HHS Public Access

Author manuscript

Struct Equ Modeling. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Struct Equ Modeling. 2020 ; 27(2): 202–218. doi:10.1080/10705511.2019.1641816.

Constrained Fourth Order Latent Differential Equation Reduces Parameter Estimation Bias for Damped Linear Oscillator Models

Steven M. Boker, **Robert G. Moulder**, **Gustav R. Sjobeck**

Department of Psychology, The University of Virginia, Charlottesville, VA 22903

Abstract

Second order linear differential equations can be used as models for regulation since under a range of parameter values they can account for return to equilibrium as well as potential oscillations in regulated variables. One method that can estimate parameters of these equations from intensive time series data is the method of Latent Differential Equations (LDE). However, the LDE method can exhibit bias in its parameters if the dimension of the time delay embedding and thus the width of the convolution kernel is not chosen wisely. This article presents a simulation study showing that a constrained fourth order Latent Differential Equation (FOLDE) model for the second order system almost completely eliminates bias as long as the width of the convolution kernel is less than two thirds the period of oscillations in the data. The FOLDE model adds two degrees of freedom over the standard LDE model but significantly improves model fit.

Introduction

One of the fundamental principals of life is that it undergoes change. Mathematically, change as an outcome has long been represented by differential equations (see, e.g., Butcher, 2016; Hotelling, 1927; Hubbard & West, 1991; Laplace, 1779). Dynamical systems are a useful way to think about differential equations as applied to many types of biological (Glass & Mackey, 1988), psychological (Nesselroade, 2002; Boker, 2002), social (Laurenceau, Barrett, & Pietromonaco, 1998), emotional (Butler, 2017), cognitive (Port & vanGelder, 1995), and developmental (Smith & Thelen, 1993) phenomena. In the past two decades there has been rapidly increasing interest in modeling psychological systems using differential equations. A few of the many examples in recent literature are developmental change (Gerstorf, Ram, Röcke, Lindenberger, & Smith, 2008), fluctuations and self-regulation of mood (Russell et al., 2011), resilience in the face of adversity (Bisconti, Bergeman, & Boker, 2004; Montpetit, Bergeman, Deboeck, Tiberio, & Boker, 2010; Ong, Bergeman, Bisconti, & Wallace, 2006), day to day changes in hunger and appetite (Klump, Keel, Burt, et al., 2013; Racine et al., 2013), monthly change in hormones (Klump, Keel, Kashy, et al., 2013), age-related change in brain functioning and cognitive performance (Gerstorf et al., 2015; Raz et al., 2005; Salthouse & Davis, 2006), and dyadic-regulation of romantic couples (Boker & Laurenceau, 2007; Laurenceau et al., 1998). Fitting differential equation models to data using Structural Equation Modeling (SEM) has become increasingly popular (e.g,

Correspondence may be addressed to Steven M. Boker, Department of Psychology, The University of Virginia, PO Box 400400, Charlottesville, VA 22903, USA; boker@virginia.edu; or browsers pointed to http://people.virginia.edu/˜smb3u.

Benson, Ram, & Stifter, 2018; Bisconti, Bergeman, & Boker, 2006; Boker, Neale, & Klump, 2014; Driver & Voelkle, 2018; Hurley, Fink, & Janata, 2018; Montpetit et al., 2010; Morales et al., 2017; Oud & Jansen, 2000).

When one considers individual differences from the perspective of dynamical systems models, it is not the measurements themselves that comprise important differences between people. We like to think of this in the context of the questions “what”, “where”, “why”, and “how”. Traditional cross-sectional individual differences designs ask the question, *what* is the difference in the observed scores? Dynamical systems theory encourages us to think about the process of change within each individual. Then individual differences can be separated into three main categories (Boker, 2013; Boker, Staples, & Hu, 2016). First, we must consider individual differences in equilibria, that is to say, *where* is the natural homeostatic condition of each individual and *how* do people differ in the state to which they tend in the absence of contextual pressure (or perturbation in the language of dynamical systems)? So the first category of dynamical systems question asks, *where* do an individual’s observations tend to return to when perturbed away from equilibrium? Second, we must find a model for the return to equilibrium, in other words, *why* does a person return to equilibrium? There might be individual differences in the model by which the dynamic return to equilibrium occurs. Third, we are interested in the question, *how* does an individual return to equilibrium? The answer to this question is manifest in the parameters of the differential equation model. Even if the same model can account for everyone in the population, individuals may still differ in the parameters of that model.

In order to test a dynamical systems theory for regulation within individuals we need time-intensive measurements of each individual. Each person in the sample must be measured a sufficient number of times in order to have power to be able to reject a hypothesis that everyone is the same. In addition, when the models are fit to the data, if some individuals’ dynamic return to equilibrium happens relatively rapidly and for others the return is slower, we must have methods that do not induce bias towards a value that is an artifact of the design of the experiment. Unfortunately, when fitting traditional discrete time auto- and cross-regressive SEM models the measurement interval is nonlinearly confounded with the estimated parameters (e.g., Montfort, Oud, & Voelkle, 2018; Oud, 2007). Bias in parameter estimates will be a function of their values and thus individual differences in parameter estimates may either be inflated or reduced depending on the data, the measurement interval, and the model. Continuous time models (i.e., differential equations models) are used so as to avoid the trap of having the measurement interval confounded with the estimated parameter.

Differential equations express dynamics as predictive relationships between variables of continuous time change (see Hubbard & West, 1991; Kaplan & Glass, 1995, for useful introductions). Suppose that an individual has been measured on a variable, X, on p occasions each separated by an interval of time, Δt, so that $X = \left\{x_0, x_{\Delta t}, x_{2\Delta t}, x_{3\Delta t}, \ldots, x_{(p-1)\Delta t}\right\}$. Now consider an individual’s score at a chosen time t (denoted x_t) and its derivatives, e.g., how fast is the score changing at time t (denoted as $dx(t)/dt$ or $\dot{x}(t)$), and how fast is this change accelerating or decelerating at the same moment of time t (denoted as $d^2x(t)/dt^2$ or $\ddot{x}(t)$). If there are predictable covariance relationships between the score and its instantaneous

derivatives then we can define a differential equation composed of predictive relations between the score and its derivatives and this differential equation defines the continuous time dynamic of the system.

The field of dynamical systems analysis is large and growing and it is beyond the scope of this article to provide a substantial introduction to the field. There are many sources (e.g., Montfort et al., 2018; Smith & Thelen, 1993; Thompson & Stewart, 1986) that may help the interested reader to become more familiar with the terms and uses of dynamical systems models. The current article will focus on second order linear differential equations that can provide an appealing model for self-regulating psychological processes.

Second order linear differential equations

Second order linear differential equations are those in which a target variable and its first and second derivatives are linearly related. This may be expressed by a regression equation where the second derivative is an outcome variable predicted by the variable and the first derivative such that,

$$d^2x(t)/dt^2 = \eta x(t) + \zeta dx(t)/dt + e(t) \quad (1)$$

where η is often termed the frequency parameter and ζ the damping parameter. The residual term, $e(t)$, can be interpreted as the degree to which the second derivative is not predicted by the displacement from equilibrium and first derivative. But this residual term may include stochastic (non-deterministic) variance as well as deterministic variance that does not conform to this linear model. We will return to this problem when we describe the constrained fourth order LDE.

A stable damped linear oscillator (DLO) occurs when $\eta < 0$ and $\eta + \zeta^2/4 < 0$. In this case, oscillation will occur with a period $\lambda = \frac{2\pi}{\sqrt{-(\eta + \zeta^2/4)}}$. If $\zeta < 0$ then if the system is perturbed away from its equilibrium, the system will over time tend towards its point equilibrium value along a negative exponential curve. If $\zeta > 0$ then the system is unstable and any perturbation away from equilibrium will induce exponential increase in displacement from equilibrium. Also, if $\eta > 0$ then the system is unstable and will exponentially diverge from equilibrium. Finally, if $\eta = 0$ or $\eta + \zeta^2/4 > 0$ then the system behaves similarly to a first order system with exponential damping or divergence depending on the sign of ζ. The simulation presented here will only consider parameter values that result in a stable DLO, the conditions most often of interest in self-regulating systems.

The flexibility in the types of curves that the second order linear differential equation can approximate is one reason that the use of this equation has become popular: Individual differences in trajectories can be modeled with a single equation even when some individuals exhibit simple damping towards equilibrium and others exhibit oscillating trajectories. Another reason is that the second order differential equation model has appealing and intuitive interpretations for its parameters. If one is interested in variability analysis, the second order differential equation model separately estimates the variance of a variable's value (amplitude of fluctuations) and the variance of its first derivative (intuitively

the “roughness” of the fluctuation trajectories) which can be independent qualities of variability in the behavior of a system (Deboeck, Monpetit, Bergeman, & Boker, 2009). If one is interested in modeling self-regulating systems, then the η parameter can give an estimate of the average cycle time inherent in an individual’s self-regulation. The ζ parameter can give an estimate of overall return to equilibrium. The two parameters together can be interpreted as being related to resilience to adversity (Boker, Montpetit, Hunter, & Bergeman, 2010; Montpetit et al., 2010; Tarter & Vanyukov, 2002). There is increasing interest in using the second order linear (i.e., DLO) model for many types of human systems (e.g., Butler, 2017; Hurley et al., 2018; Morales et al., 2017).

Parameters of second order differential equations models can be estimated from data using two main approaches. The first approach is to analytically integrate the differential equation model and then use the resulting specific integral to predict a score at a known interval of time in the future based on scores in the immediate past using state space methods (see, e.g., Chow, Hamaker, Fujita, & Boker, 2009; Ho, Shumway, & Ombao, 2005; Petris & Petrone, 2011). The exact and approximate discrete methods (Bergstrom, 1966, 1984; Singer, 1993; Oud & Jansen, 2000; Oud, 2017) and latent Kalman filtering methods (Kalman, 1960; Chow, Ferrer, & Nesselroade, 2007; So, Ott, & Dayawansa, 1994) are examples of this approach. The second approach uses convolution filtering methods such as Latent Differential Equations (LDE, Boker, Neale, & Rausch, 2004; Boker et al., 2016; McKee, Rappaport, Boker, Moskowitz, & Neale, 2018) or Generalized Local Linear Approximation (GLLA Boker, Deboeck, Edler, & Keel, 2010) to estimate the parameters of the differential equation without integrating the differential equation. The current article will focus on an improvement to the LDE method of estimation of parameters for the second order linear differential equation.

LDE Advantages and Disadvantages

The LDE method of fitting a differential equation model to data is a variant of a class of operations known as *convolution filtering* (Levinson, 1946, 1947; Wiener, 1949), a technique that multiplies a filter with windowed samples of a time series in order to obtain a result with desired characteristics. This technique has been widely used in image processing (e.g., Gonzalez & Wintz, 1977; Ozaktas, Barshan, Mendlovic, & Onural, 1994), audio processing (e.g., Reilly & McGrath, 1995), and physics (e.g., Byerly, 1965). In particular, this technique has been applied to time series data to estimate derivatives (Savitzky & Golay, 1964). Time delay embedding is a closely associated technique that was developed for nonlinear dynamical systems attractor reconstruction (Ruelle & Takens, 1971; Sauer, Yorke, & Casdagli, 1991) and was proven to be able to capture the dynamical properties of a time series (Takens, 1985) even when that series was nonlinear and had no analytic solution. Efficient calculation of a convolution filter with a time series can be produced by a matrix multiplication of the filter with a time delay embedding of a time series as is used in the LDE method (see Boker et al., 2016, for an extended discussion).

There are advantages and disadvantages to the use of the LDE method for specifying and fitting differential equations to time series data. The first advantage is that the LDE method uses time delay embedding as its data representation, which in the context of second order

linear differential equations has been shown to increase precision of parameter estimates (Oertzen & Boker, 2010) and has surprising robustness to violations of time interval homogeneity (Boker, Tiberio, & Moulder, 2018).

The second advantage of using an LDE model is that the model is easily specified in commonly understood latent variable terms. This means that it becomes easier to consider many alternative models for the same system, specify parameter constraints, create coupled systems, and even specify nonlinear systems that can lead to bifurcation. The second order linear system of differential equations is flexible and useful, but one cannot reasonably assert that no better model exists for human systems. The use of an easily specified and applied modeling method helps encourage exploration of model alternatives and model comparisons.

A third advantage to the convolution filtering in LDE becomes apparent when nonlinear systems of more than one variable are considered. In many physical systems there exist dynamics that both diverge and converge in a balanced way. When a system is converging to a known equilibrium, errors in measurement of the true signal at time t imply smaller prediction errors at time $t + \Delta t$, giving an advantage to integration forward predictive methods. But when a system is diverging, errors in measurement of the true signal at time t imply *larger* prediction errors at time $t + \Delta t$. If a system is only diverging and one has the full time series in hand it becomes easy to circumvent this problem by running time backwards, i.e., predicting the value at time t from the value at time $t + \Delta t$. So, in this case, a forward prediction method can be shown to not be at a disadvantage to a filtering method. But what happens when a system with more than one variable is both diverging and converging, i.e, when one Lyapunov exponent is positive and another is negative? The time reversal trick no longer works for prediction, and one variable is always at a built-in prediction disadvantage to another variable. Convolution filtering circumvents this problem by maintaining symmetry of the filter and thus does not advantage one variable in the system over any other variable no matter the sign of the associated Lyapunov exponents.

These advantages of the use of time delay embedding are offset by two main disadvantages of the method. First, if the time series are very short, less than about 20 observations per person, the advantages of the overlapping samples of time delay embedding disappear and one would be advised to use another method (Oud, 2017). For longer time series, time delay embedding and convolution filtering remain viable and appealing methods for estimation.

The second disadvantage is that choosing the number of columns in the time delay embedding matrix (known as the *embedding dimension*) has consequences for the estimation of the parameters. Parameter bias increases markedly when the embedding dimension is poorly chosen. Reduction of the dependence between the embedding dimension and parameter bias would be a significant improvement in the performance of the LDE method. The current article derives a constrained fourth order system of equations and demonstrates through simulation that implementing this system of equations as an LDE reduces parameter bias to near zero over a wide range of choices of embedding dimension.

Methods

Constrained Fourth Order LDE

A linear second order differential equation can be transformed into a system of two first order equations. Consider the equation

$$d^2x(t)/dt^2 = \eta x(t) + \zeta dx(t)/dt + e(t)\,. \tag{2}$$

If

$$dx(t)/dt = y(t)\,, \tag{3}$$

then

$$dx^2(t)/dt = dy(t)/dt\,, \tag{4}$$

and now we can substitute Equations 3 and 4 into Equation 2 so that the following system of linear first order equations

$$dx(t)/dt = y(t) \tag{5}$$

$$dy(t)/dt = \eta x(t) + \zeta y(t) + e(t)$$

becomes equivalent to the second order linear differential equation in Equation 2. By this well-known method, a higher order linear differential equation can be reduced to a system of first order differential equations.

Using the same logic, we can start this substitution process with higher order derivatives and end up with the second order linear differential equation. Consider that if

$$d^3x(t)/dt^3 = \eta dx(t)/dt + \zeta d^2x(t)/dt^2 + e(t) \tag{6}$$

and

$$dx(t)/dt = z(t)\,, \tag{7}$$

then by substituting Equation 7 into Equation 6 we find that

$$d^2z(t)/dt^2 = \eta z(t) + \zeta dz(t)/dt + e(t)\,, \tag{8}$$

the familiar second order differential equation for the variable z. By the same logic, two substitutions

$$dx(t)/dt = z(t) \tag{9}$$

$$dz(t)/dt = w(t) \tag{10}$$

will transform

$$d^4w(t)/dt^4 = \eta d^2w(t)/dt^2 + \zeta d^3w(t)/dt^3 + e(t) \tag{11}$$

into Equation 2.

From this logic, we propose a constrained system of equations such that the second, third, and fourth order forms hold simultaneously:

$$\begin{aligned} d^4x(t)/dt^4 &= \eta d^2x(t)/dt^2 + \zeta d^3x(t)/dt^3 + e_4(t) \\ d^3x(t)/dt^3 &= \eta dx(t)/dt + \zeta d^2x(t)/dt^2 + e_3(t) \\ d^2x(t)/dt^2 &= \eta x(t) + \zeta dx(t)/dt + e_2(t)\,. \end{aligned} \tag{12}$$

If these were entirely deterministic equations, in other words if the residual terms were zero, $e_4(t) = e_3(t) = e_2(t) = 0$, then if any one of the equations held, then all of the equations would hold. In this case, there would be nothing to be gained by estimating these equations simultaneously. However, if $e_4(t) \neq e_3(t) \neq e_2(t) \neq 0$, then each regression estimation has the opportunity to improve on the others. These two extra degrees of freedom (three residual terms versus one residual term) are, as presented in Equation 13, still mixtures of deterministic and stochastic residuals. By specifying the derivatives as latent variables, the deterministic and stochastic (non-deterministic) residual variances can be separately estimated. When a traditional second order LDE is estimated, parameter bias can be introduced either when stochastic variance is mistaken for deterministic variance or vice versa. By allowing higher order latent derivatives, the separation between stochastic and deterministic variance is improved and thus parameter bias is reduced.

A path diagram of an LDE model of the second order linear differential equation in Equation 2 applied to a time delay embedding of dimension 5 is shown in Figure 1–a. This LDE model is a simplified form that does not estimate the equilibrium or have any individual differences in parameters. Alternative forms of this second order model that estimate DLO when there are individuals differences in parameters, individual differences in equilibrium value, and within person change in equilibrium are available elsewhere as downloadable scripts (Boker et al., 2016). Figure 1–b presents the fourth order constrained approximation to the second order model that is simulated and tested in the current article against the model shown in Figure 1–a.

Model A, as shown in Figure 1–a was specified and fit in OpenMx (Neale et al., 2016) using the script in Appendix A. In the simulation condition where the data are time delay embedded with an embedding dimension of 5, a matrix $\mathbf{X}^{(5)}$ is the data to which the model is fit (see, e.g., Boker et al., 2016, for an explanation of this matrix). Then the LDE model in Figure 1–a would be specified as

$$\mathbf{X}^{(5)} = \mathbf{G}\mathbf{L}' + \mathbf{U}, \tag{13}$$

where **G** is a matrix of unobserved latent derivative scores, **U** is a matrix of unobserved unique scores and **L** is a fixed matrix that specifies the convolution filter kernel and is in this case defined as

$$\mathbf{L} = \begin{bmatrix} 1 & -2\Delta t & (-2\Delta t)^2/2 \\ 1 & -1\Delta t & (-1\Delta t)^2/2 \\ 1 & 0 & 0 \\ 1 & 1\Delta t & (1\Delta t)^2/2 \\ 1 & 2\Delta t & (2\Delta t)^2/2 \end{bmatrix}, \tag{14}$$

where Δt is the elapsed time between adjacent lagged columns in the time–delay embedded matrix $\mathbf{X}^{(5)}$.

Using the Reticular Action Model (RAM) algebra (McArdle & McDonald, 1984) and path diagram conventions (McArdle & Boker, 1990), the expected covariance of the time delay embedded matrix, $\varepsilon(\mathrm{Cov}(\mathbf{X}^{(5)}))$ is

$$\varepsilon\left(\mathrm{Cov}(\mathbf{X}^{(5)})\right) = \mathbf{L}\mathrm{Cov}(\mathbf{G})\mathbf{L}' + \mathrm{Cov}(\mathbf{U}), \tag{15}$$

where $\mathrm{Cov}(\mathbf{U})$ is a diagonal matrix in which all elements on the diagonal are constrained to be equal to one another. The covariance of the latent variables, Cov(**G**), can now be specified as

$$\mathrm{Cov}(\mathbf{G}) = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{S}(\mathbf{I} - \mathbf{A})^{-1'}. \tag{16}$$

where **I** is the identity matrix,

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \eta & \zeta & 0 \end{bmatrix}, \text{and} \tag{17}$$

$$\mathbf{S} = \begin{bmatrix} V_g & C_{g,dg/dt} & 0 \\ C_{g,dg/dt} & V_{dg/dt} & 0 \\ 0 & 0 & V_{d^2g/dt^2} \end{bmatrix}. \tag{18}$$

The parameters η and ζ are the frequency and damping parameters, V_g and $V_{dg/dt}$ are the variances of the latent variables g and dg/dt, V_{d^2g/dt^2} is the residual variance for d^2g/dt^2, and $C_{g,dg/dt}$ is the covariance between g and dg/dt.

Model B, the constrained fourth order model in Figure 1–b is specified and fit in OpenMx as shown in the script in Appendix B. The same 5 dimensional time delay embedded matrix $\mathbf{X}^{(5)}$ was again predicted by Equation 13 but now there are 5 columns in the matrix of latent

scores $\mathbf{G}$. The convolution kernel matrix, $\mathbf{L}$, for Model B is now a fourth order geometric approximation that can be defined as

$$\mathbf{L} = \begin{bmatrix} 1 & -2\Delta t & (-2\Delta t)^2/2 & (-2\Delta t)^3/6 & (-2\Delta t)^4/24 \\ 1 & -1\Delta t & (-1\Delta t)^2/2 & (-1\Delta t)^3/6 & (-1\Delta t)^4/24 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1\Delta t & (1\Delta t)^2/2 & (1\Delta t)^3/6 & (1\Delta t)^4/24 \\ 1 & 2\Delta t & (2\Delta t)^2/2 & (2\Delta t)^3/6 & (2\Delta t)^4/24 \end{bmatrix}, \tag{19}$$

where Δt is the elapsed time between adjacent lagged columns in the time–delay embedded matrix $\mathbf{X}^{(5)}$. Note that the divisors of the columns of $\mathbf{L}$ are due to the differentiation of the geometric series approximation to the derivatives used in the respective columns. If a time interval is cubed, Δt^3, then in order for it to be an appropriate weight for its corresponding third derivative, d^3g/dt, one must consider that $d(\Delta t^3)/dt = 3\Delta t^2$ and in the same way $d(3\Delta t^2)/dt = 6\Delta t$. Thus, $\Delta t^3/6$, provides the same timescale contribution of the weights in the third column as in the other columns.

The expected covariance of $\mathbf{X}^{(5)}$ is again defined as in Equation 15 except that now

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \eta & \zeta & 0 & 0 & 0 \\ 0 & \eta & \zeta & 0 & 0 \\ 0 & 0 & \eta & \zeta & 0 \end{bmatrix}, \text{and} \tag{20}$$

$$\mathbf{S} = \begin{bmatrix} V_g & C_{g,dg/dt} & 0 & 0 & 0 \\ C_{g,dg/dt} & V_{dg/dt} & 0 & 0 & 0 \\ 0 & 0 & V_{d^2g/dt^2} & 0 & 0 \\ 0 & 0 & 0 & V_{d^3g/dt^3} & 0 \\ 0 & 0 & 0 & 0 & V_{d^4g/dt^4} \end{bmatrix}. \tag{21}$$

Note that since $\mathbf{L}$ has no free parameters and every time η (or ζ) appears in $\mathbf{A}$ it is constrained to be equal to the other ηs (or ζs), there are only 2 more degrees of freedom in Model B than in Model A. Also note that if we were to start with Model B and constrain to zero the variance terms for the third and fourth derivatives as well as their corresponding η and ζ regression predictors, we would have constructed a model identical to Model A and nested within Model B. Thus the difference between the fit of Models A and B is chi square distributed with 2 degrees of freedom and the maximum likelihood ratio test can be applied to test for the difference in fit between Models A and B.

Simulation Conditions

A simulation was run in which the traditional LDE model (Model A from Figure 1–a) and the fourth order constrained approximation LDE model (Model B from Figure 1–b) were each run on the same simulated data and the results compared. The simulation was designed to test the boundaries of the applicability of each model and so we used a univariate version of the LDE, which has been shown to perform substantially worse than the multivariate version (Boker et al., 2004). Our results thus show both Model A and Model B in a particularly poor light and highlight the differences between the models and their points of failure.

Since bias in the LDE model of the DLO has been shown to be a function of embedding dimension (Hu, Boker, Neale, & Klump, 2014), the simulation has the greatest number of conditions applied to embedding dimension. In addition, LDE parameter bias in the DLO has been shown to be a function of noise in the data and so we also varied the signal to noise ratio (SNR) in the simulated data and the decay time of the DLO so that, in some conditions, the signal is only present in the part of the target data prior to when the DLO has returned to close to its equilibrium. When the true simulated signal is close to its equilibrium, it is swamped with noise. In the case of a quickly damped system, the SNR is varying markedly over the course of the time series.

In the last condition, we varied the number of observations in the time series. Since often a researcher will have some prior hypothesis about the cycle time of the system when an experimental design is constructed, we kept the total elapsed time and frequency of the simulated data constant and varied the number of observations within that total elapsed time. Note that frequency is defined as cycles per interval of time. Thus this condition varied the number of samples per cycle. A critical feature of continuous time estimation is that parameter values should not change based on the interval of time between samples. Parameters of the second order differential equation include an estimate of the frequency of oscillations. Therefore, parameters from continuous time estimation cannot depend on how many samples are collected within one cycle. The simulation conditions were constructed so that two critical conditions could be independently varied: the number of samples per cycle and the embedding dimension used to construct the time delay embedding matrix.

The simulation conditions were chosen to be representative of the range of parameters reported in articles presented in the Introduction. Reported frequency and damping parameters were converted into a common metric and number of observations were determined so that commonly reported frequencies would be contained in the conditions. In all cases, the number of observations between columns in the time delay embedding matrix (commonly referred to as the τ parameter) were set to be equal to $\tau = 1$. In practice, we have found that setting $tau = 1$ and only adjusting the embedding dimension produces more precise estimates than setting *tau* to any other value.

With the previous paragraphs in mind, the frequency parameter was fixed, $\eta = -0.50$, for all cells of the simulation. There were two conditions for the damping parameter, $\zeta = \{-0.10, -0.30\}$ corresponding to long and short decay times. The signal to noise ratio had three conditions SNR = {1, 2, 4} corresponding to poor, medium, and good SNRs.

While these do not have exact correspondence to reliability, these SNR conditions are roughly equivalent to reliabilities of .5, .65, and .8 when defined by true score over total score. There were three conditions in the length of the series, $N = \{50,100,300\}$ corresponding to a short, medium, and long time series for a single individual. Different conditions for time intervals between observations were used in order to spread the total time interval variation across one full cycle of the simulated DLO. Convolution kernels spanning a total time interval greater than the cycle length were not considered since they violate the sampling theorem (the so-called Nyquist Limit) and thus produce invalid estimates *a priori*. In the $N = 50$ condition, the embedding dimensions were set to $D = \{4,5,6,7,8,9,10,11,12,13,14,15\}$, for a total of 12 conditions when the length of the simulation was short. In the $N = 100$ condition, the embedding dimensions were set to $D = \{5,7,9,11,13,15,17,19,21,23,25,27,29\}$, for a total of 13 conditions. In the condition when $N = 300$, the embedding dimension conditions were $D = \{5,9,13,17,21,25,29,33,37,41,45,49,53,57,61\}$, for a total of 15 conditions. In this way, each condition of N had similar spread of embedding dimension conditions across a full simulated cycle interval. The full simulated data design included $2 \times 3 \times 3 \times (12 + 13 + 15) = 360$ conditions in all.

In each of the 360 condition cells 1,000 time series were simulated conforming to the parameters of the cell while for each time series randomizing the initial conditions of the DLO and the normally distributed measurement error conforming to the SNR condition. Both Model A and Model B were fit to each data set and the parameters were saved. Thus overall, the simulation ran 720,000 LDE models. Total wall-clock time for the simulation was less than 16 hours when run on a recently manufactured laptop. When strict model convergence criteria were not obtained, the reasons for non-convergence were saved, but the parameters of the non-converging models did not contribute to the analysis of the results. No attempt was made to restart non-converging models and so the non-convergence statistics can be considered to be worst-case figures: a lower bound on convergence percentages.

Results

The mean and standard deviation of each parameter was calculated for the 1000 replications in each cell. This provides the mean point estimate of each parameter and the variability of the parameter over the replications. The mean squared error (MSE) was calculated as the mean of the squared difference between each parameter point estimate and its associated simulated value. The percent bias was calculated as the mean of the difference between each point estimate and its simulated value divided by the simulated value and multiplied by 100. This measure of bias allows comparisons in bias performance independent of the chosen simulated parameter value. The percent of convergence was calculated as the percent of 1000 replications that had no convergence errors. Overall, Model A converged 79.1% and Model B converged 89.5% over all the cells in the simulation. Convergence errors that occurred were either the failure of a boundary walk by the optimizer (the SLSQP option in OpenMx) or that the optimizer exceeded the maximum allotted iterations (the default “Auto” option for SLSQP). The above calculated values are shown in Table 1 for all simulated conditions when the width of the convolution kernel covered half of one cycle of the

simulated DLO (i.e., the embedding dimension was one half the number of samples in one cycle of the simulated DLO).

It is evident from Table 1 that in every condition and for both estimated parameters, the bias as a percent of the parameter value was lower for Model B than it was for Model A. Model B's bias was in all but three conditions between −5% and +5% and the mean across all conditions was −0.5%, leading to the conclusion that, overall, Model B was performing as an unbiased estimator. On the other hand, Model A's bias ranged from −7% to −18% and the mean bias was −12.1%. Overall, Model A always estimated oscillator parameters that indicated slower cycle times and slower decay than was simulated. Thus, if percentage bias is the sole consideration, Model B is always preferred, and to a large degree.

However, in every condition the standard deviations of Model B parameter estimates were larger than those of Model A estimates and so if bias is not taken into account, Model A would be preferred. This would be a poor criterion in practice, since Model A provides estimates that are more precisely inaccurate. One way to balance precision and bias is to use a criterion such as MSE, which is a combined measure of bias and estimation variability. Here we see that in 27 out of 36 cases the MSEs of Model B are smaller than those of Model A, indicating that Model B is to be preferred 75% of the time by the criterion of MSE. The results for percentage convergence do not prefer one model over the other: Out of 18 model conditions, Model A converged more often in 9 conditions and Model B converged more often in 9 conditions.

While the results presented in Table 1 indicate that Model B gives an unbiased estimate of both the η and ζ parameters, this is under the condition that the convolution kernel is approximately one half of the period of the DLO. Since previous work has shown that bias in η is sensitive to the proportion of the period of the DLO covered by the convolution kernel, Figure 2 plots the results of estimating η and ζ with Models A and B for 12 different embedding dimensions under the simulation condition of the medium length time series ($N = 100$), medium signal to noise ratio (SNR=2:1), and longer decay ($\zeta = -0.1$).

The relationship between the bias in estimation of η and the choice of embedding dimension for Model A can be clearly seen in Figure 2–a and its suppression by Model B can be seen in Figure 2–b. Model B is unbiased for all embedding dimensions up to 22: in other words, when the convolution kernel covers less than about 2/3 of a single 30 observation cycle of the DLO. For embedding dimensions greater than 20, the bias for Model B increases rapidly until at 30 observations the Nyquist limit is reached. However, for all embedding dimensions, the bias for η in Model B is smaller than the bias in Model A.

Model B has similar relationships between the bias in η or ζ and the associated embedding dimension. That is to say, up to embedding dimension 22, Model B provides unbiased estimates of ζ. However, for embedding dimensions between 24 and 30, the bias increases markedly. Model A performs better than Model B in this regard, since Model A's bias in ζ is small over the whole range of embedding dimensions between 5 and 30.

With respect to convergence, it can be seen that both Model A and Model B have much more trouble converging when the convolution kernel covers more than 2/3 of the period of the DLO. It should be noted that this means that the parameter estimate results plotted in Figures 2–a, b, c, and d are based on substantially reduced sample cell sizes when the embedding dimension is greater than 22.

In order to understand the effects of length of the time series, two figures were created where, as in Figure 2, the SNR was set to its medium value (SNR=2:1) and simulated $\zeta = -0.1$ but the length of the time series differed. Figures 3 and 4 plot the results for the shorter ($N = 50$) and longer ($N = 300$) time series length respectively. In Figure 3, it is unsurprising that fewer observations in the time series lead to more variable estimates. The Model A results plotted in Figures 3–a and c show the previous pattern of results, but with larger standard deviations. However, Figures 3–b and d are somewhat surprising since parameter bias begins to be different from zero as the coverage of the convolution kernel is increased to be greater than 8, which is only 1/2 the period of the DLO in this simulation condition.

If we now look at the results of estimating the parameters of the longer simulated time series ($N = 300$) plotted in Figure 4, it is again unsurprising that all parameter estimates have smaller standard deviations at all embedding dimensions. Model B now performs well over all but the smallest choice of embedding dimensions whereas the bias in η for Model A is improved, but does not go away. For the longer time series, Model B is preferred in every way, giving unbiased performance with good efficiency and good convergence rates.

Since Model A and B are nested models, the difference between their respective optimized minus two log likelihood function values is distributed as chi-square with two degrees of freedom—a likelihood ratio difference test. Table 2 displays the results of calculating the difference between Model A and Model B function value for each replication in which both models successfully converged when fit using the same embedding dimension as used for Table 1. The mean, μ, and standard deviation, σ, of this difference are displayed. Since Model B has two more degrees of freedom than Model A, Model B is guaranteed to always fit at least as well as Model A. However, as can be seen from the column of means of the difference in fit values, the fit for Model B is almost always much better than would be expected by chance for two degrees of freedom. The column labeled $p < 0.05$ displays the percentage of replications for which this difference is larger than the chi-square two degree of freedom critical value of 5.991. Only when the time series are short and noisy is this percentage not greater than 95%.

Discussion

The current article presents the results of simulated second order linear differential equations that have parameters which result in damped linear oscillators, DLOs. The results of the simulation are quite clear with respect to bias. When a time delay embedding is chosen such that the elapsed time between the first and last column of the time delay embedded matrix was less than 1/2 of a cycle, a constrained fourth order LDE (Model B) had less bias in estimating the frequency parameter, η, in every tested condition than did the traditional

second order latent differential equation (Model A). In estimating ζ, the damping parameter, Model B also exhibited low bias similar to that of Model A in every condition except when the time series were short and noisy. In that case, Model A performed better than Model B in estimating ζ, although worse than Model B in estimating η. When time series included 100 or more observations per person, Model B was unbiased when the elapsed time between the first and last column of the time delay embedded matrix was less than 2/3 of a cycle.

The improvement in estimation can also be observed when comparing Model A and Model B using a likelihood ratio difference test. In all but two conditions, Model B was significantly better fitting than chance at an alpha level of $p = 0.05$ more than 95% of the time. In all but four conditions, Model B was significantly better fitting than Model A in 100% of the replications where both Model A and Model B converged.

The choice of embedding dimension has been of concern to many researchers interested in using the LDE method to estimate parameters of models for fluctuating data. The current results suggest that when there are at least 8 observations per cycle and at least 100 total observations per person, a constrained fourth order LDE can perform as an unbiased estimator of both the frequency and damping parameter of a second order linear differential equation model. Fewer observations per person will require a experimental or observational design in which there are more observations per cycle in order for the method to perform optimally.

Limitations

While the constrained fourth order LDE substantially reduces bias relative to a second order LDE in most of the simulated conditions, it did not do so in all conditions. In particular, when the embedding dimension created a convolution kernel of width more than 2/3 the cycle interval, bias was no longer attenuated. We believe that this bias may be able to be corrected up to a full cycle interval, but at present do not have a method to do so. This is a potential area of further exploration.

The simulation conditions were extensive (720 total conditions), but certainly were not comprehensive. The constrained fourth order LDE needs to be tested in a variety of other commonly used conditions. The current article only simulated stable DLOs—unstable systems and non-oscillating systems need to be tested. The current article only explored univariate LDE, but could be easily extended to multivariate LDE. We chose to explore the univariate case because univariate LDE is more prone to bias than multivariate LDE and so the simulation provided a worst-case scenario for the constrained fourth order LDE test. Future work needs to test whether multivariate LDE can be improved in the same manner.

The latent differential equation that was tested was a second order linear system with one latent variable. While this is one of the common use cases for the LDE estimation method, there are many possible differential equation models that are of use to the psychological, sociological, educational, and life sciences. One common instance is the case of coupled damped linear oscillators. In this case, the dynamics of two latent variables are linearly coupled. Estimation bias for this type of model seems to be likely to be reduced by constrained fourth order LDE, simulation work needs to be done to test this hypothesis.

Reduction of bias in the coupled oscillator model is particularly important when the two latent variables are not oscillating at the same rate. In this case, the choice of embedding dimension in second order LDE becomes difficult, and the two latent variables end up being biased to appear to be more similar than they actually are. If FOLDE exhibits similar bias suppression in coupled oscillator models, it will be particularly helpful for modeling these types of systems.

We did not test the case of estimation of time varying equilibria (e.g., Boker et al., 2016), but our intuition is that this will prove to not affect the results presented here. We also did not test the case of individual differences in parameters or equilibria, but again, the results of the current simulation are expected to apply just as well to these two cases. As for the coupled oscillator case, when there are individual differences in parameters of a DLO, the choice of embedding dimension is difficult and every choice will tend to attenuate individual differences in parameters when traditional LDE is used. Future work on these two cases would be especially useful to those interested in predicting individual differences in dynamics.

We did not test the case of time series phase resets. This is when a sudden jump occurs in a time series and then the dynamic continues from the new initial condition. This case produces bias particularly in the estimation of the damping parameter in LDEs (McKee et al., 2018). In the use of interpolation splines, fourth order splines reduce what is called “overshoot” by the spline (e.g., El Tarazi & Sallam, 1987). It seems reasonable that the fourth order LDE would similarly reduce bias in the damping parameter when the time series is interrupted by phase resets. Future simulation work is needed to explore this possibility.

We did not test the case of missing data and/or time interval misspecification. While these are important problems to consider, time delay embedding has proved to be surprisingly robust to this type of misspecification. (Boker et al., 2018) report four methods for accounting for time interval misspecification when using time delay embedding and find that imputation methods are outperformed by accounting for the time misspecification in each row using OpenMx’s definition variable feature. Simply ignoring the time misspecification produced estimates that were correct for the mean time interval and outperformed the two imputation methods. Future work needs to be done to test whether time interval misspecification and/or missing data interfere with the positive attributes of FOLDE.

Finally, the simulation was limited in that it did not include any nonlinearity or nonstationarity in the differential equation. The recently proposed Adaptive Equilibrium Regulation (AER) models for conditioning and withdrawl observed in addictive substance use (Boker, 2015; McKee et al., 2018) address nonstationarity as a dynamic in and of itself. Future research is needed to understand how the FOLDE technique might be adapted to AER models.

Given so many recommendations for future work appear here and in the interest of reproducibility, the authors include as supplemental materials the full set of simulation

scripts used in the current article in hopes of encouraging others in making rapid progress in testing additional hypotheses.

Conclusions

The use of constrained fourth order LDE reduces parameter estimation bias relative to second order LDE in almost all of the conditions simulated here. When a second order model is fit to data with at least 100 observations per person and at least 8 observations per cycle, FOLDE reduces bias and significantly improves model fit with no apparent downsides, and thus is recommended as an improvement over the use of a standard LDE on these data. When time series are short and noisy, i.e., 50 observations per person and a 1:1 signal to noise ratio, there are tradeoffs to the use of FOLDE such that its recommendation is not so clear. Given short and noisy time series, if the estimation of the damping parameter, ζ, is most critical, we recommend LDE over FOLDE. However, if the frequency parameter, η, is most critical, FOLDE is recommended even for short, noisy time series. For panel data, i.e., very short time series with 4 to 20 observations per person, the advantages of time delay embedding are much reduced and other continuous time methods such as the exact discrete (Singer, 1993), approximate discrete (Oud & Jansen, 2000) or unscented Kalman filter (Chow et al., 2007) methods are may produce better parameter estimates than convolution filtering methods such as LDE and FOLDE (Oud, 2017).

One of the surprising findings of the simulation is how quickly the large (720,000 model fits) simulation ran. The average time per model fit was less than 12 seconds over all the conditions and replications when run with OpenMx on a modern laptop. Thus there does not appear to be a computational complexity downside to FOLDE. Given that, a safe recommendation is to fit both LDE and FOLDE to ones data. In all simulated conditions when FOLDE resulted in a large improvement in fit over LDE, the parameter estimates of FOLDE were substantially less biased than LDE.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding for this work was provided in part by NIH Grant R01DA018673 and fellowships from the University of Zurich Research Priority Program on Dynamics of Healthy Aging and the Max Planck Institute for Human Development. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Max Planck Society or of the National Institutes of Health.

Appendix A – R script for MODEL A specification

```
# ————————————————————————-
# Previously created the simulation data and time delay embedded it
# to create the data “tEmbedded” that are fit by this model.
# Please see the full scripts in the supplemental material.
```

```
embedD <- 8
theTau <- 1
deltaT <- .1

# ————————————————————-
# Create the fixed 2nd order LDE loading matrix.
L1 <- rep(1,embedD)
L2 <- c(1:embedD)*theTau*deltaT-mean(c(1:embedD)*theTau*deltaT)
L3 <- (L2^2)/2
LMatrix <- cbind(L1,L2,L3)

# ————————————————————-
# Create 2nd order LDE model.

manifestVars <- dimnames(tEmbedded)[[2]]

ldeModel1 <- mxModel("LDE_Model_1",
mxMatrix("Full",
values=LMatrix,
free=FALSE,
name="L",
byrow=TRUE
),
mxMatrix("Full", 3, 3,
values=c( 0, 0, 0,
0, 0, 0,
−.2,−.2, 0),
labels=c( NA, NA, NA,
NA, NA, NA,
"eta","zeta", NA),
free=c(FALSE,FALSE,FALSE,
FALSE,FALSE,FALSE,
TRUE, TRUE,FALSE),
name="A",
byrow=TRUE
),
mxMatrix("Symm", 3, 3,
values=c( .8,
.1, .8,
0, 0, .8),
free=c( TRUE,
TRUE, TRUE,
FALSE, FALSE, TRUE),
```

```
labels=c("Vx",
"Cxdx", "Vdx",
NA, NA, "Vd2x"),
name="S",
byrow=TRUE,
lbound=c(0.00000001,
NA, 0.00000001,
NA, NA, 0.00000001)
),
mxMatrix("Diag", embedD, embedD,
values=.8,
free=TRUE,
labels="u1",
name="U",
lbound=0.000001
),
mxMatrix("Iden", 3, name="I"),
mxAlgebra(L %*% solve(I-A) %*% S %*% t(solve(I-A)) %*% t(L) + U,
name="R",
dimnames = list(manifestVars, manifestVars)
),
mxExpectationNormal(covariance="R"),
mxFitFunctionML(),
mxData(cov(tEmbedded),
type="cov",
numObs=dim(tEmbedded)[1]
)
)
# ————————————-
# Fit the LDE model and examine the summary results.
ldeModel1Fit <- mxRun(ldeModel1)
summary(ldeModel1Fit)
```

Appendix B – R script for MODEL B specification

```
# ————————————-
# Previously created the simulation data and time delay embedded it
# to create the data "tEmbedded" that are fit by this model.
# Please see the full scripts in the supplemental material.
embedD <- 8
theTau <- 1
deltaT <- .1
#####################################################
# Fourth Order model
```

```
# ————————————-
# Create the fixed 4th order LDE loading matrix.
L1 <- rep(1,embedD)
L2 <- c(1:embedD)*theTau*deltaT-mean(c(1:embedD)*theTau*deltaT)
L3 <- (L2^2)/2
L4 <- (L2^3)/6
L5 <- (L2^4)/24
LMatrix <- cbind(L1,L2,L3,L4,L5)
# ————————————-
# Create 4th order LDE model.
manifestVars <- dimnames(tEmbedded)[[2]]
ldeModel1 <- mxModel("LDE_Model_1",
mxMatrix("Full",
values=LMatrix,
free=FALSE,
name="L",
byrow=TRUE
),
mxMatrix("Full", 5, 5,
values=c( 0, 0, 0, 0, 0,
0, 0, 0, 0, 0,
-.2,-.2, 0, 0,
0,0,-.2,-.2, 0, 0,
0, 0,-.2,-.2, 0),
labels=c( NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA,
"eta","zeta", NA, NA, NA,
NA, "eta","zeta", NA, NA,
NA, NA, "eta","zeta", NA),
free=c(FALSE,FALSE,FALSE,FALSE,FALSE,
FALSE,FALSE,FALSE,FALSE,FALSE,
TRUE, TRUE,FALSE,FALSE,FALSE,
FALSE, TRUE, TRUE,FALSE,FALSE,
FALSE,FALSE, TRUE, TRUE,FALSE),
name="A",
byrow=TRUE
),
mxMatrix("Symm", 5, 5,
values=c( .8,
.1, .8,
0, 0, .8,
0, 0, 0, .8,
0, 0, 0, 0, .8),
free=c( TRUE,
```

```
TRUE, TRUE,
FALSE, FALSE, TRUE,
FALSE, FALSE, FALSE, TRUE,
FALSE, FALSE, FALSE, FALSE, TRUE),
labels=c("Vx",
"Cxdx", "Vdx",
NA, NA, "Vd2x",
"Cxd3x", NA, NA, "Vd3x",
"Cxd4x", "Cdxd4x", NA, NA, "Vd4x"),
name="S",
byrow=TRUE,
lbound=c(0.00000001,
NA, 0.00000001,
NA, NA, 0.00000001,
NA, NA, NA, 0.00000001,
NA, NA, NA, NA, 0.00000001)
),
mxMatrix("Diag", embedD, embedD,
values=.8,
free=TRUE,
labels="u1",
name="U",
lbound=0.000001
),
mxMatrix("Iden", 5, name="I"),
mxAlgebra(L %*% solve(I-A) %*% S %*% t(solve(I-A)) %*% t(L) + U,
name="R",
dimnames = list(manifestVars, manifestVars)
),
mxExpectationNormal(covariance="R"),
mxFitFunctionML(),
mxData(cov(tEmbedded),
type="cov",
numObs=dim(tEmbedded)[1]
)
)
# ————————————————-
# Fit the 4th order LDE model and save the summary results.
ldeModel1Fit <- mxRun(ldeModel1)
summary(ldeModel1Fit)
```

References

Benson L, Ram N, & Stifter CA (2018). Using fishery models to examine self-and co-regulation processes across multiple timescales. Structural Equation Modeling: A Multidisciplinary Journal, 1–18. (DOI: 10.1080/10705511.2018.1491313)

Bergstrom AR (1966). Nonrecursive models as discrete approximations to systems of stochastic differential equations. Econometrica, 34, 173–182.

Bergstrom AR (1984). Continuous time stochastic models and issues of aggregation over time. In Griliches Z & Intriligator MD (Eds.), Handbook of econometrics (Vol. 2, pp. 1145–1212). Amsterdam: North-Holland.

Bisconti TL, Bergeman CS, & Boker SM (2004). Emotion regulation in recently bereaved widows: A dynamical systems approach. Journal of Gerontology: Psychological Sciences, 59 (4), 158–167.

Bisconti TL, Bergeman CS, & Boker SM (2006). Social support as a predictor of variability: An examination of the adjustment trajectories of recent widows. Psychology and Aging, 21 (3), 590–599. [PubMed: 16953720]

Boker SM (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. Multivariate Behavioral Research, 37 (3), 405–422. [PubMed: 26751295]

Boker SM (2013). Selection, optimization, compensation, and equilibrium dynamics. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry, 26 (1), 61–73.

Boker SM (2015). Adaptive equilibrium regulation: A balancing act in two timescales. Journal of Person-Oriented Research, 1 (1–2), 99–109. (doi: 10.17505/jpor.2015.10) [PubMed: 27066197]

Boker SM, Deboeck PR, Edler C, & Keel PK (2010). Generalized local linear approximation of derivatives from time series. In Chow S-M, Ferrer E, & Hsieh F (Eds.), Statistical methods for modeling human dynamics: An interdisciplinary dialogue (pp. 161–178). Boca Raton, FL: Taylor & Francis.

Boker SM, & Laurenceau JP (2007). Coupled dynamics and mutually adaptive context. In Little TD, Bovaird JA, & Card NA (Eds.), Modeling ecological and contextual effects in longitudinal studies of human development (pp. 299–324). Mahwah, NJ: Lawrence Erlbaum Associates.

Boker SM, Montpetit MA, Hunter MD, & Bergeman CS (2010). Modeling resilience with differential equations. In Molenaar P & Newell K (Eds.), Learning and development: Individual pathways of change (pp. 183–206). Washington, DC: American Psychological Association.

Boker SM, Neale MC, & Klump KL (2014). A differential equations model for the ovarian hormone cycle. In Molenaar PC, Lerner R, & Newell K (Eds.), Handbook of relational developmental systems: Emerging methods and concepts (pp. 369–391). New York: John Wiley & Sons.

Boker SM, Neale MC, & Rausch J (2004). Latent differential equation modeling with multivariate multi-occasion indicators. In van Montfort K, Oud H, & Satorra A (Eds.), Recent developments on structural equation models: Theory and applications (pp. 151–174). Dordrecht, Netherlands: Kluwer Academic Publishers.

Boker SM, Staples A, & Hu Y (2016). Dynamics of change and changes in dynamics. Journal of Person-Oriented Research, 2 (1–2), 34–55. [PubMed: 29046764]

Boker SM, Tiberio SS, & Moulder RG (2018). Robustness of time delay embedding to sampling interval misspecification. In Montfort K. v., Oud J, & Voelkle M (Eds.), Continuous time modeling in the behavioral and related sciences (pp. 239–258). Berlin: Springer.

Butcher JC (2016). Numerical methods for ordinary differential equations New York: John Wiley & Sons.

Butler EA (2017). Emotions are temporal interpersonal systems. Current Opinion in Psychology, 17, 129–134. (DOI: 10.1016/j.copsyc.2017.07.005) [PubMed: 28950958]

Byerly PE (1965). Convolution filtering of gravity and magnetic maps. Geophysics, 30 (2), 281–283. (DOI: 10.1190/1.1439572)

Chow S-M, Ferrer E, & Nesselroade JR (2007). An unscented kalman filter approach to the estimation of nonlinear dynamical systems models. Multivariate Behavioral Research, 42 (2), 283–321. [PubMed: 26765489]

Chow S-M, Hamaker EL, Fujita F, & Boker SM (2009). Representing time-varying cyclic dynamics using multiple-subject state-space models. British Journal of Mathematical and Statistical Psychology, 62, 683–716. [PubMed: 19200409]

Deboeck PR, Monpetit MA, Bergeman CS, & Boker SM (2009). Using derivative estimates to describe intraindividual variability at multiple time scales. Psychological Methods, 14 (4), 367–386. [PubMed: 19968398]

Driver CC, & Voelkle MC (2018). Hierarchical bayesian continuous time dynamic modeling. Psychological methods (DOI: 10.1037/met0000168)

El Tarazi M, & Sallam S (1987). On quartic splines with application to quadratures. Computing, 38 (4), 355–361.

Gerstorf D, Hülür G, Drewelies J, Eibich P, Duezel S, Demuth I, et al. (2015). Secular changes in late-life cognition and well-being: Towards a long bright future with a short brisk ending? Psychology and Aging, 30 (2), 301–310. (DOI: 10.1037/pag0000016) [PubMed: 25799003]

Gerstorf D, Ram N, Röcke C, Lindenberger U, & Smith J (2008). Decline in life satisfaction in old age: Longitudinal evidence for links to distance–to–death. Psychology and Aging, 23 (1), 154–168. [PubMed: 18361663]

Glass L, & Mackey M (1988). From clocks to chaos, the rhythms of life Princeton, NJ: Princeton University Press.

Gonzalez RC, & Wintz P (1977). Digital image processing Reading, MA: Addison–Wesley.

Ho M-HR, Shumway R, & Ombao H (2005). The state space approach to modelling dynamic processes. In Walls TA & Schafer JL (Eds.), Models for intensive longitudinal data (pp. 148–175). Oxford: Oxford University Press.

Hotelling H (1927). Differential equations subject to error, and population estimates. Journal of the American Statistical Association, 22 (159), 283–314.

Hu Y, Boker SM, Neale MC, & Klump K (2014). Latent differential equations with moderators: Simulation and application. Psychological Methods, 19 (1), 56–71. [PubMed: 23646992]

Hubbard JH, & West BH (1991). Differential equations: A dynamical systems approach New York: Springer–Verlag.

Hurley BK, Fink LK, & Janata P (2018). Mapping the dynamic allocation of temporal attention in musical patterns. Journal of Experimental Psychology: Human Perception and Performance (DOI: 10.1037/xhp0000563)

Kalman RE (1960). A new approach to linear filtering and prediction problems. Transactions of the ASME Journal of Basic Engineering, 82, 35–45.

Kaplan D, & Glass L (1995). Understanding nonlinear dynamics New York: Springer Verlag.

Klump KL, Keel PK, Burt SA, Racine SE, Neale MC, Sisk CL, et al. (2013). Exploring the relationship between negative urgency and binge eating: Etiologic associations and the role of negative affect. International Journal of Eating Disorders, 46 (3), 256–263. [PubMed: 23315657]

Klump KL, Keel PK, Kashy DA, Racine S, Burt SA, Neale M, et al. (2013). The interactive effects of estrogen and progesterone on changes in binge eating across the menstrual cycle. Journal of Abnormal Psychology, 122 (1), 131–137. [PubMed: 22889242]

Laplace PS (1779). Mémoire sur les suites. Mémoires de l'Académie royale des Sciences de Paris, 207–309.

Laurenceau J-P, Barrett LF, & Pietromonaco PR (1998). Intimacy as an interpersonal process: The importance of self-disclosure, and perceived partner responsiveness in interpersonal exchanges. Journal of Personality and Social Psychology, 74, 1238–1251. [PubMed: 9599440]

Levinson N (1946). The Wiener (root mean square) error criterion in filter design and prediction. Journal of Mathematics and Physics, 25 (1–4), 261–278. (DOI: 10.1002/sapm1946251261)

Levinson N (1947). A heuristic exposition of wiener's mathematical theory of prediction and filtering. Journal of Mathematics and Physics, 26 (1–4), 110–119. (DOI: 10.1002/sapm1947261110)

McArdle JJ, & Boker SM (1990). Rampath Hillsdale, NJ: Lawrence Erlbaum.

McArdle JJ, & McDonald RP(1984). Some algebraic properties of the Reticular Action Model for moment structures. British Journal of Mathematical and Statistical Psychology, 37, 234–251. [PubMed: 6509005]

McKee KL, Rappaport LM, Boker SM, Moskowitz DS, & Neale MC (2018). Adaptive equilibrium regulation: Modeling individual dynamics on multiple timescales. Structural Equation Modeling: A Multidisciplinary Journal, 25 (6), 888–905. (DOI: 10.1080/10705511.2018.1442224) [PubMed: 30416330]

Montfort K. v., Oud J, & Voelkle M (Eds.). (2018). Continuous time modeling in the behavioral and related sciences Berlin: Springer.

Montpetit MA, Bergeman CS, Deboeck PR, Tiberio SS, & Boker SM (2010). Resilience–as–process: Negative affect, stress, and coupled dynamical systems. Psychology and Aging, 25 (3), 631–640. [PubMed: 20853969]

Morales S, Ram N, Buss KA, Cole PM, Helm JL, & Chow S-M (2017). Age-related changes in the dynamics of fear-related regulation in early childhood. Developmental Science, e12633 (DOI: 10.1111/desc.12633) [PubMed: 29193491]

Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kirkpatrick R, et al. (2016). Openmx 2.0: Extended structural equation and statistical modeling. Psychometrika, 81 (2), 535–549. [PubMed: 25622929]

Nesselroade JR (2002). Elaborating the differential in differential psychology. Multivariate Behavioral Research, 37 (4), 543–561. (DOI: 10.1207/S15327906MBR3704_06) [PubMed: 26816327]

Oertzen T. v., & Boker SM (2010). Time delay embedding increases estimation precision of models of intraindividual variability. Psychometrika, 75 (1), 158–175. [PubMed: 23335820]

Ong A, Bergeman CS, Bisconti TL, & Wallace KA (2006). Psychological resilience, positive emotions, and successful adaptation to stress in later life. Journal of Personality and Social Psychology, 91 (4), 730–749. [PubMed: 17014296]

Oud JHL (2007). Continuous time modeling of reciprocal relationships in the cross-lagged panel design. In Boker SM & Wenger MJ (Eds.), Data analytic techniques for dynamical systems (pp. 87–130). Mahwah, NJ: Lawrence Erlbaum Associates.

Oud JHL (2017). Comparison of four procedures to estimate the damped linear differential oscillator for panel data. In Longitudinal models in the behavioral and related sciences (pp. 19–39). New York: Routledge.

Oud JHL, & Jansen RARG (2000). Continuous time state space modeling of panel data by means of SEM. Psychometrika, 65 (2), 199–215.

Ozaktas HM, Barshan B, Mendlovic D, & Onural L (1994). Convolution, filtering, and multiplexing in fractional fourier domains and their relation to chirp and wavelet transforms. Journal of the Optical Society of America A, 11 (2), 547–559. (DOI: 10.1364/JOSAA.11.000547)

Petris G, & Petrone S (2011). State space models in R. Journal of Statistical Software, 41 (4), 1–25.

Port RF, & vanGelder T (Eds.). (1995). Mind as motion Cambridge, MA: MIT Press.

Racine SE, Keel PK, Burt SA, Sisk CL, Neale MC, Boker SM, et al. (2013). Exploring the relationship between negative urgency and binge eating: Etiologic associations and the role of negative affect. Journal of Abnormal Psychology, 122 (2), 433–444. [PubMed: 23356217]

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences, and modifiers. Cerebral Cortex, 15, 1676–1689. [PubMed: 15703252]

Reilly A, & McGrath D (1995). Convolution processing for realistic reverberation. In Audio engineering society convention 98 New York: Audio Engineering Society.

Ruelle D, & Takens F (1971). On the nature of turbulence. Communications of Mathematical Physics, 20, 167–192.

Russell A, Bergeman CS, Deboeck P, Baird B, Montpetit M, & Ong A (2011). Emotion control during later life: The relationship between global perceptions and daily experience. Personality and Individual Differences, 50, 1084–1088. [PubMed: 24860206]

Salthouse TA, & Davis HP (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. Developmental Review, 26 (1), 31 – 54.

Sauer T, Yorke J, & Casdagli M (1991). Embedology. Journal of Statistical Physics, 65 (3,4), 95–116.

Savitzky A, & Golay MJE (1964). Smoothing and differentiation of data by simplified least squares. Analytical Chemistry, 36, 1627–1639.

Singer H (1993). Continuous-time dynamical systems with sampled data, errors of measurement and unobserved components. Journal of Time Series Analysis, 14, 527–545.

Smith L, & Thelen E (Eds.). (1993). Dynamical systems in development Cambridge, MA: MIT Press.

So P, Ott E, & Dayawansa WP (1994). Observing chaos: deducing and tracking the state of a chaotic system from limited observation. Physical Review E, 49, 2650–2660.

Takens F (1985). Detecting strange attractors in turbulence. In Dold A & Eckman B (Eds.), Lecture notes in mathematics 1125: Dynamical systems and bifurcations (pp. 99–106). Berlin: Springer–Verlag.

Tarter RE, & Vanyukov M (2002). Differential equation models for longitudinal data. In Glantz MD & Johnson JL (Eds.), Resilience and development: positive life adaptations (pp. 85–100). New York: Kluwer.

Thompson JMT, & Stewart HB (1986). Nonlinear dynamics and chaos New York: John Wiley and Sons.

Wiener N (1949). Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications Cambridge, MA: MIT Press.

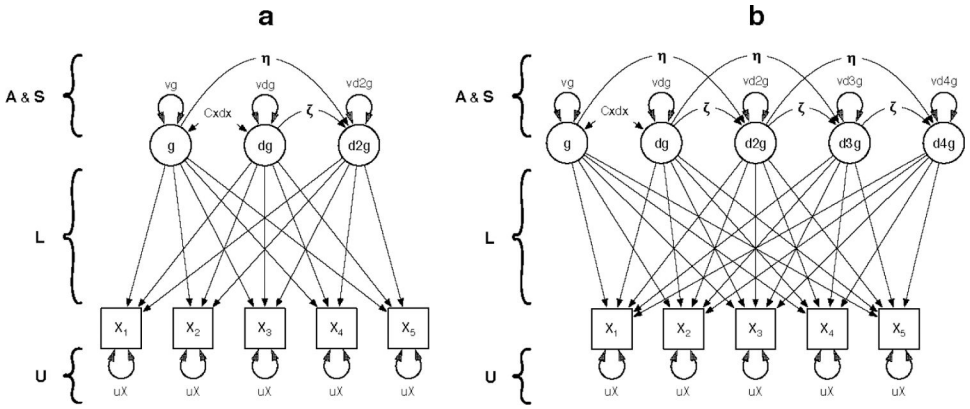

Figure 1.
Second and fourth order LDE models to fit a second order linear differential equation.

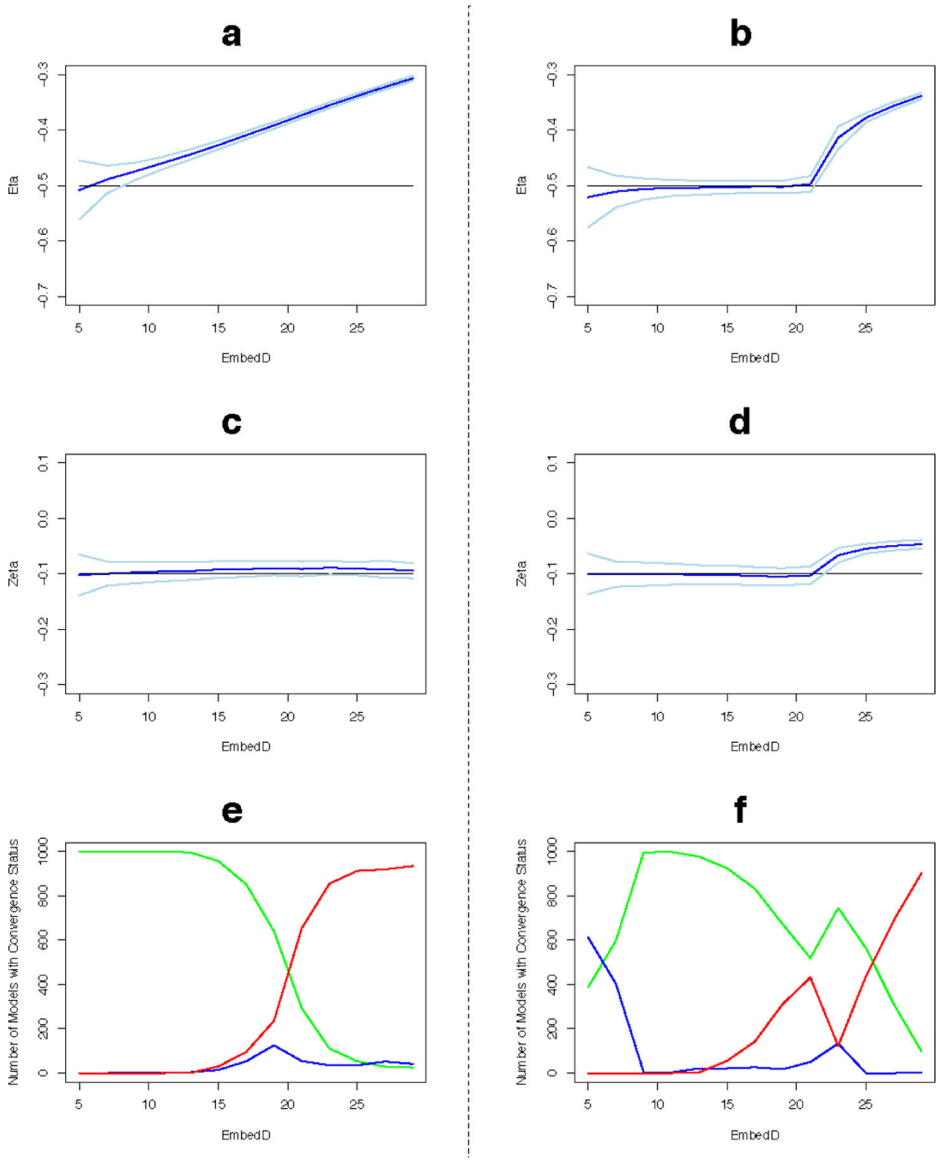

Figure 2.

Results of simulation of second order differential equations with signal to noise ratio of 2:1, $N = 100$, $\eta = -.5$, and $\zeta = -.1$ resulting in 30 observations per cycle as fit with second and fourth order models for 12 different embedding dimensions. (a) Second order approximation model estimated η parameter for each selected embedding dimension. Simulated $\eta = -.5$ is shown as horizontal line. Mean and standard deviation of the estimated parameters are plotted for all models that converged. (c) Second order approximation results for $\zeta = -.1$. (e) Percent of models that converged (green), did not converge (red), and stopped due to exceeding the maximum iterations. (b,d,f) The η, ζ, and convergence percentages for the fourth order model.

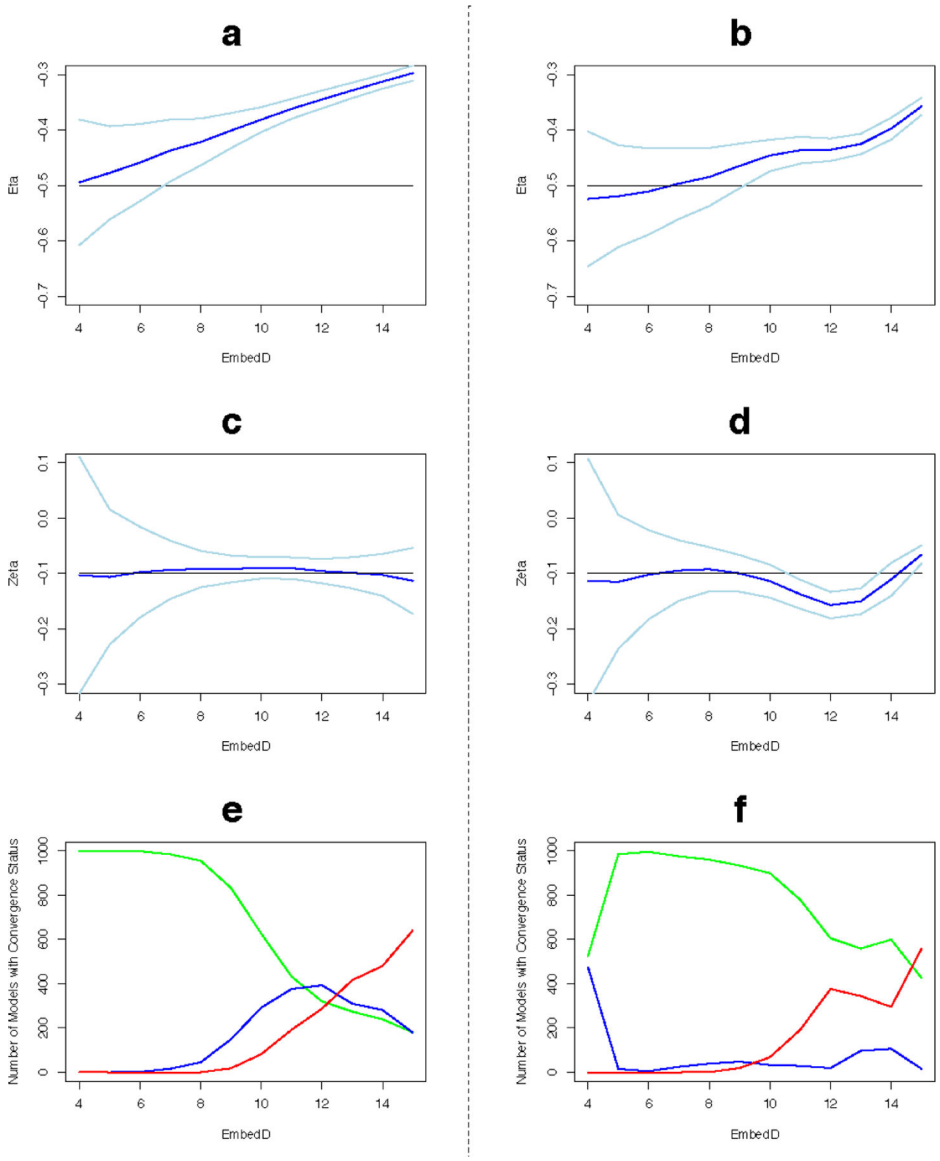

Figure 3.

Results of simulation of second order differential equations with signal to noise ratio of 2:1, $N = 50$, $\eta = -.5$, and $\zeta = -.1$ resulting in 30 observations per cycle as fit with second and constrained fourth order LDEs for 12 different embedding dimensions. (a) Second order approximation model estimated η parameter for each selected embedding dimension. Simulated $\eta = -.5$ is shown as horizontal line. Mean and standard deviation of the estimated parameters are plotted for all models that converged. (c) Second order approximation results for $\zeta = -.1$. (e) Percent of models that converged (green), did not converge (red), and stopped due to exceeding the maximum iterations. (b,d,f) The η, ζ, and convergence percentages for the constrained fourth order model.

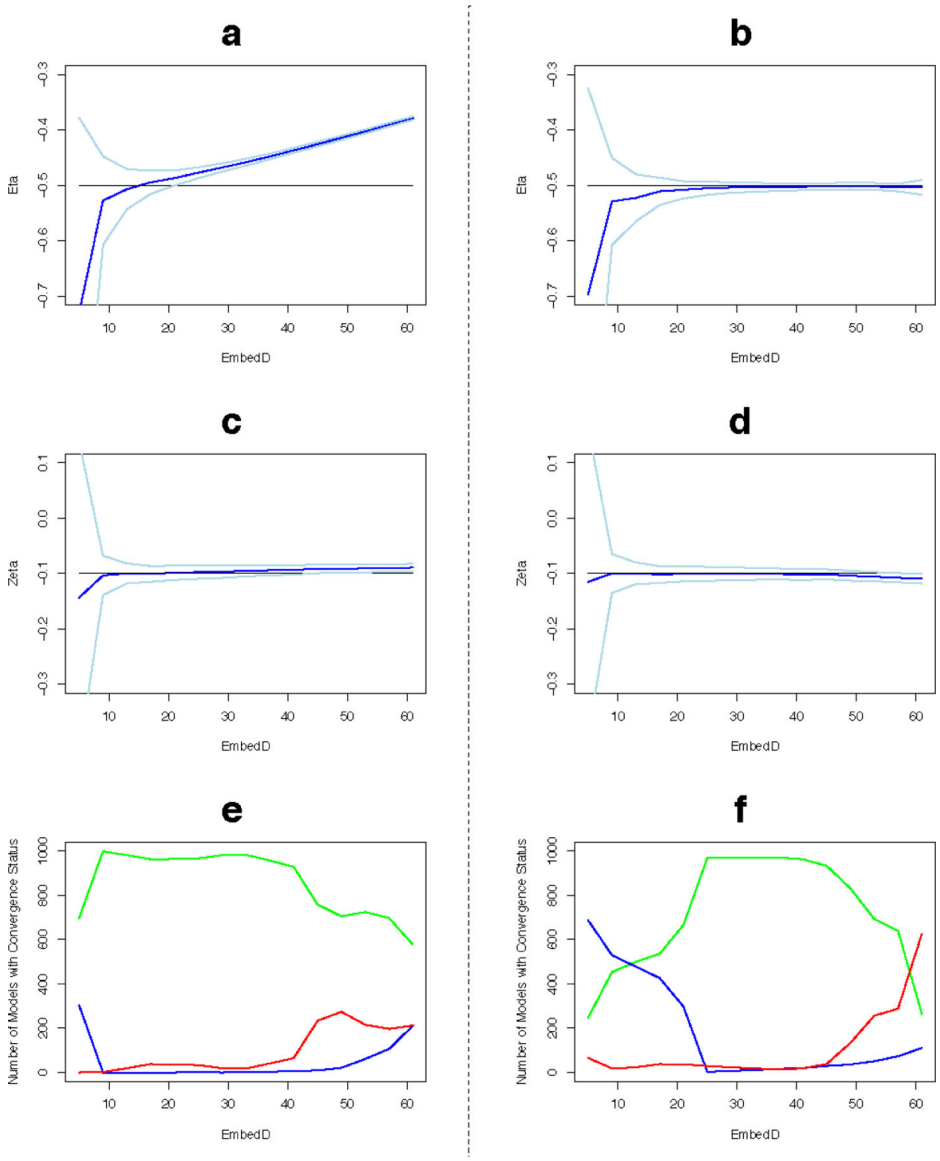

Figure 4.

Results of simulation of second order differential equations with signal to noise ratio of 2:1, $N = 300$, $\eta = -.5$, and $\zeta = -.1$ resulting in 30 observations per cycle as fit with second order and constrained fourth order models for 12 different embedding dimensions. (a) Second order model estimated η parameter for each selected embedding dimension. Simulated $\eta = -.5$ is shown as horizontal line. Mean and standard deviation of the estimated parameters are plotted for all models that converged. (c) Second order approximation results for $\zeta = -.1$. (e) Percent of models that converged (green), did not converge (red), and stopped due to exceeding the maximum iterations. (b,d,f) The η, ζ, and convergence percentages for the constrained fourth order model.

Table 1

Mean (μ) and standard deviation (σ) of point estimates, mean squared error (MSE), the percent bias (Pct Bias) and the percent convergence (Pct Cvrg) for η and ζ for both 2nd and 4th order LDEs over all simulated conditions where the convolution kernel covered 1/2 a cycle of the simulated DLO (i.e., the embedding dimension was 1/2 the number of samples per cycle).

Simulation				MODEL A				MODEL B			
				Est		Pct	Pct	Est		Pct	Pct
	N	SNR	Sim	$\mu(\sigma)$	MSE	Bias	Cvrg	$\mu(\sigma)$	MSE	Bias	Cvrg
η	50	1	−0.5	−0.425(0.095)	0.015	−14.9	96.6	−0.495(0.107)	0.012	−1.0	95.2
ζ			−0.1	−0.088(0.070)	0.005	−12.3		−0.089(0.080)	0.007	−11.0	
η	100	1	−0.5	−0.428(0.015)	0.005	−14.3	99.6	−0.507(0.023)	0.001	1.4	98.0
ζ			−0.1	−0.093(0.031)	0.001	−7.1		−0.102(0.035)	0.001	2.2	
η	300	1	−0.5	−0.427(0.009)	0.005	−14.7	99.1	−0.504(0.014)	0.000	0.9	96.9
ζ			−0.1	−0.093(0.016)	0.000	−7.2		−0.102(0.018)	0.000	2.4	
η	50	2	−0.5	−0.421(0.042)	0.008	−15.7	95.5	−0.484(0.052)	0.003	−3.1	96.0
ζ			−0.1	−0.092(0.033)	0.001	−8.0		−0.092(0.040)	0.002	−7.8	
η	100	2	−0.5	−0.427(0.008)	0.005	−14.6	95.6	−0.503(0.011)	0.000	0.6	92.4
ζ			−0.1	−0.092(0.015)	0.000	−8.0		−0.101(0.017)	0.000	1.5	
η	300	2	−0.5	−0.425(0.004)	0.006	−14.9	75.8	−0.502(0.006)	0.000	0.4	93.3
ζ			−0.1	−0.091(0.008)	0.000	−8.6		−0.101(0.009)	0.000	1.4	
η	50	4	−0.5	−0.416(0.021)	0.008	−16.9	42.2	−0.480(0.027)	0.001	−4.1	94.6
ζ			−0.1	−0.089(0.017)	0.000	−10.6		−0.092(0.020)	0.000	−7.7	
η	100	4	−0.5	−0.426(0.004)	0.005	−14.8	51.0	−0.502(0.006)	0.000	0.5	85.3
ζ			−0.1	−0.092(0.008)	0.000	−7.9		−0.102(0.008)	0.000	2.0	
η	300	4	−0.5	−0.427(0.009)	0.005	−14.5	98.8	−0.505(0.014)	0.000	1.1	96.4
ζ			−0.1	−0.093(0.016)	0.000	−7.0		−0.103(0.018)	0.000	2.8	
η	50	1	−0.5	−0.418(0.085)	0.014	−16.5	98.0	−0.493(0.105)	0.011	−1.3	95.6
ζ			−0.3	−0.276(0.082)	0.007	−8.1		−0.297(0.104)	0.011	−1.0	
η	100	1	−0.5	−0.424(0.021)	0.006	−15.2	99.5	−0.508(0.030)	0.001	1.5	97.8
ζ			−0.3	−0.273(0.031)	0.002	−8.9		−0.303(0.036)	0.001	0.9	
η	300	1	−0.5	−0.424(0.020)	0.006	−15.1	99.7	−0.508(0.029)	0.001	1.6	97.6
ζ			−0.3	−0.273(0.030)	0.002	−9.0		−0.303(0.034)	0.001	1.0	
η	50	2	−0.5	−0.414(0.040)	0.009	−17.1	94.1	−0.485(0.051)	0.003	−2.9	96.5
ζ			−0.3	−0.273(0.039)	0.002	−9.0		−0.294(0.050)	0.003	−2.1	
η	100	2	−0.5	−0.422(0.011)	0.006	−15.5	96.0	−0.503(0.015)	0.000	0.6	94.1
ζ			−0.3	−0.273(0.016)	0.001	−8.9		−0.304(0.019)	0.000	1.3	
η	300	2	−0.5	−0.423(0.014)	0.006	−15.3	66.9	−0.506(0.020)	0.000	1.3	72.1

Simulation				MODEL A				MODEL B			
				Est		Pct	Pct	Est		Pct	Pct
	N	SNR	Sim	$\mu(\sigma)$	MSE	Bias	Cvrg	$\mu(\sigma)$	MSE	Bias	Cvrg
ζ			−0.3	−0.274(0.021)	0.001	−8.7		−0.304(0.024)	0.001	1.5	
η	50	4	−0.5	−0.410(0.019)	0.008	−18.0	24.5	−0.484(0.025)	0.001	−3.2	92.5
ζ			−0.3	−0.271(0.019)	0.001	−9.6		−0.294(0.025)	0.001	−1.9	
η	100	4	−0.5	−0.421(0.005)	0.006	−15.8	63.4	−0.501(0.008)	0.000	0.1	86.3
ζ			−0.3	−0.274(0.008)	0.001	−8.6		−0.305(0.009)	0.000	1.7	
η	300	4	−0.5	−0.421(0.005)	0.006	−15.9	27.3	−0.502(0.007)	0.000	0.4	31.1
ζ			−0.3	−0.274(0.007)	0.001	−8.8		−0.305(0.009)	0.000	1.6	

Table 2

Mean (μ) and standard deviation (σ) of the difference between Model A and Model B minus two log likelihood function values over all simulated conditions where the convolution kernel covered 1/2 a cycle of the simulated DLO (i.e., the embedding dimension was 1/2 the number of samples per cycle). ($p < 0.05$ indicates the percent of comparisons where this difference exceeded a two degree of freedom chi-square critical value of 5.991 and Cnvrg indicates the number of times both models converged out of 1,000 replications so that the likelihood difference could be calculated)

Simulation					**(−2LL MODEL A) – (−2LL MODEL B)**		
N	**SNR**	η	ζ	μ	σ	$p < 0.05$	**Cnvrg**
50	1	−0.5	−0.1	12.53	6.97	83.9%	937
100	1	−0.5	−0.1	30.85	12.60	99.4%	980
300	1	−0.5	−0.1	154.51	35.75	100.0%	967
50	2	−0.5	−0.1	44.04	14.48	100.0%	937
100	2	−0.5	−0.1	118.31	27.13	100.0%	902
300	2	−0.5	−0.1	599.27	83.53	100.0%	724
50	4	−0.5	−0.1	135.48	25.93	100.0%	418
100	4	−0.5	−0.1	408.41	60.36	100.0%	492
300	4	−0.5	−0.1	155.25	36.48	100.0%	963
50	1	−0.5	−0.3	15.07	7.91	91.6%	953
100	1	−0.5	−0.3	29.77	12.33	98.5%	978
300	1	−0.5	−0.3	95.54	34.85	100.0%	976
50	2	−0.5	−0.3	52.92	15.76	100.0%	929
100	2	−0.5	−0.3	111.76	25.83	100.0%	919
300	2	−0.5	−0.3	195.25	53.13	100.0%	607
50	4	−0.5	−0.3	155.30	26.31	100.0%	245
100	4	−0.5	−0.3	377.83	56.21	100.0%	598
300	4	−0.5	−0.3	1420.39	181.39	100.0%	141