Detecting chaos by estimating entropy rate of the ε machine constructed from time series

Junna Wang

Introduction

Chaos has been found in many ecological models such as the nonlinear discrete population model (May 1976), three-species food chain model (Hastings and Powell 1991), and nutrient-plankton community model (Huisman and Weissing 1999). However, the detection of chaos in natural community remains rare (Bjørnstad and Grenfell 2001). Solid evidence of chaos in nature includes chaotic dynamics of measle disease (Sugihara and May 1990), an insect population (Costantino et al. 1997), and a long-term experiment about plankton community (Benincà et al. 2008). Calculating Lyapunov exponent from dynamical models or from long-term timeseries is the traditional method of detecting chaos (Hastings et al. 1993). A positive Lyapunov exponent is viewed as the hallmark of the existence of chaos.

Furthermore, information theory shows that entropy rate of a dynamical model is equal to the sum of positive Lyapunov exponents (lecture note of computational mechanics). If a dynamical system possesses positive Lyapunov exponents, its entropy rate should also be positive. Therefore, there exists another possible method of detecting chaos ---- construct ε machine from time series ---- calculate entropy rate of the ε machine ---- detect chaos based on the value of entropy rate. The advantage of this method is that once we have a reliable ε machine, we can calculate many properties of the ε machine such as excess entropy (mutual information between the past and the future), and structural complexity. However, the difficulty of this method is that we have to find a workable way to partition time series that are observed from continuous systems in nature. This project explores different partitioning approaches: generating partition and arbitrary partition. To test if this method of detecting chaos works, as a first attempt, I use synthetic time series generated from a dynamical system which is known to be chaotic.

Methods

Dynamical system

The chaotic dynamical system I study is a three-species food chain model (Hastings and Powell 1991). This model includes three ordinary differential equations describing predation relationships among three species: species x is eaten by species y, and species y is eaten by species z (Eq. (1)). The nonlinearity of the model arises from the nonlinear functional response terms (i.e., $f_1(x)$ and $f_2(y)$ in Eqs. (1) and (2)).

$$\frac{dx}{dt} = x(1-x) - f_1(x)y$$
$$\frac{dy}{dt} = f_1(x)y - f_2(y)z - d_1 y \quad (1)$$
$$\frac{dz}{dt} = f_2(y)z - d_2 z$$

$$f_1(x) = a_1 x/(1+b_1 x)$$
$$f_2(y) = a_2 y/(1+b_2 y) \quad (2)$$

Where t is time; x, y, and z are biomass of the three species; a_1, a_2, b_1, b_2, d_1 and d_2 are parameters, and $f_1(x)$ and $f_2(y)$ are non-linear functional response functions, which is likely the main origin of complex dynamics (Abrams and Roth 1994). It has been shown that this model exhibits complex dynamics, including stable equilibrium, chaotic attractors, and coexistence of multiple attractors when parameters vary within wide ranges. The dynamical behaviors are mostly sensitive to the variation of parameter b_1.

This model is numerically solved by the Runge-Kutta method in Matlab. I plot its bifurcation diagram over the parameter b_1 to figure out the region with chaotic dynamics. Then I choose a set of parameter values which can generate chaotic time series. The parameter values are $a_1 = 5$, $a_2 = 0.1$, $b_1 = 3$, $b_2 = 2$, $d_1 = 0.4$, $d_2 = 0.01$. I also draw Poincaré map of the dynamical system for the variable x. Fortunately, the Poincaré map shows a nice nonlinear relationship, which allows me to fit the Poincaré map with a six-order polynomial regression, and then calculate the Lyapunov exponent of the Poincaré map. This Lyapunov exponent is further compared with entropy rate of the ε machine constructed from discrete time series.

Generation of discrete time series and partitioning

I employ two methods to generate discrete time series. The first method directly uses the data points on the Poincaré map. Because the Poincaré map shows a nice nonlinear relationship, I use generating partition to partition the discrete time series into 4 alphabets. The second method generates time series from the numerical solution of the dynamical model with an arbitrary output time step. Here, I use an output time step of 10. The partition of the time series is also arbitrary. I partition them into 2 alphabets, 3 alphabets, 4 alphabets, and 5 alphabets. The purpose here is to find which type of partitioning can generate the entropy rate this is closest to Lyapunov exponent of the Poincaré map.

Construction of ε machine

I try to construct ε machine using the Bayesian inference method. Since I have time series with 4 and 5 alphabets, the Bayesian method to construct ε machine cannot run on my laptop. I end up constructing Markov chains and calculating entropy rate of these Markov chains. It is expected that entropy rate of Markov chain will be higher than entropy rate of ε machine since these Markov chains are not the minimal models representing dynamics of discrete time series.

Preliminary results

Complex dynamics of the food chain model

The bifurcation diagram of this model along the parameter b_1 is shown in Figure 1a. Dynamics of this model changes dramatically when b_1 surpasses the point 2.2 and the point 2.9. When $b_1 < 2.2$, the attractor is a fixed point; when $b_1 > 2.9$, the attractor is chaotic; and between the two regions, the attractor can switch between cycles and chaos with a subtle variation in parameter b_1. It seems that there are multiple attractors in this region. The chaotic attractor in the region $b_1 > 2.9$ is like a tea cup, as shown in Figure 1b. The biological interpretation of this attractor is that when the abundance of top species (z) increases (the trajectory goes up along the body of the cup), the abundance of middle species (y, the prey of z) declines, and the abundance of bottom species (z, the prey of y) goes up. When the number of middle species y is too low, the top species z crashes suddenly from its peak (move from the peak to the handle of the cup) and then abundance of species y starts to increase, and abundance of x starts to decline (the trajectory moves from the handle back to the body of the cup).

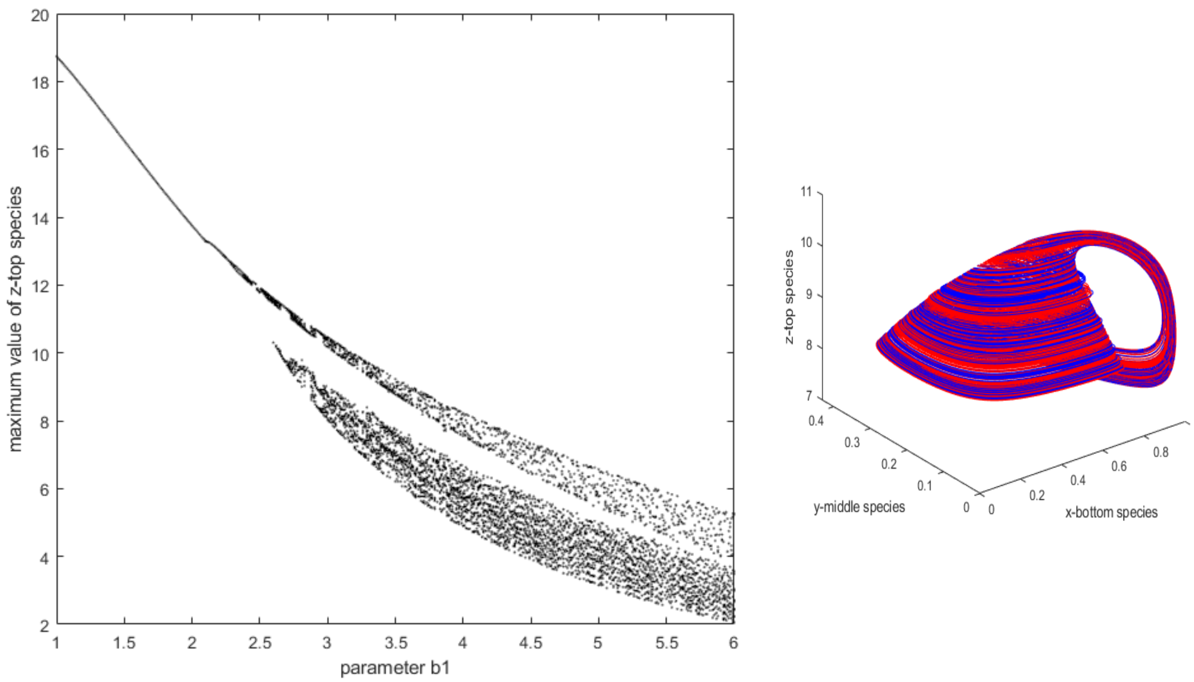

Figure 1 (a) Bifurcation diagram (left) and (b) chaotic attractor when $b_1 = 3.0$ (right).

Poincaré map and its Lyapunov exponent

I obtained a Poincaré cross-section by placing a horizontal plane with $z = 8.5$ and cutting through the handle of the cup. The trajectory passes this cross-section about every 100 time-units (here the equations are non-dimensionalized and unitless). Surprisingly, the Poincaré cross-section is like a straight line (Fig. 2a). Poincaré map, a return map of variable x is thereby obtained (see Fig. 2b). The Poincaré map has two peaks and one valley. The six-order polynomial to fit these points is given by Eq. (3). Lyapunov exponent of the Poincaré map calculated based on the fitted polynomial is 1.7132.

$$x_{n+1} = 10^9(0.4257(x_n)^6 - 2.5335(x_n)^5 + 6.2766(x_n)^4 - 8.2859(x_n)^3 + 6.1477(x_n)^2 - 2.4307x_n + 0.4001) \tag{3}$$

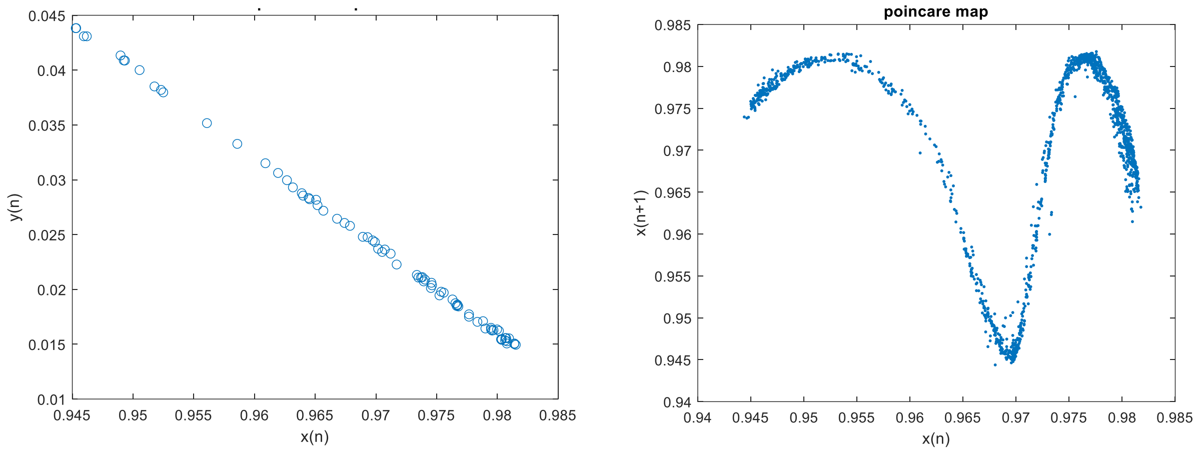

Figure 2 (a) Poincaré cross-section with z = 8.5 and (b) the Poincaré map of variable x.

Discrete time series and partitioning

As mentioned in the method, the first method to generate discrete time series is using points on the Poincaré map. I divided the range of x (Fig. 2b) into four regions using generating partition, that is partitioning at the peak and valley points: $x < 0.9529$, $0.9529 \le x < 0.9695$, $0.9695 \le x < 0.9764$, and $x \ge 0.9764$. Then, Markov chains with different numbers of states are constructed. For the second method, discrete time series of variable x (Fig. 3) is output from the dynamical model with an output time step of 10 (unitless). I then partition these points into 2 alphabets, 3 alphabets, 4 alphabets, and 5 alphabets arbitrarily. The partition points randomly chosen are listed in Table 1.

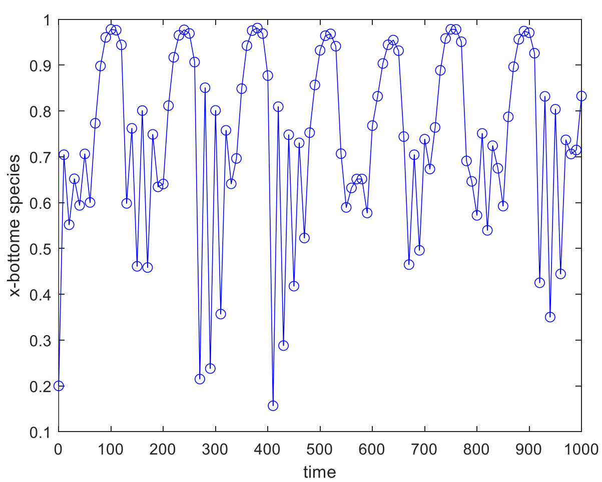

Figure 3. Discrete time series (points) directly output from the dynamical model.

Table 2 Random partition points used to generate different numbers of alphabets

Number of alphabets	Partition points
2	0.6
3	0.5, 0.8
4	0.4, 0.6, 0.8
5	0.3, 0.5, 0.65, 0.8

Entropy rate of Markov chains

Firstly, for the alphabet words generated from the Poincaré map and using generating partition, entropy rates of Markov chains with 4-5 states are between 1.683 and 1.963, quite close to the Lyapunov exponents of the Poincaré map (1.7132). This result suggests that the first method of producing discrete timeseries works as long as a return map with clear nonlinear relationship (e.g., the Poincaré map) can be obtained from observation or synthetic data.

Next, for the words generated by the second method using different numbers of alphabets (2, 3, 4, 5), entropy rates of Markov chains with number of states varying between 2 and 8 are shown in Fig. 4. When using 2 and 3 alphabets, entropy rates of Markov chains are in general very low (< 1.1), much less than the Lyapunov exponents of the Poincaré map (1.7132). This result indicates that choosing a proper number of alphabets is likely a critical step towards estimating reliable entropy rate. With the increase in number of alphabets, entropy rates mostly increase as well. Unfortunately, on my computer, when alphabet is 4, the maximum number of Markov chain states I can run is 6, and when alphabet is 5, the maximum number of Markov chain states is 5. Although I could not run Markov chain with more states, it seems that entropy rates basically increase with the number of states of Markov chains. It is expected that when numbers of alphabets are 4 and 5, entropy rate can reach values higher than the Lyapunov exponents of the Poincaré map (the black line in Fig. 4) with the increasing number of Markov chain states.

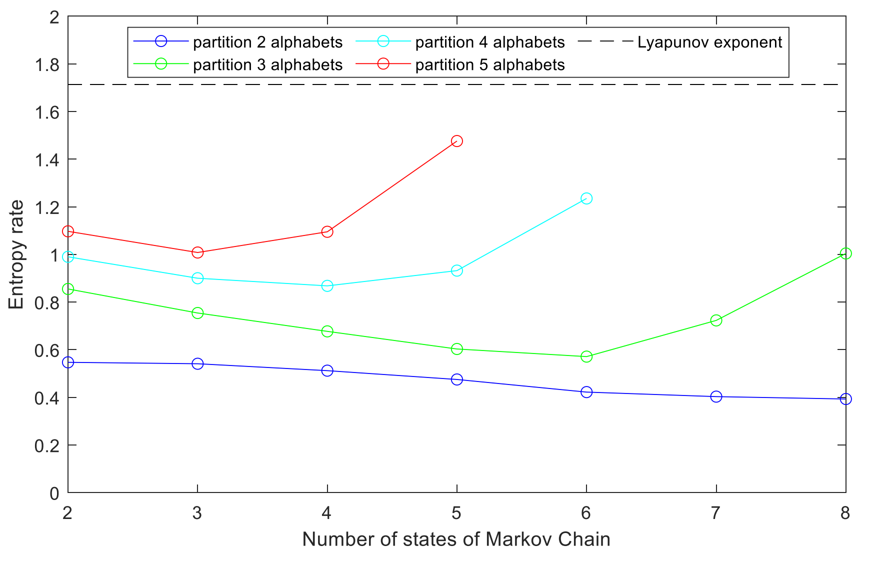

Figure 4 Entropy rate of Markov chains with different numbers of states for various partitions.

If I was able to construct ε machine and calculate its entropy rate, I would get a figure more complete and inspiring than Fig. 4: showing when entropy rate reaches its peak, and which peak of entropy rate is closer to Lyapunov exponent, much like the result shown in Strelioff and Crutchfield (2007). Such a figure would also allow me to decide if the second method of generating discrete time series and partitioning works, and to better understand how entropy rate is sensitive to number of alphabets and number of Markov chain states. This project is far from completion. It only represents a very first attempt to bridge the gap between continuous dynamical systems and computational mechanics. Computational mechanics looks promising in understanding complex systems, but there remains a great amount of work to be done to apply it to natural systems that include a large number of interacting components and plenty of observation/measurement noise.

Reference

Abrams, P. A., and J. Roth. 1994. The effects of enrichment of three-species food chains with nonlinear functional responses. Ecology **75**:1118-1130.

Benincà, E., J. Huisman, R. Heerkloss, K. D. Jöhnk, P. Branco, E. H. Van Nes, M. Scheffer, and S. P. Ellner. 2008. Chaos in a long-term experiment with a plankton community. Nature **451**:822-825.

Bjørnstad, O. N., and B. T. Grenfell. 2001. Noisy clockwork: time series analysis of population fluctuations in animals. Science **293**:638-643.

Costantino, R. F., R. Desharnais, J. M. Cushing, and B. Dennis. 1997. Chaotic dynamics in an insect population. Science **275**:389-391.

Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. Godfray. 1993. Chaos in ecology: is mother nature a strange attractor? Annual review of ecology and systematics **24**:1-33.

Hastings, A., and T. Powell. 1991. Chaos in a three-species food chain. Ecology **72**:896-903.

Huisman, J., and F. Weissing. 1999. Biodiversity of plankton by species oscillations and chaos. Nature **402**:407-410.

May, R. M. 1976. Simple mathematical models with very complicated dynamics. Nature **261**:459-467.

Strelioff, C. C., and J. P. Crutchfield. 2007. Optimal instruments and models for noisy chaos. Chaos: An Interdisciplinary Journal of Nonlinear Science **17**:043127.

Sugihara, G., and R. M. May. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature **344**:734-741.