Ecological Modelling 368 (2018) 136–146

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Causal networks reveal the dominance of bottom-up interactions in large, deep lakes

Victor Frossard[a,*], Frédéric Rimet[b], Marie-Elodie Perga[b,c]

[a] Université Savoie Mont Blanc, UMR42 CARRTEL, 73376 Le-Bourget-du-lac, France
[b] INRA, UMR42 CARRTEL, 75 Avenue de Corzent, 74203 Thonon les Bains Cedex, France
[c] Institute of Earth Surface Dynamics, University of Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 28 April 2017
Received in revised form
16 November 2017
Accepted 20 November 2017

Keywords:
Nonlinear dynamics
Convergent cross mapping
Causality
Ecological interaction
Network
Bottom-up control

ABSTRACT

Ecological dynamics often exhibit significant temporal variability and sudden shifts that characterize their non-equilibrium and nonlinear nature, challenging our ability to understand and predict their trajectories. Among a set of ecological time series originating from the long-term monitoring of three large and deep lakes, nonlinear forecasting methods (Simplex projection and S-map) indicated that most of the time series exhibited hallmarks of complex dynamics in the form of nonlinear behaviors. Convergent Cross Mapping (CCM) was used to estimate the causal relationships among these time series by considering different time lags. The significant causal relationships were then used to construct causal networks from which nodes were characterized using PageRank and CheiRank. For the three lakes, the dominance of bottom-up control was revealed and was mostly indirect (i.e., nutrient-forcing zooplankton). This result likely evidences the transitivity of the causal relationships obtained by CCM as well as the mixed phytoplankton diet of zooplankton species limiting the identification of causal relationships among these two ecological components. Complementarily, the consistence of causal relationships for the different time lags may highlight a temporal transitivity by which the instantaneous causal signal was transmitted over time. The dual representation of both PageRank and CheiRank provided a straightforward classification of each node and enabled their thorough implications in the information flow within the causal networks. The complementary use of CCM and network metrics constituted an efficient way to delineate ecological causation using a high-resolution time series, for which linear methods performed poorly, and provided insights into the dynamic hierarchy of the different ecological variables in aquatic ecosystems.

1. Introduction

The relative implication of bottom-up and top-down controls in driving ecological dynamics is of central interest in Ecology. The early investigations highlighting a strong influence of nutrients on algal dynamics (i.e., bottom-up control; Schindler, 1978) have been challenged by the evidenced role of trophic cascade on ecosystem processes (i.e., top-down control; Paine, 1980; Carpenter and Kitchell, 1985). Since then, a large number of studies have been conducted on lakes (Carpenter et al., 1995; Currie et al., 1999; Jeppesen et al., 1997), mesocosms (Sinistro, 2009) or using paleolimnological reconstructions (Perga et al., 2010) to decipher their relative implications on food-web dynamics.

Nonetheless, the quantification of the direction and strength of ecological links in food webs remains a spurious task because of their diversity (e.g., direct and indirect) and variation over time (Berlow et al., 2004; Wootton and Emmerson, 2005; Deyle et al., 2016; Lynam et al., 2017). Additionally, when extensive datasets are available, ecological dynamics (i.e., time series) can often seem erratic, possibly punctuated by sudden episodic bursts seemingly unrelated to any other putative causal variables. This particular feature lies in their nonlinearity, which is characterized by state-dependent behavior; the effect of one variable on another is dependent on the states of other variables in the system (Dixon et al., 1999; Anderson et al., 2008; Deyle et al., 2013; Glaser et al., 2013). In this context, ecological dynamics can appear correlated or not despite their causal association remaining constant, a phe-

* Corresponding author.
E-mail address: victor.frossard@univ-smb.fr (V. Frossard).

https://doi.org/10.1016/j.ecolmodel.2017.11.021
0304-3800/

nomenon called 'mirage correlation' (Sugihara et al., 2012). As a consequence, conventional correlation analyses can appear unsuitable for identifying the causal ecological links used to infer the relative role of bottom-up and top-down controls.

An alternative view to correlative approaches, one that allows the assumptions of equilibrium stability and stationarity to be relaxed, can be found in a flexible, nonparametric class of nonlinear forecasting models (Sugihara, 1994; Sugihara and May, 1990) that are based on state-space reconstructions. These models have shown impressive performance for forecasting complex ecological dynamics (Dixon et al., 1999; Glaser et al., 2013). Recently, an extension of these models, called 'Convergent Cross Mapping' (CCM), has been developed to identify causal relationships in weakly coupled nonlinear systems (Sugihara et al., 2012). CCM relies on Takens' Theorem (Takens, 1981), which states that a 'shadow-attractor' reconstruction from the lags of a single one-dimensional variable conserves the mathematical properties of the original attractor of the full dynamical system (Sugihara et al., 2012; Takens, 1981; Deyle and Sugihara, 2011). As a consequence, if two variables belong to the same dynamical system, meaning that they are causally coupled, their respective dynamics could be mutually predicted based on the local neighborhoods of their shadow attractors (Sugihara et al., 2012). Additionally, the accuracy of the forecasts (i.e., the cross mapping between the two shadow attractors) would increase along with the length of the time series (i.e., library length), thus exhibiting convergence as long as the shadow attractors became progressively denser and fulfill the historical dynamics. Due to this convergence property, CCM can further identify directional coupling: the higher a forcing dynamically constrains a subordinate variable, the more the amount of information about the forcing variable (i.e., causal imprint) will be encoded in the constrained variable (Sugihara et al., 2012). Complementarily, because causation is transitive, the causal imprint of a forcing can expand beyond its directly constrained variables to be encoded in other indirectly interacting variables (Sugihara et al., 2012), although the extent of the causal imprint decreases as the causal coupling becomes more indirect (Ye et al., 2015b). Recent valuable examples of the use of CCM can be found in Tsonis et al. (2015), which addressed the causal influence of cosmic rays on the earth's global temperature variability and in van Nes et al. (2015) which identified bi-directional causal relationships associated with lagged responses between the global earth's temperature and greenhouse gases.

In this study, we aim to couple CCM and network analyses to reveal causal links among the ecological components (i.e., nutrient/temperature, phytoplankton and zooplankton) of the three largest French lakes using time series obtained from long-term monitoring. This approach was designed to account for direct and indirect causal links; therefore, possibly evidencing the relative implication of bottom-up and top-down controls. Considering the pelagic zones of these three large lakes was expected to be especially suitable for delineating the relative controls driving the ecological dynamics because the implication of littoral and terrestrial environments have been shown to impact negligibly the pelagic zones of such large systems (Perga and Gerdeaux, 2004; Vadeboncoeur et al., 2008). We expected that because of the relatively low nutrient levels and the large size of the lakes, the ecological dynamics would be mostly constrained by bottom-up control, as suggested by Jeppesen et al. (1997), though never tested using the present methodology. Complementarily, the bottom-up control may be strong enough to be transmitted up to the highest food web level considered in this study (i.e., zooplankton), characterizing indirect bottom-up relationships. Different time lags in the causal interactions were accounted for to test the robustness of the causal relationships over time. We hypothesized a conservative structure of the causal network over time, signifying both a lag in ecological response to a driver as well as a transmission of the instantaneous causal relationship over time. Prior to performing CCM, the dynamic features (i.e., dimensionality and nonlinearity) of the time series were analyzed, providing insights about their complexity and nonlinear behavior, justifying the need for nonlinear methods to correctly model and predict their dynamics.

Table 1
Main characteristics of the three studied lakes. Numbers in brackets for Lake Bourget represent the changes in the watershed area and the flow tributary when the outlet of the lake to the Rhône river reverses in case of the river's flood, inducing temporary increase in both the watershed area and the tributary flow.

	Annecy	Bourget	Geneva
Area (km^2)	27	45	582
Watershed area (km^2)	273	560 (4600)	7395
Volume (km^3)	1.1	3.6	89
Maximal depth	65	147	309
Mean depth	41	81	152
Altitude (m a.s.l.)	446	231	372
Water turnover (year)	4	8	12
Phosphorus concentration (μg l^{-1})	8–10	8–12	8–12
Tributary flow	2.8	6.5 (365)	181

2. Methods

2.1. Study sites and ecological time series

Lakes Annecy, Geneva and Bourget are three large and deep lakes lying on the Western border of the French Alps. They lie in a similar climatic (i.e., mean annual temperature of ~11 °C) and geologic (i.e., carbonate bed rocks) context, and their main environmental features are provided in Table 1. They share a common eutrophication history, yet to different extents during the middle of the 20th century before remediation programs led to efficient re-oligotrophication by the end of the last century that has persisted to the present (Berthon et al., 2013). Despite the availability of data since the first half of the 20th century (Observatory on Alpine Lakes, www6.inra.fr/soere-ola), most environmental and ecological variables have not been continuously monitored prior to the implementation of rigorous monitoring for management at different time periods for the three lakes: in the 1970s for Lake Geneva (41 years), since 2003 for Lake Bourget (12 years) and since 2004 for Lake Annecy (11 years) (Data source © SOERE OLA-IS, INRA Thonon-les-Bains, CIPEL, SILA, CISALB, [date of download: 20/10/2015], developed by INRA Eco-informatics). For each lake, samples were collected at a single site located at the deepest part of the lake. Although biological and chemical variables can exhibit spatial variability, ongoing spatial analyses (no shown) suggest that the sampling sites were representative the actual status of the studied variables at a lake level. The parts of the time series for which nutrient, phytoplankton and zooplankton data were available but at sampling frequencies varying from weekly to monthly were considered. Among the numerous species of phytoplankton and zooplankton identified over the monitored period, those whose occurrence was <50% were excluded from the datasets. Time series were standardized and then interpolated using a monotonic Hermite spline so that they exhibited identical and regularly spaced numbers of observations (see Appendix 1 in Supplementary material for the number of initial and interpolated data) (Fig. 1). The possible effect of such an interpolation method on the nonlinear signal and the identification of causal relationships was assessed using a random and nonlinear time series by randomly deleting a variable number of time points (i.e., 1, 2, 5, 10, 15, and 20) as inputs to the spline (Appendix 2 in Supplementary material). These preliminary investigations suggested that the methods used were robust to the extent of interpolation using the monotonic Hermite spline considered in this study. The datasets for the three lakes were