A Time-Varying Causality Formalism Based on the Liang–Kleeman Information Flow for Analyzing Directed Interactions in Nonstationary Climate Systems

DANIEL FIIFI TAWIA HAGAN AND GUOJIE WANG

Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

X. SAN LIANG

School of Marine Sciences, and School of Atmospheric Sciences, Nanjing Institute of Meteorology, Nanjing, Jiangsu, China

HAN A. J. DOLMAN

Department of Earth Sciences, Faculty of Science, Free University Amsterdam, Amsterdam, Netherlands

(Manuscript received 4 January 2019, in final form 29 June 2019)

ABSTRACT

The interaction between the land surface and the atmosphere is of significant importance in the climate system because it is a key driver of the exchanges of energy and water. Several important relations to heat waves, floods, and droughts exist that are based on the interaction of soil moisture and, for instance, air temperature and humidity. Our ability to separate the elements of this coupling, identify the exact locations where they are strongest, and quantify their strengths is, therefore, of paramount importance to their predictability. A recent rigorous causality formalism based on the Liang–Kleeman (LK) information flow theory has been shown, both theoretically and in real-world applications, to have the necessary asymmetry to infer the directionality and magnitude within geophysical interactions. However, the formalism assumes stationarity in time, whereas the interactions within the land surface and atmosphere are generally nonstationary; furthermore, it requires a sufficiently long time series to ensure statistical sufficiency. In this study, we remedy this difficulty by using the square root Kalman filter to estimate the causality based on the LK formalism to derive a time-varying form. Results show that the new formalism has similar properties compared to its time-invariant form. It is shown that it is also able to capture the time-varying causality structure within soil moisture–air temperature coupling. An advantage is that it does not require very long time series to make an accurate estimation. Applying a wavelet transform to the results also reveals the full range of temporal scales of the interactions.

1. Introduction

The land surface plays an essential role in the climate system (e.g., Fischer et al. 2007; Miralles et al. 2014). A number of studies have demonstrated that the land–atmosphere interactions, especially soil moisture interactions with the atmosphere, have the potential to affect the occurrence of heat waves (Stéfanon et al. 2014), drought (Ciabatta et al. 2015; Roundy and Wood 2015), rainfall (Taylor et al. 2012; Tuttle and Salvucci 2015), and floods (Massari et al. 2014; Saini et al. 2016). Therefore, predictability of climate extremes is expected to exist in the interactions of key processes between the land and the atmosphere (Berg et al. 2014; Zheng et al. 2015). In recent years, identifying the hot spots of land–atmosphere interactions and quantifying the magnitude of such effects on climate variations and extremes have attracted a lot of attention (Casagrande et al. 2015; Papagiannopoulou et al. 2017; Tuttle and Salvucci 2015).

The approaches for quantifying land–atmosphere interactions include mainly, among others, numerical experimentation and statistical diagnosis. In recent years, many numerical studies have been conducted with the aim of improving the understanding of land–atmosphere interactions with a focus on soil moisture feedbacks on the overlying atmosphere. Koster et al. (2004) and Guo et al. (2006) identified the hot spots of soil moisture

Corresponding author: Guojie Wang, gwang_nuist@163.com

DOI: 10.1175/JCLI-D-18-0881.1

feedbacks on subsequent precipitation and air temperature over the globe based on numerical experiments of 12 general circulation models (GCMs). Such feedbacks were found to be strongest in the transitional regions between dry and wet climates with, however, rather significant discrepancies among the outputs from different GCMs. Fischer et al. (2007) simulated the European heat wave in the summer of 2003 and found the heat anomalies in the Mediterranean region reduced by about 40% if the interactions between the land and atmosphere were not adequately taken into account. There also exist a number of other numerical studies showing the importance of land–atmosphere interactions for atmospheric prediction (Koster et al. 2004; Miralles et al. 2014; Tuttle and Salvucci 2015).

Although numerical climate models offer a good alternative for studying the land–atmosphere interactions experimentally, significant discrepancies exist between the results from different models due to the use of different parameterization schemes (Flato et al. 2013; Zhao and Li 2015). Therefore, considerable efforts have also been put into statistical approaches for identifying the land–atmosphere interactions from observations: identifying where they exist as well as estimating their magnitude (Miralles et al. 2012; Notaro 2008; Zhang et al. 2008). A common practice is the use of time-lagged correlation (Ford et al. 2015; Wu and Zhang 2015), even though it is generally known that correlation does not bear the required asymmetry and directedness to indicate causality (Liang 2014). The latter implies that correlation analysis does not show the direction of the interactions. Thus, disentangling the direction of the interactions among time series has been of interest for a long time in many disciplines. Another potential problem with time-lagged correlation is that when the foreknowledge of the processes is not sufficient, it is very difficult to differentiate a lag from an advance, especially, when the processes are periodic. Therefore, several studies have been devoted to the development of reliable statistical tools for this purpose (Dajčman 2013; Casagrande et al. 2015; Notaro 2008; Paluš 2014; Papana et al. 2016). The Granger causality, first suggested by Clive Granger (Granger 1969), is one of the most widely accepted tools for studying directed interactions, in particular, the linear relations of cause and effect, among time series. It has been applied successfully to study the cause–effect relations between climate variables such as the soil moisture–precipitation interaction (Tuttle and Salvucci 2015). More recently, a nonlinear form of Granger causality based on a random forest machine learning algorithm was proposed and used to analyze the causal relations between several climate factors and vegetation (Papagiannopoulou et al. 2017).

Another method that has gained much popularity is the equilibrium feedback analysis (EFA), which was originally proposed by Frankingnoul and Hasselmann (1977) for studying coupled systems within the climate. It has since been applied to the study of soil moisture feedbacks on precipitation to investigate the hot spots of soil moisture–precipitation feedbacks over the globe using various datasets (Notaro 2008). A spatiotemporal analysis tool called the coupled manifold is also used for this purpose (Navarra and Tribbia 2005), although it is not strictly defined in the concept of causality. It seeks a solution for a system as a linear function of another system, and has also been successfully applied to several problems of land surface atmosphere interactions (Catalano et al. 2016).

Recently, Liang (2014) proposed a rigorous nonparametric formalism of causality measurement for identifying the cause–effect relations between any given time series, based on the theory of Liang–Kleeman (LK) information flow (LK causality; Liang 2016). Compared to other tools, the LK causality is rather easy to compute, involving only the sample covariance of the given time series. Additionally, since it is derived from first principles and mathematically rigorous, it avoids spurious causalities that are sometimes found in the Granger causal test and other statistical formalisms. Using a large number of simulations, Liang (2014) showed that this causality formalism is able to unambiguously obtain predetermined causalities between the simulated time series. Stips et al. (2016) used the LK causality to investigate the causal relations between global surface temperature and the radiative forcing, finding that greenhouse gases were the main causal drivers of the recent warming, but this cause–effect direction was reversed on paleoclimate time scales. Among other applications of the Liang–Kleeman information flow in recent years, Liang (2014) used it to investigate the cause–effect relation between two climate modes, El Niño and the Indian Ocean dipole (IOD), both of which have been linked to some climate extreme and socioeconomic impacts. The LK causality is hence identified to provide an easy but efficient tool for studying the land–atmosphere interactions. However, it requires stationarity for the time series in question. It is well known that the land–atmosphere interactions vary with seasons due to the changing boundary layer properties. Therefore, there is a need to further develop the LK causality in order to cope with time-varying properties of land surface atmosphere interactions.

The Kalman filter is an established approach to obtain the changing model parameters with time. It infers parameters of interest from inaccurate and uncertain observations by minimizing the mean-square error

between the estimated parameters and the observations (Grewal and Andrews 2001). It has been further developed into different forms and applied in both linear and nonlinear problems (Ghorbanidehno et al. 2015; Houtekamer and Zhang 2016; Kiani and Pourtakdoust 2014). We use the Kalman filter here to estimate the time-varying parameters for the LK information flow-based causal inference.

Many time series in geophysics contain dominant periodic components, which can vary in both amplitude and frequency over long periods. Such periodic components sometimes act in the land–atmosphere interactions, leading to signals in the frequency domain. Identifying the signals of directed interactions in the frequency domain may help to understand the underlying mechanisms of land–atmosphere interactions. Therefore, there is also a need to identify directed interactions in the frequency domain. Wavelet analysis, widely used in analyzing time series of geophysics (Casagrande et al. 2015; Wang et al. 2017), is such a tool to solve this problem by decomposing a time series into time-frequency distribution.

In this study, the Liang causality formalism is integrated with the Kalman filter and wavelet analysis to provide a statistical tool for investigating the land–atmosphere interactions in the time-frequency domain. Section 2 gives a description of the formalism, section 3 validates the formalism with synthetic models, section 4 applies it to a specific case study of land–atmosphere interaction over China, and section 5 presents a summary and discussion of the formalism and its potential applications.

2. The time-varying causal inference

a. Formalism of the information flow-based causality

Information flow, sometimes referred to as information transfer in the literature, is essentially the transfer of information through some processes between two or more entities within a dynamical system (Liang 2013). The transfer of information from a source to a receiver, which may reflect a cause–effect relationship (i.e., who drives whom) is a problem that can be found in many disciplines such as neuroscience, financial economics, climate interactions, and network dynamics, just to mention a few. After more than 30 years of research, a consensus has been reached that information flow is logically associated with causation and bears implications of predictability and uncertainty propagation (see Liang 2013 and references therein). Liang (2008, 2016) hence argued that, in this way, causality can be rigorously formulated, rather than axiomatically proposed as an ansatz; particularly, the causality between two time series can be easily obtained in a quantitative sense Liang (2014). Recently, Liang (2016) found that causality is a real physical notion and can be rigorously derived and quantified with the notion of the Liang–Kleeman information flow (Liang 2008); particularly, the causality between two time series can be easily obtained in a quantitative sense (Liang 2014). The following is a brief introduction.

Consider a d-dimensional stochastic system

$$d\mathbf{X} = \mathbf{F}(\mathbf{X}, t)\, dt + \mathbf{B}(\mathbf{X}, t)\, d\mathbf{W}, \tag{1}$$

where $\mathbf{F}$ is defined as the vector of drift coefficients, $\mathbf{B} = (b_{ij})$ is a matrix of stochastic perturbation coefficients, and $\mathbf{W}$ is a vector of a Wiener process. If two components, for instance, X_1 and X_2 are considered, the rate of information flow from X_2 to X_1, is the amount of entropy contributed by X_2 to the evolution of H_1, the marginal entropy of X_1. This can be obtained by finding the total rate of change of H_1, minus the same change rate but with the effect from X_2 instantaneously excluded. Liang (2008, 2016) proved that for a 2D system,

$$T_{2\to 1} = -E\left[\frac{1}{\rho_1}\frac{\partial(F_1\rho_1)}{\partial x_1}\right] + \frac{1}{2}E\left[\frac{1}{\rho_1}\frac{\partial^2(b_{11}^2 + b_{12}^2)\rho_1}{\partial x_1^2}\right], \tag{2}$$

where ρ_1 is the marginal probability density of X_1, and E is the mathematical expectation. For arbitrary dimensional system, the result is referred to Liang (2016). A unique property of this measure of information flow rate as obtained in Eq. (2) is its asymmetry (between $T_{2\to 1}$ and $T_{1\to 2}$), and particularly, the "principle of nil causality," which states that, if the evolution of an event is independent of another, then the causality from the latter to the former is zero. Thus, ideally, a nonzero $T_{2\to 1}$ implies that X_2 is causal to X_1, where the magnitude of $T_{2\to 1}$ is the strength of the causality. A zero value strictly implies that X_2 is not causal to X_1. (In practice, the statistical significance should be tested to infer causality). Refer to Liang (2016) for a more formally and quantitatively stated principle of nil causality, which the Granger causality test and transfer entropy analysis fail to meet in many situations (Granger 1980).

In the case of a linear system: $\mathbf{F} = \mathbf{f} + \mathbf{AX}$, where $\mathbf{f} = (f_1, f_2)^{\mathrm{T}}$, $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, then if originally (X_1, X_2) has a bivariate normal distribution, it will always be so. Given the mean $\boldsymbol{\mu} = (\mu_1, \mu_2)^{\mathrm{T}}$ and covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$, their evolutions are as follows:

$$\frac{d\boldsymbol{\mu}}{dt} = \mathbf{f} + \mathbf{A}\boldsymbol{\mu}, \quad \text{and} \tag{3}$$

$$\frac{d\boldsymbol{\Sigma}}{dt} = \mathbf{A}\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}} + \mathbf{B}\mathbf{B}^{\mathrm{T}}. \tag{4}$$

When $\mathbf{B} = (b_{ij})$ is constant, the last term of Eq. (2) disappears. Substitution of

$$F_1 = f_1 + a_{11}X_1 + a_{12}X_2, \quad \text{and} \tag{5}$$

$$\rho_1 = \frac{1}{\sqrt{(2\pi\sigma_1)}} \exp\left[-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2}\right], \tag{6}$$

into Eq. (2) for F_1 and ρ_1 produces, after a series of algebraic multiplications (Liang 2014, 2008),

$$T_{2\to1} = \frac{\sigma_{12}}{\sigma_{11}} a_{12}, \tag{7}$$

where σ_{ij} is derived by solving Eq. (4). The parameters $\mathbf{f}$, $\mathbf{A}$, and $\mathbf{B}$ can be estimated through maximum likelihood estimation; with them, the information flow from X_2 to X_1 can be obtained in terms of covariances (Liang 2014):

$$T_{2\to1} = \frac{C_{12}}{C_{11}} \frac{-C_{12}C_{1,d1} + C_{11}C_{1,d1}}{C_{11}C_{22} - C_{12}^2}, \tag{8}$$

where C_{ij} is the sample covariance between X_i and X_j, $C_{i,dj}$ the covariance between X_i and X_{dj} and a derived series from X_j using the formula $X_{dj} = [X_j(t + k\Delta t) - X_j(t)]/(k\Delta t)$, where Δt is the time step size, and k some integer; see Liang (2014) for further details. Likewise, the information flow from X_1 to X_2 can be obtained by switching the indices:

$$T_{1\to2} = \frac{C_{21}}{C_{22}} \frac{-C_{21}C_{2,d2} + C_{22}C_{1,d2}}{C_{22}C_{11} - C_{21}^2},$$

b. Kalman filter

From the above Eq. (8), estimating the causality invokes an assumption of time stationarity for the computation of the covariances. To tackle the challenge of computing a time-varying causality based on Eq. (2) and, particularly, its linear version [Eq. (7)] as rigorously established by Liang (2008), a Kalman filter is employed. The main goal of using this Kalman filter approach is to estimate the required covariances at each time step. Many studies have used the Kalman filter to obtain covariance estimates (e.g., Asl and Pourtakdoust 2007; Gonzalez et al. 2014). A Kalman filter includes two parts: the estimation part, which calculates the state estimate $\mathbf{x}$, and the gain part, which computes the covariance of the state $\mathbf{P}$.

The standard Kalman filter is formulated with a linear stochastic system, shown here as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_{k-1}, \\ \mathbf{y}_k &= \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, \end{aligned} \tag{9}$$

where $\mathbf{x}$ is the system state vector, $\mathbf{y}$ is the measurement vector, and $\mathbf{u}$ is the input vector. The parameters $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{H}$ are transition matrices of appropriate dimensions, with $\mathbf{v}$ and $\mathbf{w}$ being mutually independent and zero mean measurement and process noise, respectively in this case, and

$$\begin{aligned} \mathbf{E}[\mathbf{w}_k\mathbf{w}_k^{\mathrm{T}}] &= \mathbf{Q}, \\ \mathbf{E}[\mathbf{v}_k\mathbf{v}_k^{\mathrm{T}}] &= \mathbf{R}, \end{aligned} \tag{10}$$

where $\mathbf{Q}$ is the process noise covariance and $\mathbf{R}$ is the measurement noise covariance. Just as in previous studies, it was found that the estimation of $\mathbf{Q}$ and $\mathbf{R}$ was quite difficult, especially for $\mathbf{Q}$ (Berg et al. 2014; Odelson et al. 2006; Saha et al. 2011; Mohan et al. 2015), due to our inability to observe the process being estimated. Since the computation of the causality is based on sample covariances, it is very important to ensure that the estimated covariances are very accurate. Therefore, we do not assume them to be constant in time, but we compute them for each time step since in reality, they will not remain constant but change. A common practice, which is adopted here, is they are first obtained offline (outside of the Kalman filter iteration) for each time step; the obtained results are then used for further computations in the iteration process as seen in Eqs. (12) and (14).

An updating equation is used to compute an a posteriori state $\mathbf{x}_k$, an estimate that is a linear combination of the weighted difference between the measurement prediction $\mathbf{H}\mathbf{x}_k$ and an actual measurement $\mathbf{y}_k$, and an a priori $\mathbf{x}_k^-$. Combining the a priori estimate and the measurement data, we write $\mathbf{x}_k$ as

$$\mathbf{x}_k = \mathbf{x}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}\mathbf{x}_k^-). \tag{11}$$

The statistically optimal Kalman gain $\mathbf{K}_k$ is calculated to use in the correction stage. It minimizes the risk (loss) of the estimate error covariance, defined as $\mathbf{P} = \mathbf{E}[\mathbf{e}\mathbf{e}^{\mathrm{T}}]$, which is the value of expectation of the squared magnitude of the posterior state estimation error. The Kalman gain is computed as in the following:

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_k^-\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}. \tag{12}$$

Hence, the updated $\mathbf{P}$ becomes the solution to the equation below, using the characteristics of the measurement noise in Eq. (9):

$$\mathbf{P}_k = (1 - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^-. \tag{13}$$

The following projection equations of the state estimate $\mathbf{x}$ and covariance $\mathbf{P}$ are used in the prediction stage,

$$\begin{aligned} \mathbf{x}_k^- &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k, \\ \mathbf{P}_k^- &= \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}_k. \end{aligned} \tag{14}$$

Within the gain loop of the standard Kalman filter, the Kalman gain, which is an error correction feedback of the estimation loop, is estimated. An accurate gain, therefore, provides accurate estimations of the state parameter from round-off errors, a priori and noise estimation errors. Since no feedback exists in the gain loop, it results in unchecked accumulated computer round-off errors from the state variance–covariance computation, which happens to be a parameter of importance to this study. This creates numerical instabilities and imprecision within the Kalman (optimal) filtering process. Several studies have been committed to solving these problems, and decomposition of the covariance matrix (the process of breaking down a matrix or its expression into an equivalent product of factors) has been proposed as direction to solving the above mentioned problem (Roncero 2014; Sastry 1971).

Several methods of decomposition are available from numerous studies (Bellantoni and Dodge 1967; Gonzalez et al. 2014; Roncero 2014; Thornton 1976). In this study, we employ the Bierman–Thornton algorithm (see appendix A), which uses modified Cholesky factors of the state variance–covariance matrix, since the modified Cholesky decomposition has been found to be more robust than the classical one (Maoudj et al. 2013). It is acts as a square root filter, though it does not use any square root operations in its implementation. However, its ability to compute positive definite matrices, which have a triangular matrix square root property, makes it an optimal choice.

The LK causality formalism, as seen in Eq. (8), involves sample covariances, which are usually directly computed from given time series. We here are about to estimate the covariances with our newly preferred estimation methodology, that is, the square root Kalman filter based on the Bierman–Thornton algorithm (Bierman and Thornton 1976). In this way, Eq. (8) and its converse direction become

$$\mathbf{T}_{2\to 1t} = \frac{P_{12}}{P_{11}} \frac{-P_{12}P_{1,d1} + P_{11}P_{2,d1}}{P_{11}P_{22} - P_{12}^2},$$
$$\mathbf{T}_{1\to 2t} = \frac{P_{21}}{P_{22}} \frac{-P_{21}P_{2,d2} + P_{22}P_{1,d2}}{P_{22}P_{11} - P_{21}^2}. \tag{15}$$

where $\mathbf{P}$ is the resulting covariance matrix at each time step of the iteration $[\mathbf{P}_k = (1 - \mathbf{K}_k\mathbf{H}_k)(\mathbf{U}_k^- \mathbf{D}_k^- \mathbf{U}_k^{-\mathrm{T}})]$. We provide two formalisms of this time-varying causality, both based on the Bierman–Thornton algorithm, but in the second case, the transition matrix functions in Eq. (9) are first determined by fitting an autoregressive model to the data and performing a Kalman filter on it. Once they are determined through an iteration of this process, the covariance is estimated as shown in the previous equations. The formalism is referred to as TvLK throughout the rest of this paper.

c. The TvLK framework

The algorithm begins with an offline computation of the process and measurement noise covariances, $\mathbf{Q}$ and $\mathbf{R}$, respectively. This is because it is expected that in practice, $\mathbf{Q}$ and $\mathbf{R}$ would change with each time step. The filter then starts off with the computation of the Kalman gain $\mathbf{K}$ through to the updated covariance matrix $\mathbf{P}$, at each time step, which is where the causality is computed. At the measurement and temporal updates, the Bierman and Thornton algorithms are applied to make the computation more stable. This process is iterated through all the time steps. Because of the ability of the Kalman filter to detect change, it is expected that in using it to estimate the Liang–Kleeman causality, immediate changes along time will be preserved. As a learning algorithm the Kalman filter estimations are expected to improve along time when applied to the formalism. It should be noted that several experiments with this formalism have shown that obtaining an accurate result is highly dependent on the offline computation of $\mathbf{Q}$ and $\mathbf{R}$. A detailed diagrammatic scheme of the framework of the TvLK is given in appendix B.

d. Wavelet transforms

The wavelet transform is mainly used to decompose a signal into parts with different frequencies. Here, the continuous wavelet form is used as a postprocessing tool to decompose the output of the TvLK estimates to obtain a multi-time-scale causal structure as a distribution of time and frequency. This is especially important in instances where separated signals of a coupled system are found to occur within the same period in a time-varying distribution. The continuous wavelet is defined as the convolution of $\mathbf{x}_n$ with a translated and scaled version of

$$W_n(s) = \sum_{n'=0}^{N=1} x_{n'} \psi^* \frac{[(n'-n)\,\delta t]}{s}, \tag{16}$$

where x_n is a discrete sequence, $\psi_0(\eta)$ is a wavelet function, η is a nondimensional ''time'' parameter, the (*) symbolizes the complex conjugate, wavelet scale s and localized time index s. The obtained wavelet transform $[W_n(s)]$ is also complex, which can be divided into the real and imaginary parts.

It is expected that by applying this to the time-varying causality, it will become possible to visualize the separated causality signals over a multiscale time-frequency distribution. Following Torrence and Compo (1998), the

significance level of 5% is computed for each case using Monte Carlo simulations. There are several types of wavelet functions in use, including the most common ones, Morlet, Paul, and difference of Gaussian (DoG) functions. While the Morlet and Paul are complex functions, the DoGs are real. The choice of a wavelet function is left to the user's discretion. Readers are referred to Torrence and Compo (1998) for a guide to using the transform and its different available functions. We use the bump wavelet as part of the wavelet package of MATLAB, in this study. A detailed description can be found in Vialatte et al. (2009). The proposed formalism with the wavelet decomposition is termed TFvLK hereafter. It is important to state that this is simply the TvLK formalism with a spectral analysis extension. A summary and flow diagram of the calculation procedure is given in appendix B.

3. Simulation studies

Numerical experiments are here conducted in order to assess if the proposed TvLK and TFvLK formalisms can properly determine the directed interactions. For this purpose, two synthetic models are generated from numerical simulations. The specific aims of these synthetic model tests are to verify 1) if the proposed TvLK formalism based on Kalman filter can correctly reproduce the predefined time-varying causal structure, and 2) if this causal structure can be determined within the predefined frequencies when decomposed into the time-frequency domain by means of wavelet transform. Based on the theory of Liang–Kleeman information flow, the unit of derived causality is in nats per time.

a. *Synthetic model 1*

$$\begin{aligned}\mathbf{x}_1(t) &= 0.36\mathbf{x}_1(t-1) - 0.41\mathbf{x}_1(t-2)\\ &\quad + c_{2\to1}(t)\mathbf{x}_2(t-1) + \varepsilon_1(t),\\ \mathbf{x}_2(t) &= 0.36\mathbf{x}_2(t-1) - 0.41\mathbf{x}_2(t-2)\\ &\quad + c_{1\to2}(t)\mathbf{x}_1(t-1) + \varepsilon_2(t),\end{aligned} \tag{17}$$

We first consider a bivariate autoregressive process of $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$, as described in Eq. (17). The time series $\mathbf{x}_1$ and $\mathbf{x}_2$ have the same length of 1000 time points with added white noise of zero means and unit variance $\varepsilon_1(t)$ and $\varepsilon_2(t)$, respectively. $c_{2\to1}$ is a strength factor that indicates the influences of $\mathbf{x}_2$ on $\mathbf{x}_1$; and $c_{1\to2}$ is a strength factor that indicates the influences of $\mathbf{x}_1$ on $\mathbf{x}_2$. Both $c_{1\to2}$ and $c_{2\to1}$ are nonstationary along time. For $0 < t < 500$, $c_{2\to1}$, is set to be 0 and $c_{1\to2}$ is set to be 2; for $500 < t < 1000$, $c_{2\to1}$, is set to 2 and $c_{1\to2}$ is set to 0. By numerical control, the time series $\mathbf{x}_1$ and $\mathbf{x}_2$ are both generated at the frequency of about 0.2 Hz. This is intended to verify if the TFvLK formulism can correctly determine the directed influences at the predefined frequencies. The conceptual mode of this synthetic model is shown in Fig. 1a.

In Figs. 1b and 2b, the blue ($T_{2\to1}$) and the green ($T_{1\to2}$) lines represent the time-varying information flow signals of the two-way interaction generated, while the red ($\text{sig}_{2\to1}$) and pink ($\text{sig}_{1\to2}$) lines represent the significant test values to determine the standard error for each time step.[1] Causality is inferred here if the absolute value of the green line exceeds that of the pink, and or the absolute value of the blue line exceeds that of the red line. In the first half of the entire time period in Fig. 1b, the derived $T_{1\to2}$ is seen to exceed $\text{sig}_{1\to2}$, while the absolute value of $T_{2\to1}$ remains significantly less than $\text{sig}_{2\to1}$. Consequently, a one-way causality, indicating that $\mathbf{x}_1$ is causal to $\mathbf{x}_2$ but $\mathbf{x}_2$ is noncausal to $\mathbf{x}_1$, can be inferred as seen in the first half of the entire time period in Fig. 1a. In the second half, the reverse is observed. While $T_{2\to1}$ is consistently larger than $\text{sig}_{2\to1}$, $T_{1\to2}$ remains significantly less than $\text{sig}_{1\to2}$. Here, the obtained results imply that $\mathbf{x}_2$ is causal to $\mathbf{x}_1$ but $\mathbf{x}_1$ is noncausal to $\mathbf{x}_2$. These results are consistent with the predefined causal structure of the conceptual model described by Eq. (17) and shown in Fig. 1a, and suggest that the TvLK formalism has the capability to determine the time-varying causality within the given time series in this simple time-varying model. The missing data at the beginning of $T_{1\to2}$ and $T_{2\to1}$ are as a result of the offline estimation of the process and measurement noise covariance (details can be found in appendix B). Since the Kalman filter is in essence a learning algorithm, a sufficient amount of data is initially needed for robust estimations of these parameters. The $T_{1\to2}$ and $T_{2\to1}$ in the time domain are then subjected to the Morlet wavelet transform so that they can be decomposed into the frequency domains. The derived time-frequency representations of the directed influences between $\mathbf{x}_1$ and $\mathbf{x}_2$ are shown in Fig. 1c. Significance tests are conducted at 5% level using Monte Carlo simulations as suggested by Torrence and Compo (1998). There appears to be a significant $T_{1\to2}$ signal for the first half period, and significant $T_{2\to1}$ signal for the second half period. These signals are both obtained at the frequency band of 0.2 Hz, as predefined in the synthetic model. It is,

[1] Since the information flow signals do not represent relative values, both negative and positive information flow results can be obtained. As such, we multiply the significant test results (which are always positive) by −1 as well to determine whether or not the information flow computed exceeds the significant test results. This explains why the red and pink lines are symmetric about the x axis.

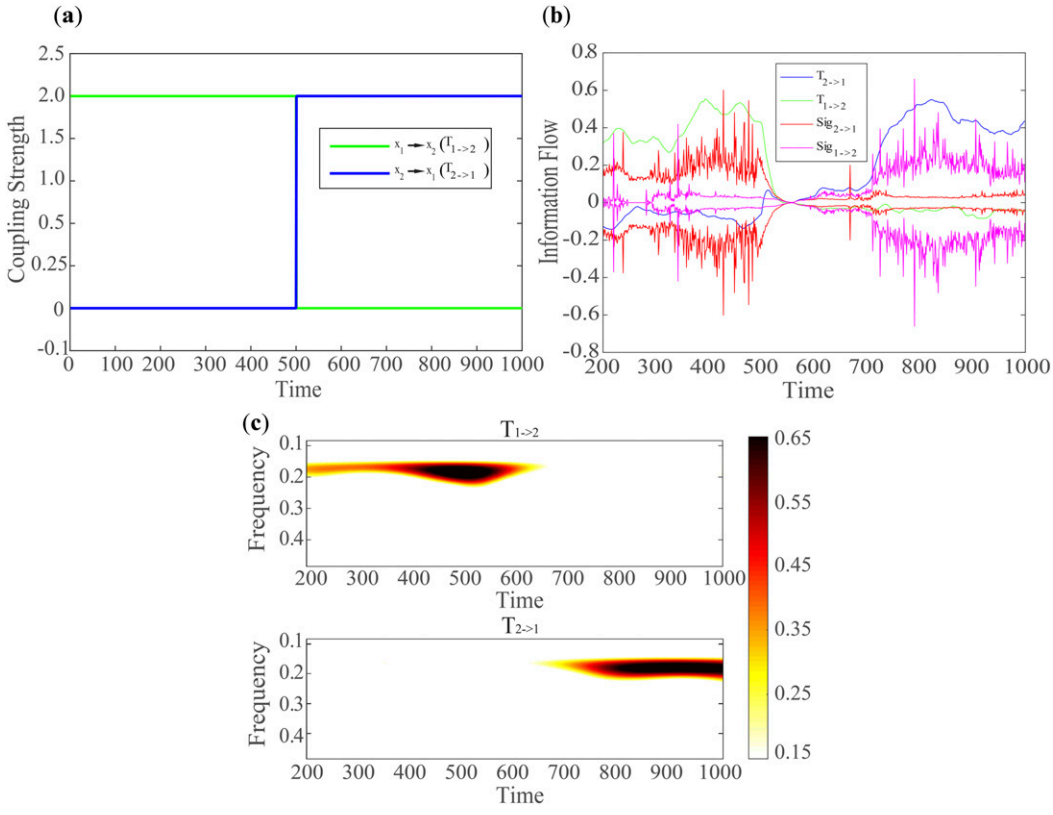

FIG. 1. (a) The conceptual synthetic model of the bivariate process as described in Eq. (17); (b) results of the time-varying coupling strengths (nats per unit time) ($T_{1\to2}$ and $T_{2\to1}$) at a 1% significance level ($\text{sig}_{1\to2}$, $\text{sig}_{2\to1}$), and (c) the time-frequency domain representations at a 5% significance level.

therefore, concluded that the TFvLK formulism is capable of correctly determining the timing and the frequency of the directed influences in the given simple conceptual model. Furthermore, significant signals in both Figs. 1b and 1c show quite similar magnitudes, implying the same strengths of influences as predefined. These results are encouraging and provide a sense of confidence for future applications.

b. *Synthetic model 2*

In the climate system, the interactions among different components may occur at different frequencies. It is thus of interest to determine if the proposed TFvLK formalism can disentangle the interactions at different frequencies. Following the first synthetic model, again we generate a bivariate autoregressive process of $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$ as described in Eq. (18). In this case, for $0 < t < 550$, both $c_{1\to2}$ and $c_{2\to1}$ are set to 0, so that neither $\mathbf{x}_1$ nor $\mathbf{x}_2$ is causal to the other. For $550 < t < 1000$, $c_{2\to1}$, is set to 0.95 and $c_{1\to2}$ is set to 0.45. By numerical control, the causality of $T_{1\to2}$ is set to occur at 0.1 Hz and that of $T_{2\to1}$ is set to occur at 0.2 Hz. Within the climate system, the chances of finding mutual causality at the same time are rather frequent. However, these may, in most cases, occur at different frequencies. Figure 2a shows a conceptual form of this synthetic model as should be obtained from the generated time series:

$$\begin{aligned}\mathbf{x}_1(t) &= 0.39\mathbf{x}_1(t-1) - 0.461\mathbf{x}_1(t-2) \\ &\quad + c_{2\to1}(t)\mathbf{x}_2(t-1) + \boldsymbol{\varepsilon}_1(t), \\ \mathbf{x}_2(t) &= 0.75\mathbf{x}_2(t-1) - 0.3017\mathbf{x}_2(t-2) \\ &\quad + c_{1\to2}(t)\mathbf{x}_1(t-1) + \boldsymbol{\varepsilon}_2(t), \end{aligned} \tag{18}$$

Applying the TvLK to the generated time series shows that it is able to yield the expected time-varying results of nil causalities in the first part of the time period since the information flow signals $T_{1\to2}$ and $T_{2\to1}$ are less than the significant values. On the other hand, mutual causalities are observed in the second part of the time period as both information flow signals are statistically significant. Additionally, the expected different coupling strengths (Fig. 2b), at 1% significance level in the second time period, are observed between the two coupling signals. While $T_{1\to2}$ shows a smaller magnitude of causality, $T_{2\to1}$ shows a much higher magnitude as

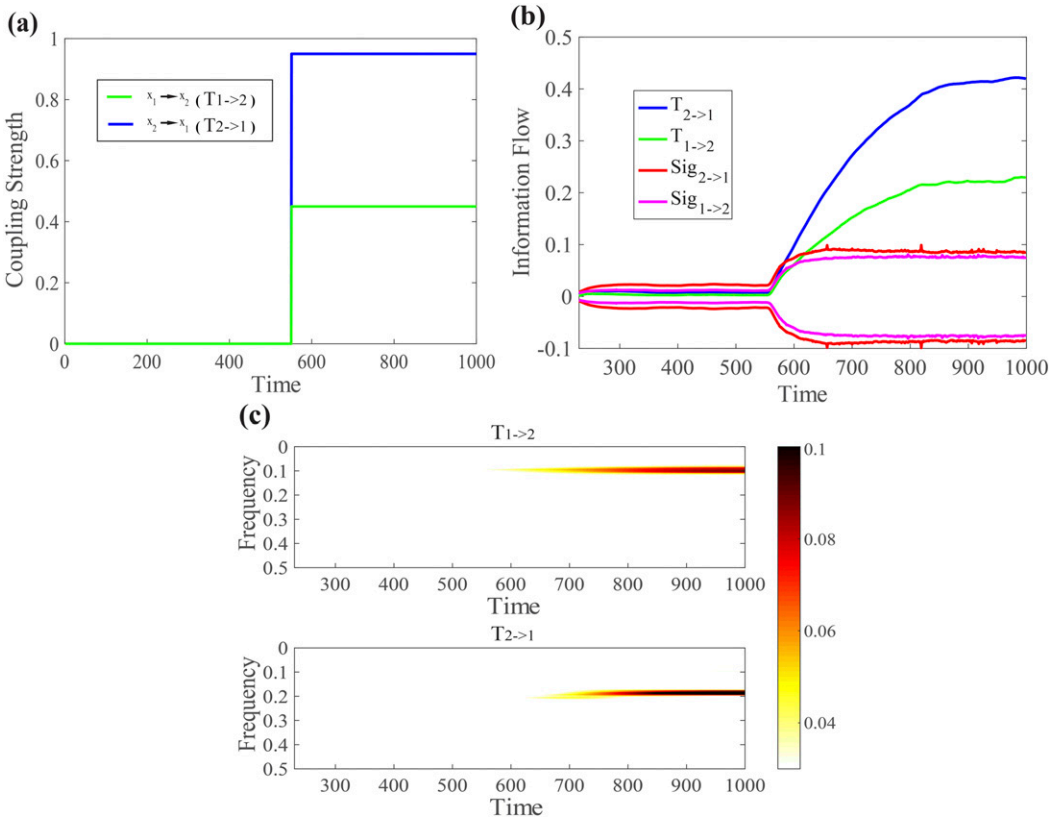

FIG. 2. (a) The conceptual synthetic model of the bivariate process as described in Eq. (18); (b) results of the time-varying coupling strengths (nats per unit time) ($T_{1\to2}$ and $T_{2\to1}$) at a 1% significance level ($\text{sig}_{1\to2}$, $\text{sig}_{2\to1}$), and (c) the time-frequency domain representations at a 5% significance level.

predefined in the bivariate model and depicted in Fig. 2b. This implies that the obtained TFvLK estimates have been able to satisfy the requirements of causal direction and magnitudes in this simulation. Decomposing $T_{1\to2}$ and $T_{2\to1}$ into the time-frequency distributions using the Bump wavelet transform shows that $T_{1\to2}$ occur around 0.1 Hz and $T_{2\to1}$ occur around a frequency of 0.2 Hz. It can also be observed from Fig. 2c that the different magnitudes of coupling strengths are also very well captured in the time-frequency domain.

4. Applications in land–atmosphere interactions

The previous section has shown the potential of the TvLK and TFvLK algorithms to capture the causality when the causal directions and magnitudes are known. The data in the real world may present scenarios that are rather more complex and chaotic and more difficult to analyze. We apply here the proposed algorithms to soil moisture and near surface air temperature coupling in China. We aim to understand their performances in disentangling the direct interactions between the top layer soil moisture and the overlaying air temperature. Soil moisture is essential in the land–atmosphere interactions. It partitions the available energy into latent heat for evaporation and sensible heat for air temperature increase (Seneviratne et al. 2010). While soil moisture may also affect the subsequent precipitation (Taylor et al. 2012; Tuttle and Salvucci 2015), it is also found to have considerable impacts on air temperature, especially on the short- to medium-term occurrence of heat waves (Miralles et al. 2012, 2014).

The territory of China features a complex topography and a diverse climate. In East China, the climate is dominated by the East Asian monsoon system, leading to cold and dry winters as well as warm and wet summers. Generally, the rainy season starts in May and lasts until October in Southeast China, and it starts in July and lasts until August in the northwest of China. Consequently, the soil moisture patterns exhibits large spatial and temporal variations (Parinussa et al. 2017). Studies have indicated the theoretical relations between soil moisture and air temperature; with limited soil moisture, evaporative cooling decreases, which leads to a rise in air temperature, and rising air temperature results in increased atmospheric demand of water vapor

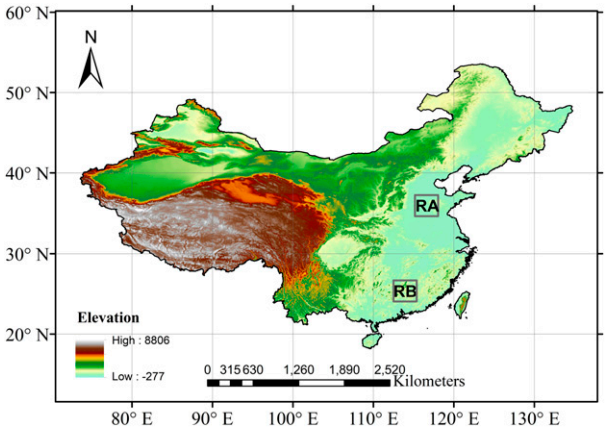

FIG. 3. An overview of the study area, mainland China. RA represents a partly wet–partly dry climate region while RB represents a wet climate region.

and thereby further reduced soil moisture (e.g., Seneviratne et al. 2010). However, such relations may vary through space and time.

We selected two regions in China for further study that are denoted as RA and RB in Fig. 3. RA is located in East China, which is a transitional climate zone between dry and wet climates; however, RB is located in South China, where it features a rather wet climate with significant precipitation. In recent years, there is more and more evidence from numerical simulations (Fischer et al. 2007) and diagnosis analysis (Casagrande et al. 2015; Miralles et al. 2012) that a significant impact of soil moisture on air temperature is most likely to occur in the transitional climate regimes. Therefore, we expect to find more significant signals of soil moisture impact in RA rather than RB. Daily data of volumetric soil moisture and 2-m air temperature from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim are used. The data are gridded at a horizontal resolution of 0.25° and span from January 2012 to December 2015. All interpretations are limited to the two parameters of the coupling.

Figures 4 and 5 show the soil moisture–temperature information flow in the time-varying and time-frequency domain over RA and RB. The significance tests of each pixel in the selected regions are computed at a 5% significance level. Here, the black solid lines present the mean of the information flow of all the individual pixels

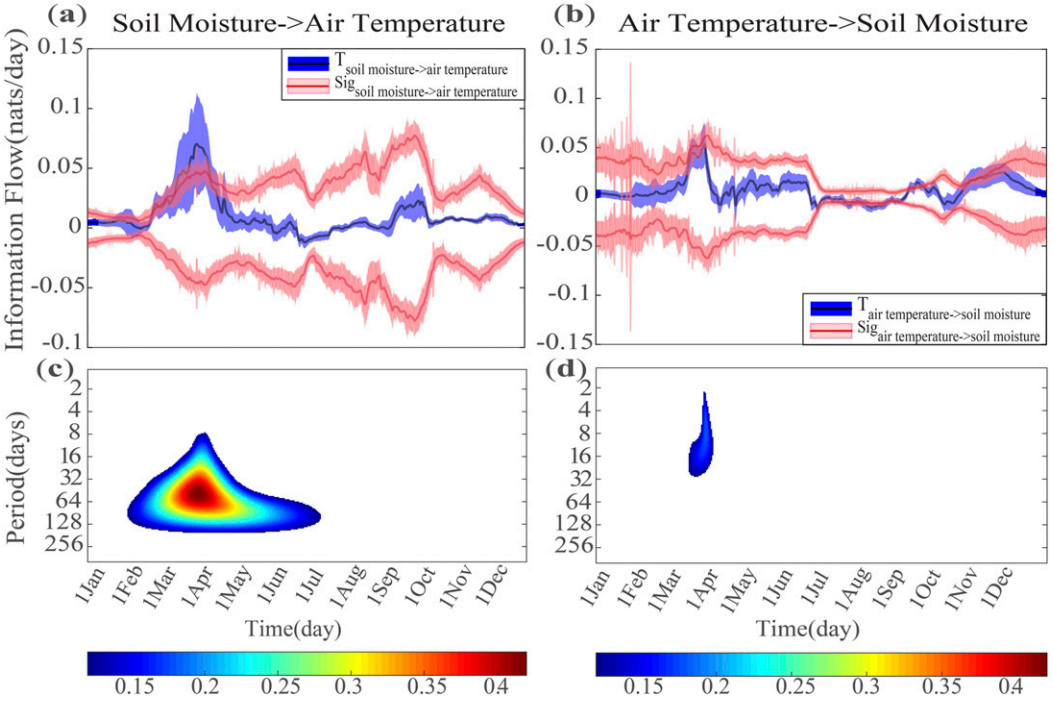

FIG. 4. Information flow between soil moisture and air temperature for RA. (a) The time-varying information flow from soil moisture to air temperature and (b) the time-varying information flow from air temperature to soil moisture. Causality is implied where the absolute value of the information flow signals (shown in blue shades and the black solid line) exceed the absolute values of the significance results computed based on Eq. (C1) (shown in red shades and red solid lines). The frequency distribution of (c) causality from soil moisture to air temperature and (d) the causality from air temperature to soil moisture. All the significance tests were done at a 5% significance level.

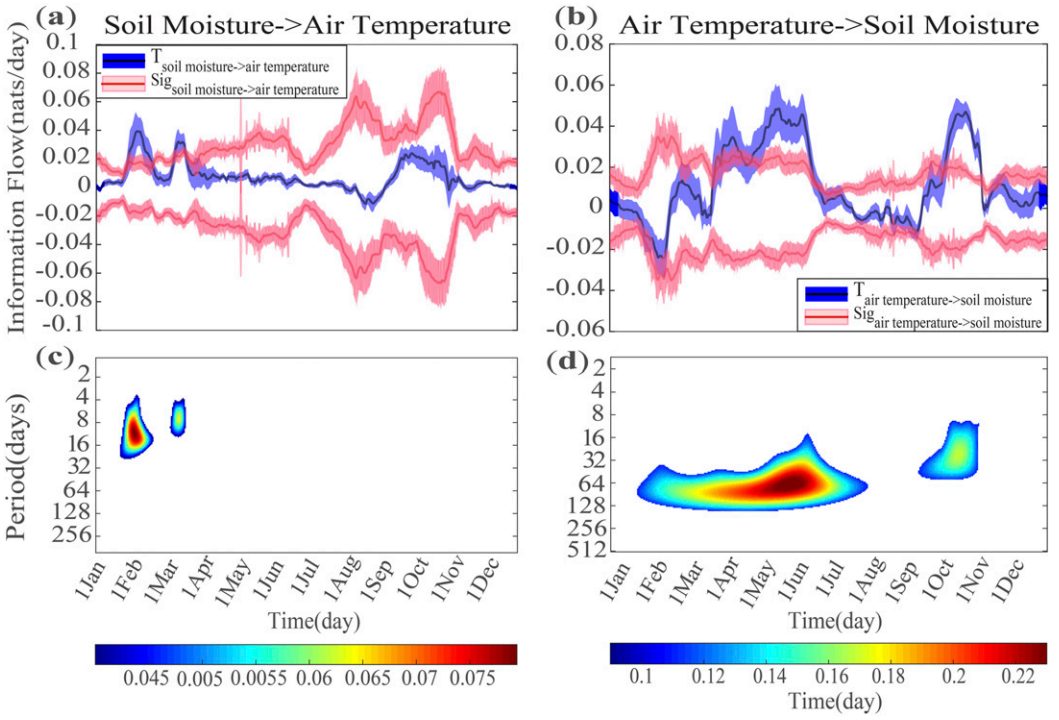

FIG. 5. Information flow between soil moisture and air temperature for RB. (a) Time-varying information flow from soil moisture to air temperature and (b) the time-varying information flow from air temperature to soil moisture. Causality is implied where the absolute value of the information flow signals (shown in blue shades and the black solid line) exceed the absolute values of the significance results computed based on Eq. (C1) (shown in red shades and red solid lines). The frequency distribution of (c) causality from soil moisture to air temperature and (d) the causality from air temperature to soil moisture. All the significance tests were done at a 5% significance level.

in the selected regions (RA in Fig. 4 and RB in Fig. 5), and the blue shades represent the standard deviation of all the computed information flow of the individual pixels in the selected region from the mean (black solid line). Similarly, the solid red line is the mean significance test values for all the pixels, and the shaded parts are the standard deviation from this mean. Causality is implied when the absolute values of the blue shades and black solid lines exceed the red shades and the red solid lines. The multiyear mean of the 4 years, 2012–15, is here used to derive the seasonal variability. No preprocessing of the datasets was done before computing the information flow results. Figure 4a shows the time-varying information flow results from soil moisture to air temperature (hereafter referred to as $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$). It can be seen that while $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ is not significant for most parts of the year in this region, a statistically significant signal occurs in spring. This implies that soil moisture has a significant feedback on air temperature in spring. Hence, soil moisture may provide some information to predict or forecast the variability of air temperature over the region. During spring, temperatures are rising relatively fast and frozen soils are generally thawed out, especially within the top soil layer. Sensible heat, therefore, becomes a very important factor within the interaction of the land and atmosphere, so that evaporation over time gradually increases. Soil moisture, thus, regulates the near surface temperature, keeping it generally cool. With a high demand of soil moisture for evapotranspiration its availability makes it a significant controlling factor. This makes $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ more significant than the information flow from air temperature to soil moisture (here after referred to as $T_{\text{airtemperature}\rightarrow\text{soilmoisture}}$) (Fig. 4b).

The $T_{\text{airtemperature}\rightarrow\text{soilmoisture}}$ (Fig. 4b) of this interaction over the region is observed to be relatively insignificant throughout the year. These separated signals of the interaction are in agreement with the current understanding of land surface atmosphere interactions. To verify that the causalities of the separated signals occurring in the same period of time are not the result of statistical pitfalls, their frequency distributions are also plotted in Figs. 4c and 4d. Applying the wavelet transform to obtain the time-frequency distributions of both signals clearly reveals that they occur at different time scales, with the latter occurring on a shorter periodicity than the former. The

significant parts of $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ are observed to occur at larger time scales in Fig. 4c, beginning from the start of spring and persisting into the early periods of summer. It is important to keep in mind that the coarse spatial and temporal resolution makes it more difficult to obtain finer signals of significance. According to Fig. 4d, if any level of $T_{\text{airtemperature}\rightarrow\text{soilmoisture}}$ is to be significant, it would be found to occur at shorter time scales between 3 and 30 days' periodicity in midspring. These results are, however, in very good agreement with previous studies (e.g., Seneviratne et al. 2010). The soil moisture to air temperature causality is here found to be stronger than the air temperature to soil moisture causality signal. An advantage of the TvLK formalism is that it is able to capture abrupt changes in the interactions over time since it does not smooth the time series estimation. Of course, since the Kalman filter is a learning algorithm, the results improve along time, as also seen in the synthetic model figures in the previous section.

Figure 5 presents the information flow results over the southeastern part of China (RB). This region is considered to be mostly wet; however, due to its position, it also receives sufficient enough solar radiation, thus also has a significant air temperature variability. A critical overview of both Figs. 5a and 5b also suggests that there is a distinct time lag between the two signals. Figure 5a demonstrates that $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ leads the coupling from late winter to early spring as expected. This period finds the region drier, but with significantly high air and soil temperatures, yet much less rainfall. As such, the beginning of the year presents the region as a soil-moisture-limited regime, causing soil moisture to constrain evaporation. The wavelet decomposition in Fig. 5c also sheds more light onto the interaction and shows that the signal significantly occurs between 4- and 30-day cycles of late winter and early spring.

Figure 5b shows the other side of the coupling in this region. From the wavelet analysis in Fig. 5d, we may say that the direct causal structures within the coupling are well separated. The periods of late spring and early to midsummer indicate strong signals in both the time-varying and time-frequency distributions. While the $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ is strongest over shorter time scales, $T_{\text{airtemperature}\rightarrow\text{soilmoisture}}$ is strongest over longer time scales. Of course, these are findings as interpreted from the ERA datasets, and as such the uncertainties within the assimilation data will surely impact the interactions. However, the results are in very good agreement with current understanding of soil moisture climate interactions (Seneviratne et al. 2010). As it is considered to be an energy-limited region, air temperature is seen to be the controlling factor in most parts of the year. With reduced temperatures in autumn, yet sufficient moisture in the soil, air temperature becomes the main controlling factor in the region and constrains the magnitude and rate of evapotranspiration. It is worth emphasizing that the results of the time-varying information flow used here are not qualitatively based, but importantly, quantitatively, in nats day^{-1}. As such, the magnitude of the signals can provide very good insights for future predictive capability.

5. Discussion

Information flow has long been recognized to be logically associated with causality and is hence ideal for understanding the cause–effect relationship within interactions (Liang 2008, 2016); it quantifies the exchange of information between two systems or events, showing both the direction and magnitude within a cause–effect relation (Hlaváčková-Schindler et al. 2007). Recently it has been reported that it actually can be rigorously derived from first principles within the context of stochastic dynamical systems (Liang 2008, 2016). Our ability to assess these qualities (magnitude, direction, and time variation) of causality in the interactions between events gives us a better insight into the behavior of the system and hence its predictive capability. Furthermore, our ability to rank interactions based on magnitudes is crucial in extreme event research as well as for a better understanding of a natural systems. This is exceedingly useful in different fields like meteorology, climatology, finance, and game theory.

In Eq. (15), the square root Kalman filter is used to estimate the information flow rate in a coupling to determine their direct causal structures. The Kalman filter as a learning algorithm is able to identify immediate changes when tracking the time-varying causality within a system. Previous studies have successfully integrated the Kalman filter into time-invariant formalisms to obtain time-varying structures (Omidvarnia et al. 2011; Punales 2011). Havlicek et al. (2010) also integrated the Kalman filter into the generalized partial directed coherence (GPDC) to obtain a dynamic Granger causality formalism to analyze causality both along frequency and time. Both the theoretical tests and application of their Kalman filter-based formalism showed significant improvements of the time-invariant GPDC. Fundamentally, interactions between any two events may vary with time; and as such, a time-invariant formalism does not offer a faithful representation of the evolution of the interactions. Here, we use the Kalman filter to estimate the covariances at each time step in Eq. (8) to obtain a time-varying structure as seen in Eq. (15) (TvLK).

We have established that the TvLK [Eq. (15)] is able to obtain a rigorous quantitative measure of their causal structure just as its time-invariant formalism [Eq. (8)]. The formalism that results in Eq. (8) has already been tested in both deterministic and stochastic systems (Liang 2016) and has been successfully applied to a variety of applications for interaction studies. These rigorous properties are clear in the TvLK, and thus offer much potential. Other methodologies like the Granger causality, directed coherence, and the directed transfer functions, which use vector autoregressive models for their parameter estimations require stationary residuals that could lead to different kinds of erroneous outcomes. However, the TvLK is able to handle both stationary and nonstationary conditions, and thus, offers advantages for wider application. Dhamala et al. (2008) identified that the parametric Granger causality techniques were better suited for low-order autoregression (AR) processes than for data that required high-order AR models. The wavelet-based Granger causality (Dhamala et al. 2008) that was proposed in the same study, though nonparametric, is computationally expensive because it has to be applied to multiple realizations. While the accuracy of this formalism relies heavily on a number of analyzed realizations, the TvLK has been shown in the above synthetic models that one realization is able to offer the needed information for a proper causality analysis. This is of great importance because real-world observations are usually taken once from a single source, and thus, may offer only one time series to capture the evolution of an event. Another important advantage of the original causality formalism in Liang (2014, 2016), and thus, the TvLK is that they are very easy to compute since they do not require AR modeling, but are composed of covariances for computation of causality.

Section 3 has shown that the principle of nil causality (Liang 2014, 2016) is detectable by the TvLK, which is not easily identified in some commonly used classical causal inference approaches. In their work, Orlowsky and Seneviratne (2010) showed with a statistical feedback parameter, the empirical feedback analysis, that covariability, although necessary, may not be sufficient to imply a causal relation between two variables. This, of course, holds for other different measures such as lagged or partial correlation, Granger causality and regression analysis. They showed that a third variable might be the reason for identifying covariability between two variables. This was illustrated in the same study by demonstrating how the influence of sea surface temperatures (SST) on precipitation persistence is easily mistaken for the impact of soil moisture on precipitation. This statistical pitfall is here tackled in two ways. First, we established through the numerical experiments that in a coupled system, the TvLK is able to detect an unambiguous one-way feedback relationship, which is essential for establishing a causal relation. In addition, where a two-way causality does not occur within the same period of time, we assume the absence of a statistical pitfall. Second, since a knowledge of the temporal scale (frequency) of causality between any two variables is essential to verifying the presence of the influence of a third variable, we extend the work flow of the TvLK to include a frequency decomposition. Using the continuous wavelet, we showed in the same synthetic model experiments that the significant causalities really do occur at their preset frequencies. This becomes even more important when the separated feedback signals tend to have a similar time-varying structure, so that we can verify the absence of a statistical pitfall by identifying if the causal relations occur at different temporal scales. This is difficult to identify just by observing the information flow time series. However, a time-frequency distribution is able to shed more light on the different temporal scales of causality. This is clearly demonstrated in the application section of this study. The results of the synthetic model experiments thus show that the TvLK does meet all the requirements expected of it for a reliable application. It is important to note that the TvLK formalism is developed to investigate the direct relationship within an n-dimensional system, and as such may not be readily applicable to investigate partial (or conditional) causal relationships. This is, however, set for future studies in line with this current one.

To understand how all these properties of the TvLK would perform in a real-world application, section 4 has demonstrated an application in the soil moisture–air temperature interaction over selected regions in China based on the first volumetric soil water layer (representing soil moisture) and the 2-m air temperature. ERA-Interim dataset is based on an assimilation scheme that considers the relationship between soil moisture and temperature based on the classical hydrological framework for defining evapotranspiration regimes as a function of soil moisture. This means we can expect the embedded causalities within the time series. $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ was found to be the more significant one over the partly wet–partly dry Huabei region. Conversely, over the Huanan region in Southeast China, $T_{\text{airtemperature}\rightarrow\text{soilmoisture}}$ was found to be more significant. These results are in agreement with the current understanding of soil moisture temperature feedbacks. Applying a spectral analysis to these results showed that the different signals occurred at different temporal scales. Over the Huabei region, $T_{\text{soilmoisture}\rightarrow\text{airtemperature}}$ was stronger and

occurred at longer time scales. The converse was found over the Huanan region.

A very significant point worth mentioning here is that the very rigorous and nice properties found in the Liang formalism [Eq. (8)] [as stated in Liang (2014, 2016)] are found to hold in the TvLK as well. This certainly would make this a preferable choice to transfer entropy and the Granger tests in many instances, as proven (Liang 2016). Another important property, which is noteworthy of mention, is that Eq. (8) was purposely derived for physical systems from nature, which is clearly demonstrated in section 4. Although the only application presented here is in land–atmosphere interactions, the TvLK, like the LK, can be applied in other problems in the natural sciences. It is expected to be especially useful in weather and intraseasonal analyses where causal relations have been suggested to change more quickly and are difficult to observe more clearly. Beyond the natural sciences, it is expected to be equally useful in the areas of finance and the social sciences.

Acknowledgments. This research was funded by National Key Research and Development Program of China (2017YFA0603701), the National Natural Science Foundation of China (41875094), the Sino-German Cooperation Group Project (GZ1447), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX17_0889).

APPENDIX A

The Bierman–Thornton Formalism

During the development of the Kalman filter in the last few decades, it was found that the standard Kalman filter has the propensity to accumulate numerical errors due to round offs during computations. As a result, the filter becomes unstable and accuracy of its estimation would thus, be compromised (Roncero 2014). In this regard, the square root family of the Kalman filter was developed to handle such instabilities, to make the filter more robust; this class of the filter has been applied in many studies and have shown significant advantages (Zhou et al. 2015). In this study, we make use of the Bierman–Thornton algorithm with modified Cholesky decomposition in the measurement and time updates of the filter (Fig. A1). This is expected to make the TvLK computation more stable and robust. The comparison of the standard Kalman filter and the square root forms are beyond the scope of this study, as such, no such experiments were included. Nonetheless, the algorithm has been developed to allow the user to choose to use either of the forms of the filter.

The Bierman–Thornton (UD filtering) methodology includes two main parts: the Bierman algorithm to update the observations (or measurements) of the $\mathbf{U}$ and $\mathbf{D}$ modified Cholesky factors of the covariance matrix $\mathbf{P}$, where $\mathbf{U}$ is a unit upper triangular matrix and the $\mathbf{D}$ is a diagonal matrix. This gives $\mathbf{P}$ in the form $\mathbf{P} = \mathbf{UDU}^{\mathrm{T}}$. The second part is the Thornton algorithm, which is mainly used for the temporal update of $\mathbf{U}$ and $\mathbf{D}$. Following Bierman and Thornton (1976) and its references thereof and as stated above, the factored for of $\mathbf{P}$ is given as

$$\mathbf{P} = \mathbf{UDU}^{\mathrm{T}}. \tag{A1}$$

Factoring Eq. (13) using Eq. (A1) yields

$$\begin{aligned} \mathbf{UDU}^{\mathrm{T}} &= \mathbf{U}^{-}[\mathbf{D}^{-} - 1/\alpha(\mathbf{D}^{-}\mathbf{U}^{-\mathrm{T}}\mathbf{a})(\mathbf{D}^{-}\mathbf{U}^{-T}\mathbf{a})^{\mathrm{T}}]\mathbf{U}^{-\mathrm{T}}, \\ 1/\mathbf{a} &= \mathbf{H}_k\mathbf{P}_k^{-}\mathbf{H}_k^{\mathrm{T}}, \end{aligned} \tag{A2}$$

where α here is defined as the innovations variance of the measurement vector and $\mathbf{a}$ is defined as an observation coefficient. Let the n vectors of $\mathbf{f}$ and $\mathbf{v}$ be defined by

$$\begin{aligned} \mathbf{f} &= \mathbf{U}^{-\mathrm{T}}\mathbf{a}, \\ \mathbf{v} &= \mathbf{D}^{-}\mathbf{f}\mathbf{a}, \end{aligned} \tag{A3}$$

and let $\overline{\mathbf{U}}$ and $\overline{\mathbf{D}}$ be the $\mathbf{UD}$ factors of $\mathbf{D}^{-} - 1/\alpha\mathbf{v}\mathbf{v}^{\mathrm{T}}$, thus

$$\mathbf{UDU}^{\mathrm{T}} = \mathbf{D}^{-} - 1/\alpha\mathbf{v}\mathbf{v}^{\mathrm{T}}. \tag{A4}$$

Substituting Eq. (A3) into Eq. (A4) gives

$$\mathbf{UDU}^{\mathrm{T}} = (\mathbf{U}^{-}\overline{\mathbf{U}})\overline{\mathbf{D}}(\mathbf{U}^{-}\overline{\mathbf{U}})^{\mathrm{T}}, \tag{A5}$$

and since $\mathbf{U}$ and $\overline{\mathbf{U}}$ are unit upper triangular, we obtain the solutions as

$$\begin{aligned} \mathbf{U} &= \mathbf{U}^{-}\overline{\mathbf{U}}, \\ \mathbf{D} &= \overline{\mathbf{D}}, \quad \text{and} \end{aligned} \tag{A6}$$

$$\mathbf{P}_{k+1}^{-} = \mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^{\mathrm{T}} + \mathbf{Q}_k. \tag{A7}$$

Substituting the Bierman observation update into Eq. (A7) gives

$$\mathbf{P}^{-} = \mathbf{A}_k(\mathbf{U}^{+}\mathbf{D}^{+}\mathbf{U}^{+})\mathbf{A}_k^{\mathrm{T}} + \mathbf{Q}_k. \tag{A8}$$

A more detailed and complete description of the Bierman–Thornton algorithm is beyond this paper. Readers are referred to Grewal and Andrews (2001) and its references therein. Table A1 below summarizes the

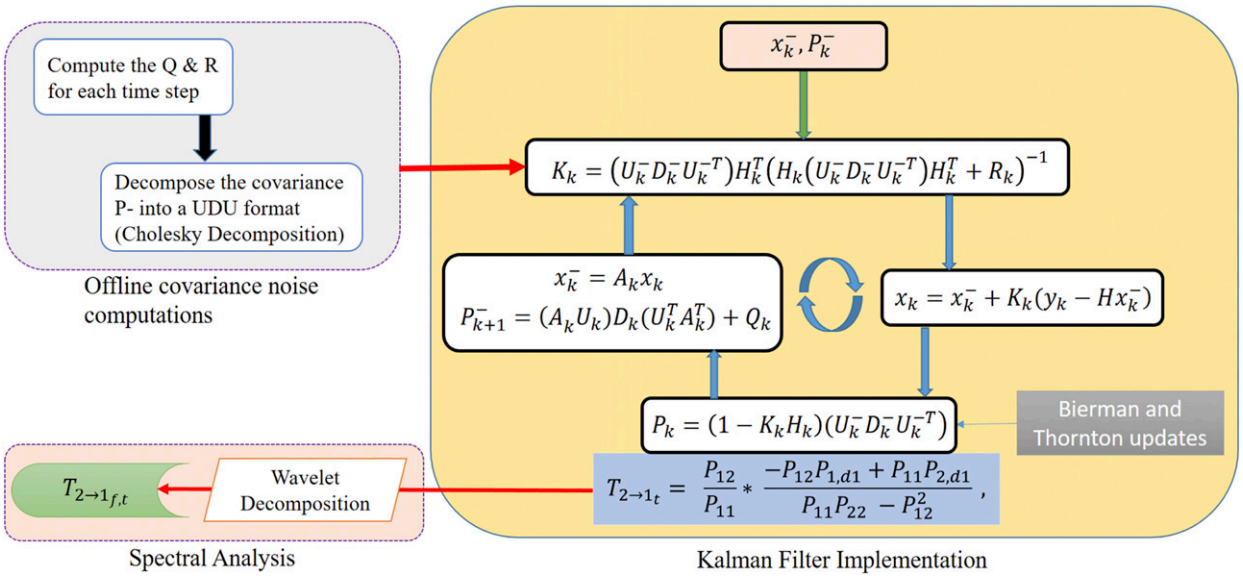

FIG. A1. A framework of the time-frequency varying causality formalism.

difference between the standard Kalman filter and its square root framework based on the Bierman–Thornton algorithm.

APPENDIX B

The Framework of the TvLK

Many studies have focused on producing accurate results for the noise covariances, and have been able to show that since we are unable to observe the process under estimation, **Q**, the process noise covariance becomes even more difficult to determine (Åkesson et al. 2008; Bavdekar et al. 2011; Odelson et al. 2006; Saha et al. 2011). In this study, we employ the use of the exponentially weighted moving average (EWMA) and unweighted moving average (UWMA) to estimate **Q** and **R** offline.

Although the Kalman filter has received a wide recognition and has been shown to provide optimal solutions where needed, its adaptive parameterization holds the greatest challenges. Previous work has identified that even though the filter is able to produce very reliable results, the estimation of the process and measurement noise covariances, **Q** and **R**, is rather difficult (Bavdekar et al. 2011; Saha et al. 2011; Mohan et al. 2015). It is therefore important to carefully parameterize these variables to obtain a more robust output. A common practice is to set **Q** and **R** to be constants, precisely, to be unit matrices. However, in reality, **Q** and **R** are expected to change at each time step. The formalism in this study uses a smoothed version of the time series as a surrogate for the true process state, which is done with EWMA or UWMA to derive **Q** and **R** offline at each time step (shown in Fig. A1). Several experiments conducted during this study have also confirmed that the Kalman gain is more accurately computed when **Q** and **R** are properly derived. Given this associative property, the formalism presented in this study is somewhat more difficult to derive because of the Kalman filter. Thus, several tests were required to obtain an optimal solution. Because of the square root form of the Kalman filter used here, computation is rather expensive. It takes about 3 times the time length to compute, as against the traditional causality in Liang (2014). The

TABLE A1. A comparison of the standard and square root Kalman filter formalisms.

Process	The standard Kalman filter	The square root Kalman filter
Compute Kalman gain	$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} + \mathbf{R})^{-1}$	$\mathbf{K}_k = (\mathbf{U}_k^- \mathbf{D}_k^- \mathbf{U}_k^{-\mathrm{T}}) \mathbf{H}_k^{\mathrm{T}} [\mathbf{H}_k (\mathbf{U}_k^- \mathbf{D}_k^- \mathbf{U}_k^{-\mathrm{T}}) \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}]^{-1}$
Updating	$\mathbf{x}_k = \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^-)$	$\mathbf{x}_k = \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^-)$
Compute updated covariance matrix	$\mathbf{P}_k = (1 - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-$	$(\mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^{\mathrm{T}}) = (1 - \mathbf{K}_k \mathbf{H}_k)(\mathbf{U}_k^- \mathbf{D}_k^- \mathbf{U}_k^{-\mathrm{T}})$
Covariance prediction	$\mathbf{P}_{k+1}^- = \mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^{\mathrm{T}} + \mathbf{Q}_k$	$(\mathbf{U}_{k+1}^- \mathbf{D}_{k+1}^- \mathbf{U}_{k+1}^{-\mathrm{T}}) = (\mathbf{A}_k \mathbf{U}_k) \mathbf{D}_k (\mathbf{U}_k^{\mathrm{T}} \mathbf{A}_k^{\mathrm{T}}) + \mathbf{Q}_k$

EWMA and UWMA chosen in this study were to facilitate easier computations. Of course, in the use of the EWMA and UWMA methods, there is the requirement of the choice of a lookback window length to compute the moving average. In an attempt to do this, an information length of the window length is lost to do the forward update of the filter. This is seen from the missing data in the time-varying results of the synthetic models in Figs. 1 and 2. An optimal linear smoother is suggested by Fraser and Potter (1969) by running the filter backward using the last estimate of the forward recursion of the filter to initialize it, and then combining both the forward and backward runs. Using this approach of forward and backward runs, which has already been tested in the course of this study, does help to curb the problem of the missing data of the initial time steps of the TvLK result, though it was eventually not adopted. Further studies will look into employing the optimal smoother in the Kalman filter run of this formalism. It can therefore be seen in the frequency decomposition of the synthetic model results, the missing data of the preceding parts of the information flow time series were not used in the wavelet decomposition.

APPENDIX C

Significance Testing

For a statistical tool such as in Eq. (8), it is rather necessary to do a significance test so as to better understand the robustness of the derived result. For this purpose, we use the Fisher information matrix equation [Eq. (C1)], following Liang (2014), since its inverse gives obtains a covariance matrix with a given significance level. Liang (2014) suggested that the normally used bootstrap method may not hold much simplicity to obtain a good significance level. With a large sample size N, the information flow approaches a normal distribution around its true value with a variance $(c_{12}/c_{11})^2\sigma^2_{a_{12}}$, which is derived from the maximum likelihood estimation. Thus, if we denote $\hat{\boldsymbol{\theta}} = (f_1, a_{11}, a_{12}, b_1)$, we obtain

$$NI_{ij} = -\sum_{n=1}^{N} \frac{\partial^2 \log\rho\,(X_{n+1}|X_n;\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}, \tag{C1}$$

where $\mathbf{I}$ in the matrix $N\mathbf{I}$ represents the Fisher information matrix. The inverse of $N\mathbf{I}$ is the covariance matrix of $\hat{\boldsymbol{\theta}}$, which has $\sigma^2_{a_{12}}$ within. Here, ρ is the marginal probability of $\{Xn\}$, a Markov process. Readers are referred to Liang (2014) for details of Eq. (C1).

REFERENCES

Åkesson, B. M., J. B. Jørgensen, N. K. Poulsen, and S. B. Jørgensen, 2008: A generalized autocovariance least-squares method for Kalman filter tuning. *J. Process Control*, **18**, 769–779, https://doi.org/10.1016/j.jprocont.2007.11.003.

Asl, H. G., and S. H. Pourtakdoust, 2007: UD covariance factorization for unscented Kalman filter using sequential measurements update. *Int. J. Aerosp. Mech. Eng.*, **1**, 629–637.

Bavdekar, V. A., A. P. Deshpande, and S. C. Patwardhan, 2011: Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *J. Process Control*, **21**, 585–601, https://doi.org/10.1016/j.jprocont.2011.01.001.

Bellantoni, J., and K. Dodge, 1967: A square root formulation of the Kalman-Schmidt filter. *AIAA J.*, **5**, 1309–1314, https://doi.org/10.2514/3.4189.

Berg, A., B. R. Lintner, K. L. Findell, S. Malyshev, P. C. Loikith, and P. Gentine, 2014: Impact of soil moisture–atmosphere interactions on surface temperature distribution. *J. Climate*, **27**, 7976–7993, https://doi.org/10.1175/JCLI-D-13-00591.1.

Bierman, G. J., and C. L. Thornton, 1976: Numerical comparison of discrete Kalman filter algorithms: Orbit determination case study. *Conf. on Decision and Control/15th Symp. on Adaptive Processes*, Clearwater, FL, IEEE, 859–872.

Casagrande, E., B. Mueller, D. G. Miralles, D. Entekhabi, and A. Molini, 2015: Wavelet correlations to reveal multiscale coupling in geophysical systems. *J. Geophys. Res.*, **120**, 7555–7572, http://doi.org/10.1002/2015JD023265.

Catalano, F., A. Alessandri, M. D. Felice, Z. Zhu, and R. B. Myneni, 2016: Observationally based analysis of land–atmosphere coupling. *Earth Syst. Dyn.*, **7**, 251–266, https://doi.org/10.5194/esd-7-251-2016.

Ciabatta, L., L. Brocca, C. Massari, T. Moramarco, S. Puca, A. Rinollo, S. Gabellani, and W. Wagner, 2015: Integration of satellite soil moisture and rainfall observations over the Italian territory. *J. Hydrometeor.*, **16**, 1341–1355, https://doi.org/10.1175/JHM-D-14-0108.1.

Dajčman, S., 2013: Interdependence between some major European stock markets—A wavelet lead/lag analysis. *Prague Econ. Pap.*, **22** (1), 28–49, https://doi.org/10.18267/j.pep.439.

Dhamala, M., G. Rangarajan, and M. Ding, 2008: Analyzing information flow in brain networks with nonparametric granger causality. *Neuroimage*, **41**, 354–362, https://doi.org/10.1016/j.neuroimage.2008.02.020.

Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, 2007: Soil moisture–atmosphere interactions during the 2003 European summer heat wave. *J. Climate*, **20**, 5081–5099, https://doi.org/10.1175/JCLI4288.1.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, and M. Rummukainen, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Ford, T. W., A. D. Rapp, S. M. Quiring, and J. Blake, 2015: Soil moisture–precipitation coupling: Observations from the Oklahoma Mesonet and underlying physical mechanisms. *Hydrol. Earth Syst. Sci.*, **19**, 3617–3631, https://doi.org/10.5194/hess-19-3617-2015.

Frankingnoul, C., and K. Hasselmann, 1977: Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, **29**, 289–305, http://doi.org/10.1111/j.2153-3490.1977.tb00740.x.

Fraser, D., and J. Potter, 1969: The optimum linear smoother as a combination of two optimum linear filters. *IEEE Trans. Autom. Control*, **14**, 387–390, https://doi.org/10.1109/TAC.1969.1099196.

Ghorbanidehno, H., A. Kokkinaki, J. Yue, E. Darve, and P. K. Kitanidis, 2015: Real-time data assimilation for large-scale systems: The spectral Kalman filter. *Advances in Water Resources*, Vol. 86, H.-J. Hendricks Franssen and I. Neuweiler, Eds., Elsevier, 260–272.

Gonzalez, R., G. Sutter, C. Sisterna, and H. Patino, 2014: FPGA-based floating-point UD filter coprocessor for integrated navigation systems. *2015 Sixth Argentine Conf. on Embedded Systems (CASE)*, Buenos Aires, Argentina, IEEE, https://doi.org/10.1109/SASE-CASE.2015.7295840.

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

——, 1980: Testing for causality: A personal viewpoint. *J. Econ. Dyn. Control*, **2**, 329–352, https://doi.org/10.1016/0165-1889(80)90069-X.

Grewal, M. S., and A. P. Andrews, 2001: *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley, 401 pp.

Guo, Z., and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview. *J. Hydrometeor.*, **7**, 611–625, https://doi.org/10.1175/JHM511.1.

Havlicek, M., J. Jan, M. Brazdil, and V. D. Calhoun, 2010: Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data. *Neuroimage*, **53**, 65–77, https://doi.org/10.1016/j.neuroimage.2010.05.063.

Hlaváčková-Schindler, K., M. Paluš, M. Vejmelka, and J. Bhattacharya, 2007: Causality detection based on information-theoretic approaches in time series analysis. *Phys. Rep.*, **441** (1), 1–46, https://doi.org/10.1016/j.physrep.2006.12.004.

Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1.

Kiani, M., and S. H. Pourtakdoust, 2014: Adaptive square-root cubature-quadrature Kalman particle filter for satellite attitude determination using vector observations. *Acta Astronaut.*, **105**, 109–116, https://doi.org/10.1016/j.actaastro.2014.08.021.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, and T. Yamada, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, https://doi.org/10.1126/science.1100217.

Liang, X. S., 2008: Information flow within stochastic dynamical systems. *Phys. Rev.*, **78E**, 031113, http://doi.org/10.1103/PhysRevE.78.031113.

——, 2013: The Liang–Kleeman information flow: Theory and applications. *Entropy*, **15**, 327–360, https://doi.org/10.3390/e15010327.

——, 2014: Unraveling the cause–effect relation between time series. *Phys. Rev.*, **90E**, 052150, https://doi.org/10.1103/PhysRevE.90.052150.

——, 2016: Information flow and causality as rigorous notions ab initio. *Phys. Rev.*, **94E**, 052201, http://doi.org/10.1103/PhysRevE.94.052201.

Maoudj, R., L. Fety, and C. Alexandre, 2013: Performance analysis of modified Gram-Schmidt Cholesky implementation. *Int. J. Comput. Digital Syst.*, **27** (1), 21–27, https://doi.org/10.12785/IJCDS/020103.

Massari, C., L. Brocca, T. Moramarco, Y. Tramblay, and J. F. D. Lescot, 2014: Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall. *Adv. Water Resour.*, **74**, 44–53, https://doi.org/10.1016/j.advwatres.2014.08.004

Miralles, D. G., M. J. V. D. Berg, A. J. Teuling, and R. A. M. D. Jeu, 2012: Soil moisture-temperature coupling: A multiscale observational analysis. *Geophys. Res. Lett.*, **39**, L21707, https://doi.org/10.1029/2012GL053703.

——, A. J. Teuling, C. C. Heerwaarden, and J. Arellano, 2014: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.*, **7**, 345–349, https://doi.org/10.1038/ngeo2141.

Mohan, M. S., N. Naik, R. M. O. Gemson, and M. R. Ananthasayanam, 2015: Introduction to the Kalman filter and tuning its statistics for near optimal estimates and Cramer Rao bound. Indian Institute of Technology Kanpur Department of Electrical Engineering Tech. Rep. TR/EE2015/401, 402 pp., https://arxiv.org/pdf/1503.04313.pdf.

Navarra, A., and J. Tribbia, 2005: The coupled manifold. *J. Atmos. Sci.*, **62**, 310–330, https://doi.org/10.1175/JAS-3345.1.

Notaro, M., 2008: Statistical identification of global hot spots in soil moisture feedbacks among IPCC AR4 models. *J. Geophys. Res.*, **113**, D09101, https://doi.org/10.1029/2007JD009199.

Odelson, B. J., M. R. Rajamani, and J. B. Rawlings, 2006: A new autocovariance least-squares method for estimating noise covariances. *Automatica*, **42**, 303–308, https://doi.org/10.1016/j.automatica.2005.09.006.

Omidvarnia, A. H., M. Mesbah, M. S. Khlif, J. M. O'Toole, P. B. Colditz, and B. Boashash, 2011: Kalman filter-based time-varying cortical connectivity analysis of newborn EEG. *2011 Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, Boston, MA, IEEE, 1423–1426, http://doi.org/10.1109/IEMBS.2011.6090335.

Orlowsky, B., and S. I. Seneviratne, 2010: Statistical analyses of land–atmosphere feedbacks and their possible pitfalls. *J. Climate*, **23**, 3918–3932, https://doi.org/10.1175/2010JCLI3366.1.

Paluš, M., 2014: Cross-scale interactions and information transfer. *Entropy*, **16**, 5263–5289, https://doi.org/10.3390/e16105263.

Papagiannopoulou, C., D. G. Miralles, S. Decubber, M. Demuzere, N. E. C. Verhoest, W. A. Dorigo, and W. Waegeman, 2017: A non-linear Granger-causality framework to investigate climate–vegetation dynamics. *Geosci. Model Dev.*, **10**, 1945–1960, https://doi.org/10.5194/gmd-10-1945-2017.

Papana, A., C. Kyrtsou, D. Kugiumtzis, and C. Diks, 2016: Detecting causality in non-stationary time series using partial symbolic transfer entropy: Evidence in financial data. *Comput. Econ.*, **47**, 341–365, https://doi.org/10.1007/s10614-015-9491-x.

Parinussa, R. M., G. Wang, Y. Y. Liu, D. F. T. Hagan, F. Lin, R. der Schalie, and R. A. M. Jeu, 2017: The evaluation of single-sensor surface soil moisture anomalies over the mainland of the People's Republic of China. *Remote Sens.*, **9**, 149, http://doi.org/10.3390/rs9020149.

Punales, A. G. S., 2011: Time-varying coefficient models and the Kalman filter. M.S. thesis, Applied Mathematics, Ryerson University, 65 pp., https://digital.library.ryerson.ca/islandora/object/RULA%3A2193/datastream/OBJ/download/Time-Varying_Coefficient_Models_And_The_Kalman_Filter___Applications_To_Hedge_Funds.pdf.

Roncero, D. F., 2014: A study of QR decomposition and Kalman filter implementations. M.S. thesis, Signal Processing, School of Electrical Engineering, Kungliga Tekniska Hgskolan, 57 pp., http://kth.diva-portal.org/smash/get/diva2:808731/FULLTEXT01.pdf.

Roundy, J. K., and E. F. Wood, 2015: The attribution of land–atmosphere interactions on the seasonal predictability of

drought. *J. Hydrometeor.*, **16**, 793–810, https://doi.org/10.1175/JHM-D-14-0121.1.

Saha, M., B. Goswami, and R. Ghosh, 2011: Two novel costs for determining the tuning parameters of the Kalman filter. Department of Instrumentation and Electronics Engineering, Jadavpur University, 8 pp., https://arxiv.org/pdf/1110.3895.pdf.

Saini, R., G. Wang, and J. S. Pal, 2016: Role of soil moisture feedback in the development of extreme summer drought and flood. *J. Hydrometeor.*, **17**, 2191–2207, https://doi.org/10.1175/JHM-D-15-0168.1.

Sastry, V., 1971: Decomposition of the extended Kalman filter. *IEEE Trans. Autom. Control*, **16**, 260–261, https://doi.org/10.1109/TAC.1971.1099709.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.*, **99**, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004.

Stéfanon, M., P. Drobinski, F. D'Andrea, C. Lebeaupin-Brossier, and S. Bastin, 2014: Soil moisture-temperature feedbacks at meso-scale during summer heat waves over Western Europe. *Climate Dyn.*, **42**, 1309–1324, https://doi.org/10.1007/s00382-013-1794-9.

Stips, A., D. Macias, C. Coughlan, E. Garcia-Gorriz, and X. S. Liang, 2016: On the causal structure between CO_2 and global temperature. *Sci. Rep.*, **6**, 21691, https://doi.org/10.1038/srep21691.

Taylor, C. M., R. A. M. D. Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo, 2012: Afternoon rain more likely over drier soils. *Nature*, **489**, 423–426, https://doi.org/10.1038/nature11377.

Thornton, C. L., 1976: Triangular covariance factorizations for Kalman filtering. NASA Tech. Memo. 33-798, 197 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19770005172.pdf.

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61–78, https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

Tuttle, S., and G. Salvucci, 2015: Empirical evidence of contrasting soil moisture–precipitation feedbacks across the United States. *Science*, **352**, 825–828, http://doi.org/10.1126/science.aaa7185.

Vialatte, F. B., J. Solé-Casals, J. Dauwels, M. Maurice, and A. Cichocki, 2009: Bump time-frequency toolbox: A toolbox for time-frequency oscillatory bursts extraction in electrophysiological signals. *BMC Neurosci.*, **10**, 46, https://doi.org/10.1186/1471-2202-10-46.

Wang, Y., T. Zhang, X. Chen, J. Li, and P. Feng, 2017: Spatial and temporal characteristics of droughts in Luanhe River basin. *Theor. Appl. Climatol.*, **131**, 1369–1385, http://doi.org/10.1007/s00704-017-2059-z.

Wu, L., and J. Zhang, 2015: The relationship between spring soil moisture and summer hot extremes over North China. *Adv. Atmos. Sci.*, **32**, 1660–1668, https://doi.org/10.1007/s00376-015-5003-0.

Zhang, J., W. C. Wang, and J. Wei, 2008: Assessing land–atmosphere coupling using soil moisture from the Global Land Data Assimilation System and observational precipitation. *J. Geophys. Res.*, **113**, D17119, https://doi.org/10.1029/2008JD009807.

Zhao, W., and A. Li, 2015: A review on land surface processes modelling over complex terrain. *Adv. Meteor.*, **2015**, 607181, http://doi.org/10.1155/2015/607181.

Zheng, Y., A. Kumar, and D. Niyogi, 2015: Impacts of land–atmosphere coupling on regional rainfall and convection. *Climate Dyn.*, **44**, 2383–2409, https://doi.org/10.1007/s00382-014-2442-8.

Zhou, Y., C. Zhang, Y. Zhang, and J. Zhang, 2015: A new adaptive square-root unscented Kalman filter for nonlinear systems with additive noise. *Int. J. Aerosp. Eng.*, **2015**, 381478, http://doi.org/10.1155/2015/381478.