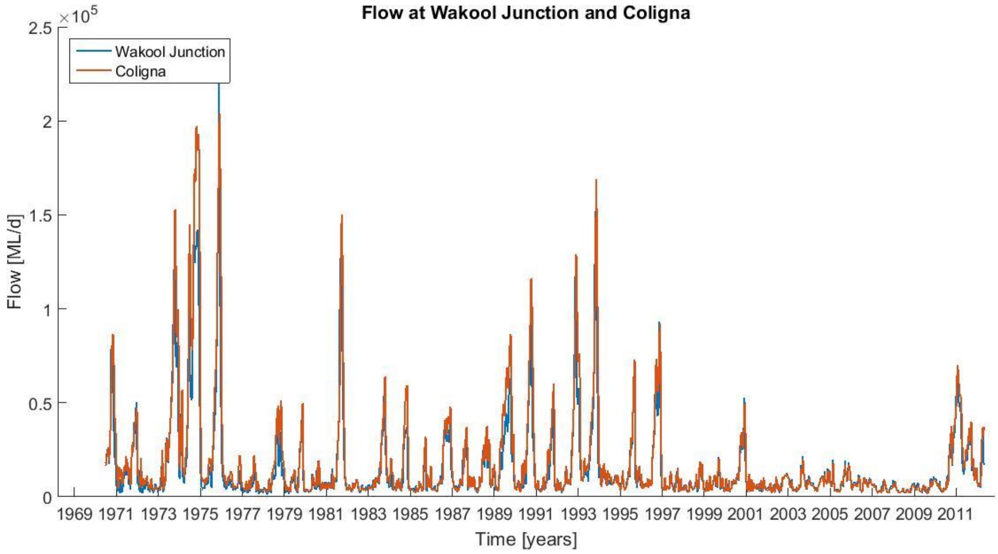

Figure 11 Flow Wakool junction-Coligna

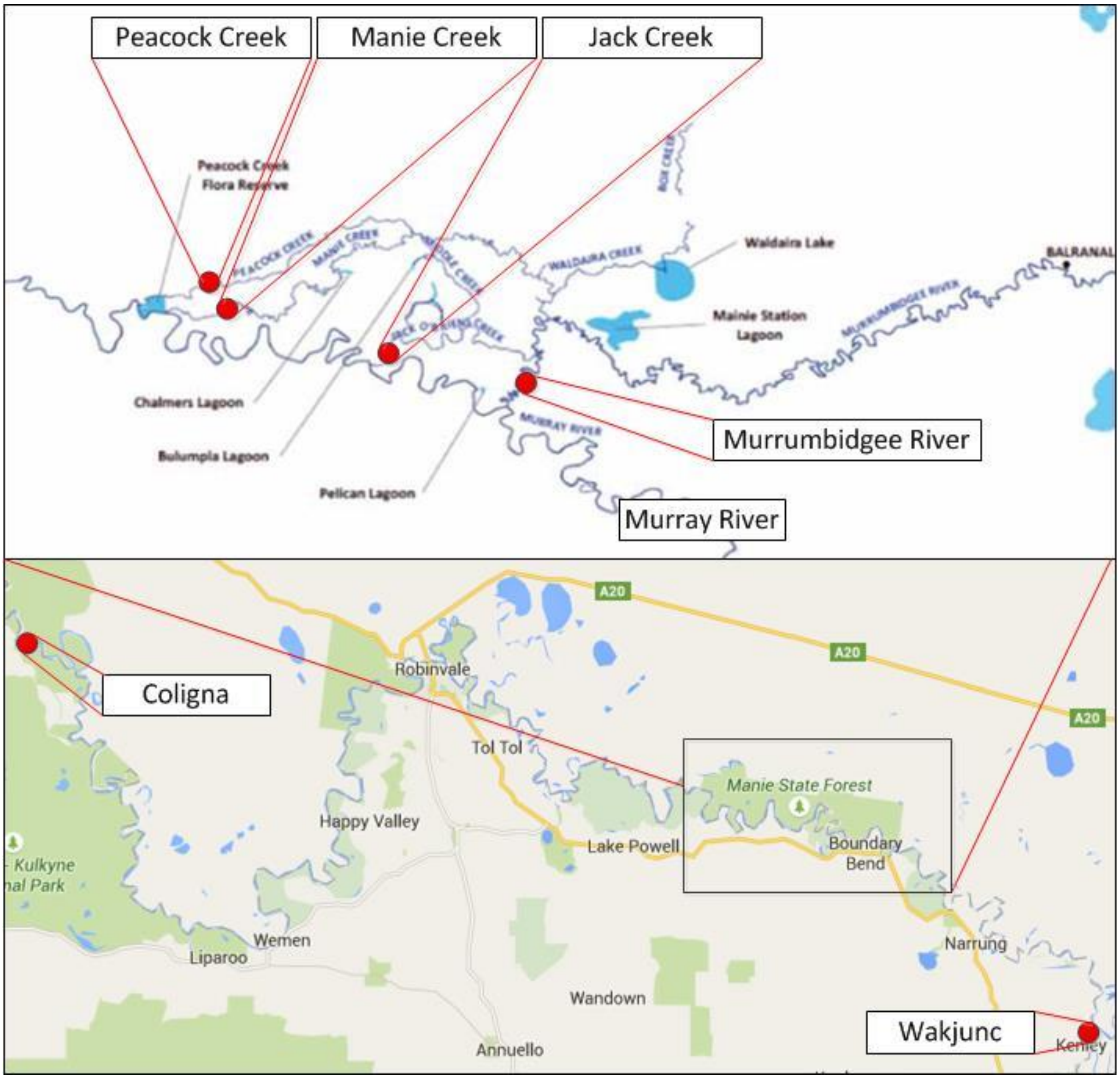

Figure 12 Murrumbidgee Valley to Murray River (Murray-Darling-Basin-Authority, Environmental Water Delivery - Murrumbidgee Valley, 2012)

Differences between Coligna (reach 2) – Lock 9 (reach 3)

Figure 13 shows the difference between the flow at Coligna and Lock 9. The mean of the flow at Coligna (upstreams) is higher than at Lock 9 (18500 ML/d respectively 16945 ML/d). Around Wentworth there is minor irrigation development (Green et al., 2012). Further the floodplain between the Darling anabranch and Lock 9 is a losing floodplain (Telfer et al., 2013). Both facts might cause the difference in mean flow. The floodplains between Coligna and Mildura are gaining (Telfer et al., 2013). Between Great Darling Anabranch and Darling River, shown in Figure 14, there is also a tributary called Walpolla Creek.

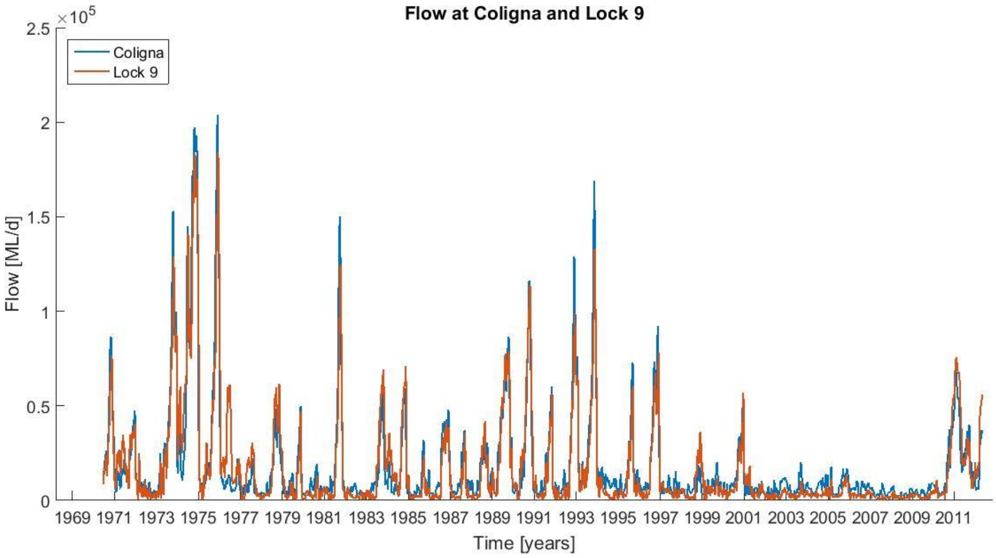

Figure 13 Flow Coligna - Lock 9

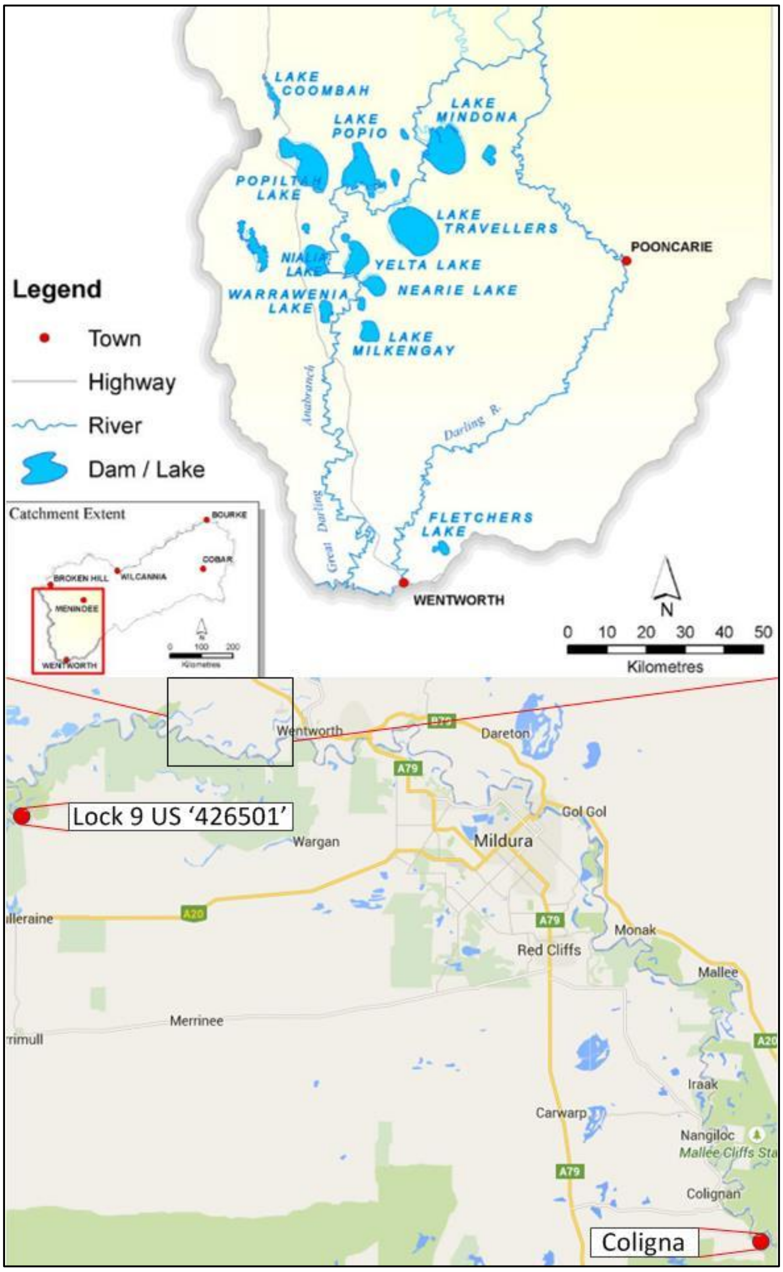

Figure 14 Lower Darling Catchment (Green et al., 2012)

Differences sites in reach 3 (Lock 9 – Lock 5)

Figure 15 shows the flow at Lock 9 to Lock 5. The average percentage change between Lock 9 and Lock 7 is 4%, between Lock 7 and Lock 6 21% and between Lock 6 and Lock 5 1%. The locks in general are used to regulate the amount of water (Telfer et al., 2012). Since the distance between Lock 9 to Lock 5 is small, the peak flow is conserved. Table 1 shows the distance between the sites and the distance between Lock 9 and Lock 5 is 203 kilometers, which is a smaller distance than the distance from Coligna to Lock 9. However, the average percentage change is still 21% between Lock 7 and Lock 6 because there are two creeks between them (Punkah Creek and Salt Creek). There are also two creeks between Lock 6 and Lock 5 (Monoman Creek and Hundee Creek). The Chowilla creek connects the Punkah Creek and the Monoman Creek.

Table 1 Distance between the sites

Reach	Distance kilometres	lock9-lock5 = 203.0
Coligna – Lock 9	228.5	(Telfer et al., 2012)
Lock 9 – Lock 7	68.5	(Telfer et al., 2012)
Lock 7 – Lock 6	76.8	(Telfer et al., 2012)
Lock 6 – Lock 5	57.7	(Telfer et al., 2012)

The floodplains between Lock 9 and Lock 6 are through-flow floodplains. Through-flow floodplains are found in reaches where the regional groundwater flow lines show that groundwater flow beneath or through the floodplain. In through-flow reaches, the floodplain alluvium is potentially gaining water from the up-gradient side, but is losing water to the regional groundwater system on the down-gradient side.

Between Lock 9 and Lock 7, Victoria Lake (Rufus River) is an important storage place that assists in regulating the flow and controlling the salinity in the Murray River just before it flows into South Australia (Murray-Darling-Basin-Authority, Lake Victoria, sd). This might explain the small differences in the amount of flow between Lock 9 and Lock 7.

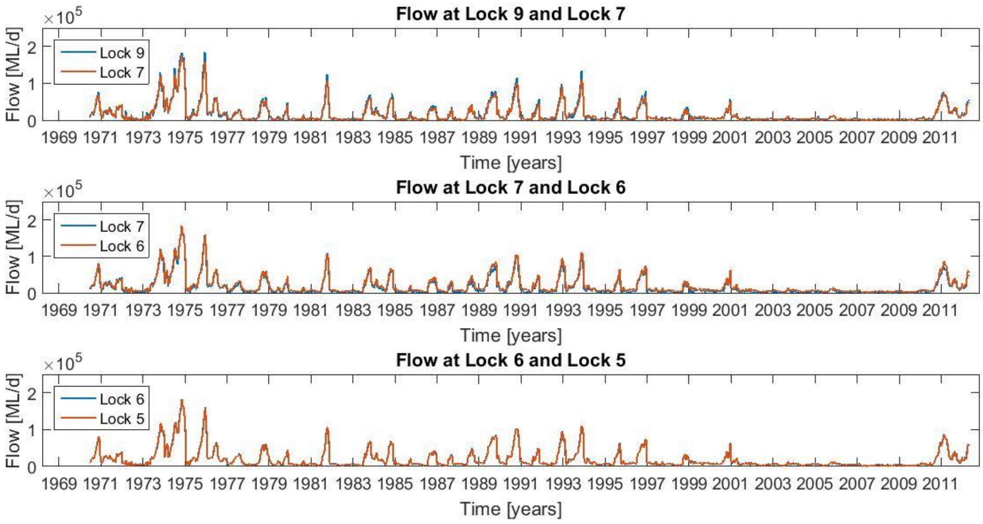

Figure 15 FLow at Lock 9, Lock 7, Lock 6 and Lock 5

Flow, Salinity and Salt load

In this section the first step will be looking at the salinity graphs to determine interesting differences in salinity. This information will then be combined with the flow and salt load data since the salinity is dependent on the flow. The salt load shows the combination of flow and salinity at a specific moment.

Figure 16, Figure 17 and Figure 18 show the amount of salinity at the nine different sites. These graphs are plotted from 1975 since the data before 1975 is incomplete. The figures show the decrease of salinity from 2000 which can be explained by several aspects. First is the drought-induced salt storage in the floodplains, second the implementation of Salt Interception Schemes, third the improved irrigation efficiencies and last low salinity surface water inputs from the Hume Dam (Telfer et al., 2012).

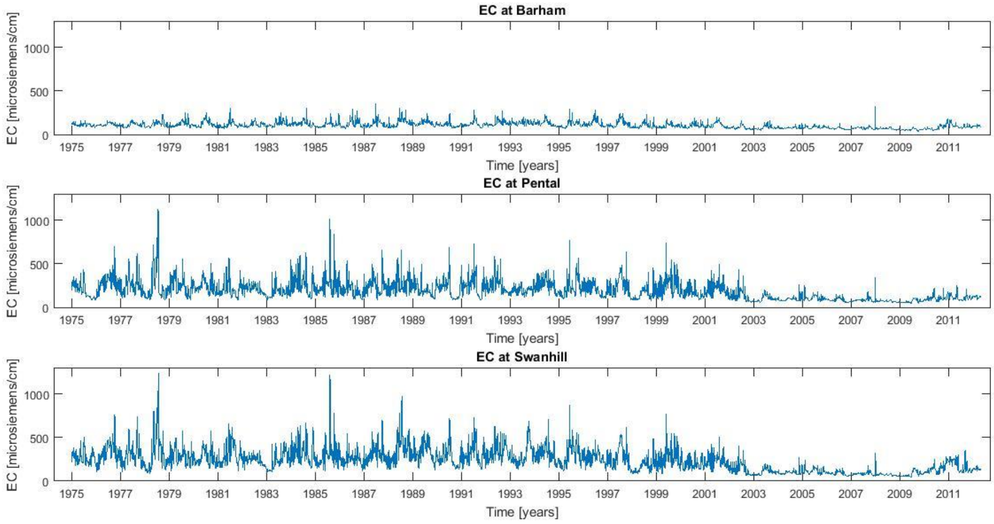

Figure 16 EC data at Barham, Pental and Swan Hill

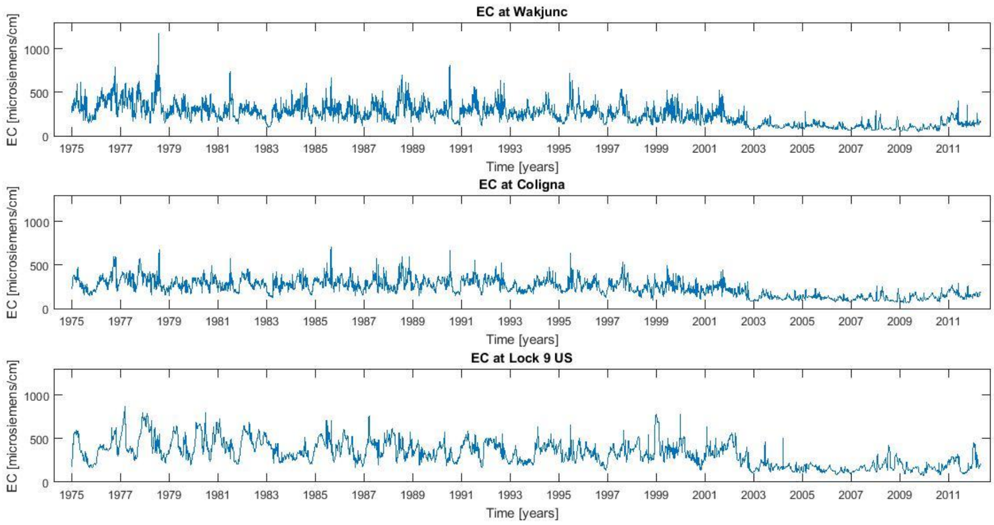

Figure 17 EC data at Wakool junction, Coligna and Lock 9

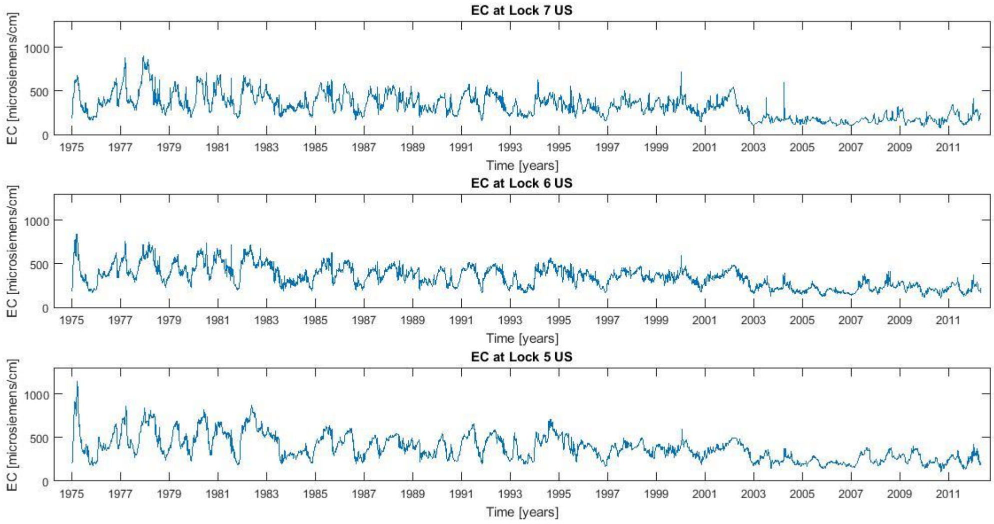

Figure 18 EC data at Lock 7, Lock 6, Lock 5

Barham, Pental and Swan Hill (reach 1)

Figure 16 shows there is a significant difference between the salinity at Barham and Pental. Figure 19 shows that the Little Murray River is an anabranch which flows in the Murray River between Barham and Pental during normal flows and flows out during high flows. The Little Murray flows into the Murray River between Pental and Swan Hill (Gippel, 2013). The regulation of the Little Murray River has resulted in significant siltation from the maintained constant water level (NCCMA, 2010).

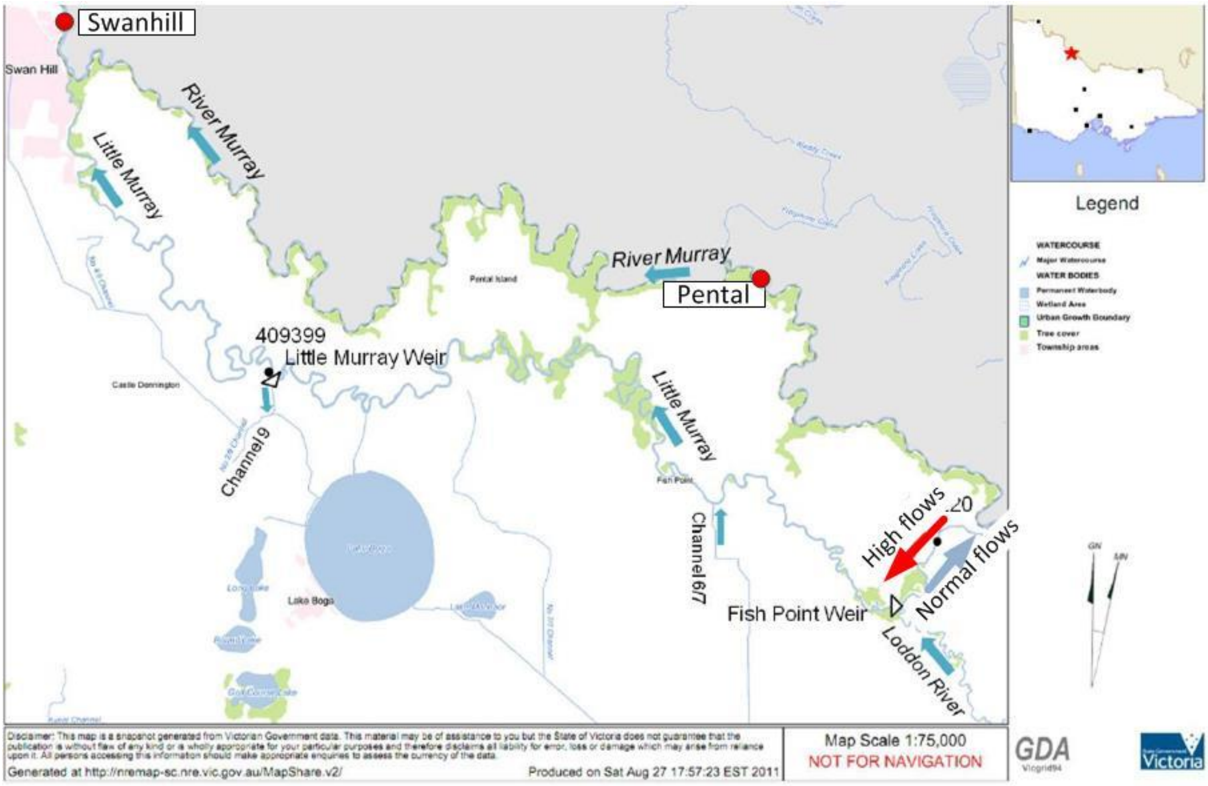

Figure 19 The Little Murray (Gippel, 2013)

When the Little Murray and the Loddon River together exceed a flow of 12200 ML/d, the Fish Point Weir (FPW) opens. During that moment the Little Murray transforms into an anabranch since the water from the Murray River flows into the Little Murray via the FDW downstream from Barham and Upstream from Pental. The Little Murray flows back into the Murray River downstream from Pental and upstream from Swan Hill. This explains the lower peak flow at Pental in comparison with Barham and Swan Hill shown in Figure 20, Figure 21 and Figure 22.

The regulation of the FDW also causes the differences in the salinity between Barham, Pental and Swan Hill. The average percentage change between Barham and Pental in salinity is 79% and between Pental and Swan Hill 24%. During lower flows, the very saline Loddon River flows into the Murray River. This gives the higher salinity peaks at Pental. Further the salt load graph at Pental, Figure 21, shows peaks which are broader than at Barham, Figure 20, but not higher. This is because the salt load is obtained by the product of flow and salinity. At Barham the flow is high and the salinity is low, at Pental both flow and salinity are mediocre. Further the salt load at Swan Hill has much higher peaks than at Barham and Pental which is caused by the fact that the peaks of the flow are much higher at Swan Hill than at Pental.

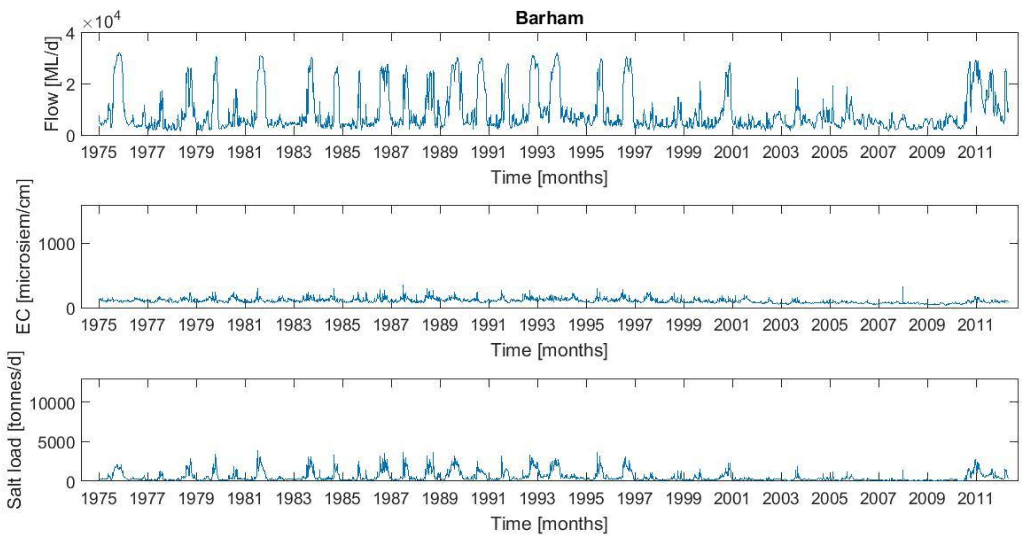

Figure 20 Barham

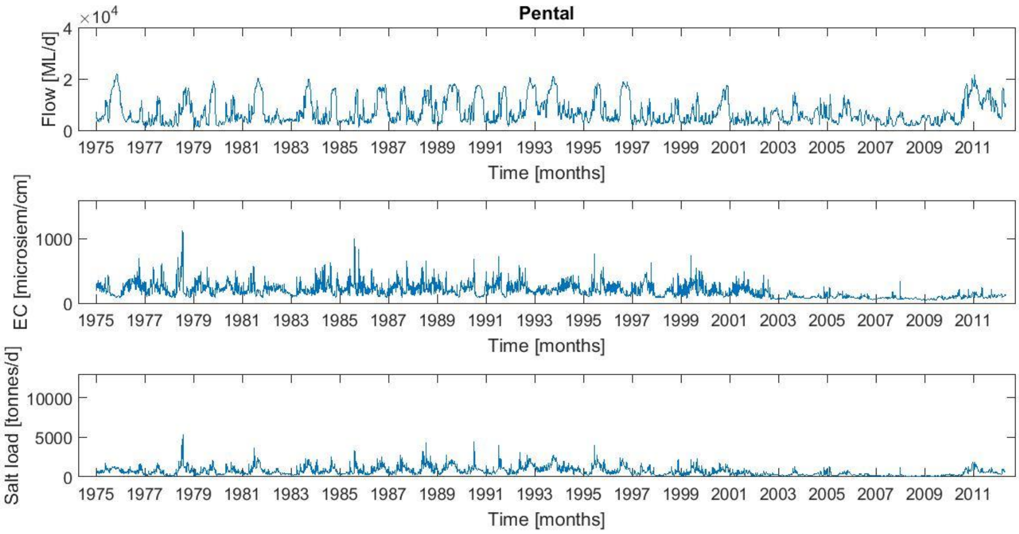

Figure 21 Pental

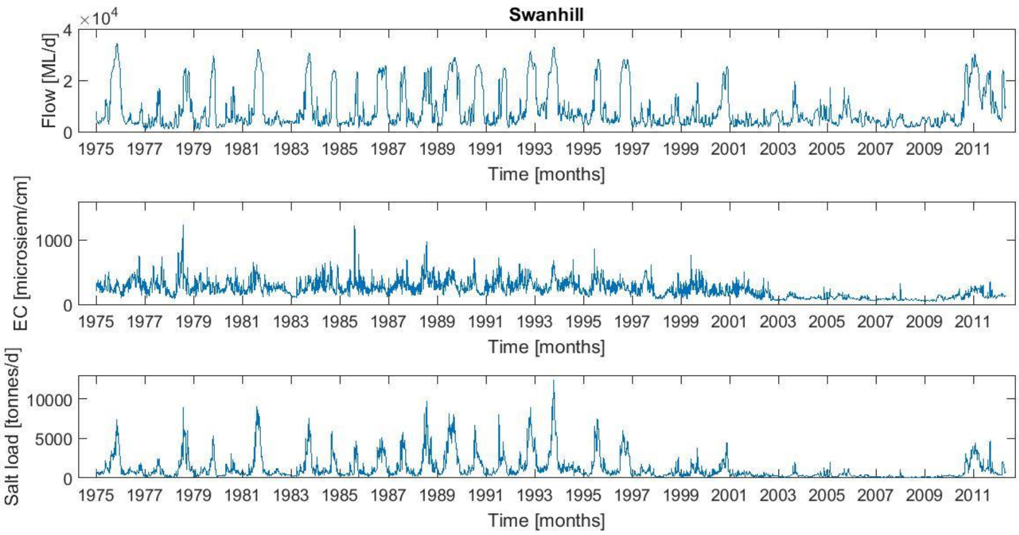

Figure 22 Swan Hill

Wakool junction (reach 2) until Lock 5 (reach 3)

Figure 23to Figure 28 are showing the flow, salinity and salt load at the different sites. The differences in flow are discussed in section 4.2, part '*flow data*'. The salinity shows that the mean of the salinity increases from Wakool junction until Lock 5 and that the peaks become broader.

Figure 24, Wakool junction, shows an average percentage change of 18% in comparison with Figure 23, Coligna when looking at the salt load. The difference in salinity is only 1% and less variable at Coligna in comparison with Wakool junction. This small difference is explained due to a higher flow at Coligna. Since the salt load is more influenced by the flow, a high salinity does not mean the salt load is high as well (nswgovernment).

Figure 25, Figure 26, Figure 27 and Figure 28 show a higher salinity at Lock 5 than upstream from Lock 5, while the flow graph and the salt load graph look the same. This situation needs a closer look.

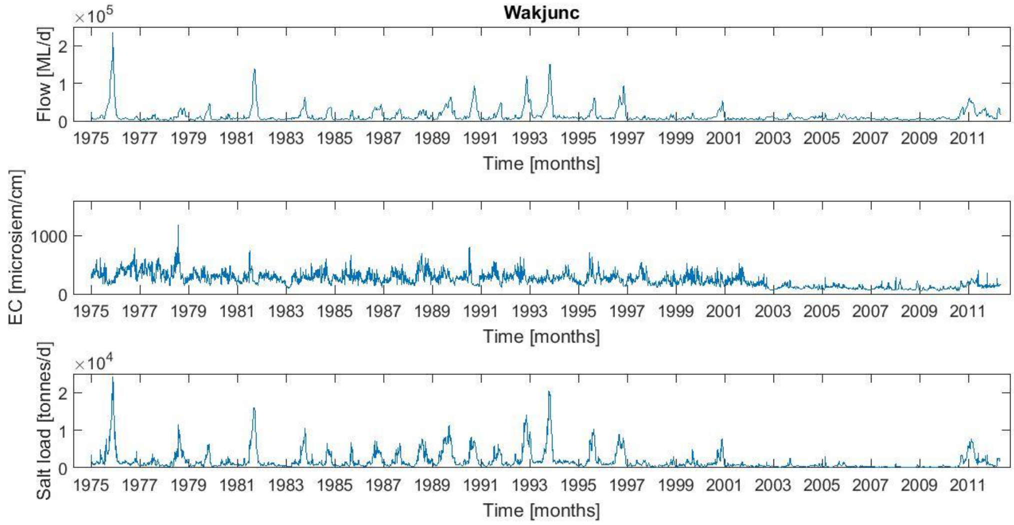

Figure 23 Wakool junction

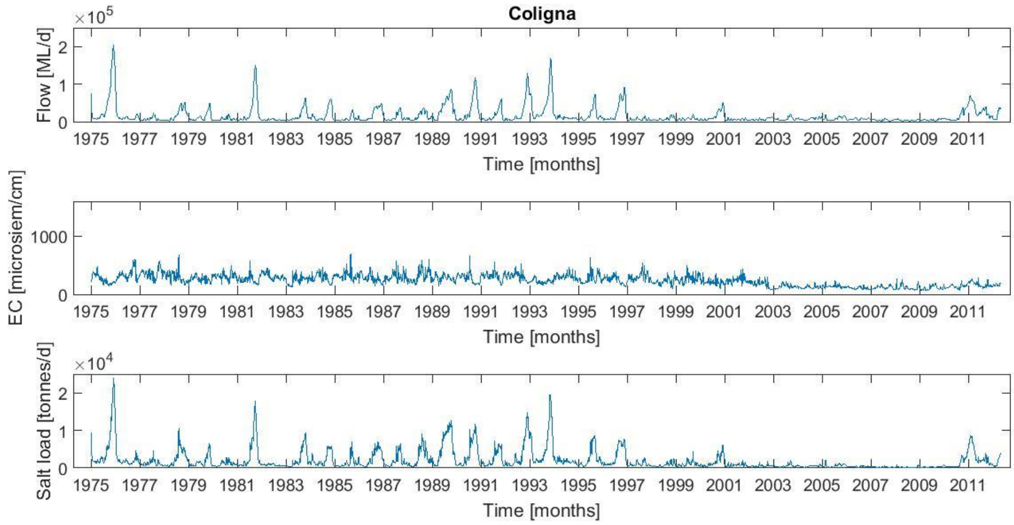

Figure 24 Coligna

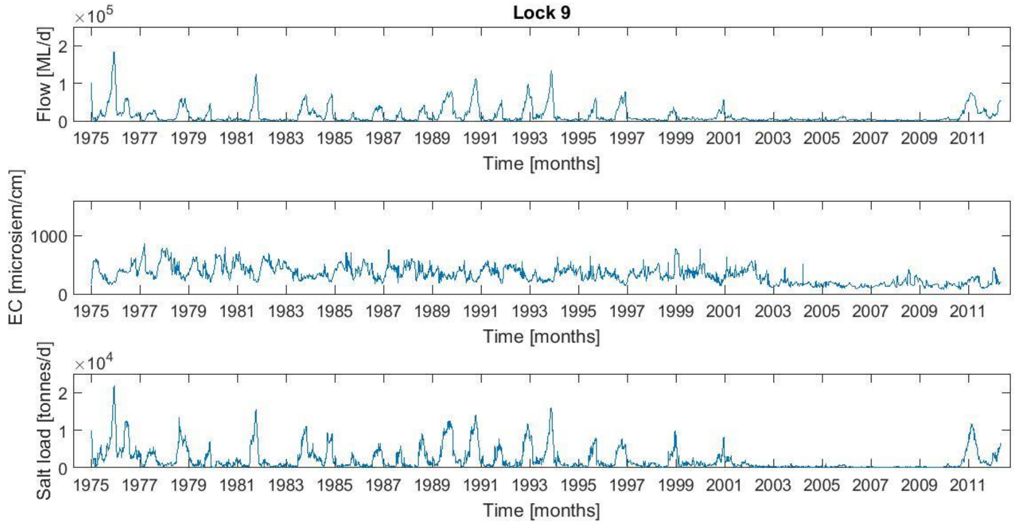

Figure 25 Lock 9

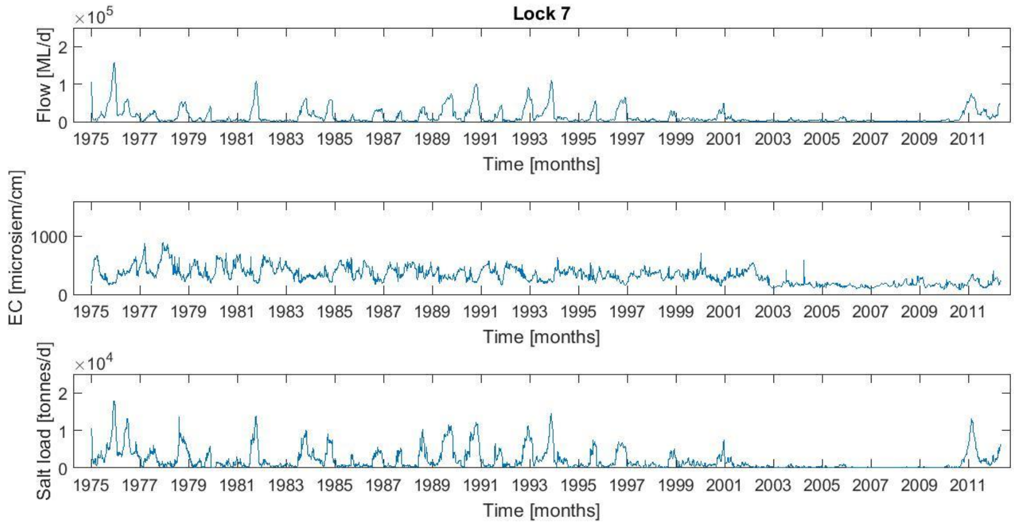

Figure 26 Lock 7

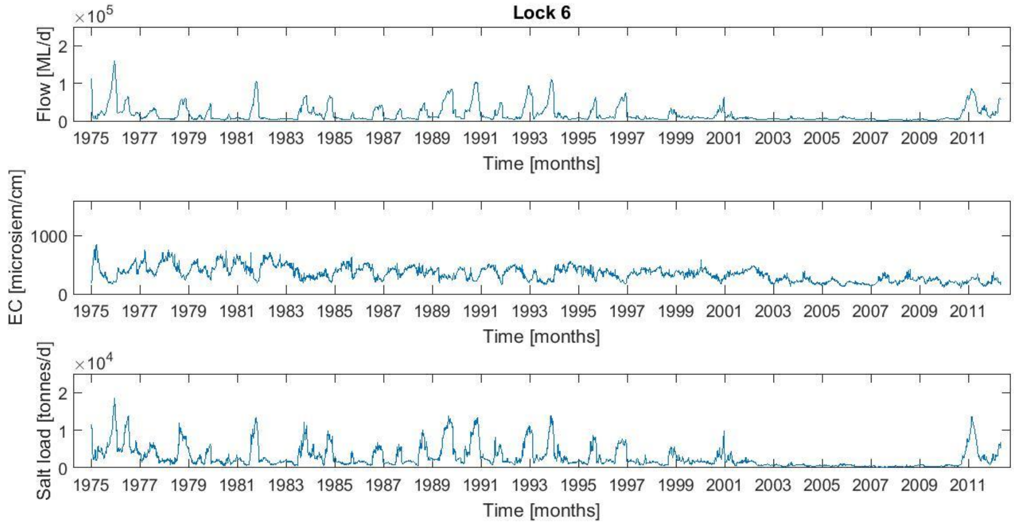

Figure 27 Lock 6

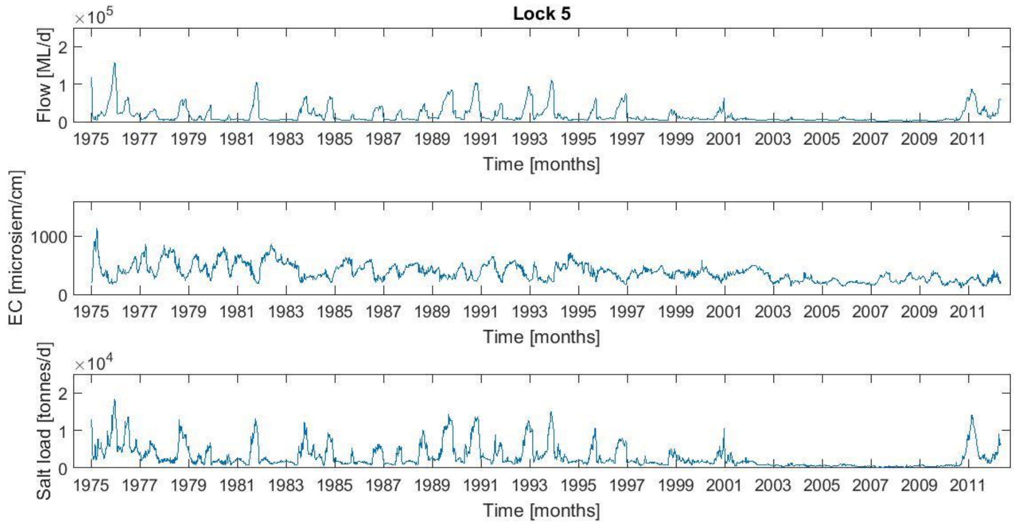

Figure 28 Lock 5

Lock 9, Lock 7, Lock 6 and Lock 5

When taking a closer look at the salinity at Lock 9 to Lock 5, Figure 29 shows that the salinity at Lock 5 is indeed higher than at the other locks. Further, Figure 29, Figure 30 and Figure 31 are showing perfectly that when the salinity is low, the flow and salt load are higher. Between Lock 9 and Lock 5 there are a lot of anabranches which increases the complexity of modelling.

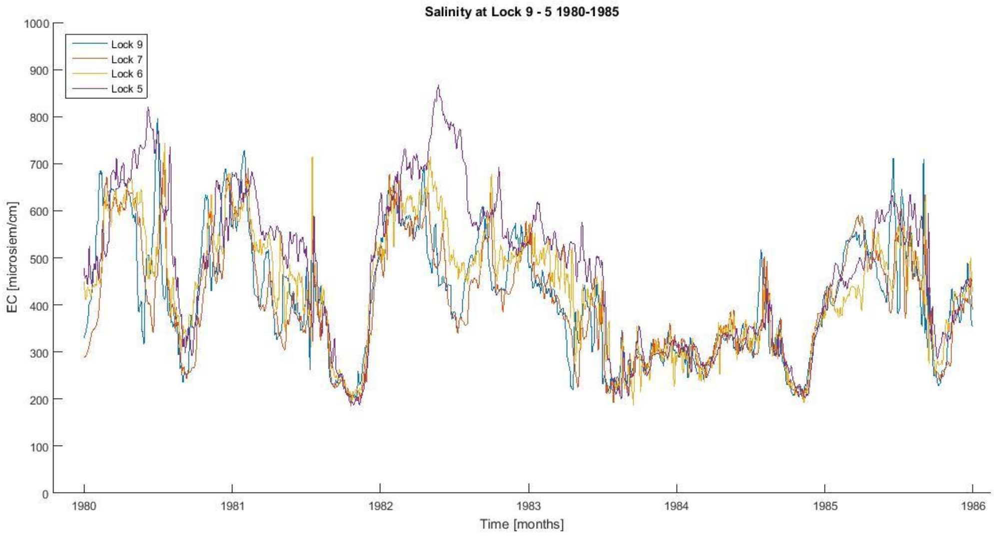

Figure 29 Salinity at Lock 9 - 5

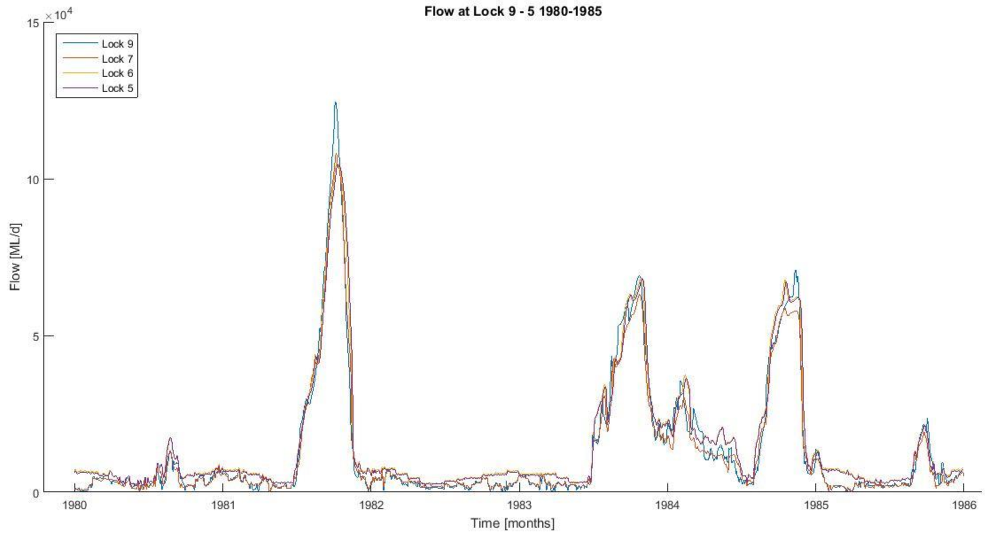

Figure 30 Flow at Lock 9 - 5

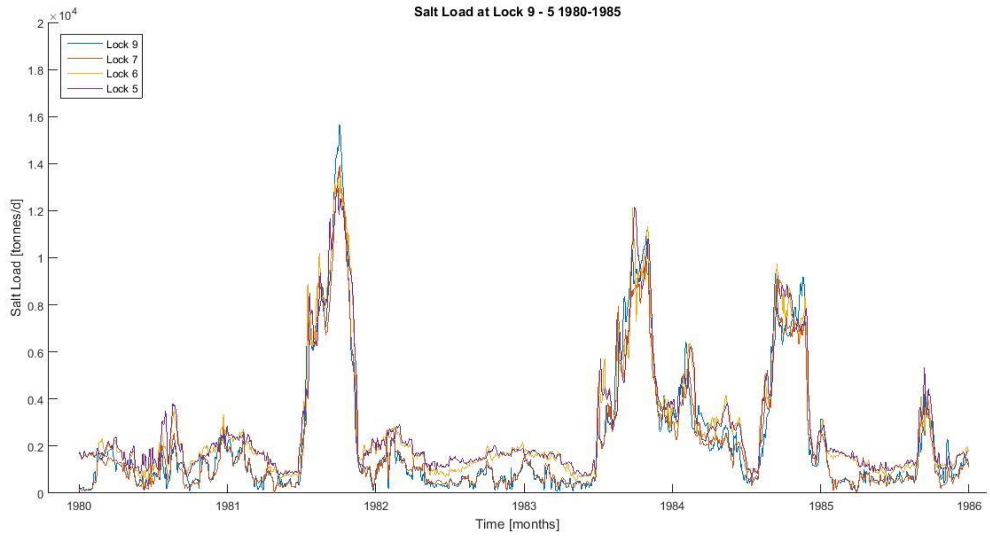

Figure 31 Salt load Lock 9 - 5

Lock 5 – Salt load, Flow and Salinity

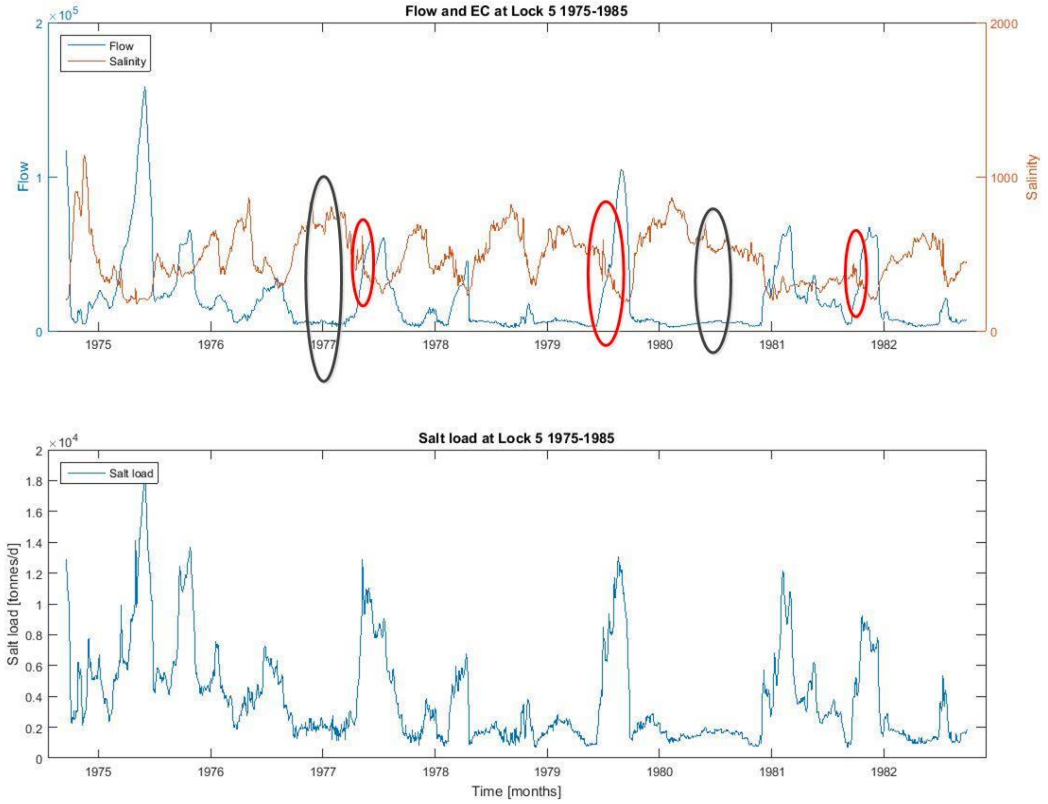

Figure 32 Lock 5

When comparing the salinity and flow data at Lock 5 over the years 1975 to 1985, Figure 32, black circles, it seems the amount of salinity is low when the flow is high and the other way around. This can be explained by the fact that salt comes predominantly from the groundwater sources and is not associated with the flow paths.

During a high flow peak, the salinity decreases. This can be explained by another process, the dilution capacity of water. Dilution is the main process for reducing the concentration of substances away from the discharge point. During high flow, the dilution capacity of the water will be higher. This means the salt concentration in the water will decrease.

An interesting observation is that during the initial days of the high flow, the salinity is less than during low flows but does not decrease very fast and still has some peaks. This can be explained by the fact that at low flows, the river becomes a continuous series of weir pools between Lock 9 and Lock 1 where Lock 5 is in between these two locks, Figure 33 (Telfer et al., 2012). There is a part of a water body in the weirs which is called the dead storage volume. Telfer et al., have made some estimations about the travel time at several moments from Lock 5 to Morgan. Morgan is located between Lock 2 and Lock 1 as shown in Figure 33. These Locks are not included in the study area of this research, but the information about the travel time and salinity peaks might explain the salinity peaks in the plotted graphs of Lock 5. Obviously, the value of the travel time might not be the same. The travel time for flow from Lock 5 to Morgan is normally 2.87 days for flows of 5.000 ML/day and 10.000 ML/day, while the dead storage volumes result in the water body travelling between Lock 5 and Morgan over a period of approximately 29 days at flows of 5.000 ML/day and 16 days at flows of 10.000 ML/day (Telfer et al., 2012). This means the initial days of high flow at Morgan contains water that has flowed through the upstream reaches at a much lower rate. The lower rate provides lower levels of dilution to the incoming salt which means the amount of salinity is higher (Telfer et al., 2012). This might explain the high salinity peaks at the beginning of a high flow.

Figure 33 Lock 1 t/m 11, 15 & 26

5 Results: Uncertainties

In Chapter 5 the results for sub-question 2 (uncertainties in flow and salinity data) and 3 (uncertainties in salt load) are given. As described in section 3.2, the uncertainties according to question 2 and 3 are 'errors associated with input data and data for calibration'.

5.1 Q2: Uncertainties in flow and salinity data

Before the uncertainties are shown, the results of the literature review about the technique to obtain the usable data is described (5.1.1). Second, the general uncertainties 'measurement errors' which occur at the execution of the measurement of both flow and salinity data are described (5.1.2). Third, the accuracy according to the available data used in this research are described (5.1.3). After that, the uncertainties in the flow data are described (5.1.4) and the uncertainties according to the salinity data (5.1.5). The way the uncertainties can be reduced, are described in the recommendations 9.2.

5.1.1 Obtaining flow and salinity data

As described in 1.2.1, the flow data is not directly measured, but obtained by using a rating curve which converts the observable quantity *stage height* (h) into the discharge rate (Q). This way of measuring is less labour intensive than repeatedly making velocity and depth measurements over the width of the channel (NationalWeatherService). The Q is the discharge; h is gauge height of the water surface; $h_0$is gauge height of zero flow for a control of regular shape; $(h - h_0)$ is depth of water on the control; a is the discharge when $(h - h_0)$ equals one and b is the slope of the rating curve (WMO, 2010b).

$$Q = a(h - h_0)^b \tag{15}$$

To define the rating curve, the cross section area A and velocity v of the river are measured:

$$Q\ (m^3/s) = A\ (m^2) * v(m/s) \tag{16}$$

As described in 1.2.1, the salinity data used in this research is obtained by using monitoring sites along the Murray River. The frequency of salinity recording varies from continuous to daily, weekly and monthly. Continuous monitoring of in-stream salinity covers more than 50 monitoring site. Generally, these sites are spaced between 20 and 30 kilometres apart at the start and end of river reaches (Jin & Close, 2012). Since the salinity data is used to estimate the salt concentration, the salinity data needs a correction. The relationship between salinity and salt concentration changes with water temperature, so salinity values are corrected to represent salinity at 25 degrees Celsius. There are no standards for calculating corrected EC. Most agencies in Australia adopt a rule of thumb where EC is increased 2% for each degree below 25 degrees Celsius (AuthorityMurrayDarlingBasin). The EC data which are used for modelling are the corrected EC value.

5.1.2 Measurement errors (flow and salinity data)

Errors which originated during the execution of the measurement are referred to as measurement errors. These measurement errors can have different causes. The estimate of the flow at a certain moment is determined by measurements of the water level h, the cross section A and the velocity v. Notice that only the water level h measurement is needed when the rating curve is already obtained and that the cross section A and velocity v are needed to create the rating curve. The measurement errors for salinity is caused by the direct measurement of salinity at a specific site.

1. Uncertainties in measured data (empirical quality)

The difference between measured discharge and actual discharge and salinity and actual salinity is caused by random and systematic errors. The random errors can be indicated as the spread in the measurement. The random errors can occur with or without a certain autocorrelation (Tillaart, 2010).

2. Uncertainties regarding the executing of the measurement (methodological quality)

There are different causes for this uncertainty. The first reason for an error in the execution of the measurements, are the conditions. During extreme conditions like high water levels, the standard procedures cannot be followed which occurs an error.

Flow: According to the flow data, another reason for a bad execution can originate if a certain flow is completely different from uniform flow, while an assumption is made that the flow is uniform (Tillaart, 2010). This has to do with the cross section of the river and the velocity of the flow. If the cross section is the same at every point in the river, the flow is considered uniform. Since the Murray River is a natural river, the cross section differs at different points. This is caused by for example sediment and vegetation growth (Baldassarre & Montanari, 2009). The velocity of the flow is also different due to locks or bridges (NSW, 2013). Depending on the location of the site, the uncertainty in the measured data is bigger.

Salinity: The second reason for a bad execution is as follows. According to Andrew Telfer et al., there is an uncertainty related to variations between the in-stream (point) salinity measurements and the salinity distribution across the river cross-section (Telfer et al., 2013). The distribution of the salinity depends on the velocity of the water which depends on the shape of the cross-section and the meander of the river.

3. Uncertainties regarding the performance of the measuring equipment

The equipment can break down or can be failing which occurs an uncertainty in the data (Tillaart, 2010). It is also possible the equipment is used in the wrong way.

5.1.3 Uncertainty in accuracy available data

In section 5.1.1 and 5.1.5 the method to obtain the data is described and the uncertainties during execution of the data are described (measurement errors). There are also uncertainties in both data sets which have an influence on the accuracy of the data.

1. Frequency of datapoints over time

The data is normally obtained at daily intervals, therefore intraday variations are not accounted for (Telfer et al., 2013). The daily data also might be the average value of that specific day. In this case one peak has a big influence on the average. By measuring the data every hour or minute, the intraday variations can be obtained. These are necessary to develop an accurate flow model.

2. Uncertainty in collected sites

Other uncertainties can be data management errors or a site location providing a non-representative sample (Telfer et al., 2013). The non-representative sample can be caused by the location of the site as described in section 'Uncertainties regarding the executing of the measurement (methodological quality).

3. Amount of datapoints in distance

Another uncertainty is caused by the fact that there is an interval between the sites. This means the salinity or flow at a certain point is known (inclusive the uncertainties during executing), but an assumption has to be made about the change in salinity or flow between two sites. This might be an anabranch or tributary which is easier to find out, but it also might be caused by saline groundwater which enters the main river flow.

5.1.4 Uncertainty in using rating curve (flow data)

The uncertainty in the flow values depends primarily on the uncertainty in the rating curve (and to a less extent on the uncertainty in the measurement of the river level h. The uncertainty in the rating curve results from uncertainty in the form of the functions used to fit the observed data, the discharge and stage measurement error during gauging, and potentially unaccounted rivers (Croke, 2009). The discharge and stage measurement errors during gauging are explained in the section 'measurement errors'.

1. <u>Uncertainty in fitting the rating curve with observed data</u>

Fitting the rating curve to the site or gauging is often carried out as a linear function through log-transformed data, using a minimized sum of least squares. Fitting the rating curve always causes uncertainties in for example the slope.

2. <u>Uncertainty due to changes in the rating curve over time</u>

The rating curve changes over time which occurs uncertainties. These changes are due to for example, vegetation growth, backwater effects and erosion or sedimentation at the gauge control site (Baldassarre & Montanari, 2009). It is unknown when the measurements to fit the rating curve are accomplished and if and when they are repeated due to changes in the river channel over time. It is also unknown how often a new rating curve is developed. When the current rating curve is often checked on reliability, the uncertainties due to changes at the site are smaller.

3. <u>Uncertainty due to the rate of rise of water level</u>

The rate of rise of the water leads to an increase in the flow velocity. This means the flow is actually higher during a fast rate of rise of the water level.

4. <u>Uncertainty during the process of updating the rating curve</u>

The decision to create a new rating curve in a specific gauge, depends on the experience and knowledge of the hydrologist. Once, when decided to create a new rating curve, the timing causes another uncertainty. The timing of the gauging's and start date of the rating curves are often not coincident with the timing of changes in the stage-discharge relationship, there is often a delay in the rating curves relative to the system. The impact of this uncertainty depends on how rapidly the river channel changes (Tomkins, 2014). Once the new rating curve is used, it might already been outdated again.

5. <u>More rating curves needed at one site</u>

Additional rating curve uncertainty arises through the problem of defining the timing of changes in the stage-discharge relationship when there is more than one rating curve at a specific site. The timing of gauging's relative to the timing of changes in the channel, and whether these changes are gradual or stage shifts, can have significant impact on the number of rating curves and the reliability of each curve.

For the sites Barham, Swan Hill and Wakool junction the rating curves are available. The rating curve for the site Wakool junction is created on 16/09/2012 and is used to 'present'. The site is last updated at 29/11/2012. The stream flow according to the water level higher than 12 m is coded as unreliable since these stream flow data are obtained by extrapolating the rating curve. Interesting is that the rating curve at Swan Hill has a maximum water level of 4.9 metres which suggest the water level there is much lower than at Wakool junction (Department-Of-Sustainability-Of-Environment, 2012). It is unknown if the current rating curves are still reliable (uncertainty 2), what the process is of creating a new rating curve (uncertainty 3) and if there is more than one rating curves at a specific site.

5.1.5 Uncertainty in correcting salinity data (salinity data)

As explained in the section 'Obtaining salinity data' a rule of thumb is used by most agencies to correct the EC values to representative EC values at 25 degrees Celsius. There is no information available about the influence of this 'rule of thumb'.

5.2 Q3: Uncertainties in salt load

In section 5.1, the causes of the uncertainties in the salinity and flow data are described, but the uncertainty in percentage is unknown since there is no detailed information available. These uncertainties in the input data are propagating in the salt load. Apart from the uncertainties in the data, there are also some uncertainties in the salt load according to the calculation of the salt load by using the salinity and flow data. In section 5.2.1 these uncertainties will be described. In section 5.2.2 the unaccounted and accounted salt loads will be compared with the unaccounted salt load from MSM-BIGMOD.

5.2.1 Uncertainties in calculating the salt load

In this section first the uncertainties in the salt load according to converting the salinity into the salt load using factor K are described. More information about this conversion is described in section 3.1. Second, the influence of the timing of the measurements and the time constant are described.

1. Conversion factor K

To convert the water salinity (EC) to salt concentration (mg/L) the conversion factor K is used. This factor K can vary between 0.4 and 0.97 $mgL^{-1}/\mu Scm^{-1}$ depending on the ionic composition of the water (nswgovernment). These values are giving a geometric mean of 0.62 with a range of 56%. Several reports are using a different value for K which is shown in Table 2, since the researchers do not know the ionic composition in the water or do not agree about the K factor. The species of the water in the study area of this research are unknown. This means there is a considerable uncertainty using the right K value.

Table 2 Different K factors

Factor K [$mgL^{-1}/ \mu Scm^{-1}$]	Argument/location	Source
0.64		(nswgovernment)
0.55	EC range of 100 to 2000 µS/cm, representative of salinity levels to the Lower Murray River	(Burnell et al., 2013)
0.65	Murray-Darling Basin	(NSWGovernment, 2000)
0.68		(NSW, 2013)
0.61 - 0.87	Border River Catchment in the Upper Darling Basin: Out of scope since too far away from study area of this research	(DepartmentofWaterandEnergy, 2008)
0.40 - 0.97		(nswgovernment)
0.55 - 0.90		(DepartmentofEnvironmentandConservation(NSW), 2004)

To get a global idea about the influence of the different K value on the salt load, a raw sensitivity analysis is done. The more extensive approach for this analysis is found in Appendix IV.

During the sensitivity analysis the K was changed and the salinity and flow values were kept constant. The K value varied between 0.3 mgL^{-1}/ µScm^{-1} and 0.9 mgL^{-1}/ µScm^{-1}. The assumption is made that there are in this case no uncertainties in the flow and salinity data. Another assumption is that the K value is constant. It is unknown if the K value changes if chemistry varies between for example high and low flows.

Figure 34 shows the salt load when changing the K factor. Since the relationship is linear, varying the K with 10% changes the salt load with 18%. However, the absolute change in salt load depends on the amount of flow and salinity since the influence from the flow on the salt load is bigger in comparison with the salinity.

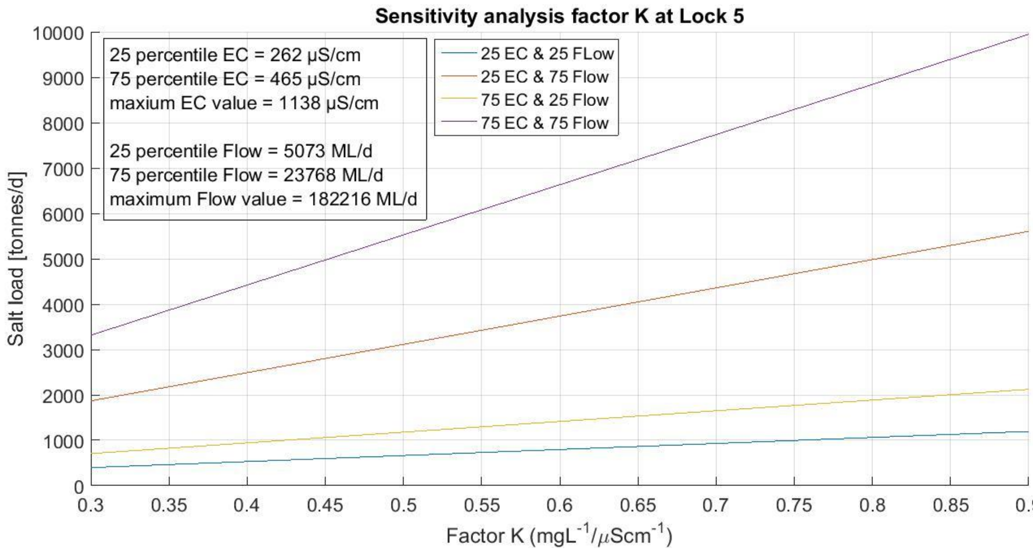

Figure 34 Sensitivity analysis factor K at Lock 5

Table 3shows the absolute change in salt load for site Barham (graph results are shown in Appendix IV) and Lock 5 (Figure 34). Since the different reports recommend a K value between 0.55 and 0.65, the output of the salt load varies with a maximum of 18% when changing the K between 0.55 and 0.65.

Table 3 Percentage change when changing K from 0.55 to 0.65

	Barham K=0.55	Barham K=0.65	Percentage change	Lock 5 K=0.55	Lock 5 K=0.65	% change
low EC & low flow	169.15	199.91	118%	731.02	863.93	118%
low EC & high flow	471.41	557.12	118%	3424.9	4047.6	118%
high EC & low flow	261.22	308.71	118%	1297.4	1533.3	118%
high EC & high flow	728	860.37	118%	6078.5	7183.7	118%

2. <u>Timing of flow and EC measurement</u>

The flow should always be measured at the same time as the dissolved salts. This is to assure calculation of total salt load can be made, although this will be dependent on the objectives of the monitoring program (NSW, 2013).

3. <u>Time constant</u>

The salt load is imputed through observations of flow and salinity. It is important to take the time constant of salt and flow changes into account. The salt load is very sensitive for the changes in flow.

MSM-BIGMOD uses a parameter called dead storage value, which is the volume of water held in the river at zero flow. This is used to differentiate between the time constant for flow changes and the time constant for salt (Telfer et al., 2012).

5.2.2 Unaccounted and Accounted salt load

In section 3.2.2 the differences between accounted and unaccounted salt load are explained. The additional unaccounted salt load which is used in MSM-BIGMOD to balance the salt load (the routed salt load at site x has to be the same as the measured salt load at site x) is compared with the unaccounted salt load from the observed data used in this research. More information about this approach is given in Appendix V.

In Figure 35 and Figure 36 is shown that the unaccounted salt load in MSM-BIGMOD (red line) is never below zero. This is due to the fact that salt load is adjusted in the model when the amount of salt load downstream is higher than upstream. No salt load is added when the salt load downstream is lower than upstream. The salt load from the observed data (blue line) shows that between Coligna and Lock 9 most of the time the amount of salt load downstream is less than the amount of salt load upstream. If the salt load upstream is bigger than downstream, it is unnecessary to adjust an amount of unaccounted salt load. When the observed data (blue) is below zero, the expectation is that the MSM-BIGMOD data is zero. However, in most situations this is not correct.

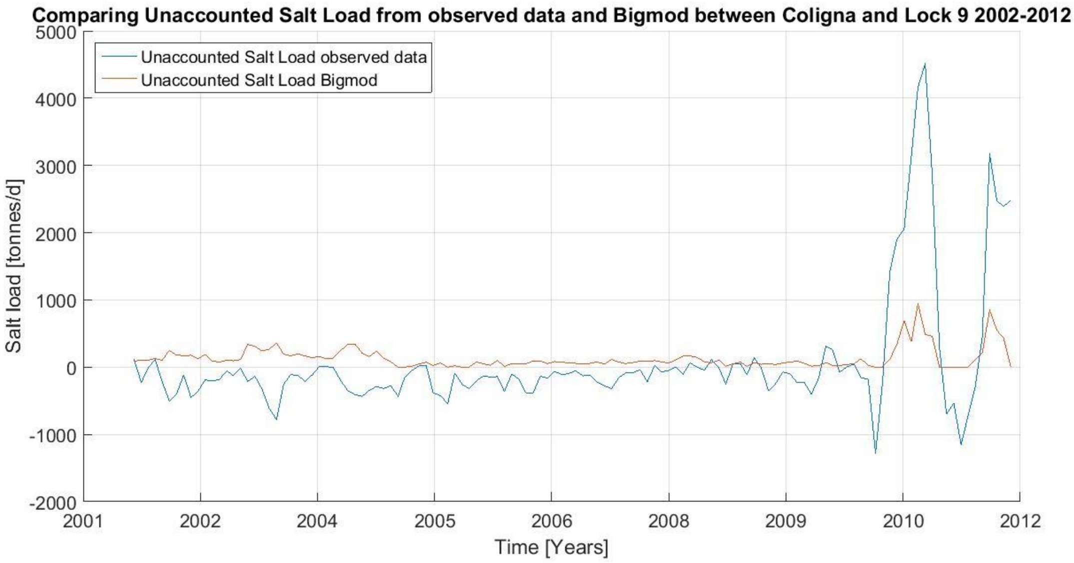

Figure 35 Comparing unaccounted salt load Coligna-Lock 9

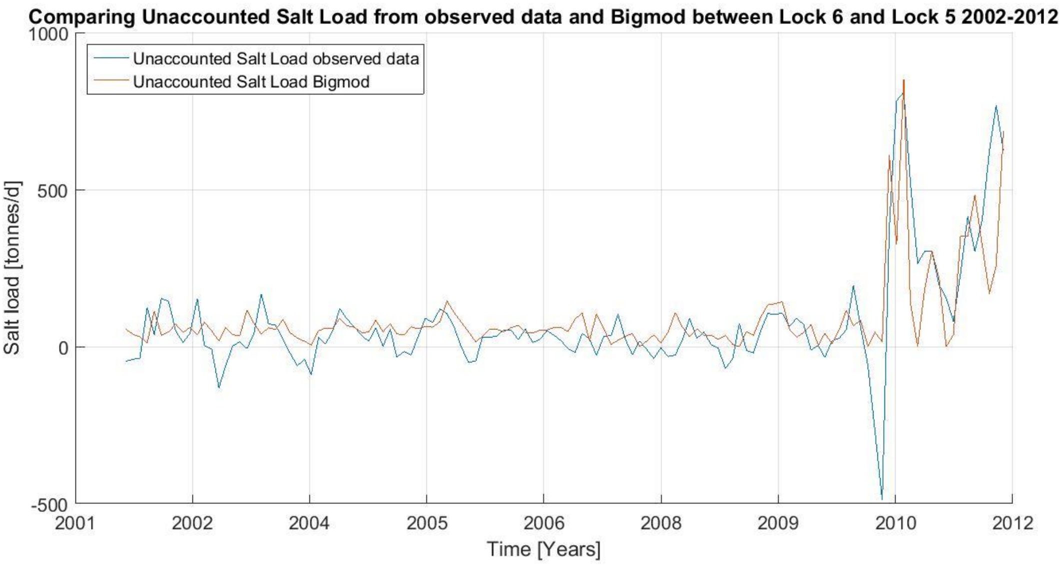

Figure 36 Comparing unaccounted salt load Lock 6-Lock 5

Figure 37 shows how this is possible. MSM-BIGMOD also includes data from tributaries and anabranches and if these extract salt and the routed salt load is less than the observed salt load, an amount of unaccounted salt load has to be added. The same observed amount of salt loads in the observed data can lead to a negative amount of salt since the tributaries and anabranches are not included. When the tributaries and anabranches add an amount of salt, it can be possible the unaccounted salt load in MSM-BIGMOD is bigger than the unaccounted salt load in the observed data. Another factor which explains the difference in the two graphs is that MSM-BIGMOD has taken the time constant of the salt and flow into account and the observed data has not.

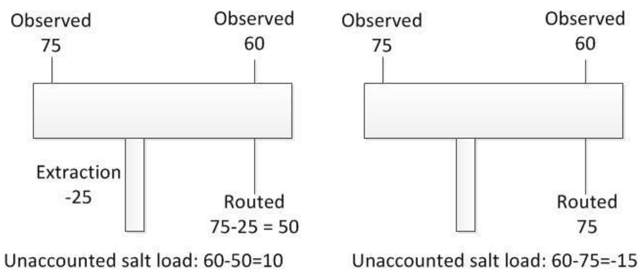

Figure 37 Left unaccounted salt load MSM-BIGMOD; Right unaccounted salt load observed data

6 Results: Model Structure

In this chapter the results for sub-question 4 'The most suitable reach for conceptualizing a flow model' and sub-question 5 'Top-down conceptual flow model' will be described. Chapter 3 explains the approach which is used to obtain the results.

6.1 Q4: The most suitable reach for conceptualizing a flow model

The reach for conceptualizing the flow model is chosen based on the data analysis in Chapter 4. The 9 sites are divided in three reaches. The second reach (Wakool Junction and Coligna) and the third reach (Lock 9 to Lock 5) are used for modelling. This is due to the significant difference between the flow from Reach 1 and Reach 2 (average percentage change of 69%). This is due to anabranches (section 4.2) which are at first not modelled. The second reason is because of the regulation of the Little Murray which is an anabranch at Reach 1. This regulation has a significant influence on the salt load within Reach 1. This means, when modelling a salt model in future research, the main challenge is to capture these regulations. In appendix X is shown what the parameter values are when using the longer reach from Coligna to Lock 5.

6.2 Q5: Top-down conceptual flow model

The extensive approach for modelling the flow model is described in section 3.3.2. To find an answer on this sub-question, an iterative approach is used. The top-down approach means the process starts with analyzing the available data. Depending on the conclusion drawn on base of the data analysis, the model structure is decided (Step 1 'Model Structure'). After every iteration step, the results are analyzed and depending on these results, the next step is formulated. The four steps of one iteration are: step 1 'Model Structure', step 2 'Calibration', step 3 'Model performance using observations and objective functions' and step 4 'Parameter uncertainties using Monte Carlo'. These four steps needed two iterations, which means they are done twice.

6.2.1 Calibration and Validation

In this research the split-sample test in combination with the idea of the k-fold partitioning test is used. The data from the years 1970-1985 is used for calibration and the data set from 1985-2012 is used for validation. Both validating and calibrating sets are divided in new sets. Since the flow and salinity data in 1972 are incomplete and since it is important to start the calibration of the model in a dry period, the calibration step is done for the period November (summer) 1976 to December 1985. The validation set starts at 1985-1996 and then 1997-2012. The validation will also be done for the whole period, but since the drought in Australia started in 1997 to 2010, it is important to validate the model for both the dry and the 'normal period'.

6.2.2 Step 1: Model Structure (Iteration 1)

As described in section 3.3.2 there are several parts to obtain the Model Structure. The results of these parts are given in this section.

Autocorrelation and cross-correlation

Figure 38 shows the autocorrelation of the input flow at Wakool junction and the cross-correlation of the flow at Wakool junction and Lock 5. The autocorrelation shows the seasonal period of the input flow. The peak of the cross-correlation coefficient shows the degree to which the input flow represents the output flow (Croke & Littlewood, 2005). The negative r means there is an anti-correlation between the input and output flow. Further, the cross-correlation graph shows the response of the catchment across the data period. This response shows the seasonal periods. The flow during the winter season is higher which means there is more resemblance during this period between the input and output flow which gives a higher correlation coefficient during winter.

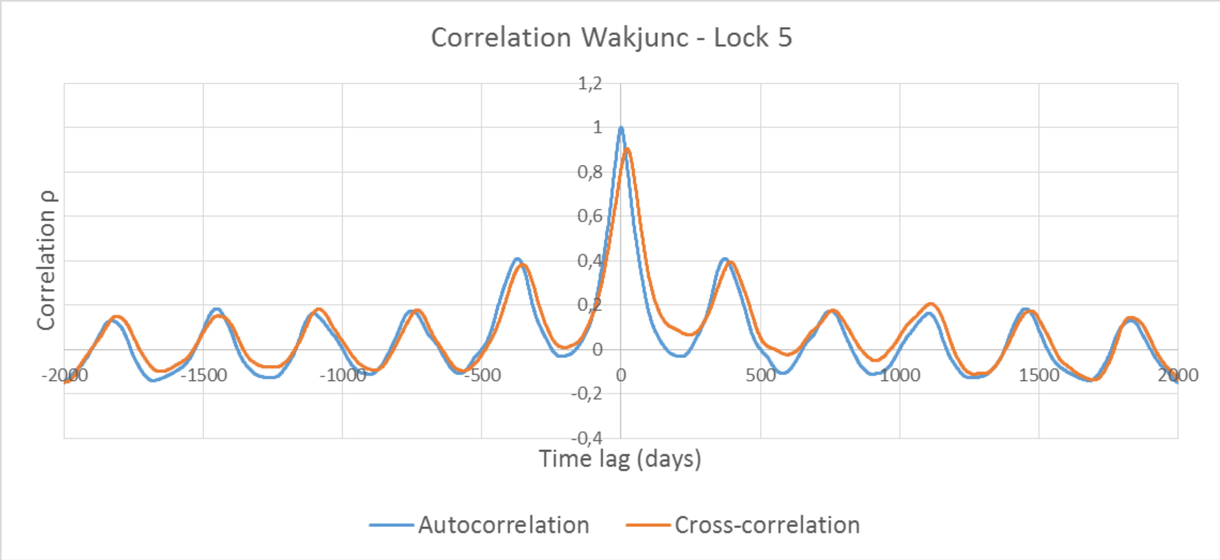

Figure 38 Correlation Flow Wakool junction-Lock 5

Figure 39 shows the correlation when zoomed in. The peak of both flows causes the biggest resemblance in the cross-correlation which in this case is around 0.87. Further, the cross-correlation function shows the influence of the catchment response function (Croke & Littlewood, 2005). The delay of the peak flow is 32 days. This means that the shapes of the two flow graphs have the most resemblance after shifting the input graph with 32 days. The analysis shows the slow response of the catchment to flow inputs.

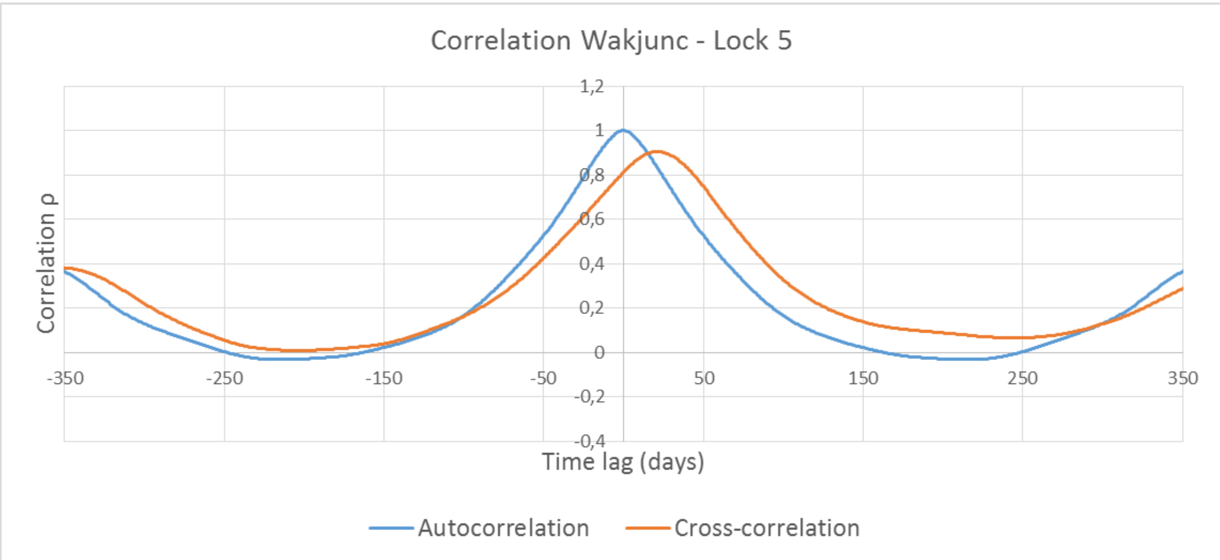

Figure 39 Zoom of correlation Flow Wakool junction-Lock 5

Shape Unit hydrograph obtained from deconvolution

Figure 40 shows the shape of the non-parametric Unit Hydrograph. There are two different peaks shown. The first peak has a fast rising limb, where the second peak has not. The first peak is referred to as quick flow pathway. This is used to infer the overland and shallow subsurface contributions to streamflow. The second peak, referred to as the slow flow pathway, interacts with the bedding of the river. It is used to infer the groundwater contributions to slow flow (Ivkovic et al., 2014). However, this division in quick and slow pathway may be an artifact of representing the transport mechanism as a combination of exponentially decaying stores, rather than physical processes. The quick flow component might have a shorter time constant than the slow flow component (Ivkovic et al., 2014).

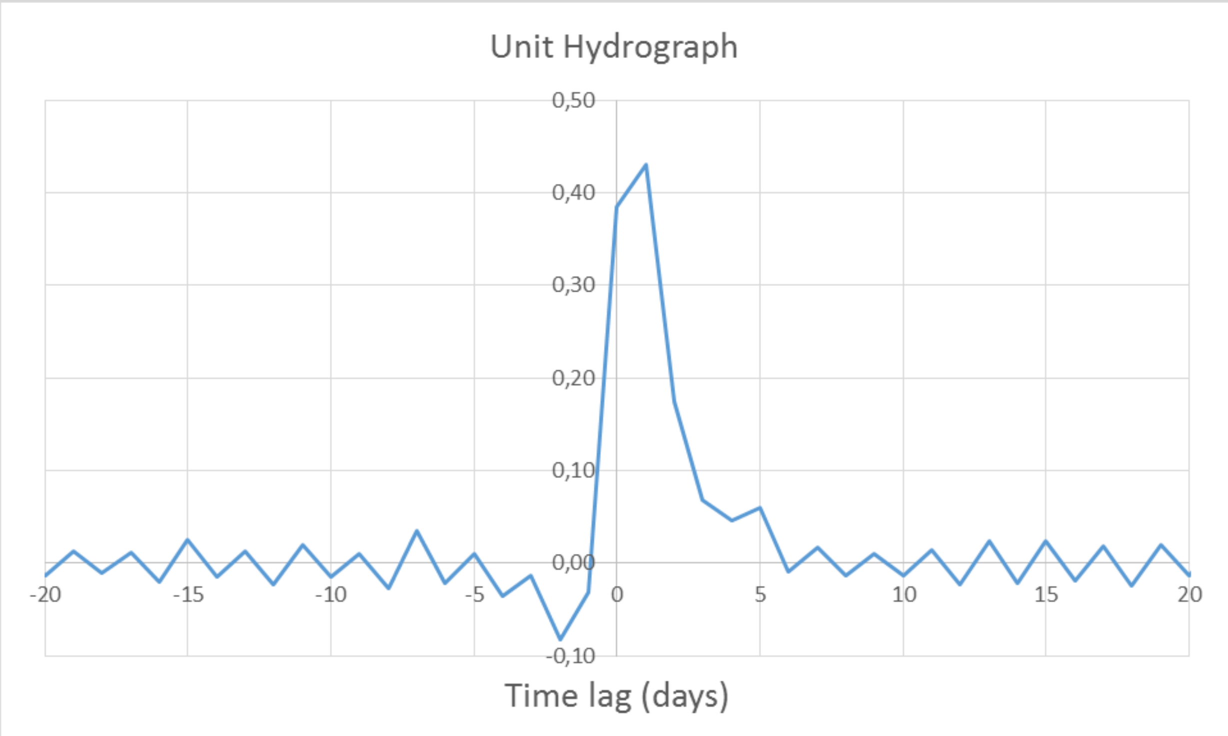

Figure 40 Shape of non-parametric empirical estimate of Unit Hydrograph

Formulation of equation Unit Hydrograph
After the shape of the Unit Hydrograph is obtained, a possible formulation of the equation of the Unit Hydrograph that has the potential to reproduce the shape of the UH of this research was formulated.

The general equation for the Unit Hydrograph is shown in equation 17 (Jakeman et al., 1990). In the UH there are two parameters, time constant τ for a linear reservoir and the parameter n for the number of stores (Nash cascade). The Nash cascade connects the linear reservoir in series. The relation with the parameters shown in this equation is given in the next session.

$$Q(\zeta) = (\frac{\beta}{1+\alpha\zeta^{-1}})^{n} * I(\zeta) \qquad (17)$$

The shape of the Unit Hydrograph represented in Figure 40 suggests a model structure with two different flow paths, the quick flow and the slow flow. The total flow input is divided over the quick flow and slow flow which is shown by using the $V\ quick$ (volume) and $V\ slow$ (volume). Further, the slow flow has a very slow response to the input flow and the peak of the slow flow is very smooth. These two characteristics suggest the Nash cascade is needed (Croke & Shin, 2015). The more n stores are used, the smoother the peak is. Figure 41 shows the model structure that has the potential to reproduce the shape of the Unit Hydrograph from Figure 40.

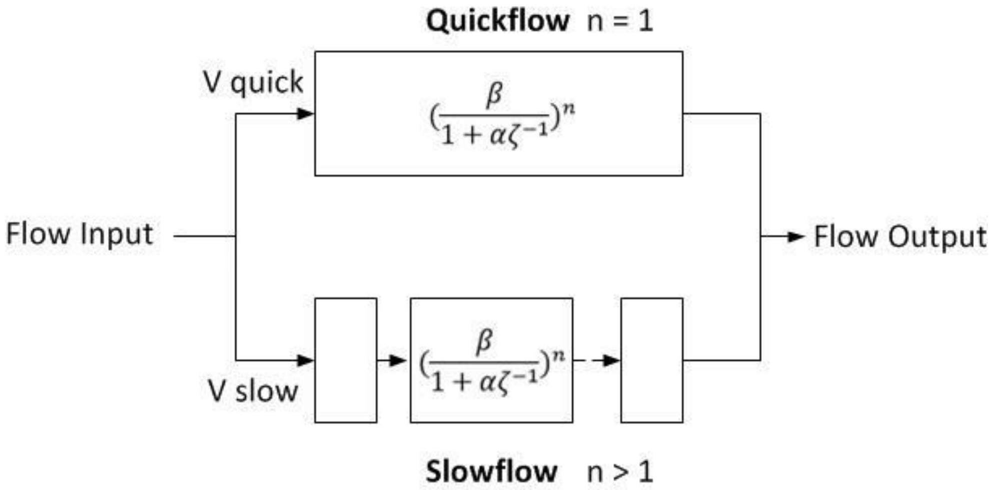

Figure 41 Model structure with general unit hydrograph equation

The parameters for the quick flow and slow flow are different. In the following part the general equation will be transformed to two different iterative equations for both quick flow and slow flow. The amount of stores in the slow flow is unknown. The situation where n is one and n is two will be given.

Quick flow

$$Q(\zeta) = \frac{\beta_q}{1 + \alpha_q \zeta^{-1}} * v_q * I(\zeta)$$

$$Q(\zeta) * \left(1 + \alpha_q \zeta^{-1}\right) = \beta_q * v_q * I(\zeta)$$

$$Q(\zeta) + \alpha_q \zeta^{-1} * Q(\zeta) = \beta_q * v_q * I(\zeta)$$

$$Q_{q,k} = -\alpha_q * Q_{q,k-1} + \beta_q * v_q * I_{q,k-\delta} \qquad (18)$$

Slow flow when n=1

$$Q_{b,k} = -\alpha_s * Q_{s,k-1} + \beta_s * v_s * I_{s,k-\delta} \qquad (19)$$

Slow flow when n=2

$$Q(\zeta) = \left(\frac{\beta}{1 + \alpha \zeta^{-1}}\right)^2 * v_s * I(\zeta)$$

$$Q(\zeta) * (1 + \alpha \zeta^{-1})^2 = \beta^2 * v_s * I(\zeta)$$

$$Q(\zeta) * (1 + 2\alpha \zeta^{-1} + \alpha^2 \zeta^{-2}) = \beta^2 * v_s * I(\zeta)$$

$$Q(\zeta) + 2\alpha \zeta^{-1} * Q(\zeta) + \alpha^2 \zeta^{-2} * Q(\zeta) = \beta^2 * v_s * I(\zeta)$$

$$Q_k + 2\alpha * Q_{k-1} + \alpha^2 * Q_{k-2} = \beta^2 * v_s * I_{s,k-\delta}$$

$$Q_k = -2\alpha * Q_{k-1} - \alpha^2 * Q_{k-2} + \beta^2 * v_s * I_{s,k-\delta} \qquad (20)$$

Parameter α

The equation of the Unit Hydrograph is a combination of a linear reservoir and a linear channel. The α represents the time constant τ which represents the time constant in a linear reservoir (Jakeman et al., 1990). In this research the delta t is one since the data is daily.

$$-\alpha = e^{-\frac{\Delta t}{\tau}} = e^{-\frac{1}{\tau}} \qquad (21)$$

Parameter δ

The δ represents the translation time. A linear channel will translate any inflow hydrograph without changing the shape (Dooge, 1959).

Parameter β

The assumption is made that the SSG (Steady State Gain) is one. This means there is no water mass which goes out or in in between the input and output of the reach. The general equation in section 3.3.2 Step 3 can be written as follows.

$$\frac{Q(\zeta)}{I(\zeta)} = \frac{\beta}{1+\alpha\zeta^{-1}} = 1 = G_{ss} \qquad (22)$$
$$\alpha = \beta - 1$$
$$\beta = \alpha + 1$$

Parameter v

The model structure has two different flow paths, the slow flow and quick flow. The total flow will be divided over these two flow paths. The total volume for both slow flow and quick flow in percentage has to be one. This means that:

$$v_s = 1 - v_q \qquad (23)$$

New equations

When expressing the β in α and the v_b in v_a, the following two equations for quick flow (n=1), respectively slow flow (n>1) are occured:

$$Q_{q,k} = -\alpha_q * Q_{q,k-1} + (1 + \alpha_q) * v_q * I_{q,k-\delta} \qquad (24)$$
$$Q_{s,k} = -\alpha_s * Q_{s,k-1} + (1 + \alpha_s) * (1 - v_q) * I_{s,k-\delta} \qquad (25)$$

At the end the total flow output can be obtained by summing $Q_{q,k} + Q_{s,k}$. Form this point forward, the α will be referred to as the τ since:

$$\tau = -e^{-\frac{1}{\alpha}} \qquad (26)$$

6.2.3 Step 2: Calibration (Iteration 1)

After calibrating using least square, the chosen reach (Reach 2 + Reach 3: Wakool junction – Lock 5), for modelling is changed to Lock 9 – Lock 5 (Reach 3). This is caused by the fact that some flow peaks are not covered due to anabranches and tributaries. These results are suggesting a gain module need to be added to the model structure. The first focus of this research is on producing a simple model which can be used to configure where this model cannot be used. Adding a gain to the model is for a later stadium. More information about changing the reach is given in Appendix VI.

Reach Lock 9 – Lock 5

The most optimized parameters are found by trying different amount of stores for the slow flow. The parameter values with the highest NSE value are shown in Figure 42. This model structure has a NSE value of 0.97 which means the modelled output is resemblance with the flow output for 97%. The appendix VII gives information about the effect of changing the different parameters for the Unit Hydrograph. This gives a better idea about the effect of changing one parameter for the model output.

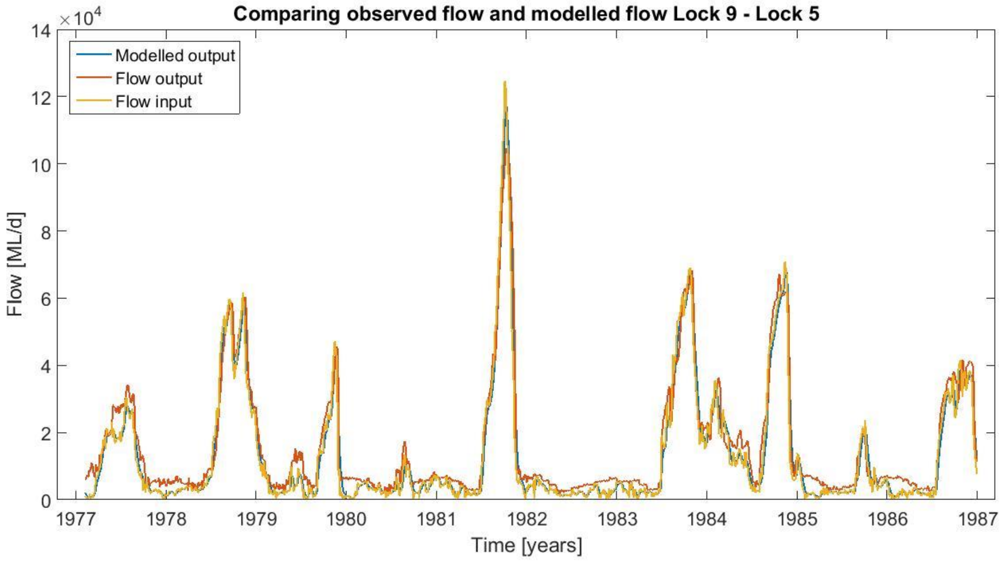

Figure 42 Modelled flow Lock 9 - Lock 5

Table 4 Parameter values

Stores	*NSE*	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,10]	0.96938	0.061804	0.025844	0.96307	0.84202	0.998

6.2.4 Step 3: Model performance using analysis and objective function (Iteration 1)

The model performance is first observed by plotting graphs and looking wat the graphs are showing (step 3.1). At the end, the model performance is tested by using objective functions (step 3.2).

Step 3.1: Observations

The first model performance test is plotting the residuals against time, Figure 44, and comparing this with Figure 43. Both graphs are showing that, although the modelled output covers the highest peaks, it does not cover the lower flows, purple circles. Further, the blue circle shows that the flow output is lower than the modelled output. This suggests there is a floodplain which gains water when the flow in the Murray River reaches a specific value. The red circles are showing the situation during average flow peaks where the modelled output does not capture the flow output. This suggests there is an extra tributary or anabranch which gains water to the Murray River.

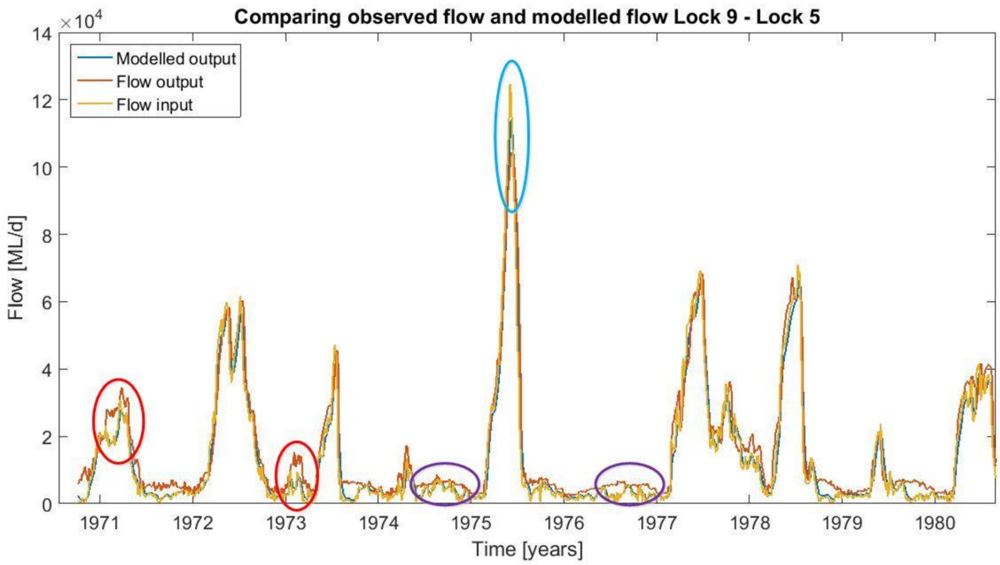

Figure 43 Modelled flow output

Table 5 Parameter values

Stores	***NSE***	***τ quick flow***	***τ slow flow***	***volume***	***δ quick flow***	***δ slow flow***
[1,10]	0.96938	0.061804	0.025844	0.96307	0.84202	0.998

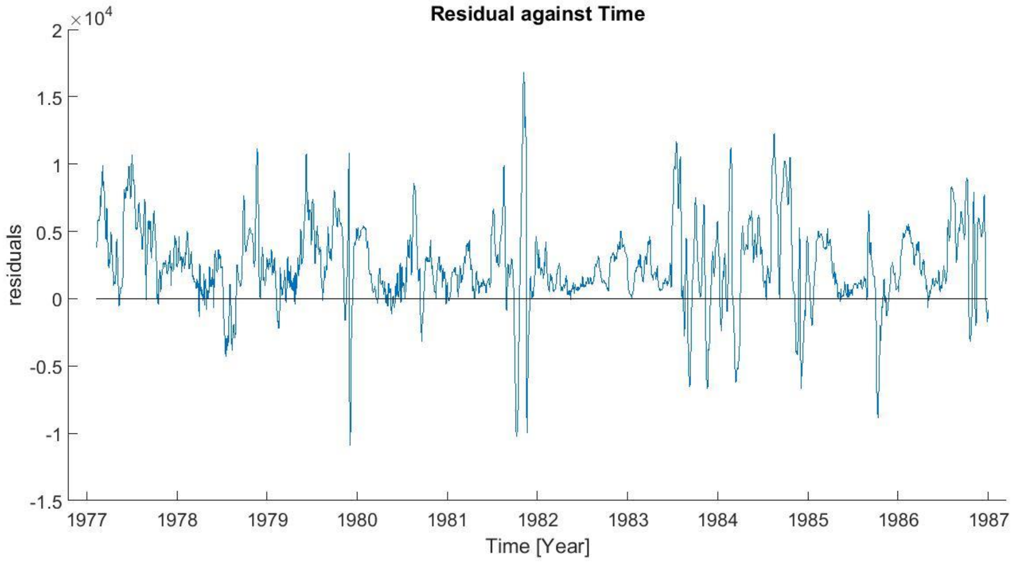

Figure 44 Residuals against time

<u>Model parameters modification</u>

Several modifications are tried to capture the lower flows (purple circle). Changing the amount of stores for the low flow while keeping the other parameters constant, does not influence the modelled output on the lower flows. Changing the time constant tau also has no influence on the capacity of the model structure to capture the lower flows.

When changing the volume parameter for the quick flow to 0.8 and thus for the low to 0.2 while keeping the other parameters constant, the lower flows are more captured as shown in Figure 45 (purple circle). Although the lower flow is more captured, the model still does not capture the flow peaks (red circle). There are even more peaks (between 1978 and 1979) which are not covered by the modelled output in this situation. There are also at least two situations where the lower flow is not captured by the modelled flow (green circles). This can be caused by or a data error, or additional slow flow which takes a longer time before it reaches the main river.

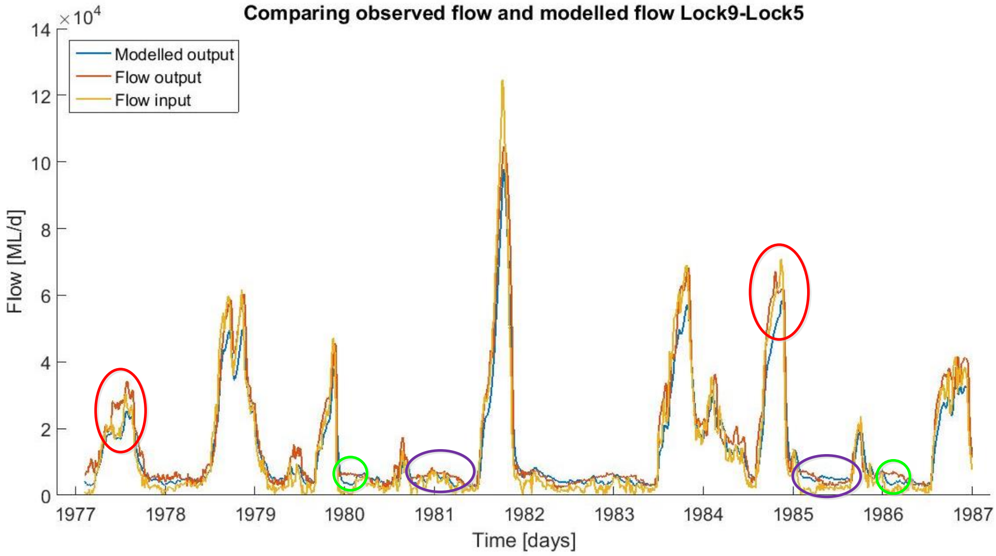

Figure 45 Modelled flow Lock 9 - Lock 5

Table 6 Parameter values

Stores	***τ quick flow***	***τ slow flow***	***volume***	***δ quick flow***	***δ slow flow***
[1,1]	0.061804	0.025844	0.8	0.84202	0.998

Since the peak flows are less captured when changing the volume than in the original situation, another solution is needed to capture the lower flows by using an additional input. This might be from the groundwater or an anabranch/tributary. By dividing the reach in smaller reaches, the location of the additional input can be obtained, which is shown in Appendix VIII. The result is that the additional inflow occurs between Lock 7 and Lock 6. Around Lock 6 the Chowilla floodplain causes the additional inflow. There is a lot of groundwater which discharges to the river or water which recharge to aquifer from the river (Telfer et al., 2013). This suggest an additional gain module is needed between Lock 7 and Lock 6 to represent the lower flow.

Step 3.2: Objective function
The results of testing the model structures with different parameter values are shown in Table 7.

Table 7 Objective function results

Model structure with 5 parameters										
	Stores	**NSE**	**Log NSE**	**RVE**	**Y**	***τ quick flow***	***τ slow flow***	***volume***	***δ quick flow***	***δ slow flow***
1	[1,1]	0.96826	0.71379	0.15037	0.84169	0.080201	0.024522	0.50749	0.81651	0.87475
2	[1,2]	0.96826	0.71379	0.15037	0.84169	0.080201	0.024522	0.50749	0.81651	0.87475
3	[2,1]	0.96826	0.71379	0.15037	0.84169	0.080201	0.024522	0.50749	0.81651	0.87475
4	[1,5]	0.96824	0.7108	0.15038	0.84167	0.019548	0.17293	0.9911	0.85017	0.75811
5	[5,1]	0.96797	0.70553	0.1505	0.84134	0.0072984	0.0076599	0.035517	1.0009	0.83983
6	[1,10]	0.96937	0.80396	0.15132	0.84197	0.061804	0.025844	0.96307	0.84202	0.998
7	[10,1]	0.96937	0.80227	0.15131	0.84197	0.026353	0.060533	0.035929	0.99804	0.84232

6.2.5 Step 4: Parameter uncertainties using Monte Carlo (Iteration 1)

In section 6.2.3 the parameters are optimized. Also the indication of the fit between a model and the system being modelled is obtained by using different performance indicators. In this section the uncertainty in the model output due to uncertainties in the parameters will be obtained by using the Monte Carlo simulation (Croke, 2009). The approach for the Monte Carlo simulation is described in section 3.3.2.

Covariance matrix

The first step for the Monte Carlo simulation, is obtaining the covariance matrix. The covariance matrix for the model parameters which are giving the highest NSE (0.96938) will be used. Table 8shows the values of the diagonal of the covariance matrix. These values are representing the variance of the parameters. The variance for the first parameter τ *quick flow* is Not a Number and for the second parameter τ *slow flow* is infinitive. Both suggest that the model has problems with identifying these parameters. During optimization, the model structure comes up with a value, but nog a good one. It suggest these parameters have no influence on the model structure. There are infinitive possible values for this parameter when fitting the least square.

Table 8 Covariance matrix optimized situation

Parameter	Parameter value	Coordinates matrix	Covariance matrix
τ *quick flow*	0.061804	(1,1)	Not a Number
τ *slow flow*	0.025844	(2,2)	Infinitive
v quick flow	0.96307	(3,3)	1.7519e-05
δ *quick flow*	0.84202	(4,4)	0.00078633
δ *slow flow*	0.998	(5,5)	1.6954e-06
n stores quick flow	1		
n stores slow flow	10		

When switching the amount of stores in such a way that the quick flow has 10 stores and the slow flow has 1 store, the variance of the τ *quick flow* is infinitive and of τ *slow flow* is 9.1992e+07. The results are shown in Table 9. Interesting is that the volume for the quick flow is 0.03 which means 3% of the flow takes this path. In Table 8 the volume for the quick flow is 0.96 which means that 96% of the flow takes this path. It seems that the most optimized situation occurs when most of the flow uses the path where the amount of stores is 1. This also explains why the variance of the τ is infinitive since changing this parameter only influences 3% of the flow.

Table 9 Covariance matrix with stores 1 for slow flow

Parameter	Parameter value	Coordinates matrix	Covariance matrix
τ *quick flow*	0.026353	(1,1)	Infinitive
τ *slow flow*	0.060533	(2,2)	9.1992e+07
v quick flow	0.035929	(3,3)	1.744e-05
δ *quick flow*	0.99804	(4,4)	1.7726e-06
δ *slow flow*	0.84232	(5,5)	0.00077914
n stores quick flow	10		
n stores slow flow	1		

When changing the amount of stores in both quick- and slow flow to 1, the variance of parameter 1 τ *quick flow* is still Not a Number (NaN) and the variance of parameter 2 τ *slow flow* is still Infinitive. This means it is unnecessary to have more flow paths which means the second flow path can be eliminated. The model structure needs to be modified, going back to step 1 'Model Structure'.

6.2.6 Step 1: Model Structure (Iteration 2)

Instead of two flow paths, the quick flow and slow flow, only one flow path remains which is shown in Figure 46. More information about the modification of the equations is described in Appendix IX.

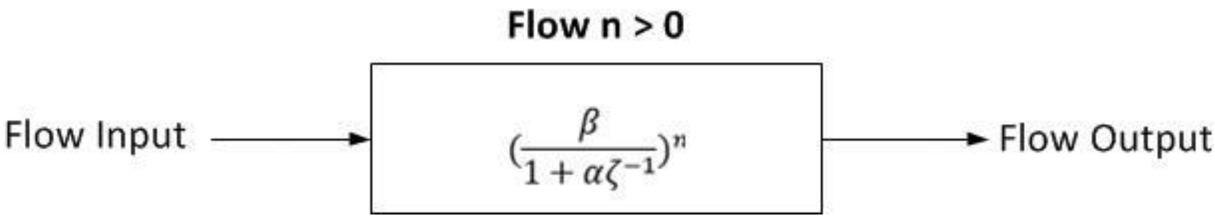

Figure 46 New model structure with one flow path

6.2.7 Step 2: Calibration 2 (Iteration 2)

Fitting the parameters for the new structure gives the modelled output as shown in Figure 47. The accompanying NSE value is 0.96824 which is less than in the situation with two flow paths (0.96938) but still very high.

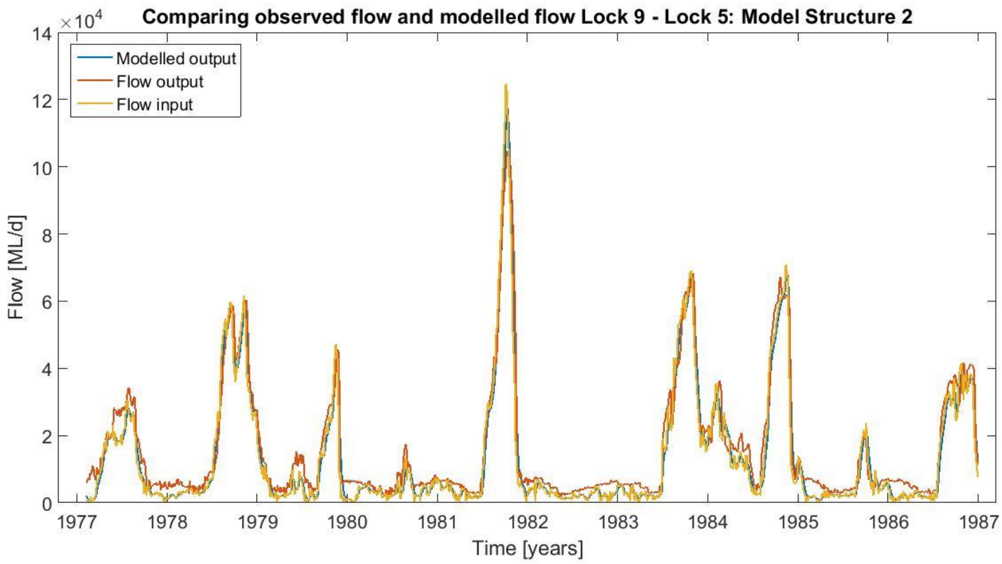

Figure 47 New modelled output

The parameter values are shown in Table 10. The parameter δ gives information about the speed of the water flow which is 2.77 m/s since the reach length is 203 metres.

Table 10 Parameter values new Model Structure

Parameter	τ	δ	*n stores*
	0.040067	0.84959	1

6.2.8 Step 3: Model performance using objective function 2 (Iteration 2)

Also during this step first observations are shown and second the results from objective functions.

Step 3.1 Observations

Figure 48 and Figure 49 showing the modelled flow with model structure two. The graphs are suggesting that still the same three processes are not captured. The red circle suggests an additional input from a tributary or anabranch, the blue circle suggests a floodplain (Chowilla floodplain) and the purple circles are suggesting an additional groundwater input. Further, the graphs are suggesting there is no difference between the modelled output in comparison with model structure 1.

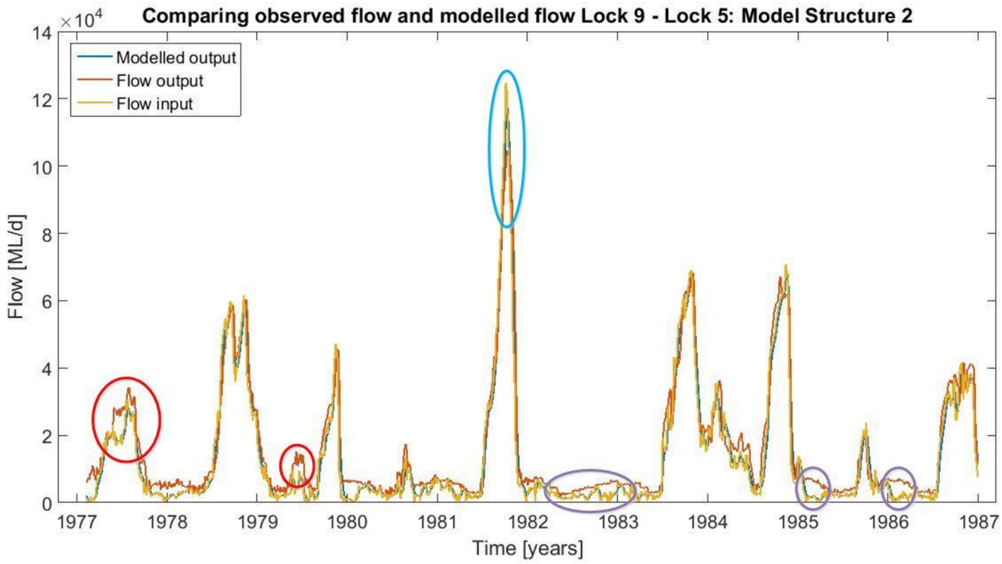

Figure 48 Model Structure 2

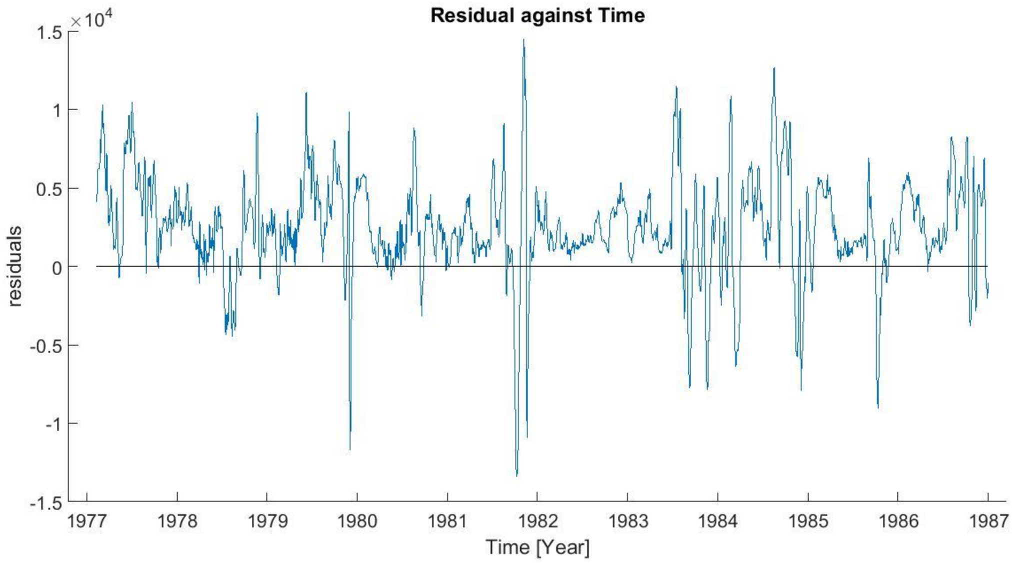

Figure 49 Residuals Model Structure 2

Step 3.2 Objective functions

Table 11shows the results from the objective functions. The amount of stores makes no difference for the different objective function results. This explains only one store is needed when the simple model structure with only two parameters is used. Further, the Y value in this situation is lower than when using Model Structure 1 (0.84197), although the RVE value is better when using Model Structure 1 (0.15132).

Table 11 Objective functions model structure with 2 parameters

Model Structure 2 parameters							
	Stores	**NSE**	**Log NSE**	**RVE**	***Y***	**τ**	**δ**
9	[1]	0.96824	0.71067	0.15039	0.84167	0.040067	0.84959
10	[2]	0.96824	0.71067	0.15039	0.84167	0.040067	0.84959
11	[5]	0.96824	0.71067	0.15039	0.84167	0.043881	0.84959
12	[10]	0.96824	0.71067	0.15039	0.84167	0.026592	0.84959

6.2.9 Step 4: Parameter uncertainties using Monte Carlo (Iteration 2)

The covariance value of the two parameters is -23.041 which means both variable values are independent. Table 12shows that the variance of the two parameters when using Model Structure 2 is not infinitive or not not a number. According to these outcomes, the values of the parameters could be:

$$\tau = 0.040067 \ \pm 72064$$
$$\delta = 0.84959 \pm 0.0030833$$

The τ could have a deviation of 179858736% where δ can have a deviation of 0.36%.

Table 12 Covariance Matrix Model Structure 2

Parameter	Parameter value	Coordinates matrix	Variance from covariance matrix	Standard deviation
τ	0.040067	(1,1)	5.1932e+09	72064
δ	0.84959	(2,2)	9.5065e-06	0.0030833
n stores	1			

Both standard deviations are not realistic. The standard deviation of the first parameter explains that there are too many possible parametervalues, where the standard deviation of the second parameter explains that the parametervalue is almost perfect. Both cannot be used for the Monte Carlo simulation; when creating 1000 random samples for parameter 1 with this standarddeviation, most parameter values become negative. When creating 1000 random samples for parameter 2, most outcomes are the same. Therefore, the Monte Carlo simulation is done fort he second parameter value since this value gives positive values, but it is important to realise the output is not realistic. The model structure needs to be modified to create a model structure where the least square method can calibrate the parameter values with more certainty.

Figure 50 shows the Monte Carlo simulation when creating 1000 random samples (normal distributed) for the second parameter. The 95% confidence boundaries are red and the mean flow output is green. There is a 95% confidence that the real flow value lies inbetween these two red boundaries.

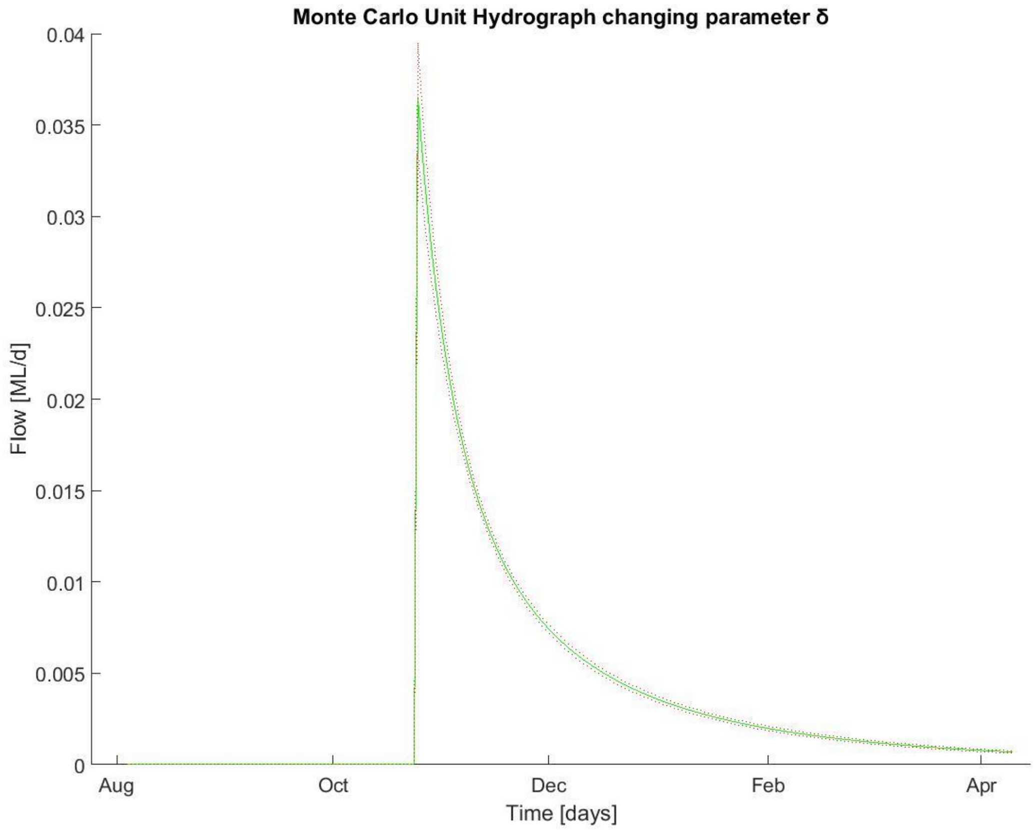

Figure 50 Monte Carlo changing parameter δ

6.2.10 Validation

For the validation Model Structure 2 is used with the following parameter values:

Table 13 Parametervalues Model Structure 2

Parameter	Parameter value
τ	0.040067
δ	0.84959
n stores	1

Table 14shows the values for the objective functions for Model Structure 2 with the calibration and validation period. Figure 51 shows the modelled output for the first validation period. The graphs for the second and third validation period are given in Appendix IX.

Table 14 Objective functions Validation and Calibration

Calibration and Validation Model Structure 2					
	Period [years]	NSE	Log NSE	RVE	*Y*
Calibration	Nov 1976 – Dec 1986	0.96824	0.71067	0.15039	0.84167
Validation 1	Jan 1987 – Dec 1996	0.97054	0.70583	0.15822	0.83796
Validation 2	Jan 1997 – April 2012	0.96401	0.78351	0.22805	0.78499
Validation 3	Jan 1987 – April 2012	0.97091	0.79051	0.1854	0.81905

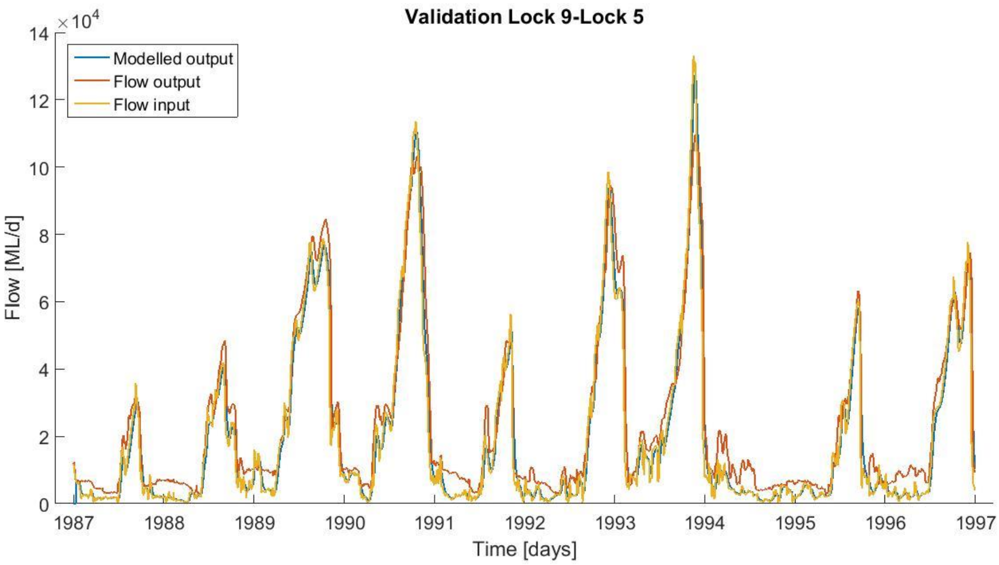

Figure 51 First validation period

6.2.11 Uncertainties using objective function NSE

Using the NSE to optimize the parameters and to test the model performance, means there are several assumptions made which makes the NSE used in this way incorrect.

<u>Identifying uncertainty 1</u>

The role of objective functions, like the NSE, is to give an accurate indication of the fit between a model and the system being modelled, called the model performance. However, failure to adequately account for variations in the errors in the observed and modelled quantities means that the objective function is only giving a measure of how well the modelled values represent the observed values, not how well the model is representing the system being modelled. The uncertainties in the model input (like flow and EC data, section 5.1) are not taken into account since there is not enough information available about these uncertainties. This is why in this research the residuals (observed output minus modelled output) instead of the uncertainties in the model input are used to calculate the NSE (Croke B. , 2012). Most researchers making the assumption that there is no uncertainty in the input data.

<u>Identifying uncertainty 2</u>

The second assumption is that the uncertainty in the flow data and other input data is often assumed to be homoscedastic (every data point has the same variance, the uncertainties are independent of the value). Instead, flow data is for example heteroscedastic. The NSE is dominated by the mismatch between observed and modelled values at high flows. As a result high flow events, which are often the most uncertain, are given too much weight (Croke B. , 2012). A lot of reports are not taken this assumption into account and are using the normal NSE equation.

<u>Reducing the errors caused by uncertainty 2</u>

The objective functions need to inform how well the model is fitting the system being modelled. In the presence of heteroscedastic uncertainties, an objective function must either account for the heteroscedasticity, or the values used in the objective function must be transformed so that the uncertainties are homoscedastic. There are several approaches which deals with the errors.

Chiew and Siriwardena (2005) opted to ignore the highest five flow values in calculating their objective functions in order to minimise the impact of the errors in the extreme high flows (Chiew & Siriwardena, December 2005). Another option is to modify the objective function. Lichty et al. (1968) used the logarithm to transform the data (Lichty et al., 1968). Another option is through introduction of weights so that the heteroscedasticity of the uncertainties is reduced. B. Croke (2012) tested the effect of the weighting average approach by using Monte Carlo. The results are that the standard deviation for the parameters are much smaller when the modified NSE is used instead of the normal. It is important to remember that such transformations are based on the assumptions of the nature of heteroscedasticity rather than from an analysis of the errors (Croke B. , 2012). Another transformation is using the Box-Cox transformation. This approach is based on the assumption that the scatter in the residuals is a good indicator of the uncertainties in the residuals. Providing this assumption is valid, this approach gives a very simple means of removing the heteroscedasticity, thereby ensuring that the objective functions give good measures of the model performance when the transformed values are used instead of the original values (Croke, 2009). There will be a value labda added which modifies the residuals by using log for instance. It is important to take into account that using the residuals and modifying these values means a part of the model error will be corrected. It is better to do the transformations on the uncertainties instead of the residuals, but in most researches the uncertainties of the data are not available.

Figure 52 shows the scatter of the variance of the residuals (modelled output minus flow output) on the y-axis and the observed flow values on the x-axis. The assumption is made that this scatter gives a good indicator of the uncertainties in the residuals. The values seems to be not very heteroscedastic which means the normal NSE can be used. Although the flow data does not look heteroscedastic, the waves shown in Figure 52 are giving the information that the variances are dependent. This means there is additional information in the data which the least square method is not using (Croke B. , 2009).

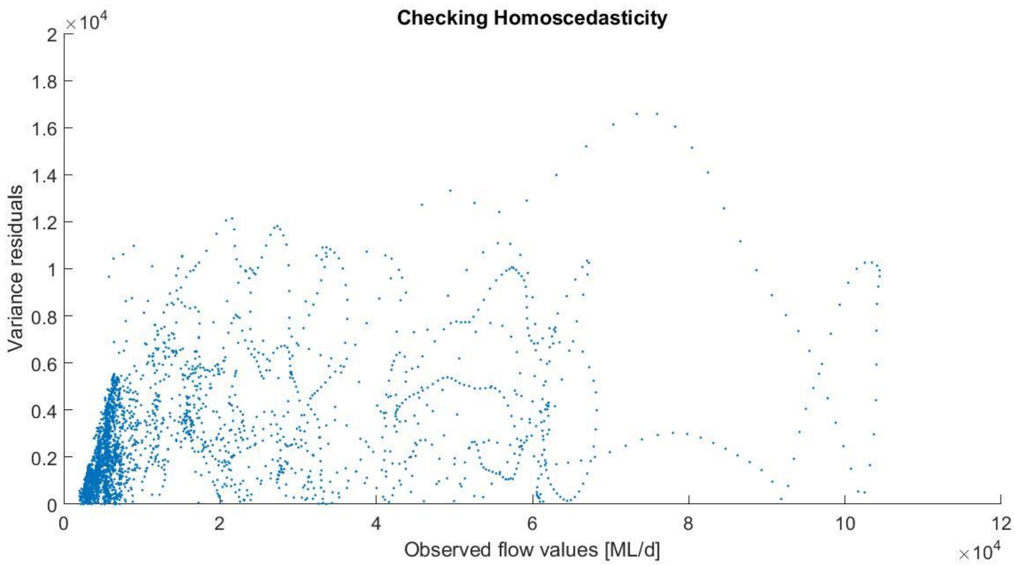

Figure 52 Homoscedasticity Flow 1976-1985

7 Discussion

in this chapter first the methodology and second the results per section (data analysis, uncertainties, model structure) of the research are critically reviewed.

7.1 Data analysis

The data analysis is based on a literature review. Past decades many strategies for regulating the irrigation and water levels in the Murray River are implemented. The Murray-Darling Basin is affected with dams that can store 103% of annual runoff (Kingsford, 2000). Another example is that in 1995 the Murray-Darling CAP was implemented which is a policy limiting the water diversions. This policy seeks to strike a balance between the amount of water available to irrigators, the security of their water supply and the sustainability of the Basin's river system (Providing security for water users and sustainable rivers, 2004). There are also a lot of management strategies implemented to reduce the salinity, such as Salt Interception Schemes and fresh water injections (Fitzpatrick et al., 2007). There are many reports about regulations which give information about the influence of the different strategies. However, these dams, locks, regulations and other strategies are making the analysis of the processes in the Murray River complex. First of all, every regulation has a different influence on the river processes which are sometimes clear but most of the time unknown. Second, it takes much time to include all the different regulations. Third, many regulations are temporary which is not always clear when reading the reports. In conclusion, for a good data analysis it would be important to create one document with an overview of the different regulations, the effect of the regulations and the location of the these regulations.

Apart from the technical and management strategies, the data measurements in the tributaries and anabranches along the Murray River are not taken into account in this data analysis due to the fact that the focus of this research lies on the main river and due to a lack of time and information. Analyzing these data helps understanding the processes and specifying which data analysis are caused by which process.

7.2 Uncertainties

The uncertainty analysis was based on a literature review. The most important uncertainties, such as the conversion factor K, the rating curve and the uncertainties according to the execution of the measurements are obtained. Still, it is important to take into account that it is unrealistic to expect all the studies about the Murray-Darling basin can be included because the available amount of time for this research project was only ten weeks.

Based on the literature review, only more detailed information about the values of the conversion factor K was found. The impact of the different uncertainties on the modelled output can only be obtained when more information about the uncertainties in the data input is obtained.

7.3 Model Structure

Chosen reach

The results of the objective function are showing a good model performance (NSE = 0.96938). It is important to realise that the input flow at the beginning of the reach is almost the same as the output flow at the end of the reach. This is due to the regulation of the Locks. If the actual output flow has the same pattern as the input flow, the modelled output flow has the same pattern as the input flow as well. This means that the objective functions (NSE et.) are showing high values.

Objective functions
As described in section 6.2.11, using the NSE as the objective function, means the assumption is made that there are no uncertainties in the input data. This assumption connects to the discussion part in the section 'Uncertainties' that it is important to obtain more information about the uncertainties in the input data.

Amount of stores
As shown in the results, the difference in parameter values and objective functions when using different amounts of stores were very small. In fact, as shown in Table 15, when optimizing the parameters of the first model structure, the volume of the quick flow was above 0.9 when the amount of stores for the quick flow was 1 and less than 0.1 when the amount of stores for the quick flow was bigger than 2. When the amount of quick flow and slow flow were both 1 or 2, the volume of both quick flow and slow flow was around 0.5. This suggest that the model structure did not use more than 1 store which means there is no Nash cascade needed.

Table 15 Volume per amount of stores

Stores [quick flow, slow flow]	***Volume [quick flow]***
[1,1]	0.50749
[1,2]	0.50749
[2,1]	0.50749
[1,5]	0.9911
[5,1]	0.035517
[1,10]	0.96307
[10,1]	0.035929

Further, when using the second model structure with only 2 parameters, only the τ (time constant) changed. Since the variance of this parameter was 5.1932e+09, this could also mean the parameter values could have been totally different (a wide range of possible parameter values) while still getting the same values for the objective functions.

The fact that the model structure did not need the Nash Cascade, can be caused by the daily data. Higher resolution of data is needed to see if the model structure captures the amount of stores.

River processes not captured by model structure
Although the objective functions showed high values for the model structure, there are still three river processes which need to be captured;

- The blue circle (Figure 48, 0) of the modelled output showed that the modelled flow had a higher flow peak than the observed flow. This is due to a floodplain along the Murray River between Lock 9 and Lock 5. This floodplain floods when the flow in the Murray River rises above a certain level:
- The purple circle (Figure 48, 0) suggest that an additional groundwater module is needed to capture the lower flows:
- The red circles (Figure 48, 0) are showing the situation during average flow peaks where the modelled output does not capture the flow output. This suggests there is an extra tributary or anabranch which gains water to the Murray River.

Monte Carlo

The covariance matrix obtained during the fourth step: ‘Parameter uncertainties using Monte Carlo’ shows the model structure needs to be modified because the standard deviation of the current parameters are not realistic. During optimization, the least square method has too many possible parameter values.

$$\tau = 0.040067 \ \pm 72064$$
$$\delta = 0.84959 \pm 0.0030833$$

8 Conclusions

In this section the conclusions per sub-question are discussed. At the end the overall conclusion, answering the main question, is given.

Q1: Obtained information from the connection between flow and salinity data

The data analysis shows the relation between the flow, salinity and salt load at the nine sites in the Murray River. The flow has a bigger influence on the salt load than the salinity. The results are showing that the Murray River is a very complex river system due to regulations (such as locks and dams), anabranches, floodplains, groundwater recharge and discharge and tributaries. These regulations and river processes are causing the differences and interaction in flow, salinity and salt load between several sites. For example, the percentage change between the site Barham and Pental is 79% in salinity, where the percentage change between Pental and Swan Hill is 24%. This is caused by the very saline Loddon River which flows into the Murray River downstream from Barham. The instream of this river is regulated which means it only streams into the Murray River when both the Loddon River and Little Murray exceed a flow of 12200 ML/d.

Q2: Uncertainties in flow and EC measurements

Based on the literature review and the first step (Identifying) from the framework to manage uncertainties, the different uncertainties are obtained shown in Chapter five. Based on these results and the results in the other chapters, a conclusion is derived how to reduce these uncertainties. This is shown in Table 16. Most researchers are making the assumption that the uncertainties in the input data are error free. Although the impact from uncertainties on the modelled output is not obtained, the amount of uncertainties are showing that it is not reliable to simply ignore the uncertainties in the input data and in the model structure.

Table 16 Framework to manage uncertainties data input

Uncertainties of both Flow and Salinity data	
Measurement Errors	How to reduce the uncertainties?
Measured data: - Random errors; - Systematic errors.	- In this situation the error can be reduced by using more accurate equipment to execute the data. There will always be an error since the actual data is unknown.
Executing measurement: - Errors due to extreme conditions; - Cross-section • Non-uniform flow; • Distribution EC across river cross-section.	- In this case it is important to determine the definition of good conditions and extreme conditions. Then the condition is known where the uncertainty is bigger. - It is difficult to reduce this uncertainty. It is important to have a good idea about the influence of this uncertainty on the data by looking at the cross section at different sites and measuring the velocity and water depth at different depth in the cross section at one site.
Performance measuring equipment: - Breaking down equipment; - Failing equipment; - Wrong use equipment.	- The error can be reduced by checking the equipment regularly and by obtaining data using different equipment. - The error can be reduced by checking the equipment regularly and by obtaining data using different equipment. - To prevent this situation, the user needs to be trained using the equipment.

Accuracy available data	How to reduce the uncertainties?
Frequency of data points over time.	By measuring the EC every hour or minute, the intraday variations can be obtained.
Uncertainty in collected sites.	A non-representative sample can be discovered by looking critically at the data during the data analysis.
Amount of data points in distance.	When decreasing the distance of the data points, the EC and flow processes between the sites can be more easily specified.
Uncertainty Flow data	
Uncertainty in using rating curve	How to reduce the uncertainties?
Fitting rating curve: - Parameter a; - Parameter b.	It is difficult to reduce this uncertainty. However, the uncertainties in the different parameters can easily be obtained by using a Monte Carlo approach.
Changes in rating curve over time: - Changes cross section.	To reduce this uncertainty, it is important to verify the rating curve more often.
Rate of rise of water level.	Research is needed to investigate what the influence of the rate of rise of the water level is on flow velocity.
Process of updating rating curve: - Qualities hydrologist; - Timing of gauging and using new rating curve.	To reduce this uncertainty, it is important to verify the rating curve more often.
More rating curves needed at one site: - Timing of changes stage-discharge relationship.	Measuring the velocity and cross section during different weather conditions will give inside in the amount of rating curves.
Uncertainties EC data	
Uncertainty in correcting EC data	How to reduce the uncertainties?
Converting EC data to EC data at 25 degrees Celsius.	It is important to investigate 'the rule of thumb' which is used. If this is the same at every site, the uncertainty will be the same.

Q3: Uncertainties in salt load

Table 17shows the results from the uncertainty analysis and gives information how to reduce these uncertainties. The comparison of the unaccounted salt load from MSM-BIGMOD with the unaccounted salt load from this research, shows big differences. This is due to the fact that the tributaries and anabranches are not taken into account in this research.

Table 17 Framework to manage uncertainties salt load

Uncertainties in salt load	
Calculating salt load	How to reduce the uncertainties?
Conversion factor K	When changing this factor with 10%, the salt load changes with 18%. Since the factor K found in literature vary between 0.45 and 0.90, the output of the salt load may vary between 0% and 180%. To reduce this uncertainty, the ionic composition of the water per reach needs to be investigated. This will reduce the range of the used factor K.
Timing of flow and EC measurement	To reduce this uncertainty, the equipment need to be tuned to make sure it measures at the same time.

Time constant	The time constant needs to be obtained during modelling. Both flow and salinity need their own time constant in the model structure.

Q4: The most suitable reach for conceptualizing a flow model

The most suitable reach for this research is the reach between Lock 9 to Lock 5, because less anabranches and tributaries are present in this reach. Since the river processes are that complex, it was important to choose a reach where the model structure can cover these processes more easily. In this situation a simple model structure is needed. When starting with less river processes, a better understanding of these processes occur since the influence of the other processes at this reach are less.

Q5: Top-down conceptual flow model

The objective functions show the model performance is very good (NSE > 0.9, RVE < 0.15). Still, there are three additional models needed, as listed in the discussion, Chapter 7, to capture all the river processes in that reach. The first is an additional floodplain module, the second is an additional groundwater module and the third is an additional gain module which covers the tributaries and anabranches.

What is the structure of a conceptual model suitable for use in modelling flow at a particular reach along the Murray river in the Murray-Darling Basin?

The main conclusion is that in this stadium it is too early to give a good answer on the main question about the structure of the conceptual model due to a lack of information. First, there is additional information needed about river processes. This research gives a detailed insight in the effect of anabranches, tributaries, floodplains and groundwater exchange. However, the knowledge about the processes of these anabranches etc. and their interaction with the main river need to be expand. More information is needed as well about the effect of the regulations implemented past decades to decrease the salinity in the River Murray. These aspects are needed when developing a conceptual model which covers the actual river processes as good as possible. The model structure developed in this research is a simple model structure with one store and one flow pathway. The model performance of this model structure is NSE 0.96 and RVE 0.15. The model structure for the specific reach, Lock 9 to Lock 5, needs an additional groundwater module, floodplain module and gain module. For the development of these modules, more information and more time is needed. Another aspect that is important when this conceptual model will be used for the development of a salt model, is information about the uncertainties in the input data and model structure. A study about uncertainties gives information about the reliability of the output of the model. There are many uncertainties in the model input and model structure. Most researchers are ignoring the uncertainties in the model input. This study shows for example the effect the uncertainty in the conversion factor K has on the model output. The method used to obtain the input data also shows there are many possible uncertainties, such as uncertainties in the rating curve. Further research is very important to make sure the most reliable model is used to study which management strategies are most effective to reduce the salinity in the Murray River.

9 Recommendations

In this chapter the recommendations are given for further research.

9.1 Data analysis

For the data analysis it is important to obtain more information about the regulations in the Murray River, anabranches, tributaries etc. This gives the opportunity to do a more detailed analysis of the differences in data at different sites. In this research the main differences are explained, but zooming in on the data, for example for specific years, will give a better understanding of the river processes. A better understanding will make it easier to understand which model structure modifications are necessary to develop a good model.

9.2 Uncertainties

As shown in this research, it is unreliable to make the assumption that uncertainties in the input data can be ignored. The first step in getting more detailed information about the uncertainties in the flow data, is gaining information about the rating curve per reach. When this information is available for researchers, the impact of using a rating curve and the impact of the assumptions made according to the rating curve can be investigated. Questions which need to be answered are: How often does the rating curve needs to be verified? How often and how changes the cross section in the Murray River?

According to the conversion factor K, the conclusion is that the ionic composition needs to be obtained to reduce the amount of used factors. The information given in the report 'Instream salinity models of NSW tributaries in the Murray-Darling Basin V1' can be used when chemistry, geology and rainfall are known for the study area. The report calculated the K factor per zone and gave corresponding information about the chemistry, geology and rainfall. It is important to keep in mind that the TDI (Total Dissolved Ions) of the catchment instead of the TDS (Total Dissolved Solids) are used to translate to EC. The difference is that the TDS also includes silica and excludes bicarbonate. The relationship between the EC and TDI is not linear and therefore it was decided to base the relationship on the percentiles of TDI/EC ratios for individual samples (DepartmentofWaterandEnergy, 2008).

Another important aspect concerns the unaccounted salt loads. MSM-BIGMOD uses this to balance the salt load which suggests there are unknown river processes. When comparing the unaccounted salt load from MSM-BIGMOD with the unaccounted salt load in this research every time a module is added to the model structure, the information about which river processes are captured by MSM-BIGMOD and which are not, will be specified.

9.3 Model Structure

More detailed information about the river processes is also necessary when improving the model structure from the flow model and when developing the salt model. When information is available where specific river processes occur, a simple example is the exact location of a tributary, this river process can be converted into an additional module for the model structure. When knowing the location of the 'module' , the model structure can be modified in such a way that the 'module' is used after the flow passes x kilometres in the Murray River.

The final purpose of the salt model is to give management advice about reducing the salinity in the Murray River. It is very important to take the uncertainties into account. For example, changing the conversion factor K with 10% leads to an 18% salt load difference. It is definitely unreliable to base management strategies on a model structure where the uncertainties in the input data are ignored.

10 References

Albritton, D., Schmeltekopf, A., & Zare, R. (1976). An Introduction of the least-squares fitting of spectroscopic data. *Molecular Spectroscopy: Modern Research II*, pp. 1-67.

Audit, N. L. (Red.). (2000). *Dryland Salinity in Australia.* Opgehaald van http://lwa.gov.au/files/products/national-land-and-water-resources-audit/pr010108/pr010108.pdf

AustralianWaterAssociation. (2000). *Australian guidelines for water quality monitoring and reporting.* Artarmon, NSW.

Authority, M.-D. B. (Red.). (2011, 02). *The Living Murray river restoration program*. Opgeroepen op 06 23, 2015, van www.mdba.gov.au: http://www.mdba.gov.au/annualreports/2011-12/chapter_02_2.html#reporttop

AuthorityMurrayDarlingBasin. (sd). *Measuring Salinity using Electrical Conductivity.* Opgehaald van http://riverdata.mdba.gov.au/includedfiles/electrical_conductivity.htm

Baldassarre, G. D., & Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences, 13*, pp. 913-921.

Bekesi, G., Telfer, A., Woods, J., Forward, P., Burnell, R., & Hatch, M. (2014). Quantitative measure of salt interception using in-river transient electromagnetic geophysics. *Australian Journal of Water Resources, Vol. 18*(No.1), 55-69. doi:10.7158/W12-030.2014.18.1.

Bennett, N. D., Croke, B., Guariso, G., Guillaume, J., Hamilton, S., Jakeman, A., . . . Andreassian. (2013). Characterising performance of environmental models. *Environmental Modelling Software*(40), 1-20.

Burnell, R., Bekesi, G., Telfer, A., Forward, P., & Porter, B. (2013). The development of a new methodology to interpret run of river salinity data to assess salt inflow to the Murray River. *Australian Journal of Water Resources, Vol. 17*(No. 1), 35-46. doi:10.7158/W12.2013.17.1

Chiew, F., & Siriwardena, L. (December 2005). Estimation Of SIMHYD Parameter Values For Application In Ungauged Catchments., (pp. 2904-2910). doi:ISBN: 0-9758400-2-9

Chiew, F., Vaze, J., N, V., Jordan, P., Perraud, J.-M., Zang, L., . . . Murphy, R. (2008). *Rainfall-runoff modelling across the Murray-Darling Basin. A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project.* Australia: CSIRO.

Chow, V., Maidment, D., & Mays, L. (1988). *Applied Hydrology.* New York: McGraw-Hill.

Croke, B. (2005). Land use impacts on hydrologic response in the Mae Chaem Catchment, Northern Thailand. *Proceedings of the 2005 International Converence on Simulation and Modeling.* Canberra: Australian National University.

Croke, B. (2009). Representing uncertainty in objective functions: extension to include the influence of serial correlation. *18th World IMACS/MODSIM Congress*, (pp. 3372-3378). Cairns.

Croke, B. (Unknown). *The Role of Uncertainty in Design of Objective Functions.* Integrated Catchment and Management Centre, Department of Mathemetics. Canberra: The Fenner school of Environment and Society.

Croke, B., & Littlewood, I. (2005). *Comparison Of Alternative Loss Modules In The IHACRES Model: An Application To 7 Catchments In Wales.*

Croke, B., & Shin, M.-J. (2015). *Modifications to a rainfall-streamflow model to handle 'non-stationarity'.* Copernicus Publications. doi:10.5194/piahs-92-1-2015

DepartmentofEnvironmentandConservation(NSW). (2004). *Use of effluent by irrigation.* Sydney: Department of ENvironment and Conservation (NSW).

Department-Of-Sustainability-Of-Environment. (2012, 11 29). *Flood Summaries by Basin*. Opgehaald van dseflood: http://203.12.195.133/dseflood/

DepartmentofWaterandEnergy. (2008). *Instream salinity models of NSW tributaries in the Murray-Darling Basin; Volume 1 - Border Rivers Salinity Integrated Quantity and Quality Model.* Sydney: NSW Government. Opgehaald van www.dwe.nsw.gov.au

Dooge, J. C. (1959, Februari). A General Theory of the Unit Hydrograph. *Journal of Geophysical Researche, Volume 64*(2), 241-256.

Environment-Climate-Change-&-Water-NSW. (2011). *Floodplain Management Plan - Edward and Wakool Rivers Stage 1 Deniliquin to Moama-Moulamein Railway.* Department of Environment, Climate Change & Water NSW.

Exell, R. (2001). *Error analysis*. Opgehaald van Practical Mathematics: http://www.jgsee.kmutt.ac.th/exell/PracMath/ErrorAn.htm

Fitzpatrick, A., Munday, T., Berens, V., Hatch, M., & Telfer, A. (2007). Mapping salt-loads of the Murray River, Australia, using airborne and in-river electromagnetic methods. *Unknown*.

Gavin, H. P. (2013). *The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems.* Duke: Department of Civil and Environmental Engineering.

Gippel, C. (2013). *Assessment of the hydraulics of the Little Murray River Weir Pool under alternative operating scenarios.* Shepparton: Fluvial Systems Pty Ltd. Stockton. Goulburn-Murray Water.

government, N. (Red.). (sd). *All about salinity*. Opgehaald van www.environment.nsw.gov.au: http://www.environment.nsw.gov.au/resources/salinity/allaboutsalinity.pdf

Green, D., Ali, A., Petrovic, J., Burrell, M., & Moss, P. (2012). *Water resources and management overwiew: Lower Darling River Catchment.* Sydney: NSW Department of Primary Industries.

Guillaume, J. H., Pierce, S. A., & Jakeman, A. (2010). *Managing uncertainty in determining sustainable aquifer yield.*

Guillaume, J. H., Qureshi, M., & J.Jakeman, A. (2012). *A structured analysis of uncertainty surrounding modeled impacts of groundwater-extraction rules.* Hydrogeology Journal. doi:10.1007/s10040-012-0864-0

Guillaume, J., Croke, B., Sawah, S. E., & Jakeman, A. (2011). *Implementing a framework for managing uncertainty holistically.* San Sebastian, Spain: Watermatex.

Hall, N., Oliver, M., Jakeman, T., Nicholson, A., & Watson, B. (2004). Land Use change for Salinity Management: A Participatory Model. *Biennial Meeting of the International Environmental Modelling and Software Society*.

Haugh, M. (2004). *The Monte Carlo Framework, Examples from Finance and Generating Correlated Random Variables.* Monte Carlo Simulation: IEOR E4703.

Herschy, R. (2009). *Streamflow Measurement.* Oxford: Taylor & Francis.

Hoekstra, A. (September 2010). *Water.* Enschede: University of Twente.

Ivkovic, K., Croke, B., & Kelly, R. (2014). Overcoming the challenges of using a rainfall-runoff model to estimate the impacts of groundwater extraction on low flows in an ephemeral stream. *Hydrology Research*, 58-72. doi:10.2166/nh.2013.204

J.C.Clarke. (sd). *Modelling uncertainty: A primer.* Department of Engineering Science. Oxford: Oxford University.

Jakeman, A. J., Post, D. A., & Beck, M. B. (1994). From data and theory to environmental model: the case of rainfall runoff. *Environmetrics, 1994*, 297-314.

Jakeman, A., Littlewood, I., & Whitehead, P. (1990). Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology, 117*, 275-300.

Jin, C., & Close, A. (2012). *Salt inflows of the lower Murray River.* Murray Darling Basin Authority.

Jin, X., Xu, C.-Y., Zhang, Q., & Singh, V. (2010). Parameter and modeling uncertainty simulated by GLUE and a formal bayesian method for conceptual hydrological model. *Journal of Hydrology*, 147-155.

Klemes, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*(31:1), 13-24. doi:10.1080/02626668609491024

L.K.Sherman. (1932). Stream flow from rainfall by the unit-graph method. *Eng. News Rec.*, 501-505.

Lichty, Dawdy, & Bergmann. (1968). Rainfall-runoff model for small basin flood hydrograph simulatin. *The use of analog and digital computers inhydrology*(81), 356-367.

Loveridge, M., Rahman, A., & Babister, M. (2013). Probabilistic flood hydrographs using Monte Carlo simulation: potential impact to flood inundation mapping. *20th International Congress on Modelling and Simulation*, (pp. 2660-2666). Adelaide.

Lyon, D. (2010). The discrete fourier tarnsform, part 6: Cross-Correlation. *Journal of Object Technology*(No. 2).

MDBA. (sd). *Geology and size*. (M. D. Authority, Redacteur) Opgehaald van www.mdba.gov.au: http://www.mdba.gov.au/about-basin/basin-environment/georgraphy/geology-and-size

Muleta, M. K. (sd). *Model Performance Sensitivity to Objective Function during Automated Calibrations.*

Murray-Darling-Basin-Authority. (2012). *Environmental Water Delivery - Murrumbidgee Valley.* Canberra: Commonwealth Environmental Water for the Australian Government.

Murray-Darling-Basin-Authority. (sd). *Lake Victoria*. Opgehaald van mdba.gov.au: http://www.mdba.gov.au/what-we-do/managing-rivers/river-murray-system/dams-weirs/lake-victoria

Nash, J. (1958). Systematic Determiniation of Unit Hydrgraph Parameters. *Journal of Geophysical Research*, 111-115.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models, i, a discussion of principles. *Journal of Hydrology*(10), 282-290.

NationalWeatherService. (sd). *Rating Tables and Curves.* Opgehaald van http://www.nws.noaa.gov/os/hod/SHManual/SHMan040_rating.htm

NCCMA. (2010). *Loddon River Environment Water Plan.* North Central Catchment Management Authority. Huntly, Victoria: Prepared for the Northern Victoria Irrigation Renewal Project. Opgehaald van http://www.g-mwater.com.au/downloads/connections/Environment/Loddon_River.pdf

NSW. (2013, October). *Types of salinity*. (N. S. government, Redacteur, & O. o. Heritage, Producent) Opgeroepen op 06 22, 2015, van environment.nsw.gov.au: http://www.environment.nsw.gov.au/salinity/basics/types.htm

NSWGovernment. (2000). *Salinity predictions for NSW rivers within the murray-Darling Basin.* Parramatta: Centre for Natural Resources.

nswgovernment. (sd). *book 1 Dryland Salinity: The Basics.* Opgehaald van www.environment.nsw.gov.au: http://www.environment.nsw.gov.au/resources/salinity/Book1DrylandSalinity.pdf

Owen, A. (2009-2013). *Ch-intro.* Stanford: Stanford University. Opgehaald van http://statweb.stanford.edu/~owen/mc/Ch-intro.pdf

Padilla, A., & Pulido-Bosch, A. (1994). *Study of hydrographs of karstic aquifers by menas of correlation and cross-spectral analysis.* Granada, Spain: Departamento de Geodinamica.

Ravalico, J., Dandy, G., & Maier, H. (2011). *MORE sensitivity Analysis of the MSM-BIGMOD Murray River Flow and Salinity Model.* The University of Adelaide, Adelaide.

Restoring the Balance in the Murray-Darling Basin. (sd). Opgehaald van www.environment.gov.au: http://www.environment.gov.au/water/rural-water/restoring-balance-murray-darling-basin

Sadoddin, A., Letcher, R., Jakeman, A., & Newham, L. (2005). A Bayesian decision network approach for assessing the ecological impacts of salinity management. *Mathematics and Computers in simulation*. doi:10.1016/j.matcom.2005.02.020

Salas, J., Delleur, J., Yevjevich, V., & Lane, W. (1980). *Applied Modeling of Hydrologic Time series.* Colorado: Highlands Ranch. Opgehaald van https://books.google.com.au/books?id=GinL-8Cc6QgC&pg=PA105&lpg=PA105&dq=autoregressive+model+hydrology&source=bl&ots=9UWcbj3U4L&sig=1dXn_l58Q5X76bKH8WseY6O2TXU&hl=nl&sa=X&ei=RgWKVbfvOuT6mQWM05OwDw&ved=0CCMQ6AEwAA#v=onepage&q=autoregressive%20model%20hydro

Scargle, J. (1989, August 15). Studies in astronomical time series analysis. III - Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly. *The Astrophysical Journal*, 874-887. Opgeroepen op 05 13, 2015, van articles.adsabs.harvard.edu: http://articles.adsabs.harvard.edu/full/1989ApJ...343..874S

Telfer, A., Burnell, R., & Alison, C. (2013). *Salt interception Schemes and instream processes.* Mildura: Mallee Catchment Management Authority.

Telfer, A., Burnell, R., & Charles, A. (2013, April). Salt Interception Schemes and instream processes. In *Salt Interception Schemes and instream processes.* Mildura: Mallee Catchment Managment Authority.

Telfer, A., Burnell, R., Charles, A., Dang, L., & Bekesi, G. (2013). *Murray River Salt Mobilisation: Characterisation of Selected Unaccounted Salt Loads.* Australian Water Environements. doi:AWE Ref: 13083

Telfer, A., Burnell, R., Woods, J., & Weir, Y. (2012). *Murray River floodplain salt mobilisation and salinity exceedances at Morgan.* Australian Water Environments. Murray Darling Basin Authority.

Telfer, A., Burnell, R., Woods, J., & Weir, Y. (2012). *Murray River floodplain salt mobilisation and salinity exceedances at Morgan.* Australian Water Environments. Murray Darling Basin Authority.

Tellinghuisen, J. (2001). *Statistical Error Propagation.* Nashville, Tennessee: Department of Chemistry.

Tillaart, S. P. (2010). *Influence of uncertainties in discharge determination on the parameter estimation and performance of a HBV model in Meuse sub basins.* Water Engineering & Management. Enschede: University of Twente.

Tomkins, K. M. (2014). Uncertainty in streamflow rating curves: methods, controls and consequences. *Hydrological Processes*, 464-481. doi:10.1002/hyp.9567

tue. (sd). *Template for parameter estimation with Matlab Optimization Toolbox; Including dynamic systems.* Opgehaald van bmi.bmt.tue.nl: http://bmi.bmt.tue.nl/sysbio/parameter_estimation/Template%20for%20parameter%20estimation%20with%20Matlab%20Optimization%20Toolbox.pdf

WMO. (2010b). *Manual on Stream Gauging Volume II - Computation of Discharge.* Geneva: World Meterological.

11 Appendix

Appendix I – Chapter 3: *The Unit Hydrograph*

The basis of the model structure, is the unit hydrograph. The theory of the unit hydrograph was first introduced by Sherman (1932) (L.K.Sherman, 1932). The original definition of the Unit Hydrograph of a catchment is defined as a direct runoff hydrograph resulting from a unit of excess rainfall generated uniformly over the drainage area at a constant rate for an effective duration (Chow et al., 1988). This definition was modified to include the slow flow component through the use of parallel flow pathways (e.g. IHACRES model, (Jakeman et al., 1990)). When the effective rainfall (u_k) for a catchment together with the Unit Hydrograph (h_{k-1}) is available, the runoff (Q_k) can be determined by using convolution. Convolution is similar to cross-correlation where in both situations one function is slid over another function (Lyon, 2010). The convolution to produce the runoff for discrete-time is as follows (Jakeman et al., 1990):

$$Q_k = h_0 u_k + h_1 u_{k-1} + h_2 u_{k-2} + \dots + h_{k-1} u_1 + \zeta_k \qquad (1)$$

The general convolution integral continuous-time is where the function $h(k)$ is well known as the instantaneous unit hydrograph:

$$Q(k) = \int_0^k h(k-s) u(s) ds \qquad (2)$$

Appendix II – Chapter 3: *Formulation of equation Unit Hydrograph*

Jakeman *et al.,* (1990) used the discrete convolution equation (1) and has written it as a transfer function model, where the u_k is the effective rainfall and ζ_k is an error compensation term assumed to represent the additive nature of all uncertainties arising from sampling, measurement and model errors.:

$$Q_k = h_0 u_k + h_1 u_{k-1} + h_2 u_{k-2} + \dots + h_{k-1} u_1 + \zeta_k \qquad (1)$$

$$Q_k = (h_0 + h_1 Z^{-1} + h_2 Z^{-2} + \dots + h_{k-1} Z^{-k+1}) u_k + +\zeta_k$$

$$Q_k = H(Z^{-1}) u_k + \zeta_k$$

$$Output = Unit\ Hydrograph * Input \qquad (3)$$

$$* = convolution$$

The transfer function is the Z transform of the impulse response which describes how the input response to the output (Jakeman, Post, & Beck, 1994). Since there are different timesteps in the equation (1), the Z^{-1}, the backward shift operator, transforms the equation to an equation without different timesteps $Z^{-1} u_k = u_{k-1}$ (Jakeman et al., 1990).

Jakeman *et al.,* (1990) have written eq (1) into an autoregressive formulation eq (4). The application of autoregressive models has been attractive mainly because the autoregressive form has an intuitive type of time dependence (the value of a variable at the present time depends on the values at previous time) and they are the simples models to use (Salas et al., 1980). It is an efficient formulation in terms of writing it down and in terms of decreasing the calculation time of the flow. This formulation also offers a power fool tool to estimate the parameters (Jakeman et al., 1990). Since in this research the effective rainfall will not be used as the input, the u_k is replaced for the flow input $I(\zeta)$.

$$Q(\zeta) = \left(\frac{\beta}{1+\alpha\zeta^{-1}}\right)^n * I(\zeta) \qquad (4)$$

Appendix III – Chapter 3: *Approach two for Monte Carlo*

Cholesky Decomposition

Another approach can be that the random noise is added to all the parameters at the same time. In that situation the covariances (correlation between parameters, non-diagonal values of matrix V) are needed to obtain the correlated random parameters. In this situation Cholesky Decomposition is used to obtain the standard deviation including the correlation of the parameters. The Cholesky decomposition obtains the matrix C which is the lower triangular of the matrix V (Haugh, 2004).

$$V = C^T C \qquad (5)$$

$$C = \begin{matrix} L_{11} & 0 & 0 \\ L_{12} & L_{22} & 0 \\ L_{13} & L_{23} & L_{33} \end{matrix} \qquad (6)$$

Appendix IV – Chapter 5: *Uncertainty conversion factor* K

To get a global idea about the influence of the different K value on the salt load, a raw sensitivity analysis is done. Two sites with big differences in flow and salinity are chosen; Barham and Lock 5. During the sensitivity analysis the K will be changed and the salinity and flow values have to stay constant. For the value of the K is chosen to use a value varying between 0.3 $mgL^{-1}/\ \mu Scm^{-1}$ and 0.9 $mgL^{-1}/\ \mu Scm^{-1}$.

Since the relationship is linear, varying the K with 10% will change the salt load with a constant percentage for every calculated salt load. However, the absolute change in salt load depends on the amount of flow and salinity. Since the flow has a bigger influence on the salt load in comparison with the salinity, the absolute change of the salt load has to be obtained by comparing high flow in combination with a low salinity, a high flow and a high salinity, a low flow and a high salinity and a low flow and low salinity. To decide which value is a 'high' flow or salinity and which is a 'low', the percentiles at the different sites are measured. It is important to take into account that the peak values are included in these percentiles, but can be much bigger than the value of the percentile. The results are shown in Figure 53 and Figure 34. The assumption is made there are in this case no uncertainties in the flow and salinity data. Another assumption is that the K value is constant. It is unknown if the K value changes if chemistry varies between for example high and low flows.

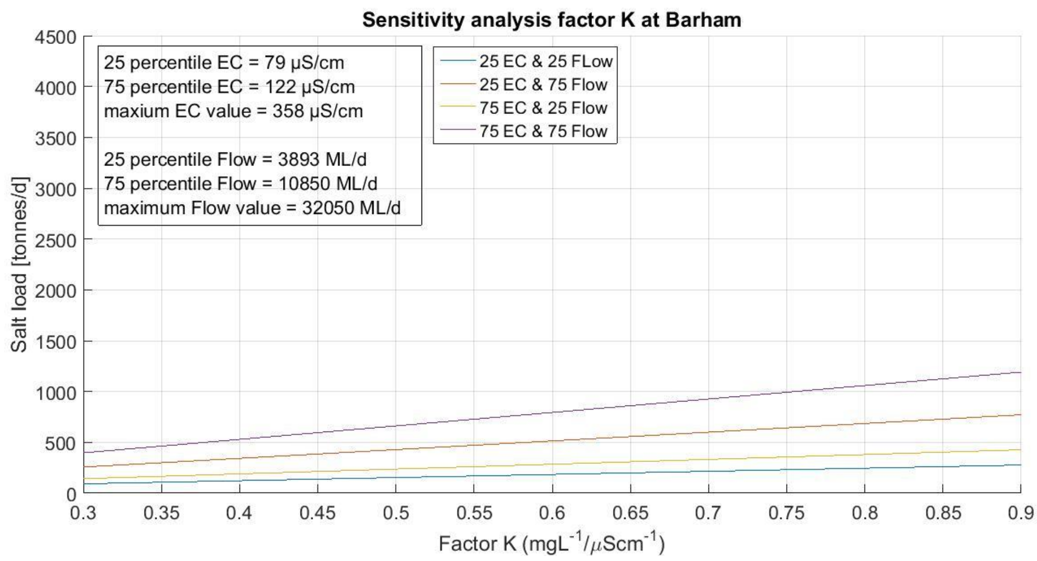

Figure 53 Sensitivity analysis factor K at Barham

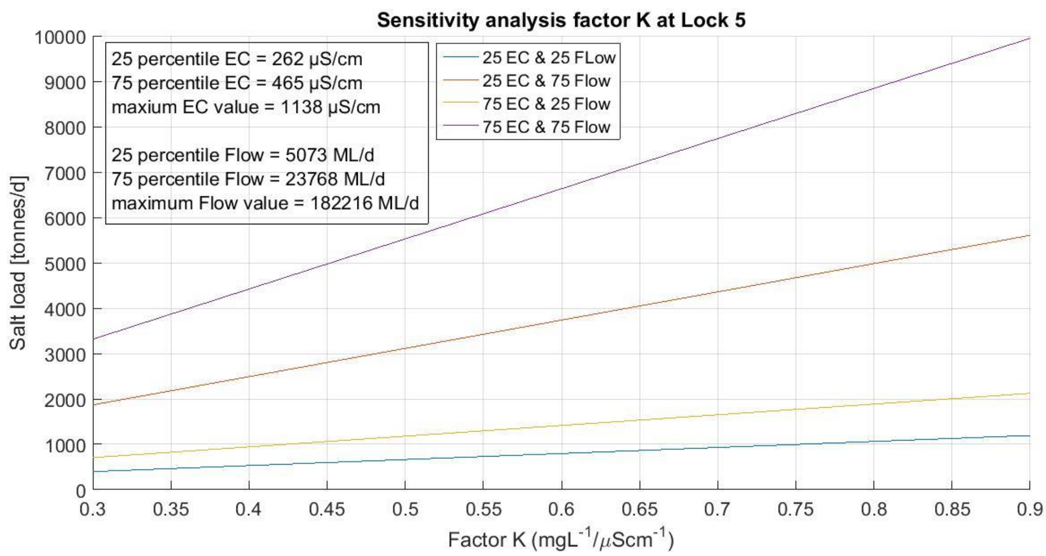

Figure 54 Sensitivity analysis factor K at Lock 5

At the different sites, changing the K factor with 0.1 causes a change of 118% for the salt load as shown in Table 18. Since the different reports recommend a K value between 0.55 and 0.65, the output of the salt load varies with a maximum of 18 % when changing the K between 0.55 and 0.65. During high flows, the absolute change in amount of salt load is bigger.

Table 18 Percentage change when changing K from 0.55 to 0.65

	Barham K=0.55	Barham K=0.65	Percentage change	Lock 5 K=0.55	Lock 5 K=0.65	% change
low EC & low flow	169.15	199.91	118%	731.02	863.93	118%
low EC & high flow	471.41	557.12	118%	3424.9	4047.6	118%
high EC & low flow	261.22	308.71	118%	1297.4	1533.3	118%
high EC & high flow	728	860.37	118%	6078.5	7183.7	118%

Appendix V – Chapter 5: U*naccounted salt load MSM-BIGMOD*

There is data available of the adjusted monthly averages of daily unaccounted salt inflow in MSM-BIGMOD. The data starts at 2002 till 2012 for a total of 35 sites. This data set will be referred to as 'unaccounted salt load MSM-BIGMOD'. The flow, EC and therefore salt load data used during this research will be referred to as 'salt load observed data'. The sites from MSM-BIGMOD are compared with the site names of the observed data. The information for the sites Coligna, Lock 9, Lock 7, Lock 6 and Lock 5 is available in both data sets. The MSM-BIGMOD data also includes the adjusted unaccounted salt loads of intermediate sites for example Coligna and Lock 9. All the adjusted unaccounted salt loads from MSM-BIGMOD between Coligna and Lock 9 are summed. Second, the daily observed salt load data is converted to monthly data between 2002 and 2012. After that step the difference between the observed salt load of Coligna and Lock 9 is calculated and compared with the adjusted unaccounted salt load from MSM-BIGMOD between Coligna and Lock 9. These steps are the same for the other sites.

Appendix VI – Chapter 6: Step 2: Calibration – changing reach

The first iterative optimization step gives the following parameter values for reach Wakool junction to Lock 5 with related graph which are shown in Figure 55. There is a big difference between the flow input (in to a reach) and the flow output (out of the reach). The modelled flow output does not cover the four peaks of the flow output. The four peaks of the flow output are not covered in the model structure since tributaries and anabranches are not taken into account. Also when changing the reach to Coligna to Lock 5, the four peaks are still not covered. Between Wakool junction and Lock 9 there are some tributaries and anabranches (section 4.2) which causes the difference between the flow output and the modelled flow. These results are suggesting a gain module need to be added to the model structure. The first focus of this research is on producing a simple model which can be used to configure where this model cannot be used. Adding a gain to the model is for a later stadium. This means the reach for developing the model will be shifted from Wakool junction to Lock 5, to Lock 9 to Lock 5.

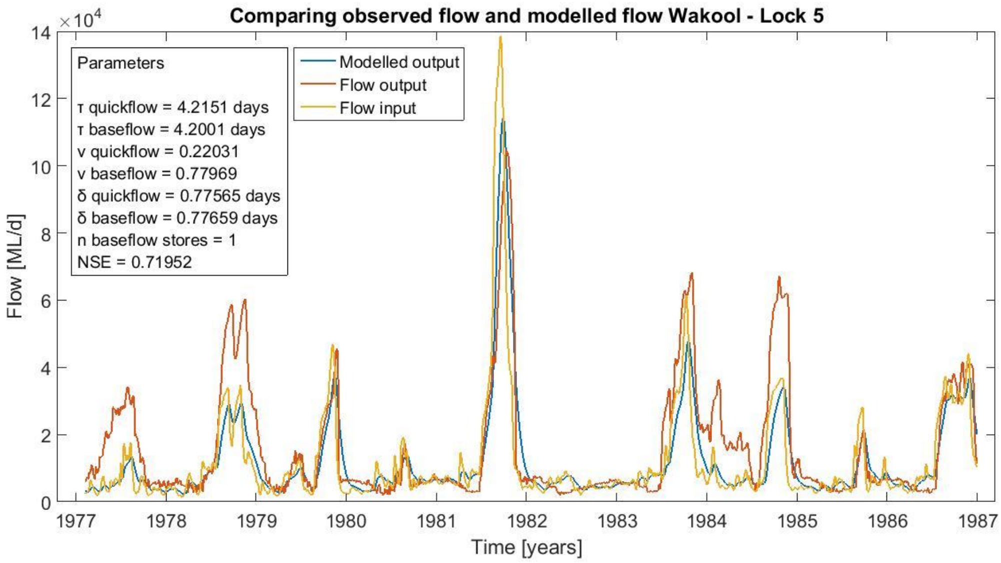

Figure 55 Modelled flow Wakool - Lock 5

Appendix VII – Ch. 6: Step 2: Calibration – changing parameters UH

In this step the Unit Hydrograph from this model structure will be plotted by using an impuls of value one as input. Changing the parameters of the model structure when plotting the UH, gives a good understanding of the effect of the changes.

Figure 56 shows the Unit Hydrograph with the parameter values when the amount of stores is 1 and the volume for the quick flow and slow flow is both set to 0.5. The parameters obtained by these requirements causes the NSE with value 0.96826. The top of this Unit Hydrographs (UH) shows there is a delay δ of about one day. Changing the amount of stores has no influence on the shape of the UH.

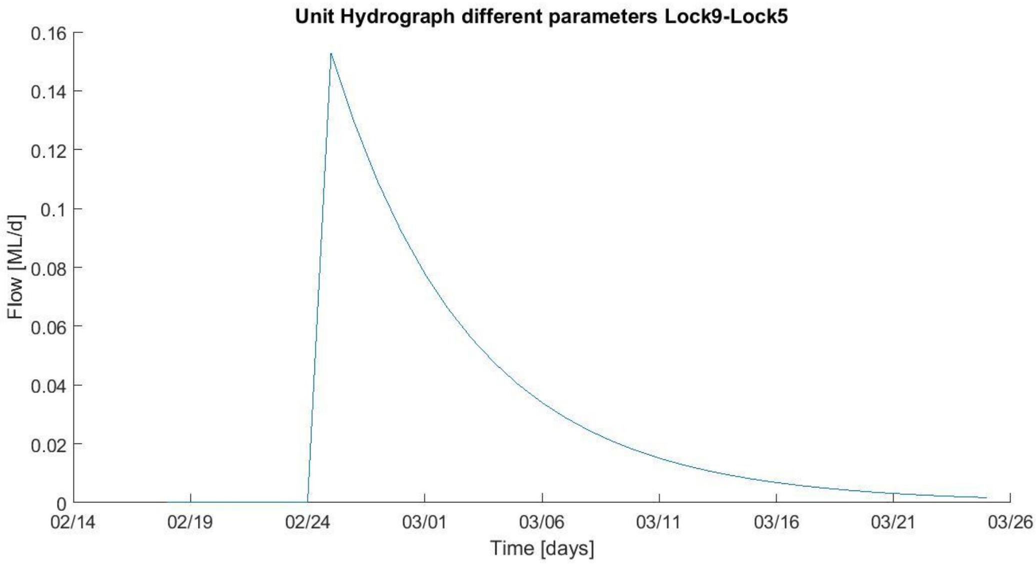

Figure 56 Impuls creating Unit Hydrograph

Table 19 Parameter values

Stores	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,1]	0.045334	0.055531	0.5	0.86932	0.82491

<u>Changing amount of stores</u>

When changing the value of the parameter tau for the slow flow to 1, changing the amount of stores has an influence as shown in Figure 57. The storage time for the slow flow in was to small that changing the stores had no influence. Figure 57 shows that the more stores for the slow flow, the lower is the peak for the quick flow and the smoother and higher is the peak for the slow flow. It causes the broadening of the UH what expected.

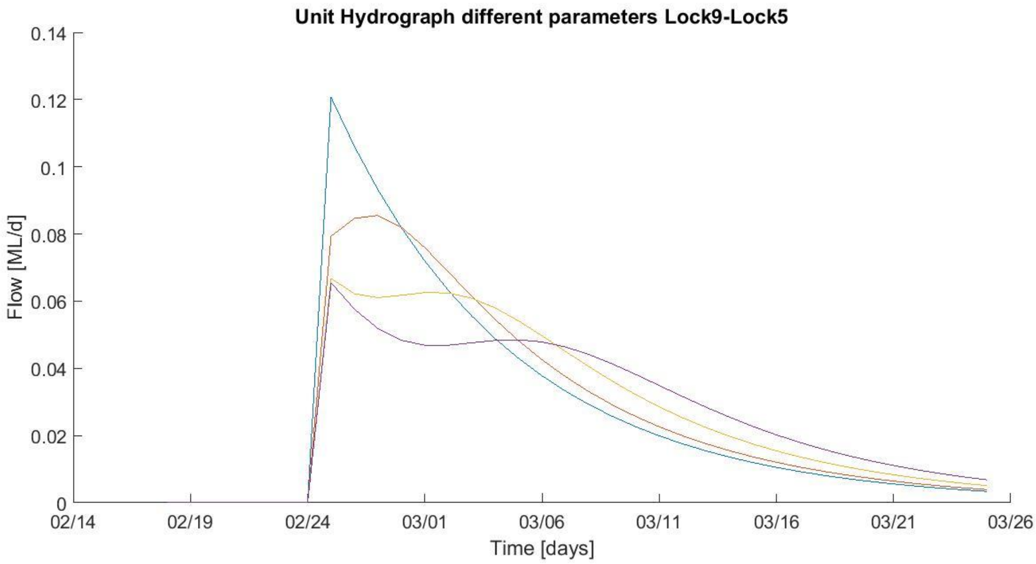

Figure 57 UH when changing amount of stores

Table 20 Parameter values and different stores

Stores	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,1] [1,5] [1,10] [1,15]	0.045334	1.0	0.5	0.86932	0.82491

<u>Changing storage time slow flow or quick flow</u>

When changing the time constant tau for the slow flow or quick flow while the amount of stores is 1, Figure 58 shows that the peak of the UH is decreasing when increasing the storage time for the slow flow or the quick flow. The quick flow gives the same shape of the UH.

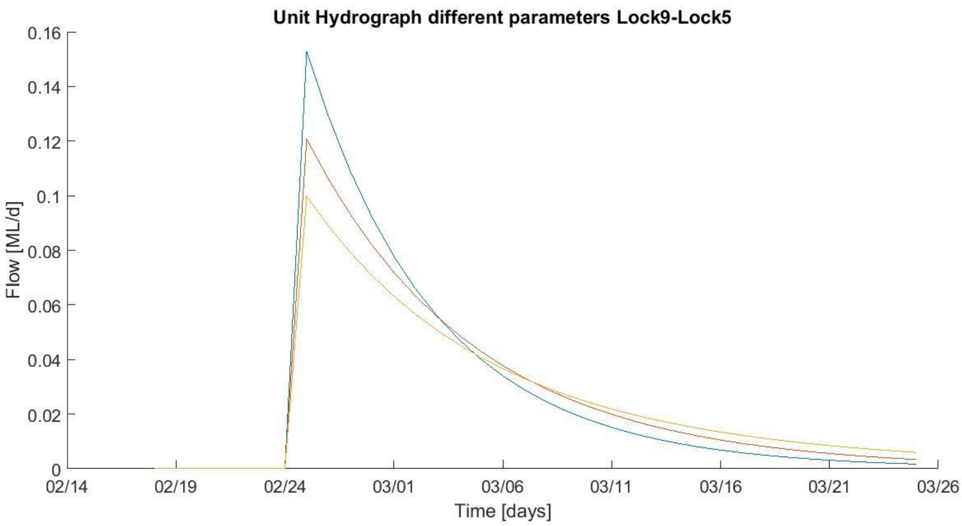

Figure 58 UH when changing time constant slow flow

Table 21 Parameter values when changing t slow flow

Stores	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,1]	0.045334	0.55531 1.0 2.0	0.5	0.86932	0.82491

Appendix VIII – Ch. 6: Step 3: Model Performance

When dividing the reach (Lock 9 to Lock 5) into smaller reaches, the location where the additional input occurs can be obtained. When looking to Figure 59, Figure 60, Figure 61 it is clear the additional slow flow input occurs between Lock 7 and Lock 6. Between Lock 9 and Lock 7 and between Lock 6 and Lock 5 no additional lower flows or flow extraction occurs, because when looking at the lower flows, the input flow is the same as the output flow.

Section 4.2 explains there are two creeks between Lock 7 – Lock 6, Punkah Creek and Salt Creek. Further, there are also two creeks between Lock 6 and Lock 5, Monoman Creek and Hundee Creek. The Chowilla creek connects the Punkah Creek and the Monoman Creek. At the Chowilla floodplain, there is a lot of groundwater which discharges to the river or water which recharge to aquifer from the river. These groundwater discharge explains the lower flows which occur between Lock 7 and Lock 6 (Telfer et al., 2013).

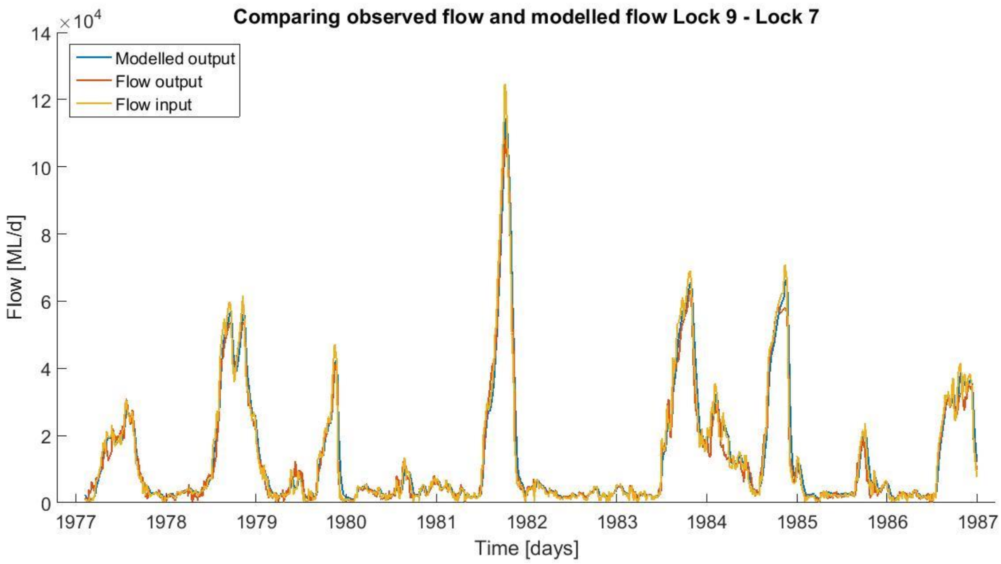

Figure 59 Modelled flow Lock 9 - Lock 7

Table 22 Parameter values

Stores	τ quick flow	τ slow flow	volume	δ quick flow	δ slow flow
[1,1]	0.061804	0.025844	0.96307	0.84202	0.998

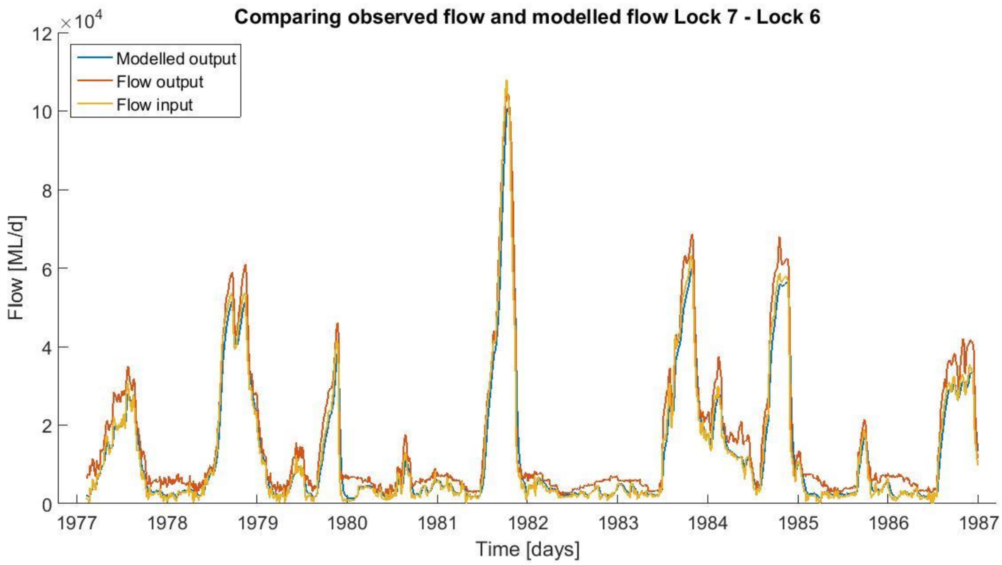

Figure 60 Modelled flow Lock 7 - Lock 6

Table 23 Parameter values

Stores	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,1]	0.061804	0.025844	0.96307	0.84202	0.998

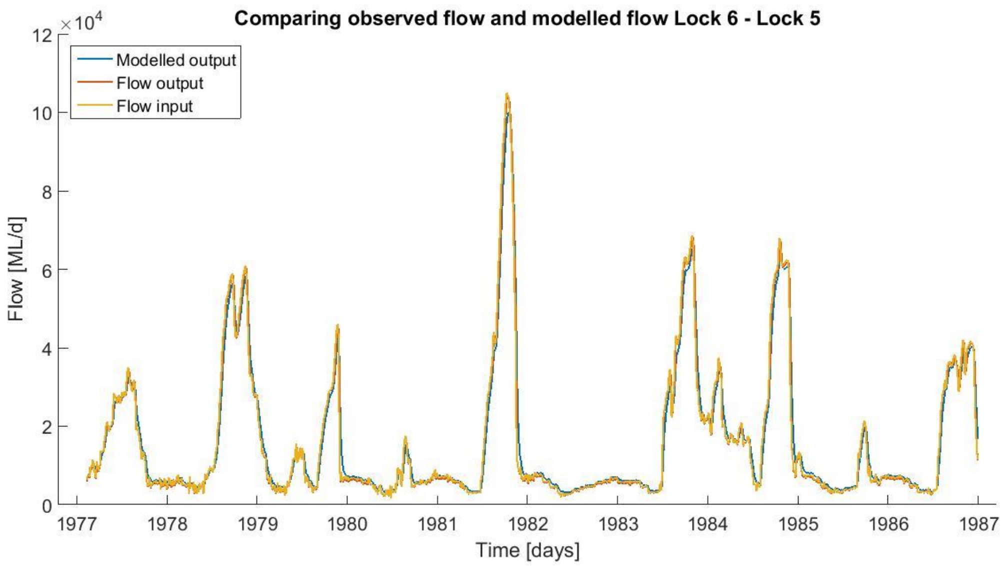

Figure 61 Modelled flow Lock 6 - Lock 5

Table 24 Parameter values

Stores	*τ quick flow*	*τ slow flow*	*volume*	*δ quick flow*	*δ slow flow*
[1,1]	0.061804	0.025844	0.96307	0.84202	0.998

Appendix VIII – Ch. 6: Step 1: Model Structure 2

Instead of two flow paths, the quick flow and slow flow, only one flow path remains which is shown in Figure 46.

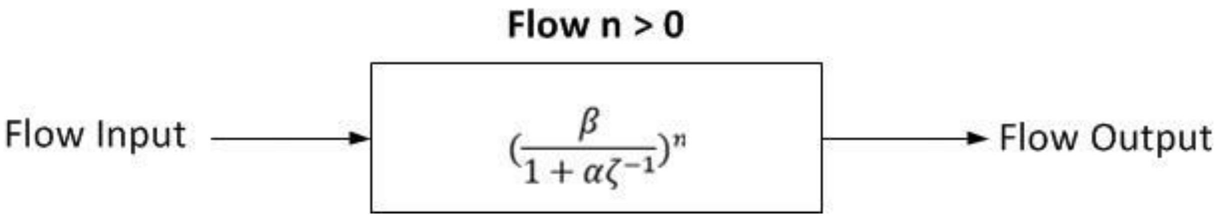

Figure 62 New model structure with one flow path

The accompanying equation is shown in 7. As described in section XX, the β can be described as the α parameter. The α parameter is referred to as the τ (section 3.3.2). There is only one τ parameter and one δ parameter since there is only one single flow path. Further, the ν is now become unnecessary since the total volume of flow will use this flow path.

$$Q(\zeta) = \frac{\beta}{1+\alpha\zeta^{-1}} * I(\zeta)$$

$$Q(\zeta) * (1+\alpha\zeta^{-1}) = \beta * I(\zeta)$$

$$Q(\zeta) + \alpha\zeta^{-1} * Q(\zeta) = \beta * I(\zeta)$$

$$Q_k = -\alpha * Q_{k-1} + \beta * I_{k-\delta} \qquad (7)$$

Appendix IX – Ch. 6: Step 5: Validation

Figure 63 and Figure 64 show the graphs for the modelled output for the validation during the second validation period and the third.

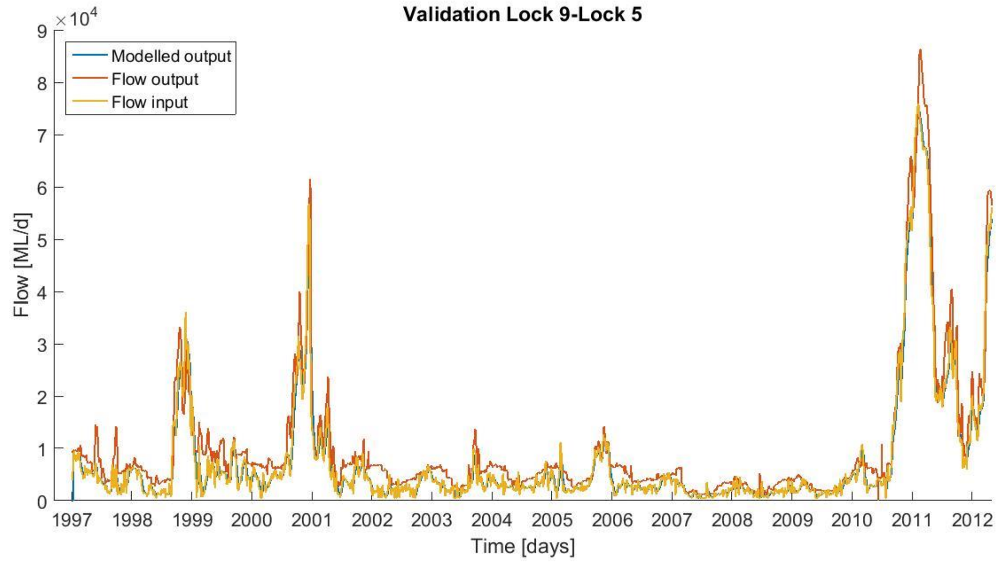

Figure 63 Validation Period 2 Januari 1997 - April 2012

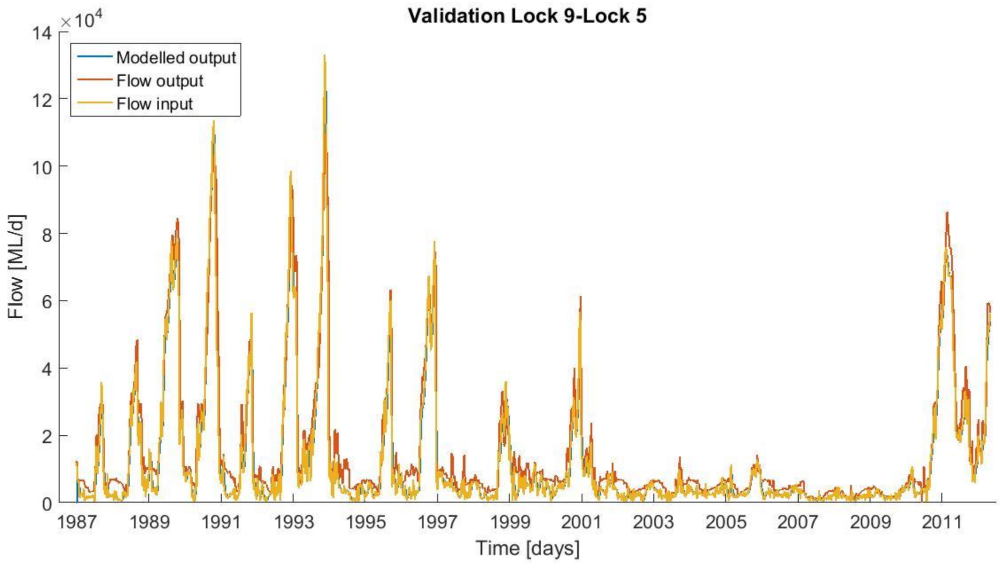

Figure 64 Validation Period 7 Januari 1987 - April 2012