Title: Decomposing predictability to identify dominant causal drivers in complex ecosystems

Author list:
Kenta Suzuki[1]*. Shin-ichiro S. Matsuzaki[2] and Hiroshi Masuya[1]
1. BioResource Research Center, RIKEN, Tsukuba, Ibaraki 305-0074, Japan.
2. Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan
*Corresponding author: kenta.suzuki.zk@riken.jp

Classification: Research reports

Keywords: Ecosystem, Ecosystem monitoring data, Causal network, Predictability, Echo state network, Granger causality

Abstract

Ecosystems are complex systems of various physical, biological, and chemical processes. Since ecosystem datasets are composed of a mixture of different levels of stochasticity and nonlinearity, handling these data is a challenge for existing methods of time-series based causal inferences. Here we show that, by harnessing contemporary machine learning approaches, the concept of Granger causality can be effectively extended to the analysis of complex ecosystem time series and bridge the gap between dynamical and statistical approaches. The central idea is to use an ensemble of fast and highly predictive artificial neural networks to select a minimal set of variables that maximizes the prediction of a given variable. It enables decomposition of the relationship among variables through quantifying the contribution of an individual variable to the overall predictive performance. We show how our approach, EcohNet, can improve interaction network inference for a mesocosm experiment and simulated ecosystems. The application of the method to a long-term lake monitoring dataset yielded new but interpretable results on the drivers causing cyanobacteria blooms, which is a serious threat to ecological integrity and ecosystem services. Since performance of EcohNet is enhanced by its predictive capabilities, it also provides an optimized forecasting of overall components in complex ecosystems. EcohNet could be used to analyze complex and hybrid multivariate time series in many scientific areas not limited to ecosystems.

Significance

Effective use of ecosystem monitoring data to resolve global environmental issues is a major challenge of ecology in the 21st century. A proposed solution to address this challenge is a time-series-based causal analysis, which can offer insight on the mechanical links between ecosystem components. In this work, a model-free framework named EcohNet is proposed that can robustly predict causal networks of ecosystem components. EcohNet utilizes ensemble predictions of echo state networks, which are known to be fast, accurate, and highly relevant for a variety of dynamical systems. It also can provide an optimized forecasting of overall components, and could be used to analyze complex and hybrid multivariate time series in many scientific areas, not only ecosystems.

Introduction

Various systems in the world, from cells, organisms, ecosystems, and our own societies, are complex and driven by many interacting components. An attempt to infer relationships among the components as a causal network is an important step in understanding the system as a whole. There has been a great body of research on inferring such networks from time series [Runge 2018, Runge et al. 2019a]. In particular, methods that can overcome the limitations of classical methods [Granger 1969, Akoglu 2018], such as CCM [Sugihara et al. 2012], which is suitable for systems under the influence of dynamic processes, and PCMCI [Runge et al. 2019b], which is suitable for stochastic systems, have been proposed and are

becoming widely used.

However, there is a problem when attempting to apply the existing time-series-based causal analyses to ecosystems. Ecosystems are complex systems of various physical, biological, and chemical processes [Durden et al. 2017]. Meteorological variables such as temperature and precipitation are under the influence of atmospheric and oceanic fluid dynamics [Vallis 2017]. Recent studies showed that these earth system dynamics are well captured as stochastic processes, and approaches based on statistical causal analysis are effective in elucidating their relationships [Kretschmer et al. 2016, Runge et al. 2019a,b]. In contrast, the density and abundance of organisms often show strong non-linear dynamics, driven by interactions between organisms. This motivates the application and success of approaches based on nonlinear dynamical systems [Sugihara et al. 2012, Peretti 2013, Deyle et al. 2016a, McGowan 2017]. While some chemicals control ecological dynamics as essential resources and are produced by organismal activities, they are also under the control of global geochemical cycles [Schlesinger et al. 2011]. These processes influence each other, but it is through rather weak causal couplings, and the dynamics of any one process may not become dominant [Sugihara et al. 2012, Tsonis et al. 2015, Deyle et al. 2016b, Rogers and Munch 2020]. Methods for revealing causal relationships among ecosystem components from time series need to be robust to dynamical complexity, i.e., the different levels of stochasticity/non-linearity. However, the assumptions of most methods do not satisfy this requirement [Baskerville and Cobey 2017, Runge 2018].

In this paper, we introduce a new method called EcohNet. It is based on the ensemble prediction of neural networks [Jaeger 2001, 2002, Jaeger and Haas 2004] that can seamlessly handle stochastic/deterministic and linear/non-linear dynamics [Bollt 2021]. Therefore, it is expected to be robust to dynamical complexity. The ensemble prediction is used to decompose relationships among variables in terms of predictability [Pennekamp et al. 2019]. Thus, the contribution of one variable to predictive performance can be evaluated apart from that of the other variables, as in the concepts of partial correlation and conditional independence. It is expected that even if some variables are driven by a strong driver, weak relationships among variables can be detected separately from the effect of the driver without special treatments [Ye et al. 2015, Matsuzaki et al. 2018, Ushio et al. 2018, Chang et al. 2020].

Although various methods have been proposed and used for descriptive purposes, the relevance of their application to actual ecosystem monitoring data has not been seriously examined. We performed benchmarking with data from long-term observation of an aquatic mesocosm, as well as a simulated dataset to test robustness to different dynamical complexity (equilibrium, equilibrium forced by an external oscillator, and intrinsic oscillatory dynamics, as well as different magnitudes of noise), and susceptibility to the often-problematic interaction topologies such as transitive relationships (e.g., in a three-species food chain $X \leftarrow Y \leftarrow Z$, a causal relationship may be identified between X and Z) and intraguild relationships (e.g., in a one-predator-two-prey-relationships $X \leftarrow Z \rightarrow Y$, a causal relationship may be identified between X and Y) [Sugihara et al. 2012, Ye et al. 2015]. We then applied our method to a long-term lake monitoring dataset that includes heterogeneous components such as meteorological, chemical, and biological variables, and interpreted the results. Here, we focused on how EcohNet provides

insight on the drivers of cyanobacteria blooms, which is often a serious threat to ecological integrity and ecosystem services [Carey et al. 2021].

Results

EcohNet combines a recurrent neural network (an echo state network (ESN)) [Jaeger 2001, 2002, Jaeger and Haas 2004, Nakajima and Fisher 2021], with a progressive selection of variables [Friedman et al. 2001] (Figure 1; see Methods for details). Initially, a target variable is selected (Fig. 1A), and the ensemble predictive performance (prediction skill) of one-time-step ahead is evaluated when the past state of itself is given as an input of ESNs (Figure 1B,C). In the next step, we choose a second variable other than the target variable one by one and adopt the one that improves the prediction skill the most. If the prediction skill does not improve, no further new variables are adopted. This process is repeated for variables that have not yet been selected, as long as new variables are adopted. Evaluation of prediction skill is done by different ESNs generated at each round. The neural connections of an ESN are randomly generated, except for the output weights (W_{out}). W_{out} is sequentially updated (as it outputs a one-step-ahead prediction of a target variable for an input) by the RLS (recursive least squares) method. This is an online implementation of ESN, which is robust to the non-stationary dynamics and easily applicable to predictive purposes [Jaeger 2002]. Here, to prevent the results from being dependent on a particular network, the evaluation of prediction skill is done by a population of ESNs. Moreover, to prevent certain contingent features of a time series to affect the prediction skill, the input to a network is replaced by a zero vector with a probability of 1/2 at each time. As a result of the above process, a set of variables $\bar{\Omega}$ that maximizes the prediction skill for the target variable is obtained. Then, $\bar{\Omega}$ is used to evaluate the *unique prediction skill*, which represents the unique contribution of each variable in Ω to the overall prediction skill (Figure 1D). If $\bar{\Omega} = \{X, Y\}$ and the target variable is X, for each variable (here, X and Y), we evaluate how much the prediction changes when one variable is excluded from it (e.g., $\bar{\Omega}^{(-Y)} = \{X\}$). For example, when denoting the prediction skill for $\bar{\Omega}$ and $\bar{\Omega}^{(-Y)}$ as, $\rho_{\bar{\Omega}}$ and $\rho_{\bar{\Omega}^{(-Y)}}$ respectively, the unique contribution of Y on X is $\rho_{Y \to X} = \rho_{\bar{\Omega}} - \rho_{\bar{\Omega}^{(-Y)}}$. By performing the above procedure with all variables as targets in turn, a prediction-skill-based causal network is obtained for the entire system.

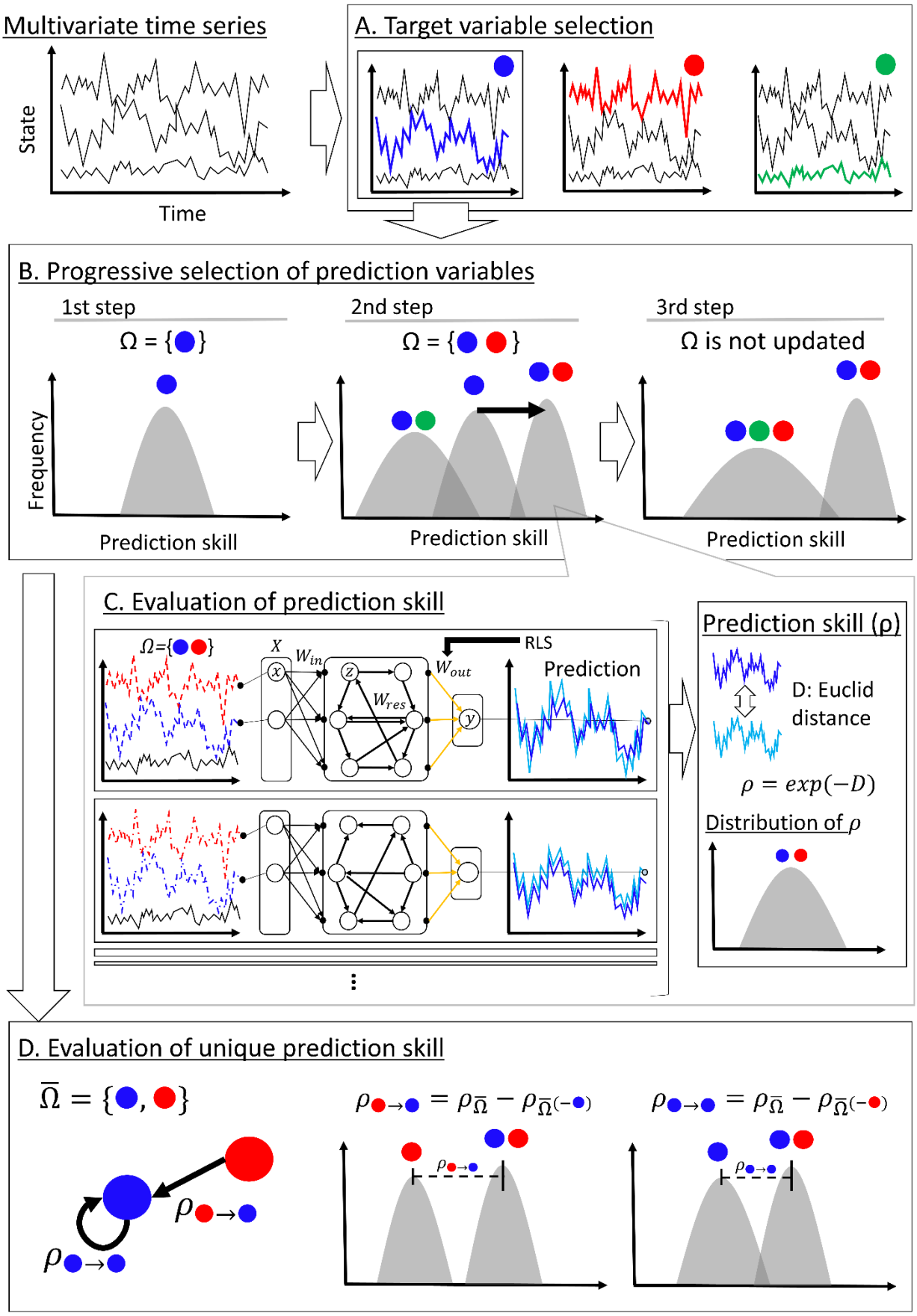

Multivariate time series
State
Time
A. Target variable selection
B. Progressive selection of prediction variables
1st step
Ω = {●}
2nd step
Ω = {● ●}
3rd step
Ω is not updated
Frequency
Prediction skill
Prediction skill
Prediction skill
C. Evaluation of prediction skill
Ω={● ●}
X
W_{in}
W_{res}
W_{out}
RLS
Prediction
Prediction skill (ρ)
D: Euclid distance
$\rho = exp(-D)$
Distribution of ρ
D. Evaluation of unique prediction skill
$\bar{\Omega} = \{●, ●\}$
$\rho_{●\to●} = \rho_{\bar{\Omega}} - \rho_{\bar{\Omega}(-●)}$
$\rho_{●\to●} = \rho_{\bar{\Omega}} - \rho_{\bar{\Omega}(-●)}$

Figure 1. Graphical explanation of EcohNet. See main body and Methods for the explanation.

Benchmarking

When applied to the data from a long-term mesocosm experiment [Beninca et al. 2008], EcohNet identified 11 out of 13 interacting pairs of components, with 3 false positives, and was superior to conventional methods (Fig. 2a-d). It outperformed conventional methods in all evaluation criteria (Fig. 2e; supporting figure S1). The results were not sensitive to different parameter values that might affect the performance of echo state networks, except when the forgetting factor λ is close to 1, which is expected from its definition (see supporting figure S2). A benchmark with datasets generated by food web models (supporting figure S3) also supported the superiority of our method. Except for $O_{S,L}$ in which the signal of interaction was considered to be strong, no other method was able to compete with EcohNet (Fig.2f-h). Performance of EcohNet did not differ from CCM and LIMITS for ROC-AUC and CCM for accuracy and F1-score in $O_{S,L}$, however the performance of CCM and LIMITS was largely reduced in both $EO_{S,L}$ and $E_{S,L}$ for all criteria. These degradations in performance were accompanied by variations in sensitivity and specificity (Fig. 3), suggesting that the fundamental sensitivity of the methods to the presence of interactions is affected by dynamic complexity. In comparison, performance of EcohNet was relatively stable, and its sensitivity and specificity were not greatly affected by the simulation conditions. Our benchmarking also showed that false positives due to transitive relationships and intraguild relationships were best suppressed in EcohNet (Fig. 2e,f).

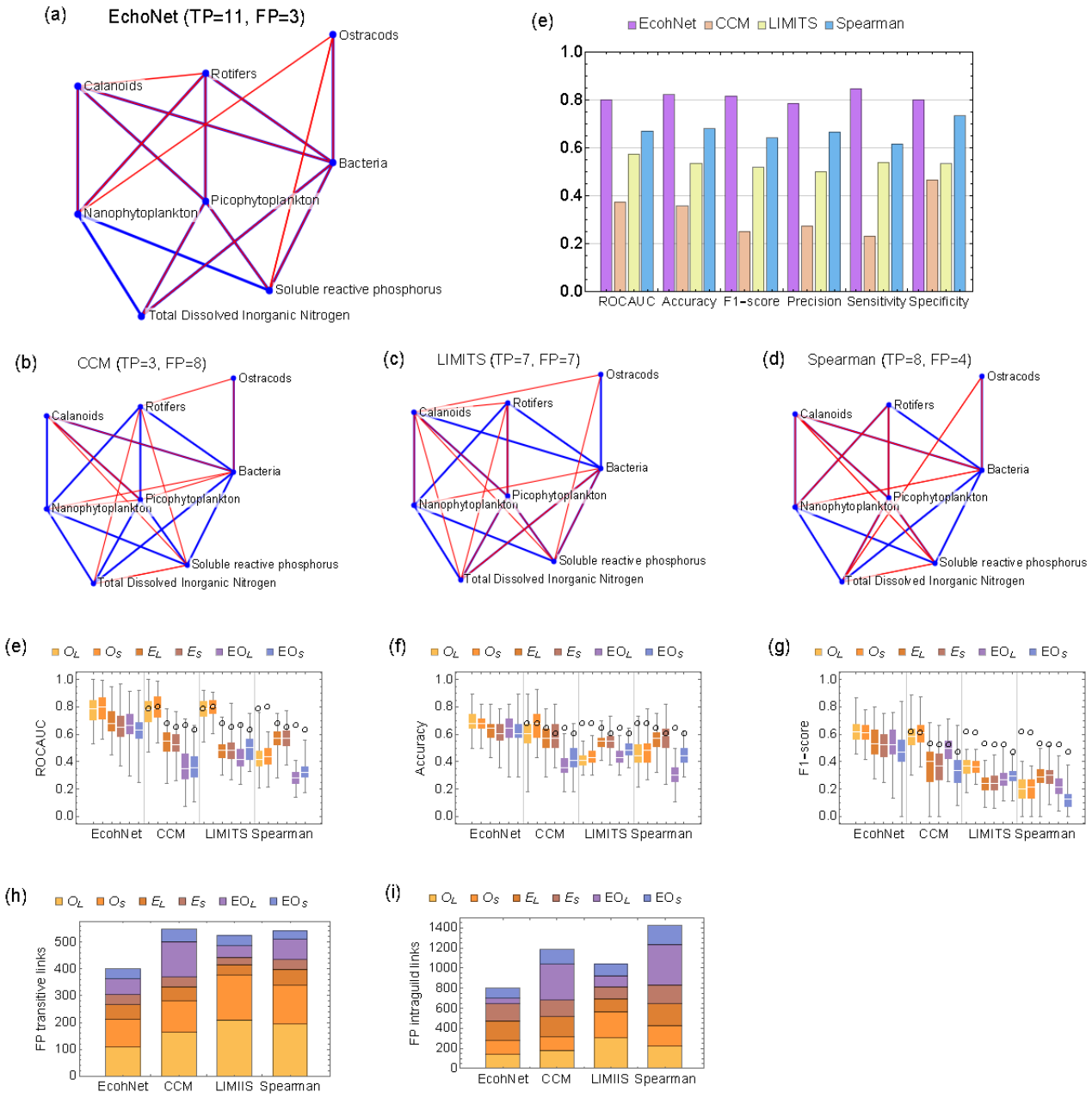


Figure 2. Benchmarking results. (a-d) Actual (blue) and inferred (red) network of EcohNet and conventional approaches (CCM, LIMITS and Spearman rank correlation) for mesocosm data. The number of true positives (TP) and false positives (FP) are shown above the panel. (e)The value of evaluation criteria (supporting figure S1) for mesocosm data. (f-h) The value of three integrative evaluation criteria (ROC-AUC, Accuracy and F1-Score) for simulation data. Here, E is equilibrium, EO is equilibrium forced by an external oscillator and O is intrinsic oscillation, and the suffixes S and L indicates small and large noise magnitude, respectively. In the box plot, white lines indicate the median, box edges indicate the first and third quartile value, and whiskers indicate maximum and minimum values. Black

circles indicate the median value of EcohNet. (i) Cumulative number of incorrectly identified transitive relationships (for $6 \times 100 = 600$ time series in total). (j) Cumulative number of incorrectly identified intraguild relationships (for $6 \times 100 = 600$ time series in total).

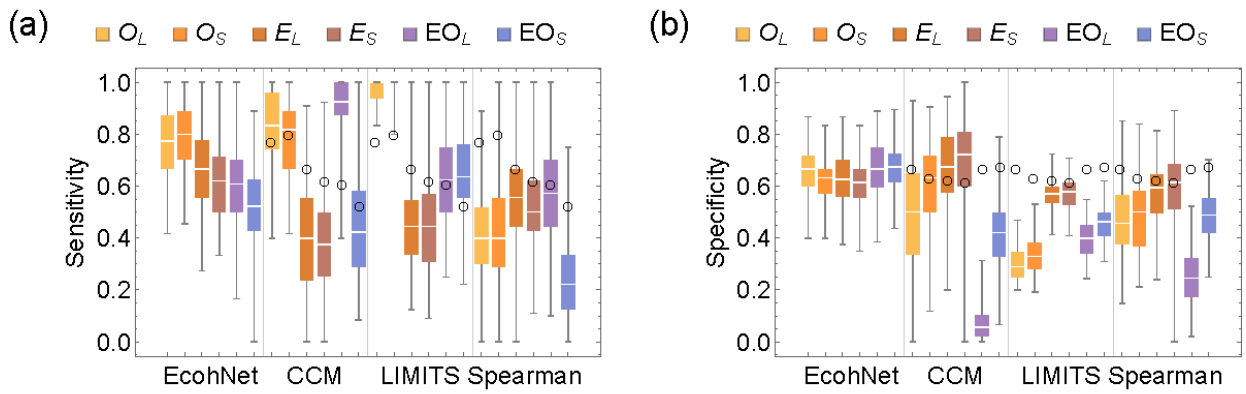


Figure 3. The value of sensitivity (a) and specificity (b) for simulation data. In the box plot, white lines indicate the median, box edges indicate the first and third quartile value, and whiskers indicate maximum and minimum values. Black circles indicate the median value of EcohNet.

Phytoplankton dynamics

We applied EcohNet to a time-series (Lake Kasumigaura Long-term Monitoring Dataset) to examine the top-down and bottom-up causal factors of phytoplankton community composition (Fig.4a). The causal network is naturally organized in a top-down structure with temperature at the top and temperature had the most numbers of interaction links across multiple trophic levels. EcohNet showed that each seven dominant phytoplankton groups were determined by different factors and those networks were complex. Four out of seven phytoplankton groups (all three cyanobacteria and Thalassiosiraceae) were forced by nitrate. We also detected the top-down control of rotifers and calanoids on not only diatoms but also cyanobacteria (Rotifers→Nitzschia, Rotifers→Oscillatoriales, Calanoida→Microcystis, Clanoida→Thalassiosiraceae, Cyclopoida→Aulacoseira), while large- and small-cladocerans did not influence dominant phytoplankton groups. In contrast, three bottom-up links from phytoplankton to zooplankton (Nitzschia→Rotifers, Oscillatoriales→Cyclopoida, Fragilaria→Small-cladocera) were also detected. Importantly, in addition to the effects of environmental variables and zooplankton, we identified three interactions among dominant phytoplankton groups (Nostocales →Oscillatoriales, Oscillatoriales→Aulacoseira, and Fragilaria→Microcystis).

An additional analysis on the direction of causation between meteorological and other components showed that the total number of unrealistic causalities was lowest in the seasonality-adjusted CCM (CCM+), but CCM+ also had fewer links with meteorological variables as factors (Fig. 4b). On the other hand, EcohNet could filter out only the unrealistic causalities. Moreover, unrealistic causal links

were likely to be detected as strong links in CCM, but appeared only as weak links in EcohNet (Fig. 4c).

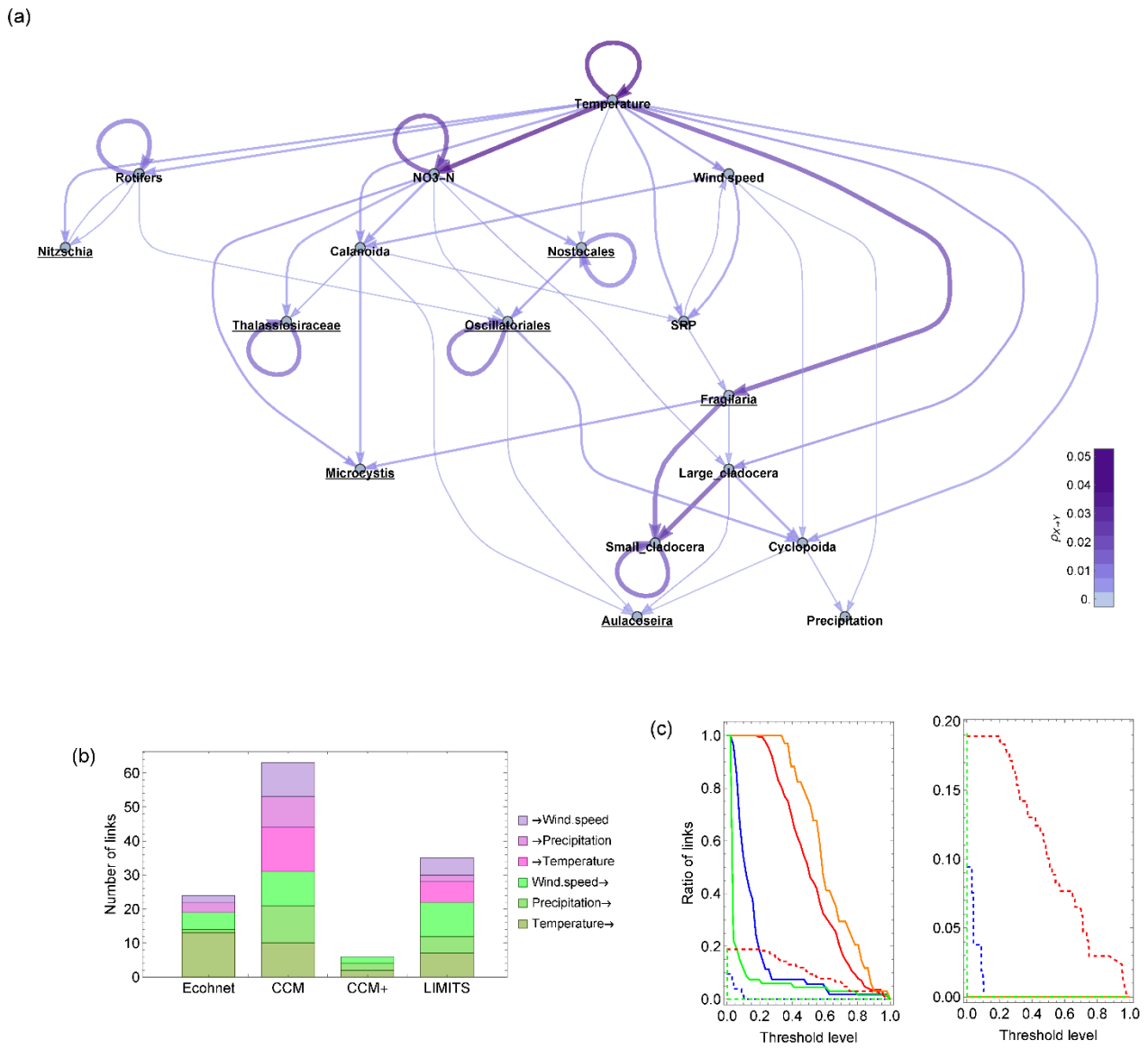


Figure 4. Analysis of Lake Kasumigaura Long-term Monitoring Dataset. (a) Causal network obtained by EcohNet. We removed causal links with a prediction skill of less than 0.002 (23 out of 70 links were removed). The magnitude of the prediction skill is indicated both by the thickness and color of arrows. Seven dominant phytoplankton groups were underlined. (b) We identified causal relationships including meteorological variables (Wind speed, Precipitation, Temperature) and examined how well unrealistic causal links (i.e., organisms/chemicals to meteorological variables) were avoided while the effect of meteorological variables were detected. (c) Ratio of causal links larger than the threshold level (x-axis) is shown for all links (solid lines) and unrealistic links (dashed lines) (blue: EcohNet, red: CCM, green:

LIMITS). The right panel is an enlarged inset of the left panel to show the low values. The faster the dashed lines converge to zero, the smaller the unique prediction skill of unrealistic causal links, and the easier it is to remove them while retaining more reliable links.

Discussion

We developed and evaluated the effectiveness of a method, EcohNet, that utilizes ensemble predictions of online echo state networks. In short, our method encompasses the scope of Granger causality [Granger 1969] and Takens' embedding theorem [Takens 1981], and implements a framework for decomposing relationships among variables in terms of predictability [Pennekamp et al. 2019]. Although ESNs have been applied to causal analysis [Huang et al. 2020, Duggento et al. 2021, Wang et al. 2022], we developed for the first time a reliable method of causal network inference by integrating adaptive online ESNs [Jaeger 2002] into an ensemble machine learning framework. In addition to its applicability to nonlinear dynamics, our approach is not affected by the reliability of nonlinear prediction methods on arbitrary underlying dynamics, which differs from previous approaches based on nonlinear time series analysis [Sugihara et al. 2012, Suzuki et al. 2017, Ushio et al. 2018, Cenci et al. 2019, Chang et al. 2021]. As we have confirmed (Fig. 2e-g; also refer to Cobey and Baskerville 2016, Monster et al. 2017), performance of CCM relies on a nonlinear forecasting method (state space reconstruction and simplex-projection) and was sensitive to the differences in dynamic complexity, although it performed better than the equation-based (LIMITS) and correlation-based (Spearman's rank correlation coefficient) approaches. Another advantage of our method is its capability to decompose relationships among variables; we confirmed that false positives due to transitive relationships and intraguild relationships were best suppressed in EcohNet (Fig. 2e,f). A similar approach has been attempted by Leng et al (2020) but, as an extension of CCM, it was based on a specific procedure of nonlinear time series analysis, and thus its robustness to dynamic complexity will be limited.

EcohNet not only outperformed other methods in ROCAUC, accuracy and F1-score, but was robust in that its sensitivity and specificity were not greatly affected by dynamical complexity (Fig. 3). This robustness is a plausible feature because it means that the error rate of an inferred network is consistently controlled by the regularization scheme (see Box 1 and Materials and Methods) irrespective of the dynamics. For example, comparison of the connectivity of ecological networks from several sites in which the dynamic complexity of components are different would not be possible if the method applied is not robust to the inherent dynamics. We also confirmed that EcohNet could identify causation between meteorological and other components while filtering out the unrealistic direction of causalities (Fig. 4b,c; Baskerville and Cobey 2017). The unrealistic causal links were likely to be detected only as weak links in EcohNet. Because of observational noise and other factors, it is hard to completely avoid detecting unrealistic causalities. Therefore, it is a desirable feature that the importance of unrealistic causality is underestimated.

EcohNet revealed that temperature determined nutrient dynamics, phytoplankton, and

zooplankton communities in Lake Kasumigaura. This result is similar to Tanentzap et al. (2020) that demonstrates strong direct and indirect impacts of temperature on phytoplankton and zooplankton. Quantifying the causal effects of temperature is an important advantage of EcohNet. CCM failed to distinguish causal relationships from seasonality-driven synchronization, leading to misidentification of causality [Kitayama et al. 2021]. Likewise, previous work on Lake Kasumigaura did not detect the causal effects of temperature on phytoplankton and zooplankton [Matsuzaki et al. 2018]. Our results suggest that temperature may determine the whole food web structure of Lake Kasumigaura.

Clearly, each seven dominant phytoplankton groups were determined by different factors (Fig. 4). Some studies reported that phytoplankton species or genera have unique physiological and ecological features and thus respond differently to environmental factors and grazing [Huber et al. 2012, Freeman et al. 2020]. These differences can result in dynamic changes in phytoplankton community composition, since it is reported that the dominant phytoplankton group changed temporally in Lake Kasumigaura [Takamura et al. 1992, Tomioka et al. 2011, Fukushima and Arai 2015]. Nitrate determined four out of seven phytoplankton groups. This is consistent with previous work on Lake Kasumigaura showing nitrogen limits phytoplankton primary production [Matsuzaki et al. 2018]. Not only diatoms but also cyanobacteria were influenced by rotifers, calanoids, and cyclopods. Although our analysis did not identify whether these are direct predation or indirect effects, it is known that some copepods can ingest and shorten the filament size of cyanobacteria and rotifers can graze or utilize decomposed cyanobacteria [Ka et al. 2012, Ger et al. 2016]. Our results suggest that these zooplankton groups may have an important role in the food web of shallow hypereutrophic lakes.

We identified significant interactions among dominant phytoplankton groups. It has been reported that phytoplankton groups may have compensatory responses to environmental factors [Jochimsen et al. 2013]. One possible mechanism for the interaction between Anabaena and Oscillatoriales could be nitrogen availability, because N2-fixing cyanobacteria, including Nostocales, compete with non-N2-fixing bacteria, including Oscillatoriales [Moisander et al. 2008]. Another mechanism could be light availability. Mixing regimes determined by climatic factors like heat exchange and wind action affect competition for light between phytoplankton species (e.g., buoyant cyanobacteria and sinking diatoms), and shading by cyanobacteria blooms also influences other phytoplankton (Huisman et al. 2004). As pointed out in Freeman et al. (2019), previous studies have not incorporated these interactions in models for identifying species-specific factors. Explicitly including not only environmental factors and grazing impacts but also community interactions in complex systems is another advantage of EcohNet.

As well as predicting the flow of causation, EcohNet enables cyanobacteria bloom forecasts for a lake (supplementary figure S4,5). The application of EcohNet to long-term monitoring data can be useful tool for lake management. Harmful cyanobacteria blooms are a serious threat to ecological integrity and ecosystem services, despite management efforts, and climate change is predicted to promote the occurrence and severity of cyanobacterial blooms [Huisman et al. 2018]. Identifying the drivers causing cyanobacteria blooms and forecasting these blooms are critically important for water quality management

[Carey et al. 2021]. As shown in our study, environmental factors, the interaction between phytoplankton-zooplankton, and the interaction between phytoplankton and phytoplankton cause cyanobacteria bloom, and temperature strongly affects the whole food-web; the system is complex.

The development of frameworks to facilitate time-series-based causal analysis would advance the field of ecology in the coming decades, where advanced techniques for ecosystem monitoring [Caporaso et al. 2010, Bohan et al. 2017, Baliint et al. 2018, Hasse et al. 2018, Taberlet et al. 2018] will lead to unprecedented increases in dataset size [Hampton et al. 2013, Mouquet et al. 2015, Durden et al. 2017]. Such analysis could promote studies to estimate the effects of global environmental change on biological communities [Chang et al. 2020], prevent regime shifts that lead to catastrophic impacts on biodiversity [Scheffer and Carpenter 2003], and support efficient management of biological resources [Chivian and Bernstein 2008]. Further development by applying EcohNet to time-series of high-frequency data, such as sensor data, could lead to real-time, near-term forecasting to prepare for or preempt future impairment of ecological functions and services. It is expected that with integration to active hypothesis testing, these studies will lead to predictive and manipulative ecology in the 21st century using a more systematic focus.

Materials and Methods

- EcohNet

-- Echo state network with a recursive least squares algorithm

An echo state network (ESN; Jaeger 2001, Jaeger and Haas 2004) implements a type of reservoir computing that uses a recurrent neural network (RNN) as a dynamical reservoir (an internal structure which is made up of individual, non-linear units, and can store information). ESNs can skillfully reconstruct and predict time series from different nonlinear dynamical systems (Pathak et al. 2017, Du et al. 2017, Lu et al. 2017, Lu et al. 2018, Pathak et al. 2018, Chattopadhyay et al. 2019, Reichstein et al. 2019). It has been demonstrated that ESNs can track the temporal evolution of time series better than the backpropagation-based artificial neural networks.

In an ESN an input signal to the dynamical reservoir RNN induces a nonlinear response, and the neural states are converted to an output signal by trainable linear weights (Fig. 1C). Assuming that $X = \{x_1, x_2, \dots, x_n\}$ is the system variables, $x_i \in X$ is a target variable, and $X^{\Omega} \subset X$ are the input variables which is a subset of X specified by a set of indices $\Omega \subset \{1,2, \dots, \mathrm{n}\}$, we use the ESN for one-step-ahead prediction of a target variable, *i.e.*, to predict $x_i(t+1)$ given $X^{\Omega}(t)$. Here, $t \in \{0,1, \dots, T-1\}$ specifies the system state at discrete time points. Following the typical implementation, we defined the ESN by three matrices and one function. The matrices are: 1) the input weight matrix, $W_{in} \in R^{n_{res} \times n_{in}}$; 2) the pool weight matrix, $W_{res} \in R^{n_{res} \times n_{res}}$ and 3) the output vector, $W_{out} \in R^{n_{res}}$, where n_{in}, n_{res} is the number of nodes in the input layer, pool, respectively. The state update equations for the ESN are as follows:

$z(t+1) = f(W_{in} X^{\Omega}(t) + W_{res} z(t))$, (1)

$y(t) = W_{out}(t) z(t+1)$. (2)

Here, $X^{\Omega}(t)$ is a vector of current inputs to the ESN, $y(t)$ is a vector of current outputs from an ESN and

$z(t)$ is a vector of current neuronal activations in an ESN pool. We assume that the pool neuron activation function as $f(z) = tanh\, z$. W_{in} and W_{res} is fixed once an ESN is generated (black allows in Fig.1C). The element of W_{in} is randomly drawn from a uniform distribution between -0.1 and 0.1. Following the standard practice in generating an ESN [Jaeger 2002], we scaled the spectral radius of W_{res} as ρ after it is generated by randomly drawing its elements from a uniform distribution between -1 and 1 (with probability 0.1, and otherwise zero). Only the output weight vector W_{out} is trainable (yellow arrows in Fig. 1C).

To evaluate the predictive performance across the full-length of a time series, we performed a recursive online update of W_{out}, using the Recursive Least Squares (RLS) algorithm [Jaeger 2002]. RLS is widely used in linear signal processing and has a desirable feature of fast convergence [Farhang-Boroujeny 2013]. It incorporates the error history of a system into the calculation of the present error compensation. The equations for recursive updating rule of RLS are as follows:

$W_{out}(t+1) = W_{out}(t) + v(t)g^T(t)$, (4)

$v(t) = x_i(t) - y(t)$, (5)

$g(t) = \frac{P(t-1)z(t)}{\lambda + z(t)^T P(t-1)z(t)}$, (6)

$P(t) = \lambda^{-1}[P(t-1) - g(t)z^T(t)P(t-1)]$. (7)

Here, λ is the forgetting factor, which determines the effect of past errors on the update, $v(t) \in R$ is the output error at time t when applying weight matrix before the update ($W_{ont}(t)$), $g(t) \in R^{n_{res}}$ is the gain vector, and $P(t) \in R^{n_{res} \times n_{res}}$ is the inverse of the self-correlation matrix of $z(k)$ ($k = 1,2, \ldots, t$) weighted by the forgetting factor. This update can be repeated multiple times at each time step, and we define the number of iterations as τ. We set the initial conditions of the update rule as,

$P(0) = \delta^{-1} I$, (8)

$W_{out}(0) = O$, (9)

where $I \in R^{n_{res} \times n_{res}}$ is the identity matrix, $O \in R^{n_{res}}$ is a zero-vector and δ is the regularization factor. This algorithm minimizes the following cost function, which is a weighted sum of the output errors at time $k = (1,2, \ldots, t)$ when applying the output weight matrix W_{out} at time t with a regularization term:

$E_{RLS}(t) = \frac{1}{2}\sum_{k=1}^{n} \lambda^{n-k} \|x_i(t) - W_{out}(t)z(k)\|^2 + \frac{\delta}{2}\lambda\|W_{out}\|^2$. (10)

Here, $\|\cdot\|$ represents the L2 norm.

The parameter values were set according to the standard recommendation for parameterization of the ESN-RLS (e.g., Ifeachor and Jervis 2002, Jaeger and Haas 2004, Lukoševičius 2012; Table 1, see also supporting figure S2). In our first benchmark, we examined the stability of the results for four representative parameters that may affect the performance of ESNs (supporting figure S2). To prevent certain contingent features of a time series to affect the prediction skill, the input to a network is replaced by a zero vector with a probability of 1/2 at each time (illustrated by the difference of line style of the input time series (indicated by red and blue) in Fig.1C).

Table 1. Parameters of ESN-RLS. For the parameters marked by *, we tested the impact on EcohNet's performance with different values. The tested values are $\tau \in \{2,4,6,8,10\}$, $\lambda \in \{0.8,0.9,0.95,0.98,0.995\}$, $\delta \in \{0.0001,0.001,0.01,0.1\}$, $\rho_0 \in \{0.7,0.8,0.9,0.95,0.98\}$.

Description	Value
Number of iterations of RLS updates (τ)	8*
Forgetting factor of RLS (λ)	0.95*
Regularization coefficient (δ)	0.001*
Spectral radius of W_{res} (ρ_0)	0.95*
Connectance of W_{in}	1
Range of values in W_{in}	[-0.01,0.01]
Number of nodes in W_{res}	32
Connectance of W_{res}	0.1
Range of values in W_{res}	[-1,1]

-- Progressive selection of input variables

For a target variable x_i (illustrated by the blue circle in Figure 1A), we used a progressive selection to obtain the set of variables Ω_{x_i} that optimize the prediction of x_i (illustrated by the set of blue and red circles in Fig. 1B). For this purpose, we first define the prediction skill in this paper as:

$$\rho_{\Omega_{x_i}} = \rho\left(x_i, y\left(X^{\Omega_{x_i}}|ESN_j\right)\right) = e^{-\left\|x_i - y\left(X^{\Omega_{x_i}}|ESN_j\right)\right\|}. \quad (11)$$

Here, x_i is the time series of target variable (blue lines in Fig.1C) and $y\left(X^{\Omega_{x_i}}|ESN_j\right)$ represents the prediction of x_i given a set of variables $X^{\Omega_{x_i}}$ and an echo state network indexed by j (light blue lines in Fig.1C). Practically, we calculated the prediction skill after truncating the first 20% of the time series. It is intended to remove the effect of W_{out} before convergence. Although not explicitly stated in the left-hand side representation, it should be noted that a single predictive skill is given by a single echo state network.

Assuming that the variable is $X = \{x_1, x_2, \ldots, x_n\}$ and the target variable is x_i, the progressive selection of variables proceeds as follows. First, the set of the index of input variables to predict variable x_i is initialized to $I_{active} = \{i\}$ because it is natural to assume that the past state of a variable can be used to predict its own (first step in Fig.1B). Correspondingly, the set of the index of the rest of the variables is initialized to $I_{inactive} = \{j\}_{j \neq i}$. Then, N ESNs are generated to perform predictions of x_i by setting $\Omega_{x_i} = I_{active}$ (Fig.1C). This returns distribution of prediction skills $D_T(\rho_{\Omega_{x_i}})$ (illustrated by the gray mountain shapes in Fig.1B). To select a next variable to be added, for each index j in $I_{inactive}$, $I_{test}^{(h)} = I_{active} \cap \{j\}$ are generated, where the suffix h indicates that j is the hth element of $I_{inactive}$, and then, for each $I_{test}^{(h)}$, N

ESNs are generated to obtain predictions of x_i by setting $\Omega_{x_i} = I_{test}^{(h)}$ (2nd step in Fig.1B). Among the distribution of the prediction skills for hs, h^* such that $I_{test}^{(h^*)}$ attains the highest median value of prediction skill is set as $D_{T+1}(\rho_{\Omega_{x_i}})$ (illustrated as the distribution corresponding to the set of blue and red circles in Fig.1B). The criterion used to determine whether or not to add the variable j corresponding to h is, $\#\{D_T(\rho_{\Omega_{x_i}})|\rho_{\Omega_{x_i}} > Q_{1/2}(D_{T+1}(\rho_{\Omega_{x_i}}))\}/N < q$, where $\#\{D|criteria\}$ is the number of elements in D that satisfies the *criteria*, $Q_{1/2}(D)$ is the median of a distribution D. This states that the median of the prediction skills of D_{T+1} must be larger than $100(1-q)\%$ of the prediction skills of D_T. If the criterion is satisfied, we replace $D_T(\rho_{\Omega_{x_i}})$ by $D_{T+1}(\rho_{\Omega_{x_i}})$, I_{active} by $I_{test}^{(h^*)}$, increment T by 1, and test whether more variables are acceptable, otherwise the procedure is stopped (3rd step in Fig.1B) with returning $\overline{\Omega}_{x_i} = I_{active}$ as the set of the indices of optimal variables that maximizes the prediction of x_i.

The number of ESNs generated for each evaluation, N, is related to the stability of the results. If N is small, the variability in results from trial to trial may be large, especially for weak causality. In this paper, we set $N = 10000$ for real-world monitoring data where the data set consists of a single time series, and $N = 2000$ for simulation data sets to consider the time required for evaluation. The threshold value for accepting a new variable, q, was set to zero throughout this paper. This number refers to the ratio of predictive skills before the addition of the new variable that exceeded the median predictive skill after the addition. Therefore, it is recommended that this value be smaller than 0.5. The closer the number is to 0.5, the more likely it is that weak links will remain. Therefore, instead of increasing the sensitivity, there is a possibility of decreasing the specificity. In this paper, the value is set close to 0.5, taking into account the evaluation by ROC-AUC.

Box 1. Summary of progressive selection of variables

Assuming that the variable is $X = \{x_1, x_2, \ldots, x_n\}$ and the target variable is x_i.

1. Set $T = 1$, $I_{active} = \{i\}$ and $I_{inactive} = \{j\}_{j \neq i}$.
2. Generate N ESNs and perform predictions of x_i by setting $\Omega_{x_i} = I_{active}$, and obtain a distribution of prediction skills $D_T(\rho_{\Omega_{x_i}})$.
3. For each index j in $I_{inactive}$, generate $I_{test}^{(h)} = I_{active} \cap \{j\}$, where the suffix h indicates that j is the hth element of $I_{inactive}$.
4. For each $I_{test}^{(h)}$ generate N ESNs and perform predictions of x_i by setting $\Omega_{x_i} = I_{test}^{(h)}$.
5. Select h^* such that $I_{test}^{(h^*)}$ attains the highest median value of prediction skills, and set the corresponding distribution as $D_{T+1}(\rho_{\Omega_{x_i}})$.
6. If $\#\{\rho_{\Omega_{x_i}} \in D_T(\rho_{\Omega_{x_i}})|\rho_{\Omega_{x_i}} > Q_{1/2}(D_{T+1}(\rho_{\Omega_{x_i}}))\}/N < q$ or $T = n$ set $D_T(\rho_{\Omega_{x_i}}) \leftarrow D_{T+1}(\rho_{\Omega_{x_i}})$, $I_{active} \leftarrow I_{test}^{(h^*)}$,

increment T by 1 and back to 3 else proceed to 7.
7. Return $\overline{\Omega}_{x_i} = I_{active}$.

-- *Calculation of unique prediction skill*

Unique prediction skill quantifies the unique contribution of a single variable to the overall prediction skill (Fig. 1D). For each $j \in \overline{\Omega}_{x_i}$, the unique prediction skill is defined as,

$$\rho_{x_j \to x_i} = Q_{1/2}(D(\rho_{\overline{\Omega}_{x_i}})) - Q_{1/2}(D(\rho_{\overline{\Omega}_{x_i}^{(-j)}})). \quad (12)$$

Here,

$$\overline{\Omega}_{x_i}^{(-j)} = \overline{\Omega}_{x_i} \backslash \{j\}, \quad (13)$$

i.e., $\overline{\Omega}_{x_i}^{(-j)}$ is obtained by removing j from $\overline{\Omega}_{x_i}$. $D(\rho_{\overline{\Omega}_{x_i}})$ and $D(\rho_{\overline{\Omega}_{x_i}^{(-j)}})$ is again calculated by N ESNs

(practically, $D(\rho_{\overline{\Omega}_{x_i}})$ only needs to be calculated once for each $\overline{\Omega}_{x_i}$). $\rho_{x_j \to x_i}$ (illustrated as the difference of the positions of two distributions in Fig.1D) represents how x_j uniquely contributes to the prediction of x_j. We identify $\rho_{x_j \to x_i}$ as the indicator of causal influence from x_j to x_i. In this step, sometimes the unique prediction skill became negative, and we replaced such relationships with 0 (i.e., no causality).

- *Conventional approaches used for comparison*

-- *Spearman Correlation Coefficient*

Spearman rank correlation coefficient [Spearman 1904] is a nonparametric measure of rank correlation. It is relevant even if both or one of the variables is non-Gaussian and thus is more broadly applicable than Pearson correlation. The correlation coefficients are frequently used for network-based analysis of biological systems [Weiss et al. 2016]. This method provides correlation coefficients between -1 and 1 for each pair of variables, accompanied by the p-values. The correlation matrix and p-values are symmetric.

-- *Convergent Cross Mapping (CCM)*

CCM [Sugihara et al. 2012] is a statistical test for a causal relationship between two time series variables and attempts to address the problem that correlation is often not an indicator of the presence or absence of actual causal relationships. This method is based on Takens' embedding theorem [Takens 1981], which states that the essential information of a multi-dimensional dynamical system is retained in the time series of any single variable of that system. Based on this theory, in a pair of time series (X, Y), if X has a high forecasting skill in predicting Y, causality will be detected in the direction of $Y \to X$. The predictive skill of CCM, denoted as $\rho_{Y \to X}$, is quantified by the Pearson correlation coefficient between

actual Y and Y predicted by X. Different from the correlation-based methods, $\rho_{Y\to X}$ is usually unequal to $\rho_{X\to Y}$. Its p-value is calculated by the bootstrapping procedure to account for the effects of the predictability inherent in the target time series (such as periodicity). To apply this method to systematic network inference, we needed to extend the implementation by the authors (Appendix A).

To perform CCM, it is necessary to determine the appropriate embedding dimension for the state space reconstruction. We used a one-step-ahead self-prediction [Sugihara et al. 2012] to determine the embedding dimension, E. Then, the strength of the causal relationship between X and Y was obtained as follows. First, we prepared the state space M_Y reconstructed by the Y's embedding dimension E_Y. Then, we predict X from M_Y using the simplex projection method [Sugihara et al. 2012], which is labeled as X^Y. The cross-map prediction skill $\rho_{X\to Y}$ is the Pearson's correlation coefficient between X and X^Y. To calculate the p-value, we followed the implementation by Clark et al. (2015); we sampled 75% of points in M_Y without replacement and repeated the same procedure as above to obtain the prediction skill. We repeated this procedure $N = 500$ times to obtain a distribution of prediction skills, C_1. We then obtained another set of prediction skills, C_2, by the same manner, but by sampling $2E$ points. We then calculated p-value as $max(q_1, q_2)$ (this is not a p-value in the strict sense, but it aims to avoid the matrix from becoming extremely dense) where q_1 is the ratio of elements in C_1 that is smaller than zero and q_2 is the ratio of elements in C_1 that is smaller than the value of 95% tile of C_2.

-- *LIMITS*

LIMITS is based on the generalized Lotka–Volterra difference equation. The algorithm combines forward stepwise regression and bootstrap aggregation (bagging) to determine, in a majority voting fashion, the pairs of interacting variables and the value of their interaction coefficient. For each species, the logarithm of the change in abundance per time step is used as the response variable. Then, this method selects variables to be explanatory variables of linear regression step by step, as long as the predictive performance is improved. The species selected as the explanatory variables are candidates for interaction partners. This process is repeated (here, 500 times) as bootstrap replicates, and the species selected in more than half of all iterations are finally determined as the interaction partners, i.e., variables that can affect the response variable. The interaction is estimated asymmetrically and in a way that includes positive and negative values. For each pair of species, the p-value was obtained as the number of bootstrap trials in which the interaction coefficient was zero divided by the number of all trials. This method returns a sparse matrix for the interaction matrix, but that of the p-values can be dense.

-- Sparsity

In this paper, we have adjusted the sparsity of the Spearman rank correlation matrix based on the p-value, i.e., we made the correlation matrix sparse by replacing elements with $p < 0.01$ with 0. In the same manner, we set the threshold p-value of CCM and LIMITS as 0.1 and 0.002, respectively. These values were selected by balancing the connectivity (whether or not the results will be close to EcohNet) when applied to the data set from a long-term mesocosm experiment we used for benchmarking and the values

often used in practice. We fixed these values throughout this paper to avoid arbitrary adjustments of sparsity.

- *Evaluation*

We evaluated the performance of EcohNet and conventional methods in detecting interaction links using the evaluation criteria of a binary classification task (Supporting Figure S1). Specifically, we considered prediction skill (EcohNet and CCM), correlation coefficient (Spearman rank correlation) and interaction coefficient (LIMITS) as a classifier to identify interactions, and evaluated their performance by ROC (receiver-operator characteristic) curve and six numerical metrics, namely, ROC-AUC (the area under the curve of a ROC curve), accuracy, F1-score, precision, sensitivity and specificity.

In predator-prey relationships species affect each other directly, whereas in a causal relationship, a relationship between species may only be detected in one direction due to a difference in the time-dependence of each effect (Sugihara et al. 2012, Ye et al. 2017, Matsuzaki et al. 2018). Therefore, in this paper, when evaluating the performance of CCM and EcohNet using the evaluation criteria of the binary classification task, we consider the ability to detect interacting pairs (pairs where at least one has a direct influence on the other). Specifically, we symmetrized the interaction matrix (which includes both the actual interaction matrix and a corresponding matrix of prediction skills inferred by CCM or EcohNet) so that $a_{ij}, a_{ji} \leftarrow max(|a_{ij}|, |a_{ij}|)$ before evaluation, and evaluated the performance of prediction skill as a classifier in detecting elements for which $i < j$ and whose values are non-zero in the (symmetrized) actual interaction matrix. The matrix of Spearman rank correlation, which is originally symmetric, was evaluated in the same manner. For LIMITS, we evaluated the performance of detecting the non-zero components of the asymmetrical interaction matrix according to its original definition. In other words, we evaluated the performance of detecting the presence of directional effects (effects from a species on another).

- *Data set*

We benchmarked our method using two datasets and then applied it to a long-term monitoring dataset from a lake ecosystem. One of the benchmarking datasets was from a real ecological system, which was obtained by long-term observation of a mesocosm (Beninca et al. 2008). The other datasets were obtained by simulations of food web models and allowed us to examine how dynamic complexity of a time series affect identification of causal relationships, and how the effects of two types of interaction topologies (transitive and intraguild relationships) can cause false positives. Finally, long-term monitoring data of an actual lake were analyzed to examine the validity of the results from the viewpoint of limnology and aquatic ecology and to develop further applications on the management of lake ecosystems.

-- Long term mesocosm experiment (Beninca et al. 2008)

We first benchmarked our method using the dataset (transformed_data_Nature2008) of a long-term mesocosm experiment by Beninca et al. 2008, which is available as a supplement of Beninca et al. 2009 (https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fj.1461-

0248.2009.01391.x&file=ELE_1391_sm_appendix+S1.xls). The mesocosm consisted of a cylindrical plastic container (74cm high, 45cm diameter), which was filled with a 10cm sediment layer and 90 ℓ of water from the Baltic Sea. The mesocosm was maintained in the laboratory at a temperature of about 20℃, a salinity of about 9‰, incident irradiation of 50 μ mol photons $m^{-1}s^{-1}$ (16:8 hour light:dark cycle) and constant aeration.

The authors identified three different strengths of interactions between species in this mesocosm. We included all of them to evaluate the performance of our selected methods. However, we only considered 9 components (Calanoids, Rotifers, Ostracods, Nanophytoplankton, Picophytoplankton, Bacteria, Total dissolved inorganic nitrogen, Solvable reactive phosphorus (SRP)) to use a time series where as many species as possible coexisted. As a result, we used data from 700 days (210 points) out of the 2319 days of the observation period. In this analysis, we set the parameters of the ESN-RLS following a standard guideline for the parameterization [Ifeachor and Jervis 2002, Jaeger and Haas 2004, Lukoševičius 2012] and examined the sensitivity of the results to the four representative parameters that can affect the performance of an ENSs (supporting figure S2).

-- *Food web model*

We used a food web model to generate a data set for benchmarking. The model is based on a generalized Lotka-Volterra equation (GLVE),

$$\frac{dx_i}{dt} = F_i(x) = x_i\left\{r_i\left(1 - \frac{\Sigma_{j=1}^{M} x_j}{K}\right) + e^{-1}\Sigma_{j=1}^{M}\frac{a_{ji}x_ix_j}{\beta_i + \Sigma_{k=1}^{M} a_{ki}x_k} - \Sigma_{j=1}^{M}\frac{a_{ij}x_ix_j}{\beta_j + \Sigma_{k=1}^{M} a_{kj}x_k} - mx_i\right\}, \quad (14)$$

which has been frequently used to benchmark methods for inferring interaction networks (e.g., Fisher and Mehta 2014; Bucci et al. 2016; Suzuki et al. 2017). Here, x_i is the abundance of species i, r_i is the intrinsic growth rate, K is the carrying capacity, e is the conversion efficiency, $a_{ij} \in A$ is the per capita grazing effect of species j on i (matrix A is an interaction matrix that represents the strength of interaction between species), $\beta_i \in \beta$ is the half-saturation constant of the interspecific interaction, and m is the mortality rate. M is the number of species and we set $M = 16$ throughout this paper.

Based on eq. (14), we generated time series by numerically solving,

$$x_i(t + \Delta t) = x_i(t) + F(x)\Delta t + \eta_i(t) + \epsilon. \quad (15)$$

Here, $\eta_i(t)$ is a random vector whose elements are drawn from a normal distribution with mean 0 and standard deviation $\sigma\sqrt{\Delta t}$, Δt is the time step for numerical simulation, and ϵ is the small size of influx introduced to avoid underflow. We solved eq.(15) for 10,000 steps with $\Delta t = 0.05$ and dropped the first 2,000 steps to remove the transient. Then, we resampled the time series with the interval of 40 and obtained a time series with 200 points. Each data set for different conditions consists of 100 time series simulated with randomly generated A and β. For a data set, a time series is accepted if the number of major species (here, it is defined as a species whose mean abundance is larger than 0.01) is equal to 8, and the abundance of species other than the major species is removed. Thus, one time point of a time series in a data set is a vector of eight elements. This corresponds to the fact that some of the minor species may not be included in the observations. Before the analysis, the time series were normalized to have a mean 0

variance of 1.

To generate the time series of different dynamical complexity, we used three different parameter settings and two different noise magnitudes. The parameter settings are represented as E, EO and O. E is the equilibrium in which $e = 0.5$ and $m = 2$, EO is the equilibrium forced by an external oscillator in which $e = 0.5$ and $m = 2$ but K is replaced by a function $K(t) = 1 + 0.8 \sin t/20\pi$, O exhibits intrinsic oscillation with $e = 0.33$ and $m = 1$. We used two different values $\sigma = 0.3$ and 0.85 for small and large noise magnitude, and indicated them by suffix S and L, e.g., E_S and E_L. Other parameter values were as follows. The interaction matrix specifies the food web by a cascade model [Cohen and Newman 1984]. Here, if the species' index j is greater than i, j preys on i with probability 0.33 ($a_{ij} = 0$ if $i \geq j$ and otherwise randomly drawn from a uniform distribution between 0.05 and 2 with probability 0.33 and otherwise 0). We set $r = 1$ and $K = 1$ for the species having no prey and $r = -1$ and $K = 10^8$ otherwise. The value of β was randomly assigned between 0.01 and 0.1.

-- Lake Kasumigaura Long-term Monitoring Dataset

We applied the EcohNet to the long-term monitoring data from Lake Kasumigaura, a hypereutrophic lake. Lake Kasumigaura is the second largest lake in Japan (167.7 km^2) and is shallow (mean depth: approximately 4 m, maximum depth: 7.4 m). The National Institute for Environmental Studies has been conducting monthly monitoring in Lake Kasumigaura since 1976 and publishing the long-term data on the Lake Kasumigaura Database. In this lake, cyanobacteria blooms occur and disappear repeatedly, and the dominant cyanobacteria group changes. Phytoplankton communities can interact with many components, such as temperature, nutrients, and zooplankton, and are complex and dynamic. Using EcohNet, we quantified the causal relationships among environmental variables, seven dominant phytoplankton groups (three cyanobacteria (Microcystis, Nostocales, Oscillatoriales) four diatoms (Thalassiosiraceae, Aulacoseira, Fragilaria, Nitzschia)), and zooplankton, and examined the interactions among these phytoplankton groups.

We analyzed the monitoring data at the center of Takahamairi Bay (Station 3), which is shallow (ca. 3.2 m) and the most eutrophic site with cyanobacteria blooms. Phytoplankton biovolume data was obtained from Takamura and Nakagawa (2016). For zooplankton, we used the abundance data of five functional groups: large cladocerans (>1.0 mm), small cladocerans (<1.0 mm), rotifers, adult calanoids, and adult cyclopoids (Matsuzaki et al. 2018). Zooplankton data were obtained from Takamura et al. (2017). Our key environmental variables were surface water temperature, soluble reactive phosphorus (PO_4-P), and nitrate nitrogen (NO_3-N) from the Lake Kasumigaura Database (https://db.cger.nies.go.jp/gem/moni-e/inter/GEMS/database/kasumi/index.html). We also included 30-day moving average of wind speed and precipitation, which were collected from the Tsukuba-Tateno Meteorological Station of the Japan Meteorological Agency (https://www.jma.go.jp/jma/menu/menureport.html). CCM requires taking consecutive time lags of observed variables, and therefore we analyzed data from April 1996 to March 2019 (276 months). There

were no missing data for any variables during this time interval. All data were square-root-transformed and the time series was normalized to have a mean 0 and variance of 1 to adjust for rapid increases in some phytoplankton species prior to analysis.

- Data availability

We used Mathematica 10.2 and 11.0 for our analysis. The computer codes used for the analysis can be downloaded from: xxx

(Before publication, we will register the files in a public repository so that it can be assigned a DOI.)

Acknowledgments

We thank all members of the Lake Kasumigaura long-term monitoring group of the National Institute for Environmental Studies. We especially thank Shinji Nakaoka for valuable discussions, and Megumi Nakagawa for counting phytoplankton and providing helpful comments. This work was supported by funding from the Management Expenses Grant for RIKEN BioResource Research Center, MEXT, and the Japan Society for the Promotion of Science (JSPS) KAKENHI JP20K06820 and JP20H03010 (to K.S.).

Author contributions

K.S. conceptualized the study, developed methodology, performed the formal analysis, and wrote the first draft. S.S.M. contributed to data curation and analysis, and wrote the paper with K.S.. All authors critically reviewed and edited the paper.

Competing interests

We declare that we have no conflict of interest.

References

1. Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28(7), 075310.
2. Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., ... & Zscheischler, J. (2019a). Inferring causation from time series in Earth system sciences. Nature communications, 10(1), 1-13.
3. Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society, 424-438.
4. Akoglu, H. (2018). User's guide to correlation coefficients. Turkish journal of emergency medicine, 18(3), 91-93.
5. Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. Science, 338(6106), 496-500.
6. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019b). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11),

eaau4996.
7. Durden, J. M., Luo, J. Y., Alexander, H., Flanagan, A. M., & Grossmann, L. (2017). Integrating "big data" into aquatic ecology: Challenges and opportunities. Limnology and Oceanography Bulletin, 26(4), 101-108.
8. Vallis, Geoffrey K. Atmospheric and oceanic fluid dynamics. Cambridge University Press, 2017.
9. Kretschmer, M., Coumou, D., Donges, J. F. & Runge, J. Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation. J. Clim. 29, 4069–4081 (2016).
10. Perretti, C. T., Munch, S. B., & Sugihara, G. (2013). Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. Proceedings of the National Academy of Sciences, 110(13), 5253-5257.
11. Deyle, E. R., May, R. M., Munch, S. B., & Sugihara, G. (2016a). Tracking and forecasting ecosystem interactions in real time. Proceedings of the Royal Society B: Biological Sciences, 283(1822), 20152258.
12. McGowan, J. A., Deyle, E. R., Ye, H., Carter, M. L., Perretti, C. T., Seger, K. D., ... & Sugihara, G. (2017). Predicting coastal algal blooms in southern California. Ecology, 98(5), 1419-1433.
13. Schlesinger, W. H., Cole, J. J., Finzi, A. C., & Holland, E. A. (2011). Introduction to coupled biogeochemical cycles. Frontiers in Ecology and the Environment, 9(1), 5-8.
14. Tsonis, A. A., Deyle, E. R., May, R. M., Sugihara, G., Swanson, K., Verbeten, J. D., & Wang, G. (2015). Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. Proceedings of the National Academy of Sciences, 112(11), 3253-3256.
15. Deyle, E. R., Maher, M. C., Hernandez, R. D., Basu, S., & Sugihara, G. (2016b). Global environmental drivers of influenza. Proceedings of the National Academy of Sciences, 113(46), 13081-13086.
16. Rogers, T. L., & Munch, S. B. (2020). Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. Proceedings of the National Academy of Sciences, 117(1), 479-485.
17. Baskerville, E. B., & Cobey, S. (2017). Does influenza drive absolute humidity? Proceedings of the National Academy of Sciences, 114(12), E2270-E2271.
18. Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. Bonn, Germany: German National Research Center for Information Technology GMD Technical Report, 148(34), 13.
19. Jaeger, H. (2002). Adaptive nonlinear system identification with echo state networks. Advances in neural information processing systems, 15, 609-616.
20. Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science, 304(5667), 78-80.
21. Bollt, E. (2021). On explaining the surprising success of reservoir computing forecaster of chaos? The universal machine learning dynamical system with contrast to VAR and DMD. Chaos: An Interdisciplinary Journal of Nonlinear Science, 31(1), 013108.
22. Pennekamp, F., Iles, A. C., Garland, J., Brennan, G., Brose, U., Gaedke, U., ... & Petchey, O. L.

(2019). The intrinsic predictability of ecological time series and its potential to guide forecasting. Ecological Monographs, 89(2), e01359.

23. Ye, H., Deyle, E. R., Gilarranz, L. J., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. Scientific reports, 5(1), 1-9.
24. Matsuzaki, S. I. S., Suzuki, K., Kadoya, T., Nakagawa, M., & Takamura, N. (2018). Bottom-up linkages between primary production, zooplankton, and fish in a shallow, hypereutrophic lake. Ecology, 99(9), 2025-2036.
25. Ushio, M., Hsieh, C. H., Masuda, R., Deyle, E. R., Ye, H., Chang, C. W., ... & Kondoh, M. (2018). Fluctuating interaction network and time-varying stability of a natural fish community. Nature, 554(7692), 360-363.
26. Chang, C. W., Ye, H., Miki, T., Deyle, E. R., Souissi, S., Anneville, O., ... & Sugihara, G. (2020). Long - term warming destabilizes aquatic ecosystems through weakening biodiversity - mediated causal networks. Global Change Biology, 26(11), 6413-6423.
27. Carey, C. C., Woelmer, W. M., Lofton, M. E., Figueiredo, R. J., Bookout, B. J., Corrigan, R. S., ... & Thomas, R. Q. (2021). Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting. Inland Waters, 1-14.
28. Nakajima, K., & Fischer, I. (2021). Reservoir Computing. Springer, Singapore.
29. Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, No. 10). New York: Springer series in statistics.
30. Benincà, E., Huisman, J., Heerkloss, R., Jöhnk, K. D., Branco, P., Van Nes, E. H., ... & Ellner, S. P. (2008). Chaos in a long-term experiment with a plankton community. Nature, 451(7180), 822-825.
31. Takens, F. (1981). Detecting strange attractors in turbulence. In Dynamical systems and turbulence, Warwick 1980 (pp. 366-381). Springer, Berlin, Heidelberg.
32. Huang, Y., Fu, Z., & Franzke, C. L. (2020). Detecting causality from time series in a machine learning framework. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(6), 063116.
33. Wang, M., & Fu, Z. (2022). A new method of nonlinear causality detection: Reservoir computing Granger causality. Chaos, Solitons & Fractals, 154, 111675.
34. Duggento, A., Guerrisi, M., & Toschi, N. (2021). Echo state network models for nonlinear granger causality. Philosophical Transactions of the Royal Society A, 379(2212), 20200256.
35. Suzuki, K., Yoshida, K., Nakanishi, Y., & Fukuda, S. (2017). An equation - free method reveals the ecological interaction networks within complex microbial ecosystems. Methods in Ecology and Evolution, 8(12), 1774-1785.
36. Cenci, S., Sugihara, G., & Saavedra, S. (2019). Regularized S-map for inference and forecasting with noisy ecological time series. Methods in Ecology and Evolution, 10(5), 650-660.
37. Chang, C. W., Miki, T., Ushio, M., Ke, P. J., Lu, H. P., Shiah, F. K., & Hsieh, C. H. (2021). Reconstructing large interaction networks from empirical time series data. Ecology Letters, 24(12), 2763-2774.
38. Cobey, S., & Baskerville, E. B. (2016). Limits to causal inference with state-space reconstruction for

infectious disease. PloS one, 11(12), e0169050.
39. Mønster, D., Fusaroli, R., Tylén, K., Roepstorff, A., & Sherson, J. F. (2017). Causal inference from noisy time-series data—testing the convergent cross-mapping algorithm in the presence of noise and external influence. Future Generation Computer Systems, 73, 52-62.
40. Leng, S., Ma, H., Kurths, J., Lai, Y. C., Lin, W., Aihara, K., & Chen, L. (2020). Partial cross mapping eliminates indirect causal influences. Nature communications, 11(1), 1-9.
41. Tanentzap, A. J., Morabito, G., Volta, P., Rogora, M., Yan, N. D., & Manca, M. (2020). Climate warming restructures an aquatic food web over 28 years. Global Change Biology, 26(12), 6852-6866.
42. Kitayama, K., Ushio, M., & Aiba, S. I. (2021). Temperature is a dominant driver of distinct annual seasonality of leaf litter production of equatorial tropical rain forests. Journal of Ecology, 109(2), 727-736.
43. Huber, V., Wagner, C., Gerten, D., & Adrian, R. (2012). To bloom or not to bloom: contrasting responses of cyanobacteria to recent heat waves explained by critical thresholds of abiotic drivers. Oecologia, 169(1), 245-256.
44. Freeman, E. C., Creed, I. F., Jones, B., & Bergström, A. K. (2020). Global changes may be promoting a rise in select cyanobacteria in nutrient - poor northern lakes. Global Change Biology, 26(9), 4966-4987.
45. Takamura, N., Otsuki, A., Aizaki, M., & Nojiri, Y. (1992). Phytoplankton species shift accompanied by transition from nitrogen dependence to phosphorus dependence of primary production in Lake Kasumigaura, Japan. Archiv für Hydrobiologie, 129-148.
46. Tomioka, N., Imai, A., & Komatsu, K. (2011). Effect of light availability on Microcystis aeruginosa blooms in shallow hypereutrophic Lake Kasumigaura. Journal of plankton research, 33(8), 1263-1273.
47. Fukushima, T., & Arai, H. (2015). Regime shifts observed in Lake Kasumigaura, a large shallow lake in Japan: Analysis of a 40 - year limnological record. Lakes & Reservoirs: Research & Management, 20(1), 54-68.
48. Kâ, S., Mendoza-Vera, J. M., Bouvy, M., Champalbert, G., N'Gom-Kâ, R., & Pagano, M. (2012). Can tropical freshwater zooplankton graze efficiently on cyanobacteria?. Hydrobiologia, 679(1), 119-138.
49. Ger, K. A., Urrutia-Cordero, P., Frost, P. C., Hansson, L. A., Sarnelle, O., Wilson, A. E., & Lürling, M. (2016). The interaction between cyanobacteria and zooplankton in a more eutrophic world. Harmful algae, 54, 128-144.
50. Moisander, P. H., Paerl, H. W., & Zehr, J. P. (2008). Effects of inorganic nitrogen on taxa - specific cyanobacterial growth and nifH expression in a subtropical estuary. Limnology and Oceanography, 53(6), 2519-2532.
51. Jochimsen, M. C., Kümmerlin, R., & Straile, D. (2013). Compensatory dynamics and the stability of phytoplankton biomass during four decades of eutrophication and oligotrophication. Ecology letters, 16(1), 81-89.
52. Huisman, J., Sharples, J., Stroom, J. M., Visser, P. M., Kardinaal, W. E. A., Verspagen, J. M., &

Sommeijer, B. (2004). Changes in turbulent mixing shift competition for light between phytoplankton species. Ecology, 85(11), 2960-2970.

53. Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M., & Visser, P. M. (2018). Cyanobacterial blooms. Nature Reviews Microbiology, 16(8), 471-483.
54. Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. Nature methods, 7(5), 335-336.
55. Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-generation global biomonitoring: large-scale, automated reconstruction of ecological networks. Trends in Ecology & Evolution, 32(7), 477-487.
56. Balint, M., Pfenninger, M., Grossart, H. P., Taberlet, P., Vellend, M., Leibold, M. A., ... & Bowler, D. (2018). Environmental DNA time series in ecology. Trends in Ecology & Evolution, 33(12), 945-957.
57. Haase, P., Tonkin, J. D., Stoll, S., Burkhard, B., Frenzel, M., Geijzendorffer, I. R., ... & Schmeller, D. S. (2018). The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity. Science of the Total Environment, 613, 1376-1384.
58. Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For biodiversity research and monitoring. Oxford University Press.
59. Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., ... & Porter, J. H. (2013). Big data and the future of ecology. Frontiers in Ecology and the Environment, 11(3), 156-162.
60. Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., ... & Loreau, M. (2015). Predictive ecology in a changing world. Journal of Applied Ecology, 52(5), 1293-1310.
61. Scheffer, M., & Carpenter, S. R. (2003). Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in ecology & evolution, 18(12), 648-656.
62. Chivian, E., & Bernstein, A. (Eds.). (2008). Sustaining life: how human health depends on biodiversity. Oxford University Press.
63. J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott (2017) "Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data," Chaos 27(12), 121102.
64. C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, and W. D. Lu, (2017) "Reservoir computing using dynamic memristors for temporal information processing," Nat. Commun. 8(1), 2204.
65. Z. Lu, J. Pathak, B. Hunt, M. Girvan, R. Brockett, and E. Ott, (2017) "Reservoir observers: Model-free inference of unmeasured variables in chaotic systems," Chaos 27(4), 041102.
66. Z. Lu, B. R. Hunt, and E. Ott, (2018) "Attractor reconstruction by machine learning," Chaos 28(6), 061104.
67. J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, (2018) "Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach," Phys. Rev. Lett. 120(2), 024102.
68. A. Chattopadhyay, P. Hassanzadeh, K. Palem, and D. Subramanian, (2019) "Data-driven prediction

of a multi-scale Lorenz 96 chaotic system using a hierarchy of deep learning methods reservoir computing, ANN, and RNN-LSTM," arXiv:1906.08829.

69. M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, and N. Carvalhais, (2019) "Deep learning and process understanding for data-driven earth system science," Nature 566(7743), 195.
70. Farhang-Boroujeny, B. (2013). Adaptive filters: theory and applications. John Wiley & Sons.
71. Ifeachor, E. C., & Jervis, B. W. (2002). Digital signal processing: a practical approach. Pearson Education.
72. Lukoševičius, M. (2012). A practical guide to applying echo state networks. In Neural networks: Tricks of the trade (pp. 659-686). Springer, Berlin, Heidelberg.
73. Spearman, C. (1904). The proof and measurement of association between two things. American journal of psychology, 15(1), 72–101.
74. Weiss, S., Van Treuren, W., Lozupone, C., Faust, K., Friedman, J., Deng, Y., ... & Knight, R. et al. (2016). Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. The ISME journal, 10(7), 1669-1681.
75. Clark, A. T., Ye, H., Isbell, F., Deyle, E. R., Cowles, J., Tilman, G. D., & Sugihara, G. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. Ecology, 96(5), 1174-1181.
76. Fisher, C. K., & Mehta, P. (2014). Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. PloS one, 9(7), e102451.
77. Benincà, E., Jöhnk, K. D., Heerkloss, R., & Huisman, J. (2009). Coupled predator–prey oscillations in a chaotic food web. Ecology letters, 12(12), 1367-1378.
78. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., ... & Gerber, G. K. (2016). MDSINE: Microbial Dynamical Systems Inference Engine for microbiome time-series analyses. Genome biology, 17(1), 1-17.
79. Cohen, J. E., & Newman, C. M. (1985). A stochastic theory of community food webs I. Models and aggregated data. Proceedings of the Royal society of London. Series B. Biological sciences, 224(1237), 421-448.
80. Takamura, N., & Nakagawa, M. (2012) Phytoplankton species abundance in Lake Kasumigaura (Japan) monitored monthly or biweekly since 1978. Ecological research, 27(5), 837-837.
81. Takamura, N., Nakagawa, M., & Hanazato, T. (2017) Zooplankton abundance in the pelagic region of Lake Kasumigaura (Japan): monthly data since 1980. Ecological Research, 32(1), 1-1.

Supporting figures

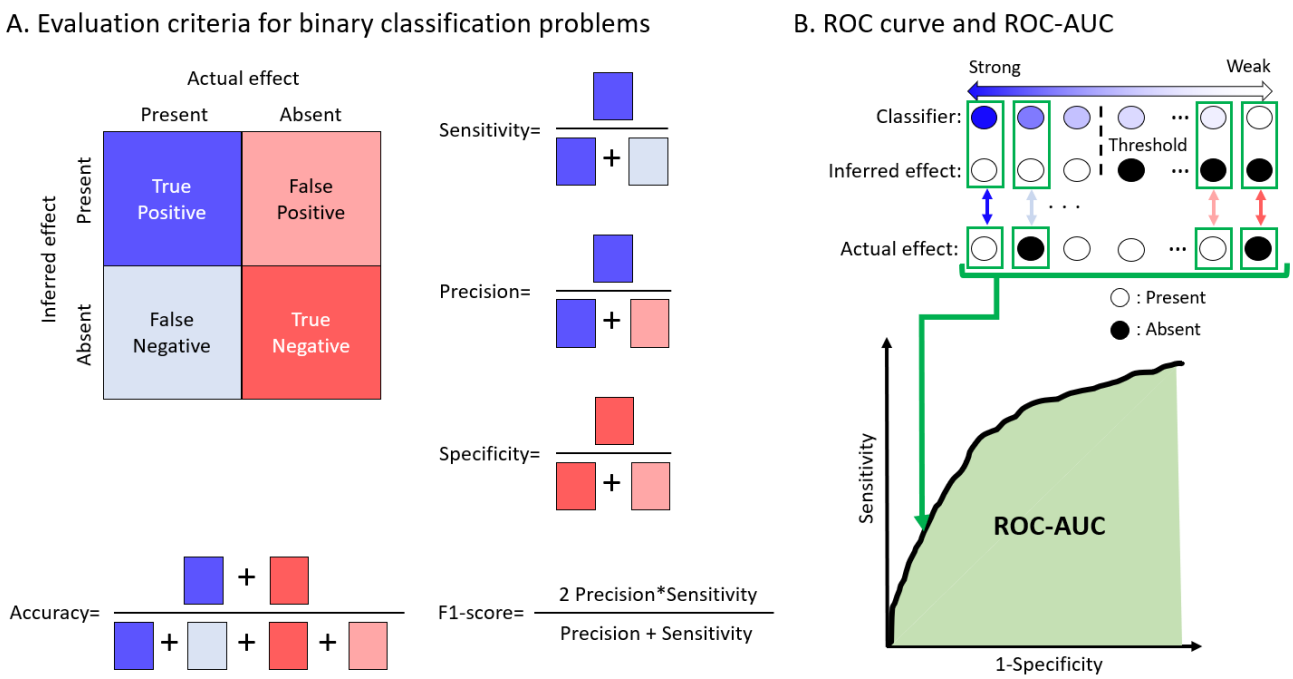
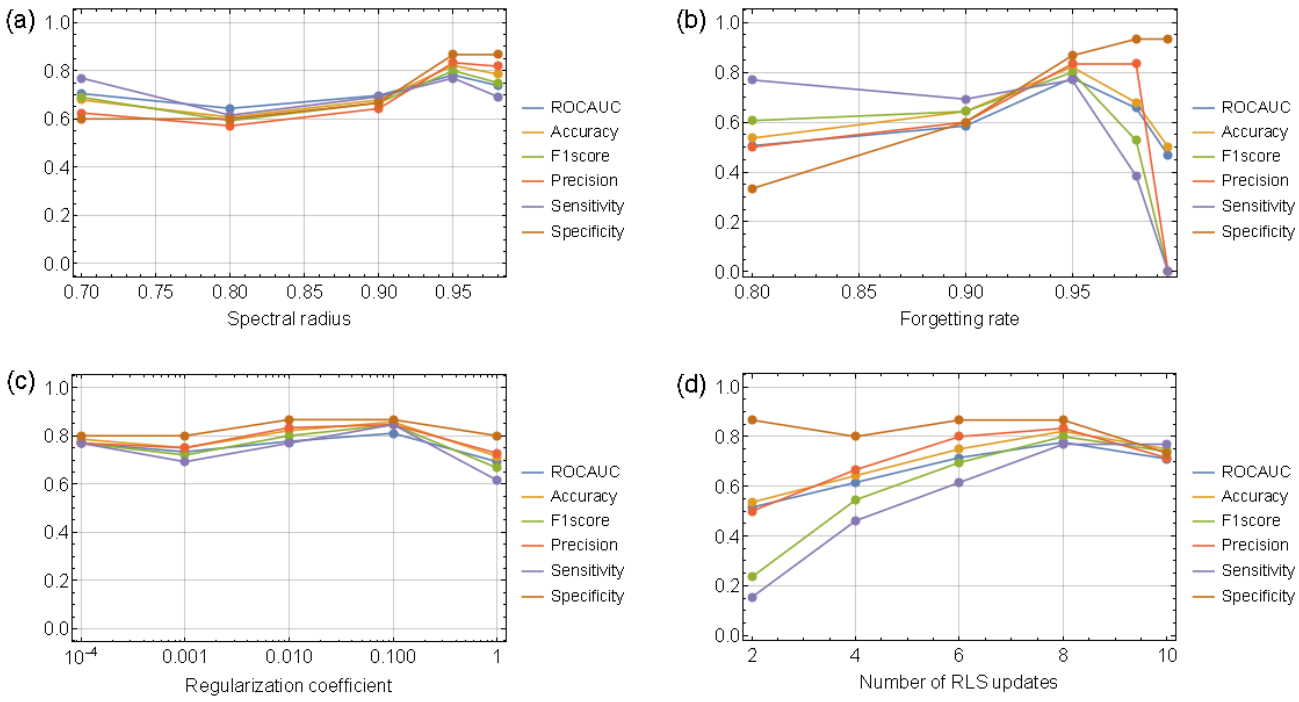
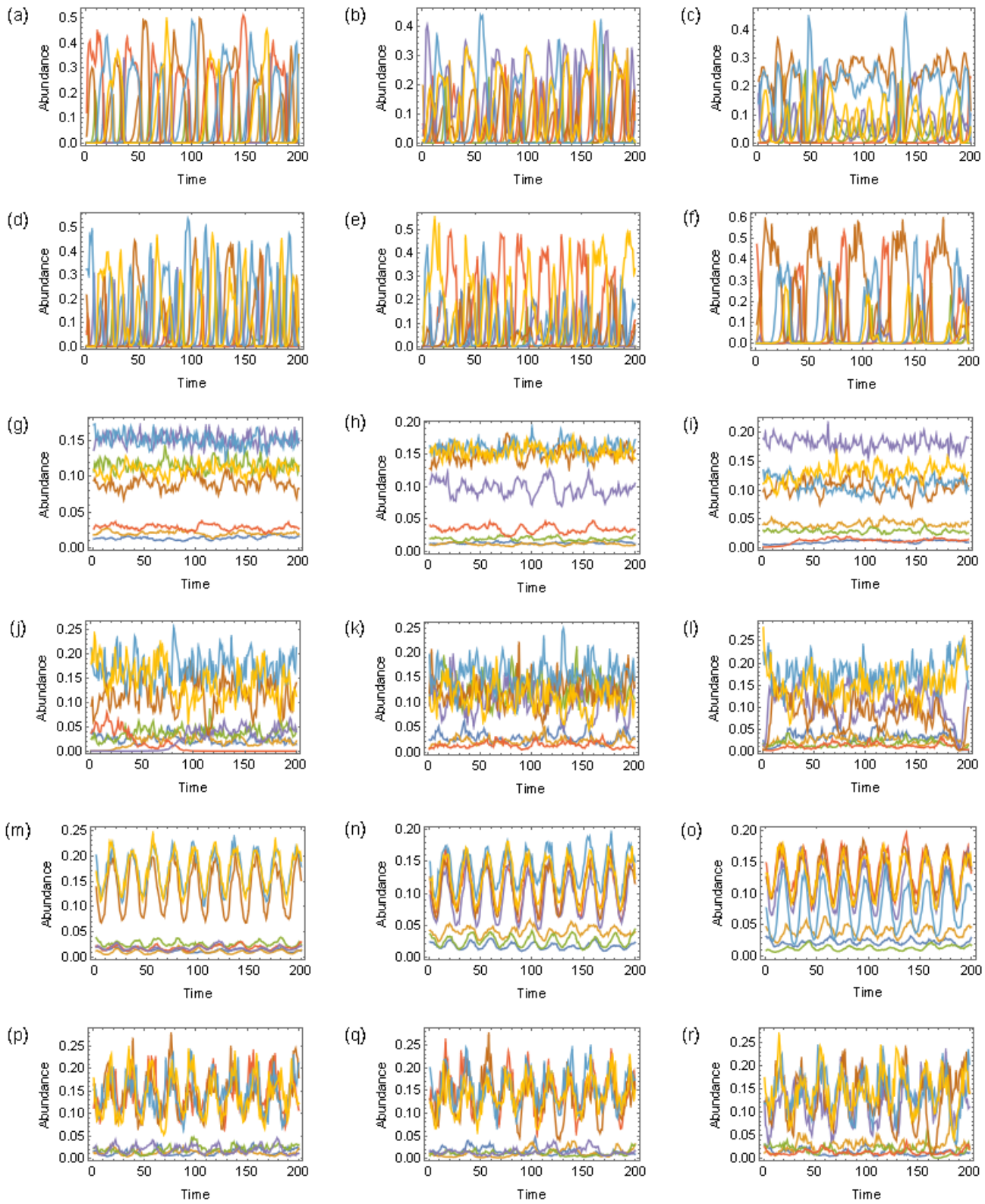


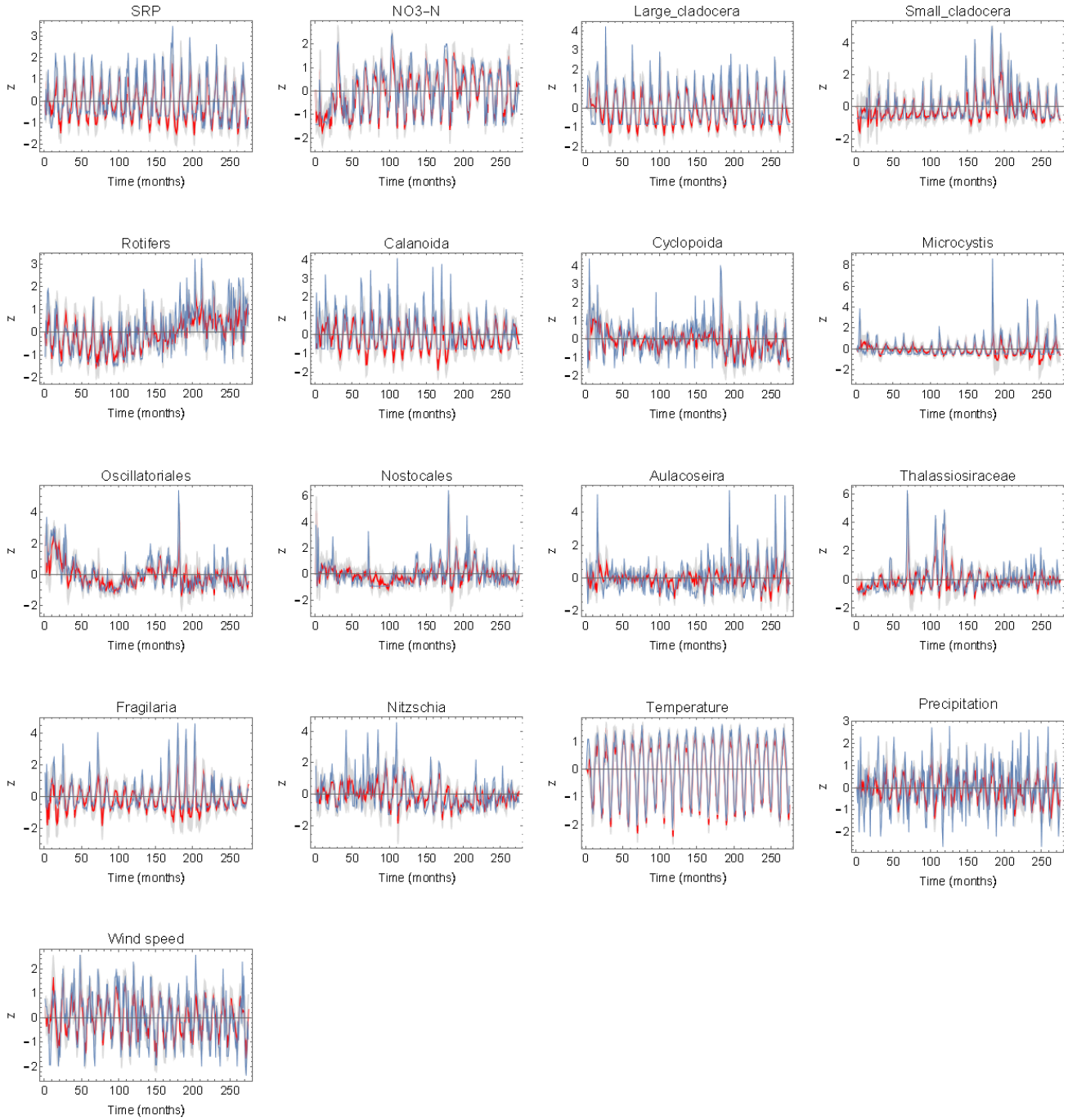
Figure S1. Evaluation criteria for binary classification problems and their relationship to ROC curves. A. evaluation criteria for binary classification problems. Here, the ability of identifying interaction links between species (represented by interacting pairs in EcohNet, CCM and Spearman, or directional effects in LIMITS) is evaluated. True positive (TP) represents a result of inference that correctly indicates the presence of links, true negative (TN) represents a result of inference that correctly indicates the absence of links, false positive (FP) represents a result of inference which incorrectly indicates presence of links, and false negative (FN) is a result of inference which incorrectly indicates the absence of links. B. Explanation of ROC curve. Here, a circle corresponds to an interaction between species. Suppose we set a threshold value for a classifier and determine that links between species above the threshold value are present, and links below the threshold value are absent. By matching this with the actual presence/absence status of interaction links, we can calculate sensitivity and specificity and plot a single point on the 2D plane. The ROC curve is obtained by repeating the calculation with changing threshold values, and the ROC-AUC is the area underneath the curve.



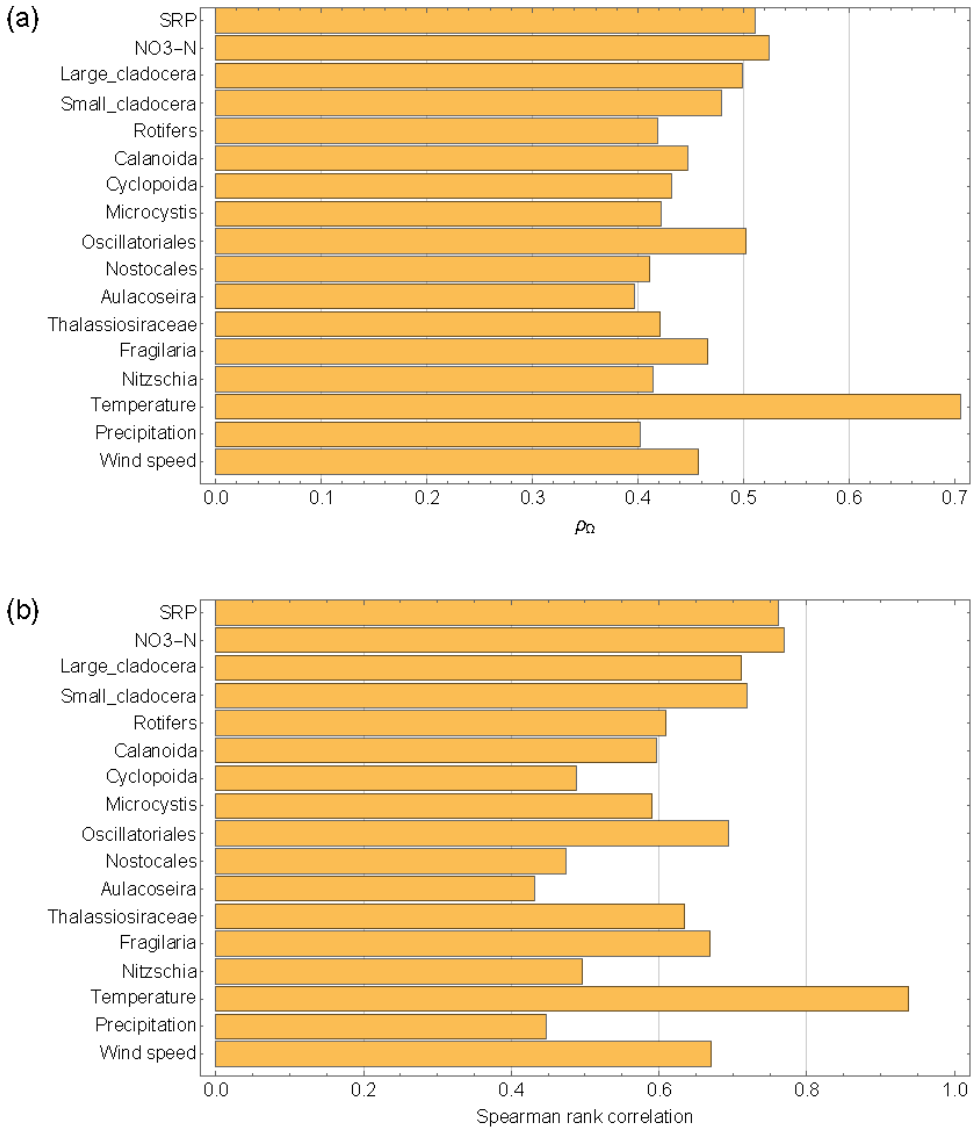
Supporting figure S2. Sensitivity of evaluation criteria for the ESN-RLS parameters. (a) spectral radius (ρ_0), (b) Forgetting factor (λ), (c) Regularization coefficient (δ), and (d) Number of RLS updates (τ). As ρ_0 becomes smaller than 1, the ability of ESN to track a time series becomes lower and it cannot adequately deal with complex time series. On the other hand, when ρ_0 becomes larger than 1, the behavior of ESN becomes unstable since $\rho_0 < 1$ ensures echo state property in most situations [Lukoševičius 2012]. Thus, it is recommended that ρ_0 be set to a value close to 1, but smaller than 1. λ determines how fast the effects of past errors decay. The closer the value is to 1, the slower the decay, and at 1, all past information is treated with the same weight. It is also recommended that λ be set to a value smaller than 1 and close to 1 [Ifeachor and Jervis 2002]. δ controls the strength of the normalization in the regression and it is recommended to set $\delta \ll 1$ [Jaeger and Haas 2004] although the results here do not vary substantially if $\delta \leq 0.1$. Since τ affects how well the ESN can respond to changes in the time series in a short time [Ifeachor and Jervis 2002], the evaluation criterion (except for specificity) was higher when there were more repetitions. Overall, our results suggested a positive relationship between the predictive performance of ESNs and the performance of network inference. This provides a criterion to judge whether the parameters are set appropriately. However, there is little concern about parameter tuning for ESN-RLS. The evaluation criteria did not vary significantly for small differences in parameters (except when λ approached 1; in this case, since the decay time is scaled by $(1-\lambda)^{-1}$, the closer λ is to 1, smaller differences in λ result in larger differences in the decay time). In fact, even without changing the parameters chosen here ($(\rho_0, \lambda, \delta, \tau) = (0.95, 0.95,\ 0.001, 8)$), EcohNet outperformed the other methods in the evaluation on simulated data.



Supporting figure S3. Representative time series of six different models in simulation dataset. (a-c) O_S, (d-f) O_L, (g-i) E_S, (j-l) E_L, (m-o) EO_S, and (p-r) EO_L.



Supporting figure S4. Results of prediction of each variables. Blue lines represent the actual time series and red lines represent the median of predicted time series with the 95% CI shown by grey areas (drawn by 10,000 time series predicted by EcohNet with optimal variables).



Supporting figure S5. Performance of forecasting by EcohNet with the optimized variables. (a) Prediction skill of the median of forecasting by EcohNet. (b) Spearman rank correlation between the actual time series and the median of forecasting.