British Journal of Mathematical and Statistical Psychology (2013), 66, 103–126

The British Psychological Society

www.wileyonlinelibrary.com

Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes

Manuel C. Voelkle[1]* and Johan H. L. Oud[2]

[1]Max Planck Institute for Human Development, Berlin, Germany
[2]Padjadjaran University, Bandung, Indonesia

When designing longitudinal studies, researchers often aim at equal intervals. In practice, however, this goal is hardly ever met, with different time intervals between assessment waves and different time intervals between individuals being more the rule than the exception. One of the reasons for the introduction of continuous time models by means of structural equation modelling has been to deal with irregularly spaced assessment waves (e.g., Oud & Delsing, 2010). In the present paper we extend the approach to *individually* varying time intervals for oscillating and non-oscillating processes. In addition, we show not only that equal intervals are unnecessary but also that it can be advantageous to use unequal sampling intervals, in particular when the sampling rate is low. Two examples are provided to support our arguments. In the first example we compare a continuous time model of a bivariate coupled process with varying time intervals to a standard discrete time model to illustrate the importance of accounting for the exact time intervals. In the second example the effect of different sampling intervals on estimating a damped linear oscillator is investigated by means of a Monte Carlo simulation. We conclude that it is important to account for individually varying time intervals, and encourage researchers to conceive of longitudinal studies with different time intervals within and between individuals as an opportunity rather than a problem.

1. Three common beliefs about longitudinal studies

In recent years, the behavioural sciences have seen an impressive increase in the number and scale of longitudinal studies. Key advantages of longitudinal data are not only a more comprehensive measurement of the phenomenon in question, but also a better assessment of causal effects and increased statistical power. Furthermore, longitudinal data are the *sine qua non* for separating between- and within-person effects (McArdle, 2009; Molenaar & Campbell, 2009; Curran & Bauer, 2011). However, when designing

*Correspondence should be addressed to Manuel C. Voelkle, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany (e-mail: voelkle@mpib-berlin.mpg.de)

DOI:10.1111/j.2044-8317.2012.02043.x

and analysing longitudinal studies, researchers often still rely on their intuition and common practice. This is true when it comes to finding cost- and power-optimal designs in terms of the number of participants, time points, and indicators (e.g., Allison, Allison, Faith, Paultre, & Pi-Sunyer, 1997; Oertzen, 2010), as well as when it comes to choosing and dealing with different time *intervals* between and within individuals. In the present paper we focus on this latter aspect by extending current continuous time models to account for individually varying time intervals for oscillating and non-oscillating processes. In addition, we show not only that equal intervals are unnecessary, but also that it can be advantageous to use unequal sampling intervals, in particular when the sampling rate is low.

Ranging from large panel surveys with many participants but comparatively few time points (e.g., the National Longitudinal Surveys in the USA) to ambulatory assessments (Fahrenberg & Myrtek, 1996; Hektner, Schmidt, & Csikszentmihalyi, 2007) with fewer participants but many time points, researchers often aim at equal intervals when designing longitudinal studies. Deviations from an equally spaced design are often considered a nuisance and are treated as a concession to practice, where different time intervals between assessment waves and different time intervals between individuals can hardly ever be avoided. We ascribe the predominance of studies with (presumably) equal time intervals to three common beliefs about longitudinal studies: first, equal intervals are important for parameter comparisons; second, equal intervals are optimal for detecting the generating process; and third, equal intervals are necessary for statistical analyses. In this section we will show that while each of these arguments contains a grain of truth, ultimately they are all wrong. Even worse, adhering to them may result in little insight and incorrect conclusions and is ultimately a waste of time and resources.

1.1. Equal intervals are important for parameter comparisons

The most naïve point of view is that the length of time intervals just does not matter. This, however, can be easily proven wrong by considering Figure 1, which shows a bivariate autoregressive cross-lagged model with time-invariant parameters. In Figure 1(a), the model was fitted to an artificial data set that was simulated with a time interval of $\Delta t = 1$, resulting in an autoregressive effect of $\eta_{1,t-1}$ on $\eta_{1,t}$ ($a_{11} = 0.52$) and a cross-lagged effect of $\eta_{2,t-1}$ on $\eta_{1,t}$ ($a_{12} = 0.25$). In Figure 1(b), the same model was fitted to a data set that differed from the first one only in the sampling interval, now $\Delta t = 2$. Even though the underlying process is exactly the same, the discrete time parameters have changed (they may even change order of magnitude and sign; cf. Oud & Delsing, 2010). For example, while the autoregressive effect became smaller ($a_{11} = 0.30$), the cross-lagged effect became larger ($a_{12} = 0.32$). As indicated by the normal distribution in Figure 1(c), time intervals may even vary across individuals (Δt_i). In this case, the discrete time parameters are different for each and every person $i = 1, \ldots, N$. Finally, one could imagine a study which combines different time intervals within *and* between individuals (Figure 1(a)–(c)). In this case it is hard to say how the individual discrete time parameters will look. Clearly, however, they will be quite different from a model with equal intervals.

A more sophisticated argument is that the intervals should be equal, because otherwise parameter estimates cannot be compared. As is apparent from Figure 1, the relationship between the parameters obtained from models with different intervals is not straightforward (as noted above, in this example the autoregressive effects go down with increasing time intervals, while the cross-lagged effects go up). As a matter of fact,

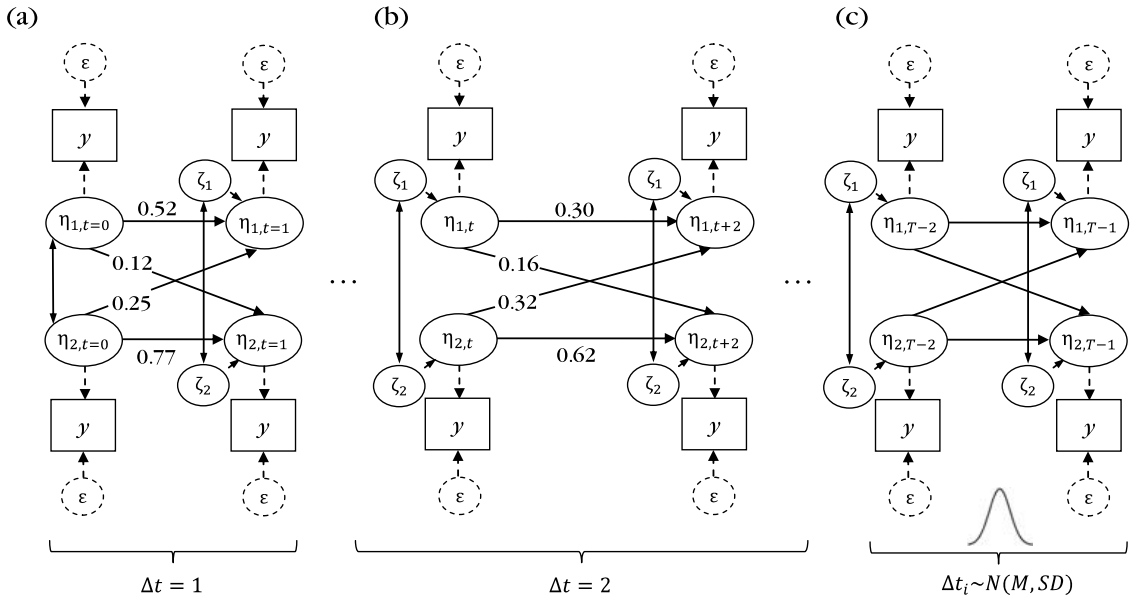

Figure 1. Bivariate autoregressive cross-lagged model with time-invariant parameters. Part (a) shows a model with discrete time intervals $\Delta t = 1$. Part (b) shows a model with discrete time intervals $\Delta t = 2$. Part (c) shows a model with individually varying time intervals. In this example, the intervals Δt_i for person $i = 1, \ldots, N$ were drawn from a normal distribution with $\Delta t_i \sim N(M, SD)$. For reasons of simplicity only two time points are shown for each (sub)model. The dots indicate that the series actually continues. Squares represent manifest variables, and ellipses latent variables. Single-headed arrows represent direct effects and double-headed arrows covariances. Each latent variable may be measured by one (as illustrated for the present paper in the figure) or multiple manifest variables. Parameters corresponding to the dashed lines (e.g., measurement errors) were not estimated in the present paper, although the model could be easily extended.

not being able to integrate findings across different studies with different time intervals is a severe limitation to scientific progress. Thus, unless there exists a method that permits the estimation of parameters which are independent of the time intervals chosen in any particular study, this is indeed a valid argument (cf. Gollob & Reichardt, 1987). Fortunately, however, one of the reasons why *continuous time modelling* has been developed is exactly to deal with this problem in time series ($N = 1$) and panel ($N > 1$) data. After quickly reviewing the approach, in the present paper we will generalize it to individually varying time intervals.

1.2. Equal intervals are optimal for detecting the generating process

Another argument for equal intervals is that they are optimal for detecting the true underlying (generating) model. For this purpose, it seems plausible to keep inter-individual differences as small as possible by assessing everyone at the same point in time (e.g., right before the summer break when assessing scholastic achievement, or every morning at 10 a.m. in a daily ambulatory assessment study). Indeed, if for any reason one is *exclusively* interested in exactly these discrete time points, this is the best thing to do. Arguably, however, researchers are interested in the underlying continuous process, and it is only due to practical limitations that the entire process cannot be observed continuously but has to be inferred from discrete 'snapshots' in time. In this

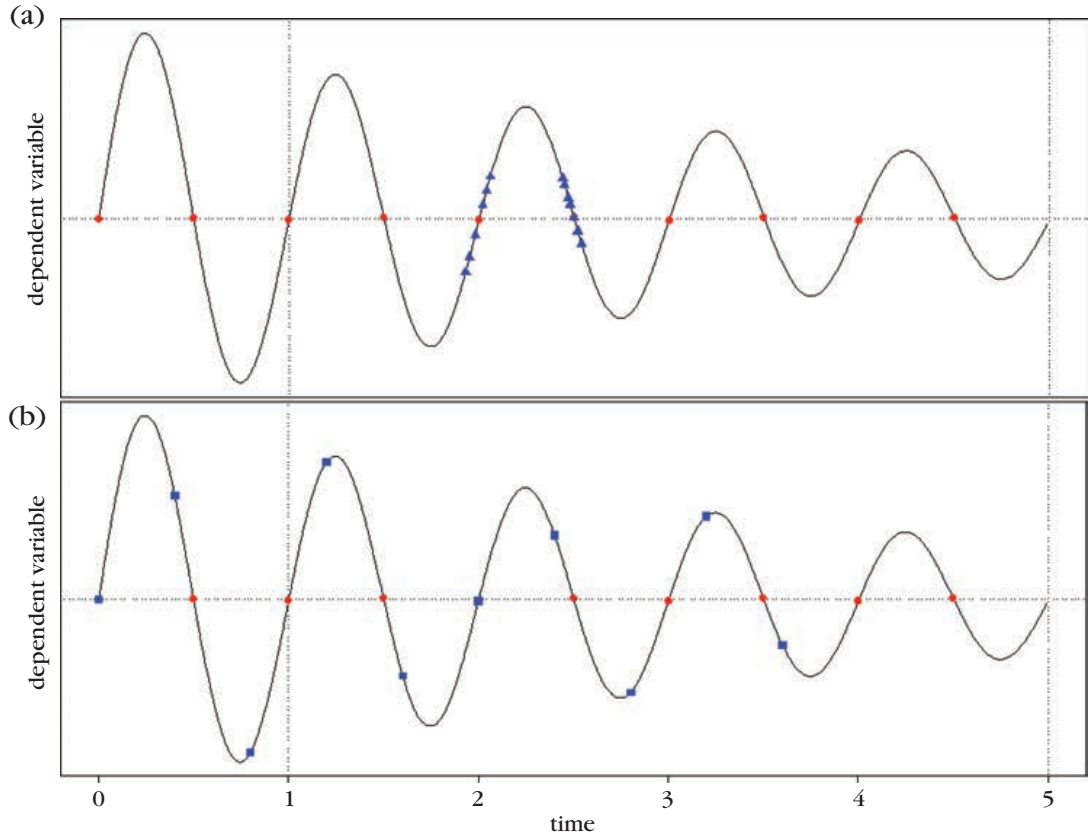

Figure 2. Damped oscillating process with different sampling intervals. The dots correspond to discrete time observations with a sampling rate of 2 Hz, which is just over two assessments per period (note that at later time points, the dots are always slightly above or below the horizontal line). The triangles in part (a) represent few selected individuals with individually varying time intervals between assessment waves 5 and 6. The squares in part (b) illustrate a higher sampling rate of 2.5 Hz.

case, however, equal intervals may be less informative of the true underlying function than unequal intervals. To illustrate this point, consider Figure 2(a), which shows a damped linear oscillator. Without the dampening this could, for example, correspond to the human circadian rhythm (Refinetti, 2006), assessed via the core body temperature over several days. The core body temperature typically goes up in the morning, peaks during the day, and reaches its lowest point in the middle of the night, before rising to the original level again.[1] With an average sampling rate of just over two assessments per period (i.e., for a frequency of 1 hertz (Hz) in Figure 2(a), the sampling rate would need to be greater than 2 Hz, which corresponds to the so-called (sub)Nyquist–Shannon sampling theorem; cf. Lueke, 1999), the generating process could be perfectly reconstructed from the discrete time observations marked as dots in Figure 2(a).[2] With a lower sampling rate,

[1] The circadian rhythm is an endogenously driven cyclic process of approximately 24 hours. It persists even in the absence of environmental cues (Takahashi & Zatz, 1982). In the present paper we limit ourselves to such endogenously driven oscillators. However, although circadian rhythms are generated endogenously, they may be adjusted by external factors (e.g., a transatlantic flight). In this case it would be necessary to represent these external factors as additional predictors in equations (7) and (8) in Section 3.1 below (the estimation of effects of exogenous variables as piecewise functions is discussed, for example, by Singer, 1992).

[2] At least for an infinite sequence of samples. As will be discussed below, when considering the dampening factor the oscillation is actually somewhat lower than 2 Hz in this example.

the true generating process could no longer be uniquely identified, but would become indistinguishable from other processes (so-called aliases). However, just as the reader will find it difficult to detect the true oscillating process based solely on the (equally spaced and identical across all subjects) dots, so will most statistical procedures. As will be shown further below, in such situations it is advantageous if the time of assessment is *not* always the same. For example, if individuals vary around the same average point of measurement (as exemplified for two time points by the triangles in Figure 2(a)), this would provide additional information on the actual process and would thus help to identify the generating process.

1.3. Equal intervals are necessary for statistical analyses

A final argument for the design of studies with equal intervals is that equal intervals are required by many common statistical methods. Violating this assumption may result in biased parameter estimates and incorrect test statistics. This, for example, applies to standard (vector-)autoregressive and cross-lagged models for time series and panel data as illustrated in Figure 1 (cf. Lütkepohl, 2005), P-technique factor analysis and recent advancements thereof (Cattell, Cattell, & Rhymer, 1947; Nesselroade, Gerstorf, Hardy, & Ram, 2007; Molenaar & Nesselroade, 2009), dynamic factor analysis, analyses employing time delayed embedding, idiographic filters, and unified structural equation modelling (Molenaar, 1985; Nesselroade *et al.*, 2007; Gates, Molenaar, Hillary, Ram, & Rovine, 2010; Gates, Molenaar, Hillary, & Slobounov, 2011). Other methods, such as mixed effects and multilevel models for longitudinal data (Pinheiro & Bates, 2000; Hox, 2002) explicitly account for individually varying time intervals in non-oscillating processes, but have not been extended to oscillating processes with unequal time intervals (see Butner, Amazeen, & Mulvey, 2005, for an example with equal intervals based on local linear approximations). Still other methods (generally referred to as latent differential equations in the sense of Boker & Nesselroade, 2002; Boker, Neale, & Rausch, 2004; Boker, 2007a, 2007b; Boker, Deboeck, Edler, & Keel, 2010; Deboeck, 2010) were designed to model oscillating processes but differ from the approach introduced in the present paper in other respects (see Oud, 2007a; Oud & Folmer, 2011, for a comparison of the two techniques).

In summary, it is true that many common statistical methods do not account for individually varying intervals. Other methods account for individually varying intervals but are not always suited to address the research questions one is interested in (e.g., fitting oscillating processes). Thus, designing longitudinal studies with equal intervals seems indeed reasonable unless a method can be found that extends existing discrete time models to account for unequal intervals. Such a method – continuous time modelling by means of differential equations – was proposed a long time ago (Bartlett, 1946; Bergstrom, 1984, 1988), but it is not well known in the psychological literature (but see Oud & Jansen, 2000). In the following we will quickly review its basic idea from a structural equation modelling (SEM) perspective, before extending the approach to individually varying intervals.

2. Continuous time modelling for non-oscillating processes

The basics of continuous time modelling by means of SEM can be summarized in five steps. A more detailed description is provided by Oud and Delsing (2010). A less technical introduction to the approach for a broader psychological readership is

provided by Voelkle, Oud, Davidov, and Schmidt (in press), who also provide a more detailed description of the five steps summarized below. Technical aspects and model identification are discussed by Oud and Jansen (2000). For an illustrative empirical example on the relationship between decoding skill and reading comprehension, see Oud (2007b). Other examples include the analysis of the relationship between externalizing and internalizing problem behaviour, and the analysis of the relationships between individualism, nationalism, and ethnocentrism in Flanders (Oud & Delsing, 2010). For an investigation of the reciprocal influence between antisocial behaviour and depressive symptoms by means of a closely related model (the so-called continuous time autoregressive latent trajectory model), see Delsing and Oud (2008).

2.1. Continuous time modelling in five steps

First, the discrete time vector-autoregressive and cross-lagged model is defined as

$$\boldsymbol{\eta}(t) = \mathbf{A}_{\Delta t}\boldsymbol{\eta}(t - \Delta t) + \mathbf{b}_{\Delta t} + \mathbf{w}(t - \Delta t), \tag{1}$$

with vector $\boldsymbol{\eta}(t) \in \mathbb{R}^{p\times 1}$ representing the p-variate outcome vector at time point t, which is regressed on $\boldsymbol{\eta}(t - \Delta t)$ at the immediately prior point in time, being separated by Δt time units from $\boldsymbol{\eta}(t)$. The matrix $\mathbf{A}_{\Delta t} \in \mathbb{R}^{p\times p}$ contains the autoregressive effects on the main diagonal and the cross-lagged effects in the off-diagonals. As illustrated above, the discrete time parameters in $\mathbf{A}_{\Delta t}$ depend on the length of the time interval, as indicated by the subscript Δt, but are otherwise assumed to be time invariant. The same applies to the intercept vector $\mathbf{b}_{\Delta t} \in \mathbb{R}^{p\times 1}$ and the discrete time error term $\mathbf{w}(t - \Delta t) \in \mathbb{R}^{p\times 1}$ with covariance matrix $\mathbf{Q}_{\Delta t} \in \mathbb{R}^{p\times p}$. The outcome vectors $\boldsymbol{\eta}(t)$, $\boldsymbol{\eta}(t - \Delta t)$ may be either directly observed or latent with measurement model

$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\epsilon}(t), \qquad \text{cov}(\boldsymbol{\epsilon}) = \boldsymbol{\Theta}. \tag{2}$$

The factor loading matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{q\times p}$ relates the q observed variables at each time point ($\mathbf{y}(t)$) to the p latent variables. Finally, $\boldsymbol{\epsilon}(t) \in \mathbb{R}^{q\times 1}$ is the corresponding vector of measurement errors with $\text{cov}(\boldsymbol{\epsilon}) = \boldsymbol{\Theta}$. The measurement model is usually, but not necessarily, assumed to be time invariant. For reasons of simplicity, we do not explicitly consider the measurement model in the present paper.

Second, the stochastic differential equation corresponding to equation (1), and from which equation (1) is derived as a solution, can be expressed as

$$\frac{\mathrm{d}\boldsymbol{\eta}(t)}{\mathrm{d}t} = \mathbf{A}\boldsymbol{\eta}(t) + \mathbf{b} + \mathbf{G}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}. \tag{3}$$

In contrast to equation (1), the coefficient matrix $\mathbf{A}$ – the so-called drift matrix – and the intercept vector $\mathbf{b}$ are no longer dependent on the time interval. $\mathbf{W}(t)$ denotes the random walk process in continuous time (the so-called Wiener process). In order to model arbitrary variances and covariances among the error terms, $\mathbf{W}(t)$ is premultiplied by the Cholesky triangle $\mathbf{G} \in \mathbb{R}^{p\times p}$ so that the resulting error covariance matrix in continuous time – the so-called diffusion matrix – is $\mathbf{Q} = \mathbf{G}\mathbf{G}^{\mathrm{T}}$, with T denoting the transposition operator.

Third, for initial value $\boldsymbol{\eta}(t_0)$ and any time interval $\Delta t = t - t_0$, the solution of the stochastic differential equation (3) is

$$\boldsymbol{\eta}(t) = e^{\mathbf{A}\cdot(t-t_0)}\boldsymbol{\eta}(t_0) + \mathbf{A}^{-1}\left[e^{\mathbf{A}\cdot(t-t_0)} - \mathbf{I}\right]\mathbf{b} + \int_{t_0}^{t} e^{\mathbf{A}\cdot(t-s)}\mathbf{G}d\mathbf{W}(s), \tag{4}$$

with

$$\text{cov}\left[\int_{t_0}^{t} e^{\mathbf{A}\cdot(t-s)}\mathbf{G}d\mathbf{W}(s)\right] = \int_{t_0}^{t} e^{\mathbf{A}\cdot(t-s)}\mathbf{Q}e^{\mathbf{A}^{\mathrm{T}}\cdot(t-s)}ds = \text{irow}\left\{\mathbf{A}_{\#}^{-1}\left[e^{\mathbf{A}_{\#}\cdot(t-t_0)} - \mathbf{I}\right]\text{row}\mathbf{Q}\right\}$$

for $\mathbf{Q} = \mathbf{G}\mathbf{G}^{\mathrm{T}}$ and $\mathbf{A}_{\#} = \mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}$. The row operator puts the elements of the $\mathbf{Q}$ matrix row-wise into a column vector, while irow represents the inverse operation (for details, see Arnold, 1974; Ruymgaart & Soong, 1985; Oud & Jansen, 2000).

Fourth, for the *observed* time intervals in a given study the discrete time coefficients in equation (1) are constrained to the underlying continuous time parameters as shown in equation (4), so that

$$\mathbf{A}_{\Delta t} = e^{\mathbf{A}\cdot\Delta t},$$

$$\mathbf{b}_{\Delta t} = \mathbf{A}^{-1}\left[e^{\mathbf{A}\cdot\Delta t} - \mathbf{I}\right]\mathbf{b},$$

with discrete time error covariance matrix

$$\mathbf{Q}_{\Delta t} = \text{irow}\left\{\mathbf{A}_{\#}^{-1}[e^{\mathbf{A}_{\#}\cdot\Delta t} - \mathbf{I}]\text{row}\mathbf{Q}\right\}.$$

Fifth, the model given in equation (1), but with the discrete time parameters constrained to the underlying continuous time parameters as defined above (i.e., the continuous time model), is translated into an SEM framework – as demonstrated in the next paragraph – with

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}, \quad \text{cov}(\boldsymbol{\zeta}) = \boldsymbol{\Psi}. \tag{5}$$

Given equations (2) and (5), it is easy to derive the model implied covariance matrix $\boldsymbol{\Sigma}$ (e.g., Bollen, 1989, p. 325). Parameter estimates are then obtained by minimizing the raw maximum likelihood function

$$RML = \sum_{i=1}^{N}\left[m_i \log(2\pi) + \log(|\boldsymbol{\Sigma}_i|) + (\mathbf{y}_i - \boldsymbol{\mu}_i)\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i)^{\mathrm{T}}\right], \tag{6}$$

with $\mathbf{y}$ representing the vector of observed variables as defined above, and $\boldsymbol{\mu}$ the corresponding mean vector. Note that missing values are easily handled by allowing the number of variables ($m = p \cdot T$), $\mathbf{y}_i$, $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ to differ across all individuals (i.e., full information maximum likelihood estimation; cf. Arbuckle, 1996). However, before returning to the definition of the matrices in equation (5), we will first extend the approach to individually varying time intervals.

2.2. *Individually varying time intervals*

Individually varying time intervals are easily introduced into the model by allowing t and Δt to be different for each individual i and time point j, that is, $t_{i,j}$ and $\Delta t_{i,j}$. As a consequence, all discrete time parameters in equation (1) may also differ across individuals (i.e., $\mathbf{A}_{\Delta t_{i,j}}$, $\mathbf{b}_{\Delta t_{i,j}}$). Only the underlying continuous time parameters remain the same. In terms of the SEM equation (5), for $j = 0, 1, \ldots, T-1$ discrete time points (t) and $i = 1, \ldots, N$ independent individuals, the final continuous time model with individually varying time intervals can thus be defined as

$$\boldsymbol{\eta}^{\mathrm{T}} = \left(\boldsymbol{\eta}(t_{i,j=0})^{\mathrm{T}} \; \boldsymbol{\eta}(t_{i,j=1})^{\mathrm{T}} \; \boldsymbol{\eta}(t_{i,j=2})^{\mathrm{T}} \; \ldots \; \boldsymbol{\eta}(t_{i,j=T-1})^{\mathrm{T}} \; 1 \right)^{\mathrm{T}},$$

with $\boldsymbol{\eta}$ being a $(p \cdot T + 1) \times 1$ vector, which contains the p-variate outcome vectors $\boldsymbol{\eta}(t_{i,j})$ that have been stacked above each other across all T time points, plus the constant 1. Likewise, $\mathbf{B}$ is defined as

$$\mathrm{B} = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 & \boldsymbol{\mu}_{\boldsymbol{\eta}(t_{i,j=0})} \\ e^{\mathbf{A}\cdot\Delta t_{i,j=1}} & 0 & \ldots & 0 & 0 & \mathbf{A}^{-1}\left[e^{\mathbf{A}\cdot\Delta t_{i,j=1}} - \mathbf{I}\right]\mathbf{b} \\ 0 & e^{\mathbf{A}\cdot\Delta t_{i,j=2}} & \ldots & 0 & 0 & \mathbf{A}^{-1}\left[e^{\mathbf{A}\cdot\Delta t_{i,j=2}} - \mathbf{I}\right]\mathbf{b} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & e^{\mathbf{A}\cdot\Delta t_{i,j=T-1}} & 0 & \mathbf{A}^{-1}\left[e^{\mathbf{A}\cdot\Delta t_{i,j=T-1}} - \mathbf{I}\right]\mathbf{b} \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

The error vector $\boldsymbol{\zeta}$ and the corresponding covariance matrix $\boldsymbol{\Psi}$ are

$$\boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\eta}(t_{i,j=0}) - \boldsymbol{\mu}_{\boldsymbol{\eta}(t_{i,j=0})} \\ \mathbf{w}(t_{i,j=1} - \Delta t_{i,j=1}) \\ \mathbf{w}(t_{i,j=2} - \Delta t_{i,j=2}) \\ \vdots \\ \mathbf{w}(t_{i,j=T-1} - \Delta t_{i,j=T-1}) \\ 1 \end{pmatrix}, \quad \boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Phi}_{\boldsymbol{\eta}(t_{i,j=0})} & & & & & \\ 0 & \mathbf{Q}_{\Delta t_{i,j=1}} & & & & \\ 0 & 0 & \mathbf{Q}_{\Delta t_{i,j=2}} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ 0 & 0 & 0 & \ldots & \mathbf{Q}_{\Delta t_{i,j=T-1}} & \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix},$$

with $\boldsymbol{\mu}_{\boldsymbol{\eta}(t_{i,j=0})}$ representing the initial mean vector and $\boldsymbol{\Phi}_{\boldsymbol{\eta}(t_{i,j=0})}$ the initial covariance matrix.

3. Continuous time modelling for oscillating processes

For asymptotically stable models of non-oscillating processes the eigenvalues of the drift matrix $\mathbf{A}$ are negative and real-valued. For positive eigenvalues the model would become unstable (i.e., would not converge to an equilibrium position), while complex eigenvalues lead to oscillating processes. If all eigenvalues have negative real part, the result is a damped linear oscillator. Accounting for complex eigenvalues, the continuous time approach with individually varying time intervals introduced above generalizes readily to oscillating processes. In this section, we will first show how to fit a linear oscillator in this way with individually varying time intervals and then discuss

oversampling (Singer, 2012) as an alternative approach to model estimation that avoids the complications in using the matrix exponential.

3.1. The damped linear oscillator

The damped linear oscillator is defined in terms of

$$\frac{\mathrm{d}^2\eta(t)}{\mathrm{d}t^2} = -\omega_0^2\eta(t) - \gamma\frac{\mathrm{d}\eta(t)}{\mathrm{d}t} + b + g\zeta(t) \tag{7}$$

with ω_0 being the angular frequency of the undamped oscillation ($\gamma = 0$) with period length $T_0 = 2\pi/\omega_0$. Thus, the frequency without dampening is $\omega_0/2\pi$.[3] The dampening parameter is denoted by γ and leads for $\gamma > 0$ to dampening, while b is the intercept, and g the strength of the random error $\zeta(t)$. An equivalent way to write equation (7) is in terms of equation (3):

$$\begin{aligned}\begin{pmatrix} \dfrac{\mathrm{d}\eta(t)}{\mathrm{d}t} \\ \dfrac{\mathrm{d}^2\eta(t)}{\mathrm{d}t^2} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{pmatrix}\begin{pmatrix} \eta(t) \\ \dfrac{\mathrm{d}\eta(t)}{\mathrm{d}t} \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & g \end{pmatrix}\begin{pmatrix} \dfrac{\mathrm{d}W_1(t)}{\mathrm{d}t} \\ \dfrac{\mathrm{d}W_2(t)}{\mathrm{d}t} \end{pmatrix} \\ &= \frac{\mathrm{d}\eta(t)}{\mathrm{d}t} = \mathbf{A}\eta(t) + \mathbf{b} + \mathbf{G}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}.\end{aligned} \tag{8}$$

As apparent from equation (8), a second-order differential equation (see equation (7)) is reformulated as a first-order differential equation by adding an extra variable. The 'trick' is to define the first-order derivative as an extra latent variable without an underlying observed variable. This is easily done by setting up a measurement model in terms of equation (2) with

$$y(t) = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} \eta(t) \\ \dfrac{\mathrm{d}\eta(t)}{\mathrm{d}t} \end{pmatrix}.$$

The same idea can be extended to multiple (coupled) oscillating processes (e.g., Oud & Folmer, 2011; Singer, 2012).

For oscillating processes, the eigenvalues of the drift matrix $\mathbf{A} = \left(\begin{smallmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{smallmatrix}\right)$ are

$$\lambda_{1,2} = -\frac{\gamma}{2} \pm \mathrm{i}\sqrt{-\frac{\gamma^2}{4} + \omega_0^2} \tag{9}$$

with $\mathrm{i} = \sqrt{-1}$. Oscillation requires the quantity under the square root sign in equation (9) to be positive. Making use of Euler's formula ($\mathrm{e}^{\mathrm{i}x} = \cos(x) + \mathrm{i}\sin(x)$), the matrix exponential provided above can also be written in terms of the sine and cosine

[3] With dampening ($\gamma > 0$), the period length is $T = 2\pi/\omega$, with ω as defined further below in the text. Thus, the frequency is $\omega/2\pi$.

function. For this purpose we first factorize **A** as in

$$e^{\mathbf{A}\cdot\Delta t_i} = e^{\begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{pmatrix}\cdot\Delta t_i} = e^{\mathbf{M}\mathbf{V}\mathbf{M}^{-1}\cdot\Delta t_i} = \mathbf{M}e^{\mathbf{V}\cdot\Delta t_i}\mathbf{M}^{-1} \tag{10}$$

with

$$\mathbf{M} = \begin{pmatrix} \lambda_1 + \gamma & \lambda_2 + \gamma \\ -\omega_0^2 & -\omega_0^2 \end{pmatrix}$$

being the matrix of eigenvectors of **A**, and **V** a diagonal matrix of the eigenvalues of **A** as defined in equation (9), resulting in

$$e^{\mathbf{A}\cdot\Delta t_i} = e^{-\frac{\gamma}{2}\Delta t_i} \times \begin{bmatrix} \frac{\gamma}{2\omega}\sin(\omega\Delta t_i) + \cos(\omega\Delta t_i) & \frac{1}{\omega}\sin(\omega\Delta t_i) \\ -\frac{\omega_0^2}{\omega}\sin(\omega\Delta t_i) & \cos(\omega\Delta t_i) - \frac{\gamma}{2\omega}\sin(\omega\Delta t_i) \end{bmatrix} \tag{11}$$

with $\omega = \sqrt{-\gamma^2/4 + \omega_0^2}$ (Oud, 2007a; Singer, 2012). The specification of the model as a path diagram is (schematically) illustrated in Figure 3(a).

3.2. Oversampling

Unfortunately, most current SEM packages do not support operations involving complex numbers. While some programs – like OpenMx (Boker *et al.*, 2011) or Mx (Neale, Boker, Xie, & Maes, 2003) – allow the computation of complex eigenvalues, these cannot be used in subsequent operations. At the same time, the estimation of oscillating processes via the eigenvalues and Euler's formula as shown in equation (11) is not only difficult but also computationally demanding. Although the approach has been successfully implemented for a univariate process (Oud, 2007a), it is quite cumbersome to extend it to multivariate processes, possibly involving multiple coupled oscillators. A way to avoid these problems is the approach of *oversampling* (Singer, 1990, 2012).

The basic idea of oversampling is to introduce intervals that are smaller than the actually observed intervals. For example, the observed, and therefore known, length of interval $\Delta t_{i,j}$ for person i and discrete time interval j, can be divided into D smaller intervals of equal size

$$\delta_{i,j,d} = \frac{\Delta t_{i,j}}{D}. \tag{12}$$

For reasons of simplicity we assume the oversampling rate (D) to be identical for all i and j, although this assumption can be easily relaxed if necessary. This allows us to rewrite the matrix exponential function as

$$e^{\mathbf{A}\cdot\Delta t_{i,j}} = \prod_{d=0}^{D-1} e^{\mathbf{A}\cdot\delta_{i,j,d}}. \tag{13}$$

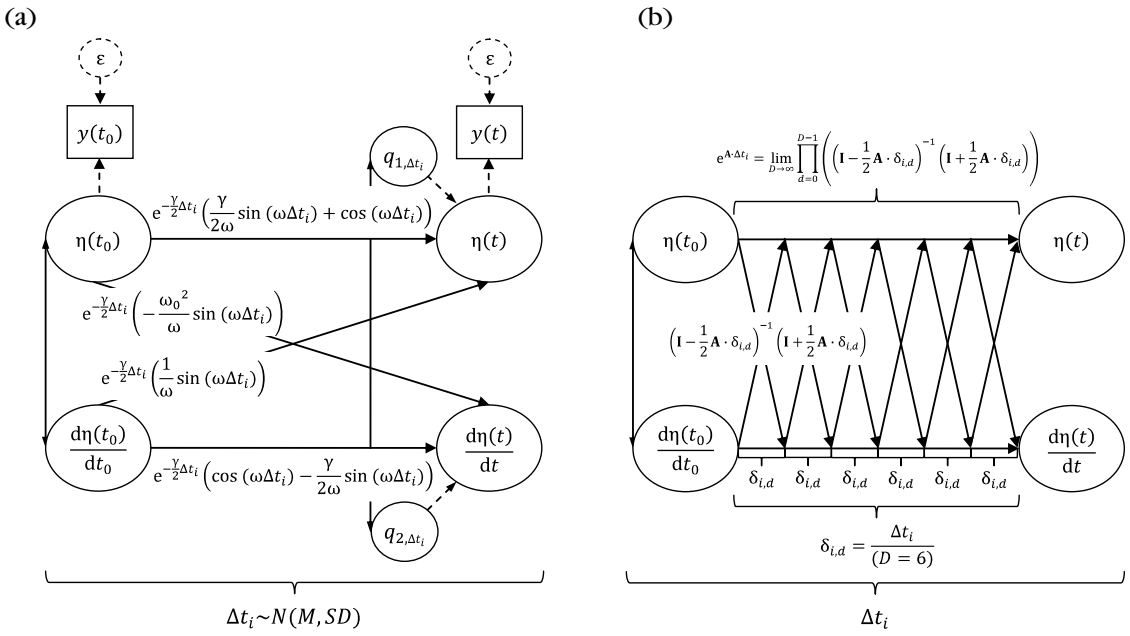

Figure 3. (a) Schematic path diagram of a damped linear oscillator as defined in equations (7) and (8). The derivative is specified as an additional latent variable. For time interval Δt_i, the autoregressive and cross-lagged parameters are constrained according to equation (11). (b) Schematic illustration of the oversampling procedure. In order to avoid computation with complex numbers or the complicated parameter constraints shown in part (a), oversampling (Singer, 2012) is used (here in a slightly different version; see text and Appendix A). The basic idea is to partition the observed time interval Δt_i into D smaller (unobserved) intervals of length $\delta_{i,d} = \Delta t_i / D$. Here $D = 6$. Instead of imposing the complicated parameter constraints on the entire interval, we can now employ much simpler parameter constraints for each of the D oversamples. These simpler constraints are given in equation (15). A detailed description of equation (15) is given in Appendix A. As also shown in Appendix A, by simply taking the product across the parameter constraints on each oversample we approximate the exact solution with arbitrary precision for $D \to \infty$. For reasons of clarity we suppress the subscript j in this figure.

For $D \to \infty$, equation (13) corresponds to

$$e^{\mathbf{A}\cdot\Delta t_{i,j}} = \lim_{D\to\infty} \prod_{d=0}^{D-1} (\mathbf{I} + \mathbf{A}\cdot\delta_{i,j,d}). \tag{14}$$

Thus, the matrix exponential function can be conveniently implemented by equation (14). Note that this holds equally true for oscillating and non-oscillating multivariate processes with any number of variables. Furthermore, since the precision depends only on D, it is up to the user to choose the appropriate level. Otherwise than suggested by Singer (2012), for this purpose it is not necessary to introduce additional latent variables, so apart from an increase in computation time, there are no costs associated with an increase in D.

In what follows we will make use of a variant of the original oversampling procedure, by employing the implicit Heun method (Singer, 1999):

$$e^{\mathbf{A}\cdot\Delta t_{i,j}} = \lim_{D\to\infty}\prod_{d=0}^{D-1}\left(\left(\mathbf{I}-\frac{1}{2}\mathbf{A}\cdot\delta_{i,j,d}\right)^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}\cdot\delta_{i,j,d}\right)\right). \tag{15}$$

Equation (15) is based on the so-called trapezoid approximation and converges much more rapidly than equation (14). For a more detailed description, see Appendix A. In our experience about $D = 30$ oversampling intervals suffice, which for many models is exact up to the precision limit of the SEM program (see also Singer, 2012, who recommends at least 20 oversampling intervals). This issue, however, should be investigated in future simulation studies. The idea of oversampling is (schematically) illustrated in Figure 3(b).

4. A quick summary

After discussing three common beliefs about the design and analysis of longitudinal studies, we reviewed the basic idea of continuous time modelling by means of SEM. We then extended the approach to individually varying time intervals and showed how to estimate oscillating and non-oscillating processes with arbitrary sampling intervals. In addition, a modified version of the oversampling approach (Singer, 2012) was discussed as an efficient way to avoid complex number arithmetic and estimation problems associated with the diagonalization method.

In Sections 5 and 6 we will return to the three common beliefs about longitudinal studies discussed in the beginning by applying the newly introduced approach to two examples. In the first example we show that there is *no* need for intervals to be equal in order to compare parameter estimates, but that it is important to account for the exact time intervals in the analysis. In the second example we show that it may be advantageous if time intervals are *not* the same across individuals. Both examples demonstrate that equal intervals are *not* necessary for statistical analyses. All analyses were conducted using the R-based OpenMx program CT-SEM developed by the authors. The program is available for download from Wiley Online Library.

5. Example 1: A bivariate continuous time model with individually varying time intervals

In this example we first simulated a bivariate continuous time auto- and cross-effects model. Second, we sampled discrete time observations from the continuous time model using eight different interval conditions. Third, we fitted a standard discrete time model and a continuous time model to each of these eight data sets. The continuous time model was estimated using the matrix exponential approach via diagonalization and the oversampling approach.

5.1. Method

A continuous time model as defined in equation (4) was simulated with drift matrix $\mathbf{A} = \left(\begin{smallmatrix}-0.7 & 0.2\\ 0.4 & -0.3\end{smallmatrix}\right)$, continuous time intercept vector $\mathbf{b} = \left(\begin{smallmatrix}0.1\\ 0.2\end{smallmatrix}\right)$, continuous time dynamic error covariance matrix $\mathbf{Q} = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$, initial mean vector $\mathbf{m}_{t_0} = \left(\begin{smallmatrix}1.5\\ 1.0\end{smallmatrix}\right)$, and initial covariance

matrix $\mathbf{\Phi}_{t_0} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. For reasons of simplicity, we assumed no measurement error. The simulated model corresponds to the model illustrated in Figure 1. From this model, data were generated for $N = 200$ cases, $T = 11$ time points, and eight different conditions of discrete time intervals. In the first three conditions, all intervals were equal across individuals: (1) $\Delta t_j = 1$ for all $j = 1, \ldots, 10$; (2) $\Delta t_j = 2$ for all $j = 1, \ldots, 10$; (3) $\Delta t_j = 1$ for $j = 1, \ldots, 5$ and $\Delta t_j = 2$ for $j = 6, \ldots, 10$. In the remaining five conditions, intervals were chosen to be individually varying by drawing them either from a normal distribution – with (4) a mean interval length of 1 and a standard deviation of 0.25 ($\Delta t_{i,j} \sim N(1, 0.25)$) for each individual i and each interval j; (5) $\Delta t_{i,j} \sim N(1, 0.5)$ for each individual i and each interval j – or from a uniform distribution with (6) lower limit 1 and upper limit 2 ($\Delta t_{i,j} \sim U(1, 2)$); (7) lower limit 1 and upper limit 3 ($\Delta t_{i,j} \sim U(1, 3)$); or (8) lower limit 1 and upper limit 10 ($\Delta t_{i,j} \sim U(1, 10)$), for each individual i and each interval j. Thus, when using enough decimal places it may well be that there is not a single interval in the entire data set that occurs twice in the last five conditions. R version 2.13 (R Development Core Team, 2011) was used for data generation and OpenMx version 1.1.1 (Boker *et al.*, 2011) for model estimation. Each data set was fitted twice using first the matrix exponential via diagonalization and second the oversampling approach. For reasons of comparison, the same seed was used. As a consequence, any differences in parameter estimates are due only to different time intervals and method, and not to sampling error.

5.2. Results

Results are presented in Table 1. For $D = 30$, the oversampling approach yielded parameter estimates that were, up to the fourth decimal place, identical to those of the diagonalization method. For this reason, parameter estimates are only reported once. For a single sample of $N = 200$ cases, the drift matrix (**A**) is reasonably close to the population matrix. Note, however, that this is not a Monte Carlo simulation. For $N \to \infty$, or multiple samples, the (average) parameter estimates would converge on the population values, as will be demonstrated in the next example. Most importantly, the drift matrix remains almost identical across all eight discrete time samples. As apparent from the first column in Table 1, this is not the case for the autoregressive cross-lagged matrix of the common discrete time analysis. Even though the generating model was exactly the same, the parameter estimates differ considerably from each other. For example, while the autoregressive effect a_{11} is 0.50 for $\Delta t = 1$, it is 0.28 for $\Delta t = 2$. In contrast to the decreasing autoregressive effect, in this example the cross-lagged effect a_{12} *increases* (e.g., from 0.26 to 0.34; see first two rows in Table 1). For larger differences in time intervals between conditions, differences in parameter estimates may – but do not necessarily have to – become even more extreme. This is apparent when comparing the first and last row in Table 1: with intervals ($\Delta t_{i,j}$) drawn from a uniform distribution with $U(1, 10)$, the discrete time autoregressive effects bear little resemblance to the autoregressive effects for $\Delta t = 1$ (0.50 versus 0.10 and 0.75 versus 0.40; see Table 1), while the cross-lagged effects are quite similar (note that the differences between cross-lagged effects were larger when differences in time intervals were smaller). Thus, there is no point in comparing or interpreting discrete time parameter estimates that are based on different time intervals. Interestingly, though, because the individually varying time intervals were drawn from a normal or uniform distribution, which are both symmetric distributions with mean M, for many different intervals the discrete time

Table 1. Discrete time and continuous time parameter estimates of eight samples with different time intervals. In the last five samples, time intervals were either drawn from a normal distribution with mean 1 and standard deviationn 0.25 or 0.5, or from a uniform distribution with minimum 1 and maximum 2, 3 or 10. Thus, all time intervals differed between persons and measurement occasions in these conditions.

Data generated with:	Discrete time analysis (ignoring Δt)	Continuous time analysis (accounting for Δt)				
	Autoregressive / crossed-lagged matrix	Drift matrix	Continuous time intercepts	Diffusion matrix	Means at t_0	Covariance matrix at t_0
$\Delta t_j = 1$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.50 & 0.26 \\ 0.11 & 0.75 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.73 & 0.17 \\ 0.43 & -0.32 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.10 \\ 0.27 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 0.96 & -0.07 \\ -0.07 & 1.01 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.34 \\ 0.90 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & 0.02 \\ 0.02 & 0.99 \end{pmatrix}$
$\Delta t_j = 2$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.28 & 0.34 \\ 0.13 & 0.61 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.73 & 0.16 \\ 0.44 & -0.30 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.09 \\ 0.25 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 0.96 & -0.07 \\ -0.07 & 1.01 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.34 \\ 0.90 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & 0.02 \\ 0.02 & 0.99 \end{pmatrix}$
$\Delta t_{j=1,\ldots,5} = 1$, & $\Delta t_{j=6,\ldots,10} = 2$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.40 & 0.28 \\ 0.09 & 0.67 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.72 & 0.17 \\ 0.43 & -0.31 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.09 \\ 0.26 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 0.94 & -0.07 \\ -0.07 & 1.01 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.34 \\ 0.90 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & 0.02 \\ 0.02 & 0.99 \end{pmatrix}$
$\Delta t_{i,j}$ with $N(1,0.25)$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.51 & 0.26 \\ 0.11 & 0.76 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.74 & 0.20 \\ 0.44 & -0.32 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.12 \\ 0.24 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 1.01 & -0.05 \\ -0.05 & 1.05 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.39 \\ 0.92 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & 0.02 \\ 0.02 & 0.98 \end{pmatrix}$
$\Delta t_{i,j}$ with $N(1,0.5)$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.54 & 0.24 \\ 0.10 & 0.77 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.76 & 0.21 \\ 0.45 & -0.32 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.12 \\ 0.23 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 1.01 & -0.05 \\ -0.05 & 1.06 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.39 \\ 0.92 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & -0.05 \\ -0.05 & 1.06 \end{pmatrix}$
$\Delta t_{i,j}$ with $U(1,2)$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.36 & 0.31 \\ 0.14 & 0.69 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.75 & 0.20 \\ 0.43 & -0.30 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.10 \\ 0.20 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 1.02 & -0.00 \\ -0.00 & 0.98 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.45 \\ 0.95 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & -0.06 \\ -0.06 & 0.86 \end{pmatrix}$
$\Delta t_{i,j}$ with $U(1,3)$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.28 & 0.33 \\ 0.14 & 0.63 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.75 & 0.19 \\ 0.43 & -0.29 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.10 \\ 0.20 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 1.02 & -0.00 \\ -0.00 & 0.97 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.45 \\ 0.95 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & -0.06 \\ -0.06 & 0.86 \end{pmatrix}$
$\Delta t_{i,j}$ with $U(1,10)$	$\hat{\mathbf{A}}_{\Delta t_j} = \begin{pmatrix} 0.10 & 0.26 \\ 0.10 & 0.40 \end{pmatrix}$	$\hat{\mathbf{A}} = \begin{pmatrix} -0.70 & 0.16 \\ 0.39 & -0.26 \end{pmatrix}$	$\hat{\mathbf{b}} = \begin{pmatrix} 0.11 \\ 0.19 \end{pmatrix}$	$\hat{\mathbf{Q}} = \begin{pmatrix} 0.99 & 0.04 \\ 0.04 & 0.95 \end{pmatrix}$	$\hat{\mathbf{m}}_{t_0} = \begin{pmatrix} 1.45 \\ 0.95 \end{pmatrix}$	$\hat{\Phi}_{t_0} = \begin{pmatrix} 1.01 & -0.06 \\ -0.06 & 0.86 \end{pmatrix}$

Note. $N = 200$; $T = 11$; Reported parameter estimates were obtained using the oversampling approach with $D = 30$ and the matrix exponential approach. Any differences occurred only after the fourth decimal place. Population parameters are: $\mathbf{A} = \left(\begin{smallmatrix} -0.7 & 0.2 \\ 0.4 & -0.3 \end{smallmatrix}\right)$, $\mathbf{b} = \left(\begin{smallmatrix} 0.1 \\ 0.2 \end{smallmatrix}\right)$, $\mathbf{Q} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\mathbf{m}_{t_0} = \left(\begin{smallmatrix} 1.5 \\ 1.0 \end{smallmatrix}\right)$, $\mathbf{\Phi}_{t_0} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Accordingly, the true discrete time autoregressive and crossed-lagged parameters for $\Delta t_j = 1$ and $\Delta t_j = 2$ are $e^{\mathbf{A}\cdot 1} = \mathbf{A}_{\Delta t_j = 1} = \left(\begin{smallmatrix} 0.52 & 0.25 \\ 0.12 & 0.77 \end{smallmatrix}\right)$, $e^{\mathbf{A}\cdot 2} = \mathbf{A}_{\Delta t_j = 2} = \left(\begin{smallmatrix} 0.30 & 0.32 \\ 0.16 & 0.62 \end{smallmatrix}\right)$.

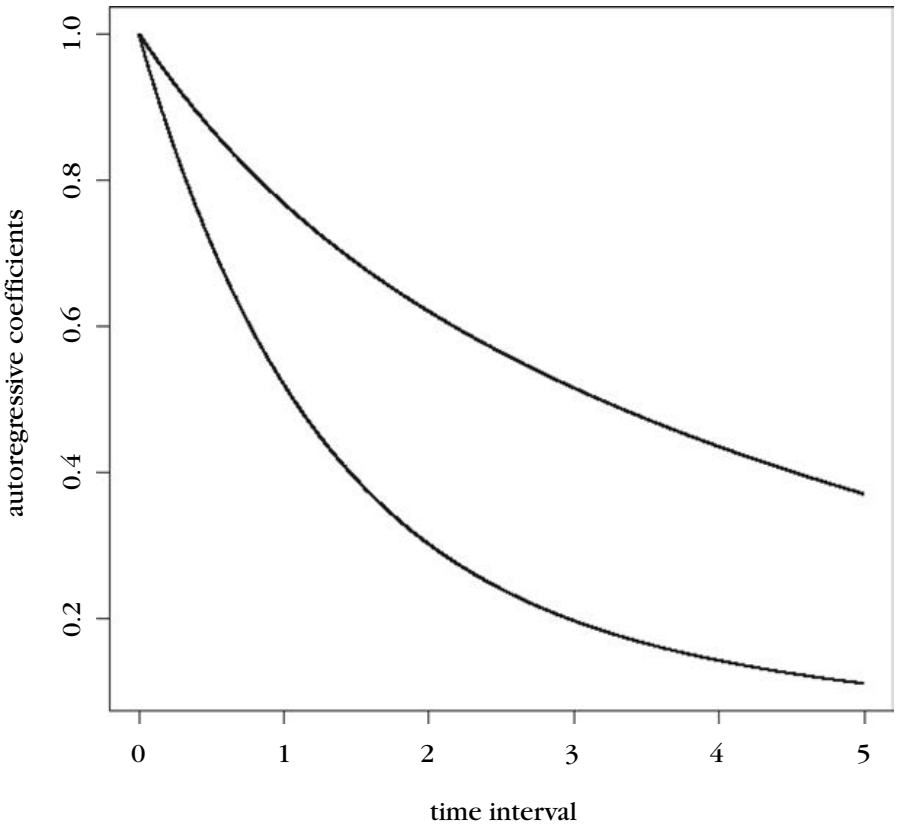

Figure 4. Discrete time autoregressive parameters as a function of the true underlying continuous time model in Example 1.

parameters get close to those of fixed time intervals with length M. This is apparent when comparing condition 1 (fixed $\Delta t = 1$) to conditions 4 and 5 ($\Delta t_{i,j}$ with $N(1, 0.25)$ versus $N(1, 0.5)$) in Table 1, in which the resulting autoregressive cross-lagged matrices ($\hat{\mathbf{A}}_{\Delta t_j}$) are quite similar. For different time intervals or different distributions, however, discrete time parameters will be quite different, as illustrated in the other conditions in Table 1. Having estimated the underlying continuous time model, it becomes possible to express the discrete time parameters as a function of the underlying continuous time model and the observed time interval in question. This is illustrated in Figure 4 for the autoregressive effects and in Figure 5 for the cross-lagged effects. As apparent from the two figures, the discrete time autoregressive and cross-lagged parameters vary as a non-linear function of Δt. With individually varying time intervals, each case has its own discrete time parameters, which eventually can all be traced back to one and the same generating model in continuous time.

To summarize, this first example shows that: (1) time intervals matter, that is, the discrete time parameters depend on the time interval chosen; (2) the continuous time parameters are independent of the sample interval, thus permitting parameter comparisons within and across studies; (3) in continuous time modelling the traditional approach for panel data with equal intervals across individuals can be extended to individually varying intervals; and (4) the oversampling approach and the matrix exponential approach via the diagonalization method yield virtually identical parameter estimates with $D = 30$.

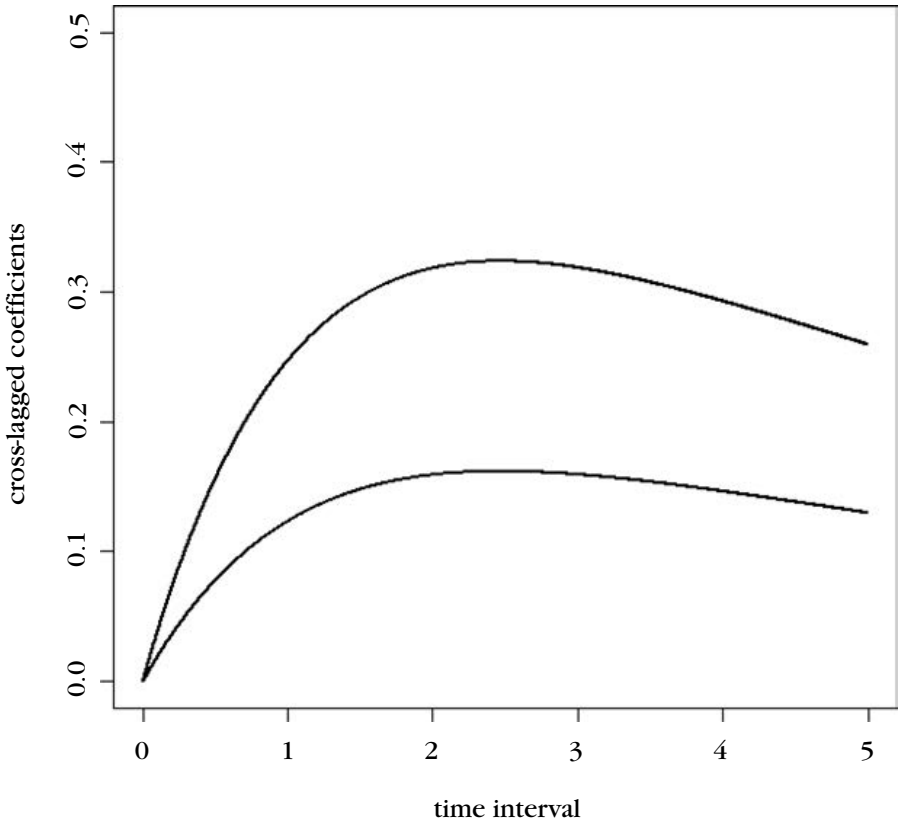

Figure 5. Discrete time cross-lagged parameters as a function of the true underlying continuous time model in Example 1.

6. Example 2: A Monte Carlo simulation on different time intervals in oscillating processes

In this example we first simulated a damped linear oscillator as defined in equation (7). Second, we sampled discrete time observations from the continuous time model using either fixed and equal intervals or individually varying intervals. The *average* sampling intervals were chosen in a way that the sampling rate was either *well above* ($\Delta t = 0.4$) the required minimum sampling rate (as defined by the Nyquist–Shannon sampling theorem), *just above* it ($\Delta t = 0.5$), or *well below* it ($\Delta t = 1.0$). Note that in the latter case (fixed intervals with a sampling rate below the Nyquist–Shannon criterion) the true generating process can no longer be uniquely identified, but is indistinguishable from other processes (aliases). Third, we fitted a (correctly specified) continuous time model to each condition. Fourth, we repeated the steps above in a Monte Carlo simulation and recorded the parameter estimates and the number of converged models.

6.1. Method

A univariate damped linear oscillator was simulated with drift matrix $\mathbf{A} = \left(\begin{smallmatrix} 0 & 1 \\ -(2\pi)^2 & -0.5 \end{smallmatrix}\right)$, continuous time dynamic error covariance matrix (diffusion matrix) $\mathbf{Q} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, initial mean vector $\mathbf{m}_{t_0} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, and initial covariance matrix $\mathbf{\Phi}_{t_0} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. As before, we assumed no additional measurement error, and for reasons of simplicity $\mathbf{b}$ was not freely estimated

but fixed to zero. The simulated process is shown in Figure 2. From this model, data were generated for $T = 11$ time points and $N = 200$ cases, with (1) fixed intervals versus individually varying intervals; (2) an average interval length of 0.4, 0.5 or 1.0; and (3) different starting values in the last condition (i.e., for $\Delta t = 1.0$).[4] In case of individually varying intervals, the interval length was drawn from a normal distribution with $\Delta t_{i,j} \sim N(0.4, \frac{1}{12})$, $\Delta t_{i,j} \sim N(0.5, \frac{1}{12})$, or $\Delta t_{i,j} \sim N(1.0, \frac{1}{2})$.[5] In addition, we included a condition without dampening (i.e., $\gamma = 0$).

For $\omega_0 = (2\pi)$, the undamped frequency $\left(1/T_0 = \omega_0/2\pi\right)$ is exactly 1. The Nyquist–Shannon sampling theorem implies that the sampling rate must be *higher* than twice the frequency of the original signal in order to be able to reconstruct this signal from discrete time observations. With a discrete time sampling interval of $\Delta t_j = 0.5$, the sampling frequency corresponds *exactly* to twice the frequency of oscillation, thus the Nyquist–Shannon condition would *not* be met. Because of the dampening factor ($\gamma = 0.5$), however, the actual (damped) frequency is somewhat lower in our example, so that the condition is just met with a sampling interval of $\Delta t = 0.5$ (illustrated by dots in Figure 2(b)), while it is well met with a sampling interval of $\Delta t = 0.4$ (illustrated by squares in Figure 2(b)). However, as mentioned above, it is clearly not met for $\Delta t = 1.0$, regardless of the dampening factor. The reason why we included these conditions is that we presume that individually varying time intervals result in unequal sampling intervals, which – due to the variation across people – may lead to a sampling scheme where the Nyquist–Shannon criterion is met for *some* of the subjects. This may help in identifying the true generating process. In contrast, with fixed intervals, the model may (1) not converge at all, or (2) converge on the true parameters of the generating process, or (3) converge on any of its aliases. However, because the generating process and its aliases are by definition indistinguishable, we suspect that whether (2) or (3) applies is just a matter of starting values. If starting values are good (i.e., close to the true parameters) the chances are high that the model (by pure chance) converges on the true parameters. In contrast, if starting values are less close to the true parameters, the model may as well converge on any other (alias) parameters. In order to support this argument, we included a condition with starting values that were 'inferior' to the starting values in the other conditions. Model simulation and parameter estimation was carried out in the same way as in the previous example. A total of 1000 Monte Carlo iterations were used in each condition.

6.2. Results

Results are presented in Table 2. Comparing the first five columns in Table 2 (fixed intervals) to the last five columns (individually varying time intervals) from left to right, we find the following. First, if the Nyquist–Shannon criterion is clearly met ($\Delta t = 0.4$),

---

[4] As 'good' starting values we chose
$\mathbf{\Phi}_{t_0} = \begin{pmatrix} 1.01 & \\ 0.1 & 1.01 \end{pmatrix}$, $\mathbf{m}_{t_0} = \begin{pmatrix} 0.05 \\ 1.01 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} \text{fixed} & \text{fixed} \\ -38 & -0.61 \end{pmatrix}$, $\mathbf{G} = \begin{pmatrix} 0 & 0 \\ 0 & 1.1 \end{pmatrix}$.
As 'inferior' starting values we chose
$\mathbf{\Phi}_{t_0} = \begin{pmatrix} 1.01 & \\ 0.1 & 1.01 \end{pmatrix}$, $\mathbf{m}_{t_0} = \begin{pmatrix} 0.01 \\ 0.01 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} \text{fixed} & \text{fixed} \\ -10 & -0.01 \end{pmatrix}$, $\mathbf{G} = \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \end{pmatrix}$.

[5] Standard deviations were chosen to represent a reasonable amount of variability around the average time intervals. All time intervals were constrained to be positive.

Table 2. Average parameter estimates and standard deviations of parameter estimates across 1000 Monte Carlo samples in each condition*

	Fixed intervals					Individually varying intervals				
$\Delta t_{i,j} =$	0.4	0.5	1.0	1.0 no dampening	1.0 inferior SV	$N(0.4, \frac{1}{12})$	$N(0.5, \frac{1}{12})$	$N(1, \frac{1}{2})$	$N(1, \frac{1}{2})$ no dampening	$N(1, \frac{1}{2})$ inferior SV
		NS criterion not met				NS criterion not met				
Valid solutions	100%	20%[a]	23%[b]	19%[c]	49%[d]	100%	100%	99%[e]	56%[f]	7%[g]
Mean(ω_0^2)	39.521	39.43	39.433	39.787	230.14	39.529	39.558	39.969	39.993	39.956
SD(ω_0^2)	0.088	0.389	0.409	0.292	630.547	0.091	0.097	0.091	0.038	0.109
Mean(γ)	0.501	0.517	0.511	0	886.22	0.501	0.502	0.51	0	0.512
SD(γ)	0.012	0.068	0.04	0.024	2381.932	0.013	0.012	0.012	0.006	0.011
Mean(g)	0.997	0.964	0.992	0.996	100.195	1	1	1.01	1.022	0.928
SD(g)	0.064	0.244	0.019	0.019	269.719	0.018	0.018	0.016	0.018	0.401
Mean(ϕ_1)	0.997	0.993	0.99	1	0.991	0.996	0.996	0.994	0.99	1.017
SD(ϕ_1)	0.102	0.09	0.102	0.099	0.098	0.099	0.098	0.101	0.107	0.098
Mean(ϕ_{12})	−0.003	−6.769	3.548	−0.302	1082.557	0.006	−0.004	0.034	0.03	0.036
SD(ϕ_{12})	0.089	129.101	83.199	47.861	17043.461	0.085	0.11	0.112	0.1	0.137
Mean(ϕ_2)	0.986	24662.027	14267.774	2932.713	13752.165	0.991	1.006	1.026	1.022	1.04
SD(ϕ_2)	0.141	35335.766	25869.343	10532.025	1726814.52	0.138	0.189	0.198	0.174	0.212
Mean(m_1)	0	−0.004	0.004	0.002	0	0.002	0.001	−0.002	−0.001	0.002
SD(m_1)	0.069	0.074	0.07	0.072	0.072	0.07	0.069	0.072	0.075	0.065
Mean(m_2)	1.001	6.897	9.528	−6.999	−1310.351	1.002	1	0.997	1.009	0.993
SD(m_2)	0.083	45.295	37.299	30.661	20913.005	0.083	0.095	0.099	0.091	0.093

Note. $T = 11$. Reported parameter estimates were obtained using the oversampling approach with $D = 30$. Population parameters are: $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -(2\pi)^2 & -0.5 \end{pmatrix}$, $\mathbf{Q} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{m}_{t_0} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{\Phi}_{t_0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; $(2\pi)^2 = 39.478$. NS = Nyquist–Shannon; SV = starting values. *Average parameter estimates are based on valid solutions only (no NPSOL error message). NPSOL error codes of excluded conditions: [a] Error 1 (code green): 27%, Error 4 (code blue): 2%, Error 6 (code red): 51%; [b] Error 1 (code green): 32%, Error 4 (code blue): 5%, Error 6 (code red): 40%; [c] Error 1 (code green): 13%, Error 4 (code blue): 2%, Error 6 (code red): 66%; [d] Error 1 (code green): 41%, Error 4 (code blue): 2%, Error 6 (code red): 8%; [e] Error 6 (code red): 1%; [f] Error 6 (code red): 44%; [g] Error 1 (code green): 1%, Error 6 (code red): 92%.

the model converges in 100% of all Monte Carlo samples.[6] This is true for fixed intervals as well as individually varying intervals (compare column 1 and 6 in Table 2).[7] In addition, all parameter estimates are unbiased and yield the same precision (i.e., the same standard deviation across 1000 Monte Carlo samples).

Second, if the Nyquist–Shannon criterion is just met ($\Delta t = 0.5$), the convergence rate drops dramatically to only 20% if the intervals are fixed. In contrast, with individually varying time intervals the model converges in 100% of all Monte Carlo samples. Furthermore, while the drift matrix (i.e., the parameters ω_0^2 and γ) seems to remain unbiased, this is not true for the mean (m_2) and variance (ϕ_2) of the first derivative, as well as the covariance between $\eta(t)$ and the first derivative (ϕ_{12}) if intervals are fixed. A similar observation has been made by Oud and Singer (2008; see their Table 5 and 6). In contrast, all parameter estimates are unbiased in the condition with individually varying intervals. Perhaps most importantly, parameter estimates are less efficient with fixed intervals than individually varying intervals. That is, the standard deviation of parameter estimates across 1000 Monte Carlo samples is substantially higher for fixed intervals and a sampling rate which just meets the Nyquist–Shannon criterion. In contrast, for individually varying intervals the precision remains about the same as in situations with a higher (less problematic) sampling rate (compare column 2 and 7 in Table 2).

Third, if the *average* sampling interval is larger than permitted by the Nyquist–Shannon criterion (i.e., $\Delta t = 1.0$), with individually varying time intervals the true generating process may still be identified. As apparent from column 8 in Table 2, the model converges in 99% of all Monte Carlo samples, all parameter estimates are unbiased, and their precision is about the same as in situations with a higher (less problematic) sampling rate. In contrast, fixed intervals result in a low convergence rate (23%; see column 3 in Table 2). Even when considering only valid solutions, several parameter estimates are biased and almost all parameter estimates exhibit a lower precision than in the condition with individually varying time intervals.

Fourth, setting the dampening parameter to zero ($\gamma = 0$) makes it even harder to identify the true generating process (column 4 and 9 in Table 2). As a matter of fact, now the number of valid solutions decreases (to 56%) even if time intervals are individually varying. With fixed intervals, the situation remains about as bad as before (i.e., with a non-zero dampening parameter). Importantly, however, even though the number of valid solutions has decreased, parameter estimates of the valid solutions remain unbiased with about the same precision as before if time intervals are individually varying.

Fifth, as pointed out above, for fixed intervals with $\Delta t = 1.0$, the generating process is not identified, that is, cannot be distinguished from other processes (its aliases). Nevertheless, as apparent from column 3 and 4 in Table 2, for about 20% of Monte Carlo samples the model converges, oftentimes on parameter estimates close to the true parameters ω_0^2 and γ. As pointed out above, we presume that apparently good parameter estimates are only due to well-chosen starting values. In order to support this argument we selected starting values that were further away from the true parameters (see footnote 3). Interestingly, for fixed intervals (column 5 in Table 2) this resulted in

[6] The default convergence criteria of OpenMx, or NPSOL, which is the optimizer employed by OpenMx (see Gill, Murray, Saunders, & Wright, 1998), were used in the present Monte Carlo simulation.

[7] Note, that not all individual intervals may meet the Nyquist–Shannon criterion.

a much higher convergence rate (49%)[8] than before, but now parameter estimates are all over the place and (with exception of the initial mean and variance of the observed variable) not even close to the true parameters. In contrast, with individually varying time intervals (column 10 in Table 2) all parameter estimates are unbiased and relatively efficient. Due to the bad starting values, however, the convergence rate drops to only 7%.

To summarize, the second example shows that continuous time models can be employed to fit oscillating processes via the use of latent derivatives on the right-hand side in equation (8). As demonstrated by the present Monte Carlo simulation, using oversampling with $D = 30$ intervals produced parameter estimates that are fairly close to the population parameters. The approach works for fixed, as well as individually varying time intervals. However, in case of a low sampling rate (close to the Nyquist–Shannon criterion), individually varying time intervals not only increased the chance of model convergence, but also resulted in better (unbiased and more efficient) parameter estimates. Even in cases, where the Nyquist–Shannon criterion is not met for the *average* sampling rate, individually varying time intervals can help to identify the generating process at least as long as some participants are measured at an interval less than the Nyquist–Shannon criterion.

7. Discussion

Many longitudinal studies in the social sciences were originally designed to be studies with equal time intervals between assessment waves. Once the data have been collected, however, it turns out that most of them are not. This is particularly true for observational studies involving many individuals. In these situations it is almost impossible to measure everyone at exactly the same time point, resulting in individually varying time intervals. In the present paper we therefore deal with two related questions. First, are equal intervals (between and within individuals) necessary – or at least recommended – when designing longitudinal studies? And, second, how do we deal with unequal time intervals (between and within individuals) once the data have been gathered?

To address these questions we reviewed the basic idea of continuous time modelling, before extending it to individually varying time intervals. The new approach permits the estimation of (possibly coupled) oscillating and non-oscillating multivariate processes with arbitrary time intervals between and within individuals.

By means of theoretical arguments and two examples, we have shown that continuous time analyses account for any time intervals, thus equal intervals are *not* necessary for statistical analyses. In addition, we have demonstrated that parameter estimates at any discrete time point are a function of the underlying continuous time parameters (see Figures 4 and 5), so that equal intervals are *not* important for parameter comparisons. Finally, we have demonstrated that equal intervals may be *sub*optimal for the estimation of oscillating processes, in particular when the sampling rate, as compared to the frequency of oscillation, is low. All analyses were carried out by CT-SEM, an OpenMx based script for continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes, which is available for download.

[8] Including solutions with NPSOL error messages 1 (code green), 90% of the models converged.

In summary, we conclude that it is important to account for individually varying time intervals, and encourage researchers to conceive of longitudinal studies with different time intervals as an opportunity rather than a problem.

Acknowledgements

We would like to thank the MPIB Grid Team (in particular Michael Krause) for their help in running the simulations, and Timothy Brick for his advice on improving the CT-SEM script.

References

Allison, D. B., Allison, R. L., Faith, M. S., Paultre, F., & Pi-Sunyer, F. X. (1997). Power and money: Designing statistically powerful studies while minimizing financial costs. *Psychological Methods*, *2*(1), 20–33. doi: 10.1037/1082-989X.2.1.20

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates.

Arnold, L. (1974). *Stochastic differential equations*. New York: Wiley.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Journal of the Royal Statistical Society (Supplement)*, 7, 27–41.

Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics* (Vol. 2). Amsterdam: Elsevier Science.

Bergstrom, A. R. (1988). The history of continuous-time econometric models. *Econometric Theory*, *4*, 365–383.

Boker, S. M. (2007a). Differential equation models for longitudinal data. In S. Menard (Ed.), *Handbook of longitudinal research* (pp. 639–652). New York: Elsevier.

Boker, S. M. (2007b). Specifying latent differential equations models. In S. M. Boker & M. J. Wenger (Eds.), *Data analytic techniques for dynamical systems in the social and behavioral sciences* (pp. 131–159). Mahwah, NJ: Lawrence Erlbaum Associates.

Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In S.-M. Chow, E. Ferrer & F. Hsieh (Eds.), *Statistical methods for modeling human dynamics : An interdisciplinary dialogue* (pp. 161–178). New York: Routledge.

Boker, S. M., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P., & Fox, J. (2011). OpenMx: An Open Source extended structural equation modeling framework. *Psychometrika*, *76*(2), 306–317. doi: 10.1007/s11336-010-9200-6

Boker, S. M., Neale, M. C., & Rausch, J. R. (2004). Latent differential equation modeling with multivariate multi-occasion indicators. In K. van Montfort, J. H. L. Oud & A. Satorra (Eds.), *Recent developments on structural equation models: Theory and applications* (pp. 151–174). Amsterdam: Kluwer.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, *37*, 127–160.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Butner, J., Amazeen, P., & Mulvey, G. (2005). Multilevel modeling of two cyclical processes: Extending differential structural equation modeling to coupled systems. *Psychological Methods*, *10*, 159–177.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika*, *12*, 267–288. doi: 10.1007/BF02288941

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583–619. doi: 10.1146/annurev.psych.093008.100356

Deboeck, P. R. (2010). Estimating dynamical systems: Derivative estimation hints from Sir Ronald A. Fisher. *Multivariate Behavioral Research*, *45*(4), 725–745. doi: 10.1080/00273171.2010.498294

Delsing, M. J. M. H., & Oud, J. H. L. (2008). Analyzing reciprocal relationships by means of the continuous-time autoregressive latent trajectory model. *Statistica Neerlandica*, *62*(1), 58–82. doi: 10.1111/j.1467-9574.2007.00386.x

Fahrenberg, J., & Myrtek, M. (Eds.) (1996). *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Seattle, WA: Hogrefe & Huber.

Gates, K. M., Molenaar, P. C. M., Hillary, F., & Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage*. doi: 10.1016/j.neuroimage.2010.08.051

Gates, K. M., Molenaar, P. C. M., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage*, *50*, 1118–1125. doi: 10.1016/j.neuroimage.2009.12.117

Gill, P. E., Murray, W., Saunders, M. A., & Wright, M. H. (1998). User's guide for NPSOL 5.0: A Fortran package for nonlinear programming. Technical Report SOL 86–2, Revised July 30, 1998. Stanford, CA: Systems Optimization Laboratory, Department of Operations Research, Stanford University.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*(1), 80–92.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lueke, H. D. (1999). The origins of the sampling theorem. *IEEE Communications Magazine*, *37*, 106–108.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605. doi: 10.1146/annurev.psych.60.110707.163612

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181–202. doi: 10.1007/BF02294246

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychology*, *18*(2), 112–117. doi: 10.1111/j.1467-8721.2009.01619.x

Molenaar, P. C. M., & Nesselroade, J. R. (2009). The recoverability of P-technique factor analysis. *Multivariate Behavioral Research*, *44*(1), 130–141. doi: 10.1080/00273170802620204

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Idiographic filters for psychological constructs. *Measurement*, *5*(4), 217–235.

Oertzen, T. von (2010). Power equivalence in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*, *63*, 257–272. doi: 10.1348/000711009X441021

Oud, J. H. L. (2007a). Comparison of four procedures to estimate the damped linear differential oscillator for panel data. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 19–39). Mahwah, NJ: Lawrence Erlbaum Associates.

Oud, J. H. L. (2007b). Continuous time modeling of reciprocal relationships in the cross-lagged panel design. In S. M. Boker & M. J. Wenger (Eds.), *Data analytic techniques for dynamical systems* (pp. 87–129): Mahwah, NJ: Lawrence Erlbaum Associates.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244). New York: Springer.

Oud, J. H. L., & Folmer, H. (2011). Modeling oscillation: Approximately or exactly? *Multivariate Behavioral Research*, *46*, 985–993.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, *65*(2), 199–215. doi: 10.1007/BF02294374

Oud, J. H. L., & Singer, H. (2008). Continuous time modeling of panel data: SEM versus filter techniques. *Statistica Neerlandica*, *62*(1), 4–28.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer.

R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Refinetti, R. (2006). *Circadian physiology* (2nd ed.). Boca Raton, FL: Taylor & Francis.

Ruymgaart, P. A., & Soong, T. T. (1985). *Mathematics of Kalman-Bucy filtering*. Berlin: Springer.

Singer, H. (1990). *Parameterschätzung in zeitkontinuierlichen dynamischen Systemen* [*Parameter estimation in continuous time dynamic systems*]. Konstanz: Hartung-Gorre.

Singer, H. (1992). The aliasing-phenomenon in visual terms. *Journal of Mathematical Sociology*, *17*(1), 39–49. doi: 10.1080/0022250x.1992.9990097

Singer, H. (1999). *Finanzmarkt-Ökonometrie: Zeitstetige Systeme und ihre Anwendung in Ökonometrie und empirischer Kapitalmarktforschung* [*Finance market econometrics: Continuous time systems and their application in econometrics and empirical capital market research*]. Heidelberg: Physica-Verlag.

Singer, H. (2012). SEM modeling with singular moment matrices part II: ML-estimation of sampled stochastic differential equations. *Journal of Mathematical Sociology*, *36*(1), 22–43. doi: 10.1080/0022250x.2010.532259

Takahashi, J., & Zatz, M. (1982). Regulation of circadian rhythmicity. *Science*, *217*(4565), 1104–1111. doi: 10.1126/science.6287576

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (in press). An SEM approach to continuous time modeling of panel data: relating authoritarianism and anomia. *Psychological Methods*.

Received 15 June 2011; revised version received 14 January 2012

Appendix A

The exponential function is defined as

$$e^{\mathbf{A}\cdot\Delta t_{i,j}} = \sum_{k=0}^{\infty} \frac{(\mathbf{A}\cdot\Delta t_{i,j})^k}{k!} = \mathbf{I} + \mathbf{A}\cdot\Delta t_{i,j} + \frac{1}{2!}\mathbf{A}^2\cdot\Delta t_{i,j}^2 + \frac{1}{3!}\mathbf{A}^3\cdot\Delta t_{i,j}^3 + \ldots. \quad \text{(A1)}$$

From the definition it is obvious that equation (14) uses only the first part of the power series (i.e., $\mathbf{I} + \mathbf{A}\cdot\Delta t_{i,j}$). Although this will approach the true matrix exponential function for $D \to \infty$, the convergence will be faster if we approximate by

$$\begin{aligned} e^{\mathbf{A}\cdot\Delta t_{i,j}} &\approx \mathbf{I} + \mathbf{A}\cdot\Delta t_{i,j} + \tfrac{1}{2}\mathbf{A}^2\cdot\Delta t_{i,j}^2 + \tfrac{1}{4}\mathbf{A}^3\cdot\Delta t_{i,j}^3 + \tfrac{1}{8}\mathbf{A}^4\cdot\Delta t_{i,j}^4 + \ldots \\ &= \left(\mathbf{I} - \tfrac{1}{2}\mathbf{A}\cdot\Delta t_{i,j}\right)^{-1}\left(\mathbf{I} + \tfrac{1}{2}\mathbf{A}\cdot\Delta t_{i,j}\right). \end{aligned} \quad \text{(A2)}$$

The only difference between equations (A1) and (A2) is that the denominators in equation (A1) follow a factorial function, while a power function of 2 is used in (A2). The advantage of (A2), however, is that it can be reformulated as shown on the second line. It is this rewrite that is used for the oversampling approach in equation (15):

$$e^{\mathbf{A}\cdot\Delta t_{i,j}} = \lim_{D\to\infty}\prod_{d=0}^{D-1}\left(\left(\mathbf{I}-\frac{1}{2}\mathbf{A}\cdot\delta_{i,j,d}\right)^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}\cdot\delta_{i,j,d}\right)\right).$$

Supporting Material

The following supporting information may be found in the online edition of this article:

1) The CT-SEM script. An R/OpenMx based program for estimating oscillating and nonoscillating processes in continuous time.
2) Two examples with individually varying time intervals:
   One example is a bivariate (coupled) nonoscillating process and one a univariate oscillating process.
   Both examples consist of a datafile (.dat) and the CT-SEM input file (.R)

The actual CT-SEM script needs to be saved in the same folder as the input file.
Usually, there is NO need to open or modify the CT-SEM.R script. The model is completely specified in the input file.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.