Modeling and prediction of phase shifts in noisy two-cycle oscillations

Vahini Reddy Nareddy[1], Jonathan Machta[1,2], Karen Abbott[3], Shadisadat Esmaeili[4] and Alan Hastings[2,4*]

[1]Department of Physics, University of Massachusetts, Amherst, Massachusetts 01003 USA.
[2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.
[3]Department of Biology, Case Western Reserve University, 10900 Euclid Ave, Cleveland, OH 44106, USA.
[4]Department of Environmental Science and Policy, One Shields Avenue, University of California, Davis, CA 95616, USA.

*Corresponding author(s). E-mail(s): amhastings@ucdavis.edu;
Contributing authors: nareddyvahini@gmail.com;
machta@physics.umass.edu; kcabbott@case.edu;
sesmaeil@fredhutch.org;

Abstract

Understanding and predicting ecological dynamics in the presence of noise remains a substantial and important challenge. This is particularly true in light of the poor quality of much ecological data and the imprecision of many ecological models. As a first approach to this problem, we focus here on a simple system expressed as a discrete time model with 2-cycle behavior, reflecting alternating high and low population sizes. Such dynamics naturally arise in ecological systems with overcompensatory density dependence. We ask how the amount of detail included in the population estimates affects the ability to forecast the likelihood of changes in the phase of oscillation, meaning whether high populations occur in odd or in even years. We adjust the level of detail by converting continuous population levels to simple, coarse-grained descriptions using two state and four state models. We also consider a cubic noisy overcompensatory model with three

parameters. The focus on phase changes is what distinguishes the question we are asking and the methods we use from more standard time series approaches. Obviously, adding observation states improves the ability to forecast phase shifts. In particular, the four state model and cubic model outperform the two state model because they include a transition state, through which the dynamics typically pass during a phase change. Nonetheless, at high noise levels the improvement in forecast skill is relatively modest. Additionally, the frequency of phase changes depends strongly on the noise level, and is much less affected by the parameter determining amplitude in the population model, so phase shift frequencies could possibly be used to infer noise levels.

Keywords: discrete-state models, phase change, two-cycle variable, alternate bearing, mutual information, forecast skill

1 Introduction

One of the great challenges in understanding complex systems is modeling and prediction in the absence of a clear understanding of underlying mechanism. Although the full behavior of a complex system is typically high dimensional and very challenging to model, there are often relatively regular features of the dynamics that are amenable to study. For example, many complex systems exhibit relatively regular oscillations that appear as an emergent phenomenon and may or may not be central to the functioning of the system. Examples in animal physiology include the beating of the heart or brain waves. In ecology, examples include masting, insect outbreaks and predator-prey oscillations. In all these cases, the oscillation is at least somewhat unpredictable both due to environmental influences and because it emerges from complex interactions which, by themselves, may not produce regular oscillations.

The simplest such oscillation is a noisy, discrete time two-cycle, a ubiquitous feature of ecological time series (May, 1976). In ecology, discreteness in time is often the product of seasonality (White and Hastings, 2020), and two-cycles naturally result from over-compensatory dynamics (meaning that when the population in year t is large, the population in year $t+1$ can actually get smaller rather than just increasing more slowly). Models in the class of quadratic maps with period doubling and noise, such as the noisy logistic and Ricker maps, display such oscillations for some ranges of parameters. These models are extremely important in dynamical systems theory and ecology. However, each such model makes assumptions about the dynamical mechanisms yielding oscillations. In this paper we take a different approach that is, as far a possible, agnostic about dynamical mechanism and seeks to build models directly from time series data with as few assumptions as possible.

Understanding the dynamics of a system is crucial for predicting future states and guiding decisions. Building a dynamical model and forecasting the

probable events involves using time series data and analysis. Time series analysis is widely used in many fields; however, ecology poses unique challenges as the accuracy and precision of obtained data is usually poor due to practical limits on how often and how completely natural populations can be censused, and how long it takes for some ecological dynamics (e.g. multi-year or even multi-decade cycles) to be captured in data (Bence, 1995; Ives et al, 2010). In addition, the ubiquitous influence of stochasticity on ecological dynamics, and complex feedbacks between stochastic and deterministic ecological processes, makes inference from ecological time series data particularly challenging (Sugihara et al, 1990; Turchin and Taylor, 1992; Ives et al, 2003).

Time series analysis has been employed to answer a wide range of questions in ecology (Powell and Steele, 1995; Auger-Méthé et al, 2021) and other disciplines, such as epidemiology, meteorology, and economics, in which temporal dynamics are of interest. For instance, time series data have been analyzed to identify populations of conservation concern (Fagan, 2001), to estimate the strength or form of density dependent regulation (Sibly et al, 2007), and for inferring demographic rates (Gross et al, 2005). Methods for analyzing time series can be crudely separated into linear approaches like ARMA models and nonlinear approaches based on state space models (see the superb review in Auger-Méthé et al, 2021). State space models are hierarchical models based on the idea of a potentially noisy observation process of a noisy underlying dynamic, with the overall approach geared to making inferences about the underlying process from the observations, and hence allowing prediction of the future.

Which of these methods is best depends on the question being asked. For example, the dimension of an ecological community (i.e. the number of interacting species or populations) can be estimated from the number of time lags in the best-fit ARMA model; precise estimates of the ARMA's parameters are not needed (Abbott et al, 2009). In contrast, ecological forecasting requires both a model that is well supported by data and good parameter estimates (Dietze, 2017). Questions about density dependence or interaction rates require a model that not only captures the statistical patterns in the time series data but also models the biological mechanisms responsible for those patterns (Turchin and Taylor, 1992). Features of the data also influence which approach is best: because measurement error and process error propagate differently through time, time series with different types and strengths of error must be treated differently (Ives et al, 2003).

In this paper, we emphasize a different, less explored question. A simple qualitative feature is the phase of the oscillation, which for a two-cycle is whether the high value occurs at an even or odd time step. A change in phase is a feature that is easy to detect even with very crude data. Thus, here we explore how much information, and what types of information, from noisy time series data should be modeled for forecasting phase shifts in ecological oscillators.

While it is appealing to describe real world data with a specific model that incorporates mechanistic dynamics, it is often the case that the detailed mechanisms are either unknown or the data too limited to decide between competing models. In this situation, it is preferable to introduce the simplest possible model to capture the salient features of the data without imputing additional features of a system that are not observed. When large observation error precludes precise estimates of population size, the data may nonetheless provide reliable information about the phase of oscillation and the probability of phase shifts, if it can reliably distinguish between a small number of discrete states. In the language of state space models, we use a very crude observation process to coarse grain the system into two, three, or four possible states at each time. We then consider several simple predictive models that are easily inferred from time series data using this approach.

Discrete time two-cycle oscillations in a single real variable can be described by an amplitude and a phase of oscillation, where the phase takes two possible values (high population size at even or odd times). Noise causes both the amplitude and phase variables to fluctuate in time as can be seen in Fig. 1A, which is a time series generated from a noisy Ricker map. If, as is often the case, the intrinsic oscillation is stable and the noise is uncorrelated or weakly correlated in time, then the amplitude variable fluctuates on short time scales. On the other hand, for weak to moderate noise strength, phase changes are rare and difficult to predict. Changes in the phase of oscillation of a system can have a dramatic effect on both the system and its environment. The data representations and dynamical models that we discuss are primarily designed to model and predict phase changes. Conversely, we also determine inferences about underlying dynamics that can be made from the frequency of such phase shifts.

One particular example of phase shifts of this kind occurs for nut production of individual trees in a pistachio orchard (Lyles et al, 2009; Rosenstock et al, 2011). These individual trees typically alternate between years with high and low pistachio yield, but not perfectly. Thus there is an oscillator, but phase shifts do occur as stochasticity plays a large role (Lyles et al, 2009). The phenomenon of alternate, but not perfectly alternate, bearing in cultivated trees is very common and widespread (Monselise and Goldschmidt, 1982). This alternate bearing behavior is not restricted to domesticated trees, but also occurs in oaks (Crawley and Long, 1995) and others, where the generated pulse in resources can be important to many species so phase shifts have a large impact. Although the actual amount of production is best described by a continuous variable, the presence of definite highs and lows suggests that even discretized data may contain useful information for predicting changes in phase.

The first step in constructing models of a continuous state two-cycle time series is to transform the data by taking a first difference of states through time multiplied by an alternating sign. This transformation is illustrated in Fig. 1A, which shows the raw time series, and Fig. 1B, which shows the alternating sign first difference. The sign of the resulting "two-cycle variable" reflects the phase

of oscillation while its absolute value is the amplitude. As is seen in Fig. 1B, this transformation makes clear the bistability of the system and shows that noisy two-cycle systems are yet another example of the broad class of bistable systems, such as those produced by the Allee effect in ecology (Courchamp et al, 1999). Changes in phase are barrier crossing events and are typically accompanied by excursions through a transition state. Here the transition state is identified with small values of the amplitude.

Most of the models we discuss are discrete state Markov processes. After transforming the data as just described, the next step in constructing such a model is to bin the continuous-state two-cycle variable data to create a data set with a discrete number of states. We propose that the bins should defined so as to maximize the mutual information between the full time series data and the discrete state representation. We therefore use bins that, at minimum, capture the phase of oscillation (i.e., two bins that correspond to population peaks at odd or even time steps). With three or more bins, we can also capture information about whether the population is in a transition state. Using mutual information to define basins of attraction for the two metastable phases of oscillation and a transition state may be more broadly applied to modeling other classes of bistable systems.

As this paper is quite complex, we provide a brief outline. In section 2, we describe how we bin the data to reduce continuous-state data to a small number of discrete states. The purpose of this binning procedure is to consider how much we can predict from crude observations containing differing amounts of information about the underlying process. Section 3 provides a description of the different models we use to describe dynamics through time with a small number of discrete states. Section 4 presents the two ways we measure how well the reduced models describe the underlying dynamics: forecast skill and mutual information. Section 5 both presents our results and provides a discussion. Our conclusions are summarized in section 6.

2 Data and Representations

2.1 Synthetic data

Our goal is to analyze noisy two-cycle dynamics. We do this using a prototypical over-compensatory ecological model, the noisy Ricker map, to generate the data (example time series in Fig. 1A). Ricker dynamics can describe the intraspecific density dependence and short-period oscillations in populations. The noisy Ricker map is given by,

$$X_{t+1} = X_t \exp\left[r(1 - X_t)\right] \exp\left(\lambda_r \zeta_t\right) \tag{1}$$

where X_t is the Ricker variable at time t representing population size, r is the population growth parameter, λ_r is the noise strength, and ζ_t is normal random variable with zero mean and unit standard deviation. Here, we study the data generated from simulations with growth parameter values $r = 2.1, 2.2$, and 2.3,

all of which fall in the two-cycle regime, and noise values $\lambda_r = 0.11 - 0.2$ in increments of 0.01. Here, the observed data, X is a continuous-state variable with period-2 oscillations around $X = 1$.

The time series of such a two-cycle system has oscillations between a high and a low states, and has two possible phases of oscillation: high state at even times or high state at odd times. We refer to these as "+" and "−" phases. The system has symmetry with respect to these two phases of oscillation. In the absence of noise, the system chooses one of the phases depending on the initial condition of the oscillation and the phase persist indefinitely breaking the symmetry between the phases. In presence of noise, the phase of oscillation changes at random times as shown in Fig. 1A. In the absence of any symmetry breaking environmental force, the system retains a statistical symmetry with respect to the two phases of oscillation. Here, we discuss both continuous and discrete data representations of noisy two-cycle oscillations and focus on the phase of oscillation and its dynamics.

While the analysis presented in this study uses simulated Ricker data with 2×10^5 time steps as the observed data, the methods and models developed are suitable for any noisy two-cycle oscillations. For the rest of the study, we assume that the dynamics of the observed data is unknown, so we use the time series data only and we pretend not to know that they were generated from the noisy Ricker map.

2.2 Continuous state representation: two-cycle variable

The observed two cycle oscillatory data shown in Fig. 1A have varying amplitude and also display changes in the phase of oscillation. Both of these features can be identified in the behavior of the first difference with alternating sign, which we refer to as the two-cycle variable, M_t,

$$M_t = (-1)^{t+1}(X_{t+1} - X_t), \tag{2}$$

as shown in Fig. 1B. The sign of the two-cycle variable is the phase of oscillation and its absolute value is the amplitude of oscillation. The phase of oscillation is displayed in Figs. 1A and B with the positive ("+") phase shaded grey.

First differences of time series are used in many prediction models (Box and Jenkins, 1976). In this study, we include the alternating sign along with the first difference in the two-cycle variable to retrieve the information about the phase of oscillation. This representation also ensures the symmetry of the two-cycle variable around zero in the long-time limit for any two-cycle oscillator without a bias toward one phase.

The two-cycle variable is a continuous-state variable just like observed time series data, but the transformation is not invertible because the initial condition is required to recover the observed data. While this representation results in a loss of information, the two-cycle variable makes the amplitude and phase of oscillation more manifest. Changes in the phase of oscillation are a salient feature of noisy two-cycle oscillators that are a primary interest of this study.

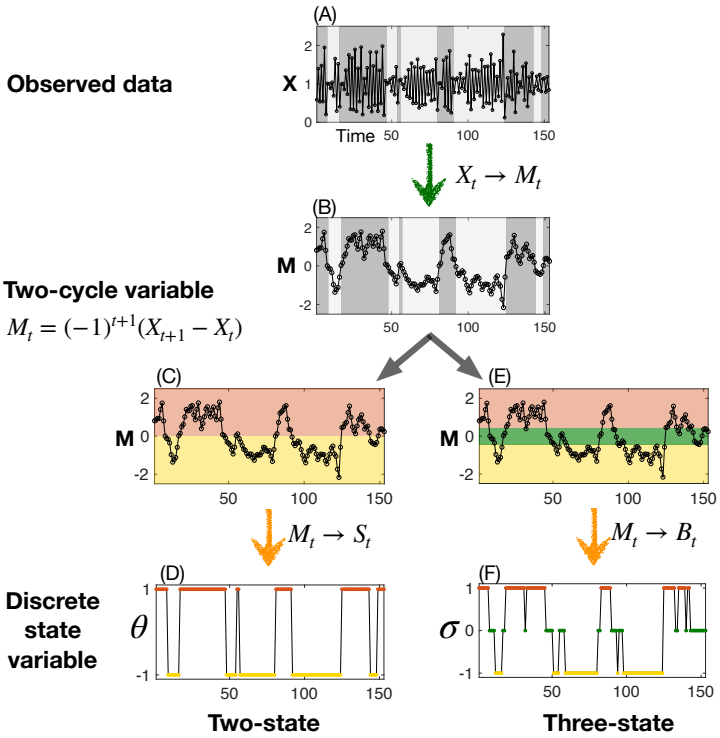

Fig. 1 Representations of noisy two-cycle data, emphasizing the phase of oscillation. (A) Observed time series. The grey highlighted region has a positive phase of oscillation and the white region has a negative phase of oscillation, as determined by which year has the high value of the oscillator. (B) The two-cycle variable, M is obtained from the observed data using Eq. (2). (C) The two-cycle variable is re-plotted highlighting two discrete bins used in the two-state representation. The two-cycle variable is symmetric around zero. (D) The two-state representation, θ, is obtained by assigning all the positive (negative) two-cycle variables as 1 (-1). (E) The two-cycle variable is re-plotted highlighting the three bins used in the three-state representation, with the central green interval representing the transition state. (F) The three-state representation, σ, is obtained by assigning all two-cycle variables in the interval around zero (green) as 0 and the rest of the positive (negative) two-cycle variables as 1 (-1).

2.3 Discrete state representations

Discrete state representations lump the continuous two-cycle variable into a small number of bins. These representations are useful when small fluctuations are not relevant or cannot be observed or the signal-to-noise ratio is low in the observed data. The discrete state representation of the data helps in understanding the essential information required for the problem. As we are interested in studying and predicting the phase changes in two-cycle oscillations, the data can minimally be represented as two states to retain the information about whether the system is in the positive or negative phase of oscillation. Here, we discuss two-state, three-state and four-state variables

where the corresponding bins are chosen to maximize the information about the two-cycle dynamics.

1. Two-state variable: The two-state variable, θ_t is the phase of oscillation and is defined as,

$$\theta_t = \text{sign}(M_t). \tag{3}$$

θ_t takes the values, $+1$ and -1, corresponding to the positive and negative phases of oscillation, respectively. This data representation (see Fig. 1D) loses all the information about the amplitude of oscillation.

2. Three-state variable: The time-series of the two-cycle variable in Fig. 1B clearly shows that the system changes its phase of oscillation when the two-cycle variable is close to zero. Hence, to better predict the phase of two-cycle oscillations, a transition state where the system is likely to change its phase of oscillations is included in the three-state representation. The three-state representation, σ_t is defined as,

$$\sigma_t = \begin{cases} \text{sign}(M_t), & \text{if } |M_t| \geq w \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The transition state ($\sigma_t = 0$) is identified as the interval where the amplitude is smaller than a cut-off w. We use mutual information to set w as described below in Sec. 5.2.

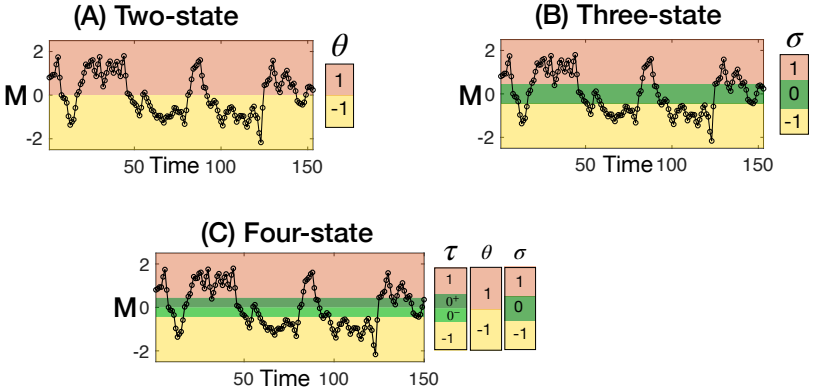

Fig. 2 Illustrations of discrete data representations of the two-cycle oscillator. **(A)** The two-cycle variable M is divided into two bins to separate the two phases of oscillation. **(B)** M is divided into three bins to separate the data into small (transition state around zero) and large amplitude oscillations (two phases of oscillations). The width of the transition state is calculated by maximizing mutual information between M and the three-state representation. **(C)** The transition state in the three-state representation is further divided into two bins symmetric around zero to form four states in total. Here, the data are separated by the two phases of oscillation as well as amplitude of the oscillation (small or large).

3. Four-state variable: Due to the symmetry of the two-cycle variable around zero it is useful to define a four-state variable that divides the transition state into two bins with positive and negative phase, respectively (see Fig. 2C). The four-state variable τ takes values in $\{+1, -1, 0^+, 0^-\}$ and is defined from the three-state variable, σ and the two-state variable, θ,

$$\tau_t = \begin{cases} \sigma_t, & \text{if } \sigma_t = \pm 1 \\ 0^+, & \text{if } \sigma_t = 0 \text{ and } \theta_t = 1 \\ 0^-, & \text{if } \sigma_t = 0 \text{ and } \theta_t = -1. \end{cases} \tag{5}$$

These four states give the information about the phase of oscillation as well as whether the system is in the transition state.

3 Dynamical Models

In this section, we discuss simple dynamical models for analyzing noisy two-cycle time series data and build probabilistic forecasting tools. The data representations studied in Sec. 2 have different states, so each requires a different model to represent its dynamics. All the data representations and corresponding dynamics discussed in this work are shown in Fig. 3. All the models studied have Markovian dynamics. More accurate predictions can potentially be obtained using non-Markovian models with the same data representations at the expense of introducing more model parameters. The trade-off between richer data representations versus non-Markovian models is an interesting topic that deserves further study but is beyond the scope of this work.

3.1 Two-state dynamics

The model of two-state dynamics is designed for a two-state system that takes values $\{\pm 1\}$. The model parameters are transition probabilities. As the model has 2 states, there are 4 transition probabilities: $P_\theta(1 \to 1)$, $P_\theta(1 \to -1)$, $P_\theta(-1 \to 1)$, $P_\theta(-1 \to -1)$ where $P_\theta(a \to b)$ is the transition probability from state $\theta = a$ to $\theta = b$.

The two states of the two-state variable have $\mathbb{Z}_2$ symmetry giving rise to conditions, $P_\theta(1 \to 1) = P_\theta(-1 \to -1)$ and $P_\theta(1 \to -1) = P_\theta(-1 \to 1)$ along with a normalization condition: $P_\theta(1 \to 1) + P_\theta(1 \to -1) = 1$. Without any loss of generality, we can choose the transition probability $P_\theta(1 \to -1)$, which measures the probability of the system to change its phase of oscillation, as the single parameter of the model.

3.2 Three-state dynamics

The model of three-state dynamics is developed for calculating transition probabilities between the three states $\{0, \pm 1\}$. In the study of two-cycle oscillations, states ± 1 correspond to the two phases of oscillation and state 0 corresponds to the transition state between the two phases of oscillation.

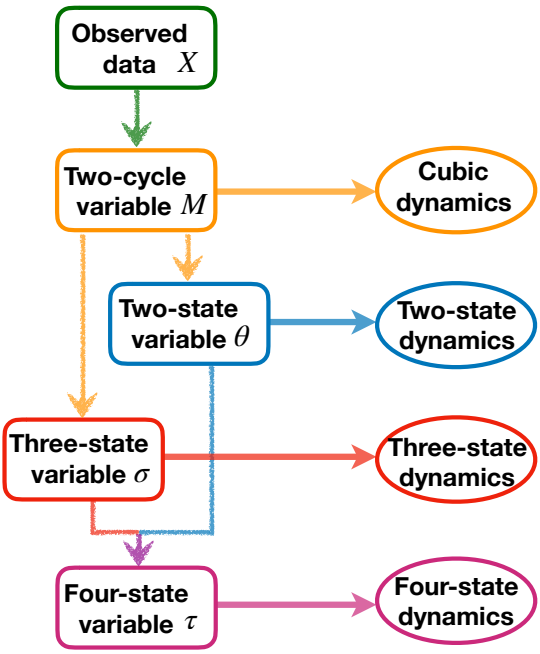

Fig. 3 Representations and dynamical models of noisy two-cycle data. **Left:** Data representations obtained from the observed data are shown in the order discussed in Sec. 2. Two-cycle variable is obtained from the observed data. Two-state and three-state variables are obtained from the two-cycle variable and are used to define the four-state variable. **Right:** Dynamical models developed to study and forecast noisy two-cycle data are discussed in Sec. 3. Data represented as the two-cycle variable are modeled with cubic dynamics. Two-state, three-state and four-state dynamics are developed using transitions between the discrete states of the two-state, three-state and four-states variables respectively.

There are nine transition probabilities between the three states. As the two-cycle variable is symmetric around zero, there are four symmetry conditions and two normalization conditions, leaving three independent probabilities. Our choice of independent transition probabilities that completely specify the dynamics are $P_\sigma(1 \to 1)$, $P_\sigma(1 \to 0)$, $P_\sigma(0 \to 1)$.

3.3 Four-state dynamics

The model of four-state dynamics is used to predict the phase of oscillation even when the system is in the transition state. There are 16 transition probabilities among the four states, $\{\pm 1, 0^+, 0^-\}$. Due to symmetry around zero, there are 8 symmetry relations and 2 normalization conditions, leaving 6 independent transition probabilities, which we choose as, $P_\tau(1 \to 1)$, $P_\tau(1 \to 0^+)$, $P_\tau(1 \to 0^-)$, $P_\tau(0^+ \to 1)$, $P_\tau(0^+ \to 0^+)$, $P_\tau(0^+ \to 0^-)$.

3.4 Cubic model

The dynamics of the two-cycle variable, M, can be developed with a transition kernel $P_M(M \to M')$. However, inferring this kernel from the data would not be tractable. So, we have chosen to represent the kernal by a three parameter cubic model that satisfies the required properties of M. The cubic model is a bistable continuous-state discrete-time model with $\mathbb{Z}_2$ symmetry. In the absence of noise it is given by,

$$M_{t+1} = aM_t - b^2 M_t^3 \tag{6}$$

where a, b are parameters of the model and M_t is the two-cycle variable at time t. The bifurcation diagram as a function of a when $b = 1$ is shown in Fig. 4A with different initial conditions (top figure, positive initial value $M_0 > 0$ and bottom figure, negative initial value $M_0 < 0$). The region of interest for $b = 1$ is $1 < a < 2$ where the system has two stable fixed points, $\pm\sqrt{a-1}$ and an unstable fixed point, 0. With additive noise, the cubic model is written as,

$$M_{t+1} = (aM_t - b^2 M_t^3) + \lambda\zeta_t \tag{7}$$

where λ is the noise strength and ζ_t is a normally distributed random variable with zero mean and unit standard deviation. The histogram of the noisy cubic model is shown in blue in Fig. 4B.

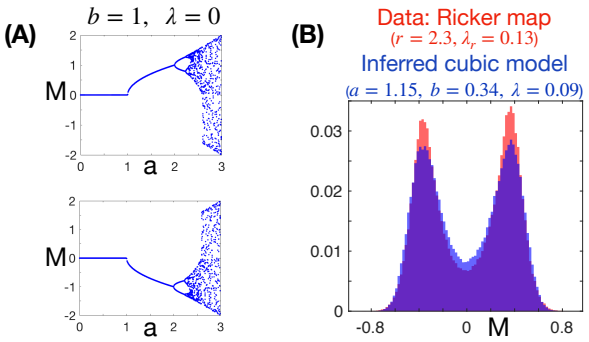

Fig. 4 The cubic model is developed to model the dynamics of the continuous-state two-cycle variable of a noisy two-cycle oscillator. **(A)** The bifurcation diagram of the deterministic cubic model is plotted with positive ($M_{t=0} > 0$, top plot) and negative initial values (bottom plot) as a function of parameter a. **(B)** The histograms of the two-cycle variable of an observed time-series data generated from the Ricker map (red) and the inferred stochastic cubic model (blue). The inferred cubic model captures relevant properties of the two-cycle data.

The parameters of the noisy cubic model, $\{a, b, \lambda\}$, can be inferred from the observed noisy two-cycle data by maximizing the likelihood function. The

likelihood function is,

$$L(a,b,\lambda;\ M_{t+1},M_t) \propto P(M_t \to M_{t+1};\ a,b,\lambda)P(M_t;\ a,b,\lambda) \tag{8}$$

where $P(M_t \to M_{t+1};\ a,b,\lambda)$ is the transition kernel of the cubic model, which can be written as a normal distribution from Eq.(7),

$$P(M_{t+1};\ M_t,a,b,\lambda) = \frac{1}{\lambda\sqrt{2\pi}} \exp\left[\frac{-(aM_t - b^2M_t^3 - M_{t+1})^2}{2\lambda^2}\right], \tag{9}$$

and $P(M_t;\ a,b,\lambda)$ is the stationary probability density, which is the fixed point of the dynamical equation,

$$P(M_{t+1};\ a,b,\lambda) = \int P(M_{t+1};\ M_t,a,b,\lambda)P(M_t;\ a,b,\lambda)dM_t. \tag{10}$$

Using results from Noble et al (2017), the solution for the stationary distribution to fourth order in M is,

$$P(M_t;\ a,b,\lambda) = A(a,b,\lambda)\exp\left[\frac{2(a-1)M_t^2 - b^2M_t^4}{2\lambda^2}\right] \tag{11}$$

where the normalization constant $A(a,b,\lambda)$ is,

$$A(a,b,\lambda) = \frac{2b\exp(-\kappa)}{(\pi\sqrt{a-1})\left[I_{\frac{1}{4}}(\kappa) + I_{-\frac{1}{4}}(\kappa)\right]} \tag{12}$$

with

$$\kappa = \frac{(a-1)^2}{4b^2\lambda^2} \tag{13}$$

and $I_n(\kappa)$ is the modified Bessel function of the first kind.

Using Eq.(8), Eq.(9) and Eq.(11), the likelihood function is written as,

$$L(a,b,\lambda;\ M_{t+1},M_t) = \frac{A(a,b,\lambda)}{\lambda\sqrt{2\pi}} \exp\left[\frac{2(a-1)M_t^2 - b^2M_t^4 - (aM_t - b^2M_t^3 - M_{t+1})^2}{2\lambda^2}\right]. \tag{14}$$

Given pairs of two-cycle variables (M_t, M_{t+1}), the parameters (a,b,λ) can be estimated by maximizing the likelihood function.

The cubic model is inferred from the two-cycle variable from the observed data to obtain the parameters a,b,λ. The histogram of the two-cycle data is compared with the inferred cubic model in Fig. 4B. The cubic model captures the dynamics of the two-cycle observed data well but overestimates the probability of two-cycle variable being near zero and underestimates the probability of being near the peaks of the distribution.

4 Analytical tools

In this section, we discuss mutual information and forecast skill as they are used to analyze the data representations and calculate prediction quality of the dynamical models. In Sec. 4.1, we introduce mutual information which is used for two purposes in this work: 1) to set the width of the transition state for the three and four state representations and 2) to assess the quality of the representations in capturing the information present in the time series. In Sec. 4.2, we introduce forecast skill and Brier score, and discuss the analytical calculations of the Brier Score for the dynamical models.

4.1 Mutual Information

Mutual information is the amount of information shared between two random variables (Cover and Thomas, 2005). Mutual information is measured between two variables and is symmetric, $I(A,B) = I(B,A)$. The mutual information between two continuous variable such as M_t and M_{t+1} is given by,

$$I(M_t, M_{t+1}) = \int_{-\infty}^{\infty} dm_1 \int_{-\infty}^{\infty} dm_2 P_{M_t,M_{t+1}}(m_1,m_2) \log_2 \left(\frac{P_{M_t,M_{t+1}}(m_1,m_2)}{P_{M_t}(m_1) P_{M_{t+1}}(m_2)} \right) \tag{15}$$

where P_{M_t} and $P_{M_{t+1}}$ are marginal probability densities, and $P_{M_t,M_{t+1}}$ is the joint probability density. The mutual information between discrete variables such as σ_t and σ_{t+1} is given as,

$$I(\sigma_t, \sigma_{t+1}) = \sum_{d_1} \sum_{d_2} P_{\sigma_t,\sigma_{t+1}}(d_1,d_2) \log_2 \left(\frac{P_{\sigma_t,\sigma_{t+1}}(d_1,d_2)}{P_{\sigma_t}(d_1) P_{\sigma_{t+1}}(d_2)} \right) \tag{16}$$

where, in this case, d_1 and d_2 take the discrete values $\{-1, 0, +1\}$, $P_{\sigma_t,\sigma_{t+1}}$ is the joint probability mass function of σ_t and σ_{t+1}, and P_{σ_t}, $P_{\sigma_{t+1}}$ are the marginal probability mass functions. The mutual information between a discrete and continuous variable can be written in a similar way with an integral and sum over the domain of variables.

4.2 Forecast skill

Forecast skill measures the ability of a model to predict the probability of an event (Hamill and Juras, 2006). We use the Brier score to characterize the ability of the dynamical models discussed in Sec. 3 to predict the observed data one time step in the future. The Brier score, B is defined as the mean squared error in a probabilistic prediction, $\tilde{P}_t$, of a binary observable, $\mathcal{O}_t \in \{0,1\}$ at time t, averaged over many times,

$$B = \frac{1}{T} \sum_{t=1}^{T} (\tilde{P}_t - \mathcal{O}_t)^2. \tag{17}$$

For example, in weather forecasting, where the concept of forecast skill originated, a binary observable might be rain ($\mathcal{O}_t = 1$) or no rain ($\mathcal{O}_t = 0$) and $\tilde{P}_t$ the probabilistic forecast (e.g. 20% chance of rain).

The Brier forecast skill, F compares the Brier score of the probabilistic prediction to a simple reference forecast,

$$F = 1 - \frac{B}{B_{\text{ref}}} \tag{18}$$

where B_{ref} is the Brier score of the reference forecast (e.g. the overall climatological probability of rain). Reference forecasts relevant for our study are introduced at the end of this subsection. Forecast skill is 1 for a perfect forecast, greater than zero for a forecast that improves on the reference, and negative for forecast with less skill than the reference.

Here, we present forecast skill results for the event that the system changes its phase of oscillation at time $t+1$ given the two-cycle variable M_t as shown in Fig. 5 where D_t is any of the representations discussed in Sec. 2 and Fig. 3 and $\tilde{D}_{t+1}$ the corresponding model prediction. The forecast skill results exclude the three-state model since the three-state variable does not have information about the phase of oscillation when the system is in the transition state, $\sigma_t = 0$. For every D_t except the three-state variable, we calculate the probability $\tilde{P}_t$ that the system changes its phase of oscillation using its dynamics with inferred parameters (discussed in Sec. 3). The predicted probability of a phase change for each model is given by,

$$\tilde{P}_t = \Pr(\text{sign}(M_t \tilde{D}_{t+1}) = -1). \tag{19}$$

Note that in the four state model, we interpret $\tau = 0^{\pm}$ as being a small, non-zero numbers with signs corresponding to superscript.

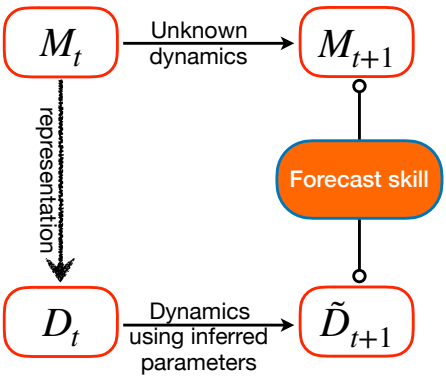

Fig. 5 The forecast skill of a model is obtained from the conditional probabilities of M_{t+1} on M_t and of $\tilde{D}_{t+1}$ on M_t. The properties of the Ricker time series data (on which M is based) are obtained as a function of noise and the growth parameter r.

Brier Score calculation:
The forecast skill is measured by calculating the Brier score between the calculated probability and observed data using Eq. (17) over many times t. The forecast event is whether the system changes its phase of oscillation at time $t+1$. The time series data used to infer the parameters (discussed in Sec. 3) is analyzed to calculate the forecast skill.

For the two-state dynamics, the predicted probability for phase change $\tilde{P}_t$ is the same as the model parameter, $P_\theta(1 \to -1)$. As $\mathcal{O}_t^2 = \mathcal{O}_t$ and $\langle \mathcal{O} \rangle = P_\theta(1 \to -1)$, the Brier score for the two-state dynamics is,

$$B_\theta = P_\theta(1 \to -1)\big[1 - P_\theta(1 \to -1)\big]. \tag{20}$$

In the case of four-state dynamics, there are two predicted probabilities depending on whether the system is in the transition state or not. The predicted probabilities are given by,

$$\tilde{P}_t = \begin{cases} Q_0 \equiv 1 - \Big[P_\tau(0^+ \to 0^+) + P_\tau(0^+ \to 1)\Big], & \text{if } \tau_t = 0^\pm \\ \\ Q_1 \equiv 1 - \Big[P_\tau(1 \to 0^+) + P_\tau(1 \to 1)\Big], & \text{if } \tau_t = \pm 1. \end{cases} \tag{21}$$

Using the probability, $P_\tau(0^+)$ that the system is in the transition state $\tau = 0^+$, the Brier score for four-state dynamics is,

$$B_\tau = \big[2P_\tau(0^+)\big]Q_0(1-Q_0) + \big[1 - 2P_\tau(0^+)\big]Q_1(1-Q_1). \tag{22}$$

For the cubic model, the predicted probability of phase change is a function of M_t and the inferred parameters $\{a,\ b,\ \lambda\}$, which follows from the normal distribution in Eq. (7),

$$\tilde{P}_t = \frac{1}{2}\mathrm{erfc}\Big(\frac{\alpha}{\sqrt{2}\lambda}\Big), \tag{23}$$

where

$$\alpha = a|M_t| - b^2|M_t|^3, \tag{24}$$

and erfc is the complementary error function.

The Brier score for the cubic model does not have a analytical form and must be calculated numerically using the above expressions for the predicted probability in Eq. (17).

Along with the dynamical models, a reference forecast is needed, as shown in Eq. (18), to measure forecast skill. As the probability of the system to change its phase of oscillation is close to zero (see Fig. 6) in the range of $\lambda_r = [0.11, 0.2]$, one of the simple reference probabilities for phase change is, $\tilde{P}_{\mathrm{ref}} = 0$. So, the reference Brier score in this case is,

$$B_0 = \langle \mathcal{O} \rangle = P_\theta(1 \to -1). \tag{25}$$

We will also consider the predictions of the two-state model as a reference for the more complex representations, as described below.

5 Results

The results discussed in this section are obtained using synthetic data generated from noisy Ricker maps as the observed time series. Although the underlying dynamics is known in this study, we treat the two-cycle time series data as though the true dynamics is unknown, as would be the case with real-world data. In Sec. 5.1, we present dynamical properties of the data by studying the changes in the phase of oscillation. In Sec. 5.2, we calculate the width of the transition states for the three-state and four-state variables using mutual information. Finally, in Sec. 5.3, we compare the prediction quality of the dynamical models using mutual information and forecast skill under various conditions.

5.1 Dynamical properties: phase changes

In the presence of noise, two-cycle oscillators change their phase of oscillation at random times. Using the two-cycle or two-state representation of the given time series data, the binary observable, $\mathcal{O}_t$,

$$\mathcal{O}_t = \frac{1 - \text{sign}(M_t M_{t+1})}{2} = \frac{1 - \theta_t \theta_{t+1}}{2} \tag{26}$$

is one if the system changes phase from time t to $t+1$ and zero otherwise. The probability of the system to change its phase of oscillation is estimated from the average over a long time series,

$$\langle \mathcal{O} \rangle = \frac{1}{T} \sum_{t}^{T} \mathcal{O}_{t=1}. \tag{27}$$

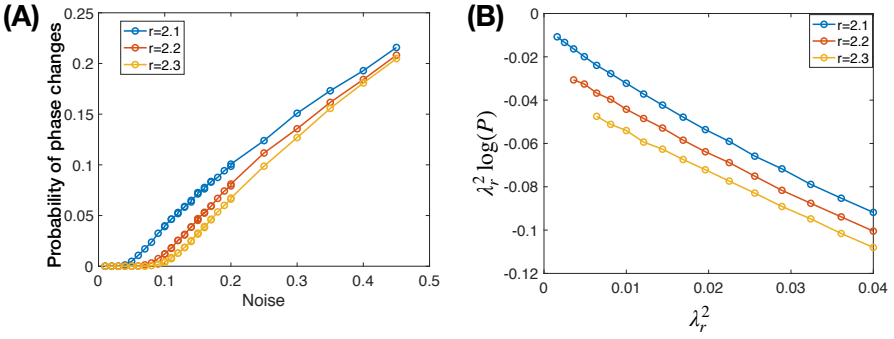

Fig. 6 Phase changes in a noisy two-cycle oscillator. **(A)**Probability of phase changes, $P = \langle \mathcal{O} \rangle$, as a function of noise, λ_r is plotted for two-cycle data obtained from the noisy Ricker map (see Eq. (1)) with growth parameter r. In the low noise limit, phase changes are very rare and the probability approaches zero. **(B)** The data is re-plotted with rescaled axes suggested by Eq. (28) with λ_r as noise and P as probability of phase changes. The trend of the plot is linear as expected, with $\log(k)$ as the slope and H as the y-intercept.

The probability of phase change is measured as a function of noise and growth parameter for the noisy Ricker map data in the two-cycle regime and shown in Fig. 6A. At high noise, phase changes are random and the phase change probability approaches 1/2. In the low noise limit, phase changes are very rare and the probability of phase change approaches zero. In this work, except for this plot, we analyze the time series data in the intermediate noise range $\lambda_r \in [0.11, 0.2]$.

The phase changes in noisy two-cycle oscillatory systems can be related to the state changes in bistable systems such as barrier crossing in the presence of thermal noise. Using results from the Kramer's reaction-rate theory (Hänggi et al, 1990; Kramers, 1940), the probability for phase change has the form,

$$\text{Probability} = k \exp\left[\frac{-H}{\lambda^2}\right] \tag{28}$$

where λ^2 is the variance of the noise, H is the barrier potential and k is a proportionality constant. The probability of phase changes shown in Fig. 6A is re-plotted in Fig. 6B to demonstrate how well the above equation obtained from the theory of bistable systems explains the dynamical property of phase changes in two-cycle oscillations.

The barrier potential, H can be calculated approximately using the cubic model as,

$$H = \frac{(a-1)^2}{2b^2} \tag{29}$$

where a, b are the parameters of the cubic model. Equations (28) and (29) follow from the idea that the transition probability is proportional to the ratio of the maximum to minimum stationary state probabilities, which, in the cubic model are obtained from Eq. (11). The linear trend in Fig. 6B is expected to have H as the y-intercept and $\log(k)$ as the slope.

5.2 The transition state

The three-state and four-state representation of the observed data require setting the width of the transition state. The objective in choosing the width is to minimize the loss of predictive information about the future due to the coarse-graining involved in the going from the continuous variable, M_t, to the discrete variable σ_t. Here, we use mutual information to set the width of the transition state.

Since we have assumed both the data and the models are Markovian, we only consider mutual information between two consecutive time steps. The width parameter of the three-state representation, w, is thus calculated by minimizing the difference between the mutual information of successive two-cycle variables $I(M_t, M_{t+1})$ and the mutual information of successive three-state variables $I(\sigma_t, \sigma_{t+1})$ as shown in Fig. 7. This is equivalent to maximizing $I(\sigma_t, \sigma_{t+1})$, but we present the results in terms of this difference for ease of interpretation.

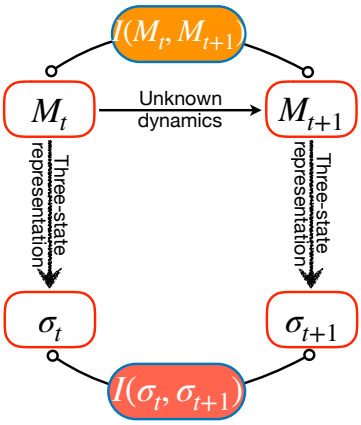

Fig. 7 The width parameter, w, for the three-state representation is calculated using mutual information. Pairs of two-cycle variables (M_t, M_{t+1}) are obtained from the observed data and their mutual information $I(M_t, M_{t+1})$ is calculated. Pairs of three-state variables (σ_t, σ_{t+1}) are calculated directly from (M_t, M_{t+1}) (not from three-state dynamics) and the loss of mutual information, $I(M_t, M_{t+1}) - I(\sigma_t, \sigma_{t+1})$ is minimized as a function of width, w.

The mutual information difference $I(M_t, M_{t+1}) - I(\sigma_t, \sigma_{t+1})$ as a function of width w are shown Fig. 8A for Ricker time-series data with growth parameter $r = 2.2$ and various values of noise λ_r. For a given noise λ_r, the width corresponding to the minimum information loss is chosen as the width of the transition state. The resulting values of the width of the transition state are shown in Fig. 8B as a function of λ_r for $r = 2.1$, 2.2 and 2.3. For any given observed time series data, calculation of the width of the transition state by minimizing the information loss completes the definition of the three-state representation.

Since the two and three-state representations are identical when width is zero, Fig. 8A also shows the loss in mutual information due to representing M_t with the two-state variable, θ_t. The results show that the three-state representation contains at least 0.2 bits more information than the two-state representation for the noise values shown here.

The two-cycle variable M_t is course-grained into three bins by the three-state representation and these bin are superimposed on histograms of M in Fig. 8C for three values of noise. As noise increases, transitions between the phase of oscillations increase and the probability of being in the transition state, plotted in Fig. 8D, increases. At high noise, the histogram approaches a pure quartic potential and the probability of being in the transition state saturates to 1/3.

In the very low noise limit the joint distribution of M_t and M_{t+1} is concentrated in the two regions, $M_t \approx M_{t+1} \approx \pm(a^+ - a^-)$ where a^+ and a^- are the two fixed points of the deterministic two-cycle Ricker map and the map is nearly deterministic. Thus, the mutual information $I(\sigma_t, \sigma_{t+1})$ is near one bit for $w < (a^+ - a^-)$ since the system is almost always in the one of the states

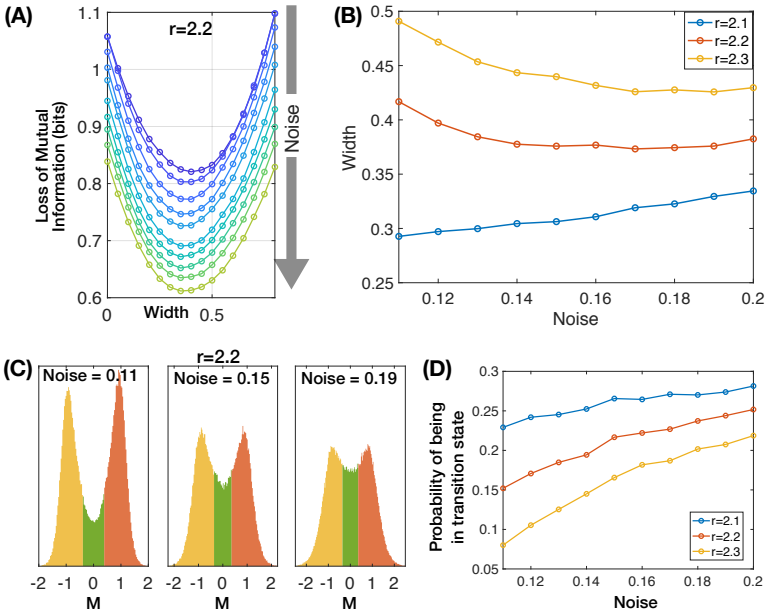

Fig. 8 Calculation of width w of the transition state in the three-state representation using mutual information. **(A)** Loss of mutual information (bits), $I(M_t, M_{t+1}) - I(\sigma_t, \sigma_{t+1})$ is calculated as function of width with observed data from the noisy Ricker map with $r = 2.2$ and various noise values. **(B)** The width that minimizes the information loss is plotted for noisy Ricker data as function of noise λ_r for several values of the growth parameter r. **(C)** The distributions of M are shown for $r = 2.2$ by highlighting the three states of σ where the green (middle) bin corresponds to the transition state ($\sigma = 0$). **(D)** The probability of being in the transition state is plotted and has an increasing trend as expected from the increase in phase changes with noise.

$\sigma_t = \pm 1$ state and almost always $\sigma_{t+1} = -\sigma_t$. Similarly, the mutual information is near zero for $w > (a^+ - a^-)$. Furthermore, $I(\sigma_t, \sigma_{t+1})$ increases as w approaches $(a^+ - a^-)$ from below so that $w \nearrow (a^+ - a^-)$ as $\lambda_r \to 0$. Although the width of the transition state increases as the noise decreases, the stationary probability of being in the transition state goes to zero, $P(\sigma = 0) \to 0$ as $\lambda_r \to 0$. The weak noise limits are shown as arrows in Fig. 9.

It should be noted that the bins defining the two-state, three-state and four-state representations have implicit properties that have been set by symmetry. For example, in the two-state representation there is a hidden parameter that sets the value of M separating $\theta = +1$ and $\theta = -1$. Symmetry arguments show this parameter is zero but the mutual information method yields the same result and could be used more generally when no symmetry argument is available.

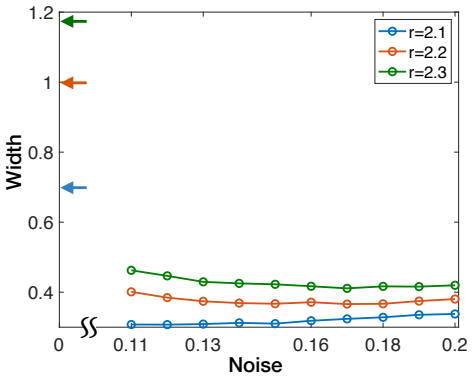

Fig. 9 In the low noise limit, the width of σ approaches the difference between the stable fixed points of oscillation. The low noise limits of the width are shown as arrows on the y-axis.

5.3 Prediction quality

In the following two subsections we assess the predictive ability of the dynamical models introduced above using mutual information and forecast skill.

5.3.1 Mutual Information results

In this subsection, we study the amount of information the model predictions share with the observed data. The strategy is shown in Fig. 10A. First, pairs of two-cycle variables M_t and M_{t+1} are obtained from the observed data. Next, the two-cycle variable at the initial time M_t is represented as D_t Here D_t is either M_t itself for the cubic model, or, for the discrete state models, θ_t, σ_t or τ_t for the two-state, three-state or four-state models, respectively. The subsequent value, $\tilde{D}_{t+1}$ is obtained from the corresponding models with the inferred parameters and then the mutual information is calculated between M_{t+1} and $\tilde{D}_{t+1}$.

Mutual information results are shown in Fig. 10B for the cubic, two-state, three-state and four-state models. Mutual information is highest for the cubic model, which is the representation that retains full information from M_t, and decreases as the number of discrete states decreases as more information is lost by representing M_t as D_t. Mutual information decreases roughly linearly as a noise increases.

In the limit of low noise, all the models converge to one bit of mutual information. The oscillator can be in one of two states so the information content of its state is one bit. Furthermore, all models incorporate the property that no phase changes occur at zero noise so the predictions are perfect.

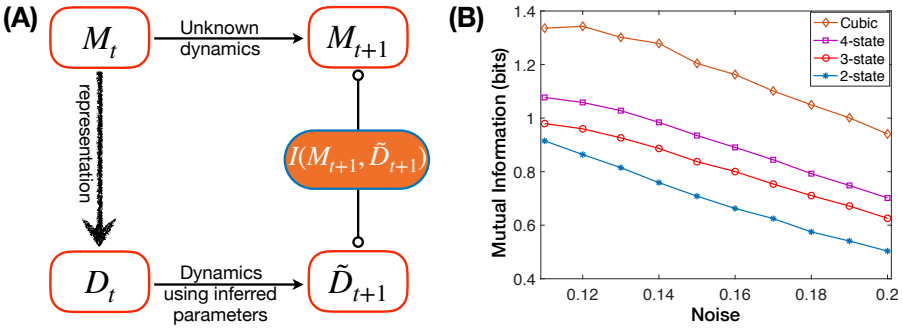

Fig. 10 Mutual information method and results. **(A)** The top row shows the two-cycle variables M_t and M_{t+1} obtained from the observed data. M_t is represented as D_t which can be any of the data representations discussed in Sec. 2. Using the models discussed in Sec. 3, corresponding realizations of $\tilde{D}_{t+1}$ are calculated. **(B)** The mutual information $I(M_{t+1}, \tilde{D}_{t+1})$ between M_{t+1} and $\tilde{D}_{t+1}$. The results shown here are obtained from the data generated from the Ricker map with $r = 2.3$. The more fine-grained models contain more information about the underlying data and all models deteriorate with increasing noise.

5.3.2 Forecast skill results

In this section we study the ability of the dynamical models to predict changes in the phase of oscillation. We use forecast skill with respect to two reference models to quantify this ability. The first reference model is the simple no-noise assumption that phase changes do not occur. The second, more accurate, reference model is that phase changes occur at an average rate independent of the current state. This reference is none other than the transition probability from the two-state model.

The forecast skill relative to the no-noise reference is shown in Fig. 11A. Using Eqs. (17), (20), and (25), one sees that the forecast skill of the two-state model is the same as the probability of the system changing its phase of oscillation, which is small even for intermediate noise. The addition of the transition state in the four-state model increases the forecast skill as shown in the plot. The cubic model does a much better job of predicting phase changes at low noise because it is sensitive to the rare situations that M is close to zero when transitions can occur. However, for higher noise the cubic model and the four-state model perform similarly.

A more accurate choice of reference is the two-state model with $P_\theta(1 \to -1)$ (see Sec. 3) as the reference probability for change in phase of oscillations. Results for the cubic and four-state models are shown in Fig. 11B and, again, the cubic model does much better at low noise because it detects the very small values of M associated with transitions. while the transition state of the four-state representation is quite wide here. On the other hand, as noise gets larger, being in the transition state is a good predictor of an immediate phase change and knowing the exact value of M is less helpful. The four-state model improves with noise over the range studied here because it becomes more likely that a phase change occurs following a single visit to the transition state.

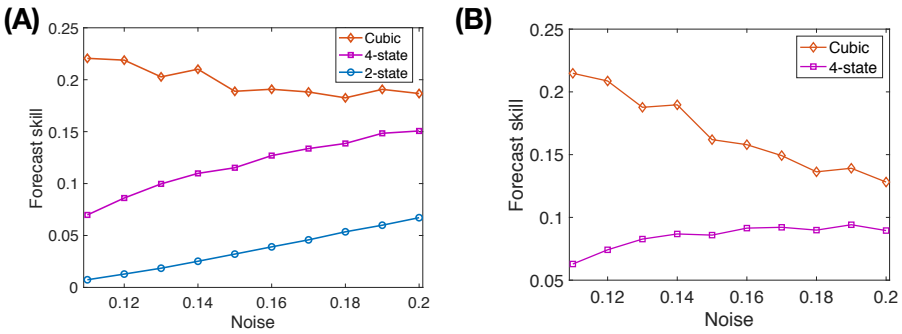

Fig. 11 Forecast skill results. **(A)** Forecast skill for predicting a change in the phase of oscillation for the cubic, four-state, and two-state models where the reference probability of change in phase of oscillation is 0. **(B)** Forecast skill for the cubic and four-state models with the two-state model as the reference. Note that all of the models have positive forecast skill.

In the high noise limit, phase changes occur with probability 1/2 and, relative to the two-state reference, the more complicated models have no improved forecast skill. In the low noise limit, the no-noise, two-state and four-state models all have Brier scores of the form $P_\theta(1 \to -1)(1+\epsilon)$ with ϵ depending on the model but, in all cases, $\epsilon \to 0$. Thus, none of these models have forecast skill relative to the other. The proof follows from Eqs. (20), (22) and the fact that $P_\theta(1 \to -1)$, Q_0 and Q_1 all vanish in the low noise limit.

The transition state plays an important role in the forecast skill measurement as most of the changes in phase of oscillation happen when the system is in the transition state. Forecast skill results shown in Fig. 11B with the two-state dynamics as reference are shown as solid lines in Fig. 12 along with the results obtained by conditioning on σ_t (that is, conditioning on whether or not the system is in the transition state). The forecast skill measured when the two-cycle variable M_t is in the transition state, $\sigma_t = 0$ is shown as dashed lines and when M_t is not in the transition state, $\sigma_t = \pm 1$ is shown as dotted lines. The accuracy of prediction increases when the system is in the transition state for both cubic and four-state dynamics. This result confirms that the forecast skill about change in phase of oscillations can be improved with the information of transition state.

6 Conclusions

Any effort at ecological modeling should begin with a clear understanding of the question being asked. Here we use the idea of asking what information can be gleaned from a focus on phase shifts in the relatively common two-cycle oscillations in ecology using only crude data. By a phase shift, we mean a change in the expected alternation of high and low states, so the following high or low states occur at a different time than expected. Although there is a long literature on time series analysis in ecology, much of this has been focused on fitting data that is relatively high resolution.

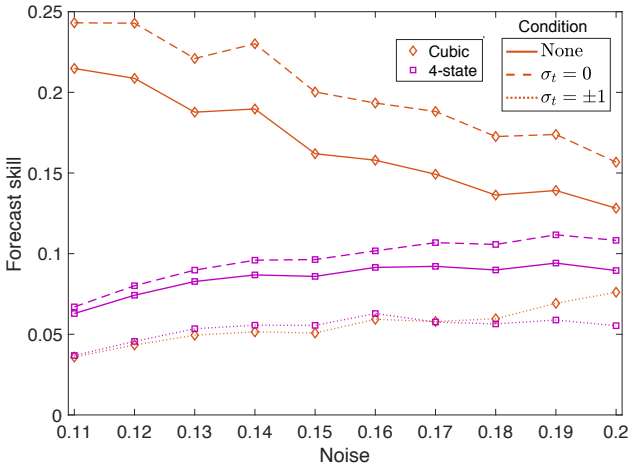

Fig. 12 Forecast skill is measured with the two-state model as the reference for predicting a change in the phase of oscillation. The solid lines are the same as shown in Fig. 11B. For solid lines, the state D_t is observed and D_{t+1} is predicted. The same measurement of forecast skill is obtained by conditioning on the observed σ_t and is shown as dashed and dotted lines for the cubic and four-state model. The dashed forecast skill is measured when the observed D_t is present in the transition-state of the three-state model (Condition: $\sigma_t = 0$ or $\tau_t = 0^+, 0^-$). The dotted forecast skill is obtained when D_t is not present in the transition-state (Condition: $\sigma_t = \pm 1$).

Yet much ecological data is rather crude, so methods and approaches that use only coarsely binned data can have advantages. For example, another typical limitation is that high resolution time series are short while very long time series that can be obtained from different kinds of historical records which of necessity produce very rough estimates, such as high or low abundance.

Even if the actual observations may be crude, the underlying system being observed will typically have continuous states. So a key question is how much is lost by using only a coarse grained representation with a small number of states to describe the system. One challenge is to be able to predict the likelihood and occurrence of future phase shifts.

We consider how much information is lost when relying on a discrete state representation, relative to having knowledge of the full system, using mutual information (Cover and Thomas (2005)) and forecast skill. In many ways this view is completely within the standard state space approach used in time series analysis (e.g. Auger-Méthé et al (2021)) where there is an underlying process and an observation process and inferences must be made using only the observation. What is different here is our focus is on phase shifts and on systems where the observation process leads to a very small number of states.

An intriguing outcome of our study is the possibility of inferring process from pattern: what does knowing the frequency of phase shifts tell us about the underlying dynamics. Here, the result illustrated in Fig. 6A, where the relationship between noise level and frequency of phase shifts is only slightly

affected by the growth parameter of the Ricker model, suggests that noise level might be estimated by frequency of phase shifts.

We have limited our study to Markovian models. It would be interesting to see to what extent forecast skill is improved by adding some history to the model. For example, especially in the low noise regime, the system must visit the transition state for several time steps before a phase change occurs so a model with history may perform better than a Markov model, though at the expense of more parameters that need to be inferred. In addition, the role of observation noise is an important topic for future study and would help decide between models. Intuitively, the performance advantage of the more complex models is expected to disappear as observation noise increases. Finally, the symmetry between the two phases of oscillation can be broken, either explicitly or via state dependent noise. The same can models can be applied but with more parameters and different criteria for coarse-graining into discrete states.

The paper as a whole focuses on a very simple system, a single oscillator. However, even for this system, we have shown how a focus on phase shifts for a noisy cyclic system can provide new insights. This is in the spirit of the idea that studying cyclic systems provides much more information than studying systems that approach an equilibrium. The obvious steps of considering more complex systems will have to wait for future research.

Acknowledgments

This research was supported by NSF Grant DMS-1840221.

References

Abbott KC, Ripa J, Ives AR (2009) Environmental variation in ecological communities and inferences from single-species data. Ecology 90(5):1268–1278

Auger-Méthé M, Newman K, Cole D, et al (2021) A guide to state–space modeling of ecological time series. Ecological Monographs 91(4):e01,470

Bence JR (1995) Analysis of short time series: correcting for autocorrelation. Ecology 76(2):628–639

Box G, Jenkins GM (1976) Time Series Analysis: Forecasting and Control. Holden-Day

Courchamp F, Clutton-Brock T, Grenfell B (1999) Inverse density dependence and the allee effect. Trends in ecology & evolution 14(10):405–410

Cover TM, Thomas JA (2005) Elements of Information Theory. John Wiley & Sons, Ltd

Crawley M, Long C (1995) Alternate bearing, predator satiation and seedling recruitment in quercus robur l. Journal of Ecology pp 683–696

Dietze MC (2017) Ecological Forecasting. Princeton University Press

Fagan WF (2001) Characterizing population vulnerability for 758 species. Ecology Letters 4:132–138

Gross K, Ives AR, Nordheim EV (2005) Estimating time-varying vital rates from observation time series: a case study in aphid biocontrol. Ecology 86:740–752

Hamill TM, Juras J (2006) Measuring forecast skill: is it real skill or is it the varying climatology? Quarterly Journal of the Royal Meteorological Society 132(621C):2905–2923

Hänggi P, Talkner P, Borkovec M (1990) Reaction-rate theory: fifty years after kramers. Rev Mod Phys 62:251–341

Ives AR, Dennis B, Cottingham KL, et al (2003) Estimating community stability and ecological interactions from time-series data. Ecological Monographs 73(2):301–330

Ives AR, Abbott KC, Ziebarth NL (2010) Analysis of ecological time series with ARMA(p, q) models. Ecology 91(3):856–871

Kramers H (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 7(4):284–304

Lyles D, Rosenstock TS, Hastings A, et al (2009) The role of large environmental noise in masting: general model and example from pistachio trees. Journal of Theoretical Biology 259(4):701–713

May RM (1976) Models for single populations, in "theoretical ecology"(rm may, ed.). Blackwell, Oxford 150:4–25

Monselise S, Goldschmidt E (1982) Alternate bearing in fruit trees. Horticultural reviews 4(1):128–173

Noble A, Karimeddiny S, Hastings A, et al (2017) Critical fluctuations of noisy period-doubling maps. Eur Phys J B 90(7)

Powell TM, Steele JH (eds) (1995) Ecological Time Series. Springer

Rosenstock TS, Hastings A, Koenig WD, et al (2011) Testing moran's theorem in an agroecosystem. Oikos 120(9):1434–1440

Sibly RM, Barker D, Hone J, et al (2007) On the stability of populations of mammals, birds, fish and insects. Ecology Letters 10:970–976

Sugihara G, Grenfell B, May RM (1990) Distinguishing error from chaos in ecological time series. Philosophical Transactions of the Royal Society B 330:235–251

Turchin P, Taylor AD (1992) Complex dynamics in ecological time series. Ecology 73(1):289–305

White ER, Hastings A (2020) Seasonality in ecology: Progress and prospects in theory. Ecological Complexity 44:100,867