**Available online at www.fooddynamics.org**

**Int. J. Food System Dynamics 3 (2010) 184-193**

INTERNATIONAL JOURNAL ON FOOD SYSTEM DYNAMICS

# Phase Space Reconstruction from Economic Time Series Data: Improving Models of Complex Real-World Dynamic Systems

Ray Huffaker

*University of Florida, USA*
*rhuffaker@ufl.edu*

*Received March 2010, accepted September 2010, available online October 2010*

## ABSTRACT

Failure of economic models to anticipate the global financial crisis illustrates the need for modeling to better capture complex real-world dynamics. Conventional models—in which economic variables evolve toward equilibria or fluctuate about equilibria in response to exogenous random shocks—are ill-equipped to portray complex real-world dynamics in which economic variables may cycle aperiodically along low-dimensional 'strange attractors'. We present a method developed in the physics literature—'phase space reconstruction'—that reconstructs strange attractors present in real-world dynamical systems using time series data on a single variable. Phase space reconstruction provides pictures of real-world dynamics that can guide model specification.

*Keywords: phase space reconstruction, time series data, economic dynamics*

## 1 Introduction

Economic time series data may conceal complex recurring patterns. For example, system variables may cycle aperiodically along low-dimensional 'strange attractors'. Strange attractors are difficult to detect directly from time series data—which may appear mathematically random. Consider, for example, the time series data shown in Figures 1(a)-(c). The data in Figures 1(a) and (b) are generated by deterministic dynamic models. In particular, Figure 1 (a) plots one of the three variables from the Ikeda dynamical system, and system 1(b) plots one of the three variables from the Henon dynamical system. Alternatively, the time series in Figure 1(c) is generated as white noise. The data from the Ikeda and Henon dynamical systems are not easily distinguished from the white noise—their underlying deterministic structures are concealed.

'Phase space reconstruction' is a method for uncloaking deterministic structure in time series data. Specifically, it reconstructs attractors present in real-world dynamical systems using time series data on a single variable (Broomhead, King, 1985; Schaffer, Kott, 1985; Kott et al, 1988; Williams, 1997). The underlying intuition is that the dynamics of the entire system are embedded in the history of each variable. The second panels in Figures 1(a) and (b) show the strange attractors reconstructed for the Ikeda and Henon systems using this method. The value of phase space reconstruction is to provide pictures of the systematic patterns occurring in real-world dynamic systems that can be used to guide model specification.

We begin by introducing the concept of phase space attractors within the context of a dynamic ISLM model. We next demonstrate how phase space reconstruction faithfully reproduces one of the model's attractors. Finally, we discuss how phase space reconstruction fits into a more general 'diagnostic' modeling approach that relies on historical data to guide the deterministic formulation of theoretical dynamical models. As an example of diagnostic modeling, we examine how the dynamic ISLM model can be adapted to generate behavior approximating an aperiodic attractor reconstructed from time series data on real-world interest rates.

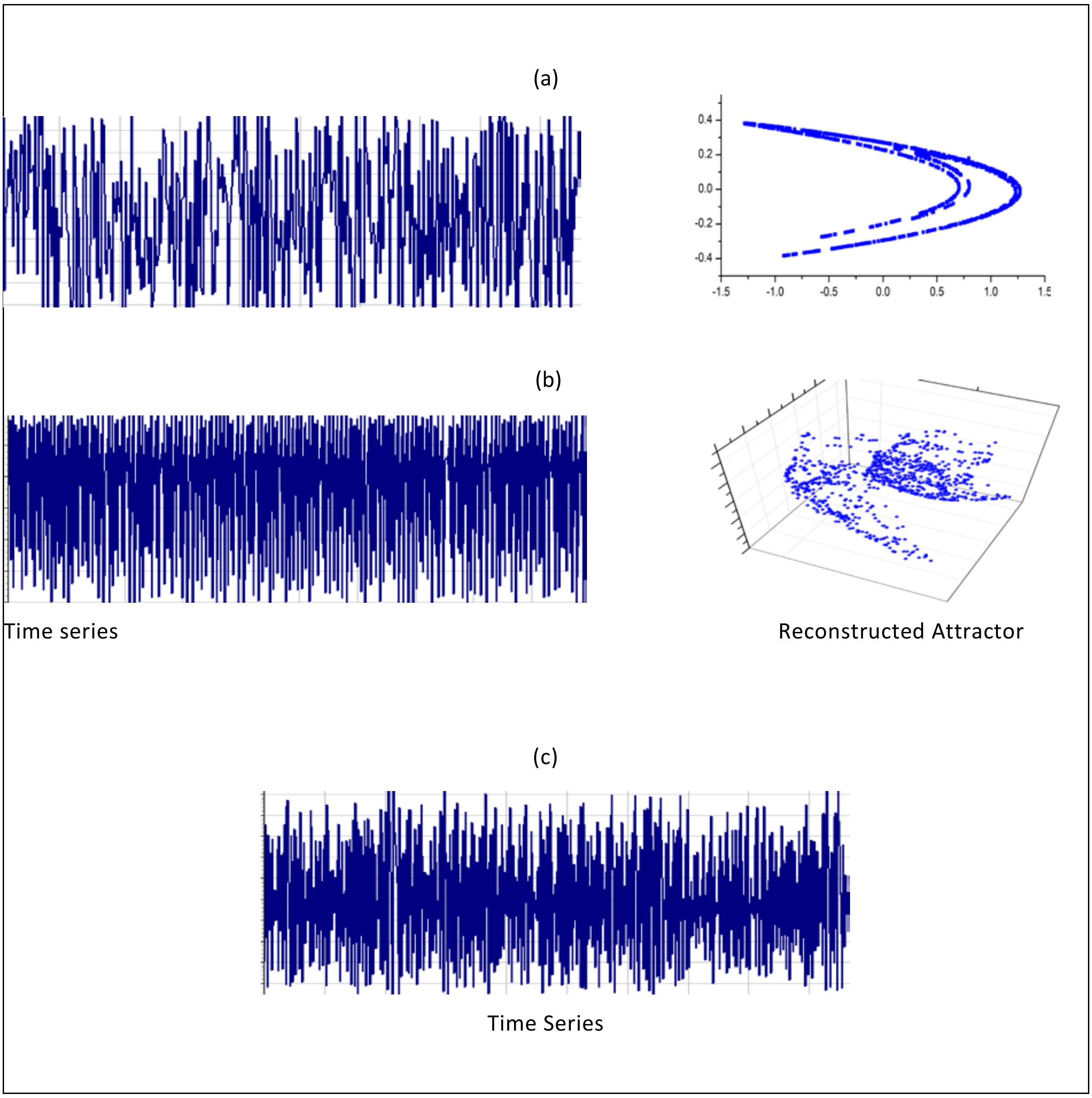


**Figure 1.**
Time series data and reconstructed attractors: (a) Ikeda Map, (b) Henon Map, (c) White noise

## 2 A phase space attractor

Consider the following dynamic ISLM model (Shone, 2002, chapter 10):

$$\frac{dr}{dt} = \beta \underbrace{\left[ m^d \left( r, y; \Omega_1 \right) - m_0 \right]}_{\text{excess demand in money market}} \qquad (1a)$$

$$\frac{dy}{dt} = \alpha \underbrace{\left[ e \left( r, y; \Omega_2 \right) - y \right]}_{\text{excess demand in goods market}} \qquad (1b)$$

where $r$ is the nominal interest rate, $y$ is real income, $m^d(r, y)$ measures the demand for real money balances, $m_0$ is the exogenous real money supply, $e(r, y)$ measures real expenditures, and underlying parameters $\Omega_1$ and $\Omega_2$ are constant. The model exhibits a wide range of dynamic behaviors depending on how rapidly the money and goods markets adjust to excess demand. One possible behavior is for interest rates and incomes to oscillate toward an equilibrium that balances demand and supply in each market [Figures 2(a),(b)]. Figure 2(c) shows the solution in phase space obtained by plotting $r(t)$, and $y(t)$ for each time $t$, and connecting the points (a 'scatter' diagram). System dynamics are characterized by a 'stable-focus' attractor.

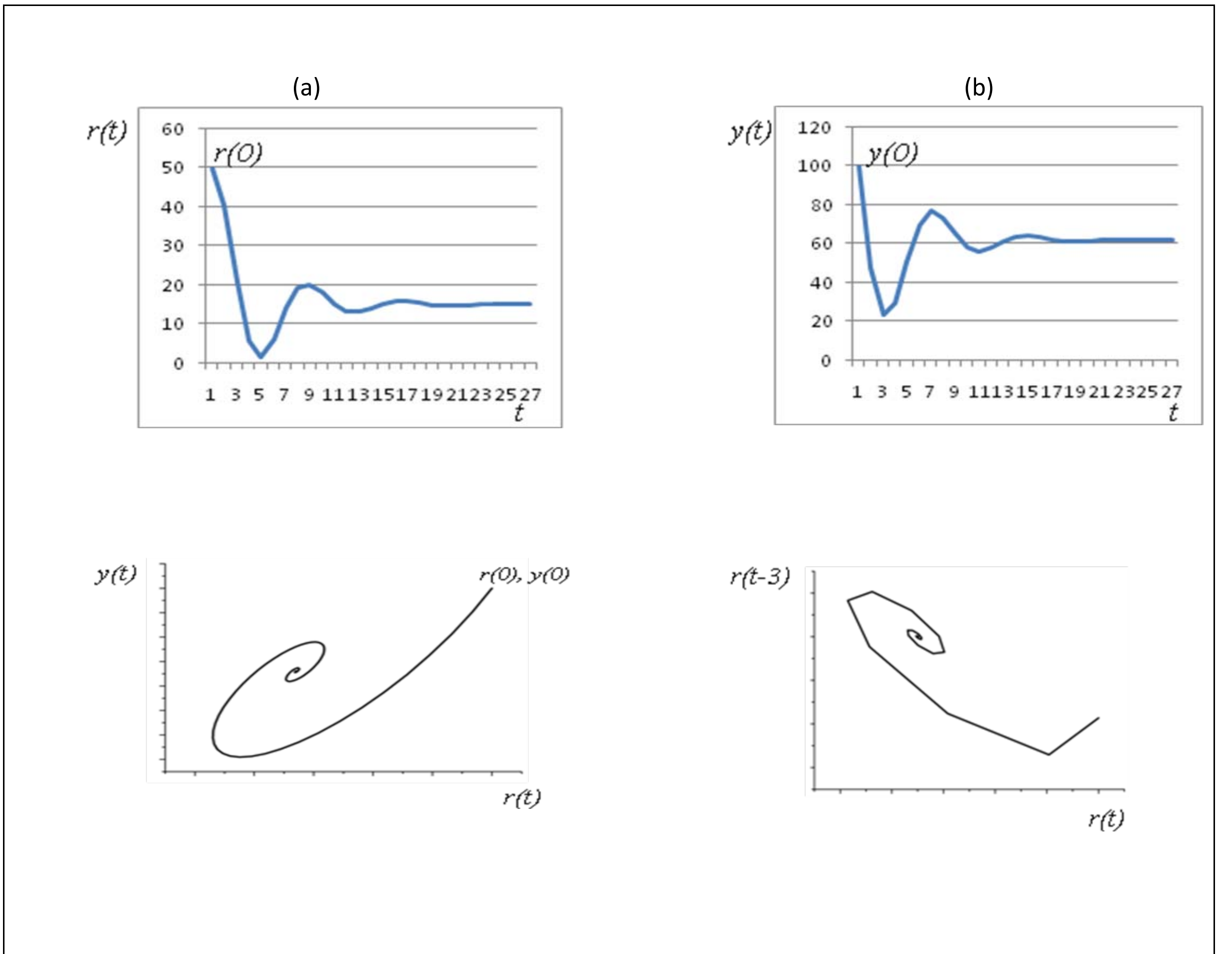


**Figure 2.**
The spiral-node atrractor in the dynamic ISLM model: (a) Time series of interest rates; (b) Time series of incomes; (c) Phase diagram solution where $r(0)$, $y(0)$ are values in the initial time period; (d) Attractor reconstruction. We solved the model numerically in Excel using an Euler approximation

## 3 Attractor reconstruction

To demonstrate how attractor reconstruction works, we reconstruct the stable-focus attractor in the ISLM model using only one of the system variables, in this case, $r(t)$.

### 3.1 Time-delay embedding procedure

The time-delay embedding procedure of attractor reconstruction is firmly rooted in mathematical topology theory thanks to Takens' theorems (1980). In general, suppose that the solution of a real world dynamical system asymptotically contracts onto an unobservable attractor contained within a lower-dimensional submanifold of unknown dimension *N*. The observed time series, call it *x(t)*, lies on this attractor. The time-delay embedding procedure uses *x(t)* to 'reconstruct' the attractor with a 'scatter' plot of the following lagged coordinate vectors: $x(t), x(t-\tau), x(t-2\tau), ..., x(t-(M-1)\tau)$ where $\tau$ is the length of the lag, and *M* is the number of lagged coordinate vectors (called the 'embedding dimension').

For example, assume that we observe the time series $x(t) = (4,2,7,4,9,6,10,3)^T$ [*T* is the transpose]. Assume a lag length $\tau = 2$ and an embedding dimension *M* = 3. This generates the following lagged coordinate vectors:

$$x(t) = \begin{pmatrix} 4 \\ 2 \\ 7 \\ 4 \\ 9 \\ 6 \\ 10 \\ 3 \end{pmatrix}, \quad x(t-2) = \begin{pmatrix} 7 \\ 4 \\ 9 \\ 6 \\ 10 \\ 3 \end{pmatrix}, \quad x(t-4) = \begin{pmatrix} 9 \\ 6 \\ 10 \\ 3 \end{pmatrix}$$

We next detect whether an underlying attractor exists with a 'scatter plot' of the elements in the first four rows of the lagged coordinate vectors. The first point on the attractor is composed of the first elements of each vector (4,7,9). Similarly, the second point is (2,4,6); the third point is (7,9,10); and the fourth point is (4,6,3). Elements in the remaining rows are lost in the lagging process. The reconstruction for this example is shown in Figure 3.

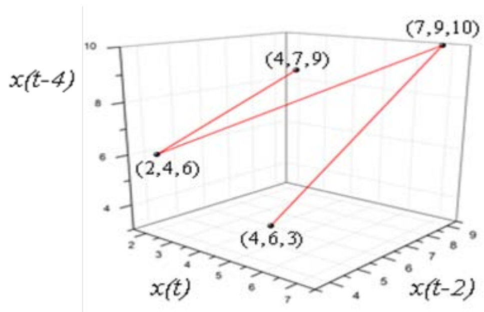


**Figure 3.**
Phase diagram for lagged coordinate vectors
$x(t) = (4,2,7,4,9,6,10,3)^T$, $x(t-2) = (7,4,9,6,10,3)^T$, $x(t-4) = (9,6,10,3)^T$.

Takens' theorems give the following sufficient condition for a reconstructed attractor to have the same dynamical properties as the unobservable attractor from the real-world system: $M \geq (2N+1)$. In words, the embedding dimension *M* has to be at least as large as twice the dimension of the real-world solution manifold *N* plus 1. Three notes are in order: (1) Since this is a sufficient condition, fewer embedding dimensions may be necessary; (2) Takens' theorems do not imply that the reconstructed attractor is identical to the real-world attractor—just that it has the same dynamical (topological) properties; and (3) Since *N* is unobservable, the embedding dimension *M* cannot be directly calculated from the sufficient condition.

### 3.2 Selecting lag length ( $\tau$ ) and embedding dimension (M)

Selecting the reconstruction parameters $\tau$ and *M* is like focusing a camera—some pairs result in much clearer reconstructed attractors than others. A 'brute force' approach would simply search over a number of combinations until an attractor comes into focus. Fortunately, the literature suggests various diagnostic procedures to narrow the search (Williams, 1997, pp. 275-284).

The lag length $\tau$ is conventionally chosen as the first minimum of the 'mutual information function'—a probabilistic measure of the extent to which $x(t+\tau)$ is related to $x(t)$ at a given $\tau$ . The rationale behind the approach is to introduce the needed statistical independence between successive lagged values. For example, the problem with overly short lags is that successive values contain redundant mutual information. Consequently, the reconstructed attractor is restricted to rest on a $45^{\circ}$ diagonal in the embedding space. Alternatively, the problem with overly long lags is that successive lagged values contain too little mutual information to reconstruct the attractor.

The embedding dimension *M* is conventionally chosen using the 'false nearest neighbors' method. This method measures the percentage of close neighboring points in a given dimension that remain so in the next highest dimension. The minimum embedding dimension capable of containing the reconstructed attractor is that for which the percentage of false nearest neighbors drops to zero for a given tolerance level. For example, assume that an overcoat is crammed into small box. If you gaze into the top of the box, you see the material, but can't tell if it's part of a coat or just a lump of cloth. Take the coat out of the box, unfurl it into its proper 'dimension', and you see plainly that it's a coat. When the coat is wadded up, the end of the sleeve might be touching the collar. When the coat is unfurled, the end of the sleeve is distant from the collar. Thus, when the coat is wadded up, the end of the sleeve and the collar are false nearest neighbors. When the coat is completely unfurled, these two points (and all other points on the coat) no longer move away from each other—the percentage of false nearest neighbors is zero. Similarly, an attractor embedded into too few dimensions does not have room to fully express itself. Alternatively, an attractor embedded into too many dimensions goes out of focus.

Software is available to compute lag length and embedding dimension. For example, Visual Recurrence Analysis (VRA) software is free for downloading, use, copying, and distribution at http://visual-recurrence-analysis.software.informer.com/ . We use it below to compute lag length and embedding dimension.

### 3.3 What constitutes a 'good' reconstruction?

The test of a 'good' reconstruction is largely visual. Does the reconstructed attractor exhibit apparent structure? Recent research in *recurrence quantification analysis* is adding objective numerical measures to identify recurring patterns in data (see, e.g., Webber, Zbilut, 2004). Moreover, inferential statistics can be developed to test whether a reconstruction from a given time series is significantly different from a reconstruction from a randomized (surrogate) data set artificially constructed to mimic the same statistical properties as the given time series (i.e., probability distribution, power spectrum, and correlation function) (Sprott, 2003, pp. 232-235).

### 3.4 Reconstructing the ISLM stable-focus attractor

The stable-focus attractor reconstructed from the single time series $r(t)$ generated in the dynamic ISLM model is shown in Figure 2(d). The first minimum of the mutual information function occurs at lag length $\tau = 3$ [Figure 4(a)]. The embedding dimension for which the percentage of false nearest neighbors drops effectively to zero is *M* = 2 [Figure 4(b)]. The attractor reconstructed with these parameters (Figure 2d) faithfully reproduces the stable-spiral attractor from the original system (Figure 2c).

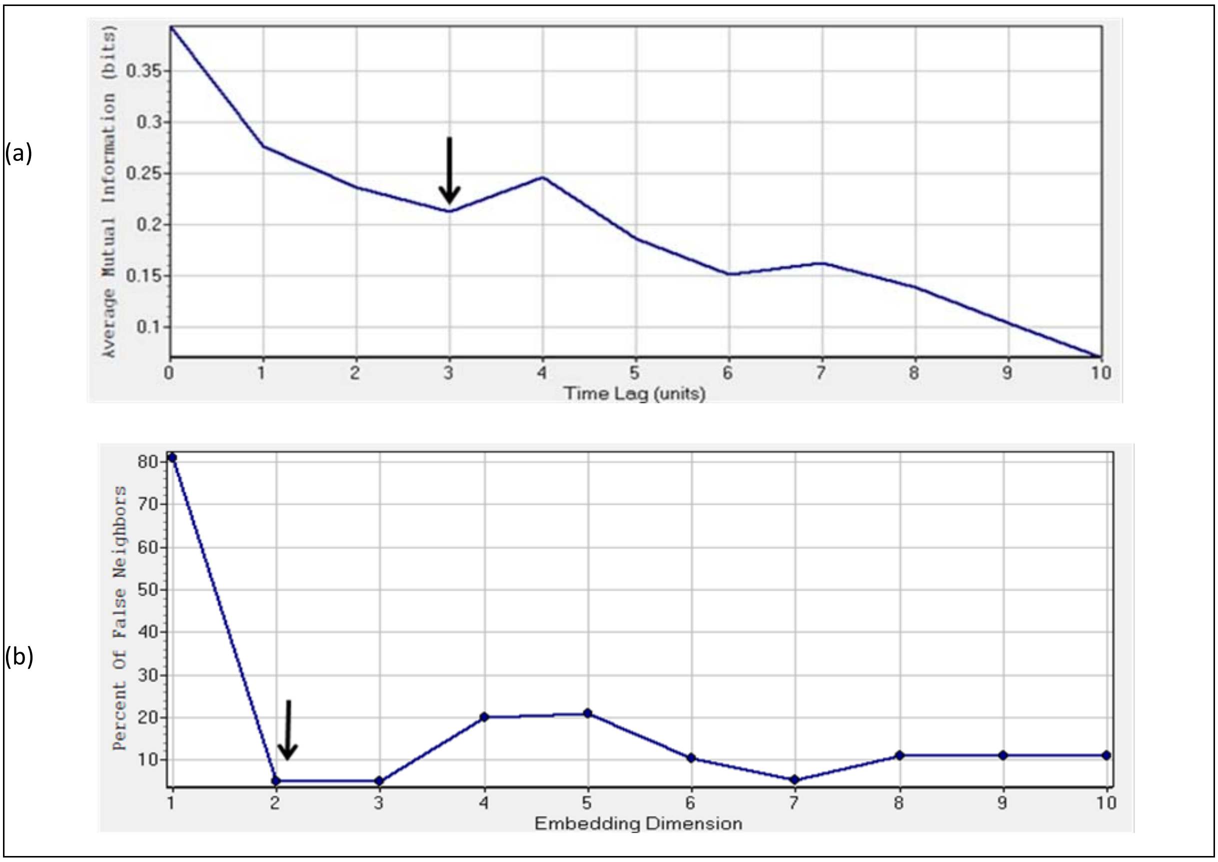


**Figure 4.**
Phase Space Reconstruction Parameters: (a) Mutual information function; (b) False nearest neighbors test

### 3.5 Problems in application: Small and noisy data sets

Attractor reconstruction may fail for a couple of reasons (Williams, 1997, pp. 274-275). First, the real world dynamical system may not evolve toward a low-dimensional attractor that can be visualized in two or three dimensions. In other words, there is no low-dimensional attractor to reconstruct, and the false nearest neighbors test gives a minimum embedding dimension greater than what can be visualized (i.e., > 3 dimensions). Second, attractor reconstruction has difficulty with small and noisy data sets—common characteristics of economic time series. Principal components analysis may remedy both problems (Williams, 1997, pp. 284-287).

Principal components analysis filters noise from data without distorting the underlying dynamical properties. As a result, the attractor may become more clearly focused in fewer embedding dimensions. A brief outline of the procedural steps follows (Theil, 1971):

1. Construct the 'embedded date' matrix $X$ from the observed time series $x(t)$ with $n$ observations. Set lag length to $\tau = 1$ and embedding dimension to the level $M$ computed from the false nearest neighbors test:

$$\underset{\substack{[n-\tau(M-1)]\times M \\ \text{matrix in rectangle}}}{X} = \begin{bmatrix} x(t) & x(t-1) & x(t-2) & x(t-3) & \bullet & \bullet & x(t-M+1) \\ x(t-1) & x(t-2) & x(t-3) & x(t-4) & \bullet & \bullet & x(t-M) \\ x(t-2) & x(t-3) & x(t-4) & x(t-5) & \bullet & \bullet & x(t-M-1) \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & x(t-n+1) \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \\ \bullet & \bullet & \bullet & \bullet & \bullet & & \\ \bullet & \bullet & \bullet & x(t-n+1) & & & \\ \bullet & \bullet & x(t-n+1) & & & & \\ \bullet & x(t-n+1) & & & & & \\ x(t-n+1) & & & & & & \end{bmatrix}$$

2. Compute the eigensystem of $X'X$ : $\left(X'X - \lambda_i I\right)a_i = 0$ where $\lambda_i\ (i = 1, M)$ are eigenvalues and $a_i$ are the associated eigenvectors.

3. Compute the principal components: $\underset{M\times 1}{p_i} = \frac{1}{\lambda_i} X a_i \quad i = 1, M$

4. Compute the proportion of the variance of the embedded data accounted for by each principal component: $V = \frac{\lambda_i}{Trace(X'X)}$.

5. Reconstruct phase space with a scatter plot of the principal components matrix: $P = \left(p_1, p_2, p_3\right)$ where $p_1, p_2, p_3$ are the principal components associated with the three largest eigenvalues and corresponding eigenvectors.

This procedure filters noise by using the principal components explaining the greatest proportion of the variance in the embedded data (the principal components associated with smaller eigenvalues contain more noise). The procedure is better justified as the proportion *V* associated with the three largest eigenvalues approaches 100%. Reducing the noise in the data increases the possibility that a reconstructed attractor can be visualized in two or three dimensions with a scatter plot of the principal components associated with the two or three largest eigenvalues

## 4 Diagnostic modeling using real-world interest rate data

Diagnostic modeling is a two-pronged procedure for guiding and testing the specification of dynamic models (Figure 5). The left-ward prong relies on reconstruction of a real-world attractor from historic data on a single observed variable. The right-ward prong formulates a dynamic model that attempts to explain the dynamic forces characterizing the reconstructed attractor. A useful specification test for the model's attractor is whether it is visually similar to the real-world attractor reconstructed from the historic data.

We now consider an example of diagnostic modeling. Figure 6 (upper left corner) shows a plot of time series data on U.S. interest rates (Moody's Seasoned AAA Corporate Bond Yields, January 1, 1919 to January 5, 2010, Board of Governors of the Federal Reserve System). Directly below is the graph of the attractor reconstructed from these data using the time-delay method and principle components. The attractor exhibits aperiodic cycling that is more complex than the stable-focus attractor (and other possible attractors) generated by the specification of the dynamic ISLM model in equations 1(a),(b). This indicates that the model should be re-specified so that it can generate an aperiodic attractor. For purposes of this example, we re-specified the model by making the marginal propensity to consume in the real expenditure function seasonally forced, i.e., an explicit function of time: $e(r, y, t)$ (upper right corner of Figure 6). Specifically, we modeled the controlling parameter as a periodic cosine function. Similar to the historic data, the attractor generated by the re-specified model exhibits aperiodic cycling (lower right corner of Figure 6).

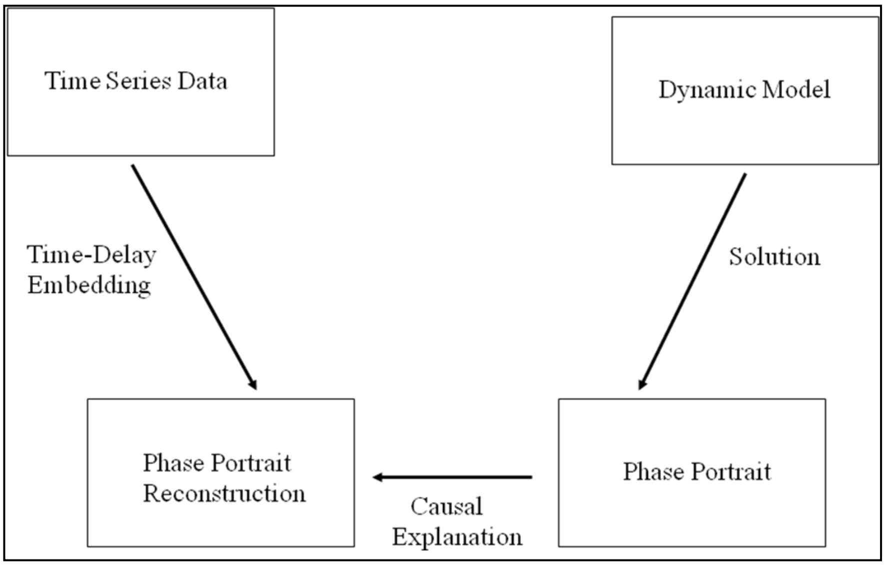


**Figure 5.**
Diagnostic Modeling

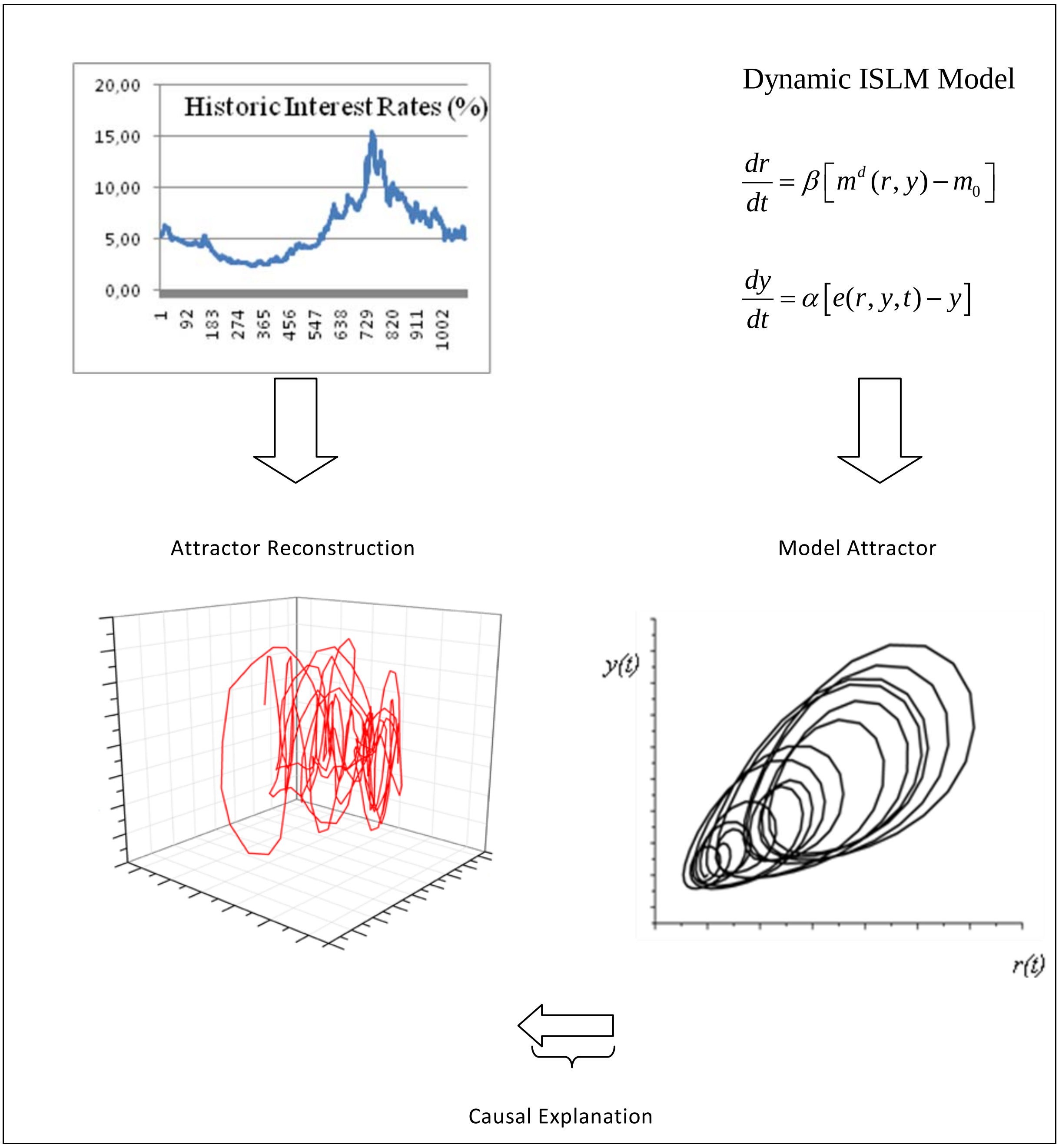


**Figure 6.**
Example of diagnostic modeling with the dynamic IS-LM model: To simulate the aperiodic cycling of the attractor reconstructed fromthe interest rate data (Moody's seasoned AAA Corporate Bond Yield), the marginal propensity to consumer in the dynamic IS-LM model was assumed to adjust seasonally

## 5. Concluding comments

Phase space reconstruction, along with emerging methods in recurrence quantification analysis, is a potentially useful method for detecting dynamic structure in apparently random historical data. As a part of diagnostic modeling, it can improve the formulation and performance of dynamical models.

## References

Broomhead, D.S., King, G.P. (1985). Extracting qualitative dynamics from experimental data. Physica 20D: 217-236.

Kot, M., Schaffer, W., Truty, G., Graser, D., Olsen, L. (1988). Changing criteria for imposing order. *Ecological Modeling* 43: 75-110.

Schaffer, W., Kot, M. (1985). Nearly one dimensional dynamics in an epidemic. J. Theor. Biol. 112: 403-427.

Shone, R. (2002). Economic Dynamics. Cambridge: Cambridge University Press.

Sprott, J. (2003). Chaos and Time-Series Analysis. New York: Oxford University Press.

Takens, F. (1980). Detecting strange attractors in turbulence. In: Rand, D., Young, L. (eds). Dynamical Systems and Turbulence (pp. 366-381). New York: Springer.

Theil, H. (1971). Principles of Econometrics. New York: John Wiley & Sons.

Weber, C. and Zbilut, J. (2004). Introduction to recurrence quantification analysis. In: C. Weber (ed). Recurrence Quantification Analysis of Nonlinear Dynamical Systems. New York: Springer-Verlag.

Williams, G. (1997). Chaos Theory Tamed (Chapter 19: Attractor Reconstruction). Washington D.C.: Joseph Henry Press.