Nonlin. Processes Geophys. Discuss., 2, C649–C652, 2016
www.nonlin-processes-geophys-discuss.net/2/C649/2016/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Brief Communication: Breeding vectors in the phase space reconstructed from time series data" *by* E. Lynch et al.

E. Lynch et al.

elynch@umd.edu

Received and published: 26 January 2016

Referee Comment:

The authors present a purely data driven method to extract from a given time series dynamical information about the underlying dynamical system. To this extend they combine the bred vector method with the time delay embedding method to construct the phase space of the dynamical system. Within this reconstructed phase space pairs of nearby trajectories between a control and an initially nearby trajectory are piecewise followed over a specified time interval to measure the final separation distance. At this point the bred vector idea comes into play. After a rescaling of the final separation vector a local search in phase space for trajectories close to the final control phase space point shifted by the rescaled final separation vector of the previous iteration is initiated. After a proper identification to avoid points on the control trajectory, the next piecewise tracking is initiated. The growth rates at the end of each interval are the basic information derived from the analysis. The method is applied to the Lorenz 1963 L63 model. Three set ups are compared: (1) the standard breeding using the explicit L63, the proposed model, and an intermediate one where the time delay embedding is not used in favor of the original three dim L63 structure. The methods are compared by monitoring the local growth rates along the control trajectory and using positive large values as predictors for the regime shifts between the two leaves of the L63 attractor. With respect to this metric the new method performs reasonably well and the authors conclude that the new method provides a purely data-driven way to diagnose regime shifts for dynamical systems not well or not at all described by a set of equations.

In principle the paper is worth to be published in NPGP. It contains new information e.g. the proposed method and offers (some) help in interpreting the results. However, there is no clear conclusion or message especially with respect to data requirements and/or the dimensionality of the dynamics. The authors simply state that they ensured sufficient data density in three dimensions. But it would be worth to see how the contingency table statistics degrade when the actual data density is reduced. Although it is only a result for the idealized L63 it can give hints about the performance using real data. Another nice-to-know information would be on the frequency distribution (estimated probability density) of the calculated growth rates for the three test beds. This would again give more confidence into the new method than the simple thresholding of looking at large growth rates. So my major suggestion before publication of this paper is that the authors should provide a clear message to the reader and potential user: is it worth to apply the method to other types of (real) data because the method is generic or are the results specific to the chosen setup?

Author Comment:

NPGD

2, C649–C652, 2016

Interactive Comment

We thank the referee for taking the time to review our manuscript and provide us with these thoughtful comments. We appreciate the suggestion that our manuscript is appropriate for publication in this journal. Below we address the points raised.

The key objective of our paper is to demonstrate that the breeding vector technique can be used to determine the stability of low-dimensional dynamical systems using the time series data of one of its variables. For this proof-of-principle, a simple and well-known Lorenz63 is used and the parameters used in the analysis are identified. It is expected that this technique may not be effective for some systems and data sets, but a comprehensive analysis is needed to address many related issues, including the case of data density. Like most early papers on a new technique, including Packard et al.(1980) that showed time delay embedding using well-known systems, the emphasis in the paper is to demonstrate the new technique. Further studies on different systems and data sets would, over time, provide answers to the practical questions and the direct message to the reader is that the breeding technique for time series data works for the chosen well-known system (Lorenz63) and thus shows promise as a new technique.

The data density required for accurate results will depend on the individual system and the availability of data – including the sampling rate, the length of the time series, the time delay and dimension required for a proper embedding, and the recurrence time of the system. One way to visualize this is to look at a recurrence plot for the embedding time and dimension selected. To form this plot, the pairwise distance between each point in the dataset is computed. If this distance places the points within a neighborhood of given radius, a symbol is plotted at the corresponding location. This method depends on having segments of trajectories that lie near the control trajectory in question. Taking the desired initial magnitude of separation between the control and perturbed trajectory as the neighborhood radius, this is represented on the recurrence plot by diagonal segments of neighboring points. For a deterministic, nonlinear system, the trajectory will return to a given neighborhood given sufficient time. One can increase the density of the points covering the attractor by observing the time series for a longer time. On the recurrence plot this can be thought of as taking successively larger square segments. For shorter the duration of the time series, resulting in a data set with lower density, there are few neighboring trajectories within the specified radius. This means that it will be difficult to identify neighboring trajectories to serve as perturbations for the control trajectory. As the duration of the time series is increased, more neighboring trajectories within the specified radius become available and the contingency statistics improve. The density we selected was sufficient such that we were able to find analogues to our perturbed initial conditions that met our target perturbation size of $\alpha = 0.1$, on average.

The application of this method to real data will require some knowledge of the characteristics of underling system and analysis of the properties of the time series. Applying the techniques of nonlinear time series analysis outlined in the referenced articles will allow one to assess an appropriate embedding and whether such an approach is suitable for the data in question. Further, with the size limitations of a brief communication, we felt it best to restrict our analysis to the Lorenz attractor. Applications of this and other similar techniques to more complicated real systems are in progress and be forthcoming in future publications.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 1301, 2015.