

## Supplementary Methods

To determine whether a time series reflects linear or nonlinear processes we compare the out-of-sample forecast skill of a linear model versus an equivalent nonlinear model. To do this, we apply a two-step procedure: 1) we use simplex-projection<sup>1</sup> to identify the best embedding dimension, and 2) we use this embedding in the S-map procedure<sup>2</sup> to assess the nonlinearity of the time series. In both cases, model performance is evaluated out-of-sample with the time series divided into equal halves. The first half (library set,  $X$ ) is used to build the model, while the second half (prediction set,  $Y$ ) is reserved to judge the out-of-sample performance of model forecasts originating from the library. This forecast protocol is a rigorous standard that avoids model over-fitting or arbitrary fits to the data.

### Simplex projection

Simplex projection is a nearest-neighbor forecasting algorithm that involves tracking the forward evolution of nearby points in an embedding (a lagged coordinate state space reconstruction).<sup>3,4,5</sup> Thus, similar past events are used to forecast the future, with the important caveat that the dimensionality of the embedding determines what past events are similar (nearby) to the predictee.

Given a library set of  $n$  points, we generate an  $E$ -dimensional embedding from the time series data using lagged coordinates so that the library vectors  $X_t, \{x_t, x_{t-\tau}, x_{t-2\tau} \dots x_{t-(E-1)\tau}\}$  and prediction vectors  $Y_t, \{y_t, y_{t-\tau}, y_{t-2\tau} \dots y_{t-(E-1)\tau}\}$  are points in  $E$ -dimensional state space (where  $\tau=1$ , and  $t \in [1, 2, \dots, n]$ ). We then choose the  $E+1$  neighbors of  $Y_t$  from the library set  $X$  that form the vertices of the smallest simplex that contains  $Y_t$  as an interior point. The forecast is based on how these nearby library points (domain simplex) move forward in time (range simplex). Again, the key is in the definition of “nearby” which depends critically on the dimension of the embedding. Thus, the domain simplex is projected forward into its range by incrementing the time index of

each neighbor in the domain simplex by 1. The forecast is simply a weighted average of the values given by the range simplex. The weights depend exponentially on proximity of the library vectors in the domain simplex to the predictee vector  $Y_t$  (illustrated in Supplementary Figure 1). For this study an exploratory series of embedding dimensions ( $E$ ) ranging from 1 to 20 (or higher) are used to evaluate the prediction, and the best  $E$  is chosen based on prediction skill. This embedding is then used in the S-map procedure.

### **S-map algorithm**

S-maps are an extension of standard linear autoregressive models in which the coefficients depend on the location of the predictee  $Y_t$  in an  $E$ -dimensional embedding. New coefficients are recalculated (from the library set  $X$ ) by singular value decomposition (SVD) for each new prediction. In this calculation, the weight given to each vector in the library depends on how close that vector  $X_i$  is to the predictee  $Y_t$ . The extent of this weighting is determined by the parameter  $\theta$ .

As above, we generate an  $E$ -dimensional embedding from points in the library using lagged coordinates to obtain an embedded time series with vectors  $X_t \in R^{E+1}$ , where  $X_t(0) = I$  is the constant term in the solution of Eq. (S2) below. Let the time series observation in the prediction set  $T_p$  time steps forward be  $Y_{t+T_p}(I) = Y(t)$ .

Then the forecast for  $Y(t)$  is 
$$\sum_{j=0}^m C_t(j) X_t(j) . \quad (S1)$$

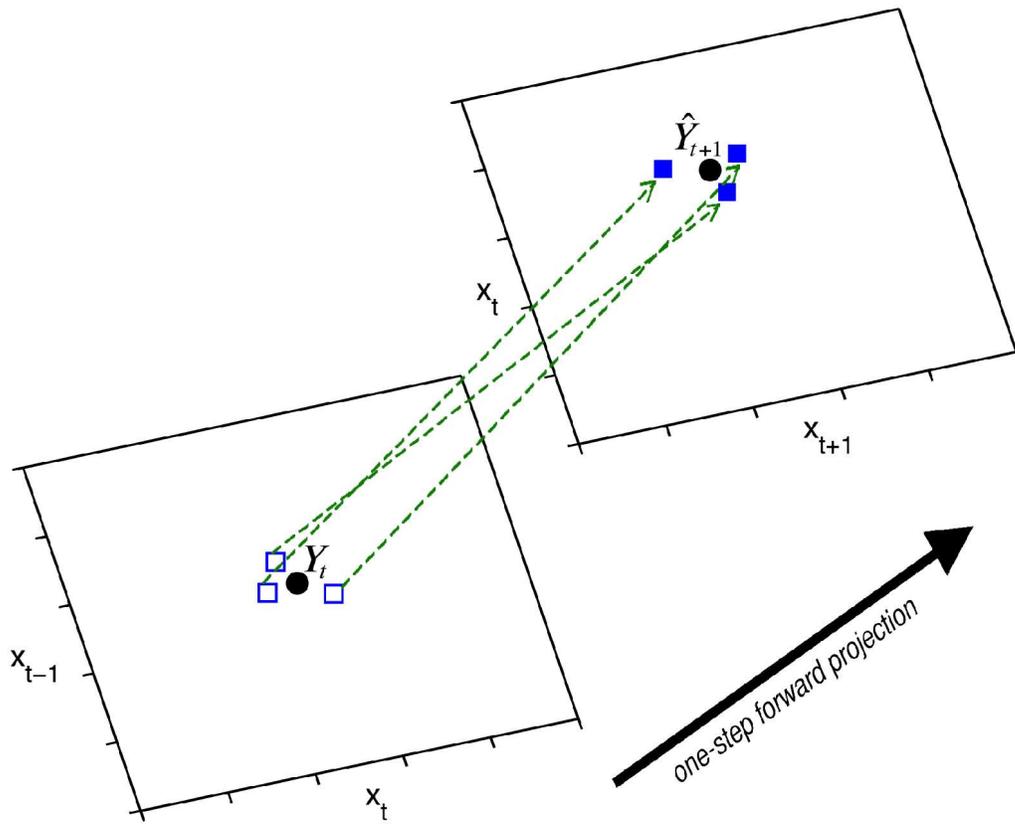
For our analysis, we chose  $T_p = I$ . For each  $E$ -dimensional predictee vector  $Y_t$ ,  $C$  is solved by SVD using the library set as follows:

$$B = AC, \quad (S2)$$

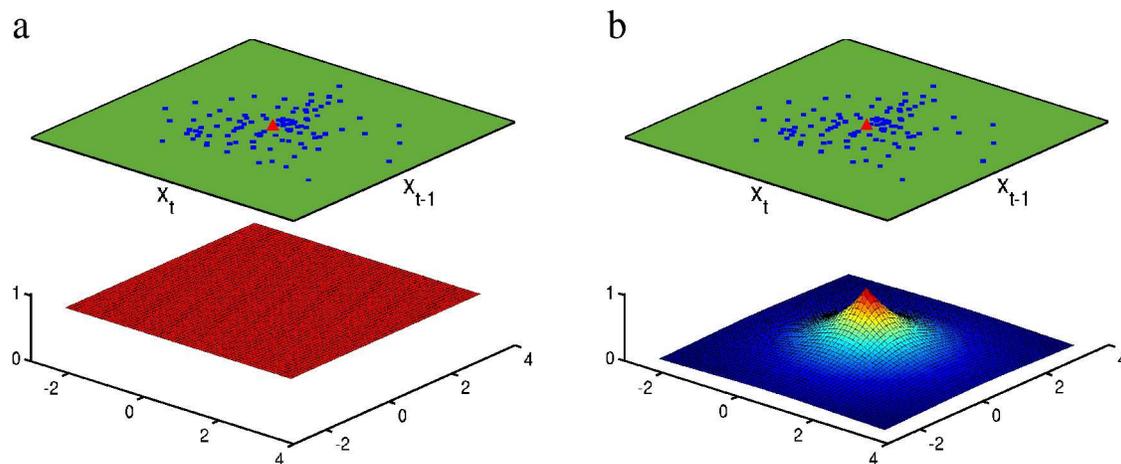
where  $B_i = w(\|X_i - Y_t\|) Y_i$ ,  $A_{ij} = w(\|X_i - X_t\|) X_i(j)$ , and  $w(d) = e^{-\theta d / \bar{d}}$

$\theta \geq 0$ ,  $d_{it}$  is the distance between  $Y_t$  and the  $i^{\text{th}}$  neighbor vector  $X_i$  in the library embedding, and the scale vector,  $\bar{d}$ , is the average distance between neighbors in the library. Note that

$A$  has dimensions  $n_i * (m + 1)$ , where  $n$  = size of the library. Again, a different map is generated for each forecast, with the weightings in each map depending on the location of the predictee in the  $E$ -dimensional state space. This weighting procedure is governed by the tuning parameter  $\theta$ , where  $\theta = 0$  gives a global linear map, and increasing values of  $\theta$  give increasingly local or nonlinear mappings. Note that when  $\theta = 0$ , all vectors are weighted equally so a single (global) linear map can be used for all predictions. In the case where  $\theta > 0$ , vectors closest to the predictee in state-space are weighted more heavily in the SVD solution. Such forecasts emphasize local information in the library set, and are therefore nonlinear. The method is illustrated in the Supplementary Figure 2 for a 2-dimensional embedding. The upper planes (in green) show a predictee (red triangle) and the library vectors (blue squares). The lower surfaces illustrate the weighting functions for a linear map (Supplementary Figure 2a) and a nonlinear map (Supplementary Figure 2b). The time series reflects nonlinear dynamics when nonlinear mappings (Supplementary Figure 2b) outperform the corresponding linear map (Supplementary Figure 2a) in out-of-sample forecasts.



**Supplementary Figure 1.** An example illustrating simplex projection for a time series embedded in a two dimensional ( $E = 2$ ) lagged-coordinate space. The figure shows a one-step forward forecast using nearby neighbors from the library. The predictee  $Y_t$  is a two dimensional vector formed from points in the prediction set. The three ( $E+1$ ) nearest neighbors ( $\square$ ) from the library set that form the domain simplex are projected one step forward to yield the range simplex ( $\blacksquare$ ). The forecast ( $\hat{Y}_{t+1}$ ) is a weighted average of the values given by the range simplex ( $\blacksquare$ ) with weights depending exponentially on proximity of the neighboring domain vectors ( $\square$ ) to the predictee  $Y_t$ .



**Supplementary Figure 2.** Examples illustrating the S-map procedure for a linear map (a) and a nonlinear map (b) with embedding dimension equal to two. The upper plane in each frame represents the lag coordinate embedding of the library file where  $E = 2$ . The lower panel is a geometric representation of the weighting function  $w(d) = e^{-\theta d|\bar{d}|}$ , where  $\theta = 0$  (left panel) and  $\theta > 0$  (right panel). In the linear map, all library vectors (blue squares) are weighted equally. In the nonlinear map, the points closest to the predictee (red triangle) contribute most heavily to the forecast.

## References

1. Sugihara, G. & May, R. M. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time-series. *Nature* **344**, 734-741 (1990).
2. Sugihara, G. Nonlinear forecasting for the classification of natural time-series. *Philos. T. Roy. Soc. A* **348**, 477-495 (1994).
3. Crutchfield, J. P. Prediction and stability in classical mechanics. Bachelor's Thesis, University of California, Santa Cruz (1979).
4. Takens, F. Detecting strange attractors in turbulence. *Lect. Notes Math.* **898**, 366-381 (1981).

5. Farmer, J. D. & Sidorowich, J. J. in *Evolution, learning and cognition* (ed. Lee., Y. C.)  
277-304 (World Scientific Press, New York, 1989).