STOCHASTIC DIFFERENTIAL EQUATION MODELS WITH TIME-VARYING PARAMETERS

A Thesis in
Human Development and Family Studies
by
Meng Chen

© 2017 Meng Chen

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2017
The thesis of Meng Chen was reviewed and approved* by the following:

Sy-Miin Chow  
Associate Professor of Human Development and Family Studies  
Thesis Advisor

Zita Oravecz  
Assistant Professor of Human Development and Family Studies

Lisa Gatzke-Kopp  
Associate Professor of Human Development and Family Studies  
Professor-in-Charge

*Signatures are on file in the Graduate School.
Abstract

Self-organization occurs when a system shows distinct shifts in dynamics due to variations in the parameters that govern the system. Relatedly, many human dynamic processes with self-organizing features comprise subprocesses that unfold across multiple time scales. Incorporating time-varying parameters (TVPs) into a dynamic model of choice provides one way of representing self-organization as well as multi-time scale processes. Extant applications involving models with TVPs have been restricted to formulation in discrete time. Related work for representing time-varying parameters in continuous-time models remains scarce. We propose a stochastic differential equation (SDE) modeling framework with TVPs as a way to capture self-organization in continuous time. We present several examples of SDEs with TVPs, including a stochastic damped oscillator model with hypothesized functional shifts in both set-points and damping. Furthermore, we discuss plausible models that may be used to approximate changes in the TVPs in the absence of further knowledge concerning their true change mechanisms.
# Table of Contents

List of Figures v

List of Tables vi

Acknowledgments vii

Chapter 1
Introduction 1
1.1 Self Organization in Human Behaviors 1
1.2 Modeling in Continuous Time with Time-varying Parameters (TVPs) 2

Chapter 2
Stochastic Differential Equation Models with Time-varying Parameters 4
2.1 A Stochastic Differential Equation Framework 4
2.2 Estimation Procedures 5
2.3 Motivating Model 7
2.4 Approximation Functions for TVPs 8
2.5 Model Fitting via dynr 10

Chapter 3
Illustrative Examples 11
3.1 Example 1: Stagewise Shifts in Set-Point 11
3.2 Example 2: Logistic Growth in Set-Point 13
3.3 Example 3: Stagewise Shifts in Damping 15

Chapter 4
Discussion 19

Appendix
Sample dynr Code for Fitting the Models in the Illustrative Examples 22
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Simulated observed and latent trajectory for one hypothetical subject generated using the damped oscillator model with stage-wise shifts in set-point considered in Illustrative Example 1.</td>
<td>12</td>
</tr>
<tr>
<td>3.2</td>
<td>Estimated TVP trajectories for one hypothetical subject in Illustrative Example 1: (a) estimated and true trajectories, with the corresponding RMSEs; (b) biases in the estimated TVP values in comparison to the true values.</td>
<td>14</td>
</tr>
<tr>
<td>3.3</td>
<td>A simulated set of observed and latent trajectories generated under the scenario of logistic growth as in Example 2.</td>
<td>15</td>
</tr>
<tr>
<td>3.4</td>
<td>Estimated TVP trajectories for one hypothetical subject in Illustrative Example 2: (a) estimated and true trajectories, with the corresponding RMSEs; (b) biases in the estimated TVP values in comparison to the true values.</td>
<td>16</td>
</tr>
<tr>
<td>3.5</td>
<td>Simulated observed and latent trajectory for one hypothetical subject generated using the damped oscillator model with stage-wise shifts in $\zeta$ as considered in Illustrative Example 3. Damping = the damping/amplification parameter, $\zeta(t)$.</td>
<td>17</td>
</tr>
<tr>
<td>3.6</td>
<td>Estimated TVP trajectories for one hypothetical subject in Illustrative Example 3: (a) estimated and true trajectories, with the corresponding RMSEs; (b) biases in the estimated TVP values in comparison to the true values.</td>
<td>18</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Newell’s Table: Time Scale of Human Action. . . . . . . . . . . . . . . . . . . . 2
3.1 Parameter Recovery in Illustrative Example 1 Obtained from Fitting the True Data Generation Model and the Two Approximation Models. . . . . . . . . . . . . . 12
3.2 Parameter Recovery in Illustrative Example 2 Obtained from Fitting the True Data Generation Model and the Two Approximation Models. . . . . . . . . . . . . . 15
3.3 Parameter Recovery in Illustrative Example 3 Obtained from Fitting the True Data Generation Model and the Two Approximation Models. . . . . . . . . . . . . . 18
Acknowledgments

Funding for this study was provided by NSF grant SES-1357666, NIH grant R01GM105004, Penn State Quantitative Social Sciences Initiative and UL TR000127 from the National Center for Advancing Translational Sciences awarded to Sy-Miin Chow. Special thanks go to Dr. Sy-Miin Chow for her support and encourage towards the completion of this master thesis, and also Dr. Michael Hunter at University of Oklahoma Health Sciences Center and Dr. Zita Oravecz for their help and comments along the way. I appreciate all the efforts and time they invested in this project. I also thank Lu Ou, Linying Ji, Dr. Tim Brick, Dr. Nilam Ram, and other members of QuantDev for suggestions and thoughts that have led to improvements. This thesis will be further edited and submitted to be a chapter in the book Continuous Time Modeling in the Behavioral and Related Sciences, to be published by Springer in Spring 2018 (editors Kees van Montfort, Johan Oud and Manuel Voelkle).
Chapter 1 | Introduction

1.1 Self Organization in Human Behaviors

The world we live in is full of complexities, and we as humans are no exception. Our bodies are composed on trillions of cells, and yet we each function as an integrated unit. We move, think, learn, socialize, and thrive. Our actions could reflect a conscious effort, or an emergent reaction to — or interaction with — changes in the environment (Newtson et al., 1987; Newtson, 1993; Gottlieb et al., 1996). In other words, our minds and bodies are capable of self-organization.

Self-organization is a process through which orderliness emerges from seeming disorderliness (Lewis and Ferrari, 2001). The idea of self-organization is not new to the research of the human mind and behavior (Bosma and Kunnen, 2011; Kelso, 1995a; Magnusson and Cairns, 1996). Taking human movement as an example, a simple motion involves approximately $10^2$ muscles, $10^3$ joints, and $10^{14}$ cells, yet the physical movement of the human body can be characterized by relatively few dimensions (Turvey, 1990; Bertenthal, 2007). That is, our body organizes complex systems into simple patterns described by comparatively few elements.

Changes happening in the human body can be captured at multiple levels and across multiple time scales: from neural networks that fire every few milliseconds to cognitive and behavioral processes that progress in seconds, to even longer-term processes like personality traits that develop through the life span. To this end, Newell (1990) proposed a time scale of human actions that organized topics of interest in the field of psychology into a hierarchy of levels of analysis (see Table 1.1). According to Newell (1990), the biological band consists of neural activities, which happen on the scale of milliseconds. In the cognitive band, cognitive operations — such as recognizing a person approaching in order to engage in a conversation — happen on the scale of seconds. The next hierarchy consists of the rational band with tasks such as decision making, which span minutes or hours. At the highest level, the social band encompasses social processes such as interactions with others and forming relationships, which may unfold over days, weeks, months, or even longer (Bertenthal, 2007). Individual and lifespan development, too, can be conceived as a multi-time scale process that interweaves more gradual developmental changes and short-term fluctuations or intraindividual variability (Baltes and Nesselroade, 1979; Kelso, 1995b; Ram and Gerstorf, 2009) — a process described by Nesselroade (1991) as the warp and
woof of the developmental fabric.

Table 1.1: Newell’s Table: Time Scale of Human Action.

<table>
<thead>
<tr>
<th>Scale (Sec)</th>
<th>Time Units</th>
<th>System</th>
<th>World (theory)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^7$</td>
<td>Months</td>
<td></td>
<td>Social Band</td>
</tr>
<tr>
<td>$10^6$</td>
<td>Weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^5$</td>
<td>Days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^4$</td>
<td>Hours</td>
<td>Task</td>
<td>Rational Band</td>
</tr>
<tr>
<td>$10^3$</td>
<td>10 min</td>
<td>Task</td>
<td></td>
</tr>
<tr>
<td>$10^2$</td>
<td>Minutes</td>
<td>Task</td>
<td></td>
</tr>
<tr>
<td>$10^1$</td>
<td>10 sec</td>
<td>Unit task</td>
<td></td>
</tr>
<tr>
<td>$10^0$</td>
<td>1 sec</td>
<td>Operations</td>
<td>Cognitive Band</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>100 ms</td>
<td>Deliberate act</td>
<td></td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>10 ms</td>
<td>Neural circuit</td>
<td>Biological Band</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>1 ms</td>
<td>Neuron</td>
<td></td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>100 µs</td>
<td>Organelle</td>
<td></td>
</tr>
</tbody>
</table>

1.2 Modeling in Continuous Time with Time-varying Parameters (TVPs)

Dynamical systems modeling offers a framework for studying processes that undergo self-organization. The key characteristic of any dynamical system is that the complex behaviors that the system manifests can be captured with relatively simple rules or patterns (Nowak and Lewenstein, 1994). This line of thinking has considerable appeal to social and behavioral scientists. Unfortunately, due in part to the scarcity of data that span multiple bands and methodological difficulties in integrating multiple processes with distinct time scales, applications of dynamical systems concepts in the behavioral sciences have hitherto been limited to representation of phenomena and activities within a particular band in Newell’s table (see e.g., Vallacher and Nowak, 1994).

Data availability, however, is no longer an insurmountable hurdle. The past decade has evidenced an increased popularity of studies aimed at collecting ecological momentary assessments (EMA). In these studies, participants are measured in real-time in their natural environment, and often for many repeated occasions (Shiffman et al., 2008). This, along with advances in data collection technology such as wearable assessment tools and smartphone-based surveys, affords renewed opportunities for modeling multi-time scale processes as they unfold in continuous time.

In continuous-time models, time is denoted by real numbers, as opposed to integers. Although conceptualizing and representing dynamic phenomena in discrete time has many practical advantages from a computational standpoint, modeling a dynamical process in continuous time offers several other advantages (Molenaar and Newell, 2003; Oud and Singer, 2008; Voelkle and Oud, 2013). One of these advantages is that a continuous representation of time mirrors the phenomena of interest: time and human behaviors that unfold in time are continuous in nature.
That is, even though empirical measurements of human dynamics are always observed at discrete
time points, human behaviors do not cease to exist between successive measurements. Changes
are unfolding at any particular time point even when observed data may not be available at those
time points. Another important advantage of continuous-time modeling is that it can readily
accommodate irregularly spaced observations (Oud, 2007b; Chow et al., 2016; Voelkle and Oud,
2013).

Here, a stochastic differential equation (SDE) modeling framework with time-varying param-
eters (TVPs) is proposed as a way to capture self-organization in continuous time. Incorporating
TVPs into a dynamic model of choice provides one way of representing self-organization as well as
multi-time scale processes. Extant applications involving models with TVPs have been restricted
to formulation in discrete time. Examples of models with TVPs in the discrete-time modeling
framework include time-varying autoregressive moving average (ARMA) model (Tarvainen et al.,
2006; Weiss, 1985), the local linear trend model (Harvey, 2001), stochastic regression model (Pa-
gan, 1980), time-varying cyclic models (Chow et al., 2009), and dynamic factor analysis models
with TVPs (Chow et al., 2011; Del Negro and Otrok, 2008; Molenaar, 1994; Stock and Watson,
2008). Related work for representing time-varying parameters in continuous-time models remains
scarce.

The work in this thesis thus extends previous work on dynamical systems modeling in a
number of ways. First, it extends the damped oscillator model — a benchmark dynamical
systems model of human behaviors with time-invariant parameters (Boker and Graham, 1998;
Oud, 2007a) — by allowing the parameters to be time-varying. Second, several approximation
functions are proposed and discussed that can be used to represent the dynamics of TVPs even
when their underlying change mechanisms are unknown. Third, a continuous-time framework is
proposed and several illustrative examples of how TVPs may be represented in continuous time
are provided. Fourth, a set of algorithms based on the continuous-discrete Kalman filter (Bar-
Shalom et al., 2001; Chow et al., 2017; Kulikov and Kulikova, 2014; Kulikova and Kulikov, 2014)
is proposed that allow the simultaneous estimation of the dynamics of a system and all associated
TVPs and time-invariant parameters. Finally, how the proposed models can be fitted using a
statistical package, dynr (Dynamic Modeling in R; Ou et al., 2016, 2017), in R is illustrated.

The reminder of this thesis is organized as follows. It begins by a description of the SDE
framework on which the proposed continuous-time models with TVPs are based. Then several
time-varying extensions of the stochastic damped linear oscillation model (Oud and Singer, 2008;
Voelkle and Oud, 2013) are presented, with some hypothetical scenarios of why and how TVPs
may be used to capture multi-time scale oscillatory processes in continuous time. Illustrative
simulation results are then used to highlight the utility of the proposed approach and how infor-
mation criterion measures such as the Bayesian Information Criterion (Schwarz, 1978) and the
Akaike Information Criterion (Akaike, 1973) can be used to perform selection of models with
TVPs. Finally, a brief illustration of the code for specifying the proposed TVP models in dynr
is included.
Chapter 2  
Stochastic Differential Equation Models with Time-varying Parameters

2.1 A Stochastic Differential Equation Framework

Differential equation models capture a system’s change processes in terms of rates of change involving variables and their derivatives. The hypothesized system of SDEs is assumed to be in the general form of:

\[ d\eta_i(t) = f(\eta_i(t), t, x_i(t)) dt + G dw(t). \]  

(2.1)

This system is represented by a deterministic part, \( f(\eta_i(t), t, x_i(t)) dt \), and a stochastic part, \( G dw(t) \), where \( i \) is the smallest independent unit of analysis (e.g., person, dyad; \( i = 1, \ldots, n \)), and \( t \) is an index for time that can take on any real number; \( \eta \) is a vector of latent system variables, \( x \) represents the set of covariates that may influence the system, and \( f \) is a vector of drift functions that govern the deterministic (i.e., predictable given perfect knowledge of the previous states of the system and the true parameters) portion of the changes in the system variables. Without the stochastic part, the system reduces to an ordinary differential equation (ODE) model. The term \( w \) denotes a vector of standard Wiener processes (referred to herein as the process noise) whose differentials, \( dw(t) \), over the time interval, \( dt \), are Gaussian distributed, and are characterized by variances that increase linearly with increasing \( dt \). \( G \) is a diffusion matrix containing matrix square-root of the process noise variance-covariance matrix. Thus, the SDE model in (2.1) can be used to represent latent processes that develop and fluctuate over time.

The initial conditions (ICs) for the SDEs are defined explicitly to be the latent variables at an initial time point, \( t_i \) (e.g., the first observed time point), denoted as \( \eta_i(t_{i,1}) \), and are specified to be normally distributed with means \( \mu_{\eta_i} \) and covariance matrix, \( \Sigma_{\eta_i} \). The time interval between time \( t_{i,j} \) and \( t_{i,j+1} \) is denoted by \( \Delta t_{i,j} = t_{i,j+1} - t_{i,j} \).
The corresponding measurement model, which specifies the relations between the underlying latent processes and their observed manifestations measured at discrete and person-specific occasion \( j, t_{i,j} \ (j = 1, \ldots, T_i; i = 1, \ldots, n) \), is expressed as:

\[
\begin{align*}
y_i(t_{i,j}) &= \tau + \Lambda \eta_i(t_{i,j}) + A x_i(t_{i,j}) + \epsilon_i(t_{i,j}) \\
\epsilon_i(t_{i,j}) &\sim N(0, R),
\end{align*}
\] (2.2)

where \( \tau \) is a vector of intercepts, \( \Lambda \) is the factor loading matrix, \( A \) is the regression matrix associated with the vector of covariates \( x \), and \( \epsilon_i(t_{i,j}) \) is a vector of Gaussian distributed measurement errors.

Although both follow a Gaussian distribution, the process noise in \( \omega \) is different from the measurement noise in \( \epsilon \) in several important ways. As noted, the former contains “dynamical” or “process” noise, and it captures stochastic shocks to the system variables that continue to influence the underlying states of these variables over time. The latter comprises “measurement” errors whose influences are restricted to one and only one measurement occasion, namely at time \( t_{i,j} \). As such, the process noise — but not the measurement noise — help capture additional sources of uncertainties that affect the latent dynamics of the system at the current as well as future time points. In the study of individual development and more broadly in behavioral sciences as a whole, it is unreasonable to assume that we have collected information about all the factors that influence a certain dynamic process. Referring back to Table 1.1, any change on any level of analysis is subjected to the influence of all the other levels in the system. The additional process noise component under the SDE framework, compared to the ODE framework, allows for the possibility of incorporating the influence of other unknown and unmodeled (by the deterministic drift functions, \( f(.) \)) sources of disturbances to the system.

2.2 Estimation Procedures

The proposed SDE models with TVPs were fitted using a set of algorithms based on the continuous-discrete extended Kalman filter (CDEKF; Kulikov and Kulikova, 2014) available in \texttt{dynr} (Ou et al., 2017). The CDEKF is a nonlinear, continuous-time counterpart of the Kalman filter which assumes that the underlying dynamic processes unfold in continuous-time but the associated empirical measurements are taken at discrete time points. Following the general procedure of the Kalman filter, the CDEKF involves a prediction step and a correction step. Unlike the traditional prediction step in the discrete-time EKF, the first step in the CDEKF integrates the SDE model from time \( t_{i,j-1} \) to the next measured time \( t_{i,j} \) by means of numerical integration. Let \( \eta_i(t_{i,j-1}|t_{i,j-1}) \triangleq E[\eta_i(t_{i,j-1}) | Y_i(t_{i,j-1})] \), \( P_i(t_{i,j-1}|t_{i,j-1}) \triangleq \text{Cov}[\eta_i(t_{i,j-1})|Y_i(t_{i,j-1})] \), and \( Y_i(t_{i,j-1}) \triangleq \{y_i(t_{i,1}), \ldots, y_i(t_{i,j-1})\} \). The numerical integration is performed by solving the following ODEs using numerical solvers with \( \eta_i(t_{i,j-1}|t_{i,j-1}) \) and \( P_i(t_{i,j-1}|t_{i,j-1}) \) as the respective initial conditions:
\[
\frac{d\hat{\eta}(t)}{dt} = f(\hat{\eta}(t), t, \mathbf{x}_i(t)),
\]
\[
\frac{dP(t)}{dt} = F(t, \hat{\eta}(t))P(t) + P(t)F(t, \hat{\eta}(t))^T + Q.
\]

Here $\hat{\eta}(t)$ is the vector of predicted latent variables for subject $i$ at time $t$; $P(t)$ is the covariance matrix of prediction errors on the latent variables at time $t$, $\text{Cov}[\eta_i(t) - \hat{\eta}_i(t)]$. $F(t, \hat{\eta}(t))$ is the Jacobian matrix that consists of first-order partial derivatives of $f$ with respect to each element of the latent variable vector, $\eta_i(t)$, evaluated at $\eta_i(t_{i,j-1} | t_{i,j-1})$, and $Q = GG^T$.

The correction step follows the equations below:

\[
\mathbf{v}_i(t_{i,j}) = \mathbf{y}_i(t_{i,j}) - \left(\mathbf{r} + \Lambda \eta_i(t_{i,j} | t_{i,j-1}) + \mathbf{Ax}(t_{i,j})\right),
\]
\[
\mathbf{V}_i(t_{i,j}) = \mathbf{AP}_i(t_{i,j} | t_{i,j-1})\mathbf{A}^T + \mathbf{R}.
\]
\[
\eta_i(t_{i,j} | t_{i,j}) = \eta(t_{i,j} | t_{i,j-1}) + \mathbf{K}(t_{i,j})\mathbf{v}_i(t_{i,j}),
\]
\[
\mathbf{P}_i(t_{i,j} | t_{i,j-1}) = \mathbf{P}_i(t_{i,j} | t_{i,j-1}) - \mathbf{K}(t_{i,j})\mathbf{AP}_i(t_{i,j} | t_{i,j-1}),
\]

where $\mathbf{v}_i(t_{i,j})$ is a vector of prediction errors and $\mathbf{V}_i(t_{i,j})$ is the prediction error covariance matrix. $\eta(t_{i,j} | t_{i,j})$ and $\mathbf{P}_i(t_{i,j} | t_{i,j})$ are the filtered estimates of $\eta_i(t_{i,j})$ and $\mathbf{P}_i(t_{i,j})$; finally, $\mathbf{K}(t_{i,j}) = \mathbf{P}_i(t_{i,j} | t_{i,j-1})\mathbf{A}^T[\mathbf{V}_i(t_{i,j})]^{-1}$ is the Kalman gain function.

Parameter estimation is performed by optimizing a log likelihood function known as the prediction error decomposition function, computed using $\mathbf{v}_i(t_{i,j})$ and $\mathbf{V}_i(t_{i,j})$ as: (Schweppe, 1965; Chow et al., 2007):

\[
\log[f(\mathbf{Y}; \theta)] = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{T_i} \left[p \log(2\pi) + \log |\mathbf{V}_i(t_{i,j})| + \mathbf{v}_i(t_{i,j})^T\mathbf{V}_i(t_{i,j})^{-1}\mathbf{v}_i(t_{i,j})\right].
\]

where $p$ is the number of observed variables. Information criteria measures such as the AIC (Akaike, 1973) and BIC (Schwarz, 1978) can then be computed using $\log[f(\mathbf{Y}; \theta)]$ as (see Harvey, 2001, p. 80):

\[
\text{AIC} = -2\log[f(\mathbf{Y}; \theta)] + 2q
\]
\[
\text{BIC} = -2\log[f(\mathbf{Y}; \theta)] + q \log(\sum_{i} T_i),
\]

where $q$ is the number of time-invariant parameters in a model.

Kalman filtering occurs recursively forward in time by predicting and updating estimates. However, doing so only produces conditional latent variable estimates at a particular time point, $t_{i,j}$, and their corresponding covariance matrix, based on observed information up to time $t_{i,j}$. These filtered estimates can be further refined using information from the entire observed time series, namely, $\{\mathbf{y}_i(t_{i,j}); j = 1, \ldots, T_i\}$. The fixed interval smoother, in particular, occurs recursively backward in time by incorporating all available information from person $i$ up to time $T_i$. 
The parameter \( \omega \) damping parameter, model with TVPs. The TVPs considered include both the set-point parameter, illustrative examples in this chapter, we consider SDE variations of the damped linear oscillator mood regulation in recent widows (Bisconti et al., 2004), among many other examples. In the (Chow et al., 2005), self-regulation in adolescence substance use (Boker and Graham, 1998), and intraindividual variability. It has been used to represent dynamic processes such as emotions (2001; Bar-Shalom et al., 2001; Shumway and Stoffer, 2000; Chow et al., 2017):

\[
\eta_i(t_i|T_i) = \eta_i(t_i|t_{i,j}) + \tilde{\mathbf{P}}(t_i)(\eta_i(t_{i,j+1}|T_i) - \eta_i(t_{i,j+1}|t_{i,j})), \\
\mathbf{P}_i(t_i|T_i) = \mathbf{P}_i(t_i|t_{i,j}) + \tilde{\mathbf{P}}(t_i)(\mathbf{P}_i(t_{i,j+1}|T_i) - \mathbf{P}_i(t_{i,j+1}|t_{i,j}))\tilde{\mathbf{P}}(t_i),
\]

where \( \tilde{\mathbf{P}}(t_i) = \mathbf{P}_i(t_i|t_{i,j})\mathbf{F}(t, \hat{\eta}_i(t))^T[\mathbf{P}_i(t_{i,j+1}|t_{i,j})]^{-1} \).

### 2.3 Motivating Model

One of the better-known and widely applied differential equation models in psychology is the damped linear oscillator model, a second-order ODE expressed as (Boker and Graham, 1998):

\[
\frac{d^2 \eta_i(t)}{dt^2} = \omega [\eta_i(t) - \mu] + \zeta \frac{d\eta_i(t)}{dt},
\]

where \( \eta_i(t) \) denotes a univariate, latent process of interest for unit \( i \) at any arbitrary time \( t \); \( \omega \) denotes the process’s first derivative at time \( t \) and \( \frac{d^2 \eta_i(t)}{dt^2} \) denotes the second derivative at time \( t \). The parameter \( \omega \) governs the frequency of the oscillations; \( \mu \) is the set-point or equilibrium around which the system oscillates (usually set to 0), and \( \zeta \) is a damping (when \( \zeta < 0 \)) or amplification (when \( \zeta > 0 \)) parameter that governs changes in the magnitude of the oscillations over time. The damped oscillator model is an ODE model known for its oscillatory properties representing intraindividual variability. It has been used to represent dynamic processes such as emotions (Chow et al., 2005), self-regulation in adolescence substance use (Boker and Graham, 1998), and mood regulation in recent widows (Bisconti et al., 2004), among many other examples. In the illustrative examples in this chapter, we consider SDE variations of the damped linear oscillator model with TVPs. The TVPs considered include both the set-point parameter, \( \mu \), as well as the damping parameter, \( \zeta \).

Every higher-order ODE (i.e., involving higher-order derivatives than just the first derivatives) can be expressed as a system of first-order ODEs. As a special case, the damped oscillator model can also be expressed as as system of two first-order ODEs. In particular, the SDE form of Equation 2.11 is typically expressed as two first-order differential equations involving \( \eta_i(t) = [\eta_{1i}(t) \ \eta_{2i}(t)]' \) in Itô form as:

\[
d\eta_{1i}(t) = \eta_{2i}(t)dt \\
d\eta_{2i}(t) = \left(\omega \left(\eta_{1i}(t) - \mu(t)\right) + \zeta \eta_{2i}(t)\right)dt + \sigma_p dw(t)
\]

where \( d\eta_{1i} \) is the first-order differential and \( d\eta_{2i} \) is the second-order differential of the scalar latent process \( \eta_1 \) associated with unit \( i \). \( dw(t) \) is the differential of a univariate standard Wiener process (i.e. the process noise component in this model) and \( \sigma_p \) is the standard deviation of the
The dynamic functions in (2.12) are assumed in the present context to be identified using a single manifest indicator measured at discrete but possibly irregularly spaced time points as:

\[ y_i(t_{i,j}) = \eta_i(t_{i,j}) + \epsilon_i(t_{i,j}), \quad \epsilon_i(t_{i,j}) \sim N \left(0, \sigma^2_\epsilon\right), \]  

where \( \eta_i(t_{i,j}) \) is the observed manifest indicator for unit \( i \) at discrete time, \( t_{i,j} \); and \( \epsilon_i(t_{i,j}) \) is the corresponding measurement error with variance \( \sigma^2_\epsilon \).

To model the dynamics of the TVPs, we treat the TVPs as unknown latent variables to be inserted into \( \eta_i(t) \) and estimated with other latent variables in the system. Doing so would, in most cases, entail SDEs that are nonlinear because they involve the interaction between at least two latent variables. For instance, in illustrative example 3, we consider the scenario of time-varying damping/amplification parameter \( \zeta(t) \). In such a scenario, even when the true model is known and is fitted directly to the data, the latent variable vector now includes a TVP as \( \eta_i(t) = \begin{bmatrix} \eta_1_i(t) & \eta_2_i(t) & \zeta_i(t) \end{bmatrix}' \). Consequently, the revised Equation 2.12, when written in the form of Equation (2.1), now involves the interaction between two latent variables, \( \zeta_i(t) \) and \( \eta_2_i(t) \). Thus, the dynamic model of interest now becomes nonlinear in \( \eta_i(t) \), even though the damped oscillator function, conditional on \( \zeta_i(t) \), is linear in form. The CDEKF algorithm as implemented in dynr, the R package used for model fitting purposes in the present study, can readily handle this form of nonlinearity.

The CDEKF algorithm can also handle any linear or nonlinear multivariate latent processes as alternatives to the damped oscillator model shown in Equations (2.11) and (2.12), provided that the functions are at least first-order differentiable with respect to the latent variables, and are characterized by Gaussian distributed process noises. These latent processes, in turn, are allowed to be apprehended through multiple observed indicators, but the measurement functions have to be linear, with Gaussian distributed measurement errors. The fact that the measurement model is linear with Gaussian distributed measurement noise also renders the one-step-ahead prediction errors, \( v_i(t_{i,j}) \), Gaussian distributed conditional on values of \( \eta_i(t_{i,j}) \). Thus, estimation of the remaining time-invariant parameters in the extended damped linear oscillator model with TVPs can be performed by optimizing the prediction error decomposition function in (2.9).

### 2.4 Approximation Functions for TVPs

Across many scientific disciplines, it is often impossible to know the true underlying model for the phenomenon of interest (MacCallum, 2003). Devising models for TVPs is even more challenging due to the scarcity of knowledge from the literature on plausible forms of their functional dynamics. Here we propose two flexible models as approximations for any TVP. The approximation functions considered include an Ornstein-Uhlenbeck (O-U) model and a stochastic
The O-U process, which is represented by a first-order SDE (Equation 2.14), has been utilized to represent processes such as emotion regulation (Oravecz et al., 2009), which have the tendency to show exponential return to an equilibrium or "home base" after they are moved away from it by some random shocks or process disturbances. The O-U model for a TVP $\theta(t)$ is expressed as

$$d\theta_i(t) = \beta(\theta_i(t) - \mu_{ou})dt + \sigma_{ou}dw(t)$$ (2.14)

where $\mu_{ou}$ is the equilibrium or attractor the O-U process, $\beta$, constrained to be greater than or equal to 0, is the velocity that the process returns to $\mu_{ou}$, and $\sigma_{ou}$ is the diffusion parameter that governs the standard deviation of the random process noise.

Another approximation function considered in the present chapter is the stochastic noise model, written as:

$$d\theta_i(t) = \sigma_s dw(t).$$ (2.15)

In this model, no deterministic drift function is included. Instead, any changes in $\theta$ are posited to be driven completely by the random shocks captured in $w(t)$. Thus, this function may be suited for capturing processes that show ongoing, noise-like shifts without the tendency to return to an equilibrium.

Both the O-U model and the stochastic noise model have specialized features that render them well-suited for representing change phenomena with particular characteristics. First, both functions are relatively simple. Second, both are built on the notion that the true value of the TVP, though unknown and possibly varying over time, is likely to vary much more gradually than the other latent processes in $\eta_i(t)$. In particular, they both include as a special case the possibility of a completely stable, time-invariant parameter when there is no process noise in the system (namely, when $\sigma_{ou} = \sigma_s = 0$). In the O-U model, if $\theta_i(t)$ is a time-invariant parameter, its value may be set to equal its initial condition (IC) value, $\theta_i(t_{i,1})$, with $\mu_{ou}$, $\beta$, $\sigma_{ou}$, and the corresponding variance for $\theta_i(t_{i,1})$ in the IC covariance matrix, $\Sigma_{\eta_1}$, all set to 0. Consequently, $\theta$ would just be a freely estimated (time-invariant) parameter, denoted herein as $\mu_{\theta_1}$. $\mu_{\theta_1}$ is one of the entries in the IC latent variable vector, $\mu_{\eta_1}$, and its value is estimated by optimizing the prediction error decomposition function. Other less restrictive special cases of the O-U process with $\beta > 0$ allow for transient deviations from an otherwise stable value as captured by $\mu_{ou}$. In a similar vein, a time-invariant parameter would be represented in the stochastic noise model by the value of $\mu_{\theta_1}$ in $\mu_{\eta_1}$, with $\sigma_s$ and the corresponding IC variance of $\theta_i(t_{i,1})$ in $\Sigma_{\eta_1}$ set to 0.

Despite the fact that the O-U model and the stochastic noise model are likely imperfect or misspecified functions of the true change mechanisms of the TVPs, the smoothing procedures summarized in (2.10) can still be used to provide smoothed estimates of the TVPs conditional on the observed data. Visual inspection of such smoothed estimates can, in turn, provide insights into the nature and over-time dynamics of the TVPs, as we show in the context of three illustrative examples.
2.5 Model Fitting via dynr

The model fitting was done in R (R Core Team, 2016) through the package dynr (Ou et al., 2016). The dynr package provides an interface between R and the C language, where the models can be specified in R and the computations are carried out in C. The following components need to be provided for successful model fitting: the measurement model (with \texttt{prep.measurement()}), the measurement and dynamic noise components (with \texttt{prep.noise()}), the initial values of the latent state variables and their covariance matrix (with \texttt{prep.initial()}), and the dynamic model (with \texttt{prep.formulaDynamics()}). Once all these components are specified, they are combined into a dynr model object with \texttt{dynr.model()} for parameter estimation with this estimation done using the \texttt{dynr.cook()} function. There are also options to include transformation functions for parameters with restrictive ranges and upper/lower bounds of plausible estimated values. Once the estimation and smoothing procedures are complete, summary statistics can be extracted using the \texttt{summary()} command. It is also possible to extract parameter estimates using \texttt{coef()} and AIC or BIC measures using \texttt{AIC()} or \texttt{BIC()}. A sample script for one of the illustrative examples is included in the Appendix.
Chapter 3
Illustrative Examples

3.1 Example 1: Stagewise Shifts in Set-Point

For the first example, let us consider a variation of the damped oscillator model with distinct stage-like shifts in the set-point parameter \( \mu \) as the true data generation model. That is, \( \mu \) in Equation (2.11) or (2.12) is now expressed as \( \mu(t) \) and its value is hypothesized to undergo two stage-wise shifts. Denoting the first stage as the “reference stage”, we consider a three-stage model for \( \mu(t) \) as:

\[
\mu_i(t) = \mu_1 + c_2(t)\mu_2 + c_3(t)\mu_3
\]  

(3.1)

where \( c_2(t) \) and \( c_3(t) \) are time-dependent, binary indicators of stages 2 and 3 of the TVP values, while \( \mu_2 \) and \( \mu_3 \) are the differences in equilibria of stage 2 and 3 compared to that of stage 1 (\( \mu_1 \)). If the stage indicators, \( c_2(t) \) and \( c_3(t) \), are known and observed with perfect knowledge, then modeling of the time-varying set-point, \( \mu(t) \), is relatively trivial. However, such knowledge is not always available, and other approximation functions such as the O-U model and the stochastic noise model may have to be used.

Figure 3.1 illustrates the proposed continuous-time process with stage-wise changes in set points around which the system of interest oscillates. One example of such a process would be mood swings in individuals with bipolar I or II disorder. Bipolar disorder (I or II) is a severe chronic disorder that is characterized by clear shifts in mood and energy levels between mood episodes, namely alternating episodes of depression and (hypo)mania. During a depressive episode, an individual would feel sad, unmotivated, and have low levels of energy. During a manic episode, the individual would feel elevated, excited, and have high levels of energy (The National Institute of Mental Health, 2016). Thus, on the episodic level, there is a sharp, stage-wise change in mood (set point); within each episode, however, the individual’s mood would fluctuate from moment to moment (or on an hourly or daily basis) around a set point as a result of environmental stimuli or daily events (Larson et al., 1990; Silk et al., 2011).

Data were simulated for 10 subjects, each with 241 time points \( (t_{i,0} = 0, t_{i,T} = 30, \Delta t = t_{i,j+1} - t_{i,j} = 0.125 \) for all \( i \in \{1, \ldots, 10\} \) and \( j \in \{1, \ldots, 240\} \)\), using the parameter values specified in Table 3.1 under the assumption that the process of interest follows three stages.
Figure 3.1: Simulated observed and latent trajectory for one hypothetical subject generated using the damped oscillator model with stagewise shifts in set-point considered in Illustrative Example 1.

Table 3.1 also summarizes the model fitting results with the true and approximation models. As expected, fitting the true data generation model using the CDEKF and associated algorithms led to satisfactory recovery of all time-invariant parameters. The two approximation models performed reasonably well in recovering most of the parameters, except for an overestimation of the process noise variance, $\sigma^2_p$.

Table 3.1: Parameter Recovery in Illustrative Example 1 Obtained from Fitting the True Data Generation Model and the Two Approximation Models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Estimate (SE)</th>
<th>Stage-wise Model</th>
<th>O-U Model</th>
<th>Stochastic Noise Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0 -0.043 (0.040)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>6 6.003 (0.039)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>10 9.940 (0.040)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\eta$</td>
<td>-2 -2.011 (0.011)</td>
<td>-1.793 (0.035)</td>
<td>-1.807 (0.038)</td>
<td></td>
</tr>
<tr>
<td>$\zeta$</td>
<td>-0.1 -0.102 (0.007)</td>
<td>-0.109 (0.026)</td>
<td>-0.116 (0.027)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_p$</td>
<td>0.25 0.250 (0.044)</td>
<td>1.191 (0.431)</td>
<td>0.764 (0.405)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_n$</td>
<td>1 1.038 (0.031)</td>
<td>1.035 (0.032)</td>
<td>1.038 (0.032)</td>
<td></td>
</tr>
<tr>
<td>$\mu_\text{ou}$</td>
<td>-</td>
<td>-</td>
<td>16.161 (6.004)</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma^2_\text{ou}$</td>
<td>-</td>
<td>-</td>
<td>1.114 (0.205)</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma^2_\text{s}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.508 (0.229)</td>
</tr>
<tr>
<td>AIC</td>
<td>7205.86 7811.90 7840.73</td>
<td>7246.37 7852.42 7869.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>7246.37 7852.42 7869.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computational Time:</td>
<td>20.23 sec 15.80 sec 9.50 sec</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The AIC and BIC both indicated that the true data generation model was preferred over
the other approximation models. This finding highlights, for the first time (to my knowledge), the plausibility of using the AIC and BIC for modeling selection purposes in conditionally linear models with TVPs. The smoothed estimates of the TVP trajectory for one arbitrary subject based on the three fitted models are plotted in Figure 3.2a. It can be seen that the O-U and stochastic noise models, despite being misspecified compared to the true data generation model, still do a reasonable job in capturing the overall patterns of change in $\mu(t)$. For comparison purposes, the root-mean-square error (RMSE) obtained from each function for the plotted subject is also included in Figure 3.2a. RMSE with respect to a TVP of interest $\theta(t)$ is calculated as follows for model $m$ and subject $i$:

$$\text{RMSE}_{i,m} = \sqrt{\frac{\sum_{j=1}^{T_i} [\hat{\theta}_m(t_{i,j}) - \theta(t_{i,j})]^2}{T_i}}$$

(3.2)

where $\hat{\theta}_m$ is the estimated process of $\theta$ by model $m$. Figure 3.2b shows the biases of the TVP estimates, $(\hat{\theta}(t_{i,j}) - \theta(t_{i,j}))$, for that particular subject at each available time point.

Interestingly, even though the stagewise changes in $\mu(t)$ posited in the data generation model were somewhat distinct from the dynamics postulated in the O-U model (e.g., dictating a return to the home base), the smoothed estimates from the O-U model still yielded reasonable approximations of the true trajectory of $\mu(t)$, with only a minor increase in the RMSE in the estimates for $\mu(t)$. Inspection of the parameter estimates from the O-U model in Table 3.1, particularly the high estimated value of $\mu_{\text{OU}}$, suggested the time-varying $\mu(t)$ was estimated as a process that was trying to approach its equilibrium but was not quite there yet by the end of the observations. Because both of the approximation models have stochastic components, the sudden jumps were captured as part of the influence of the process noise, thus resulting in over-estimation of the process noise variance, $\sigma_p^2$ (see Table 3.1).

### 3.2 Example 2: Logistic Growth in Set-Point

For the second example, we model the set-point parameter, $\mu(t)$, as a logistic growth curve:

$$d\mu_i(t) = r\mu_i(t) \left(1 - \frac{\mu_i(t)}{k}\right)dt.$$  

(3.3)

A logistic growth curve takes an elongated “S” shape and has the following two characteristics. First, the system develops in a monotonically increasing fashion. That is, the process does not regress. Second, the system has identifiable upper and lower bounds (Grimm and Ram, 2009). An illustration of a damped oscillator whose equilibrium is undergoing logistic growth is shown in Figure 3.3. Given their characteristics, the logistic curve can thus serve as a reasonable model for many learning and lifespan developmental processes (Grimm and Ram, 2009). Take the scenario of hormonal changes around puberty as an example. There is a shift of overall level of sex hormones from before to after the onset of puberty, though the change may unfold in a gradual way during puberty. On a faster time scale, however, the level of sex hormones has been shown to fluctuate from moment to moment, as subjected to the influence of factors such as time.
Figure 3.2: Estimated TVP trajectories for one hypothetical subject in Illustrative Example 1: (a) estimated and true trajectories, with the corresponding RMSEs; (b) biases in the estimated TVP values in comparison to the true values.

of day and day in the menstrual cycle (Liening et al., 2010; Marceau et al., 2011).

Similar to the previous example, data were simulated for 10 subjects, each with 241 time points \((t_i, 0 = 0, t_i, T = 30, \Delta t = 0.125)\), using the parameter values specified in Table 3.2. As in illustrative example 1, both the AIC and BIC preferred the true model to the other approximation models (see Table 3.2). The true parameters were recovered very well when the true data generation model was fitted; the two approximation models also performed reasonably well in recovering all other parameters except for the process noise variance \(\sigma_p^2\). In contrast to Illustrative Example 1, \(\sigma_p^2\) was under- as opposed to over-estimated in this particular case. Whether this under-estimation in process noise variance was the result of sampling variability or a systematic under-estimation requires further validation via a full Monte Carlo study. The smoothed estimates of the TVP for one arbitrary subject obtained using the three models are plotted in Figure 3.4a. Compared to the smoothed estimates obtained from fitting the true data generation model, the estimated set point trajectories from the two approximation models still captured the overall sigmoid-shaped changes in \(\mu(t)\) reasonably accurately. Greater biases were observed in the TVP estimates obtained using the approximation models between \(t_{i,i}\) of 0 and 15 (see Figure 3.4b), during which the set-point showed more pronounced rises in value as dictated by the logistic growth curve function (see Figure 3.4a). The overall patterns of changes in set-point were well approximated, however, despite these biases in TVP estimation.
Figure 3.3: A simulated set of observed and latent trajectories generated under the scenario of logistic growth as in Example 2.

Table 3.2: Parameter Recovery in Illustrative Example 2 Obtained from Fitting the True Data Generation Model and the Two Approximation Models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Logistic Model</th>
<th>O-U Model</th>
<th>Stochastic Noise Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta )</td>
<td>-1</td>
<td>-0.997 (0.003)</td>
<td>-0.992 (0.016)</td>
<td>-0.981 (0.020)</td>
</tr>
<tr>
<td>( \zeta )</td>
<td>-0.1</td>
<td>-0.098 (0.003)</td>
<td>-0.141 (0.018)</td>
<td>-0.167 (0.025)</td>
</tr>
<tr>
<td>( r )</td>
<td>0.1</td>
<td>0.100(0.000)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( k )</td>
<td>0.5</td>
<td>0.499 (0.002)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( \sigma^2_p )</td>
<td>0.01</td>
<td>0.005 (0.003)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>( \sigma^2_\epsilon )</td>
<td>1</td>
<td>1.046 (0.031)</td>
<td>1.037 (0.031)</td>
<td>1.036 (0.031)</td>
</tr>
<tr>
<td>( \mu_{ou} )</td>
<td>–</td>
<td>–</td>
<td>13.141 (1.255)</td>
<td>–</td>
</tr>
<tr>
<td>( \sigma^2_{ou} )</td>
<td>–</td>
<td>–</td>
<td>0.339 (0.045)</td>
<td>–</td>
</tr>
<tr>
<td>( \sigma^2_w )</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.677 (0.086)</td>
</tr>
<tr>
<td>AIC</td>
<td>6987.08</td>
<td>7404.27</td>
<td>7495.35</td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>7021.80</td>
<td>7444.78</td>
<td>7524.29</td>
<td></td>
</tr>
</tbody>
</table>

Computational Time: 15.49 sec 17.97 sec 11.68 sec

3.3 Example 3: Stagewise Shifts in Damping

In the last example, we allow the damping/amplification parameter, \( \zeta \), to undergo time-varying dynamics as:

\[
\zeta(t) = \zeta_1 + c_2\zeta_2 + c_3\zeta_3
\]  

(3.4)

This model assumes two stage-wise shifts in the TVP as in Example 1, but in the damping/amplification as opposed to the set-point parameter. \( c_2 \) and \( c_3 \) are again indicators of stage 2 and
stage 3. \( \zeta_1, \zeta_2, \) and \( \zeta_3 \) represent, respectively, the value of the damping/amplification parameter in stage 1, and differences between the stage-1 value and the values in the two subsequent stages. Figure 3.5 shows a simulated example of such a case, in which the process goes through an oscillatory phase with no alterations in amplitude (\( \zeta_1(t) = 0 \)), followed by a phase with amplification (\( \zeta_2(t) > 0 \)), and ending with a phase with damping (\( \zeta_3(t) < 0 \)).

Allowing the damping/amplification parameter to be time-varying provides a mathematical model for describing processes that undergo transient periods of increased/decreased fluctuations or instabilities. Consider again the bipolar disorder scenario. Some researchers have noted that the transitions between the depressive and manic episodes are often interspersed with a period of mood instability (Bonsall et al., 2011). The proposed model with stagewise shifts in the damping/amplification parameter, as illustrated in Figure 3.5, provides one way of representing such a transitional phase of increased mood instability. We did not allow the set-point parameter to show concurrent stagewise shifts in the present illustration to ease presentation, but we should add that allowing for stagewise shifts in both the set-point as well as the \( \zeta \) parameter would be more consistent with the scenario of bipolar disorder described here.

For illustration purposes, data were simulated for 10 subjects, each with 241 time points \((t_{i,0} = 0, \ t_{i,T} = 30, \Delta t = 0.125)\), using the parameter values specified in Table 3.3. The first stage has no damping effect (\( \zeta_1 = 0 \)), followed by an amplification stage (\( \zeta_2 = 0.3 \)) and a damping stage (\( \zeta_3 = \ -0.2 \)). Based on the model-fitting results in Table 3.3, the true parameter values were recovered well when the correctly specified model was fitted to the data. The O-U model performed slightly better than the stochastic noise model in yielding smaller parameter biases,
Figure 3.5: Simulated observed and latent trajectory for one hypothetical subject generated using the damped oscillator model with stage-wise shifts in $\zeta$ as considered in Illustrative Example 3. Damping = the damping/amplification parameter, $\zeta_i(t)$.

RMSE, AIC, and BIC compared to the stochastic noise model.

Consistent with the results from the previous illustrative examples, the AIC and BIC selected the true model over the other approximation models as the preferred model. The smoothed estimates of the TVP trajectory for one arbitrary subject are shown in Figure 3.6a. Fitting the correctly specified model recovers the three values of damping. By comparison, estimates from the two approximation models show less salient transitions across the stages. Rather, the TVP trajectories as estimated using the approximation functions convey the shifts from stage 1 to stage 2 more as a gradually increasing linear trend as opposed to a discrete shift. This set of results, compared to that from Example 1, suggests that time-varying damping/amplification parameter may be more difficult to estimate than time-varying set-point parameter, and requires more replications of complete cycle to reliably distinguish discrete from other more gradual shifts in parameter value.
Table 3.3: Parameter Recovery in Illustrative Example 3 Obtained from Fitting the True Data Generation Model and the Two Approximation Models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Stage-wise Model</th>
<th>O-U Model</th>
<th>Stochastic Noise Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>ζ_1</td>
<td>0</td>
<td>-0.019 (0.011)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ζ_2</td>
<td>0.3</td>
<td>0.302 (0.004)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ζ_3</td>
<td>-0.2</td>
<td>-0.200 (0.004)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>η</td>
<td>-2</td>
<td>-2.002 (0.004)</td>
<td>-1.813 (0.033)</td>
<td>-1.98 (0.003)</td>
</tr>
<tr>
<td>σ_σ^2</td>
<td>0.25</td>
<td>0.183 (0.039)</td>
<td>0.108 (0.063)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>σ_η^2</td>
<td>1</td>
<td>0.970 (0.032)</td>
<td>1.038 (0.032)</td>
<td>1.049 (0.032)</td>
</tr>
<tr>
<td>η_{ou}</td>
<td>-</td>
<td>-</td>
<td>-0.002 (0.046)</td>
<td>-</td>
</tr>
<tr>
<td>σ_{η_{ou}}</td>
<td>-</td>
<td>-</td>
<td>0.026 (0.005)</td>
<td>-</td>
</tr>
<tr>
<td>σ_{ε}^2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.024 (0.003)</td>
</tr>
</tbody>
</table>

AIC       | 7227.36 | 7539.15 | 7569.57 |
BIC       | 7262.09 | 7573.88 | 7592.72 |

Computational Time: 14.63 sec 12.81 sec 9.42 sec

Figure 3.6: Estimated TVP trajectories for one hypothetical subject in Illustrative Example 3: (a) estimated and true trajectories, with the corresponding RMSEs; (b) biases in the estimated TVP values in comparison to the true values.
Chapter 4  |
Discussion

In this thesis, an SDE modeling framework with TVPs was proposed as a way to capture multitime scale and self-organizing processes in continuous time. Several variations of the damped oscillator model were considered: two with time-varying set-points and one with time-varying damping/amplification. Furthermore, the use of two functions — the O-U model and the stochastic noise model — is presented and evaluated as plausible functions for approximating changes in the TVPs in the absence of further knowledge concerning their true change mechanisms.

All three of the illustrative examples considered showed that the TVPs can be recovered very well and the smoothed estimates provided accurate reflections of the underlying dynamic processes when the true models were fitted to the simulated data. Approximating the true models of TVP with the O-U model and the stochastic noise model yielded satisfactory estimates of other time-invariant parameters; the smoothed estimates of the TVP trajectories also revealed the general change patterns of the true TVP processes reasonably well, though not as accurately as the smoothed estimates obtained from fitting the true models. Overall, results confirmed the plausibility of using these functions as approximations for model and data exploration purposes. It is further showed that the smoothed latent parameter estimates can be inspected to facilitate later confirmatory modeling of the trajectories of the TVP. Finally, model comparison results using the AIC and BIC provided an initial validation of the utility of these information criterion measures as tools in selecting continuous-time models that involve mild forms of nonlinearity (i.e., nonlinearity due to the TVPs).

Although not considered in one of the illustrative examples in this chapter, it is also possible to allow the frequency parameter, \( \omega \), to be modeled as a TVP. Preliminary results (not shown here) suggested that similar results and modeling conclusions would still hold in scenarios involving a single, time-varying frequency parameter. Relatedly, even though only examples that are extensions of the well-known damped oscillator process were considered, the procedures proposed in this thesis are, in principle, readily generalizable to other dynamic processes of similar characteristics to those considered in the chapter. The relative flexibility of the CDEKF algorithms and the modeling interface provided as part of the \texttt{dynr} package make such generalizations viable, and potentially very doable.

The illustrations and explorations in this thesis of SDEs with TVPs are far from comprehen-
sive. In particular, a full Monte Carlo study should be conducted to evaluate the performance of the CDEKF algorithms, the AIC, and the BIC in fitting and facilitating the selection of SDEs with TVPs. Such a study would likewise help to establish more detailed guidelines in fitting such models. Also worth investigation is the signal-to-noise ratio this proposed data exploration scheme can tolerate. In fitting an approximation model such as the O-U model or stochastic noise model, the “true” noise in the data can be manifested in more than one place through one or several of the process noise parameters in the approximation models. Hence, it is important to know to what extent the approximation models could still capture the underlying changes in the TVPs with increased process and measurement noise.

Several limitations of the proposed estimation procedures should be noted. First, due to the use of a Jacobian matrix $F$ in the prediction step in (2.4), only drift functions that are at least first-order differentiable with respect to elements in $\eta_i(t)$ may be used in the dynamic model. This requirement, however, is not met for some dynamical systems — for instance, examples that show sudden shifts in the TVP, similar to the stagewise true model used in illustrative examples. When preexisting indicators of the exact change points are available (e.g., as reflected in how the true model was fitted to the simulated data), estimation can be handled in a straightforward manner. When direct indicators of these change points are not available, researchers may need to resort to using other differentiable functions, such as the O-U model, as an approximation for such discrete shifts. For better approximation results, we may need to utilize some other filtering methods, such as iterated particle filtering (Ionides et al., 2015), the unscented Kalman filter (Chow et al., 2007), or other regime-switching extensions (Chow et al., 2017) as alternatives for systems that are not first-order differentiable.

A fourth-order Runge-Kutta procedure is used to solve the two ODEs in Equations (2.3) and (2.4) in the prediction step of the CDEKF – namely, to obtain conditional expectations of the latent variables for time $t_{i,j}$ given the estimates from up to time $t_{i,j-1}$, that is, $E(\eta_i(t_{i,j}|Y_i(t_{i,j-1}))$ and $\text{Cov}(\eta_i(t_{i,j}|Y_i(t_{i,j-1})))$. Using numerical solvers such as the Runge-Kutta procedure opens up more possibilities of fitting ODE/SDE models that do not have closed-form solutions. In exchange for this added flexibility, however, is the need to select appropriate $\Delta t$ or time step for the numerical integration as the truncation errors embedded in the numerical solutions are direct functions of $\Delta t$. $\Delta t$ that is too large would yield poor estimation results; in contrast, $\Delta t$ that is too small can be computationally inefficient and may also lead to other numerical problems. For the examples given in this chapter, $\Delta t$ was fixed at 0.125 throughout all the examples. The exact same observations with a $\Delta t$ of 1 would return inaccurate inferred model dynamics and associated parameter estimates, even with the same model fitting procedures and specifications. In some cases, rescaling of the time units — for example, recoding time from seconds to minutes — may be necessary. Other alternatives, such as adaptive ODE solvers (Kulikov and Kulikova, 2014; Kulikova and Kulikov, 2014) and multiple shooting methods (Kiehl, 1994; Gander and Vandewalle, 2007; Stoer and Bulirsch, 2013), may also be used in place of the fourth-order Runge-Kutta procedure to improve the robustness of the estimation procedures to choices of $\Delta t$. In addition, the CDEKF algorithms as currently implemented in dynr rely on the use of first-order Taylor series expansion to approximate the uncertainties implicated in
When the system is highly nonlinear, retaining only the first-order terms in the Taylor series expansion may be inadequate and higher-order terms may have to be incorporated instead (Bar-Shalom et al., 2001).

The smoothing algorithm used in the present study can also be improved in a number of ways. First, in the illustrative examples considered, the smoothed estimates of the TVP obtained from the approximation models generally recovered the true patterns of the TVPs reasonably well. However, there were noticeable delays in recovering the true values of the TVPs around all the transition points in the stagewise examples. One way to circumvent this estimation delay is to reverse the time ordering of the data (so that the last time of observation $T$ now becomes the first time point $t_1$ and so on), repeat the model-fitting procedure, and generate a set of refined smoothed estimates as a weighted average of the smoothed estimates obtained from forward- and reverse-smoothing. This revised smoothing procedure is similar in principle to the zero-phase noncausal filtering utilized in the field of engineering (Powell and Chau, 1991; Kormylo and Jain, 1974). Finally, the prediction error decomposition function used for parameter optimization purposes in (2.9) uses only the one-step-ahead prediction errors and associated covariance matrix from the prediction step. One alternative, which replaces the PED function with a penalized log likelihood function (Fuhrmeir and Wagenpfeil, 1997) constructed using an iterated extended Kalman filter (e.g., Bar-Shalom et al., 2001, p. 404), may be used to improve the quality of the parameter estimates, and consequently, the quality of the smoothed estimates.

In conclusion, modeling in continuous time using SDEs with TVPs provides renewed possibilities for capturing self-organization across different levels of analysis and time scales. In cases where the true underlying mechanisms of interest are unknown, this thesis has illustrated the viability of several approximation models for representing TVP processes that unfold on much slower time scales compared to the changes manifested by other system variables. I hope my subset of examples provides a starting point for researchers to pursue other modeling ideas and applications involving self-organizing phenomena in continuous time.
Appendix
Sample dynr Code for Fitting the Models in the Illustrative Examples.

########## True Model (Three-stage damping parameter) ##########
meas <- prep.measurement(
  values.load=matrix(c(1, 0), 1, 2),
  params.load=matrix(c('fixed', 'fixed'), 1, 2),
  state.names=c("x","y"),
  obs.names=c("obsy"))

ecov <- prep.noise(
  values.latent=diag(c(0, .1), 2),
  params.latent=diag(c('fixed', 'pnoise'), 2),
  values.observed=diag(.1, 1),
  params.observed=diag('mnoise', 1))

initial <- prep.initial(
  values.inistate=c(-5, 0),
  params.inistate=c('fixed', 'fixed'),
  values.inicov=diag(1, 2),
  params.inicov=diag('fixed', 2))

formula3stages<-list(
  x~y,
  y~eta*x+(p1*b1+p2*b2+p3*b3)*y
)
dynamics <- prep.formulaDynamics(formula=formula3stages,
    startval=c(eta=-.5,b1=0.1,b2=0.5,b3=-0.5),
    isContinuousTime=TRUE)

data3stages <- dynr.data(simdata3stages, id="id", time="times",
    observed="obsy", covariates=c("p1","p2","p3"))

model3stages <- dynr.model(dynamics=dynamics, measurement=meas,
    noise=ecov, initial=initial,
    data=data3stages, outfile="phases.c")

model3stages@ub <- rep(10, 6)
model3stages@lb <- rep(-10, 6)
model3stages@lb[model3stages$param.names %in% c("pnoise","ounoise")]
    = -30

res3stages <- dynr.cook(model3stages)
summary(res3stages)

######### O-U Model #########

measOU <- prep.measurement(
    values.load=matrix(c(1, 0, 0), 1, 3),
    params.load=matrix("fixed", 1, 3),
    state.names=c("x","y","z"),
    obs.names=c("obsy"))

ecovOU <- prep.noise(
    values.latent=diag(c(0,0.1,0.1), 3),
    params.latent=diag(c('fixed', 'pnoise','ounoise'), 3),
    values.observed=diag(.3, 1), params.observed=diag('mnoise', 1))

initialOU <- prep.initial(
    values.inistate=c(-5, 0,.1),
    params.inistate=c('fixed', 'fixed','fixed'),
    values.inicov=diag(.01, 3),
    params.inicov=diag('fixed', 3))

formulaOU <- list(
    x~y,
    y~eta*x+z*y,
z - beta*(mu - z)

```r
dynamicsOU <- prep.formulaDynamics(formula=formulaOU,  
                                  startval=c(eta=-.5, beta=.1, mu=10),  
                                  isContinuousTime=TRUE)
```

```r
modelOU <- dynr.model(dynamics=dynamicsOU, measurement=measOU,  
                       noise=ecovOU, initial=initialOU,  
                       data=data3stages, outfile="OU.c")
```

```r
modelOU@ub <- rep(10,6)  
modelOU@lb <- rep(-10,6)  
modelOU@lb[modelOU$param.names %in% c("pnoise","ounoise")]= -30  
modelOU@ub[modelOU$param.names %in% c("mu")]= 20  
resOU <- dynr.cook(modelOU)  
summary(resOU)
```

```r
######### Stochastic Noise Model ##########
formulaOU2<-list(  
                  x~y,  
                  y~eta*x+z*y,  
                  z~0)
```

```r
dynamicsOU2<-prep.formulaDynamics(formula=formulaOU2, startval=c(eta=-.5),  
                                  isContinuousTime=TRUE)
```

```r
# Other model specifics are the same with those in the O-U model above.
modelOU2 <- dynr.model(dynamics=dynamicsOU2, measurement=measOU,  
                       noise=ecovOU, initial=initialOU,  
                       data=data3stages, outfile="StocNoise.c")
```

```r
modelOU2@ub <- rep(10,4)  
modelOU2@lb <- rep(-10,4)  
modelOU2@lb[modelOU2$param.names %in% c("pnoise","ounoise")]= -30
```

```r
resOU2 <- dynr.cook(modelOU2)  
summary(resOU2)
```
Bibliography


equation models with random effects and unknown initial conditions using the Stochastic

(accepted with revisions, 2017). Differential equation modeling approaches to representing
sudden shifts in intensive dyadic interaction data. *Psychometrika*.


measuring changes in international business cycles. Staff Reports 326, Federal Reserve Bank
of New York, New York.

smoothing for non-Gaussian dynamic regression models. *Computational Statistics & Data


Equation Modeling*, 16(4):676–701.


and latent variable models via iterated, perturbed bayes maps. *Proceedings of the National

Press, Cambridge, MA.

of brain and behavior*, chapter 1. The MIT Press, Cambridge, MA.


The National Institute of Mental Health (2016). Bipolar disorder.


