Causality and time-lagged dependencies at the watershed scale

Kalyl Gomes Calixto[a], Jaqueline Vígolo Coutinho[a], Edson Wendland[a]

kalyl.calixto@gmail.com; jaqueline.vigolo@gmail.com; ew@sc.usp.br

[a] São Carlos School of Engineering, University of São Paulo, Av. Trabalhador são-carlense 400, 13566-590, São Carlos SP, Brazil

Abstract

Investigating watershed hydrology from a data-driven causal perspective is an attractive opportunity to characterize and understand relationships between water storages and fluxes. Here we assess integrally how the water balance components interact with themselves, aiming to find relevant time-lags or dependency patterns. Granger's causality test and time-lagged mutual information were used in a pairwise approach to examine cause-effect relationships between precipitation, streamflow, groundwater levels under different land-covers, and evapotranspiration data (daily timescale) from 2009 to 2019 in a Brazilian watershed (52 km²), located in a recharge area of the Guarani Aquifer System. A verification assessment using synthetic datasets shows that the methods are effective to identify the underlying generating mechanisms. Statistically significant causal connections were confirmed in practically all pairs of observed data. Granger's causality indicates that groundwater and streamflow responses are influenced by precipitation even with a lag of 1-day, while evapotranspiration can take more than 200 days to influence groundwater responses, depending on the water table depth and surrounding land-cover. Mutual information curves show dependency patterns between hydrological processes that are different from the ones obtained by cross-correlation functions. The causal analysis provides a complementary view of the hydrological system's functioning and may lead us to develop predictive models that reproduce not only the target variables but also the diverse time-lagged dependencies observed in environmental data.

1 Introduction

Applying causal inference methods in hydrology finds motivation in the complexity of the multiple interactions and interdependencies between hydrological, climatic, environmental, and human systems, in both space and time domains (Kumar, 2015). The analysis of observational data, which are becoming increasingly available from satellite remote sensing, station-based, and field site measurements (Runge et al., 2019), has been one of the most feasible alternatives to investigate hydrological variability and causality (Blöschl et al., 2019; Ombadi et al., 2020). Mechanisms of how the water cycle components are being impacted by natural and human-induced changes, including climate change, land-use changes, and increasing water demands, may be unraveled with causal frameworks (Goodwell et al., 2020). Similarly, the environmental conditions that control hydrologic responses of interest for management purposes (e.g., extreme events) can be better understood (Pelletier and Tucotte, 1997; Dey and Mujumdar, 2018).

Ruddell and Kumar (2009) used transfer entropy (an information-theoretic statistical measure) for quantifying several properties of information flow and interactions between pairs of variables measured at an eddy flux tower. Goodwell and Kumar (2017) investigated complex dependencies and proposed a partitioning method to characterize how two source variables jointly influence a third (target) variable. The results were illustrated with an application to data measured at a weather station. In more recent applications of causal methods in hydrology: Ombadi et al. (2020) compared four methods (Granger causality, transfer entropy, PC algorithm, and convergent cross mapping) and examined pairwise causal relationships in the evapotranspiration process using data from a flux tower; Franzen et al. (2020) characterized time-lagged dependencies between precipitation and streamflow data observed in a large river basin using mutual information; and

Bennett et al. (2019) used a lag 1-day transfer entropy to quantify and compare the intensity of interaction between simulated hydrologic data.

As suggested by the scope of related studies, applications in which the underlying causality mechanisms are reasonably well-known represent opportunities to test and explore causal methods before moving to the analysis of more complex interactions. To date, no studies have focused on an exploratory and comprehensive causal analysis in a monitored watershed aiming to quantify the time-lagged dependencies between precipitation, evapotranspiration, streamflow, groundwater levels, and vegetation indices data, not even in a simple pairwise setup.

As prediction is one of the main goals of science, the presence of properties in time series data that facilitate to predict future terms from past observed behavior and patterns has a fundamental value. Memory and persistence (dependences between past and future states) are properties widely found in hydrologic data. Hurst was the precursor in identifying that the hypothesis of serial independence of hydrological data results in statistical inconsistencies, triggering a series of studies that formalized the theory of long memory processes (Amblard & Michel, 2013; Graves et al., 2017). As practical examples, Zimmermann et al. (2006) and Tomasella et al. (2008) identified significant memory effects in studies related to deforestation and interannual variability of precipitation in the Amazon rainforest, suggesting that the history of land-use and the groundwater system were influencing later hydrological responses. To detect such cause-and-effect relationships from observational data, several methods have been developed (e.g., Granger's method and variations, causal maps, causal networks, algorithms based on information theory) (Runge et al., 2019). The selection of the most appropriate method depends on the prior knowledge of the system, on the nature of the variables involved and on the intended objectives.

The Granger causality test has been one of the most used statistical tools to determine the presence of causal relationships between random variables (Barnett & Seth, 2015). Papagiannopoulou et al. (2017), McGraw and Barnes (2018), Singh and Borrok (2019) and Huang et al. (2019) applied the test to investigate, respectively, the dynamics between climate and vegetation, the climate variability, the association between groundwater reserves and food production on a global scale, and the dependence between climatic and hydrological variables. The advantage of the method proposed by Granger (1969) over traditional methods based on correlations or regressions with lagged explanatory variables is the consideration of the memory effects of the response variables (McGraw & Barnes, 2018). Granger (1969) proposed that a process S causes, in Granger's sense, another process, T, if future values of T can be better predicted using values of S and T instead of just values of T. An assumption of the classic Granger causality test is the linearity premise for the underlying system. Despite that, some studies have reported the capability to detect, with some limitations, even nonlinear interactions (Barnett & Seth, 2014; Ombadi et al., 2020).

Along with the Granger test, methods derived from the information theory (IT) – which are based on the entropy measure (Shannon, 1948) and operate on probability distribution functions (pdf) – are attracting increasing attention in hydrologic research and Earth system sciences (Weijs et al., 2010; Rinderer et al., 2018; Goodwell et al., 2020; Kumar & Gupta, 2020). The principal reasons are found in the capability of information-theoretic methods to provide us with stronger and more robust conclusions with respect to data interaction and connectivity (Ruddell & Kumar, 2009; Goodwell & Kumar, 2017; Jiang & Kumar, 2019). Moreover, the methods do not rely on specific data properties or on the nature of dependencies (linear or

nonlinear). Data limitations – for instance, significant changing behaviors (Ombadi et al., 2020) and insufficient sample lengths (Li et al., 2018) – still hinder the application of IT methods.

Two specific measures based on informational entropy have been widely applied for causal inference: time-lagged mutual information (TLMI) and transfer entropy (TE). The first, TLMI, is a more practicable approach, which measures the general dependence (linear and nonlinear) between two variables (Fraser & Swinney, 1986), and requires considerably shorter sample lengths of data. TLMI is not capable to eliminate data memory effects (like the Granger causality test does) so that static dependencies are not ignored (Li et al., 2018). To address this limitation, Schreiber (2000) proposed the transfer entropy (TE), a conditioned mutual information, which measures the amount of directional information transferred between variables, excluding those memory effects induced by the response (target) time series. Barnett et al. (2009) demonstrated that the TE metric is proportional to the likelihood ratio of the Granger causality test for Gaussian random variables. Despite the advantages, the estimation of TE is still a challenging problem and an active area of research, due to numerical issues, high dimensionality (determined by the number of time lags between the variables), and dependence on accurate estimates of probability distributions (Gençağa, 2018).

Here we explore time-lagged dependencies between hydrologic variables measured in a small watershed. The Granger causality test and the normalized time-lagged mutual information metric (NMI) were selected to perform this study due to their suitability to assess the connectivity and dependency throughout long time windows. We expected to find patterns of interactions and estimate memory time scales associated with the hydrological processes, which may support, directly or not, the predictive modeling and the system characterization. This study represents an opportunity to identify the potentials and limitations of applying causal methods in headwater or

small watersheds, which are widely acknowledged for their importance in the context of water management, and to discuss the innovative or relevant information we can obtain from them.

2 Material and Methods

2.1 Study Area

The study domain is limited to the Onça Creek watershed (OCW) (~65 km²), located in an agricultural area of the state of São Paulo (Brazil) (47°54' – 48°00'W, and 22°09' – 22°15'S) (Figure 1). The watershed entirely lies on a recharge area of the Guarani Aquifer System (GAS), which is one of the most important groundwater reserves in South America, responsible for supplying water to more than 90 million people in Argentina, Brazil, Paraguay, and Uruguay (Araújo et al., 1999; Kirchheim et al., 2019).

The GAS is composed of sandstone layers from the Jurassic (Botucatu Formation) and Triassic (Piramboia Formation) periods and is widely (~ 90%) confined by basaltic spills that occurred in the Cretaceous period (Serra Geral Formation). The study area, instead of presenting such a confining layer, presents a permeable deposit of Cenozoic sediments with a thickness of tens of meters, and a specific yield varying from 0.08 to 0.16 (Wendland et al., 2015; Coutinho et al., 2020).

The watershed has an average terrain surface slope of 8 m/km, and elevations between 825 and 655 m above sea level (a.s.l.). Sandy soils, which dominate the watershed, along with the mild surface slope favor the occurrence of high infiltration rates and low overland flow. The Köppen-Geiger climate classification (Alvares et al., 2013) indicates a humid subtropical climate, with dry winters and rainy summers (Cwa). From 1979 to 2014, the mean annual precipitation was

1486 mm (about 65% in the rainy seasons), and the mean annual temperature, 21.6 °C (Cabrera et al., 2016).

Agricultural and livestock activities have dominated the area during the last decades. In 1990, eucalyptus plantations covered about 30% of the total area, pastures, 15%, and the native vegetation (Cerrado, tropical savanna), 30% (Pompeo, 1990). In 2017, eucalyptus, citrus, sugarcane, and pasture activities summed-up approximately 35%, 30%, 10%, and 4%, respectively, with the remaining native vegetation corresponding to only 10% (Figure 1).

Data from the monitoring stations and wells indicated in Figure 1 were used in this study. The reference period covered December 2008 to September 2019. Basic quality control procedures were followed to ensure data consistency.

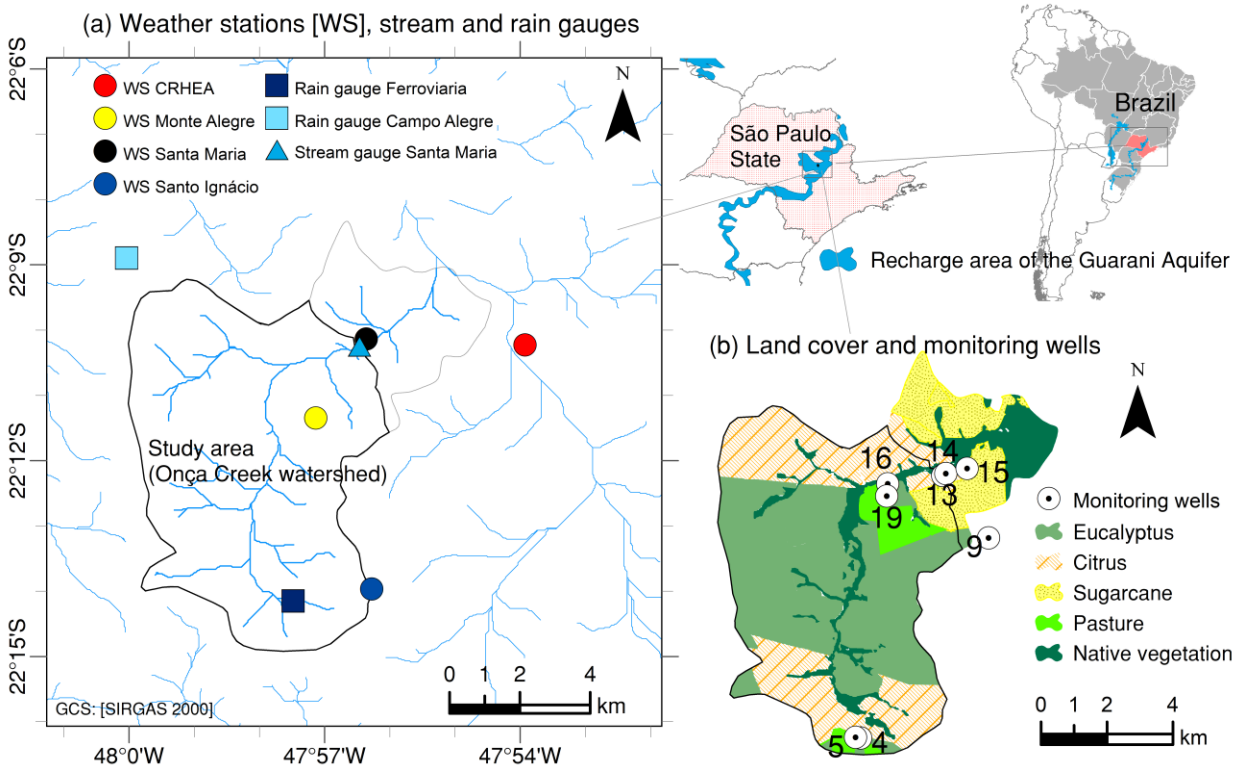

Figure 1. Location of the study area. (a) hydrometeorological stations and (b) dominant land-covers and monitoring wells. Data sources: 1. Recharge areas of the Guarani Aquifer: http://geoserver.ourinhos.unesp.br/; 2. DEM: https://www.infraestruturameioambiente.sp.gov.br/

2.2 Streamflow and Baseflow Data

The discharge time series was estimated at the stream gauge Santa Maria, in a section that drains an area of 52.1 km² (Figure 1), and located immediately upstream to a crossroad with two culvert pipes with a diameter of 1.50 m. Discharge measurements have been carried out by current meters twice a month since 2004. The daily discharges (long-term mean = 0.65 m³/s) were estimated from the transformation of sub-hourly water level data, using the classic power-law stage-discharge transformation function. The level data are recorded every 15 minutes, since December 2008, by water level loggers maintained in a stilling well hydraulically connected to the stream. The parameter for zero flow condition was set based on the invert elevation of the culverts, while the other rating curves parameters were calculated by the ordinary least square method. The parameters are variable over time due to the channel section instability (sandy stream bed) and due to the downstream impoundment with controlled water release during the dry seasons. Considering the errors associated with the rating-curves and the water level data, the median uncertainty in the estimated discharges was 11.9%, whereas the mean uncertainty, 13.0%.

The unconfined aquifer holds a high interaction with the stream, maintaining its perennial regime (Wendland et al., 2015). To estimate the daily baseflow time series, the two-parameter separation method proposed by Duncan (2019) was applied. This method comprises a backward filtering operation to fit an exponential master recession curve, followed by the original Lyne and Hollick digital filter to smooth the resulting curve. The filter parameters in the first and second pass were set equal to 0.983. The baseflow index was estimated as 0.86, which is comparable to the value of 0.83 obtained by Batista et al. (2018) through isotopic mass balance calculations, in nearby catchments with similar meteorological and hydrogeological conditions (recharge areas of the GAS).

2.3 Meteorological Data

The operation periods up to September 30, 2019, of the four nearby weather stations (WS, circles in Figure 1) and the two rain gauges (RG, squares in Figure 1) defined the meteorological datasets used here. All weather stations are equipped with rain gauges, air temperature and humidity probes, pyranometers, and anemometers.

Reference evapotranspiration rates were calculated by the Penman-Monteith method (PM FAO-56) (Allen et al., 1998). Given the dimensions of the study area and the spatial variability of meteorological variables, daily precipitation data and the reference evapotranspiration calculated at each WS were interpolated. The deterministic method of inverse distance weighting (IDW), with power 2 (Dirks et al., 1998), was used to obtain daily grid surfaces (50 x 50 m) from simultaneous records. The average values over the monitored drainage area formed the final time series of precipitation and reference evapotranspiration.

2.4 Spectral Vegetation Indices and Evapotranspiration

The vegetation dynamics can be relevant to describe the hydrological behavior of catchments, especially in small spatial and temporal scales (Area < 500 km² and 1 - 5 years) (Donohue et al., 2007; Wegehenkel, 2009). Studies have shown that the relationships between actual (AET) and reference (RET) evapotranspiration are reasonably well estimated, in diverse biomes, as a function of remotely sensed vegetation indices (Glenn et al., 2008; Glenn et al., 2011; Kamble et al., 2013; Nagler et al., 2013).

The empirical method proposed by Nagler et al. (2013), based on experimental studies in the state of Arizona (USA), was adapted here to provide an estimate of the actual evapotranspiration. The method follows Equation 1, which is based on the Beer-Lambert law to

determine the absorption of light by a canopy and takes the enhanced vegetation index (EVI) as an indicator of the density of light-absorbing particles.

$$\mathrm{AET} = \left[a\left(1 - \exp\left(-b.\,\mathrm{EVI}\right)\right) - c\right] \times \mathrm{RET} \quad (1)$$

in *a*, *b*, and *c* are parameters to be calibrated against observed data. As a simplification, motivated by the lack of observed evapotranspiration in the study area (e.g., by flux towers in representative land covers), we adopted $a = 1.65$ and $b = 2.25$, the same values estimated by Nagler et al. (2013), and left *c* as a free parameter to be adjusted in the water balance assessment. This is equivalent to allow for slight translations in the transformation curve, maintaining the sensibility of the ratio AET/RET with respect to EVI.

The EVI data, referring to the surrounding areas of the monitoring wells and to the watershed area, were taken from the product MOD13Q1 (Didan, 2015), generated every 16 days in 250 m spatial resolution, and made available on the NASA Land Processes Distributed Active Archive Center (LP DAAC). Daily values of EVI were obtained from linear interpolation, then individual time series of actual evapotranspiration for each well and for the watershed were calculated.

2.5 Groundwater Level

The surrounding typical land covers, the ranges of water level variation of the eight monitoring wells considered in this study (Figure 1), and the respective specific yield (drainable porosity) values are shown in Table 1. The wells have been monitored every 15 days by water level meters, and twice a day by dataloggers (Levellogger® Edge 3001 or Diver® DI501). The datasets from the manual and automatic measurements were adjusted, resampled, and merged to

form the final daily time series for each well. Also, an average groundwater level variable (named 'GWL') was defined as the arithmetic mean of the levels observed in the monitoring wells.

Table 1. Description of the groundwater monitoring wells (m b.g.l. = meters below ground level). Specific yield (drainable porosity) values were obtained from Wendland et al. (2015).

ID	Land cover	Water level depth (m b.g.l.)			Specific yield
		min	mean	max	
W4	Pasture	14.4	16.5	17.9	15.9%
W5	Pasture	3.6	6.3	9.1	15.9%
W9	Eucalyptus	15.6	20.2	24.4	15.1%
W13	Sugarcane	5.5	9.8	11.0	15.1%
W14	Sugarcane	2.8	6.6	7.6	15.1%
W15	Citrus	4.6	7.8	9.5	8.5%
W16	Eucalyptus	2.7	5.2	7.2	12.3%
W19	Eucalyptus	8.9	15.1	18.7	12.3%

2.6 Water Balance

We examined the water balance dynamics in the watershed from October 01st, 2009 to September 30th, 2019, which corresponds to ten complete water years in the study area. The control volume of inputs and outputs was defined as the region from the upper canopy layer to the layer where deep regional recharge (outflow to the GAS) occurs.

$$\Delta S = PP - AET - Q - DR \quad (2)$$

in which ΔS: water storage change, PP: precipitation, AET: actual evapotranspiration, Q: discharge, and DR: deep recharge. Based on the study by Wendland et al. (2007), the deep recharge was assumed as 3.5% of the average annual precipitation (~ 50 mm/a in the study period). The accumulated water storage changes are supposed to follow the average behavior of the groundwater levels because the soil moisture has negligible annual variations in the study area (Pompeo, 1990).

2.7 Granger Causality Test (GC)

The Granger causality definition is based on the increase of predictive power of an autoregressive model by including an additional variable, candidate to present a causal relationship, with a certain time lag (Granger, 1969). Some considerations to make such definition applicable are that the data are generated according to a linear and stochastic process. In the simplest, bivariate case, two models are considered: the unrestricted (Equation 3), with the past values of the variables X and Y; and the restricted (Equation 4), only with the past values of the variable Y. Here we assumed that hydrological processes can be modeled by vector autoregressive models, despite their nonlinearities.

$$y_{u_t} = c_u + \sum_{j=1}^{p} \left[\phi_{1j} y_{t-j} + \phi_{2j} x_{t-j}\right] + \epsilon_t, \quad \epsilon_t \sim N(0, v_u) \tag{3}$$

$$y_{r_t} = c_r + \sum_{j=1}^{p} \phi_{1j} y_{t-j} + \epsilon_t, \quad \epsilon_t \sim N(0, v_r) \tag{4}$$

in which c_u, c_r, ϕ_{1j} e ϕ_{2j} are model parameters, and ϵ_t is a Gaussian error, with variance v.

The null hypothesis is accepted when $\phi_{2j} = 0$ for j =1, 2, …, p, meaning that X does not cause Y in the Granger sense. Conversely the null hypothesis is rejected when $\phi_{2j} \neq 0$ for a j between 1 and p.

The open source library Statsmodels 0.9.0 (Seabold & Perktold, 2010) was used for the computational implementation of the method. The algorithm executes the test for multiple time lags (up to the maximum lag length) between pairs of variables and examines the corresponding statistical significance of the causal relationships based on the likelihood ratio test (Equation 5).

$$\mathrm{LR} = -2\log\left[\frac{\mathcal{L}_{\mathrm{r}}(\hat{\theta}_0)}{\mathcal{L}_{\mathrm{u}}(\hat{\theta}_1)}\right] \sim \chi^2\left(\mathrm{d.f.} = \mathrm{p_u}\text{-}\mathrm{p_r}\right) \quad (5)$$

in which LR is the likelihood ratio, $\hat{\theta}_i$ indicates the maximum likelihood estimation of θ_i (model parameters) under the hypothesis i, $\mathcal{L}$ indicates the likelihood function, and the subscripts u and r refer to the unrestricted and restricted model, respectively. The LR follows approximately a chi-square distribution, with degrees of freedom (d.f.) equal to the difference between the number of parameters of the unrestricted (p_u) and restricted (p_r) models (Wilks, 1938).

2.8 Time-lagged normalized mutual information (NMI)

The mutual information I (X; Y) is defined as the relative entropy between the joint distribution and the product distribution (Cover & Thomas, 2005):

$$\mathrm{I(X;Y)} = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathrm{p_{XY}}(x,y)\log_2\frac{\mathrm{p_{XY}}(x,y)}{\mathrm{p_X}(x)\mathrm{p_Y}(y)} = \mathrm{H(X)} + \mathrm{H(Y)} - \mathrm{H(X,Y)} \quad (6)$$

in which $\mathrm{p_X}$, $\mathrm{p_Y}$ and $\mathrm{p_{XY}}$ are marginal and joint probability mass functions, and H(X), H(Y) and H(X,Y) are the entropy and joint entropy of the discrete random variables X and Y. The mutual information I (X; Y) measures the general dependence (linear and non-linear) between two variables and can be seen as the reduction of the uncertainty of variable X (or Y) due to knowledge of variable Y (or X). By adopting some time lag between the variables under analysis, mutual information can be used to detect the direction and intensity of interaction between linear or non-linear processes (Li et al., 2018).

$$\mathrm{TLMI(X;Y,\tau)} = -\sum\sum \mathrm{p_{XY}}(x_t, y_{t+\tau})\log_2\frac{\mathrm{p_{XY}}(x_t, y_{t+\tau})}{\mathrm{p_X}(x_t)\mathrm{p_Y}(y_{t+\tau})} \quad (7)$$

in which TLMI is the time-lagged mutual information, and τ is the time lag between the cause (X) and effect (Y) variable. A normalized metric, presented in Equation 8, was used here.

$$0 \leq \ \mathrm{NMI}(\mathrm{X},\mathrm{Y};\tau) = \frac{\mathrm{TLMI}(\mathrm{X},\mathrm{Y};\tau)}{\min[\mathrm{H}(\mathrm{X}),\mathrm{H}(\mathrm{Y})]} \leq 1 \tag{8}$$

in which H(X) and H(Y) are the entropy of X and Y.

The marginal and joint probability distributions were estimated from histograms (Li et al., 2017). For each pair of variables (X, Y), the number of bins was set as the smallest between 30 and the geometric average of the quantities individually determined by the Freedman-Diaconis rule (Freedman; Diaconis, 1981).

2.9 Verification Assessment

The verification strategy presented here aimed to identify errors in the computational implementation and to understand the results for situations in which the temporal dependencies between cause and effect are well-known. Synthetic time series, generated according to Equation 9 (W_t, a simple summation function) and Equation 10 (Z_t, a discrete convolution operation), were investigated by both the Granger causality test and the normalized time-lagged mutual information.

$$\mathrm{W_t} = \max\left(0\,;0.20 + 1\times10^{-4}\times\sum_{\mathrm{j}=-\mathrm{m}}^{\mathrm{m}}\left[\mathrm{PP}_{\mathrm{t}-\tau^{*}+\mathrm{j}}\right]^{\mathrm{a}} + \epsilon_{\mathrm{t}}\right), \qquad \epsilon_{\mathrm{t}} \sim \mathrm{N}(0,\mathrm{d}^2\sigma_{\mathrm{PP}}^2) \tag{9}$$

$$\mathrm{Z_t} = \sum_{\mathrm{j}=-\infty}^{\infty}\mathrm{PP}_{\mathrm{t}-\mathrm{j}}\mathrm{UH_j} + \epsilon_{\mathrm{t}},\ \mathrm{UH_t} = \frac{(\mathrm{t/k})^{\mathrm{n}-1}\exp(-\mathrm{t/k})}{\mathrm{k}\,(\mathrm{n}-1)!},\ \mathrm{E}[\mathrm{UH_t}] = \mathrm{n.\,k};\ \ \epsilon_{\mathrm{t}} \sim \mathrm{N}(0,\mathrm{d}^2\sigma_{\mathrm{PP}}^2) \tag{10}$$

in which PP_t [mm] is the mean areal daily precipitation time series in the study area, σ_{PP} is the sample standard deviation of PP (~9 mm), $E[UH_t]$ is the expected value of the gamma distribution function UH_t (Besbes & de Marsily, 1984), and ϵ_t is an uncorrelated Gaussian noise. The parameter arrays, (a, m, τ^*, d) for W_t and (n, k, d) for Z_t, were set as shown in Table 2. The ranges

intended to assess the methods under diverse conditions. The premise is that the methods were capable to identify the lags $L_W = \tau^*$ for W_t, and $L_Z = n \times k$ for Z_t.

Table 2. Parameters adopted to generate the W and Z synthetic time series.

Time series	Group ID	Fixed Parameters	Variable Parameter
W	1	a = 1 d = 1 τ^* = 200	m_j = 1; 10; 100
	2	m = 1 d = 1 τ^* = 200	a_j = 0.1; 0.5; 2
	3	a = 1 m = 1 τ^* = 200	d_j = 0.01; 0.10; 10
	4	a = 1 m = 1 d = 1	τ_j^* = 200; 500; 800
Z	1	n = 10 d = 0.01	k_j = 10; 20; 30
	2	k = 10 d = 0.01	n_j = 5; 20; 30
	3	n = 14 k = 14	d_j = 0.001; 0.1; 1.0

2.10 Analysis of Observed Data

The Granger causality test and the normalized lagged mutual information were calculated in a pairwise setup. The pairs were defined as: in a first set, precipitation and evapotranspiration as source ('cause') variables, whereas streamflow, baseflow, and groundwater level as target ('response') variables; and in a second set, reference evapotranspiration, enhanced vegetation index and groundwater level as sources, while streamflow and baseflow as targets. Incremental lags, from 1 up to 1000 days, were considered. These pairs were defined based on likely or possible cause-effect relationships in the hydrological system. For convenience, seasonally adjusted data were not used. Then, the results referred to time lags greater than one year are affected by seasonal components.

All the observed time series were affected by uncorrelated Gaussian noise to compensate for uncertainties in the deterministic estimations of the hydrological processes. Due to data insufficiency and for the purpose intended here, the statistical properties of errors were taken as fixed approximations. A coefficient of variation (relative uncertainty) of 15% was adopted for streamflow, baseflow, and evapotranspiration estimates, whereas fixed standard deviations of 10 mm, 0.02 m, and 0.06 were set for precipitation, groundwater level and EVI data, respectively.

The statistical significance of the mutual information values was verified using a shuffled surrogate method (Ruddell & Kumar, 2009; Franzen et al., 2020). In this method, the time series data are shuffled to destroy time dependencies between the variables, and the mutual information is calculated using both the shuffled and the original data. Thirty iterations of shuffled data were used to compute a critical value associated with a 95% confidence level. Gaussian distribution was assumed. When the mutual information calculated for the observed data is greater than the critical value (calculated based on the shuffled sequences), the mutual information value is considered statistically significant.

The cross-correlation function, a usual method to measure the similarity between time series (Gómez et al. 2018), was also computed here for each case. This aimed to question the relevance of the results obtained from causal methods when compared with the results from a simpler method: are they the same?

For convenience and in view of the purpose of this study, all processes were assumed stationary. There are no significant shifts or trends in the observed time series (Figure 2) that motivate the use of more complex statistical models (Koutsoyiannis, 2006).

3 Results

3.1 Water Balance Dynamics

Figure 2 shows the time series from 2009 to 2019 of the streamflow (total, Q and baseflow, Qb), precipitation (PP), estimated actual evapotranspiration (AET), enhanced vegetation index (EVI) and average groundwater level (GWL). There is a groundwater dominance in the streamflow dynamics, with a baseflow index greater than 0.85. The ratio of reference evapotranspiration by annual precipitation is 0.87, and the ratio of AET by annual precipitation, 0.70. Trend analysis was not the focus here, however, it is possible to see a slight downwards trend in Q, Qb, and GWL time series data, and also an apparent upwards trend in the AET.

The water year 2013-2014 was marked by a meteorological drought (Coelho et al., 2016; Marengo et al., 2015), with annual precipitation 20% lesser than the average. The impacts on groundwater levels and streamflow can be readily observed. Minimum groundwater levels were reached in January 2015, and minimum streamflow, in September 2015.

From October 2009 to September 2019, the average groundwater level experienced a reduction from 716.13 to 714.90 m a.s.l. Considering a mean specific yield of 0.10, this groundwater level change represents a water storage reduction of 123 mm in the phreatic zone. In the water balance, when the parameter c used to estimate the AET is taken as 0.220, the accumulated water storage changes in the watershed varied from -4 to -125 mm in the same period. This reasonable water balance closure suggests that the estimates presented in Figure 2 are consistent and can be used in our causal analyses.

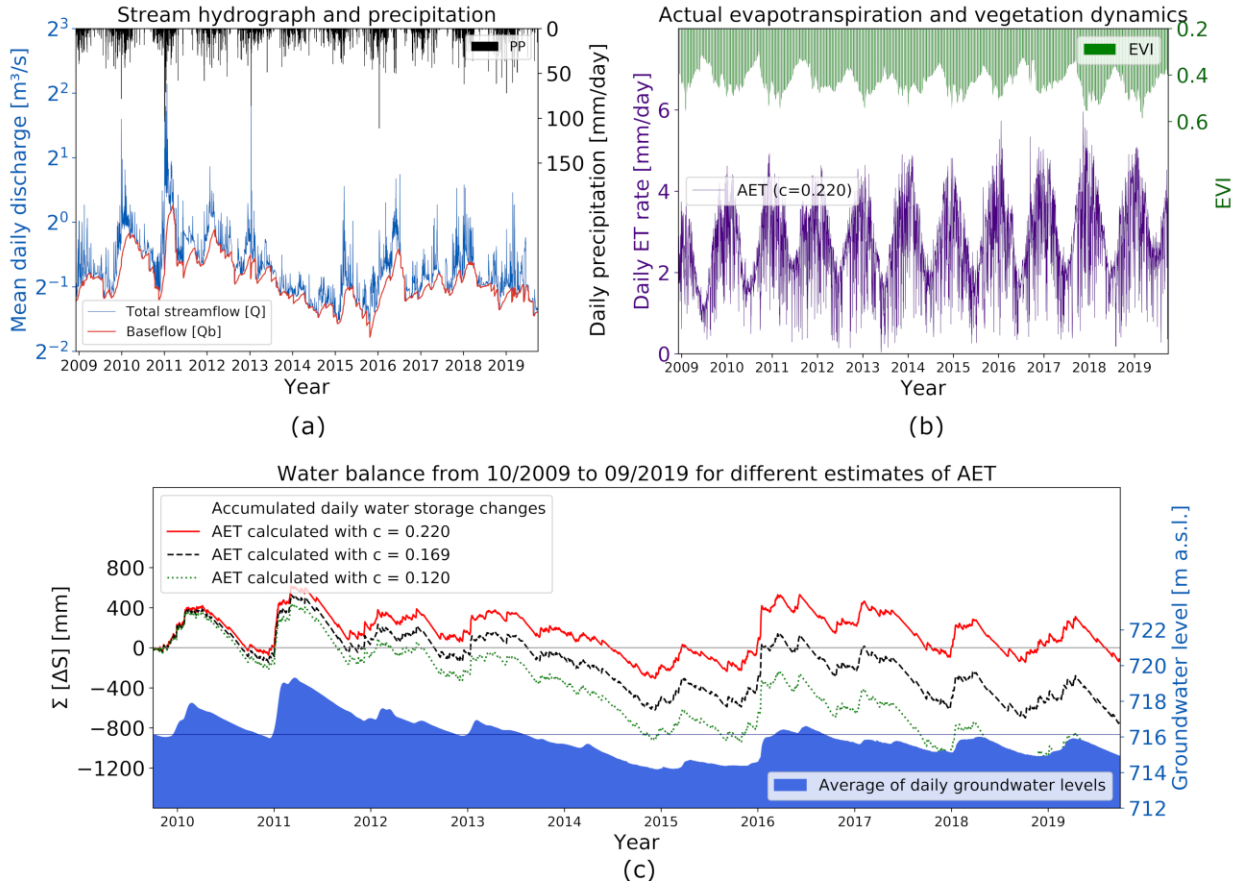

Figure 2. Water balance components and vegetation index time series in the study area. The actual evapotranspiration calculated with c = 0.220 (continuous red line) led to a water balance consistent with the groundwater level variation from October 2009 to September 2019. Mean annual rates in the study period: precipitation = 1403 mm/a; streamflow = 384 mm/a (0.63 m³/s); baseflow = 324 mm/a (0.54 m³/s); reference evapotranspiration = 1220 mm/a; and actual evapotranspiration = 985 mm/a. Horizontal lines indicate initial values.

3.2 Synthetic Time Series

The methods were effective to characterize the time delay mechanisms that generated the synthetic time series W (derived from a summation function) and Z (derived from a convolution operation), as shown in Figures 3 and 4. When the p-value (GC) is greater than 0.05 (adopted significance level), there is statistical evidence that the lagged time series does not present a causal relationship in the Granger sense. Complementary, the intensity of the functional connectivity, or causal interaction, between the lagged time series is proportional to the NMI metric. When the results are analyzed simultaneously, we realize that they were capable to identify the time lags in

practically all instances. Exceptions happened for the cases in which large noise variances were used to generate the response time series (Figure 3i, Figure 4i), affecting especially the mutual information-based method. GC worked well for most of the cases. The most relevant inconsistencies were found in Figure 3c and Figure 3l, situations in which the causal relationships in the Granger sense were identified for time lags about 100 and 400 days, respectively, shorter than they should be.

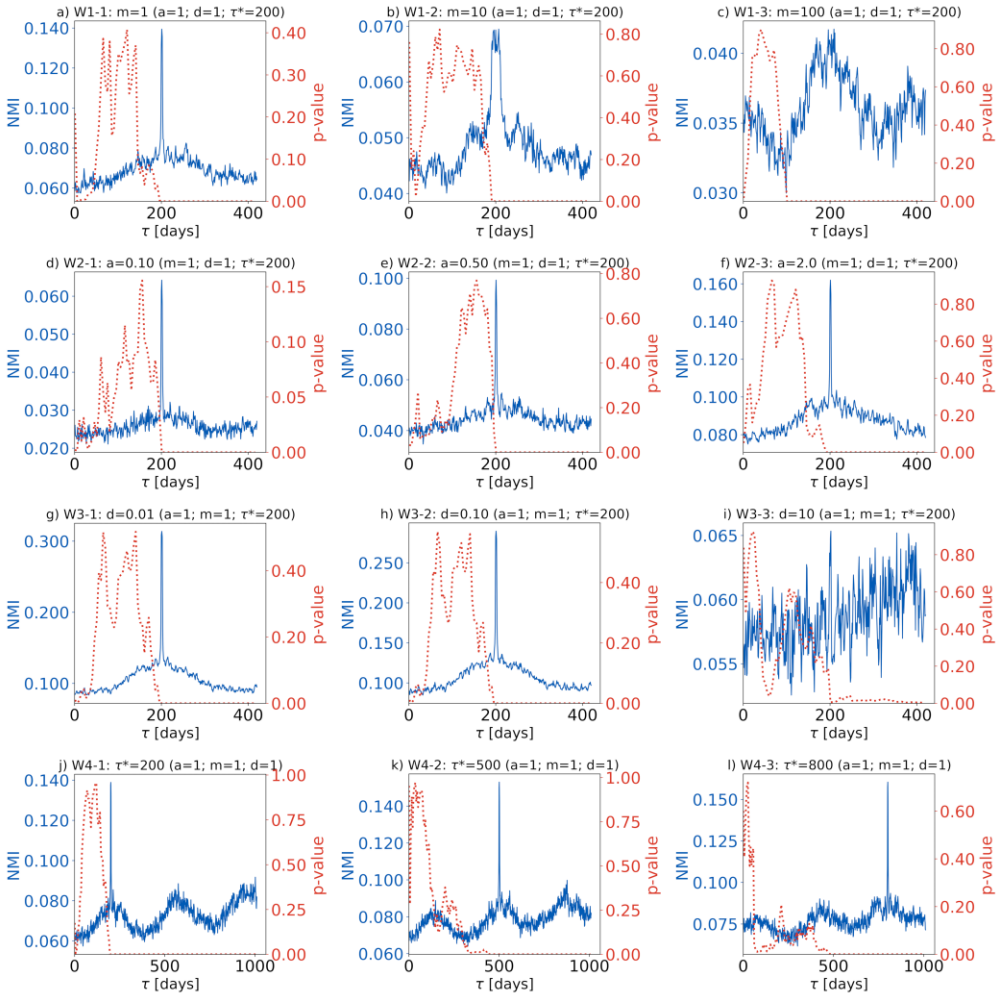

Figure 3. Causal detection in the synthetic time series W. For the Granger causality analysis (GC, results indicated by the dotted red lines), the null hypothesis (absence of causal interaction in the Granger sense) is rejected when p-value < 0.05. The normalized time-lagged mutual information NMI (indicated by the blue lines) shows the functional connectivity dynamics between the time series (PP and W_j). The time lags in which global peaks occur correspond to the time lags (τ^*) used to generate the synthetic time series. In most of the instances, the GC and NMI methods correctly detected the time lags τ^*. The exceptions are found in plots *c*, *i* and *l*.

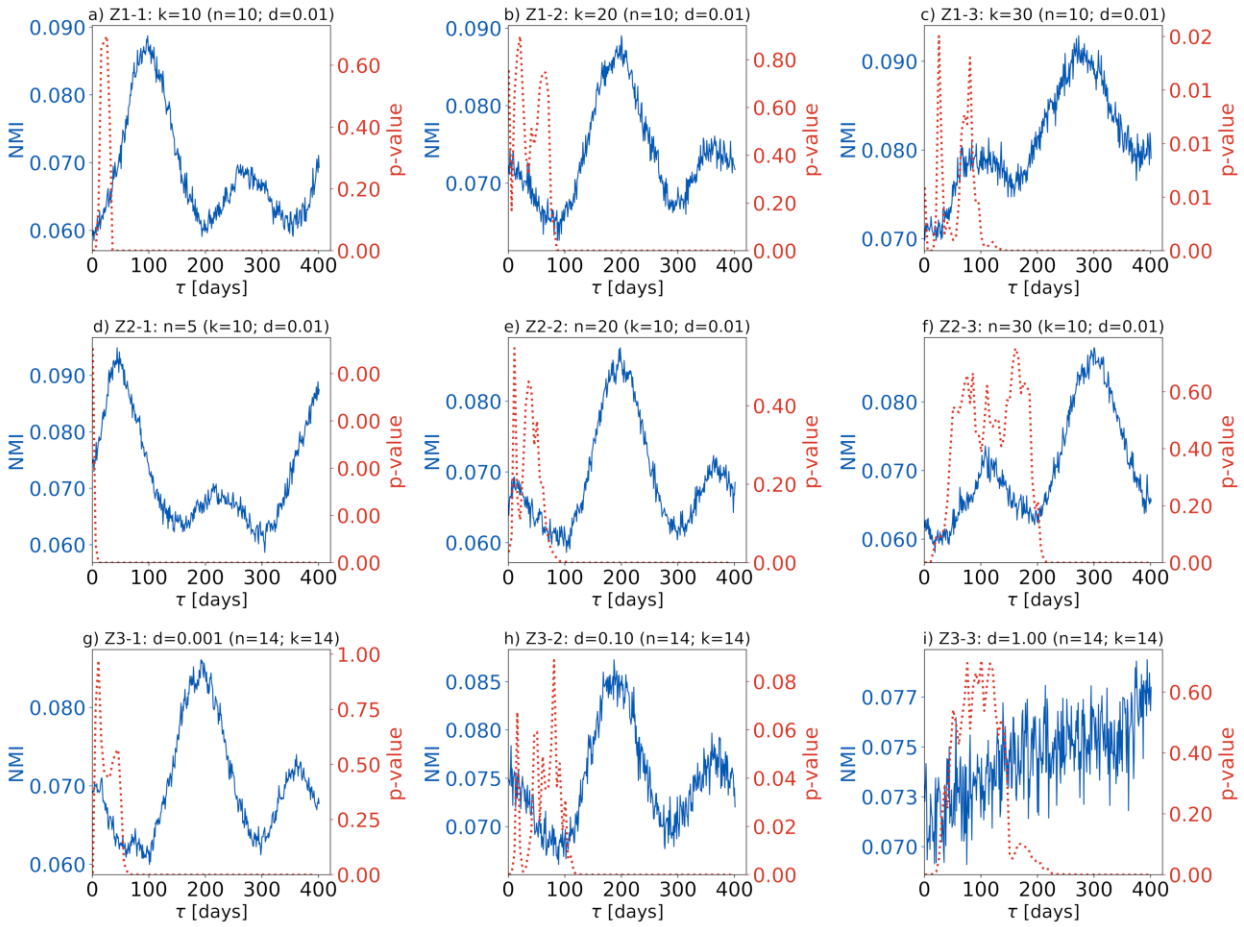

Figure 4. Causal detection in the synthetic time series Z. The NMI metric correctly detected in most of the cases the mean time-delay (n.k) of the unit hydrographs used to generate the time series, while the Granger causality analysis detected the causal relationships (p-value < 0.05) with some anticipation (between 50 and 150 days).

The Granger causality test, despite being designed to detect linear causal relationships, was capable to identify the lags in series W2-1, W2-2 and W2-3, in which non-linear functions of PP were used.

3.3 Observed Time Series

The Granger causality (GC) test confirmed that meteorological variables are influencing groundwater levels (GWL) and streamflow (total or baseflow) for practically all cases, with different time lags. A summary of the results is presented in Table 3. The precipitation affects (p-

value < 0.05) the streamflow and the water level at practically all monitoring wells, even for short time lags (starting from 1 day). A singular behavior was found in the deepest water table (well W9), in which the causal interaction between precipitation and water level started to be confirmed only for time lags greater than 150 days.

Each well presented a unique GC response to evapotranspiration, regardless of the surrounding land cover, suggesting that the water table depth and the soil properties are together controlling the minimum time lag required to detect a causal interaction.

When analyzed as a response of the precipitation or as a cause of groundwater level changes, the EVI presented two disconnected intervals with p-value < 0.05. GC analysis only determines whether a causal interaction exists between time series, identifying the time-delay associated with them. As an additional source of information, the NMI measure shows details related to the strength and dynamics of functional connectivities between processes.

Table 3. Results of the Granger causality test.

Causal interaction		**Time lag intervals with p-value < 0.05** (M = 1000 days)
Cause	Response	
PP →	Q or Q_b	[1, M]
	GWL_j (j = 4, 5,13, 14, 15, 16, 19)	[1, M]
	GWL_9	[150, M]
	EVI	[1, 50] ∪ [265, M]
AET →	Q	[1, M]
	Q_b	[100, M]
	GWL_{14}	[1, M]
	$GWL_{4,16}$	[10, M]
	GWL_5	[20, M]
	$GWL_{13, 15}$	[30, M]
	GWL_{19}	[100, M]
	GWL_9	[260, M]
EVI →	Q	[300, M]
	Q_b	[370, M]
	GWL_{mean}	[1, 30] ∪ [270, M]
GWL →	Q or Q_b	[1, M]

Figures 5 and 6 present NMI curves considering precipitation and evapotranspiration as causes and the groundwater levels as responses. To facilitate the visualization, simple moving averages of the NMI values (period of 5 days) were plotted. The upwards trend observed in most of the curves occurs because the time-lagged mutual information, unlike GC analysis, does not eliminate the effects induced by the own response time series memory (dependency on previous states and seasonality).

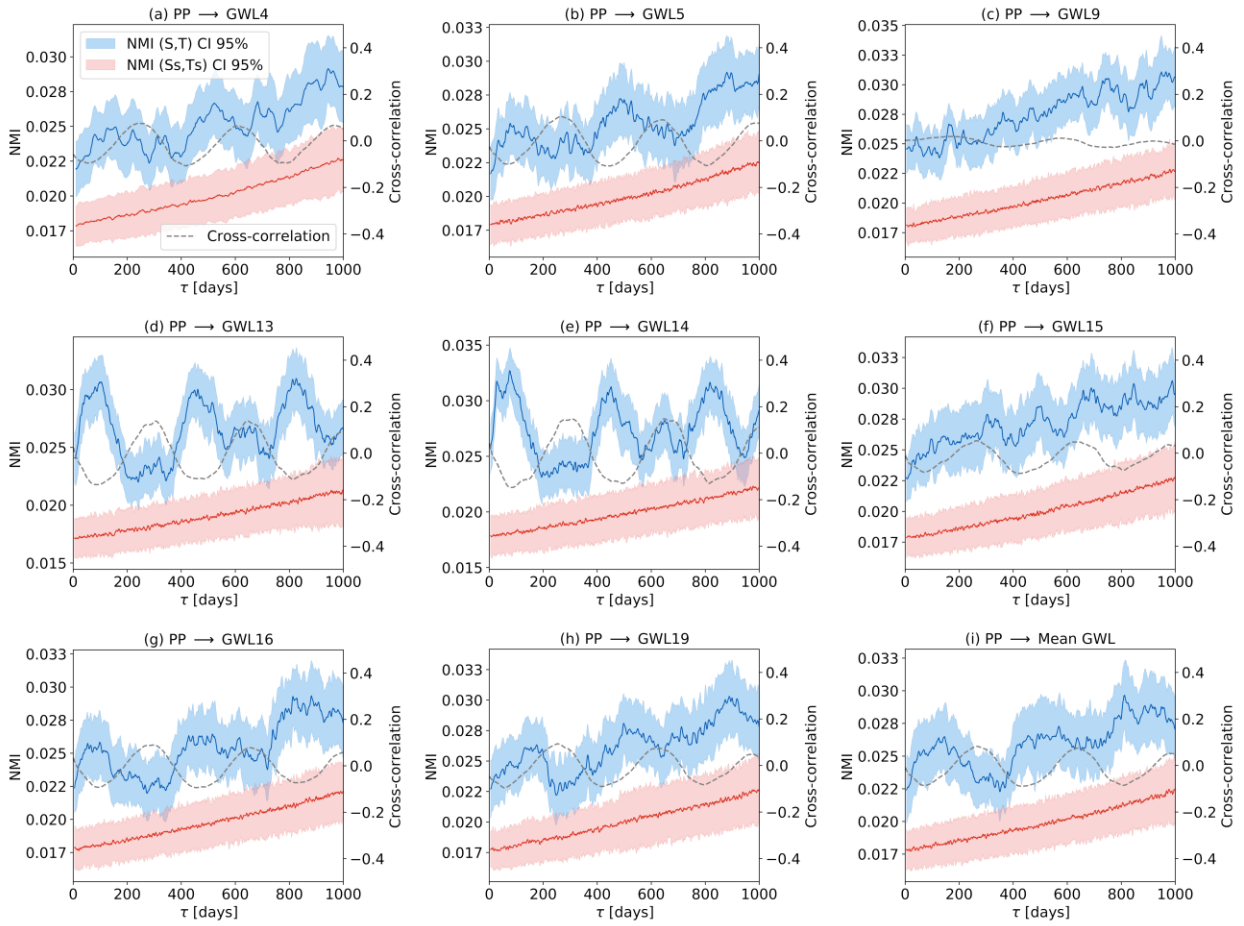

Figure 5. Cross-correlation and normalized time-lagged mutual information (NMI) between precipitation (PP) and groundwater level at the monitoring wells (GWL_j) for time delays (τ) up to 1000 days. (S, T) (blue): NMI curves and confidence intervals for the source (cause) and target (response) variables (variables generated randomly from mean and uncertainty values). (Ss, Ts) (red): NMI curves and confidence intervals for the shuffled time series.

The results were examined individually. Considering first the precipitation, the wells W4 and W5, which are 30 m apart and installed in a pasture area, presented a similar pattern despite the depth difference (~10 m, Table 1). This suggests that location and soil characteristics may be a critical factor to understand the results because they are monitoring different aquifers (mean water level equal to 16.5 m b.g.l. at W4, and 6.3 m b.g.l. at W5). The earliest local NMI peaks occurred at 240 and 140 days for W4 and W5, respectively. These time lags seem to be related to the mean time required to the water reach the respective water tables.

Only the well W9 did not present statistically significant NMI. That well is located out of the watershed, 500 m away from the water divide and surrounded by a Eucalyptus plantation. The great depth of the water table (20.2 m b.g.l., Table 1) may be one of the reasons that contributed to the singular behavior.

The wells W13 and W14, 100 m apart in a sugarcane crop area, followed a similar general shape, with the earliest NMI peaks at 100 and 75 days.

The well W15, located in a citrus orchard, presented a trending NMI curve, without any relevant peaks. For the wells W16 and W19, which are in a Eucalyptus plantation area and located 150 and 500 m away from the Onça Creek (main stream), the NMI curves showed similar patterns, with earliest local (and global) peaks at 115 and 175 days, respectively.

When examining the interaction between groundwater levels and evapotranspiration (AET) (Figure 6), the water levels at wells close to each other (W4 and W5; W13 and W14; W16 and W19) exhibited similar NMI patterns. Such patterns are characterized by multiple local peaks, which overall did not coincide with the peaks found in the analyses with precipitation.

When comparing the time lags in which local NMI peaks occurred with those lags detected by the GC analysis (Table 3), no wells showed inconsistency, that is, all peaks occurred in intervals with causal interaction confirmed in the Granger sense.

Overall, the NMI curves and cross-correlation functions do not indicate the first peaks at the same lag times. For example, considerable differences can be observed in the groundwater level responses at the wells W13 and W14. The linearity premise of the correlation method probably leads to the observed deviations.

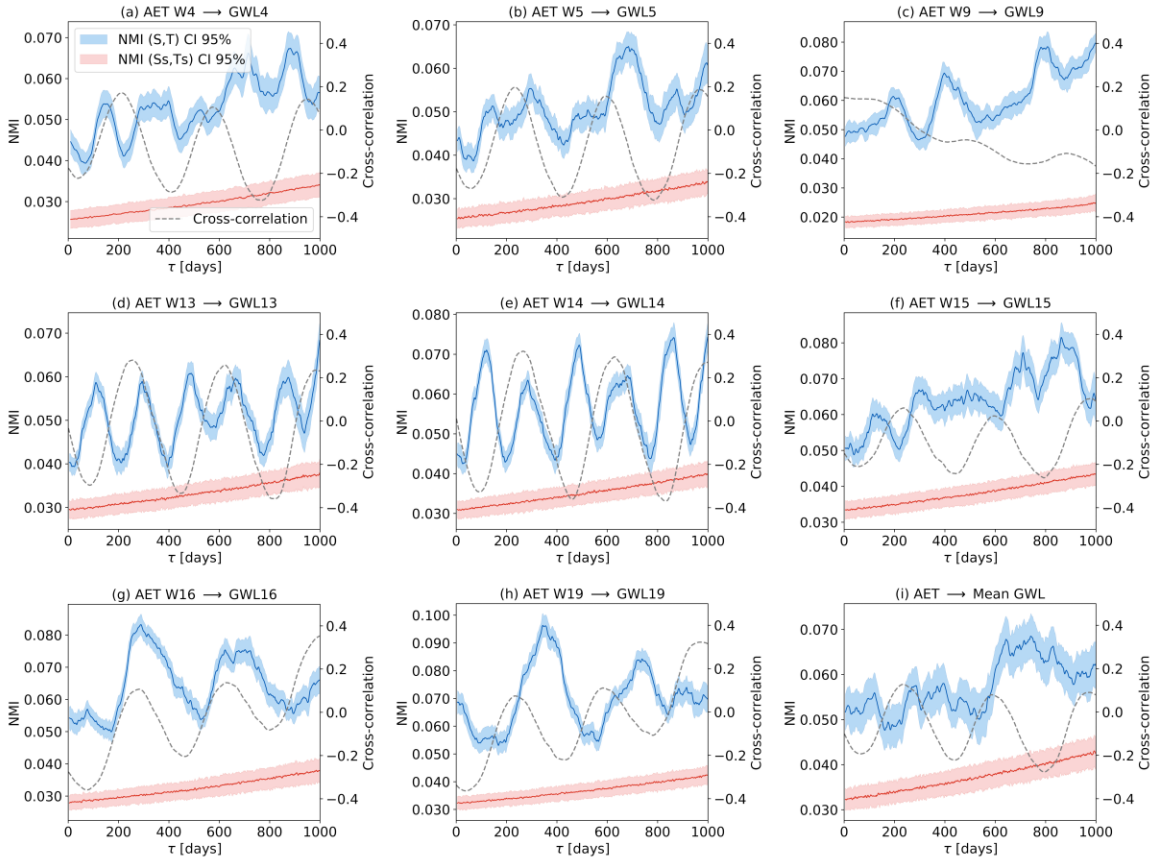

Figure 6. Cross-correlation and normalized time-lagged mutual information (NMI) between evapotranspiration (AET) and groundwater level at the monitoring wells (GWL_j) for time delays (τ) up to 1000 days. (S, T): NMI curves for the source (cause) and target (response) variables. (Ss, Ts): NMI curves for the shuffled time series.

Figure 7 shows how precipitation, evapotranspiration, EVI and mean GWL interact with total streamflow (a) and baseflow (b) over different time delays. The hydraulic connectivity

between the aquifer and the stream is evidenced by the high NMI values for early time lags. The NMI curve for the relationship between precipitation and streamflow drops rapidly in the first 5 days of delay, and it develops two local peaks at 380 and 765 days. The NMI curve for the baseflow as a response of the precipitation presented local maxima at 50, 415, 780, and 845 days. The evapotranspiration presented slight peaks every ~ 180 days. Again, divergences between NMI curves and cross-correlation functions are evidenced.

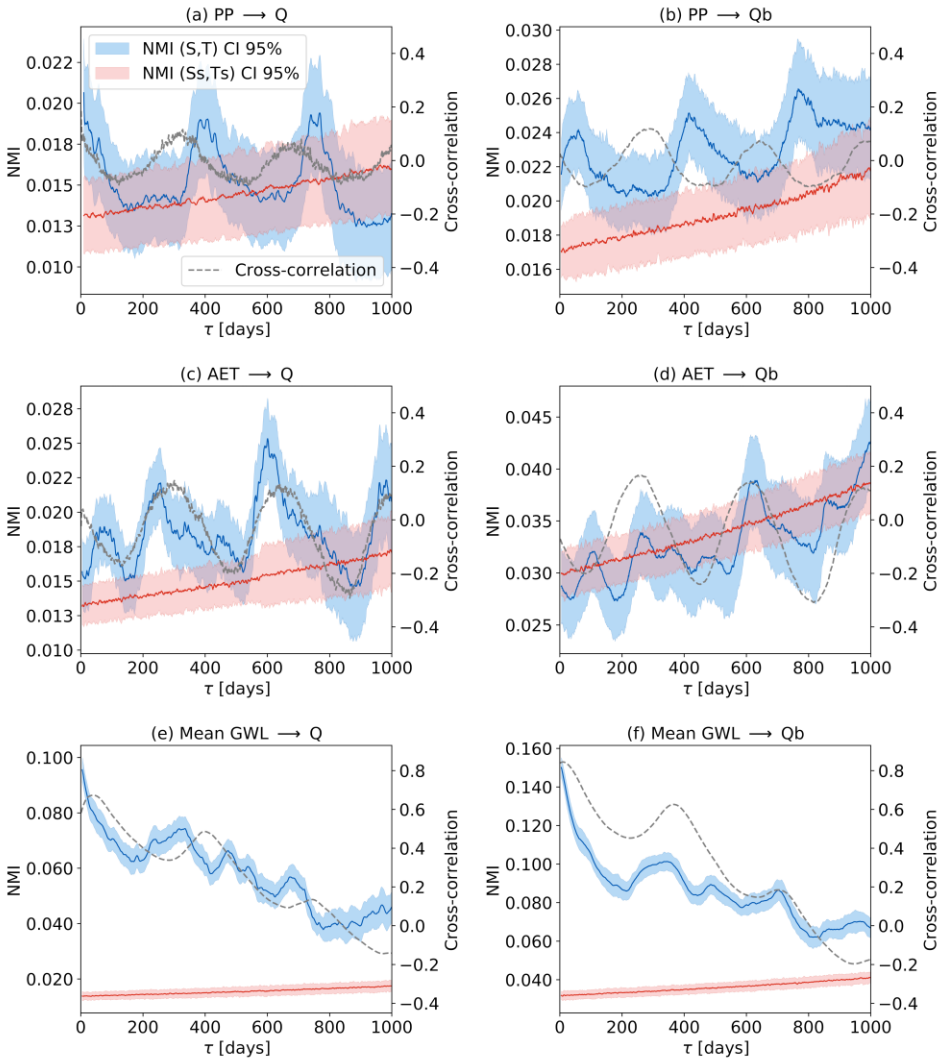

Figure 7. Cross-correlation and NMI between meteorological variables (precipitation-PP, actual evapotranspiration-AET), groundwater states (mean groundwater level-GWL), and streamflow (Q) and baseflow (Q_b) responses. (S, T): NMI curves for source (cause) and target (response) variables. (Ss, Ts): NMI curves for the shuffled time series.

To support the interpretation of the results showed in Figures 5, 6 and 7, especially to better know the memory timescale and the persistence of the hydrological processes, the time-lagged mutual information was also calculated for each variable, determining a 'self-information' (Figure 8). Notably, the mean groundwater level (GWL) has the highest persistence, followed by the baseflow and streamflow. The memory timescale, if estimated by the first local minimum of the time-lagged mutual information, was about 200 days for GWL, Q, and Q_b, and about 100 days for EVI and AET. The precipitation did not present statistically significant results.

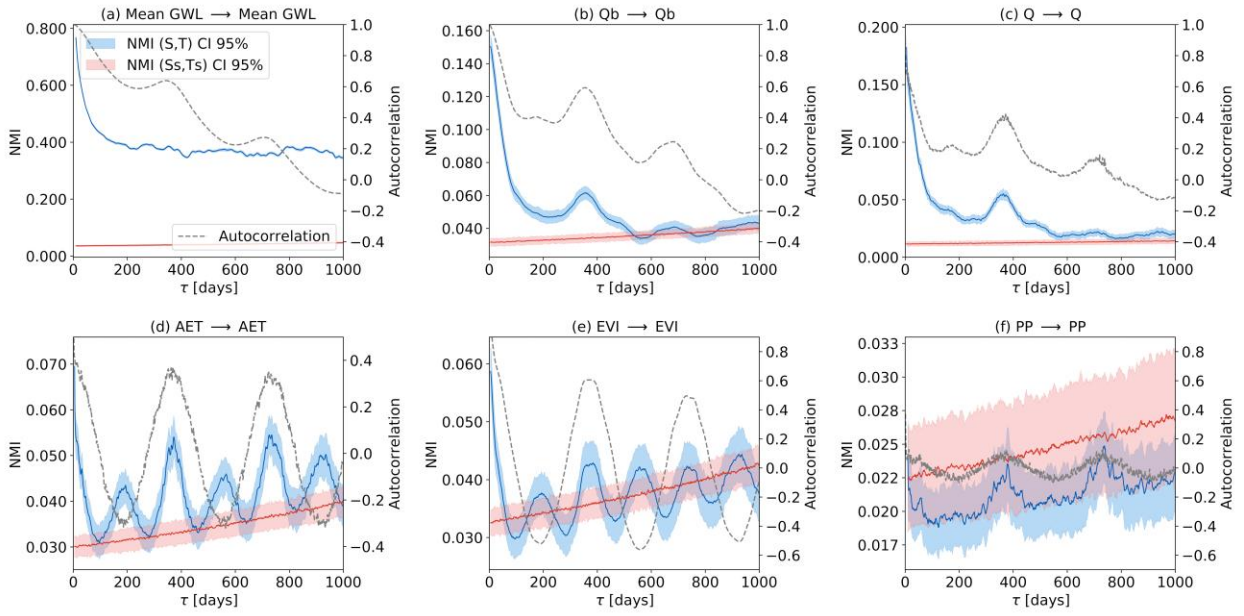

Figure 8. Autocorrelation and time-lagged (auto) mutual information series. (S, T): NMI curves for source (cause) and target (response) variables. (Ss, Ts): NMI curves for the shuffled time series.

4 Discussion

The time-lagged mutual information and the Granger causality test proved to be effective to detect relevant time lags in synthetic time series. Simple and convergent interpretations were possible in our assessment. When analyzing the watershed data, however, the results were diverse,

without a clear pattern. The Granger test seemed to be useful to detect processes that demand a large time delay to establish a connection. Since our study area is relatively small (~ 52 km²), in most of the cases, the causal interaction was detected even for 1-day lag (Table 3). These results could be expected since all local processes are structurally connected, sometimes with an intermediate reservoir (e.g., soil). Nevertheless, the innovative information we obtain from them is very limited. The NMI curves presented multiple and statistically significant local peaks, which are likely attributed to the memory of the response time series itself (Li et al., 2017). The causal information flow may be limited only to the first local peak, mainly when we try to give a physical interpretation. For example, the results from the pairs formed by precipitation and groundwater level (Figure 5) suggest that the first peak is the mean time required for the rainwater to reach the water table, while the second and third peaks are products of the seasonality of the variables. The diversity of patterns in the NMI curves also suggest that location, soil characteristics and land cover are all contributing to the responses of groundwater levels. Streamflow is highly dependent on the groundwater (Figure 7), validating the physical knowledge (or conceptual model) we have about the system (Machado et al., 2016; Wendland et al., 2015).

For the same watershed studied here, the study by Gómez et al. (2018) applied correlation and wavelet transform based techniques and found a response time of approximately 2 years of the baseflow to precipitation events. This result seems to correspond to that presented in Figure 7b (global NMI peak at ~750 days). Nevertheless, when considering the general upwards trend of the mutual information observed in practically all pairwise analyses, we understand that only the first local peak has a potential physical meaning related to response time, whereas the other peaks are associated with the auto dependency (and seasonal effects) found in the precipitation data (Figure 8f) or in other driving variables. This interpretation is reinforced by the groundwater level

responses to precipitation (Figure 5) once the first NMI peaks occur in the same range of 50 to 100 days of time lag.

Although the interactions in the hydrologic system are generally nonlinear, the Granger causality test revealed the existence of causal interactions in most of the pairwise analyses, even for very short time lags. Similar capability to detect nonlinear dependencies in hydrometeorological systems using the Granger test was reported by Ombadi et al. (2020). In our case, the detections may have been facilitated by the high connectivity between the hydrological processes. In some instances, the Granger test was useful to detect low connectivity between causes and responses (e.g., GWL9). We did not find a clear pattern in the results that could be associated with a physical interpretation of the lags detected. Although the Granger causality test seems to not be as useful as other information theoretic-based methods when assessing the connectivity between processes, it still provides some information. One utility we defend is precisely the opportunity to reduce uncertainties related to the interpretations. For example, a local peak in the NMI curve between the variables S and T for a time lag in which there is no causality in the Granger sense possibly will not have a relevant meaning.

4.1 Limitations and Future Directions

This study consists of a practical application of causal analysis to a hydrologic system. The meaning of the multiple peaks of the NMI measure and the possible physical interpretation associated with the results require further clarification. Even in a scenario with doubts about these meanings and interpretations, we understand that the methods have potential to show relevant characteristics of the hydrologic system behavior and could be useful for the development and evaluation of models.

The reproduction of hydrologic responses and their time-lagged dependencies (connectivity) with driving variables seems to be an attractive criterion to be considered in future research. Time lagged mutual information (with the issues related to the memory effects), or even a more complex measure such as the transfer entropy (Bennett et al., 2019), may serve as an additional metric to evaluate the consistency and adequacy of hydrologic models.

Another potential application based on the results here presented refer to the construction of tools for real-time streamflow forecasts because many of these tools use a combination of lagged precipitation and streamflow data to perform the predictions (Gómez et al., 2019; Jahandideh-Tehrani et al., 2020; Lv et al., 2020). Mutual information measures can be useful to define the 'optimal' time lags to be considered.

Our dataset is relatively short (10 years of daily data) to detect eventual variations in the way that the hydrologic variables interact with each other over time, such as done in the study by Franzen et al. (2020). Undoubtedly, understanding how connectivity measures change over time and how environmental changes (e.g., land use, climate variability) influence the connectivity between hydrological processes are relevant topics to be addressed in future studies.

5 Conclusions

Two causal discovery methods (Granger's causality and mutual information) were assessed and used to characterize pairwise time-lagged dependencies of daily data observed in a small sub-tropical watershed. The water balance closure was verified considering the groundwater storage dynamics. Unsurprisingly, statistically significant causal interactions were confirmed between most of the water balance components. The analysis conducted allowed us to characterize temporal

interdependencies with long time windows, to identify some patterns, to explore the strength of connectivity between hydrological processes, and to estimate the memory timescale of variables.

Despite these capabilities, further studies are required to constrain the possible interpretations and to create a connection between statistical results and the hydrologic system dynamics. The option of using data from a watershed, an open system with many associated uncertainties, and insufficient characterization, made it unfeasible to advance in that sense. This is an opportunity, however, to advance combined field hydrology and modeling studies to move from abstract statistical results to objective physical interpretations.

Throughout this paper, the potential of causal methods in characterizing the connectivity between variables was evidenced. Real-world applications, with examples of how such methods can contribute to hydrological science and applied hydrology, considering data limitations, seem to be essential to engage the community. In future studies, one can test, for instance, if a model accurately reproduces the connectivity patterns found in observed data, even when the reasons behind the patterns are unknown. Similar approaches can improve the adequacy and performance of predictive tools and constrain uncertainties.

Data Availability

Hydrologic data used in this research is publicly available at

https://github.com/kalylgc/causebro/tree/Data/

Funding Information

K.G.C. was supported by the São Paulo Research Foundation FAPESP, grant # 2018/19222-7. This research was financed in part by the FAPESP, Process # 2015/03806-1, and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Finance Code 001.

References

Allen, R. G. (1998). FAO Irrigation and Drainage Paper Crop by. *Irrigation and Drainage*, *300*(56), 300. https://doi.org/10.1016/j.eja.2010.12.001

Alvares, C. A., Stape, J. L., Sentelhas, P. C., De Moraes Gonçalves, J. L., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, *22*(6), 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Amblard, P. O., & Michel, O. J. J. (2013). The relation between granger causality and directed information theory: A review. *Entropy*, *15*(1), 113–143. https://doi.org/10.3390/e15010113

Aráujo, L. M., França, A. B., & Potter, P. E. (1999). Hydrogeology of the Mercosul aquifer system in the Paraná and Chaco-Paraná Basins, South America, and comparison with the Navajo-Nugget aquifer system, USA. *Hydrogeology Journal*, *7*(3), 317–336. https://doi.org/10.1007/s100400050205

Barnett, L., & Seth, A. K. (2015). Granger causality for state-space models. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *91*(4), 1–5. https://doi.org/10.1103/PhysRevE.91.040101

Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy Are equivalent for gaussian variables. *Physical Review Letters*, *103*(23), 2–5. https://doi.org/10.1103/PhysRevLett.103.238701

Batista, L. V., Gastmans, D., Sánchez-Murillo, R., Farinha, B. S., dos Santos, S. M. R., & Kiang, C. H. (2018). Groundwater and surface water connectivity within the recharge area of Guarani aquifer system during El Niño 2014–2016. *Hydrological Processes*, *32*(16), 2483–2495. https://doi.org/10.1002/hyp.13211

Bennett, A., Nijssen, B., Ou, G., Clark, M., & Nearing, G. (2019). Quantifying Process Connectivity With Transfer Entropy in Hydrologic Models. *Water Resources Research*, *55*(6), 4613–4629. https://doi.org/10.1029/2018WR024555

Besbes, M., & De Marsily, G. (1984). From infiltration to recharge: use of a parametric transfer function. *Journal of Hydrology*, *74*(3–4), 271–293. https://doi.org/10.1016/0022-1694(84)90019-2

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

Cabrera, M. C. M., Anache, J. A. A., Youlton, C., & Wendland, E. (2016). Performance of evaporation estimation methods compared with standard 20 m2 tank. *Revista Brasileira de Engenharia Agricola e Ambiental*, *20*(10), 874–879. https://doi.org/10.1590/1807-1929/agriambi.v20n10p874-879

Coelho, C. A. S., de Oliveira, C. P., Ambrizzi, T., Reboita, M. S., Carpenedo, C. B., Campos, J. L. P. S., et al. (2016). The 2014 southeast Brazil austral summer drought: regional scale mechanisms and teleconnections. *Climate Dynamics*, *46*(11–12), 3737–3752. https://doi.org/10.1007/s00382-015-2800-1

Coutinho, J. V., Porsani, J. L., Elis, V. R., Santos, V. R. N., Ustra, A. T., & Wendland, E. (2020). Applications of geophysical techniques to improve a groundwater conceptual model in an outcrop area of the Guarani Aquifer System, in Brazil. *Environmental Earth Sciences*, *79*(18), 1–11. https://doi.org/10.1007/s12665-020-09163-4

Cover, T. M., & Thomas, J. A. (2005). *Elements of Information Theory* (2nd ed.), Hoboken, NJ: John Wiley & Sons Inc. https://doi.org/10.1002/047174882X

Dey, P., & Mujumdar, P. P. (2018). Multiscale evolution of persistence of rainfall and streamflow. *Advances in Water Resources*, *121*(August), 285–303. https://doi.org/10.1016/j.advwatres.2018.08.018

Didan, K. (2015). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2020-11-23 from https://doi.org/10.5067/MODIS/MOD13Q1.006

Dirks, K. N., Hay, J. E., Stow, C. D., & Harris, D. (1998). High-resolution studies of rainfall on Norfolk Island Part II : Interpolation of rainfall data. *Journal of Hydrology, 208(3-4)*, 187–193. https://doi.org/10.1016/S0022-1694(98)00155-3

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2007). On the importance of including vegetation dynamics in Budyko's hydrological model. *Hydrology and Earth System Sciences, 11*(2), 983–995. https://doi.org/10.5194/hess-11-983-2007

Duncan, H. P. (2019). Baseflow separation – A practical approach. *Journal of Hydrology*, *575*, 308–313. https://doi.org/10.1016/j.jhydrol.2019.05.040

Franzen, S. E., Farahani, M. A., & Goodwell, A. E. (2020). Information Flows: Characterizing Precipitation-Streamflow Dependencies in the Colorado Headwaters With an Information Theory Approach. *Water Resources Research*, *56*(10), 1–17. https://doi.org/10.1029/2019WR026133

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

Freedman, D., & Diaconis, P. (1981). On the histogram as a density estimator:L2 theory. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, *57*(4), 453–476. https://doi.org/10.1007/BF01025868

Gençağa, D. (2018). Transfer entropy. *Entropy*, *20*(4), 1–4. https://doi.org/10.3390/e20040288

Glenn, E. P., Neale, C. M. U., Hunsaker, D. J., & Nagler, P. L. (2011). Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrological Processes*, *25*(26), 4050–4062. https://doi.org/10.1002/hyp.8392

Gómez, D., Melo, D. C. D., Rodrigues, D. B. B., Xavier, A. C., Guido, R. C., & Wendland, E. (2018). Aquifer Responses to Rainfall through Spectral and Correlation Analysis. *Journal of the American Water Resources Association*, *54*(6), 1341–1354. https://doi.org/10.1111/1752-1688.12696

Gómez, D., Wendland, E., & Melo, D. C. D. (2020). Empirical rainfall-based model for defining baseflow and dynamical water use rights. *River Research and Applications*, *36*(2), 189–198. https://doi.org/10.1002/rra.3565

Goodwell, A. E., & Kumar, P. (2017). Temporal information partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *Water Resources Research*, 53(7), 5920-5942. http://doi.org/10.1002/2016WR020216

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. J. W. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Graves, T., Gramacy, R., Watkins, N., & Franzke, C. (2017). A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA, 1951-1980. *Entropy*, *19*(9), 1–21. https://doi.org/10.3390/e19090437

Huang, X., Maçaira, P. M., Hassani, H., Cyrino Oliveira, F. L., & Dhesi, G. (2019). Hydrological natural inflow and climate variables: Time and frequency causality analysis. *Physica A: Statistical Mechanics and Its Applications*, *516*, 480–495. https://doi.org/10.1016/j.physa.2018.09.079

Jiang, P., & Kumar, P. (2019). Information transfer from causal history in complex system dynamics. *Physical Review E*, *99*(1), 1–13. https://doi.org/10.1103/PhysRevE.99.012306

Kamble, B., Kilic, A., & Hubbard, K. (2013). Estimating crop coefficients using remote sensing-based vegetation index. *Remote Sensing*, *5*(4), 1588–1602. https://doi.org/10.3390/rs5041588

Kirchheim, R. E., Gastmans, D., Chang, H. K., & Gilmore, T. E. (2019). The use of isotopes in evolving groundwater circulation models of regional continental aquifers: The case of the Guarani Aquifer System. *Hydrological Processes*, *33*(17), 2266–2278. https://doi.org/10.1002/hyp.13476

Koutsoyiannis, D. (2006). Nonstationarity versus scaling in hydrology. *Journal of Hydrology*, 324(1–4), 239–254. https://doi.org/10.1016/j.jhydrol.2005.09.022

Kumar, P., & Gupta, H. V. (2020). Debates—Does Information Theory Provide a New Paradigm for Earth Science? *Water Resources Research*, 56(2), 1–10. https://doi.org/10.1029/2019WR026398

Li, S., Xu, J., Chen, G., Lin, L., Zhou, D., & Cai, D. (2017). The characterization of hippocampal theta-driving neurons-A time-delayed mutual information approach. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-05527-2

Li, S., Xiao, Y., Zhou, D., & Cai, D. (2018). Causal inference in nonlinear systems: Granger causality versus time-delayed mutual information. *Physical Review E*, *97*(5), 1–9. https://doi.org/10.1103/PhysRevE.97.052216

Lv, N., Liang, X., Chen, C., Zhou, Y., Li, J., Wei, H., & Wang, H. (2020). A long Short-Term memory cyclic model with mutual information for hydrology forecasting: A Case study in the xixian basin. *Advances in Water Resources*, *141*(May). https://doi.org/10.1016/j.advwatres.2020.103622

Machado, A. R., Wendland, E., & Krause, P. (2016). Hydrologic Simulation for Water Balance Improvement in an Outcrop Area of the Guarani Aquifer System. *Environmental Processes*, *3*(1), 19–38. https://doi.org/10.1007/s40710-016-0128-4

Marengo, J. A., Nobre, C. A., Seluchi, M. E., Cuartas, A., Alves, L. M., Mendiondo, E. M., et al. (2015). A seca e a crise hídrica de 2014-2015 em São Paulo. *Revista USP*, (106), 31. https://doi.org/10.11606/issn.2316-9036.v0i106p31-44

McGraw, M. C., & Barnes, E. A. (2018). Memory matters: A case for granger causality in climate variability studies. *Journal of Climate*, *31*(8), 3289–3300. https://doi.org/10.1175/JCLI-D-17-0334.1

Nagler, P. L., Glenn, E. P., Nguyen, U., Scott, R. L., & Doody, T. (2013). Estimating riparian and agricultural actual evapotranspiration by reference evapotranspiration and MODIS enhanced vegetation index. *Remote Sensing*, *5*(8), 3849–3871. https://doi.org/10.3390/rs5083849

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. lin. (2020). Evaluation of Methods for Causal Discovery in Hydrometeorological Systems. *Water Resources Research*, *56*(7), 1–22. https://doi.org/10.1029/2020WR027251

Papagiannopoulou, C., Miralles, Di. G., Decubber, S., Demuzere, M., Verhoest, N. E. C., Dorigo, W. A., & Waegeman, W. (2017). A non-linear Granger-causality framework to investigate climate-vegetation dynamics. *Geoscientific Model Development*, *10*(5), 1945–1960. https://doi.org/10.5194/gmd-10-1945-2017

Pelletier, J. D., & Turcotte, D. L. (1997). Long-range persistence in climatological and hydrological time series: Analysis, modeling and application to drought hazard assessment. *Journal of Hydrology*, *203*(1–4), 198–208. https://doi.org/10.1016/S0022-1694(97)00102-9

Pompeo, C. A. (1990). Balanço hídrico da zona não-saturada do solo na bacia do Ribeirão da Onça (SP), (Doctoral dissertation). São Carlos, SP: University of São Paulo.

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

Ruddell, B. L., & Kumar, P. (2009). Ecohydrologic process networks: 1. Identification. *Water*

Resources Research, *45*(3), 1–23. https://doi.org/10.1029/2008WR007279

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), 1–13. https://doi.org/10.1038/s41467-019-10105-3

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*(2), 461–464. https://doi.org/10.1103/PhysRevLett.85.461

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, (Scipy), 92–96. https://doi.org/10.25080/majora-92bf1922-011

Singh, N. K., & Borrok, D. M. (2019). A Granger causality analysis of groundwater patterns over a half-century. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-49278-8

Tomasella, J., Hodnett, M.G., Cuartas, L.A., Nobre, A.D., Waterloo, M.J., & Oliveira, S.M. (2008). The water balance of an Amazonian micro-catchment: the effect of interannual variability of rainfall on hydrological behaviour. *Hydrological Processes*, 22(13), 2133-2147. https://doi.org/10.1002/hyp.6813

Wegehenkel, M. (2009). Modeling of vegetation dynamics in hydrological models for the assessment of the effects of climate change on evapotranspiration and groundwater recharge. *Advances in Geosciences*, *21*, 109–115. https://doi.org/10.5194/adgeo-21-109-2009

Weijs, S. V., Schoups, G., & Van De Giesen, N. (2010). Why hydrological predictions should be evaluated using information theory. *Hydrology and Earth System Sciences*, *14*(12), 2545–2558. https://doi.org/10.5194/hess-14-2545-2010

Wendland, E., Barreto, C., & Gomes, L. H. (2007). Water balance in the Guarani Aquifer outcrop zone based on hydrogeologic monitoring. *Journal of Hydrology*, *342*(3–4), 261–269. https://doi.org/10.1016/j.jhydrol.2007.05.033

Wendland, E., Gomes, L. H., & Troeger, U. (2015). Recharge contribution to the Guarani aquifer system estimated from the water balance method in a representative watershed. *Anais Da Academia Brasileira de Ciencias*, *87*(2), 595–609. https://doi.org/10.1590/0001-3765201520140062

Wilks, S. S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, *9*(1), 60–62. https://doi.org/10.1214/aoms/1177732360

Zimmermann, B., Elsenbeer, H., & Moraes, J. M. (2006). The influence of land-use changes on soil hydraulic properties: Implications for runoff generation. *Forest Ecology and Management*, *222*(1–3), 29–38. https://doi.org/10.1016/j.foreco.2005.10.070