The Pennsylvania State University

The Graduate School

College of Health and Human Development

KALMAN FILTER MODELS FOR ECOLOGICAL MOMENTARY ASSESSMENT DESIGNS

A Thesis in

Human Development and Family Studies

by

Lawrence L. Lo

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

August 2014

The thesis of Lawrence L. Lo was reviewed and approved* by the following:

Steven H. Zarit

Professor of Human Development and Family Studies

Department Head

Peter C.M. Molenaar

Professor of Human Development and Family Studies

Thesis Adviser

Michael J. Rovine

Professor of Human Development and Family Studies

*Signatures are on file in the Graduate School

ABSTRACT

Ecological momentary assessment (EMA) designs attempt to assess individuals in real-time within real-world settings in order to provide a stronger sense of ecological validity. EMA uses temporal sampling schemes that either occur randomly or follow targeted behaviors of interest. Both of these temporal schemes produce data which contain unequal temporal spacing, violating the assumptions of many longitudinal modeling techniques. The current study examines the hybrid Kalman filter (HKF), an extension of the Kalman filter that may be particularly suitable for EMA designs. The HKF is compared with the standard discrete Kalman filter (DKF) approach in two simulation studies reflecting the different temporal schemes. The HKF demonstrated superior performance in these simulation studies; however, the DKF was robust to violations of temporal spacing and exhibited satisfactory performance in certain conditions. Finally, an implementation of the HKF with real data is demonstrated in order to provide an example of the types of results and inferences that the HKF can yield.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my friends and family for their support during my work on this thesis. I am also grateful to my advisors in helping me complete this thesis.

Chapter 1. Introduction

Behavior scientists often aim to make inferences about psychological processes that occur in real-world settings. A clinical psychologist may be interested in how an individual's functional impairments are expressed when facing day-to-day difficulties while a social psychologist may focus on interpersonal interactions that occur in a person's daily life. Traditional research methodology may be limited in both design and statistical approaches when attempting to make inferences about psychological processes in the real world. First, traditional research methods have relied on designs in which behavior either occurs within a lab setting or is retrospectively summarized. Second, traditional statistical approaches have focused on differences between people rather than processes that may occur within people. Inferences from these between-person analyses are often implicitly generalized to make claims about within-person processes. In both design and modeling scenarios, a similar theme occurs where inferences may be incorrectly generalized.

Ecological Momentary Assessment

Ecological momentary assessment (EMA; Stone & Shiffman, 1994; Shiffman, Stone, & Hufford, 2008) is a research design framework that places emphasis on studying dynamic processes that unfold in real time, occur in real-world settings, and vary in context. The key feature of EMA designs are repeated momentary assessments. Repeated assessments allow for investigation of dynamic processes; such inquiry requires multiple time points within an individual. Momentary assessments focus on events or states that are currently or recently experienced. This strategy can avoid problems associated with recall, such as memory distortion and heuristic bias (Clark & Teasdale, 1982; Bradburn, Rips, & Shevell, 1987). EMA designs have been implemented in a variety of substantive domains; reviews of particular domains have appeared for research in clinical treatment (Theile et al., 2002), psychopharmacology (Moskowitz & Young, 2006), personality (Tennen et al., 2005), and industrial psychology (Beal & Weiss, 2003).

Rather than being a single research method, EMA is an approach that combines several methodological strategies. Self-monitoring designs and experience sampling methods, both within the framework of naturalistic observation, are the two primary research strategies in EMA. Self-monitoring designs involve persons reporting experiences directly following a targeted behavior or event (Korotitsch & Nelson-Gray, 1999). The goal of this strategy is to focus on particular behaviors or events of interest; for example, a substance abuse researcher may target times directly following drug use. In contrast, experience sampling designs involve measurements that are triggered by the measurement instrument (e.g., a mobile device) at random time points rather than being triggered by a particular behavior or event (Larson & Csikszentmihalyi, 1983). The goal of experience sampling designs is analogous to that of random sampling from a population in conventional statistical design: the researcher seeks to obtain a representative sample of a person's daily experiences. Although these two strategies utilize different temporal schemes, they both provide repeated momentary assessments. Each strategy has advantages depending on the psychological or behavioral construct of interest (Stone et al., 2007).

Although these temporal sampling schemes may be advantageous from a design perspective, they present unique analytic modeling challenges. Most models for repeated measures data involve the assumption that there are equal temporal spacing intervals between the repeated measurements. Mixed-effects modeling (Laird & Ware, 1982), also known as hierarchical linear modeling (Bryk & Raudenbush, 1992) and multi-level modeling (Singer & Willet, 2003), and time series modeling (Box & Jenkins, 1976; Shumway & Stoffer, 2000) are the main techniques for modeling repeated measures data. Both types of modeling procedures contain the assumption of equal temporal spacing intervals, which is mainly reflected in conditions where variance and autocovariance (i.e., serial dependency or covariation of variables at different time points) are invariant across time.

Several methodological developments within the mixed effects framework have addressed this problem of unequally spaced data, specifically in the context of EMA designs. Hedeker, Mermelstein, and Demirtas (2008) developed extensions of the mixed-effects model to allow for variance structures that can accommodate EMA designs. Schwartz and Stone (1998, 2007) also provided extensions of the mixed-effects model that incorporated variance and serial dependency structures that address the unequal temporal spacing issues within EMA. Although these extensions of the mixed-effects model have benefited EMA research, they maintain some general problems of the mixed-effects model. These problems mainly involve the ecological fallacy, where inferences of models based on aggregates are invalidly generalized to the nature of individuals.

Molenaar (2004) addresses these problems within concepts based on ergodic theory, a branch of mathematics that provides conditions that are necessary for generalizing inferences from pooled aggregates to a single individual within this pooling procedure. These conditions (under Gaussian assumptions) are stationarity and homogeneity. Stationarity refers to the equivalence across time of statistical characteristics such as means, covariances, and autocovariances. Homogeneity refers to the equivalence of statistical characteristics across individuals within the sample to be pooled. This latter condition of homogeneity is of primary concern when considering the mixed effects model. The insensitivity of these models to heterogeneity in individual processes has been demonstrated in both simulation studies (Molenaar, 1997; Molenaar, Huizenga, & Nesselroade, 2002) and formal mathematical proofs (Kelderman & Molenaar, 2007). Empirical demonstrations of heterogenous processes and the insensitivity of mixed-effects models to detect these forms of heterogeneity have occurred in domains such as personality (Borkenau & Ostendorf, 1998) and affective processes (Rovine, Molenaar, & Corneal, 1999; Molenaar et al., 2009)

Person-specific analysis via individual time-series models avoids these problems of the mixed-effect model and have been argued to be necessary for investigating any dynamic or developmental process (Molenaar, 2007; Molenaar and Campbell, 2009). The primary motivation for person-specific analysis is to avoid the strict condition of homogeneity across individuals. Rather than construct a model based on pooled data of aggregates that are assumed to be homogenous, the person-specific framework involves constructing individual time series models separately. The individual models are allowed to be different for each individual without any homogeneity assumption. A practical issue for these individual time series models is the requirement of more time points per person compared to standard approaches such as the mixed-effects model; however, EMA usually involves intensive repeated measures designs that can provide the necessary number of time points for these person-specific methods (Schiffman, Stone, & Hufford, 2008).

The Kalman Filter Framework

The Kalman filter (KF; Kalman, 1960) is one of several time series modeling approaches used in person-specific analysis. Although there are several types of KF models, these different types have common and essential features. A KF consists of (i) a stochastic dynamic system model of state (or latent) variables with Gaussian noise, (ii) a measurement model relating these unobservable states to manifested observations, and (iii) an estimation scheme based on recursive methods. Various extensions and types of the KF have been developed for different model types (i.e., linear vs. nonlinear) and temporal types (i.e., discrete time vs. continuous time); overviews of these different KF types can be found in Simon (2006) and Gibbs (2011). Since the main problem with EMA data involves a problematic temporal scheme, the current paper focuses on different temporal types of KF models.

Recursive methods in the KF framework involve a sequence of prediction and updating stages that can be summarized in a series of steps. First, the dynamic model predicts the states at the current time point based on initial or a priori values. Second, these predicted states are updated based on deviance of these predictions from the observed data at the current time as well as model components that contain uncertainty about the dynamic and measurement models. Next, these updated states are used to make predictions at the subsequent time point (first step repeated), which are updated based on data from the observed data at the subsequent time point (second step repeated). This prediction and update process repeats through each time point, starting at the beginning of the time series and ending at the final time point. In addition, predictions and updates of the state error covariance matrix occur alongside predictions and updates of the previously mentioned actual states. The KF recursion is based on recursive least-squares estimation; a clear demonstration is provided in Simon (2006). Additional details regarding parameter estimation and maximum-likelihood (ML) approaches are given in Harvey (1990) and Gibbs (2011). The Appendix contains details on the recursive estimators of the different models as well as the basic steps and equations for ML parameter estimation.

The most basic form of the KF is the discrete Kalman filter (DKF; Kalman, 1960). The DKF uses difference equations to model processes that are measured at discrete time points which are assumed to be equally temporally spaced. The dynamic model for the DKF is

$$\eta_t = B\eta_{t-1} + \zeta_t \tag{1}$$

where η_t is a q-variate state vector at time *t*, B is a (q,q)-dimensional matrix containing transition (or autoregressive) coefficients, and ζ_t is a q-variate Gaussian process noise vector at time *t*. The measurement model for the DKF is

$$y_t = \Lambda\eta_t + \varepsilon_t \tag{2}$$

where y_t is a p-variate observed variable vector at time *t*, Λ is a (p,q)-dimensional matrix containing measurement loading coefficients, and ε_t is a p-variate Gaussian measurement error vector at time *t*. Both of the stochastic components are assumed to have normal distributions, with $\zeta_t \sim N(0, \Psi)$ and $\varepsilon_t \sim N(0, \Theta)$. An example diagram of a Kalman filter model (or state-space model) with 2 state variables and 4 observed variables is depicted in Figure 1. The recursion equations for the DKF are provided in the Appendix.

Although the DKF model contains the equal interval assumption, some forms of unequal spacing may be handled via methods used for handling missing data (Shumway and Stoffer, 2000). If the unequal intervals have a common positive integer denominator, the transition matrix, B, may be raised to the power corresponding to the interval length. For example, in the case of unequal temporal intervals with a common denominator of one hour, measurements 1 hour apart would use the original B, measurements 2 hours apart would use B^2, measurements 7 hours apart would use B^7. This type of unequal spacing follows a restrictive form where there must be a common denominator.

The DKF model is similar to time series structural equation models (SEM), another popular approach in person-specific time series modeling; however, the two approaches use different estimation schemes. Time series SEM involves fitting similar models to block-Toeplitz covariance matrices (Molenaar, 1985) that incorporate the autocovariance structure of the data. Modeling these covariance structures involves pooling data across time; such procedures assume equal temporal spacing and neglect any data regarding the time intervals. In contrast, the recursive estimation schemes with the KF framework involve raw data procedures (analogous to full information maximum likelihood procedures in SEM). Information on the temporal spacing can be used in the missing data procedure mentioned above. Other conceptual and practical comparisons of the DKF and time series SEM have been addressed in the literature (Molenaar, 2003; Zhang, Hamaker, & Nesselroade, 2008; Chow et al., 2010).

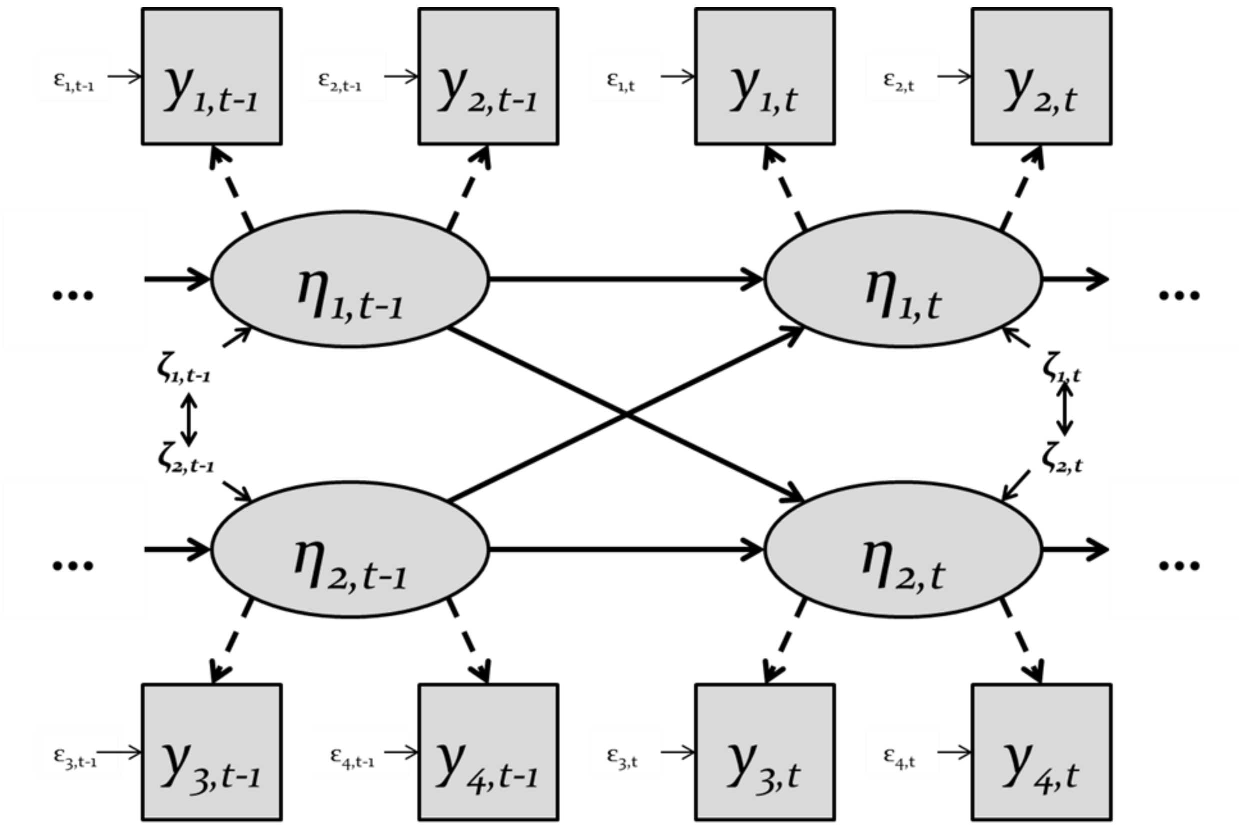

Figure 1. *Depiction of an example state-space model.* The two state variables (η_1, η_2) at time *t-1* predict their values at the subsequent time point *t* with each state also being influenced by a Gaussian white noise component (ζ_1, ζ_2) that can covary concurrently. These unobserved state variables map onto observed variables (y_1, y_2, y_3, y_4) at concurrent time points and contain measurement errors (ε_1, ε_2, ε_3, ε_4).

Although initially developed in the engineering literature, the DKF has recently been applied in time series modeling of psychological processes. Some example applications include research on emotion-cognition linkages (Chow, Hamagani, &

Nesselroade, 2007), dyadic relations (Chow, Ferrer, & Nesselroade, 2007), and affective processes (Molenaar et al., 2009). These example applications involved regular temporal schemes (e.g., daily measurements) and thus were appropriately handled by the DKF model.

The continuous Kalman filter (CKF; Kalman & Bucy, 1961) uses differential equations to model continuous processes that are assumed to be measured continuously. As the name implies, the CKF is the continuous-time analog of the DKF and has some similarities. The dynamic model for the CKF is

$$\dot{\eta}(t) = \mathrm{B}_C\eta(t) + \zeta(t) \quad (3)$$

where $\dot{\eta}(t)$ is a q-variate vector of state derivatives at time *t*, B_C is a (q,q)-dimensional matrix containing continuous transition coefficients, $\eta(t)$ is a q-variate state vector, and $\zeta(t)$ is a q-variate Gaussian process noise vector at time *t*. Here, $\eta(t)$ and $\zeta(t)$ are continuous processes. The measurement model for the CKF is

$$y(t) = \Lambda_C\eta(t) + \varepsilon(t) \quad (4)$$

where $y(t)$ is a p-variate observed variable vector at time *t*, Λ_C is a (p,q)-dimensional matrix containing measurement loading coefficients, and ε_t is a p-variate Gaussian measurement error vector at time *t*. Similar to the dynamic model, $y(t)$ and $\varepsilon(t)$ are continuous processes. Again, both of the stochastic components are assumed to have normal distributions, with $\zeta(t){\sim}N(0,\Psi_c)$ and $\varepsilon(t){\sim}N(0,\Theta_c)$. The recursion equations for the CKF are also provided in the Appendix.

The CKF has a conceptual advantage over the DKF, yet has a similar practical problem. The continuous dynamic process modeled by the CKF may provide a more realistic model of psychological phenomena that change gradually rather than in discrete shifts. The practical problem of the CKF is its assumption of continuous measurement; this temporal condition may be feasible with biological measures of interest to psychologists, yet such a measurement model will be unsuitable for most EMA designs. Although the DKF and CKF make different assumptions about the temporal spacing of the data, both models are similarly unsuitable for EMA designs because of these assumptions.

The hybrid Kalman filter (HKF; Simon, 2006) combines the dynamic model from the CKF (3) with the measurement model from the DKF (2). Likewise, the recursion for the HKF will also be a hybrid of the prediction equations from the CKF and the updating equations from the DKF (see Appendix). The HKF can be characterized as a model that contains a continuous unobservable state process that is measured at discrete time points that can be arbitrarily spaced. A key feature of the HKF is that there is no equal temporal spacing assumption. Likewise, unlike the DKF which can only handle strict forms of temporal irregularities (the previously mentioned common denominator condition), the HKF has no restrictions on the types of temporal irregularities that can occur.

The fundamental difference between the DKF and HKF pertaining to the temporal spacing assumption can be found within the prediction equations for the recursive estimators. For illustration, we focus on the prediction of the states (although the same reasoning applies to the prediction of the state error covariance matrices). The prediction equations for the states within each model correspond to the deterministic parts of the dynamic models. The prediction equation for the DKF is

$$\eta_t = B\eta_{t-1} \tag{5}$$

while the prediction equation for the HKF is

$$\dot{\eta}(t) = \mathrm{B}_C\eta(t) \quad (6).$$

Again, notice that (5) is the deterministic portion of (1) and (6) is the deterministic portion of (3).

The DKF prediction equation (5) makes a next-step prediction by multiplying the previous value by a constant matrix; therefore, a predicted value of the state can only be obtained for a constant time interval. Although there are extensions of the DKF with time-varying transition matrices (Molenaar et al., 2009; Yang & Chow, 2010), these extensions don't explicitly incorporate irregular temporal spacing into the time-varying nature of these models. In contrast, the HKF prediction equation (6) makes a next-step prediction by solving a differential equation that incorporates the time interval. The previous state values are used as the initial values used for the differential equation solution. The key feature is that solving this differential equation involves numerical integration across bounds that specifically include the temporal spacing interval between the time points of interest; this bound can adjust to whatever temporal interval that may occur in the time series. As a result, (6) can produce predicted values that can be obtained for any time interval.

The temporal flexibility of the HKF may make it particularly suitable for EMA designs. Although the DKF may handle simple cases of unequal temporal spacing (situations with a common interval denominator), EMA designs typically produce temporal irregularities that don't follow this form. In contrast, the HKF can handle temporal spacing of any time length, thus being able to handle any irregularity that an EMA design may produce. As previously mentioned, the continuous measurement assumption of the CKF makes it unsuitable for most EMA designs. Although these conceptual differences of the KF model types point towards the HKF being advantageous for EMA designs, no empirical demonstration of differential performance has been conducted.

Discrete and Continuous Noise Relations

There are conflicting accounts on the relationships between discrete and continuous noise process. Although it is agreed upon that a Wiener noise process (continuous time) is the limiting case of Random Walk (discrete time), there are several accounts of the relationships between the discrete and continuous noise variance components in KF models. Most of these proposed relations involve functions that include the time interval leading to the measurement at time t, δ_t. Smith and Roberts (1978) claimed that

$$\Psi_t = \Psi_c \delta_t \quad (7a)$$

$$\Theta_t = \Theta_c \delta_t \tag{7b}$$

where the discrete process (Ψ_t) and measurement (Θ_t) noise components are proportional to the continuous analogs and can vary in proportion to the time interval, δ_t. While (7a) is consistent with Simon (2006), Harvey (1990) and Gibbs (2011) give a different account for the process noise variance relation than includes the transition matrix:

$$\Psi_t = \int_{t-1}^{t} \left[e^{(\tau \mathrm{B} c)}\right] \Psi_c \left[e^{(\tau \mathrm{B} c)}\right]^T d\tau \tag{8}$$

where the time interval is incorporated into the integral bounds. Both Simon and Gibbs give a form for the measurement error variance relation that conflicts with (7b):

$$\Theta_t = {}^{\Theta_c}\!/_{\delta_t} \tag{9}$$

where the discrete measurement error is proportional to the reciprocal of the time interval instead of the original time interval as in (7a). It is generally accepted that time-invariant forms in one representation will produce time-varying forms in the other; i.e., time-invariant continuous variance implies time-varying discrete variance, and likewise, time-invariant discrete variance implies time-varying continuos variance

The question that naturally arises is the choice of which noise components to use for the hybrid model. In Simon's (2006) depiction of the HKF, the process noise is continuous and follows the original continuous form without any discrete approximation.

What is unclear is the choice of the representation of the measurement error variances: should they be (i) the original time-invariant measurement noise variance (Θ) or (ii) the discrete approximations of the continuous measurement noise variances that vary with time interval (Θ_t) given by (7b) or (9)?

The latter option (ii) implies that measurement error variance in the HKF is a function of time. As the distance from the last measurement occasion increases, measurement precision will either increase by (7a) or decrease by (9). Both cases seem substantially questionable, as it is peculiar to think of why measurement precision should be related to temporal proximity of the previous measurement (although counterexamples to this are plausible, they would be exceptions rather than norms in substantive applications). The first option (i) of time-invariant measurement noise variance was chosen in this study because of these conceptual problems with the time-varying measurement noise variance. Although it is possible that the measurement noise variance may be time-varying due to some psychologically relevant process (e.g., fatigue), it is questionable as to why it would be time-varying due specifically to the temporal proximity of the previous measurement.

The Current Study

The current study attempted to examine the performance of the KF framework with data that would follow typical EMA temporal schemes. The DKF and HKF were fit to simulated data following an EMA format in order to address two questions: (i) how much better does the HKF perform in comparison to the DKF in designs with irregular temporal intervals, and (ii) how robust is the DKF to violations of the equal temporal spacing assumption. Two simulation studies were conducted to correspond to the two main EMA temporal sampling strategies. The first simulation corresponded to the experience sampling (randomly triggered) design while the second corresponded to the self-monitoring (behavior-linked) design. These two temporal sampling schemes can produce different types of temporal irregularities that may differentially affect performance of the KF models. Finally, the HKF was applied to real data in order to provide a demonstration of the model with descriptions of the types of results that can be obtained with the HKF.

Chapter 2. Methods

Simulation 1

The first simulation examined performance of the KF types with data following a random temporal scheme. A typical time sampling would involve a common interval that would have some random deviation around it (Shiffman, Stone, & Hufford, 2008); for example, a person could be assessed within an hour window every two hours. The simulation generated data in such a format, with time increments being every 2 hours with a *U*(-0.5, 0.5) random number drawn to provide the time point. This would occur 7 times, yielding roughly a 12 hour daytime sampling period; this would be followed by a 12 hour overnight period that would contain no assessments. There were two time length conditions in this simulation, with one condition containing a shorter time series of one week (49 observations) and the other containing a longer time series of three weeks (147 observations).

The second main factor for this simulation was signal to noise ratio (SNR), and resulted in two different generating models corresponding to low and high SNR conditions. Both models contained three state variables with three within-variable coefficients (diagonal components of B, also known as an autoregressive parameter) and one between-variable coefficient (off-diagonal component of B, also known as a cross-lag parameter); a structural depiction is given in Figure 2. The models used hours as the time-interval unit and were simulated according to the HKF in order to accommodate the irregular time interval scheme. Measurement errors were generated to be uncorrelated and equal, thus yielding a diagonal Θ with equal elements. For the low SNR condition, the coefficients of B were generated to be 0.7 (discrete-time scale) while the measurement errors were 1.0; for the high SNR condition, the coefficients of B were generated to be 0.7 while the measurement errors were 0.5. For the purposes of highlighting potential differences between the HKF and the DKF, relatively simple models were generated. The measurement loading matrix was set to an identity matrix and the process noise covariance matrix was set to diagonal with variance scaled at 0.1.

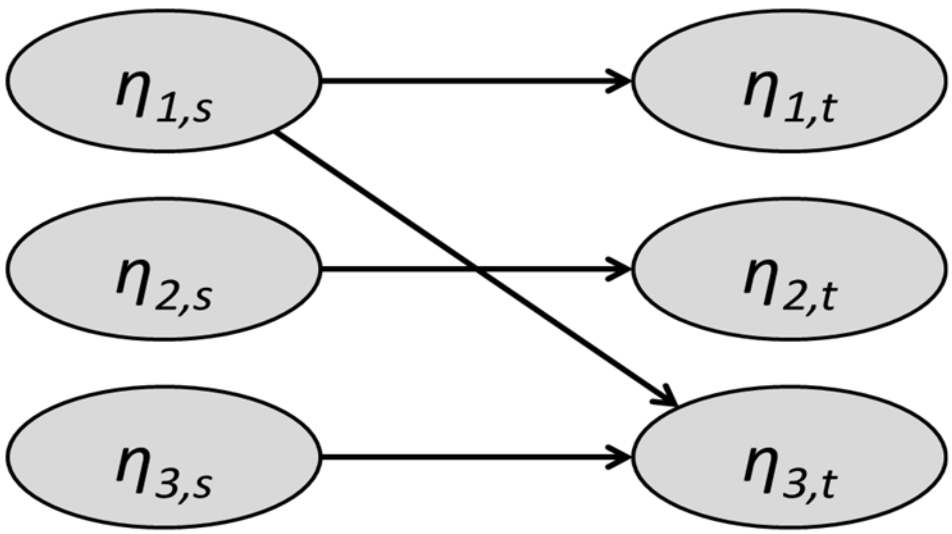

Figure 2. *Structural Model for Simulations.* The three state variables (η_1, η_2, η_3) at time *t are* predicted by themselves at a previous time point *s* that varies according to the simulation design. All of these structural relations have coefficients of 0.7 and have continuous process noise variances of 0.1. In addition, these state variables map onto observed variables with loadings of 1.0 and measurement errors of 0.5 (high SNR) or 1.0 (low SNR).

This resulted in a fully crossed factorial design with parameter bias as the outcome variable. The factors each had two conditions, which were filter type (DKF, HKF), time series length (49, 147 observations), and SNR (low, high). A fully crossed design was implemented in order to examine if there possible interactions of these factors; e.g., if the difference in performance between the filters varied across time series length or SNR. There were 100 replications within each design cell, each using a different random number generating seed value.

Simulation 2

The second simulation used the self-monitoring (event-linked) design as opposed to the random time scheme design used in the first simulation. The time points, rather than being randomly generated, were chosen randomly from real empirical EMA data. This real EMA data set (further described later in the real data demonstration) implemented an event-linked design; the simulation randomly drew people from this study and used their time point data to generate simulated data with features similar to that of the first simulation. The time intervals in these real data sequences were, in general, more irregular than the time-sampling scheme of simulation one. While the minimum and maximum time intervals of the first simulation were constrained to be 1.5 hours and 12.5 hours, the minimum and maximum intervals of the real data set were x and y. Also, while a distribution of time intervals would produce a bimodal distribution (with peaks at 2 and 12) for the first simulation, the second simulation produced a distribution following Figure 3.

Besides this difference in the time-sampling scheme, other simulation factors were generally similar. The generating model and SNR conditions were similar to the first simulation. For comparison of results across simulations, equal time lengths were used for the short and long time series length conditions; thus only the first 49 or 147 of the time intervals were used from the randomly selected data sets. Only time series that exceeded 147 observations were used in order to avoid inconsistencies in time series length. The same fully crossed factorial design described in simulation one was used in this simulation; again, there were 100 replications within each design cell.

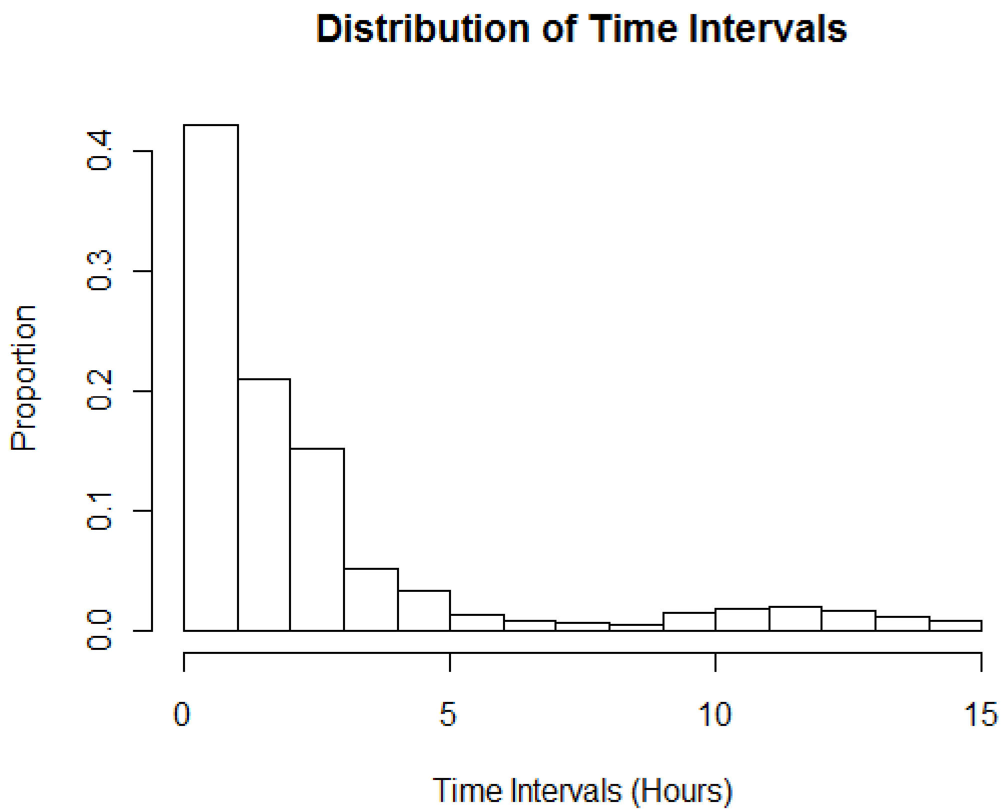

Figure 3. *Distribution of Time Intervals for Simulation 2.* These are the sampled from empirical time-series from the ISAHIB data set. Although the histogram breaks correspond to integer-hours, the time intervals are actually non-integer valued that occur between these integer intervals. Time intervals above 15 hours have been truncated due to low proportions.

Real Data Example

The data for the real data example came from the Intraindividual Study of Aging, Health, and Interpersonal Behavior (ISAHIB). The study implemented an EMA design that followed people for up to a year period with people having 3 sessions (otherwise known as 'bursts') of 21 day periods that involved event-linked assessment via mobile devices. The particular events of interest in this study were social interactions with other people that the person would encounter in their day-to-day situations. Each person was assessed on a wide range of constructs including interpersonal behavior, well-being, and health. For sake of space, two people from this first burst were chosen randomly and analysis was conducted on a subset of 4 variables that formed one of the primary hypotheses of the study. The four variables were (i) interpersonal communication (IC), the coldness/warmth of the person in the interaction, (ii) interpersonal agency (IA), the submissiveness/dominance of the person in the interaction, (iii) affective valence (AV), the negative/positive feeling of this interaction, and (iv) affective arousal (AA), the uninterested/aroused element of this interaction.

A hypothesis of this study was that the two interpersonal perception variables (IC & IA) would have predict subsequent states of the affective variables (AV & AA). A model was fit that contained the following specifications: dynamic relationships from variables IC and IA to variables AV and AA were treated as free(cross-lags), diagonals of the matrix were also treated as free to allow for within-variable dynamic relationships (autoregressions). In addition, the process noise variances and covariances were allowed to be freely estimated. Measurement loadings were fixed to an identity matrix for purposes of model identification and scaling.

Chapter 3. Results

Simulation 1

An ANOVA model was implemented in order to look at the effects of filter type, time series length, and SNR on parameter estimate bias. Three different parameter types were analyzed: a (i) autoregressive component (diagonal element of B), (ii) cross-lag component, and (iii) measurement error variance component. For sake of space, only one of the autoregressive parameters, B(1,1), is presented; the remaining diagonal components had similar results as this component. The ANOVA results, including F statistics and eta squared effect sized estimates are presented for all three parameters in Table 1. The least square means from these ANOVA analyses are presented in Table 2; these least square means reflect the average bias across 100 replications in each design condition. Values near zero signify accuracy, negative values signify a bias toward underestimating the parameter while positive values signify bias towards overestimation, and distance from zero (absolute value) signifies the magnitude or degree of bias.

As predicted, there was a difference in performance between the DKF and HKF filters across parameter types. Although the F statistics in Table 1 indicate significant differences in performance across parameters, the effect size attributable to filter type was more pronounced for the B(3,1) and Θ parameters, with filter type explaining roughly half of the variance in parameter bias for B(3,1) and 2/3 the variance for the parameter bias for Θ. SNR condition was also significant for these two parameter types, although the effect size was a meager .02 for B(3,1) while .l4 for Θ. For the Θ parameter, there was also a significant filter by SNR interaction. For both the B parameters, there was a significant main effect for time series length; however the effect size was small in both situations.

Table 1. *Simulation 1: ANOVA Results*

	B(1,1)		B(3,1)		Θ	
Term	F	Eta^2	F	Eta^2	F	Eta^2
Filter type	15.77*	.02	762.48*	.49	1458.11*	.64
SNR	2.95	.00	12.92*	.02	133.35*	.14
Length	9.74*	.01	22.23*	.03	0.86	.03
Filter x SNR	0.05	.00	3.44	.00	141.77*	.15
Filter x Length	2.49	.00	2.88	.00	1.27	.00
SNR x Length	0.56	.00	6.36*	.01	1.81	.00
3 way int.	0.02	.00	0.22	.00	0.20	.00

Examination of the Table 2 provides a supplementary picture of the comparative performance of the filter types. For the B(1,1) parameter, both the DKF and HKF are negative and thus downward biased, yet the magnitude of bias for the DKF estimates are higher than the HKF estimates. The 3 week time series length is also systematically less biased than the shorter 1 week time series length across the other factors. The B(3,1) parameter yielded strong bias in from both filters in opposite directions; the HKF underestimated this parameter while the DKF overestimated it. The same trend can be seen for time series length, longer time series length yielded more accurate results. The Θ parameter was drastically underestimated by the DKF while only slightly overestimated by the HFK. As consistent with the significant filter by SNR interaction, the bias for of the DKF was more pronounced in the low SNR condition.

Table 2. *Simulation 1: LSM Results*

Filter	SNR	Weeks	B(1,1)	B(3,1)	Θ
HKF	Low	3	-.08	-.19	.04
HKF	Low	1	-.09	-.29	.05
HKF	High	3	-.05	-.21	.03
HKF	High	1	-.08	-.21	.06
DKF	Low	3	-.11	.11	-.48
DKF	Low	1	-.16	.25	-.50
DKF	High	3	-.08	.24	-.22
DKF	High	1	-.15	.31	-.24

Several methods were attempted within the DKF framework. An initial method that completely disregarded the unequal time intervals performed very poorly. A second method that rounded time intervals to the nearest hour and used these lag times of the exponent of B (following the DKF procedure previously mentioned using B^{lag}) to incorporate information on the unequal spacing. A final method used the raw, unrounded time intervals as the B exponent components. This final method produced the best results for the DKF and was therefore used in the HKF-DKF comparison ANOVA models shown in Tables 1 and 2.

Simulation 2

The same ANOVA model was implemented to assess performance, with results for simulation 2 presented in a format similar to that of simulation 1. The ANOVA F and eta-squared statistics appear in Table 3 while the least square means appear in Table 4. There were significant main effects for filter type only for the B(1,1) and B(3,1) parameters; the effect size was more pronounced for the B(1,1) parameter, explaining nearly a third of the variance in parameter bias for this statistic. SNR was a significant main effect across parameters, yet the effect size was most pronounced for the Θ parameter, explaining roughly ¾ of the parameter bias variance. There were no main effects for time series length, however there was a filter type by time series length interaction for B(3,1) and a SNR by time series length interaction for B(1,1). In addition, there were filter type by SNR interactions for both B(1,1) and B(3,1).

Again, the least square means give details on the nature of these effects on parameter bias. The B(1,1) parameter was underestimated across conditions; however the HKF had less magnitude of bias than the DKF. The magnitude of bias was decreased in the high SNR condition across filter types, with the HKF estimates in the high SNR scenario producing quite accurate estimates. Similar findings were observed for B(3,3) with HKF and high SNR producing more accurate estimates; however, the estimates for B(3,3) were generally underestimated with a strong magnitude. Θ was accurately estimated only in the high SNR conditions and was strongly underestimated in the low SNR condition. Unlike the B parameters, estimates of Θ were unaffected by filter type.

Table 3. *Simulation 2: ANOVA Results*

	B(1,1)		B(3,1)		Θ	
Term	F	Eta^2	F	Eta^2	F	Eta^2
Filter type	365.94*	.32	52.51*	.06	2.60	.00
SNR	126.82*	.14	30.02*	.04	2495.33*	.76
Length	1.45	.00	0.59	.00	2.73	.00
Filter x SNR	10.49*	.01	4.32*	.01	0.52	.00
Filter x Length	0.22	.00	8.40*	.01	2.00	.00
SNR x Length	7.38*	.01	2.16	.00	0.06	.00
3 way int.	2.95	.00	0.22	.00	0.02	.00

Table 4. *Simulation 2: LSM Results*

Filter	SNR	Weeks	B(1,1)	B(3,1)	Θ
HKF	Low	3	-.18	-.30	-.40
HKF	Low	1	-.19	-.43	-.40
HKF	High	3	-.05	-.17	.01
HKF	High	1	-.09	-.21	.01
DKF	Low	3	-.53	-.49	-.39
DKF	Low	1	-.48	-.47	-.42
DKF	High	3	-.26	-.44	.00
DKF	High	1	-.33	-.37	-.02

Similar to the first simulation method, several strategies were attempted for obtaining optimal DFK estimates. The same three methods were attempted (no lag shift, nearest integer lag shift, raw lag shift), with the third method same method again producing the best performing results for the DKF. These estimates were used in the ANOVA models presented in Tables 3 and 4.

Real Data Example

Separate HKF models were fit to two random individuals form the ISAHIB data. Estimates from each individual are presented side by side in order to facilitate comparisons between the individual models. The models and parameter estimates are presented in Figure 4. These models focused on the structural components of the models proposed above; thus only structural regressions, covariances, and variances were estimated. The estimates represent the coefficients that have been transformed into discrete time metric with the time interval corresponding to hours.

There are several structural relationships that are of particular interest. The cross-lag relationship from IA to AA is among the largest of structural relationships for both individuals and also shows the greatest similarity for both individuals (.66 for person 1 and .68 for person 2). Thus, interpersonal agency strongly predicts future affective arousal for both persons. Notable structural differences in cross-lag relationships are IA to AV, which is weak for person 1 (.07) while moderately strong for person 2 (.38), and IC to AV, which is strong for person 1 (.68) while moderately strong for person 2 (.39). For affective valence, previous interpersonal communication is strongly predictive for person 1 while both interpersonal communication and interpersonal agency are moderately predictive. The autoregressive relationships are generally weak except for AV (.20 for person 1 and .26 for person 2); thus only affective valence shows a moderate degree of stability for both persons.

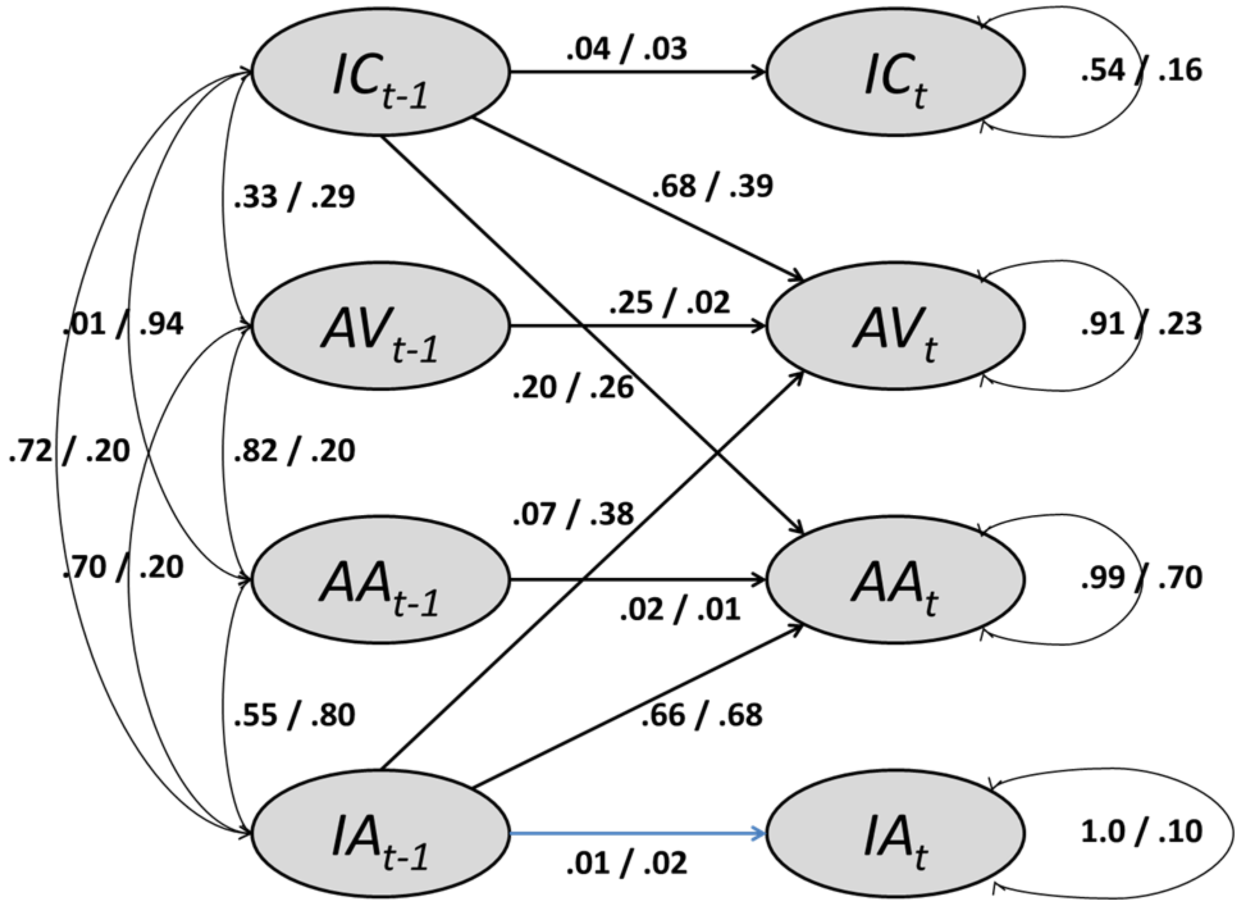

Figure 4. *Real Data Demonstration Results.* The four variables were interpersonal communication (IC), interpersonal agency (IA), affective valence (AV), and affective arousal (AA). For spacing purposes, process noise covariances are depicted by double-headed arrows between variables on the left side of the diagram while process noise variances are depicted on the right by double-headed headed arrows within variables. Estimates for the first individual are initially given, followed by a '/' and the second person's estimates.

There are different patterns of process noise variances and covariances between the individuals. The process noise variances for the first person are larger than those of the second across all four variables. This suggests that the first person shows more intraindividual variability in general. As for process noise covariances, the relationship between IC and AV shows the greatest similarity across the individuals (.33 for person 1 and .29 for person 2) while the relationship between IC and AA shows the greatest difference (.01 for person 1 and .94 for person 2). Such process noise covariances describe the concurrent relationships between variables; such relationships may be due to endogenous (direct relationships between the variables) or exogenous (relationships due to some unmeasured factor) sources.

Modeling procedures for these real data demonstrations involved a more extensive modeling procedure than the methods used in the simulations studies. Initial efforts using ordinary ML routines would converge near their starting values, even after trying several different starting values. This suggests that likelihood functions generated by these individuals contain several local minima that would attract nearby starting values in the parameter space. In order to confront this problem, a global routine (Hu et al., 1994) was used based on simulation procedures in order to account for local minima. First, multivariate uniform distributions were generated for each parameter that were bounded by values that corresponded to the bounds of stationary processes. 10,000 randomly generated sets of starting values were sampled from these distributions and had their likelihood values evaluated. Second, the 5 best sets of starting values were chosen based on evaluated likelihood values and were subsequently used for 5 separate ML runs. Finally the best of these 5 models (also chosen by likelihood value criteria) were chosen as the final model for the individual. This provided ML estimates that weren't based on an a priori choice of starting values.

Chapter 4. Discussion

As predicted, the HKF generally provided more accurate parameter estimates compared to the DKF. The particular conditions where this was most evident were autoregressive parameters in both time sampling schemes and measurement errors in the random time sampling time scheme. The HKF either was comparable or superior to the DKF; no scenario showed a relative disadvantage for the HKF. The HKF, which is specifically designed to handle irregular spaced data, demonstrated its advantages in these simulation studies.

The DKF was either accurate or comparable in performance within several conditions. Similar to the HKF, the DKF yielded accurate estimates of measurement error variances in the high SNR condition in the event-linked simulation design. The autoregressive parameters were only slightly more biased than the HKF estimates in the random time sampling simulation. These findings demonstrate certain situations when the DKF may be robust to violation of its assumption of equal time spacing, yet this only occurred when the DKF was assisted by using the previously mentioned adjusted lag scheme to account for the unequal time spacing.

The cross-lag parameter appeared to present trouble to both filter types. In the random time sampling scenario within simulation 1, both filter types were biased in opposite directions (DKF: upward, HKF: downward) with similar bias magnitudes. In the event-linked time sampling scenario within simulation 2, both filter types were downward biased with similar magnitudes. It could be the case that these the types of time irregularities occurring in these simulations could produce difficulty in capturing cross-lagged relationships regardless of the estimation scheme.

The real data demonstration generally confirmed the above-mentioned hypotheses and displayed several similarities and differences between people. With the exception of one parameter for one person (cross-lag from IA to AV for person 1), the two interpersonal perception variables (IC & IA) either moderately or strongly predicted subsequent states of the affective variables (AV & AA), as consistent with the research hypotheses. The main similarity between the individuals was the cross-lag relationship from IA to AA, suggesting a strong predictive relationship from interpersonal agency to subsequent affective arousal for both persons. The notable differences between individuals were the predictors of affective valence; interpersonal communication was the sole predictor for person 1 while both interpersonal variables were moderately predictive for person 2.

As demonstrated in this real data example, the process of modeling persons individually allows researchers to observe similarities or differences as they appear in the modeling procedure; no prior assumptions are needed about equivalences or differences between people. The individual parameters can subsequently be used in different types of interindividual analyses. One potential question of interest is examining if there are grouping patterns in these individual models; interindividual modeling procedures such as cluster analysis can be used to investigate whether there are groups of individuals that share particular model patterns. Another interidividual procedure could examine how these parameters to relate to other interindividual variables of interest (e.g., sex, age, clinical status). This sequence of analysis initially models intraindividual variability in persons and subsequently examines interindividual differences in these intraindividual models.

Limitations and Future Directions

There were several limitations in the current study. First, the generating models in the simulations were relatively simple; theses models only included parameters corresponding to dynamic relationships and measurement variances (which were set to be equivalent). Process noise variances were fixed and process noise covariances were all simultaneously generated and fixed at 0; as the real data scenarios demonstrated, these particular parameters may be important because they reflect contemporaneous relationships between the variables. Another simplification of the simulation models was the measurement structure: there was a 1-to-1 state to observed variable correspondence as opposed to the common situation of many observed variables mapping onto a smaller dimensional set of state variables (as depicted in Figure 1). A second limitation was the limited range of SNR conditions, there were only two broad conditions (low, high) used that may not capture an adequate range of SNR conditions that would occur in practice.

There are also limitations concerning the ability to apply formal statistical tests to the parameter estimates of the HKF. In the real data example, Wald tests based upon the numerically approximated Hessian matrix provided questionable results and were not reported due the unreliability of the Hessian approximation. There were also attempts at constructing likelihood ratio tests by constraining parameters to zero and comparing likelihood values between the constrained and unconstrained models; however, models with parameters constrained to zero failed to converge, even after adjusting ML tolerance and iteration criteria. Gibbs (2011) and Harvey (1989) provide methods for obtaining the information matrix of the parameters analytically for the DKF; such methods would provide better procedures for conducting Wald tests for parameters. Future work should examine if there are similar or analogous methods for analytically obtaining the information matrix for the HKF in order to facilitate the use of Wald tests.

In addition to addressing these limitations, there are several promising directions for future research in this area. The above limitations can be addressed by an expanded set of simulations to cover a breadth of scenarios encompassing different model types, a wider range of parameter types, and more broad SNR conditions. Another potential area for research would be to examine exploratory modeling techniques that aid in identifying relevant model components; such methods could either be generated with no a priori hypothesis or can expand on such hypotheses. Such methods have been developed in the time series SEM framework (Gates et al., 2010; 2011). Candidate methods for doing this could involve techniques involving sequential likelihood ratio tests; these could involve Lagrange multiplier methods to estimate likelihood ratio statistics during parameter estimation, as implemented in the Gates studies, or brute-force methods that carry out empirical likelihood ratio tests.

Other fruitful areas of expansion on the HKF involve effects on longer time scales. The current model only examines effects that are proximal in time. There are potential psychological constructs where effects may occur on longer time scales, such as across days, weeks, or months. For example, the ISAHIB data has subsequent bursts of data 4-6 months apart that include an additional 3 week periods of data collection; possible research questions could involve effects that could change across measurement bursts. Possible methods for incorporating these types of effects could include design variables as external input components of Kalman filter models (see Simon, 2006) or multi-sample methods that treat bursts as separate samples with likelihood ratio tests to examine differences between burst periods.

Conclusions

The HKF may be ideal for EMA designs for multiple reasons. First they can best handle the irregular time spacing problems that occur in EMA research; this can account for the time irregularities produced by both event-linked and random-sampling time sampling schemes. Second, the HKF can capitalize on the rich within-individual data structures to provide for person-specific inferences on dynamics. As previously mentioned, this allows for examination of more intricate dynamic processes as well as possible heterogeneity within these dynamic processes.

EMA research is gaining in popularity and is becoming more feasible with advances in mobile devices and related data collection software. Similarly, time series and person-specific methodological approaches are gaining popularity in the statistical analysis field. Both EMA designs and person-specific analysis are quite complementary to each other: the EMA design method demands new analytical devices while the person-specific analytical method demands designs that produce richer data structures. Both of these methodological approaches used in tandem could provide valuable insights for a variety of substantive investigations.

REFERENCES

Beal D.J. & Weiss H.M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods 6,* 440‑464.

Borkenau, P. & Ostendorf, F. (1998). The Big Five as states: How useful is the five-factor model to describe intraindividual variations over time? *Journal of Personality Research, 32,* 202–221.

Box, G. & Jenkins, G. (1976), *Time series analysis: forecasting and control*. Oakland, California: Holden-Day

Bradburn N, Rips L, & Shevell S. 1987. Answering autobiographical questions: the impact of memory and inference on surveys. *Science 236,* 157–61.

Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Chow, S-M., Ferrer, E. & Nesselroade J. R. (2007). An unscented Kalman filter approach to the estimation of nonlinear dynamical systems models. *Multivariate Behavioral Research, 42 (2),* 283-321.

Chow, S-M., Hamagami, F, & Nesselroade, J.R (2007). Age differences in dynamical emotion-cognition linkages. *Psychology and Aging, 22 (4),* 765-780.

Chow, S-M., Ho, M.H.R., Hamaker, E., & Dolan, C. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling, 17,* 303-332.

Clark D.M. & Teasdale J.D. (1982). Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *Journal of Abnoral. Psychology 91,* 87‑95.

Gates, K. M., Molenaar, P. C. M., Hillary, F., Ram, N., & Rovine, M. (2010). Automatic search in fMRI connectivity mapping: An alternative to Granger causality using formal equivalences between SEM path modeling, VAR, and unified SEM. *NeuroImage, 53,* 1118-1125.

Gates, K. M., Molenaar, P. C. M., Hillary, F., & Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage, 54,* 1151-1158.

Gibbs, B.P. (2011). *Advanced Kalman filtering, least-squareds and modeling: a practical handbook.* Hoboken, N.J.: Wiley.

Harvey, A.C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hedeker D., Mermelstein R.J., & Demirtas H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics, 64(2),* 627-634.

Hu, X. and Shonkwiler, R. & Spruill, M.C. (1994). *Random Restarts in Global Optimization*, Georgia Institute of technology, Atlanta.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering, 82,* 35-45.

Kalman, R.E. & Bucy, R.S. (1961). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering, 83,* 95-108.

Korotitsch W.J., Nelson-Gray R.O. (1999). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment 2,* 415‑425.

Laird, N.M. &Ware, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics, 38 (4)*, 963–974.

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social and Behavioral Science*, *15,* 41-56.

Kelderman, H. & Molenaar, P.C.M. (2007). The effect of individual differences in factor loadings on the standard factor model. *Multivariate Behavioral Research*, *42*, 435-456.

Molenaar, P.C.M. (1997). Time series analysis and its relationship with longitudinal analysis. *International Journal of Sports Medicine, 18*, 232–237.

Molenaar, P.C.M., Huizenga, H.M., & Nesselroade, J.R. (2002). The relationship between the structure of inter-individual and intraindividual variability: A theoretical and empirical vindication of Developmental Systems Theory. In U.M. Staudinger & U. Lindenberger (Eds.), *Understanding human development* (pp. 339–360.). Dordrecht, The Netherlands: Kluwer.

Molenaar, P. C. M. (2003). *State space techniques in structural equation modeling: Transformation of latent variables in and out of latent variable models.* Unpublished manuscript, University of Amsterdam. Retrieved from http://www.hhdev.psu.edu/hdfs/faculty/pubs/StateSpaceTechniques.pdf

Molenaar, P.C.M. (2004). A manifesto on Psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2 (4),* 201-218.

Molenaar, P.C.M. (2007). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation. *Developmental Psychobiology, 50,* 60-69.

Molenaar, P.C.M., & Campbell, C.G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychology, 18 (2),* 112-117.

Molenaar, P.C.M., Sinclair, K.O., Rovine, M.J., Ram, N.,&Corneal, S.E. (2009). Analyzing developmental processes on a individual level using non-stationary time series modeling. *Developmental Psychology, 45,* 260–271.

Moskowitz D, Young SN. 2006. Ecological Momentary Assessment: what it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry and Neuroscience. 31,* 13‑20.

Rovine, M.J., Molenaar, P.C.M., & Corneal, S.E. (1999). Analysis of emotional response patterns for adolescent stepsons using P-technique factor analysis. In R. K. Silbereisen & A. von Eye (Eds.), *Growing up in times of social change* (pp. 261–285). Berlin: De Gruyter.

Schwartz J.E. & Stone A.A. (1998). Data analysis for EMA studies. *Health Psychology, 17,* 6‑16.

Schwartz J.E. & Stone A.A. (2007). Analysis of real-time momentary data: a practical guide. See Stone et al. 2007, pp. 76‑113.

Simon, D. (2006). *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience.

Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Shiffman, S., Stone, A.A., Hufford, M.R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4,* 1-32.

Shumway, R.H. & Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. New York: Springer.

Smith, M.W.A., & Roberts, A.P. (1978). An exact equivalence between the discrete- and continuous-time formulations of the Kalman filter. *Mathematics and Computers in Simulation, 20,* 102-109.

Stone A.A. & Shiffman S. (1994). Ecological momentary assessment in behavioral medicine. *Annals of Behavioral Medicine 16,* 199–202.

Stone A.A., Shiffman S., Atienza A., & Nebeling L. (2007). *The Science of Real-Time Data Capture: Self-Reports in Health Research*. New York: Oxford Univ. Press

Tennen H., Affleck G., & Armeli S. (2005). Personality and daily experience revisited. *Journal of Personality. 73,* 1–19.

Thiele, C., Laireiter, A-R., & Baumann, U. (2002). Diaries in clinical psychology and psychotherapy: a selective review. *Clinical Psychology and Psychotherapy, 9,* 1–37.

Yang, M-S. & Chow, S-M. (2010). Using state-space models with regime switching to represent the dynamics of facial electromyography (EMG) data. *Psychometrika, 74(4),* 744-771

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparisons of four methods for estimating dynamic factor models. *Structural Equation Modeling, 15 (3),* 377-402.

APPENDIX

Recursion

The recursion for the DKF consists of the prediction equations:

$$\eta_{t|t-1} = \mathrm{B}\eta_{t-1|t-1}$$
$$P_{t|t-1} = \mathrm{B}P_{t-1|t-1}\mathrm{B}^T + \Psi \qquad \text{(A1)}$$

where $\eta_{t|t-1}$ is the predicted state estimated at time current time t while $\eta_{t-1|t-1}$ is the updated state estimated at the previous time point $t-1$, and likewise $P_{t|t-1}$ is the predicted state covariance matrix estimated at time current time t while $P_{t-1|t-1}$ is the updated state covariance matrix estimated at the previous time point $t-1$. the . The updating equations for the DKF recursion are:

$$\varepsilon_t = y_t - \Lambda\eta_{t|t-1}$$
$$S_t = \Lambda P_{t|t-1}\Lambda^T + \Theta$$
$$K_t = P_{t|t-1}\Lambda^T S_t^{-1} \qquad \text{(A2)}$$
$$\eta_{t|t} = \eta_{t|t-1} + K_t\varepsilon_t$$
$$P_{t|t} = (I - K_t\Lambda)P_{t|t-1}(I - K_t\Lambda)^T + K_t\Theta K_t^T$$

where ε_t is the residual at time t, y_t is the observed data at time t, and S_t is the residual covariance matrix at time t, and K_t is the Kalman gain matrix at time t.

The recursion for the CKF consists of the equations:

$$\dot{\eta}(t) = B_c\eta(t) + K(t)\varepsilon(t)$$
$$\dot{P}(t) = B_cP(t) + P(t)B_c^T + \Psi_c + K(t)S(t)K^T(t)$$
$$\varepsilon(t) = y(t) - \Lambda_c\eta(t) \qquad \text{(A3)}$$
$$S(t) = \Theta_c$$
$$K(t) = P(t)\Lambda_c^T S(t)^{-1}$$

where $\dot{P}(t)$ is the derivative of the state covariance matrix at point t on the continuous timeline while $P(t)$ is the state covariance matrix at this time point; $\varepsilon(t), y(t)$, $S(t)$, and $K(t)$ are the continuous analogues of the residuals, observed data, residual covariance matrices, and Kalman gain matrices, respectively. Unlike the DKF and HKF, there is no distinction between the prediction and update steps in the CKF. Similar to the HKF, these recursions are obtained by numerically solving these differential equations with the numerical integration bounds corresponding to the time interval between measurements.

As previously mentioned, the HKF combines the prediction equations from the CKF (A3) with the updating equations from the DKF (A2). There is a slight modification of the CKF prediction equations to account for the hybrid nature of the model.

$$\dot{\eta}(t) = B_c\eta(t)$$
$$\dot{P}(t) = B_cP(t) + P(t)B_c^T + \Psi_c \qquad \text{(A4)}$$

which are slightly abridged from (A3). Solving the differential equation for the state $\dot{\eta}(t)$ yields the predicted state estimate $\eta_{t|t-1}$ that is used for the update equation of (A2). Similarly, solving the differential equation for the state covariance matrix $\dot{P}(t)$ yields the predicted state estimate $P_{t|t-1}$ that is used for the update equation of (A2). The updated values $\eta_{t|t}$ and $P_{t|t}$ obtained by the (A2) are subsequently used as initial values for solving the differential equations for predictions at the subsequent time point.

Maximum likelihood estimation

The recursion equations given in (A1-A4) only give state estimates with given model parameters $(\mathrm{B}, \Psi, \Lambda, \Theta)$. The recursion can be embedded in a likelihood function in order to yield maximum likelihood estimates by some numerical optimization scheme. The log-likelihood function for the DKF, CKF, and HKF take an equivalent form:

$$LL(\theta) = \frac{1}{2}\sum_{t=1}^{N}[pln(2\pi) - ln|S_t| - \varepsilon_t^T S_t^{-1} \varepsilon_t] \qquad \text{(A5)}$$

where p is the number of observed variables and N is the number of time points. The discrete-time indexed components S_t and ε_t are converted to their continuous time analogues $S(t)$ and $\varepsilon(t)$ for the continuous model log-likelihood function. Wald tests for the parameters can be obtained by the Hessian matrix evaluated at the ML parameter values. Standard errors for the Wald tests can be obtained from the following transformation of the Hessian:

$$SE(\theta) = \{diag[H(\theta)^{-1}]\}^{1/2} \qquad \text{(A5)}$$

where $H(\theta)$ is the Hessian matrix evaluated at the ML parameter values. The Hessian matrix can be obtained numerically alongside most optimization routines; however, this numerical approximation may be unreliable and provide inadequate standard errors.