University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Bonotto, G;Peterson, TJ;Fowler, K;Western, AW

Title:
Identifying Causal Interactions Between Groundwater and Streamflow Using Convergent Cross-Mapping

Date:
2022-08-01

Citation:
Bonotto, G., Peterson, T. J., Fowler, K. & Western, A. W. (2022). Identifying Causal Interactions Between Groundwater and Streamflow Using Convergent Cross-Mapping. Water Resources Research, 58 (8), https://doi.org/10.1029/2021WR030231.

Persistent Link:
https://hdl.handle.net/11343/322061

License:
CC BY

Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR030231

Identifying Causal Interactions Between Groundwater and Streamflow Using Convergent Cross-Mapping

Giancarlo Bonotto[1], Tim J. Peterson[1,2], Keirnan Fowler[1], and Andrew W. Western[1]

[1]Department of Infrastructure Engineering, University of Melbourne, Parkville, VIC, Australia, [2]Department of Civil Engineering, Monash University, Clayton, VIC, Australia

Key Points:

- Convergent cross-mapping (CCM) is a nonparametric method to identify causality in nonlinear dynamic systems that is widely used in ecology
- CCM results suggest heterogeneous causal links that are largely consistent with perceptual understanding of processes
- CCM seems widely applicable in hydrology, but also has limitations and thus should be used and interpreted with care

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

G. Bonotto,
g.bonotto@unimelb.edu.au

Citation:

Bonotto, G., Peterson, T. J., Fowler, K., & Western, A. W. (2022). Identifying causal interactions between groundwater and streamflow using convergent cross-mapping. *Water Resources Research*, *58*, e2021WR030231. https://doi.org/10.1029/2021WR030231

Received 23 APR 2021
Accepted 10 AUG 2022

© 2022. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract Groundwater (GW) is commonly conceptualized as causally linked to streamflow (SF). However, confirming where and how it occurs is challenging given the expense of experimental field monitoring. Therefore, hydrological modeling and water management often rely on expert knowledge to draw causality between SF and GW. This paper investigates the potential of convergent cross-mapping (CCM) to identify causal interactions between SF and GW head. Widely used in ecology, CCM is a nonparametric method to identify causality in nonlinear dynamic systems. To apply CCM between variables the only required inputs are time-series data (stream gauge and bore), so it may be an attractive alternative or complement to expensive field-based studies of causality. Three upland catchments across different hydrogeologic settings and climatic conditions in Victoria, Australia, are adopted as case studies. The outputs of the method seem to largely agree with a priori perceptual understanding of the study areas and offered additional insights about hydrological processes. For instance, it suggested weaker SF-GW interactions during and after the Millennium Drought than in the previous wet periods. However, we show that CCM limitations around seasonality, data sampling frequency, and long-term trends could impact the variability and significance of causal links. Hence, care must be taken while physically interpreting the causal links suggested by CCM. Overall, this study shows that CCM can provide valuable causal information from common hydrological time-series, which is relevant to a wide range of applications, but it should be used and interpreted with care and future research is needed.

Plain Language Summary It is common in hydrology to assume that the water below the ground (in aquifers) interacts with the rivers. However, it is challenging to identify where and how it occurs with the available data and financial resources. Therefore, we often use computer models that represent these interactions according to the knowledge of experts. This paper investigates the potential of convergent cross-mapping (CCM), a method that is widely used in ecology, to identify these interactions. This method only requires time-series data (e.g., streamflow and groundwater level data) and represents an attractive alternative or complement to expensive field-based studies. The results of the analysis appear to agree with an initial understanding of the aquifer-river interactions in three study areas in Victoria, Australia, and offered additional insights. For instance, it suggests weaker interactions during and after the Millennium Drought than in the previous wet periods. However, variability and significance of the causal links inferred by CCM can be due to some limitations of the method. Hence, care must be taken while physically interpreting the results. Overall, CCM can unlock valuable information from common data, which is relevant to hydrology, but it should be used and interpreted with care and future research is needed.

1. Introduction

Due to multiple interacting factors, global water scarcity is projected to increase in the coming decades. As projected by the United Nations, the world's population will reach 9.8 billion by 2050 (U.N., 2017), which combined with income growth of Asian (led by China and India) and African countries, will lead to a 50% increase in demand for agricultural output mostly in water-stressed regions and developing countries (FAO, 2017). Moreover, as climate change is expected to increase variability in rainfall and temperatures (IPCC, 2014; Sheffield et al., 2012; Wasko & Nathan, 2019), reliable water sources such as groundwater (GW) are likely to represent an increasing share of global water consumption. Therefore, understanding the role of GW usage on the availability of both subsurface and surface water is of paramount importance for global food, water, and environmental security.

Although it is common to conceptualize GW as causally linked to streamflow (SF), it remains challenging to identify where and how this interaction occurs (e.g., strength and direction). Interactions between GW and

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

surface water are dictated by multiple processes that are often difficult to measure, insufficiently monitored, and inaccurately conceptualized (Barthel, 2014; Fleckenstein et al., 2010; Krause et al., 2011; Peterson & Western, 2018; Ross, 2012; Staudinger et al., 2019). For instance, hydrologic models often assume unidirectional gaining streams and the importance of the hydrogeologic setting is often neglected by considering horizontal homogeneous isotropic GW formations (Barthel & Banzhaf, 2016; Rinderer et al., 2018). Also, powerful experimental methods, like water balance with geochemical signatures (Atkinson et al., 2015; Cartwright & Gilfedder, 2015; Cartwright & Morgenstern, 2015; Cartwright et al., 2014; Gilfedder et al., 2013; Hartmann et al., 2013; Rimmer & Hartmann, 2014) have the ability to constrain the spatial heterogeneity of GW response and sources of baseflow (Ali et al., 2015; Seibert et al., 2003, 2011), but they have their own inherent uncertainty (Atkinson et al., 2015; Barthel, 2014; McDonnell, 2014) and are sufficiently costly to often limit their application to small study areas and short periods of time (Barthel & Banzhaf, 2016). This is the case particularly in the developing world, which is under-represented in hydrological studies (T. P. Burt & J. J. McDonnell, 2015) and where financial resources are scarce and increasing water demands are managed with limited hydrological data (Davtalab et al., 2017; FAO, 2017; Hulsman et al., 2021).

Therefore, hydrological models, which provide an indirect way to identify dominant hydrological processes, arguably remain the most practical and prevalent method to understand process interaction and estimate SF-GW interactions at the catchment scale. Yet, the original development of a model often relies on limited data and the perceptual understanding of the developers to draw causal links and define the model structure that is usually static, therefore, representing a single subset of working hypothesis that might not be appropriate across hydrologically distinct catchments (Fenicia et al., 2011; Kavetski & Fenicia, 2011; Staudinger et al., 2019). In fact, the choice of model structure can represent a large source of uncertainty that hinders the ability of hydrological models to characterize the hydrological processes in the catchment of interest.

In the context of hydrologic models, distinguishing model structures from rainfall-runoff (RR) response is challenging. For instance, Knoben et al. (2020) evaluated 36 conceptual RR models having different structures (1–15 parameters) on daily observed SF across 559 catchments in the United States of America and have shown that in most catchments up to 28 different models were very close to the best performing model. This illustrates the difficulties in using RR models to identify hydrological processes, as these 28 models are effectively multiple working hypotheses (Clark et al., 2011; Fenicia et al., 2011) with different conceptualizations of processes within the catchments. For example, from this study it would be difficult to infer which catchments have significant SF-GW interactions, as only 10 models had a GW store. Moreover, even if all models had a representation of GW processes, the different structures could result in internal fluxes and storage behavior that lack realism when compared with independent complementary data. Therefore, the nonidentifiability of model structure contributes significantly to the challenge of having models "working for the right reasons" (Kirchner, 2006), despite the advances represented by frameworks like Flux Mapping (Khatami et al., 2019), SUPERFLEX (Fenicia et al., 2011; Kavetski & Fenicia, 2011), SUMMA (Clark et al., 2015a, 2015b), and MARRMoT (Knoben et al., 2018). Given the difficulties of inferring hydrological processes from models as well as the high cost of detailed field monitoring, there is a clear need for more direct data-driven approaches to investigate SF-GW interactions from commonly available monitoring data.

Causal inference methods (CIMs) seek to identify causality between variables directly from the data. Although these methods are gaining attention in earth sciences (Goodwell et al., 2020; Runge et al., 2019), their application in hydrological sciences remains mostly focused on linear methods, particularly cross-correlation and Granger causality (Granger, 1969). Linear CMIs are faster to compute and easier to relate, however, they assume linear and separable systems which reflects the common conceptualization that variables behave linearly and information about a causative variable is not present in the affected variable. Yet, it is widely known that correlation is not sufficient to infer causality, nor does lack of correlation infers lack of causation. In addition, the linear conceptualization might be inadequate for catchments as they can have an underlying nonlinear dynamic (Kirchner, 2009; Sivakumar, 2000, 2009; Sivakumar & Singh, 2012) with finite resilience to disturbances (Peterson, 2009; Peterson & Western, 2014a; Peterson et al., 2014, 2021; Saft et al., 2015).

The theory governing this nonlinear dynamic, called *dynamical system theory* (DST), has been widely applied in different areas of hydrology (Sivakumar, 2000, 2009, 2017). For instance, Rodriguez-Iturbe et al. (1989) has investigated the chaotic dynamics and limits of predictability of rainfall. Jayawardena and Lai (1994) applied it to predict rainfall and SF and Delforge et al. (2020) to analyze and forecast SF recessions. Erkyihun et al. (2016),

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Rajagopalan et al. (2019), and Sangoyomi et al. (1996) utilized it for hydroclimatic forecasting and simulation. Also, studies that identified low-dimensional dynamics across different hydrological processes showed that complexity can emerge from a small set of governing equations, which underpins the idea that simple deterministic hydrological models can adequately describe complex hydrological processes (Procaccia, 1988). Nonlinear CIMs have been used in hydrology, for instance, to investigate interactions between soil moisture and precipitation (Wang et al., 2018), ecohydrological interactions (Ruddell & Kumar, 2009), dynamics of treatment wetlands (Medina et al., 2019), stream chemistry dynamics (Jiang & Kumar, 2019), drivers of evapotranspiration (ET) (Ombadi et al., 2020), hydrological connectivity in karst systems (Delforge et al., 2022), in rivers deltas (Sendrowski & Passalacqua, 2017), and between GW and SF (Rinderer et al., 2018). Also, recent studies have investigated the potential and limits of some nonlinear CIMs, particularly the impact of noise, sample size, and the occurrence of spurious causal links (Delforge et al., 2022; Ombadi et al., 2020). However, the application of these methods remains incipient and further exploration is important to foster their use in hydrology.

The nonlinear CIM investigated in this paper, called convergent cross-mapping or simply CCM (Sugihara et al., 2012), is based on time-delay embedding proposed by Takens (1981). Here, it was used to investigate SF-GW interactions from the perspective of empirical causal inference: that is, are GW head dynamics causing SF dynamics, or vice versa, or both? This paper explores the potential of the method to contribute to the long-standing challenge of identifying hydrological processes and understanding heterogeneity by comparing the CCM results with a priori perceptual understanding of processes in three upland catchments located in different hydrogeologic and topographic settings of southeast Australia. In light of a high number of spurious causation identified by CCM and reported in the literature (Delforge et al., 2022; Ombadi et al., 2020), we also explore some limitations of the method that are yet not well covered, particularly the role of seasonality, data sampling frequency, and long-term trends. This is done with multiyear GW head and SF data encompassing a wet and a dry period, including a long and severe multiyear drought that has produced complex hydrological behaviors in the study areas.

In the next sections we describe the causality framework applied in this paper (Section 2), present a general account of the study areas along with a priori perceptual understanding of hydrological processes (Section 2.2), and describe the input data for the CCM analysis (Section 2.2.2). The CCM results are presented in Section 3 and discussed in Section 4.

2. Methods and Data

2.1. Dynamic Systems, State-Space Reconstructions, and Attractors

Nonlinear dynamic systems, notably observed by Lorenz (1963) while running weather simulations, are not stochastic, meaning that they follow certain sets of rules to change from one state to the other (Deyle & Sugihara, 2011; Tsonis, 1992). In fine, the temporal behavior of a dynamical system is often pictured as a trajectory within the state-space diagram which represents its evolution from some unknown initial state (Ivancevic & Ivancevic, 2008). Dynamical system's states are solutions of the equations that describe how the system evolves along the state-space trajectories and how the D-explanatory variables interact over time. The temporal behavior of these systems produces trajectories that are recurrent and converge to a portion of the state-space, which is called an "attractor." This sub-space of attraction can be multi-dimensional, but its dimension is less than the D-dimension of the system's state-space driven by D-variables. Deterministic systems with long-term predictability have attractors of integer dimension. In the case of dynamic systems highly sensitive to initial conditions, attractors have fractal (or non-integer) dimension, and the long-term predictability of the system is not certain. Such attractors have a manifold geometry and are called "strange attractors," terminology coined by Ruelle and Takens (1971).

State-space reconstructions using a single time-series often rely on Takens' theorem (Takens, 1981). Essentially, Takens considered time-series as a unidimensional projection of the attractor of a dynamical system (Deyle & Sugihara, 2011). Figuratively, in the context of watersheds, the state-space could be tri-dimensional and explained by rainfall, GW head, and SF, each representing one dimension of the manifold. Therefore, projecting the state-space onto one of its dimensions would yield the time-ordered value of one of the explanatory variables. In other words, time-series are considered a sequential observation of the system's state-space. Accordingly, Takens' theorem suggests that a system's state-space can be reconstructed, or embedded, using enough lagged

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

coordinates of a single time-series to evaluate the characteristics of a dynamic system. Therefore, even if the dynamic system behaves according to multiple variables, information about the system behavior should still be encoded in the state-space reconstructed from a single time-series, or in the so called shadow manifold. Takens (1981) and Deyle and Sugihara (2011) describe how a reconstruction of the system state can be done and for a more thorough explanation of state-space reconstructions and attractors the reader is directed to Ivancevic and Ivancevic (2008), Sugihara et al. (2012), and Tsonis (2018).

2.1.1. Convergent Cross-Mapping (CCM)

Convergent cross-mapping (CCM) is a non-parametric method to infer causality between two variables in nonlinear dynamical systems. This method is based on Takens' theorem of manifold reconstruction. To identify nonlinear interactions, CCM tests for topological similarities between reconstructed attractor manifolds (Sugihara et al., 2012; Tsonis et al., 2018). In other words, CCM detects causality by checking if two variables behave consistently while revisiting similar states.

This is done by measuring the ability of increasingly large fractions of the historical time-series (or library size) of one variable to forecast states of the other variable. Essentially, stronger correlations between observed and predicted states indicate stronger interactions. In detail, to test the hypothesis that variable X causes variable Y, the first step is to construct a shadow manifold from Y (M_Y) where the trajectories M_Y are given by $\{Y_t, Y_{t-\text{tau}}, \ldots, Y_{t-(E-1)\text{tau}}\}$. Here, E is the embedding dimension, or the number of lags used to reconstruct the shadow manifold and tau is the embedding delay which was set at 1 time step of the time-series. For a single forecast at a time of reference, $E+1$ nearest neighbors are identified in M_Y according to their Euclidean distance to Y considering a specific library size. The library size (or sample length) defines the subsample length of the shadow manifold M_Y where nearest neighbors can be identified from. Then, the time-indices of the identified nearest-neighbors are used to map points in X which are then averaged considering the Euclidean distances in M_Y to compute the forecast of X^*. For all time indices, CCM cross-map skill is then quantified as the Pearson's correlation between the observed X and forecasted X^*. As for each library size the M_Y is bootstrapped (randomly and without replacement) a number of times (set to one hundred in this study), the cross-map skill for each unique library size is computed as the mean Pearson's correlation.

Concisely, variable X causes Y and pertains to the same system if it is possible to cross-map X from the shadow manifold of Y, as the causal variable leaves an information signature in the time-series of the affected variable. Hence, the causal variable can be estimated from its signature in the affected variable, which is counterintuitive and different to the popular GC framework (Granger, 1969) and those derived from it, such as conditional mutual information and transfer of entropy (Fraser, 1989; Fraser & Swinney, 1986; Schreiber, 2000), which assume the separability of the causative from the affected variable.

Another important concept is that the affected variable can be used to estimate the states of the causal variable, but not the other way around, as the dynamics of the causal variable does not depend on the affected variable. To illustrate, for a catchment with significant two-way SF-GW interactions, cross-mapping would be possible from SF to GW and vice-versa. However, in cases where SF-GW interactions are primarily one way (i.e., GW drives SF), recovery of information would be similarly one-sided (i.e., only GW's causal relationship on SF can be estimated). Also, CCM estimates the average interactions over the analyzed time-series rather than for a specific subset. Thus, when analyzing multiyear hydrographs, the results refer to the entire time period and not, for example, to a single day or over a season. For a more thorough and visual explanation of the CCM fundamental ideas and algorithm, the reader is directed to Sugihara et al. (2012), Ye et al. (2015), Delforge et al. (2022), and to the following animation: http://tinyurl.com/EDM-intro.

In applying CCM, the SF and GW head used in the CCM analysis were not transformed, therefore, they are the original hydrographs. The optimal embedding dimension (E) to reconstruct the shadow manifolds (Table S4 in Supporting Information S1) was identified by finding the E values (from 2 to 10) that gives the highest cross-map skill (ρ). The CCM analysis was carried out assuming a contemporaneous causality given that a lag time of zero invariably produced the higher cross-correlation in our data. The rEDM 1.2.3 package (Ye et al., 2020) for the programming language R 4.0.0 (R Core Team, 2020) was the software package used in the data analysis.

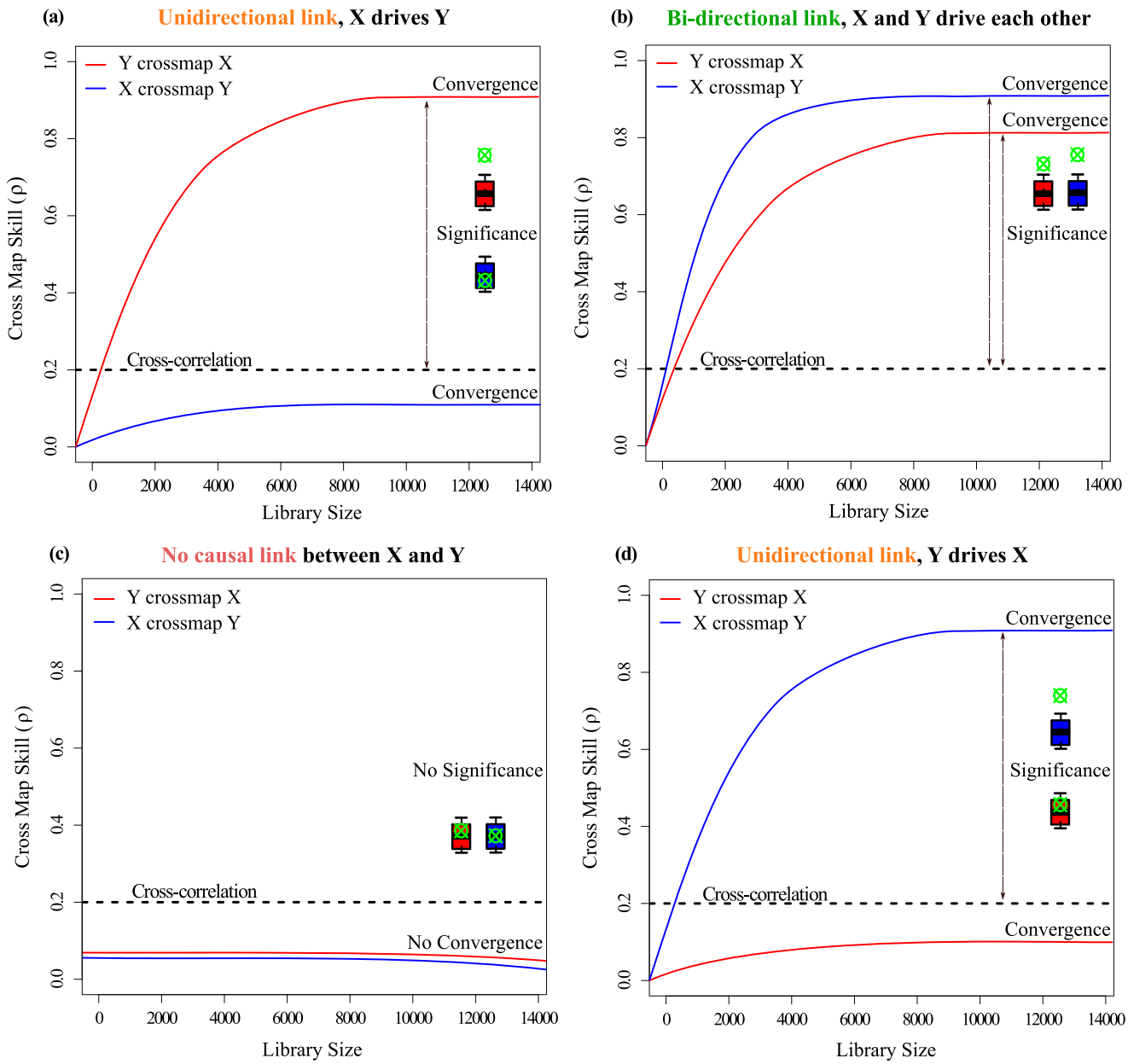


Figure 1. Explanatory example of convergent-cross mapping results and interpretation. (a) A unidirectional causal link as only variable *X* drives *Y*, without feedbacks, because only the cross-map skill (ρ) from *Y* to *X* significantly exceeds the cross-correlation at full library size and at the same time rejects the null hypothesis of a strong common forcing or seasonality as the driver of the results (see cross-correlation lines and green marker in relation to the boxplots in the figure). (b) A strong bi-directional link between *X* and *Y*, with a somewhat weaker effect of *X* on *Y*. (c) Shows no causal interaction between *X* and *Y*, and (d) suggests a unidirectional effect of *Y* on *X*. As indicated in the text, the library size refers to the number of data points considered.

2.1.2. Convergence and Significance to Distinguish Unidirectional, Bidirectional, and Non-Causal Interactions

To distinguish causation from spurious causality using CCM, it is necessary to observe convergence in the cross-map skill, which indicates information flow from the causal to the affected variable. This is done by checking if larger library seizes yield increasingly larger cross-map skills. To illustrate, Figure 1(a) shows that larger library sizes produce higher cross-map skill from Y to *X* (hypothetical variables). The idea is that longer samples generate denser manifolds with closer nearest neighbors that allow more precise cross-map estimates.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Moreover, for this convergence to be significant, the cross-map skill (mean Pearson's correlation) at full library size should be significant to avoid being confounded as spurious causality. To assess the significance of the results in this study ($p < 0.01$), we tested whether the CCM results could be due to seasonality in the hydrographs or to a strong common forcing (i.e., seasonal rainfall or ET) using Fisher's Z test to check whether the cross-map skill (ρ) at full library size was significantly higher than the highest lagged cross-correlation (0–30 time steps) as well as stochastic replicates of SF and GW head. The seasonal method available in Ye et al. (2020) was used to create $k = 200$ stochastic replicates, that is, surrogate time-series for each hydrograph. Here, the method computes the seasonality of the annual cycle with a smoothing spline and shuffles the residuals without additive Gaussian noise. The p-value is calculated according to $(n + 1)/(k + 1)$, where k is the total number of surrogates and n represents the amount of replicates that produced higher cross-map skill than the actual data.

Figure 1 explains the interpretation of the different types of interaction using the convergence and significance criteria. This figure uses horizontal lines to illustrate the highest lagged cross-correlation between the variables and boxplots to show the difference between the CCM results obtained from stochastic replicates and the one from the original data. The blue boxplots represent the distribution of CCM results obtained from stochastic replicates of the variable X cross-mapping the observed variable Y. The red boxplots show the results from surrogate time-series of Y cross-mapping the variable X. The green target circles show the CCM results obtained with the original datasets. Accordingly, Figures 1a and 1d show unidirectional causal links between the variables, as the significance criteria is achieved only for one direction of interaction (X causing Y) even though convergence is achieved in all cases. Figure 1b illustrates a bi-directional causal link since the convergence and significance criteria are achieved for both cross-mapping directions. Figure 1c shows no causal interaction between the variables as both the convergence and significance criteria were not achieved.

2.1.3. Exploring the Role of Data Length and Frequency, Seasonality, Long-Term Trends, and Climatic Conditions on the CCM Analysis

As mentioned, in addition to use CCM to test an a priori conceptual understanding of processes in the study areas, we seek to explore the role of seasonality, long-term trends and data sampling frequency on the CCM results. Seasonality potentially leads to spurious causality as well as represent the dynamic of a third common cause. Data sampling frequency affects both seasonality and length of the data available for the analysis.

To evaluate the dynamics of SF-GW interactions at different time steps as well as the role of data sampling frequency on the significance of the results, the CCM analysis was done using the original daily data and using monthly aggregated hydrographs. The monthly SF hydrographs were derived as the monthly summation of daily SF and the monthly GW head time-series as the head observed in the last day of the month. The same significance tests were applied to both daily and monthly CCM results and a comparison between the daily and monthly analysis in the context of the stochastic replicates was used to evaluate the role of seasonality in the CCM results, as at the monthly time step both SF and GW hydrographs are expected to be more seasonal.

Regarding long-term trends, they can be problematic for the CCM framework as they can violate the recurrence assumption of dynamic trajectories, thus, reducing the capacity to verify whether two time series behave consistently in the state-space. This means that nearest neighbors in the state-space could be irrelevant, a behavior that would contradict the CCM framework. To avoid this issue, one could use different data filtering techniques to eliminate trends, however, as trends in nature are real (Gleick, 1987) and filtering the data usually changes its dimensionality (Badii et al., 1988; Chennaoui et al., 1990; Jayawardena & Lai, 1994; Mpitsos et al., 1988), we investigated the role of trends in our CCM results using the original datasets. To examine the role of long-term trends jointly with different climatic conditions, we divided the daily CCM analysis for bores S_bore1 and S_bore2 in Sunday Creek into three subperiods of approximately 10 years each. These subperiods encompassed different climatic conditions as well as remarkably different trends in both SF and GW head hydrographs. The 1,500 days moving average (MA1500, roughly 4 years) was used to make the long-term trends more noticeable in the SF hydrograph of Sunday Creek.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

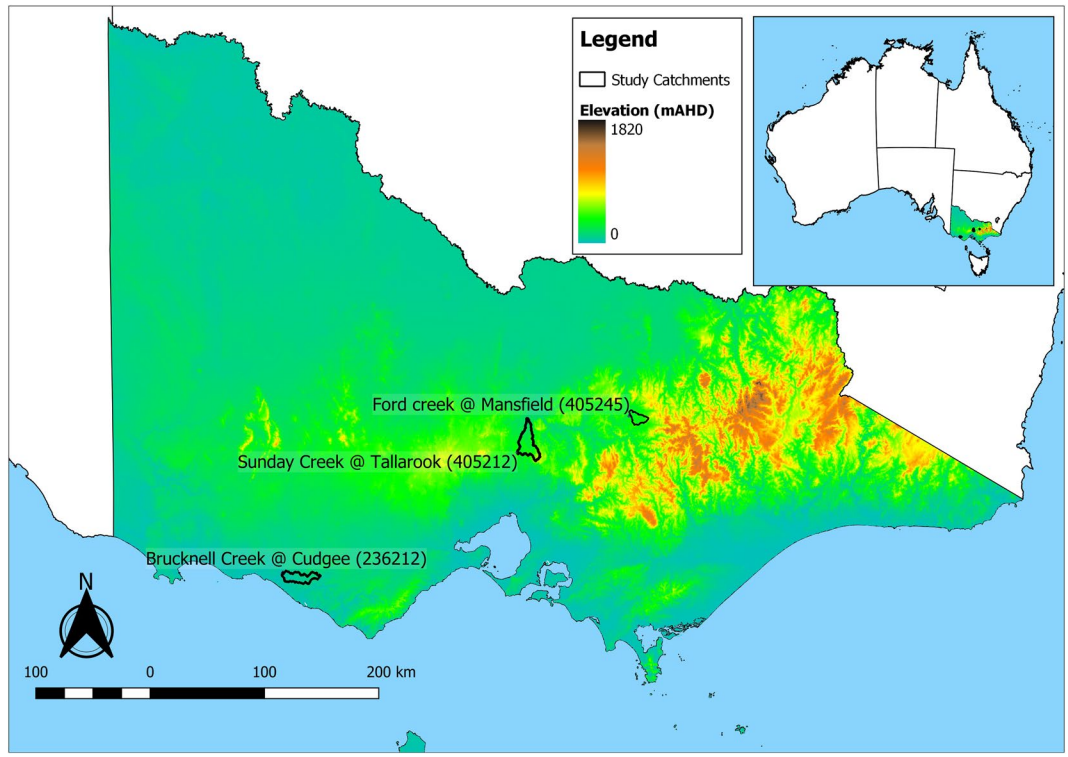


Figure 2. Map of Victoria, Australia, emphasizing the study catchments and the topography of the region.

2.2. Study Site and Data

2.2.1. Study Areas and *A priori* Perceptual Understanding of Hydrological Processes

The study area encompasses three upland catchments located in different hydrogeological settings in the state of Victoria, Australia (Figure 2), where SF and GW interactions are important management topics. Victoria's climate is classified as temperate (Peel et al., 2007) with high intra-annual seasonality in SF and rainfall (Peel et al., 2001). Moreover, the region experiences severe multiyear droughts (Potter et al., 2010) and the recent Millennium Drought (1997–2009) was the most severe on record (CSIRO, 2010; Van Dijk et al., 2013). Meteorologically, this drought ended in 2010 with large floods caused by a strong La Niña event (National Climate Centre - Bureau of Meteorology, 2011; Van Dijk et al., 2013). However, the hydrological drought appears to have continued, producing complex RR non-recovery behavior that challenges the assumption that SF always recovers from drought (Peterson et al., 2021), and there is much speculation as to whether changes in ET and the "long memory" of GW dynamics play a part in the persistence of the hydrological drought (Hughes et al., 2012; Peterson et al., 2021; Petrone et al., 2010; Saft et al., 2015, 2016a, Saft, Peel, Western, & Zhang, 2016; Scaife & Band, 2017). For instance, a recent multi-disciplinary study suggests the duration of the drought as the main driver of the RR non-recovery together with interconnected GW processes, namely: declining GW levels and recharge which are associated with unsaturated zone expansion and less interaction between surface and underground water (Fowler et al., 2022). At the same time, the same study highlights other plausible causes, such as increased evaporative demand, and the need for more long-term field monitoring of internal and subsurface processes to consolidate a physical explanation for the multiyear dynamics.

The Australian Bureau of Meteorology (Doody et al., 2017) considers all study areas as ground water dependent ecosystems (GDE) and therefore we expect GW to be driving SF in all three catchments. Various characteristics of the three catchments are provided in Table 1 and, according to the joint analysis and discussion that is presented below, our conceptual understanding suggests that GW drives SF more significantly in Brucknell Creek than in Sunday or Ford Creek. Additionally, the causal relationship in the opposite direction (flow driving GW) is conceptualized to be less significant in general, as the study areas are upland catchments where recharge from the stream is conceptualized to be less active than rainfall infiltration. Within the catchments, we expect bores

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Hydrological Characteristics of the Study Areas

Catchment	Brucknell creek @ cudgee	Ford creek @ mansfield	Sunday creek @ tallarook
Catchment area (km^2)	223	116	332
Runoff (mm/year)	117	91	97
Rainfall (mm/year)	813	878	761
PET Morton-CRAE (mm/year)	1,017	1,141	1,093
Rainfall-runoff ratio during the Millennium Drought (1997–2009)	−30% to −50% than expected	−50% to −75% than expected	>−75% than expected
Runoff ratio	0.14	0.10	0.13
Cease to flow ratio	0	0.03	0.11
Intra-annual monthly flow variation (Max-Min)/Mean	10	16	13
Monthly flow variance explained by seasonal cycle (R^2)	0.39	0.18	0.01
Flow Daily CV (%)	264	508	416
Baseflow Index (Lynne and Hollick)	0.27	0.08	0.12
Number of observation bores used in the study	5	2	5
Average number of monitoring years	14	8	21
Average number of observations per bore per year	40	22	15
Permissible consumptive volume of groundwater	85,000 ml/yr.	1,496 ml/yr.	3,000 ml/yr.
Hydrogeology	Regional porous limestone and local fractured basalt aquifers	Local fractured sedimentary and local low-grade metamorphic rock aquifers	Local fractured sedimentary and local low-grade metamorphic rock aquifers

Note. GW PCV as per DELWP's Groundwater Resource Reports.

in valley bottoms to have stronger links with SF, as they tend not only to be closer to the stream but also to be in areas that are more porous and permeable given sedimentary formations and/or deeper soils.

In terms of the hydrological characteristics provided in Table 1, at all three catchments the annual RR ratio shifted in response to the Millennium Drought but to different degrees (Saft, Peel, Western, & Zhang, 2016) and appears to have complex non-recovery behavior (Peterson et al., 2021). Brucknell Creek was the one less affected, where the RR ratio was 30%–50% lower during the drought. Ford and Sunday Creek were more strongly impacted, with decreases of 50%–75%, and >75%, respectively (Saft, Peel, Western, & Zhang, 2016). As previously mentioned, changes in GW might explain these shifts, however, there is no detailed monitoring of the hydraulic gradient between the river and the surrounding water table that could describe changes in the gaining and loosing stream processes. Moreover, classic RR models cannot reproduce this behavior given their representation of SF-GW interactions (Fowler et al., 2020) and therefore the use of a data driven approach to shed light on the SF-GW interactions in the study areas is of paramount importance for future research.

Focusing on possible anthropogenic causes of RR shifts, the volume of GW extractions within the catchments could help explain these different behaviors, but the catchment with the smaller shift in the RR ratio was the one that has the highest annual permissible consumptive volume (PCV) of GW, namely Brucknell Creek. In this catchment, the PCV represents approximately 330% of the mean annual flow, while in Ford and Sunday Creek it represents only around 14% and 9% of the mean annual flow for the entire record, respectively.

At the same time, while rainfall and potential evapotranspiration (PET) are different between the catchments, Brucknell Creek does not have the larger mean annual rainfall. Neither has Sunday Creek the highest PET demand. In addition, all three catchments have seasonal SF (low flows in summer and high flows in winter, as seen in Figure S1 in Supporting Information S1), which is mostly driven by seasonality in evaporative demand

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Figure S2 in Supporting Information S1) followed by rainfall (Figure S3 in Supporting Information S1). Brucknell Creek has a stronger seasonality as 39% of its monthly flow variance is explained by the seasonal cycle computed with a smoothing LOESS (Savitzky & Golay, 1964) whereas in Ford and Sunday Creek this percentage decreases to 18% and 1%, respectively.

Brucknell Creek is a perennial stream, even during the Millennium Drought, and has the highest RR ratio (0.14), a baseflow index (BFI as per Lyne & Hollick, 1979; Nathan & McMahon, 1990) of 0.27 while in the other catchments BFI is below 0.13. Also, the intra-annual variation of monthly SF is more than 30% smaller in Brucknell Creek than in Ford or Sunday Creek. With regard to hydrogeology, Brucknell Creek sits mostly on a regional high productivity porous limestone aquifer, having a small portion of the catchment located on local low productivity fractured basalt aquifers (Brodie et al., 1998; Dahlhaus et al., 2002). Conversely, Ford and Sunday Creek are on local low productivity aquifers that are composed of fractured sedimentary and low-grade metamorphic rocks (Brodie et al., 1998; Walker et al., 2003). Figure S4 in Supporting Information S1 presents a detailed hydrogeological map of each catchment emphasizing the distinct types of aquifers and the location of the observation bores.

Although the RR shift magnitudes cannot be easily explained, as previously mentioned, Brucknell Creek is expected to have more significant GW and SF interactions that potentially could explain the smaller shift in the RR ratio during the long and severe Millennium Drought, as its regional high-productivity aquifer could offer extra buffer to sustain SF-GW connectivity thanks to its larger recharge area and larger GW storage. At the same time, all these parameters do not shed light on the strength and direction of this interaction across the catchments, or on how different areas of the aquifer might interact differently with the SF. Only using commonly available monitoring data, this study offers a methodology that seeks to derive more granular analysis of SF-GW interactions and a way to further develop and test conceptual models.

2.2.2. Streamflow and Groundwater Data

The daily observed SF data was sourced from Victoria's Department of Land, Environment, Water and Planning (DELWP) using the Water Measurement Information System (Figure S8 in Supporting Information S1). The GW head observation data was from DELWP's State Observation Bore Network (SOBN), available at the same information system. There are 34 SOBN in the catchment areas, however, in this study we consider only those that had (a) multiyear monitoring (>5 years of data), (b) a relatively high number of observations per year of monitoring (>9 obs./year), and (c) monitoring period fully or partially covering the Millennium Drought (1997–2009, approximately). In total, 12 bores fulfilled these criteria in the study areas, two in Ford Creek, and five bores in each of Brucknell and Sunday Creeks. Table 1 and Figure S5 in Supporting Information S1 summarize the raw GW data for each catchment.

Given the paucity of the observed GW head time-series in Victoria, the raw observations were temporally interpolated on a daily time-step using the freely available HydroSight software (https://github.com/peterson-tim-j/HydroSight/wiki). HydroSight uses the modeling approach of Peterson & Western (2014b) to interpolate between each observation based on climate drivers (Peterson & Western, 2018). In this study, we first identified and removed outlier observations as per Peterson et al. (2017) and then interpolated the hydrographs based on area weighted daily rainfall and PET at the location of each bore. These climate forcings were sourced and derived using the freely available R-package AWAPer (https://github.com/peterson-tim-j/AWAPer, Peterson et al., 2019). The PET data were determined using results from Morton's complementary relationship areal model (Guo et al., 2016; Morton, 1983) with daily maximum and minimum temperature and vapor pressure data from the Australian Water Availability Project (AWAP) database, and net solar radiation calculated with empirical equations from the Food and Agriculture Organization (Allen et al., 1998). The reliability of the resulting interpolated hydrographs was then evaluated using leave-every-other-point-out cross-validation. The average performance across all bores was 0.84 (Nash-Sutcliffe coefficient of efficiency) and the cross-validation results are presented in Table S1 and Figure S7 in Supporting Information S1. The observed and interpolated GW head data are shown in Figure S5 in Supporting Information S1. The SF and GW head time-series used in the CCM analysis are presented in their transformed form (zero mean, unit variance) in Figure S9 in Supporting Information S1.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3. Results

The CCM results indicate statistically significant causal relationships between GW head and SF in all three upland catchments at both daily and monthly time steps. Figures 3–5 present the daily and monthly CCM results and their interpretation for all 12 bores evaluated in Brucknell, Sunday, and Ford Creek, respectively, along with the location within the hydrogeologic setting of each catchment. The daily results are plotted with contiguous lines whereas the monthly results are shown with dashed lines. For a more detailed bore location and description of their hydrogeology, the reader is directed to Figure S4 in Supporting Information S1 and to Dahlhaus et al. (2002) and Walker et al. (2003).

For both daily and monthly time scales, three types of SF-GW causal interactions were identified across the study catchments, namely (a) GW is causal to SF, marked in red as GW- > SF; (b) SF is causal to GW, marked in blue as SF - > GW; and (c) GW and SF are causal to each other, marked in green as GW <-> SF. For instance, at the daily time step, Figure 3(a) indicates a significant bidirectional link between GW head and SF, as the cross-map skill (ρ) at maximum library size substantially exceeds the cross-correlation in both directions and is significant when compared to the surrogate hydrographs. At the same time, as the cross-map skill from head to flow is smaller than in the opposite direction, the results indicate a slightly weaker effect of flow on head. In terms of unidirectional link, Figure 3d shows GW head significantly driving SF without feedbacks, as only the cross-map skill (ρ) from flow to head significantly exceeds the cross-correlation at maximum library size although it is significant when compared to the stochastic replicates. Conversely, Figure 4c indicates SF unidirectionally affecting GW head, because only the cross-map skill (ρ) from head to flow is significant when compared to the cross-correlation and surrogate time-series. As mentioned in the Section 2, the direction of causation is contrary to the cross-mapping direction since only the causal variable can leave footprints on the manifold of the affected variable, making the backward cross-mapping possible.

Table 2 summarizes the daily SF-GW interactions interpreted from the CCM results for the three study areas. It shows the presence of bidirectional interaction in at least one bore location in each catchment. Furthermore, at least one type of unidirectional link was identified in all study areas. Across all bores, there is a significant variability in SF-GW causality strength and direction across and between catchments. Ford Creek presented both unidirectional and bidirectional causal links, where GW head in both bores is affecting SF with similar strength. In Brucknell Creek the interaction strength and direction vary within the catchment with both bidirectional and unidirectional causal links. Sunday Creek presents the more heterogeneous combination of interactions, as the CCM results suggest the presence of all three causal links with variable strengths. In addition to Table 2, the interpretation of the CCM results in terms of the SF-GW interaction for each bore in the three study is presented as color-coded annotations within Figures 3–5 and 7.

Figure 6 uses boxplots to show the significance of both the daily and monthly CCM results in relation to the null hypothesis that the results are driven by a common strong forcing or seasonality ($p < 0.01$). The blue boxplots represent the distribution of CCM results obtained from stochastic replicates of GW cross-mapping the original SF. The red boxplots show the results from surrogate time-series of SF cross-mapping actual GW. The green target circles show the CCM results obtained with the original hydrographs and their relative distance to the ones obtained with stochastic replicates. The boxplots for the daily and monthly analysis are grouped together for each bore, which are grouped by catchment. Figure 6a shows the results obtained for Brucknell Creek, while Figures 6b and 6c present the surrogate analysis for Sunday and Ford Creek, respectively. Figure 6d shows the surrogate analysis for three subperiods with different trends in the hydrographs and different climatic characteristics for two bores (S_bore1 and S_bore2) in Sunday Creek. Figures 6a–6c show that the monthly stochastic replicates tend to produce higher median cross-map skills, regardless of the direction of interaction, than the daily stochastic replicates, particularly in Brucknell Creek. Figures 6a–6d show that the daily CCM results are significant for most of the bores in all catchments when compared to the surrogate time-series.

Figure 7 and Table 3 examine in more detail the role of long-term trends jointly with different climatic characteristics on the CCM results. Figures 7b, 7d and 7f show the original daily data emphasizing the three subperiods evaluated and the trends in the data. Figure 7(a) shows the daily CCM results for the first wet period of 1975–1985. Figure 7b presents the CCM results for the wet period just before the Millennium Drought, from 1985 to 1996. Figure 7c shows the CCM analysis for during and after the Millennium Drought, from 1996 to 2013. The daily results for S_bore1 are plotted with contiguous lines whereas the results for S_bore2 are shown

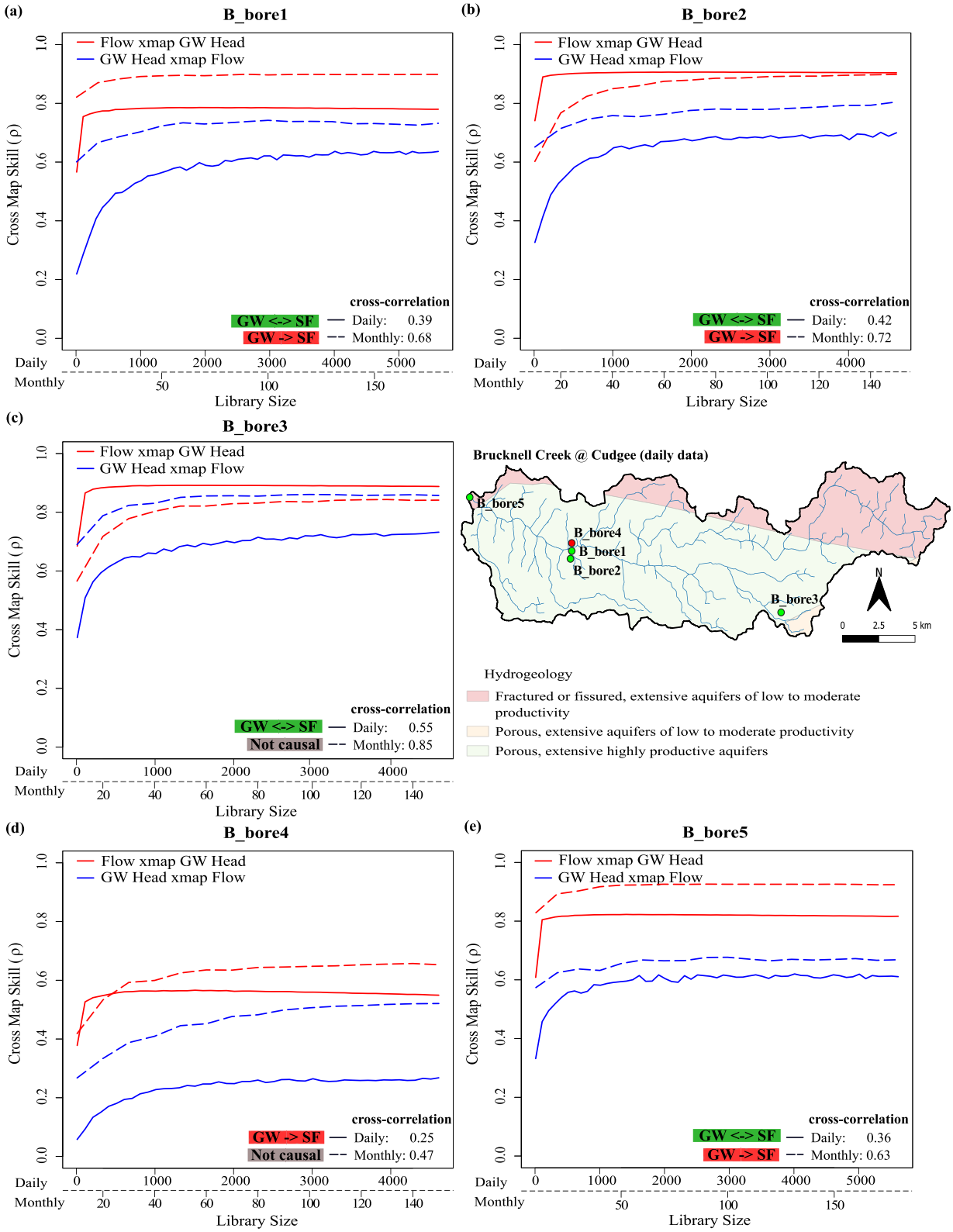


Figure 3. Daily and monthly convergent cross-mapping (CCM) results and the causal links between the streamflow of Brucknell Creek at Cudgee and the groundwater head in five state-observation bores, including their location in the hydrogeologic setting of the catchment. For the interpretation of the daily results for each bore, the reader is directed to Table 2 and the color-coded annotations within the figure (red = GW causes SF; blue = SF causes GW; green = bidirectional SF-GW interaction). For the monthly results, the reader is directed to Table S3 in Supporting Information S1.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

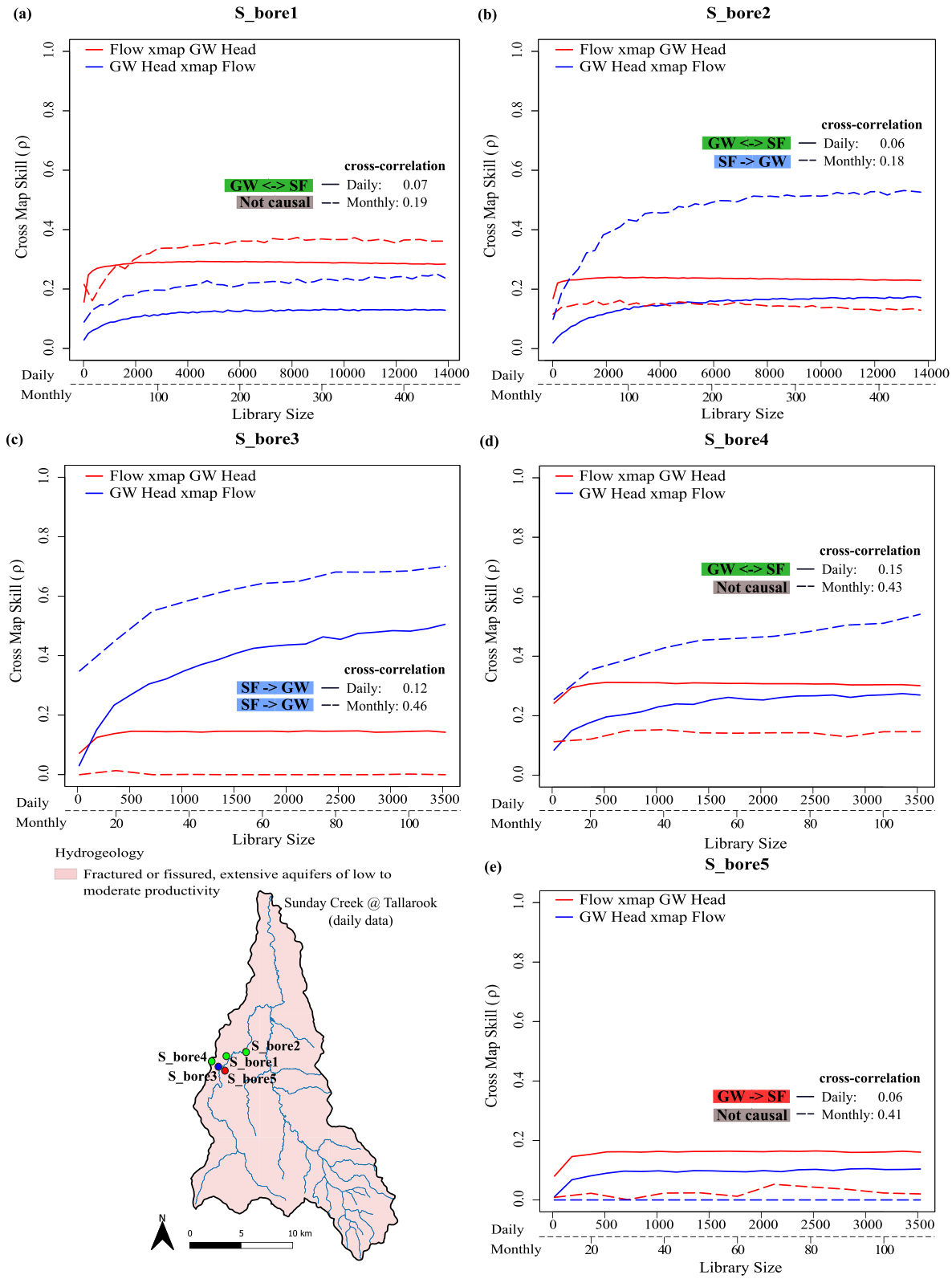


Figure 4. Daily and monthly convergent cross-mapping (CCM) results and the causal links between the streamflow of Sunday Creek at Tallarook and the groundwater head in five state-observation bores, including their location in the hydrogeologic setting of the catchment. For the interpretation of the daily results for each bore, the reader is directed to Table 2 and the color-coded annotations within the figure (red = GW causes SF; blue = SF causes GW; green = bidirectional SF-GW interaction). For the monthly results, the reader is directed to Table S3 in Supporting Information S1.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

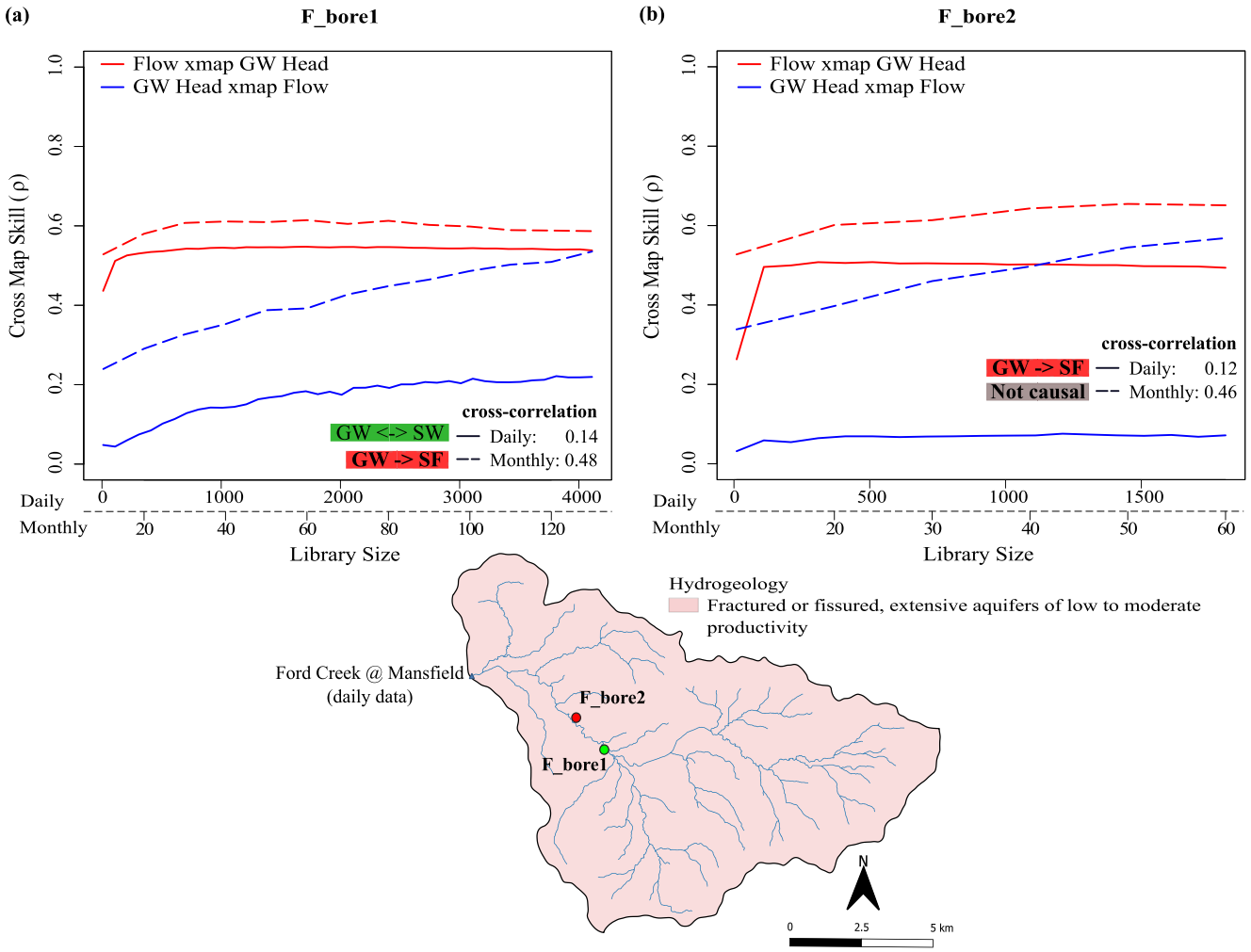


Figure 5. Daily and monthly convergent cross-mapping (CCM) results and the causal links between the SF of Ford Creek at Mansfield and the groundwater head in two state-observation bores, including their location in the hydrogeologic setting of the catchment. For the interpretation of the daily results for each bore, the reader is directed to Table 2 and the color-coded annotations within the figure (red = GW causes SF; blue = SF causes GW; green = bidirectional SF-GW interaction). For the monthly results, the reader is directed to Table S3 in Supporting Information S1.

with dashed lines. As mentioned in Section 2, despite the time steps used in the analysis, the causation results relate to the average interactions over the period evaluated rather than to a specific point or subset of this period, therefore, relating to the perceptual understanding of predominant hydrological processes within the study areas over the period analyzed. Tables S2 and S3 in Supporting Information S1 show the daily and monthly CCM results, respectively, along with their statistical significance in the context of both cross-correlation and surrogate hydrographs.

4. Discussion

4.1. CCM Results and the A priori Perceptual Understanding of Processes

The literature on CCM in hydrology remains incipient and, as such, this paper offers new insights into the potential of this method, hopefully contributing to the long-standing challenge of identifying dominant hydrological processes for a given catchment. At the same time, given the limited range of environments considered here, we would like to highlight the empirical nature of our findings. Not only the sample size needs to be increased, preferably using hundreds of more diverse and longer hydrographs, but detailed field monitoring should be implemented to confirm the results. Furthermore, the CCM technique itself is relatively young and in addition to the

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2
Daily CCM Results at Full Library Size Along With the SF-GW Causal Links Identified Across Brucknell, Ford, and Sunday Creek

Catchment	Observation bore ID	Flow cross-map GW head	GW head cross-map flow	Linear cross-correlation	Causal link type	Causal link details
Brucknell Creek @ Cudgee	B_bore1	0.78	0.63	0.39	Strong bidirectional link	Slightly weaker effect of SF on GW head
	B_bore2	0.90	0.69	0.42	Strong bidirectional link	Weaker effect of SF on GW head
	B_bore3	0. 89	0.73	0.55	Strong bidirectional link	Weaker effect of SF on GW head
	B_bore4	0.55	0.27	0.25	Medium unidirectional link	GW head driving SF
	B_bore5	0.82	0.61	0.36	Strong bidirectional link	A weaker effect of SF on GW head
Sunday Creek @ Tallarook	S_bore1	0.28	0.14	0.07	Weak bidirectional link	Weaker effect of SF on GW head
	S_bore2	0.22	0.18	0.06	Weak bidirectional link	Slightly weaker effect of SF on GW head
	S_bore3	0.15	0.50	0.12	Medium unidirectional link	SF driving GW head dynamics
	S_bore4	0.32	0.27	0.15	Weak bidirectional link	Slightly weaker effect of SF on GW head
	S_bore5	0.16	0.10	0.06	Weak unidirectional link	GW head weakly driving SF
Ford Creek @ Mansfield	F_bore1	0.54	0.21	0.14	Medium bidirectional link	Weaker effect of SF on GW head
	F_bore2	0.49	0.07	0.12	Medium unidirectional link	GW head driving SF

limitations investigated and discussed in this study (Section 4.2), several aspects need further exploration, such as the effect of error and intermittency in the data.

Besides the limitations of the method, the daily and monthly CCM results (all bores in Figures 3–5) offer significant statistical evidence that suggests a robust estimation of the long-term average causal pathways in the study areas which would not be evident from correlation in our data (e.g., F_bore2 in Table 2). Moreover, the results are consistent with the a priori understanding of SF-GW interactions Specifically, they suggest that Brucknell Creek has stronger SF-GW interactions than in Sunday and Ford Creek, which is in accordance with the understanding that this catchment has more significant and persistent SF-GW processes that led to smaller shifts in RR ratio during the Millennium Drought.

The SF-GW causal links presented large variability in strength and direction among bores and between catchments, identifying three types of causal links. The variability in the causality results are in agreement with experimental work that suggests that SF and GW interactions are complex and take place with different strengths and directions in response to a multitude of factors (Barthel, 2014; Barthel & Banzhaf, 2016). In this study, the meteorological forcings, namely rainfall and PET, differed by only 16% and 12%, respectively, between Brucknell, Ford and Sunday Creek (see Table 1, Figures S2 and S3 in Supporting Information S1) and, therefore, might not be the major factors driving the variability in the SF-GW interactions between catchments. Conversely, the different topography (hilliness) and hydrogeology (aquifer type) of the study areas seem to be good indicators of the differences in the long-term average causal pathways in the catchments.

In order to explore the role of topography, Figure 8 shows the 1-s Multi-resolution Valley Bottom Flatness (MrVBF) of the catchment areas (Gallant & Dowling, 2003), which is a topographic index used to map depositional/erosional areas and infer valleys and hills. Zero MrVBF values indicate steep and erosional terrain, whereas larger values indicate increasingly larger depositional areas, which are often associated with deeper soils (Gallant & Dowling, 2003). Hence, in Figure 8 we assume that higher MrVBF values relate to sedimentary formations

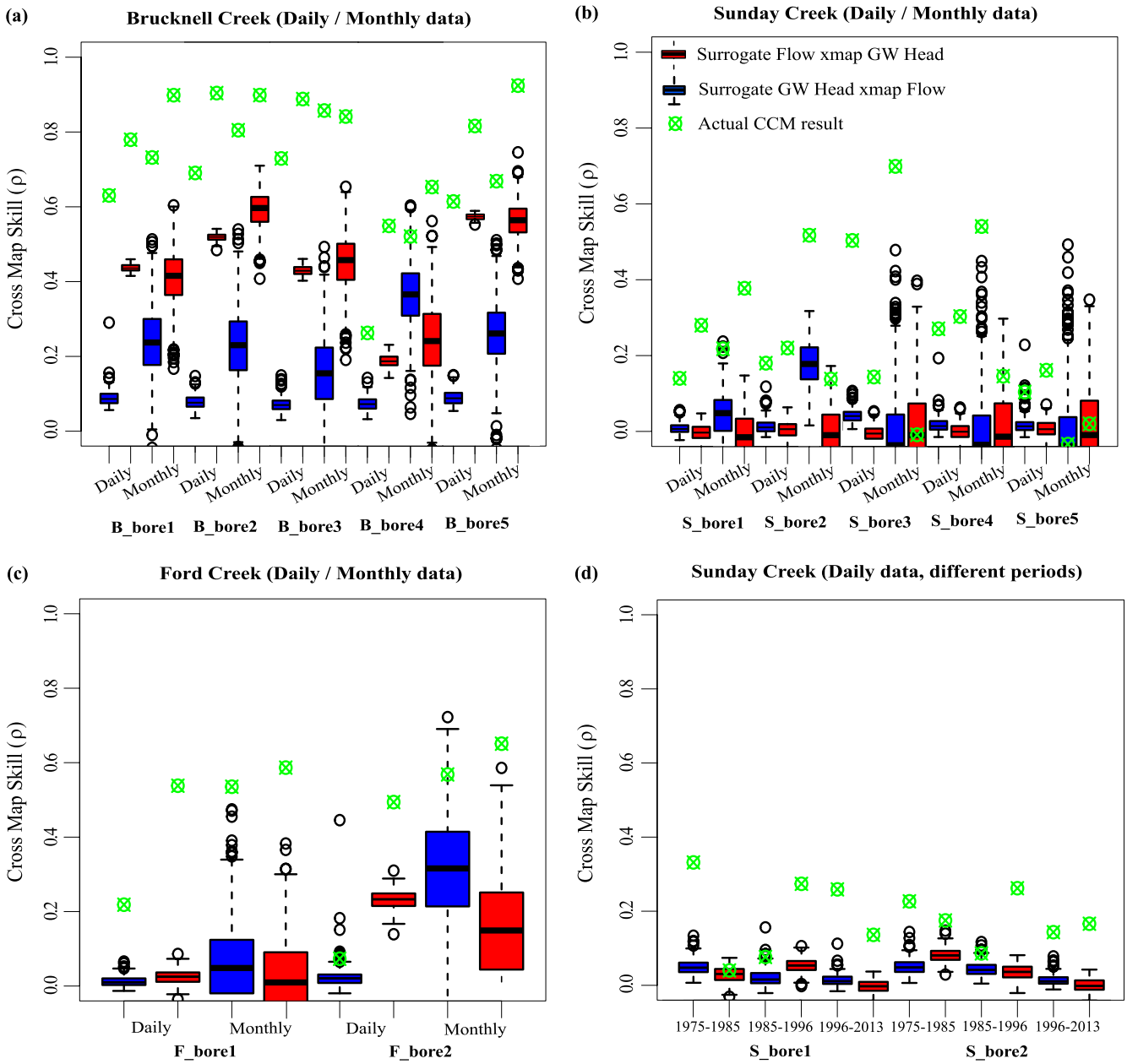


Figure 6. Boxplots showing the significance of the daily and monthly convergent cross-mapping (CCM) results of the actual hydrographs (green markers) when compared to the results obtained for stochastic replicates (blue boxplots for groundwater cross-mapping streamflow (SF) and red boxplots for the opposite) in order to test the null hypothesis of a common strong forcing or seasonality as the driver of the results. (a–c) present the significance of the daily and monthly results for Brucknell, Sunday and Ford Creek. (d) Presents the significance of the daily results for three different periods for bores S_bore1 and S_bore2 in Sunday Creek.

with higher storage and permeability. Carefully looking at the daily CCM results (see Table 2, Figures 3–5) jointly with the MrVBF values across the catchments, it is possible to observe that GW head in bores away from valley bottoms tend to be less causal on SF and have unidirectional interaction, mostly from head to flow. Moreover, bores located in the valley bottoms seem to interact more strongly with SF, and this interaction can go either way: GW affecting flow, or SF driving GW dynamics.

In terms of how the hydrogeologic settings impact the daily SF-GW interactions, by comparing the CCM results with the hydrogeology of the catchments (Figure S4 in Supporting Information S1), one can notice that GW head in the regional porous and productive aquifer of Brucknell Creek appear to drive SF more strongly than in the local fractured and low productivity aquifers found in Ford and Sunday Creek, where SF eventually ceases during

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

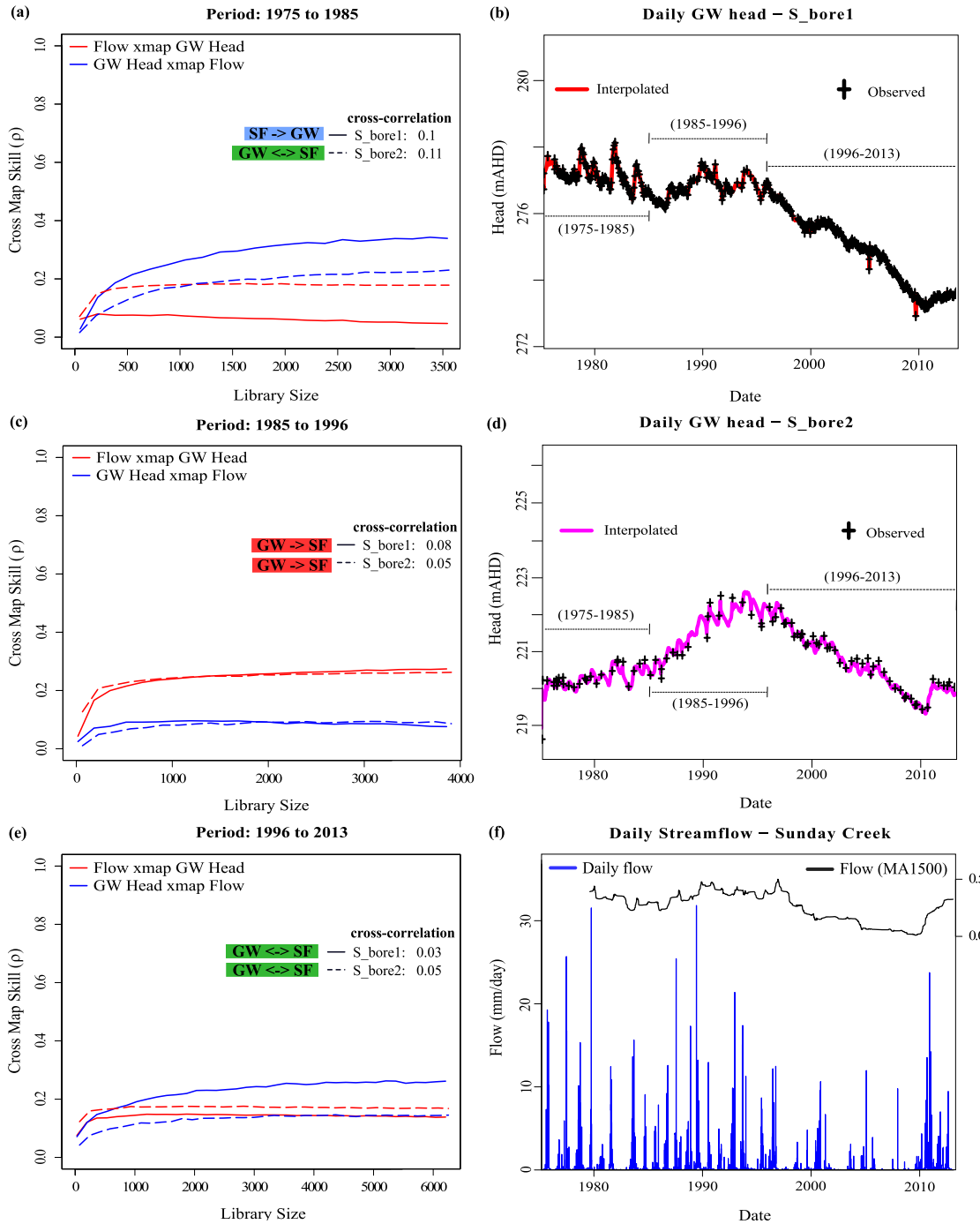


Figure 7. Daily convergent cross-mapping (CCM) results and the causal links between the streamflow (SF) and groundwater (GW) head for three subperiods, namely (a) 1975–1985, (c) 1985–1996, and (e) 1996–2013, where there were remarkable trends in the hydrographs and climatic conditions were different (before, and during and after the Millennium Drought). (b) and (d) present the daily GW head hydrographs for bores S_bore1 and S_bore2, respectively. (f) Shows the daily SF of Sunday Creek at Tallarook and its moving average of 1,500 days to emphasize multiyear trends. For the interpretation of the results for each bore, the reader is directed to Table 3 and the color-coded annotations within the figure.

prolonged dry periods. When further cross-referencing the CCM results with the hydrologic characteristics of the catchments (Table 1), we see that the higher runoff ratios, stronger seasonality, and lower intra-annual variation of monthly flow occur in the catchment with greater influence of GW head on SF, namely Brucknell Creek. This reinforces the understanding that SF-GW interactions are more significant in this catchment.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Daily CCM Results Along With the SF-GW Causal Links for Three Different Subperiods With Remarkable Trends in the Hydrographs of Bores S_bore1 and S_bore2 and Sunday Creek at Tallarook Encompassing Different Climatic Conditions (Before, During and After the Millennium Drought)

Catchment	Observation bore ID	Flow cross-map GW head	GW head cross-map flow	Linear cross-correlation	Causal link type	Causal link details
Sunday Creek @ Tallarook (different trends and climatic conditions)	S_bore1 (1975–1985)	0.04	0.33	0.14	Weak unidirectional link	Streamflow driving GW head dynamics
	S_bore1 (1985–1996)	0.28	0.08	0.08	Weak unidirectional link	GW head driving streamflow
	S_bore1 (1996–2013)	0.14	0.26	0.03	Weak bidirectional link	Weaker effect GW head on flow
	S_bore2 (1975–1985)	0.18	0.23	0.11	Weak bidirectional link	Slightly weaker effect of GW head on flow
	S_bore2 (1985–1996)	0.26	0.09	0.05	Weak unidirectional link	GW head driving streamflow
	S_bore2 (1996–2013)	0.17	0.14	0.05	Weak bidirectional link	Slightly weaker effect of flow on GW head

The variability in our results are in accordance with previous experimental studies that also found spatial heterogeneity in SF-GW interactions (Bishop et al., 2011; Cartwright & Gilfedder, 2015; Cartwright & Morgenstern, 2015; Grayson et al., 1997; Rosenbaum et al., 2012; Seibert et al., 2003, 2011), however, at the same time, our results offer complementary insights. For instance, Cartwright and Gilfedder (2015) mapped GW inflows to 62 km of the Deep Creek, southeast Australia, using Radon surveys and found that gaining sections of the river were in reaches closer to steep valleys or in the edge of its floodplain (close to hills with basement rock), suggesting that steeper and less permeable areas were the major source of baseflow. At first glance, this seems to contradict our findings. However, it is important to notice that our CCM results represent the long-term average causality of the GW dynamics of each bore to the catchment SF and, although our study found that bores in valley bottoms are more causal on SF, it does not invalidate the idea of steep and less permeable areas as sources of baseflow. In fact, what it suggests is that GW head variations in steeper terrain affect the catchment SF less intensely than those in valley bottoms. If we consider that GW moves in a continuum from top to bottom of a hillslope until it discharges to the stream and, at the same time, assume that the underground formation of valley bottoms tend to be more porous and permeable, we can conceptualize that head fluctuations in the valley bottom represent larger displacements of water and, consequently, it would affect the SF more strongly.

Similarly, Bishop et al. (2011) identified a strong and spatially structured differentiation of local flow within a small till catchment, as during runoff events most of the SF was coming from the interflow in the more porous and permeable upper level of the soil. As our bores are all relatively shallow and have a shallow depth to water table, it can be that GW head in the valley bottoms are responsible for triggering strong interflow in areas where soils tend to be deeper and more porous.

Additionally, GW head in the valleys could also be linked to soil saturation levels across wider areas of the catchment and therefore it would significantly trigger different runoff generation mechanisms in the catchment (Ali et al., 2015; Hughes et al., 2012; Petrone et al., 2010; Saffarpour et al., 2016; Seibert et al., 2003). In this case, conceptually, higher GW levels in the valley bottoms would imply the occurrence of saturation excess runoff across larger areas of the catchment, potentially making it the major runoff mechanism. In other words, higher GW levels would not only be related to more baseflow but when close to the surface they could also influence saturation excess runoff and interflow. Thus, our results suggest the existence of a balance between catchment hilliness and villainess that yields the strongest SF-GW interaction. Therefore, not necessarily the river flowing through the steeper catchment will present the stronger contribution of GW to SF.

In summary, the CCM results suggests that (i) GW head in the regional porous/productive aquifer of Brucknell Creek is driving SF more strongly than in the local fractured aquifers of Sunday and Ford Creek and that (ii) there is significant variability in SF-GW causality strength and direction within a catchment. Bores in the valley bottoms tend to be more causal on SF, and this interaction can go either way (GW causal on flow, or vice versa). These insights also suggest that CCM could allow more granular and rigorous testing of perceptual models and

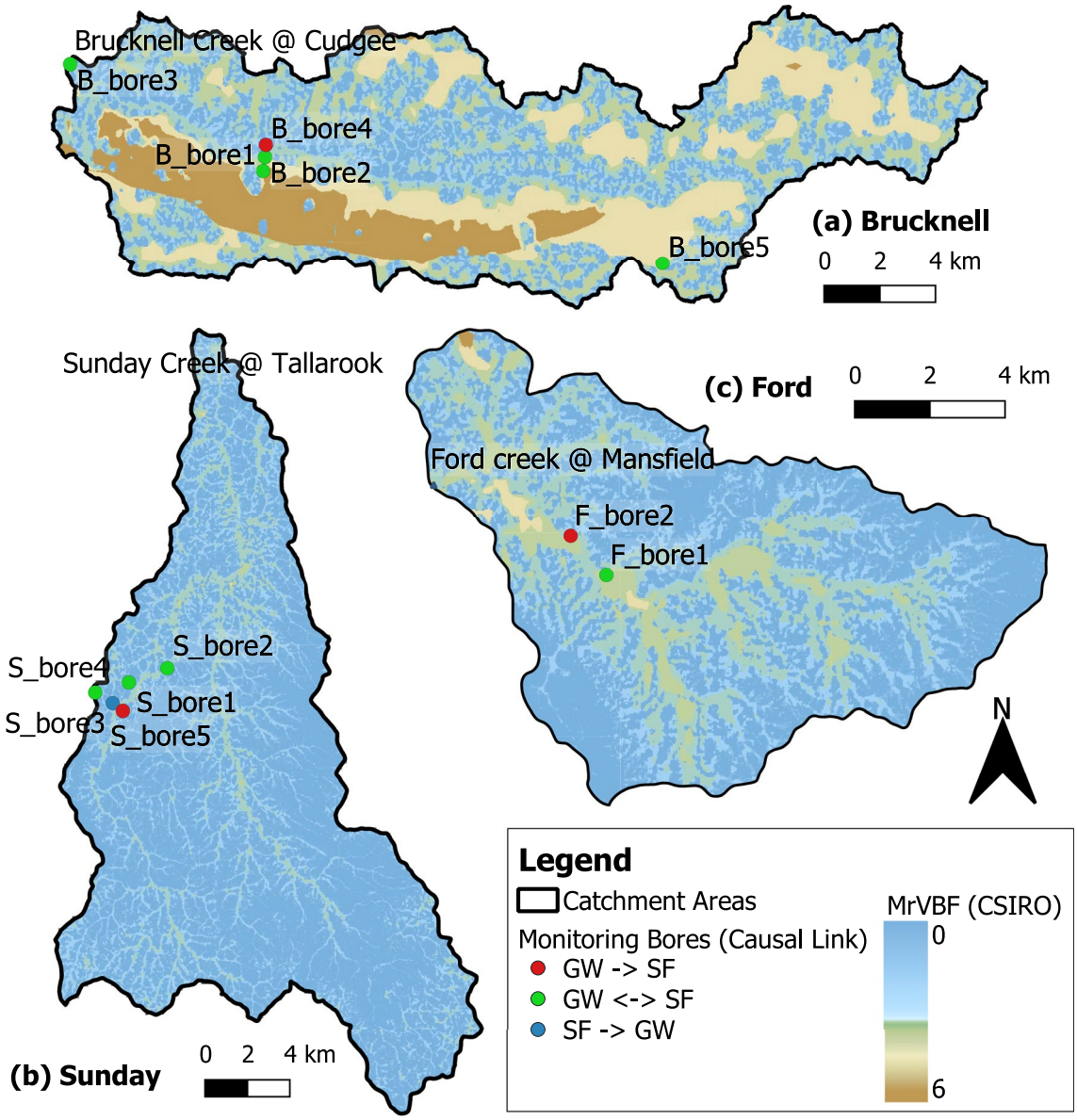


Figure 8. One-second Multi-resolution Valley Bottom Flatness (MrVBF) for (a) Brucknell, (b) Sunday, and (c) Ford Creek, along with the location of the bores used in this study color coded according to its daily causal link. High MrVBF values refer to valley bottoms and low values indicate steep and erosional areas. MrVBF data set developed by CSIRO and extracted from https://data.gov.au/data/dataset.

contribute to the structural design of hydrologic models and reduce the uncertainty of their projections (Anderson & Woessner, 1992; Beven, 1993; Blöschl et al., 2019; Gallart et al., 2007; Konikow & Bredehoeft, 1992; Oreskes et al., 1994).

4.2. Exploring Limitations of CCM

While the results described above are insightful, CCM has limitations that reportedly produce a high false positive rate when identifying causal links in hydrology (Delforge et al., 2022; Ombadi et al., 2020). Hence we have explored the role of seasonality, sampling frequency, and long-term trends in our results. The significance of these aspects on the CCM results are discussed in the sections below.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.2.1. The Role of Seasonality

As mentioned in Section 4.1, the meteorological forcings (PET, rainfall) are not remarkably different between the study areas, however, together with SF and, to a lesser degree, GW head, they present remarkable seasonality (Figures S1, S2, S3, and S6 in Supporting Information S1), which is problematic for causal inference studies as cycles are both deterministic and predictable and are therefore a likely confounding factor. In addition, CCM is a bi-variate causality framework and hence the identified causal links might be the result of a third (or more) strong common forcing, a hypothesis that is inevitable in hydrology when studying SF and GW interactions. Both SF and GW are forced by rainfall and PET (and their seasonality) and hence some degree of association might always be present. In fact, a synthetic study has shown that disconnected SF and GW reservoirs could always exhibit CCM convergence due to the common meteorological forcing (Delforge et al., 2022). This issue is known as the Moran effect or Reichenbach principle of the common cause (Moran, 1953; Reichenbach, 1956). Both the impact of seasonality and Moran effect on the significance of our results was evaluated by comparing the CCM results for the real data against those obtained for surrogate time-series that preserved the seasonality and shuffled the residuals of the original hydrographs. The results of this analysis are presented in Figure 6.

Regarding the null hypothesis of a third common cause, the considerable variability in causal direction and strength within and between catchments and the significance of the CCM results (Figures 3–7) suggest that it would not be a major issue, particularly for the daily analysis. If the CCM results obtained from the original hydrographs were mainly due to a strong common forcing, we would expect homogeneous results across all sites with CCM presenting convergence and significance in a similar fashion. This expectation derives from the fact that the catchments are relatively small and the climate forcing is fairly similar across them. At the same time, the fact that eight out of the twelve causal links in Table 2 are bi-directional raises the hypothesis that the variables could be strongly synchronized. With a strong common forcing or a strong driver (i.e., GW strongly driving SF), the dynamics of the forced variable could become subordinate to the forcing and CCM would indicate a bi-directional link when in fact the interaction is unidirectional. The lagged version of CCM (Ye et al., 2015) could help resolve this issue of synchrony, but it was outside the scope of this study that assumed a contemporaneous causality given the higher cross-correlation at zero time lag in our data.

Similarly, one can notice that the cross-map skill of both real and surrogate hydrographs was generally higher in Brucknell Creek (Figure 6a) than in Sunday and Ford Creek (Figures 6b and 6c), even more so when looking at the monthly CCM results, when seasonality becomes more pronounced in the hydrographs (Figures S1 and S6 in Supporting Information S1). In fact, the results in Figures 3–6 suggest that the monthly analysis often produces cross-map skills higher than for the daily data. This indicates that increased variance captured by the seasonal cycle can strengthen the SF-GW interaction suggested by CCM. Therefore, it is possible that the differences in cross-map skill between catchments is related to the different degrees of variance explained by seasonality, particularly the higher seasonality of SF and GW head in Brucknell Creek (Table 1, Figure S6a in Supporting Information S1), and not to different processes that emerge from the topography and hydrogeology of the study areas. Ultimately, seasonality hinders the significance and the physical interpretability of the causal links suggested by the CCM analysis.

4.2.2. The Role of Data Length and Sampling Frequency

The issue of data length in CCM was explicitly explored by Ombadi et al. (2020) which concluded that CCM is applicable to short daily time-series (<100 days). The results from Delforge et al. (2022) also converge to the same conclusion and suggest that convergence is not worth checking for daily hydrological time-series. However, as mentioned, DST assumes infinite, noise and error free datasets, even though different authors suggest different lengths of data as the minimum to properly reconstruct a system manifold (Kurths & Herzel, 1987; Nerenberg & Essex, 1990; Procaccia, 1988; Smith, 1988; Sugihara et al., 2012). Moreover, the dynamics of the system can vary according to the time scale.

Therefore, longer observation time scales, as in the case of our monthly time-series, could affect not only the significance of the results in relation to both stochastic replicates and cross-correlation (due to shorter time-series), but also the ability to properly reconstruct the state-space that underpins the method. Although the general convergence of our CCM results suggest that short daily hydrological time-series are likely appropriate (Delforge et al., 2022; Ombadi et al., 2020), some monthly curves did not finish converging, particularly, those for GW cross-mapping SF in bores S3, S4, F1, F2. Even though these are significant convergence values, they

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

illustrate the effect of data length and the spatial variability of convergence rate from site to site. On the same note, difficulties to unfold the dynamics of hydrological processes at longer time scales is reported in the literature. For instance, the seminal work of Rodriguez-Iturbe et al. (1989) managed to unfold the chaotic dynamics of a rainfall event using 1990 data points of 15 s interval but the same was not possible for 148 years of weekly rainfall data (around 7,696 data points), even though the authors suggested the possibility of a weekly chaotic behavior in a large dimensional space.

The issue of sampling frequency and data length can be exacerbated by the application of CCM to the entire SF hydrograph instead of using only the baseflow. The higher dimensionality or randomness of the daily flood hydrograph can be harder to cross-map from GW head because of its smoother and more deterministic behavior, possibly even more so due to the paucity of the observed GW head and the use of interpolated data. These aspects could be causing the CCM assimetry between the monthly and daily analysis, that is, monthly data generally producing higher cross-map skills. Using baseflow hydrographs could mitigate this issue, however, not only it is hard to estimate baseflow (Su et al., 2016) but it directly represents how much of the SF is coming from GW which can defeat the purpose of using a casual inference method. Additionally, GW head dynamics are potentially linked to runoff processes, for instance, by activating saturation excess runoff, muting the response to rainfall and possibly contributing to changes in RR states. Therefore, applying CCM to the entire SF hydrograph tests the link between GW head and all these processes, not only low-flow dynamics. Anyway, baseflow filters can and should be used to complement and interpret the CCM results, as done in this study. To summarize, the issues of sampling frequency and data length remain sources of uncertainty in this study and highlights the need to have longer, better, and diverse hydrological monitoring to substantiate insights derived from CCM analysis.

4.2.3. The Role of Long-Term Trends and Climatic Conditions

To examine the role of long-term trends jointly with different climatic conditions, Figure 7 and Table 3 show the daily CCM analysis for bores S_bore1 and S_bore2 in Sunday Creek for three subperiods emcompassing different climatic conditions and different trends in both SF and GW head hydrographs. Figure 7 also shows the original data emphasizing these subperiods.

For the first period, from 1975 to 1985 (Figure 7a), the GW head hydrographs for both bores and the SF had little trend during a relatively wet period. Even though the trends were similar in this period, the CCM results were quite different for the two observations bores. For S_bore1, the CCM analysis indicates a unidirectional causal link where SF was driving GW head, which corroborates with the signatures typical of strong recharge events that are present in the GW head hydrograph. Whereas for S_bore2, CCM indicates a relatively weaker bidirectional causal link that supports the smoother and more seasonal characteristics of its GW head hydrograph.

For the second period, from 1985 to 1996 (Figure 7b), the GW head hydrographs presented remarkably different trends, where S_bore1 displayed little trend while S_bore2 experienced a strong upward trend. In this period, SF had little trend while the climate was relatively wet in comparison to the subsequent Millennium Drought. Although S_bore1 and S_bore2 had very different trends in their hydrographs, the CCM results were rather similar and indicated a unidirectional causal link where GW head was driving SF, suggesting a shift to a dominantly gaining stream before the Millennium Drought (van Dijk et al., 2009).

For the last period, from 1996 to 2013 (Figure 7c), hence during and after the Millennium Drought, the GW head hydrographs for both bores and the SF presented similar downward trends, which would change only in 2010 with the end of the drought. Both bores presented similar trends in the hydrographs and the CCM results indicate a bidirectional interaction between SF and GW head, where SF drives GW head more strongly in S_bore1 than S_bore2. Potentially, S_bore1 had a stronger connection to the stream during the recharge event observed in 2010 toward the end of the drought. However, this contradicts the sharper GW rise observed in S_bore2 during the same period.

As mentioned, pronounced trends can be problematic for CCM as it violates the recurrence hypothesis that underpins the idea of identifying causality by checking if shadow manifolds behave consistently while revisiting similar states. Hence, the differences between the CCM results for the entire observation period and for the three subperiods could primarily be caused by different trends in the data. Trends can also produce artificially high cross-map skills because of temporal autocorrelation rather than recurrence, which was not controlled in this study with a Theiler window (Theiler, 1986). This parameter (exclusion radius in the rEDM package) ignores from the universe of possible nearest neighbors the points whose time indices are too close to the predicted. As Sunday

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Creek has generally lower cross-map skills than in Brucknell and Ford Creek (Figures 3–,5 and 7), it seems that trends in its GW data (Figure S5 in Supporting Information S1) are potentially diminishing CCM's predictive skill. To help solve this conundrum, a Theiler window could be systematically applied to investigate the role of autocorrelation and the current data could be detrended before the application of CCM. However, this could make the results questionable as we know that trends in nature are real (Gleick, 1987) and filtering the data can change its dimensionality (Badii et al., 1988; Chennaoui et al., 1990; Jayawardena & Lai, 1994; Mpitsos et al., 1988). Hence, the application of CCM in Sunday Creek would require longer time-series to confirm whether the trends are part of a long time scale dynamics and would disappear into the future.

At the same time, one can look at the results in light of the different climatic conditions during the three subperiods. Particularly, the CCM results suggest weaker SF-GW interactions, regardless of the direction, during and after the Millennium Drought (1996–2013) than in the previous wet periods. At first glance, this seems inconsistent with the general understanding that GW (baseflow) should contribute more significantly to SF during droughts. However, this long drought seems have strongly affected both GW and SF (Fowler et al., 2022) and produced generally lower SF-GW connectivity triggered by declining GW levels and soil moisture (Grayson et al., 1997; Hughes et al., 2012; Petrone et al., 2010). This behavior would be consistent with the observation of similar SF hydrograph shapes (Trotter et al., 2022) and lower RR ratios during and after the Millennium Drought (Peterson et al., 2021; Saft, Peel, Western, Perraud, & Zhang, 2016; Saft, Peel, Western, & Zhang, 2016). Additionally, between 1975 and 1985 the GW head in S_bore1 and S_bore2 presents remarkably different trends, if not opposite, while the CCM results are quite similar. This raises the question of whether the long-term trends are strong enough to be impacting the results.

In summary, although it is possible to interpret the CCM results in light of the different climatic conditions, care must be taken while physically interpreting the results because of the impact of long-term trends in CCM. Thus, future research is necessary to investigate the role of trends using a larger and diverse sample of time-series with long trends under different climatic conditions while systematically applying a Theiler window.

5. Concluding Remarks

We have examined the potential and some limitations of CCM to identify causal links between SF and GW head in three upland catchments located in different hydrogeologic settings using multiyear daily and monthly time-series encompassing both wet and dry periods. Overall, this study shows that CCM has the potential to unlock valuable causal information from SF and GW head hydrographs that would not be evident from cross-correlation in our data. It allowed to test an a priori perceptual understanding of processes in the study areas and offered additional insights about hydrological processes, both spatially and temporally. Specifically, the CCM results suggest that: (a) GW head in the regional porous/productive aquifer is driving SF more strongly than in the local fractured aquifers; (b) there is significant variability in SF-GW causality strength and direction within a catchment; (c) Bores in the valley bottoms tend to be more causal on SF, and this interaction can go either way (GW causal on flow, or vice versa); (d) SF-GW interactions were weaker during and after the Millennium Drought than in the previous wet periods possibly due to weaker hydraulic connectivity. This is important to understand catchment dynamics as CCM offered more granular insights that would not be evident from cross-correlation. Given the often-limited resources for environmental monitoring, we argue that commonly available monitoring data and CCM can help to guide and optimize the implementation of detailed fieldwork for different purposes, for instance, in water contamination studies. At the same time, care must be taken while physically interpreting the causal links suggested by this method, as CCM has limitations that can lead to spurious causality.

This paper investigated how the CCM results were affected by seasonality, data frequency, and long-term trends. Besides the use of surrogate data to test our results more rigorously, these aspects impact both the significance and the physical interpretability of the causal links. This corroborates with the high false positive rate of CCM in the hydrology literature (Delforge et al., 2022; Ombadi et al., 2020). Therefore, future research is needed to further investigate these and other issues, such as the role of data error and intermittency, ideally using a large and heterogeneous sample of catchments and time-series. Ultimately, it is hard to prove causality from statistical methods and they should be coupled with detailed field monitoring to further substantiate the relationships suggested by CCM or other CIMs.

6. Perspectives on CCM in Hydrology

Finally, the application of dynamic system theory to identify causality in hydrology remains limited. Therefore, we would like to touch upon some perspectives that shall represent exciting interdisciplinary research opportunities, for instance, among hydrologists, mathematicians, and meteorologists.

6.1. How Does CCM Compare to Other Causal Inference Frameworks to Infer SF-GW Causality at Different Climatic, Hydrogeologic, and Topographic Settings?

CCM was particularly developed to account for state-dependent nonlinear interactions (Runge et al., 2019; Sugihara et al., 2012), however, it can present high false positive rates (Delforge et al., 2022; Ombadi et al., 2020). Therefore, using large and diverse datasets of catchments/hydrographs which could be clustered according to hydrological similarity (Papacharalampous et al., 2021; Rinderer et al., 2019), how would the CCM results compare with other nonlinear CIMs, such as Transfer Entropy (Schreiber, 2000) and multivariate PC algorithm (Spirtes et al., 2000)? Would the causal links remain attributable to differences in hydrogeology, topography, and climate?

6.2. How Sensitive Is the Significance of the CCM Results to the Stochastic Replication Technique and Other Adjustable Parameters of CCM?

Here we used standard CCM and a stochastic replication technique in rEDM; however, using a global sensitivity analysis approach, future research should investigate the sensitivity of the results to stochastic techniques commonly used in hydrology and also to adjustable parameters of CCM, such as the prediction horizon (lagged version of CCM), the embedding delay (tau), and the Theiler window to help understand confounding factors (e.g., seasonality, strong synchrony) and the role of data autocorrelation.

6.3. What Is the Role of Seasonal Dynamics in the Long-Term SF-GW Interactions?

Given the complex RR response in the study areas during the Millennium Drought, we analyzed the average long-term causal interactions between SF and GW by using the entire time-series. However, a strong seasonal dynamic shall be present in the long-term interactions. This could be further investigated with a CCM analysis focused on different seasons. For instance, is GW more significant to SF in dry months or in the wet season?

6.4. What Is the Role of Droughts and Persistent Shifts in Rainfall-Runoff on the CCM Results?

We analyzed three catchments during and after the Millennium Drought. From our daily CCM results, the catchments with higher shifts in runoff where those with the lower SF-GW interactions. However, there are many more catchments in Australia that shifted and have been presenting different recovery behaviors. Therefore, looking at a larger sample size, could the shift in runoff and persistence in a lower state affect the CCM results due to long-term trends in the hydrographs?

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Streamflow and groundwater data used in this study were from the Water Measurement Information System of DELWP, https://data.water.vic.gov.au/. Climate data were from the AWAP project, www.bom.gov.au/jsp/awap/. The computational resources used in the data analysis were kindly provided by Spartan, the high-performance computational cluster of the University of Melbourne (Lafayette et al., 2016). We thankfully acknowledge the valuable comments from five anonymous reviewers, whose feedback greatly improved the manuscript.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Acknowledgments
Giancarlo Bonotto was supported by the Melbourne Research Scholarship, Australian Government Research Training Program Scholarship, Melbourne School of Engineering Travelling Scholarship, DELWP, and Goulburn Broken Catchment Management Authority (TP 707158). The Australian Research Council supported Keirnan Fowler (LP170100598) and Tim Peterson (LP180100796).

References

Ali, G., Tetzlaff, D., Mcdonnell, J. J., Soulsby, C., Carey, S., Laudon, H., et al. (2015). Comparison of threshold hydrologic response across northern catchments. *Hydrological Processes*, *29*(16), 3575–3591. https://doi.org/10.1002/hyp.10527

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration — guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56.

Anderson, M. P., & Woessner, W. W. (1992). The role of the postaudit in model validation. *Advances in Water Resources*, *15*(3), 167–173. https://doi.org/10.1016/0309-1708(92)90021-S

Atkinson, A. P., Cartwright, I., Gilfedder, B. S., Hofmann, H., Unland, N. P., Cendón, D. I., & Chisari, R. (2015). A multi-tracer approach to quantifying groundwater inflows to an upland river; assessing the influence of variable groundwater chemistry. *Hydrological Processes*, *29*, 1–12. https://doi.org/10.1002/hyp.10122

Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., & Rubio, M. A. (1988). Dimension increase in filtered chaotic signals. *Physical Review Letters*, *60*(11), 979–982. https://doi.org/10.1103/PhysRevLett.60.979

Barthel, R. (2014). HESS Opinions "integration of groundwater and surface water research: An interdisciplinary problem?". *Hydrology and Earth System Sciences*, *18*(7), 2615–2628. https://doi.org/10.5194/hess-18-2615-2014

Barthel, R., & Banzhaf, S. (2016). Groundwater and surface water interaction at the regional-scale – a review with focus on regional integrated models. *Water Resources Management*, *30*, 1–32. https://doi.org/10.1007/s11269-015-1163-z

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. https://doi.org/10.1016/0309-1708(93)90028-E

Bishop, K., Seibert, J., Nyberg, L., & Rodhe, A. (2011). Water storage in a till catchment. II: Implications of transmissivity feedback for flow paths and turnover times. *Hydrological Processes*, *25*, 3950–3959. https://doi.org/10.1002/hyp.8355

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., et al. (2019). Twenty-three unsolved problems in hydrology (UPH)–a community perspective. *Hydrological Sciences Journal*, *64*(10), 1141–1158. https://doi.org/10.1080/02626667.2019.1620507

Brodie, R. S., Kilgour, B., Jacobson, G., & Lau, J. E. (1998). Hydrogeology map of Australia.

Burt, T. P., & McDonnell, J. J. (2015). Whither field hydrology? The need for discovery science and outrageous hydrological hypotheses. *Water Resources Research*, *51*(8), 5919–5928. https://doi.org/10.1002/2014WR016839

Cartwright, I., & Gilfedder, B. (2015). Mapping and quantifying groundwater inflows to Deep Creek (Maribyrnong catchment, SE Australia) using 222Rn, implications for protecting groundwater-dependant ecosystems. *Applied Geochemistry*, *52*, 118–129. https://doi.org/10.1016/j.apgeochem.2014.11.020

Cartwright, I., Gilfedder, B., & Hofmann, H. (2014). Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers. *Hydrology and Earth System Sciences*, *18*(1), 15–30. https://doi.org/10.5194/hess-18-15-2014

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 Discussions*, *12*, 5427–5463. https://doi.org/10.5194/hessd-12-5427-2015

Chennaoui, A., Pawelzik, K., Liebert, W., Schuster, H. G., & Pfister, G. (1990). Attractor reconstruction from filtered chaotic time series. *Physical Review A*, *41*(8), 4151–4159. https://doi.org/10.1103/physreva.41.4151

Clark, M. P., Kavetski, D., & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, *47*(9), 1–16. https://doi.org/10.1029/2010WR009827

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., et al. (2015a). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, *51*(4), 2498–2514. https://doi.org/10.1002/2015WR017198

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., et al. (2015b). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*, *51*(4), 2515–2542. https://doi.org/10.1002/2015WR017200

CSIRO. (2010). Climate variability and change in south-eastern Australia: A synthesis of findings from phase 1 of the south eastern Australian climate initiative (SEACI), climate variability and change in south-eastern Australia: A synthesis of findings from phase 1 of the south eastern Australian climate initiative (SEACI).

Dahlhaus, P., Heislers, D., & Dyson, P. (2002). *Glenelg Hopkins CMA groundwater flow systems*. Buninyong.

Davtalab, R., Mirchi, A., Khatami, S., Gyawali, R., Massah, A., Farajzadeh, M., & Madani, K. (2017). Improving continuous hydrologic modeling of data-poor river basins using hydrologic engineering center's hydrologic modeling system: Case Study of Karkheh River Basin. *Journal of Hydrologic Engineering*, *22*(8), 05017011. https://doi.org/10.1061/(asce)he.1943-5584.0001525

Delforge, D., De Viron, O., Vanclooster, M., Van Camp, M., & Watlet, A. (2022). Detecting hydrological connectivity using causal inference from time series: Synthetic and real karstic case studies. *Hydrology and Earth System Sciences*, *26*(8), 2181–2199. https://doi.org/10.5194/hess-26-2181-2022

Delforge, D., Muñoz-Carpena, R., Van Camp, M., & Vanclooster, M. (2020). A parsimonious empirical approach to streamflow recession analysis and forecasting. *Water Resources Research*, *56*(2), 1–19. https://doi.org/10.1029/2019WR025771

Deyle, E. R., & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS One*, *6*(3), e18295. https://doi.org/10.1371/journal.pone.0018295

Doody, T. M., Barron, O. V., Dowsley, K., Emelyanova, I., Fawcett, J., Overton, I. C., et al. (2017). Continental mapping of groundwater dependent ecosystems: A methodological framework to integrate diverse data and expert opinion. *Journal of Hydrology: Regional Studies*, *10*, 61–81. https://doi.org/10.1016/j.ejrh.2017.01.003

Erkyihun, S. T., Rajagopalan, B., Zagona, E., Lall, U., & Nowak, K. (2016). Wavelet-based time series bootstrap model for multidecadal streamflow simulation using climate indicators. *Water Resources Research*, *52*(5), 4061–4077. https://doi.org/10.1002/2016WR018696

FAO. (2017). *The future of food and agriculture: Trends and challenges*. United Nations.

Fenicia, F., Kavetski, D., & Savenije, H. H. G. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, *47*(11), 1–13. https://doi.org/10.1029/2010WR010174

Fleckenstein, J. H., Krause, S., Hannah, D. M., & Boano, F. (2010). Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics. *Advances in Water Resources*, *33*(11), 1291–1295. https://doi.org/10.1016/j.advwatres.2010.09.011

Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., et al. (2020). Many commonly used rainfall-runoff models lack long, slow dynamics: Implications for runoff projections. *Water Resources Research*, *56*(5). https://doi.org/10.1029/2019WR025286

Fowler, K., Peel, M., Peterson, T., Saft, M., Band, L., Bonotto, G., et al. (2022). Explaining changes in rainfall-runoff relationships during and after Australia's Millennium Drought. *Hydrology and Earth System Sciences Discussions*, 1–56.

Fraser, A. M. (1989). Information and entropy in strange attractors. *IEEE Transactions on Information Theory*, *35*(2), 245–262. https://doi.org/10.1109/18.32121

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinatesfor strange attractors from mutual information. *Physical Review A*, *33*(2), 1134–1140. https://doi.org/10.1103/PhysRevA.33.1134

Gallant, J. C., & Dowling, T. I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, *39*(12). https://doi.org/10.1029/2002WR001426

Gallart, F., Latron, J., Llorens, P., & Beven, K. (2007). Using internal catchment information to reduce the uncertainty of discharge and baseflow predictions. *Advances in Water Resources*, *30*(4), 808–823. https://doi.org/10.1016/j.advwatres.2006.06.005

Gilfedder, B. S., Hofmann, H., & Cartwright, I. (2013). Novel instruments for in situ continuous Rn-222 measurement in groundwater and the application to river bank infiltration. *Environmental Science & Technology*, *47*(2), 993–1000. https://doi.org/10.1021/es3034928

Gleick, J. (1987). *Chaos: Making a new science*. Viking Penguim Inc.

Goodwell, A. E., Jiang, P., Ruddell, B. L., & Kumar, P. (2020). Debates—Does information theory provide a new paradigm for Earth science? Causality, interaction, and feedback. *Water Resources Research*, *56*(2), 1–12. https://doi.org/10.1029/2019WR024940

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Grayson, R. B., Western, A. W., Chiew, F. H. S., & Blöschl, G. (1997). Preferred states in spatial soil moisture patterns: Local and nonlocal controls. *Water Resources Research*, *33*(12), 2897–2908. https://doi.org/10.1029/97WR02174

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

Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H., & Weiler, M. (2013). Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research*, *49*(6), 3345–3358. https://doi.org/10.1002/wrcr.20229

Hughes, J. D., Petrone, K. C., & Silberstein, R. P. (2012). Drought, groundwater storage and stream flow decline in southwestern Australia. *Geophysical Research Letters*, *39*(3), 1–6. https://doi.org/10.1029/2011GL050797

Hulsman, P., Hrachowitz, M., & Savenije, H. H. G. (2021). Improving the representation of long-term storage variations with conceptual hydrological models in data-scarce regions. *Water Resources Research*, *57*(4), 1–31. https://doi.org/10.1029/2020WR028837

IPCC. (2014). Climate change 2014 synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change.

Ivancevic, V. G., & Ivancevic, T. T. (2008). *Complex nonlinearty: Chaos, phase transitions, topology change, and path integrals*. Springer Series: Understaning Complex Systems.

Jayawardena, A. W., & Lai, F. (1994). Analysis and prediction of chaos in rainfall and stream flow time series. *Journal of Hydrology*, *153*(1–4), 23–52. https://doi.org/10.1016/0022-1694(94)90185-6

Jiang, P., & Kumar, P. (2019). Using information flow for whole system understanding from component dynamics. *Water Resources Research*, *55*(11), 8305–8329. https://doi.org/10.1029/2019WR025820

Kavetski, D., & Fenicia, F. (2011). Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research*, *47*(11), 1–19. https://doi.org/10.1029/2011WR010748

Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and flux mapping: A new approach to model evaluation and process representation under uncertainty. *Water Resources Research*, *55*(11), 8922–8941. https://doi.org/10.1029/2018WR023750

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), 1–5. https://doi.org/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. https://doi.org/10.1029/2008WR006912

Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., & Woods, R. A. (2018). Modular assessment of rainfall-runoff models toolbox - Matlab code for 46 conceptual hydrologic models 1–26. https://doi.org/10.5194/gmd-2018-332

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), 1–23. https://doi.org/10.1029/2019wr025975

Konikow, L. F., & Bredehoeft, J. D. (1992). Ground-water models cannot be validated. *Advances in Water Resources*, *15*(1), 75–83. https://doi.org/10.1016/0309-1708(92)90033-X

Krause, S., Hannah, D. M., Fleckenstein, J. H., Heppell, C. M., Kaeser, D., Pickup, R., et al. (2011). Inter-disciplinary perspectives on processes in the hyporheic zone. *Ecohydrology*, *130*(4), 481–499. https://doi.org/10.1002/eco176

Kurths, J., & Herzel, H. (1987). An attractor in a solar time series. *Physica D: Nonlinear Phenomena*, *25*(1–3), 165–172. https://doi.org/10.1016/0167-2789(87)90099-6

Lafayette, L., Sauter, G., Vu, L., & Meade, B. (2016). Spartan performance and flexibility: An HPC-cloud chimera. In *OpenStack summit*. https://doi.org/10.4225/49/58ead90dceaaa

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Lyne, V., & Hollick, M. (1979). Stochastic time variable rainfall-runoff modelling. In *Proceedings of the hydrology and water resources symposium* (pp. 89–92). Institution of Engineers.

McDonnell, J. J. (2014). The two water worlds hypothesis: Ecohydrological separation of water between streams and trees? *Wiley Interdisciplinary Reviews: Water*, *1*(4), 323–329. https://doi.org/10.1002/wat2.1027

Medina, M., Huffaker, R., Jawitz, J. W., & Muñoz-Carpena, R. (2019). Nonlinear dynamics in treatment wetlands: Identifying systematic drivers of nonequilibrium outlet concentrations in everglades STAs. *Water Resources Research*, *55*(12), 11101–11120. https://doi.org/10.1029/2018WR024427

Moran, P. A. P. (1953). The statistical analysis of the Canadian lynx cycle. *Australian Journal of Zoology*, *1*(3), 291–298. https://doi.org/10.1071/ZO9530291

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

Mpitsos, G. J., Creech, H. C., Cohan, C. S., & Mendelson, M. (1988). Variability and chaos: Neurointegrative principles in self-organization of motor patterns. *Dynamic Patterns in Complex Systems - World Scientific*, 162–190.

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for baseflow and recession analyses. *Water Resources Research*, *26*(7), 1465–1473. https://doi.org/10.1029/90wr00136

National Climate Centre - Bureau of Meteorology. (2011). Frequent heavy rain events in late 2010/early 2011 lead to widespread flooding across eastern Australia, Special Climate Statement 24.

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Nerenberg, M. A. H., & Essex, C. (1990). Correlation dimension and systematic geometric effects M. *Physical Review A*, *42*(12), 7065–7074. https://doi.org/10.1103/physreva.42.7065

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. l. (2020). Evaluation of methods for causal discovery in hydrometeorological systems. *Water Resources Research*, *56*(7). https://doi.org/10.1029/2020WR027251

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, *263*(5147), 641–646. https://doi.org/10.1126/science.263.5147.641

Papacharalampous, G., Tyralis, H., Papalexiou, S. M., Langousis, A., Khatami, S., Volpi, E., & Grimaldi, S. (2021). Global-scale massive feature extraction from monthly hydroclimatic time series: Statistical characterizations, spatial patterns and hydrological similarity. *Science of the Total Environment*, *767*, 144612. https://doi.org/10.1016/j.scitotenv.2020.144612

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Peel, M. C., McMahon, T. A., Finlayson, B. L., & Watson, F. G. R. (2001). Identification and explanation of continental differences in the variability of annual runoff. *Journal of Hydrology*, *250*(1–4), 224–240. https://doi.org/10.1016/S0022-1694(01)00438-3

Peterson, T., Wasko, C., Saft, M., & Peel, M. (2019). AWAPer: An R package for area weighted catchment daily meteorological data anywhere within Australia. *Hydrological Processes*, *34*(5), 1–6. https://doi.org/10.1002/hyp.13637

Peterson, T. J. (2009). *Multiple hydrological steady states and resilience. PhD thesis, Engineering - Civil and environmental engineering*. The University of Melbourne.

Peterson, T. J., Saft, M., Peel, M. C., & John, A. (2021). Watersheds may not recover from drought. *Science*, *372*(6543), 745–749. https://doi.org/10.1126/science.abd5085

Peterson, T. J., & Western, A. W. (2014a). Multiple hydrological attractors under stochastic daily forcing: 1. Can multiple attractors exist? *Water Resources Research*, *50*(4), 1–17. https://doi.org/10.1002/2012WR013004.Key

Peterson, T. J., & Western, A. W. (2014b). Nonlinear time-series modeling of unconfined groundwater head. *Water Resources Research*, *50*(10), 1–26. https://doi.org/10.1002/2013WR014800

Peterson, T. J., & Western, A. W. (2018). Statistical interpolation of groundwater hydrographs. *Water Resources Research*, *54*(7), 1–18. https://doi.org/10.1029/2017WR021838

Peterson, T. J., Western, A. W., & Argent, R. M. (2014). Multiple hydrological attractors under stochastic daily forcing: 2. Can multiple attractors emerge? *Water Resources Research*, *50*(4), 3010–3029. https://doi.org/10.1002/2012WR013004

Peterson, T. J., Western, A. W., & Cheng, X. (2017). The good, the bad and the outliers: Automated detection of errors and outliers from groundwater hydrographs. *Hydrogeology Journal*, *26*(2), 371–380. https://doi.org/10.1007/s10040-017-1660-7

Petrone, K. C., Hughes, J. D., Van Niel, T. G., & Silberstein, R. P. (2010). Streamflow decline in south Western Australia, 1950-2008. *Geophysical Research Letters*, *37*(11), 1–7. https://doi.org/10.1029/2010GL043102

Potter, N. J., Chiew, F. H. S., & Frost, A. J. (2010). An assessment of the severity of recent reductions in rainfall and runoff in the Murray-Darling Basin. *Journal of Hydrology*, *381*(1–2), 52–64. https://doi.org/10.1016/j.jhydrol.2009.11.025

Procaccia, I. (1988). Complex or just complicated? *Nature*, *333*(6173), 498–499. https://doi.org/10.1038/333498a0

Rajagopalan, B., Erkyihun, S. T., Lall, U., Zagona, E., & Nowak, K. (2019). A nonlinear dynamical systems-based modeling approach for stochastic simulation of streamflow and understanding predictability. *Water Resources Research*, *55*(7), 6268–6284. https://doi.org/10.1029/2018WR023650

R Core Team. (2020). R: A language and environment for statistical computing.

Reichenbach, H. (1956). *The direction of time*. University of California Press.

Rimmer, A., & Hartmann, A. (2014). Optimal hydrograph separation filter to evaluate transport routines of hydrological models. *Journal of Hydrology*, *514*, 249–257. https://doi.org/10.1016/j.jhydrol.2014.04.033

Rinderer, M., Ali, G., & Larsen, L. G. (2018). Assessing structural, functional and effective hydrologic connectivity with brain neuroscience methods: State-of-the-art and research directions. *Earth-Science Reviews*, *178*, 29–47. https://doi.org/10.1016/j.earscirev.2018.01.009

Rinderer, M., van Meerveld, H. J., & McGlynn, B. L. (2019). From points to patterns: Using groundwater time series clustering to investigate subsurface hydrological connectivity and runoff source area dynamics. *Water Resources Research*, *55*(7), 5784–5806. https://doi.org/10.1029/2018WR023886

Rodriguez-Iturbe, I., Febres De Power, B., Sharifi, M. B., & Georgakakos, K. P. (1989). Chaos in rainfall. *Water Resources Research*, *25*(7), 1667–1675. https://doi.org/10.1029/WR025i007p01667

Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J. A., Peterson, T. J., Weuthen, A., et al. (2012). Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. *Water Resources Research*, *48*(10), 1–22. https://doi.org/10.1029/2011WR011518

Ross, A. (2012). Easy to say, hard to do: Integrated surface water and groundwater management in the Murray-Darling Basin. *Water Policy*, *14*(4), 709–724. https://doi.org/10.2166/wp.2012.129

Ruddell, B. L., & Kumar, P. (2009). Ecohydrologic process networks: 1. Identification. *Water Resources Research*, *45*(3), 1–23. https://doi.org/10.1029/2008WR007279

Ruelle, D., & Takens, F. (1971). On the nature of turbulence. *Communications in Mathematical Physics*, *20*(3), 167–192. https://doi.org/10.1007/BF01646553

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., et al. (2019). Inferring causation from time series in Earth system sciences. *Nature Communications*, *10*, 1–13. https://doi.org/10.1038/s41467-019-10105-3

Saffarpour, S., Western, A. W., Adams, R., & Mcdonnell, J. J. (2016). Multiple runoff processes and multiple thresholds control agricultural runoff generation. *Hydrology and Earth System Sciences*, *20*(11), 4525–4545. https://doi.org/10.5194/hess-20-4525-2016

Saft, M., Peel, M. C., Western, A. W., Perraud, J. M., & Zhang, L. (2016a). Bias in streamflow projections due to climate-induced shifts in catchment response. *Geophysical Research Letters*, *43*(4), 1574–1581. https://doi.org/10.1002/2015GL067326

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. https://doi.org/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. https://doi.org/10.1002/2014WR015348

Sangoyomi, T. B., Lall, U., & Abarbanel, H. D. I. (1996). Nonlinear dynamics of the Great Salt lake: Dimension estimation. *Water Resources Research*, *32*(1), 149–159. https://doi.org/10.1029/95WR02872

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Scaife, C. I., & Band, L. E. (2017). Nonstationarity in threshold response of stormflow in southern Appalachian headwater catchments. *Water Resources Research*, *53*(8), 6579–6596. https://doi.org/10.1002/2017WR020376

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*(2), 461–464. https://doi.org/10.1103/PhysRevLett.85.461
Seibert, J., Bishop, K., Nyberg, L., & Rodhe, A. (2011). Water storage in a till catchment. I: Distributed modelling and relationship to runoff. *Hydrological Processes*, *25*, 3937–3949. https://doi.org/10.1002/hyp.8309
Seibert, J., Bishop, K., Rodhe, A., & McDonnell, J. J. (2003). Groundwater dynamics along a hillslope: A test of the steady state hypothesis. *Water Resources Research*, *39*, 1–9. https://doi.org/10.1029/2002WR001404
Sendrowski, A., & Passalacqua, P. (2017). Process connectivity in a naturally prograding river delta. *Water Resources Research*, *53*(3), 1841–1863. https://doi.org/10.1002/2016WR019768
Sheffield, J., Wood, E. F., & Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature*, *491*(7424), 435–438. https://doi.org/10.1038/nature11575
Sivakumar, B. (2000). Chaos theory in hydrology: Important issues and interpretations. *Journal of Hydrology*, *227*(1–4), 1–20. https://doi.org/10.1016/S0022-1694(99)00186-9
Sivakumar, B. (2009). Nonlinear dynamics and chaos in hydrologic systems: Latest developments and a look forward. *Stochastic Environmental Research and Risk Assessment*, *23*(7), 1027–1036. https://doi.org/10.1007/s00477-008-0265-z
Sivakumar, B. (2017). *Chaos in hydrology: Bridging determinism and stochasticity*. Springer. https://doi.org/10.1007/978-90-481-2552-4
Sivakumar, B., & Singh, V. P. (2012). Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. *Hydrology and Earth System Sciences*, *16*(11), 4119–4131. https://doi.org/10.5194/hess-16-4119-2012
Smith, L. A. (1988). Intrinsic limits on dimension calculations. *Physics Letters A*, *133*(6), 283–288. https://doi.org/10.1016/0375-9601(88)90445-8
Spirtes, P., Glymour, C. N., Scheines, R., & Heckerman, D. (2000). *Causation, prediction, and search* (2nd ed.). The MIT Press.
Staudinger, M., Stoelzle, M., Cochand, F., Seibert, J., Weiler, M., & Hunkeler, D. (2019). Your work is my boundary condition!: Challenges and approaches for a closer collaboration between hydrologists and hydrogeologists. *Journal of Hydrology*, *571*, 235–243. https://doi.org/10.1016/j.jhydrol.2019.01.058
Su, C. H., Costelloe, J. F., Peterson, T. J., & Western, A. W. (2016). On the structural limitations of recursive digital filters for base flow estimation. *Water Resources Research*, *52*(6), 4745–4764. https://doi.org/10.1002/2015WR018067
Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/science.1227079
Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L. Young (Eds.), *Dynamical systems and turbulence* (pp. 366–381). Springer. https://doi.org/10.1007/BFb0091924
Theiler, J. (1986). Spurious dimension from correlation algorithms applied to limited time-series data. *Physical Review A*, *34*(3), 2427–2432. https://doi.org/10.1103/PhysRevA.34.2427
Trotter, L., Saft, M., Peel, M. C., & Fowler, K. J. A. (2022). Symptoms of performance degradation during multi-annual drought: A large-sample, multi-model study. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10509844.3
Tsonis, A. A. (1992). *Chaos: From theory to applications*. Springer Science+Business Media. https://doi.org/10.1007/978-1-4615-3360-3
Tsonis, A. A. (2018). Advances in nonlinear geosciences. *Advances in Nonlinear Geosciences*. https://doi.org/10.1007/978-3-319-58895-7
Tsonis, A. A., Deyle, E. R., Ye, H., & Sugihara, G. (2018). Convergent cross mapping: Theory and an example. In *Advances in nonlinear geosciences*. https://doi.org/10.1007/978-3-319-58895-7
U.N. (2017). World population prospects: The 2017 revision, key findings and advance tables. Working Paper No. ESA/P/WP/248.
van Dijk, A., Renzullo, L., King, E., & Guerschman, J.-P. (2009). The Australian water resources assessment system. Blending water cycle observations and models at local to continental scale. In *GEWEX/ILEAPS water in a changing climate*. GEWEX/iLEAPS. Retrieved from http://hdl.handle.net/102.100.100/112808?index=1
Van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., De Jeu, R. A. M., Liu, Y. Y., Podger, G. M., et al. (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. https://doi.org/10.1002/wrcr.20123
Walker, G., Gilfedder, M., Evans, W. R., Dyson, D., & Stauffacher, M. (2003). Groundwater flow systems framework.
Wang, Y., Yang, J., Chen, Y., De Maeyer, P., Li, Z., & Duan, W. (2018). Detecting the causal effect of soil moisture on precipitation using convergent cross mapping. *Scientific Reports*, *8*, 1–8. https://doi.org/10.1038/s41598-018-30669-2
Wasko, C., & Nathan, R. (2019). Influence of changes in rainfall and soil moisture on trends in flooding. *Journal of Hydrology*, *575*, 432–441. https://doi.org/10.1016/j.jhydrol.2019.05.054
Ye, H., Clark, A., Deyle, E., & Sugihara, G. (2020). *rEDM: an R package for empirical dynamic modeling and convergent cross-mapping* (pp. 1–53). R-Project Org. https://doi.org/10.18637/jss.v000.i00
Ye, H., Deyle, E. R., Gilarranz, L. J., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, *5*, 1–9. https://doi.org/10.1038/srep14750

19447973, 2022, 8, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021WR030231 by Readcube (Labtiva Inc.), Wiley Online Library on [20/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License