

Published in final edited form as:

Multivariate Behav Res. 2008 October 1; 43(4): 497–523. doi:10.1080/00273170802490616.

Modeling Individual Damped Linear Oscillator Processes with Differential Equations:

Using Surrogate Data Analysis to Estimate the Smoothing Parameter

Pascal R. Deboeck,
University of Notre Dame

Steven M. Boker, and
University of Virginia

C. S. Bergeman
University of Notre Dame

Abstract

Among the many methods available for modeling intraindividual time series, differential equation modeling has several advantages that make it promising for applications to psychological data. One interesting differential equation model is that of the damped linear oscillator (DLO), which can be used to model variables that have a tendency to fluctuate around some typical, or equilibrium, value. Methods available for fitting the damped linear oscillator model using differential equation modeling can yield biased parameter estimates when applied to univariate time series. The degree of this bias depends on a smoothing-like parameter, which balances the need for increasing smoothing to minimize error variance but not smoothing so much as to obscure change of interest. This article explores a technique that uses surrogate data analysis to select such a parameter, thereby producing approximately unbiased parameter estimates. Furthermore the smoothing parameter, which is usually researcher-selected, is produced in an automated manner so as to reduce the experience required by researchers to apply these methods. Focus is placed on the damped linear model; however, similar issues are expected with other differential equation models and other techniques in which parameter estimates depend on a smoothing parameter. An example using affect data from the Notre Dame Longitudinal Study of Aging (2004) is presented, which contrasts the use of a single smoothing parameter for all individuals versus use of a smoothing parameter for each individual.

Psychological variables are comprised of many different sources of intraindividual variability (Nesselroade, 1991; Nesselroade & Boker, 1994). One way to study intraindividual variability is through the collection of multiple measurements of the same variables on several individuals over time, that is, *time series*. There are many techniques available for the analysis of interindividual differences in intraindividual variability. One method of growing importance is *differential equation modeling*. The following presents an introduction to differential equation modeling and outlines some of its advantages as a technique for modeling psychological time series. Problems currently encountered with differential equation modeling in psychology are then considered. This is followed by a presentation of a methodological solution using surrogate data analysis as well as consideration of the importance of such a method for applied research. The method presented is intended to allow each individual's time

series to be modeled in a manner optimal for that individual. This more ideographic approach is contrasted with a more nomothetic modeling approach using affect data from the Notre Dame Longitudinal Study of Aging (2004).

DIFFERENTIAL EQUATION MODELING

There are many methods to model data from an individual as they change over time. One could first consider long-term trends in a time series, such as linear or quadratic trends. These trends are frequently examined using techniques such as latent growth curve modeling (Duncan, Duncan, Strycker, Li, & Alpert, 1999) and hierarchical linear modeling (Raudenbush & Bryk, 2001). Following these types of analyses, residual variability usually remains that cannot be expressed using low polynomial trends due to its complexity. This variability is frequently treated as error, that is, independent observations with some distribution. Intuitively, however, it seems likely that day-to-day observations are not independent of each other, even following the removal of some monthly or annual trend. Increasingly, researchers are suggesting that this *intraindividual variability* can yield productive psychological insights (Nesselroade & Ram, 2004). One method that allows for the study of intraindividual variability is differential equation modeling.

A *differential equation* is any equation with a variable that is a derivative (Kaplan & Glass, 1995). A *derivative* is an expression of the change in a variable with respect to another variable. To study intraindividual variability over time, the derivatives of interest will often be changes in some psychological variable x with respect to changes in time. For example, the first derivative \dot{x} could express the change in affect with respect to time, that is, the speed at which affect changes. The second derivative \ddot{x} could express the change in the first derivative with respect to time, that is, how affect accelerates. Differential equation modeling allows one to model the relationships between a person's current state and how she or he changes over time. This modeling, also called *state space modeling*, differentiates differential equation modeling from many other methods that model trajectories of observed scores.

To highlight advantages of differential equation modeling, consider a simple model that may be informative for many psychological contexts: the damped linear oscillator. The prototypical example of a damped linear oscillator is a pendulum, the characteristics of which can be described with two features: frequency and damping. As a pendulum swings back and forth it oscillates at a particular *frequency*. The amplitude of a pendulum may change over time, a characteristic called *damping*. The damped linear oscillator is a versatile model that can be used to describe many different types of change over time, for instance those shown in Figure 1. Psychological constructs that would be well fit by the damped linear oscillator model are ones in which individuals are expected to move back and forth around some "typical" value (i.e., the equilibrium value). It should be noted that this equilibrium does not have to be constant over time (e.g., Bisconti, Bergeman, & Boker, 2004).

In most psychological contexts it would be unlikely to expect observed scores to literally change over time in a manner similar to a linear oscillator. For example, it is unlikely that one would postulate that individuals always reach a peak mood on the weekend and the opposite extreme in the middle of the week. Fitting the damped linear oscillator model using techniques that model the observed time series trajectory, rather than the state space, would literally require postulating a perfect-oscillator hypothesis, which seems unlikely for most psychological variables.

Using differential equation modeling, one can hypothesize that a system has certain relationships between its current state and how it is changing without having to postulate unrealistic hypotheses, such as perfect oscillations over time. Figures 2a and 2b present two

linear oscillators, one of which is a true linear oscillator and the other of which is a linear oscillator with random phase resetting (i.e., at randomly selected times the phase of the oscillation is randomly redrawn from a uniform distribution bounded by 0 and 2π ; this is called process error, or alternatively dynamic error). This phase resetting can be imagined to correspond to random events, such as an unexpected accident or an overdue visit from a good friend, which would alter a trajectory away from the path that would have occurred had the event been absent. The state space of the linear oscillator can be expressed by plotting the observed scores x versus the second derivative of the scores \ddot{x} . The state spaces of the two oscillators are presented in Figures 2c and 2d. Examination of the observed score trajectories (a & b) suggest dramatically different changes over time. Linear regressions of the state spaces, however, suggest that both oscillations have a similar relationship between the observed score and its second derivative.

By modeling the state space one can examine variables that have similar relationships between derivatives but may have trajectories that upon visual inspection appear to be very different. A consequence of examining the relationship between derivatives is that individuals are not required to be aligned at some $t = 0$, as is the case with many other techniques that model the observed trajectory (Boker, Neale, & Rausch, 2004). The practical implication is that differential equation modeling can be used to model everyday situations, such as daily measures of affect during a typical month, rather than only situations following a specific event. Furthermore, as demonstrated in Figure 2, random events can perturb the state space, but the relationships governing how observed values are changing can still be recovered. In psychology this is particularly useful, as events that temporarily perturb how a participant responds may occur—particularly if one examines intraindividual change over an extended period of time. Neither the timing nor the effect of the random events needs to be specified and additional model parameters are not required.

Models, such as the differential equation for a damped linear oscillator, also have the advantage of parameters that have meaningful interpretations (to contrast this, consider trying to describe a time series using a high-order polynomial). The differential equation for a damped linear oscillator is characterized by a frequency parameter and a damping parameter. The *frequency* parameter provides information regarding how quickly the observed scores move toward and away from equilibrium. This change may not be literally as an oscillation; however, a time series with a high frequency will tend to return and move away from equilibrium more rapidly than a low frequency. The *damping* parameter can similarly be interpreted as a change in amplitude over time, although not necessarily a change in amplitude over the course of the time series. As the amplitude can be randomly reset, just as the phase of oscillation, the damping parameter gives an idea as to whether the system has a tendency to converge or diverge from equilibrium.

It is also important to note that differential equation modeling is a form of continuous time modeling. The consequence of fitting a continuous time model, rather than a discrete time model, is that estimated parameters are not dependent on the times at which data are collected. Conversely, the estimates produced by discrete time models frequently depend on the time elapsed between samples, which may lead to different conclusions in the literature regarding how a system changes over time. It should be noted that continuous time modeling is not a cure-all for this problem, however, as one must ensure sufficient sampling of the system of interest.

The combination of advantages presented in the previous paragraphs differentiate differential equation modeling from most other techniques of modeling intraindividual change (e.g., time series analysis; Gottman, 1981; analyses based on the Fourier transform; Gasquet & Witomski, 1999; or wavelets; Gasquet & Witomski, 1999; Polikar, 2004; functional data analysis; Ramsay

& Silverman, 2005; latent growth curve modeling; Duncan et al., 1999; hierarchical linear modeling; Raudenbush & Bryk, 2001). A few examples of differential equation modeling using the damped linear oscillator model have included the mental health scores of recent widows (Bisconti et al., 2004; Boker & Bisconti, 2005), the use of cigarettes and alcohol by adolescents (Boker & Graham, 1998), infant and adult postural control (Boker & Ghisletta, 2001; Boker, 1998), dance movements with ambiguous and unambiguous rhythms (Boker, Covey, Tiberio, & Deboeck, 2005), and intimacy/disclosure in couples (Boker & Laurenceau, 2005). These applications do not seek to definitively determine whether psychological variables oscillate but rather model data as a damped linear oscillator and examine whether logically related variables predict how individuals change around their equilibrium or change their amplitude over time.

LOCAL LINEAR APPROXIMATION

There are several methods that can be used to fit the differential equation model for a damped linear oscillator to an observed time series (e.g., Boker et al., 2004; Boker & Nesselroade, 2002; Oud & Jansen, 2000; Ramsey, Hooker, Campbell, & Cao, 2007). One conceptually and computationally simple technique that has been informative in the development of other techniques is local linear approximation (LLA; Boker, 2001; Boker & Ghisletta, 2001; Boker & Nesselroade, 2002). LLA uses observed values of a time series to estimate derivatives. To estimate derivatives using LLA at a given time t , an integer τ is first selected. This integer is used to specify a triplet of observations that are used to calculate the first and second derivatives at each time t . For instance, LLA might use the triplet $\{x_7, x_{10}, x_{13}\}$ to estimate derivatives at x_{10} . In this triplet $\tau = 3$ and the triplet could be written $\{x_{(t-\tau)}, x_t, x_{(t+\tau)}\}$, or more generally $\{x_{(t-\tau)}, x_t, x_{(t+\tau)}\}$. For all observations for which the required triplet is observed, the first and second derivatives can be calculated as

$$\dot{x} \approx \frac{x_{(t+\tau)} - x_{(t-\tau)}}{2\tau\Delta t}, \text{ and}$$

$$\ddot{x} \approx \frac{x_{(t+\tau)} - x_{(t-\tau)} - 2x_{(t)}}{\tau^2\Delta t^2},$$

respectively. In these equations Δt is the elapsed time between equally spaced observations, and τ is an observation distance as previously described.

LLA and the DLO Model

The differential equation for a damped linear oscillator has the linear form

$$\ddot{x} = \eta x + \zeta \dot{x} \quad (1)$$

and describes the relationship between the second derivative \ddot{x} , the first derivative \dot{x} , and the observed scores x using two parameters. One parameter is related to frequency (η) and one related to damping (ζ). As this is a linear equation, if the observed time series and first two derivatives are estimated using LLA, the parameters η and ζ can be estimated using multiple regression, such that

$$\ddot{x}_t = \eta x_t + \zeta \dot{x}_t + e_t. \quad (2)$$

In order to solve this equation as a multiple regression, standard regression assumptions must be made regarding the distribution of e_t (independent, identically, and normally distributed errors). The appropriateness of this assumption is unlikely in many cases, due to the misspecification between one's model and the true underlying process. Work presented later in this article examines one case, that of dynamic error, in which the relationship in Equation (1) is the true model for a process, but the errors, e_t , do not fully meet the i.i.d. and normality assumptions.

The primary difficulty when implementing LLA is the selection of the parameter τ , which can result in severely biased derivative estimates if poorly selected (Boker & Nesselroade, 2002). Although selection of τ can be relatively straight-forward in long, well-sampled stationary oscillating systems with little noise, psychological data sets do not typically have these characteristics. The selection of τ without biasing parameter estimates in relatively short time series (less than a few hundred observations) and with relatively low sampling rates (e.g., daily measures of affect) is challenging for novice and experienced researchers alike. The bias in parameter estimates is partially due to the fact that a single observed time series is being used to estimate several derivative time series, resulting in correlated errors. A simple solution to this bias is the use of multivariate time series collected on a single construct (Boker et al., 2004); unfortunately these types of data tend to be rare, relative to univariate time series, due to participant load constraints.

A common approach of many related methods is to smooth a time series prior to estimation of the derivatives, so as to minimize error variance (e.g., Ramsay & Silverman, 2005). To smooth data, however, assumptions must be made regarding the time over which features of interest occur; too little smoothing will leave large error effects whereas too much smoothing can obscure or obliterate features of interest. Frequently, smoothing also assumes that sampling rates are high relative to the rate of change. These assumptions and requirements are currently problematic for many areas of psychology in which *a priori* knowledge about attributes of systems (e.g., daily change in affect) are limited.

One consequence of the difficulties of applying LLA to psychological data is that applied research has often selected a single value of τ for all individuals using the data from across all individuals. A single value of τ for all individuals makes an assumption that all individuals change at similar rates. Like smoothing in the same way over time series that change at different rates, a single τ will be just right for some individuals, not smooth over enough error for other individuals (slow-changing trajectories), and (worst of all) smooth over change of interest in yet other individuals (quickly changing trajectories). This means that the variance in estimated parameters is potentially dramatically reduced, leaving little variance to explain when using methods such as multilevel modeling. Furthermore, the η (frequency) parameter should often not be directly interpreted because η is almost directly a function of τ . This means that the researcher-selected value of τ will almost invariably produce a very narrow, predictable range of η values. To most accurately model data, preserve parameter variability, and allow for interpretation of η , it is imperative that value of τ is estimated for each individual's time series without researcher involvement.

PROPOSED METHOD

Boker and Nesselroade (2002) suggest two plots that can be used to determine optimal values of τ ; examples of each are shown in Figure 3. Both plots are based on fitting the damped linear oscillator model using a range of τ s and examining the pattern that is produced (i.e., estimating derivatives using a range of τ values and repeating the multiple regression using the data produced by a particular value of τ). Figure 3a shows the estimated η values for each value of τ . In this plot, the optimal value for τ occurs just before the asymptote. Figure 3b shows the

variance explained by the model, R^2 , for each value of τ . In this plot the optimal τ occurs just as the explained variance reaches its asymptote. The latter plot has been more strongly recommended for the selection of τ . The difficulty with real data is selecting the point *just before the asymptote or just as a plot reaches the asymptote*, as these instructions can be ambiguous and difficult to implement in practice, particularly for researchers unfamiliar with these methods. Even when researchers are familiar with these methods, selection of τ is difficult to perform for a large number of time series, and it is nearly impossible to evaluate whether a researcher selects values of τ in an unbiased and optimal manner.

We propose that the selection of τ can be automated, thus (1) allowing values of τ to be estimated for individual time series, (2) producing approximately unbiased selections of τ , and (3) reducing the expertise required to apply these methods. Boker and Nesselroade (2002) suggest that in plots of explained variance versus τ , the optimal value of τ is the first value at which the variance explained begins to asymptote (i.e., the first value at which the variance explained seems to become maximal). Automation of this technique can be accomplished using two steps: first examining how unlikely the observed explained variance is given a null hypothesis (established using surrogate data analysis) and then modeling when the explained variance appears to asymptote.

In the following section the proposed method is described. The application of the method is then considered over three different scenarios. In order for the method to be useful, it must be first considered whether (1) the automation of the method works and (2) whether values of τ can be selected in an unbiased manner. This scenario is addressed in the first set of simulations. One of the primary advantages of this method is the ability to select values of τ for each individual's time series. In the second application, the method is applied to a real set of data so as to examine the differences that would occur using a single value of τ versus individually estimated values of τ . One potential problem that occurs in these first two sections is that the proposed method will produce a value of τ for any time series, regardless of whether the time series is well described by the selected model. The third application of the proposed method therefore considers the scenario that method is applied to time series consisting of white noise (i.e., independent, normally-distributed observations). This application allows one to contrast the results expected for a true signal, measured with error, against the results for time series with no true signal, while also indicating a potential future direction for certain types of model selection.

METHOD FOR ESTIMATION OF τ

In the estimated R^2 versus τ plot, the optimal value for τ occurs just as the R^2 values reach asymptote (Boker & Nesselroade, 2002). Figure 4a (black line) shows an example of this plot for an oscillating system with measurement errors. We propose that identification of the asymptote, and thereby τ , can be accomplished in an automated fashion using a two-step procedure. The first step is to estimate how likely the estimated R^2 values are for each value of τ . The reference for determining how likely the observations are will be a null hypothesis distribution that is established using surrogate data analysis procedure. This first step draws out the values for τ at which the estimated R^2 values are relatively low or high. In the second step, the point at which the variance explained begins to reach asymptote is estimated, which then allows a value for τ to be selected that should correspond to the optimal value for τ (Boker & Nesselroade, 2002). Each step is described in the following section in greater detail. It is assumed that prior to applying this method, LLA has been used to calculate the estimated R^2 over a range of τ values using the original time series.

First, null hypothesis time series are created to establish a reference for determining the extent to which an observed estimate of R^2 is unlikely. The null hypothesis selected here is that there

are no relationships between observations of the time series that are dependent on time. That is, the null we are using postulates that the ordering of the observations is not important, as it carries no information. Time series conforming to the null can be created by randomly permuting the observations of the original time series. These time series are called *surrogates*. The random permutation of the observations removes the information in the time series that is dependent on time (Schreiber & Schmitz, 2000; Theiler, Eubank, Longtin, Galdrikian, & Farmer, 1992). This technique, commonly called surrogate data analysis in nonpsychological literature, may be considered a subcategory of the randomization tests known in psychology (Edgington, 1995). The distinguishing feature of surrogate data analysis is its application to time series, although like randomization tests the full sample is sampled without replacement (Rodgers, 1999).

Each of the null hypothesis time series is analyzed using LLA in the same manner as the original time series, such that R^2 is estimated over a range of τ values for each surrogate time series. Figure 4a shows an example of the variance explained over a range of τ values for an oscillating time series (black) and multiple null hypothesis time series (grey). Given a large number of surrogate time series, within each value of τ a distribution of R^2 values can be established. The variance explained, at a particular value of τ , in the original time series can then be compared with the null distribution to estimate its likelihood. Figure 4b shows an example of a plot of the likelihood of the explained variance versus τ , given the data in Figure 4a. Rather than assume a particular distributional form for the null hypothesis, the likelihood was based on the proportion of values estimated from the surrogate time series smaller than the value estimated from the original time series.

In the second step, the point at which the variance explained asymptotes is estimated. One half square-waves (*pulses*) are fit to the proportion data (Figure 4b). These pulses were produced by taking a continuous half cycle of a square-wave (with an initial value of 0, maximum of 1, and transition at the midpoint), dividing it into a number of parts, and taking the average for each part. The result is a vector in which the first half equals zero and the second half equals one. That is, except for vectors with an odd number of values, in which case a value of 0.5 occurs between the two halves.

Pulses of varying lengths were compared with the results from the previous step, examples of which can be seen in Figure 4c and 4d. The fit of each pulse was assessed by calculating the difference (D) between each pulse and the likelihood data. The mean squared difference

($MSD, \frac{1}{N} \sum D^2$) was calculated for each pulse. The pulse that produced the minimum mean squared difference ($MMSD$) was selected as the best estimate of the asymptote. Given the best fitting pulse, the value of τ was selected to be the first pulse value corresponding to 1, which reflects the beginning of the asymptote. Fitting pulses, rather than selecting the first statistically unusual observation, was found to best recover parameter estimates, especially when examining very noisy data (Deboeck, 2005).

The selection of the pulses as one half square-waves was based on observations from data. Time series for the null hypothesis tend to show relatively little variance in estimates of R^2 (relative to oscillating time series). See Figure 4a for an example. Because of this, most estimated likelihoods are either very small (i.e., $P(data|null) \ll 0.05$) or very large $P(data|null) \gg 0.95$, with few likelihoods represented in the transition between extremes. This suggested that a square-wave was sufficient for modeling these likelihoods.

Due to the large variety of frequencies, damping parameters, and sampling intervals that are possible with the damped linear oscillator, several limitations were placed on the scope of this article. First, all systems are assumed to be underdamped, such that they do not damp to zero

within one cycle or amplify so rapidly that no oscillations occur. Systems that are underdamped satisfy the equations $\eta < 0$ and $\sqrt{-\eta} > |\zeta|/2$. It is also assumed that observations are equally spaced in time. Furthermore, because it is not recommended for studies to be designed to have a $\tau = 1$, the frequencies of measurement and oscillation were selected such that the optimal τ for any time series is not likely to equal 1.

APPLICATION I: UNBIASED SELECTION OF τ

The following sections consider different applications of the method just described. The first application considered addresses whether the method described can automate the selection of τ and if it can be done in an unbiased manner. These questions are addressed by first simulating oscillating time series, applying the method previously described, and then examining the estimated parameters produced by the method. The simulated time series included both true oscillating time series and oscillating trajectories with random resets. In practice this would correspond to psychological constructs that may literally oscillate as well as constructs that have the dynamics of an oscillator but experience events that may perturb individuals from literally having an oscillating trajectory.

Method

Time series were generated using Mathematica (Mathematica 5.2, 2005). A long time series (1,000 observations) was generated for each possible η value. Initial pairs of x and \dot{x} were randomly selected from this time series in order to quasi-randomly select the initial observation of each time series. Subsequent observations of the time series were generated one observation at a time. A criterion value was selected for each time series, which determined how frequently the system would be random perturbed from the trajectory of a damped linear oscillator. A random number was generated using a uniform distribution for each observation. When the random number was greater than the criterion value, Runge-Kutta fourth order numerical integration of the differential equation for a damped linear oscillator was used to generate the next observation. When the random value was less than the criterion value, a new pair of x and \dot{x} was randomly selected from the time series used to generate the starting values of the time series.

Simulation parameters were as follows: Seven values were selected for the η parameter, $\eta = \{-0.4, -0.2, -0.1, -0.05, -0.025, -0.0125, -0.00625\}$. In a true linear oscillator, these η values correspond approximately to 9.9, 14.0, 19.9, 28.1, 39.7, 56.2, and 79.5 observations measured per cycle, assuming $\Delta t = 1$. Five values were selected for ζ , $\zeta = \{-0.02, -0.01, 0.00, 0.01, 0.02\}$. Three different variances of independent, normally distributed observations (i.e., white noise) were added to each data set, corresponding to signal to noise ratios of 2:1, 1:1, and 1:2. Three different lengths of time series were examined. These lengths correspond to the number of observations per cycle for a given η , multiplied by 1, 2, or 3, rounded up to the nearest whole number. Twenty additional observations were added to each time series to ensure that sufficient observations were available for the LLA regression. Three values were selected for the probability of systems resetting at any given observation, $p(reset) = \{0.000, 0.025, 0.050\}$. All conditions were crossed (945 cells) and 100 data sets were generated per cell.

Fifty surrogate data null hypothesis data sets were created for each time series. The values of τ for which null hypothesis and original data time series results were estimated varied based on the length (l) of the time series. The maximum τ examined for a given time series was equal to $(l - 20)/2$. This is the maximum value of τ possible, assuming one requires at least 20 derivative triplets for the LLA regression at the maximum value of τ . The minimum value of τ was 1. The values of τ consisted of all integers between the maximum and minimum values.

Optimal selection of τ is defined as a τ value that most accurately recovers the true parameter values of η and ζ . It is anticipated that the parameter estimates will not necessarily be normally distributed or even symmetric. Therefore, medians and quartiles for parameter estimates are examined. Means are reported for completeness. The mean squared error (*MSE*), based on the difference between estimated values and true parameter values, is also reported.

Results

For most conditions the estimates of the parameter η were nearly unbiased. The mean and median results for η , under a wide range of conditions, are shown in Tables 1 and 2. There was a general tendency for true high frequency signals to be estimated as slightly lower frequencies and for true low frequencies to be estimated as slightly higher frequencies. Estimates reflected the true value of η as the time series length increased and the signal to noise ratio increased. The effect of phase resetting was mixed, although generally of small magnitude. There was some tendency to estimate low frequencies with frequent phase resets as biased toward slightly higher frequencies. The interquartile ranges and *MSEs* are also reported in Table 1 and 2. Results produced using a fixed τ of 2 are also provided as a contrast to the results.

The recovery of the ζ parameter follows the expected trend, although the results are biased toward zero. The bias toward zero was not unexpected, as this occurs in simulations even when an ideal τ is specified. Frequent phase resetting diminishes the magnitude of the estimated ζ s, as do very large proportions of error variance (e.g., signal to noise ratio of 1:2). The means, medians, and interquartile range for the estimated values of ζ are reported in Table 3. Results produced using a fixed τ of 2 are also provided as a contrast to the results.

Discussion

The results suggest that the proposed method can automatically select values of τ that result in nearly unbiased recovery of model parameters. In particular, the central tendency of η reflects the true values of η , which is important as τ and the estimated values of η are a function of each other (Boker & Nesselrode, 2002). These results held under a wide range of signal to noise ratios, a moderate range of time series lengths, and a moderately large range of probabilities for phase resets. The phase reset results were particularly important for demonstrating modeling of the state space, a feature that differentiates differential equation modeling from other methodologies that seek to model the observed score trajectories.

This method is very promising for the estimation of damped linear oscillator model parameters for individual time series. There are, however, some caveats that should be highlighted. The first caveat, expressed earlier, is that this technique is intended for analysis of time series that is best recovered with a τ equal to or greater than 2. This means that for a system that literally oscillates, it is desirable that a minimum of eight observations are collected per cycle. As stated earlier, reduction of error variance, increasing of time series length, and selection of systems with only low or moderate phase resetting will tend to produce more accurate and precise parameter estimates. It also should be noted that this method will always produce an estimated value of τ , so the production of results should not be interpreted as the presence of oscillation in a time series. We suggest that the best use of this modeling technique is work in which the parameters or derivatives estimated using individual τ values are predicted using other logically related variables. Examples of these applications, without the individually estimated τ values provided by this method, include Bisconti et al. (2004) and Boker and Laurenceau (2005).

APPLICATION II: APPLIED EXAMPLE

The next application of the method is to a set of real data. The data consist of affect time series from a sample of older adults. Using the method described, we estimate τ values for each

individual's time series. Results are contrasted against the results that are produced when a fixed τ of 2 is used for all individuals—a choice made in several published articles (e.g., Bisconti et al., 2004; Boker & Laurenceau, 2005).

Method

Participants consisted of a subsample of the Notre Dame Longitudinal Study of Aging (2004) who agreed to participate in a 56-day daily diary study. The original subsample consisted of 101 participants who responded to a survey questionnaire, with 61 of those further agreeing to provide daily data. Due to the difficulties of collecting complete data, not all of the 61 participants have sufficient data for analysis. Thus a subset of participants was selected for these analyses. The criteria for selection was that participants had to have complete data on more than 75% of both the positive and negative affect assessments. Furthermore, due to low variability in the negative affect assessments, individuals who had 90% or more assessments with the same negative affect score were not analyzed because variability is required to model intraindividual variability. There were a total of 26 participants analyzed, who ranged in age from 65 to 92, with an average age of 79.2. All participants lived around a mid-size midwestern city, 71% were female, 29% were married at the time of the study, and 55% completed some post-high school education. Given the limited sample, the self-selection that occurred, and the limited demographics, the results presented are intended primarily as an example of application of the method described in this article.

Self-reported daily positive and negative affect was collected for 56 days using the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegan, 1988). Chronbach's α for positive and negative affect on day 1 was 0.85 and 0.92, respectively. Daily data were collected using 1-, 2-, and 3-week packets that were counterbalanced within and between subjects to circumvent concerns that the manner in which participants received and mailed back packets could contribute to observed periodicity in the data. Participants were asked to fill out the questionnaire every evening before they went to bed (and reflect back on their day). The specific time at which the data were collected was not recorded.

Estimates of τ and parameters were made for the positive and negative affect time series separately within each individual using the method described previously. Within each time series, all possible values of τ were examined when at least 10 sets of derivatives could be estimated. Because trends were not expected, equilibrium values were estimated as the individual mean of each time series. As variation in results can be observed due to random sampling of the null hypothesis data sets, the number of null hypothesis data sets was increased to 500 and analyses were conducted twice to ensure stability of estimates. We present the results of the estimated parameters using the method presented in this article as well as the results if the data were analyzed with a fixed value of 2 for all individuals.

Results

The results for positive affect are examined first. Using the method presented, the individual τ s that were automatically selected ranged from 2 to 13, with a mean of 7.04 and median of 6. Estimates of η ranged from -0.600 to -0.011, with mean of -0.129 and median of -0.051. For a true linear oscillator the range of η estimates corresponds to cycles of between 8.1 and 59.9 days. When τ was fixed to equal 2, estimates of η ranged from -0.614 to -0.155, with mean of -0.407 and median of -0.425. For a true linear oscillator this range of η estimates corresponds to cycles of between 8.0 and 16.0 days. The correlation between η estimates using fixed and individual τ s was 0.53. The ζ estimates with individually selected τ s ranged from -0.177 to 0.074, with mean of -0.025 and median of -0.015. When τ was fixed to equal 2, estimates of ζ ranged from -0.490 to -0.271, with mean of -0.037 and median of -0.007. The correlation between ζ estimates using fixed and individual τ s was 0.57.

For negative affect, the individual τ s that were automatically selected ranged from 2 to 12, with a mean of 5.89 and median of 5. Estimates of η ranged from -0.580 to -0.013, with mean of -0.152 and median of -0.076. For a true linear oscillator, the range of η estimates corresponds to a cycle of between 8.3 and 55.1 days. When τ was fixed to equal 2, estimates of η ranged from -0.580 to -0.251, with mean of -0.441 and median of -0.450. For a linear oscillator this range of η estimates corresponds to cycles of between 8.3 and 12.5 days. The correlation between η estimates using fixed and individual τ s was 0.46. The ζ estimates with individually selected τ s ranged from -0.279 to -0.440, with mean of -0.005 and median of -0.017. When τ was fixed to equal 2, estimates of ζ ranged from -0.423 to 0.440, with mean of -0.059 and median of -0.005. The correlation between ζ estimates using fixed and individual τ s was 0.69.

Discussion

The applied example highlights several items of importance for applied researchers. First, the estimation of frequency η , as this is the parameter that has been shown to be biased depending on selection of τ and therefore the parameter that necessitates the methods presented (Boker & Nesselroade, 2002; Deboeck, 2005). Considered in absence of additional information, the frequency estimates using a fixed value of τ produce a much narrower range of estimates than when estimating τ for each individual. The additional variability of the estimates produced with the proposed method are likely to be most beneficial when researchers are interested in using exogenous variables to predict these parameter estimates (Bisconti, Bergeman, & Boker, 2006).

It is important to consider which range of results is likely to more closely reflect reality. The results from the previous simulation, as well as work by Boker and Covey (2002), suggest that for a fixed τ of 2 the resulting frequency estimates have an expected cycle of 8.88 days ($\eta = -0.5$). This is precisely what is observed in these applied results (mean cycle length of 9.8 days), suggesting that it would be unwise to directly interpret the results from the fixed τ case. That is, an applied researcher would need to avoid claiming that most people have dynamics of a cycle of 8.3 to 16.0 days, as the central tendency of this range will change depending on the value of τ selected. From the previous simulation, however, it is likely that with individual τ estimates direct interpretation of the η estimates is reasonable. The individual τ results suggest that there are substantial differences in how people move toward and away from their equilibrium state and that this sample has a mean cycle length of 17.5 days.

The aspect of this method that has the greatest impact for applied researchers is that the expertise required for application of these methods is reduced. Selection of τ values can be difficult even for experienced researchers, and this only becomes more difficult if one wishes to do so for multiple time series. Furthermore, it is difficult to evaluate how well any one researcher selects τ or which researcher selects τ better in the case of researchers with differing opinions as to the best value of τ . The analyses presented are fully automated once a time series is provided and a maximum value of τ is selected (often dictated by the data available). Code for the statistical program R (R 2.1.0, 2005) is available from Pascal R. Deboeck. Naturally one must still consider issues related to properly sampling the system of interest, such as how long one must sample and what is an appropriate sampling rate.

Note that when these methods are applied in practice, a subsequent step is to predict the estimated model parameters using other variables of theoretical interest. Figure 5 plots the estimated cycle length for positive versus negative affect for each individual. This figure suggests that most people have a faster negative affect frequency than positive affect frequency. This corresponds to quicker returns to equilibrium and away from equilibrium for negative than positive affect. Ideally, one would use exogenous variables (e.g., personality, coping, social support variables) to predict which individuals tend to have higher/lower positive/negative affect frequencies. It would also be interesting to examine whether the differences

observed in the estimated parameters were predictive of future outcome, such as future mental or physical health outcomes. These goals are supported in multiple ways with the current method by allowing researchers to better model data within one individual before considering interindividual differences (Emmerich, 1968; Molenaar, 2004).

APPLICATION III: NONOSCILLATING TIME SERIES

From the design of the method, it is clear that it will estimate a value of τ , and consequently the frequency and damping parameters, even if a time series of interest has no oscillatory component. In this section consider the application of this method to nonoscillating time series, specifically white noise. Naturally, if one were examining nonoscillating time series in practice, by trying to relate the time series to exogenous variables one should have a difficult time finding statistical relationships. The benefit of this analysis, however, is to suggest one method that may help to distinguish whether a time series of interest is likely to contain an oscillatory component or whether the observations may be more random.

Method

Time series consisting of white noise with mean of 0 and variance of 1 were created using the statistical program R (R 2.1.0, 2005). Time series lengths corresponded to the 21 different lengths used in the previous simulation, which ranged in length from 30 to 258 observations. One thousand white noise time series were created for each possible length. Analysis of the data was conducted using the same methods previously described, including the same number of surrogate time series and range of τ values.

Results

As expected, a wide range of η and ζ values were produced, despite the time series having no true signal variance. Although it is unlikely that random time series could be distinguished from their oscillating counterparts based on estimates parameters, differences were observed in the distribution of *MMSD* values. The percentiles of each of the signal to noise ratio conditions from the first simulation (Application I) and those of the white noise simulation are presented in Table 4. This table suggests that typical *MMSD* values for oscillating time series, even with relatively high signal to noise ratios, are very likely to be smaller than those observed for a time series consisting entirely of white noise. Of oscillating time series with a 1:1 signal to noise ratio, 94.8%, 80.1%, and 67.8% fell below the 5th percentile of the results for the white noise time series when the chance of phase resetting was 0%, 2.5%, and 5%, respectively. Even when the signal to noise ratio was small (1:2) and the chance of phase resets was large (5%), about 38.5% of oscillating time series fell below the 5th percentile *MMSD* of the white noise time series.

Discussion

The results emphasize the fact that the *MMSD* method can be used regardless of whether a time series truly oscillates. Furthermore, parameter estimates are quite diverse, suggesting that these alone cannot be used to identify oscillating time series. These results suggest, however, that when the method is used on data created using the same model, the distribution of *MMSDs* tends to consist of values that are much smaller than those for white noise. Large amounts of measurement error, or a high probability of phase reset, rapidly diminished the percentage of true-signal time series with an *MMSD* smaller than the white noise time series.

At this time it is not recommended that these results be used for statistical testing as to whether time series oscillate. For example, Figure 6 plots the densities of the *MMSD* for the positive and negative affect, the results for a true oscillator with a signal to noise ratio of 1:2, and the results for white noise time series. We would not suggest interpreting the model fit as good or

bad based on this figure. The results do suggest, however, that the affect data are not behaving in a manner that would be expected for white noise. This is supported by the fact that 86.1% and 85.1% of negative and positive affect time series produced *MMSDs* smaller than the median result for white noise time series, a statistically unlikely result if due to chance alone. Although not behaving like white noise, however, the data are not distributed like a damped linear oscillator measured with 1:2 signal to noise ratio, where the error consists of white noise. This could be suggestive of several conditions, including: model misspecification, need for a higher sampling rate, very high amount of measurement error, frequent phase resetting, or perhaps other conditions.

CONCLUSIONS

This article presented a method that allows for the modeling of individual time series using differential equation modeling by providing an automated method to select the parameter τ (an R script for applying this method can be obtained from Pascal R. Deboeck). The simulation results suggest that the method can produce nearly unbiased parameter estimates. This is an important step for differential equation modeling as it allows estimation of individual τ values for time series in an automated, objective manner, thereby improving estimates of the damped linear oscillator frequency parameter.

Proper selection of τ will have several consequences in practice, including reducing the proportion of error in the derivatives while preserving change of interest in the time series and increasing interindividual parameter variability. When model parameters or derivatives are related to other variables, individual τ s will provide researchers with additional power to detect relationships. On a practical level, these methods also help to reduce the expertise required to fit the damped linear oscillator model and thus help to make this methodology more accessible.

In this article, only the damped linear oscillator was examined, as it was a model that seems likely to be one of interest in the psychological sciences. Despite this limitation, similar difficulties with modeling are likely to occur in many differential equation models. It should be noted that the problems that arise from the linear regression of the derivatives, using the estimated derivatives, do so because of the correlations between error estimates of the derivatives. In the case of the linear oscillator, the errors of the observed values of x and the estimated values of \dot{x} are correlated. These methods are likely to be of use for differential equations in which the estimation of derivatives leads to correlated errors, and it is this correlation that drives the relationship observed by Boker and Nesselroade (2002).

One limitation of the current work is that it only considers the possibility of random perturbations to a system. Perturbations, however, may not be random for many psychological systems, for example, the effects of a weekly cycle. These types of processes might be described using an alternative model, such as one with a nonlinear damping term. It should be noted, however, that if the nonlinear damping term is independent of the values of x or \dot{x} it is unlikely that such a term will dramatically change the relationship between the variance explained by the model and the selection of τ . Consider Figure 7, which plots two examples with nonlinear damping (or forcing). The first system is a type of damping associated with a block attached to a spring as it slides across a table. The second is a power curve that resets to zero every seven observations; this is intended to mimic the possibility of a weekly forcing term. Both systems in this figure produce plots of variance explained versus τ that are very similar to those for a damped linear oscillator, as the errors of the nonlinear damping term are independent of x and \dot{x} (the two terms with correlated errors). This method, therefore, is likely to be applicable to a wide variety of interesting models, including those with nonlinear damping or forcing terms.

The simulation results also suggest that *MMSD* might be a useful tool for distinguishing certain types of models. The white noise results suggest that *MMSD* may be able to be used to distinguish an oscillating time series from one that is random. The systems considered in Figure 7, however, suggest that *MMSD* is unlikely to distinguish efficiently between models that have different nonlinear damping or forcing terms. The difference between these cases is whether the relationship between x and \dot{x} has been changed, that is, whether the relationship the derivatives with correlated errors has changed. Models that do not alter the relationship between derivatives with correlated errors resulting from the estimation process do not seem likely to produce dramatically different *MMSDs*. Further research is needed to have good objective criteria for selecting between differential equation models.

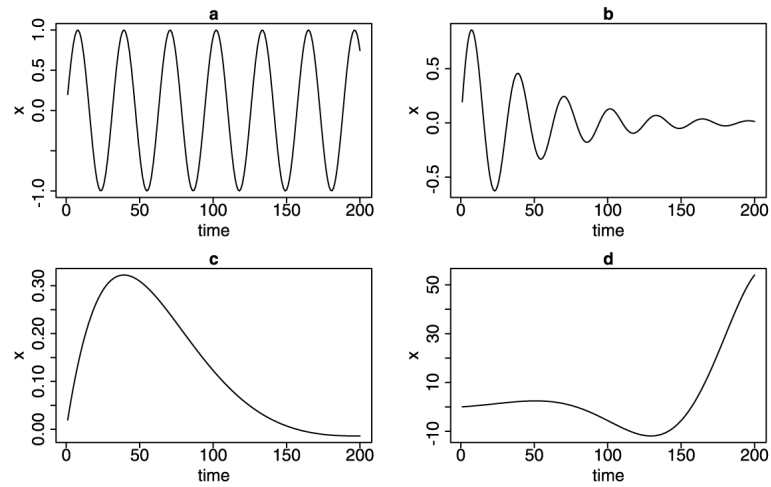
Regardless of the objective criteria used, *MMSD* or otherwise, it seems unwise for many psychological applications to seek curve fitting as an end goal. Many psychological applications are unlikely to produce sufficient data to distinguish between models that produce similar trajectories within a narrow window of time. Psychological time series are frequently short (less than a few hundred observations), measurement error is frequently high, and sampling rates are relatively low. Given these conditions, there are likely to be many different functions that fit the same set of data equally well. A better strategy is to estimate the effects of exogenous variables that are theoretically interesting and statistically related to the parameters estimated from time series.

In order to model time series, one must consider the balance between smoothing enough to reduce error but not so much as to obscure change. To smooth data using a common smoothing parameter across individuals, however, risks ignoring some of the rich interindividual differences in intraindividual variability. To avoid such a risk and best model intraindividual variability, we must remember that one τ does not fit all.

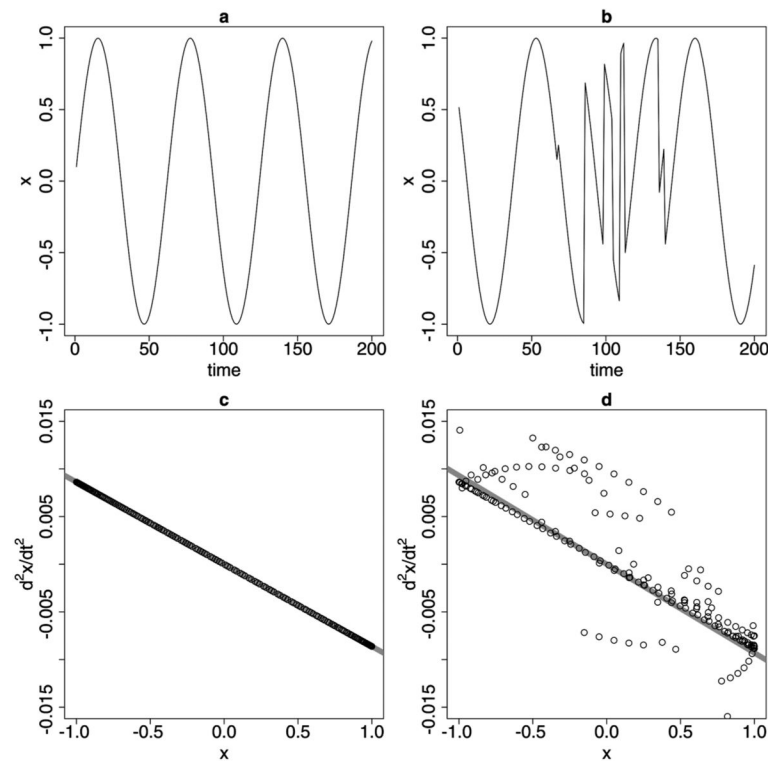
REFERENCES

- Bisconti TL, Bergeman CS, Boker SM. Emotional well-being in recently bereaved widows: A dynamical systems approach. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 2004;59:158–167.
- Bisconti TL, Bergeman CS, Boker SM. Social support as a predictor of variability: An examination of the adjustment trajectories of recent widows. *Psychology and Aging* 2006;21(3):590–599. [PubMed: 16953720]
- Boker, SM. Age-based comparisons of nonlinear dependency in postural control; Annual Meeting of the Gerontological Society; Philadelphia. 1998;
- Boker, SM. Differential structural equation modeling of intraindividual variability. In: Collins, LM.; Sayer, AG., editors. *New methods for the analysis of change*. American Psychological Association; Washington, DC: 2001. p. 5-27.
- Boker, SM.; Bisconti, TL. Dynamical systems modeling in aging research. In: Boker, SM.; Bergeman, CS., editors. *Quantitative methods in aging research*. Erlbaum; Mahwah, NJ: 2005.
- Boker SM, Covey ES. Two recent advances in estimating and testing differential equations models. *Society of Multivariate Experimental Psychology*. October;2002
- Boker, SM.; Covey, ES.; Tiberio, SS.; Deboeck, PR. Synchronization in dancing is not winner-takes-all: Ambiguity persists in spatiotemporal symmetry between dancers. *North American Association for Computational, Social, and Organizational Science*; 2005. Available online at http://www.casos.cs.cmu.edu/events/conferences/2005/conference_papers.php
- Boker SM, Ghisletta P. Random coefficients model for control parameters in dynamical systems. *Multilevel Modelling Newsletter* 2001;13(1):10–17.
- Boker SM, Graham J. A dynamical systems analysis of adolescent substance abuse. *Multivariate Behavioral Research* 1998;33(4):479–507.

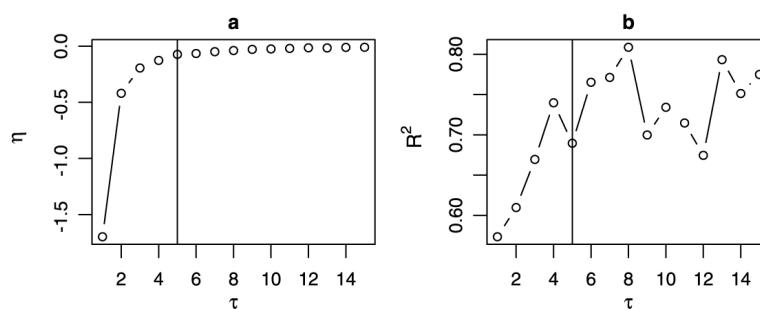
- Boker, SM.; Laurenceau, JP. Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. In: Walls, TA.; Schafer, JL., editors. Models for intensive longitudinal data. Oxford University Press; Oxford, UK: 2005. p. 195-218.
- Boker, SM.; Neale, MC.; Rausch, JR. Latent differential equation modeling with multivariate multi-occasion indicators. In: Montfort, KV.; Oud, J.; Satorra, A., editors. Recent developments on structural equation models: Theory and applications. Kluwer; Amsterdam: 2004. p. 151-174.
- Boker SM, Nesselroade JR. A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research* 2002;37(1):127–160.
- Deboeck, PD. Unpublished master's thesis. University of Notre Dame; Notre Dame, IN: 2005. Using surrogate data analysis to estimate τ for local linear approximation of damped linear oscillators.
- Duncan, TE.; Duncan, SC.; Strycker, LA.; Li, F.; Alpert, A. An introduction to latent variable growth curve modeling: Concepts, issues, and applications. Erlbaum; Mahwah, NJ: 1999.
- Edgington, ES. Randomization tests. Vol. 3rd ed.. Marcel Dekker; New York: 1995.
- Emmerich W. Personality development and concepts of structure. *Child Development* 1968;39:671–690. [PubMed: 4879843]
- Gasquet, C.; Witomski, P. Fourier analysis and applications: Filtering, numerical computation, wavelets. Springer; New York: 1999.
- Gottman, JM. Time-series analysis: A comprehensive introduction for social scientists. Cambridge University Press; New York: 1981.
- Kaplan, D.; Glass, L. Understanding nonlinear dynamics. Springer; New York: 1995.
- Computer software. Wolfram Research; Champaign, IL: 2005. MathematicaWebsite: <http://www.wolfram.com>
- Molenaar PC. A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement* 2004;2(4):201–218.
- Nesselroade, JR. Interindividual differences in intraindividual change. In: Collins, LM.; Horn, JL., editors. Best methods for the analysis of change. American Psychological Association; Washington, DC: 1991. p. 92-105.
- Nesselroade, JR.; Boker, SM. Assessing constancy and change. In: Heatherton, TF.; Weinberger, JL., editors. Can personality change?. American Psychological Association; Washington, DC: 1994. p. 121-147.
- Nesselroade JR, Ram N. Studying intraindividual variability: What we have learned that will help us understand lives in context. *Research in Human Development* 2004;1(1 2):9–29.
- Oud JH, Jansen RA. Continuous time state space modeling of panel data by means of sem. *Psychometrika* 2000;65:199–215.
- Polikar, R. The engineer's ultimate guide to wavelet analysis: The wavelet tutorial. Oct. 2004 Available from <http://users.rowan.edu/polikar/WAVELETS/WTtutorial.html>
- Computer software. Apr. 2005 R. Available from <http://www.r-project.org/>
- Ramsay J, Hooker G, Campbell D, Cao J. Parameter estimation for differential equations: A generalized smoothing approach. *Journal of the Royal Statistical Society* 2007;69:741–770.
- Ramsay, J.; Silverman, BW. Functional data analysis. Springer; New York: 2005.
- Raudenbush, S.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. Vol. 2nd ed.. Sage; Newbury Park, CA: 2001.
- Rodgers JL. The bootstrap, the jackknife, and the randomization test: A sampling taxonomy. *Multivariate Behavioral Research* 1999;34(4):441–456.
- Schreiber T, Schmitz A. Surrogate time series. *Physica D* 2000;142:346–382.
- Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD. Testing for nonlinearity in time series: The method of surrogate data. *Physica D* 1992;58:77–94.
- Watson D, Clark LA, Tellegan A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 1988;54:1063–1070. [PubMed: 3397865]

**FIGURE 1.**

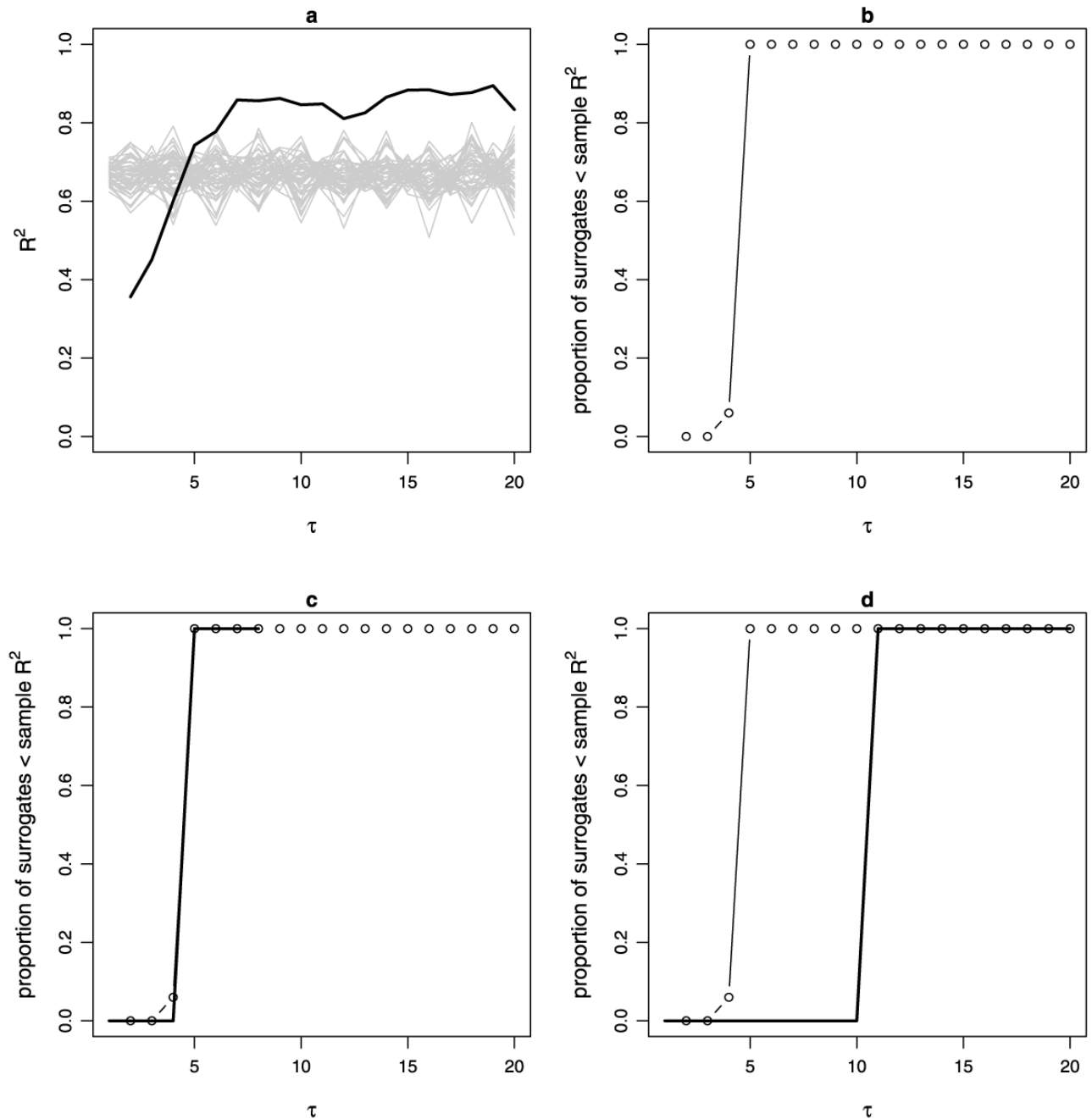
Examples of trajectories that can be modeled using the damped linear oscillator model. These examples include high (a & b) and low (c & d) frequencies, no damping (a), negative damping (b & c), and positive damping (d, i.e., amplification).

**FIGURE 2.**

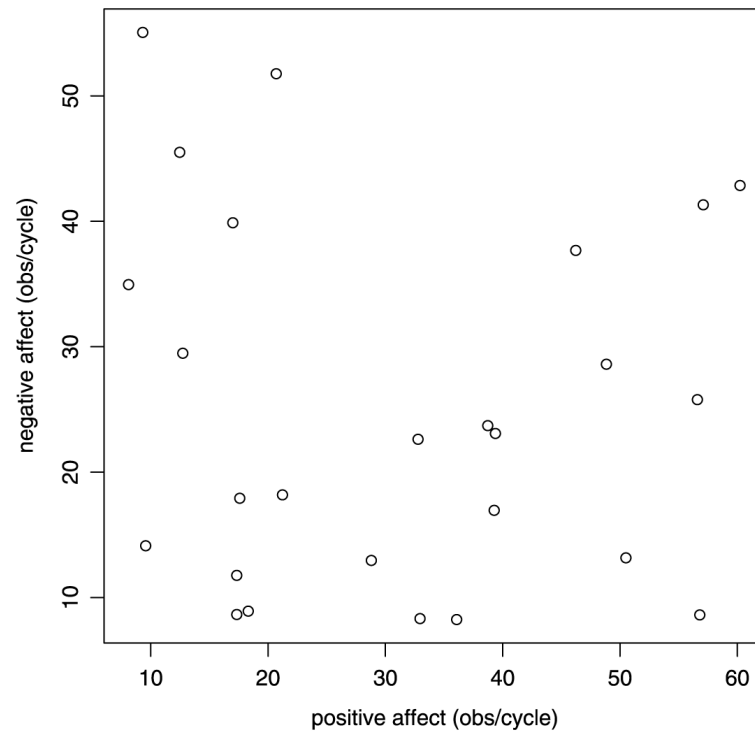
Plots of linear oscillator trajectories without (a) and with (b) phase resetting. The state space for linear oscillators a and b are shown in figures c and d, respectively. The state space plots include a line produced by a linear regression of x on \ddot{x} .

**FIGURE 3.**

Graphs suggested by Boker and Nesselroade (2002) for the selection of τ . In both plots, the optimal τ value is marked by a vertical line.

**FIGURE 4.**

(a) Plot of R^2 versus τ suggested by Boker & Nesselroade (2002). (b) Plot of the proportion of surrogate time series that produced an R^2 estimate smaller than that of the original data set. (c & d) Fitting pulses to a plot such as that shown in (b). By varying the pulse length, pulses that described the asymptote of values could be identified.

**FIGURE 5.**

Plot of the estimated η values for positive and negative affect. Due to the nonlinear nature of η estimates, the estimates have been converted into the estimated number of observations measured per cycle.

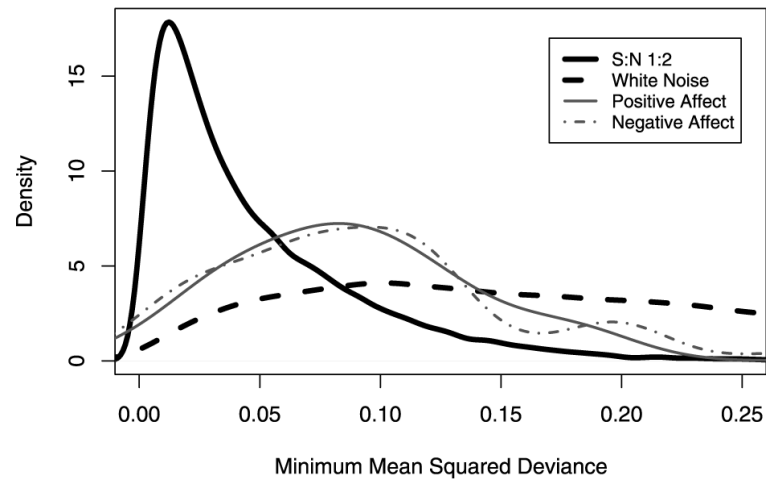
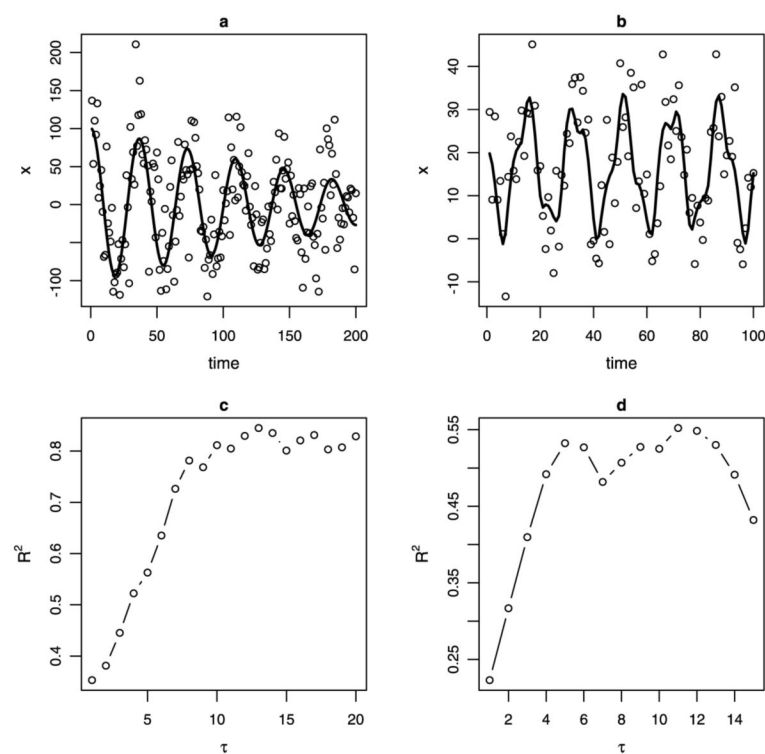


FIGURE 6.

Plot of the minimum mean squared deviance of the positive and negative affect time series relative to results produced in the simulations presented earlier.

**FIGURE 7.**

Plots of two systems with nonlinear damping: $\ddot{x} = \eta x + \zeta \text{sng}(\dot{x})$ (a) and $\ddot{x} = \eta x + \zeta \text{remainder}(t/7)^4$ (b). The line is the true system, whereas the points are the system with measurement error. The estimated variance versus τ for each system is shown in c and d. Note the similarity to earlier figures of a damped linear oscillator [$\eta_a = -0.03$, $\zeta_a = -0.1$, $\eta_b = -0.12$, $\zeta_b = 10$].

TABLE 1

Estimated Values of η

$P(\text{reset})$	$\text{True } \eta$	<i>Signal to Noise Ratio of 2:1 (moderate measurement error)</i>					
		<i>Individual τs</i>			<i>Fixed $\tau = 2$</i>		
		<i>Median</i>	<i>M</i>	<i>25-75%ile</i>	<i>MSE</i>	<i>Median</i>	<i>MSE</i>
0%	-0.4	-0.3811	-0.3610	{-0.4139, -0.2775}	6.78e-03	-0.3986	1.47e-03
	-0.2	-0.1866	-0.1840	{-0.1979, -0.1743}	8.64e-04	-0.289	1.03e-02
	-0.1	-0.0973	-0.0964	{-0.1046, -0.0897}	1.86e-04	-0.2338	2.08e-02
	-0.05	-0.0502	-0.0509	{-0.0548, -0.0461}	4.28e-05	-0.198	2.49e-02
	-0.025	-0.0269	-0.0271	{-0.0290, -0.0249}	1.50e-05	-0.1792	2.60e-02
	-0.0125	-0.0139	-0.0140	{-0.0150, -0.0129}	4.99e-06	-0.1714	2.72e-02
	-0.00625	-0.0071	-0.0072	{-0.0077, -0.0067}	1.46e-06	-0.1688	2.75e-02
	-0.4	-0.3399	-0.3377	{-0.4139, -0.2775}	1.37e-02	-0.4077	3.12e-03
	-0.2	-0.1835	-0.1759	{-0.1979, -0.1743}	3.04e-03	-0.3077	1.61e-02
	-0.1	-0.0967	-0.0974	{-0.1046, -0.0897}	1.20e-03	-0.2529	2.81e-02
5%	-0.05	-0.0501	-0.0537	{-0.0548, -0.0461}	4.40e-04	-0.2269	3.41e-02
	-0.025	-0.0275	-0.0304	{-0.0290, -0.0249}	2.81e-04	-0.2048	3.69e-02
	-0.0125	-0.0148	-0.0181	{-0.0150, -0.0129}	1.78e-04	-0.1953	3.70e-02
	-0.00625	-0.0098	-0.0124	{-0.0077, -0.0067}	1.32e-04	-0.19	3.67e-02

Note. Estimates average over all values of ζ for the set of conditions presented. The results are from the shortest time series length; central tendency estimates were similar for other lengths. $P(\text{reset})$ corresponds to the probability of a phase reset at any observation.

TABLE 2

Estimated Values of η (continued)

Signal to Noise Ratio of 1:2 (large amount of measurement error)							
P(reset)	True η	Individual τ s					
		25-75%ile			Fixed $\tau = 2$		
		Median	M		MSE	Median	MSE
0%	-0.4	-0.2639	-0.3137	{-0.4337, -0.2313}	2.39e-02	-0.4505	7.66e-03
	-0.2	-0.1846	-0.1911	{-0.2235, -0.1348}	7.48e-03	-0.3934	4.52e-02
	-0.1	-0.0940	-0.1139	{-0.1194, -0.0777}	6.32e-03	-0.3745	8.07e-02
	-0.05	-0.0523	-0.0629	{-0.0644, -0.0451}	2.77e-03	-0.3509	9.61e-02
	-0.025	-0.0286	-0.0325	{-0.0338, -0.0233}	1.14e-03	-0.3452	1.05e-01
	-0.0125	-0.0153	-0.0164	{-0.0181, -0.0127}	1.49e-04	-0.3338	1.08e-01
	-0.00625	-0.0078	-0.0082	{-0.0093, -0.0066}	1.08e-05	-0.3318	1.10e-01
5%	-0.4	-0.2600	-0.3111	{-0.4337, -0.2313}	2.62e-02	-0.459	9.22e-03
	-0.2	-0.2011	-0.2125	{-0.2235, -0.1348}	1.45e-02	-0.4017	5.00e-02
	-0.1	-0.1064	-0.1273	{-0.1194, -0.0777}	9.04e-03	-0.3716	8.21e-02
	-0.05	-0.0561	-0.0768	{-0.0644, -0.0451}	5.13e-03	-0.3584	1.05e-01
	-0.025	-0.0313	-0.0454	{-0.0338, -0.0233}	3.16e-03	-0.3493	1.12e-01
	-0.0125	-0.0182	-0.0277	{-0.0181, -0.0127}	1.99e-03	-0.3467	1.14e-01
	-0.00625	-0.0113	-0.0166	{-0.0093, -0.0066}	4.32e-04	-0.336	1.12e-01

Note. Estimates average over all values of ζ for the set of conditions presented. The results are from the shortest time series length; central tendency estimates were similar for other lengths. *P(reset)* corresponds to the probability of a phase reset at any observation.

TABLE 3

Estimated Values of ζ

Individual τ s							Fixed $\tau = 2$	
Probability of Phase Reset 0%, S:N Ratio 1:2								
True ζ	Median	M	25-75%ile		MSE	Median	MSE	
-0.02	-0.0160	-0.0158	{-0.0222, -0.0100}		6.20e-04	-0.0095	1.50e-03	
-0.01	-0.0081	-0.0082	{-0.0138, -0.0028}		6.60e-04	-0.0063	1.30e-03	
-0.00	-0.0001	-0.0002	{-0.0049, -0.0056}		5.72e-04	-0.0006	1.32e-03	
0.01	0.0081	0.0080	{ 0.0025, 0.0141}		6.41e-04	0.0054	1.38e-03	
0.02	0.0159	0.0162	{ 0.0101, 0.0221}		6.64e-04	0.0102	1.58e-03	
Probability of Phase Reset 5%, S:N Ratio 1:2								
-0.02	-0.0031	-0.0046	{-0.0222, -0.0100}		9.91e-04	-0.0024	1.72e-03	
-0.01	-0.0016	-0.0022	{-0.0138, -0.0028}		7.76e-04	-0.0006	1.32e-03	
0.00	0.0000	0.0002	{-0.0049, -0.0056}		7.08e-04	-0.0005	1.16e-03	
0.01	0.0012	0.0021	{0.0025, 0.0141}		7.07e-04	0.0012	1.37e-03	
0.02	0.0032	0.0051	{0.0101, 0.0221}		9.17e-04	0.0027	1.60e-03	

Note. Estimates values of ζ average over all values of η as well as time series lengths.

TABLE 4

Observed MMSD Percentiles

S:N Ratio	Percentile						
	5th	10th	25th	50th	75th	90th	95th
2:1	0.0004	0.0010	0.0028	0.0063	0.0128	0.0264	0.0493
1:1	0.0013	0.0025	0.0056	0.0118	0.0242	0.0523	0.0791
1:2	0.0041	0.0069	0.0139	0.0293	0.0591	0.0988	0.1278
0:∞	0.0306	0.0471	0.0887	0.1525	0.2291	0.2900	0.3234

Note. The presented values average over all conditions not otherwise specified. The signal to noise ratios are as listed in the previous simulation, except for the ratio 0:∞, which represents white noise.