Storage in south-eastern Australian catchments

Alexander J.V. Buzacott [1] **R. Willem Vervoort** [1]

[1]Sydney Institute of Agriculture, School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Key Points:

- We adopt a multi-method and multi-catchment approach to estimate storage in south eastern Australia
- Storage in the study catchments is related to variables that indicate slow subsurface movement of water
- The results reinforce the idea that storage is a useful metric for catchment comparison

Corresponding author: Alexander Buzacott, `alexander.buzacott@sydney.edu.au`

Abstract

The storage and subsequent release of water is a key function of catchments and provides a buffer against meteorological and climate extremes. While catchment storage sits at the intersection of the main hydrological processes and largely controls them, it is difficult to quantify due to catchment heterogeneity and the paucity of hydrogeological data. We adopt a multi-method approach to estimate the dynamic and extended dynamic storages using hydrometric data in 75 catchments across the south east of Australia that span across the largest mountain range in the country. The results are compared to hydrological and physical characteristics to determine the main controls of catchment storage. Each of the methods produced a wide range of storage estimates for each catchment, but estimates from each of the methods were largely ranked consistently across the study catchments. Consistent and robust relationships between catchment characteristics and estimates of storage were difficult to establish, however the results suggest that streamflow is derived from slow storage release and long flow paths while a substantial portion of storage is reserved for evapotranspiration. This study highlights some limitations with the current methodology and reinforces the need to collect data that can validate storage estimates at the catchment scale.

1 Introduction

The hydrological system is perhaps best characterised by the volume of water stored within a catchment (McNamara et al., 2011). Storage directly influences the runoff response (Spence, 2007), stream water chemistry (Hrachowitz, Fovet, Ruiz, & Savenije, 2015; Kirchner & Neal, 2013), drought severity (Van Loon & Laaha, 2015) and transpiration behaviour (Dawson, 1996; Jackson, Sperry, & Dawson, 2000). While as early as 1967, the seminal variable source area work of Hewlett and Hibbert (1967) highlighted the importance of storage, the topic has been mostly neglected by hydrologists. Instead, much of the work on understanding the hydrological system has focussed on quantifying catchment fluxes (Soulsby, Tetzlaff, & Hrachowitz, 2009). Much of the neglect stems from the elusive nature of storage. Storage is difficult to characterise or observe at the catchment scale (Seyfried, Grant, Marks, Winstral, & McNamara, 2009), owing to its large spatial heterogeneity and the limited inference than can be drawn from individual observations (Soulsby et al., 2008).

An improved sense of how and how much water is retained in catchments will in turn provide a greater understanding of how water is released from catchments (McNamara et al., 2011). As an example, a major goal of catchment hydrologists has been to accurately predict streamflow for scenario analysis and forecasts. However, to achieve good model performance, the water balance and other hydrological processes are often violated (Kirchner, 2006) which results in a poor simulation of the temporal storage and release of water. This is exemplified in the study by K. Fowler et al. (2020), which showed five conceptual models failing to reproduce long term declines in water storages over an extended drought. Rather, the models prioritise seasonal cycles of water storage in a more dynamic fashion. Beyond water yield from catchments, storage also strongly controls water quality. Many biogeochemical reactions depend on subsurface contact time (Hornberger, Scanlon, & Raffensperger, 2001; Kirchner, 2003) and this subsequently affects the persistence of pollutants (Hrachowitz et al., 2016).

More recently, the role of storage within the hydrological cycle has received greater attention (e.g. Buttle, 2016; Fan, 2019; McNamara et al., 2011; Soulsby et al., 2009; Spence, 2007, 2010; Tetzlaff, McNamara, & Carey, 2011). This recognises that storage is understudied (Soulsby et al., 2009) but it is also driven by novel methods that describe catchment storage, such as through recession analysis (Kirchner, 2009), tracer applications (Gleeson, Befus, Jasechko, Luijendijk, & Cardenas, 2016; Soulsby et al., 2009) and remote sensing methods at a larger scale (Ramillien, Famiglietti, & Wahr, 2008). McNamara et al. (2011) proposed using standardised methods and comparative investigations of storage across a range of environments to yield better insights into relationships between catchment processes and storage dynamics. Since then, a few studies have employed a multi-method and multi-catchment approaches to investigate storage (e.g. Peters & Aulenbach, 2011; Sayama, McDonnell, Dhakal, & Sullivan, 2011; Staudinger et al., 2017), however globally such studies are still sparse.

Much of the current research has been devoted to understanding storage in headwater catchments. Headwater catchments are often located in mountainous regions and provide high volumes of river flows to lowland areas, such that they are considered the "water towers of the world" (Viviroli, Dürr, Messerli, Meybeck, & Weingartner, 2007). These flows are important for maintaining hydrologic connectivity and ecological integrity of regional hydrologic systems (Freeman, Pringle, & Jackson, 2007) and are important sources of water for downstream human water demands. Headwater catchments in montane areas are particularly vulnerable to climate change and other anthropogenic developments (Immerzeel et al., 2020; Viviroli et al., 2011) and a lack of a deep understanding of catchment storage threatens global water security. In the south-east of Australia, forested catchments along the Great Dividing Range are responsible for large inflows into the Murray-Darling Basin, which is Australia's largest food bowl (Wheeler, 2014) and a region of significant ecological importance. In Australia, as well as in other temperate to semi-arid regions across the globe, droughts are a frequent phenomenon and are often severe. Water stored and later released by headwater catchments serve as a buffering mechanism that can reduce the impacts of drought and understanding the role catchment storage plays in sustaining streamflows and evapotranspiration is therefore crucial.

In this study, we build on the multi-method and multi-catchment approaches and evaluate the different levels of storage in catchments across south eastern Australia. Subsequently we are interested in what landscape and climate factors may be associated with storage, such that specific catchments can be protected and managed effectively. Catchment characteristics may also reveal common controls on catchment storage (Geris, Tetzlaff, & Soulsby, 2015; Saft, Peel, Western, & Zhang, 2016; Wagener, Sivapalan, Troch, & Woods, 2007). We assess the relationship between the estimates of storage from the different approaches to fundamental hydrological and physical catchment characteristics. In addition, we evaluate if the methods here allow storage to be used as an appropriate metric for catchment comparison (Buttle, 2016). Specifically, the aims of this paper are to (1) Estimate and evaluate the dynamic storage, extended dynamic storage and total storage of catchments in the south east of Australia (2) Determine if there are robust relationships with catchment characteristics and if the approach is useful as a metric for catchment comparison and (3) Evaluate the results with respect to the study area and calculate useful metrics, such as the turnover time, and discuss the significance.

2 Materials and Methods

2.1 Defining catchment storage

Water storage can be considered the sum of the individual stores of water that exist within catchments, such as groundwater, soil moisture, vegetation, surface water and snow. Generally, the term storage is used inconsistently in hydrology and may include or omit some of these features (Condon et al., 2020; McNamara et al., 2011), largely owing to the diverse applications and specific domains of hydrological studies. We follow the suggestion of McNamara et al. (2011) and evaluate a wide range of catchments using standardised methods to investigate the relationship between storage dynamics and catchment processes. Staudinger et al. (2017) created a scheme that distinguishes different perceptual catchment storages (Figure 1). The different conceptual storages are: total storage, immobile storage, mobile storage, extended dynamic storage and dynamic storage. The partitions are based on specific methodologies that derive them and are of practical interest. Total storage can be considered the sum of all water stored in the catchment, including both mobile and immobile water. Total storage can be estimated through an aggregation of hydrogeological assessment of aquifers, groundwater and soil moisture information. Immobile water is water that does not participate in the hydrological cycle and may be found in bedrock with poor permeability (Staudinger et al., 2017). Mobile water is water that participates in the hydrological cycle and is connected to catchment fluxes. Mobile water can comprise of water of a variety of ages, such soil moisture (young), shallow groundwater and deep groundwater (old) passing through fractured rock systems. Estimates of mobile water can be obtained using tracer methods (Birkel, Soulsby, & Tetzlaff, 2011; Cartwright & Morgenstern, 2016; Howcroft, Cartwright, & Morgenstern, 2018) or through hydrological transport models (Rinaldo et al., 2015; van der Velde, Torfs, van der Zee, & Uijlenhoet, 2012), however these models need verification with tracer data.

Dynamic storage is the storage that controls streamflow dynamics (Birkel et al., 2011; Kirchner, 2009; Spence, 2007) and can be linked to evapotranspiration dynamics, such as diurnal streamflow variation (Gribovszki, Kalicz, Szilágyi, & Kucsara, 2008; Mutzner et al., 2015; Teuling, Lehner, Kirchner, & Seneviratne, 2010). Dynamic storage can be estimated from streamflow data alone using, for example, streamflow recession analysis (Kirchner, 2009) or by hydrological modelling (K. Fowler et al., 2020; Staudinger et al., 2017). For non-perennial streams, such as intermittent or ephemeral streams, there are periods when there is no streamflow yet storage continues to decrease due to subsurface water flow and evapotranspiration. 'Extended dynamic storage' (Staudinger et al., 2017) estimates this storage when all catchment fluxes cease and the storage approaches zero. Extended dynamic storage can be estimated using modelling or the cumulative water balance.

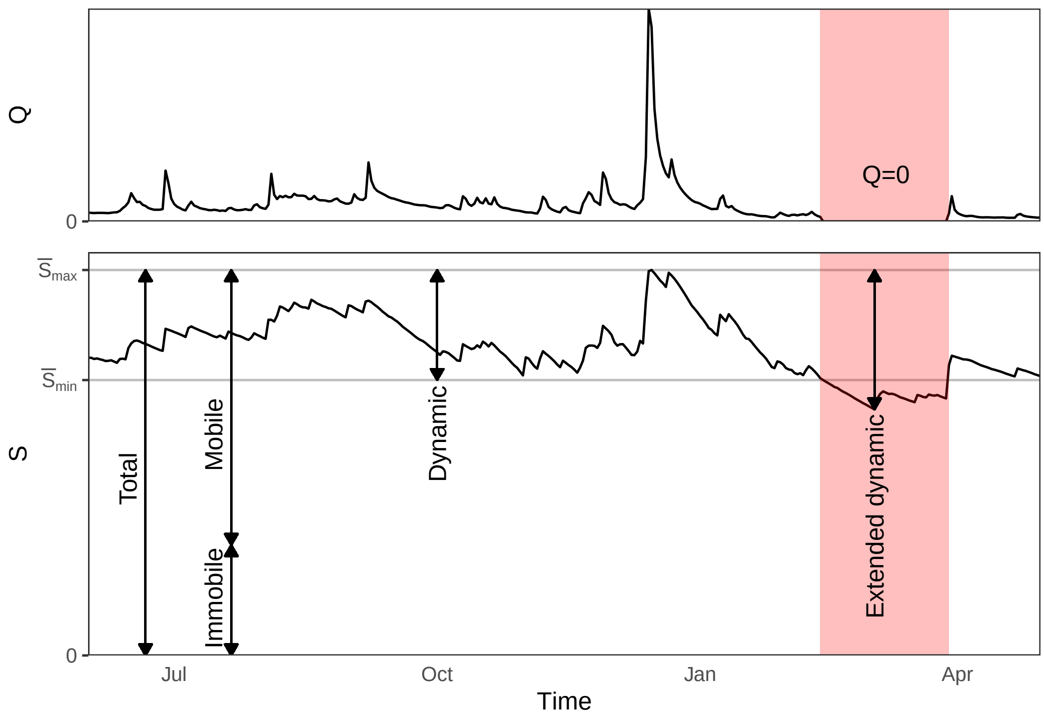

Figure 1. Illustration of different conceptual ideas of storage within a catchment, as adapted from Staudinger et al. (2017; Figure 1). The top panel shows catchment streamflow (Q) and the bottom panel catchment storage (S) through time. The red shaded area indicates a period when streamflow ceases yet catchment storage still decreases.

2.2 Study catchments and data

A subset of catchments located in south eastern Australia was selected from the Australian Bureau of Meteorology (BOM) Hydrological Reference Station (HRS) project (X. S. Zhang et al., 2016). HRS are catchments with high quality streamflow records and are located in areas with minimal land use change and impacts of water resource development and are ideal for long term analysis. Catchments were selected in the states of New South Wales, Victoria and Tasmania and in the Australian Capital Territory. We limited selection to catchments in these states/territories for two reasons. Firstly, the majority of the catchments are located along the Great Dividing Range, a mountain range that runs along the east coast of Australia. Secondly, the choice restricts the climatic and geographic diversity of included catchments. This provides a greater chance that storage can be robustly estimated and improve comparability. Catchment selection was further refined by the availability of high quality data, as described later.

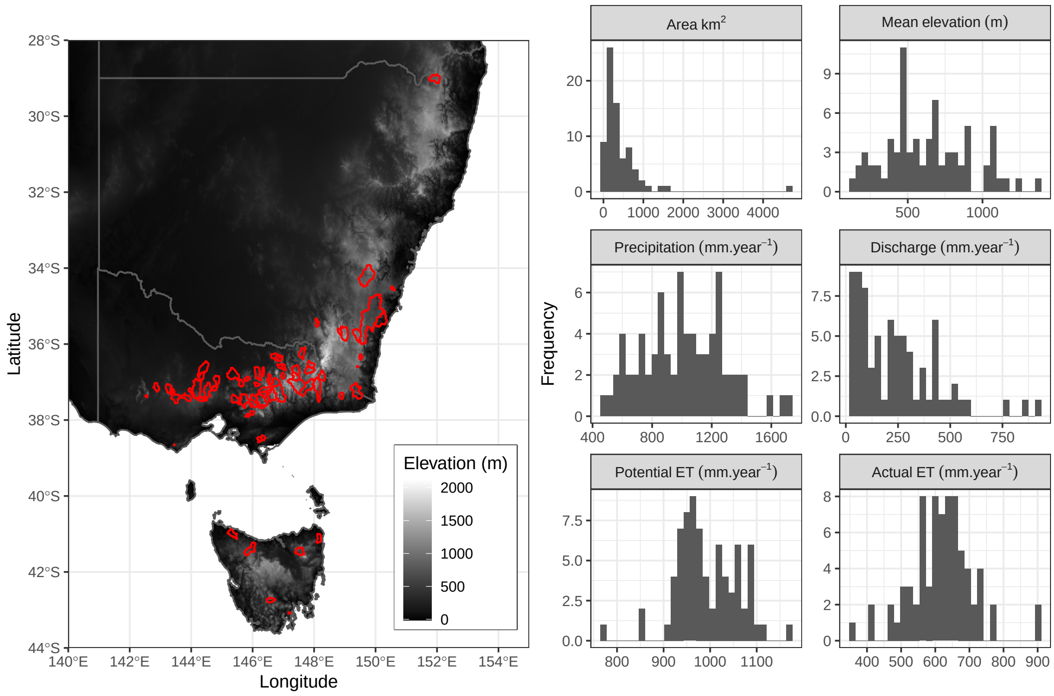

Figure 2. Study catchments included in this study and histograms of key hydrological and physical variables. Potential and actual evapotranspiration (ET) are calculated using Morton's models.

The study period focuses on 1990-2018. This time range includes both distinct wet and dry periods. The 1990s, early 2010s and 2016 are notably wetter years, while the Millennium Drought (van Dijk et al., 2013) was a severe drought that extended over much of the 2000s.

2.3 Hydrological data

Streamflow data was obtained from the BOM HRS portal (`http://www.bom.gov.au/water/hrs/`). There is no missing data in the records as the BOM gap fills the data using the GR4J model. Gauges needed to have more than 70% of data classified with the highest quality code (A) over the study period. After the geographical and data quality filtering, 75 catchments were selected for analysis in this study (Figure 2).

Daily catchment means of precipitation, maximum temperature, minimum temperature, vapour pressure and solar radiation were extracted from the Australian Water Availability Project (AWAP) dataset (Jones, Wang, & Fawcett, 2009) via the AWAPer

R package (Peterson, Wasko, Saft, & Peel, 2020). Monthly potential evapotranspiration (PET) and actual evapotranspiration (ET) are calculated using Morton's model (Morton, 1983) of wet areal environment evapotranspiration and actual areal evapotranspiration, respectively, using the R package Evapotranspiration (Guo, Westra, & Maier, 2016). Daily estimates of PET are obtained by linear interpolation of the monthly estimates within the AWAPer R package. The choice of Morton's models is motivated by the suitability of the models to calculate catchment water balances and within rainfall-runoff modelling (McMahon, Peel, Lowe, Srikanthan, & McVicar, 2013). A summary of the main catchment hydrological forcings is presented in Figure 2.

2.4 Dynamic storage: Storage discharge relationship

The first method uses the storage discharge (SD) relationship to estimate dynamic storage. The storage discharge relationship is derived by examining the relationship of streamflow recession (-dQ/dt) and discharge (Q) during minimal flux periods of precipitation (P) and evapotranspiration (ET). Kirchner (2009) showed that during these periods, storage is theoretically a function of discharge (i.e. $S = f^{-1}(Q)$) and several studies have applied the method (Ajami, Troch, Maddock, Meixner, & Eastoe, 2011; Birkel et al., 2011; Staudinger et al., 2017; Teuling et al., 2010; Yeh & Huang, 2019). Dynamic storage is estimated as the difference between maximum (S_{max}) and minimum storage (S_{min}) corresponding to some maximum (Q_{max}) and (Q_{min}) discharge rates. We estimate dynamic storage using the means of the annual maxima and minima of flows for each catchment, as done by Kirchner (2009).

$$S_{max} - S_{min} = \int_{Q_{min}}^{Q_{max}} \frac{1}{g(Q)} dQ \tag{1}$$

where g(Q) is:

$$g(Q) = \frac{dQ}{dS} = \frac{dQ/dt}{dS/dt} \approx \frac{-dQ/dt}{Q} \mid_{P<<Q,ET<<Q} \tag{2}$$

Daily data was used to estimate the storage discharge relationships. While hourly data was used in his original study, Kirchner (2009) also demonstrated that daily data could yield similar estimates of storage with a sufficient amount of data points. Kirchner (2009) selected days in the recession where P and ET were less than 10% of discharge.

In south east Australia, the latter condition of ET being less than 10% of discharge is rarely met as high rates of ET are possible even in cooler seasons. This results in an insufficient amount of data points to calculate robust storage discharge relationships. To minimise the effect of catchment fluxes, days on and after precipitation occur are excluded, and the months between June and August are used to calculate storage discharge relationships. This may result in some of the storages being underestimated due to the effects of ET and this will be discussed later.

2.5 Extended dynamic storage: water balance

The second method uses the cumulative running water balance to calculate the extended dynamic storage:

$$S(t) = S_0 + \Delta t \sum_{i=1}^{i=t} P_i - Q_i - ET_i.s_{ET} \tag{3}$$

where $S(t)$ is the storage at time step t, S_0 is the initial storage at time step $t = 0$, P is precipitation, Q is streamflow, ET is actual evapotranspiration and s_{ET} is the evapotranspiration scaling factor. P, Q and ET are in mm per timestep, which is monthly as Morton's actual ET (AET) is calculated monthly. ET was scaled for each catchment using a scaling factor s_{ET} to ensure the water balances closed (equivalent to f_{WB} in Equation 2, Staudinger et al., 2017). s_{ET} is calculated as:

$$s_{ET} = \frac{\overline{P} - \overline{Q}}{\overline{ET}} \tag{4}$$

where $\overline{P}$, $\overline{Q}$ and $\overline{ET}$ are mean annual precipitation, discharge and actual evapotranspiration, respectively. Extended dynamic storage is calculated as the difference between the maximum and minimum storage volumes observed over the study period (1990-2018).

2.6 Extended dynamic storage: Budyko framework

A second estimate of extended dynamic storage is obtained using the Budyko framework (Budyko, 1974) to estimate annual evapotranspiration and subsequently the water balance. The Budyko framework relates the index of dryness (PET/P) and the evaporative index (ET/P) on the basis that water availability and atmospheric demand are the primary constraints on the equilibrium water balance (J. Y. Zhang, Wang, & Wei, 2008). The Budyko curve therefore captures the interactions and feedbacks between the atmosphere, vegetation and soil within the hydrological cycle (van der Velde et al., 2014). The Fu-Zhang equation (Fu, 1981; L. Zhang et al., 2004), a Budyko-like equation, is used in this study and is defined as:

$$\frac{\overline{ET}}{\overline{P}} = 1 + \frac{\overline{PET}}{\overline{P}} - \left[1 + \left(\frac{\overline{PET}}{\overline{P}}\right)^{w}\right]^{1/w} \tag{5}$$

where w is an adjustable catchment parameter. The implementation of the w parameter allows for representation of geographical variation of the Budyko curve and the integrated effects of vegetation cover, soil properties and catchment topography (L. Zhang et al., 2004). In long term water balances ET is estimated to equal $\overline{ET} = \overline{P} - \overline{Q}$, assuming negligible changes in catchment storage (i.e. $\Delta S = 0$). The catchment w parameter is optimised using the least-squares approach and values $1 < w \leq 10$ are evaluated.

This approach yields the average annual evapotranspiration, however we are interested in the inter-annual variation in evapotranspiration and subsequently the water balance to derive storage. ET is limited by water availability and energy, but water availability can be carried through time via storage and is not simply a result of annual precipitation. Zeng and Cai (2015) showed that the water balance (ΔS) can be integrated into the Fu-Zhang equation to obtain an estimate of inter-annual ET:

$$ET_i = P'_i\left[1 + \frac{PET_i}{P'_i} - \left[1 + \left(\frac{PET_i}{P'_i}\right)^{w}\right]^{1/w}\right] \tag{6}$$

where i is the timestep (annual in this case), $P'_i = P_i + \Delta S_{i-1}$, and w is the optimised catchment parameter from equation (5). The annual running water balance is calculated using P, Q and the estimations of ET from equation (6) and the extended dynamic storage is estimated as the difference between the maximum and minimum observed level of ΔS.

2.7 Dynamic and extended dynamic storage: conceptual model

The last approach estimates dynamic and extended dynamic storages using a conceptual hydrological model using the same approach as Staudinger et al. (2017) (Section 3.3). In this method, the calibration of model parameters controls the sizes of the storage state variables in the model. The variation of the storage state variables over time is then used to calculate the dynamic and extended dynamic storage. An adaptation of the HBV-light model as described in Seibert and Vis (2012) is used within the R package `hydromad` (Andrews, Croke, & Jakeman, 2011). HBV model parameters are calibrated using the Shuffled Complex Evolution - University of Arizona (SCE-UA) algorithm (Duan, Sorooshian, & Gupta, 1992) and the Nash-Sutcliffe Efficiency objective function (Nash & Sutcliffe, 1970) with the Linsdtröm penalty for volume error (R_V^2) (Lindström, 1997). The parameter ranges used in calibration are presented in Supporting Information S1 Table S1. The full study period (1990-2018) is used to calibrate the model for each catchment. The state variables within HBV that store water are: snow depth, soil moisture, upper groundwater storage and lower groundwater storage. The extended dynamic storage is estimated as the sum of the maximum size of the HBV state variables. The dynamic storage is estimated as the sum of the differences between the maximum and minimum sizes of the HBV state variables.

2.8 Catchment characteristics

Several catchment physical characteristics are used to explore the controls on catchment storage. The BOM Geofabric V2.1 product (`http://www.bom.gov.au/water/geofabric/`), a stream and nested catchment framework for Australia (Stein, Hutchinson, & Stein, 2014), is used to extracted several characteristics including: mean elevation, elevation range, stream density, stream length, slope, saturated hydraulic conductivity for the A horizon (KSat) and the proportion of catchment grid cells that are valley bottoms (henceforth named PVB). Three geological attributes are also extracted; the catchment areal percent proportion of igneous rocks, sedimentary rocks and metamorphic rocks. The catchment average of the Silica Index (Gray, Bishop, & Wilford, 2016; Gray, Bishop, & Yang, 2015), a broad classification of soil parent material that focuses on chemical composition rather than the formation process, is an additional measure included to evaluate the effect of lithology on storage. Catchment average soil depth and clay content in the top metre of soil are extracted from the Soil and Landscapes Grid of Australia (Grundy et al., 2015) and Plant Available Water Capacity (PAWC) within the top metre is extracted from the Australian Soil Resource Information System (Johnston et al., 2003). The Australian Woody Vegetation Cover product (Gill et al., 2017) is used to calculate two variables: proportion of forest cover and foliage projective cover.

Three additional catchment characteristics were calculated using hydrometric data: the coefficient of annual streamflow variability (Q_{cv}), the mean annual aridity index (P/PET), annual runoff ratio (Q/P), the baseflow index (BFI) and the lag-1 day autocorrelation coefficient (AC) (Winsemius, Schaefli, Montanari, & Savenije, 2009). The BFI has been shown to represent the storage and release properties of catchments (Salinas et al., 2013; Van Loon & Laaha, 2015) and was calculated using the lfstat R package (Koffler & Laaha, 2013). The lag-1 autocorrelation is a measure of smoothness of the hydrograph and can provide insights into water release properties of a catchment, where a higher autocorrelation coefficient indicates a slower release of water from the catchment. It is also considered one of the key hydrological signatures (Euser et al., 2013).

The study catchments cover a wide range of catchment physical properties and characteristics (Table 1). The catchment areas range from 4.5 to 4660 km^2. The vast majority of the catchments are forested (mean 86%), while the woody fractional cover varies from 7% to 86%. Igneous and sedimentary rocks are the most common underlying geologies of the catchments. Soils are moderately deep (mean depth is 0.73 - 1.13 m) with a range of clay fractions (22-44%).

Mean annual precipitation ranges from 473 mm/year to 1721 mm/year. Mean annual discharge ranges from 20 mm/year to 909 mm/year and highly variable, where the annual variance (Qcv) ranges from 724 mm/year to 125,095 mm/year. Greater precipitation and discharge are mildly correlated to latitude, with Pearson's correlations (r) of -0.31 and -0.41, respectively. PET increases with latitude ($r = 0.64$), while AET decreases ($r = -0.32$) but increases with longitude ($r = 0.4$), suggesting a limit of water availability. Two catchments are semi-arid ($0.20 < \text{P/PET} < 0.5$), eight are dry subhumid ($0.50 < \text{P/PET} < 0.65$) and the remaining catchments ($n = 70$) are temperate. No catchments are considered "cold" catchments as the minimum mean annual PET (778 mm/year) exceeds the 400 mm/year threshold.

Spearman's *rho* statistic (ρ) is used to evaluate the association between the different storage estimates and catchment properties. Significance ($P < 0.05$) of the relationship is evaluated using Spearman's rank correlation test Algorithm AS 89 (Best & Roberts, 1975).

Table 1. Numerical summary of the catchment characteristics.

Characteristic	Min	1st quartile	Median	Mean	3rd quartile	Max
Lat (°)	-43.07	-37.46	-37.11	-37.17	-36.59	-29.03
Lon (°)	142.51	145.39	146.32	146.54	147.69	151.73
Area (km^2)	4.50	126.10	284.10	417.72	525.95	4660.00
Elev mean (m)	108.07	451.85	604.73	634.25	821.85	1351.28
Elev range (m)	141.30	487.36	795.22	848.36	1215.73	1750.86
Slope (°)	0.81	4.47	7.66	7.83	11.25	14.98
Stream length (km^2)	0.28	1.32	2.10	2.66	3.46	9.76
Stream density (km/km^2)	0.49	0.74	0.83	0.82	0.89	1.14
Soil depth (m)	0.73	0.90	0.96	0.95	1.01	1.13
Clay (%)	22.45	28.07	30.38	30.94	33.60	44.08
KSat (mm/hr)	30.00	80.27	174.85	159.96	223.27	300.00
PAWC (mm/m)	58.35	101.20	124.04	123.65	151.70	176.81
Forest Cover (%)	25.58	73.75	95.33	85.50	99.95	100.00
Foliage Cover (%)	7.09	27.12	44.01	44.18	57.68	85.89
Silica Index	57.34	67.02	68.43	68.52	71.14	80.00
PVB (%)	0.00	0.00	0.04	2.40	1.23	31.23
Regolith depth (m)	1.41	2.19	3.66	4.49	5.48	23.94
Igneous rocks (%)	0.00	8.99	25.29	33.69	56.31	100.00
Sedimentary rocks (%)	0.00	27.82	54.02	52.22	83.67	100.00
Metamorphic rocks (%)	0.00	0.00	0.00	5.43	0.00	90.63
Qcv (mm/year)	724.19	4921.64	13414.17	19509.47	28811.39	125095.71
P (mm/year)	473.79	810.11	1000.69	1005.91	1223.37	1721.20
Q (mm/year)	19.39	78.47	211.37	240.30	351.27	909.22
PET (mm/year)	777.62	949.69	979.87	993.07	1043.57	1176.77
AET (mm/year)	355.73	564.91	619.40	618.64	668.80	900.20
P/PET	0.43	0.80	1.02	1.03	1.23	1.91
Q/P	0.03	0.11	0.19	0.21	0.30	0.55
BFI	0.09	0.28	0.49	0.45	0.60	0.81
AC	0.40	0.57	0.68	0.71	0.84	0.97

3 Results

3.1 Storages

Estimates of storage covered wide ranges of values for each of the methods across all catchments (Figure 3). Robust storage discharge relationships were found for the vast majority of the catchments, where the mean and standard deviation of the coefficient of determination was (mean ± standard deviation) $R^2 = 0.93 \pm 0.07$. The minimum R^2 was 0.48. Storage values for the storage discharge method range from $4 - 493$ mm with a mean storage value of 42 mm. One station located near the Sydney basin (AWRC ID: 212209) is an outlier for the storage discharge method with an estimated dynamic storage of 493 mm, where the next largest storage is estimated to be 108 mm (AWRC ID: 405264). Recession plots and plots of the storage-discharge relationships are presented in Supporting Information S1 Figure S2 and Figure S3, respectively.

All HBV models obtained reasonable calibration scores ($R_V^2 = 0.70 \pm 0.11$). All catchments obtained a NSE component above 0 (minimum $R_V^2 = 0.35$), which is often used to distinguish good and bad performance (Knoben, Freer, & Woods, 2019). Dynamic storages derived from calibration of the HBV model are generally larger (mean 303 mm) and have a narrower distribution (range 90-576 mm) than those derived from the storage discharge relationships.

HBV extended dynamic storage estimates covered a range from 112-1114 mm and had a mean storage of 596 mm. Extended dynamic storages estimated by the water balance method range from 577-2993 mm and had the highest mean value of 1483 mm. The water balance scaling factor, s_{ET}, had a mean and standard deviation of 1.27 ± 0.21, highlighting that most catchments required greater actual evapotranspiration than estimated using Morton's relationship to close the water balance. The relation of s_{ET} to storage is described later. Budyko curve derived water balance storage estimates covered the widest range from 76-3631 mm, but were on average smaller with a mean value of 598 mm. The Fu-Zhang parameter w had a mean and standard deviation of 4.63 ± 2.27 across all catchments and the parameter's relationship to storage is discussed later.

There is little agreement on the size of the storage for each catchment across the methods. Using the dynamic and extended dynamic HBV storage estimates as the reference level for either method, the differences were calculated (Figure 3). The mean and

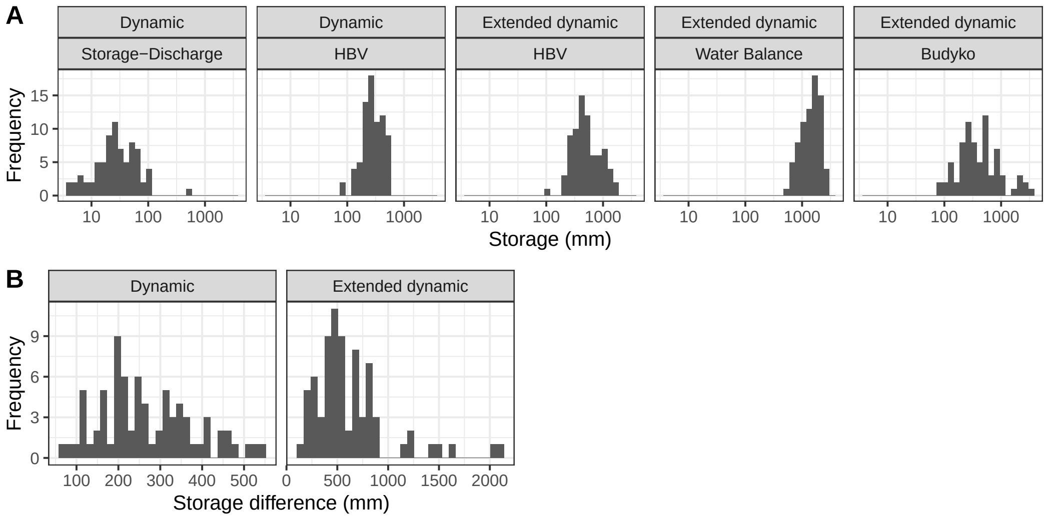

Figure 3. (A) Distributions of storage values for each the method. (B) Mean differences of the storage estimates for the two methods: dynamic and extended dynamic. The differences are calculated using HBV storage estimates as the reference level for both methods.

Table 2. Rankings of the storage size across all catchments and methods. Rank 1 represents the smallest storage and Rank 5 the largest storage.

		Rank				
Storage	Method	1	2	3	4	5
Dynamic	Storage-Discharge	74	0	1	0	0
	HBV	1	54	20	0	0
Extended dynamic	HBV	0	1	19	55	0
	Budyko	0	20	35	18	2
	Water Balance	0	0	0	2	73

standard deviation of the storage differences are 269±114 mm for the dynamic method and 627±385 mm for the extended dynamic storage method. To determine if there are consistent storage size differences between the methods, we ranked the sizes of the estimated storage for each catchment. The rankings of the storage sizes are consistent, except for the Budyko method (Table 2). The storage discharge approach consistently yielded the smallest storages and the water balance approach the largest, with HBV in between. The Budyko method had the largest spread of storage ranks. This result is also indicated by the Spearman correlation matrix for the different storage estimates (Table 3). All methods significantly correlate to each other, with the only exception being the storage discharge and the extended dynamic HBV methods.

Table 3. Spearman correlation coefficients for the storage estimates. Significant correlations ($P < 0.05$) are bold.

		Dynamic		Extended-dynamic		
		SD	HBV	HBV	WB	Budyko
Dynamic	SD	**1.00**	**0.39**	**0.33**	**0.47**	**0.69**
	HBV		**1.00**	**0.93**	**0.58**	**0.62**
Extended dynamic	HBV			**1.00**	**0.48**	**0.56**
	WB				**1.00**	**0.70**
	Budyko					**1.00**

3.2 Physical characteristics

Significant Spearman correlations ($P < 0.05$) were found for several characteristics across all storage methods (Table 4). Greater mean catchment slope is associated with greater storage. This is also indicated with PVB, where the greater the proportion of valley bottoms in catchments indicate less storage. This is in line with previous suggestions that steeper catchments have more vertical infiltration and longer groundwater flow paths (Jasechko, Kirchner, Welker, & McDonnell, 2016), which in turn suggests more groundwater storage. Moreover, steeper catchments tend to have areas of detention storage where water may be stored.

Greater soil depth, unsurprisingly, indicates greater water storage. This is despite clay, the particle size fraction responsible for the greatest water storage potential, having no significant correlation with storage. However, the PAWC was significant for all methods and this may indicate that the PAWC captures some of the water retention properties of catchments. The hydraulic conductivity of the A horizon positively correlates with storage, suggesting that free drainage to lower soil profiles and groundwater increases storage. Mean annual precipitation and the aridity index indicate that the wetter catchments have more storage potential. The BFI has strong correlations with all storage methods, suggesting the digital low pass filter is capturing some aspect of storage and release properties.

No geological variables had consistent and strong relationships to the different storage estimates. Sedimentary rocks, which are the dominant geological rock across the catchments in this study, only had a weak relationship to the storage discharge derived storage estimates, while metamorphic rocks had a weak association to HBV dynamic storage estimates. Igneous rocks had no significant relationships to any of the storage estimates. For the hydrometric variables, storage was significantly correlated with Q and Qcv for all methods except for the extended dynamic HBV estimates, effectively meaning that catchments with greater mean annual flow and variance have greater storage capacity.

3.3 HBV partitioning

The HBV model has conceptual stores for snow, soil water and groundwater and can provide insights into the simulated partitioning of water storage in the study catchments. The calibrated models show that soil storage is simulated as the largest storage for most catchments (Figure 4). Groundwater storage is the next largest storage, but the distribution is long tailed and some catchments have large simulated groundwater storages (maximum 1139 mm). Snow storage is minimal with most catchments having zero simulated snowfall. Predicted soil water storage has a moderate association to soil depth ($\rho = 0.60$), BFI ($\rho = 0.52$), and mean annual P ($\rho = 0.47$) (Supporting Information S1 Table S2). Of the other soil characteristics, predicted soil water storage has significant correlations to PAWC ($\rho = 0.40$) and KSat ($\rho = 0.29$). There is surprisingly an insignificant relationship with clay content ($\rho = 0.07$). Groundwater storage had the greatest association with the BFI ($\rho = 0.65$), P/PET ($\rho = 0.65$) and PVB ($\rho = -0.64$). The positive associations between the BFI and the conceptual storages are likely to be due to the BFI and the model calibration identifying the same low pass signal.

Table 4. Spearman correlation coefficients between storage estimates and the catchment characteristics. Bolded values are significant ($P < 0.05$) correlations.

	Dynamic		Extended dynamic		
Characteristic	SD	HBV	HBV	WB	Budyko
Lat (°)	-0.16	-0.11	-0.15	0.11	-0.06
Lon (°)	**0.36**	-0.04	0.02	0.16	**0.4**
Area (km^2)	-0.03	-0.22	**-0.24**	-0.2	**-0.29**
Elev mean (m)	**0.47**	**0.32**	**0.32**	0.14	**0.4**
Elev range (m)	**0.45**	**0.37**	**0.32**	0.2	**0.32**
Slope (°)	**0.71**	**0.55**	**0.52**	**0.44**	**0.6**
Stream length (km^2)	-0.03	-0.08	-0.07	-0.2	-0.11
Stream density (km/km^2)	-0.07	0.03	-0.02	0.14	-0.07
PVB (%)	**-0.58**	**-0.56**	**-0.6**	**-0.44**	**-0.62**
Regolith depth (m)	**-0.31**	-0.08	-0.1	-0.1	-0.23
Soil depth (m)	**0.24**	**0.69**	**0.57**	**0.62**	**0.38**
Clay (%)	0.12	0.18	0.16	0.07	0.07
KSat (mm/hr)	**0.63**	**0.53**	**0.53**	**0.41**	**0.63**
PAWC (mm/m)	**0.49**	**0.55**	**0.51**	**0.59**	**0.58**
Forest Cover (%)	**0.69**	0.21	0.23	**0.35**	**0.59**
Foliage Cover (%)	**0.66**	**0.28**	**0.28**	**0.34**	**0.63**
Igneous rocks (%)	-0.22	-0.11	-0.08	-0.21	-0.05
Sedimentary rocks (%)	**0.3**	-0.03	-0.03	0.12	0.04
Metamorphic rocks (%)	-0.03	**0.34**	**0.35**	0.17	0.19
Silica Index	0.16	0.09	0.1	**0.28**	0.21
Qcv (mm/year)	**0.86**	**0.36**	**0.26**	**0.52**	**0.75**
P (mm/year)	**0.82**	**0.71**	**0.64**	**0.57**	**0.83**
Q (mm/year)	**0.9**	**0.46**	**0.39**	**0.46**	**0.77**
PET (mm/year)	**-0.5**	-0.21	**-0.3**	0.05	**-0.46**
AET (mm/year)	**0.63**	**0.24**	**0.25**	0.23	**0.58**
P/PET	**0.83**	**0.66**	**0.63**	**0.49**	**0.83**
Q/P	**0.88**	**0.33**	**0.26**	**0.39**	**0.7**
BFI	**0.66**	**0.73**	**0.67**	**0.55**	**0.7**
AC	**0.67**	**0.53**	**0.43**	**0.4**	**0.49**

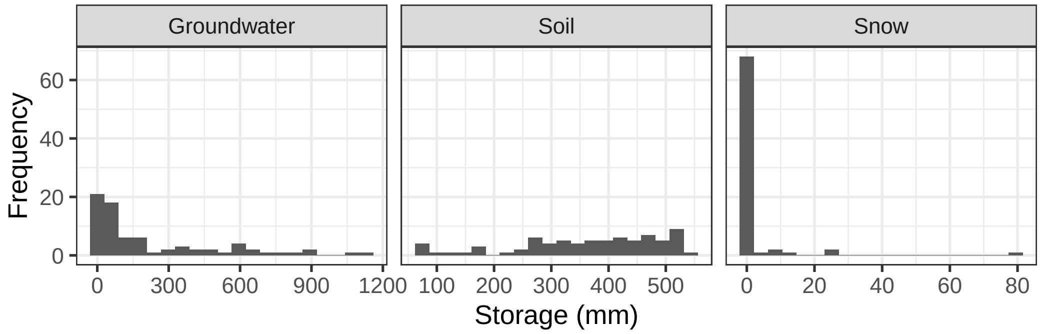

Figure 4. Distributions for each HBV conceptual storage component derived from the extended dynamic storage method.

3.4 Water balance

The water balances were scaled by the scaling factor s_{ET} to ensure the water balances closed over the study period. As mentioned, s_{ET} has a mean and standard deviation of 1.27 ± 0.21 across all the catchments and the minimum and maximum values are 0.77 and 1.81, respectively. The factor has a significant positive correlation with water balance derived storage estimates ($\rho = 0.54$), essentially meaning that the larger the storage the greater scaling factor required. We evaluated whether s_{ET} has any relationships to the catchment characteristics to determine if the scaling factor is representative of any characteristics or has any spatial relationships (Supporting Information S1 Table S3). Soil depth ($\rho = 0.56$), the BFI ($\rho = 0.42$), PAWC ($\rho = 0.41$) had the greatest Spearman correlations. Annual precipitation ($\rho = 0.34$) and P/PET ($\rho = 0.28$) have a significant and positive relationship. These characteristics together all relate to water availability and suggest that evapotranspiration is underestimated.

s_{ET} also has a significant relationships with the percentage of metamorphic rocks ($\rho = 0.41$), slope ($\rho = 0.34$), PVB ($\rho = -0.31$) and elevation range ($\rho = 0.24$). This suggests that there may be terrain and geological factors that influence s_{ET} or water loss from the catchment. Spatially, there was only a mild but significant correlation with latitude ($\rho = 0.26$).

3.5 Budyko approach

The relationships of the catchment storage and the calibrated Fu-Zhang curve parameter w in the Budyko space is presented in Figure 5. Greater catchment storage is associated with a lower aridity index ($\rho = -0.83$) and a lower evaporative index ($\rho = -0.70$). The calibrated parameter w had a weak association with storage ($\rho = 0.22$, $r = 0.49$) and was not significant at the 95% level ($P = 0.053$). There are significant ($P < 0.05$) associations with metamorphic rocks, annual runoff ratio, mean elevation, Qcv, old rocks and PVB.

The distribution of points in Figure 5 show the catchments generally respecting the Budyko water and energy limits. Notably, there are a few catchments that plot left of the energy limit and have high storage values. As suitable w parameter values cannot be found for these catchments, this results in higher storage as it restricts annual AET. Potential reasons why those catchments plot left of the energy limit are (1) overestimation of AET (2) underestimation of PET.

Fitting a w parameter for all catchments in the study by minimising the sum of squared error results in a value of 3.81. This number is greater than the fitted w parameter of 2.84 and 2.55 found by L. Zhang et al. (2004) for forested and grassed Australian catchments, respectively. The higher average value of w in this study suggests a greater amount of ET than in the study by L. Zhang et al. (2004), as expected from the structure of equation (5). The inclusion of the Millennium Drought period is also influential in our study, a period of increased evaporative demand and lower water availability, where the L. Zhang et al. (2004) study period only covered up to the year 2000.

4 Discussion

4.1 Storage and catchment characteristics

Our study catchments tended to have small dynamic storages and relatively large extended dynamic storages. Dynamic storage represents the storage that directly contributes to streamflow and the fact the storages were estimated to be small is a reflection of the study environment, where evapotranspiration dominates catchment losses. The difference between the sizes of dynamic and extended dynamic storage sizes can be interpreted that a large proportion of catchment storage is "reserved" for evapotranspiration (Brooks, Barnard, Coulombe, & McDonnell, 2010). This behaviour has been ob-

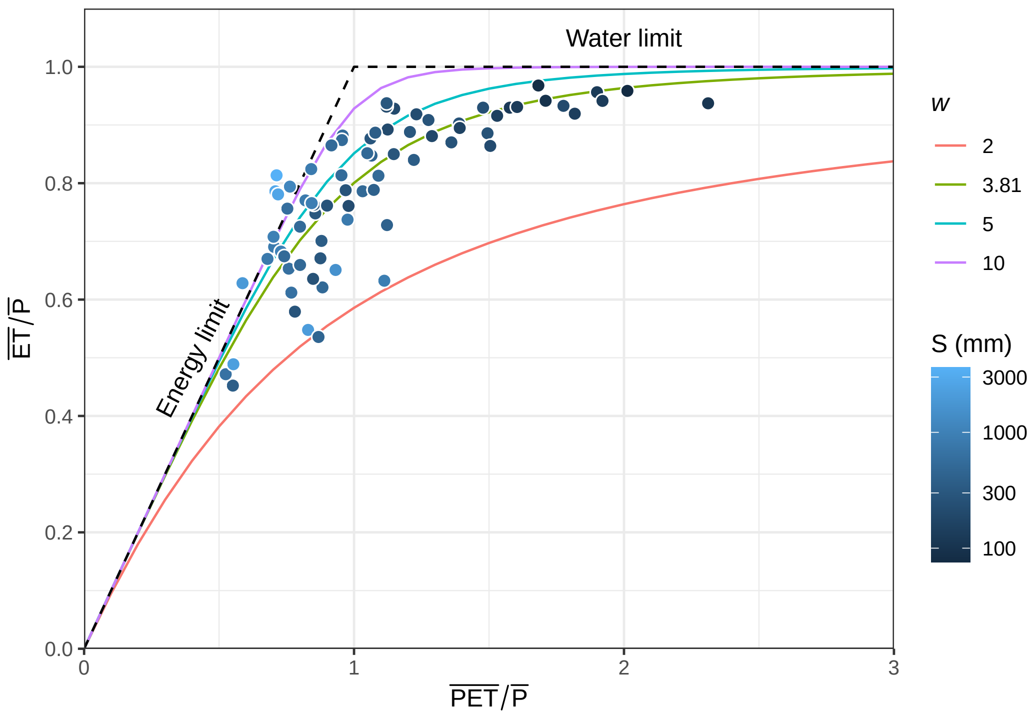

Figure 5. Ratio of the aridity index ($\overline{PET}/\overline{P}$) and evaporative index ($\overline{ET}/\overline{P}$) and Fu-Zhang curves. Each point represents one catchment.

served in *Eucalyptus* forested catchments where transpiration continues at normal rates even during extended dry periods (Talsma & Gardner, 1986). While the dynamic storages are small, the fact that the headwater catchments along the south east of Australia continue to flow in prolonged dry periods and have long travel times (Cartwright et al., 2020) suggests that these stores are deeper in the subsurface and are connected to long groundwater flowpaths (Howcroft et al., 2018).

We used several physical catchment characteristics to assess the controls on catchment storage, as well as to assess if the storage estimation methods possess physical realism. Larger storages were strongly linked with topographic characteristics. Catchments with greater slope and a lower percentage of valley bottoms were significantly correlated to storage across all methods. Higher saturated hydraulic conductivity is also significantly and positively correlated to storage. These characteristics together express a physical system where water can readily drain to subsurface stores. This is in line with other findings that catchment topographic characteristics are pivotal to water storage (Jencso & McGlynn, 2011). Soil storage was found to be important, with soil depth significantly correlated to all the storage estimates. This is also highlighted in the simulated partitioning of water by the HBV model, where soil water represented the greatest store for most catchments.

The BFI and stream AC also significantly correlated to water storage, in line with other studies that have found the BFI captures storage and release properties of catchments (Salinas et al., 2013). A greater BFI relates to higher stream autocorrelation, and for the study catchments there is a Pearson's correlation of 0.76 between the two characteristics. A physical interpretation of this result is that greater autocorrelation, and therefore greater memory in the streamflow signal, suggests a slower storage release and slower flow paths.

The geological characteristics were not found to be a strong indication of storage, with no consistent significant correlations across the storage estimates from the different methods. This may be a result of the coarseness of the parent data (1:1M) and the uncertainty of spatial mapping of geology. Subsurface geology and the geology-soil interface are important to hydrological storage (Jencso & McGlynn, 2011; Sklash & Farvolden, 1979; Sophocleous, 2002), however other evidence has shown that the physical arrangement of these features (e.g. McGuire et al. (2005)) is more important than the simple geological rock constituencies. Staudinger et al. (2017) also did not find a significant relationship between their geological indicator (average quaternary depth) and derived storage. This raises a broader issue of what the ideal geological indicators and measures are when determining broad scale storage controls.

This study did not find a strong spatial pattern in the results. There was no significant relationship between latitude and storage size and longitude was only significant for the storage discharge and Budyko approaches (Table 4), where there is a slight west-east gradient of increasing storage. This finding also strongly suggests that local catchment characteristics and physiography play a large role in water storage potential (Berghuijs, Sivapalan, Woods, & Savenije, 2014).

4.2 Turnover times and buffering capacity

A useful metric that can be calculated once storage is known is the turnover time. The turnover time expresses storage relative to the flow rate (Małoszewski & Zuber, 1982; McGuire & McDonnell, 2006), which is ordinarily the mean annual flow rate. The turnover time also serves as a reference for catchment mean travel time where there is no direct observations of water age (McGuire & McDonnell, 2006). The length of time it takes for water to travel through a catchment is controlled by the catchment geology, soils, vegetation and topography, and a powerful feature of travel times is that they integrate these spatial heterogeneities (Botter, Bertuzzo, & Rinaldo, 2011). While travel time distributions are more informative of catchment hydrological processes, the mean travel time is useful as a broad-scale measure to compare catchments and is often related to catchment characteristics (McDonnell et al., 2010).

The distribution of turnover times for each of the storage methods are presented in Figure 6. Across all methods, turnover times range from 0.07 to 44.5 years. Predictably, the methods that yielded smaller storage estimates resulted in shorter turnover times. Mean transit times (MTTs) estimated using tracers show that flows tend to be from years and decades to greater than a 100 years in south eastern Australia (e.g. (Buzacott, van der Velde, Keitel, & Vervoort, 2020; Cartwright & Morgenstern, 2015, 2016; Cartwright et al., 2020; Duvert, Stewart, Cendon, & Raiber, 2016; Howcroft et al., 2018)). If the dynamic storage is assumed to be the storage that contributes to discharge, using the turnover time is an unsuitable approximation of the MTT given the large disparity between the estimates in this study and the results from tracer studies. However, the results here suggest that the extended dynamic storage may provide a rough approximation of the MTT and that the size of those storages may be realistic.

Given that water in extended dynamic storages can be removed by evapotranspiration and streamflow, an additional measure to consider is the buffering capacity of a catchment. In other words, how long can a catchment sustain the mean behaviour from its maximum storage potential. This is more relevant to the study catchments given the high rates of evapotranspiration. Here we calculate the total catchment turnover time relative to mean annual evapotranspiration and streamflow. The buffering capacity of the study catchments ranges from less than one to approximately three years (Figure 6). This range shows that catchments can withstand drought periods for several years. Recent droughts have exposed the vulnerability of the study catchments, such as the Millennium Drought which spanned a decade (Potter & Chiew, 2011; van Dijk et al., 2013). As future droughts are expected to become more severe, an insufficient buffering is likely to be offered by these catchments and flows downstream will be impacted. Another factor to consider is what happens when catchments are pushed past their buffering capac-

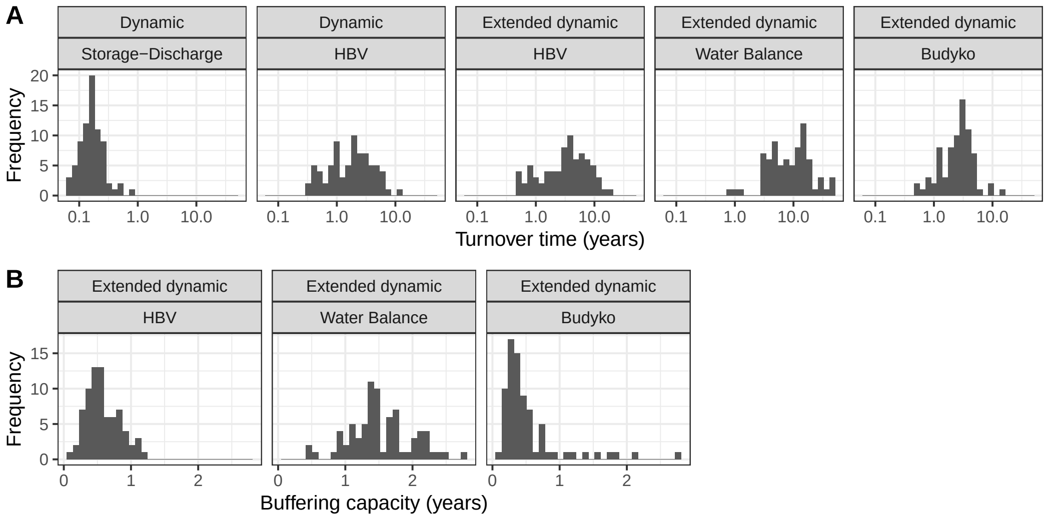

Figure 6. (A) Distributions of turnover times for each method and (B) distributions of buffering capacity for the extended dynamic storage methods.

ities. Significant changes in rainfall-runoff behaviour have been observed from extended droughts (K. J. A. Fowler, Peel, Western, Zhang, & Peterson, 2016; Saft, Western, Zhang, Peel, & Potter, 2015), indicating that storage behaviour is non-linear and storage return might be hysteretic. It is possible that the change in rainfall-runoff behaviour coincides with the exhaustion of the ability to sustain average behaviour.

4.3 Methodology

In this study we adopted the recommendation of McNamara et al. (2011) and adopted a common methods to evaluate storage. Staudinger et al. (2017) refined and clarified the definitions of storage commonly used within hydrology and we adopted their framework to investigate storage in catchments in the south east of Australia. The benefit of this approach is that allows a comparison of methods, a comparison of the study catchments, and comparisons to other studies to be easily made. We also closely adopted Staudinger et al. (2017) methods, including the storage discharge approach, HBV methods and the water balance approach. We added the Budyko method to include another approach to evaluate the water balance that is based on rigorous water-energy balance approach.

The five methods we applied all yielded different results, but like Staudinger et al. (2017) we found that the methods had similar rankings. That is, the methods are consistently estimated relatively smaller or larger storages for the same catchment. All methods were significantly correlated with each other, with the exception of the storage discharge method and the HBV extended dynamic method. Moreover, the multi-method and multi-catchment approach demonstrates the difficulty of quantifying catchment storage. The strong correlations to the physical characteristics show the methods are capturing some aspect of catchment storage behaviour that match conceptual ideas of catchment storage, however the inconsistencies of the correlations to some of the methods creates uncertainty if simple rules about what govern catchment storage can be established. A potential source of this inconsistency is the fact that, despite using the most up to date sources of data that covered the study region, many of the physical characteristics are spatially modelled values derived from other landscape level data.

Each of the methods have their own relative strengths (and weaknesses) and are discussed in subsections below. A general problem that applies to all the methods in this study is that none of the methods are direct observations of storage, rather they have been inferred from catchment fluxes. Without some direct measure of storage there is a reciprocal problem: it is difficult to define storage without defining it from fluxes, when storage itself is defining or controlling those processes.

4.3.1 Storage discharge

The storage discharge method provides a clever way of estimating the storage size by analysing times when streamflow is a function of storage. This behaviour can be observed during low flux hours, i.e. when there is negligible precipitation and evapotranspiration, and the stream is in recession. This proved challenging to implement in this study using daily data. In his original study, Kirchner (2009) provided a limited demonstration of the use of daily data to establish storage discharge relationships. In his example, he minimised the effect of P and ET on the relationship by only selecting days where the daily flux of P and ET was less than 10% of Q. In our study, the use of a long time series allowed days with any rain to be excluded, however excluding days with ET less than 10% of Q resulted in an insufficient amount of data points to yield robust relationships of of Q and g(Q). This is likely to be an issue with this method for warmer environments. An additional complication is that catchments in Australia tend to be larger due to the flatter topography. This typically results in low yields of water and it is rarely the case that streamflow is substantially larger than evapotranspiration.

The effects of P and ET are minimised in this study by removing the day of and after precipitation from the analysis and limiting analysis to cooler months of June to August. However, ET can still be considerable during these months in south eastern Australia and there is almost certainly an effect on the calculated storage sizes. Improperly excluding ET results in storage being underestimated (Kirchner, 2009) and this is a caveat of the results here where the storage discharge method estimated the smallest storages. The use of hourly data is one opportunity to improve the reliability of storage estimates using this method. This comes with other challenges, including (1) long timeseries of hourly data for many catchments are not widely available (2) nocturnal transpiration can still be considerable in the Australian environment (Buckley, Turnbull, Pfautsch, & Adams, 2011).

4.3.2 HBV

As Staudinger et al. (2017) identified, the HBV model can consider different sources of storage and their relative contributions to the total dynamic or extended dynamic storage. These storages are simulated and are not based on any real observations of groundwater, soil water or snow storage. While they are simulated storages, our results show these conceptual stores are significantly correlated to many physical characteristics that are representative of these stores. Model structure and the choice of the objective function are likely to have an impact on the partitioning of water and model performance (Knoben, Freer, Peel, Fowler, & Woods, 2020). This source of uncertainty was not assessed in this study, but could be examined by comparing the results of multiple conceptual models and objective functions to evaluate the consistency of water partitioning and subsequently storage size. Additionally, there is always uncertainty that derives from the chosen initial parameter ranges and model calibration routine (Butts, Payne, Kristensen, & Madsen, 2004). We used parameter ranges that are consistent with the literature (Lidén & Harlin, 2000; Seibert, 1997; Seibert & Vis, 2012). Parameter ranges will have a large effect on the partitioning of water between the different stores. The ranges of calibrated parameters did not indicate that there was limiting behaviour preventing further increases or decreases of the sizes of storages. To reduce some calibration uncertainty, Staudinger et al. (2017) averaged 100 parameter set runs using the Genetic Algorithm and Powell optimisation in their study. We consider the use of SCE-UA to perform as optimally due to its combination of random global and local searches and evolutionary process, and has been shown to be regularly be a robust calibration algorithm (Boyle, Gupta, & Sorooshian, 2000; Kuczera, 1997; Wang, Yu, & Yang, 2010).

4.3.3 Water balance

The water balance approach should theoretically provide the optimal measure of extended-dynamic catchment storage. It can be run at different temporal scales and tries to directly relate changes in storage with fluxes. However, a clear source of uncertainty for the water balance approach is the use of the scaling factor s_{ET}. The use of this scaling factor was necessary as without this factor sensible water balances could not be computed with the data for most catchments. Most catchments had a positive s_{ET}, indicating there are greater catchment losses to ET than what is estimated by the Morton's actual areal evapotranspiration. This raises a few possibilities: poor estimation of actual evapotranspiration, inaccurate spatial estimation of precipitation or inaccurate gauging of streamflow. Small errors in any of these variables accumulate over time and cause the water balance not to close. This raises a broader issue in that we cannot close the water balance from the best datasets we have available. Moreover, despite the ubiquity of the cumulative water balance equation (i.e. $\Delta S = P - Q - ET$) in hydrology, the equation excludes other losses, such as inter-catchment flows which are often (and potentially falsely) assumed to be negligible (Bouaziz et al., 2018; Fan, 2019). This also gives rise to another common assumption, employed here, that long term average AET can be estimated using $\overline{AET} = \overline{P} - \overline{Q}$. This term could actually be considered the mean loss term that excludes Q, as any losses to other sources are attributed to ET.

4.3.4 Budyko

The Budyko approach simplifies the complex processes and interactions and expresses the controls of actual evapotranspiration by the availability of energy and water and has been validated globally (Choudhury, 1999; Koster & Suarez, 1999; L. Zhang, Dawes, & Walker, 2001). We added this method due to the limitations of the water balance approach, where it is suspected poor evapotranspiration estimates may hinder an accurate simulation of the water balance. The advantages here come at the expense of temporal resolution, where the Budyko approach is ordinarily computed annually. This effectively cuts the extremes of the storage estimates, as shown in many of the results where an overall smaller storage was estimated compared to the water balance approach.

Despite the temporal coarseness of this approach, the Budyko method still highlighted the larger inter-annual storage changes in the study catchments.

This approach required the use of $AET = P - Q$ to estimate the w parameter that is subsequently used for the calculation of annual AET. The use of effective precipitation (P'), allows for an annual estimate of ET that is dependent on excess precipitation from past years that has not been consumed, as well as that years precipitation. However, several catchments in the study plotted to the left of the so-called 'energy-limit' (Figure 5). There are two probable causes: that AET is overestimated (causing the points to move up) or that PET is underestimated (causing the points to move to the left). Alternatively, it could be a combination of both factors. Despite the apparent suitability of Morton's estimates of evapotranspiration to calculate the water balance (McMahon et al., 2013), Morton's estimates of evapotranspiration do not factor in effects from wind, which can cause large differences in PET and AET calculations (Donohue, McVicar, & Roderick, 2010). Alternatively, these catchments have other losses of water that overestimates long term actual evapotranspiration from the water balance. A potential improvement to the method applied here is to incorporate dynamic vegetation data into the Budyko formulation to yield more accurate AET estimates (Donohue et al., 2010), and therefore improve estimates of storage.

4.4 Implications and future research

This study builds on the global push to understand water storages in catchments by using common storage definitions (McNamara et al., 2011) and estimation methods (Staudinger et al., 2017). In our study catchments, the multi-method and multi-catchment approach did not tightly constrain the sizes of dynamic or extended dynamic storages. Further research is required to obtain physical estimates of storage to validate the approaches used here. This includes, and is not limited to, using tracers to characterise mobile storage, and using satellite products and groundwater level data to study storage. While remotely sensed data and groundwater level data may not directly reveal storage, they can be indicators of catchment wetness and could be useful to determine varying states of catchment storage.

Many of the results here indicate that groundwater and slow flow processes are important to water storage and release from catchments. Hydrological models poorly simulate these features and are likely a reason why performance outside their calibration windows is lacking. Our results reinforce the call to improve conceptual models to better account for slow flow processes (e.g. K. Fowler et al. (2020)). It is likely an incomplete understanding of the underlying mechanisms can be attained without grasping the mobile storage. Much of the underlying hydrological processes likely occur in the mobile storage domain where there is an important distinction between particle velocities and celerities (Beven & Davies, 2015; McDonnell & Beven, 2014). Mobile storage was not assessed as it cannot be determined from hydrometric data alone. Rather, it is usually inferred with the assistance of tracers. There are several studies that have evaluated MTTs using tracer data within the study region and these could be pooled to evaluate mobile storage. However, the physical controls on MTTs in some of these catchments have not been readily identified (Cartwright et al., 2020; Howcroft et al., 2018) and the estimates of MTTs often carry considerable uncertainty due to the assumptions required to estimate recharge rates (e.g. Li, Jasechko, and Si (2019)). Despite the clear challenges, further work focussing on water age behaviour could lead to breakthroughs in the understanding of the controls on catchment storage.

It is largely unknown how temporally variable storage is. Storage is often assumed to be constant through time, as was assumed in this study to derive storage from the long term water balance. It is possible that dynamic and extended dynamic storages behave non-linearly, as indicated by the research by Saft et al. (2015) and Saft et al. (2016) which shows that drought induces changes to the land system which are likely to influence water storage and release properties. Storage could be evaluated using rolling windows that encompass wet and dry periods to evaluate if there are changes to the size of storage through time and if changes are trending in a particular direction. Beyond naturally induced changes to catchment storage, human interventions can have large impacts on groundwater-surface water exchange (e.g. Yang et al. (2017)) and there is a clear need to understand how these manifest in terms of water storage capacity.

5 Conclusions

Storage sits at the intersection of the main hydrological processes and advances in the understanding of catchment storage will provide greater insight into catchment functioning. While in hydrology the focus is often on the fluxes, flux behaviour can be more precisely quantified within hydrological boundary conditions if that boundary can be established. We adopted the multi-method and multi-catchment, which have been proposed as a clear way to advance the case of storage (e.g. McNamara et al., 2011, and Staudinger et al. (2017)), and evaluated the results against key catchment characteristics to evaluate the controls on storage size. The results of this study highlight the challenge of investigating catchment storage and the ongoing need to further refine the approach. Future research directions in the study area should consider evaluating mobile storage, investigating the potentially transient nature of upper storage. With impending challenges such as climate change and large scale land use change, it is critical to understand the role storage plays in catchments from a water resource and management perspective. This is particularly the case for the study region, which already exhibits severe interannual hydrological variability relative to the world (Peel, McMahon, & Finlayson, 2004).

In relation to our original aims, (1) we successfully estimated the storages across our study area. While the different methods were generally ranked consistently, the estimates of dynamic and extended dynamic storage could vary substantially depending on the catchment. (2) It was difficult to determine robust catchment characteristics that control storage across the methods, but several key characteristics highlighted the nature of the storage. To that end, this supports the idea that storage is a useful metric for catchment comparison (McNamara et al., 2011). Our results indicate that slow flow processes are important sources of catchment storage for streamflow and that catchment physical arrangement, rather than purely spatial location, proved to be better indicators of storage. The geological characteristics used in the study did not strongly relate to the storage estimations and further work is required to identify useful geological measures that relate to storage. (3) We calculated the turnover and buffering capacities of the catchments. The turnover times are comparable to the mean transit times of regional studies. The buffering capacities indicate that while the study catchments have some resistance to drought, they are vulnerable to harsher droughts that are anticipated with future climate projections.

Acknowledgments

Alexander Buzacott acknowledges the support of the Research Training Program scholarship provided by the Australian Government. This study was conducted with the support of the Australian Research Council (LP130101183). The authors acknowledge the Sydney Informatics Hub and The University of Sydney's high performance computing cluster Artemis for providing the high performance computing resources that have contributed to the research results reported within this paper. Rainfall data and data used to calculate evapotranspiration are from the Australian Bureau of Meteorology's (BOM) Australian Water Availability Project (Jones et al., 2009) www.bom.gov.au/jsp/awap/ and were obtained via the AWAPer R package (Peterson et al. 2020). Streamflow data were from the BOM Hydrologic Reference Station project website (`http://www.bom.gov.au/water/hrs/`) and BOM Geofabric products were retrieved from `http://www.bom.gov.au/water/geofabric/` (Stein et al., 2014). The accompanying Geofabric National Environmental Stream Attributes product was downloaded from `http://pid.geoscience.gov.au/dataset/ga/73045`. The Soil and Landscape Grid of Australia (Grundy et al., 2015) can be retrieved from `https://www.clw.csiro.au/aclep/soilandlandscapegrid/index.html`. The Australian woody vegetation cover product was retrieved from `http://auscover.org.au/purl/landsat-persistent-green-2000-2010`. The analysis code for this study is available on GitHub (`https://github.com/buzacott/StorageSEAus`).

References

Ajami, H., Troch, P. A., Maddock, T., Meixner, T., & Eastoe, C. (2011). Quantifying mountain block recharge by means of catchment-scale storage-discharge relationships. *Water Resources Research*, *47*(4). doi: 10.1029/2010WR009598

Andrews, F. T., Croke, B. F. W., & Jakeman, A. J. (2011). An open software environment for hydrological model assessment and development. *Environmental Modelling & Software*, *26*(10), 1171–1185. doi: 10.1016/j.envsoft.2011.04.006

Berghuijs, W. R., Sivapalan, M., Woods, R. A., & Savenije, H. H. G. (2014). Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales. *Water Resources Research*, *50*(7), 5638–5661. doi: 10.1002/2014WR015692

Best, D. J., & Roberts, D. E. (1975). Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *24*(3), 377–379. doi: 10.2307/2347111

Beven, K., & Davies, J. (2015). Velocities, celerities and the basin of attraction in catchment response. *Hydrological Processes*, *29*(25), 5214–5226. doi: 10.1002/hyp.10699

Birkel, C., Soulsby, C., & Tetzlaff, D. (2011). Modelling catchment-scale water storage dynamics: Reconciling dynamic storage with tracer-inferred passive storage. *Hydrological Processes*, *25*(25), 3924–3936. doi: 10.1002/hyp.8201

Botter, G., Bertuzzo, E., & Rinaldo, A. (2011). Catchment Residence and Travel Time Distributions: The Master Equation. *Geophysical Research Letters*, *38*(11), L11403. doi: 10.1029/2011GL047666

Bouaziz, L., Weerts, A., Schellekens, J., Sprokkereef, E., Stam, J., Savenije, H., & Hrachowitz, M. (2018). Redressing the balance: Quantifying net intercatchment groundwater flows. *Hydrology and Earth System Sciences*, *22*(12), 6415–6434. doi: 10.5194/hess-22-6415-2018

Boyle, D. P., Gupta, H. V., & Sorooshian, S. (2000). Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*, *36*(12), 3663–3674. doi: 10.1029/2000WR900207

Brooks, J. R., Barnard, H. R., Coulombe, R., & McDonnell, J. J. (2010). Ecohydrologic separation of water between trees and streams in a Mediterranean climate. *Nature Geoscience*, *3*(2), 100–104. doi: 10.1038/ngeo722

Buckley, T. N., Turnbull, T. L., Pfautsch, S., & Adams, M. A. (2011). Nocturnal water loss in mature subalpine Eucalyptus delegatensis tall open forests and adjacent E. pauciflora woodlands. *Ecology and Evolution*, *1*(3), 435–450. doi: 10.1002/ece3.44

Budyko, M. (1974). *Climate and Life* (D. H. Miller, Ed.). New York: Academic Press.

Buttle, J. M. (2016). Dynamic storage: A potential metric of inter-basin differences in storage properties. *Hydrological Processes*, *30*(24), 4644–4653. doi: 10.1002/hyp.10931

Butts, M. B., Payne, J. T., Kristensen, M., & Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, *298*(1), 242–266. doi: 10.1016/j.jhydrol.2004.03.042

Buzacott, A. J. V., van der Velde, Y., Keitel, C., & Vervoort, R. W. (2020). Constraining water age dynamics in a south-eastern Australian catchment using an age-ranked storage and stable isotope approach. *Hydrological Processes*, *34*(23), 4384–4403. doi: 10.1002/hyp.13880

Cartwright, I., & Morgenstern, U. (2015). Transit times from rainfall to baseflow in headwater catchments estimated using tritium: The Ovens River, Australia. *Hydrology and Earth System Sciences*, *19*(9), 3771–3785. doi: 10.5194/hess-19-3771-2015

Cartwright, I., & Morgenstern, U. (2016). Using tritium to document the mean transit time and sources of water contributing to a chain-of-ponds river system: Implications for resource protection. *Applied Geochemistry*, *75*, 9–19. doi: 10.1016/j.apgeochem.2016.10.007

Cartwright, I., Morgenstern, U., Howcroft, W., Hofmann, H., Armit, R., Stewart, M., … Irvine, D. (2020). The variation and controls of mean transit times in Australian headwater catchments. *Hydrological Processes*, *34*(21), 4034–4048. doi: 10.1002/hyp.13862

Choudhury, B. (1999). Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model. *Journal of Hydrology*, *216*(1), 99–110. doi: 10.1016/S0022-1694(98)00293-5

Condon, L. E., Markovich, K. H., Kelleher, C. A., McDonnell, J. J., Ferguson, G., & McIntosh, J. C. (2020). Where Is the Bottom of a Watershed? *Water Resources Research*, *56*(3), e2019WR026010. doi: 10.1029/2019WR026010

Dawson, T. E. (1996). Determining water use by trees and forests from isotopic, energy balance and transpiration analyses: The roles of tree size and hydraulic lift. *Tree Physiology*, *16*(1-2), 263–272. doi: 10.1093/treephys/16.1-2.263

Donohue, R. J., McVicar, T. R., & Roderick, M. L. (2010). Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. *Journal of Hydrology*, *386*(1), 186–197. doi: 10.1016/j.jhydrol.2010.03.020

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, *28*(4), 1015–1031. doi: 10.1029/91WR02985

Duvert, C., Stewart, M. K., Cendon, D. I., & Raiber, M. (2016). Time series of tritium, stable isotopes and chloride reveal short-term variations in groundwater contribution to a stream. *Hydrology and Earth System Sciences*, *20*(1), 257–277. doi: 10.5194/hess-20-257-2016

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., &

Savenije, H. H. G. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences*, *17*(5), 1893–1912. doi: 10.5194/hess-17-1893-2013

Fan, Y. (2019). Are catchments leaky? *WIREs Water*, *6*(6), e1386. doi: 10.1002/wat2.1386

Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., … Western, A. (2020). Many Commonly Used Rainfall-Runoff Models Lack Long, Slow Dynamics: Implications for Runoff Projections. *Water Resources Research*, *56*(5), e2019WR025286. doi: 10.1029/2019WR025286

Fowler, K. J. A., Peel, M. C., Western, A. W., Zhang, L., & Peterson, T. J. (2016). Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resources Research*, *52*(3), 1820–1846. doi: 10.1002/2015WR018068

Freeman, M. C., Pringle, C. M., & Jackson, C. R. (2007). Hydrologic Connectivity and the Contribution of Stream Headwaters to Ecological Integrity at Regional Scales. *JAWRA Journal of the American Water Resources Association*, *43*(1), 5–14. doi: 10.1111/j.1752-1688.2007.00002.x

Fu, B.-p. (1981). On the calculation of the evaporation from land surface. *Chinese Journal of Atmospheric Sciences*, *5*(1), 23–31.

Geris, J., Tetzlaff, D., & Soulsby, C. (2015). Resistance and resilience to droughts: Hydropedological controls on catchment storage and run-off response. *Hydrological Processes*, *29*(21), 4579–4593. doi: 10.1002/hyp.10480

Gill, T., Johansen, K., Phinn, S., Trevithick, R., Scarth, P., & Armston, J. (2017). A method for mapping Australian woody vegetation cover by linking continental-scale field data and long-term Landsat time series. *International Journal of Remote Sensing*, *38*(3), 679–705. doi: 10.1080/01431161.2016.1266112

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., & Cardenas, M. B. (2016). The global volume and distribution of modern groundwater. *Nature Geoscience*, *9*(2), 161–167. doi: 10.1038/ngeo2590

Gray, J. M., Bishop, T. F. A., & Wilford, J. R. (2016). Lithology and soil relationships for soil modelling and mapping. *Catena*, *147*, 429–440. doi: 10.1016/j.catena.2016.07.045

Gray, J. M., Bishop, T. F. A., & Yang, X. (2015, March). Pragmatic models for the prediction and digital mapping of soil properties in eastern Australia. *Soil Research*, *53*(1), 24–42. doi: 10.1071/SR13306

Gribovszki, Z., Kalicz, P., Szilágyi, J., & Kucsara, M. (2008). Riparian zone evapotranspiration estimation from diurnal groundwater level fluctuations. *Journal of Hydrology*, *349*(1), 6–17. doi: 10.1016/j.jhydrol.2007.10.049

Grundy, M. J., Rossel, R. A. V., Searle, R. D., Wilson, P. L., Chen, C., & Gregory, L. J. (2015). Soil and Landscape Grid of Australia. *Soil Research*, *53*(8), 835–844. doi: 10.1071/SR15191

Guo, D., Westra, S., & Maier, H. R. (2016). An R package for modelling actual, potential and reference evapotranspiration. *Environmental Modelling & Software*, *78*, 216–224. doi: 10.1016/j.envsoft.2015.12.019

Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. In W. E. Sopper & H. W. Lull (Eds.), *Forest Hydrology* (pp. 275–291). New York: Pergamon Press.

Hornberger, G. M., Scanlon, T. M., & Raffensperger, J. P. (2001). Modelling transport of dissolved silica in a forested headwater catchment: The effect of hydrological and chemical time scales on hysteresis in the concentration–discharge relationship. *Hydrological Processes*, *15*(10), 2029–2038. doi: 10.1002/hyp.254

Howcroft, W., Cartwright, I., & Morgenstern, U. (2018). Mean transit times in headwater catchments: Insights from the Otway Ranges, Australia. *Hydrol. Earth Syst. Sci.*, *22*(1), 635–653. doi: 10.5194/hess-22-635-2018

Hrachowitz, M., Benettin, P., van Breukelen, B. M., Fovet, O., Howden, N. J., Ruiz, L., ... Wade, A. J. (2016). Transit times—the link between hydrology and water quality at the catchment scale. *Wiley Interdisciplinary Reviews: Water*, *3*(5), 629–657. doi: 10.1002/wat2.1155

Hrachowitz, M., Fovet, O., Ruiz, L., & Savenije, H. H. G. (2015). Transit time distributions, legacy contamination and variability in biogeochemical 1/fα scaling: How are hydrological response dynamics linked to water quality at the catchment scale? *Hydrological Processes*, *29*(25), 5241–5256. doi: 10.1002/hyp.10546

Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., ... Baillie, J. E. M. (2020). Importance and vulnerability of the world's water towers. *Nature*, *577*(7790), 364–369. doi: 10.1038/s41586-019-1822-y

Jackson, R. B., Sperry, J. S., & Dawson, T. E. (2000). Root water uptake and transport: Using physiological processes in global predictions. *Trends in Plant Science*, *5*(11), 482–488. doi: 10.1016/S1360-1385(00)01766-0

Jasechko, S., Kirchner, J. W., Welker, J. M., & McDonnell, J. J. (2016). Substantial proportion of global streamflow less than three months old. *Nature Geoscience*, *9*(2), 126–129. doi: 10.1038/ngeo2636

Jencso, K. G., & McGlynn, B. L. (2011). Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, *47*(11). doi: 10.1029/2011WR010666

Johnston, R. M., Barry, S. J., Bleys, E., Bui, E. N., Moran, C. J., Simon, D. a. P., ... Grundy, M. (2003). ASRIS: The database. *Soil Research*, *41*(6), 1021–1036. doi: 10.1071/sr02033

Jones, D., Wang, W., & Fawcett, R. (2009). High-Quality Spatial Climate Data-Sets for Australia. *Australian Meteorological and Oceanographic Journal*, *58*(4), 233–248.

Kirchner, J. W. (2003). A double paradox in catchment hydrology and geochemistry. *Hydrological Processes*, *17*(4), 871–874. doi: 10.1002/hyp.5108

Kirchner, J. W. (2006). Getting the Right Answers for the Right Reasons: Linking Measurements, Analyses, and Models to Advance the Science of Hydrology. *Water Resources Research*, *42*(3), W03S04. doi: 10.1029/2005WR004362

Kirchner, J. W. (2009). Catchments as Simple Dynamical Systems: Catchment Characterization, Rainfall-Runoff Modeling, and Doing Hydrology Backward. *Water Resources Research*, *45*(2), 1–34. doi: 10.1029/2008WR006912

Kirchner, J. W., & Neal, C. (2013). Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proceedings of the National Academy of Sciences*, *110*(30), 12213–12218. doi: 10.1073/pnas.1304328110

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., & Woods, R. A. (2020). A Brief Analysis of Conceptual Model Structure Uncertainty Using 36 Models and 559 Catchments. *Water Resources Research*, *56*(9), e2019WR025975. doi: 10.1029/2019WR025975

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. doi: 10.5194/hess-23-4323-2019

Koffler, D., & Laaha, G. (2013). LFSTAT - Low-Flow Analysis in R. In *Geophysical Research Abstracts 15, EGU General Assembly.* Vienna, Austria.

Koster, R. D., & Suarez, M. J. (1999). A Simple Framework for Examining the Interannual Variability of Land Surface Moisture Fluxes. *Journal of Climate*, *12*(7), 1911–1917. doi: 10.1175/1520-0442(1999)012⟨1911:ASFFET⟩2.0.CO;2

Kuczera, G. (1997). Efficient subspace probabilistic parameter optimization for catchment models. *Water Resources Research*, *33*(1), 177–185. doi: 10.1029/96WR02671

Li, Z., Jasechko, S., & Si, B. (2019). Uncertainties in tritium mass balance models for groundwater recharge estimation. *Journal of Hydrology*, *571*, 150–158. doi: 10.1016/j.jhydrol.2019.01.030

Lidén, R., & Harlin, J. (2000). Analysis of conceptual rainfall–runoff modelling performance in different climates. *Journal of Hydrology*, *238*(3), 231–247. doi: 10.1016/S0022-1694(00)00330-9

Lindström, G. (1997). A Simple Automatic Calibration Routine for the HBV Model. *Hydrology Research*, *28*(3), 153–168. doi: 10.2166/nh.1997.0009

Małoszewski, P., & Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. *Journal of Hydrology*, *57*(3), 207–231. doi: 10.1016/0022-1694(82)90147-0

McDonnell, J. J., & Beven, K. (2014). Debates-The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. *Water Resources Research*, *50*(6), 5342–5350. doi: 10.1002/2013WR015141

McDonnell, J. J., McGuire, K., Aggarwal, P., Beven, K. J., Biondi, D., Destouni, G., … Wrede, S. (2010). How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis. , *24*(12), 1745–1754. doi: 10.1002/hyp.7796

McGuire, K. J., & McDonnell, J. J. (2006). A Review and Evaluation of Catchment Transit Time Modeling. *Journal of Hydrology*, *330*(3–4), 543–563. doi: 10.1016/j.jhydrol.2006.04.020

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., & Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research*, *41*(5). doi: 10.1029/2004WR003657

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., & McVicar, T. R. (2013). Estimating Actual, Potential, Reference Crop and Pan Evaporation Using Standard Meteorological Data: A Pragmatic Synthesis. *Hydrol. Earth Syst. Sci.*, *17*(4), 1331–1363. doi: 10.5194/hess-17-1331-2013

McNamara, J. P., Tetzlaff, D., Bishop, K., Soulsby, C., Seyfried, M., Peters, N. E., ... Hooper, R. (2011). Storage as a Metric of Catchment Comparison. *Hydrological Processes*, *25*(21), 3364–3371. doi: 10.1002/hyp.8113

Morton, F. I. (1983). Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, *66*(1), 1–76. doi: 10.1016/0022-1694(83)90177-4

Mutzner, R., Weijs, S. V., Tarolli, P., Calaf, M., Oldroyd, H. J., & Parlange, M. B. (2015). Controls on the diurnal streamflow cycles in two subbasins of an alpine headwater catchment. *Water Resources Research*, *51*(5), 3403–3418. doi: 10.1002/2014WR016581

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. doi: 10.1016/0022-1694(70)90255-6

Peel, M. C., McMahon, T. A., & Finlayson, B. L. (2004). Continental differences in the variability of annual runoff-update and reassessment. *Journal of Hydrology*, *295*(1), 185–197. doi: 10.1016/j.jhydrol.2004.03.004

Peters, N. E., & Aulenbach, B. T. (2011). Water storage at the Panola Mountain Research Watershed, Georgia, USA. *Hydrological Processes*, *25*(25), 3878–3889. doi: 10.1002/hyp.8334

Peterson, T. J., Wasko, C., Saft, M., & Peel, M. C. (2020). AWAPer: An R package for area weighted catchment daily meteorological data anywhere within Australia. *Hydrological Processes*, *34*(5), 1301–1306. doi: 10.1002/hyp.13637

Potter, N. J., & Chiew, F. H. S. (2011). An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray-Darling Basin using rainfall-runoff models. *Water Resources Research*, *47*(12), W00G10. doi: 10.1029/2010WR010333

Ramillien, G., Famiglietti, J. S., & Wahr, J. (2008). Detection of Continental Hydrology and Glaciology Signals from GRACE: A Review. *Surveys in Geophysics*, *29*(4), 361–374. doi: 10.1007/s10712-008-9048-9

Rinaldo, A., Benettin, P., Harman, C. J., Hrachowitz, M., McGuire, K. J., van der Velde, Y., ... Botter, G. (2015). Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes. *Water Resources Research*, *51*(6), 4840–4847. doi: 10.1002/2015WR017273

Saft, M., Peel, M. C., Western, A. W., & Zhang, L. (2016). Predicting shifts in rainfall-runoff partitioning during multiyear drought: Roles of dry period and catchment characteristics. *Water Resources Research*, *52*(12), 9290–9305. doi: 10.1002/2016WR019525

Saft, M., Western, A. W., Zhang, L., Peel, M. C., & Potter, N. J. (2015). The Influence of Multiyear Drought on the Annual Rainfall-Runoff Relationship: An Australian Perspective. *Water Resources Research*, *51*(4), 2444–2463. doi: 10.1002/2014WR015348

Salinas, J. L., Laaha, G., Rogger, M., Parajka, J., Viglione, A., Sivapalan, M., & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins – Part 2: Flood and low flow studies. *Hydrology and Earth System Sciences*, *17*(7), 2637–2652. doi: 10.5194/hess-17-2637-2013

Sayama, T., McDonnell, J. J., Dhakal, A., & Sullivan, K. (2011). How much water can a watershed store? *Hydrological Processes*, *25*(25), 3899–3908. doi: 10.1002/hyp.8288

Seibert, J. (1997). Estimation of Parameter Uncertainty in the HBV Model. *Hydrology Research*, *28*(4-5), 247–262.

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*(9), 3315–3325. doi: 10.5194/hess-16-3315-2012

Seyfried, M. S., Grant, L. E., Marks, D., Winstral, A., & McNamara, J. (2009). Simulated soil water storage effects on streamflow generation in a mountainous snowmelt environment, Idaho, USA. *Hydrological Processes*, *23*(6), 858–873. doi: 10.1002/hyp.7211

Sklash, M. G., & Farvolden, R. N. (1979). The role of groundwater in storm runoff. *Journal of Hydrology*, *43*(1), 45–65. doi: 10.1016/0022-1694(79)90164-1

Sophocleous, M. (2002). Interactions between groundwater and surface water: The state of the science. *Hydrogeology Journal*, *10*(1), 52–67. doi: 10.1007/s10040-001-0170-8

Soulsby, C., Neal, C., Laudon, H., Burns, D. A., Merot, P., Bonell, M., … Tetzlaff, D. (2008). Catchment data for process conceptualization: Simply not enough? *Hydrological Processes*, *22*(12), 2057–2061. doi: 10.1002/hyp.7068

Soulsby, C., Tetzlaff, D., & Hrachowitz, M. (2009). Tracers and transit times: Windows for viewing catchment scale storage? *Hydrological Processes*, *23*(24), 3503–3507. doi: 10.1002/hyp.7501

Spence, C. (2007). On the relation between dynamic storage and runoff: A discussion on thresholds, efficiency, and function. *Water Resources Research*, *43*(12). doi: 10.1029/2006WR005645

Spence, C. (2010). A Paradigm Shift in Hydrology: Storage Thresholds Across Scales Influence Catchment Runoff Generation. *Geography Compass*, *4*(7), 819–833. doi: 10.1111/j.1749-8198.2010.00341.x

Staudinger, M., Stoelzle, M., Seeger, S., Seibert, J., Weiler, M., & Stahl, K. (2017). Catchment water storage variation with elevation. *Hydrological Processes*, *31*(11), 2000–2015. doi: 10.1002/hyp.11158

Stein, J. L., Hutchinson, M. F., & Stein, J. A. (2014). A new stream and nested catchment framework for Australia. *Hydrology and Earth System Sciences*, *18*(5), 1917–1933. doi: 10.5194/hess-18-1917-2014

Talsma, T., & Gardner, E. A. (1986). Soil water extraction by a mixed eucalypt forest during a drought period. *Soil Research*, *24*(1), 25–32. doi: 10.1071/sr9860025

Tetzlaff, D., McNamara, J. P., & Carey, S. K. (2011). Measurements and modelling of storage dynamics across scales. *Hydrological Processes*, *25*(25), 3831–3835. doi: 10.1002/hyp.8396

Teuling, A. J., Lehner, I., Kirchner, J. W., & Seneviratne, S. I. (2010). Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment. *Water Resources Research*, *46*(10). doi: 10.1029/2009WR008777

van der Velde, Y., Torfs, P. J. J. F., van der Zee, S. E. A. T. M., & Uijlenhoet, R. (2012). Quantifying catchment-scale mixing and its effect on time-varying travel time distributions. *Water Resources Research*, *48*(6), 1–13. doi:

10.1029/2011WR011310
van der Velde, Y., Vercauteren, N., Jaramillo, F., Dekker, S. C., Destouni, G., & Lyon, S. W. (2014). Exploring hydroclimatic change disparity via the Budyko framework. *Hydrological Processes*, *28*(13), 4110–4118. doi: 10.1002/hyp.9949

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., … Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. doi: 10.1002/wrcr.20123

Van Loon, A. F., & Laaha, G. (2015). Hydrological drought severity explained by climate and catchment characteristics. *Journal of Hydrology*, *526*, 3–14. doi: 10.1016/j.jhydrol.2014.10.059

Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., … Woods, R. (2011). Climate change and mountain water resources: Overview and recommendations for research, management and policy. *Hydrology and Earth System Sciences*, *15*(2), 471–504. doi: 10.5194/hess-15-471-2011

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*(7). doi: 10.1029/2006WR005653

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment Classification and Hydrologic Similarity. *Geography Compass*, *1*(4), 901–931. doi: 10.1111/j.1749-8198.2007.00039.x

Wang, Y.-C., Yu, P.-S., & Yang, T.-C. (2010). Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. *Hydrological Processes*, *24*(8), 1015–1026. doi: 10.1002/hyp.7543

Wheeler, S. A. (2014). Insights, lessons and benefits from improved regional water security and integration in Australia. *Water Resources and Economics*, *8*, 57–78. doi: 10.1016/j.wre.2014.05.006

Winsemius, H. C., Schaefli, B., Montanari, A., & Savenije, H. H. G. (2009). On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Research*, *45*(12). doi: 10.1029/2009WR007706

Yang, Z., Zhou, Y., Wenninger, J., Uhlenbrook, S., Wang, X., & Wan, L. (2017). Groundwater and surface-water interactions and impacts of human activities in the Hailiutu catchment, northwest China. *Hydrogeology Journal*, *25*(5), 1341–1355. doi: 10.1007/s10040-017-1541-0

Yeh, H.-F., & Huang, C.-C. (2019). Evaluation of basin storage–discharge sensitivity in Taiwan using low-flow recession analysis. *Hydrological Processes*, *33*(10), 1434–1447. doi: 10.1002/hyp.13411

Zeng, R., & Cai, X. (2015). Assessing the temporal variance of evapotranspiration considering climate and catchment storage factors. *Advances in Water Resources*, *79*, 51–60. doi: 10.1016/j.advwatres.2015.02.008

Zhang, J. Y., Wang, W. C., & Wei, J. F. (2008). Assessing Land-Atmosphere Coupling Using Soil Moisture from the Global Land Data Assimilation System and Observational Precipitation. *Journal of Geophysical Research-Atmospheres*, *113*(D17).

Zhang, L., Dawes, W. R., & Walker, G. R. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resources Research*, *37*(3), 701–708. doi: 10.1029/2000WR900325

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., & Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, *40*(2). doi: 10.1029/2003WR002710

Zhang, X. S., Amirthanathan, G. E., Bari, M. A., Laugesen, R. M., Shin, D., Kent, D. M., ... Tuteja, N. K. (2016). How streamflow has changed across Australia since the 1950s: Evidence from the network of hydrologic reference stations. *Hydrology and Earth System Sciences*, *20*(9), 3947–3965. doi: 10.5194/hess-20-3947-2016