Optimizing Hydrologic Model Selection for Low-Flows

By:
John “Jack” Tarricone
Dept. of Geography Honors Thesis
University of Colorado Boulder

Defended April 4, 2016

Thesis Advisor
Dr. John Pitlick, Dept. of Geography

Committee Members
Dr. Holly Barnard, Dept. of Geography
Dr. Michael Gooseff, Dept. of Civil, Environmental, and Architectural Engineering
Abstract

Conceptual rainfall runoff models (CRRM) are used to predict the flow characteristics of waterways by correctly classifying the hydrologic exchanges occurring in the natural environment. Their practical applications range from predicting catchment yields to filling in gaps in stream flow data. Many of these models were developed for predicting and quantifying mean and peak flows and have shown difficulty in replicating the low-flow hydrology of a waterway.

We collected stream flow, rainfall, and evapotranspiration data from 16 gauges across Australia over a 30 year period. The gauges displayed four different types of flow characteristics. We then tested the ability of three commonly used CRRMs: SimHyd, Sacramento, and Australian Water Balance Model (AWBM) to predict the low-flow hydrology of the 16 different catchments.

We found that the predictive performance of AWBM was consistently better than the two other models. While SimHyd fit well to many flow duration curves, it struggled to model low-flows correctly. The information gathered from this study suggests that AWBM should be used when trying to predict low-flow hydrology. While AWBM was the best model, many improvements can still be made in consistently replicating the low-flow hydrology of intermittent systems.

Keywords: Hydrology, Hydrologic Modeling, Model Optimization
# TABLE OF CONTENTS

Abstract........................................................................................................................................... ii

Table of Contents.............................................................................................................................. iii-iv

List of Figures and Tables................................................................................................................ v

List of Abbreviations and Definitions........................................................................................... vi

1. Introduction/Background............................................................................................................. 1
   1.1 Hydrology............................................................................................................................. 1
   1.2 Hydrologic models................................................................................................................ 1
   1.3 Justification of Research...................................................................................................... 2
   1.4 Study Goals........................................................................................................................ 3

2. Literature Review....................................................................................................................... 3
   2.1 Low-Flow Hydrology.......................................................................................................... 3
      2.1.1 Modeling Gaps............................................................................................................... 4
      2.1.2 Importance to Water Management.............................................................................. 5
   2.2 RRL eWater RRL............................................................................................................... 6
      2.2.1 AWBM......................................................................................................................... 7
      2.2.2 SimHyd......................................................................................................................... 8
      2.2.3 Sacramento.................................................................................................................. 10
   2.3 Optimizers and Objective Functions.................................................................................. 11
      2.3.1 Global Optimizer: Shuffle Complex Evolution......................................................... 12
      2.3.2 Objective Functions: Nash-Sutcliffe and Flow Duration Curve............................. 12

3. Methods..................................................................................................................................... 12
   3.1 Gauge Selection by Hydrologic Classification.................................................................. 13
   3.2 Data Acquisition................................................................................................................ 15
      3.2.1 Daily Stream Flow Time Series.................................................................................. 15
      3.2.2 Average Daily Potential Evapotranspiration............................................................. 15
   3.3 Selecting Models and Optimizers..................................................................................... 17
   3.4 Testing Model/Optimizer Combinations......................................................................... 17
   3.5 Processing Modeled Data................................................................................................. 18

4. Results...................................................................................................................................... 15
   4.1 Normalized Class Flow Duration Curves.......................................................................... 19
   4.2 Observed vs. Modeled Flow Duration.............................................................................. 21
   4.3 Log root-mean-square error (RMSE) Tables................................................................... 23
   4.4 Cease to Flow Statistics..................................................................................................... 25

5. Discussion................................................................................................................................. 30
   5.1 Results Overview................................................................................................................. 30
   5.2 Model Structure Deficiencies......................................................................................... 32
   5.3 Practical Implications........................................................................................................ 32

6. Conclusion................................................................................................................................. 33
7. References

8. Appendices
List of Figures and Tables

Table 1 – Original gauge information data generated in R………………………….12
Table 2 – Log RMSE values for Class One……………………………………………….23
Table 3 – Log RMSE values for Class Four……………………………………………….23
Table 4 – Log RMSE values for Class Six……………………………………………….24
Table 5 – Log RMSE values for Class Eight…………………………………………….24

Figure 1 - Structure of the AWBM CRRM……………………………………………….6
Figure 2 – Structure of the SimHyd CRRM……………………………………………….7
Figure 3 – Structure of the Sacramento CRRM…………………………………………8
Figure 4 – Map of Australia with location of flow class of all 16 gauges………………12
Figure 5 - Screenshot of the Swan River Catchment drawn in ArcGIS…………………14
Figure 6 - Flow Duration Curve for the Class One……………………………………17
Figure 7 - Flow Duration Curve for Class Four……………………………………17
Figure 8 - Flow Duration Curve for Class Six………………………………………..18
Figure 9 – Flow Duration Curve for Class Eight………………………………………..18
Figure 10 – Flow Duration Curve for the Flowerdale River…………………………20
Figure 11 - Flow Duration Curve for the Tia River……………………………………..20
Figure 12 - Flow Duration Curve for the Tone River ……………………………..21
Figure 13 - Flow Duration Curve for the Mandagery Creek …………………………21
Figure 14 - Graph of Observed vs. Modeled Percent Time CTF for Class One……25
Figure 15 - Graph of the average CTF Spell Length for Class One……………………25
Figure 16 - Graph of Observed vs. Modeled Percent Time CTF for Class Four……26
Figure 17 - Graph of the average CTF Spell Length for Class Four …………………26
Figure 18 - Graph of Observed vs. Modeled Percent Time CTF for Class Six………27
Figure 19 - Graph of the average CTF Spell Length for Class Six ……………………27
Figure 20 - Graph of Observed vs. Modeled Percent Time CTF for Class Eight…..28
Figure 21 - Graph of the average CTF Spell Length for Class Eight …………………28
Figure 22 – Graph of Tone observed, AWBM_FDC, SIM_FDC, SAC_FD………..30
List of Abbreviations and Definitions

PET: Potential Evapotranspiration
ArcGIS: Commonly used Geographic Informational Science mapping program
R: Open source statistical analysis and data processing software
ML: Megalitre
ML/day: Megalitres per day
mm/day: millimeters per day
RRL: The Rainfall Runoff Library Modeling Software
AHGF: Australian Hydrological Geospatial Fabric
SCE: Shuffle Complex Evolution Global Optimizer
NSE: Nash-Sutcliffe Efficiency
FDC: Flow Duration Curve Objective Function
CTF: Cease to Flow
AWBM: Australian Water Balance Model
AWBM_FDC: AWBM model with the Flow Duration Curve objective function
AWBM_NSE: AWBM model with Nash-Sutcliffe Efficiency objective function
SIM_FDC: SimHyd model with the Flow Duration Curve objective function
SIM_NSE: SimHyd model with the Nash-Sutcliffe Efficiency objective function
SAC_FDC: Sacramento model with the Flow Duration Curve objective function
SAC_NSE: Sacramento model with the Nash-Sutcliffe Efficiency objective function
BFI: Base Flow Index
RMSE: Root-mean-square error
1. Introduction

1.4 Study Goals

The goal of this study is to identify which of three commonly used conceptual rainfall runoff models (CRRM): Australian Water Balance Model, SimHyd, and Sacramento, best replicate low-flow and cease-to-flow (CTF) events in a range of different flow regimes. The knowledge obtained will help to improve water management strategies in times of drought and water stress.

2. Literature Review

2.1 Low-Flow Hydrology

Low-flow hydrology is a sub-discipline in hydrology that focuses on the seasonal phenomena in which unregulated rivers and streams naturally have a lower flow. These low-flow events usually occur in areas with a pronounced dry season, which is a prolonged period with very little precipitation. During low-flow periods, runoff is usually derived from groundwater infiltrating onto the surface and surface water discharge from lakes, marshes, and even glacial melt in high altitudes (Smakhtin, 2011). Low-flow hydrology also focuses on periods in which a river stops flowing, also known as a cease-to-flow (CTF) spells. These CTF events put stress on the natural ecosystem and the human population surrounding the waterway. As the climate becomes warmer and more variable, low-flow hydrology has received increasing interest from hydrologists and water management groups. As the hydrologic cycle changes due to the effects of climatic changes, many areas that were once water secure are seeing less available water. This is due to changes in rainfall timing, less predictable precipitation amounts, increased evapotranspiration, and less snow fall in the winter (Murphy and Timbal, 2008). Quantifying water yields and
documenting low-flow event behavior has now become crucial for properly managing water resources.

2.1.1 Modeling Gaps

There is a fundamental difference in modeling mean and peak flows and modeling low-flows because of a difference in primary runoff drivers. Mean and peak flow indicators are much more homogenous throughout the catchment because they rely more on rainfall from which runoff can more easily be predicted. In contrast, low-flows are influenced primarily by subsurface flow or groundwater infiltrating into surface water systems (Barma and Lowe, 2012). Due to the high spatial variability of soil moisture content and groundwater throughout a catchment, these inputs are much more difficult to quantify for hydrologic models.

Many models calibrate soil moisture content by using a small time series of data in the beginning of the calibration period. The model will then produce its best estimate of soil moisture and base the rest of its discharge values, especially the low-flows off these estimated numbers. The ground water amounts in intermittent systems are naturally much more variable than perennial systems, which again is challenging for the models. This has inherently more room for error than modeling for rivers with perennial flow and steady base flow indexes that have relatively constant groundwater inputs. This modeling knowledge gap is exemplified by CTF events. These events occur in prolonged dry periods where soil moisture decreases by constant reduction from evapotranspiration. Even when a rainfall event does happen discharge is not always created because of the low moisture content of the soil, and CRRMs struggle to correctly replicate this phenomena.

More issues arise when modeling ungauged catchments (i.e. areas for which no stream flow data is available). A method hydrologist’s use to model ungauged and
intermittent streams is called transposition. This involves the use of characteristics acquired from a nearby gauged stream and applying them to the ungauged stream. Characteristics taken into consideration for similarities are physical geography, soil characteristics, vegetation, and climate. Transposition is currently based on factors related to mean flow rather than low-flow characteristics. There are currently no methods available that enable easy and accurate assessment of the hydrological similarity between waterways with respect to low-flows. These factors result in unreliable predictions when modeling low-flow streams, especially when trying to transpose models onto ungauged catchments (Barma and Lowe, 2012).

There are also anthropogenic influences that create challenges when modeling low-flows. Groundwater extraction, farm dams, lands use changes, and wastewater discharges all impact on low-flow streams. Quantifying these impacts is typically difficult. It would mean independently defining the impact they have and incorporating that impact into a CRRM (Barma and Lowe, 2012, p25-28).

### 2.1.2 Importance to Water Management

Understanding low-flows is crucial to managing water resources especially in more arid climates. Presently dams and reservoirs are used to level out the fluctuations in flow regimes of rivers. They allow for a more balanced supply of water year round, which is suitable for societal needs. Therefore knowledge of flow during low-flow periods is vital for determining water allocations and designing reservoirs (Gustard, 2008).

In Australia where droughts are a recurring problem, many water management strategies have been designed to cope. These strategies also rely on wet years to recharge the reservoirs, so issues arise in extended periods of drought. Southeast Australia experienced a 10-year drought from 1996 – 2006, which caused serious
water stress on the agricultural sector and therefore the economy. Looking at large scale climatic trends there is little doubt that Southeast Australia is being coming warmer and rainfall amounts are diminishing, meaning that these drought events are only going to become more common place (Murphy & Timbal, 2008). Being able to better model waterways during these drought periods becomes more challenging as low-flows and CTF events become more prominent. This why improving modeling abilities of low-flow events will especially help in periods of drought and water stress. Understanding how waterways will react in times drought before they happen allow for better water management strategies to be specially tailored to these time periods.

2.2 eWater Toolkit Rainfall Runoff Library

The eWater Toolkit is a package of modeling software designed by eWater, an Australian non-profit organization focusing on hydrologic modeling, water management issues, and overall ecological health (http://ewater.com.au/). The Rainfall Runoff Library (RRL) is a widely used rainfall-runoff modeling program. It was designed to simulate runoff in a catchment by using daily rainfall and potential evapotranspiration (PET) values. The models can be applied to catchments ranging in size from 10 km² to 10,000 km² (http://ewater.com.au/). The models in the RRL will typically be used to fill in gaps of stream flow records when there is an error in data acquisition, estimate discharge in similar ungauged catchments, and creating an extended stream flow time series to further predict stream yields (Podger, 2004). The RRL comes equipped with five different models, three of which are used in this study: Australian Water Balance Model, Sacramento, and SimHyd. The models range in complexity from the simplistic structure of AWBM to the complex structure of Sacramento, which has 19 parameters to optimize (Podger, 2004).
2.2.1 Australian Water Balance Model

The Australian Water Balance Model is a catchment water balance model that can relate rainfall runoff to rainfall on a daily time step. The model uses three surfaces stores which each simulate different areas of partial runoff through the catchment (Podger, 2004). The stores represent the different types of surfaces, infiltration rates, and saturation capacities that different materials have within a catchment. Podger then details how the model begins by calculating a moisture balance independently for each of the three stores. On a daily time step rainfall is added to each of the three stores and PET is subtracted. If the PET exceeds the rainfall amount the value will bottom out at zero because you can’t have negative moisture amounts. If the rainfall amount exceeds to PET amount in a store the excess will become runoff and the store will remain full.

Every time a runoff event occurs from any of the stores, part of that runoff becomes base flow recharge if the stream has a consistent base flow. The fraction of runoff used to recharge is calculated by multiplying the base flow index (BFI), which is the ratio of base flow to total flow in a stream, by the amount of runoff in the event. The base flow store is depleted at the constant rate of the base flow recession constant. All runoff that doesn’t enter base flow recharge becomes surface runoff or stream flow (2004, p. 47).
2.2.2 SimHyd

Podger states SimHyd is a CRRM that estimates daily stream flow using daily rainfall and areal PET as inputs. SimHyd produces runoff from four sources: impervious runoff, infiltration excess runoff, saturation excess runoff, and base flow.

The model starts with a rainfall event that fills up its interception store with rainfall, or hits an impervious threshold where it immediately becomes impervious runoff. Some this initial rainfall is lost to PET and water that entered the inception store becomes through flow. This runoff then goes through an infiltration function determining whether the rainfall event rate is greater than the infiltration rate. If the rainfall rate is greater excess runoff will commence, while the rest of the infiltration enters the next store of the model.
Figure 2 – This image shows the structure of the SimHyd CRRM (Podger, 2004)

The water is then subjected to a soil moisture function that either diverts water to the stream, to the groundwater store, or the soil moisture store. Water diverted to the stream in this store, or interflow, is estimated by using a linear function of soil moisture level divided by soil moisture capacity. The greater the initial soil moisture level, the more saturation excess runoff it will produce. The water that enters the soil is then subjected to groundwater recharge function, which estimates groundwater infiltration rate from initial soil moisture much like the previous two functions. The
infiltration that enters the groundwater store will contribute to base flow at a linear rate (2004, p. 54-55).

2.2.3 Sacramento
The Sacramento CRRM uses soil moisture characteristics to estimate the water balance and therefore the runoff within a catchment. Soil moisture increases during rainfall events and decreases by PET and natural flow out of the catchment. The moisture levels of the different stores in the soil determine rainfall infiltration and PET rates. The model defines five soil stores: Upper zone tension water (UZTW), Upper zone free water (UZFW), Lower zone tension water (LZTW), Lower zone primary free water (LSFWP), and Lower zone supplementary free water (LZFWS) (2004, p. 47). The two tension stores represent the water that is stored in the soil by its capillary tension, which can only be removed from the soil by evapotranspiration. In the three other stores water can move vertically between them, or laterally out of a store. If water leaves laterally through the upper zone it’s referred to as interflow, and an exit through the lower zone is referred to as base flow. Surface runoff is generated when the UZTWS is filled to capacity and the rainfall rate is greater than the infiltration rate.
Sacramento requires 16 parameters to be calibrated to simulate a water balance. 5 define the size of each the soil moisture stores, which is used to calculate moisture levels and therefore infiltration rates of each store. 3 calculate the rate of lateral outflow for the three different types of runoff, which helps determines the reduction rates of the stores. 3 others calculate the rate percolation, or drainage of water, from the upper stores to the lower stores. The last 3 calculate losses to the system from ground water recharge, and evapotranspiration.

2.3 Optimizers and Objective Functions
Optimization is a diverse and complex branch of applied mathematics that deals with minimizing or maximizing certain objective function within the general framework of a model (Neumaier, 2004). The RRL has many choices of optimizers
and objective functions for the CRRM models, three of which will be detailed as they were chosen to be used in the study

2.3.1 Global Optimizer: Shuffle Complex Evolution
Shuffle complex evolution (SCE) was a global optimization method developed by (Duan et al., 1993) promising to provide a robust, effective, and efficient way for function optimization. The SCE method is based on the combination of four concepts that were successful in applied optimization: combination probabilistic and deterministic approaches, clustering, systematic evolution of points, and competitive evolution. This global optimizer was chosen because of consistency of high quality results and common use throughout the hydrologic modeling community.

2.3.2 Objective Functions: Nash-Sutcliffe and Flow Duration Curve
Each global optimizer requires an objective function to be optimized. Nash-Sutcliffe is efficiency coefficient (NSE) is used to assess how well a hydrologic model can predict discharge. It produces a coefficient value, the closer to one that value is the better it does at predicting runoff (Gupta et al., 2009). A flow duration curve is a plot that shows a percentage of time a stream is likely to exceed a certain discharge value. The flow duration curve (FDC) function tries to align the modeled stream flows curve as closely as it can to the observed values.

3. Methods

3.1 Gauge Selection by Hydrologic Classifications
Flow regimes were classified by Kennard et al. (2010) across Australia, with the aim of supporting environmental flow management. Kennard et al. (2010) identified twelve natural flow regime classes, reflecting the varying levels of
perenniality, seasonality, and variability. The data base used to develop the classification made use of daily stream flow data from 830 gauges, which all contained at least 15 years of record. From the twelve classes we selected four to be included in this study. They are Class One: Stable base flow, Class Four: Unpredictable base flow, Class Six: Predictable winter intermittent, and Class Eight: Unreliable winter intermittent. These classes show a range of flow characteristics to test the ability of the CRRMs across a diverse set of circumstances. All streams located in Class One are perennial or flow for the entire year. The streams in classes Four, Six, and Eight all show difference levels of intermittency, with Six and Eight being winter-dominated systems, with predictable and unpredictable flow respectively.

Using R and the dplyr package we randomly selected four gauges from each of four flow classes specified above. We further specified that the minimum catchment area for selection would be 100 square kilometers to allow for a slight lag in the time series of stream flow data and PET and precipitation data. The data generated for each randomly selected gauge included the latitude, longitude, flow class, and gauge number. The data table was saved as a .csv (comma separated value) to be imported into ArcGIS for further data collection.
Table 1 – This table contains the original data generated for each of the 16 gauges.

<table>
<thead>
<tr>
<th>Flow Class</th>
<th>Gauge Location</th>
<th>Catchment Area (km²)</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Start Date</th>
<th>End Date</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class One: Stable base flow</td>
<td>Tanjil River</td>
<td>299</td>
<td>-37.981</td>
<td>146.194</td>
<td>1/1/66</td>
<td>31/12/1994</td>
<td>VIC</td>
</tr>
<tr>
<td></td>
<td>Latrobe River</td>
<td>558</td>
<td>-38.089</td>
<td>146.16</td>
<td>1/1/71</td>
<td>31/12/1999</td>
<td>VIC</td>
</tr>
<tr>
<td></td>
<td>Flowerdale River</td>
<td>162</td>
<td>-40.969</td>
<td>145.609</td>
<td>1/1/77</td>
<td>31/12/1995</td>
<td>TAS</td>
</tr>
<tr>
<td></td>
<td>Murrindindi River</td>
<td>108</td>
<td>-37.414</td>
<td>145.564</td>
<td>1/1/71</td>
<td>31/12/1999</td>
<td>VIC</td>
</tr>
<tr>
<td>Class Four: Unpredictable base flow</td>
<td>Christmas Creek</td>
<td>160</td>
<td>-28.171</td>
<td>152.984</td>
<td>1/1/68</td>
<td>31/12/1987</td>
<td>QLD</td>
</tr>
<tr>
<td></td>
<td>Burra Creek</td>
<td>533</td>
<td>-33.839</td>
<td>139.079</td>
<td>1/1/74</td>
<td>31/12/1992</td>
<td>SA</td>
</tr>
<tr>
<td></td>
<td>Oxley Creek</td>
<td>215</td>
<td>-28.352</td>
<td>153.295</td>
<td>1/1/71</td>
<td>31/12/2000</td>
<td>NSW</td>
</tr>
<tr>
<td></td>
<td>Tia River</td>
<td>899</td>
<td>-31.188</td>
<td>151.829</td>
<td>1/1/79</td>
<td>31/12/1995</td>
<td>NSW</td>
</tr>
<tr>
<td>Class Six: Predictable winter intermittent</td>
<td>Swan River</td>
<td>7614</td>
<td>-31.754</td>
<td>116.066</td>
<td>1/1/85</td>
<td>31/12/2000</td>
<td>WA</td>
</tr>
<tr>
<td></td>
<td>Rocky Creek</td>
<td>197</td>
<td>-35.955</td>
<td>136.697</td>
<td>1/1/74</td>
<td>31/12/2000</td>
<td>SA</td>
</tr>
<tr>
<td></td>
<td>Denmark River</td>
<td>534</td>
<td>-34.867</td>
<td>117.315</td>
<td>1/1/71</td>
<td>31/12/2000</td>
<td>WA</td>
</tr>
<tr>
<td></td>
<td>Tone River</td>
<td>982</td>
<td>-34.25</td>
<td>116.679</td>
<td>1/1/84</td>
<td>31/12/2000</td>
<td>WA</td>
</tr>
<tr>
<td>Class Eight: Unpredictable winter intermittent</td>
<td>Cockburn River</td>
<td>265</td>
<td>-31.064</td>
<td>151.126</td>
<td>1/1/71</td>
<td>31/12/2000</td>
<td>NSW</td>
</tr>
<tr>
<td></td>
<td>Copes Creek</td>
<td>250</td>
<td>-29.916</td>
<td>151.114</td>
<td>1/1/71</td>
<td>31/12/2000</td>
<td>NSW</td>
</tr>
<tr>
<td></td>
<td>Tuena Creek</td>
<td>325</td>
<td>-34.019</td>
<td>149.334</td>
<td>1/1/73</td>
<td>31/12/1995</td>
<td>NSW</td>
</tr>
<tr>
<td></td>
<td>Mandagery Creek</td>
<td>1697</td>
<td>-33.376</td>
<td>148.444</td>
<td>1/1/67</td>
<td>31/12/1981</td>
<td>NSW</td>
</tr>
</tbody>
</table>

Figure 4 – This map shows the location and flow class for each of the 16 gauges.
3.2 Data Acquisition

The RRL requires input data in a daily time series for average areal PET, average areal rainfall, and stream flow to start the modeling process for a catchment. Acquiring these required the use of ArcGIS in combination with R. The data for PET and rainfall were sourced from an Australia wide five by five kilometer gridded daily time series that dates back to the start of the 20th century (http://www.bom.gov.au/climate/data/).

3.3.1 Daily Stream Flow Time Series

The data for the daily stream flow time series was the same data Kennard et al. (2010) used for their hydrologic classification research. This dataset consisted of a date and a flow amount for that day reported in megalitres a day (ML/day). The data in each time series ranges between 1/1/1966 to 31/12/2000. The original files were changed from .csv format to .cdt (comma delineated time series), so they could be imported into the RRL.

3.3.2 Average Daily Potential Evapotranspiration/Rainfall

Using ArcGIS we imported the previously displayed .csv file of the latitude and longitude coordinates for each of the 16 gauges. These data points were positioned on top of the Australian Hydrological Geospatial Fabric (AHGF). This is a geospatial data set that consists of the entire hydrologic network across Australia. We used the layers Surface Network, which displays all documented streams and rivers across Australia and Surface Catchment, which shows the catchment boundaries for all of the documented streams. The stream gauges were all located on a segment of stream, each of which has a segment number. Each stream gauge’s segment number was documented.
Using R and an Upstream Catchment Aggregation Function (Bond N., Walsh C., pers. comm. 2015), we calculated upstream drainage area for the specified local subcatchment the gauge was positioned in. We then dissolved the inner boundaries giving an exact catchment area and location for each stream gauge. These files were saved as shape files to be loaded into R for further computations.

Figure 5 – This figure shows the area of the Swan River catchment drawn in ArcGIS. The red dot on the left side of the map is the location of the gauge and the outlined light green area all drains through that point. This process was repeated for all 16 different gauges.

For the next step we located a data set that reports monthly PET values on a five by five kilometer grid across all of Australia. These data were loaded into R along with the previously created shape files for each of the 16 gauges (http://www.bom.gov.au/climate/data/). We used the raster package, which analyses and models large scale spatial gridded data. The raster package computes the data by using the latitude and longitude coordinates of each catchment shape file and overlaying it on the gridded PET data we provided. The code trims the grid down to the outline of the shape file entered, sums the different PET values by weighting for
value area, and then finds the average value for the whole catchment. The final product is an average monthly time series of PET with the units of mm/month, or the amount of water in millimeters that will potentially be lost in a month due to evaporation and transpiration. The time series was then formatted to daily values by dividing each monthly value by the corresponding number of days in that month, to end with a PET time series in mm/day.

Average daily rainfall was calculated using the raster package and exact same steps used to generate the PET values. The final rainfall data produced was the average areal catchment rainfall reported in mm/day. Unlike the PET data the rainfall data was already formatted in a daily time series so no data conversion was necessary.

### 3.3 Selecting Models and Optimizers

Three commonly used CRRMs across Australia were selected for testing: AWBM, SimHyd, and Sacramento. Then using information outlined in Vaze et al. (2011) we then selected Shuffle Complex Evolution (SCE) as a global optimizer because of its common use in calibrating models across the hydrologic community. Two separate objective functions were selected to test in combination with the optimizer: Flow Duration Curve (FDC) and Nash-Sutcliffe Criterion (NSE). They were shown to work well at replicating low flows by Vaze et al. (2011).

### 3.4 Testing Model/Optimizer Combinations

Before initiating the testing, calibration periods were created for each gauge. The calibration time scale consisted of exactly two-thirds of the total number of days in the time series was calculated for each of the 16 gauges. The two-thirds date range was decided on by the model calibration information in the RRL User Guide (Podger,
The models would be calibrated for the first two thirds of the time series and then the calibration would be verified for the last third of the data. The modeling process began for each gauge by inputting the daily time series data of PET, stream flow, and rainfall into the RRL. The catchment area of the selected gauge was entered in square kilometers, calibration dates were set, and the AWBM model was chosen. SCE was then chosen as the optimizer with FDC as the objective function, changing the weight to 100% on FDC. The model was then calibrated, checked for errors, and saved as daily time series of stream flow data in ML/day. Each modeled stream flow was saved with a title identifying the location of the gauge, the fact it was model-produced data, the model that was used, and objective function that was used.

This calibration processed was then repeated for the AWBM model and NSE objective function. The FDC and NSE combinations of SimHyd and Sacramento were then calibrated using the same calibration techniques. The process then was repeated for all 15 other gauges until each one had an observed stream flow time series along with six different modeled generated stream flow time series.

**3.5 Processing Modeled Output**

The modeled and observed stream flow data were then input all the observed back into R for further analysis. Using the hydrostats package (Bond, 2014) flow duration curves were generated for each of the 16 gauge’s 7 different stream flow time series. The package also produced low-flow and CTF statistics which were used in the creation of the graphs in the results section.
4. Results

4.1 Normalized Class Flow Duration Curves

Figure 6 – This plot shows the Class One flow duration curves normalized through the 50th percentile. The x-axis has a normal probability scale while the y-axis uses a log scale.

Figure 7 – This plot shows the Class Four flow duration curves normalized through the 50th percentile. The x-axis has a normal probability scale while the y-axis uses a log scale.
**Figure 8** – This plot shows the Class Six flow duration curves normalized through the 50\textsuperscript{th} percentile. The x-axis has a normal probability scale while the y-axis uses a log scale.

**Figure 9** – This plot shows the Class Eight flow duration curves normalized through the 50\textsuperscript{th} percentile. The x-axis has a normal probability scale while the y-axis uses a log scale.
Flow duration curves of observed data were created for each class with each of the four gauges plotted in the same area. These plots are used to visually assess the hydrologic differences in each class defined by Kennard et al. (2010), and show clear patterns that validate Kennard’s method of classification. Figure 6 shows Class One: Stable base flow. It has the smallest range in flow values and almost normal distribution shown by the nearly straight lines. These four gauges located in southern Victoria and northern Tasmania demonstrates their stable hydrologic regime, with predictable base flow never dropping below a certain threshold. The lack of high flows shows a low reactivity to precipitation events and a large dependence on groundwater infiltration.

Figure 7 is Class Four: Unpredictable base flow and it shows high flows one order of magnitude greater than Class one, and low-flows around two orders of magnitude. This is with the exception of Burra Creek, which is located in South Australia thousands of kilometers away from the other gauges in Queensland and New South Wales. Burra displays base flows similar to Class One that are much greater than the rest of its class. This is the only large discrepancy between FDC’s in any of the classes shown.

Figure 8 and 9 respectively show Class Six: Predictable winter intermittent and Class Eight: Unpredictable winter intermittent. These two flow duration curves have steeper slopes than the previous two classes. They’re both intermittent and stop flowing for parts of the year, but still display high flows in the same relative magnitude.
4.2 Observed vs. Modeled Flow Duration Curves

**Figure 10** – This graph displays the observed vs. modeled flow duration curve for the Flowerdale River. The y-axis uses a logarithmic scale while the x-axis uses a probability scale.

**Figure 11** – This graph displays the observed vs. modeled flow duration curve for the Tia River. The y-axis uses a logarithmic scale while the x-axis uses a probability scale.
Figure 12 – This graph displays the observed vs. modeled flow duration curve for the Tone River. The y-axis uses a logarithmic scale while the x-axis uses a probability scale.

Figure 13 – This graph displays the observed vs. modeled flow duration curve for the Mandagery Creek. The y-axis uses a logarithmic scale while the x-axis uses a probability scale.
Of the 16 flow duration curves generated; one example curve from each class was included. Figure 10 displays the Flowerdale River FDCs from Class One. This curve shows all the models doing a relatively good job estimating mean flows, with AWBM_FDC doing the best job estimating low-flows. Figure 11 shows the Tuena River FDCs from Class Four. These FDCs show a much larger variation in model performance again with AWBM_FDC doing the best overall job compared to the observed. Figure 12 shows the Tone River FDCs from Class Six. This class displays CTF events, again with AWBM predicting low-flows with the most accuracy. Finally Figure 13 is the Mandagery FDCs from Class Eight. A large variety of model performance is seen with no models doing a particularly good job with low-flows.

4.3 Assessment of Model Performance

To evaluate a model’s goodness of fit compared to the observed runoff, log root-mean-square error (RMSE) values were calculated. The logQ was used because the flow duration curves also were plotted in log space. The first step in computing the RMSE values was to order all modeled and observed Q data from greatest to lowest. The log of all daily values was taken, and then each of the six modeled logQ values was subtracted from the observed logQ value. Each of the difference values were squared, and then summed to obtain the sum of squares. The next step was dividing sum of squares by the number of observations (excluding 0 from log transformation), and the square root was taken to calculate a final RMSE value.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

These values represent a models total mean variation from the observed curve. The lowest values represent the lowest amounts of variation, and there for the best
fitting model. This metric provides useful insight when considering a model's overall performance through the range of flow classes. Provided below are the tables for each RMSE calculation with one example FDC from each class.

Table 2 shows the results of the RMSE calculations for Class one with AWBM having three of the four lowest values. Table 3 contains the values for Class Four with AWBM_FDC and SimHyd_FDC each have two of the lowest values. Table 4 displays the Class Six values with SimHyd_FDC and Sacramento_FDC having a low value, and AWBM_FDC having two. Finally Table 5 shows the results for Class Eight with AWBM having three of the four lowest RMSE values.

### Table 2

<table>
<thead>
<tr>
<th>Flow Class</th>
<th>Gauge Location</th>
<th>AWBM FDC</th>
<th>AWBM NSE</th>
<th>Sacramento FDC</th>
<th>Sacramento NSE</th>
<th>SimHyd FDC</th>
<th>SimHyd NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class One: Stable base flow</td>
<td>Tanjil River</td>
<td>0.141</td>
<td>0.178</td>
<td>0.551</td>
<td>0.527</td>
<td></td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>Latrobe River</td>
<td>0.275</td>
<td>0.205</td>
<td>0.504</td>
<td>0.773</td>
<td>0.804</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td>Flowerdale River</td>
<td>0.041</td>
<td>0.399</td>
<td>0.195</td>
<td>0.536</td>
<td>0.233</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>Murrindindi River</td>
<td>0.178</td>
<td>0.064</td>
<td>0.302</td>
<td>0.875</td>
<td>0.182</td>
<td>0.217</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Flow Class</th>
<th>Gauge Location</th>
<th>AWBM FDC</th>
<th>AWBM NSE</th>
<th>Sacramento FDC</th>
<th>Sacramento NSE</th>
<th>SimHyd FDC</th>
<th>SimHyd NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Four: Unpredictable base flow</td>
<td>Christmas Creek</td>
<td>0.200</td>
<td>0.389</td>
<td>0.706</td>
<td>0.629</td>
<td>0.075</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>Burra Creek</td>
<td>0.715</td>
<td>0.356</td>
<td>0.367</td>
<td>0.426</td>
<td>0.133</td>
<td>0.600</td>
</tr>
<tr>
<td></td>
<td>Oxley Creek</td>
<td>0.065</td>
<td>1.390</td>
<td>0.797</td>
<td>0.736</td>
<td>1.170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tia River</td>
<td>0.204</td>
<td>0.284</td>
<td>0.686</td>
<td>0.448</td>
<td>1.086</td>
<td>0.826</td>
</tr>
</tbody>
</table>
Table 4 – This table shows the difference in log RMSE values for observed vs. modeled data in Class Six. Data for the Tone Sacramento FDC model was omitted due to computational error.

<table>
<thead>
<tr>
<th>Flow Class</th>
<th>Gauge Location</th>
<th>AWBM FDC</th>
<th>AWBM NSE</th>
<th>Sacramento FDC</th>
<th>Sacramento NSE</th>
<th>SimHyd FDC</th>
<th>SimHyd NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Six: Predictable</td>
<td>Swan River</td>
<td>0.252</td>
<td>1.093</td>
<td>0.367</td>
<td>0.977</td>
<td>0.538</td>
<td>0.847</td>
</tr>
<tr>
<td></td>
<td>Rocky Creek</td>
<td>0.250</td>
<td>0.860</td>
<td>0.224</td>
<td>0.331</td>
<td>0.160</td>
<td>0.589</td>
</tr>
<tr>
<td>Predictable winter</td>
<td>Denmark River</td>
<td>0.305</td>
<td>0.430</td>
<td>0.160</td>
<td>0.231</td>
<td>0.246</td>
<td>0.654</td>
</tr>
<tr>
<td>intermittent</td>
<td>Tone River</td>
<td>0.141</td>
<td>0.531</td>
<td>0.681</td>
<td>0.808</td>
<td>0.641</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 – This table shows the difference in log RMSE values for observed vs. modeled data in Class Eight.

<table>
<thead>
<tr>
<th>Flow Class</th>
<th>Gauge Location</th>
<th>AWBM FDC</th>
<th>AWBM NSE</th>
<th>Sacramento FDC</th>
<th>Sacramento NSE</th>
<th>SimHyd FDC</th>
<th>SimHyd NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Eight: Unpredictable winter intermittent</td>
<td>Cockburn River</td>
<td>0.116</td>
<td>0.368</td>
<td>0.449</td>
<td>0.842</td>
<td>0.363</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>Copes Creek</td>
<td>0.154</td>
<td>0.892</td>
<td>0.463</td>
<td>0.576</td>
<td>0.320</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>Tuena Creek</td>
<td>0.301</td>
<td>0.846</td>
<td>0.245</td>
<td>0.431</td>
<td>0.675</td>
<td>0.224</td>
</tr>
<tr>
<td></td>
<td>Mandagery Creek</td>
<td>0.421</td>
<td>0.490</td>
<td>0.873</td>
<td>0.460</td>
<td>0.527</td>
<td>0.636</td>
</tr>
</tbody>
</table>

4.4 Cease to Flow Statistics

The graphs below show results of the cease-to-flow analysis for each stream. For each flow class two graphs are presented showing percent of CTF days and average CTF span. The percentage days CTF function adds up all the days in the time series the stream is reporting <.01 ML/day. 0 is not used for calculation because modeled data typically reports extremely small near zero decimal values. The function then divides the CTF days by the total number days in the time series to get a percentage of days there is no stream flow. The average CTF spell function identifies each CTF event and the number of days that event occurs for. It then adds up all the events and divides by the total number of events to obtain the average length in days.
for a CTF event. These CTF values show a models ability to replicate discharge
during low-flow events. Combining this data with the RMSE values calculated before
gives a clear view of model performance with respect to low-flows.

**Class One**

![Figure 14](image1)

**Figure 14** – A) The percentage of CTF time for all modeled and observed stream flow in
Class One.

![Figure 15](image2)

**Figure 15** – B) The average CTF spell length for all modeled and observed stream flow in
Class One.
Class Four

Figure 16 – A) The percentage of CTF time for all modeled and observed stream flow in Class Four.

Figure 17 - B) The average CTF spell length for all modeled and observed stream flow in Class Four.
Class Six

**Figure 18** – A) The percentage of CTF time for all modeled and observed stream flow in Class Six.

**Figure 19** - B) The average CTF spell length for all modeled and observed stream flow in Class Six.
Class Eight

Figure 20 – A) The percentage of CTF time for all modeled and observed stream flow in Class Eight.

Figure 21 - B) The average CTF spell length for all modeled and observed stream flow in Class Eight.
Figures 14 and 15 show the CTF statistics for Class One, and Figure 16 and 17 show the CTF statistics for Class Four. Each of these classes contain only perennial streams, so any modeled CTF value is doing a poor job representing the low-flows in the waterways. In these classes, the statistics for both average time and spell length that are closest to zero are doing the best job representing low-flows.

Figures 18 and 19 show the CTF statistics for Class Six, while Figure 20 and 21 show them Class Eight. Both of these classes contain intermittent streams, so they will have non-zero values for both types of CTF statistics. This means the model that predicts a percentage of time and number of days closest to that of the observed is doing the best job at representing the low-flows on a particular waterway.

5. Discussion

5.1 Results Overview

This study shows that out of the three models tested, AWBM was overall the most consistent in predicting low-flows across the 16 gauges that were tested. Specifically AWBM_FDC and AWBM_NSE combined had 10 out of 16, or 62.5% of the lowest RMSE values. The lowest value in a gauge represents the smallest amount of variation between modeled and observed stream flows. This can be attributed AWBM’s tendency to react less to rain events; showing the best consistency in correctly modeling low-flow amount and duration. A low RMSE value for AWBM also correlated well with the CTF statistics which were generated indicting overall good low-flow performance. AWBM did display cases where its modeled discharge was considerably different than the observed, but they were much less frequent than both Sacramento and SimHyd. Sacramento had the worst overall performance with only one lowest RMSE value, and a tendency to predict modeled discharge that was
sustainably different than the observed. SimHyd also did well fitting to some flow duration curves and had five of the lowest RMSE values.

Upon further investigation by graphing calculated and model runoff amounts for the gauges (Fig. 22), SimHyd tended to show high reactivity to rainfall events even during low-flow spells.

![Tone River Modeled vs Observed Flow](image)

**Figure 22** – This graph shows the observed runoff amount on the Tone River gauge in comparison to the modeled AWBM_FDC, SIM_FDC, and SAC_FDC.

Shown in Figure 12 is flow duration curve for the Tone River. Flow duration curves represent a model’s ability to fit to the overall discharge distribution of the graph, not primarily focusing on the low-flows. The above graph in Figure 22 shows a prolonged low-flow period for the Tone River with observed and modeled flow. As shown by the RMSE value along with the flow duration curve, AWBM_FDC is the best performing model even though it tends to underestimate high flows. SimHyd tends to spike at any small rain event to peaks that are way too high, and zero while it should be representing low-flow amounts. This was a consistent finding for the overall study when looking at observed and comparing it to SimHyd’s modeled discharge. While the magnitude SimHyd predicted of low-flows is close to the
observed, the tendency it has to spike and predict huge flows while the observed data flows are low is major issue the model has. This is another reason why AWMB is the clear favorite when trying to model for low-flows. Sacramento does a good job representing mean and peak flows, but predicts no flow for a much larger percentage of time compared to the observed.

5.2 Model Structure Deficiencies

While AWBM was the most consistent, there is still much room for improvement in its accuracy of predicting low-flows. Issues arise for all three models when trying to predict for low-flow and intermittent streams. In many of these catchments there are long stretches of little to no precipitation, and the vast majority of the water entering the channel is from groundwater. None of the models tested have a parameter for groundwater data; they rely on the calibration period to establish and estimate groundwater conditions from evapotranspiration, precipitation, and the discharge data that is inputted.

The estimated groundwater conditions have huge room for error because of the lack of physical data used to generate them. This is a large factor in the inaccuracy in most rainfall runoff models when predicting low-flows. When waterways enter periods of low-flows the physical processes to generate flow are outside of the physical data fed to the model. Without this data, the model doesn’t have enough information to accurately estimate groundwater. As Barma & Lowe state, there is not yet a model in use that combines surface water hydrology and hydrogeology (2012). The extreme spatial variably and lack of physical data groundwater data provides a serious issue for hydrologists when trying to combine use it for improvements in low-flow modeling. Until a systemic method is created to combine accurate
hydrogeological data with hydrologic models, deficiencies in modeling low-flow situations will still exist.

5.3 Practical Implications

While Vaze (et al. 2011) discuss optimizer and objective function selection and weighting to best model low-flow in the RRL, they do not discuss model selection for low-flows. No previous study has been done to determine the best model to select for low-flows. This study clearly shows that choosing AWBM over SimHyd will allow for much better low-flow modeling based on the model structure, not by what objective functions are choosing.

This study was performed on three widely CRRM’s, one optimizer, and two separate primary objective functions. In the RRL alone there is the choice of six different models, 7 optimizers, 8 primary objective functions, and four secondary objective functions. The objective functions can used in combination by setting the weight to your desired amount. The possible model, optimizer, objective function combinations are endless in just the RRL. A more in depth study encompassing all models and many commonly used optimizer and function combinations could be done to gain a fuller understanding of modeling for low-flows with the RRL.

6. Conclusion

Through our study and testing of three widely used CRRMs, we found the AWBM model to perform the best for low-flows for different types of flow regimes across Australia. Its overall ability to model low-flow and CTF events was much better than the other models due to its tendency not to spike at any rainfall event. These findings provide a small piece in the puzzle of expanding our knowledge of water resources to be able to manage them in a more sustainable fashion, and
predicting future changes to the hydrologic cycle as our climate warms. In regions in which low-flow and cease to flow events are common, more research into water security is imperative as their water sources are much more limited.

While we found AWBM preformed the best out of all the models, in many cases no models came close to accurately predicting runoff for high, mean, or low-flow situations. Gains in low-flow modeling structure will be challenging because it would require the input of collected soil moisture data of the desired catchment. The extreme spatial variability of soil moisture and the lack of easy large-scale data collection strategies pose large challenges for the implementation of this process. Continued research into modeling is pivotal in furthering our knowledge and therefore conservation of water resources for future generations.
7. References

http://cran.r-project.org/web/packages/hydrostats/hydrostats.pdf


http://link.springer.com/article/10.1007/BF00939380#page-1


Gupta V., Harald Kling, Koray K. Yilmaz, Guillermo F. Martinez,
(http://www.sciencedirect.com/science/article/pii/S0022169409004843)


8. Appendices

Appendix 1 – Flow Duration Curve plots for Class One.
Appendix 2 – Flow Duration Curve plots for Class Four
Appendix 3 – Flow Duration Curve plots for Class Six.
Appendix 4 – Flow Duration Curve plots for Class Eight.