T E C H N I C A L

R E P O R T

# 0694

# A HIERARCHICAL ORNSTEIN-UHLENBECK DIFFUSION PROCESS FOR MULTIVARIATE LONGITUDINAL DATA

ORAVECZ, Z. and F. TUERLINCKX

I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Running head: A HIERARCHICAL ORNSTEIN-UHLENBECK

# A hierarchical Ornstein-Uhlenbeck diffusion process for multivariate longitudinal data

Zita Oravecz & Francis Tuerlinckx

Department of Psychology, University of Leuven

Address all correspondence to:

Zita Oravecz

Department of Psychology, University of Leuven

Tiensestraat 102

3000 Leuven

Belgium

Tel.: ++32 16 32 61 36

Fax : ++32 16 32 59 93

Abstract

In this paper, a diffusion model for the analysis of multivariate longitudinal data with continuous and possibly unbalanced measurement times is presented. The central idea is to model the data from a single person with an Ornstein-Uhlenbeck diffusion process. We extend it hierarchically by allowing the parameters of the diffusion process to vary randomly over different persons. With this approach, both intra- and interindividual differences can be analyzed simultaneously. The individual difference parameters can be regressed on covariates, thereby providing an explanation. We apply the method to data from an experience sampling study to investigate changes in the core affect. It can be concluded that some Big Five factors (extraversion, agreeableness, neuroticism) are related to features of the trajectories in the core affect space, such as average pleasantness, autocorrelation, variability of the movements.

A hierarchical Ornstein-Uhlenbeck diffusion process for multivariate longitudinal data

Many characteristics are constantly subject to change. Emotions and mood are obvious examples, but also measures that are intuitively believed to be very stable such as personality characteristics (e.g., Borkenau & Ostendorf, 1998) reveal their changing nature when repeated measures are taken. The analysis of a multivariate or vector-valued time series measured for a single individual is well developed. The dynamic factor analysis model (Ferrer & Nesselroade, 2003; Molenaar, 1985; Wood & Brown, 1994) - which is an improvement upon the older P-technique factor analysis (Cattell, 1963) - provides a widely used example of a method for such a multivariate time series.

Besides variation within a single individual, people also tend to differ from each other with respect to almost every conceivable psychological characteristic. A single characteristic of this nature is often measured repeatedly for a sample of individuals. The analysis of such a relatively short univariate time series collection (also called longitudinal data) is commonly carried out with multilevel models (Goldstein, 2003; Raudenbush & Bryk, 2002). Such models cope in a straightforward way with unbalanced data (i.e., the number of measurements and measurement times can differ over persons) and with intra- and interindividual difference variables.

The analysis of multivariate longitudinal data is not as well developed as the univariate case (Dubin & Müller, 2005; Weiss, 2005). In the case of multivariate longitudinal data, several response variables are measured at multiple time points for a sample of individuals. The methodology for multivariate longitudinal data is a currently active field of research, both in psychology (e.g., Blozis, 2004; Bollen & Curran, 2006; MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997; Thum, 1997) and in biostatistics (e.g., Becket, Tancredi, & Wilson, 2004; Fang, Tian, Xiong, & Tan, 2006; Fieuws & Verbeke, 2004; O'Brien & Fitzmaurice, 2005; Thiébaut, Jacqmin-Gadda, Babiker, & Commenges, 2005).

In this paper, we present a random-effects model for the analysis of multivariate longitudinal profiles with continuous and possibly unbalanced measurement times. The main idea is to model the sequence of multivariate measurements within a person with an Ornstein-Uhlenbeck (OU) diffusion process. A diffusion process is a continuous-time continuous state Markov process. Both the cross-correlation (dependence between the several response variables) and the serial correlation structure are represented in a natural way in the diffusion process-based model. In addition, we allow the driving parameters of the OU process to differ over persons (i.e., they are random effects). The mean as well as the variance and covariance parameters will be turned into random effects. In this way, not only the interindividual but also the intraindividual variability can be described and explained (by regressing the random effects onto predictors or covariates; see De Boeck & Wilson, 2004).

Because the OU process has a stationary distribution, the proposed model is especially useful for phenomena that exhibit a centralizing tendency. Moods and emotions are prime examples of domains in which a centralizing tendency can be found (Larsen, 2000; Lykken & Tellegen, 1996). The idea is that every individual has their own ideal set point with respect to a certain emotion or mood. It can be assumed that people adjust their current emotional state towards this ideal set point whenever there is a deviation, and the larger the deviation, the stronger the adjustment. Evidence for such a homeostasis effect in emotions has been recently found by Chow, Ram, Boker, Fujita and Clore (2005) using a damped oscillator model, which shows some similarities with the OU model proposed in this paper.

The OU process and some variants (e.g., the integrated OU process) have been proposed as a models for the analysis of longitudinal profiles in several fields. For example, single time series from measurements of animal movement have been modeled as an OU process by Blackwell (1997, 2003), Brillinger (2004) and Dunn and Gipson (1977). The position of a tagged animal is recorded at regular times and the resulting trajectories are modeled with an OU process. However, a major difference between our psychological approach and this biological application is that in the latter one there is no interest in describing and explaining differences between subjects. In the context of longitudinal data analysis, Taylor, Cumberland, and Sy (1994) have proposed a univariate integrated OU process for the serial correlation process in a linear mixed model, besides the regular Gaussian measurement error (see Verbeke & Molenberghs, 2000). In the case of an integrated OU process, it is the first derivative that follows an OU process (for details, see below). A possible disadvantage of the integrated OU process is its non-stationarity (the variance of the process depends on the measurement time) and therefore Taylor et al. (1994) model the difference between consecutive measurements. Sy, Taylor, and Cumberland (1997) use the bivariate integrated OU process for the modeling of bivariate longitudinal data. De la Cruz-Mesía and Marshall (2006) propose the OU model as a model for serially correlated error in a nonlinear mixed-effects model.

In addition, it has to be noted that diffusion models have been applied extensively in psychological research; most often as reaction time models in the area of elementary decision tasks and information processing paradigms. Commonly the well-known Wiener process with absorbing boundaries has been used extensively (Ratcliff, 1978, 2002; Ratcliff & Tuerlinckx, 2002), but Diederich (1997) and Diederich and Busemeyer (2003) apply an OU process to data from elementary decision tasks. Comparing this field with the type of application in the current paper, the major difference lies in the nature of reaction time (and choice) modeling, i.e. the realized sample path of the diffusion model is considered unobservable. As a model for reaction times and choice responses, the only observable features are the time to absorption and possibly the boundary of absorption and from this information, some basic features of the process have to be determined.

The structure of the remainder of the paper is as follows. In the next section, we explain the OU diffusion process together with the interpretation of its parameters. Subsequently, we discuss a hierarchical extension. An application to the modeling of core affect trajectories is presented afterwards and the paper ends with a discussion.

## The Ornstein-Uhlenbeck Diffusion Process

In this section, we want to give a non-technical and self-contained account of the OU diffusion process. A more detailed explanation can be found in Cox and Miller (1972), Dunn and Gipson (1977), Karlin and Taylor (1981) and Blackwell (2003). At this point, we do not yet present an application, so our explanation will be in general terms without reference to a substantive area.

Let us assume that the state of an individual at time $t$ ($t \geq 0$) can be represented as a point $\boldsymbol{Y}(t) = (Y_1(t), Y_2(t), \ldots, Y_q(t))^T$ in a continuous $q$-dimensional space. In general, an OU diffusion process is a continuous-time Gaussian process $\{\boldsymbol{Y}(t) : t \geq 0\}$ defined on this $q$-dimensional space such that, given that the process was in state $Y(t)$ at time $t$, the conditional distribution of the position $\boldsymbol{Y}(t+d)$ $d$ time units later is:

$$\boldsymbol{Y}(t+d) \big| \boldsymbol{Y}(t) = \boldsymbol{y}(t) \sim \boldsymbol{N}_q(\boldsymbol{\mu} + e^{-\mathbf{B}d}(\boldsymbol{y}(t) - \boldsymbol{\mu}), \boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}'d}), \tag{1}$$

where $\boldsymbol{\mu}$ is a $q$-dimensional vector and $\mathbf{B}$ and $\boldsymbol{\Gamma}$ are $q \times q$ matrices. The function $e^M$ (with $M$ a square matrix) is the matrix exponential [1].

If $\mathbf{B}$ is a positive definite matrix, then the Ornstein-Uhlenbeck process is stationary, because as $d \to \infty$, the matrix exponential part $e^{-\mathbf{B}d} \to 0$ and the process has the following equilibrium or stationary distribution:

$$\boldsymbol{Y}(t) \sim \boldsymbol{N}_q(\boldsymbol{\mu}, \boldsymbol{\Gamma}). \tag{2}$$

From Equation 2, it can be deduced that $\boldsymbol{\mu}$ is the mean of the equilibrium distribution and $\boldsymbol{\Gamma}$ is its covariance matrix (hence, $\boldsymbol{\Gamma}$ is positive definite). The assumption of stationarity implies that if the process runs for an infinite long period of time, this equilibrium density is the density function of the visited points in the $q$-dimensional space. It will also be assumed in this paper that the distribution of $Y(0)$ (the first measurement, when time $t$ equals 0) is the equilibrium distribution.

From the formulation of the conditional or instantaneous (if $d$ is very small) mean of the process $\boldsymbol{\mu} + e^{-\mathbf{B}d}(\boldsymbol{y}(t) - \boldsymbol{\mu})$ (see Equation 1), we can see that the conditional mean depends on the previous position $\boldsymbol{y}(t)$, on the time difference $d$ between the two measurements and on two parameters of the process: $\boldsymbol{\mu}$ and $\mathbf{B}$. The interpretation of these two parameters will be explained now. As explained above, the parameter $\boldsymbol{\mu}$ is the centre of the equilibrium distribution and can thus be seen as a point of attraction in the $q$-dimensional space. Because of this property, $\boldsymbol{\mu}$ is also denoted as the homebase of the process.

The $q \times q$ matrix $\mathbf{B}$ controls the strength of the centralizing force, which keeps the process in the vicinity of the homebase. This matrix represents a mean reverting or centralizing tendency, since it impedes the process to diffuse away from the homebase. To interpret the centralizing tendency, it is easier to start with its explanation for a one-dimensional model ($q$=1). In that case, the conditional mean of $Y(t+d)$ given $y(t)$ is equal to

$\mu + e^{-\beta d}(y(t) - \mu)$ (with β >0). From the latter equation, it can be easily derived that if β is large, the conditional mean is very close to the homebase μ. When β approaches zero, the homebase takes a value close to the previous position (with the Wiener process as limiting case if β→0, provided that γβ remains constant where γ is the standard deviation of the univariate equilibrium distribution).

To simplify the interpretation in the general $q$-dimensional case, we only deal in this paper with the subset of isotropic $\mathbf{B}$ matrices. The matrix $\mathbf{B}$ is said to be isotropic if $\mathbf{B} = \beta \boldsymbol{I}$, where $\boldsymbol{I}$ is the $q \times q$ identity matrix (because $\mathbf{B}$ is required to be positive definite, β >0). If $\mathbf{B}$ is restricted to be isotropic, it can be shown that $e^{-\mathbf{B}d} = e^{-\beta d}\boldsymbol{I}$. The characteristics of the conditional mean $\boldsymbol{\mu} + e^{-\beta d}(\boldsymbol{y}(t) - \boldsymbol{\mu})$ reveal that if $\beta$ becomes large (i.e., strong centralizing tendency), the exponential factor goes to 0, meaning that the next point comes from a normal distribution with the homebase $\boldsymbol{\mu}$ as a mean. Alternatively, when $\beta$ is small (i.e., weak centralizing tendency), the exponential factor is close to 1 and the next position is a draw from a normal distribution with the previous point as a mean. Thus, as in the one-dimensional case, the matrix exponential part $e^{-\beta d}$ behaves as a weight function, taking values between 0 and 1, and adding a certain proportion of the distance between the previous point and the homebase to the homebase.

Given a previous point $\boldsymbol{y}(t)$ and a mean $\boldsymbol{\mu}$, Figure 1 depicts the 0.5 probability contour curves of three conditional distributions from Equation 1 for a two-dimensional ($q$=2) model. In all three cases, all parameters are kept constant, and just the centralizing tendency differs (small, medium, large). With a small $\beta$ value, the instantaneous mean of the distribution of the next point is close to the previous point but as $\beta$ increases the instantaneous mean is moving closer to the homebase. Note that the instantaneous variance also depends on the centralizing tendency parameter: a larger $\beta$ value involves a larger variability (see later).

INSERT FIGURE 1 ABOUT HERE

For isotropic $\mathbf{B}$ matrices, it holds that for any dimensionality of the space, the conditional mean lies somewhere on the line connecting $\boldsymbol{\mu}$ and $\boldsymbol{y}(t)$. A related advantage with isotropic matrices is that the centralizing tendency does not give any special importance to the chosen coordinate system (the centralizing tendency matrix is invariant under rotation and reflection; see also Blackwell, 1997). Thus, the centralizing tendency is controlled only by the distance from the homebase of the process and not its direction [2].

In Appendix A, it is shown that the (continuous) autocorrelation function $\rho(d)$ of an OU process equals $e^{-\beta d}$. Because the autocorrelation function is exponentially decaying as a function of continuous time, the OU process is the continuous time variant of what is known in the time series literature as the autoregressive process of order 1 (denoted as AR(1); see Brockwell & Davis, 2002). Figure 2 shows the change in the autocorrelation function as a function of time in the case of four different $\beta$ values. Although the range of the $\beta$ parameter is relatively small, it leads to a remarkable effect on the slope of the autocorrelation function. High (low) $\beta$ values lead in general to a low (high) autocorrelation because the serial correlation function decreases more (less) steeply.

INSERT FIGURE 2 ABOUT HERE

The matrix $\boldsymbol{\Gamma}$ is the covariance matrix of the stationary distribution in Equation 2 and plays a role in the conditional covariance as well. The matrix $\boldsymbol{\Gamma}$ is a positive definite symmetric $q \times q$ matrix containing the variances for each dimension as diagonal elements (denoted with the corresponding row index) and the covariances as off-diagonal elements:

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_1 & \cdots & \gamma_{1q} \\ \vdots & \ddots & \vdots \\ \gamma_{q1} & \cdots & \gamma_q \end{pmatrix}.$$

Large variance values imply that the process can go through dramatic changes (i.e., it is very volatile), while small variances presume smoother trajectories. The covariances represent the extent to which changes in one dimension tend to covary with changes in another dimension. It is often referred to as the volatility parameter.

Examining the features of the instantaneous variance $\boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}'d}$ we can see that as the exponential part goes to 0 (i.e., a large centralizing tendency and/or time difference), the instantaneous variance converges to the variance of the stationary distribution. As the exponential part goes to 1 (i.e., small centralizing tendency and/or time difference), the instantaneous variance becomes very small. To illustrate this, assume for simplicity that $q$ = 2. Then $\boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}'d}$ equals:

$$\begin{pmatrix} \gamma_1\left(1-e^{-2\beta d}\right) & \gamma_{12}\left(1-e^{-2\beta d}\right) \\ \gamma_{12}\left(1-e^{-2\beta d}\right) & \gamma_2\left(1-e^{-2\beta d}\right) \end{pmatrix},$$

from which we can see the effect of the centralizing tendency $\beta$ for a higher dimensional case (given a constant time difference). On one hand, when $\beta$ is large, $\gamma_1$ and $\gamma_2$ are multiplied by a number close to one and hence the instantaneous variances are near the variances of the stationary distribution. On the other hand, when $\beta$ is small, the conditional variances are close to 0. The same reasoning applies to the covariances. However, when considering the instantaneous cross-correlations, it can easily be seen that they are independent of the centralizing tendency and the time difference. Moreover, they are equal to the correlations of the stationary distribution ( $\rho = \frac{\gamma_{12}}{\sqrt{\gamma_1} \times \sqrt{\gamma_2}}$ ).

An equivalent and convenient way of representing the OU process is through a stochastic differential equation (SDE). SDEs provide an intuitive appealing way of understanding diffusion processes and they have a direct link with the more familiar field of deterministic differential equations. However, a full and rigorous treatment of SDEs would require the length of a monograph. Arnold (1974), Karlin and Taylor (1981) and Smith (2000) are good starting points for the study of SDEs. To simplify matters, we will assume that $q = 1$, such that the vector $\boldsymbol{\mu}$ and the matrices $\mathbf{B}$ and $\boldsymbol{\Gamma}$ reduce to the scalars $\mu$, $\beta$ and $\gamma$, respectively. For the one-dimensional OU process considered in this paper, the corresponding SDE equals:

$$dY(t) = \beta\left[\mu - Y(t)\right]dt + \sqrt{2\beta\gamma}dB(t)\,, \tag{3}$$

where $dY(t)$ is the random change in the process $Y(t)$ in a small time interval $dt$ and $B(t)$ represents a univariate standard (i.e., driftless and variance equal to one) Brownian motion process. As can be seen from Equation 3, the change in $Y(t)$ consists of two components: a deterministic part and a stochastic part (embodied by the Brownian motion term). If $\gamma \to 0$, the stochastic part disappears and we are left with a simple first-order linear differential equation. One can also see that the magnitude of change in the deterministic part depends on the difference between the homebase $\mu$ and the current position $Y(t)$. If there were no stochastic disturbance in Equation 3, the solution of the deterministic differential equation would be equal to (given an arbitrary initial value at time 0 of $Y(0)$):

$$Y(t) = \mu + \left[Y(0) - \mu\right] e^{-\beta t}.$$

This solution is exactly the scalar version of the mean of the conditional distribution in Equation 1. Given the initial value $Y(0)$, the exact position of the deterministic process can be found for every time difference $t$. Moreover, if the time difference becomes large, the process converges to the homebase μ. However, when the stochastic disturbance term is added again, the OU process is retrieved and after a very long time difference, the exact position of the process is unpredictable because of the inherent stochastic nature of the process.

## A hierarchical extension of the OU diffusion process

For the case where longitudinal data are collected for a random sample of persons, it is natural to consider a hierarchical extension of the OU diffusion process in order to describe and explain interindividual differences. Let us first fix some notation. A specific person $p$ ($p$=1,…,$P$) is measured $n_p$ times at the sequence of time points: $t_{p1}, t_{p2}, \ldots, t_{ps}, \ldots, t_{p,n_p}$. Note that we do not require that persons are measured at regular time intervals or that they are measured at exactly the same time points. The measured sequence of positions in the multidimensional space is denoted as $y(t_{p1}), y(t_{p2}), \ldots, y(t_{ps}), \ldots, y(t_{p,n_p})$. In the hierarchical extension of the OU diffusion process, it will be assumed that the individual sets of OU parameters are drawn from a common population distribution.

For all persons, the model for the first observation of the chain of measurements is the person-specific equilibrium distribution:

$$\boldsymbol{Y}(t_{p1}) \sim \boldsymbol{N}_q(\boldsymbol{\mu}_p, \boldsymbol{\Gamma}_p).$$

This assumption can be justified, since in many applications the process has been diffusing long enough to have converged to its stationary distribution and forgotten its initial position. For the subsequent points, the conditional distribution of a single person $p$ at time $t_s$ given its position in the previous position $\boldsymbol{y}(t_{p,s\text{-}1})$ is normal with instantaneous mean $\boldsymbol{\delta}_{ps}$ and variance $\boldsymbol{\Lambda}_{ps}$:

$$\boldsymbol{Y}(t_{ps}) \big| \boldsymbol{Y}(t_{p,s-1}) = \boldsymbol{y}(t_{p,s-1}) \sim N_q\left(\boldsymbol{\delta}_{ps}, \boldsymbol{\Lambda}_{ps}\right) \tag{4}$$

where $\boldsymbol{\delta}_{ps} = \boldsymbol{\mu}_p + e^{-\mathbf{B}_p(t_{ps}-t_{p,s-1})}(\boldsymbol{y}(t_{p,s-1}) - \boldsymbol{\mu}_p)$ and $\boldsymbol{\Lambda}_{ps} = \boldsymbol{\Gamma}_p - e^{-\mathbf{B}_p(t_{ps}-t_{p,s-1})}\boldsymbol{\Gamma}_p e^{-\mathbf{B}_p'(t_{ps}-t_{p,s-1})}$.

The individual parameters on the population level come from different distributions. The homebase $\boldsymbol{\mu}_p$ is assumed to follow a multivariate normal distribution. The mean vector $\boldsymbol{\mu}$ is the population mean of the means of the individual stationary distributions. The covariance matrix $\boldsymbol{\Sigma}_{\mu}$ represents the variations and associations that exist in the population between the individual means of the stationary distributions. It should be noted once more that each person has their own equilibrium distribution but the means come from a $q$-dimensional normal population distribution.

Not only is the mean of the stationary distribution allowed to be person-specific, but also its covariance matrix is assumed to be person specific:

$$\mathbf{\Gamma}_p = \begin{pmatrix} \boldsymbol{\gamma}_{1p} & \cdots & \boldsymbol{\gamma}_{1qp} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\gamma}_{q1p} & \cdots & \boldsymbol{\gamma}_{qp} \end{pmatrix}.$$

Therefore, we need to propose a population distribution for these covariance matrices, one that ensures the positive definiteness of $\mathbf{\Gamma}_p$. Moreover, we would like to regress the variances and covariances on covariates. The latter requirement rules out the inverse-Wishart distribution as candidate population distribution, since it does not allow in a natural way the regression of variances and covariances on covariates. One possible solution is to decompose the covariance matrix (Barnard, McCulloch, & Meng, 2000), usually into standard deviations and correlation matrices and assume proper distributions for the latter in order to ensure the positive definiteness. In this paper, we will split the covariance matrix into variances and a correlation. Next, the logarithms of the variances are sampled from a normal distribution. Using the Fisher transformation of the correlation coefficient, we are able to sample this transformed value from a normal distribution, which also provides the possibility of regressing the mean of this distribution on covariates. However, this solution for the decomposition is restricted to the two-dimensional case (and given the application we have in mind, that is sufficient). To regress (co)variances on predictors in the general multidimensional solution, we refer to Daniels and Pourahmadi (2002), who use a Cholesky decomposition for the covariance matrix.

For the special case where $q = 2$ (as in the application of this paper), the simple solution looks as follows. The logarithm of the variances $\gamma_{1p}$ and $\gamma_{2p}$ can be sampled from a multivariate normal distribution:

$$\begin{pmatrix} \log(\gamma_{1p}) \\ \log(\gamma_{2p}) \end{pmatrix} \sim \boldsymbol{N}_2(\boldsymbol{\omega}_{\log(\gamma)}, \boldsymbol{\Sigma}_{\log(\gamma)}),$$

where $\boldsymbol{\Sigma}_{\log(\gamma)} = \begin{pmatrix} \sigma^2_{\log(\gamma_1)} & \sigma_{\log(\gamma_1\gamma_2)} \\ \sigma_{\log(\gamma_1\gamma_2)} & \sigma^2_{\log(\gamma_2)} \end{pmatrix}$ is the population variance-covariance matrix of the log-gammas. The correlation of the population distribution of the variances represents the dependence between the log-variances of the two dimensions.

The covariance parameter $\gamma_{12p}$ of $\boldsymbol{\Gamma}_p$ can be expressed in terms of standard deviations and the correlation $\gamma_{12p} = \sqrt{\gamma_{1p}} \times \sqrt{\gamma_{2p}} \times \rho_p$. Instead of proposing a population distribution for the covariance parameter, it will be assumed that the Fisher-*z* transformed (or *z*-transformed in short) individual-specific cross-correlation coefficient $\mathrm{F}(\rho_p)$ is drawn from a normal population distribution:

$$\mathrm{F}(\rho_p) = \frac{1}{2}\log\left(\frac{1+\rho_p}{1-\rho_p}\right) \sim \boldsymbol{N}(\mathrm{F}(\rho), \sigma^2_{\mathrm{F}}).$$

The cross-correlation for a person *p*, $\rho_p$ indicates the extent to which changes in one dimension tend to correlate with changes in the other dimensions for person *p*. From the mean of the population distribution of the *z*-transformed $\rho_p$, it can be learned whether there is on average (i.e., in the population) a positive, negative or zero correlation between the changes in two dimensions.

Finally, because the centralizing tendency matrix is isotropic ($\mathbf{B}_p = \beta_p \boldsymbol{I}$), we need to assume a population distribution only for the single parameter $\beta_p$. It is assumed in this paper that $\beta_p$ is sampled from a lognormal distribution (it has to be larger than zero) so that:

$$\log(\beta_p) \sim \boldsymbol{N}(\omega_{\log(\beta)}, \sigma^2_{\log(\beta)}).$$

In a next step, covariates can be introduced into the model by regressing the individual OU parameters onto these covariates. This allows for explaining the individual differences in characteristics of the trajectories. Suppose *k* covariates are measured and $x_{jp}$ denotes the score of person *p* on covariate *j* ($j = 1,\ldots,k$). Then we can collect all covariate scores into a vector (together with a constant 1 for the intercept) $\boldsymbol{x}'_p = \left(1, x_{1p}, x_{2p} \ldots, x_{kp}\right)$. Moreover, let $\boldsymbol{\alpha}_{\mu_1}$ be the vector with regression coefficients for the regression of $\mu_1$ (i.e., the population homebase of the first dimension). The vectors with regression coefficients for the other parameters are given analogous names (e.g., $\boldsymbol{\alpha}_{\mu_2}$ for $\mu_2$, $\boldsymbol{\alpha}_{\gamma_1}$ for $\gamma_1$, etc.) .

For the homebase, we assume the following linear regression model:

$$\boldsymbol{\mu}_p \sim N_2\left(\begin{pmatrix} \boldsymbol{x}'_p \boldsymbol{\alpha}_{\mu_1} \\ \boldsymbol{x}'_p \boldsymbol{\alpha}_{\mu_2} \end{pmatrix}, \boldsymbol{\Sigma}_\mu\right)$$

where $\boldsymbol{\Sigma}_\mu = \begin{pmatrix} \sigma^2_{\mu_1} & \sigma_{\mu_1\mu_2} \\ \sigma_{\mu_1\mu_2} & \sigma^2_{\mu_2} \end{pmatrix}$ is the covariance matrix.

The logarithms of the diagonal elements of the volatility covariance matrix $\boldsymbol{\Gamma}_p$ are regressed on the covariates in the following way:

$$\log\begin{pmatrix} \gamma_{1p} \\ \gamma_{2p} \end{pmatrix} \sim \boldsymbol{N}_2\left(\begin{pmatrix} \boldsymbol{x}'_p \boldsymbol{\alpha}_{\gamma_1} \\ \boldsymbol{x}'_p \boldsymbol{\alpha}_{\gamma_2} \end{pmatrix}, \boldsymbol{\Sigma}_{\log(\gamma)}\right).$$

The *z*-transformed correlation coefficient $\mathrm{F}(\rho_p)$ is also regressed on covariates:

$$\mathrm{F}(\rho_p) \sim \boldsymbol{N}(\boldsymbol{x}'_p \boldsymbol{\alpha}_{\mathrm{F}}, \sigma^2_{\mathrm{F}}) .$$

The logarithm of the centralizing tendency is sampled from a normal distribution in the following way:

$$\log\left(\beta_p\right) \sim \boldsymbol{N}\left(\boldsymbol{x}'_p \boldsymbol{\alpha}_\beta, \sigma^2{}_{\log(\beta)}\right) .$$

## Application to core affect trajectories

The hierarchical OU model as described in the previous sections will serve as a model for the trajectories of individuals in the core affect space (Russell, 2003). According to Russell (2003), the core affect lies at the heart of a person's emotional experience and can be characterized as a compound of hedonic (pleasure-displeasure) and arousal (deactivated-activated) values. This core effect is always part of the human psyche as a consciously accessible neurophysiological state and it is changing continuously over time. The core affect space is defined by two dimensions: activation (vs. deactivation) and pleasantness (vs. unpleasantness). Consequently, the emotional experience at a particular moment can be represented as a single point in the two-dimensional plane, and the itinerary of a person's emotional experience is the core affect trajectory.

Our goal with modeling the core affect trajectories with an OU model is twofold. First, we want to describe the individual and population characteristics of movement throughout the core affect space. By making use of a stochastic model approach such as the OU model, we are able to treat the elapsed time as continuous and model the two dependent variables (pleasantness and activation) simultaneously (together with their cross-correlation). Moreover, we are able to evaluate the strength of the centralizing tendency and individual difference herein. As a second goal, we want to explain the individual differences in the characteristics of individual trajectories. In doing this, we could try to answer such questions as: Can the average position (mood) of an individual be predicted from some of their major personality dimensions?

The most common method for collecting data about such trajectories is experience sampling (Bolger, Davis & Rafaeli, 2003; Csikszentmihalyi & Larson, 1987; Larson & Csikszentmihalyi, 1983; Russell & Feldman-Barrett, 1999). Persons are surveyed repeatedly at random chosen time points with respect to their position in the core affect space and this assessment takes place in the natural environment of the participants. From data obtained by experience sampling, we are able to investigate the intra- and interindividual variation in core affect position.

As an illustration of the hierarchical OU model for the core affect trajectories, we analyze part of the data described in Kuppens, Van Mechelen, Nezlek, Dossche, and Timmermans (2006). In that study, 56 students from the University of Leuven were paid to give systematic self-reports about their emotional state in the core affect space during one week. A wristwatch beeped nine times per day. The participants were provided with a booklet in which they could indicate their positions on the Affect Grid (Russell, Weiss & Mendelssohn, 1989; see the used format of the grid on Figure 3). When the participants missed a beep, they had to indicate why they missed it, and most of the time the reason was that they did not hear it. In the analysis the missing data were considered to be missing at random (MAR; see Little & Rubin, 2002) and hence it was assumed that there was no observation at that time. The elapsed time interval between the measurements was semi-random. The participants were asked to give information about the time when they were awake, this interval was divided into equal periods, and a random beep was scheduled into each period. As a consequence of this procedure, we do not have measurements for the evenings.

INSERT FIGURE 3 ABOUT HERE

In addition to the experience sampling, in an introductory session the participants also completed the Dutch version of NEO-FFI (Hoekstra, Ormel, & de Fruyt, 1996), which is a questionnaire to measure the dimensions of the Five Factor model of personality (Big Five). The NEO-FFI consists of 60 items divided equally into five scales which asses Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness. All items are rated on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). The five factors will be used as covariates to explain individual differences in the characteristics of core affect trajectories.

The average age of the participants was 21 years (*SD* = 1.9) and 70% of them were women. The maximum number of measurements for a single person was 63 and on average there were 60 measurements per person (*SD* = 2.9).

An exploratory data analysis was carried out to investigate the main characteristics of the measurements. First, we present two typical person profiles from the data set (see Figure 4). The subsequent measurements are connected with straight lines. We can see from these profiles that there is variability in the occupied positions in the core affect space but we can also notice that its extent may differ somewhat between the individuals.

INSERT FIGURE 4 ABOUT HERE

Figure 5 shows a smoothed heat map of the visit frequencies in the core affect space, based on the aggregated data. We can clearly see a central area, where most of the visits are concentrated. We expect the location of the population distribution of homebase to be somewhere in that area.

INSERT FIGURE 5 ABOUT HERE

Figure 6 shows the estimated vector-field of the core affect grid with the data pooled together for all individuals in the dataset. For each cell in the core affect grid, we calculated the average distance traveled when leaving the cell. (The last measurements for each day were not taken into account because the transition from the last measurement of the previous day and the first of the current day is not necessarily the same process as transitions within a day; see also below). The length of the vector is proportional to the estimated escape velocity from that cell and its angle corresponds to the direction of escape. A central location in the grid can be observed from where the average velocity to move away is very small. As we are moving away from this central point, we can notice a tendency to be pulled towards this center: the direction of most of the vectors are more or less pointed to the central location and with increasing distance from the central point, the vector length tends to increase. (At the border cells, we notice more irregularity - these are due to sampling variability because there are much less visits to these outer cells; see also the heat map in Figure 5.)

INSERT FIGURE 6 ABOUT HERE

Because of the obvious day-night problem (i.e., transitions within a day do not necessarily result from the same process as transitions overnight; see also above), we extend the hierarchical OU model with an additional level. Instead of having measurements (level 1) nested within persons (level 2), we will now consider a three-level model: measurements nested within days nested within persons. In this way, we avoid making strong assumptions about the process of overnight transitions. To illustrate how the equations change, assume that the days are denoted by an index $r$ ($r$ = 1,…,7). Then the conditional distribution of a single person $p$, on day $r$, at time $t_s$ given its position in the previous position $\boldsymbol{y}(t_{p,r,s\text{-}1})$ on the same day, is normal with instantaneous mean $\boldsymbol{\delta}_{prs}$ and variance $\boldsymbol{\Lambda}_{prs}$ (an extension of Equation 4):

$$\boldsymbol{Y}(t_{pr,s})\,\big|\,\boldsymbol{Y}(t_{pr,s-1}) = \boldsymbol{y}(t_{pr,s-1}) \sim N_2\left(\boldsymbol{\delta}_{prs}, \boldsymbol{\Lambda}_{prs}\right),$$

where $\boldsymbol{\delta}_{prs} = \boldsymbol{\mu}_{pr} + e^{-\mathbf{B}_p(t_{pr,s}-t_{pr,s-1})}(\boldsymbol{y}(t_{pr,s-1}) - \boldsymbol{\mu}_{pr})$ and $\boldsymbol{\Lambda}_{prs} = \boldsymbol{\Gamma}_p - e^{-\mathbf{B}_p(t_{pr,s}-t_{pr,s-1})}\boldsymbol{\Gamma}_p e^{-\mathbf{B}_p'(t_{pr,s}-t_{pr,s-1})}$.

The first observation of each day comes from a person- and day-specific equilibrium distribution (corresponding to the defined equilibrium distribution of the process, see Equation 2), in the following manner:

$$\mathbf{Y}_{pr,1} \sim N_2(\boldsymbol{\mu}_{pr}, \boldsymbol{\Gamma}_p)\,.$$

Comparing the equations of the three-level model to the two-level model formulation, we can notice that each person has seven day-specific homebases. The time difference also has a day index $r$. The parameter $s$ runs from the second observation of day $r$ until the last observation of day $r$. The new notation ensures that the transition between the last measurement of day $r$ and the first measurement of day $r$+1 is not taken into account. The centralizing tendency and the volatility matrix parameters correspond to the 2-level model formulation, presented in the previous chapter.

The person-day homebases are considered as random samples from a person distribution, and the person-specific homebase provides the mean of this distribution. This person-specific homebase comes from a common population distribution. These sample distributions look as follows:

$$\boldsymbol{\mu}_{pr} \sim N_2\left(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_{\mu day}\right),$$
$$\boldsymbol{\mu}_p \sim N_2\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mu}\right).$$

We have allowed the homebase parameters to vary over days, but not the other parameters, for three reasons. First, as already mentioned above, we do not want to make assumptions about the process during the night. Second, it is plausible that the homebase differs systematically over days, following e.g. a weekly cycle (the most well-know phenomenon is called as “Blue Monday”, see further references in the paper of Ram et al., 2005). Third, making the other parameters also day-specific would increase the complexity of the model enormously with obvious consequences for ease and duration of computation on the one hand and interpretation on the other hand.

Two three-level hierarchical models and their application on a dataset will be presented. In a first model (from now on referred to as the empty model), the person-specific OU parameter sets are not regressed on covariates, only their population distribution is investigated. In a second model (from now on referred to as the full model), the OU parameters are regressed on these covariates. Such a model not only takes interindividual differences for granted but also tries to explain them (see also De Boeck & Wilson, 2004).

Because of the complexity of both the empty and full model and the large number of parameters in each case, a Bayesian approach is taken to estimate the parameters so that we can make use of Markov Chain Monte Carlo (MCMC) methods to explore the posterior distribution (Gelman, Carlin, Stern, & Rubin, 2003; Robert & Casella, 2004). The MCMC computations were carried out using WinBUGS (Spiegelhalter, Thomas, Best, & Lunn, 2003).

The choice of a Bayesian estimation procedure implies that we had to assign priors to the parameters of the models. Let us consider first the sampling distributions of the empty model. When the person-specific diffusion parameters were constrained to be positive, their logarithms were sampled from normal population distributions. The *z*-transformed correlation value came also from a normal population distribution. The means of these normal distributions were assigned a normal prior distribution with mean zero and precision 0.1 (i.e., a variance of 10). The inverted variances of the population distribution were given a gamma prior distribution with parameters 0.1 and 0.1. (Sensitivity had been checked by allowing normal priors with precision 0.01 and gamma priors with parameters 0.01 and 0.01, but that did not change the results.) The prior for the correlation of the log of the variances came from a uniform distribution between -1 and 1.

The person specific homebases were sampled from a bivariate normal population distribution. The prior distribution of the population homebase was a bivariate normal distribution with a zero mean vector and a precision matrix $0.001 \times I$ (where $I$ is the two-by-two identity matrix). The person specific homebase created the mean of the normal distribution of the person-day specific homebases. For the inverse of their covariance matrices $\mathbf{\Sigma}_{\mu}$ and $\mathbf{\Sigma}_{\mu day}$ non-informative Wishart-distributions were assigned. With regard to the full model, intercepts and regression coefficients come into play. For those parameters, normal priors have been used.

The results we will present below are based on 60000 draws from the posterior distribution, which come from 10000 iterations with six chains. Apart from the selected 10000 iterations, each chain started with a discarded burn-in period, which was chosen to be 10000 iterations in the case of the empty model, and 15000 iterations in the case of the full model (the convergence was somewhat slower in the full model). The initial values of the chains are randomly perturbed rational values derived from the data (e.g., the sample average for the homebases). The evaluation of the convergence is based on a visual assessment of the trace plots and on the values of the Gelman-Rubin (1992) $\sqrt{\hat{R}}$ diagnostic as modified by Brooks and Gelman (1998).

Table 1 shows a summary of the results for the empty model containing the posterior mean and standard deviation and the endpoints of the 95% posterior credibility interval. The estimated means of the homebase population distribution are (5.88, 5.20), corresponding to the findings of previous research (Russell, Weiss, & Mendelssohn, 1989), which show that on average the emotional state of persons is slightly pleasant and rather activated than deactivated. The correlation in the population between the homebases of the dimensions is estimated to be 0.28, which means that across persons, the homebases of the pleasantness and the activation dimensions are only moderately correlated. If we look at the standard deviations of the homebases ($\sigma_{\mu_1}, \sigma_{\mu_2}$) on the two dimensions, it appears that there is an approximately equal amount of variability. Comparing this interindividual standard deviations to the standard deviations of the day-person specific homebases ($\sigma_{\mu_1 day}, \sigma_{\mu_2 day}$), which represent an intraindividual change in the position of homebases during a week, we can see that the latter one is smaller (with 37% and 39%, respectively). With regard to the day specific part of the model, averaging the person-day specific homebases across days showed no remarkable daily variation.

The average log-variance of the activation-deactivation dimension in the stationary distribution is somewhat larger than the average log-variance of the pleasantness-unpleasantness dimension, although the 95% credibility intervals show considerable overlap. The correlation between the logarithms of the variances is relatively high (i.e., 0.68), hence it seems that there is a tendency for people with rather low/high variability in the pleasantness dimension to show correspondingly low/high variability in the activation dimension as well and vice versa. Considering the posterior mean of the *z*-transformed correlation in the stationary distribution, we see that the latter is rather small (i.e., 0.01) which suggests that on average, there is not much cross-correlation (i.e., on average, the bivariate person-specific stationary distributions have zero correlation). However, there is considerable variability in the estimated person-specific cross-correlation parameters: Although the average is 0.01, the 95% confidence interval varies between -0.34 and 0.35. It must be emphasized here that there is a conceptual difference between the population correlation of the homebase distribution and the average correlation of the stationary distribution and that they are unrelated with each other.

INSERT TABLE 1 ABOUT HERE

The mean of the population distribution of the logarithm of the centralizing tendency (i.e., $\omega_{\log(\beta)}$) is estimated to be -4.19. Based on that, the posterior mean of the population average centralizing tendency is 0.019. We may also look at the posterior estimates for the person-specific centralizing force (i.e., $e^{\log(\beta_p)}$). To illustrate graphically the interpretation of this average and the individual variation in centralizing tendency, we convert the posterior $\beta_p$-values to the corresponding autocorrelation functions. Figure 7 shows the person-specific autocorrelation function together with the autocorrelation function based on the average population value (thick line). For most participants, the autocorrelation between subsequent core affect positions separated 2 hours in time had already fallen below 0.2.

INSERT FIGURE 7 ABOUT HERE

In the full model, each of the six person-specific Ornstein-Uhlenbeck parameters (i.e., the homebases and the log-variances of the pleasantness-unpleasantness and the activation-deactivation dimensions, the *z*-transformed cross-correlation between the dimensions and the log-centralizing tendency) are regressed onto the Big Five personality dimensions. Table 2 summarizes posterior mean and standard deviations for the eight regression coefficients for which the 90% posterior credibility intervals do not contain zero. (We do not use the traditional 95% confidence interval because in this way we are able to present some rather fascinating associations.)

The results show that the locations of the individuals' homebases are related to neuroticism and extraversion: the homebase of the neurotic individuals is less pleasant and the extravert people tend to occupy homebase with a higher activation level. The intraindividual variabilities in the occupied positions in the core affect space on the two dimensions (i.e. parameters $\log(\gamma_{1p})$ and $\log(\gamma_{2p})$ ) show a relation with extraversion, agreeableness and neuroticism. Specifically, the results suggest that the extraverts show higher variability in the pleasantness dimension. Moreover, the variability of the trajectories is also related to the agreeableness in both dimensions: Less agreeable people tend to show higher variability with respect to pleasantness-unpleasantness and activation-deactivation as well. Additionally, people with a high neuroticism score show higher variability in the pleasantness-unpleasantness dimension. Hence, agreeableness and neuroticism have opposite effects on the variability of the pleasantness position.

Not only the trajectory variability, but also the centralizing tendency parameter is associated with neuroticism: Neurotic individuals in general have a smaller centralizing tendency, which implies that they are more likely to leave their homebase region and they are not eager to return quickly. In combination with the observation that neurotics have a larger variability on the pleasantness dimension, one might conclude that neurotic individuals do show less emotional stability. Finally, the positions in the core affect space visited by agreeable people are somewhat stronger correlated across dimensions. Thus, for agreeable persons, changes in one dimension tend to covary more strongly with changes in the other dimension in the same direction.

INSERT TABLE 2 ABOUT HERE

The empty as well as the full model handle the two repeatedly measured variables pleasantness and activation together. However, it brings forth the question, whether it is necessary to add cross-correlation parameters between dimensions. Moreover, every person has a unique value to express their dependence level between pleasantness and activation. We decided to test whether this person-specific parameter is really necessary in the model. Furthermore, it is also useful to investigate whether the three-level model formulation is appropriate; hence we need to compare the three-level model with a two-level one, in which the observations are nested just within persons. In addition, we can also examine whether the full model provided a better fit than the empty model.

To carry out the model selection, we opted for the Deviance Information Criterion (DIC) statistic (Spiegelhalter, Best, Carlin, & van der Linde, 2002). The DIC takes into account two important features of the model: the complexity (based on the number of the parameters) and the fit (typically measured by a deviance statistic). DIC examines the two features together and gives a measure which balances between the two. Its formula is the sum of the effective number of parameters and the posterior mean of the deviance (defined as -2 times loglikelihood). Theoretically, the model with smaller DIC would better predict a replicate dataset of the same structure. We compared two alternative models, namely a reduced three-level model and a two-level model to the three-level empty and full models. In the case of the reduced three-level model, all the person-specific cross-correlation parameters were set to zero. In the two-level model, the day level in the homebases was omitted, so that the person-specific OU parameter sets could not vary among the days. From the results it turns out that the three-level full model performs better than any of the alternative ones. The DIC value of this model is 25210, which is relatively close to the value of the three-level empty model, which is 25213. Nevertheless, both models do better than the two-level model (DIC = 25266) and the three-level reduced model (DIC = 25325), suggesting that the three-level model formulation and the simultaneous modeling of the two longitudinal variables with person-specific cross-correlation are good choices.

## Discussion

In this paper, we have introduced a model for the analysis of multivariate longitudinal profiles with continuous and possibly unbalanced measurement times. The core of our approach is a multivariate Ornstein-Uhlenbeck diffusion process, which serves as a model for the intraindividual behavior. The model is hierarchically extended to handle data from many persons drawn randomly from a population. Thus, it becomes a convenient framework to describe and explain interindividual differences in the characteristics of individual trajectories.

Our hierarchical diffusion modeling approach offers several distinctive characteristics. First of all, we are able to regress in principle all the diffusion parameters onto predictors (and not only mean parameters as is the standard in traditional multilevel or structural equation modeling). The diffusion parameters represent aspects such as intraindividual variability, autocorrelation in the positions, and cross-correlation, of which interindividual differences can both be described and explained with the help of covariate information. Secondly, the version of the OU process used admits the compelling feature of stationarity, allowing researchers to model not only subsequent measurements (as does the integrated multivariate OU process of Sy, Taylor, & Cumberland, 1997), but also the first measurement time. Additionally, the equlibrium distribution shows a general picture of the time-varying phenomena. However, in future applications it would be nice to compare the non-stationary variant in an empirical way with the stationary one. Thirdly, the model can be fit in the framework of differential equations, which offers an advantageous way to compare the OU model with other time-dependent models, like the damped oscillator model of Chow, Ram, Boker, Fujita, and Clore (2005) or the non-linear coupled oscillator model of Butner, Amazeen, and Mulvey (2005). The stochastic differential equation representation opens the possibility of casting many more interesting models in a hierarchical framework as presented here.

Additionally, the joint modeling of several time-varying psychological phenomena represents a great potential. Although this paper focused on the two-dimensional case, our general model formulation is multidimensional and the model can easily be implemented for the multidimensional case as well: the only difference resides in the decomposition of the covariance matrix, but for that problem a solution is available in the paper of Daniels and Pourahmadi (2002).

Finally, the Bayesian implemantation offers an easy way for estimating the parameters, moreover it makes it possible to use information from previous findings about the same topic to improve the parameter estimation.

We applied the model to a data set from a study that tracked the trajectories of persons in the core affect space. The hierarchical modeling approach pointed out some intriguing aspects of emotion change. With the OU model, we can regress the intraindividual variability parameter of the pleasantness-unpleasantness and activation-deactivation dimensions onto known personality dimension in order to examine this issue. It turned out, that for instance, agreeable people experience less change in their hedonic and arousal level. Besides the intraindividual changes, we were able to describe the degree of a centralizing force, which motivates the individual to readjust their current mood to their homeostatic homebase. Not only did we get a general picture about this centralizing tendency in the emotion experience, but we also learnt that the neurotic individuals tend to have a lower level of it. Taking into account the relationship between the longitudinal variables of pleasantness-unpleasantness and activation-deactivation also proved useful in the current application. Besides the improving model fit, we could observe that agreeable people show more consistency in the changes of the hedonic and arousal values.

As a drawback to the current application, we need to mention the discreteness of the core affect grid. Since the offered change process is continuous, a more precise measure of the occupied position in the core affect space would be desirable. Also just representing the core affect space without the grid (and putting a very subtle grid afterwards to code the obtained self-reports) could be a nice improvement for decreasing the measurement error. However, for such data it may also be possible to assume a continuous-time discrete state space Markov chain as a model for the intraindividual data. This type of model might investigated in future research.

## References


Arnold, L. (1974). *Stochastic differential equations: Theory and applications*. New York: John Wiley & Sons.

Barnard, J., McCulloch, R. E., & Meng, X. L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with applications to shrinkage. *Statistica Sinica*, *10*, 1281– 1311.

Beckett, L. A., Tancredi D. J., & Wilson R. S. (2004). Multivariate longitudinal models for complex change processes. *Statistics in Medicine*, *23*, 231-239.

Blackwell, P. G. (1997). Random diffusion models for animal movement. *Ecological Modelling*, *100*, 87-102.

Blackwell, P. G. (2003). Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*, *90*, 613-627.

Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*, 334-353.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579-616.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach.* New Jersey, NJ: John Wiley & Sons.

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 Research in Personality*, *32*, 202-221.

Brillinger, D. R., Preisler, H. K., Ager, A. A., & Kie, J. G. (2004). An exploratory data analysis (EDA) of the paths of moving animals. *Journal of Statistical Planning and Inference*, *122*, 43-63.

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd. ed.). New York: Springer-Verlag.

Brooks, S. P., & Gelman, A. (1998). Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434-455.

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

Cattell, R. B. (1963). The structuring of change by P-technique and incremental R-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison: University of Wisconsin Press.

Chow, S-M., Ram, N., Boker, S. M., Fujita, F., & Clore, G. (2005). Emotion as a thermostat: Representing emotion regulation using a damped oscillator model. *Emotion*, *5*, 208-225.

Cox, D. R., & Miller, H. D. (1972). *The Theory of Stochastic Processes*. London: Chapman and Hall.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method*.* *The Journal of Nervous and Mental Disease*, *175*, 526-536.

Daniels, M., & Pourahmadi, M. (2002). Bayesian analysis of covariance matrices and dynamic models for longitudinal data. *Biometrika*, *89*, 553-566.

De Boeck, P., & Wilson, M. (Eds.) (2004). *Explanatory item response models. A generalized linear and nonlinear approach*. New York: Springer.

De la Cruz-Mesía, R., & Marshall, G. (2006). Non-linear random effects models with continuous time autoregressive errors: a Bayesian approach. *Statistics in Medicine*, *25*, 1471-1784.

Diederich, A. (1997). Dynamic stochastic models for decision making under time pressure. *Journal of Mathematical Psychology*, *41*, 260-274.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time and simple response time. *Journal of Mathematical Psychology*, *47*, 304 - 322.

Dubin, J. A., & Müller, H. G. (2005). Dynamical correlation for multivariate longitudinal data. *Journal of the American Statistical Association*, *100*, 872-881.

Dunn, J. E., & Gipson, P. S. (1977). Analysis of radio telemetry data in studies of home range. *Biometrics*, *33*, 85-101.

Fang, H. B., Tian, G. L., Xiong, X., & Tan, M. (2006). A multivariate random-effects model with restricted parameters: Application to assessing radiation therapy for brain tumours. *Statistics in Medicine*, *25(11)*, 1948-59

Ferrer, E., & Nesselroade, J. R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion*, *3*, 344-360.

Fieuws, S., & Verbeke, G. (2004). Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach. *Statistics in Medicine*, *23*, 3093–3104.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2003). *Bayesian data analysis* (2nd ed.). New York: Chapman & Hall.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457-472.

Goldstein, H. (2003). *Multilevel Statistical Models.* London: Arnold.

Hoekstra, H. A., Ormel, J., & De Fruyt, F. (1996). *NEO PI-R, NEO FFI Big Five Persoonlijkheidsvragenlijsten : Handleiding* [*NEO PI-R, NEO FFI Big Five Personality Questionnaire: Manual*] . Lisse, The Netherlands: Swets & Zeitlinger B.V.

Karlin, S., & Taylor, H. (1981). *A Second Course in Stochastic Processes*. New York: Academic Press.

Kuppens, P., Van Mechelen, I., Nezlek, J. B., Dossche, D., & Timmermans, T. (2006). Individual differences in core affect variability and their relationship to personality and adjustment. Manuscript submitted for publication.

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, *11*, 129-141.

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

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed.). New York: John Wiley.

Lykken, D. T., & Tellegen, A. (1996). Happiness is a stochastic phenomenon. *Psychological Science*, *7*, 186-189.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215-253.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*, 181-202.

O'Brien, L. M., & Fitzmaurice, G. M. (2005). Regression models for the analysis of longitudinal Gaussian data from multiple sources. *Statistics in Medicine*, *24*, 1725-1744.

Ram, N., Chow, S.-M., Bowles, R. P., Wang, L., Grimm, K., Fujita, F., & Nesselroade, J. R. (2005). Examining interindividual differences in cyclicity of pleasant and unpleasant affects using spectral analysis and item response modeling, *Psychometrika*, *70,* 773-790.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Ratcliff, R. (2002). A diffusion model account of reaction time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin and Review*, *9*, 278-291.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating the parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin and Review*, *9*, 438-481.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods* (2nd ed.). New York: Springer.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145-172.

Russell, J. A., & Feldman-Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819.

Russell, J. A., Weiss, A., & Mendelssohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*, 493–502.

Schach, S. (1971). Weak convergence results for a class of multivariate Markov processes. *The annals of mathematical statistics*, *42*, 451-465.

Smith, P. L. (2000). Stochastic, dynamic models of response times and accuracy: A foundational primer", *Journal of Mathematical Psychology*, *44*, 408-463.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*, *Series B*, *6*, 583-640.

Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. (2003). *WinBUGS Version 1.4 Users Manual*. Cambridge: MRC Biostatistics Unit. Retrieved from: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf

Sy, J. P., Taylor, J. M. G., & Cumberland, W. G. (1997). A stochastic model for the analysis of bivariate longitudinal AIDS data. *Biometrics*, *53*, 542-555.

Taylor, J. M. G., Cumberland, W. G., & Sy, J. P. (1994). A stochastic model for analysis of longitudinal AIDS data. *Journal of the American Statistical Association*, *89*, 727-736.

Thiébaut, R., Jacqmin-Gadda, H., Babiker, A., & Commenges, D. (2005). Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of CD4+ cell count and HIV RNA viral load in response to treatment of HIV infection. *Statistics in Medicine*, *24*, 65-82.

Thum, Y. M. (1997). Hierarchical linear models for multivariate outcomes. *Journal of Educational and Behavioral Statistics*, *22*, 77-108.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

Weiss, R. E. (2005). *Modeling longitudinal data*. New York: Springer.

Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of Dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin*, *116*, 166-186.

Appendix A : Derivation of the autocorrelation function for a multivariate OU process

For simplicity and without loss of generality let us assume that $\boldsymbol{\mu}$ is a $q\times 1$ vector of zeros. Then the covariance matrix function of the OU process with a general $\mathbf{B}$ and $\boldsymbol{\Gamma}$ can be derived as follows (see also Schach, 1971):

$$
\begin{aligned}
E\left[\boldsymbol{Y}(t)\boldsymbol{Y}'(t+d)\right] &= E\left[E\left(\boldsymbol{Y}(t)\boldsymbol{Y}'(t+d)\middle|\boldsymbol{Y}(t)\right)\right] \\
&= E\left[\boldsymbol{Y}(t)E\left(\boldsymbol{Y}'(t+d)\middle|\boldsymbol{Y}(t)\right)\right] \\
&= E\left[\boldsymbol{Y}(t)\boldsymbol{Y}'(t)e^{-\mathbf{B}'d}\right] \\
&= E\left[\boldsymbol{Y}(t)\boldsymbol{Y}'(t)\right]e^{-\mathbf{B}'d} \\
&= \boldsymbol{\Gamma}e^{-\mathbf{B}'d}.
\end{aligned}
$$

If $\mathbf{B}$ equals $\beta\boldsymbol{I}$ then the covariance matrix function becomes $\boldsymbol{\Gamma}e^{-\beta d}$. Because $\mathrm{var}\left[\boldsymbol{Y}(t)\right] = \mathrm{var}\left[\boldsymbol{Y}(t+d)\right] = \boldsymbol{\Gamma}$, the autocorrelation function $\rho(d)$ is equal to $e^{-\beta d}$ in the isotropic case.

Footnotes

[1] The matrix exponential $e^M$ is defined as follows:

$$e^M = I + \sum_{j=1}^{\infty} \frac{M^j}{j!}.$$

If $M$ is a diagonal matrix with elements $m_1,\ldots,m_q$, it is equal to;

$$e^M = \begin{pmatrix} e^{m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{m_q} \end{pmatrix},$$

where $e^{m_i}$ is the scalar exponential function value of $m_i$.

[2] The nature of the currently used isotropic $\mathbf{B}$ matrices implies that the overall behavior of the process around the homebase can always be depicted with straight lines tending towards the concentration point. If we allowed for a wider class of $\mathbf{B}$ matrices with different values in the dimensions, we could achieve different curvatures around the homebase, which would provide different interpretations of the relationship among the dimensions by giving a special meaning to the used co-ordinate system. On the one hand taking into account this possibility could improve the quality of the model, on the other hand it would significantly increase its complexity, as well as its implementation and the interpretation of its parameters. The implementation is more complicated by the fact that not just $\mathbf{B}$ but the $\mathbf{B\Gamma}+\mathbf{\Gamma B}'$ matrix must be also always positive definite, see more in Dunn and Gipson (1977). For more details on the effect of the isotropic constraint see Blackwell (1997)

Table 1

*Summary measures of the posterior distributions for the most important parameters*

| Parameters | Description | Posterior mean | 95% posterior credibility interval | | Posterior SD |
|---|---|---|---|---|---|
| | Pleasantness | | | | |
| $\mu_1$ | Average homebase | 5.88 | 5.68 | 6.08 | 0.10 |
| $\sigma_{\mu_1}$ | SD of the average homebase | 0.68 | 0.54 | 0.85 | 0.08 |
| $\sigma_{\mu_1 day}$ | Daily SD of the average homeb. | 0.43 | 0.32 | 0.56 | 0.06 |
| $\omega_{\log(\gamma_1)}$ | Average log-variability | 0.94 | 0.79 | 1.09 | 0.08 |
| | Activation | | | | |
| $\mu_2$ | Average homebase | 5.20 | 5.02 | 5.38 | 0.09 |
| $\sigma_{\mu_2}$ | SD of the average homebase | 0.59 | 0.46 | 0.75 | 0.07 |
| $\sigma_{\mu_2 day}$ | Daily SD of the average homeb. | 0.36 | 0.27 | 0.47 | 0.05 |
| $\omega_{\log(\gamma_2)}$ | Average log-variability | 1.18 | 1.03 | 1.32 | 0.07 |
| $\rho_{\mu_1\mu_2}$ | Correlation between homebases | 0.28 | -0.05 | 0.57 | 0.16 |
| $\rho_{\log(\gamma_1),\log(\gamma_2)}$ | Correlation between log-variances | 0.68 | 0.43 | 0.88 | 0.12 |
| $F(\rho)$ | Average *z*-transf. cross-correlation | 0.01 | -0.06 | 0.08 | 0.04 |
| $\omega_{\log(\beta)}$ | Average log-centralizing tendency | -4.19 | -4.39 | -3.97 | 0.11 |

Table 2

*Summary of the regression coefficients with a 90% posterior credibility not containing 0*

| Parameters | Description | Covariates | Posterior mean | 90% posterior credibility interval | | Posterior SD |
|---|---|---|---|---|---|---|
| | Pleasantness | | | | | |
| $\mu_{1p}$ | Homebase | neuroticism | -0.28 | -0.56 | -0.01 | 0.17 |
| $\omega_{\log(\gamma_{1p})}$ | Variability | neuroticism | 0.24 | 0.03 | 0.46 | 0.13 |
| $\omega_{\log(\gamma_{1p})}$ | Variability | extraversion | 0.33 | 0.02 | 0.64 | 0.19 |
| $\omega_{\log(\gamma_{1p})}$ | Variability | agreeableness | -0.45 | -0.73 | -0.17 | 0.17 |
| | Activation | | | | | |
| $\mu_{2p}$ | Homebase | extraversion | 0.36 | 0.02 | 0.69 | 0.20 |
| $\omega_{\log(\gamma_{2p})}$ | Variability | agreeableness | -0.38 | -0.65 | -0.11 | 0.17 |
| $\omega_{\log(\beta_p)}$ | Centralizing tendency | neuroticism | -0.34 | -0.65 | -0.03 | 0.19 |
| $F(\rho_p)$ | Cross-correlation | agreeableness | 0.21 | 0.07 | 0.35 | 0.08 |

Figure Caption

*Figure 1*. The 0.5 probability contour curves of conditional distributions with three different values for $\beta$ (0.0005; 0.005; 0.03) and with $\mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, $\boldsymbol{\Gamma} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $\boldsymbol{y}(t-d) = \begin{pmatrix} 7 \\ 7 \end{pmatrix}$ and $d$ = 100.

*Figure 2*. Four different $\beta$ values and corresponding autocorrelation functions.

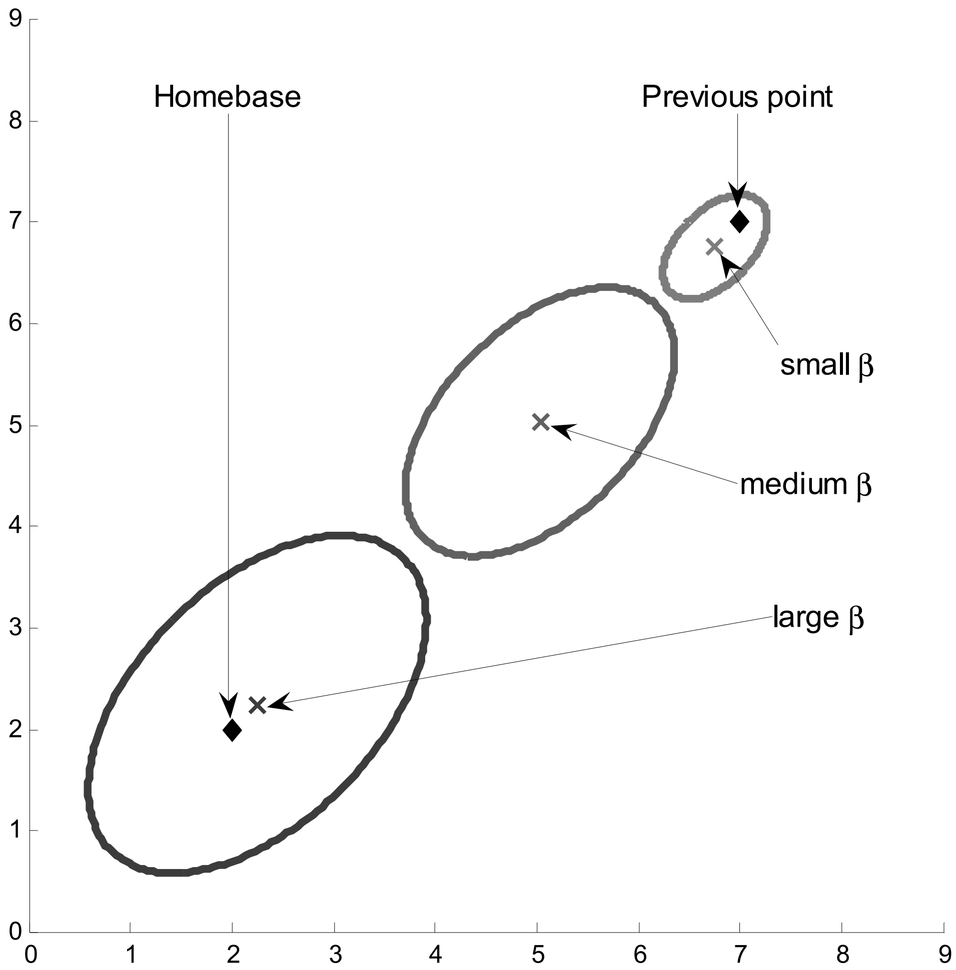
*Figure 3*. The Affect Grid used in the application.

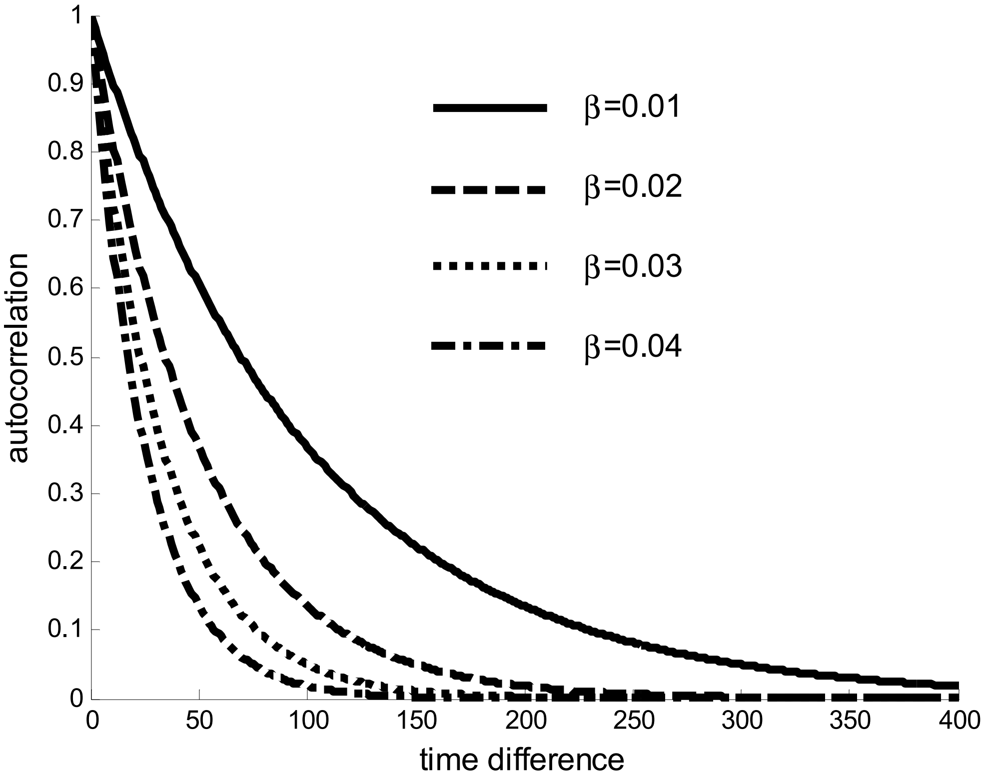
*Figure 4*. Person profiles in the core affect grid.

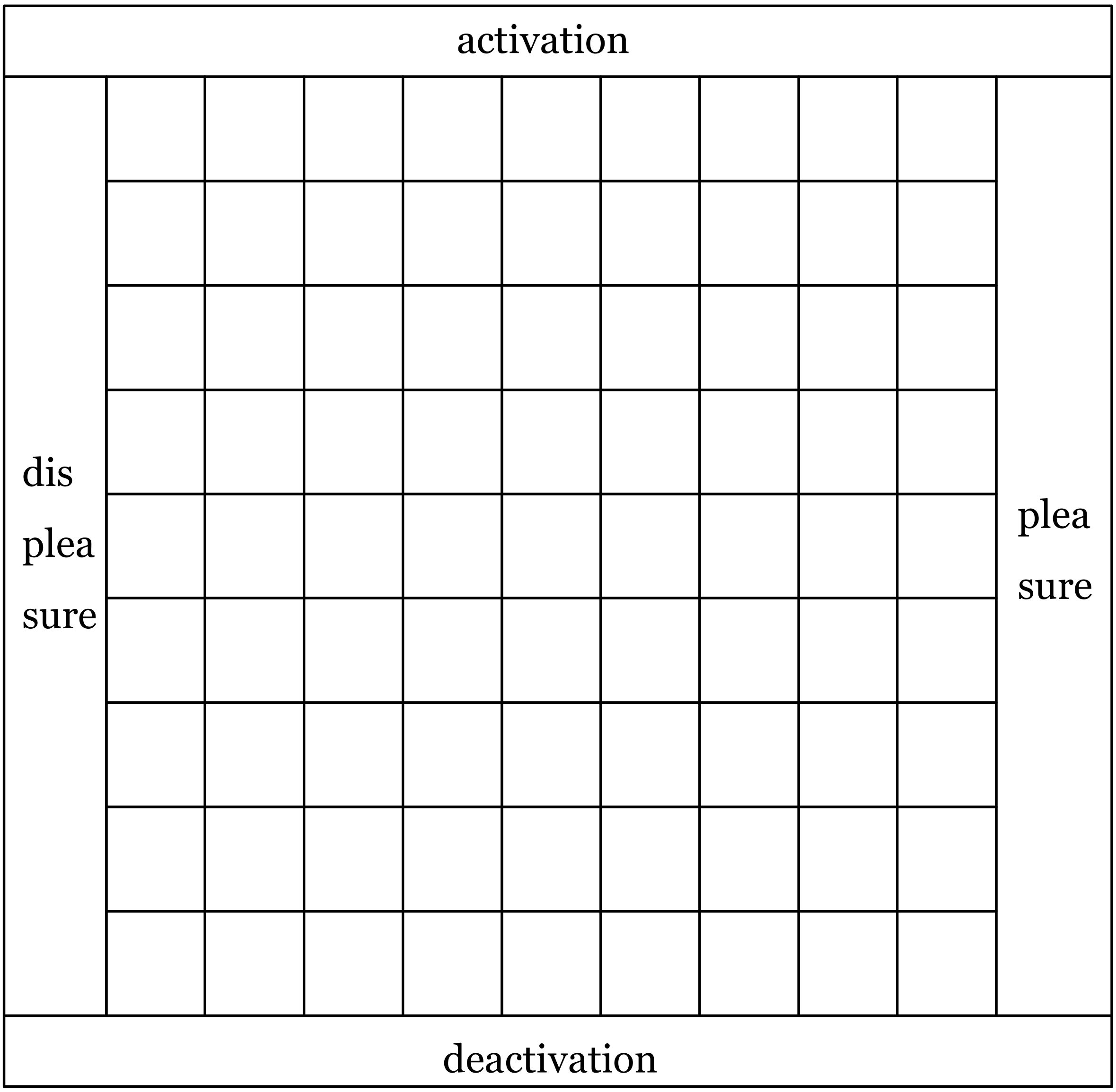
*Figure 5*. Smoothed heat map of the visit frequencies in the core affect space.

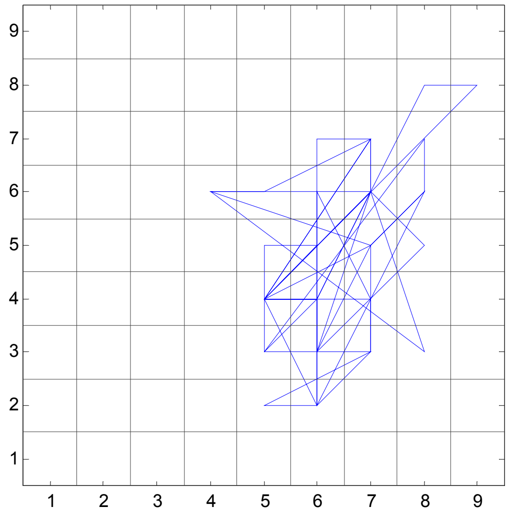
*Figure 6*. Estimated vector field of the core affect grid.

*Figure 7*. Change in the autocorrelation function according to the estimated $\beta$ parameters of the individuals.

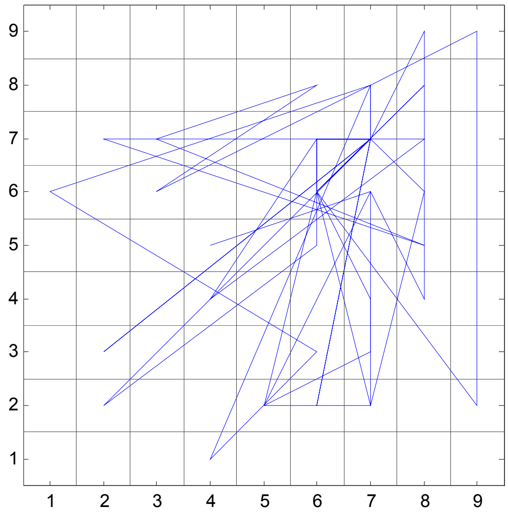

Homebase
Previous point
small β
medium β
large β
0
1
2
3
4
5
6
7
8
9


β=0.01
β=0.02
β=0.03
β=0.04
autocorrelation
time difference
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
250
300
350
400

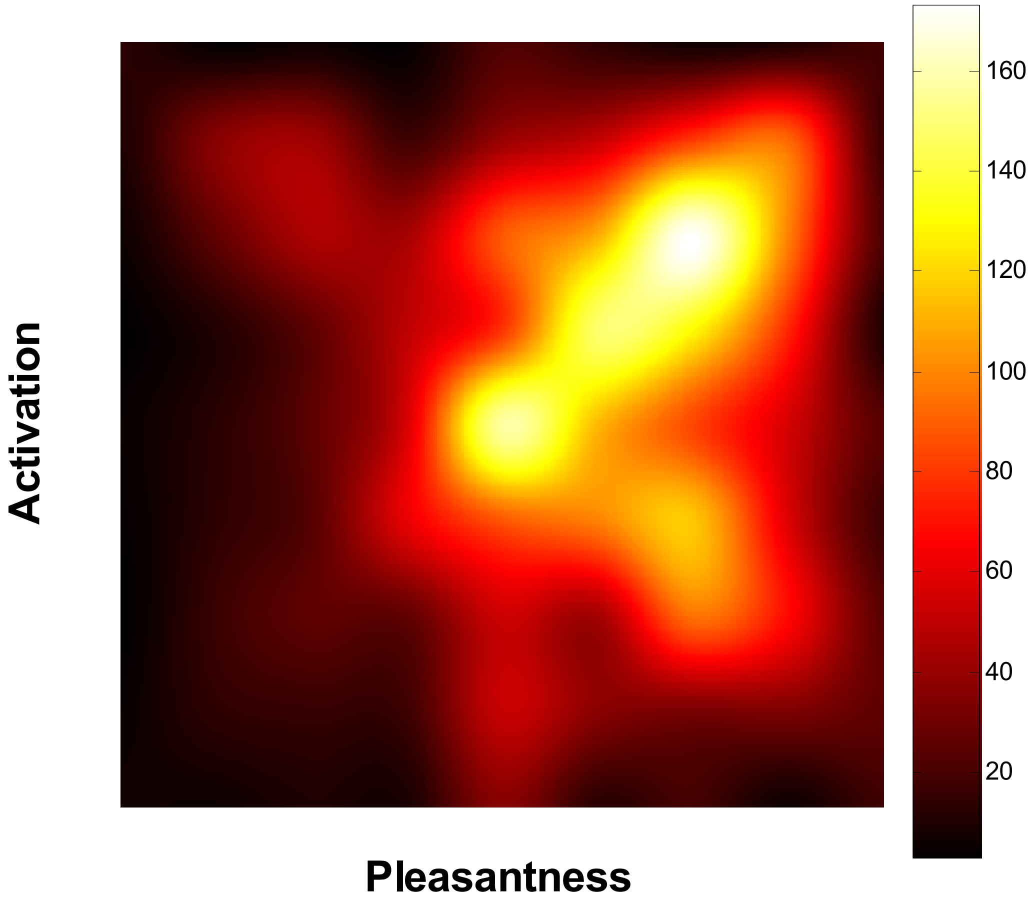

activation
dis
plea
sure
plea
sure
deactivation

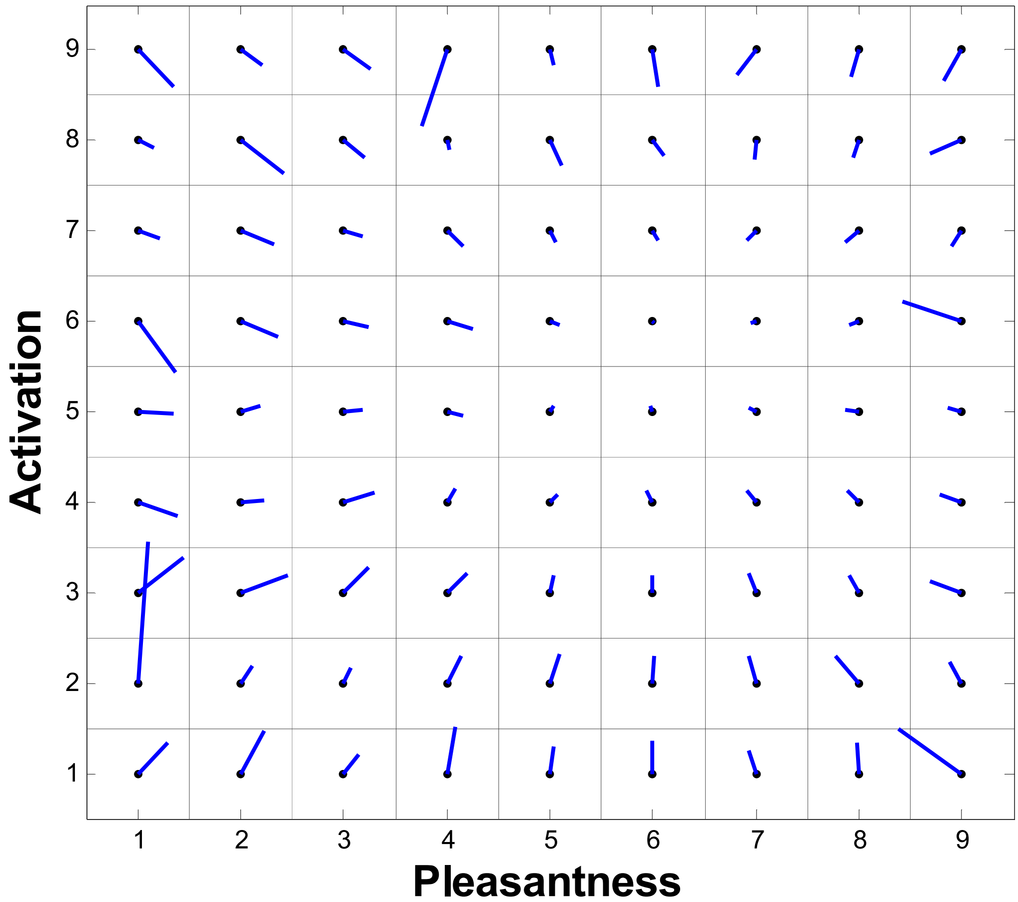


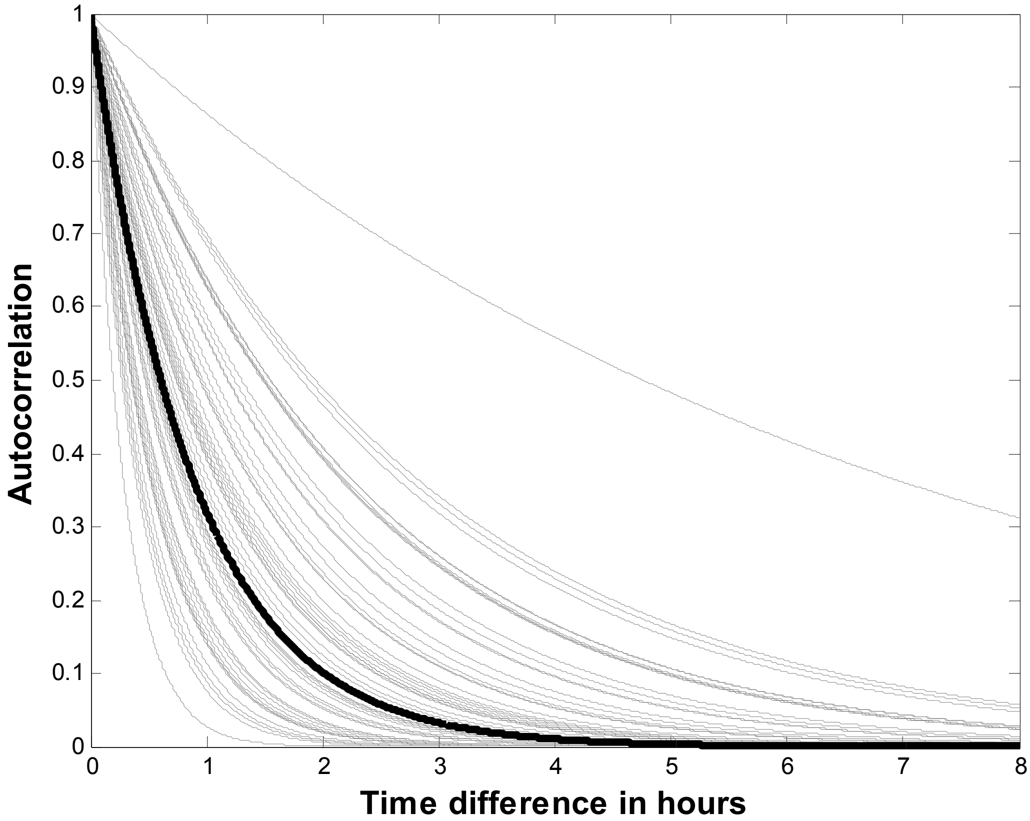

**Participant nr.7 (61 self-reports)**




**Participant nr.4 (61 self-reports)**


160
140
120
100
80
60
40
20
Activation
Pleasantness


Activation
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
Pleasantness


Autocorrelation
Time difference in hours
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8