



THE UNIVERSITY  
*of* ADELAIDE

Quantifying Intermittent Streamflow  
with Low-cost Sensors and Fully  
Integrated Hydrological Modelling of  
Upstream Runoff

Alicja Natalia Makarewicz

Thesis submitted in fulfilment of the requirements for the degree of Doctorate of Philosophy

The University of Adelaide  
Faculty of Engineering, Computer and Mathematical Sciences  
School of Civil, Environmental and Mining Engineering

April 2020





## Abstract

Water policy, planning and design problems can be challenging. They are often nested across multiple scales such that downstream issues cannot be addressed without detailed understanding of upstream processes. Complexity of catchment processes, their interdependence and their feedback mechanisms represent significant catchment heterogeneity that is not easily represented in hydrological models. As a result, many problems are bounded or idealized to focus on a subset of processes restricted to a specific scale. For instance, spatially lumped hydrological models—commonly used for policy, planning and design decisions—are often calibrated to gauging data at a catchment outlet, with limited regard for the hillslope processes leading to runoff generation.

Traditional approaches to streamflow monitoring are limited by relatively high costs of gathering observations leading to sparse geographical coverage. This can be problematic when attempting to understand the non-linear and complex runoff behaviour present on hillslopes and reaches throughout an entire river network. The representation of runoff in hydrological models can be critical for supporting a range of spatially distributed problems. Many practical water management problems focus on the aggregation of processes to an outlet which can be inadequate for some questions that are nested across scales, such as management decisions relating to upstream land use, small farm dams and environmental flows. In the absence of streamflow monitoring at local scales, traditional approaches are less likely to be scalable and pose challenges in effectively representing surface and subsurface runoff generation at finer scales.

Increasingly sophisticated low-cost and low-maintenance sensing technologies are an accessible means of addressing data gaps in the spatial coverage of streamflow. Environmental sensor technologies are continually being developed with miniaturization, wireless communication and reduced costs, enabling automated electronic data loggers with reliable high frequency measurements. New initiatives in hydrological and environmental monitoring provide opportunities for high-density and widespread environmental data collection providing opportunities to better understand water management questions which span multiple scales.

This study combined process modelling with field data collected from low-cost distributed sensors to improve the representation of local scale flow processes. While process modelling can, in part, address some deficiencies, they are limited by known scaling issues of theoretical understanding and data availability of parameters that cannot be adequately measured. To compensate for these limitations, an alternative hydrological modelling approach for calibration was presented. The approach was supported with inexpensive data on the presence of water in the streambed, providing additional information on hillslope intermittency.

A small 10km<sup>2</sup> South Australian catchment, located in the Mount Lofty Ranges, was selected and instrumented across twelve sites with paired in-stream and on-bank temperature data loggers and pressure transducers required to evaluate results. A two-state hidden Markov model was applied to temperature data to classify whether the stream was 'wet' (flowing) or 'dry' (not flowing) for a given day. The accuracy of classifications was between 89% to 99% during calibration and 82% to 97% during evaluation of algorithm performance. The binary 'wet'-'dry' classifications were used to calculate a number of intermittency signatures for each site (e.g. number of zero

flow days, number of zero flow periods and average duration of zero flow periods) with the results demonstrating a high degree of heterogeneity of flow permanence within the small catchment area.

A physically based hydrological model, HydroGeoSphere, was calibrated exclusively to discharge at the outlet to represent four conceptual models of runoff generation. The four competing conceptualisations were: (1) saturation excess dominated, (2) saturation excess and groundwater dominated, (3) groundwater dominated and (4) groundwater dominated but containing 17% infiltration excess. The conceptual models dominated by groundwater discharge showed a 20% increase in low flow days directly below at point of interception compared to upstream. This highlighted that conceptual assumptions about runoff generation mechanisms in intermittent river systems have significant implications on locally dependent water management problems.

The four conceptual models were evaluated with classified 'wet'-'dry' binary data showing that no single candidate calibration performed consistently well at all upstream sites. This result demonstrating that the high heterogeneity in headwaters means that catchment-average process simulations were not able to capture localized variability. Distributed data of flow intermittency in headwater catchments was able to improve process-based hydrological model calibration with eight out of nine sites showed significant improvement in performance, with only a small deterioration to the outlet.

Headwaters, which are typically considered to include first to third order streams, are important for understanding intra-catchment fluxes and how these fluxes accumulate out of the catchment. With improved model performance there can be many benefits of getting the internal processes right, such as: simulating conditions that are outside a calibration period or infilling and completing data sets. As sensing and communications technologies continue to improve, there will be increasing opportunities to use information sources such as local-scale intermittency to supplement reliable streamflow records for representing hydrological processes across scales for more accurate water accounting and improve planning of water allocations for consumptive and environmental water needs.

## Statement of originality

I, Alicja Natalia Makarewicz, certify that this thesis contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and, where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works:

Makarewicz, A., Leonard, M., Westra, S. 2020. Quantifying streamflow intermittency and signatures with high-resolution temperature data. *J Hydrol*, in preparation.

Makarewicz, A., Leonard, M., Westra, S., Partington, D. 2020. Does conceptualization of local-scale catchment runoff production matter in representation of intermittent streams across Mediterranean catchments? *Env. Modelling & Software*, in preparation.

Makarewicz, A., Leonard, M., Westra, S., Partington, D. 2020. Does intermittency data from headwater catchments improve multi-site hydrological model calibration? *Env. Modelling & Software*, in preparation.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

)  
9<sup>th</sup> April 2020

**Alicja Natalia Makarewicz**

**Date**



## Acknowledgements

I am thankful for the privilege of being able to experience the pleasure of research. For this, I am grateful to my supervisors, Dr Michael Leonard, Professor Seth Westra and Dr Daniel Partington, for their guidance and support throughout my Ph.D. Your feedback, combined skills and inspiration have been invaluable to me throughout my candidature.

I acknowledge the support of the South Australian government Department for Environment and Water for identifying the case study and providing catchment data. The South Australian Government, Department for Environment and Water (DEW) also assisted with the development of sensor deployment methods. Additionally, I would like to thank the landowners who participated in this study for their co-operation in providing us with site access for the deployment of sensors and collection of data.

I would like to dedicate this thesis to my son, Jae Adams, who motivates me daily with renewed inspiration to be the best version of myself. Every day I have continually strived to set a good example by showing gratitude and a good work ethic. I would also like to thank my family and friends for their unwavering support and encouragement.



# Table of contents

<b>Abstract</b> .....	i
Statement of originality .....	iii
Acknowledgements.....	v
Table of contents .....	vi
List of Tables .....	ix
List of Figures .....	xi
List of symbols.....	xiv
Chapter 1.....	1
1.1 Introduction .....	1
1.2 Dynamics of intermittency flows .....	2
1.3 Anthropogenic influences .....	4
1.4 Ecological significance of intermittency.....	6
1.5 Review of modelling tools.....	6
1.6 Research gaps and questions.....	7
1.7 Thesis organisation .....	8
Chapter 2.....	10
Abstract.....	13
2.1 Introduction .....	14
2.2 Field measurements.....	17
2.2.1 Description of study area .....	17
2.2.2 Experimental setup and data collection .....	18
2.2.3 Characteristics of in-stream temperature data for intermittency classification ...	21
2.3 Methods of classifying streamflow and intermittency signature .....	23
2.3.1 Standard deviation method for intermittency classification .....	25
2.3.2 Two-state hidden Markov model for intermittency classification .....	26
2.3.3 Calibration and evaluation of intermittency classification methods.....	27
2.3.4 Quantification of intermittency signatures.....	28
2.4 Results and discussion .....	29
2.4.1 Calibration and evaluation of the standard deviation method .....	29
2.4.2 Calibration and evaluation of the two-state hidden Markov model .....	31
2.4.3 Comparison of intermittency classification methods.....	33
2.4.4 Calculated number of 'wet' flow days and intermittency signatures.....	34

2.5	Conclusions.....	37
	Acknowledgements.....	39
	References.....	40
	Supplementary material.....	45
Chapter 3.....		65
	Abstract .....	69
3.1	Introduction.....	70
3.2	Case study.....	72
3.3	Methodology .....	73
3.3.1	Model set-up: explicit simulation of surface/subsurface flows .....	73
3.3.2	Calibration of scenarios.....	75
3.3.3	Quantifying runoff generation mechanisms .....	77
3.4	Results and discussion.....	78
3.4.1	Runoff response comparison at the catchment outlet .....	78
3.4.2	Comparison of runoff response across the catchment.....	81
3.4.3	Analysis of saturation and groundwater behaviour for the runoff mechanisms... 83	
3.4.4	Comparison of flow at the local scale .....	85
3.5	Conclusions.....	88
	Acknowledgements.....	90
	References.....	91
	Appendix 3-A .....	93
Chapter 4.....		97
	Abstract .....	101
4.1	Introduction.....	102
4.2	Case study and data .....	103
4.2.1	Catchment characteristics.....	103
4.2.2	Meteorological data .....	104
4.2.3	Discharge and intermittency data .....	105
4.2.4	Calibration and evaluation time-series .....	105
4.3	Methodology .....	106
4.3.1	Model parameterisation.....	106
4.3.2	Outlet-only model calibration and evaluation .....	107
4.3.3	Multi-site model calibration and evaluation .....	107
4.3.4	Quantification of model performance with binary observations, censored simulations and discharge.....	110
4.4	Results and discussion.....	112

4.4.1	Upstream site performance of outlet-only model calibrations .....	112
4.4.2	Upstream site performance of multi-site calibration .....	113
4.4.3	Outlet performance of multi-site calibration.....	116
4.4.4	Summary of single site versus multi-site calibration .....	117
4.5	Conclusions .....	118
	Acknowledgements.....	120
	References .....	121
	Appendix 4-A.....	123
	Supplementary material .....	124
Chapter 5.....		126
5.1	Summary of research objectives.....	126
5.2	Key research findings and contributions .....	126
5.3	Research limitations and challenges.....	129
5.4	Further recommendations for future work .....	131
5.5	Conclusions .....	133
Appendices.....		134
	Appendix A: Field trip photos .....	134

## List of Tables

Table 1-1: Comparison of models and their capabilities in simulating infiltration, evapo-transpiration and subsurface flow processes. ....	7
Table 2-1: Comparison of existing methods for on-ground intermittency monitoring and estimation. Cost were estimated based on commercial rates, and are presented as a ratio relative to temperature loggers as this is likely to be more stable over time compared to absolute costs. Costs do not include cost of field installation and collection as well as data processing and time required to develop models. ....	15
Table 2-2: Site list and descriptions as well the length of data and the split used for calibration and evaluation of the classification method and method the classification were evaluated (indirectly or directly) is presented. The sensor configuration a-f refers to the sensors shown in Figure 2-2. ....	21
Table 2-3: List of the inputs iterated for the two classification methods, where the application of in-stream temperature variance alone provides a baseline performance accuracy for comparison of developed methods.....	24
Table 2-4: Definitions of intermittency signatures.....	28
Table 2-5: Results for the HMM with inputs (1) in-stream variance, (2) 30-day rainfall and (3) F(t) ratio. ...	32
Table 2-6: Comparison for the two classification methods assessed and observation inputs investigated .	34
Table 2-7: Annual intermittency signatures were calculated for each individual site of data collection.....	35
Table 2-8: The summed length of first to seventh stream orders for the South Australian region .....	45
Table 2-9: Detailed description of each site .....	46
Table 2-10: Results of standard deviation method: Variance approach .....	56
Table 2-11: Results of standard deviation method: F(t) ratio approach .....	56
Table 2-12: HMM observation inputs: in-stream variance.....	58
Table 2-13: HMM observation inputs: in-stream variance and 30 day rainfall.....	58
Table 2-14: HMM observation inputs: in-stream variance and difference in daily mean.....	59
Table 2-15: HMM observation inputs: in-stream variance and temperature profile lag .....	59
Table 2-16: HMM observation inputs: in-stream variance and ratio of on-bank to in-stream variance .....	60
Table 2-17: HMM observation inputs: in-stream variance, 30 day rainfall and difference in daily mean ....	60
Table 2-18: HMM observation inputs: in-stream variance, 30 day rainfall and lag in temperature profile .	61
Table 2-19: HMM observation inputs: in-stream variance, 30 day rainfall and F(t) .....	61
Table 2-20:HMM observation inputs: in-stream variance, difference in mean and lag in temperature profile .....	62

Table 2-21: HMM observation inputs: in-stream variance, difference in mean and $F(t)$ .....	63
Table 2-22: HMM observation inputs: in-stream variance, temperature profile lag and $F(t)$ .....	63
Table 2-23: The results of the five alternative inputs iterated in the two-state hidden Markov model which has been limited to a maximum of three inputs and with in-stream daily variance included on each calibration .....	64
Table 3-1: Surface and subsurface parameters for the catchment model, parameter values follow (Partington et al., 2010; Li et al., 2013).....	77
Table 3-2: Comparison of flow and storage behaviour for the four scenarios .....	88
Table 3-3: Definitions of the unique fractions tracked in the HMC method that are updated at each time step. ....	94
Table 4-1: Description of data available for each site and the calibration and evaluation periods applied to analysis.....	105
Table 4-2: Model parameters used to represent the catchment physical characteristics and the soil parameters applied to each sub-catchment.....	110
Table 4-3: Description of performance measures used to evaluation model simulations. ....	111
Table 4-4: Comparison of calculated performance metrics for four scenarios conditioned to the outlet. The yellow boxes indicate metrics that pass the criteria.....	112
Table 4-5: Comparison of multi-site simulated results to the observed data for each site are illustrated .	116
Table 4-6: Details of the single-site calibrations for four Scenarios representing alternative conceptualizations of runoff generation. ....	123
Table 4-7: Threshold values applied to censor simulated flows for binary 'wet'-'dry' classification to determine stream flow intermittency .....	124
Table 4-8: Comparison of multi-site conditioned simulation to simulations calibrated to the outlet only. The yellow, colours indicate where the individually calculated metric have passed the criteria. ....	125

## List of Figures

- Figure 1-1: The Dunne diagram illustrating catchment characteristics that favour different runoff generation mechanisms (Dingman, 2015; Mirus & Loague, 2013). ..... 3
- Figure 1-2 illustrates the processes of infiltration excess and saturation excess runoff generation according to well-known conceptualisations (Dunne, 1983) ..... 4
- Figure 1-3: Illustrating the (a) stream length and stream orders for a catchment in the Mount Lofty Ranges in South Eastern Australian and boxplots of storage size by stream order of the 7000 farm dams. .... 5
- Figure 1-4: Conceptual overview of research motivation, each motivation 1-3 corresponds to chapters 2-4. The large black dot represents discharge data as the outlet and the small orange dots represent additional low-cost data used to improve the simulation of catchment processes. .... 7
- Figure 2-1: The location of the 10 km<sup>2</sup> case-study catchment in a Mediterranean region of southern Australia (a), and partially vegetated, stream network and instrumentation at nine sites on third and fourth order streams (b). Circle and square markers illustrate sites where temperature sensors (sites S001, S002, S003, S004, S005, S006, S008, S010 and S011) and the square markers indicate additional pressure sensors (sites S001, S003, S006). Note that sites S007 and S009 are located in a different catchment and are not part of this case study. A high quality gauge is located at S001 and used for additional evaluation of data. Experimental setup and data collection ..... 18
- Figure 2-2: (A) Different configurations were explored in this study to evaluate the effectiveness of temperature sensors to infer intermittency. Four types of data were collected: in-stream and on-bank pressure sensors used to validate results from the paired in-stream and on-bank temperature measurements. Configuration (e) was used at site S004 where the flow onset threshold was defined as the base-flow threshold (see Table 2). Photos from the outlet (S006) showing installed (B) on-bank and (C) in-stream sensors. .... 20
- Figure 2-3: Example characteristics of paired in-stream and on-bank temperature profiles. (a) A similar mean temperature but where variation of the in-stream temperature was significantly lower, 2°C, compared to the on-bank temperature range of 20°C; (b) a difference in the mean between in-stream and on-bank profiles; and (c) similar temperature range but where the in-stream (wet) profile was lagged. 22
- Figure 2-4: Example of time series of (a) water level at an intermittent site used as reference data for the true stream state and (b) in-stream and on-bank temperatures. .... 22
- Figure 2-5: Schematic illustration of the method applied for the classification of intermittency using a hidden Markov algorithm. **O** are the observation variables, **S** are the unobserved states of the stream and **Reference** are the observed reference states used to evaluate the performance of the algorithm. .... 26
- Figure 2-6: Illustration of the standard deviation method for site S008 where: (a) the in-stream daily temperature variance is used to delineate ‘wet’ and ‘dry’ periods by applying a threshold deviation shown with the horizontal red line, where variance values below the threshold were classified as ‘wet’ and above the threshold classified as ‘dry’; (b) the ratio of on-bank to in-stream daily temperature variance with the threshold deviation difference shown with the horizontal grey line, where **F(t)** values above the threshold were classified as ‘wet’ and below the threshold ‘dry’. .... 30
- Figure 2-7: Showing site S008 over the ‘wet’ period of 18<sup>th</sup> to 26<sup>th</sup> August, where blue indicates a ‘wet’ classification and grey indicated a false ‘dry’ classification. The panels illustrate the misclassification of the standard deviation methods using the (a) variance and (b) F(t) ratio approaches compared to (c)

temperature measurements. The dashed lines show an example of daily peak in-stream temperature lag compared to the on-bank temperature. ....	31
Figure 2-8: Shows the accuracy of the two-state HMM, with the red line showing the reference state and the grey line showing the state of algorithm as classified by the HMM for site S001. Additional observations were added to the algorithm; where the (a) daily in-stream temperature variance was applied, (b) daily in-stream temperature variance and 30-day antecedent rainfall was applied, and (c) daily in-stream temperature variance, 30-day antecedent rainfall and the $F(t)$ ratio was applied....	33
Figure 2-9: Map of the study catchment, with three main tributaries, one of which is largely un-vegetated, with nine sites where data was collected for a periods of 11 to 26 months. The bar at each site shows the inferred intermittency for each site. The wet panel for sites S001 and S004 shows the presence of base-flow, with the sensor was installed and classified above the base-flow.....	37
Figure 2-10: Stream orders using the Strahler (1957) ordering method for the South Australian stream network, ranging from first to seventh order streams. ....	45
Figure 3-1: Illustration of two main runoff generation mechanisms: (a) infiltration excess runoff; and (b) saturation excess runoff. Black arrows represent general processes in the hydrological cycle and red arrows represent processes directly relevant to each associated runoff mechanism and flow pathways around a small storage. ....	71
Figure 3-2: (a) catchment location superimposed on Koppen climate classification and (b) a catchment map illustrating the main tributaries, stream orders, vegetated areas and surface areas of local storages that were represented in the model DEM.....	72
Figure 3-3: (a) Modelled catchment geometry and discretization showing smaller elements located in the channels with numbers 1-3 indicating the three main tributaries and the red dots indicating probe points to compare streambank and hillslope saturation behaviour, and (b) simulated overland flow channels (grey lines) and ponded water present in catchment dams (blue shading) with the dam investigated in this study shown with a red circle. ....	74
Figure 3-4: Schematic of the method for identifying contrasting simulations that have a comparable discharge performance at the outlet by adjusting the lateral and vertical hydraulic conductivities $K_{xyi}$ and $K_{zi}$ respectively. Four resulting scenarios were selected: (1) saturation excess dominant, (2) saturation excess & groundwater dominant, (3) groundwater dominant, and (4) groundwater dominant but with infiltration excess.....	76
Figure 3-5: The different fractions of runoff contributing to total flow at the catchment outlet over a two year continuous period for the four simulated scenarios. ....	79
Figure 3-6: Showing model outputs over the calibration and evaluation periods for the (a) applied rainfall and hydrographs of the four simulated scenarios; (b) scenario 1: saturation excess dominated; (c) scenario 2: saturation excess and groundwater dominated; (d) scenario 3: groundwater dominated; and (e) groundwater dominated with infiltration excess. ....	80
Figure 3-7: Simulated spatial behaviour at day 250 of the surface saturation (a) and depth to groundwater table (b) for Scenario 1, surface saturation (c) and depth to groundwater table (d) for Scenario 2, surface saturation (e) and depth to groundwater table (f) for Scenario 3, and surface saturation (g) and depth to groundwater table (h) for Scenario 4. ....	82
Figure 3-8: Comparison of hillslope and on-bank saturation for (a-b) Scenario 1, (c-d) Scenario 2, (e-f) Scenario 3 and (g-h) Scenario 4, with daily rainfall shown in the top panel.....	84

Figure 3-9: Comparison of scatter plots of the saturation behaviour on the hillslope and on the bank to 7 day antecedent moisture for (a-b) Scenario 1, (c-d) Scenario 2, (e-f) Scenario 3, and (g-h) Scenario 4. ....	85
Figure 3-10: Comparison of the small storage (dam) inflows, outflows and water level depth for Scenario 1 (a-c), Scenario 2 (d-f), Scenario 3 (g-i) and Scenario 4 (j-l). The blue shading represents periods where flow was present and the grey shading represents periods when flow was not present in-stream....	87
Figure 3-11: Illustrates soil column of a single cell in the model where (a) shows the tagging of infiltration excess (IE) fractions and (b) shows the tagging of saturation excess (SE) fractions. The return flow for IE and SE can re-emerge in any downstream cell providing that deep infiltration does not occur, causing the fraction of water to be tagged as groundwater. ....	93
Figure 3-12: For each cell in the model a fluid mass balance is applied to determine the fractions of runoff generation mechanisms with the example of a surface cell shown .....	95
Figure 4-1: The case study is located in (a) South Australia, South East of Adelaide with 10 sites instrumented within the catchments three main tributaries and outlet. The 10 km <sup>2</sup> headwater catchment contains (b) loam and clay soil parameters and is (c) partially vegetated. ....	104
Figure 4-2: (a) Modelled catchment geometry and discretization showing smaller elements located in the channels with numbers 1-3 indicating the three main tributaries. (b) Simulated overland flow channels (grey lines) and ponded water in the catchment (blue shading).....	107
Figure 4-3: Flow chart illustrating method for the calibration and evaluation of the model to multiple sites .....	109
Figure 4-4 Comparison of binary simulations to observed intermittent flow, showing improved representation of ‘wet’-‘dry’ patterns across all site, where the blue shading represents periods on flow. Site S010 and S011 contain only one year of observations, the first year which does not contain data is shown as grey.....	114
Figure 4-5: Percentage matching binary observations for the calibration and evaluation period for the multi-site conditioned simulation. Sites S010 and S011 were reserved exclusively for evaluations of performance across the catchment. The coloured markers compare the percentage matching binary observations for the four scenarios over a two year period. ....	115
Figure 4-6: Outlet hydrograph showing the calibration (Mar 2017-Mar 2018) and evaluation (Mar 2018-Mar 2019) periods with the (a) rainfall forcings shown and the (b) simulated discharge closely following the observed data and (c) comparison of simulated and observed intermittent flow where 0=‘no-flow’ and 1=‘flowing’. ....	117
Figure 4-7: Comparison of multi-site conditioned simulation and simulations calibrated to the outlet only i.e. Scenarios 1, 2, 3 and 4. ....	118
Figure 4-8: Showing (a) applied rainfall over two years and (b) the hydrographs of the four single-site calibrations compared to the observed discharge. ....	123



## List of symbols

Symbol	Description
$\sigma$	Standard deviation
$F(t)$	The ratio of on-bank to in-stream temperature variance
$O$	Observed variables
$S(t)$	The unobserved state of a stream, either 'wet' – flowing or 'dry' – not flowing
$S_{reference}$	Observed (reference) state of the stream
$\theta$	Model parameters
$\mu_W, \mu_D$	Mean vectors
$\Sigma_W, \Sigma_D$	Covariance matrices
$Pr$	Probability of stream state being 'wet' – flowing
$p_{DW}$	Transition probability of 'dry' to 'wet'
$p_{WD}$	Transition probability of 'wet' to 'dry'
$K_{xy}, K_z$	Lateral and vertical hydraulic conductivity
$Q_{in}, Q_{out}$	Inflow and outflow of water
$\Delta S$	Change in storage
$V$	Volume of water stored in a cell
$IE$	Infiltration excess
$SE$	Saturation excess
$f_{IE - OL}$	The fraction of water of infiltration excess runoff that is overland runoff
$f_{SE - OL}$	The fraction of saturation excess runoff that is overland runoff
$f_{IE - infil}$	The fraction of water of infiltration excess ponding that is infiltrated in a cell
$f_{SE - infil}$	The fraction of water of saturation excess ponding that is infiltrated in a cell
$f_{IE - OL - RF}$	The fraction of water of infiltration excess infiltration that re-emerges downstream as overland return flow
$f_{SE - OL - RF}$	The fraction of water of saturation excess infiltration that re-emerges downstream as overland return flow
$f_{GW}$	The fraction of water of groundwater discharge
$f_{river\_flux}$	The fraction of water that enters a channel while ponded
$f_{initial}$	The fraction of water that was present at the time of simulation initialisation.
$BC_{in}$ or $BC_{out}$	Inflow and outflow boundary conditions

# Chapter 1.

## 1.1 Introduction

Watercourses in semi-arid and Mediterranean catchments are typically intermittent and experience many periods of low- to no-flows (*Arthington et al., 2014; Datry et al., 2016; Gallart et al., 2016, Wani et al., 2017*). The high variability of intermittent flows is important to consider because more than 17% of the world's land mass falls within semi-arid climates and flow intermittency has significant ecological functioning within these climates (*Peel et al., 2007; Essenwager, 2001*). Despite their significance, there is poor understanding of flow intermittency due to traditional approaches of monitoring and modelling that focus on discharge at a catchment outlet.

A significant proportion of hydrology is at the catchment scale, with attention typically given to conceptual models that have predefined model structures and aggregate catchment characteristic based on calibration to flows at the outlet (*Wellen et al., 2015*). This method of analysis is traditionally used due to data restrictions on collecting hillslope information, large computational requirements for complex models as well as incomplete knowledge and simplified scientific assumptions. These challenges are present for all modelling methods inclusive of blackbox to physical and lumped to distributed simulation approaches.

Aggregated understanding of catchment behaviour is inadequate for many scientific and engineering problems that are distributed throughout a catchment, such as:

- **Environmental flows** – which are sensitive due to changes within a catchment and water extractions that are relevant for the ecological health of individual reaches;
- **Water quality** – influenced by sediment transport and multiscale processes controlled by runoff generation;
- **Low-flows** – which have a disproportionate impact of ecosystem health due to their low magnitude which are required on individual reaches during times of stress for native species;
- **Land use change** – which influences subsurface and surface runoff due to interception and local scale transpiration processes; and
- **Agricultural storages** – where strategically placed low-flow bypasses on farm dams are distributed in catchments and include subsurface flow pathways (i.e. base-flows).

An aggregated approach to understanding water management questions is limited because of complex interactions and feedbacks in runoff generation processes. These processes are required for understanding local fluxes and are rarely represented in commonly applied modelling approaches. Accurate representation of hydrological processes throughout a catchment is critically important due to heterogeneity in catchment features in space and time. Reaches on the hillslope (herein defined as first to third order streams) are highly variable and intermittent, and become more permanent with greater degrees of aggregation. The complexity of physical behaviour of local scale runoff generation, flow pathways and runoff response pose significant challenges in hydrological understanding, scaling and runoff partitioning.

Physically based models (*Freeze, 1969; Abbott et al., 1986; VanderKwaak et al., 2001; Brunner et al., 2012*) are able to represent processes at the local or process scale, but are limited by the ability to identify parameters at all locations due to data scarcity. The outcome of limited data is that models can be calibrated to an outlet, but the internal dynamics may not be correctly represented (*Li et al., 2015*). One option to remedy this issue is to collect additional data. Such data need not be

of the same nature as traditional long-term sources such as streamflow measurements at the outlet, but can be collected for a limited time from cheap sensors.

There are numerous alternative methods for measuring catchment properties such as intermittency (Datry *et al.*, 2016; Gallart *et al.*, 2016). Alternative methods include:

- (i) on-ground surveys and 'wet'-'dry' mapping (e.g. Stanley *et al.*, 1997; Hunter *et al.*, 2005; Turner & Richer, 2011; Larned *et al.*, 2011; Datry *et al.*, 2016).
- (ii) remote sensing (e.g. Hamada *et al.*, 2016; Puntennet *et al.*, 2017)
- (i) digital images and time lapse photography (e.g. Bradley *et al.*, 2002; Young *et al.*, 2014; Kaplan *et al.*, 2019); and
- (ii) low-cost data loggers as flow surrogates such as temperature and EC measurements (e.g. Constantz *et al.*, 2001; Blasch *et al.*, 2004; Chapin *et al.*, 2014; Arismendi *et al.*, 2017; Hofer *et al.*, 2018).

Real-time data collection from unconventional sources can be exploited in urban areas (Kerkez *et al.*, 2016) such as the use CCTV footage to investigate flood events (Le Coz *et al.*, 2016). However, image-based forms of data collection are not as readily available in rural areas due to their scale and distance from population centres. In contrast, low-cost sensor technologies can infer the occurrence of runoff events in the form of binary ('wet'-'dry') observations (Wani *et al.*, 2017).

## 1.2 Dynamics of intermittency flows

Mechanisms that generate low flows and the drivers of low flow periods are different to mechanisms generating high and medium flows. Low flows are typically derived from storage within a catchment (McMahon & Finlayson, 2003) and can originate from groundwater, shallow stormwater flow, bank storage and delayed surface flow (Smakhtin, 2001). Key factors controlling low flows are: (i) soil properties, (ii) the macropore network, (iii) the depth at which flow processes occur (shallow or deep subsurface), (iv) the magnitude of deep storage (i.e. fractured rocks or porous aquifers feeding the stream) and (v) aquifer characteristics.

Low flow processes can vary considerably, both spatially and temporally (Ouarda *et al.*, 2008). Understanding these processes is critical, such as distinguishing between Horton overland runoff, shallow subsurface storm flow or Dunne saturated overland flow (Figure 1-1). For example, a catchment with an aquifer may continue to flow during a low flow period, while a similar catchment residing on different bedrock may cease to flow in the same period. It is important to identify and understand these dominant processes because it avoids model over-parameterisation (Grayson & Blöschl, 2000).

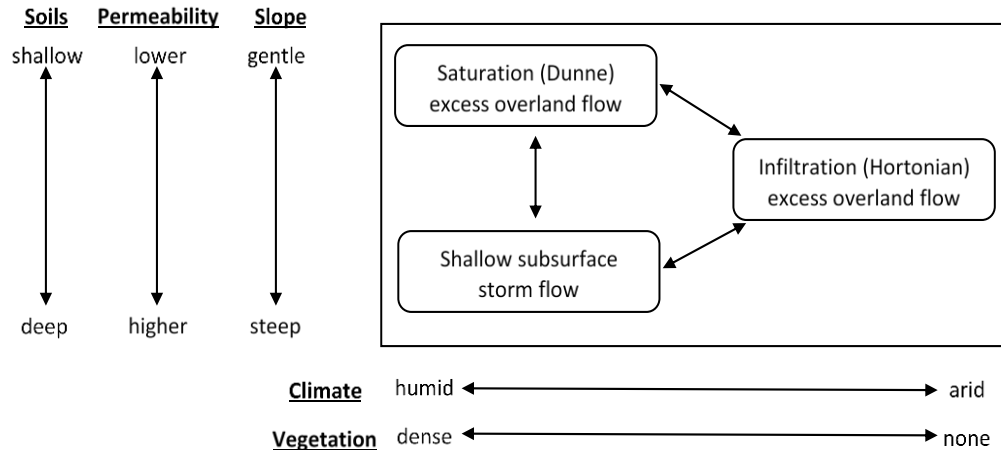
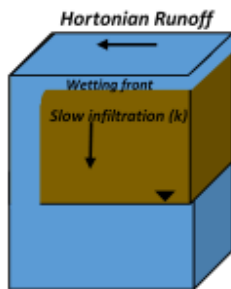
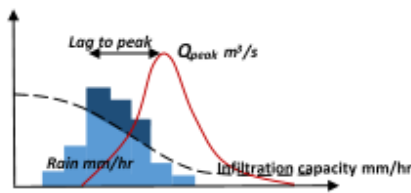


Figure 1-1: The Dunne diagram illustrating catchment characteristics that favour different runoff generation mechanisms (Dingman, 2015; Mirus & Loague, 2013).

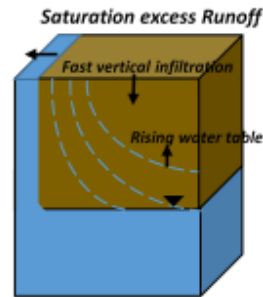
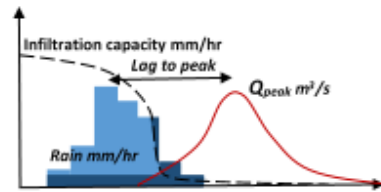
Understanding runoff generating mechanisms is necessary for responding to anthropogenic influences within catchments, setting water policy and managing flows. Figure 1-2 illustrates the processes of infiltration excess and saturation excess runoff generation according to well-known conceptualisations (Dunne, 1983). Infiltration excess runoff is produced by saturation of the soil surface (from above) when the intensity of precipitation exceeds the rate of infiltration. Saturation excess runoff occurs when the entire soil profile is saturated (from below) resulting in return flow and runoff from subsequent precipitation on this area. Runoff mechanisms are non-uniform and highly variable for a variety of reasons including: variable infiltration capacities and depression storage (Esteves and Lapetiti, 2003); physical and chemical properties of the soil surface such as soil crusting (Vaezi et al., 2010); hydraulic conductivity and slopes (Li et al., 2012); the size of a precipitation events (Newman et al., 1998); antecedent conditions (Singh, 1997); whether a catchment is water-limited or energy-limited (Trancoso et al., 2016); and subsurface flow path distribution, transit times and distribution of shallow storage (Vivoni et al., 2007). Catchment runoff response is often nonlinear and scale-dependent, with multiple mechanisms occurring simultaneously in a catchment (Vivoni et al., 2007).

### Hortonian (HOF)



Rainfall intensity exceeds the infiltration capacity  
 $R > I$

### Saturation excess (SOF)



Low soil water storage capacity  
i.e. perched or high water table or wet initial antecedent conditions

Figure 1-2 illustrates the processes of infiltration excess and saturation excess runoff generation according to well-known conceptualisations (Dunne, 1983). Where  $R$  is rainfall intensity and  $I$  is infiltration capacity.

## 1.3 Anthropogenic influences

Anthropogenic impacts have a significant impact on a basin's flow regime. Growing populations in many regions result in increasing pressure and demand on freshwater resources, with increasing pressure for human consumption, industry, agriculture as well as environmental requirements. These pressures are further amplified in drier water-limited catchments in arid and semi-arid regions. Human impacts on the landscape affect low flow frequencies and magnitudes. Flow regulation, surface water abstraction, groundwater abstraction and land use all influence gains or losses to low flow discharges. At present, over two million agricultural dams are distributed on hillslopes in Australia alone, each with limited storage volume but collectively storing over 8,000 GL of water (Land and Water Australia, 2010)—a volume equivalent to the total volume of water stored in large reservoirs supplying Australia's major capital cities.

A desktop study conducted in an area in the Mount Lofty Ranges showed that 90% of streamflow lengths are first to third order streams (Figure 1-3a). Similarly, of the more than 2,500 km<sup>2</sup> area and 7000 farm dams investigated, 90% were located on headwater catchments (first to third order). Additionally, 70% of the storage capacity was located on the headwaters, with larger dams typically located downstream (Figure 1-3b). On average the density across the entire area was determined as 6.6 ML/km<sup>2</sup>, however some sub-catchments had surface storage densities as high as 25.0 ML/km<sup>2</sup>. Assuming a runoff co-efficient of 10% and an average rainfall of 800 mm per year, these densities imply that as much as 33% of surface runoff is being detained in dams across the area. This is a significant volume, especially for the water-limited catchments that typify Southern Australia.

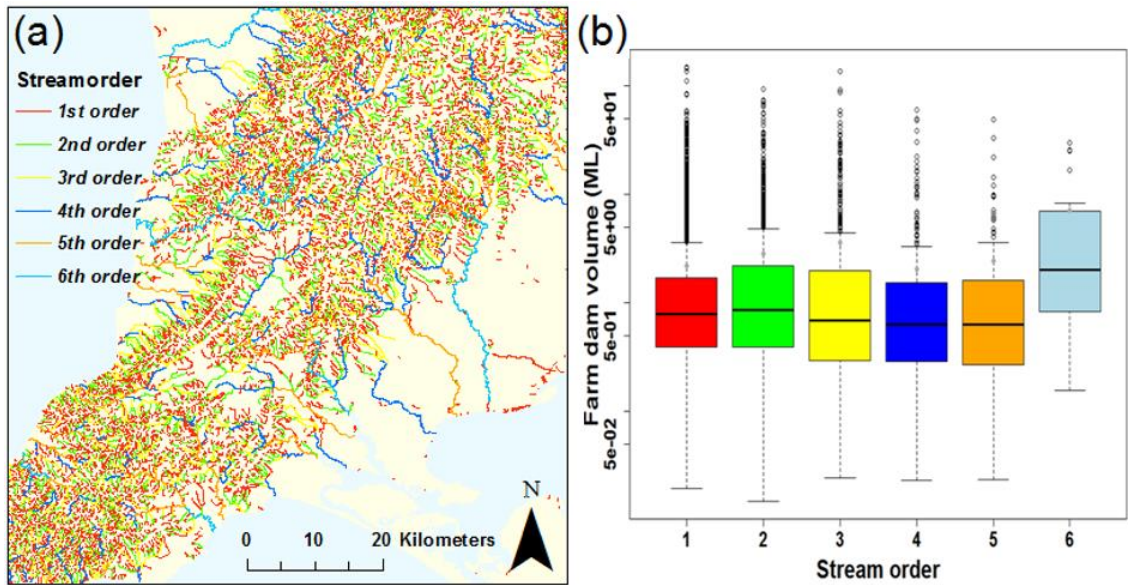


Figure 1-3: Illustrating the (a) stream length and stream orders for a catchment in the Mount Lofty Ranges in South Eastern Australian and boxplots of storage size by stream order of the 7000 farm dams.

There are many possible causes for interception of runoff across a catchment relating to land use change, including: fire, increased urbanisation, remediation, development of wetlands, forestry and farm dams. Farm dams are highlighted here because they represent a common feature of the landscape in agricultural regions that are distributed throughout a catchment. Farm dams can range in size from a few kilolitres to hundreds of mega litres, and in recent years the cumulative impacts to catchment health of these dams have been recognized (Nathan & Lowe, 2012). Farm dams are located upstream on intermittent watercourses and can be responsible for altering seasonal flow patterns when present in high densities. Farm dams can also be responsible for impacting low flow periods by extending natural cease of- and low- flow durations (Grantham *et al.*, 2010). To understand their impact requires spatially explicit representation and not just an aggregated model that relies on flows at the outlet. However, understanding their impacts is challenging because they are dispersed in a catchment and there is a lack of data availability on hillslopes and on individual reaches.

## 1.4 Ecological significance of intermittency

The low flow end of the hydrograph is vulnerable to alteration caused by anthropogenic changes to the flow regimes (*Smakhtin, 2001*), however the mechanistic effects of human-induced water reduction on riverine ecosystems is poorly understood (*Rolls et al., 2012; Boulton and Lake, 2008*). The low flow end of the hydrograph is vulnerable to alteration caused by anthropogenic changes to the flow regimes (*Smakhtin, 2001*). However, the mechanistic effects of human-induced water reduction on riverine ecosystems is poorly understood (*Rolls et al., 2012; Boulton and Lake, 2008*). A key challenge for water policymakers is balancing the trade-offs between environmental water requirements and economical water demands (e.g. agricultural). This challenge is amplified during dry periods, as it is recognised that a small percent change in low flow magnitude may cause a disproportionately large change in ecological response (*Rolls et al., 2012*). In Australia low flow periods are a natural feature of river systems, but they are also periods of high stress for ecosystems. Therefore, to better address water management questions, it is critical to understand and acknowledge the ecological characteristics and responses to low flow regimes.

Surface-flow and groundwater-level fluctuations drive expansion-contraction cycles of intermittent river networks (*Stanley et al., 1997*), which generate alternating patterns of wetting and drying of reaches. These cycles result in the shifting of habitats at a network scale between lotic, lentic and terrestrial habitats, forming a shifting aquatic-terrestrial habit mosaic (*Datry et al., 2014*). Hydrological drivers of magnitude, frequency and duration of cease-to-flow and low flow periods control the diversity, spatial arrangement, turnover and connectivity of ecosystems (*Stanley et al., 1997; Bunn et al., 2006*).

## 1.5 Review of modelling tools

A majority of hydrological models widely used by policy decision makers represent low flow and intermittency flow processes poorly. Very few models provide reliable information on subsurface flow pathways and are not able to adequately represent low flows. Current modelling frameworks focus on reproducing peak and total flow volumes, and have significant structural weaknesses with respect to reproducing low flow behaviour. For example, the GR4J conceptual model (*Perrin et al., 2001*) simulates flows as a function of storage, making it impossible to replicate zero flow periods without some form of threshold parameter. Furthermore, models that have been developed for one catchment or time period can perform poorly when applied to another catchment or time period. The uniqueness of flow mechanisms makes it difficult, if not impossible, to develop a single model structure suitable in all circumstances and locations. Therefore, process modelling and/or flexible modelling structures are better for understanding hydrological systems.

There are a number of models capable of investigating subsurface flow paths as well as testing alternative model configurations (or model hypothesis based on alternative runoff mechanisms). Some of these include SUMMA (*Clarke et al., 2015*), SUPERFLEX (*Fenicia et al., 2011*), MIKE SHE (*Abbott et al., 1986*) and HydroGeoSphere (*Therrien et al., 2006*) (Table 1-1). HydroGeoSphere (HGS) was selected as the best candidate model and was used for all modelling tasks required for the project because it was able to represent the required physical processes while also being able to quantify proportions of runoff mechanisms. HGS is a fully coupled surface-subsurface model that represents 3D variably saturated flow using the modified Richard's equation and 2D surface flow using the diffusion wave approximation to the St Venant equations (*Therrien et al., 2009*).

Table 1-1: Comparison of models and their capabilities in simulating infiltration, evapo-transpiration and subsurface flow processes.

	Structure	Infiltration	Evapo-transpiration	Unsaturated Flow	Saturated flow	Groundwater contribution	Preferential flow
<b>SUMMA</b>	Flexible/modular	Yes-Green and Ampt	Yes, various representation	Yes-Mixed form of Richards equation using Van Genuchten closure relations – vertical redistribution	Yes, Function of water store with power-law hydraulic conductivity profiles	Yes, Conceptual power law representation	No
<b>SUPER FLEX</b>	Flexible/modular	Yes	Yes	Conceptual bucket representation	Conceptual bucket representation	Yes	No
<b>MIKE SHE</b>	Fixed	Yes-Green and Ampt	Yes, Rutter Accounting method	Yes, (1) multilayer with Richards equation or gravity flow; or (2) two-layer root zone	Yes, 3D Bousinesq	Yes	No
<b>Hydro Geo Sphere</b>	Fixed	Yes-Richards Equation	Yes- Rutter Accounting method	Yes, Full 3D Richards	Yes, Full 3D Richards	Yes	Yes, dual porosity & permeability

## 1.6 Research gaps and questions

There are many processes distributed throughout a catchment, yet often only catchment outlet data is used to represent the catchment. With the improvement of technology, it is possible to increase data collection on the hillslope with low-cost sensors. It is necessary to demonstrate the impact and benefits of additional data and how it can inform intermittent flow processes at the reach scale. There are many model formulations and assumptions, and it is important to show that a small amount of additional data can improve the simulation of internal processes. These points are summarized in Figure 1-4 which outlines the three motivations of this research.

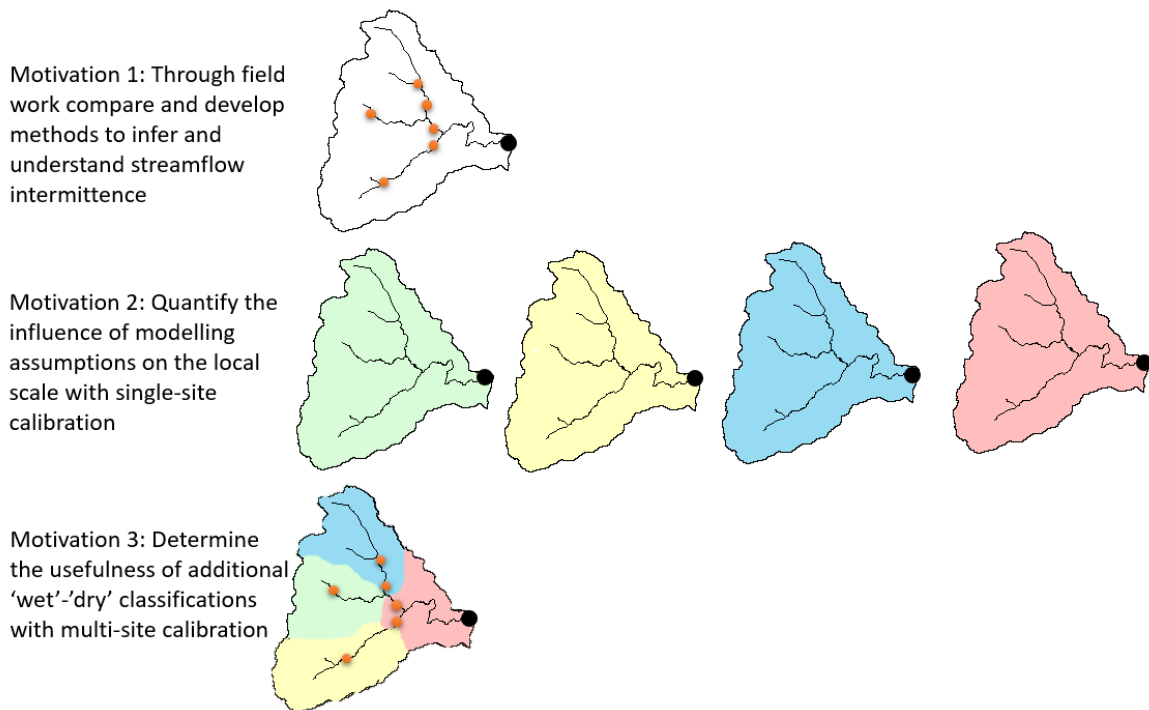


Figure 1-4: Conceptual overview of research motivation, each motivation 1-3 corresponds to chapters 2-4. The large black dot represents discharge data as the outlet and the small orange dots represent additional low-cost data used to improve the simulation of catchment processes.



The overall objective of this research was to demonstrate the value of additional data for characterising intermittent flow processes within a temperate catchment. The corresponding research questions were addressed by combining relatively underutilised methods of data collection via low-cost environmental sensors with a physically-based model in order to improve the representation of local scale processes.

Three specific research objectives have been identified:

**Objective 1: Quantifying streamflow intermittency and signatures with low-cost sensing technology:** To assess the effectiveness of low-cost temperature sensors in their ability to infer streamflow intermittency through the application of a two-state hidden Markov model applied across multiple locations (*Paper 1*).

*Objective 1.1:* Compare and develop methods for inferring streamflow intermittency with binary 'wet'-'dry' classifications derived from continuous temperature measurements.

*Objective 1.2:* Quantify intermittency signatures for individual reaches and data collection locations to illustrate differences between the sites.

**Objective 2: Implications of modelling assumptions on the representation of local-scale intermittent streamflow:** A fully-coupled process model, HydroGeoSphere, was applied to develop alternative simulated scenarios calibrated to discharge at the outlet while representing alternative runoff mechanism on the hillslope. (*Paper 2*).

*Objective 2.1:* Understand the effect of spatial variability of alternative simulated runoff mechanisms calibrated exclusively to outlet discharge and;

*Objective 2.2:* To investigate the influence of different runoff mechanisms on localized flow pathways.

**Objective 3: Representing intermittent streamflow in headwaters using additional data in a multi-site calibration:** To determine the extent to which data collected in a headwater catchment can be applied to improve calibration (*Paper 3*).

*Objective 3.1:* To quantify the upstream performance of multiple candidate calibrations based on an outlet-only calibration method.

*Objective 3.2:* To illustrate a multi-site calibration method for a physically based model that utilizes additional collected data to improve the representation of upstream physical flow processes.

*Objective 3.3:* To compare the performance of the outlet-only and multi-site calibrated models.

The outcome of this thesis was a method to generate daily timeseries of 'wet'-'dry' values and associated intermittency signatures using low-cost environmental sensors. The additional intermittency data was used to calibrate a physically based model and improve process representation throughout the catchment, and in so doing improve the ability to address water management questions relating to multiple reaches of the catchment.

## 1.7 Thesis organisation

The thesis contains five chapters, with the main contributions presented in **Chapter 2** to **Chapter 4**. The contents of each of these chapters map directly to each of the three research objectives in Section 1.6, and are presented in the form of a journal paper. The focus of these chapters is as follows:

- **Chapter 2** (Objective 1, Paper 1) presents the proposed approach for inferring streamflow intermittency and quantifying intermittent signatures with low-cost sensing technologies.
- **Chapter 3** (Objective 2, Paper 2) explores the implications of modelling assumption on local-scale intermittent stream flow by investigating alternate model parameterisations using a fully-coupled process model, HydroGeoSphere.
- **Chapter 4** (Objective 3, Paper 3) demonstrates the benefits representing intermittent streamflow using low-cost and additional data shown with in a multi-site calibration method.

Conclusions are provided in **Chapter 5**, which includes a discussion of the contributions, limitations and future directions of the research.

## Chapter 2.

# Quantifying streamflow intermittency and signatures with high-resolution temperature data

Alicja Makarewicz<sup>a</sup>; Michael Leonard<sup>a</sup>; Seth Westra<sup>a</sup>

<sup>a</sup>School of Civil, Environmental and Mining Engineering, the University of Adelaide, North Terrace, Adelaide SA 5005, Australia.


*Journal of Hydrology*, in preparation



# Statement of Authorship

Title of Paper	Quantifying streamflow intermittency and signatures with high-resolution temperature data
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Invitation pending to resubmit to the Journal of Hydrology


## Principal Author


Name of Principal Author (Candidate)	Alicja Makarewicz		
Contribution to the Paper	Contributed to the method design. Performed analysis, interpreted results and wrote the manuscript		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Diploma by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/03/2020

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Michael Leonard		
Contribution to the Paper	Contributed to the method design, analysis and interpretation of results and editing the manuscript		
Signature		Date	26/03/2020

Name of Co-Author	Seth Westra		
Contribution to the Paper	Contributed to the method design, analysis and interpretation of results and editing the manuscript		
Signature		Date	25/03/2020

Please cut and paste additional co-author panels here as required.

## Abstract

Headwaters generally lack continuous long-term data collection of streamflow intermittency. Headwaters, which are typically considered to include first to third order streams, are important for understanding intra-catchment fluxes and how these fluxes accumulate out of the catchment. The heavy reliance on streamflow gauges as the sole calibration dataset for hydrological models is increasingly recognized as a limiter to predictive accuracy and fidelity of simulated small scale hydrological processes. Low-cost and low-maintenance sensing technologies are an accessible means of addressing spatial data gaps which complement other high quality data sources (e.g. streamflow and rain gauges) and to provide high resolution spatial representation of upstream runoff behaviour. Intermittent streamflow signatures, based on flow surrogates at multiple points in a catchment, were developed to provide additional understanding of streamflow regimes within individual reaches. Paired low-cost temperature sensors at nine headwater sites were implemented within a 10 km<sup>2</sup> catchment. A two-state hidden Markov model was applied to classify whether the stream was 'wet' or 'dry' for a given day with the binary classifications used to determine a range of intermittency signatures. For the nine sites, the accuracy of the classifications was between 89 to 99% during calibration and 82 to 95% during evaluation. Differences between the tributaries were demonstrated in terms of onset and secession of flow, number of zero flow days, number of zero flow periods, and percentage of flow permanence annually and by season. The study quantifies the high degree of variability in intermittent patterns throughout the small catchment area, demonstrating the benefits of alternative data collection methods that are not otherwise accessible with the existing network of streamflow gauge data alone. With ongoing development of low-cost sensor technologies, it is likely that the use of this information as complementary data to traditional streamflow gauging's provides a valuable opportunity to improve hydrological understanding across catchments.

## 2.1 Introduction

Streams in Mediterranean and semi-arid headwater catchments are often characterized by high variability and intermittency (Arthington *et al.*, 2014; Datry *et al.*, 2016; Gallart *et al.*, 2016; Wani *et al.*, 2017). Intermittent flow events are significant for understanding solute and sediment fluxes and how flows change throughout a catchment (Gomi *et al.*, 2002). Intermittent flows are also important in semi-arid regions and provide many benefits, such as: the effect of intermittent flows on riparian vegetation (Stromberg *et al.*, 2005) and macroinvertebrate communities (Vidal-Abarca *et al.*, 2013). Headwater catchments, which are defined herein as comprising first to third order streams, can often be difficult to access, especially during rain periods. Moreover, unlike lower reaches, these catchments often lack continuous long-term data collection of streamflow onset and cessation. Methods to estimate intermittent flows have typically been restricted to field observations, such as 'wet'-'dry' mapping at a coarse time scale (Stanley *et al.*, 1997; Datry *et al.*, 2016; Godsey & Kirchner, 2014; Jensen *et al.*, 2017), models or regressions to infer properties of ungauged catchments (Engeland & Hisdale, 2009; Snelder *et al.*, 2013) or through probabilistic models (Ogtrop *et al.*, 2011).

Headwaters make up approximately 80% of the total stream length within a catchment, draining 70-80% of the total land surface area (Sidle *et al.*, 2000; Meyer and Wallace, 2001). In southeast Australia, for every kilometer of fourth order stream there are over 30 kilometres of first order river network (see supplementary information). These catchments are usually sparsely monitored, with streamflow gauges typically located in higher-order streams (Leigh *et al.*, 2016) where flows are more regular and total flow volumes are greater. This limits the understanding of flow intermittency for low order streams. However, intermittency is an important driver of river ecosystems across multiple scales (Larned *et al.*, 2010), and recognition of the importance of understanding local-scale intermittent processes is increasing (Leigh *et al.*, 2016).

To characterize localized streamflow patterns at the headwater scale requires a distributed sensor network. A network of sensors is required because of the significant heterogeneity often characterized at localized scales. The costs of streamflow gauging stations comprising a dense network of observation locations is prohibitive. There are further practical limitations and measurement challenges with this option, for example intermittent channel scour and deposition during flashy events can continuously change channel geomorphology and lead to non-stationary stage-discharge curves (Constantz *et al.*, 2001). For these reasons, a different approach that is not reliant on traditional streamflow gauging technologies is required.

Progress in environmental sensor technology has resulted in the miniaturization of electronic devices, improvements in wireless communication (particularly in remote regions), and reduced costs (Ruiz-Garcia *et al.*, 2009). This combination of technological improvements creates new opportunities for automated electronic data loggers and reliable high frequency measurements (Wickert *et al.*, 2018), with potential applications for widespread hydrological and environmental data collection (Lovett *et al.*, 2007). Recent studies have investigated the potential of on-ground data loggers as alternative methods of flow detection (Table 2-1). In particular, temperature sensors have been used to determine: river-aquifer interaction (McCallum *et al.*, 2014), flow permanence based on streambed thermographs (Constantz *et al.*, 2001; Blasch *et al.*, 2004, Arismendi *et al.*, 2017, Hofer *et al.*, 2018) and flow intermittency with modified loggers measuring conductivity (Blasch *et al.*, 2001; Chapin *et al.*, 2014). Alternatively, image based technologies have shown the benefit of time lapse photography (Young *et al.*, 2014; Kaplan *et al.*, 2019) and video

imagery (Bradley *et al.*, 2002) to estimate discharge or closed-circuit TV footage for flood hydrology (Tsubaki *et al.*, 2011; Le Coz *et al.*, 2016).

Table 2-1: Comparison of existing methods for on-ground intermittency monitoring and estimation. Cost were estimated based on commercial rates, and are presented as a ratio relative to temperature loggers as this is likely to be more stable over time compared to absolute costs. Costs do not include cost of field installation and collection as well as data processing and time required to develop models.

Sensor type	Cost (Ratio of cost relative to temperature loggers)	Data processing	Limitations	References
Temperature sensors	Very-low 1:1	High	<p>Precipitation can mask the presence of streamflow</p> <p>Large requirement of data interpretation</p> <p>Deposit of sediment around sensors can affect analysis</p> <p>Diurnal temperature range is required for analysis. Therefore, flow conditions that last for short periods (i.e. minutes to hours) as missed</p>	<p><i>Constantz et al.</i>, 2001</p> <p><i>Blasch et al.</i>, 2004</p> <p><i>Chapin et al.</i>, 2014</p> <p><i>Arismendi et al.</i>, 2017</p> <p><i>Hofer et al.</i>, 2018</p> <p><i>Paillex et al.</i>, 2019</p>
Pressure transducers	High 1:10	High	<p>Scour and deposit in channels from flash events often results in time varying stage-discharge curve.</p> <p>Can be costly for equipment setup compared to other low-cost sensor methods while providing data of flow magnitudes.</p> <p>Barometric pressure correction required</p>	<p><i>Kröger et al.</i>, 2008</p> <p><i>Dugan et al.</i>, 2009</p> <p><i>Carling et al.</i>, 2012</p> <p><i>Magnusson et al.</i>, 2012</p>
'Wet'-'dry' mapping Surveying	Low to high -	Low	<p>Flashy nature of intermittent streams means flowing conditions can last for short periods (i.e. minutes to days) resulting in missed events</p> <p>Accuracy is limited by frequency of site visits</p> <p>Large overhead of time required for observations</p>	<p><i>Hunter et al.</i>, 2005</p> <p><i>Turner &amp; Richer</i>, 2011</p> <p><i>Larned et al.</i>, 2011</p> <p><i>Datry et al.</i>, 2016</p>
Electrical conductivity sensors	Low 1:3	Low	<p>Deposit of sediment around sensors can lead to false signals</p> <p>Method is unable to determine whether the sensor is in a stagnant pool or a flowing stream</p>	<p><i>Blasch et al.</i>, 2001</p> <p><i>Goulsbra et al.</i>, 2009</p> <p><i>Jaeger &amp; Olden</i>, 2012</p>



			Conversion of temperature sensors are required for low-cost solutions, e.g. removing the thermistor and implementing electrical probes in the logger  Clogging of electrical contacts can disturb measurements	<i>Chapin et al., 2014</i>
Image analysis	Low  1:4	High	Battery and memory storage requirements  Large effort required for data analysis  Shade and texture changes between the bank and water are not strong or consistent resulting in ambiguous detection of water levels  Lens can be affected by weather conditions such as rain or fog.	<i>Bradley et al., 2002</i> <i>Tsubaki et al., 2011</i> <i>Young et al., 2015</i> <i>Le Coz et al., 2016</i> <i>Kaplan et al., 2019</i>

While there are examples of studies where alternative technologies, such as temperature and electrical conductivity sensors, have been applied (*Blasch et al., 2001; Constantz et al., 2001; Le Coz et al., 2016*), there has been significantly more investment to advance modelling approaches (*Mishra, 2009*). Nonetheless, there is a general consensus within the hydrological community that field work and additional monitoring are required to advance hydrological understanding and assist modelling advancements (*Blume et al., 2017*) because upscaling of runoff behaviour is rarely strictly additive, with headwaters behaving as complex heterogeneous systems that have localized impacts and influences (*Kirchner, 2006*). Alternative and low-cost monitoring approaches provide an opportunity to address data gaps but more research is required to investigate the quality of the measurements and the benefits they can provide. For example, intermittency signatures can be useful in constraining models and improving the representation of hydrological processes with advanced model calibration. Intermittent signatures are also useful in providing eco-hydrological insight, such as the fluctuation of ‘wet’ (flowing) and ‘dry’ (not flowing) states (*Boulton et al, 2017*), local-scale remnant pools (*Dell et al, 2017*) and surface-subsurface water interactions.

This paper presents a method for inferring streamflow intermittency in Mediterranean headwater streams using temperature sensors. Temperature sensors were used as they are a cost effective method of flow detection on intermittent reaches. Unlike other low-cost sensor options using EC sensors (*Blasch et al., 2001; Goulsbra et al., 2009; Jaeger & Olden, 2012; Chapin et al., 2014*), the temperature sensors do not need modification (i.e. replacing the thermistor in temperature sensors for EC probes) to be applicable. Additionally, continuous measurements of temperature are frequently collected in headwater catchments by government agencies for water quality purposes because water temperature is an important parameter for ecological organisms (*Webb et al., 2007*).

To assess the effectiveness of low-cost temperature sensors in their ability to provide improved estimates of streamflow intermittency in Mediterranean reaches, the paper addresses the following objectives:

- (1) To develop and compare methods for inferring streamflow intermittency with binary 'wet'- 'dry' classifications from continuous temperature measurements; and
- (2) To quantify intermittency signatures for individual sites and demonstrate differences between them.

To achieve the objectives, continuous 15-minute in-stream and on-bank temperature measurements were collected at nine sites within a small 10 km<sup>2</sup> catchment located in South East, South Australia (Section 2.2.2). The thermal signatures of the data were assessed (Section 2.2.3), and a two methods of streamflow classification were compared: (1) a standard deviation approach where a threshold deviation was applied (*Constantz et al.*, 2001; *Blasch et al.*, 2004); and (2) a two-state hidden Markov model was applied for stream state classification (*Arismendi et al.*, 2017) (Section 2.3). The inferred binary ('wet'-'dry') classifications were used to interpret the state of multiple locations (Section 2.4). The discussion explores potential uses of this data (e.g. development of a physical model) and the wider value of environmental sensors for policy, planning and management of water resources (Section 2.5).

## 2.2 Field measurements

### 2.2.1 Description of study area

A small warm and temperate catchment in South Australia, GPS coordinates (-35.266420, 138.731021), was selected to assess the feasibility of temperature sensors as a measure for flow intermittency (Figure 2-1a). Due to resource constraints a single catchment was selected for the pilot study. The case study catchment, approximately 10 km<sup>2</sup>, was selected based on the diversity of catchment features as well as the availability of high quality hydrometric gauging data. The catchment has three main tributaries with a total length of approximately 5 km and has vegetated areas that cover a large proportion of two of the tributaries. The vegetated areas contain woody trees (e.g. eucalypts with a dense understory), with the remaining catchment covered with grass (8%), pastures (50%) and sparse woody trees (2%). The elevation ranges from 175 m to 420 m above sea level, and contains fractured rock aquifers with a shallow to moderately thick topsoil layer of acidic, sandy loam and clayey soils. The reaches range from first- to fourth-order streams and mostly have seasonal flow in the winter/spring months (Figure 2-1b). Based on site visits and informal discussions with local residents, it was known that some sites have significant recharge to groundwater while other sites retain a permanent flow due to bedrock exfiltration. The region is also an important water resource for domestic water use, irrigation of crops and stock as well as providing water for environmental purposes, with environmentally significant assets such as Fleurieu Peninsula swamps and numerous pools/springs located downstream.

The channels have varying profiles, for example site S005 is shallow and only 36.5 cm wide, Site S008 is 118.6 cm wide and deeply incised due to high velocities, site S002 is deeply incised but narrow (64.0 cm) while the outlet location had a width of 251.8 cm (see supplementary material). The streambed at all sites contain dark brown soils with high loads on organic material present in-stream and on-bank (clay soils with high organic matter content). Reach 1 contains pools, springs and riffles, with sections of the channel containing continuous base-flows all year round. Upstream of Reach 1 is rocky with outcrops of fractured rock observed in and around the channel. Reach 2 and Reach 3 are seasonal, while Reach 5 is densely vegetated and steep (slope approximately 0.045 m/m) and Reach 3 is cleared of dense vegetation and relatively flat by comparison (0.020 m/m).

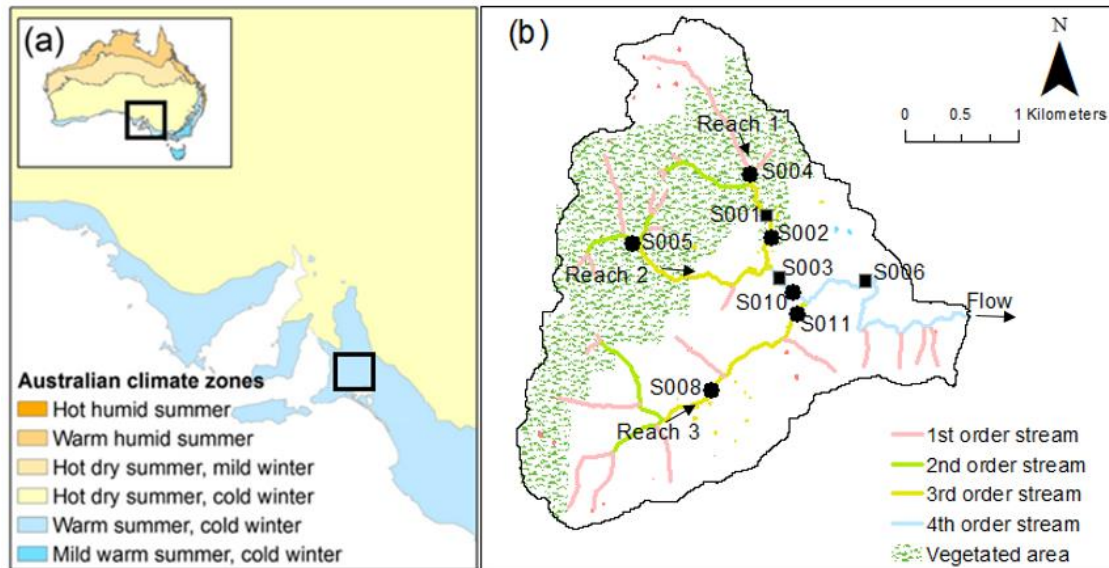


Figure 2-1: The location of the 10 km<sup>2</sup> case-study catchment in a Mediterranean region of southern Australia (a), and partially vegetated, stream network and instrumentation at nine sites on third and fourth order streams (b). Circle and square markers illustrate sites where temperature sensors (sites S001, S002, S003, S004, S005, S006, S008, S010 and S011) and the square markers indicate additional pressure sensors (sites S001, S003, S006). Note that sites S007 and S009 are located in a different catchment and are not part of this case study. A high quality gauge is located at S001 and used for additional evaluation of data.

## 2.2.2 Experimental setup and data collection

High velocity flows, sediment transport and floating debris conditions can present a challenge for long-term accurate sensor readings. For example, the steel post at site S008 was bent during a high-flow period and showed significant sediment build-up (supplementary material). Knowing these challenges, robust and commercially available sensors were deployed at all sites. The two types of instruments selected to classify and then evaluate the state of the stream as either ‘wet’ or ‘dry’ were:

- (1) Temperature sensors with the ability to log temperature measurements – Onset Hobo Pendant<sup>®</sup> UA-001-64 temperature data logger covered in a waterproof casing; and
- (2) Pressure sensors with the ability to log water levels – Onset Hobo Water level U20L-04 used as reference data for the true (of reference) state of the stream.

The pressure sensors in this study were used for evaluation purposes. All sensors were deployed to record at 15-minute intervals, which allows for approximately 12 months of data to be stored on the sensor. The battery life of the sensors was rated at typically over 12 months. In this study, sensor batteries lasted the duration of the study (i.e. up to 26 months).

The network of sensors was designed to obtain measurements for the three main tributaries, with additional sensors used at some sites to compare temperature readings. Vegetated and non-vegetated sites were included to investigate the effects on streamflow timing. Nine sites were selected (Figure 2-1b) including the outlet and both upstream and downstream locations for each tributary. Four kinds of data were collected using two kinds of sensors: (1) on-bank temperature, (2) on-bank barometric pressure, (3) in-stream temperature, and (4) in-stream water pressure. The on-bank barometric pressure was used to obtain the water level for the in-stream pressure measurements at sites S001, S003 and S006. The pressure sensors installed at S001, S003 and S006 provide the ability to directly evaluate classifications at these sites. All other sites (S002,

S004, S005, S008, S010 and S011), located upstream of the pressure sensors, contain only temperature data loggers. At sites S010 and S011 the sensors were installed in the second year of data collection because they were immediately upstream of the confluence of the two largest tributaries and were used to identify whether there were similarities in the respective intermittent patterns.

To establish consistency when applying temperature sensors to infer intermittency, two different sensor configurations were implemented: (1) on-bank and in-stream temperature sensors; and (2) on-bank temperature and dual in-stream temperature sensors on a single post (Figure 2-2). The second configuration was applied because pressure sensors were not deployed at every site; therefore, using dual sensors provided an additional check on the thermal data used to classify the stream state. Additionally, the dual sensor configuration was applied to site S004 where the stream is fed by groundwater all year round and thus rarely dries out. For this case, one sensor installed below the base-flow depth and a second sensor installed above the base-flow depth to record response to rainfall events. A threshold for the base-flow depth was defined as the depth of flow that was present at the time of sensor installation during the annual dry period (December to March). This was only relevant for sites S001 and S004 where year round base-flow was present. All flows below a depth threshold (Table 2-2) were described as 'dry' (or 'zero-flow') and all recorded flow responses were defined as 'wet'. The reason for installing the sensors just above the streambed (e.g. approximately 0.10 m) was to minimize the risk of silt and sediment build up on and around the sensor which can cause errors in readings.

The on-bank temperature sensor was used to reduce errors as a result of micro-climate variations across the sites (i.e. shade or thermoclines in valleys). On-bank temperature sensors were installed in shaded locations (i.e. under a tree) to avoid UV light from impacting readings. Locations with a well-defined channel were selected and sensors were attached to steel posts installed within the channel (at the lowest point in the streambed). Each sensor was encased in PVC piping and capped at the bottom for protection from UV and floating debris. The PVC pipe and cap had several drilled holes (20 mm in diameter) to allow for the free flow of air and water around the sensor. Site visits were conducted on an approximate monthly basis to collect data from the sensors, and maintain the site by clearing debris and silt when necessary. During these site visits additionally data was recorded, such as the state of the stream (i.e. 'wet' or 'dry') and landowners provided additional qualitative information about the site's recent flow behaviour.

The true state of the stream (or reference state used to evaluate the classified state) was determined using water level data for sites S001, S003 and S006. This approach to evaluation was defined as 'direct evaluation' (Table 2-2). The remaining sites were evaluated using the 'true' in-stream state that was indirectly determined using visual inspection of auxiliary data (defined as indirect evaluation). Auxiliary data used to determine the true state of a stream included: (i) the recorded state of the stream from sites visits; (ii) upstream and/or downstream flow data (which in some cases was located approximately 100m away) and (iii) visual inspection of temperature and rainfall data obtained from a gauge less than 5 km from the site (023799 Prospect Hill SA). The data collection period varies between 11 months (from March 2018 to February 2019) at sites S010 and S011 to 26 months (from December 2016 to February 2019) at sites S001 and S002. For the other sites S003, S004, S005, S006 and S008 between 23 and 24 months of data (Table 2-2).

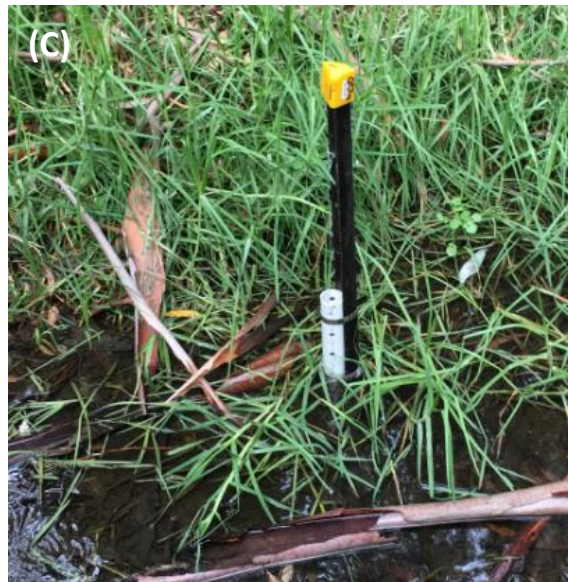
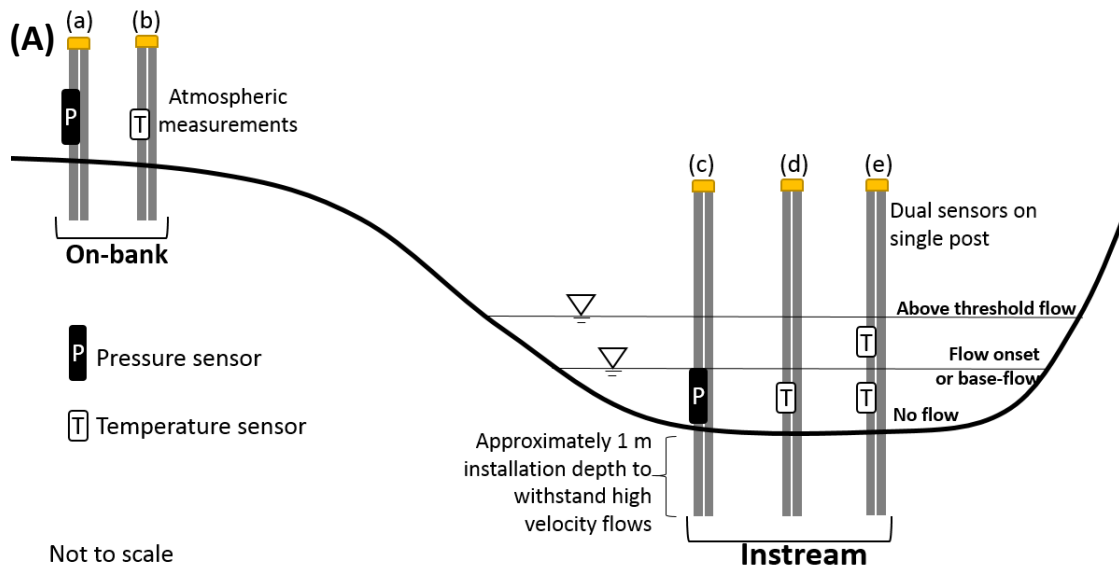


Figure 2-2: (A) Different configurations were explored in this study to evaluate the effectiveness of temperature sensors to infer intermittency. Four types of data were collected: in-stream and on-bank pressure sensors used to validate results from the paired in-stream and on-bank temperature measurements. Configuration (e) was used at site S004 where the flow onset threshold was defined as the base-flow threshold (see Table 2). Photos from the outlet (S006) showing installed (B) on-bank and (C) in-stream sensors.

Table 2-2: Site list and descriptions as well the length of data and the split used for calibration and evaluation of the classification method and method the classification were evaluated (indirectly or directly) is presented. The sensor configuration a-f refers to the sensors shown in Figure 2-2.

	Site	Location	Sensor depth above streambed	Data length and split		Evaluation method	Sensor configuration				
				Calibration	Evaluation		(a)	(b)	(c)	(d)	(e)
1	S001	Downstream Tributary 1	0.10 m	19 Dec '16-19 Dec '17	20 Dec '18-20 Feb '19	Direct	✓	✓	✓	✓	
2	S002	Downstream Tributary 1	0.15 m	19 Dec '16-19 Dec '17	20 Dec '18-20 Feb '19	Indirect		✓		✓	
3	S003	Tributary 1 and 2 outlet	0.10 m	1 Mar '17-1 Mar '18	2 Mar '17-20 Feb '19	Direct		✓	✓	✓	
4	S004	Upstream Tributary 1	0.15 m	22 Mar '17-22 Mar'18	23 Mar'18-18 Mar'19	Indirect		✓			✓
5	S005	Upstream Tributary 2	0.05 m	28 Feb'17-28 Feb'18	1 Mar'18-20 Feb'19	Indirect		✓		✓	
6	S006	Tributary 1 and 3 outlet	0.05 m	15 Mar'17-15 Mar'18	16 Mar'18-3 May'19	Direct		✓	✓	✓	
7	S008	Upstream Tributary 3	0.1 m	21 Mar'17-21 Mar'18	22 Mar'18-20 Feb'19	Indirect		✓			✓
8	S010	Downstream Tributary 1	0.05 m	1 Mar'18-20 Feb'19	NA	Indirect		✓		✓	
9	S011	Downstream Tributary 3	0.05 m	1 Mar'18-20 Feb'19	NA	Indirect		✓			✓

### 2.2.3 Characteristics of in-stream temperature data for intermittency classification

The difference in temperature thermal profiles have previously been used to delineate flow with streambed thermographs (Blasch *et al.*, 2004; Sowder and Steel, 2012; Arismendi *et al.*, 2017). There were a number of temperature-based characteristics that can be used to inform the state of a stream as either 'wet' or 'dry' and were used as the conceptual basis of the proposed classification algorithm. Periods of streamflow and zero-flow can be identified visually in a number of temperature time series when comparing flowing water (i.e. in-stream) and air (i.e. on-bank) temperature measurements (Figure 2-3). The temperature data when flow was present in the channel illustrate: (a) differences in the daily range of temperatures, (b) differences in the mean temperature and (c) lagged peak values and smoother fluctuations (Figure 2-3). A distinct feature was the reduced diurnal temperature variations of water compared to air (Figure 2-3a). For this study site, the in-stream temperature range was approximately 2 °C while the air temperature range was 15 °C—more than seven times larger. This feature was observed in winter flowing periods (July-October) and for this site the average ratio of in-stream temperature (when flow was present) to on-bank temperature range was 1:8.

While changes in the range of the in-stream temperature were the strongest feature, differences in the mean can also be observed. Figure 2-3b provides an example where the mean air temperature was approximately 7.5 °C whereas the mean stream temperature was approximately 11 °C, because bodies of water can take longer to cool down in response to changing weather conditions (Webb and Zhang, 1999). Additionally, the daily peak temperature for a sensor submerged in water can present a delay compared to air temperature (Figure 2-3a) due to the

specific heat of water being greater than air (Halliday and Resnick, 2013). Figure 2-3c also illustrates that the temperature sensor submerged in water had fewer and smaller temperature fluctuations due to flow dampening the atmospheric variations relative to the on-bank air temperature readings. Figure 2-4 summarises these observations via a comparison to water levels, showing that at times that streamflow was present (Figure 2-4a, highlighted with the grey shading based on information from the co-located pressure sensor), the diurnal range of in-stream temperature was dampened (Figure 2-4b).

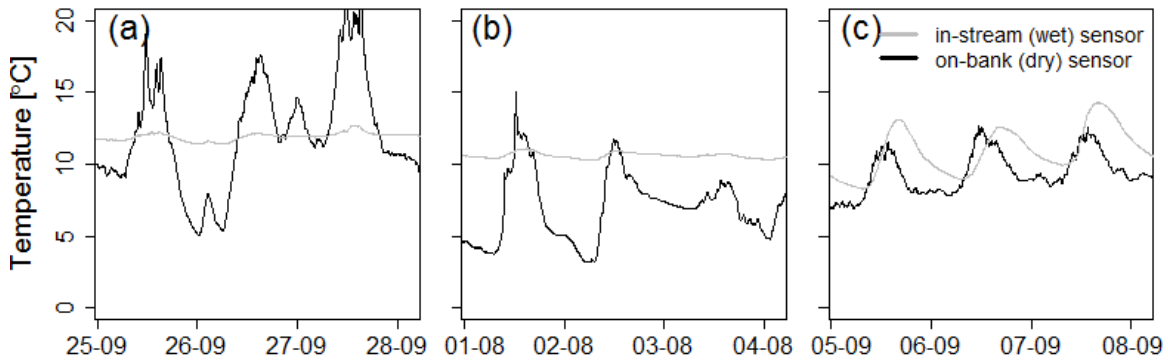


Figure 2-3: Example characteristics of paired in-stream and on-bank temperature profiles. (a) A similar mean temperature but where variation of the in-stream temperature was significantly lower, 2°C, compared to the on-bank temperature range of 20°C; (b) a difference in the mean between in-stream and on-bank profiles; and (c) similar temperature range but where the in-stream (wet) profile was lagged.

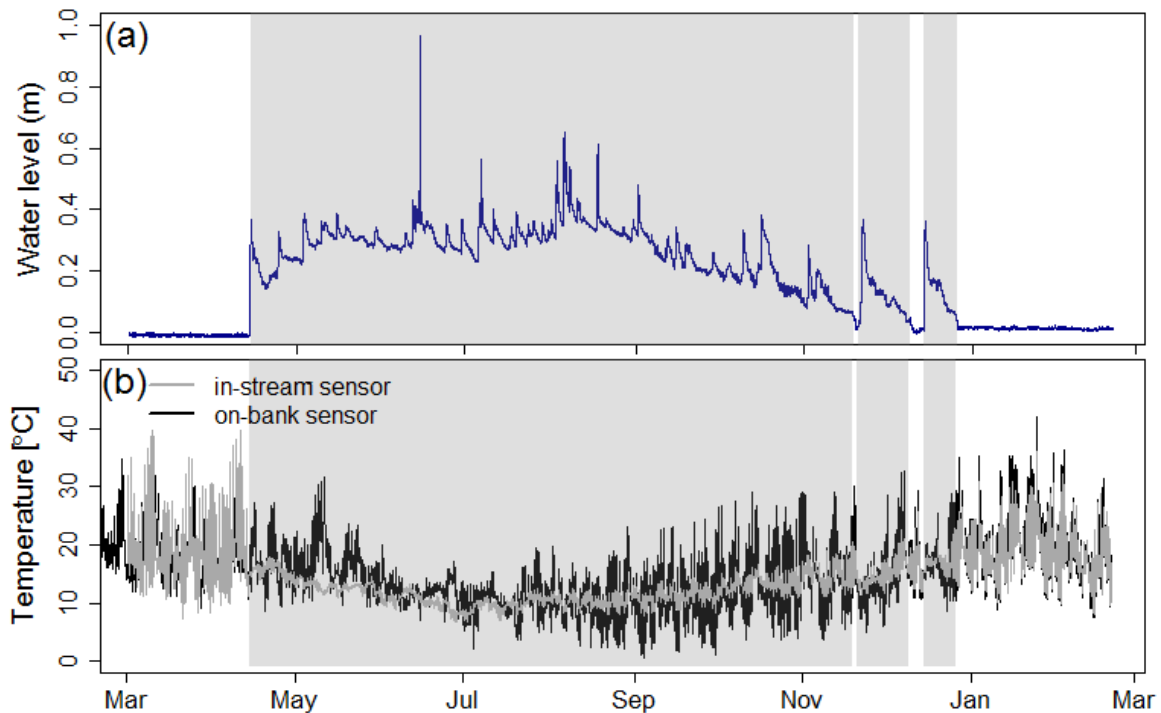


Figure 2-4: Example of time series of (a) water level at an intermittent site used as reference data for the true stream state and (b) in-stream and on-bank temperatures.

### 2.3 Methods of classifying streamflow and intermittency signature

To assess the effectiveness of low-cost temperature sensors deployed in a Mediterranean environment in their ability to detect streamflow, two classification methods of inferring streamflow intermittency were compared. Table 2-3 illustrates the approaches: (1) the standard deviation method, where a threshold deviation was applied (*Constantz et al., 2001; Blasch et al., 2004*) and; (2) the two-state hidden Markov model (HMM), used as a signal detection method (*Arismendi et al., 2017*). To ensure consistency between the approaches, a daily timestep was used and each of the methods were further developed to investigate whether improved performance could be achieved.

The standard deviation method was assessed using two approaches; the first was to delineate events using the in-stream daily temperature variance (*Constantz et al., 2001; Blasch et al., 2004*) and was used as a benchmark for comparison purposes. The second approach, based on the HMM methodology, also included on-bank temperature information. The inclusion of on-bank temperature was able to reduce false 'wet' classifications due to cold atmospheric fluctuations that can mimic 'wet' in-stream temperature data. Importantly, this removes the requirement of calibrating a minimum flow duration parameter that had been applied for previous studies (*Blasch et al., 2004*). The approach involves calculating the ratio of on-bank to in-stream daily temperature variance, and a threshold of this ratio was selected to obtain maximum classification performance when compared to the true state of the stream.

The effectiveness of the HMM method was determined by investigating six alternative algorithm inputs (e.g. single and multiple inputs) (Table 2-3). A benchmark was established by using the singular input of the in-stream daily temperature variance, as previously presented by *Arismendi et al., 2017*. The variables considered in the algorithm were selected based on thermal characteristics outlined in Section 2.2.3 and findings from previous studies (*Constantz et al., 2001; Blasch et al., 2004; Arismendi et al., 2017*). It had been previously shown that in some cases the HMM identified multiple (more than seven) state changes over a time period where the state was constant (*Arismendi, et al. 2017*). In order to minimize this error, 30-day antecedent daily rainfall data was collected from a local weather station to provide the algorithm with additional seasonal statistical characteristics of streamflow persistence. The maximum number of inputs implemented in any single implementation of the algorithm was limited to three to ensure model parsimony, and the in-stream daily variance was included for all iterations as it was the strongest signal identified in the data. In total there were five possible observation inputs: (1) in-stream temperature variance; (2) ratio of on-bank to in-stream temperature variance; (3) difference of on-bank to in-stream daily temperature mean; (4) in-stream to on-bank temperature profile lag; and (5) 30-day antecedent rainfall. Table 2-3 shows eleven combinations of these inputs that were considered.



Table 2-3: List of the inputs iterated for the two classification methods, where the application of in-stream temperature variance alone provides a baseline performance accuracy for comparison of developed methods

			Combination of observation inputs, $O$ , used in algorithm					
			Daily in-stream temperature variance	Daily rainfall antecedent moisture	Difference in daily in-stream to on-bank temperature mean	Lag of in-stream to on-bank temperature profile	$F^{(t)}$ : Ratio of on-bank to in-stream daily temperature variance	
Classification approach	Method of application	Iteration for method						
Standard deviation method	Calibration of threshold to delineation 'wet' (flow) and 'dry' (no-flow) periods	1 (baseline)	✓					
		2					✓	
Two-state hidden Markov model (HMM)	Unsupervised application of algorithm to determine the statistical parameters of the input data to categorize observations into two states (i.e. 'wet' or 'dry').	1 (baseline)	✓					
		2	✓	✓				
		3	✓		✓			
		4	✓				✓	
		5	✓					✓
		6	✓	✓	✓			
		7	✓	✓			✓	
		8	✓	✓				✓
		9	✓		✓		✓	
		10	✓		✓			✓
		11	✓				✓	✓

The collected observation inputs were split into training data, defined as the first year of data, and testing data, defined as the remaining data (March 2018 to March 2019), additional detail provided in Section 2.3.3. The effectiveness of each method was determined by calculating three performance measures: accuracy in binary classifications, comparison of number flow transitions (e.g. ‘wet’ to ‘dry’ and ‘dry’ to ‘wet’) and the comparison in the number of ‘wet’ days. Accuracy in binary classifications was defined as the number of times the daily classification matches the reference state of the stream divided by the number of daily time steps, i.e.  $T=365$  days. The comparisons on the number of ‘wet’ days was determined as an error where the percentage difference of the classified number of ‘wet’ days to calculated against the true number of ‘wet’ days.

### 2.3.1 Standard deviation method for intermittency classification

The standard deviation method of stream intermittency classification had been previously applied (Constantz *et al.* 2001; Blasch *et al.*, 2004). Using a daily timestep, the approach determined the in-stream temperature standard deviation,  $\sigma_s(t)$ , which was calculated as:

$$\sigma_s(t) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{s_i}(t) - \bar{x}_s(t))^2} \quad (1)$$

where each day  $t$  has  $i = 1, \dots, n$  measurements,  $x_{s_i}(t)$  is the  $i^{\text{th}}$  measurements of the in-stream temperature on day  $t$ , and  $\bar{x}_s(t)$  is the daily mean in-stream temperature. For 15-minute samples, there were  $n = 96$  measurements per day. Rather than solely rely on in-stream variation, the method is extended here to allow for the ratio of in-stream to on-bank variance at each timestep, referred to here as  $F(t)$ . The  $F(t)$  ratio is defined as the ratio of on-bank to in-stream variance and can be calculated as:

$$F(t) = \frac{\sum_{i=1}^n [x_{b_i}(t) - \bar{x}_b(t)]^2}{\sum_{i=1}^n [x_{s_i}(t) - \bar{x}_s(t)]^2} \quad (2)$$

where  $x_{b_i}(t)$  is the  $i^{\text{th}}$  measurement of the on-bank temperature on day  $t$ , and  $\bar{x}_b(t) = \frac{1}{n} \sum_{i=1}^n x_{b_i}(t)$  is the daily mean on-bank temperature.

To identify periods of streamflow and no-flow, a deviation threshold parameter was required to determine the daily stream state. The threshold parameter,  $\tau_\sigma$ , was defined as the magnitude of the standard deviation separating flowing and non-flow periods. The state of the streambed,  $S(t)$  was determined as:

$$S(t) \sim \begin{cases} S_t = W & \sigma_s(t) \leq \tau_\sigma \\ S_t = D & \sigma_s(t) > \tau_\sigma \end{cases} \quad (3)$$

Where a ‘wet’ (flow) state is  $\sigma_s(t) \leq \tau$  and was given a value of 1 and a ‘dry’ (no-flow) is when  $\sigma_s(t) > \tau$  and was given a value of 0.

The  $F(t)$  threshold parameter,  $\tau_F$ , was determined by:

$$S(t) \sim \begin{cases} S_t = D & \sigma_s(t) \leq \tau_F \\ S_t = W & \sigma_s(t) > \tau_F \end{cases} \quad (4)$$

where a ‘dry’ (no-flow) state is  $\sigma_{s_i}(t) \leq \tau$  and was given a value of 0 and a ‘wet’ (flow) is when  $\sigma_{s_i}(t) > \tau$  and was given a value of 1. Similarly, the threshold parameter was calibrated for the first 12 months of data at each site and was optimized to maximize the accuracy of binary classifications. Minimum flow duration and minimum inter-event duration parameters were not applied here, because of the daily scale of classifications, with single day events and inter-event durations possible for this case study, particularly during seasonal transitions.

### 2.3.2 Two-state hidden Markov model for intermittency classification

A two-state hidden Markov model (HMM) was applied to classify streamflow intermittency. The algorithm has the ability to handle multiple data inputs, including temperature and climate measurements, to determine the state of the stream as either ‘wet’ or ‘dry’. The advantage of the approach is that it is an unsupervised method, where temperature data can be classified without the requirement of optimizing the algorithm parameters to the observed stream state. This is a significant contrast to the standard deviation method which requires the calibration of one or more parameters using the ‘true’ state of the stream (Constantz et al., 2001; Blasch et al., 2004). This advantage allows low-cost temperature sensors to be deployed across a catchment without the requirement of calibration data (e.g. water levels).

Figure 2-5 illustrates the applied method of intermittency classifications. The observed data,  $O$ , consists of on-bank and in-stream temperature measurements and daily precipitation data collected at a local weather station. The data was split into training data (the first year of data), and testing data (the remaining data).

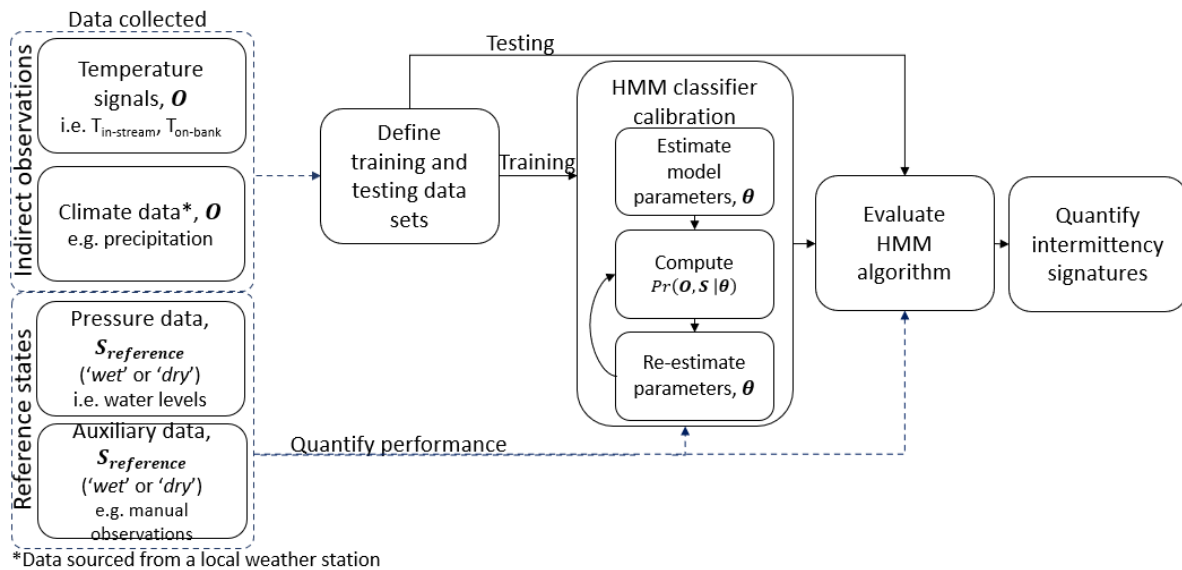


Figure 2-5: Schematic illustration of the method applied for the classification of intermittency using a hidden Markov algorithm.  $O$  are the observation variables,  $S$  are the unobserved states of the stream and  $S_{reference}$  are the observed reference states used to evaluate the performance of the algorithm.

The two-state hidden Markov model (HMM) is an unsupervised classification approach that uses multiple inputs. The R function depmixS4 was the primary means of method application (Visser & Speekenbrink, 2010). Relevant daily statistics of observed data are denoted as  $O = \{O_{j,t}, j = 1, \dots, M, t = 1, \dots, T\}$  for  $M$  input dimensions and  $T$  time-steps. In the following, the vector  $O_t$  is used as a

shorthand for  $\{O_{j,t}, j = 1, \dots, M\}$ . The state of the stream is the vector  $\mathbf{S} = \{S_t, t = 1, \dots, T\}$  with two possible states,  $S_t \in \{S_1, S_2\}$ . While the HMM delineates the observed inputs (listed in Table 2-3) into two states, the algorithm itself was ignorant of what ‘wet’ and ‘dry’ means. An additional step to determine which of the two states correspond to ‘wet’ and ‘dry’ was required to obtain  $\{W, D\}$ , where  $W$  was defined as a ‘wet’ state when the probability of it being ‘wet’ was greater than or equal to 0.50 and  $D$  was defined as the ‘dry’ state when the probability of it being ‘wet’ was less than 0.50.

The hidden state of the streambed was assumed to follow a first-order Markov model describing a sequence of states in which the probability of each state depends on the previous state. The transition probabilities  $p_{DW}$  and  $p_{WD}$  govern the transition between states, and by corollary, the persistence in each state. The distribution of observations in a given state was assumed to be from an  $M$ -dimensional Gaussian distribution:

$$O_t \sim \begin{cases} N_M(\boldsymbol{\mu}_W, \boldsymbol{\Sigma}_W) & S_t = W \\ N_M(\boldsymbol{\mu}_D, \boldsymbol{\Sigma}_D) & S_t = D \end{cases} \quad (5)$$

where  $\boldsymbol{\mu}_W$  and  $\boldsymbol{\mu}_D$  are the mean vectors and  $\boldsymbol{\Sigma}_W$  and  $\boldsymbol{\Sigma}_D$  are the covariance matrices for each state  $S_t$  at time-step  $t$ . The vector of unknown parameters is:

$$\boldsymbol{\theta} = \{\boldsymbol{\mu}_W, \boldsymbol{\Sigma}_W, \boldsymbol{\mu}_D, \boldsymbol{\Sigma}_D, p_{WD}, p_{DW}\} \quad (6)$$

which includes the vector of means and covariance matrix for the Gaussian distributions in the wet and dry states, as well as the two scalar transition probabilities. Given these parameters, the HMM defines the probability of the observations along with the hidden state variables as  $Pr(\mathbf{O}, \mathbf{S} | \boldsymbol{\theta})$ . The marginal likelihood of the observations was constructed using the forward-backward algorithm (Rabiner, 1989), and parameters were estimated using the expectation-maximization algorithm (Dempster et al., 1977). For this application, rather than the distribution of observed values, the main interest was in the inferred best estimate of state variables that were used to determine intermittency statistics.

In this study,  $M = 1$  to 3 inputs are considered and typically  $T = 365$  daily time-steps for calibration (except for S008 due to a short period of missing data). The inputs used for the HMM were selected based on physical and hydro-climatic characteristics of the site and insights from previous studies (Constantz et al., 2001; Blasch et al., 2004; Arismendi et al., 2017) in order to provide the strongest relationship to streamflow state (Table 2-3).

### 2.3.3 Calibration and evaluation of intermittency classification methods

Using a split sample approach, the standard deviation and the two-state HMM methods were calibrated at nine sites and evaluated at seven sites (see Table 2-2). The first 12 months of data was used for calibration and the remaining data (ranging from 11 to 14 months) was used to evaluate the predictions against the collected reference data. The two sites S010 and S011 installed during the second year of data collection were calibrated against 12 months of data.

For performance evaluation, reference data ( $S_{reference}$ ) were used to define the true state of the streambed. Reference data were determined in two ways: pressure data and auxiliary data. Pressure measurements were used to provide direct measurements of water levels at locations where available. Sites S001, S003 and S006 contained water levels and were directly evaluated using this approach. Auxiliary data were developed as an indirect method for determining the true state of the

streambed for remaining sites. The auxiliary estimates of streamflow were developed from multiple sources, including (1) records of stream state from monthly site visits; (2) qualitative information from landowners; (3) upstream and downstream water level data; and (4) visual inspection of temperature and rainfall data (see Section 2.2.2).

The results from the two-state HMM were defined such that if the probability of a 'wet' state from the HMM was greater than 0.5, then it was classified as 'wet'. An accurate prediction was if the classifier was equal to the observed reference state of the stream, where accuracy in binary classifications was defined as the sum of correctly classified states divide by the number of days in the period. The accuracy in binary classifications was further analysed by comparing the number of false 'wet' states and the number of false 'dry' states. Additionally, performance of the methods was evaluated by comparing the number of transitions over a period to the true number of transitions in terms of 'wet' to 'dry' and 'dry' to 'wet' as well as the error in the number of 'wet' days predicted compared the observed number of 'wet' days.

### 2.3.4 Quantification of intermittency signatures

Using the most accurate approach based on the methodology above, the intermittency was defined at each site and enables comparison to better understand catchment processes and heterogeneity. Specifically, seven intermittency signatures were calculated to determine streamflow intermittency at the nine monitored sites. The selected signatures use the binary data to characterize different aspects of streamflow intermittency. Here the applied intermittency signatures refer to 'wet'-'dry' sequencing with flow below any previously applied threshold defined as 'zero-flow' (Table 2-4).

Table 2-4: Definitions of intermittency signatures

Low flow signature	Definition
(1) Number of zero flow days	The annual average number of days per annum that there was 'zero-flow'.
(2) Average duration of zero flow periods	The annual average duration of 'zero-flow' periods. Calculated as the number of zero-flow days divided by the number of contiguous zero-flow periods
(3) Annual percentage of time flow was present continuous	The annual average percentage of days there was flow in a channel. Calculated as the number of flow days divided by 365.
(4) Percentage of time flow was present by season	The average percentage of days there was flow in a channel by season. Calculated as the number of flow days in a season divided by the total number of days in the season. Summer was defined between Dec-Feb, autumn between Mar-May, winter between Jun-Aug and spring between Sep-Nov.
(5) Timing of flow permanence	The Julian start and end date where flows were permanent or continuous in a channel, including short duration stops of less than 3 days.
(6) Flow variability	The coefficient of variation in flow occurrence, calculated as the standard deviation divided by the mean of the binary classifications (modified from <i>Jowett and Duncan, 1990</i> )
(7) 30-day antecedent rainfall to flow permanence	The cumulative rainfall occurring 30 days prior to the day which flow in a channel becomes permanent.

## 2.4 Results and discussion

### 2.4.1 Calibration and evaluation of the standard deviation method

The standard deviation method was assessed by calibrating a threshold deviation for daily temperature variance for the nine sites and evaluated at seven sites (evaluation data was not available for site S010 and S011). The temperature variation approach yielded an average calibration accuracy (fraction of time with the correct state) for the nine sites of 90% (ranging from 68% to 97%). The average accuracy for the evaluation period was 79% (ranging from 52% to 99%). The threshold deviation varied across the sites from 0.35°C to 5.07°C (see supplementary information, Table 2-10). The standard deviation method where a threshold deviation was applied to the  $F(t)$  ratio showed an average calibration accuracy of 91% (ranging from 77% to 98%) and an average evaluation accuracy of 78% (ranging from 50% to 98%). The performance of this approach, in terms of average accuracy was similar to the baselined variance approach. The threshold applied to  $F(t)$  across the sites ranged between 2.5 and 11.0 (see supplementary information, Table 2-11).

The results show, given the simple nature of the method, both approaches perform well under calibration. However, both standard deviation methods (variance and  $F(t)$  ratio) showed a deterioration in performance over the evaluation period of 11% and 13% respectively. Figure 2-6 illustrates an example of performance variations for the two approaches from 70% to 77% for the standard deviation method and from 82% to 67% for the  $F(t)$  ratio approach. The performance at a number of sites was consistently good, but the deterioration in performance at other sites shows a lack of robustness in the method.

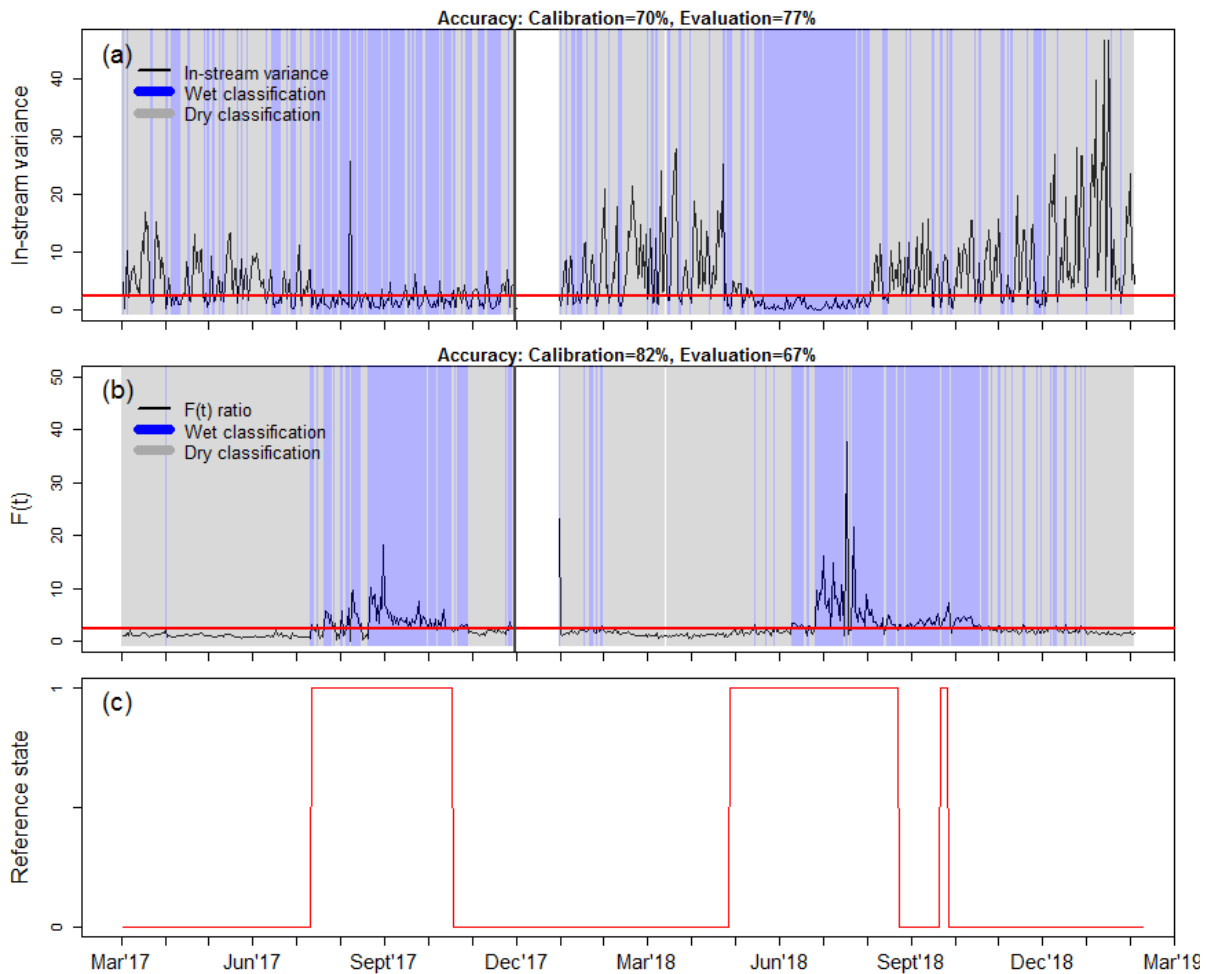


Figure 2-6: Illustration of the standard deviation method for site S008 where: (a) the in-stream daily temperature variance is used to delineate 'wet' and 'dry' periods by applying a threshold deviation shown with the horizontal red line, where variance values below the threshold were classified as 'wet' and above the threshold classified as 'dry'; (b) the ratio of on-bank to in-stream daily temperature variance with the threshold deviation difference shown with the horizontal grey line, where  $F(t)$  values above the threshold were classified as 'wet' and below the threshold 'dry'; (c) illustrates the reference (or observed) state of the stream where, 'wet'=1 and 'dry'=0.

To further investigate the misclassification of days for site S008, Figure 2-7 illustrates an eight day period of flow (or 'wet' days) between 18<sup>th</sup> to 26<sup>th</sup> August. The figure shows that seven days for the variance approach and four days for the  $F(t)$  ratio approach were misclassified, where all days should be classified as 'wet' (Figure 2-7a and b). Figure 2-7c shows that this error can be attributed to the presence of alternative temperature signals that can indicate the presence of flow. For example, a profile lag was illustrated with the vertical dashed lines on day 19/09 in Figure 2-7c. While there was flow present in-stream, the daily in-stream peak was lagged approximately two hours after the daily on-bank daily temperature. However, given that the in-stream daily variance was large, this day had been classified as a false 'dry' for the standard deviation method using the in-stream variance. In contrast, the on-bank daily variance used to calculate the  $F(t)$  ratio was large enough to correctly identify this day as 'wet'.

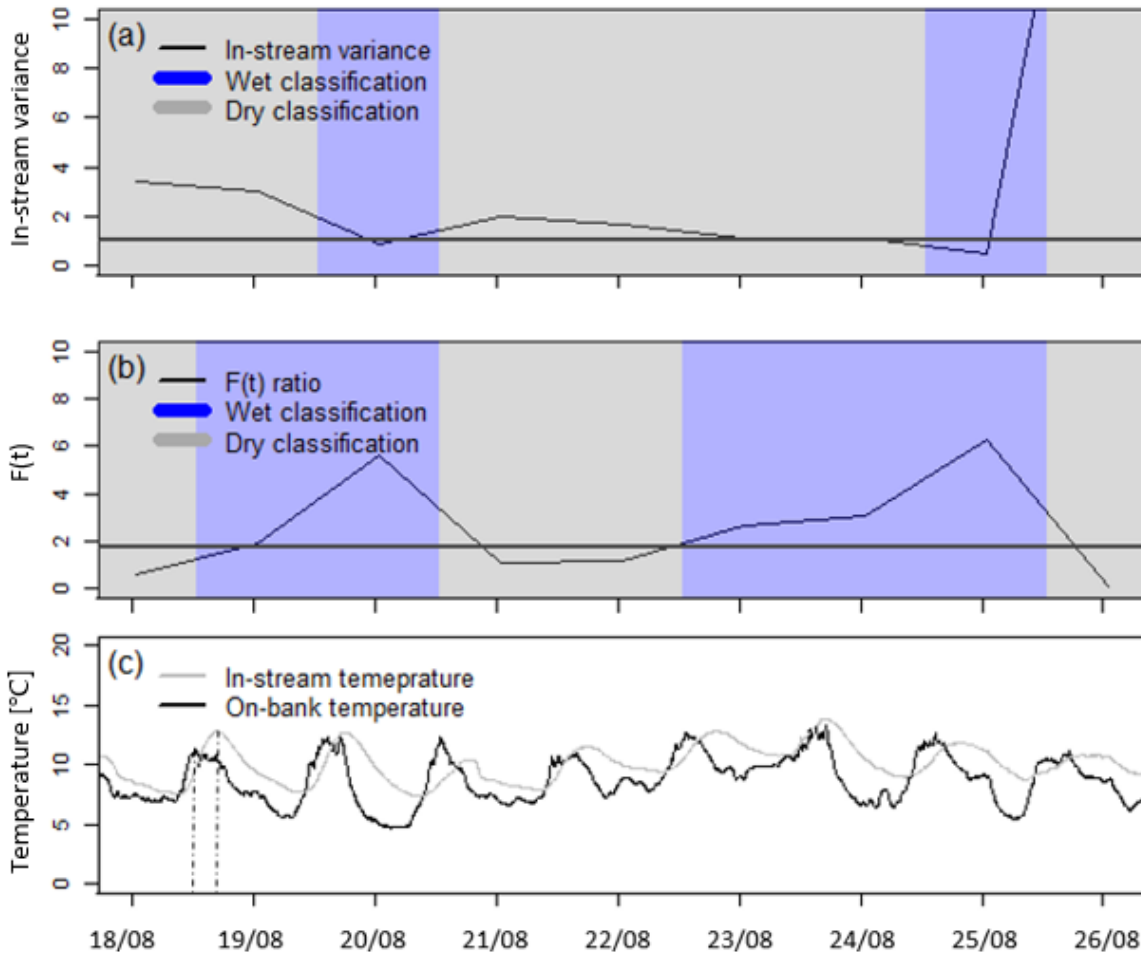


Figure 2-7: Showing site S008 over the 'wet' period of 18<sup>th</sup> to 26<sup>th</sup> August, where blue indicates a 'wet' classification and grey indicated a false 'dry' classification. The panels illustrate the misclassification of the standard deviation methods using the (a) variance and (b)  $F(t)$  ratio approaches compared to (c) temperature measurements. The dashed lines show an example of daily peak in-stream temperature lag compared to the on-bank temperature.

#### 2.4.2 Calibration and evaluation of the two-state hidden Markov model

A baseline for comparison was established by applying the HMM with a single input for in-stream temperature variance. The baseline average accuracy of binary classifications for the calibration and evaluation period was 83% and 81% respectively, a deterioration of compared to the simple standard deviation approach. The alternative combinations of inputs yielded a range of results, varying across sites as well as the type of input variables. For the best performing method, Table 2-5 shows the performance at each site along with the median performance across all sites. The median performance was used as the basis of discussion and shows, the highest performing approach resulted in a calibration and evaluation accuracy of 93% and 92% respectively, outperforming both benchmarked results. The observation variables for this implementation were: (1) the daily in-stream temperature variance; (2) the ratio of on-bank to in-stream temperature variance; and (3) 30-day antecedent rainfall. The error in the number of 'wet' days across all sites was relatively low, with a calibration and evaluation error of 10% and 5%, respectively. The number of false 'wet' and false 'dry' classifications were on average 3% and 4% respectively across the eleven approaches over the entire period. This results shows that the inclusion of alternative identified signals (Figure 2-3) did not to improve the average accuracy of 'wet' and 'dry' prediction for the two-state hidden Markov model when tested with the 11 alternative combinations (see supplementary material).



The number of predicted transitions compared to the observed transitions performed well across the sites, both during training and evaluation with a 22 to 14 transition comparison during calibration and an 8 to 8 transition comparison during evaluation (Table 2-5). The exception was site S001, where the probability of transitioning was significantly under-estimated by the HMM algorithm during the calibration period with a comparison of 40 to 9 transitions. It was also observed that sites S004 and S008 significantly under-estimate the number of predicted ‘wet’ days with a resulting error of -26% and -40% respectively during calibration.

Table 2-5: Results for the HMM with inputs (1) in-stream variance, (2) 30-day rainfall and (3)  $F(t)$  ratio.

Site	Calibration					Evaluation				
	Accuracy in binary classifications	False ‘wet’ classification error	False ‘dry’ classification error	Error in the number of ‘wet’ days	Number of transitions $T_{truth}:T_{inferred}$	Accuracy in binary classifications	False ‘wet’ classification error	False ‘dry’ classification error	Error in the number of ‘wet’ days	Number of transitions $T_{truth}:T_{inferred}$
S001	90%	5%	5%	5%	40:9	93%	2%	5%	-5%	12:8
S002	83%	7%	7%	-6%	38:49	89%	7%	4%	5%	18:11
S003	98%	0%	2%	-3%	1:5	94%	5%	1%	5%	10:13
S004	93%	0%	7%	-26%	14:8	91%	3%	5%	-3%	6:14
S005	98%	1%	1%	3%	10:10	99%	0%	1%	-7%	4:2
S006	99%	0%	1%	0%	5:6	92%	2%	7%	-11%	6:4
S008	83%	0%	17%	-40%	2:7	84%	15%	1%	45%	4:3
S010	99%	0%	1%	-2%	4:8					
S011	96%	0%	4%	-6%	8:10					
Median	93%	2%	5%	10%	22:14	92%	5%	3%	5%	8:8

Figure 2-8 illustrates how the addition of 30 day rainfall and the  $F(t)$  ratio improves the predictive capabilities of the two-state HMM. For site S001, the singular input variable, in-stream daily variance, results in a median accuracy of binary classifications of approximately 69%. The number of predicted transitions over the period was 131 compared to the observed 52. Similarly the error in the percentage of ‘wet’ flow days was significantly over-estimated with a calibration and evaluation error of 94% and 34% respectively (Figure 2-8a). Following the addition of the 30 day antecedent rainfall, the predictive performance of the HMM algorithm improves to 73%. The positive bias in classified transitions was also significantly improved reducing to 61 (previously 131 classified transitions). Significant improvement in predictive performance was identified with the addition of the  $F(t)$  ratio to the in-stream daily variation and the 30 day rainfall inputs. The error in predicted ‘wet’ days during calibration and evaluation performed well, with -10% and 4% respectively.

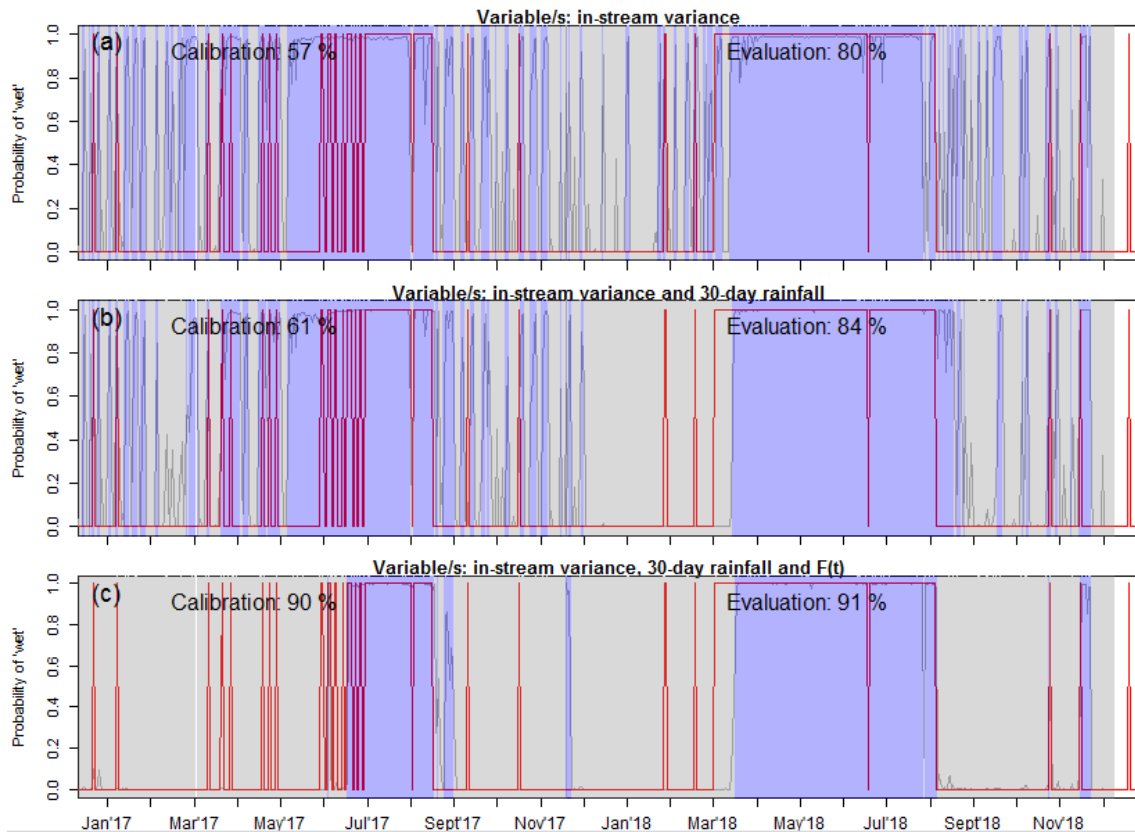


Figure 2-8: Shows the accuracy of the two-state HMM, with the red line showing the reference state and the grey line showing the state of algorithm as classified by the HMM for site S001. Additional observations were added to the algorithm; where the (a) daily in-stream temperature variance was applied, (b) daily in-stream temperature variance and 30-day antecedent rainfall was applied, and (c) daily in-stream temperature variance, 30-day antecedent rainfall and the  $F(t)$  ratio was applied.

### 2.4.3 Comparison of intermittency classification methods

A comparison of the standard deviation methods and the best performing two-state HMM method shows that one of the most significant improvements was the predicted number of transitions when applying the HMM (Table 2-6). That is, the average number of transitions observed was 30, while the standard deviation methods predict 85 and 44 transitions for the variance and  $F(t)$  ratio approach compared to 19 transitions for the HMM. The HMM maintained a good evaluation accuracy of 91% compared to the standard deviation approaches of 79% and 78% for the variance and  $F(t)$  ratio approach. The median accuracy of binary classifications achieved with the standard deviation method was 85% for the variance approach and 85% for the  $F(t)$  ratio approach. In contrast, the median accuracy achieved with the HMM was 92% [i.e.  $(93+91)/2$ ], an improvement of 7%, while also reducing the error in predicted number of 'wet' days by 10% and reducing the over-estimation of flow transitions across all sites.

Table 2-6: Comparison for the two classification methods assessed and observation inputs investigated

Classification Method	Approach	Calibration accuracy in binary classifications	Evaluation Accuracy in binary classifications	Calibration error in number of 'wet' days	Evaluation error in number of 'wet' days	Calibration error in the number of 'wet' days predicted	Evaluation error in the number of 'wet' days predicted	Calibration error in the number of 'dry' days predicted	Evaluation error in the number of 'dry' days predicted	Number of transitions $T_{truth} : T_{inferred}$
Standard deviation method	in-stream variance	90%	79%	15%	35%	4%	6%	8%	11%	30:85
	$F(t)$ ratio	91%	78%	17%	34%	4%	5%	11%	11%	30:44
Two-state hidden Markov model (HMM)	in-stream variance, 30-day rainfall, $F(t)$ ratio	93%	92%	5%	5%	1%	5%	5%	3%	30:19

#### 2.4.4 Calculated number of 'wet' flow days and intermittency signatures

Seven intermittency signatures were calculated for the individual sites to determine the statistical properties of low flow regimes within the catchment. The percentage of 'dry' and 'wet' days, determined by the average of 'wet' and 'dry' days over two years, shows the differing proportions of flow permanence (Table 2-7). The outlet, S006, had flow present for approximately 65% of the year, with locations upstream having different annual flow permanence. For example, site S005 flows 22% of the year (approximately three months), while site S008, which is cleared of all vegetation, flows for 33% of the year. Further downstream of S008, site S011 flows for 60% of the year demonstrating the variety of flow behaviour for the nine sites. This highlights that a single gauge at the outlet is not capable of effectively representing local-scale details within the catchment.

Table 2-7 illustrates that site S004 was flashier with an average zero flow duration of 21 days compared to site S001 with 47 days. The calculated signatures illustrate that sites S002 and S003 (tributary 1) dry out on average for 136 days and 47 days per year respectively. However, the pools do not remain dry for long, with average durations of dry spells being 11 days and 13 days respectively. This behaviour can be related to subsurface processes occurring within the sub-catchment (i.e. groundwater contributions at S001 and S004). In addition to this the contributing area of this sub-catchment is densely vegetated.

The intermittent site S005 (tributary 2) had an average 47 zero-flow-day duration with a majority of runoff occurring in the middle of winter and receding in spring. While the contributing area is densely vegetated, there was no evidence of groundwater processes contributing to flow during dry periods and was likely to be a contributing factor of the longer dry spell periods, contrasting tributary 1. Conversely while site S008 (tributary 3) also flows for 50% of the winter periods, it recedes quickly in early spring and was less flashy with an average 100 zero-flow-day duration. While similar to tributary 2, in that there was no evidence of subsurface flow processes contributing to flow during dry periods, the average period of dry days was doubled by comparison. This may be due to the contributing area being cleared of dense vegetation and containing grass, pastures and sparse woody trees. These results indicate that the presence of dense vegetation and subsurface processes may reduce the length of average zero-flow-day durations

Table 2-7: Annual intermittency signatures were calculated for each individual site of data collection

Site	Description of sub-catchments and local conditions							Intermittency signatures						
	Local description	Upstream contributing area (km <sup>2</sup> )	Average slope (m/m)	Contributing area with dense vegetation (%)	Channel width (cm)	Average in-stream temperature mean when flow present (°C)	Average in-stream temperature variance when flow present (°C)	Number of zero flow days	Average duration of zero flow periods	Percentage of ' wet' flow days per annum	Percentage of ' wet' days by season <sup>x</sup>	Time seasonal flow becomes permanent and recedes	Flow variability determined as the coefficient of covariance	30 day antecedent moisture to flow onset
<b>S001<sup>^</sup></b>	Sensor on reach 1. Site is clear of dense vegetation.	1.8	0.04	75%	95.0	10.9	1.1	250 days	47 days	31%	Su=5% Au=43% Wi=90% Sp=10%	2017: 19 Jun-10 Sept 2018: 2 Apr-4 Sept	1.37	2017: 54.8 mm 2018: 14.8 mm
<b>S002</b>	Sensor on reach 1. Site is clear of dense vegetation.	1.9	0.04	75%	64.0	11.3	0.63	177 days	11 days	51%	Su=23% Au=63% Wi=100% Sp=49%	2017: 15 Apr-29 Sept 2018: 1 Apr-19 Oct	0.79	2017: 17.5 mm 2018: 13.6 mm
<b>S003</b>	Sensors located on reach 1. Site is clear of dense vegetation.	4.4	0.036	70%	132.5	13.1	1.28	47days	13 days	87%	Su=31% Au=78% Wi=99% Sp=88%	2017: 16 Apr-29 Sept 12 Oct-3 Jan'18 2018: 15 Apr-23 Dec	0.57	2017:17.5 mm 2017:49.4 mm 2018:50.4 mm
<b>S004<sup>^</sup></b>	Sensors located upstream of Reach 1. Site is densely vegetated channel	0.9	0.05	80%	31.5	10.9	2.72	182 days	21 days	50%	Su=0% Au=6% Wi=85% Sp=21%	2017: 15 Jun-9 Sept 2018: 7 Jun-2 Sept	1.59	2017: 49.6 mm 2018: 48.2 mm

	contains rock and outcrop of rock observed at site.													
<b>S005</b>	Sensor upstream on tributary 2. Site is densely vegetated.	1.4	0.04	100%	36.5	10.9	0.19	285 days	47 days	21%	Su=1%, Au=0% Wi=47% Sp=40%	2017: 30 Jul-26 Oct 2018: 18 Jul-16 Sept	1.90	2017: 153.2 mm 2018: 55.1 mm
<b>S006</b>	The site is the catchment outlet with three tributaries upstream. Site is clear of dense vegetation.	8.6	0.036	40%	251.8	12.4	2.13	143 days	14 days	65%	Su=3% Au=70% Wi=100% Sp=83%	2017: 10 Apr-6 Nov 2018: 15 Apr-19 Nov	0.73	2017: 25.1 mm 2018: 50.4 mm
<b>S008</b>	Site is upstream reach 3, Site is clear of dense vegetation. Channel is deeply incised.	2.6	0.02	20%	118.6	11.5	1.99	244 days	81 days	33%	Su=0% Au=0% Wi=50% Sp=19%	2017: 19 Jul-2 Oct 2018: 14 Jul-3 Sept	2.26	2017: 131.6 mm 2018: 81.3 mm
<b>S010 *</b>	Sensors located downstream reach 1. Site is clear of dense vegetation.	4.1	0.036	20%	174.0	12.4	0.73	107 days	27 days	70%	Su=36%, Au=56% Wi=100% Sp=99%	2018: 14 Apr-12 Dec	0.59	2018: 33.6 mm
<b>S011 *</b>	Sensors located downstream reach 3. Site is clear of dense vegetation.	4.5	0.02	70%	75.0	12.1	1.48	145 days	18 days	60%	Su=14% Au=35% Wi=100% Sp=87%	2018: 4 May-9 Nov	0.81	2018: 68.4 mm

\* 12 months of data from the second year used to calculate signatures

^ All signatures calculated as flow response above a continuous flow threshold

\* Breakdown of the percentage of 'wet' days by: Su = summer, Au = Autumn, Wi = Winter and Sp= Spring.

Figure 2-9 shows the spatially variable flow behaviour and complexity of flow regimes across the catchment area. The binary classifications illustrate that some sites transitioned to ‘wet’ or ‘dry’ states at the same or similar times. In contrast, other sites shifted to a ‘wet’ state later and a dry state sooner (see also Table 2-7). As previously discussed, reach 1 (i.e. sites S001 and S004) was fed by continuous base-flow year round, while downstream flow had seasonal behaviour (i.e. S010 and the outlet S006). The study shows that given the relatively small sub-catchment area, the streamflow across the landscape was heterogeneous with detailed seasonal and annual flow patterns being illustrated based on the classified temperature data.

The binary classification of stream-flow intermittency has been shown to be useful for describing important hydrological features and ecological responses, such as, timing, duration and frequency of drying within the stream network. The unique signatures at the individual sites have a number applications, such as: (i) evaluation or improved calibration of models, (ii) informing ecological management of river networks at the reach scale; (iii) applying regressions to extrapolate the probability of reaches drying and wetting; or (iv) evaluating local scale risks of drought and climate change impacts where studies are typically based on with outlet data.

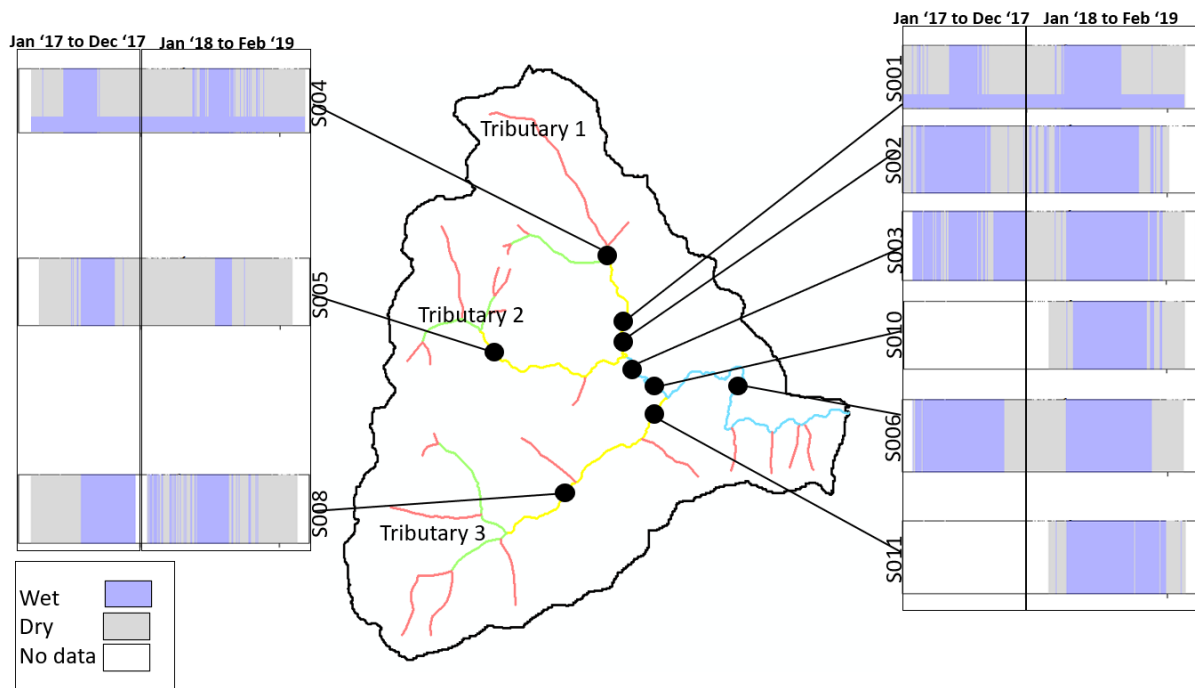


Figure 2-9: Map of the study catchment, with three main tributaries, tributary 3 which is largely un-vegetated, with nine sites where data was collected for a periods of 11 to 26 months. The bar at each site shows the inferred intermittency for each site. The wet panel for sites S001 and S004 shows the presence of base-flow, with the sensor was installed and classified above the base-flow.

## 2.5 Conclusions

This research demonstrated that temperature sensors can be successfully used to infer streamflow intermittency by comparing alternative methods of delineating flow permanence within a headwater stream network. The advantage being, the method was low-cost and unsupervised, was reliable under evaluation, had an accuracy of over 90% and was able to provide detailed intermittency signatures at the local scale. The study compared the standard deviation and two-state HMM methods with a total of 13 variations investigated which depended on input parameters into each approach. The best performing model over an independent evaluation period that was not used for parameter estimation

was the HMM model using inputs: (1) the daily in-stream temperature variance; (2) the ratio of on-bank to in-stream temperature variance; and (3) 30-day antecedent rainfall. The algorithm quantitatively determined 'wet' (or 'flowing') and 'dry' (or 'no-flow') transitions for the paired in-stream and on-bank sub-daily temperature readings with achieve an average accuracy of 91% in evaluation (93% in calibration), an improvement of 7% compared to the standard deviation approach. The daily temperature variance was the strongest temperature signal and due to differences across the sites (e.g. slope, vegetation cover and groundwater contributions), other temperature characteristics were identified such as differences in daily mean temperature and lag to daily peak (Figure 2-3). The study showed that the inclusion of alternative identified signals did not to improve the accuracy of 'wet' and 'dry' prediction for the two-state hidden Markov model when tested with the 11 alternative combinations. The approach was used to quantify intermittent signatures across all the sites and demonstrated a high degree of heterogeneity in the headwater catchment.

Although methods based on temperature sensing provide significant additional information relative to a single downstream streamflow gauge, they have a number of limitations. When a sensor touches the streambed, the moisture may persist and cause inference, leading to false 'wet' day classifications. On average false 'wet' classifications for the standard deviation and the HMM methods accounted for 5%, and 3% of the error, respectively. Alternatively, when placed marginally above the streambed, a trickle of very low flows can pass under the sensor. When the sensor was not touching the base, sedimentation may cause rivulets at very low flows to bypass flow around the sensor, or equally, may cause build-up so that the sensor was buried and has spurious readings. This was observed at site S008, where high flows and a deep channel resulted in in-stream temperature sensors being buried under sediment (observed 19<sup>th</sup> July 2018, see supplementary material for photos). For these reasons, the sensors do not detect truly 'zero' flow conditions, but represent the intermittency of the stream above some very minor base threshold. This was also the case when the sensor was placed underwater in a standing pool – even though the pool may rarely dry out, it was possible to use temperature variation to infer the onset of flow – and the classification of 'wet' flowing and 'dry' not-flowing was not identical to whether remnant water was present at the site. As a result of these differences in site conditions, each location will have unique diurnal temperature patterns. Other practical issues to consider in locating sensors include the accessibility of the site, strength of the stake and foundation on which it was posted, avoiding direct sunlight (Sowder & Steel, 2012) having similar shade conditions for the on-bank and in-stream temperature sensors and avoiding significant vegetation which may cause the build-up of debris and shelter a sensor.

This paper demonstrated a low-cost monitoring technique that can be used to complement other high quality data sources (e.g. streamflow). The advantage being that, the technique has the ability to provide higher resolution spatial representation of headwater runoff behaviour. Data of this type has the potential to improve methods of model calibration and evaluation for temperate catchments. In particular the representation of upstream processes in physical models that are commonly difficult to calibrate and evaluate due to their high data requirements can be addressed and is applicable for alternative catchments. The calculated intermittency signatures for each site illustrate highly variable intermittent patterns within the small study area, ranging between highly ephemeral, seasonal to continuous annual base-flow. It was clear that there was significant variability of headwater catchments, even over small scales (Figure 2-9). With ongoing development of low-cost sensor technologies, it is likely that the use of this information as complementary data to traditional streamflow gauging's provides a valuable opportunity to improve hydrological understanding in headwater catchments.

## Acknowledgements

We would like to acknowledge the South Australian Government, Department for Environment and Water (DEW) for their assistance in selecting a study location and the development of sensor deployment methods. We are grateful to Joshua Larson for initial conversations on detecting stream state and sensor technologies. Additionally, we would like to thank the landowners who participated in this study for their co-operation in providing us with site access for the installation and collection of data.



## References

- Acuña, V., Datry, T., Marshall, J., Barcelo, D., Dahm, C.N, Ginebreda, A., G. McGregor, S. Sabater, K. Tockner, M.A. Palmer. 2014. Why should we care about temporary waterways? *Science* 2014, 343, 1080–1081. 10.1126/science.1246666
- Alsdorf, D.E., Rodriguez, E., Lettenmaier, D.P. 2007. Measuring surface water from space. *Rev. Geophys.*, 45 (2007), pp. 1-24, 10.1029/2006RG000197.1
- Arismendi, I., Dunham, J. B., Heck, M. P., Schultz, L. D., Hockman-Wert, D. 2017. A statistical method to predict flow permanence in dryland streams from time series of stream temperature. *Water* 2017, 9, 946. 10.3390/w9120946.
- Arthington, A. H., Bernardo, J. M., Ilheu, M. 2014. Temporary rivers: linking ecohydrology, ecological quality and reconciliation ecology. *River Res. And App.* 30: 1209-1215.
- Blasch, K. W., Ferre, T. P. A., Christensen, A. H., Hoffmann, J. P. 2002. New field method to determine streamflow timing using electrical resistance sensors. *Vadose Zone Journal.* 1, 289-299.
- Blasch, K. W., Ferre, T. P. A., Hoffmann, J. P. A statistical technique for interpreting streamflow timing using streambed sediment thermographs. *Vadose Zone Journal.* 3, 936-946.
- Blume, T., van Meerveld, I., Weiler, M. 2017. The role of experimental work in hydrological sciences – insights from a community survey. *Hydrological Sciences Journal*, 62 (3), 334–337
- Boulton, A., Rolls, R., Jaeger, K., Datry, T. 2017. Hydrological Connectivity in Intermittent Rivers and Ephemeral Streams. Intermittent Rivers and Ephemeral Streams: *Ecology and Management.* doi:10.1016/B978-0-12-803835-2.00004-8
- Bradley, A. A., Kruger, A., Meselhe, E. A., Muste, M. V. I. 2002. Flow measurement in streams using video imagery, *Water Resour. Res.*, 38(12), 1315, doi:10.1029/2002WR001317
- Carling, G.T., Mayo, A. L., Tingey, D., Brunthans, J. 2012. Mechanisms, timing and rates of arid region mountain front recharge. *J. Hydrol*, 428-429: 15-31. doi.org/10.1016/j.jhydrol.2011.12.043
- Caruso, B.S., Haynes. J. 2011. Biophysical-regulatory classification and profiling of streams across management units and ecoregions. *J. Am. Water Resour. Assoc.* 447, 386-407.
- Chapin, T.P., Tood, A.S., Zeigler, M.P. 2014. Robust, low-cost data loggers for stream temperature, flow intermittency, and relative conductivity monitoring. *Water Resour. Res.* 50, 6542-6548.
- Constantz, J., Stonestorm, D., Stewart, A.E., Niswonger, R., Smith, T.R. 2001. Analysis of streambed temperatures in ephemeral channels to determine streamflow frequency and duration. *Water Resour. Res.* 37, 317-328.
- Datry, T., Pella, H., Leigh, C., Bonada, N., Hugueny, B. 2016. A landscape approach to advance intermittent river ecology. *Freshw. Biol.* 61, 1200-1213.
- David, J. C., Rutten, M. M., Pandey, A., Devkota, N., van Oyen, W. D., Prajapati, R., van de Giesen, N. 2018. Citizen science flow – an assessment of citizen science streamflow measurement methods. *Hydrol. Earth Syst. Sci.* doi.org/10.5194/hess-2018-425.

Dell, A.I., Alford, R.A., Pearson, R.G. 2014. Intermittent Pool Beds Are Permanent Cyclic Habitats with Distinct Wet, Moist and Dry Phases. *PLoS ONE* 9(9): e108203. doi:10.1371/journal.pone.0108203

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum Likelihood for Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B.* 39(1): 1-38

Dugan, H. A., Lamoureux, S. F., Lafrenière, M. J., Lewis, T. 2009. Hydrological and sediment yield response to summer rainfall in a small high arctic watershed. *Hydrological Processes.* 23: 10. 1514-1526.

Engeland, K., Hisdal, H. 2009. A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model. *Water Res. Mgmt.* 23: 12. 2567-2586.

Gallart, F., Llorens, P., Latron, J., Cid, N., Rieradevall, M., Prat, N. 2016. Validating alternative methodologies to estimate regime of temperate rivers when flow data are unavailable. *Sci. Total Environ.* 565, 1001-1010.

Godsey S.E. & Kirchner J.W. 2014. Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes,* 28, 5791– 5803.

Gomi, T., Sidle, R. C., Richardson, J. S. 2002. Understanding processes and downstream linkages of Headwater systems: Headwaters differ from downstream reaches by their close coupling to hillslope processes, more temporal and spatial variation, and their need for different means of protection from land use. *BioScience.* 52: 10, 905-916.

Goulsbra, C. S., Lindsay, J. B., Evans, M. G. 2009. A new approach to the application of electrical resistance sensors to measuring the onset of ephemeral streamflow in wetland environments. *Water. Res. Research.* 45: W09501.

Gungle, B. 2007. Timing and duration of flow in ephemeral streams of the Sierra Vista subwatershed of the Upper San Pedr Basin, Cochise County, southeastern Arizona: U.S. Geological Survey Scientific Investigations Report 2005-519. Pubs.usgs.gov/sir/2005/5190/.

Halliday, D., Resnick, R., Walker, J. 2005. *Fundamentals of Physics.* 7<sup>th</sup> Edition. Wiley.

Hansen, W.F. 2001. Identifying stream types and management implications. *For Ecol. Manage.* 143, 39-46.

Hofer, T., Gruber, G., Montserrat, A. Gamerith, V. 2014. A robust and accurate surrogate method for monitoring the frequency and duration of combined sewer overflows. *Env. Monitoring & Assessment.* 190(4). 10.1007/s10661-018-6589-3

Hunter, M. A., Timothy, Q., Hayes, M. P. 2005. Low Flow Spatial Characteristics in Forested Headwater Channels of Southwest Washington. *Journal of the American Water Resources Association (JAWRA)* 41(3):503-516

Jensen, C. K., McGuire, K. J., and Prince, P. S. 2017. Headwater stream length dynamics across four physiographic provinces of the Appalachian Highlands. *Hydrol. Process,* 31, 3350–3363, <https://doi.org/10.1002/hyp.11259>, 2017.

Jowett, I., Duncan, M. 1990. Flow Variability in New Zealand Rivers and Its Relationship to In-Stream Habitat and Biota. *New Zealand Journal of Marine and Freshwater Research* 24(3):305-317. 10.1080/00288330.1990.9516427

Kaplan, N. H., Sohrt, E., Blume, T., Weiler, M. 2019. Monitoring ephemeral, intermittent and perennial streamflow: a dataset from 182 sites in the Attert catchment, Luxembourg. *Earth Syst. Sci. Data*, 11, 1363–1374, 2019. doi.org/10.5194/essd-11-1363-2019

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O., Poresky, A., Pak, C. 2016. Smarter stormwater systems. *Environ. Sci. Technol.* 50, 7267-7273. dx.doi.org/10.1021/acs.est.5b05870

Kirchner, J. W. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water. Res. Research.*, 42(3).

Kirchner, J. W. 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backwards. *Water. Res. Research.*, 45. doi:10.1029/2008WR006912

Kröger, R., Holland, M. M., Moore, M. T., Cooper, C. M. 2008. Agricultural drainage ditches mitigate phosphorous loads as a function of hydrological variability. *J. Env. Quality.* 37: 1. 107-113.

Larned, S. T., Schmidt, J., Datry, T., Konrad, C. P., Dumas, J. K. 2010. Longitudinal river ecohydrology: flow variation down lengths of alluvial rivers. *Ecohydrol.* 4. 532-548. 10.1002/eco.126

Le Coz, J., A. Patalano, D. Collins, N.F. Guillen, C.M Garcia, G.M. Smart, J. Bind, A. Chiaverini, R. Le Boursicaud, G. Dramais, I. Braud. 2016. Crowdsourced data for flood hydrology: feedback from recent citizen science projects in Argentina, France and New Zealand. *J. Hydrol.* 541, 766-777. dx.doi.org/10.1016/j.jhydrol.2016.07.036

Leigh, C., Boulton, A. J., Courtwright, J. L., Fritz, K., May, C. L., Walker, R. H., Datry, T. 2015. Ecological research management of intermittent rivers: a historical review and future directions. *Freshwater biology.* 66: 8. 1181-1199.

Lovett, G. M., D. Burns, C. T. Driscoll, J. C. Jenkins, M. J. Mitchell, L. Rustad, J. B. Shanley, G. E. Likens, R. Haeuber. 2007. Who needs environmental monitoring? *The ecological society of America.* 5: 253-260.

Magnusson, J., Jonas, T., Kirchner, J. W. 2012. Temperature dynamics of a proglacial stream: Identifying dominant energy balance components and inferring spatially integrated hydraulic geometry. *Water Resour. Res.*, 48. W06510, doi:10.1029/2011WR011378

Meyer, J. L., Wallace, J. B. 2001. Lost linkages and lotic ecology: Rediscovering small streams. Pages. 295-317. in Press MC Huntly NJ Levin S, eds. *Ecology: Achievement and Challenge.* Oxford (United Kingdom): Blackwell Scientific.

McCallum, A. M., Andersen, M. S., Rau, G. C., Larsen, J. R., Acworth, I. R. 2014. River-aquifer interactions in a semiarid environment investigated using point and reach measurements, *Water Resour. Res.*, 50, 2815–2829, doi:10.1002/2012WR012922.

Mishra, A.K., Coulibaly, P. 2009. Developments in hydrometric network design: A review. *Reviews of Geophysics*, 47 (2). doi:10.1029/2007RG000243

van Ogtrop, F. F., Vervoort, R. W., Heller, G. Z., Stasinopoulos, D. M., Rigby, R. A. 2011. Long-range forecasting of intermittent streamflow, *Hydrol. Earth Syst. Sci.*, 15, 3343-3354, [www.hydrol-earth-syst-sci.net/15/3343/2011/](http://www.hydrol-earth-syst-sci.net/15/3343/2011/), doi:10.5194/hess-15-3343-2011

Paillex, A., Siebers, A. R., Ebi, C., Mesman, J., Robinson, C. T. 2019. High stream intermittency in an alpine fluvial network: Val Roseg, Switzerland. *Limnol. Oceanogr.* 9999, 2019, 1–12. doi: 10.1002/lno.11324

L. Ruiz-Garcia, L. Lunadei, P. Barreiro, J. I. Robla. 2009. A review of wireless sensor technologies and applications in agriculture and food industry: State of the air and current trends. *Sensors*. 9(6): 4728-4750.

Sidle, R. C., Tsuboyama, Y., Noguchi, S., Hosoda, I., Fujieda, M., Shimizu, T. 2000. Streamflow generation in steep headwaters: A linked hydro-geomorphic paradigm. *Hydrological Processes*. 14: 369-385.

T. H. Snelder, T. Datry, N. Lamouroux, L. T. Larned, E. Sauquet, H. Pella, C. Catalogne. 2013. Regionalization of patterns of flow intermittence from gauging station records. *Hydrology & Earth System Sciences Discussions* 17: 2685–2699.

Sowder, C., Steel, E.A. A note on the collection and cleaning of water temperature data. *Water* 2012, 4, 597-606

Stromberg, J., Bagstad, K., Leenhouts, J., Lite, S., Makings, E. 2005. Effects of Stream Flow Intermittency on Riparian Vegetation of a Semiarid Region River (San Pedro River, Arizona). *River Res. And App.* doi.org/10.1002/rra.858

Tauro, F., Selker, J., van de Giesen, N., Abrate, T., Uijlenhoet, R., Porfiri, M., Manfreda, S., Caylor, K., Moramarco, T., Benveniste, J., Ciraolo, G. 2018. Measurements and Observations in the XXI century (MOXXI): innovation and multidisciplinary to sense the hydrological cycle. *Hydrological Sciences Journal*, 63(2), pp.169-196.

Tsubaki, R. Tsutsumi, S. Fujita, I. Measurement of the flood discharge of a small-sized river using an existing digital video recording system. *Journal of Hydro-environment Research*, 5, 4. 313-321. 2011.

Turner, D. S., Richter, H. E. 2011. Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions. *Env. Mgmt* 47: 497–505

Vidal-Abarca, M. Sánchez-Montoya, M. Guerrero Romero, C. Gómez, Rarce, M. García-García, V. Suárez Alonso, M. 2013. Effects of intermittent stream flow on macroinvertebrate community composition and biological traits in a naturally saline Mediterranean stream. *Journal of Arid Environments*. 99.10.1016/j.jaridenv.2013.09.008.

Visser, I. Speekenbrink, M. depmixS4: An R-package for hidden Markov models, *Journal of Statistical Software*, Vol VV, II.

Wani, O., Scheidegger, A., Carbajal, J.P., Rieckermann, J., Blumensaat, F. 2017. Parameter estimation of hydrologic models using likelihood function for censored and binary observations. *Water Res.* 121, 290-301. 10.1016/j.watres.2017.05.038

Webb, B. W., Hannah, D. M., Moore, R. D., Brown, L. E., Nobilis, F. 2008. Recent advances in stream and river temperature research. *Hydrol. Process.* 22, 902 – 918. 0.1002/hyp.6994

Wickert, A. D., Sandell, C. T., Schulz, B., Ng, G. H. C. 2018. Open-source Arduino-derived data loggers designed for field research. *Hydrol. Earth Syst. Sci.* doi.org/10.5194/hess-2018-591

Young, D., Hart, J., Martinez, K. 2015. Image analysis techniques to estimate river discharge using time-lapse cameras in remote locations, *Computers & Geosciences*, 76: 1-10.

## 1. Quantifying length of stream orders

The calculated sums of stream orders are presented in Table 2-8. The shape file of South Australian waterways which contained stream order data (Figure 2-10) was accessed from the Government of South Australia website *WaterConnect*.

Table 2-8: The summed length of first to seventh stream orders for the South Australian region

	First order	Second order	Third order	Fourth order	Fifth order	Sixth order	Seventh order	Total
Total length	173,929 km	21,792 km	11,117 km	5,547 km	2,768 km	1,297 km	321 km	216,771 km

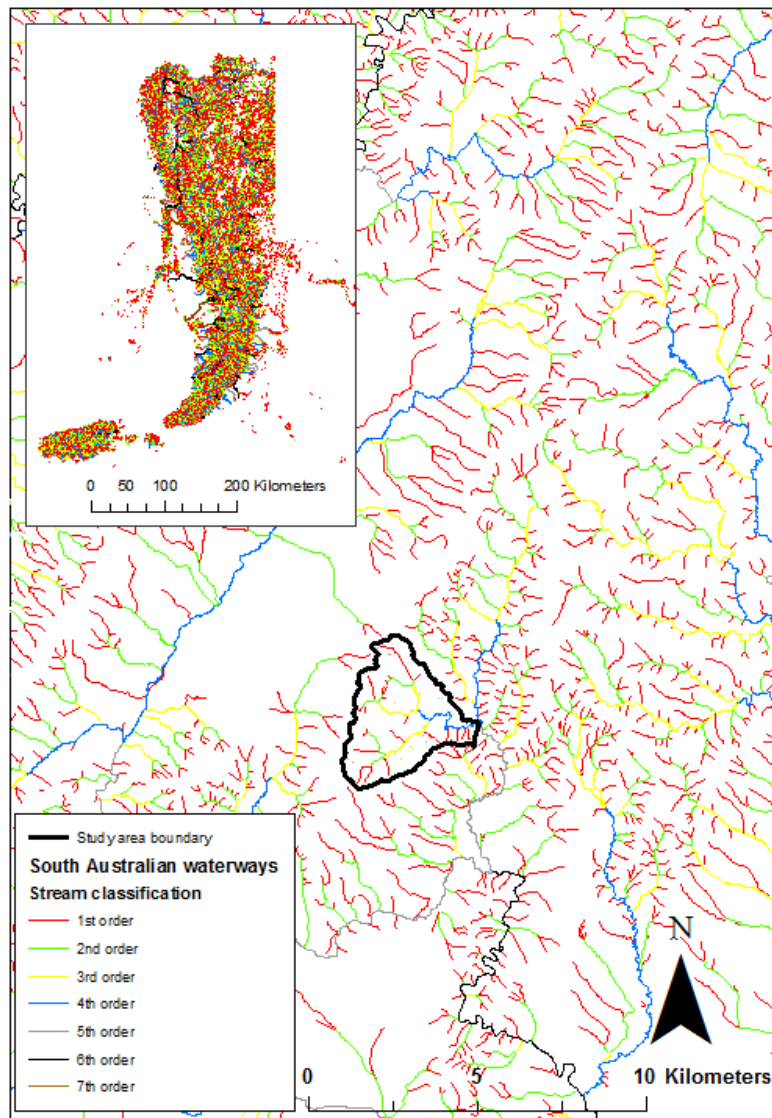


Figure 2-10: Stream orders using the Strahler (1957) ordering method for the South Australian stream network, ranging from first to seventh order streams.

## 2. Detailed descriptions of site details and contributing areas

Table 2-9: Detailed description of each site



Site	Reach width	Site Description	Photos
S001	95.0 cm	Sensors located on reach 1, approximately 1.5 km upstream of the outlet. The sensors are installed 0.10m above the stream bed and above the continuous base-flow which typically flows all year round. The base-flow receded in January 2019 during data collection which is uncharacteristic of the area. Average channel slope is 0.04 m/m	 <p>Photo taken on the 25<sup>th</sup> January 2017 with the temperature and pressure sensor sitting just above the water level. The true classification of these sensors is 'dry' the HMM probability of it being in a 'wet' state is <math>1.36 \times 10^{-4}</math></p>
S002	64.0 cm	Sensors located on reach 1, approximately 1.5 km upstream of the outlet. The sensors are installed in a pool of water. Average channel slope is 0.04 m/m	



			Photo taken on the 9 <sup>th</sup> August 2018 with a single temperature sensor installed on the post. The 'true' state of the stream is 'wet' the HMM probability of 'wet' state is 0.998
S003	132.5 cm	Sensors located on reach 1, approximately 1.0 km upstream of the outlet. The sensor is located in a pool, the reach section does not flow year round and is seasonal. The average channel slope is 0.036 m/m. The site contains grass, with dense vegetation not present in the area. The channel contains dark brown soils, with leaves, grasses and organic matter present in-stream and on-bank. The channel contains pool and springs, with no presence of rock in the channel or outcropping of fractured rock observed in the vicinity.	 <p>Photo taken on the 11<sup>th</sup> October 2017 with temperature and pressure sensor recording measurements. The true state of the stream has been classified as 'wet' and the HMM probability of 'wet' state is 0.022 which has been incorrectly classified by the classifier.</p>
S004	31.5 cm	Sensors located upstream of Reach 1, approximately 2.0 km upstream of the outlet. Site is vegetated with clayey/loamy stream bed (dark brown soil with organic material overlaying i.e. leaves). Clay soil high in organic matter content. The channel is very rocky. Fracture rock outcrop is observed at the site (see photos) and is underlying approximately 0.5m below the sensor location. Pools, springs and riffles located along the channel. Average slope of the channel is 0.05 m/m.	





Photo taken on the 18<sup>th</sup> March 2019. The post contains dual temperature sensor s one for the base-flow measurements and one temperature sensor to measure response above the 0.15m threshold. The photo illustrates the stream which has completely receded which is not typical for this reach. The true state of the stream is 'dry' the HMM probability of 'wet' state is 0.0



Examples of fractured rock at the site



Photo illustrating rocky channel, with vegetation debris also observed.

S005	36.5 cm	<p>Sensor was deployed upstream of tributary 2. The contributing area is densely vegetated.</p> <p>The sensor was installed in a well defined swallow channel. A majority of the section of reach is not incised.</p> <p>Streambed and surrounding soil is medium to dark brown with lots of vegetation surrounding and organic matter overlaying soil. (clay soil high in organic matter content). No springs or pools present. No underlying rock (outcrop) observed. The overall catchment and site is steep (slope 0.04 m/m).</p>	 <p>Photo taken on the 11<sup>th</sup> April 2017. The true state of the stream is 'dry' the HMM probability of 'wet' state is 0.0</p>  <p>Photo taken on the 28<sup>th</sup> August 2017. The post contains a single temperature sensor. The true state of the stream is 'wet' the HMM probability of 'wet' state is 1.0</p>
------	---------	---	--



			 <p data-bbox="1070 683 2029 753">Photo taken directly below the sensor showing the stream channel is less defined (not incised) looking like puddles of water.</p>
S006	251.8 cm	<p data-bbox="479 778 1043 1062">The site is the catchment outlet with three tributaries upstream. The site contains mainly grasses and some trees. The channel is well defined and wide. The soil is dark brown (clay/loam) with high loads of organic matter. The average slope of the catchment of 0.036 m/m. No rocky outcrops are observed in the area with the channel containing vegetation debris.</p>	



Photo taken on the 11<sup>th</sup> April 2017. The true state of the stream is 'wet' the HMM probability of 'wet' state is 1.0.



Photo taken on the 23<sup>th</sup> August 2017. The catchment outlet contains sensor which measures temperature and pressure. The true state of the stream is 'wet' the HMM probability of 'wet' state is 1.0.







			 <p data-bbox="1070 635 2031 735">Photo of the confluence located upstream of the outlet. On the left side is Reach 1 and the right side is Reach 3, both reaches flowing. Photo taken on 9<sup>th</sup> August 2018.</p>
S008	118.6 cm	<p data-bbox="479 767 1043 1118">Average slope of the catchment is 0.02 m/m. At the site there is non-dense vegetation present, the catchment overall is cleared of vegetation containing pastures and grasses. The channel soil is dark brown (clay with high content of organic matter). The channel is deeply incised due to high velocity flows. The reach contains some pools along the channel, with flow being seasonal. Fracture rock is not observed and the channel contains on rock.</p>	 <p data-bbox="1070 1187 2031 1326">Photo taken on the 24<sup>th</sup> August 2017. High velocity flows have bent the post which had dual temperature sensors attached. The true classification of the site is 'wet' and the HMM probability of 'wet' state is 1.0. The sensor was observed to be bent due to high velocity flows.</p>



Illustration of high velocity flows at site. A large branch was observed to be attached to the sensor. Owner at this property has described flows being able to pick up large concrete blocks (Photo taken 24<sup>th</sup> August 2017)



			<p>Post where sensor is located became bent during 2017 winter high flows (June-August 2017).</p>  <p>Sensor was buried as a result of sediment build up around the sensor. This sensor was accumulated over two months (between visits). Photo taken on the 19<sup>th</sup> July 2018.</p>
S010	174.0 cm	<p>Sensors located on reach 1 downstream of site S003. Average channel slope is 0.036 m/m. The site contains grass, with dense vegetation not present in the area. The channel contains dark brown soils, with leaves, grasses and organic matter present in-stream and on-bank. The channel is seasonal with no presence of flow during dry periods. There is no presence of rock in the channel or outcropping of fractured rock observed in the vicinity.</p>	 <p>Photo taken on the 1<sup>st</sup> March 2018. The post contains a single temperature sensor. The true state of the stream is 'dry' the HMM probability of 'wet' state is 0.0</p>

			 <p data-bbox="1077 644 2029 711">Photo taken on the 9<sup>th</sup> November 2018. The true state of the stream is 'wet'. The HMM probability of 'wet' state is 1.0</p>
S011	75.5	<p data-bbox="479 743 1046 1058">Sensors located downstream of site S008 on tributary 3. The site contains mainly grasses and some trees. The channel is well defined, although the channel is less incised and deep as upstream. The soil is dark brown (clay/loam) with high loads of organic matter. The average slope of the catchment of 0.02 m/m. No rocky outcrops are observed in the area with the channel containing vegetation debris.</p>	 <p data-bbox="1070 1147 2029 1249">Photo taken on the 9<sup>th</sup> November 2018. The post contains dual temperature sensors. Debris has accumulated around the post. The true state of the stream is 'wet'. The HMM probability of 'wet' state is 1.0.</p>



### 3. Results of the Standard deviation method

The result of the standard deviation method are shown for the variance approach (Table 2-10) and the  $F(t)$  ratio approach (Table 2-11). The performance across sites varies and diminished accuracy (less than 90%) was observed at sites S001, S002, S004 and S008 for both approaches.

Table 2-10: Results of standard deviation method: Variance approach

Site	Threshold deviation (°C)	Calibration					Evaluation				
		Accuracy in binary classifications	Error in the number of 'wet' days predicted	False 'wet' classification	False 'dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error in the number of 'wet' days predicted	Number of transitions $T_{ref}: T_{inferred}$	False 'wet' classification	False 'dry' classification
S001	0.5	88%	57%	3%	9%	40:32	77%	59%	12:54	3%	20%
S002	2.1	83%	12%	12%	5%	38:66	66%	78%	18:62	34%	0%
S003	5.7	96%	-3%	4%	0%	1:17	93%	-1%	4:2	5%	2%
S004	1.5	90%	16%	7%	3%	14:46	59%	67%	6:58	0	41%
S005	0.6	97%	10%	0%	3%	10:13	99%	7%	4:2	0%	1%
S006	3.6	97%	-1%	2%	1%	5:20	92%	9%	7:79	0%	8%
S008	2.3	70%	-46%	23%	7%	2:85	77%	-20%	4:79	15%	8%
S010	2.7	95%	-4%	4%	1%	4:30					
S011	5.1	94%	0%	3%	3%	8:24					
Median		90%	15%	4%	6%	22:37	80%	35%	8:48	8%	11%

Table 2-11: Results of standard deviation method: F(t) ratio approach

Site	Threshold deviation	Calibration					Evaluation				
		Accuracy in binary classifications	Error in the number of 'wet' days predicted	False 'wet' classification	False 'dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error in the number of 'wet' days predicted	Number of transitions $T_{ref}: T_{inferred}$	False 'wet' classification	False 'dry' classification
S001	14.9	88%	54%	1%	11%	40:20	89%	26%	12:38	0%	11%
S002	8.7	77%	29%	21%	3%	38:21	50%	105%	18:28	48%	2%
S003	4.4	98%	0%	1%	1%	1:11	94%	-2%	10:20	5%	1%
S004	5.7	88%	36%	1%	10%	14:28	53%	66%	6:46	0%	47%
S005	7.1	98%	3%	0%	2%	10:12	98%	-7%	4:6	1%	1%
S006	2.7	98%	0%	1%	1%	5:10	94%	7%	7:11	1%	5%
S008	2.3	82%	25%	5%	7%	2:29	67%	-7%	4:49	20%	13%
S010	10.9	95%	3%	6%	1%	4:24					
S011	4.3	96%	0%	1%	3%	8:12					
Median		91%	17%	4%	5%	22:18	78%	36%	8:26	11%	11%

## 4. Results of the HMM method

The two-state hidden Markov model (HMM) was applied by investigating five alternative algorithm inputs: (1) in-stream temperature variance; (2) ratio of on-bank to in-stream temperature variance; (3) difference of on-bank to in-stream daily temperature mean; (4) in-stream to on-bank temperature profile lag; and (5) 30-day antecedent rainfall. Eleven different combinations of inputs were trialed and the results are presented in the preceding sections. The eleven inputs were:

1. Daily in-stream temperature variance (used as a benchmark)
2. Daily in-stream temperature variance and daily rainfall antecedent moisture
3. Daily in-stream temperature variance and difference in daily in-stream to on-bank temperature mean
4. Daily in-stream temperature variance and in-stream to on-bank temperature lag
5. Daily in-stream temperature variance and ratio of on-bank to in-stream temperature variance –  $F(t)$
6. Daily in-stream temperature variance, daily rainfall antecedent moisture and difference in daily in-stream to on-bank temperature mean
7. Daily in-stream temperature variance, daily rainfall antecedent moisture and in-stream to on-bank temperature lag
8. Daily in-stream temperature variance, daily rainfall antecedent moisture and ratio of on-bank to in-stream temperature variance –  $F(t)$
9. Daily in-stream temperature variance, difference in daily in-stream to on-bank temperature mean and in-stream to on-bank temperature lag
10. Daily in-stream temperature variance, difference in daily in-stream to on-bank temperature mean and ratio of on-bank to in-stream temperature variance –  $F(t)$
11. Daily in-stream temperature variance, in-stream to on-bank temperature lag and ratio of on-bank to in-stream temperature variance –  $F(t)$

Table 2-12: HMM observation inputs: in-stream variance

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	57	212	42	1	40:67	81	34	17	2	12:64
S002	82	-4	8	10	38:60	73	55	26	1	18:50
S003	97	-4	0	3	1:7	84	-2	3	4	10:14
S004	81	6	17	2	14:27	88	-1	4	5	6:19
S005	69	116	31	0	10:23	60	218	40	0	4:27
S006	98	-1	1	1	5:12	91	-14	0	9	6:8
S008	68	65	29	3	2:51	74	46	20	6	4:67
S010	98	-2	1	1	4:12					
S011	93	-8	1	6	8:16					
Median	83	42	14	3	22:31	78	48	16	4	8:36

Table 2-13: HMM observation inputs: in-stream variance and 30 day rainfall

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	70	54	25	5	35:63	86	3	8	6	16:25
S002	82	-6	7	11	38:50	84	12	11	5	18:20
S003	96	-5	0	4	1:11	93	5	5	2	10:9
S004	79	77	21	0	14:20	89	1	4	3	6:4
S005	70	12	30	0	10:14	65	-7	17	18	4:10
S006	97	-3	0	3	5:12	91	-14	0	9	6:8
S008	86	-34	0	14	2:8	66	102	33	1	4:11
S010	97	-2	1	2	4:10					
S011	94	-7	1	5	8:16					
Median	86	10	9	5	22:23	82	15	11	6	8:12

Table 2-14: HMM observation inputs: in-stream variance and difference in daily mean

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	59	197	40	1	40:85	77	41	20	3	12:78
S002	82	-3	8	10	38:64	73	33	21	6	18:92
S003	95	-4	1	4	1:19	93	-1	3	4	10:16
S004	76	82	22	2	14:42	86	1	7	7	6:27
S005	65	133	35	0	10:33	66	186	34	0	4:45
S006	98	0	1	1	5:15	93	-10	1	6	6:9
S008	66	49	27	7	2:72	70	67	26	4	4:95
S010	95	0	2	3	2:24					
S011	90	-13	1	9	8:32					
Median	81	49	15	4	22:43	80	45	16	4	8:52

Table 2-15: HMM observation inputs: in-stream variance and temperature profile lag

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	67	146	31	2	40:73	84	-14	5	11	12:38
S002	81	-6	7	12	38:58	80	-10	8	12	18:52
S003	97	-4	0	3	1:9	94	-2	2	4	10:14
S004	69	-57	8	23	14:47	88	-4	5	7	6:16
S005	72	105	28	0	10:19	76	132	24	0	4:13
S006	98	-1	1	1	5:12	90	-16	0	10	6:14
S008	79	4	11	1	2:33	67	69	26	6	4:101
S010	87	-14	2	11	4:70					
S011	87	-19	1	12	8:62					
Median	82	17	10	7	22:43	83	22	10	7	8:35

Table 2-16: HMM observation inputs: in-stream variance and ratio of on-bank to in-stream variance

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}:T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}:T_{inferred}$
S001	86	-11	6	8	40:29	95	3	3	2	12:14
S002	83	-4	7	10	38:49	89	19	10	1	18:25
S003	98	-3	0	2	1:5	94	1	3	3	10:12
S004	92	-9	3	5	14:6	77	-27	0	23	6:18
S005	97	1	2	1	10:10	99	-7	0	1	4:2
S006	100	0	0	0	5:5	93	-8	1	6	6:11
S008	86	-1	7	7	2:13	73	11	15	12	4:5
S010	98	-2	0	2	4:8					
S011	95	-7	0	5	8:8					
Median	93	-4	3	4	22:15	89	-1	4	7	8:12

Table 2-17: HMM observation inputs: in-stream variance, 30 day rainfall and difference in daily mean

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}:T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}:T_{inferred}$
S001	59	200	40	1	40:85	78	44	20	2	12:70
S002	82	-3	8	10	38:62	75	26	18	7	18:88
S003	97	-3	0	3	1:9	93	4	5	2	10:11
S004	77	82	22	1	14:42	89	-4	4	7	6:17
S005	66	129	34	0	10:31	68	175	32	0	4:39
S006	96	-1	1	2	5:18	91	-15	0	9	6:10
S008	69	55	27	4	2:71	70	69	26	4	4:91
S010	97	-2	1	2	4:14					
S011	90	-14	1	9	8:32					
Median	81	49	15	4	22:40	80	43	15	4	8:47

Table 2-18: HMM observation inputs: in-stream variance, 30 day rainfall and lag in temperature profile

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	83	49	13	4	40:49	84	18	11	4	12:34
S002	81	-9	7	12	38:54	83	8	10	7	18:34
S003	98	-3	0	2	1:5	93	5	5	2	10:11
S004	80	64	18	2	14:28	87	-3	6	7	6:10
S005	70	113	30	0	10:13	65	195	35	0	4:17
S006	95	-7	0	5	5:12	90	-16	0	10	6:10
S008	83	-31	2	15	2:17	69	93	30	1	4:17
S010	90	-14	0	10	4:56					
S011	87	-19	1	12	8:60					
Median	85	16	8	7	22:33	82	43	14	4	8:19

Table 2-19: HMM observation inputs: in-stream variance, 30 day rainfall and F(t)

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	90%	5%	5%	5%	40:9	93%	2%	5%	-5%	12:8
S002	83%	7%	7%	-6%	38:49	89%	7%	4%	5%	18:11
S003	98%	0%	2%	-3%	1:5	94%	5%	1%	5%	10:13
S004	93%	0%	7%	-26%	14:8	91%	3%	5%	-3%	6:14
S005	98%	1%	1%	3%	10:10	99%	0%	1%	-7%	4:2
S006	99%	0%	1%	0%	5:6	92%	2%	7%	-11%	6:4
S008	83%	0%	17%	-40%	2:7	84%	15%	1%	45%	4:3
S010	99%	0%	1%	-2%	4:8					
S011	96%	0%	4%	-6%	8:10					
Median	93%	2%	-8%	10%	22:14	92%	5%	3%	4%	8:8

Table 2-20:HMM observation inputs: in-stream variance, difference in mean and lag in temperature profile

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	60	160	39	1	40:85	79	28	16	5	12:78
S002	82	-3	8	10	38:62	75	26	18	7	18:90
S003	96	-4	2	4	1:19	93	-1	3	4	10:16
S004	70	98	28	2	14:71	84	-3	5	7	6:30
S005	65	133	35	0	10:33	67	182	33	0	4:47
S006	98	0	1	1	5:15	91	-14	0	9	6:11
S008	68	43	25	8	2:75	70	60	24	6	4:85
S010	95	0	2	3	4:24					
S011	90	-14	1	9	8:32					
Median	80	46	16	4	22:46	79	40	14	5	8:51

Table 2-21: HMM observation inputs: in-stream variance, difference in mean and F(t)

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	51	136	38	11	40:95	82	32	16	2	12:68
S002	82	0	9	9	38:63	73	48	24	3	18:87
S003	97	-3	0	3	1:9	94	-2	3	4	10:16
S004	83	56	16	1	14:44	88	2	7	5	6:22
S005	98	3	1	1	10:10	99	2	1	0	4:4
S006	99	1	1	0	5:7	93	-8	1	6	6:11
S008	70	15	18	12	2:73	76	30	17	7	4:75
S010	98	0	1	1	4:16					
S011	93	-8	1	6	8:14					
Median	86	25	9	5	22:37	86	18	10	4	8:40

Table 2-22: HMM observation inputs: in-stream variance, temperature profile lag and F(t)

Site	Calibration					Evaluation				
	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$	Accuracy in binary classifications	Error predicted number of ' wet' days	False ' wet' classification	False ' dry' classification	Number of transitions $T_{ref}: T_{inferred}$
S001	56	143	36	8	40:84	93	6	5	2	12:20
S002	82	-6	7	11	38:49	87	21	11	2	18:37
S003	98	-3	0	2	1:5	94	1	4	3	10:12
S004	86	25	10	4	14:26	67	-42	0	33	6:24
S005	97	2	2	1	10:11	97	14	3	0	4:7
S006	90	-12	2	8	5:62	93	-10	1	7	6:9
S008	69	-69	1	29	2:29	73	11	15	12	4:17
S010	88	-14	1	11	4:64					
S011	85	-22	1	14	8:64					
Median	83	33	7	10	22:44	86	15	6	8	8:18



## 5. Summarised results

Table 2-23: The results of the five alternative inputs iterated in the two-state hidden Markov model which had been limited to a maximum of three inputs and with in-stream daily variance included on each calibration

iteration #	Observation response inputs, <i>O</i> (reference state used during algorithm training only)	Calibration accuracy	Evaluation Accuracy	On-bank measurement required?
1	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> </ul>	83%	78%	X
2	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Daily rainfall antecedent moisture</li> </ul>	86%	82%	X
3	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Difference in daily in-stream to on-bank temperature mean</li> </ul>	81%	80%	✓
4	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Lag of in-stream to on-bank temperature profile</li> </ul>	82%	83%	✓
5	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Ratio of on-bank to in-stream daily temperature variance</li> </ul>	93%	89%	✓
6	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Daily rainfall antecedent moisture</li> <li>Difference in daily in-stream to on-bank temperature mean</li> </ul>	81%	80%	✓
7	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Daily rainfall antecedent moisture</li> <li>Lag of in-stream to on-bank temperature profile</li> </ul>	85%	82%	✓
8	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Daily rainfall antecedent moisture</li> <li>Ratio of on-bank to in-stream daily temperature variance</li> </ul>	93%	92%	✓
9	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Difference in daily in-stream to on-bank temperature mean</li> <li>Lag of in-stream to on-bank temperature profile</li> </ul>	80%	79%	✓
10	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Difference in daily in-stream to on-bank temperature mean</li> <li>Ratio of on-bank to in-stream daily temperature variance</li> </ul>	86%	86%	✓
11	<ul style="list-style-type: none"> <li>Daily in-stream temperature variance</li> <li>Lag of in-stream to on-bank temperature profile</li> <li>Ratio of on-bank to in-stream daily temperature variance</li> </ul>	83%	86%	✓

## Chapter 3.

# Does conceptualization of local-scale catchment runoff production matter in representation of intermittent streams across Mediterranean catchments?

Alicja Makarewicz<sup>a</sup>, Michael Leonard<sup>a</sup>, Seth Westra<sup>a</sup>, Daniel Partington<sup>b</sup>

<sup>a</sup> School of Civil, Environmental and Mining Engineering, the University of Adelaide, North Terrace, Adelaide SA 5005, Australia.

<sup>b</sup> National Centre for Groundwater Research & Training, Flinders University, Adelaide SA 5001, Australia

*Environmental modelling and software*, in preparation



# Statement of Authorship

Title of Paper	Does the conceptualisation of local-scale catchment runoff production matter in representation of intermittent streams across Mediterranean catchments?
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Intention to submit to Environmental Modelling and Software

## Principal Author

Name of Principal Author (Candidate)	Alicja Makarewicz		
Contribution to the Paper	Built and calibrated model, analysis and interpretation of results, wrote the manuscript		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/3/2020

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Michael Leonard		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	26/3/2020

Name of Co-Author	Selh Westra		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	25/3/2020

Name of Co-Author	Daniel Partington		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	2/4/2020

Please cut and paste additional co-author panels here as required.

## Abstract

The requirement to quantify the intermittency of streamflow on reaches distributed across Mediterranean catchments is crucial for the management of water resources with multi-scale implications, for example, decisions relating to environmental flows and small distributed dams. The scale of processes being represented is important because many methods and processes that are applicable at one scale are not applicable at another scale. Not capturing spatial variations and relying on a single metric downstream for the evaluation of a rainfall-runoff models could potentially result in the misrepresentation of runoff and hence less effective water management outcomes. This study investigated the implications of local-scale runoff production and the functioning of a small dam when different conceptualizations of hydrological models are calibrated to a single point estimate of streamflow variation. Using the 3D surface-subsurface flow code HydroGeoSphere a partly forested 10 km<sup>2</sup> catchment with intermittent streamflow was calibrated to four competing conceptualisations of runoff-generation behaviour exhibiting near equivalent Nash-Sutcliffe coefficients for streamflow at the outlet. The four competing conceptualisations were: (1) saturation excess dominated, (2) saturation excess and groundwater dominated, (3) groundwater dominated and (4) groundwater dominated but containing 17% infiltration excess. The results demonstrated that subsurface pathways influence the behaviour of flow in and around the dam with groundwater dominant scenarios showing that simulated flow days downstream of the dam increased more than 20% compared to upstream. These differences show that more meaningful quantitative tools for process representation of streamflow intermittency are required. To resolve these catchment complexities, additional data on headwater reaches is required to better inform model parameterizations and to allow for the rejection of some competing conceptual models. Additionally, future research should investigate added model complexities, additional low-cost data and how to best balance the faithful representation on streamflow intermittency within a practical framework.

### 3.1 Introduction

Understanding localized runoff generation is important for predicting intermittent flows and runoff responses that are nested across many scales. For many water management questions, such as low flows, land use management, distributed storages and environmental flows, gauged discharge data are typically exclusively used for policy, planning and design decisions. The aggregation of processes to the outlet can be limited in the representation of catchment-wide runoff production that can be highly variable due to the heterogeneities of surface and subsurface characteristics. Catchment slope, topography, soil, vegetation features, as well as hydroclimatic features depend on fluxes at scales that span multiple orders of magnitude. To represent this heterogeneity, many hydrological problems are bounded or idealized to focus on a subset of processes manifested at a specific scale, such as the microphysical modelling of soil properties (*Li et al.*, 2013), ecohydraulic representation of individual reaches (*Bockelmann et al.*, 2003) or catchment-scale assessment of outlet flows.

There are many hydrological problems that require accurate representation across multiple scales, where larger-scale analyses cannot easily be decoupled from process scales orders of magnitude smaller. For example, over two million agricultural dams are distributed on hillslopes in Australia alone, each with limited storage volume but collectively storing over 8,000 GL of water (*Land and Water Australia*, 2010)—a volume equivalent to the total volume of water stored in large reservoirs supplying Australia's major capital cities. Understanding the hydrological implications of small distributed storages therefore requires accurate representation of small reservoir dynamics at the hillslope scale, and how these dynamics scale up across larger drainage basins. Similarly, ecohydrological problems require the understanding of network intermittency which shapes local and catchment-wide aquatic ecosystems (*Larned et al.*, 2010). Finally, understanding the generation of sediment and associated water quality issues requires multiscale process knowledge, where factors controlling runoff generation and associated sediment transport vary according to the spatial scale (*Inoubli et al.*, 2017).

To address these problems, it is necessary to accurately represent hydrological processes that span both the hillslope scale and larger catchment scale. This is challenging because of vast heterogeneity spanning both space and time scales, non-additivity of surface flows (with local-scale infiltration often re-emerging further downstream) and limited or no availability of streamflow measurements in smaller-scale reaches (*Kirchner*, 2006). The hydrology of hillslopes is very different from the hydrology of catchments, whereby runoff at the hillslope is characterized by greater intermittency and variability than when flow is aggregated across the catchment. For example, preferential flows (*Weiler and McDonnell*, 2007) and threshold responses (*Graham et al.*, 2010) observed at the hillslope can be poorly generalized at larger scales due to complexity in the processes as well as the lack of local scale data. The outcome is that upstream reaches which are typically more variable with intermittent flow and cease-of-flow periods having less data, whereas downstream reaches having more data are less variable due to flow aggregation.

This paper demonstrates simulated differences in process representation at the hillslope scale, which can have implications for multi-scale decisions. This is represented conceptually in Figure 3-1, where two extreme runoff generation mechanisms—filtration excess and saturation excess runoff—are shown to influence drainage patterns around localized storages. Infiltration excess runoff is produced by saturation of the soil surface (saturation from above) when the intensity of precipitation exceeds the rate of infiltration. Figure 3-1 (a) illustrates how infiltration excess runoff travels through a storage with limited interaction with the subsurface. This conceptualization

implies high levels of additivity in surface flows at the hillslope and catchment scales, thereby allowing for idealized assumptions of flow behaviour (typical in node-link conceptual models) with little to no surface-subsurface interaction. Conversely, saturation excess runoff occurs when the entire soil profile is saturated (saturation from below) resulting in return flow as well as runoff from subsequent precipitation on the saturated area, and implies very different processes at the hillslope scale (e.g. lower proportion of saturated area; high infiltration rates; lower proportions of runoff) and the larger scale (higher proportion of saturated area; potential for return flow; higher proportions of runoff). Figure 3-1 (b) shows how the saturation excess mechanism increases the potential of flow pathways around the storage that include explicit interactions with the subsurface. How infiltration excess and saturation excess evolve over time will also vary. For example, infiltration excess could occur over the entire hillslope simultaneously whereas saturation excess might start in low-lying areas and work its way up the hillslope.

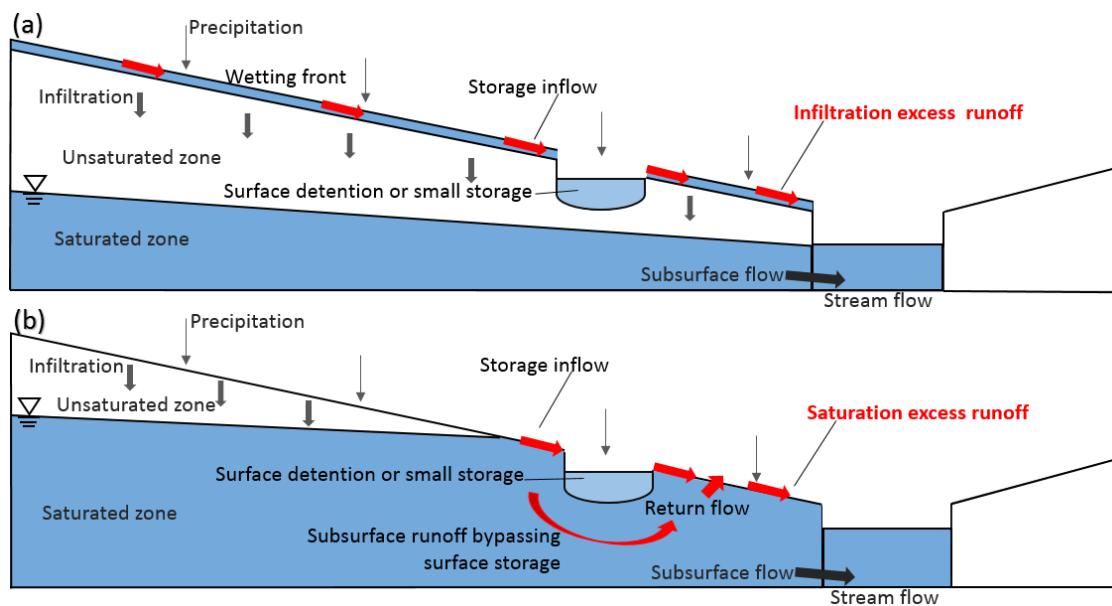


Figure 3-1: Illustration of two main runoff generation mechanisms: (a) infiltration excess runoff; and (b) saturation excess runoff. Black arrows represent general processes in the hydrological cycle and red arrows represent processes directly relevant to each associated runoff mechanism and flow pathways around a small storage.

Although the presence of different runoff mechanisms at the hillslope is widely known (Horton, 1933; Dunne, 1983), hydrological modelling at the catchment scale has been dominated by conceptual models representing aggregate characteristics calibrated to flows at the catchment outlet, with limited physical process realism at the local scale. Although conceptual models are often able to equal or out-perform more complicated physically-based distributed models in a standard calibration/validation setting (Smith et al., 2012), they are typically limited or unable to address intra-catchment behaviour. For example, even where there is the representation of an upstream process in a conceptual model (such as groundwater recharge or quickflow), it is not necessarily useful for interpretation (Partington et al., 2012; Li et al., 2013).

Therefore, it is necessary to better understand the implication of process representation in hydrological models for problems that require faithful representation of hydrological stores and fluxes at both the hillslope and catchment scales. Where, commonly applied models, do not consider soil moisture and water table information. To this end, this paper aims to:



1. Understand the effect of spatial variability of alternative simulated runoff mechanisms calibrated exclusively to outlet discharge and;
2. To investigate the influence of different runoff mechanisms on localized flow pathways.

To achieve these aims a numerical modelling study of a 10 km<sup>2</sup> South Australian catchment was used (Section 3.2). The fully integrated hydrological model, HydroGeoSphere, was selected to simulate multiple runoff scenarios while preserving the model discharge performance at the outlet (Section 3.3). The simulations were analysed to illustrate the spatial features of runoff generation mechanisms and the influence runoff production had on localized flow pathways for small distributed storages (Section 3.4). Discussion and implications of the comparison are outlined along with conclusions of the study (Section 3.5).

### 3.2 Case study

The implications of process representation at the hillslope scale and larger catchment scale were investigated using a small (10 km<sup>2</sup>) Mediterranean catchment in South Australia, GPS coordinates (-35.266420, 138.731021). Due to resource constraints a single catchment was selected. The catchment geometry, physical features and hydro-climatic characteristics were equally implemented for all simulated scenarios. The sub-catchment has three main tributaries (Figure 3-2b), and contains 1<sup>st</sup> to 4<sup>th</sup> order streams that range from perennial to intermittent, with various sources and sinks located along the channels. The catchment ranges from 175 m to 420 m above sea level and is partially vegetated with closed shrub-land and open wood-land. Non-vegetated locations were made up of grasses and pastures. Soils in the area are primarily shallow, made up of loam and sandy loam over clay on rock. Fractured rock is common throughout the area. There are 50 dam storages ranging in size from 1 to 10 mega litres and were represented in the model DEM.

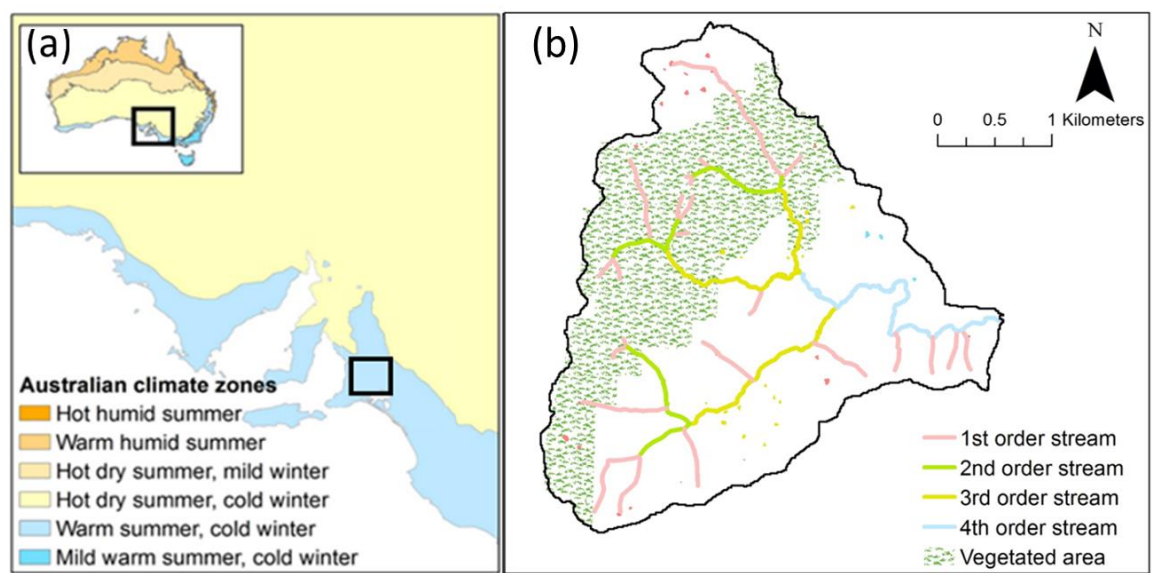


Figure 3-2: (a) catchment location superimposed on Koppen climate classification and (b) a catchment map illustrating the main tributaries, stream orders, vegetated areas and surface areas of local storages that were represented in the model DEM

Subdaily rainfall and daily potential evapotranspiration were obtained from a weather station less than 20 km southeast of the catchment, and given the small catchment size, were assumed to be uniform over the region. A two year period from 2017 to 2019 had been selected, with an average of 700 mm/year rain and 1800 mm/year potential evapotranspiration over this time. The region

is highly seasonal with hot dry summers, and while large rainfall events are possible, in summer there is typically little flow at many sites, other than immediately following the events. The majority of the rainfall arrives in the winter months, and the highest flows occur in the spring period.

### 3.3 Methodology

To investigate the potential influence of modelling assumptions on water management decisions, this study seeks to represent alternate flow pathways through alternative proportions of runoff generation mechanisms—infiltration excess, saturation excess and groundwater runoff—with similar discharge performance at the catchment outlet. The runoff mechanisms were simulated by adjusting the hydraulic conductivity in both the vertical and horizontal dimensions, as well as differences in the vertical discretization of the soil profile. This determines the infiltration rate (controlling the degree of ‘saturation from above’) as well as groundwater level and lateral flow (determining the degree of ‘saturation from below’).

The selected soil properties were within the range of parameters expected for sandy loam and silt soil characteristics, which are present in the catchment being simulated. The parameters were adjusted to achieve a minimum Nash-Sutcliffe efficiency coefficient (NSE) of 0.5 at the outlet and to reflect typical aggregate runoff behaviour in Mediterranean catchments. Where the aim was to investigate the influence of assumption of commonly applied conceptual models that do not consider soil moisture and groundwater data.

A Nash-Sutcliffe efficiency coefficient of 0.5 was deemed appropriate for this study given the simulation time constraints and the time consuming nature of model calibration with alternative HGS studies having produced similar NSE's (*Partington et al.*, 2013; *Li et al.* 2015; *Glaser et al.* 2016; *Tang et al.* 2017). To quantify different proportions of runoff generation, the hydraulic mixing-cell (HMC) method (*Partington et al.*, 2013) was applied within a fully-coupled surface-subsurface flow model, HydroGeoSphere (HGS). Further details on the HMC method are outlined in Appendix A. A scenario was considered to be dominated by a particular runoff mechanism when that mechanism contributes the largest proportion to total flow at the outlet. The identified flow proportions were used to address the research aims regarding the spatial variability of runoff generation and their influence on local scale flow pathways.

#### 3.3.1 Model set-up: explicit simulation of surface/subsurface flows

The process-based hydrological model HydroGeoSphere (HGS) (*Therrien et al.* 2009) was used as the basis for this study, as it is capable of providing insight into physical processes at scales not otherwise viable for empirical studies. The fully integrated surface/subsurface model can simulate surface and subsurface flows at high spatial and temporal resolution, allowing interrogation of flow pathways and runoff generation mechanisms. HGS simulates both 3D variably saturated subsurface flow using a modified Richards' equation, and 2D surface flow (*Brunner et al.* 2012) using the diffusion wave approximation to the Saint Venant equations. HGS solves all governing flow equations simultaneously to simulate streamflow and groundwater discharge to the stream as a function of the catchment physical characteristics and hydrological inputs. The equations were solved using a control volume finite element method with Newton-Raphson linearization.

The model domain was discretised as an irregular triangular mesh to simulate small-scale runoff in and around the channel. The element lengths range from 250m at the outer boundary to 25m in and around the catchment streams and reservoirs. The spatial discretization of the modelled domain was represented with 12,015 elements and 6,115 nodes (Figure 3-3a). In the z direction

the model domain consists of seven layers with a discretization of 0.3 m for the top 1.5 m. The soil depth ranges from 5-24 m beneath the surface domain, with greatest depths at the boundary and linearly becoming shallower towards the outlet. The maximum soil thickness was 26 m and the minimum was 6.5 m, selected to represent the relatively shallow soils characteristic of the area represented in soil maps. The groundwater depths of simulations were different for each scenario and were dependent on the soil profile parameters and the resulting dynamic equilibrium reached. The depth to water table around the catchment boundary ranged between 1-2 m (saturation excess) and 10-26 m (containing infiltration excess) depending on the scenario. The digital elevation model (DEM) features 50 on-stream and off-stream storages ranging from 1 to 10 mega litres in size (Figure 3-3b). The outlet was set as the critical depth boundary and a no-flow boundary was fixed for the bottom and lateral subsurface domain (i.e. water can only leave the model domain through a critical depth boundary at the outlet). The critical depth boundary simulates the transition from sub-critical to supercritical flow, e.g. as one sees with flow over a weir (Therrien *et al.*, 2009).

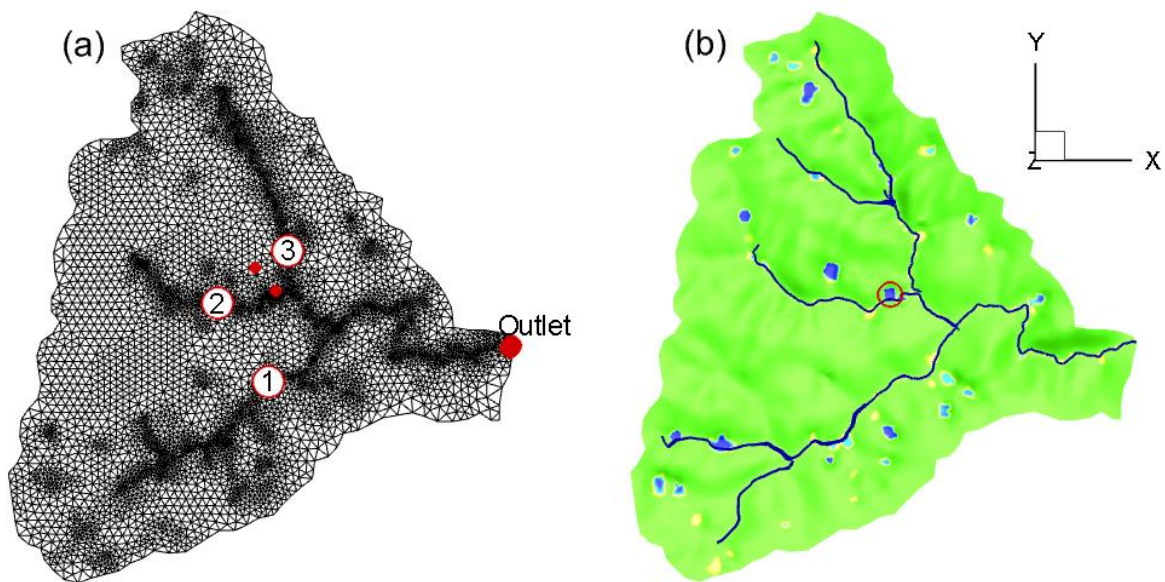


Figure 3-3: (a) Modelled catchment geometry and discretization showing smaller elements located in the channels with numbers 1-3 indicating the three main tributaries and the red dots indicating probe points to compare streambank and hillslope saturation behaviour, and (b) simulated overland flow channels (grey lines) and ponded water present in catchment dams, blue - deep and yellow - shallow shading, with the dam investigated in this study shown with a red circle.

To allow the alternative scenario simulations to reach a dynamic equilibrium, the soil ranges used to develop the scenarios were initialized with a 20-30 year spin up period, depending on soil type. The spin up allows the water table distribution to adjust to the catchment conditions. The models were initialized with the water table at the land surface and then allowed to drain under gravity. The spin up was initialized with the catchment full saturated, and allowed to drain for 6 months. After this, 20-30 years of observed rainfall and evaporation time-series were applied, allowing for the transient adjustment of the subsurface flow paths and gradients relative to the catchment geomorphology. The simulation period for the studied scenarios reflects the length of discharge data available and was March 2016 to March 2019 (two years) of continuous simulation. The first 12 months of data was used for calibration, and the second year was reserved for simulation evaluation.

The catchment was parameterized by dividing the model domain into two units based on vegetation; (1) densely vegetated and (2) grass and pastures (Figure 3-2b). The evapotranspiration (ET) properties were chosen to represent these two groups of vegetated sections. Selected root and evapotranspiration depth for the vegetated and clear areas were 3.5 m and 0.2 m respectively, and were deemed reasonable given the type of vegetation present in the catchment. The values were taken from *Canadell et al., 1996* and *Specht and Rayson 1957* to reflect native trees in the area (e.g. Banksia and Eucalyptus). These parameters were selected to reflect physical characteristics for the location and were kept identical across the all scenarios (Table 3-1).

### 3.3.2 Calibration of scenarios

Each scenario was calibrated by adjusting soil properties to produce a minimum discharge NSE of 0.5 at the outlet (Figure 3-4). The aim of model calibration was to simulate the scenarios with the required performance at the outlet while producing different proportions of runoff mechanisms on the hillslope (quantification of mechanisms is outlined in Section 3.3.3). The relationship between rainfall intensity and saturated hydraulic conductivity was used as the primary control of the dominant runoff generation mechanism (*Mirus & Loague, 2013*). Vertical and lateral saturated hydraulic conductivity were used to control runoff response and surface-subsurface flow pathways. For all scenarios, the applied soil properties remained within realistic values for the catchment soil types, which ranged from clay to sandy loam and were based on parameter ranges in *Puhlmann et al. (2009)*.

The infiltration excess runoff was achieved by impeding infiltration. This required a low vertical saturated hydraulic conductivity with silt/clay properties. This resulted in high velocity runoff events that quickly recede. Runoff was controlled with the lateral saturated hydraulic conductivity (x- and y-direction) to slow or accelerate runoff. The saturation excess runoff was parameterized by increasing the infiltration rate (i.e. fast saturated hydraulic conductivity in the z-direction) with sandy-loam soil properties. While a higher hydraulic conductivity in the vertical than horizontal direction is uncommon, this feature was plausible for the region given the shallow soils and fractured rock common in the study area, since vertical cracking would allow for rapid vertical flow, retarding horizontal flow. The applied soil properties were homogenous and were considered adequate given the small catchment size and soil map descriptions.

To calibrate each of the scenarios a grid-based search was used where the range of each parameter (i.e.  $K_{xy}^i$  and  $K_z^i$ ) was discretised into 10 points. Given there were two parameters to calibrate, this results in 100 instances of parameter combinations to evaluate per iteration of the grid-search. For each of the 100 simulations the Nash-Sutcliffe efficiency and the proportions of runoff mechanisms were calculated for the discharge at the catchment outlet. An evaluation was made whether the identified region had suitably converged based on this performance criteria. When the parameter range had not converged, subsequent iterations of the grid-search were performed by selecting the best performing subdomain of the parameter space and further discretising that region. The initial conditions (e.g. soil moisture and groundwater state) were updated with each iteration using the best performing subdomain (Figure 3-4). The second year of the final simulation was used solely for performance evaluation. The selected parameters for the model calibrated to multiple sites are presented in Table 3-1.

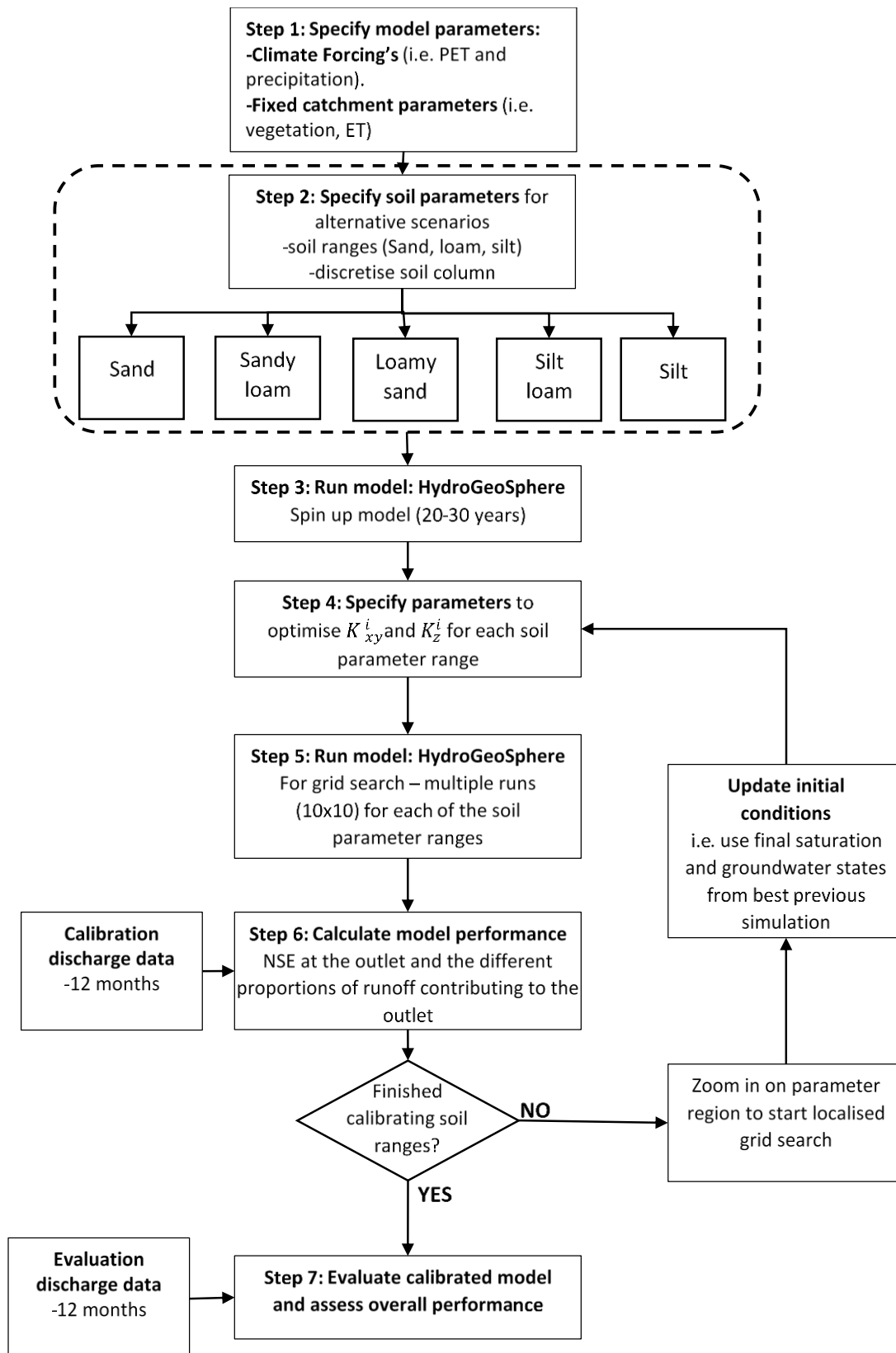


Figure 3-4: Schematic of the method for identifying contrasting simulations that have a comparable discharge performance at the outlet by adjusting the lateral and vertical hydraulic conductivities  $K_{xy}^i$  and  $K_z^i$  respectively. Four resulting scenarios were selected: (1) saturation excess dominant, (2) saturation excess & groundwater dominant, (3) groundwater dominant, and (4) groundwater dominant but with infiltration excess.

Table 3-1: Surface and subsurface parameters for the catchment model, parameter values follow (Partington et al., 2010; Li et al., 2013)

Parameter	Value			
	Scenario 1	Scenario 2	Scenario 3	Scenario 4
<b>Description</b>	Saturation excess dominant	Saturation excess and groundwater dominant	Groundwater dominant	Groundwater dominant and containing infiltration excess
Lateral and vertical hydraulic conductivity (m/s)	$K_{xy}=2.0 \times 10^{-6}$ $K_z=1.3 \times 10^{-4}$	$K_{xy}=1.3 \times 10^{-6}$ $K_z=1.5 \times 10^{-6}$	$K_{xy}=4.8 \times 10^{-5}$ $K_z=7.0 \times 10^{-6}$	$K_{xy}=9.8 \times 10^{-5}$ $K_z=2.8 \times 10^{-7}$
Surface Properties (uniform)				
Manning's roughness (channel) ( $s/m^{1/3}$ )	0.15			
Manning's roughness (overland) ( $s/m^{1/3}$ )	0.05			
Rill storage height (m)	0.001			
Transpiration fitting parameter c1	0.3			
Transpiration fitting parameter c2	0.2			
Transpiration fitting parameter c3	10.0			
Wilting point	0.1			
Field capacity	0.15			
Oxic limit	0.9			
Anoxic limit	1.0			
Limiting saturation (minimum)	0.2			
Limiting saturation (maximum)	0.32			
Surface – Vegetated (uniform)				
Leaf area index	2.08			
Root Depth (m)	3.5			
Canopy storage parameter (mm)	0.5			
Initial interception storage (mm)	0.5			
Evaporation depth (m)	3.5			
Surface – Cleared (uniform)				
Leaf area index	0.9			
Root Depth (m)	0.2			
Canopy storage parameter (m)	0.0			
Initial interception storage (m)	0.0			
Evaporation depth (m)	0.2			
Coupling length (m)	0.00001			

### 3.3.3 Quantifying runoff generation mechanisms

To quantify the components of runoff generation across the catchment model, the hydraulic mixing-cell (HMC) method was applied (Partington et al., 2011, 2013). The method had the ability to track surface and subsurface flows (Schilling et al., 2018) based on inflow boundary conditions such as rainfall. The method 'tags' inputs from boundary domains by delineating the inflow between infiltration excess and saturation excess with flow labelled as a function of the model cell's surface saturation and depth to groundwater (e.g. saturation from below or saturation from

above). The tagged water volume was tracked as a fraction of the total water volume in each cell through the model domain using a modified mixing-cell approach (after *Campana & Simpson, 1984*). The reason for tracking the fraction of water attributed to the runoff generation mechanism was to quantify the amount of flow contributing to a point of interest, such as the catchment outlet, as some flow remains stored within the catchment (*Partington et al., 2013*).

The previous method of *Partington et al. (2011, 2013)* has been extended here by using additional HMC fractions to those previously implemented, which only considered direct rainfall to a cell without further delineation into saturation excess or infiltration excess. All in-stream and overland flow generation mechanisms were delineated by user defined model surface nodes as either overland or in-stream nodes, defined as nodes which were located overland (catchment surface) or within a stream network channel. The model provides unique runoff fractions of: (1) infiltration excess flow as return flow; (2) saturation excess flow as return flow; (3) direct rainfall as infiltration excess flow; (4) direct rainfall as saturation excess flow; (5) direct rainfall to the channel; and (6) groundwater discharge. Existing water within the catchment at the start of the simulation was delineated as “initial” which had an unknown origin. To check for any errors within the HMC analysis and ensure sensible results, a reset fraction and error were also computed (see *Partington et al., 2013*). Appendix 3-A outlines details of the HMC method.

### 3.4 Results and discussion

#### 3.4.1 Runoff response comparison at the catchment outlet

The four calibrated scenarios show different proportions of runoff generation mechanisms: infiltration excess, saturation excess and groundwater (Figure 3-5). The four calibrated scenarios selected for analysis were:

- (1) Scenario 1: dominated by saturation excess runoff
- (2) Scenario 2: similarly dominated by saturation excess and groundwater runoff
- (3) Scenario 3: dominated by groundwater runoff
- (4) Scenario 4: dominated by groundwater runoff but containing the greatest proportion of infiltration excess achieved while maintaining performance at the outlet.

Figure 3-5 illustrated the breakdown in the runoff generation mechanisms showing that Scenario 1 simulates 57% saturation excess runoff, with groundwater contributing to 30% of runoff at the outlet. Scenario 2 was similarly dominated by saturation excess and groundwater runoff, with 44% and 47% contributing to flow at the outlet respectively. In contrast, Scenario 3 was dominated by groundwater runoff with 57% contributing to the outlet and 30% contributed from saturation excess runoff. Additionally, 0.2% of runoff was infiltration excess overland flow and 3.5% of infiltration excess was infiltrated, re-emerging at the outlet. Finally, Scenario 4 while also dominated by groundwater runoff (51%) contains both saturation excess (21%) and infiltration excess (17%) runoff at the outlet, noting that 4% was overland runoff and 13% was infiltrated, re-emerging at the outlet (i.e. return flow). While every attempt was made by the authors to produce a scenario dominated by infiltration excess the task proved to be difficult because some infiltration excess that was simulated on the hillslope was infiltrated into the groundwater table (being re-tagged as groundwater) and was not flowing to the outlet. Scenarios 1 and 2 simulated negligible infiltration excess runoff and infiltration. All scenarios produced less than 0.5% of saturation excess that was infiltrated and later re-emerging.

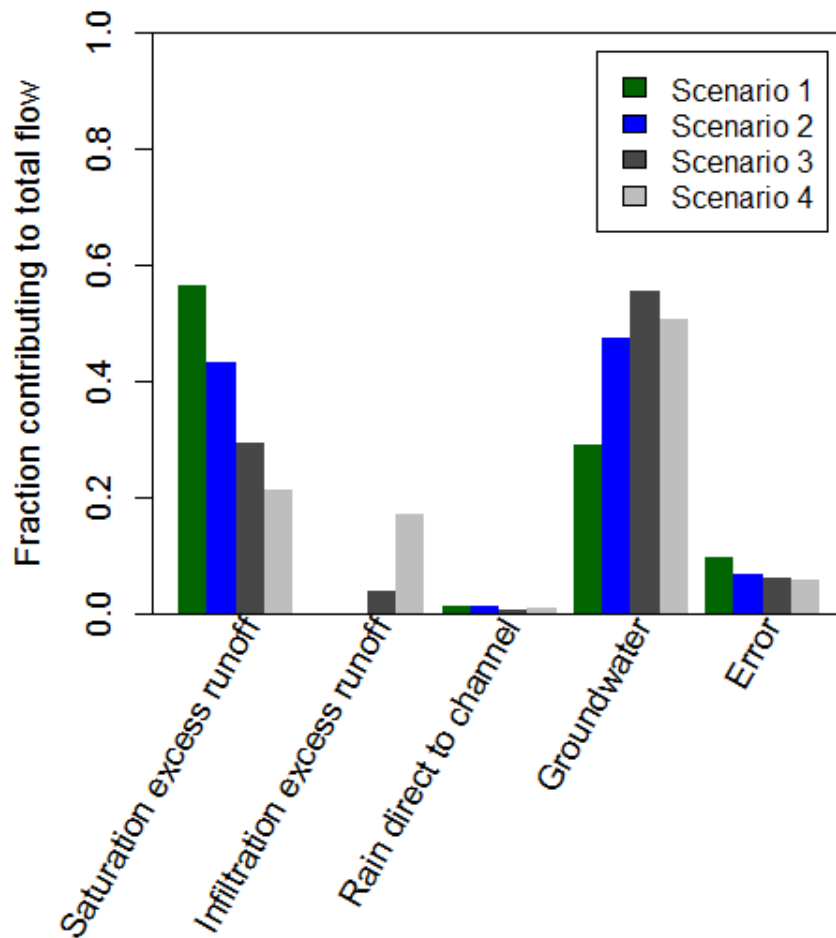


Figure 3-5: The different fractions of runoff contributing to total flow at the catchment outlet over a two year continuous period for the four simulated scenarios.

The four scenarios were calibrated with a minimum Nash-Sutcliffe efficiency coefficient (NSE) of 0.50 at the outlet over the first year of simulations. The simulations were evaluated with outlet discharge over the second year of simulations are shown in Figure 3-6. Scenarios 1, 2, 3 and 4 each achieved a calibration NSE of 0.52, 0.54, 0.62 and 0.69 respectively. All four scenarios reduced in performance over the evaluation period achieving an NSE of 0.32, 0.48, 0.29 and 0.15 respectively. Difference in timing of flow onset and cessation for the four scenarios were identified. Scenarios 1 and 2, dominated by saturation excess and saturation excess and groundwater runoff respectively, had continuous flow for the majority of the year. In contrast, Scenarios 3 and 4, both dominated with groundwater but containing 3.5% and 17% infiltration excess respectively, simulated seasonal flow characteristics where continuous flow began in early April and late June over the first year of simulations, respectively. Scenarios 3 and 4 streamflow receded in mid-December and early-December respectively. All simulations had a negative total flow bias with Scenario 1, 2, 3 and 4 each under-estimating total flow by 25%, 7%, 8% and 29%.



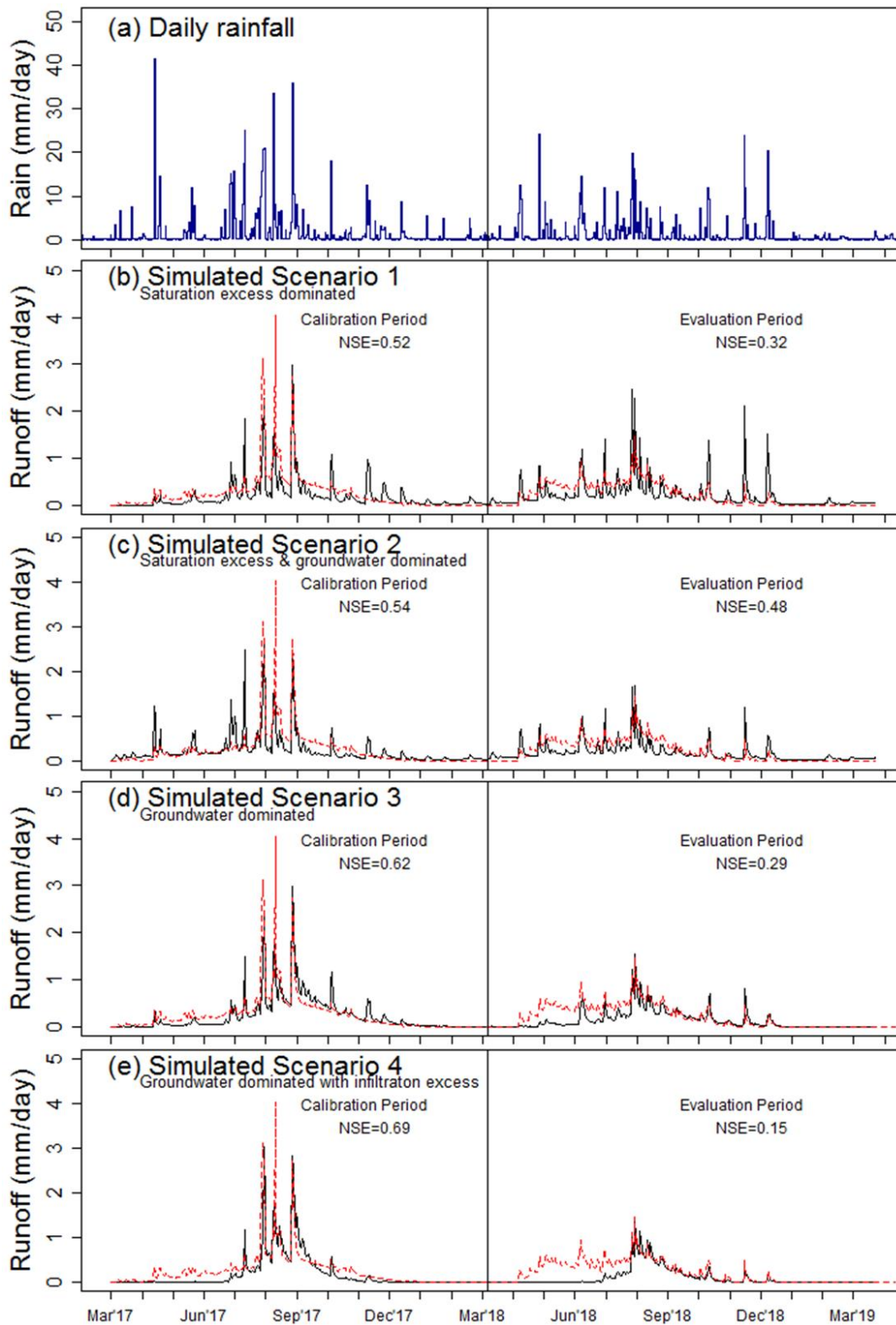


Figure 3-6: Showing model outputs over the calibration and evaluation periods for the (a) applied rainfall and hydrographs of the four simulated scenarios; (b) scenario 1: saturation excess dominated; (c) scenario 2: saturation excess and groundwater dominated; (d) scenario 3: groundwater dominated; and (e) groundwater dominated with infiltration excess.

### 3.4.2 Comparison of runoff response across the catchment

The spatial distribution of moisture stores and fluxes during the first major rain event of this series (day 250) was shown in Figure 3-7. Scenario 1, dominated by saturation excess, shows that surface saturation (Figure 3-7a) occurs in areas where the groundwater table was at the surface (Figure 3-7b), illustrating 'saturation from below'. The results also indicate that the groundwater table in unsaturated areas on the hillslope remains close to the surface (i.e. less than 0.5m below the surface). In contrast, Scenario 4, containing 17% infiltrated excess runoff, shows that the surface of the catchment was saturated (Figure 3-7g) while the depth to the groundwater table was significantly below the surface elevation on the hillslope, (Figure 3-7h), illustrating 'saturation from above'. Figure 3-7h shows that the depth to groundwater table ranged from 26.0 m below the surface at the boundary to approximately 2.0 m within the catchment and closer to the channel. The results show that the depth to groundwater table within the channel was at the surface for some sections and below the channel surface (less than 1.0 m below) in other sections of the channel. Scenarios 2 and 3 show behaviour that ranges between Scenario 1 and 4, where the surface saturation increased from Scenarios 1 to 4, the hillslope saturation approximately 17%, 53%, 77% and 95% respectively (at day 250 of the simulation). The groundwater table does not show a similar pattern of behaviour with the depth to groundwater table in Scenario 3 (groundwater dominated) than Scenario 2 (saturation excess and groundwater dominated).

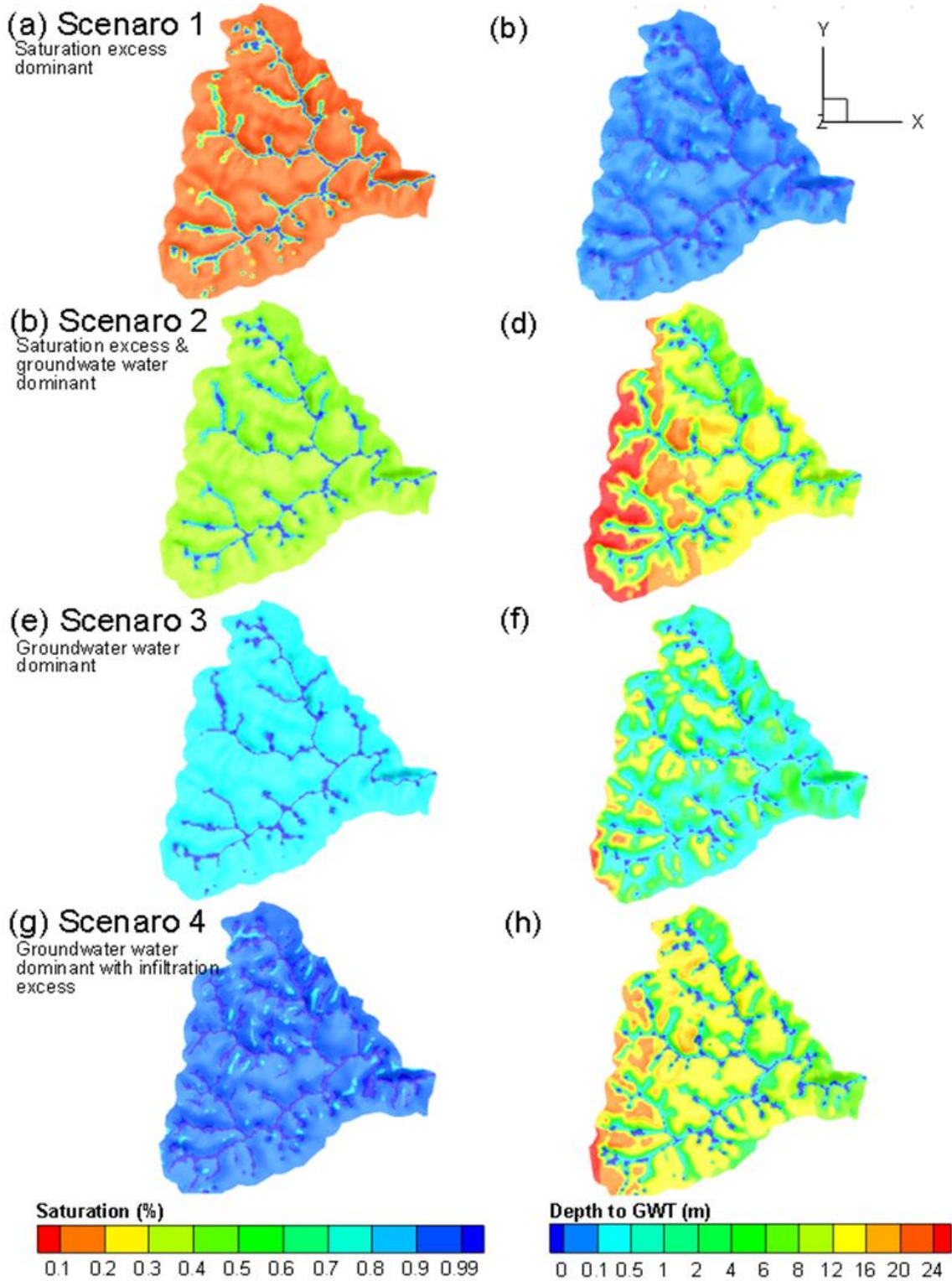


Figure 3-7: Simulated spatial behaviour at day 250 of the surface saturation (a) and depth to groundwater table (b) for Scenario 1, surface saturation (c) and depth to groundwater table (d) for Scenario 2, surface saturation (e) and depth to groundwater table (f) for Scenario 3, and surface saturation (g) and depth to groundwater table (h) for Scenario 4.

### 3.4.3 Analysis of saturation and groundwater behaviour for the runoff mechanisms

To compare and investigate the on-bank and hillslope saturation and groundwater behaviour for the scenarios, a daily time series of surface saturation was compared at two points (Figure 3-8). A point of the hillslope and an on-bank (directly next to the channel) were selected (Figure 3-4). The on-bank point was located directly next to the channel of downstream tributary 3 and the hillslope point was located approximately 300 m uphill. The results illustrated for Scenario 1, show that on the hillslope surface saturation remains less than 20% over the one year period with an average depth to groundwater depth of 1.98 m (Figure 3-8a), while near the channel the saturation remains above 80% during the same period, with saturation becoming 100% during high flow periods (Figure 3-8b). The average depth to groundwater table was 0.03 m reflecting the behaviour of Scenario 1 shown in Figure 3-7.

Scenario 2 shows the hillslope becoming more saturated, where during dry periods saturation remains at approximately 10%, with jumps in saturation, up to 56%, observed during rainfall events. The average depth to groundwater table was 5.19 m (Figure 3-8c). Conversely, saturation next to the channel was approximately 60% during dry periods and becomes completely saturated during the high flow months, with a relatively shallow depth to groundwater table of 0.04 m (Figure 3-8d). Scenario 3, shows that on the hillslope the saturation reached a maximum of approximately 78% while average depth to groundwater table was 2.43 m (Figure 3-8e). Next to the channel, Scenario 3 illustrates that during wet periods the bank completely saturated, while during dry periods saturation ranged between 40% to 20%, with an average depth to groundwater table of 0.09 m.

Finally, Scenario 4 shows that surface saturation was highly responsive to rainfall with maximum saturation of 95% simulated at the end of July (note that the plot was at an instantaneous daily scale and that 100% saturation occurring at a sub-daily scale). The average depth to groundwater table was determined at 5.88 m. On-bank Scenario 4 behaves differently compared to other simulations, where continuous 100% saturation was not simulated over the high flow period. Additionally the period of high saturation (i.e. greater than 80%) was significantly shorter compared to other scenarios. For example, Scenario 4 had high saturation for approximately three months, while 100% saturation was simulated for Scenario 3 for over nine months.

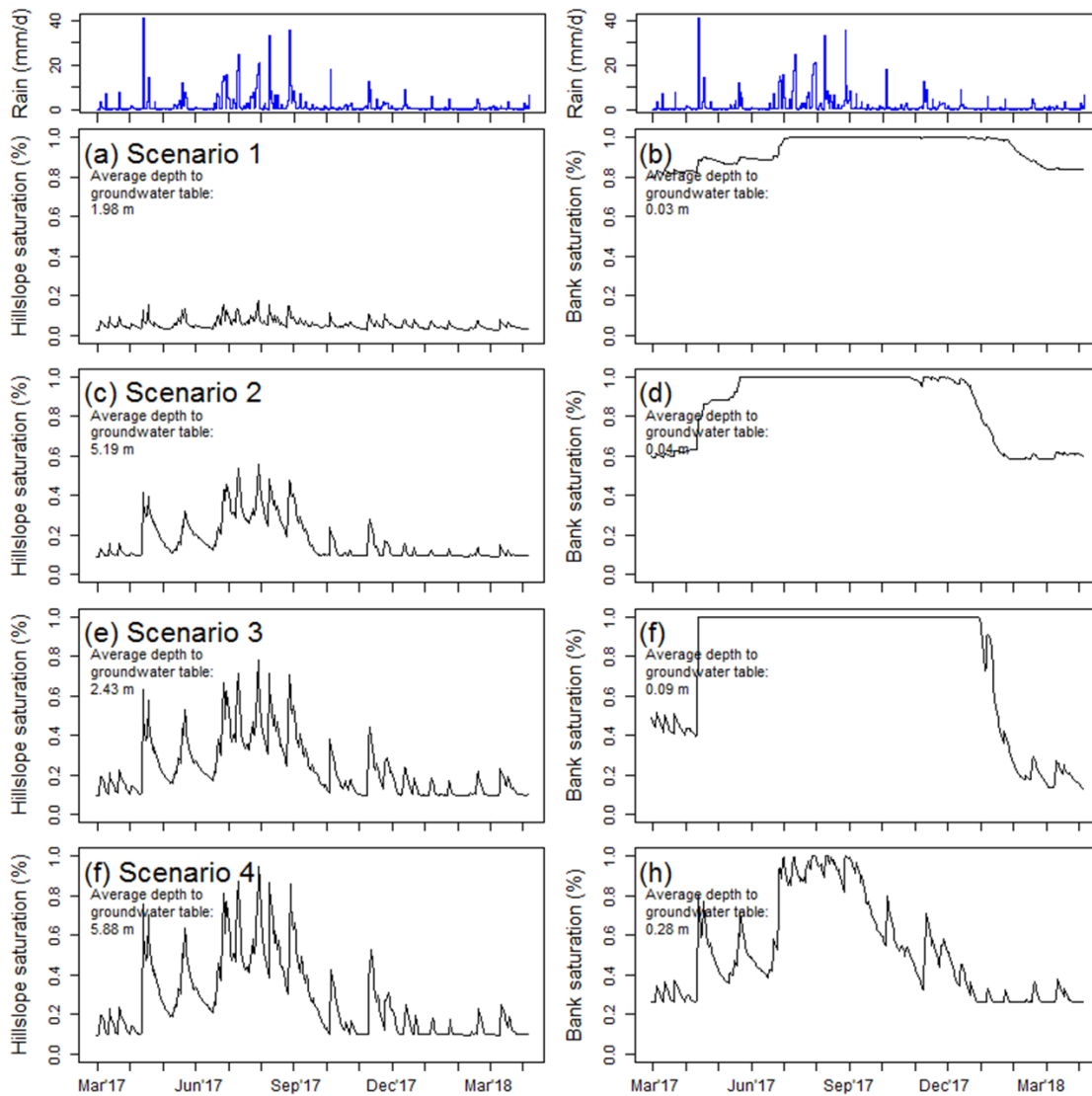


Figure 3-8: Comparison of hillslope and on-bank saturation for (a-b) Scenario 1, (c-d) Scenario 2, (e-f) Scenario 3 and (g-h) Scenario 4, with daily rainfall shown in the top panel.

To further compare the saturation behaviour of each of the scenarios, scatter plots are illustrated in Figure 3-9 showing the relationship of saturation (which leads to runoff) and 7-day antecedent rainfall. The 7-day antecedent rainfall was selected because saturation excess was influenced by long-term rainfall (e.g. 30-day rainfall), with infiltration excess was influenced by daily (or sub-daily) rainfall intensity and the 7-day rainfall provides a mid-range of comparative purposes. Figure 3-9 shows that there was a strong linear rainfall-saturation (or runoff) relationship of the hillslope ( $\sim 0.71$ ) for Scenario 4 (containing infiltration excess), while in contrast Scenario 1 (dominated by saturation excess), illustrates a weak linear relationship (0.35). This linear relationship strengthens for Scenario 2 ( $\sim 0.74$ ) and for Scenario 3 ( $\sim 0.68$ ).

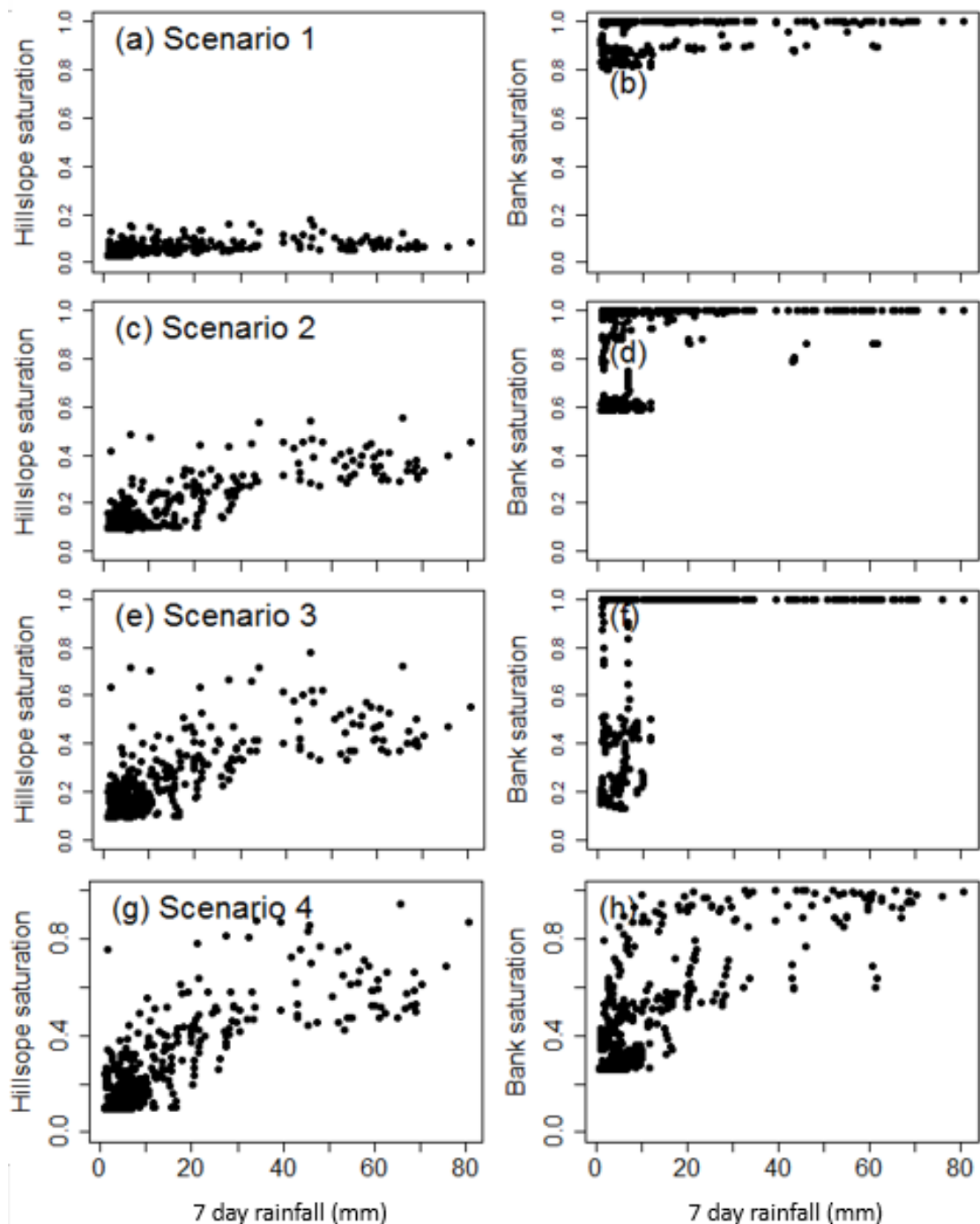


Figure 3-9: Comparison of scatter plots of the saturation behaviour on the hillslope and on the bank to 7 day antecedent moisture for (a-b) Scenario 1, (c-d) Scenario 2, (e-f) Scenario 3, and (g-h) Scenario 4.

#### 3.4.4 Comparison of flow at the local scale

There were differences in storage dynamics as well as inflows and outflows in and around the dam under the four scenarios (Figure 3-10). Each simulation was initialized with the 10ML in-stream storage at full capacity (depth of 2.6 m). A volume of 0.08 ML/day, approximately 0.02 m/day, was extracted from the dam on days where the 30 day antecedent rainfall was less than 50 mm. These conditions equate to a total volume of approximately 18 ML/year being extracted from the storage and 230 extraction days per year. These conditions were implemented in order to

realistically simulate water extraction behaviour, which would typically occur during dry periods of the year.

The total annual simulated inflow into the storage for scenarios 1, 2, 3, and 4 was 112 ML, 140 ML, 123 ML and 87 ML (Table 3-2). Despite having similar inflows into the dam, the total number of inflow days differ. For example, above the dam, Scenario 2 flows all year, while Scenario 4 flows for approximately 40% of the year. A comparison of the number of spill days per year shows that the number of days flow was present above and below the dam for Scenario 1 and 2, both dominated by saturation excess, were the same (325 days and 365 days respectively). However the results in Table 3-2 show that for Scenario 3, which was dominated by groundwater runoff the number of flow days below the dam increase by 38% compared to the number of flow days above the dam (257 days to 355 days). Similarly, for Scenario 4, which was also dominated by groundwater but also contained infiltration excess runoff, the number of flow days increased by 22.5% (141 days to 182 days). Therefore, simulated flow days for groundwater dominated scenarios increased more than 20% compared to upstream. This was significant because this infers that the groundwater component of runoff was bypassing flow around the dam via subsurface flow pathways in order to produce low flows. For example, the onset of streamflow below the dam in Scenario 3 (dominated by groundwater runoff) begins in late March, this was despite there being no spill from the dam, i.e. below full capacity (Figure 3-10).

The number of days per year for which the storage in Scenarios 1, 2, 3 and 4 was at capacity (>2.6 m depth) was 173 days, 332 days, 146 days and 77 days respectively (Figure 3-10). The results demonstrate a difference of 33% for the volume spilled for every spill day, 0.10 ML/spill day for Scenario 1 and 0.15 ML/spill day for Scenario 2 (Table 3-2). Scenario 3 shows that the dam became empty for 74 days of the year (Figure 3-10), while all other scenarios remain above 0.5 m full. Therefore, the volume of water extract from the storage for Scenario 4 was 17% less (15 ML extracted/year) compared to Scenarios 1 to 3 (18 ML extracted/year).

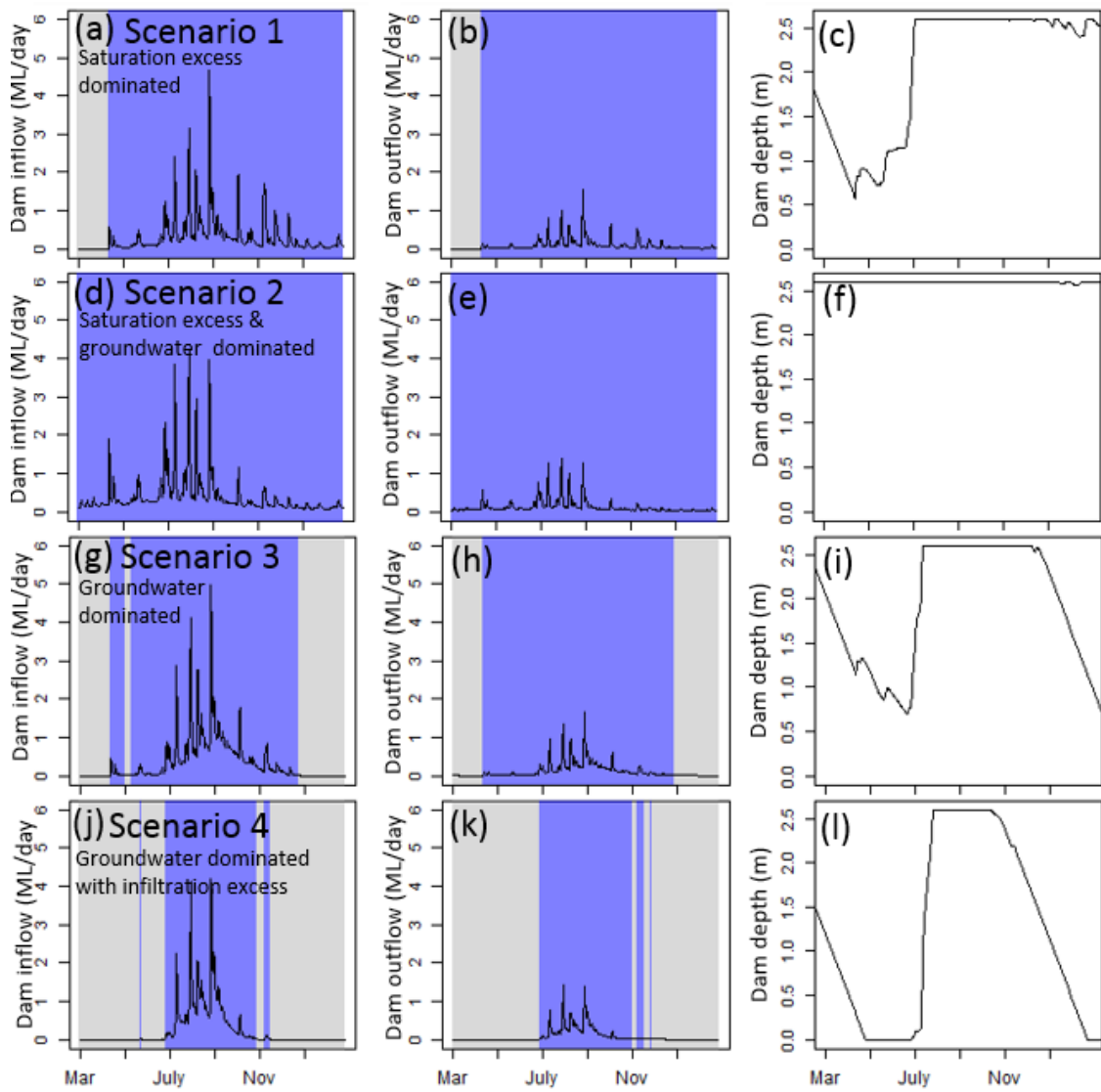


Figure 3-10: Comparison of the small storage (dam) inflows, outflows and water level depth for Scenario 1 (a-c), Scenario 2 (d-f), Scenario 3 (g-i) and Scenario 4 (j-l). The blue shading represents periods where flow was present and the grey shading represents periods when flow was not present in-stream.



Table 3-2: Comparison of flow and storage behaviour for the four scenarios

	<b>Scenario 1</b>	<b>Scenario 2</b>	<b>Scenario 3</b>	<b>Scenario 4</b>
	<b>SE dominated</b>	<b>SE&amp;GW dominated</b>	<b>GW dominated</b>	<b>GW dominated with IE</b>
Annual inflow (ML)	112	140	123	87
Number of annual inflow days	325	365	257	141
Annual outflow (ML)	34	40	43	28
Number of annual outflow days (spill days)	325	365	355	182
Percentage difference in inflows and outflows (%)	-69	-71	-65	-68
Percentage difference of flows days above and below dam (%)	0	0	+38	+22.5
Volume spilled per spill day (ML/day)	0.10	0.11	0.12	0.15
Day per year dam was at capacity (>2.6m)	173	332	146	77
Number of days dam was empty (days)	0	0	0	74
Number of days period year water was extracted	230	230	230	188
Total volume extract per year (ML)	18	18	18	15

### 3.5 Conclusions

This paper has compared four scenarios having very different proportion of runoff generation mechanisms (infiltration-excess, saturation excess and groundwater), yet retaining similar discharge performance at the outlet ( $NSE > 0.5$ ). Despite having similar NSE's at the outlet, differences were shown in numerous aspects of internal representation, including the spatial representation of saturation as well as the behaviour of inflows, outflows and storage levels of a small storage. The observed difference in flows for the four scenarios demonstrates the potential for hydrological modelling assumptions to arrive at different approaches for managing water resources, particularly those relating to environmental and low flows.

In terms of spatial variability, the results suggest that model parameterization and structures with a priori process assumptions (i.e. surface or subsurface flow-dominated) influence the estimation of flows throughout a catchment. This is significant because many catchment management practices require accurate predictions in both surface and subsurface domains. Examples include: increasing or decreasing flow, retarding flows, enhancing recharge and preventing recharge. The study showed that there were large differences in the catchment state of including saturation and depth to water table across the catchment. These processes influence how both lumped and distributed model structures should be developed and parameterized to effectively estimate runoff. Commonly applied conceptual models cannot adequately account for these differences in mechanism because of aggregation of processes to an outlet, as well as their inability to explicitly simulate subsurface flows (Li et al., 2013).

The influence on local flow pathways was shown to affect surface and subsurface flow volumes in and around an in-stream storage. The groundwater dominant scenarios illustrated that there was the potential of more than a 20% increase in the number of flow days directly below a storage compared to directly above due to subsurface generated runoff. The implication is that the representation of processes can be limited by the selection of the model. While process-based models are one way to address this challenge, the significant parameterization required for these

models mean that challenges remain validating these models at the local scale due to data limitations in upstream reaches. Further studies could investigate alternative data sources that provide information at the hillslope and on dam volumes (e.g. low-cost sensors, remote sensing, crowd sourcing) to augment data from downstream streamflow gauges.

Further studies could use the results from this investigation as a baseline to determine the influence that different volumes and patterns of extraction have on the dynamics of the storage. Alternative storage configurations, such as multiple storages in series and parallel can also be compared as well as the storage behaviour of in-stream versus off-stream. Research questions around the influence that increasing storage density per unit area has on flows upstream and downstream using process models is also significant in managing water resources across scales. While the example of in-stream storages was used in this study, there are other management questions that are influenced at multiple scales, such as changes in upstream land use or changes to the characterization of stream intermittency and low flows.

A possible limitation of this study was the assumption that the coupled HGS model is a realistic representation of reality. However, the benefit of physically based models is the ability to investigate processes across a wide range of scales, which is not otherwise feasible from empirical studies. To address this limitation, additional data could be collected and used to validate simulations or potentially eliminate candidate hypotheses. The study was conducted in a small 10 km<sup>2</sup> catchment, and there would be benefit from additional studies over larger regions or different catchment types. To better represent alternative runoff mechanisms, flexible model structures (Clark et al., 2015) could also be investigated that explicitly represent groundwater flows and coupled interactions.

The results highlight the need for greater attention to process modelling and better representation of flow dynamics at the local scale than aggregated models calibrated at an outlet. Additionally, the results show that the parameter uncertainty is so large that it allows equifinality of the model, where different processes emerged based on the fully integrated model structure providing the same downstream model performance. distinction is important for policy planning, as many engineering, ecological and water management problems exist at the local scale but can have catchment wide implications. With improved representation of hydrological complexity in space and time, planning and implementation of environmental water policies can be based on more accurate representation of flow stores and fluxes across all relevant scales.

## Acknowledgements

The authors acknowledge the support of the South Australian government Department for Environment and Water for identifying the case study and providing catchment data.

## References

- Abbott, M.M., Bathurst, J.C., Cunge J.A. O'Connell, P.E., and Rasmussen, J. 1986. A introduction to the European Hydrological System – Ststene Hydrologique European, 1: History and philosophy of a physically-based distributed modelling system, *J. Hydrol.*, 87, 45-59.
- Barré de Saint-Venant, A. J. C. 1871, Théorie du Mouvement Non Permanent des Eaux, avec Application aux Crues de Rivières et à l'Introduction des Marées dans leur Lit, C. R. Séances Acad. *Sci., Paris, Fr.*, 73(4), 237–240.
- Bockelmann, A. Zamfirescu, D. Ptak, T. Grathwohl, P. Teutsch, G. 2003. Quantification of mass fluxes and natural attenuation rates at an industrial site with a limited monitoring network: A case study. *Journal of contaminant hydrology*. Vol 60(1-2): 97-121. doi:10.1016/S0169-7722(02)00060-8
- Brunner, P., Simmons, C. T. 2012. HydroGeoSphere: A Fully Integrated, Physically Based Hydrological Model. *Groundwater*. 50 (2): 170–176. doi:10.1111/j.1745-6584.2011.00882.
- Campana, M. E., and E. S. Simpson 1984, Groundwater residence times and recharge rates using a discrete-state compartment model and C-14 data, *J. Hydrol.*, 72, 171–185, doi:10.1016/0022-1694(84)90190-2.
- Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski. D., Rupp, D. E. Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., Rasmussen, R. M. 2015. A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.* doi:10.1002/2015WR017200.
- Dunne, T. 1983. Relation of field studies and modeling in the prediction of storm runoff. *J. Hydrol.*, 65 (1–3) (1983), pp. 25-48, 10.1016/0022-1694(83)90209-3
- Graham, C. B., Woods, R. A. and McDonnell, J. J. 2010. Hillslope threshold response to rainfall: (1) A field based forensic approach. *J. Hydrol.*, 393: 65-76. doi:10.1016/j.jhydrol.2009.12.015.
- Hewlett, J.D. and Nutter, W.L., 1970. The varying source area of streamflow from upland basins. Interdisciplinary aspects of watershed management. Proc. Syrup. Bozeman, Mont., *Am. Soc. Civ. Eng.*, pp. 65--83.
- Horton, R.E. 1933. The role of infiltration in the hydrological cycle. *Trans AGU*, 14, 446-460, 193
- Kennard, M.J.; Pusey, B.J.; Olden, J.D.; Mackay, S.J.; Stein, J.L.; Marsh, N. 2010. Classification of natural flow regimes in Australia to support environmental flow management. *In Freshwater Biology*; Blackwell Publishing Ltd: Queensland, Australia; Volume 55, pp.171–193.
- Kirchner, J. W. 2016. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, 42(3).
- Inoubli, N., D. Raclot, I. Mekki, R. Moussa, and Y. Le Bissonnais. 2017. A spatiotemporal multiscale analysis of runoff and erosion in a Mediterranean marly catchment. *Vadose Zone J.* 16(12). doi:10.2136/vzj2017.06.0124 3.

Ivanov, V.Y, Vivoni, E.R., Bras, R.L., and Entekhabi, D. 2004. Catchment hydrological response with a fully-distributed triangulated irregular network model, *Water Resour. Res.*, 40(11), W11102, doi:10.1029/2004WR003218

Grantham, T. E., Merenlender, A. M., & Resh, V. H. 2010. Climatic influences and anthropogenic stressors: an integrated framework for streamflow management in, *Freshwater Biology* 55, 188–204. <https://doi.org/10.1111/j.1365-2427.2009.02379.x>

Larned, S.T., Datry, T., Arscott, D.B., Tockner, K. 2010. Emerging concepts in temporary-river ecology. *Freshwater Biology* 55: 717–738. DOI: 10.1111/j.1365-2427.2009.02322.x

Li, H., Sivapalan, M., & Tian, F. 2012. Comparative diagnostic analysis of runoff generation processes in Oklahoma DMIP2 basins: The Blue River and the Illinois River. *J. Hydrol.*, 418–419, 90–109. <https://doi.org/10.1016/j.jhydrol.2010.08.005>

Li, L. Maier, H. R. Lambert, M. F. Simmons, C. T. Partington, D. 2013. Framework for assessing and improving the performance of recursive digital filters for base-flow estimation with application to the Lyne and Hollick filter. *Env. Modelling & Software*. 41(2013) 163-175.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., Fleckenstein, J. H. 2013. Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, *Water Resour. Res.*, 49, 5501–5519. <https://doi.org/10.1002/wrcr.20405>

Puhlmann, H., Von Wilpert, K., Lukes, M., Droge, W. 2009. Multistep outflow experiments to derive a soil hydraulic database for forest soils. *European Journal of Soil Science* 60 (5), 792e806.

Richards, L. A. 1931. Capillary conduction of liquids through porous mediums, *J. Appl. Phys.*, 1(5), 318–333.

Schmocker-Fackel, P. Naef, F. Scherrer, S. 2007. Identifying runoff processes on the plot and catchment scale. *Hydrol. Earth Syst. Sci.*, European Geosciences Union, 2007, 11 (2), pp.891-906

Smith, M. B., K. P. Georgakakos, and X. Liang. 2004. The distributed model intercomparison project (DMIP), *J. Hydrol.*, 298, 1–3.

Therrien, R. McLaren, R. G. Sudicky, E. A. Panday, S. 2009. HydroGeoSphere. A Three-dimensional Numerical Model Describing Fully-integrated Subsurface and Surface Flow and Solute Transport (Manual). *Groundwater Simulations Group*, University of Waterloo.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S., & Srikanthan, S. 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resour. Res.*, 45(12), 1-22.

VanderKwaak, J.E. and Loague, K. 2001. Hydrologic-response simulation for the R-5 catchment with a comprehensive physical-based model. *Water Resour. Res.*, 37(4), 999-1013, 2001

Weiler, M. McDonnell J. J. 2007. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes. *Water Resour. Res.*, 43 (2007), p. W03403, 10.1029/2006WR004867

## Appendix 3-A

### HMC method

The hydraulic mixing-cell (HMC) method developed by *Partington et al.* (2011, 2013) allows for delineation of simulated runoff mechanisms through an entire catchment. For the purpose of delineating runoff mechanisms, any water that enters as rainfall was tracked throughout the model domain. In this study, when rainfall enters the model domain during each time-step of the simulation, ponded water (water in the rill storage) was identified as either being infiltration excess (IE) or saturation excess (SE) (e.g. *Gutierrez-Jurado et al.*, 2020). It was at this point that incoming water from boundary conditions (i.e. rainfall) was tagged as either infiltration excess or saturation excess water to be further tracked and delineated during simulations. Figure 3-11 illustrates that to define the ponded fraction of water, the model checks the state of groundwater in the cell, at that time step, if the water table was below the surface domain, the ponded water was tagged as infiltration excess ponding (i.e. IE-ponding), if the water table was at the surface, the ponded water was tagged as saturation excess ponding (i.e. SE-ponding). Tagging of the inflowing water within a simulation timestep as IE-ponding or SE-ponding was a local temporary tagging to classify the water in the particular inflowing cell only. Any movement of that tagged water to an adjacent cell will result in a reassignment of the water to something more meaningful with respect to runoff as follows.

The IE-ponding and SE-ponding fractions were further tracked over each time step and delineated as either infiltration (i.e. IE-infiltration and SE-infiltration) or overland flow (IE-overland flow and SE-overland flow). If the IE-infiltration and SE-infiltration fractions of water discharge to the overland domain after tracking a shallow subsurface pathway, they were delineated as return flow (i.e. IE-return flow and SE-return flow). Alternatively, if these fractions of infiltrated water undergo deep infiltration, they were delineated as groundwater. Water that was present within the model domain at simulation initialization was assumed to be groundwater. All cells in the model were defined as subsurface cells, overland cells or channel cells. For ponding that occurs in a cell defined as a channel cell, the fraction of water was defined as 'rainfall direct to channel'.

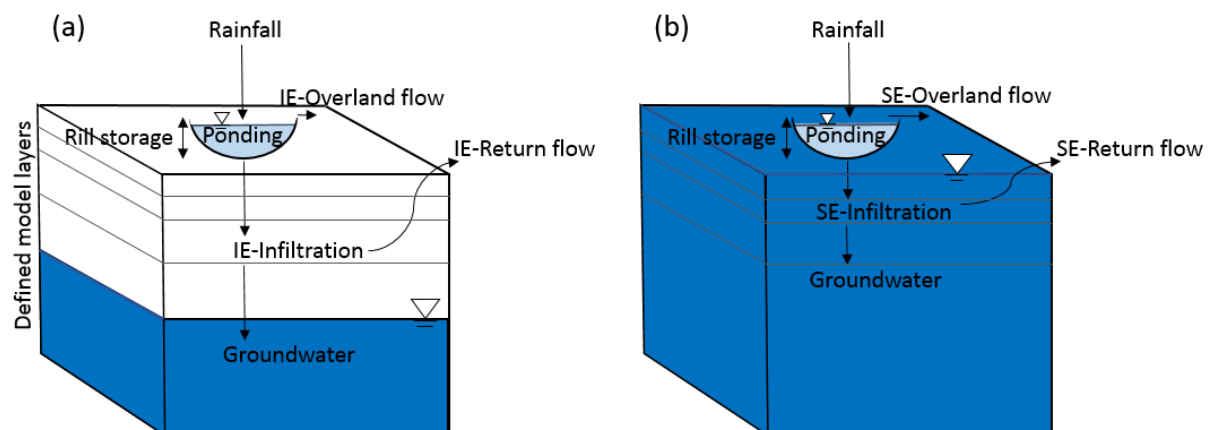


Figure 3-11: Illustrates soil column of a single cell in the model where (a) shows the tagging of infiltration excess (IE) fractions and (b) shows the tagging of saturation excess (SE) fractions. The return flow for IE and SE can re-emerge in any downstream cell providing that deep infiltration does not occur, causing the fraction of water to be tagged as groundwater.

## HMC fraction definitions

The unique fractions used in this study to determine the contributions of different flow generation mechanisms to total runoff are defined in Table 3-3.

Table 3-3: Definitions of the unique fractions tracked in the HMC method that are updated at each time step.

Flow generation mechanism delineated fractions	Unique fraction	Definition
Infiltration excess overland flow	f_IE-OL	The fraction of water of infiltration excess runoff that is overland runoff
Saturation excess overland flow	f_SE-OL	The fraction of saturation excess runoff that is overland runoff
Infiltration excess infiltration	f_IE-infil	The fraction of water of infiltration excess ponding that is infiltrated in a cell
Saturation excess infiltration	f_SE-infil	The fraction of water of saturation excess ponding that is infiltrated in a cell
Infiltration excess return flow	f_IE-OL-RF	The fraction of water of infiltration excess infiltration that re-emerges downstream as overland return flow
Saturation excess return flow	f_SE-OL-RF	The fraction of water of saturation excess infiltration that re-emerges downstream as overland return flow
Groundwater flow	f_GW	The fraction of water of groundwater discharge
Rainfall direct to channel	f_river_flux	The fraction of water that enters a channel while ponded
Initial water	f_initial	The fraction of water that was present in the model domain at the time of simulation initialisation.

## HMC mathematical formulation

For each package of water entering the model defined from boundary conditions (e.g. rainfall), the water volume is assigned a unique fraction  $f$  (defined in Table 3-3). For each cell and over each time step, fractions assigned to the boundary condition flows (i.e.  $BC_{in}$  or  $BC_{out}$ ), the inflow and outflow from subsurface (i.e. *subsurface*  $Q_{in}$  or *subsurface*  $Q_{out}$ ) and the inflow and outflow from neighboring surface cells (i.e. *neighbour*  $Q_{ij}$ ) are summed for all fractions. The sum of all fractions is equal to 1 for an error free fluid mass balance (Figure 3-12).

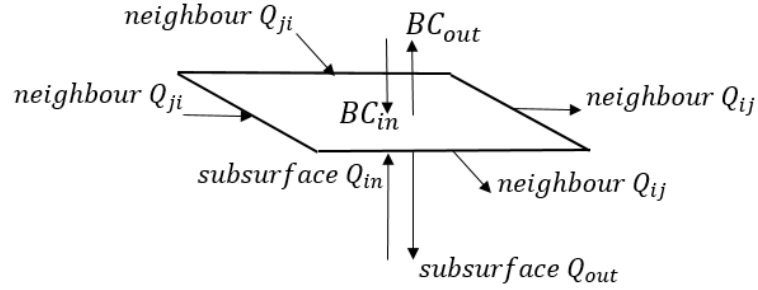


Figure 3-12: For each cell in the model a fluid mass balance is applied to determine the fractions of runoff generation mechanisms with the example of a surface cell shown

The method begins with the conservation of mass equation calculated for each cell:

$$\frac{\Delta S}{\Delta t} = Q_{in} - Q_{out} \quad (1)$$

$$\Delta S = Q_{in}\Delta t - Q_{out}\Delta t \quad (2)$$

$$V^{N+1} = V^N + Q_{in}\Delta t - Q_{out}\Delta t \quad (3)$$

where at time step  $\Delta t$ , the change in storage  $\Delta S$  is the sum of all the inflows and outflows,  $Q_{in}$  and  $Q_{out}$ . The volume of water storage  $V$  in a cell for a given time period  $N$  is determined by the addition of the change in storage to the volume of the cell in the previous time step.

The inflows and outflows for a cell  $i$  include all the fluxes from neighboring cells  $Q_{ij}$  and  $Q_{ji}$ , and boundary conditions  $Q_{BC}$ .

$$V^{N+1} = V^N + \sum_{\forall j \in n} Q_{ji}\Delta t + \sum_{\forall BCin} Q_{BCin}\Delta t - \sum_{\forall j \in m} Q_{ij}\Delta t - \sum_{\forall BCout} Q_{BCout}\Delta t \quad (4)$$

where  $n$  denotes the set of inflowing neighboring cells into cell  $i$  and  $m$  denotes the set of outflows from cell  $i$  into neighboring cells. The  $ij$  subscripts denotes the volume into cell  $j$  leaving cell  $i$  and subscript  $ji$  denotes the volume of neighbor cell  $j$  into cell  $i$ .

Applying the modified mixing-cell approach of Campana and Simpson (1984), each fraction  $f$  for the runoff generation mechanism  $k$  in cell  $i$  is calculated as:

$$f_{i(k)}^{N+1}V^{N+1} = f_{i(k)}^N V^N + \sum_{\forall j \in n} Q_{ji}\Delta t f_{j(k)}^N + \sum_{\forall BCin} Q_{BCin}\Delta t - f_{i(k)}^N \sum_{\forall j \in m} Q_{ij}\Delta t - f_{i(k)}^N \sum_{\forall BCout} Q_{BCout}\Delta t \quad (5)$$



where  $f_{i(k)}^N$  denotes fraction  $k$  at time  $N$  in the cell  $i$  and  $f_{j(k)}^N$  in the neighboring cell  $j$ . The superscript of the volume  $V$  denotes the time state.

The volume  $V$  of water for the current time step  $N + 1$  is calculated by summing the volume of the cell in the previous times step, denoted as  $V^N$  and the inflow and outflow volumes from neighboring cells  $V_{ji}^{n+1}$  and  $V_{ij}^{n+1}$  and the inflows and outflows from the boundary conditions  $V_{BC}^{n+1}$ .

$$V^{N+1} = V^N + V_{BCin}^{N+1} + V_{ji}^{N+1} - V_{ij}^{N+1} - V_{BCout}^{N+1} \quad (6)$$

Therefore, for any given fraction,  $f_{i(k)}^{N+1}$  is calculated as:

$$f_{i(k)}^{N+1} = \frac{f_{i(k)}^N V^N}{V^{N+1}} + \frac{\sum_{\forall j \in n} Q_{ji} \Delta t f_{j(k)}^N}{V^{N+1}} + \frac{\sum_{\forall BCin} Q_{BCin} \Delta t}{V^{N+1}} - f_{i(k)}^N \left( \frac{\sum_{\forall j \in m} Q_{ij} \Delta t}{V^{N+1}} + \frac{\sum_{\forall BCout} Q_{BCout} \Delta t}{V^{N+1}} \right) \quad (7)$$

## References

Campana, M. E., and E. S. Simpson (1984), Groundwater residence times and recharge rates using a discrete-state compartment model and C-14 data, *J. Hydrol.*, 72, 171–185, doi:10.1016/0022-1694(84)90190-2.

Gutierrez-Jurado, K. Y., Partington, D., Batelann, O., Cook, P., Shnafield, M. 2019. What triggers streamflow for Intermittent Rivers and Ephemeral Streams in Low-Gradient Catchments in Mediterranean Climates. *Water Res. Res.* doi.org/10.1029/2019WR025041

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., Fleckenstein, J. H. (2013). Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, 49, 5501–5519. *Water Res. Res.* doi.org/10.1002/wrcr.20405

## Chapter 4.

# Does intermittency data from headwater catchments improve multi-site hydrological model calibration?

Alicja Makarewicz<sup>a</sup>, Michael Leonard<sup>a</sup>, Seth Westra<sup>a</sup>, Daniel Partington<sup>b</sup>

<sup>a</sup> School of Civil, Environmental and Mining Engineering, the University of Adelaide, North Terrace, Adelaide SA 5005, Australia.

<sup>b</sup> National Centre for Groundwater Research & Training, Flinders University, Adelaide SA 5001, Australia

*Environmental modelling and software*, in preparation



# Statement of Authorship

Title of Paper	Does Intermittency data from headwater catchments, applied to multi-site model calibration, improve simulated catchment internal functioning?
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Intention to submit to Environmental Modelling and Software

## Principal Author

Name of Principal Author (Candidate)	Alicja Makarewicz		
Contribution to the Paper	Built and calibrated model, analysis and interpretation of results, wrote the manuscript		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	03/3/2020

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Michael Leonard		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	26/03/2020

Name of Co-Author	Seth Westra		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	25/3/2020

Name of Co-Author	Daniel Partington		
Contribution to the Paper	Analysis and interpretation of results, editing the manuscript		
Signature		Date	2/4/2020

Please cut and paste additional co-author panels here as required.

## Abstract

Calibrating physically-based hydrological models is difficult due to the significant heterogeneity of many catchments, leading to complex non-linear dynamics across nested spatial and temporal scales. Without sufficient data it can be difficult to resolve this multi-scale variability, and/or distinguish between multiple competing hypotheses of runoff generation. Beyond conventional streamflow measurements, inexpensive data collection methods can provide valuable ancillary data that can be used to constrain model parameters and allow the model to more faithfully represent key hydrological processes. The value of additional data was demonstrated on a case study catchment in South Australia. A multi-hypothesis framework was adopted where initially four candidate calibrations of the physically based model HydroGeoSphere—each representing a different conceptualization of runoff but with reasonable performance at the catchment outlet—were identified. When evaluated at nine upstream sites, it was shown that no single candidate calibration performed consistently well. The results were compared with a locally calibrated model that also incorporated additional upstream information on binary ‘wet’-‘dry’ (flowing-not flowing) state using low-cost temperature sensors. The outcomes were that eight out of nine sites showed significant improvement in performance, with only a small deterioration to the calibration results at the outlet. These results highlight that as sensing and communications technologies continue to improve, there will be increasing opportunities to use information sources such as local-scale intermittency to supplement reliable streamflow records for representing hydrological processes across scales. With improved model performance there can be many benefits of getting the internal processes right, such as: simulating conditions that are outside a calibration period or infilling and completing data sets.

## 4.1 Introduction

Water management decisions such as those related to environmental flows, water quality, land use and dispersed water stores often require information on hydrological processes throughout a catchment, not just at scales that are routinely measured using streamflow gauging's. Accurate representation of hydrological processes across multiple scales is critically important for local scale water resource problems that can have cascading implications across the catchment, from local reaches and hillslopes through to more flows further downstream. Because physically based models explicitly solve the fundamental equation of water movement overland and through soils, they are in theory able to represent processes across multiple scales. However, data scarcity means that in practice they are often limited in their capacity to identify parameters at all locations. The outcome is that although these models can be calibrated to streamflow at a particular gauge, the dynamics upstream responsible for streamflow generation may not be correctly represented (*Gaukroger and Werner, 2011; Li et al., 2015; Wellen et al., 2015*). One option to remedy this issue is to collect additional data or exploit data already available but under-utilized. Such data need not be of the same nature as traditional long-term sources such as continuous streamflow measurements at the outlet, but can be collected for a limited time from low-cost sensors.

Physically based models have the potential to add a lot of value because they are able to represent heterogeneous processes with variations at all scales influencing flow pathways and runoff production due to characteristics such as catchment size, slope, topography, soil and vegetation features (*Mirus & Loague, 2013*). However, physically based models are limited in their ability to deal with heterogeneity due to the 'curse' of dimensionality and the impossibility of parameter calibrations (*Beven, 1989*). This is because catchment-scale models only have a partial view of the physical processes and miss essential characteristics of the functioning of the catchment that is influenced by small scale complexity, non-stationarity and non-linearity of water stores and fluxes (*Soulsby et al., 2015; Blumstock et al., 2015; Kirchner, 2006*). For example, ecohydrological understanding of network intermittency, which shapes local and catchment-wide aquatic ecosystems (*Larned et al., 2010*). Similarly, understanding the generation of sediment and associated water quality issues requires multiscale process knowledge, where factors controlling runoff generation and associated sediment transport vary according to the spatial scale (*Inoubli et al., 2017*). The representation of these flow patterns remains a challenge because many numerical parameters cannot be adequately measured, with complex models having high uncertainty and prone to equifinality (*Beven, 1993*).

The scarcity of data at sites distributed throughout a catchment, let alone the vast number of ungauged catchments, causes significant uncertainty in the representation of hydrological processes. In order for physically based models to capture catchment heterogeneity, they need local scale data. Some datasets such as high resolution topography using lidar (*Brubaker et al., 2013; Thomas et al., 2017*), or other spatial data such as vegetation (*Brubaker et al., 2014; Chance et al., 2016*) now are commonly available, but datasets that reflect catchment fluxes (e.g. flows) are still not available at an appropriate scale. For example, stream flow intermittency is rarely observed or measured on headwater reaches (typically first to third order streams) despite their ecological importance in many arid and semi-arid environments (*Larned et al., 2010*).

Recent technological advances present opportunities to strategically collect additional data using increasingly cost effective means enabling high-density and widespread environmental data collection. For example, low-cost environmental sensors (*Lovett et al., 2007; Ruiz-Garcia et al.,*

2009; Wickert *et al.*, 2018) and citizen science data (Turner *et al.*, 2011; Le Coz *et al.*, 2016; David *et al.*, 2018), while often less accurate, have the potential to improve model calibration. This data is complementary to existing high-quality data sources such as gauged records, and thus calibration methods that balance the relative strengths of the alternative data sources are likely to outperform those that rely on a single data source in isolation. This is especially important for representation in headwaters where data scarcity is more pronounced.

While there is a clear imperative and opportunities for additional data, there is a strong parallel need to improve calibration methods to manage and exploit all available data. A significant challenge is the demonstration of multi-objective and multi-site calibration methods that can account for trade-offs between overall catchment performance and local biases, since improved spatial representation can risk a deterioration of performance at the outlet. Techniques for efficient optimization are important given the computational demand of many models and the potential burden of managing increased volumes of data within calibration. The provenance and quality of data is an important consideration, especially when there are multiple sources, and the errors, biases and level of uncertainty in each source can have a significant impact on the calibration process.

To avoid unwanted effects of equifinality and to fully utilize additional data requires the application and improvement to methods for multi-site calibration. To this end, this study seeks to demonstrate the extent to which data collected in a headwater catchment, in particular low-cost intermittent data, can be applied to improve model calibration. The specific objectives were:

1. To quantify the upstream performance of multiple candidate calibrations based on an outlet-only calibration method;
2. To illustrate a multi-site calibration method for a physically based model that utilizes additional collected data to improve the representation of upstream physical flow processes; and
3. To compare the performance of the outlet-only and multi-site calibrated models.

To achieve these aims a numerical modelling study of a 10 km<sup>2</sup> South Australian catchment was used where a two year data collection campaign had provided daily 'wet'-'dry' binary classifications of stream-flow intermittency at multiple locations (Section 4.2). Here, the data's usefulness was evaluated by complementing outlet flow measurements and constraining a model's solution space in order to identify a more representative model configuration. The fully integrated hydrological model, HydroGeoSphere, was selected to simulate flow processes and was calibrated to multiple upstream sites while preserving the Nash-Sutcliffe coefficient at the outlet (Section 4.3). The results are used to illustrate the potential benefit of additional data for process understanding (Section 4.4). Discussion and implications of the comparison are outlined along with conclusions of the study (Section 4.5).

## 4.2 Case study and data

### 4.2.1 Catchment characteristics

A Mediterranean headwater catchment located in South Australia was selected to assess the usefulness of additional data for more accurate representation of hydrological processes (Figure 4-1). The catchment, approximately 10 km<sup>2</sup> in size, has three main tributaries with a total length of approximately 5 km and with vegetated areas that cover a large proportion of two of the tributaries. The elevation ranges from 175 m to 420 m above sea level and contains mainly fractured rock aquifers with a shallow to moderately thick topsoil layer of acidic, sandy loam soils.



The reaches range from first- to fourth-order streams and flow mostly in the winter/spring months, with limited or no flow in the summer and autumn. Based on fieldtrips and interviews with landowners, it was qualitatively known that some sites have significant recharge while other sites retain a permanent trickle due to fractured rock prominent in the area. The wider region is an important water resource to local properties for domestic water use, irrigation of crops and stock as well as providing water for environmental purposes, with environmentally significant assets such as Fleurieu Peninsula swamps and numerous pools/springs located downstream.

The catchment contains shallow soils with acidic loam over clay on rock. The vegetated area is 40%, with closed and open woody trees, with the remaining catchment covered with grass (8%), pastures (50%) and sparse woody trees (2%). A majority of woody trees have a root depth of at least 2 m (Canadell *et al.*, 1996) with native tree of the area (e.g. Banksia) growing as much as 2.4 m (Specht and Rayson, 1957) and Eucalyptus growing more than 15 m in depth (Kimber, 1974). The average slope in the catchment is 5.8% and two of the main tributaries have an average slope of 6% or greater due to the steepness around the top boundary. The average channel slope ranges between 2% and 4%.

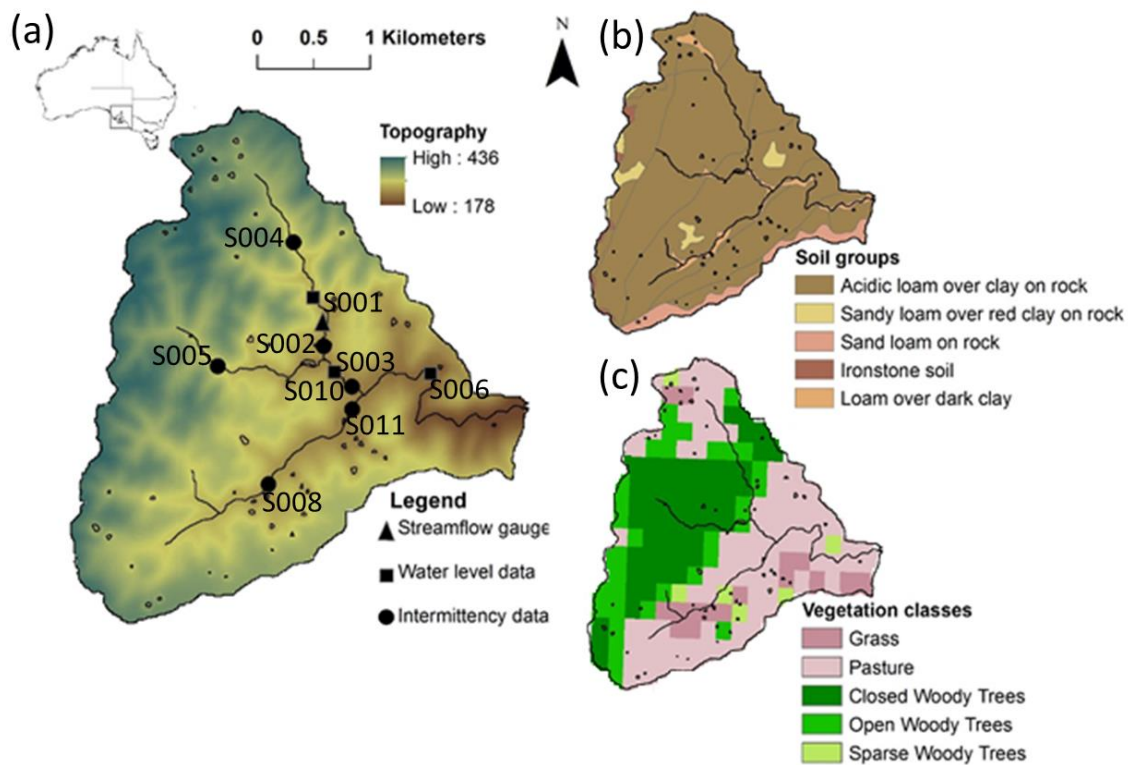


Figure 4-1: The case study is located in (a) South Australia, South East of Adelaide with 10 sites instrumented within the catchments three main tributaries and outlet. The 10 km<sup>2</sup> headwater catchment contains (b) loam and clay soil parameters and is (c) partially vegetated.

#### 4.2.2 Meteorological data

Rainfall and potential evapotranspiration data were obtained from the nearest weather station, 20 km southeast of the catchment. Given the small catchment size, the climatic data were assumed to be uniform over the catchment. A two-year period from March 2017 to March 2019 was selected for analysis, where the first year was used for calibration and the second for evaluation. This period was selected as it corresponds to a field campaign of data collection (Makarewicz *et al.*, 2020a). The average rainfall over the period was 780 mm/year rain and

1800 mm/year average potential evapotranspiration. The region is highly seasonal with hot dry summers, and while large rainfall events are possible, in summer there is negligible flow at most sites, other than immediately following rainfall events. The majority of the rainfall arrives in the winter months, and the highest flows occur during the spring period.

#### 4.2.3 Discharge and intermittency data

There were seven upstream sites and the outlet within the catchment with a majority of sites containing two years of continuous measurements between 01/03/2017 to 01/3/2019 (Figure 4-1a). Additional sites S010 and S011 each contain one year of data, available for the second year of the simulation period. Discharge data was available for sites S001, S003 and S006 (the outlet). Additional data (binary ‘wet’-‘dry’ classifications) was available for sites S004, S005, S008, S010 and S011.

Data was collected using low-cost temperature sensors installed on streambeds to determine when a stream was flowing (‘wet’) or not flowing (‘dry’). The observations were previously classified using a two-state hidden Markov model to derive a time series of binary ‘wet’-‘dry’ values, with a 92% classification accuracy (*Makarewicz et al., 2020a*). The emphasis of the data collection was to use low-cost techniques, with the trade-off that flow magnitude was not available for calibrating the model.

#### 4.2.4 Calibration and evaluation time-series

To calibrate and evaluate simulations, two years of observations (01/03/2017 to 01/3/2019) were used in the study. Two types of data were applied: discharge data and additional binary ‘wet’ – ‘dry’ classifications. The calibration strategy was designed such that both temporal and spatial performance of simulations were evaluated. A total of five sites were used for model calibration. This comprised four upstream sites that contained information on binary classifications (S001, S004, S005 and S008). This data was split with the first year used for calibration (01/03/2017 to 01/3/2018) and the second year reserved for evaluation (01/03/2018 to 01/3/2019).

To evaluate how effectively the use of only binary data upstream is at estimated flow volumes, discharge data available upstream (i.e. S001 and S003) was used exclusively for evaluation. Additionally, binary classification as sites S003, S010 and S011 were reserved exclusively for model evaluation to determine the spatial performance of the simulations. A detailed list and summary of data used for model calibration and evaluation is illustrated in Table 4-1.

*Table 4-1: Description of data available for each site and the calibration and evaluation periods applied to analysis*

Site	Data Type	Calibration range	Evaluation range
Outlet (S006)	Discharge	March 2017-March 2018	March 2018-March 2019
Outlet (S006)	Binary	March 2017-March 2018	March 2018-March 2019
S001	Binary	March 2017-March 2018	March 2018-March 2019
S001	Discharge	-	March 2017-March 2019
S002	Binary	March 2017-March 2018	March 2018-March 2019
S003	Binary	-	March 2017-March 2019
S003	Discharge	-	March 2017-March 2019
S004	Binary	March 2017-March 2018	March 2018-March 2019
S005	Binary	March 2017-March 2018	March 2018-March 2019
S008	Binary	March 2017-March 2018	March 2018-March 2019
S010	Binary	-	March 2018-March 2019
S011	Binary	-	March 2018-March 2019

## 4.3 Methodology

### 4.3.1 Model parameterisation

HydroGeoSphere (HGS) (Therrien *et al.*, 2009) is a fully integrated surface-subsurface model that can simulate surface and subsurface flows at high spatial and temporal resolutions. HGS simulates 3D variably saturated subsurface flow using a modified Richards' equation, and 2D surface flow (Brunner *et al.*, 2012) using the diffusion wave approximation to the Saint Venant equations. HGS solves all governing flow equations simultaneously to simulate surface and groundwater flows to the stream as a function of the catchment physical characteristics and hydrological inputs. The equations were solved using a control-volume finite element method with Newton-Raphson linearization.

The model domain was discretised as an irregular triangular mesh to simulate small-scale runoff in and around the channels. The element lengths range from 250m at the outer boundary to 25m in and around the catchment streams. The spatial discretization of the modelled domain was represented with 12,015 elements and 6,115 nodes (Figure 4-2). In the z-direction the model domain consists of seven layers with a discretization of 0.6 m for the top 3.0 m. Beneath the top 3.0 m, the soil extends a further 5.0 m to 24.0 m below the surface to represent the characteristics of the area, with greatest depths at the boundary and linearly becoming shallower towards the outlet. The outlet was set as the critical depth boundary and a no-flow boundary was fixed for the bottom and lateral subsurface domain (i.e. water can only leave the model domain through a critical depth boundary at the outlet). The model simulation uses sub-daily adaptive time steps.

The vegetation and evapotranspiration parameters were zonally applied using vegetation zones shown in Figure 4-1c. These parameters were grouped into two categories, (1) vegetated area and (2) cleared area (Figure 4-2). The vegetated areas correspond to locations where closed and open woody trees are located within the catchment. The remaining catchment area contains grass and pastures and was categorized as cleared. The selected root and evapo-transpiration depth for the vegetated and clear areas are 3.5 m and 0.2 m based on the type of vegetation present in the catchment (discussed in Section 4.2.1). The overland properties were adopted from Panday and Huyakorn (2004), with the exception of the channel Manning's roughness coefficient. The rill storage, which provides a threshold of flow, was set as 1.0 mm and the overland Manning's roughness was set as  $0.15 \text{ s/m}^{1/3}$ . The channel roughness parameter was set to  $0.05 \text{ s/m}^{1/3}$  which is typical for a natural channel, with light vegetation and pools.

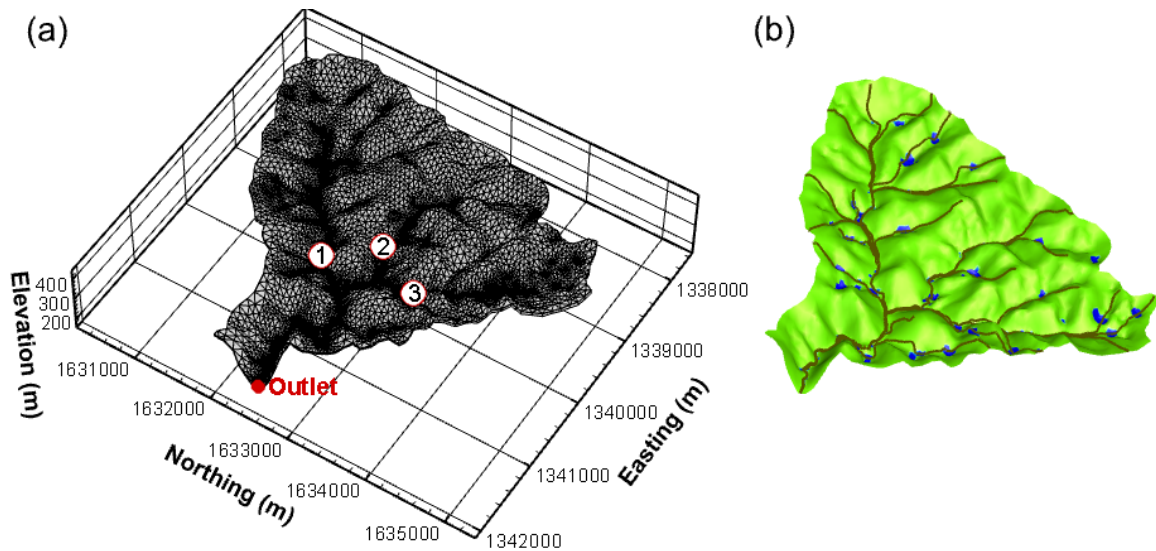


Figure 4-2: (a) Modelled catchment geometry and discretization showing smaller elements located in the channels with numbers 1-3 indicating the three main tributaries. (b) Simulated overland flow channels (grey lines) and ponded water in the catchment (blue shading).

#### 4.3.2 Outlet-only model calibration and evaluation

To investigate the performance of multiple conceptual models (i.e. different parameterizations that can represent different process conceptualizations of the system) based on an outlet-only calibration, four alternative calibrations were applied. The four parameterisations, here referred to as ‘*Scenarios*’, were previously demonstrated to achieve a Nash-Sutcliffe efficiency coefficient of 0.5 at the outlet while having different process representations on the hillslope (Makarewicz *et al.*, 2020b), comprising different proportions of saturation excess, infiltration excess and groundwater runoff. This led to the following parameterisations:

- **Scenario 1** was dominated with saturation excess runoff and contains groundwater;
- **Scenario 2** was dominated by both saturation excess and groundwater runoff;
- **Scenario 3** was dominated by groundwater runoff and contains saturation excess runoff; and
- **Scenario 4** was dominated by groundwater runoff and contains both infiltration excess and saturation excess runoff (outlined in Appendix 4-A).

See Makarewicz *et al.* (2020b) for further details on the model parameterization and calibration process.

#### 4.3.3 Multi-site model calibration and evaluation

To demonstrate the influence of additional ‘wet’-‘dry’ binary data on the calibration performance and flow representation of upstream reaches, the catchment was divided into five sub-catchments. The outlets of the sub-catchments each correspond to the sites used for calibration: S001, S004, S005, S008 and the overall catchment outlet (see Section 4.2.1). Figure 4-3 illustrates the method for calibration to multiple sites, and comprises nine steps and two iteration points.

*Step 1: Determine which parameters should remain ‘fixed’ and which parameters should be ‘varied’:* This decision was based on previous work (Makarewicz *et al.*, 2020b) and all parameters were fixed except for the soil characteristics (i.e. lateral and vertical hydraulic conductivity. Catchment characteristics such as vegetation parameters were ‘fixed’ on estimated based on available data as outlined in Section 4.2.2, and were not modified during model calibration.

*Step 2: Identify initial values and bounds for the calibration parameters:* The lateral ( $K_{xy}^i$ ) and vertical ( $K_z^i$ ) hydraulic soil conductivity were identified as the two key parameters for calibration in each sub-catchment since they strongly influence runoff response (Mirus & Loague, 2013). To reflect the realistic conditions in the catchment, the bounds on parameter values were based on soil properties ranging from clay to sandy loam (Puhlmann et al., 2009). The initial soil properties for each sub-catchment were based on the preferred values from modelling conducted for outlet-only calibration (Makarewicz et al., 2020b).

*Step 3: Spin-up model with selected parameters until dynamic equilibrium is reached:* For physically based models a long spin-up is generally required to allow the initial and boundary conditions to run forward and stabilize. In this case, a 20 year period was selected and initialized with soil properties with the preferred values. The spin-up contained observed climate data for the region (1985-2005) and was used to allow for the water table to adjust to the soil properties applied for each sub-catchment. Whenever the soil properties were altered during calibration, the model required an additional spin up period to allow for some adjustment of the subsurface state. A one year spin-up period was used due to the similarity of soil hydraulic conductivities and simulation time restrictions.

*Step 4 to Step 8: Calibrate parameters for each sub-catchment:* A single model evaluation for sub-catchment  $i$  ( $i=1$  to 5 sub-catchments) proceeds by specifying an instance of the two hydraulic conductivity parameter values (all other parameters fixed). A grid-based search was used where the range of each parameter was discretised into 10 points uniformly between bounds. Given there were two parameters to calibrate, this results in 100 parameter combinations to evaluate. For each of the 100 simulations, streamflow intermittency was calculated for comparison to observed streamflow intermittency. An objective function was used to determine the best performing parameter combinations from the 100 instances. An evaluation was made whether the identified region has suitably converged based on the performance criteria outlined in the next section (Section 3.4). When the parameter range has not converged, subsequent iterations of the grid-search were performed by selecting the best performing subdomain of the parameter space and further discretising that region. Once a parameter set for a sub-catchment was calibrated, they were 'fixed' for all remaining calibration runs.

*Step 9: Evaluate the model:* The second year of the final simulation was used solely for performance evaluation (Table 4-1). The selected parameters for the model calibrated to multiple sites are presented in Table 4-2. While a higher hydraulic conductivity in the vertical than horizontal direction is uncommon, this feature was plausible for the region given the clayey shallow soils common in the study area, since vertical cracking, due to drying, would allow for rapid vertical flow.

All simulations were conducted on a High-Performance Computer (HPC). The approximate runtime for a two year simulation was approximately 48-72 hours and was dependent on soil properties being simulated. The simulations required for calibration comprised 100 parameter combinations multiplied by several grid search iterations for each sub-catchment, and took approximately 3000 days of CPU time.

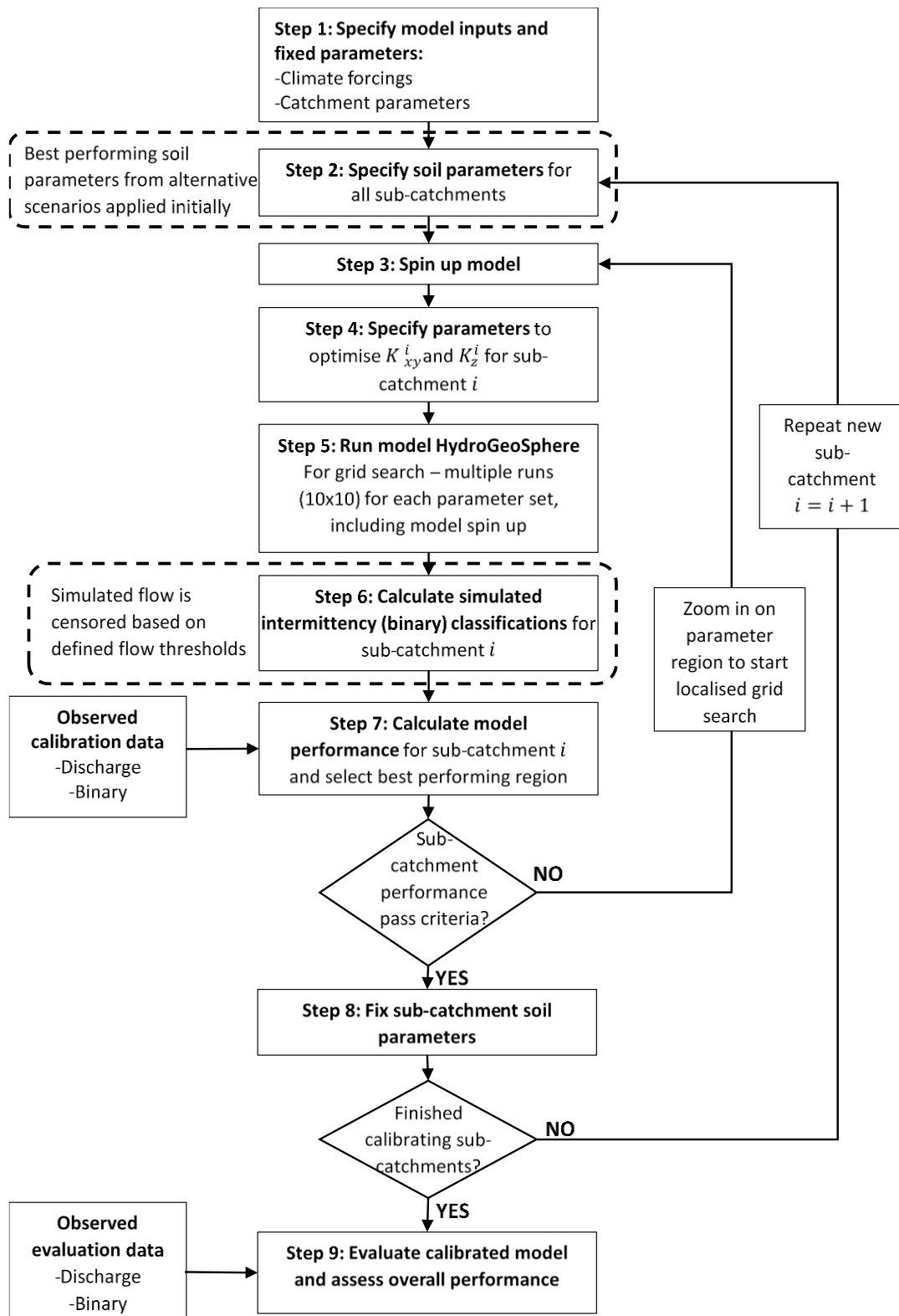


Figure 4-3: Flow chart illustrating method for the calibration and evaluation of the model to multiple sites

Table 4-2: Model parameters used to represent the catchment physical characteristics and the soil parameters applied to each sub-catchment.

Parameter	Value				
<b>Porous media (soil)</b>					
	<b>S001</b>	<b>S004</b>	<b>S005</b>	<b>S008</b>	<b>Outlet</b>
Lateral hydraulic conductivity (m/s)	$4.6 \times 10^{-6}$	$4.6 \times 10^{-6}$	$5.0 \times 10^{-7}$	$4.5 \times 10^{-5}$	$4.6 \times 10^{-5}$
Vertical hydraulic conductivity (m/s)	$8.4 \times 10^{-6}$	$8.4 \times 10^{-6}$	$1.3 \times 10^{-5}$	$4.7 \times 10^{-5}$	$8.4 \times 10^{-5}$
<b>Evapotranspiration</b>					
Vegetated area Evaporation depth (m)	3.5				
Cleared area Evaporation depth (m)	0.2				
Vegetated area Root depth (m)	3.5				
Cleared area Root depth (m)	0.2				
Vegetated area Canopy storage (mm)	0.5				
Cleared area Canopy storage (mm)	0.0				
Vegetated area Initial interception (mm)	0.5				
Cleared area Initial interception (mm)	0.0				
<b>Overland</b>					
Rill storage height (mm)	1.0				
Obstruction storage height (mm)	0.0				
x and y friction ( $s/m^{1/3}$ )	0.05				
<b>Channel</b>					
Rill storage height (mm)	1.0				
Obstruction storage height (mm)	0.0				
x and y friction	0.15				

#### 4.3.4 Quantification of model performance with binary observations, censored simulations and discharge

To evaluate the calibrated simulations, five metrics of model performance were used. The five performance metrics were applied depending on whether the comparison data corresponded to the binary streamflow state or to the combined streamflow state and magnitude (Table 4-3).

- The Nash-Sutcliffe efficiency (NSE) (*Nash and Sutcliffe, 1970*) provides an overall measure of the goodness of fit for discharge data. The minimum criteria of performance for the NSE was set to be 0.5, based on typical performance of several previous studies applying HGS (*Partington et al., 2013; Li et al., 2015; Glaser et al., 2016; Tang et al., 2017*).
- The correlation coefficient provides a measure of the strength of relationship between the observed and simulated discharge values. The minimum performance criteria applied for calculations of the correlation coefficient was defined as 0.65.
- The bias in flow volume was used to determine whether simulations over- or under- estimated total discharge. The maximum bias in flow volume was defined as 15%.
- Bias in number of flow days was defined as the percentage of days within a year where there is flow present in a streambed (e.g. number of flow days divided by 365), and was selected to determine the effectiveness of simulated total flow days. The maximum criteria for the calculated bias in flow days (using binary data) was 15% .

- The percentage matching observations was defined at the count of days that observed and simulated flow classifications (wet-dry) match. The minimum criteria for the percentage matching observations was defined as greater than or equal to 80%.

The bias in simulated flow permanence and percentage matching binary observations performance metrics all depend on the simulated streamflow dry/wet state, but this can be sensitive to the definition of zero flows. Therefore, to determine the simulated streamflow intermittency, the simulated discharge data was censored using flow thresholds. The thresholds were defined for each monitoring location using Manning’s equation with inputs of the sensor height above the streambed, the channel cross-section and the average slope of the channel (see Supplementary material). The flow threshold values are kept constant for all performance comparisons (Step 6, Figure 4-3). To benchmark the model calibrations across multiple locations within the catchment, the four alternative scenarios representing plausible yet differing runoff generation mechanisms were evaluated (see Section 4.3.2).

Table 4-3: Description of performance measures used to evaluate model simulations.

Performance measure	Data type	Definition	Accepted performance criteria
Daily Nash-Sutcliffe efficiency coefficient	Continuous	$NSE = 1 - \frac{\sum_{t=1}^T (Q_{m_t} - Q_{o_t})^2}{\sum_{t=1}^T (Q_{o_t} - \bar{Q}_o)^2}$ where $Q_{m_t}$ is the modelled flow at time, $t$ , $Q_{o_t}$ is the observed flow at time, $t$ , and $\bar{Q}_o$ is the mean of the observed data over the total time period, $T$ .	$NSE \geq 0.5$
Correlation coefficient	Continuous	$r = \frac{\sigma_{sim,obs}}{\sigma_{sim}\sigma_{obs}}$ where $\sigma_{sim,obs}$ is the covariance of the binary or flow data, $\sigma_{sim}$ is the standard deviation of the binary or flow simulations and $\sigma_{obs}$ is the standard deviation of the binary classifications or flow of the observations.	$r \geq 0.65$
Bias in flow volume	Continuous	$bias = \frac{\sum_{t=1}^T (Q_{m,t} - Q_{o,t})}{\sum_{t=1}^T Q_{o,t}} * 100$ where $Q_{m,t}$ is the modelled flow at time $t$ , $Q_{o,t}$ is the observed flow at time $t$ , and the total time period is $T$ .	$bias \leq 15\%$
Bias in number of flow days	Binary	$bias = \frac{\sum_{t=1}^T (S_{m,t}^{binary} - S_{o,t}^{binary})}{S_{o,t}^{binary}}$ where $S_{m,t}^{binary}$ is the simulated flow censored for the binary state, no-flow = 0 and flowing = 1, $S_{o,t}^{binary}$ is the observed time-series of intermittent (binary) flow.	$bias \leq 15\%$
Percentage matching binary observations	Binary	$\%match = \frac{\sum_{t=1}^T S_{m,t}^{binary} \text{ in } S_{o,t}^{binary}}{T} * 100$ where $S_{m,t}^{binary}$ is the simulated flow censored and classified in binary form, no-flow = 0 and flowing = 1, $S_{o,t}^{binary}$ is the observed time-series of intermittent (binary) flow and $T$ is the total length of the series.	$\%match \geq 80\%$



## 4.4 Results and discussion

### 4.4.1 Upstream site performance of outlet-only model calibrations

To determine effectiveness of the simulated upstream dynamics for the four scenarios calibrated at the catchment outlet, performance metrics were calculated both at the outlet and at the upstream sites (Table 4-4). The results show there was no single simulated scenario that performs well at all upstream sites. In particular, each scenario outperforms the other scenarios for at least a single site, highlighting that no single scenario can be removed based on it being inferior to the other scenarios at all the sites. For instance, Scenario 4 shows relatively high performance metrics at site S005 and low performance at site S004, whereas Scenario 2 shows relatively high performance at site S004 and low performance at site S005. The percentage matching binary observations for simulated flow at the outlet ranged between 14% and 89%. Of the 25 metrics calculated across the sites, the best performing was Scenario 2 with 52% of calculated metrics passing the specified criteria.

Table 4-4: Comparison of calculated performance metrics for four scenarios conditioned to the outlet. The yellow boxes indicate metrics that pass the criteria.

		Scenario 1	Scenario 2	Scenario 3	Scenario 4
<b>Outlet Performance</b>					
1	NSE	0.52,0.32	0.54,0.48	0.62,0.29	0.69,0.15
2	Correlation coefficient	0.74,0.70	0.71,0.77	0.74,0.70	0.80,0.68
3	Total flow volume bias (%)	45	-4	-33	-46
4	Percentage matching binary observation (%)	89	22	14	72
5	Number of flow days bias (%)	11	87	89	-55
<b>S001</b>					
6	NSE	0.28, -0.90	0.27,0.24	0.31, -0.83	0.34, -0.48
7	Correlation coefficient	0.71,0.64	0.67,0.70	0.71,0.64	0.77,0.45
8	Total flow volume bias (%)	-24	-57	-70	-75
9	Percentage matching binary observation (%)	88	91	89	45
10	Number of flow days bias (%)	28	11	28	46
<b>S003</b>					
11	NSE	0.29, -0.77	0.33, -0.06	0.27, -0.90	-0.23, -0.30
12	Correlation coefficient	0.79,0.74	0.68,0.76	0.79,0.74	0.79,0.66
13	Total flow volume bias (%)	60	-3	-32	-43
14	Percentage matching binary observation (%)	81	92	81	56
15	Number of flow days bias (%)	20	24	13	-56
<b>S004</b>					
16	Percentage matching binary observation (%)	80	82	80	76
17	Number of flow days bias (%)	50	-3	54	-21
<b>S005</b>					
18	Percentage matching binary observation (%)	66	75	66	90
19	Number of flow days bias (%)	228	139	178	11
<b>S008</b>					
20	Percentage matching binary observation (%)	76	77	76	66
21	Number of flow days bias (%)	24	51	24	-45
<b>S010</b>					
22	Percentage matching binary observation (%)	77	83	38	66
23	Number of flow days bias (%)	37	26	26	-54
<b>S011</b>					
24	Percentage matching binary observation (%)	71	80	79	40
25	Number of flow days bias (%)	-54	34	39	36
	<b>Percentage of metrics passing criteria</b>	<b>35%</b>	<b>52%</b>	<b>32%</b>	<b>24%</b>

#### 4.4.2 Upstream site performance of multi-site calibration

Whereas the previous section described the upstream performance of four alternative scenarios that were calibrated to the catchment outlet, we now explore the performance of a the multi-site calibration that included the additional streamflow intermittency information using the approach described in Figure 4-3. Figure 4-4 shows the simulated versus observed 'wet'-'dry' flow patterns across both the calibration and evaluation periods. Despite some discrepancy at sites S004 and S005, the multi-site calibration was able to achieve the objective of significantly improving the metric for the timing of flow onset and secession. At site S001 (reach 1) intermittent-flow was shown to begin in early March 2017 with flow becoming continuous in June 2017. Similarly, for site S004, located on reach 1 and upstream of site S001, intermittent-flow becomes continuous in June 2017. Sites S001 and S004 both recede in early to mid-September. Site S005 (upstream of reach 2) is a relatively steep intermittent channel with dense vegetation located upstream and at the site of sensor deployment. The stream becomes continuous in mid- July 2017 in response to large rainfall events and recedes in mid- September 2017 similar to sites S001 and S004.

There were many obvious differences between the sites, for example S004 and S005 have only 33% and 27% 'wet' days respectively, while sites S010 and S011 nearer the outlet have much more permanent flow of 60% 'wet' days. Figure 4-4 shows that, overall, the onset and period of permanent flow was reproduced well and the high degree of variation between sites was well represented in the model. This was not otherwise the case when the model was calibrated only at the outlet and therefore had the same parameters for all sub-catchments. The use of binary data provides significant improvement to the representation of the permanent flow period at the upstream sites. Some sites, notably S004 and S005, show differences in the scattered days of flow.

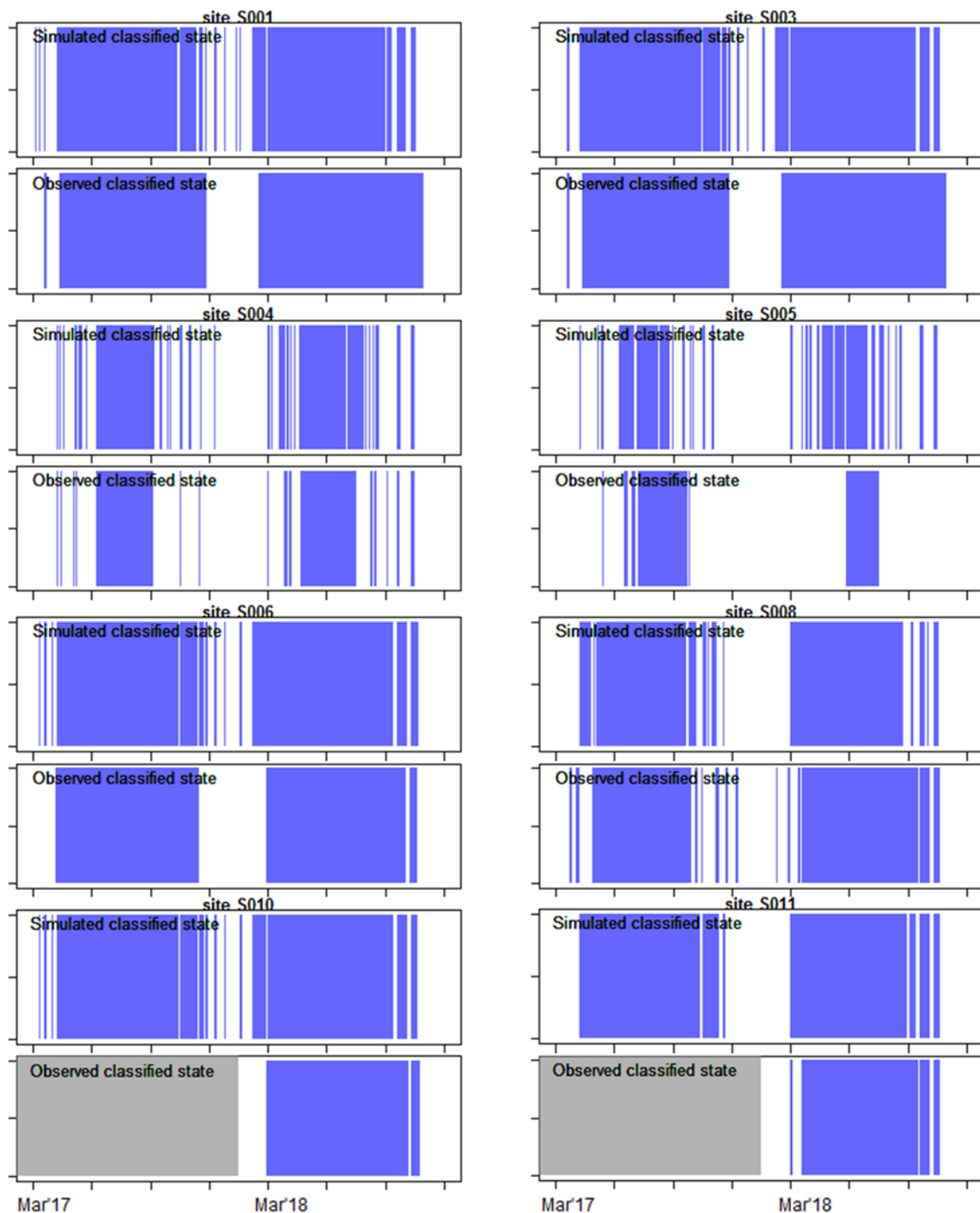


Figure 4-4 Comparison of binary simulations to observed intermittent flow, showing improved representation of 'wet'- 'dry' patterns across all site, where the blue shading represents periods on flow. Site S010 and S011 contain only one year of observations, the first year which does not contain data is shown as grey.

Figure 4-5 shows the percentage of days that the simulated binary data matches observed binary classification for the calibration and evaluation period, and demonstrates significant improvements in the capacity of the multi-site calibration to represent upstream processes. In particular, Figure 4-5 shows that the percentage of matching states exceeds 80% during the calibration and evaluation periods for all sites. In contrast, the outlet-only calibrations showed results ranging between 14% and 92%, with no single scenario producing percentage of simulated binary data matches observations results that exceed 80% across all sites (Table 4-4). The best performing site was S003 (the outlet for tributary 1 and 2), where the percentage matching was

88% with only a small decrease in performance during the evaluation period (87%). Sites S010 and S011, which were reserved exclusively for spatial evaluation, also performed well with percentage matching binary observations of 85% and 83%.

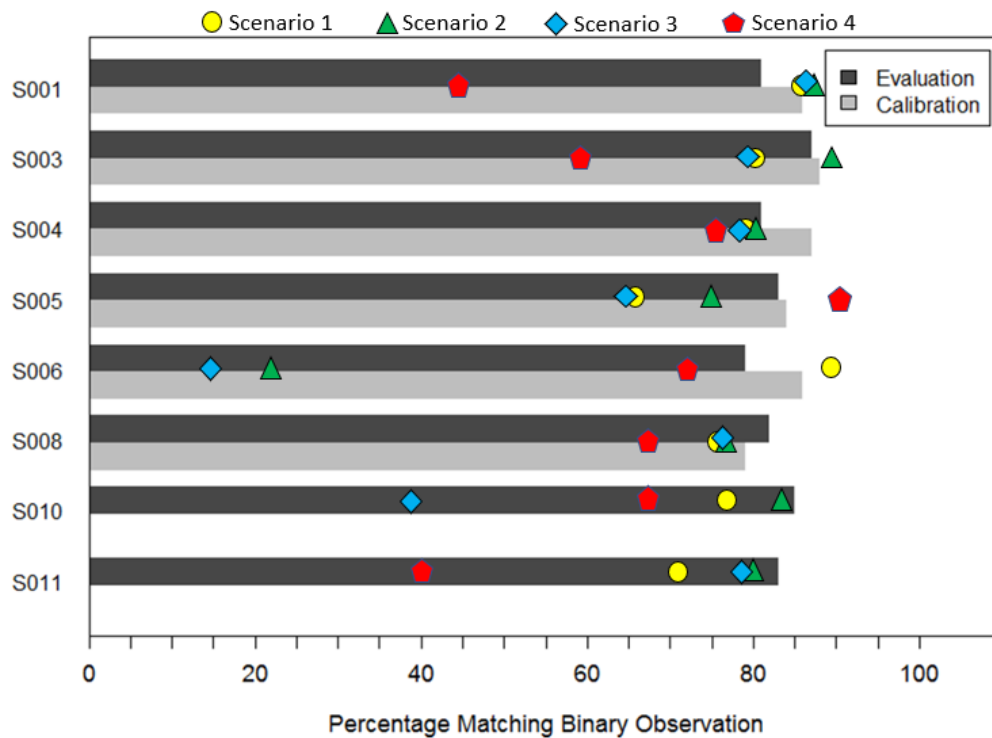


Figure 4-5: Percentage matching binary observations for the calibration and evaluation period for the multi-site conditioned simulation. Sites S010 and S011 were reserved exclusively for evaluations of performance across the catchment. The coloured markers compare the percentage matching binary observations for the four scenarios over a two year period.

The remaining performance metrics were calculated and presented for each location in Table 4-5. The table shows that performance had significantly improved at upstream sites when compared to the outlet-only calibration. For example, the outlet-only calibration had a correlation coefficient ranging from 0.36 to 0.76 across all sites (Table 4-4), whereas the multi-site calibration had 0.61 to 0.79 across all sites (Table 4-5). A comparison of binary and non-binary performance metrics also shows there's a trade-off between calibrating to intermittency and getting volumes right, in that, while stream flow intermittency maybe be well represented, discharge volumes may be less accurately represented. For example, site S001 flow data (which was reserved for evaluation) shows that there was a negative bias of 50% over the two year period, underestimating the total flow.

Table 4-5: Comparison of multi-site simulated results to the observed data for each site are illustrated

Site	Outlet	S001	S003	S004	S005	S008	S010 *	S011 *
<b>Performance measures</b>								
r	0.75,0.77	0.74,0.70	-0.59, -0.52	-	-	-	-	-
NSE	0.50,0.49	0.54,0.32	-0.08, -0.13	-	-	-	-	-
Observe zero flow days	130	105	86	248	290	160	119	143
Simulated zero flow days	104	141	79	244	262	178	145	149
Error in zero flow days (%)	-20	35.5	-11.6	-1.6	-9.7	11.2	21.8	4.0
Number of flow days bias (%)	11.6	-14.2	2.4	2.9	34.9	-9.0	-10.9	-2.6
Flow volume bias (%)	1.5	-50	-12	-	-	-	-	-
<b>Simulated flow characteristics</b>								
Simulated flow volume	614	132	300	203	103	203	322	285
Q <sub>20</sub> (high flow threshold) [Simulated] (ML/d)	2.15	0.47	1.07	0.24	0.33	0.73	1.14	1.01
Q <sub>20</sub> (high flow threshold) [Observed] (ML/d)	3.5	1.24	1.34	-	-	-	-	-
Q <sub>70</sub> (low flow threshold) [Simulated] (ML/d)	0.22	0.06	0.14	0.03	0.04	0.02	0.15	0.07
Q <sub>70</sub> (low flow threshold) [Observed] (ML/d)	0.0	0.43	0.45	-	-	-	-	-
* indicates site data is used for evaluation only								

#### 4.4.3 Outlet performance of multi-site calibration

The multi-site calibration approach utilizing the additional ‘wet’-‘dry’ information was able to capture the main characteristics of the outlet hydrograph over the two year simulation period (Figure 4-6). Over the calibration period, the multi-site calibrated model had slightly reduced performance at the outlet (NSE=0.5) when compared to the outlet-only calibrations (NSE=0.52 to 0.69). In contrast, for the evaluation period the multi-site calibration retains or improves its performance (NSE=0.49) relative to the outlet (NSE=0.15 to 0.48). Importantly, the multi-site calibration had improved upstream performance at all sub-catchment sites.

Figure 4-6 illustrates that the simulation calibrated to multiple sites reasonably simulated peak flows compared to the observed data. The baseflow was underestimated in the first year of simulations, resulting in an under-estimation of flows over the two years (24 mm or negative bias of 15%). Figure 4-6c illustrates observed streamflow state and simulated flow, censored to represent flow intermittency at the catchment outlet. The figure shows that while the timing of the onset of flow was accurately simulated in the first year, flow onset begins approximately a month sooner than was observed in the classified data for the second year.

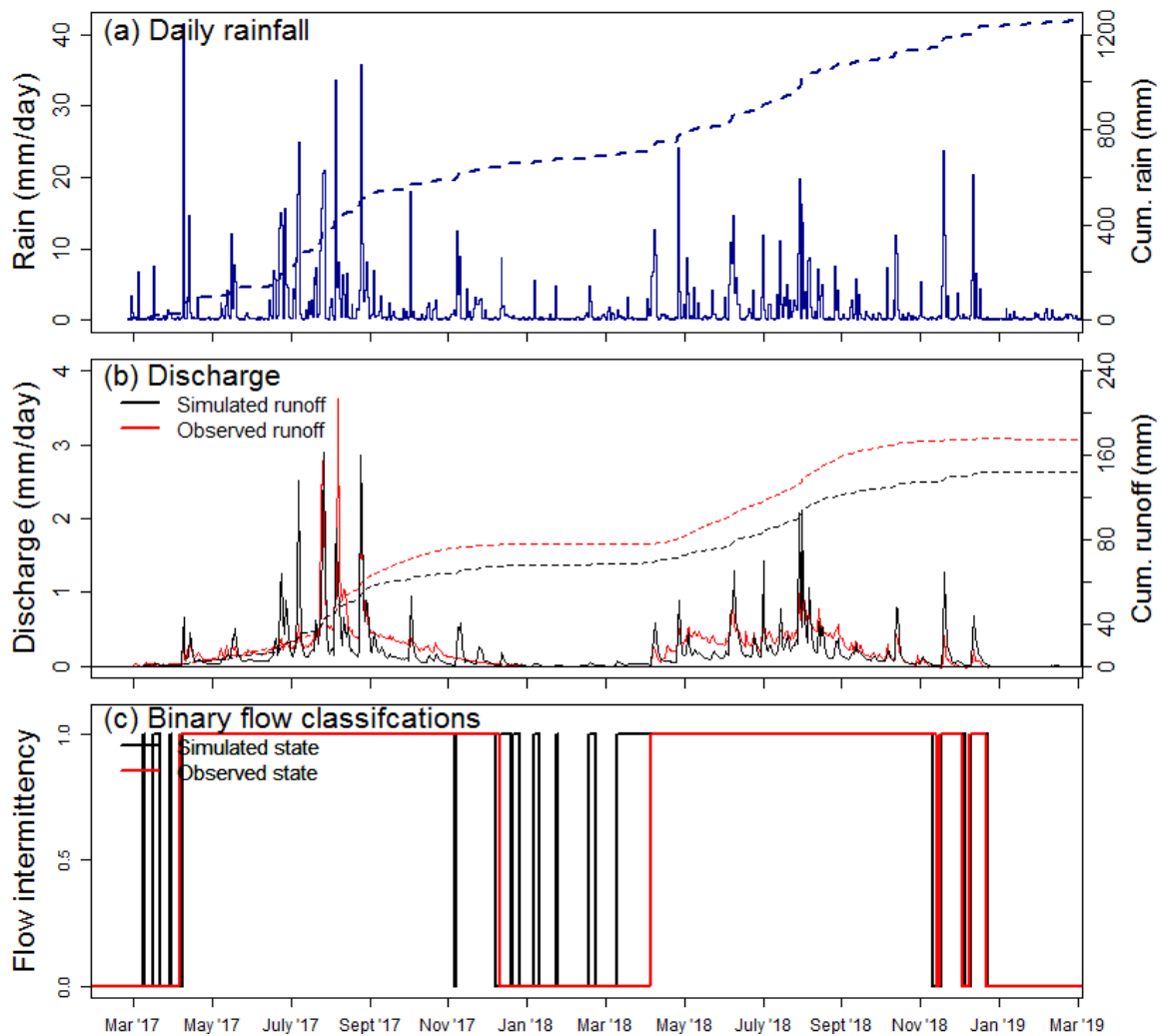


Figure 4-6: Outlet hydrograph showing the calibration (Mar 2017-Mar 2018) and evaluation (Mar 2018-Mar 2019) periods with the (a) rainfall forcings shown and the (b) simulated discharge closely flowing the observed data and (c) comparison of simulated and observed intermittent flow where 0='no-flow' and 1='flowing'.

#### 4.4.4 Summary of single site versus multi-site calibration

The multi-site calibration showed more than a 60% (i.e. [84%-52%]/52%) improvement in upstream catchment performance compared to the best performing single-site calibration. That is, the multi-site calibration had 84% of metrics passing criteria compared to the single-site calibrations with 24% to 52% of metrics passing the criteria (see Supplementary material, Table 4-8). Figure 4-7 visually illustrates a summary of performance across the sites within the catchments and refers to metrics at the outlet as well as metrics from sites S001, S004, S005 and S008. The metrics calculated were determined and illustrated in a panel where yellow refers to the performance criteria passing and orange refers to metrics that fail the criteria. Figure 4-7 shows that regardless of metric and site, the multi-site model performs better than the outlet-only calibrations for upstream sites, while having comparable performance at the outlet. The exception to this is site S005 where Scenario 4 outperforms all other calibrations for this site.

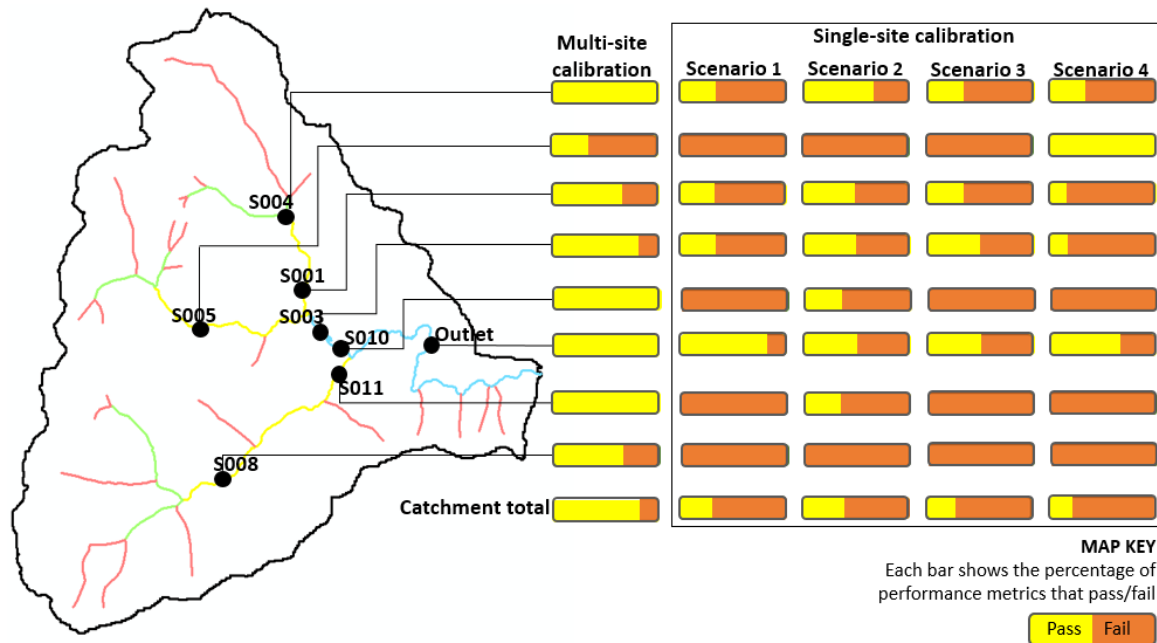


Figure 4-7: Comparison of multi-site conditioned simulation and simulations calibrated to the outlet only i.e. Scenarios 1, 2, 3 and 4.

#### 4.5 Conclusions

The goal of this study was to illustrate the benefit of low-cost data collected in a headwater catchment in the context of physically based hydrological model calibration. The upstream performance of four conceptual models of runoff generation mechanisms simulated using a surface-subsurface flow model, HydroGeoSphere, and producing similar streamflow at the outlet were assessed and compared using the additional intermittency data. The data was applied to illustrate a multi-site calibration method where the overall performance of all candidate simulations compared.

The challenge of quantifying high resolution intermittent streamflow across scales can be addressed with insights from improved modelling approaches which incorporate additional low-cost data. There is significant benefit in additional data to test hypotheses around alternative model configurations and structures to assess the most plausible runoff mechanisms behaviour present in the studied catchment. Table 4-4 shows that of the four calibrated outlet-only simulates there was no single scenario that was applicable over the whole catchment. This indicates that model methodologies that perform single-site calibrations are likely to have significant deterioration in flow performance upstream of the outlet. An outcome of placing a lumped parameter set across the catchment.

With the multi-site calibration approach, the study was able to improve the representation of performance of all sites across the catchment. In addition, the evaluation sites (S003, S010 and S011) show significant improvements compared to the outlet-only simulations (Figure 4-7) showing that additional data, even if it is of binary streamflow state, provides significant additional information for constraining a models solution space. An outcome of applying five lumped parameters sites, each representing an individual section of reach. By increasing the resolution of the model, there is still the question of whether the five sub-catchments are the adequate spatial resolution or whether the data points are in the correct locations to capture the heterogeneous behaviour of runoff, given that upstream on each data point location at catchment is

homogeneous. These questions in part can be answered by considering the scale of the water management question and the influence of decisions being made. If the water management question is at the outlet, does it matter that the model used it is lumped? Alternatively, if the water management question is at the reach scale, is the scale which the temperature sensors deployed the correct scale to lump parameter sets?

As is the case with all modelling studies, the simulations were not a true representation of reality. However, the aim of this study was to illustrate the usefulness of additional low-cost data to improve spatial representation within a catchment. The lack of groundwater metering in the catchment area is also an acknowledged limitation of this work. While the data may have served in assisting with hypothesis testing the scenarios, again the aim of the study was to illustrate the benefits that a short-term data campaign may have a model representation, which may not require the explicit representation of groundwater in the model being applied.

The censoring of simulated discharge data was also dependent on the threshold flow rating being applied to the data. The calculation results of simulation performance were sensitive to these thresholds. However, to overcome this, the flow threshold was applied consistently across all sites. Long simulation times and the vast number of parameters in fully-coupled processes models also restricted the calibration method applied, with a semi-manual approach adopted. This approach means that the model is not truly optimized and was conditioned to the best possible representation of the data. Additional calibration of the model would see improvement of simulated outputs given that performance criteria for calibration parameters were pre-defined in this study. There is significant opportunity to investigating alternative model structures as well as compared process, semi-distributed and conceptual models which have been calibrated with the additional data.

The study serves as a caution that calibrating a hydrological model against a single point can produce results that may not reproduce catchment functioning which may be required for water management questions that are nested across scales. This is particularly the case when there is a large amount of heterogeneity throughout the catchment (e.g. vegetated vs cleared areas and sources vs sinks) as showed with the small catchment used in this study. In contrast, the consideration of additional data to supplement high-quality streamflow records can improve the ability to constrain the model's solution space in simulating in-stream fluxes of water. The study has highlighted the benefit of combining alternative data sources, collected through a short-term data campaign that encompasses both hillslope-scale hydrological features as well as flow further downstream.



## Acknowledgements

The authors acknowledge the support of the South Australian government Department for Environment and Water for identifying the case study and providing catchment data. We would also like to thank the individual land owners in the study area who provided us the site access.

## References

- Bergstrom, A., Jencso K., McGlynn B., 2016. Spatiotemporal processes that contribute to hydrologic exchange between hillslopes, valley bottoms, and streams, *Water Resour. Res.*, 52, 4628–4645, doi:10.1002/2015WR017972
- Beven, K.J., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling, *Adv Water Resour.*, 16 (1993), pp. 41-51
- Birkel, C., Soulsby C., and Tetzlaff, D., 2014. Developing a consistent process-based conceptualization of catchment functioning using measurements of internal state variables, *Water Resour. Res.*, 50, 3481–3501, doi:10.1002/2013WR014925.
- Blumstock, M., Tetzlaff, D. I. A., Malcolm Nuetzmann G., Soulsby., C., 2015. Baseflow dynamics: Multi - tracer surveys to assess groundwater contributions to montane streams under low flows, *J. Hydrol.*, 527, 1021-1033, doi:10.1016/j.jhydrol.2015.05.019.
- Brubaker, K. M, Myers, W.L., Drohan, P. J., Miller, D. A., Boyer, E. W. 2013. The Use of LiDAR Terrain Data in Characterizing Surface Roughness and Microtopography. *App. & Env. Soil Sci.* doi.org/10.1155/2013/891534
- Chance, C. M., Coops, N. C., Plowright, A. A., Tooke, T. R., Christen, Aven, N. 2016. Invasive Shrub Mapping in an Urban Environment from Hyperspectral and LiDAR-Derived Attributes. *Front. Plant Sci.* doi.org/10.3389/fpls.2016.01528
- David, J. C., Rutten, M. M., Pandey, A., Devkota, N., van Oyen, W. D., Prajapati, R., van de Giesen, N., 2018. Citizen science flow – an assessment of citizen science streamflow measurement methods. *Hydrol. Earth Syst. Sci.* doi.org/10.5194/hess-2018-425.
- Gaukroger, A. M., Werner, A. D., 2011. On the Panday and Huyakorn surface–subsurface hydrology test case: analysis of internal flow dynamics. *Hydrol Processes.* 25: 2085-2093. doi.org/10.1002/hyp.7959
- Glaser, B., J. Klaus, S. Frei, J. Frentress, L. Pfister, and L. Hopp., 2016. On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum, *Water Resour. Res.*, 52, 8317–8342, doi:10.1002/2015WR018414.
- Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377(1-2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.
- Kirchner, J. W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42, W03S04, doi:10.1029/2005WR004362.
- Kuppel, S., Tetzlaff, D., Maneta, M. P., Soulsby, C., 2018. What can we learn from multi-data calibrate of a process-based ecohydrological model? *Env. Modelling & Soft* 101, 301-316. doi.org/10.1016/j.envsoft.2018.01.001
- Le Coz, J., Patalano, A., Collins, D., Guillen, N.F., Garcia, C.M, Smart, G.M., Bind, J., Chiaverini, A., Le Boursicaud, R., Dramais, G., Braud, I., 2016. Crowdsourced data for flood hydrology: feedback from recent citizen science projects in Argentina, France and New Zealand. *J. Hydrol.* 541, 766-777. dx.doi.org/10.1016/j.jhydrol.2016.07.036

Lovett, G. M., Burns, D., Driscoll, C. T., Jenkins, J. C., Mitchell, M. J., Rustad, L., Shanley, J. B., Likens, G. E., Haeuber, R., 2007. Who needs environmental monitoring? *The ecological society of America*. 5: 253-260.

Panday S, Huyakorn PS. 2004. A fully - coupled spatially distributed model for evaluating surface/subsurface flow. *Advances in Water Resources* 27: 361-382.

Ruiz-Garcia, L., Lunadei, L., Barreiro, P., Robla, J. I., 2009. A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. *Sensors*. 9(6): 4728-4750.

Turner, D. S., Richter, H. E., 2011. Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions. *Environmental Management* 47: 497-505

Li, L., Lambert, M. F., Maier, H. R., Partington, D., Simmons, T., 2015. Assessment of the internal dynamics of the Australian Water Balance Model under different calibration regimes. *Env. Modelling & Soft.* 66: 57-68.

Maxwell, R. M., Condon, L. E., 2016. Connections between groundwater flow and transpiration partitioning. *Science* 353, 377-380

Soulsby, C., C. Birkel, J. Geris, J. Dick, C. Tunaley, and D. Tetzlaff. 2015. Stream water age distributions controlled by storage dynamics and nonlinear hydrologic connectivity: Modeling with high-resolution isotope data, *Water Resour. Res.*, 51, 7759-7776, doi:10.1002/2015WR017888.

Tang, Q., Kurtz, W., Schilling, O.S., Brunner, P., Vereecken, H., Hendricks Franssen, H.-J., 2017. The influence of riverbed heterogeneity patterns on river-aquifer exchange fluxes under different connection regimes. *Journal Hydro*, 554: 383-396. doi.org/10.1016/j.jhydrol.2017.09.031

Thomas, I.A., Jordan, P., Shine, O., Fenton, O. Mellander, P.-E., Dunlop, P., Murphy, P.N.C. 2017. Defining optimal DEM resolutions and point densities for modelling hydrologically sensitive areas in agricultural catchments dominated by microtopography. *Int. Journal of App. Earth Obs. & Geoinformation*. 54: 38-52.

Wellen, C., A.R. Kamran-Disfani, G.B. Arhonditsis., 2015. Evaluation of the current state of distributed watershed nutrient water quality modeling. *Environ. Sci. Technol.*, 49: 3278-3290.

Wickert, A. D., Sandell, C. T., Schulz, B., Ng., G. H. C., 2018. Open-source Arduino-derived data loggers designed for field research. *Hydrol. Earth Syst. Sci.* doi.org/10.5194/hess-2018-591

## Appendix 4-A

The results of four simulated scenarios are plotted against the observed discharge data. The simulated scenarios were calibrated to a single point at the outlet (Figure 4-8). The breakdown of different runoff proportions is detailed in Table 4-6. Additional details on simulations and results are given in *Chapter 2*.

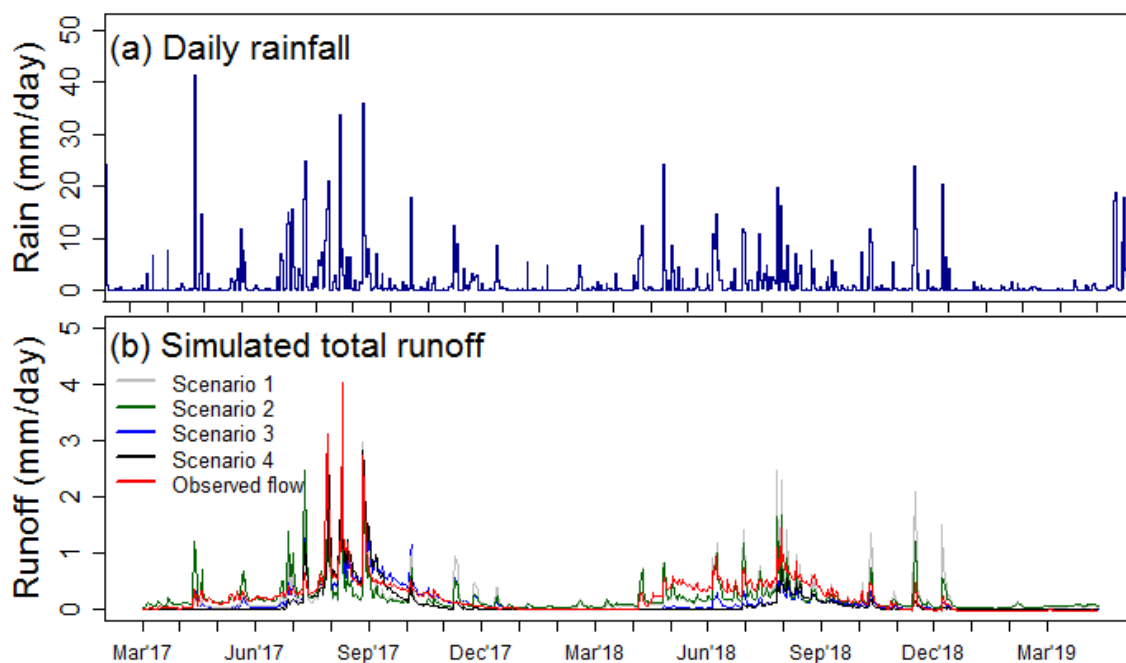


Figure 4-8: Showing (a) applied rainfall over two years and (b) the hydrographs of the four single-site calibrations compared to the observed discharge.

Table 4-6: Details of the single-site calibrations for four Scenarios representing alternative conceptualizations of runoff generation.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
	SE with GW	SE & GW	GW with SE	GW with IE&SE
<b>Lateral hydraulic conductivity</b>	$2.0 \times 10^{-6}$	$1.3 \times 10^{-6}$	$4.8 \times 10^{-5}$	$9.8 \times 10^{-5}$
<b>Vertical hydraulic conductivity</b>	$1.3 \times 10^{-4}$	$1.5 \times 10^{-6}$	$7.0 \times 10^{-7}$	$2.8 \times 10^{-7}$
<b>Percentage saturation excess runoff</b>	61%	51%	31%	69%
<b>Percentage infiltration excess runoff</b>	0%	0%	0%	7%
<b>Percentage groundwater runoff</b>	35%	46%	57%	19%
<b>Percentage error</b>	4%	3%	2%	5%

## Supplementary material

### Threshold for censoring simulated discharge

The threshold flow response for each site used to censor the simulated flow data was determined using Manning's equation. Individual site details are outlined in Table 4-7.

*Table 4-7: Threshold values applied to censor simulated flows for binary 'wet'-'dry' classification to determine stream flow intermittency*

<b>Site</b>	<b>Contributing area (km<sup>2</sup>)</b>	<b>Height above streambed (m)</b>	<b>Average channel slope</b>	<b>Manning's n applied</b>	<b>Threshold flow (m<sup>3</sup>/s)</b>
S001	1.8	0.02	0.04	0.4	0.001
S003	4.4	0.03	0.036	0.3	0.002
S004	0.9	0.05	0.05	0.4	0.001
S005	1.5	0.10	0.04	0.4	0.002
S006	8.6	0.02	0.036	0.4	0.002
S008	2.6	0.03	0.02	0.2	0.002
S010	4.1	0.03	0.036	0.4	0.002
S011	4.5	0.04	0.02	0.2	0.002

## Detailed results comparing all parameterised simulations

The detailed results for each site for the calculated performance metrics were compared and presented in Table 4-8. The yellow shading indicated when a metric had passed the defined criteria (outlined in Section 4.4.4, Figure 4-7). The results show that 82% of the metrics pass the performance criteria for the multi-site calibration. In contrast less than 50% of all the outlet-only scenarios pass the performance criteria.

Table 4-8: Comparison of multi-site conditioned simulation to simulations calibrated to the outlet only. The yellow, colours indicate where the individually calculated metric have passed the criteria.

		Multi-site calibration	Scenario 1	Scenario 2	Scenario 3	Scenario 4	
CALIBRATION SITES	<b>Outlet Performance</b>						
	1	NSE (continuous)	0.50	0.52	0.54	0.62	0.69
	2	Correlation coefficient (continuous)	0.76	0.72	0.74	0.72	0.74
	3	Total flow volume bias (%)	1	45	-4	-33	-46
	4	Percentage matching binary observation (%)	83	89	22	14	72
	5	Flow permanence bias (%)	3	11	87	89	-55
	<b>S004</b>						
	6	Percentage matching binary observation (%)	84	80	82	80	76
	7	Flow permanence bias (%)	4	50	-3	54	-21
	<b>S005</b>						
	8	Percentage matching binary observation (%)	83	66	75	66	90
9	Flow permanence bias (%)	38	228	139	178	11	
<b>S008</b>							
10	Percentage matching binary observation (%)	81%	76	77	76	66	
11	Flow permanence bias (%)	-8%	24	51	24	-45	
EVALUATION SITES	<b>S001</b>						
	12	NSE (continuous)	0.23	-0.31	0.26	-0.26	-0.07
	13	Correlation coefficient (continuous)	0.72	0.65	0.69	0.68	0.65
	14	Total flow volume bias (%)	-50	-24	-57	-70	-75
	15	Percentage matching binary observation (%)	84	88	91	89	45
	16	Flow permanence bias (%)	-14	28	11	28	46
	<b>S003</b>						
	17	NSE (non-binary)	-0.08	-0.24	0.27	-0.31	-0.26
	18	Correlation coefficient (continuous)	0.75	0.77	0.72	0.77	0.73
	19	Total flow volume bias (%)	-12	60	-3	-32	-43
	20	Percentage matching binary observation (%)	92	81	92	81	56
	21	Flow permanence bias (%)	0	20	24	13	-56
	<b>S010</b>						
	22	Percentage matching binary observation (%)	91	77	83	38	66
	23	Flow permanence bias (%)	12	37	26	26	-54
<b>S011</b>							
24	Percentage matching binary observation (%)	94	71	80	79	40	
25	Flow permanence bias (%)	2	-54	34	39	36	
<b>Percentage of metrics passing criteria</b>		<b>84%</b>	<b>36%</b>	<b>52%</b>	<b>32%</b>	<b>24%</b>	

## Chapter 5.

### 5.1 Summary of research objectives

Knowledge of intermittent flows and the variation of flow throughout a catchment is critical for managing water resources in semi-arid and Mediterranean environments. There are multiple examples where understanding local scale processes can significantly improve efficiencies in water allocation, for example the interaction of farm dams and environmental flows or the impact of land-use changes throughout a catchment. Current modelling methodologies typically aggregate catchment processes to a single point-estimate of discharge at the outlet. As a result, approaches that aggregate catchment function are less likely to represent upstream catchment dynamics adequately.

This thesis was focused on investigating intermittent flow processes and understanding the implications of commonly applied modelling assumptions on water management within the catchment, for example the limitation of aggregate representation of catchment flow at the outlet. The research combined relatively underutilised methods of data collection such as low-cost environmental sensors with a physically-based model, HydroGeoSphere.

The overall objective of this research was to demonstrate the value of additional data for characterising intermittent flow processes within a Mediterranean catchment. The specific objectives were:

**Objective 1: To quantify streamflow intermittency and signatures with low-cost sensing technology.** Low-cost temperature sensors were used to collect in-stream data on individual reaches. Methods of analysing data were compared and developed to classify the stream state as either 'wet' (flowing) or 'dry' (not flowing). Timeseries and statistics of streamflow intermittency were developed to address the needs of subsequent model calibration to account for upstream catchment dynamics.

**Objective 2: To assess implications of modelling assumptions on the representation of local-scale intermittent streamflow.** Four alternative simulations were calibrated exclusively to discharge at the outlet and were parameterised to represent different conceptualisations of runoff generation. The four simulations were calibrated to represent a minimum Nash-Sutcliffe efficiency co-efficient of 0.5 while simulating different proportions of runoff mechanisms on the hillslope (i.e. infiltration excess, saturation excess and groundwater). This objective demonstrated the significance of internal model misspecification.

**Objective 3: To represent intermittent streamflow in headwaters using additional data in a multi-site calibration.** To demonstrate the value of additional low-cost data for the representation of catchment processes in hydrological model, the binary 'wet'-'dry' classifications of streamflow intermittency were applied to investigate the effectiveness of the four aggregated conceptualisations of catchment runoff (simulations from Objective 2). The data was used for multi-site calibration of the catchment and illustrated the improved representation of intermittent flow processes at the local scale.

### 5.2 Key research findings and contributions

**Objective 1 – Quantifying streamflow intermittency and signatures with low-cost sensing technology**

Headwaters, here defined as first to third order streams, can often be difficult to access, especially during raining periods, and unlike lower reaches, often lack continuous long-term data collection of streamflow onset and secession. Intermittent streamflow signatures, based on flow surrogates at multiple points in a catchment, were used to provide additional understanding of streamflow regimes within individual reaches. Paired low-cost temperature sensors at nine headwater sites were implemented within a 10 km<sup>2</sup> catchment.

Chapter 2 outlined an approach to collect continuous time-series of in-stream and on-bank temperature measurements at 15-minute intervals across nine sites within a small 10 km<sup>2</sup> catchment. Thermal properties of the data were assessed and compared to investigate their effectiveness of streamflow delineation as either a 'wet' or 'dry' state. Two methods of flow detection were compared and developed to determine their effectiveness. The two methods of classification were: (1) the standard deviation method, where a threshold deviation was applied (*Constantz, et al. 2001; Blasch, et al. 2004*) and; (2) the two-state hidden Markov model (HMM), used as an unsupervised signal detection method (*Arismendi, et al. 2017*). Both methods were further developed to maximise classification accuracy.

A two-state hidden Markov model (HMM) improving the overall accuracy of 'wet' – 'dry' stream state classification compared to a naïve streamflow classification, applying a threshold variance for delineation. Alternative configurations of the HMM were investigated to maximise overall accuracy. The HMM method was able to achieve an average accuracy of 92% across the sites and also has the advantage of being an unsupervised method of flow detection. The study enabled the unsupervised HMM method of flow detection with three required inputs including:

- (1) **In-stream temperature variance** was selected based on thermal characteristics of the data showing temperature variance as the strongest signal of flow detection and of previous studies (*Constantz et al. 2001; Blasch et al., 2004; Arismendi, et al. 2017*).
- (2) **30 day antecedent rainfall** was used to provide the algorithm with additional seasonal statistical characteristics of streamflow persistence.
- (3) **The ratio of in-stream to on-bank variance (defined as the  $F(t)$ ratio)** was developed as an additional signal of flow detection.

This approach allowed the inference of binary ('wet'–'dry') classifications which were used to interpret the properties of the small headwater catchment across multiple locations. Differences between the tributaries were demonstrated in terms of onset and secession of flow, number of zero flow days, number of zero flow periods, and percentage of flow permanence annually and by season. The study quantified the high degree of variability in intermittent flow throughout the small catchment area, demonstrating the benefits of alternative data collection methods that are not otherwise accessible with the existing network of streamflow gauges. Such data has the potential to be applied to model calibration and evaluation approaches for improved local scale flow representation.

## **Objective 2 – Implications of modelling assumptions on the representation of local-scale intermittent streamflow**

Many water management questions rely on a single point of gauged discharge data for policy, planning and design decisions. However, the hydrology of hillslopes is very different from the hydrology of catchments, whereby runoff at the hillslope is characterized by greater intermittency and variability than when flow is aggregated across the catchment. The key research motivation here was to understand the implications of commonly made modelling assumptions (such as



calibration of hydrological models to a single discharge point at the outlet) on the representation of processes at the hillslope and how this may influence water management decisions.

Chapter 3 outlined the approach of calibrating four candidate simulations to a single point estimate of streamflow using the process-based model HydroGeoSphere. Four alternative conceptualisations of runoff-generation were calibrated to represent different proportions of runoff mechanisms on the hillslope and near equivalent Nash-Sutcliffe coefficients for streamflow at the outlet. The four competing conceptualisations were: (1) saturation excess dominated, (2) saturation excess and groundwater dominated, (3) groundwater dominated and (4) groundwater dominated but containing 17% infiltration excess. Despite having similar performance at the outlet, differences were shown in numerous aspects of upstream representation, including the spatial representation of catchment saturation. The results demonstrated that subsurface pathways influence the behaviour of flow in and around a small in-stream dam with groundwater dominant scenarios showing that low flows were greater than 20% more likely to be simulated directly downstream compared to upstream.

The results suggested that model parameterisation and structures with a priori process assumptions influence the estimation of flows throughout a catchment. The results highlight the need for greater attention to process modelling and better representation of flow dynamics than aggregated models calibrated at an outlet. This distinction is important for policy planning, as many engineering, ecological and water management problems exist at the local scale but can have catchment wide implications. With improved representation of hydrological complexity in space and time, planning and implementation of environmental water policies can be based on more accurate representation of flow stores and fluxes across all relevant scales.

### **Objective 3 – Representing intermittent streamflow in headwaters using additional data in multi-site calibration**

Calibrating physically-based models is difficult given the complexity of catchments, having non-linear dynamics across nested spatial and temporal scales. A significant advantage of fully integrated hydrological models, is that no process hypotheses have to be chosen. The data 'selects' the process and without sufficient it can be difficult to distinguish between multiple competing hypotheses of catchment processes. Inexpensive data collection on the hillslope can provide supplementary information of intermittent streamflow that can be used to constrain a model parameter space. While there are opportunities for the application of additional data, there is a strong parallel with the need to improve calibration methods to manage and exploit all available data.

Chapter 4 demonstrated the extent to which data collected in a headwater catchment can be applied to improve model calibration. The analysis showed that for the four candidate simulations, calibrated exclusively to discharge at the outlet, there was no single simulation that was able to represent upstream intermittency across all sites. Binary 'wet'-'dry' classifications of streamflow intermittency at multiple sites were applied to calibrate the model. The multi-site calibration was evaluated showing that the simulation was able to significantly improve upstream representation of the catchment.

The multi-site calibrated model improved performance at eight out of nine sites, with only a small deterioration to the calibration results at the outlet. The approach increased the overall performance of the model with 84% of performance metrics across the catchment passing a pre-defined criterion compared to the best performing single-site calibration with 52% of performance

metrics passing. The study demonstrated that applying constraints to the model configurations allowed for greater confidence in outputs. The results suggest that as sensing and transmission technologies continue to improve, there will be increasing opportunities to use information sources such as local-scale intermittency to supplement reliable streamflow records in order to faithfully represent hydrological processes across scales. With improved representation of runoff across scales there are many benefits of getting the internal processes right. For example, simulating conditions that are outside the range of calibration or even reanalysis datasets that utilise 3D simulations.

### 5.3 Research limitations and challenges

This thesis presents an investigation of intermittent flow processes within a Mediterranean environment. This research had a number of limitations relating to data collection, data analysis methods, and data availability, modelling and simulation challenges. These limitations are further discussed with respect to each research objective.

#### **Objective 1 – Quantifying streamflow intermittency and signatures with low-cost sensing technology**

The applications of low-cost sensors introduce the potential for poor quality and accuracy of measurements. There were also limitations in the methods used to deploy sensors. These were:

- Sensors touching the streambed when the soil moisture may persist and cause inference, leading to false 'wet' day classifications.
- Sensors placed marginally above the streambed; a trickle of very low flows can pass under the sensor. When the sensor is not touching the base, sedimentation may cause rivulets at very low flows to bypass flow around the sensor.
- Build-up of sediment may cause the sensor to become buried and result in spurious readings.
- Sensors do not detect truly 'zero' flow conditions, but represent the intermittency of the stream above some very minor base threshold.
- Differences in site conditions could result in unique diurnal patterns, the use of paired sensors over a single in-stream sensor has partly addressed this limitation.
- Cloudy days and cold days can result in a reduced diurnal temperature range and can mimic the in-stream signal when flow is present. This can result in a false 'wet' classification.
- The time and man power required to collect data were the largest resource required for the low-cost method applied in this study. That is, monthly visits were conducted to multiple sites with multiple people. This is achievable for a research study, but for wider adoption, low cost telemetry would be required to support a similar field campaign.
- Site access was a significant limitation. Permission from property owners was required to access reaches located in the study area. Additionally, some sites could only be accessed on foot. Preferred sites within the catchment (e.g. the confluence of two reaches) were inaccessible and not able to be used for data collection.
- The battery life of low-cost sensors can limit the time the sensors are deployed. The battery life also influences the resolution of measurements taken. That is, increased resolution decreased battery life.
- Telemetry is a significant barrier when applying such sensors to headwater areas. Typically, telephone coverage in rural areas is limited, meaning the manual nature of data collection may not be completely eliminated. Alternatives could be investigated such as Bluetooth to limit the disturbance of deployed sensors.

- Temperature may not be the best measurement for data collection. There are other potential methods for inferring streamflow intermittency, for example, EC probes, video or time-delayed images.
- The lack of other data is a significant limitation of this study. Data such as groundwater, soil moisture measurements, tracers, and pluviometer gauges within the study location are just some example of additional data that could further inform the study.

### **Objective 2 – Implications of modelling assumptions on the representation of local-scale intermittent streamflow**

There were a number of modelling limitation as a result of the selection of the complex fully coupled surface-subsurface groundwater model. These include:

- Long simulation times and the large number of parameters required to simulate the model restricted the parameters used to calibrate the model. The primarily parameters applied were the lateral and vertical hydraulic conductivity to limit the search space to two dimensions.
- The semi-manual method of calibration meant that the models were not exhaustively optimised.
- The discretisation of the model domain was selected to simulate fine scale processes in and around the stream (irregular elements approximately 25m) while the elements became larger around the catchment boundary (~250m). There was no analysis conducted on what the best resolution of elements was. The selection of the size of elements was based on the DEM resolution with the aim of reducing simulation run times.
- Fast and shallow storm-water runoff (or interflow) were not represented in the selected model, this runoff process has the potential to have a significant influence on how fluxes move in and around agricultural storages and in individual reaches in the landscape.
- The spin-up of the models was a limiting factor of the calibration process, where the models may not have entirely reached equilibrium for each of the calibration runs, with simulation runtimes a limiting factor.
- The region is characterised by fractured rock. This feature was not explicitly modelled for the catchment and instead the hydraulic conductivity was used to represent an equivalent porous media.

### **Objective 3 – Representing intermittent streamflow in headwaters using additional data in multi-site calibration**

The study's modelling methods and the application of data had a number of limitations. These include:

- The lack of soil moisture or groundwater data meant that these elements of the model could not be evaluated and that alternative model conceptualisation could not be eliminated.
- Limitations of the computational budget and long simulation runtimes resulted in a semi-manual (using a grid search) calibration method, meaning an optimiser was not used.
- Long simulation times and the vast number of parameters in fully-coupled process models restricted the calibration method, with a semi-manual approach adopted. This approach means that the model was not exhaustively optimized.
- Spin-up dependences resulted due to the difficulty of truly spinning up the model with all updated parameters while also calibrating. This was a challenge due to simulation and time constraints as well as convergence issues. Therefore, simulations were initiated with initial conditions updated from the previous simulation outputs for each subsequent grid search.

- Step-wise approach to the multi-site calibration was a simplified approach to model calibration and allowed for a two-dimensional search space. With improved computer power, alternative approaches can be investigated with multidimensional search spaces, e.g. all sites calibrated at once.
- Data issues - model performance results were dependent on the flow threshold rate applied to simulated data and the threshold applied to the observed binary classification where a zero flow day was not a true zero flow day and was dependent on the location of the sensor used for data collection.
- The censoring of simulated discharge data was dependent on the threshold flow rating being applied to the data. The calculation results of simulation performance were sensitive to these thresholds.
- Reliance on super-computing facilities means that the approach is not yet able to be adopted more widely and that further investment is needed in numerical methods for computational efficiency.

#### 5.4 Further recommendations for future work

There are significant opportunities for further work, with the current limitations also presenting potential options for future developments.

##### **Objective 1 – Quantifying streamflow intermittency and signatures with low-cost sensing technology**

- *Investigating and comparing alternative environmental sensors:* There is significant potential to investigate the viability and accuracy of other on-ground environmental sensor technologies as well as their potential for upscaling. While work has been conducted in this field for two decades (*Constantz et al., 2001; Bradley et al., 2002; Blasch et al., 2004; Goulsbra et al., 2009; Tsubaki et al., 2011; Jaeger & Olden, 2012; Chapin et al., 2014; Hofer et al., 2018; Paillex et al., 2019*) there is still a disproportionate focus of improving modelling methods. Additionally, investigations could include: direct comparisons of electrical conductivity sensors, temperature sensors and other water quality measurements.
- *Image analysis as a low-cost option:* Adding to the field of research, investigating low-cost methods of image analysis to provide flow depths or velocities presents opportunity to provide additional data on headwater streamflow intermittency. Examples include: time-delay photography, video and thermal imaging analysis.
- *Telemetry and similar technologies:* The application of radio transmitters on sensors means that studies can be dramatically up-scaled without the requirement of large human resources for data collection.
- *Upscaling sensor deployment:* This research was a pilot study to illustrate what could be achieved with a short-term data campaign. There is significant potential to upscale the study to quantify intermittent streamflow across larger catchments and across wider regions.
- *Regression of additional data:* There is scope to determine the relationship between downstream (outlet) flow data and upstream intermittency and catchment characteristics which can be applied to delineate flow events upstream. Therefore, regression techniques could be developed that can be applied in similar catchment when only outlet data and catchment features are available.

##### **Objective 2 – Implications of modelling assumptions on the representation of local-scale intermittent streamflow**

- *Agricultural storage dynamics*: Additional studies can determine the influence that different volumes and patterns of extraction have on the dynamics of the storage. Alternative storage configurations, such as multiple storages in series and parallel can also be compared as well as the storage behaviour of in-stream versus off-stream. Research questions around the influence that increasing storage density per unit area has on flows upstream and downstream using process models is also significant in managing water resources across scales.
- *Sophisticated calibration approaches*: Opportunities for further work to include the application of more sophisticated calibration approaches using the collected data. For example, a likelihood method that accounts for binary wet-dry classifications within model calibration and uncertainty.
- *Alternative model structures*: The investigation and comparison of alternative model structures, such as conceptual, semi-distributed and flexible models provide an opportunity to further develop a framework for the application of additional data.
- *Hypothesis testing*: Further analysis using hypothesis testing methods and additional data such as groundwater depth to eliminate some or all of the candidate simulations.
- *More case studies*: Alternative catchments and/or paired catchments provide an additional insight to understand how modelling assumptions may impact the representation of flow processes at the local scale. With further insights modelling frameworks could be updated to better reflect the climate and catchment characteristics
- *Alternative model resolutions*: Investigation of the model resolutions could provide additional information on what resolutions are required to effectively simulate flow for alternative problems
- *Virtual laboratories*: Virtual laboratories allows for additional model structures and processes equations to be investigated. For example, comparing alternative infiltration equations.
- *More direct validation of method quantifying runoff mechanisms*: There is a requirement to directly validate the Hydraulic mixing cell method (*Partington et al., 2011*). For example, tracer studies could provide data to track the accuracy of the simulated runoff generation mechanisms.

### **Objective 3 – Representing intermittent streamflow in headwaters using additional data in multi-site calibration**

- *Alternative calibration approaches*: There are numerous alternative approaches for calibration (this study used a grid search). For example, with increased computing power optimisers can be used. Additionally, parameters can also be investigated for calibration such as the vertical discretization of the soil profile depths.
- *Lumping of parameters*: Additional research is required to investigate the lumping of parameters used for the study. Whether the discretization to five sub-catchments provides adequate parameter representation or whether distributed vegetation and soil properties could provide further insights.
- *Investigate model resolution*: By increasing the resolution of the model, there is still the question of whether the five sub-catchments are the adequate spatial resolution or whether the data points are in the correct locations to capture the heterogeneous behaviour of runoff, given that upstream region of each data location was assumed to be homogeneous. More research is need to determine the resolution required to adequately capture the required processes.

- *Local vs global tradeoffs*: An analysis of the tradeoffs between improved representation of processes upstream and the deterioration of performance at the outlet can provide further answers for the development of frameworks that require coarse resolution modelling.
- *Parameter uncertainty*: An investigation of the impact and measurement and sampling error and how they propagate through the model.

## 5.5 Conclusions

This thesis has illustrated the importance of representing of intermittent flows and the variation of flow throughout a catchment. The adequate representation and understanding local scale processes can significantly improve efficiencies in water allocation, such as the interaction of farm dams and low flows and impacts of land-use changes throughout a catchment. The results of this work have highlighted that current modelling methodologies which typically aggregate catchment processes to a single point-estimate of discharge at the outlet, can have significant implications for water management outcomes. The research has showed that by combining relatively underutilised methods of data collection such as low-cost environmental sensors with a fully-integrated modelling the characterising intermittent flow processes within a catchment were improved.

Finally, technological advancements are increasing at exponential rates and providing unprecedented opportunities to learn and gather data from our environment. Unconventional sources of data, if managed correctly, can significantly assist in making more informed water decisions. The availability of new data streams (remote sensing, drones, infrared, video imagery, etc) will mean that more studies are needed to demonstrate the potential benefits of alternative data and how to blend them into modelling frameworks alongside established sources of measurements from streamflow gauges.

## Appendices

### Appendix A: Field trip photos



**Photo taken of the study catchment headwaters**

# Site S001



25/01/2017 (10:32AM)



25/01/2017 (10:32AM)



25/01/2017 (10:33AM)



25/01/2017 (10:34AM)



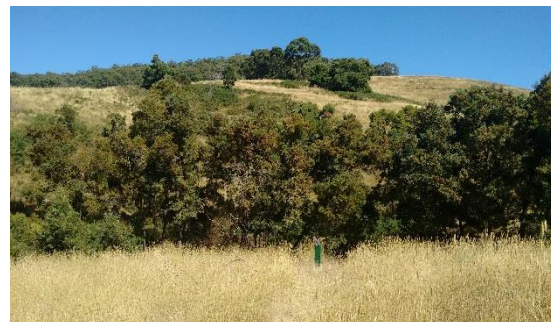
25/01/2017 (10:35AM)



25/01/2017 (10:59AM)



25/01/2017 (10:57AM)



25/01/2017 (10:57AM)





25/01/2017 (11:00AM)



25/05/2018 11am



9/08/2018 (10:43am)

S002



25/01/2017 (10:35AM)



25/01/2017 (10:35AM)



9/08/2018 (10:59am)



25/05/2018

S003



11/10/2017 (12pm)



11/10/2017 (12pm)



11/10/2017 (12pm)



11/10/2017 (12pm)



9/11/2019 (11:26)

S004



23/08/2017 11am



18/03/2019 (12:14)



18/03/2019 (12:14)



18/03/2019 (12:14)



9/08/2018 (10:59am)



9/08/2018 (10:59am)



9/08/2018 (10:59am)

# S006



23/08/2017 (10:40 AM)



23/08/2017 (10:40 AM)



23/08/2017 (10:40 AM)





9/08/2018 (11:23am)



9/08/2018 (11:23am)



# S008



19/07/2018 (12:23)



23/08/2017 (10:54AM)



23/08/2017 (15:56AM)



23/08/2017 (11:02AM)



9/11/2018 (11am)



9/11/2018 (11am)



9/11/2018 (11am)



9/11/2018 (11am)

# S005



23/08/2017 (10:19AM)



23/08/2017 (10:19AM)



23/8/2018 (11:01am)



23/8/2018 (11:14am)



23/8/2018 (11:08am)



23/8/2018 (11:06am)

## S010



8/11/2018 (11:30am)

## S011



8/11/2018 (11:20am)



8/11/2018 (11:20am)



9/08/2018 (11.23am)