A multivariate statistical model for emotion dynamics

Tanja Krone[1], Casper J. Albers[1], Peter Kuppens[2] and Marieke E. Timmerman[1]

[1]University of Groningen

[2]KU Leuven-University of Leuven

Accepted for publication by EMOTION on 11 September 2017

© 2017, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission. The final article will be available, upon publication, via its DOI: 10.1037/emo0000384

Author Note

Corresponding author:

Tanja Krone,

Grote Kruisstraat 2/1,

9712TS Groningen, The Netherlands.

email: tanja.krone@gmail.com

This work is funded by the Netherlands Organization for Scientific Research (NWO), grant number 406-11-018.

Abstract

In emotion dynamic research one distinguishes various elementary emotion dynamic features, which are studied using intensive longitudinal data. Typically, each emotion dynamic feature is quantified separately, which hampers the study of relationships between various features. Further, the length of the observed time series in emotion research is limited, and often suffers from a high percentage of missing values. In this paper we propose a vector autoregressive Bayesian dynamic model, that is useful for emotion dynamic research. The model encompasses six elementary properties of emotions, and can be applied with relatively short time series, including missing data. The individual elementary properties covered are: within person variability, innovation variability, inertia, granularity, cross-lag regression and average intensity. The model can be applied to both univariate and multivariate time series, allowing to model the relationships between emotions. One may include external variables and non-Gaussian observed data. We illustrate the usefulness of the model on data involving 50 participants self-reporting on their experience of three emotions across the period of one week using experience sampling.

Keywords: Bayesian, vector autoregressive, experience sampling, longitudinal data

A multivariate statistical model for emotion dynamics

Emotions are an important part of our daily lives. The importance of emotions for our health and well-being is recognized more and more (Tugade, Fredrickson, & Feldman Barrett, 2004; Grühn, Lumley, Diehl, & Labouvie-Vief, 2013; Lewis, Haviland-Jones, & Feldman Barret, 2008, Ch. 29). Unlike, perhaps, personality and values, emotions fluctuate across time, changing both within and between days under the influence of external events and internal regulation. As such, a central function of emotions is to alert us to important events and changes, and to motivate us to deal with them (Larsen, 2000; Frijda, 2007; Scherer, 2009; Kuppens & Verduyn, 2015). Understanding the dynamics of emotions is therefore important, not in the least because it provides a window on how emotions may become dysregulated, which is considered a central feature of several mental disorders (Houben, Van Den Noortgate, & Kuppens, 2015; Wichers, Wigman, & Myin-Germeys, 2015).

To study emotion dynamics, intensive longitudinal data are used, sampled sufficiently frequently to characterize the dynamics of interest (Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015). Technological advantages facilitate the collection of such data, both in an experimental setting in a lab and in daily life. In lab studies, for instance, video mediated recall and physiological recording can provide information on the dynamics of emotional episodes. In daily life, the widespread availability of mobile devices, first palmtops and now smartphones, enables researchers to collect multiple measurements per day in so-called ecological momentary assessment (EMA, also known as experience sampling) studies (Larson & Csikszentmihalyi, 1983; Shiffman, Stone, & Hufford, 2008; Bolger & Laurenceau, 2013; Bos, Schoevers, & Aan het Rot, 2015).

When intensive longitudinal data is gathered with a structure as complex as found in emotion data, the choice of a proper analysis is of paramount importance. This choice is far from straightforward, given the diversity of techniques available (Hamaker et al., 2015). The analysis typically focuses on identifying particular elementary features of emotion dynamics, with the aim to reveal distinct information on affective functioning and regulation (Kuppens & Verduyn, 2015). For instance, one may be interested in the level of variability emotions display within an individual, or in how different emotions covary across time. However, choosing a proper analysis is hampered by the fact that these elementary features can often be quantified in different ways. Further, the quantifications of these elementary features are typically considered separately. This implies that relationships between these features remain hidden.

To provide a good picture of emotion dynamics, we propose to use a single model of which most parameters have a clear interpretation in terms of a number of key features that are considered central to emotion dynamics. To this end, we propose to use Bayesian dynamic modelling (West & Harrison, 1997). The Bayesian Dynamic Model (BDM) we propose offers a representation of multivariate time series, and may be applied to multiple individuals. Furthermore, the BDM as proposed in this paper can be conveniently interpreted in terms of six important emotion dynamic features. This offers insight into the dynamics of single emotions, as well as the dynamics between multiple emotions within an individual. By applying the model to data from multiple individuals, one can achieve insight into interindividual differences in emotion dynamics. Using Bayesian estimation offers flexibility with regard to the distributions used in specifying the model.

Emotion Dynamic Features

The patterns and regularities of an individual's experienced emotions across time can be captured by various elementary properties. We denote these elementary properties as emotion dynamic features (EDFs). There is a vast range of EDFs discussed in the literature (Houben et al., 2015; Kuppens & Verduyn, 2015; Grühn et al., 2013; Carstensen, Pasupathi, Mayr, & Nesselroade, 2000; Brose, de Roover, Ceulemans, & Kuppens, 2015). The taxonomy as discussed by Kuppens and Verduyn (2015) organizes these EDFs into four categories: emotional variability, emotional inertia, emotional cross-lag, and emotional granularity. If we complement these four with the average emotional intensity, we provide a fairly complete picture of an individual's experience of emotions across time. Our aim is to propose a way to succinctly capture the EDFs from the five categories in a single model. In the following section, we discuss each category and how it is captured in our model.

Emotional variability. Emotional variability reflects to what extent the intensity of an emotion as experienced by an individual, varies across time. Emotional variability has been found to increase with increasing stress levels (Scott, Sliwinski, Mogle, & Almeida, 2014) and decrease with increasing age (Carstensen et al., 2000; Scott et al., 2014; Brose et al., 2015). High emotional variability has been linked with lower emotional well-being and higher prevalence and severity of mood disorders (Houben et al., 2015). To quantify emotional variability, the within person variance or standard deviation is typically used (Carstensen et al., 2000; Röcke, Li, & Smith, 2009; Grühn et al., 2013; Scott et al., 2014; Kuppens & Verduyn, 2015).

The within person variance can be seen as a global summary of the degree of emotional variability. This variability can be decomposed into various elements (Jahng, Wood, & Trull, 2008). In this context, it is useful to distinguish the predictable part from the

random part. The predictable part can be interpreted in terms of the emotional inertia and cross-lag, as will be discussed in the next paragraphs. The random part covers the instantaneous change, and consists of the innovation variance and the white noise variance.

The innovation variance and the white noise variance express the sizes of the instantaneous changes at each measurement point. The key difference between the two is that the innovation variance captures the part of the change that is carried through to the next measurement point, while the white noise variance captures the part of the change that is not carried through to the next measurement point. Thus, a person showing a high white noise variance and low innovation variance is characterized by large variability at successive measurement points. Vice versa, a low noise variance with high innovation variance also indicates a large variability across the whole time span observed, but at a much slower rate, yielding much fewer oscillations in scores at successive measurement points.

As the global summary measure of within person variability, we will use the within person variance. Though our model includes both the innovation variance and white noise variance, we will only interpret the innovation variance, as a measure of innovation variability. We leave aside the white noise variance in our interpretation, because change due to white noise is typically attributed to measurement error.

Emotional inertia. Emotional inertia refers to the tendency of an emotion to carry over from one moment to the next, reflecting resistance to change (Cook et al., 1995; Suls, Green, & Hills, 1998; Kuppens, Allen, & Sheeber, 2010). High inertia has been linked to impaired emotion regulation (Kuppens, Allen, & Sheeber, 2010; Suls et al., 1998; Koval et al., 2015; Gross, 2015), inflexibility in adapting emotions (Kashdan & Rottenberg, 2010) and rumination (Koval, Kuppens, Allen, & Sheeber, 2012). Emotional inertia is generally

quantified as the autoregression between successive measurements of an emotion. Note that although the concept of inertia is linked to the autocorrelation, it is generally quantified as a (autoregressive) regression variable. Therefore, autoregression is the adequate name for this variable, although it has also been referred to as autocorrelation in literature (e.g., Kuppens, Allen, & Sheeber, 2010; Kuppens & Verduyn, 2015).

Emotional cross-lag. Emotions can be regulated through feedback-loops: the increase of one emotion may infer an increase or decrease in another emotion (Gross, 2015; Kuppens & Verduyn, 2015; Pe & Kuppens, 2012). Although few studies have yet been conducted on emotional cross-lag, it is an important part of emotion regulation (Gross, 2015; Kuppens & Verduyn, 2015). For example, it has been found to be increased in major depression patients in terms of higher levels of overall emotion network density (Pe et al., 2015). Emotional cross-lag is quantified via the cross-lag regression; the lagged regression between two emotions (Pe & Kuppens, 2012). Analogously to the term autoregression, it is sometimes called the cross-lag correlation (e.g., Kuppens and Verduyn, 2015), while the cross-lag regression is generally used to quantify the emotional cross-lag. When the cross-lag regression is positive, this is called augmentation: the experience of one emotion increases the strength of another emotion on a later time point. A negative cross-lag regression is called blunting: the experience of one emotion decreases the strength of another emotion on a later time point (Pe & Kuppens, 2012).

Emotional granularity. Emotional granularity refers to the ability of differentiating between different emotions and identifying emotions with specificity and precision. This is also known as emotional differentiation and is often measured in terms of emotional covariation (Feldman, 1995; Barrett, Gross, Christensen, & Benvenuto, 2001; Barrett & Gross, 2001; Kuppens & Verduyn, 2015). Higher emotional granularity is linked to increased

emotion regulation (Barrett et al., 2001) and different, more effective coping mechanisms (Tugade et al., 2004). Furthermore, higher emotional granularity is associated with lower levels of neuroticism (Carstensen et al., 2000), and lower incidence of social anxiety disorder (Kashdan & Farmer, 2014) and depression (Erbas, Ceulemans, Pe, Koval, & Kuppens, 2014).

Emotional granularity has been quantified in different ways. For example, the differentiation index equals the number of components found by a principal component analysis (PCA) on the covariances between emotions of a single individual (Grühn et al., 2013; Brose et al., 2015). Related to this is the concept of the unshared variance: the percentage of variance unexplained by the first component of such a PCA (Grühn et al., 2013). In practice, the choice between the differentiation index and the unshared variance is a pragmatic one. For a large number of emotions, the differentiation index appears to be more informative, and for a small number of emotions, the unshared variance. For these measures, higher scores indicate a higher granularity.

Other quantifications can be calculated directly from the observed data: the covariance between two emotions within a person (Grühn et al., 2013; Erbas et al., 2014), the correlation between two emotions (Barrett et al., 2001), and the intraclass correlation (ICC) between all emotions (Tugade et al., 2004; Erbas et al., 2014). A higher covariance, correlation and ICC indicate a lower level of differentiation between emotions, and thus a lower granularity. The correlation has the advantage of being standardized, allowing for a direct comparison between pairs of emotions both within and between individuals. However, a low within person variance reduces the size of the absolute correlation, which renders interpretation difficult. This issue is not encountered when using the covariance (Scott et al., 2014). As all named quantifications measure the covariation between

emotions, we do not need to include them all. In our model, the granularity will be quantified via the correlation.

Average emotional intensity. The EDFs discussed thus far capture the dynamics of emotions over time. In addition, how strong an emotion is felt on average may also differ, both between emotions within an individual, and between individuals, and can provide important information on people's emotional lives. The average emotional intensity for positive emotions is positively related to emotion regulation (Barrett et al., 2001), as well as with extraversion, agreeableness and conscientiousness, but negatively related with neuroticism (Carstensen et al., 2000). The average emotional intensity of negative emotions is higher for individuals with social anxiety disorder (Kashdan & Farmer, 2014) as well as for individuals with depression, high scores on neuroticism, and low scores on self-esteem (Erbas et al., 2014). To assess the average emotional intensity across time, we take into account the average intensity (Carstensen et al., 2000; Barrett et al., 2001), quantified as the mean score over time (Kashdan & Farmer, 2014; Erbas et al., 2014).

This paper. Each of these features provides unique information on how emotions (co)vary, carry over from one moment to the next, or mutually influence each other, and together they provide insight into many crucial aspects of emotional functioning and flexibility. As such, we propose a BDM that captures within person variability, innovation variability, inertia, cross-lag, granularity, and average intensity for multiple emotions and individuals in a single model. First, we will introduce the model and its possibilities. Then, we will present an empirical application of the model. We will conclude with a discussion on the model, its advantages and disadvantages, and recommendations for future research.

Model

To combine the aforementioned concepts for multiple variables, we will use a Bayesian interpretation of a State Space Model, called the Bayesian Dynamic Model (BDM) (West & Harrison, 1997). We will use the BDM to estimate a vector autoregressive model, which can be rewritten into a State Space Model (Harvey, 1990; Durbin & Koopman, 2012).

The BDM has two equations: the observation equation and the system equation. For univariate data, the inclusion of so-called white noise in the observation equation improves estimation of the autoregression (Schuurman, Houtveen, & Hamaker, 2015). Following this, we include white noise in our multivariate BDM as well.

Single individual. For each individual, a separate model can be formulated. We model $y_{\mathrm{i,t,n}}$, the score on emotion i (i = 1,2,...,I), at time point t (t = 1,2,...,T_n), for individual n (n = 1,2,...,N). The first equation, the observation equation, links the observed score $Y_{\mathrm{i,t,n}}$ to the latent variable $\boldsymbol{\theta}_{\mathrm{i,t,n}}$. The observation equation for the score vector $\mathbf{y}_{\mathrm{t,n}} = \left[y_{1,\mathrm{t,n}}, y_{2,\mathrm{t,n}}, \ldots, y_{\mathrm{I,t,n}}\right]'$ is as follows:

$$\mathbf{y}_{\mathrm{t,n}} = \boldsymbol{\mu}_{\mathrm{n}} + \boldsymbol{\theta}_{\mathrm{t,n}} + \boldsymbol{\varepsilon}_{\mathrm{t,n}}, \qquad \boldsymbol{\varepsilon}_{\mathrm{t,n}} \sim N(\mathbf{0}, \mathbf{H}_{\mathrm{n}}), \tag{1}$$

where $\boldsymbol{\mu}_{\mathrm{n}}$ $(I \times 1)$ denotes the mean vector of the *I* emotions, $\boldsymbol{\theta}_{\mathrm{t,n}}$ $(I \times 1)$ the latent variable vector, $\boldsymbol{\varepsilon}_{\mathrm{t,n}}$ $(I \times 1)$ the white noise vector and $\mathbf{H}_{\mathrm{n}}$ the covariance matrix of $\boldsymbol{\varepsilon}_{\mathrm{t,n}}$. As $\boldsymbol{\varepsilon}_{\mathrm{i,t,n}}$ is assumed to be independent across emotions, $\mathbf{H}_{\mathrm{n}}$ is a $I \times I$ diagonal matrix with $\sigma^2_{\varepsilon_{\mathrm{i,n}}}$ as diagonal elements.

The system equation models the autoregression and cross-lag regression and the innovation over time of the latent variable $\boldsymbol{\theta}_{\mathrm{t,n}}$:

$$\boldsymbol{\theta}_{\mathrm{t,n}} = \boldsymbol{\Phi}_{\mathrm{n}} \times \boldsymbol{\theta}_{\mathrm{t-1,n}} + \boldsymbol{\eta}_{\mathrm{t,n}}, \qquad \boldsymbol{\eta}_{\mathrm{t,n}} \sim N(\mathbf{0}, \mathbf{Q}_{\mathrm{n}}), \tag{2}$$

where $\boldsymbol{\Phi}_{\mathrm{n}}$ $(I \times I)$ is the autoregression and cross-lag regression matrix, $\boldsymbol{\eta}_{\mathrm{t,n}}$ $(I \times 1)$ is the innovation vector and $\mathbf{Q}_{\mathrm{n}}$ $(I \times I)$ the covariance matrix of the innovation. All error terms,

$\boldsymbol{\eta}_{\mathrm{t,n}}$ and $\boldsymbol{\varepsilon}_{\mathrm{t,n}}$, are assumed to be mutually independent. The model includes one lag, meaning that the autoregression is with respect to the previous time point only, and this model is denoted as the VAR(1)-BDM. A graphical representation of the model can be seen in Figure 1.

We remark that $\boldsymbol{\Sigma}_n$ $(I \times I)$ is the model implied variance-covariance matrix of the observed scores for individual *n*, which is computed through

$$\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{n}}) = (\boldsymbol{I} - \boldsymbol{\Phi}'_{\mathbf{n}} \otimes \boldsymbol{\Phi}'_{\mathrm{n}})^{-1}\mathrm{vec}(\mathbf{Q}_{\mathrm{n}} + \mathbf{H}_{\mathrm{n}}), \quad (3)$$

where $\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{n}})$ is the vectorized version of $\boldsymbol{\Sigma}_{\mathrm{n}}$ and $\otimes$ denotes the Kronecker product. This is an adaptation of the Lyapunov equation used for the traditional vector autoregressive (VAR) model with only one error term, as discussed in Hamilton (1994, p. 265).

The VAR(1)-BDM includes dynamic parameters that are set to be constant across time. This implicit assumption seems most likely to be met when the dynamics of the individual under study do not change drastically (e.g., as a result of an intervention or life event) and the time interval between observations is roughly the same. A study using simulated data showed that the VAR(1)-model, and with that the VAR(1)-BDM, is very robust against violations of the assumption of equal interval width (Albers, in preparation). If time intervals would differ widely (e.g., ranging from one hour up to a week), one may use an adapted model that explicitly allows for non-equidistant time points (see Kuppens, Oravecz, and Tuerlinckx, 2010; Oravecz, Tuerlinckx, and Vandekerckhove, 2011).

VAR(1)-BDM modeling can be done also in the presence of incidental missing measurements. To this, a link function is used that links the observed $y_{\mathrm{i,t,n}}$ to a latent $y^*_{\mathrm{i,t,n}}$. In the VAR(1)-BDM model, $y_{\mathrm{i,t,n}} = y^*_{\mathrm{i,t,n}}$, because $y_{\mathrm{i,t,n}}$ and $y^*_{\mathrm{i,t,n}}$ are assumed to be equally distributed. At time points with missing data, the latent variable $y^*_{\mathrm{i,t,n}}$ is not linked to the observed variable $y_{\mathrm{i,t,n}}$. The price to be paid for the missingness is that the uncertainty

increases with more missing data points in a row, since the estimation cannot be checked against the observed data anymore. Further, the generalizability of the model to the 'population of time points' is assured only if the missing data are missing completely at random or missing at random (see e.g. Schafer & Graham, 2002).

The parameters of this statistical model are immediate translations of the EDFs, providing a direct link with emotion theory and application to empirical data. The linkage between each of the discussed EDFs and the model is summarized in Table 1. The within person variability for individual *n* and emotion *i* is expressed via $\mathbf{\Sigma}_{\mathrm{ii,n}}$ and the innovation variability via $\mathbf{Q}_{\mathrm{ii,n}}$. The autoregression for individual *n* and emotion *i* is $\mathbf{\Phi}_{\mathrm{ii,n}}$, and the cross-lag regression is $\mathbf{\Phi}_{\mathrm{ij,n}}$ for $i \neq j$. The correlation, used for the granularity, is obtained via $\mathbf{\Sigma}_{\mathrm{ij,n}}$ for $i \neq j$. The average intensity for individual *n* on emotion *i* is the mean $\mu_{\mathrm{i,n}}$. Hence, the model enables the simultaneous study of all discussed emotion dynamics.

Multiple individuals. The model can be applied without any difficulty for multivariate data collected among multiple individuals. This can be done by estimating the model of each individual separately. This implies that no assumptions are made with regard to the sampling of the individuals.

As an alternative, one may assume that the individuals are drawn at random from a certain population. As such, the parameters of the individuals can be assumed to be drawn randomly from the population distribution of the parameter concerned. These assumptions may be expressed in the model via a level 2 model, for example by assuming, as is standard in multilevel modeling, that each $\mathbf{\Phi}_{\mathrm{ij,n}}$ is drawn from a normal distribution: $\mathbf{\Phi}_{\mathrm{ij,n}} \sim N(\mathbf{\Phi}_{\mathrm{ij}}, \sigma_{\Phi_{\mathrm{ij}}})$ (Lodewyckx, Tuerlinckx, Kuppens, Allen, & Sheeber, 2011). The potential advantage is that the parameters can be estimated with more precision, provided that the

distributional assumptions on the individual parameters are met. In practice these assumptions may be too strict, yielding the first approach more attractive.

Both approaches have a heavy computational burden, with second approach being even worse than the first, as modeling all individuals jointly takes more time per iteration and more iterations to reach convergence. This is why we decided upon the first approach in this paper.

When the dynamics of a large number of individuals are studied, it typically will be of interest to examine relations between the different individual parameters. As will be illustrated in our empirical example, this can be for instance done via a cluster analysis on the individual parameters.

Possible extensions of the VAR(1)-BDM model

The VAR(1)-BDM model as defined in Equations (1) and (2) offers a rather flexible model for stationary individual time series with about normally distributed fluctuations that are constant dynamics across time. If the nature of the data requires a more flexible approach, the model can be extended in various ways.

Time varying parameters. In the VAR(1)-BDM model, parameters are assumed to be equal over time, implying that the emotion dynamics are assumed to be constant across the time span measured. In case this assumption would be too rigid, alternative models are available. For example, in regime switching or threshold models, the autoregression may change as the state changes (Hamaker & Grasman, 2012; De Haan-Rietdijk, Gottman, Bergeman, & Hamaker, 2014). Such an extension would be useful for modelling significant personal changes, for instance as the consequence of clinical intervention. In time-varying autoregressive models, more gradual changes in parameters across time can occur

(Bringmann, Hamaker, Vigo, Aubert, Borsboom, Tuerlinckx, 2016), reflecting that the dynamics of psychological processes are not stationary in the long run.

Non-normal distributions. The VAR(1)-BDM model can be extended to non-normal distributions in two ways. First, when the observations are assumed to be non-Gaussian realizations of an underlying Gaussian process, a link function can be used to transform the latent Gaussian scores into estimated observed non-Gaussian scores. Examples are a probit-link for ordinal data (Chaubert, Mortier, & Saint André, 2008), or the log-link for count data (assuming a Poisson distribution) (Terui, Ban, & Maki, 2010; Krone, Albers & Timmerman, 2016a). However, this adds more complexity to the model, requiring a larger sample size to estimate the model parameters with reasonable precision. Second, when the underlying process is assumed to be non-Gaussian, the distributions used in the model, for example the white noise and innovation distributions, can be adjusted accordingly (Durbin & Koopman, 2012; West & Harrison, 1997, Ch. 13, 14).

More than a single lag. The VAR(1) model includes an autoregression on the previous time point only. The model can be extended with autoregressive effects from earlier time points as well, yielding a VAR(p) model, with p the number of previous time points regressed upon. To detect such effects, one would need very intensively sampled data. Therefore we deem this extension to be of possible use to model for instance psychophysiological measures, but not for emotion ratings, as this implies a very heavy burden on respondents.

External variables. Research question often probe how dynamic features of emotions are related to, or a function of, other variables, such as experimental manipulation or individual differences reflecting personality or well-being. Due to the flexible nature of the model, external variables can be dealt with in two ways. First, the external variable can be included in the model as an active covariate. This can be done as a

direct effect, for example letting $\mathbf{y}_{t,n}$ being dependent on $\boldsymbol{\mu}_n$, $\boldsymbol{\theta}_{t,n}$ and a covariate, and as a moderator effect, for example letting elements of $\boldsymbol{\Phi}_n$ be dependent on the level of a covariate. Second, inactive covariates can be implemented post-hoc, by examining the relation between any model parameter and an external variable after the model estimation. This can be done, for instance, by using partial correlations or linear regression, thereby accounting for confounding variables.

Modeling empirical time series

Model estimation. The VAR(1)-BDM is a Bayesian model that can be estimated using Bayesian Markov Chain Monte Carlo (MCMC). To this end, we use Hamiltonian Monte Carlo (HMC), a generalization of the Metropolis-Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953; Hastings, 1970) that allows for an efficient estimation of the parameters (Gelman et al., 2013). This is incorporated in the software RStan (Stan Development Team, 2014; R Core Team, 2015). Example R-code and Stan-code for the VAR(1)-BDM can be found in Appendices A and B, respectively.

In Bayesian modelling, prior distributions, quantifying the a priori degree of belief in parameter values, have to be specified. In empirical situations where there is relevant context information (such as results of a pilot study), this information can be incorporated by specifying informative priors. Alternatively, one can aim for weak-informative priors, to reduce the influence of the choice of priors on the estimates.

Estimating the model to empirical data may yield convergence problems. In general, these problems may be due to the identifiability of the model and/or a lack of data. The VAR(1)-BDM as expressed in Equations 1 and 2, is identified. Therefore, when estimation issues arise with this model, this is due to a lack of data. It is impossible to offer a general

guideline on the minimal number of measurements required, because this would depend on the specific values of the population parameters and the required precision.

To check the convergence, two methods may be used: assessment of the potential scale reduction factor, $\hat{R}$, and visual inspection of the trace plots. The $\hat{R}$ shows the ratio of how much the estimation may change when the number of iterations is doubled, with an (ideal) value of 1 indicating that no change is expected (Gelman & Rubin, 1992; Stan Development Team, 2014). Trace plots show the MCMC estimates for each parameter at each iteration. If a parameter reaches convergence, the estimates over iterations are highly similar across chains. As a result, the trace plot will look like a fat caterpillar where all chains completely overlap, except at the fringe of the caterpillar (as shown in Figure 3).

Model selection. In modeling empirical emotion dynamics, one might ask whether to use the VAR(1)-BDM, or a constrained version thereof, or even a more extended model. As in any statistical model specification (see Snijders & Bosker, 1999, p. 91), the steering wheels for model selection are substantive and statistical considerations.

With respect to the substantive considerations, the core advantage of the VAR(1)-BDM to characterize changes in emotions over time, is that its parameters can be interpreted directly in terms of emotion dynamics features. Subject-matter related considerations may indicate the need for an extended model, as described above. For example, when the outcome variables pertain to counts, the assumption of a normal distribution would be too far off and a Poisson distribution would be a better choice.

With respect to the statistical considerations, it is useful to distinguish tests for individual parameters from measures of model fit. A test for an individual parameter can be used to assess the evidence for a particular value of a specific parameter. In this way, one could assess whether all parameters of the VAR(1)-BDM would be actually needed to

describe the time series of a single subject, or whether one could do with a constrained version of the VAR(1)-BDM, for example by fixing an innovation term for a specific variable at 0, or by setting the cross-lag between two specific emotions at 0.

Measures of model fit indicate the fit in a global way, instead of focusing on a single parameter. As an absolute fit measure, we consider the predictive value of the model for the next measurement, via the root mean squared error for emotion i of individual n ($RMSE_{i,n}$), as

$$\mathrm{RMSE}_{\mathrm{i,n}} = \sqrt{\sum_{\mathrm{t=2}}^{\mathrm{T_n}} \left(y_{\mathrm{i,t,n}} - \hat{y}_{\mathrm{i,(t-1),n}}\right)^2}, \quad (4)$$

with $\hat{y}_{\mathrm{i,(t-1),n}}$ the model predicted score at time point t, on the basis of the observed score at time point t-1. An RMSE-value of 0 indicates perfect prediction, and larger values imply a lower predictive value.

In case substantive considerations would yield various competing models (e.g., VAR(1)-BDM versus VAR(2)-BDM), information criteria such as the Bayesian information criterion (BIC; Schwarz, 1978) or the Watanabe-Akaike information criterion (WAIC; Watanabe, 2010) are of use. Such criteria add a penalty for model complexity to the model fit (as expressed via minus two times the log likelihood of the model). A smaller BIC/WAIC value points at better model, thus the model with the lowest BIC/WAIC is favored. Which criterion to select is an ongoing debate between statisticians (Gelman, Hwang, Vehtari, 2014), but in practice the criteria usually point into the same direction.

Empirical example

In this paper, we re-analyse data described in Brans, Koval, Verduyn, Lim, and Kuppens (2013) and Erbas et al. (2014). As part of a larger study, the emotions of 50

individuals were self-reported using experience sampling. The data collection process consisted of three parts. In a first lab session the participants signed an informed consent form and were handed out a palmtop which they would use to record their emotions, along with instructions for its use. Second, during at least seven days (extending to a maximum of 10 days for some participants), participants carried the palmtops and self-reported their emotions using the Experience Sampling Program (ESP) (Barrett & Barrett, 2000). The waking hours of the individuals on each day were divided into ten intervals. During each interval, at a random time point, the ESP would ask them to rate their emotions in terms of how angry, depressed and stressed, for example, they felt at that moment. This yielded observations at roughly similar time intervals. Each emotion was rated on a 6-point Likert scale, ranging from 0 to 5, with higher values indicating a stronger feeling of that emotion. Finally, in a second lab session, the participants returned the palmtops and were each given AU$ 40,- for their participation.

Sample data. We selected the data on the three negative emotions 'angry', 'depressed' and 'stressed' for re-analysis. The number of time points observed of the 50 individuals ranged from 20 to 90, with incidental missing data. To illustrate the sample data, the observed scores of three individuals are depicted in Figure 2. As can be seen, the first individual shows missing data all through the sample, the second shows large patches of complete and missing data, and the third shows little missing data.

Model specification. To model the trivariate (i.e., of angry, depressed, stressed) time series of the 50 individuals we applied the VAR(1)-BDM, for each of the 50 individuals separately. This implies that there is no assumption with regard to the sampling of the individuals. In view of the limited length of the observed time series, we refrained from considering more complicated models (e.g., VAR(2)-BDM, or explicitly modeling the discrete

nature of the dependent variable), as this would involve even more parameters to estimate yielding unstable results.

The priors for the parameters were specified as follows. For the elements of $\mathbf{\Phi}_n$ we used a symmetrized reference prior (Berger & Yang, 1994); for the scale parameter of $\mathbf{H}_n$ a half-Cauchy(0, 2.5) prior; for the correlation matrix of $\mathbf{H}_n$ a Lewandowski-Kurowicka-Joe correlation prior (Lewandowski, Kurowicka, & Joe, 2009); for the diagonal elements of $\mathbf{Q}_n$ a $\Gamma(3,3)$; and for $\mu_{i,n}$ a normal prior with mean 0 and variance 4.

Results

We start by presenting results concerning the estimation (i.e., convergence and computation time), and the absolute fit of the individual time series by as expressed via the RMSE. Then, we summarize the parameter estimates related to the emotion dynamic features across the 50 individuals.

Estimation: Convergence and computation time. In our analysis, we used four MCMC-chains of 30.000 iterations each. To check whether convergence was reached, we used the $\hat{R}$ and the trace plots. All elements of the model parameters ($\mathbf{\Phi}_n$, $\mathbf{Q}_n$ and $\boldsymbol{\mu}_n$) reached convergence for all individuals and emotions, with an $\hat{R}$ below 1.02 for all estimated model parameters. One of the trace plots, representable for all relevant trace plots, is given in Figure 3 which, as can be seen, gives the expected fat caterpillar.

The total computation time (computer with 24 Intel Xeon 2.5 GHz cores) was 20.7 days, with a mean (SD) computation time per individual of 9.93 (4.70) hours. Note that parallel computing can be used in order to get the waiting time considerably smaller than the computing time.

The RMSE value had a mean (SD) across the 50 individuals of .20 (.21) for angry, .44 (.47) for depressed and .50 (.45) for stressed. Given that each of the emotions are rated on a 6-point Likert scale this indicates a reasonable absolute predictive fit of the models.

Summary of emotion dynamic features. The VAR(1)-BDM of the emotions angry, depressed and stressed of each individual involved 21 parameters of interest. For each of the 21 parameters of the 50 individuals, we computed the mean posterior parameter estimate, briefly denoted as parameter estimate in what follows. In Table 2 (columns 3, 4 and 5), the means, standard deviations and 95% confidence intervals of all parameter estimates across all individuals are presented. As can be seen, the standard deviations are rather large for all parameters, implying that there is a large variability in individual dynamics.

To achieve an insightful summary of the individual similarities and differences of emotion dynamic features, we performed a *K*-means cluster analysis (MacQueen, 1967) on the parameter estimates (in R, using 100 random starts), after scaling to ensure an equal weight of the variables to the cluster solution. As variables in the *K*-means analysis we included all parameters except for the within-person variability ($\mathbf{\Sigma}_{\mathrm{ij,n}}$), because of some outlying values (e.g., 7 values above 6). These outliers are likely due to instable model implied estimates. Based on a scree-plot of the total within-cluster sum of squares versus the number of clusters, we selected five clusters. The cluster size ranged from 9 to 12 individuals; for each cluster the mean parameter estimate and the within cluster standard deviation of each of the 21 parameters related to the emotion dynamic features are presented in Table 2 (columns 5 to 9); we ordered clusters 1 to 5 according to their average intensities (i.e., $\mu_{\mathrm{i,n}}$) for angry, depressed and stressed, with clusters 1 and 2 showing low, clusters 3 and 4 medium, and cluster 5 high levels of intensity.

To depict the nature of the time series of the VAR(1)-BDM for each of the five clusters, we simulated time series, using the mean parameter estimates of that cluster a parameter value of the simulated model. The simulated time series for 70 time points for each of the variables and clusters are depicted in Figure 4. The appearance of the time series differs between the series, but the nature of the differences in dynamics is, of course, expressed in more detail via the differences in the model parameters. Close inspection of Figure 4 reveals that the intensity increases from Cluster 1 to 5 (by definition, as we ordered them as such), and more importantly, that the short-term and long-term dynamics characterizing the clusters differ.

We will discuss the emotion dynamic features in the same order as in the introduction, i.e., the parameters related to the within person variability, innovation variability, inertia, cross-lag, granularity and intensity.

Emotional variability is quantified using the EDFs within person variability ($\mathbf{\Sigma}_{\mathrm{ii,n}}$) and innovation variability ($\mathbf{Q}_{\mathrm{ii,n}}$). As can be seen in Table 2, the mean within person variability and innovation variability differs substantially across emotions and clusters. The within person variability ($\mathbf{\Sigma}_{\mathrm{ii,n}}$), indicating the overall variability across all measurements, of angry is smallest, where depressed and stressed show roughly equal variability across clusters. In contrast, the computed innovation variability, expressing the amount of change that carries over to the next measurement, is about similar for angry and depressed in all clusters, and in cluster 3 substantially larger than in the other clusters.

To see to what extent within person variability and innovation variability would be related across all individuals, we computed Spearman's rho between the within person variability ($\mathbf{\Sigma}_{\mathrm{ii,n}}$) and innovation variability ($\mathbf{Q}_{\mathrm{ii,n}}$), per emotion. The values were 0.80 for angry, 0.43 for depressed and 0.40 for stressed. This suggests that the individual differences

in dynamics for these three emotions are to some extent timescale-invariant, i.e., individuals with a high variability over the whole time period also show high variability between two time points. This has been found elsewhere as well (Kuppens, Oravecz, & Tuerlinckx, 2010). This effect is much larger for angry than for depressed and stressed. This indicates that short-term anger fluctuations are more informative for its long-term fluctuations as compared to those of depression and stress feelings, perhaps because the latter typically change more slowly (or last longer; Verduyn & Lavrijsen, 2015).

Emotional inertia is quantified using the autoregression. As can be seen in Table 2, the mean autoregression differs substantially across emotions and clusters. Generally, the mean values are positive, indicating that a relative highly scored emotion (relative in deviation from the overall mean of the time series for that individual) is followed by a relatively high scored emotion on the next measurement. The highest mean values are found for stressed for clusters 3 and 4 (about .50), suggesting a relative large stability of reported stress levels over time. Taken together, these results provide evidence that emotions are self-related over time. This is reflected in the literature. Indeed, most previous research has found that emotional states tend to be mildly or strongly predictive over time in daily life (e.g., Suls et al., 1998; Kuppens, Allen, & Sheeber, 2010; Koval et al., 2012). In exceptional cases (e.g., depressed in cluster 3), the autoregressive parameter is negative, indicating individuals whose emotions seem to contrast themselves from one moment to the next. An interesting avenue for future research would consist of understanding what contributes to such dynamic patterns in these individuals.

Emotional cross-lag. The impact of one emotion on another is measured through the cross-lag regressions. As can be seen in Table 2, the mean cross-lag regressions differ across emotions and clusters, including their directions (i.e., positive and negative). This suggests

large individual differences in degree of augmenting and blunting effects of emotions. The substantial individual differences that are found in the extent to which different emotions augment or blunt each other across time, are consistent with previous research (Pe & Kuppens, 2012). The effects seem largest in clusters 3 and 4, that have a medium intensity in emotion ratings.

A useful tool in interpreting the autoregression matrices – both the emotion inertia as the emotion cross-lag – is by network visualization (cf. Bringmann, Pe, Vissers, Ceulemans, Borsboom, Vanpaemel, Tuerlinckx, Kuppens, 2016, for an example from emotion psychology). Figure 5 displays the mean autoregressive relations across all 50 individual models: out of the three emotions, stress clearly has the strongest inertia. The strongest relation between emotions is a positive relation between stress and depressed: higher levels of stress precede higher levels of depression, on average. Lagged relations between angry and depressed are virtually absent.

Where Figure 5 displays the network for the average of the 50 participants, Figure 6 provides the network graph for each of the five clusters. This visualization demonstrates where the differences between the clusters are most prominent. Interestingly, even though the overall average network (Figure 5) only has positive connections, four out of five clusters display at least one negative relation, and thus instances where one emotion blunts the subsequent experience of another emotion. The first cluster can be classified as one where the ‘stressed → depressed’ relation is reciprocated by a ‘depressed → stressed’ relation. The second cluster mainly has stronger relations than the average individual. Cluster 3 can be typified as having negative outgoing relations from depression to all three emotions. Cluster 4, on the other hand, has negative outgoing relations from angry. Cluster 5 has strong auto-regressions but relatively small cross-regressions. Note that the first two clusters are closer

to the average cluster than the other three: The value of the distance metric $\sum_{i,j}\left|\boldsymbol{\Phi}_{c;i,j} - \bar{\boldsymbol{\Phi}}_{i,j}\right|$, with $\boldsymbol{\Phi}_c$ denoting the autoregression matrix for cluster c, and $\bar{\boldsymbol{\Phi}}$ the autoregression matrix of the average, is 0.62 for cluster 1, 0.78 for cluster 2, 1.38 for cluster 5, 1.87 for cluster 4, and 1.99 for cluster 3.

Emotional granularity. Granularity expresses the covariation between emotions and is quantified via the bivariate correlation between the different couples of emotions per individual. Here, higher values indicate a lower granularity. The correlations between the paired emotions are generally positive (see Table 2). Except for cluster 1, the correlations are moderate (mean .25 for angry, depressed in cluster 4) to strong (mean .70 for depressed, stressed in cluster 2). This strongly resonates with previous research showing generally positive relations between like-valenced emotional states across time within individuals (e.g., Vansteelandt, Van Mechelen, & Nezlek, 2005; Brose et al., 2015; Carstensen et al., 2000). The variation in correlations across individuals indicate individual differences in the level of emotion differentiation or granularity, thought to be indicative of differences in emotion regulation and functioning (e.g., Barrett et al., 2001; Erbas et al., 2014).

Emotional Intensity, quantified as $\mu_{i,n}$, appears to differ substantially across clusters (and thus individuals), but their order seems rather similar, with angry having the lowest intensity, and depressed and stressed about equal intensity. The intensity (see Table 2) is lowest for clusters 1 and 2, medium for clusters 3 and 4, and highest cluster 5.

In earlier studies, a higher average intensity over all emotions seemed related to a lower average correlation over all emotion pairs (Carstensen et al., 2000; Kashdan & Farmer, 2014; Erbas et al., 2014). This finding was not replicated in our sample, with Spearman's rho

values between the average correlation over all emotion pairs and the intensity of .02 for angry, .06 for depressed and .05 for stressed.

Discussion

In this paper we proposed to use a BDM to analyze intensive longitudinal data to capture the patterns and regularities of an individual's experience of emotions across time. To accommodate for missing data, a link function was introduced, linking the observed to the latent variable only when the data is indeed observed. With this VAR-BDM we analyzed a data set consisting of three emotions for 50 individuals. Using cluster analysis we subsequently constructed five clusters of typical emotion dynamics. Our data set consisted of self-report data, but the VAR-BDM model is also applicable to behavioural and physiological measures of emotions over time, as well as to data sets containing a mixture of such variables.

The study of emotion dynamics and the relation between EDFs is an important topic in psychological research. Most of the earlier studies regarding emotion dynamics have computed summary statistics for several EDFs (e.g., Carstensen et al., 2000; Barrett et al., 2001; Erbas et al., 2014; Scott et al., 2014). A complete model that encompasses all EDFs capturing the essential dynamics of multivariate data and that can be applied to EMA data with its inevitable limitations, was lacking so far. Models that do capture multiple EDFs at once, often require rather large sample sizes of at least 100 time points without missing values (e.g., Hamaker & Grasman, 2012; De Haan-Rietdijk et al., 2014). Another option, which does deal with the small sample size, is the use of a multilevel model. However, these models generally require a large number of individuals measured, and are based on distributional assumptions on the individual parameters (e.g., Kuppens, Allen, & Sheeber, 2010; Lodewyckx et al., 2011; Bringmann et al., 2013, 2016). The multilevel VAR model of

Schuurman, Ferrer, de Boer-Sonnenschein, & Hamaker (2016) lacked an important element, namely the white noise. Further, they did not link the model elements to the EDFs as we did.

The results found in our empirical study largely concur with the results in previous literature. With our model, we strive at aiding in the research on emotion dynamics, allowing for testing the proposed theories on empirical data. The theories on emotion dynamics are growing in number, but also in complexity (Gross, 2015; Kuppens & Verduyn, 2015). Due to the expansiveness of our model, it is possible to identify relations between emotions within individuals and between individuals, but also among EDFs, such as the relation between inertia and variability, or between EDFs and external variables, such as the relation between the average inertia and the level of mood disorder or other forms of psychopathology. Note that to quantify such relationships, a data set with a large number of individuals is needed. This allows for a more direct way to validating theories stating relations between EDFs.

Further advantages of the model lies in its flexibility. The model, as shown in Equations 1 and 2, can be used with non-Gaussian data through the implementation of the link function or through adjustment of the distributions used. The link function also allows for data with missing values, without the need to adapt the model. Furthermore, the model can be used to implement external variables, both as active and inactive covariates. The final advantage lies in the wide range of data for which the model is applicable: one individual for which one emotion is measured, is enough for estimation. However, more individuals and emotions can be added.

Although the expansiveness of the proposed model is an advantage, it also introduces issues with estimation. While the proposed VAR-BDM (Equations 1 and 2) can be estimated with relatively little data points, the estimated credible intervals, and thus the uncertainty,

is large for data sets with short time series. The estimation of the cross-lag regression carries another issue. In empirical data it may occur that the cross-lag regressions of the different variables are substantial. This introduces multicollinearity, which complicates estimation.

Further studies should focus on the estimation properties of the model. The conditions under which the model reaches convergence while estimating, are unknown, as are the conditions needed to estimate the parameters with reasonable precision. Factors which may be of influence on the estimation and convergence of the model are the number of time points, the number of emotions, the values in $\mathbf{\Phi}_n$ and the number and pattern of missing data in the data set. These same factors may influence the precision with which the parameters may be estimated.

The cluster analysis clearly shows the multidimensional nature of emotion dynamics. By focusing on a single (set of) variable(s) only, important distinctions would be missed. For instance, clusters 1 and 2 are both characterized by having low values for the emotional intensity, yet differ on inertia, emotional cross-lag and variability. Clusters 3, 4 and 5 have similar values on the white noise variance per emotion, yet are different on the other characteristics. This demonstrates the need for the multivariate model with the ingredients as employed in this paper.

A relevant future research question is how much measurements are needed to obtain accurate parameter estimates in the VAR(1)-BDM, given certain settings. Research on related models (Krone, Albers, & Timmerman, 2016b, 2017; Schuurman et al., 2015) suggest that at least 50-100 measurements are required to estimate the parameters of an individual accurately. However, in the present study we mainly focus on clusters of 9-12 individuals each. Parameter estimation for the clusters can draw power from all individuals, thus requiring fewer measurements per person for accurate estimates.

632 References

633 Albers, C. J. (in preparation). Robustness of discrete time models against violation of
634 equidistant time point assumption, in preparation.
635 Barrett, L. F. & Barrett, D. (2000). The Experience Sampling Program. Retrieved from
636 http://www.experience-sampling.org
637 Barrett, L. F. & Gross, J. J. (2001). Emotional intelligence: A process model of emotion
638 representation and regulation. In T. J. M. G. A. Bonanno (Ed.), *Emotions: currrent*
639 *issues and future directions* (pp. 286–310). Emotions and social behavior. New York:
640 The Guilford Press.
641 Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're
642 feeling and knowing what to do about it: Mapping the relation between emotion
643 differentiation and emotion regulation. *Cognition and Emotion*, *15*(6),
644 713–724. doi:10.1080/02699930143000239
645 Berger, J. O. & Yang, R.-Y. (1994). Noninformative Priors and Bayesian Testing for the AR(1)
646 Model. *Econometric Theory*, *10*(3-4), 461–482. doi:10.1017/S026646660000863X
647 Bolger, N. & Laurenceau, J.-P. (2013). *Methodology In The Social Sciences: Intensive*
648 *Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*.
649 New York: The Guildford Press.
650 Bos, F. M., Schoevers, R. A., & Aan het Rot, M. (2015). Experience sampling and ecological
651 momentary assessment studies in psychopharmacology: A systematic review.
652 *European neuropsychopharmacology*, *25*(18), 1853–1864.
653 doi:10.1016/j.euroneuro.2015.08.008
654 Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative
655 and positive affect in daily life. *Emotion*, *13*(5), 926–939. doi:10.1037/a0032400

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., ... Tuerlinckx, F. (2013). A network approach to psychopathology: new insights into clinical longitudinal data. *PloS one*, *8*(4), e60188. doi:10.1371/journal.pone.0060188

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., Tuerlinckx, F. (2016). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological methods*, doi:10.1037/met0000085

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., Tuerlinckx, F., & Kuppers, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, 23(4), doi: 10.1177/1073191116645909

Brose, A., de Roover, K., Ceulemans, E., & Kuppens, P. (2015). Older adults' affective experiences across 100 days are less variable and less complex than younger adults'. *Psychology and Aging*, *30*(1), 194–208. doi:10.1037/a0038690

Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. R. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology*, *79*(4), 644–655. doi:10.1037/0022-3514.79.4.644

Chaubert, F., Mortier, F., & Saint André, L. (2008). Multivariate dynamic model for ordinal outcomes. *Journal of Multivariate Analysis*, *99*(8), 1717–1732. doi:http://dx.doi.org/10.1016/j.jmva.2008.01.011

Cook, J., Tyson, R., White, J., Rushe, R., Gottman, J., & Murray, J. (1995). Mathematics of marital conflict: Qualitative dynamic mathematical modeling of marital interaction. *Journal of Family Psychology*. Methodological Advances in Family Psychology, *9*(2), 110–130. doi:10.1037/0893-3200.9.2.110

De Haan-Rietdijk, S., Gottman, J. M., Bergeman, C. S., & Hamaker, E. L. (2014). Get Over It! A Multilevel Threshold Autoregressive Model for State-Dependent Affect Regulation. *Psychometrika*. doi:10.1007/s11336-014-9417-x

Durbin, J. & Koopman, S. J. (2012). *Time Series Analysis by State Space Models*. Oxford, UK: Oxford University Press.

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schittmann, V. D., Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of statistical software*, 48(4): 1-18. http://www.jstatsoft.org/v48/i04

Erbas, Y., Ceulemans, E., Pe, M. L., Koval, P., & Kuppens, P. (2014). Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion*, *28*(7), 1196–1213. doi:10.1080/02699931.2013.875890

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*(1), 153–166. doi:10.1037/0022-3514.69.1.153

Frijda, N. H. (2007). *The laws of emotion*. Mahwah, NJ: Erlbaum.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd). Roca Baton, FL: Chapman and Hall/CRC.

Gelman, A., Hwang, J. & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6), 997-1016. doi:10.1007/s11222-013-9416-2

Gelman, A. & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457–472. doi:10.1214/ss/1177011136

Gross, J. J. (2015). The Extended Process Model of Emotion Regulation: Elaborations,

Applications, and Future Directions. *Psychological Inquiry*, *26*(1), 130–137. doi:10.1080/1047840X.2015.989751

Grühn, D., Lumley, M. A., Diehl, M., & Labouvie-Vief, G. (2013). Time-based indicators of emotional complexity: Interrelations and correlates. Time-based indicators of emotional complexity: interrelations and correlates. *Emotion, 13*(2), 226-237. doi:10.1037/a0030363

Hamaker, E. L., Ceulemans, E., Grasman, R. P. P. P., & Tuerlinckx, F. (2015). Modeling Affect Dynamics: State of the Art and Future Challenges. *Emotion Review*, *7*(4), 316–322. doi:10.1177/1754073915590619

Hamaker, E. L. & Grasman, R. P. P. (2012). Regime Switching State-Space Models applied to psychological processes: Handling missing data and making inferences. *Psychometrika*, *77*(2), 400–422. doi:10.1007/s11336-012-9254-8

Hamilton, J. D. (1994). *Time series analysis*. Princeton, N.J., Princeton University Press.

Harvey, A. C. (1990). Forecasting, structural time series models, and the Kalman filter. Cambridge, UK: Cambridge University Press.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. doi:10.1093/biomet/57.1.97

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The Relation Between Short-Term Emotion Dynamics and Psychological Well-Being: A Meta-Analysis. *Psychological Bulletin*. doi:10.1037/a0038822

Jahng, S., Wood, P., & Trull, T. J. (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychological methods*, *13*(4), 354–375. doi:10.1037/a0014173

Kashdan, T. B. & Farmer, A. S. (2014). Differentiating emotions across contexts: Comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion*, *14*(3), 629–638. doi:10.1037/a0035796

Kashdan, T. B. & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review Clinical Psychology Review*, *30*(7), 865–878. doi:10.1016/j.cpr.2010.03.001

Koval, P., Brose, A., Pe, M. L., Houben, M., Erbas, Y., Champagne, D., & Kuppens, P. (2015). Emotional inertia and external events: The roles of exposure, reactivity, and recovery. *Emotion*, *15*(5), 625–636. doi:10.1037/emo0000059

Koval, P., Kuppens, P., Allen, N. B., & Sheeber, L. (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition & Emotion*, *26*(8), 1412–1427. doi:10.1080/02699931.2012.667392

Krone, T., Albers, C. J., & Timmerman, M. E. (2016a). Bayesian dynamic modelling to assess differential treatment effects on panic attack frequencies. *Statistical Modelling*, 16(5), doi: 10.1177/1471082X16650777

Krone, T., Albers, C. J., & Timmerman, M. E. (2016b). Comparison of estimation procedures for multilevel AR(1) models. *Frontiers in Psychology*, 4(486), doi:10.3389/fpsych.2016.00486

Krone, T., Albers, C. J., & Timmerman, M. E. (2017). A comparative simulation study of AR(1) estimators in short time series. *Quality & Quantity*, 51(1): 1-21. doi:10.1007/s11135-015-0290-1

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional Inertia and Psychological Maladjustment. *Psychological Science*, *21*(7), 984–991. doi:10.1177/0956797610372634

749 Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual
750 differences in the temporal dynamics of affect. *Journal of Personality and Social*
751 *Psychology*, *99*(6), 1042–1060. doi:10.1037/a0020962
752 Kuppens, P. & Verduyn, P. (2015). Looking at Emotion Regulation Through the Window of
753 Emotion Dynamics. *Psychological Inquiry*, *26*(1), 72–79.
754 doi:10.1080/1047840X.2015.960505
755 Larsen, R. J. (2000). Toward a Science of Mood Regulation. *Psychological Inquiry*, *11*(3), 129–
756 141. doi:10.1207/S15327965PLI1103_01
757 Larson, R. & Csikszentmihalyi, M. (1983). The Experience Sampling Method. In *New*
758 *directions for methodology of social and behavioral sciences* (Vol. 15, pp. 41–56).
759 Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices
760 based on vines and extended onion method. *Journal of MultivariateAnalysis*, *100*(9),
761 1989–2001. doi:10.1016/j.jmva.2009.04.008
762 Lewis, M., Haviland-Jones, J. M., & Feldman Barret, L. (Eds.). (2008). *Handbook of Emotions*
763 (3rd edition). New York, NY: The Guildford Press.
764 Lodewyckx, T., Tuerlinckx, F., Kuppens, P., Allen, N. B., & Sheeber, L. (2011). A hierarchical
765 state space approach to affective dynamics. *Journal of Mathematical*
766 *Psychology*, *55*(1), 68–83. doi:10.1016/j.jmp.2010.08.004
767 MacQueen, J. (1967). Some methods for classification and analysis of multivariate
768 observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical*
769 *Statistics and Probability*, eds L. M. Le Cam & J. Neyman, **1**, pp. 281–297. Berkeley, CA:
770 University of California Press.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, *21*, 1087–1092. doi:10.1063/1.1699114

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods, 16*(4), 468-490. doi:10.1037/a0024375

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., ... Gotlib, I. H. (2015). Emotion-Network Density in Major Depressive Disorder. *Clinical Psychological Science Clinical Psychological Science*, *3*(2), 292–300. doi:10.1177/2167702614540645

Pe, M. L. & Kuppens, P. (2012). The dynamic interplay between emotions in daily life: augmentation, blunting, and the role of appraisal overlap. *Emotion*, *12*(6), 1320–1328. doi:10.1037/a0028262

R Core Team. (2015). *R: A Language and Environment for Statistical Computing.* Vienna, Austria. Retrieved from http://www.r-project.org/

Röcke, C., Li, S.-C., & Smith, J. (2009). Intraindividual variability in positive and negative affect over 45 days: Do older adults fluctuate less than young adults? *Psychology and Aging*, *24*(4), 863–878. doi:10.1037/a0016276

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods*, *7*(2), 147.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, *23*(7), 1307–1351. doi:10.1080/02699930902928969

794 Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to
795 Compare Cross-Lagged Associations in a Multilevel Autoregressive Model.
796 *Psychological Methods, 21(*2), 206-221. doi:10.1037/met0000062
797 Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement
798 error in n=1 psychological autoregressive modeling. *Frontiers in Psychology*, *6*(1038).
799 doi:10.3389/fpsyg.2015.01038
800 Scott, S. B., Sliwinski, M. J., Mogle, J. A., & Almeida, D. M. (2014). Age, stress, and emotional
801 complexity: Results from two studies of daily experiences. *Psychology and Aging*,
802 *29*(3), 577–587. doi:10.1037/a0037282
803 Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment.
804 *Annual review of clinical psychology*, *4*, 1–32.
805 doi:10.1146/annurev.clinpsy.3.022806.091415
806 Snijders, T.A.B., & Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and*
807 *applied multilevel analysis.* London: Sage.
808 Stan Development Team. (2014). Rstan: the R interface to Stan, version 2.4. Retrieved from
809 http://mc-stan.org/rstan.html
810 Suls, J., Green, P., & Hills, S. (1998). Emotional reactivity to everyday problems, affective
811 inertia and Neuroticism. *Personality and social psychology bulletin*, *24*(2), 127–136.
812 doi:10.1177/0146167298242002
813 Terui, N., Ban, M., & Maki, T. (2010). Finding market structure by sales count dynamics -
814 multivariate structural time series models with hierarchical structure for count data.
815 *Annals of the Institute of Statistical Mathematics*, *62*(1), 91–107. doi:10/1007/s10463-
816 009-0244-2

Tugade, M. M., Fredrickson, B. L., & Feldman Barrett, L. (2004). Psychological Resilience and Positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality*, *72*(6), 1161–1190. doi:10.1111/j.1467-6494.2004.00294.x

Vansteelandt, K., Van Mechelen, I., & Nezlek, J. B. (2005). The co-occurrence of emotions in daily life: A multilevel approach. *Journal of Research in Personality*, *39*(3), 325–335. doi:10.1016/j.jrp.2004.05.006

Verduyn, P., & Lavrijsen, S. (2015). Which emotions last longest and why: The role of event importance and rumination. *Motivation and Emotion*, *39*(1), 119-127.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571-3594.

West, M. & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. New York, NY: Springer Verlag.

Wetzels, R. & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064. doi:10.3758/s13423-012-0295-x

Wichers, M., Wigman, J. T. W., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review*, *7*(4), 362–367. doi:10.1177/1754073915590623

838

Concept	Quantification	Parameter
Within person variability	Variance of $Y_{\mathrm{i,n}}$	$\mathbf{\Sigma}_{\mathrm{ii,n}}$
Innovation variability	Innovation of $Y_{\mathrm{i,n}}$	$\mathbf{Q}_{\mathrm{ii,n}}$
Inertia	Autoregression	$\mathbf{\Phi}_{\mathrm{ii,n}}$
Emotional cross-lag	Cross-lag regression	$\mathbf{\Phi}_{\mathrm{ij,n}}$
Granularity	Covariance of $y_{\mathrm{i,n}}$	$\mathbf{\Sigma}_{\mathrm{ij,n}}$
	Correlation of $y_{\mathrm{ij,n}}$	$\mathrm{Cor}(\boldsymbol{y}_{\mathrm{ij,n}})$
Intensity	Mean estimated score	$\mu_{\mathrm{i,n}}$

839 Table 1

840 Quantification of emotion dynamics features for emotion i, in relation to emotion j where

841 applicable, in the notation of Equations 1, 2 and 3.

Cluster				**1**	**2**	**3**	**4**	**5**
Number of individuals		**50**		**9**	**12**	**9**	**10**	**10**
		Mean (SD)	**95%CI**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Within person variability $\boldsymbol{\Sigma_{ii,n}}$	**Angry**	1.34 (2.58)	[0.61,2.07]	0.34 (0.43)	2.32 (5.05)	1.36 (1.15)	0.91 (0.57)	1.47 (0.92)
	Depressed	2.05 (3.25)	[1.12,2.97]	1.02 (0.86)	3.41 (6.34)	1.81 (0.89)	2.06 (1.57)	1.53 (0.71)
	Stressed	2.56 (4.33)	[1.33,3.79]	1.48 (1.36)	4.49 (8.44)	2.61 (0.86)	2.23 (2.14)	1.51 (0.85)
Innovation variability $\boldsymbol{Q_{ii,n}}$	**Angry**	0.37 (0.27)	[0.29,0.45]	0.21 (0.26)	0.21 (0.13)	0.51 (0.28)	0.36 (0.23)	0.58 (0.24)
	Depressed	0.37 (0.23)	[0.31,0.44]	0.27 (0.18)	0.19 (0.12)	0.53 (0.23)	0.38 (0.17)	0.52 (0.23)
	Stressed	0.42 (0.22)	[0.35,0.48]	0.42 (0.21)	0.25 (0.13)	0.70 (0.19)	0.34 (0.13)	0.43 (0.19)
Inertia $\boldsymbol{\Phi_{ii,n}}$	**Angry**	0.16 (0.30)	[0.07,0.24]	0.05 (0.13)	0.27 (0.32)	0.22 (0.26)	-0.08 (0.26)	0.29 (0.32)
	Depressed	0.07 (0.29)	[-0.02,0.15]	0.15 (0.25)	0.17 (0.19)	-0.33 (0.24)	0.12 (0.26)	0.17 (0.23)
	Stressed	0.25 (0.37)	[0.14,0.36]	0.15 (0.22)	0.27 (0.37)	0.51 (0.31)	0.50 (0.21)	-.18 (0.29)
Emotional cross-lag $\boldsymbol{\Phi_{ij,n}}$	**A_{t-1} on D_t**	0.01 (0.24)	[-0.06,0.08]	0.02 (0.11)	0.06 (0.24)	-0.21 (0.25)	0.10 (0.08)	0.06 (0.31)
	A_{t-1} on S_t	0.10 (0.22)	[0.04,0.17]	0.03 (0.09)	0.08 (0.27)	0.17 (0.21)	0.25 (0.17)	-0.01 (0.23)
	D_{t-1} on A_t	0.04 (0.36)	[-0.06,0.14]	-0.05 (0.38)	0.27 (0.29)	0.26 (0.26)	-0.30 (0.24)	0.00 (0.30)
	D_{t-1} on S_t	0.19 (0.34)	[0.09,0.28]	0.13 (0.19)	0.16 (0.40)	0.32 (0.25)	0.49 (0.21)	-0.15 (0.23)
	S_{t-1} on A_t	0.12 (0.32)	[0.03,0.21]	0.11 (0.30)	0.29 (0.27)	0.33 (0.25)	-0.23 (0.23)	0.07 (0.26)
	S_{t-1} on D_t	0.08 (0.33)	[-0.01,0.18]	0.18 (0.26)	0.12 (0.28)	-0.34 (0.34)	0.19 (0.14)	0.21 (0.33)
Granularity $\boldsymbol{Cor(y_{ij,n})}$	**A,D**	0.24 (0.24)	[0.17,0.31]	0.02 (0.10)	0.45 (0.23)	0.28 (0.26)	0.25 (0.18)	0.14 (0.19)
	A,S	0.31 (0.23)	[0.25,0.38]	-0.03 (0.07)	0.47 (0.19)	0.35 (0.14)	0.38 (0.13)	0.34 (0.21)
	D,S	0.52 (0.24)	[0.45,0.59]	0.37 (0.28)	0.70 (0.20)	0.46 (0.10)	0.58 (0.16)	0.44 (0.27)
Intensity $\boldsymbol{\mu_{i,n}}$	**A**	1.47 (0.65)	[1.29,1.66]	1.12 (0.13)	1.09 (0.47)	1.45 (0.60)	1.64 (0.53)	2.09 (0.82)
	D	2.06 (0.83)	[1.82,2.30]	1.33 (0.23)	1.51 (0.62)	2.24 (0.69)	2.27 (0.71)	3.01 (0.58)
	S	2.09 (0.85)	[1.85,2.33]	1.45 (0.26)	1.45 (0.67)	2.29 (0.41)	2.39 (0.82)	2.95 (0.76)

Table 2. Mean parameter estimates per EDF, across all individuals (*N* = 50) and for each of the 5 clusters, resulting from the *K*-means analysis. A, D, and S stand for Angry, Depressed and Stressed, respectively.

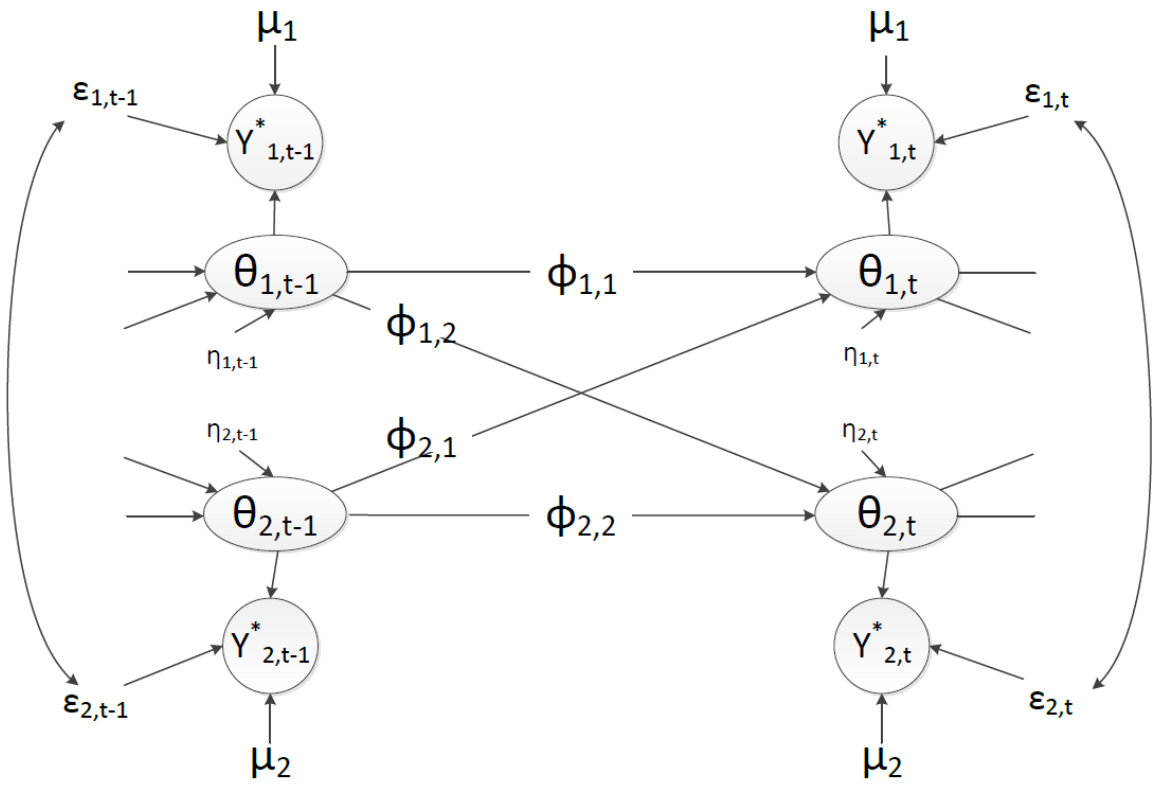

Figure 1. Schematic representation of the model as expressed in Equations 1 and 2 for a single individual and *i* = 2 emotions. As the focus lies on a single individual, all variables would have 'n' in their subscript. For clarity, this is omitted. We use standard notation for referring to elements of vectors and matrices. For instance, $\theta_{1,t-1}$ represents the first element of the vector $\boldsymbol{\theta}_{t-1}$.

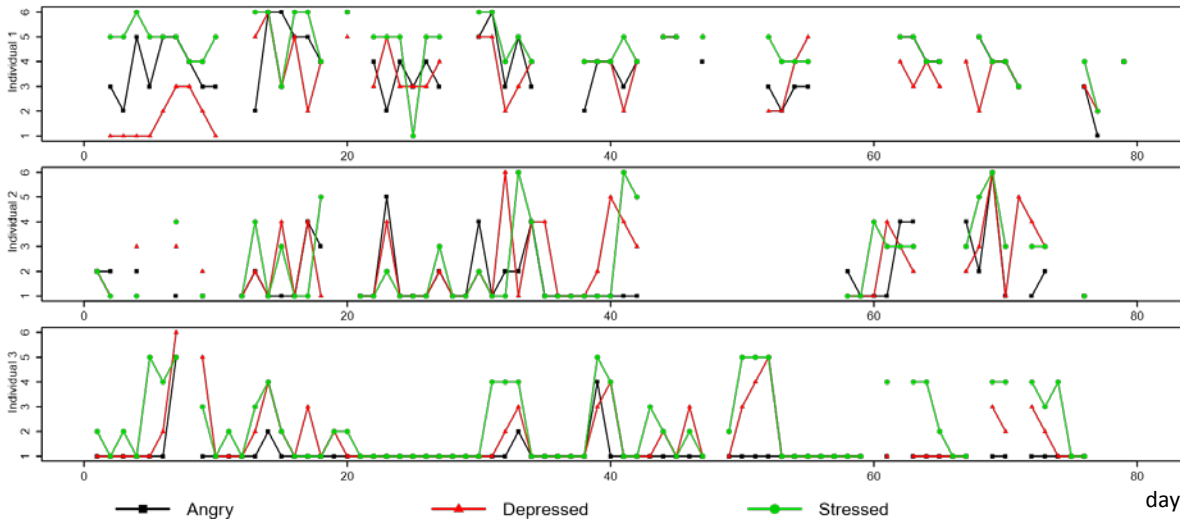

Figure 2. Observed time series for three of the individuals.

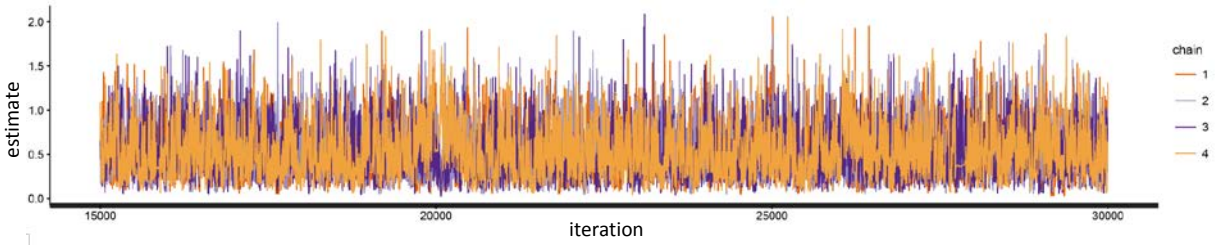

Figure 3. Trace plot of $Q_{11,1}$ as estimated in the empirical data, using six chains and 10.000 iterations (5.000 warm-up).

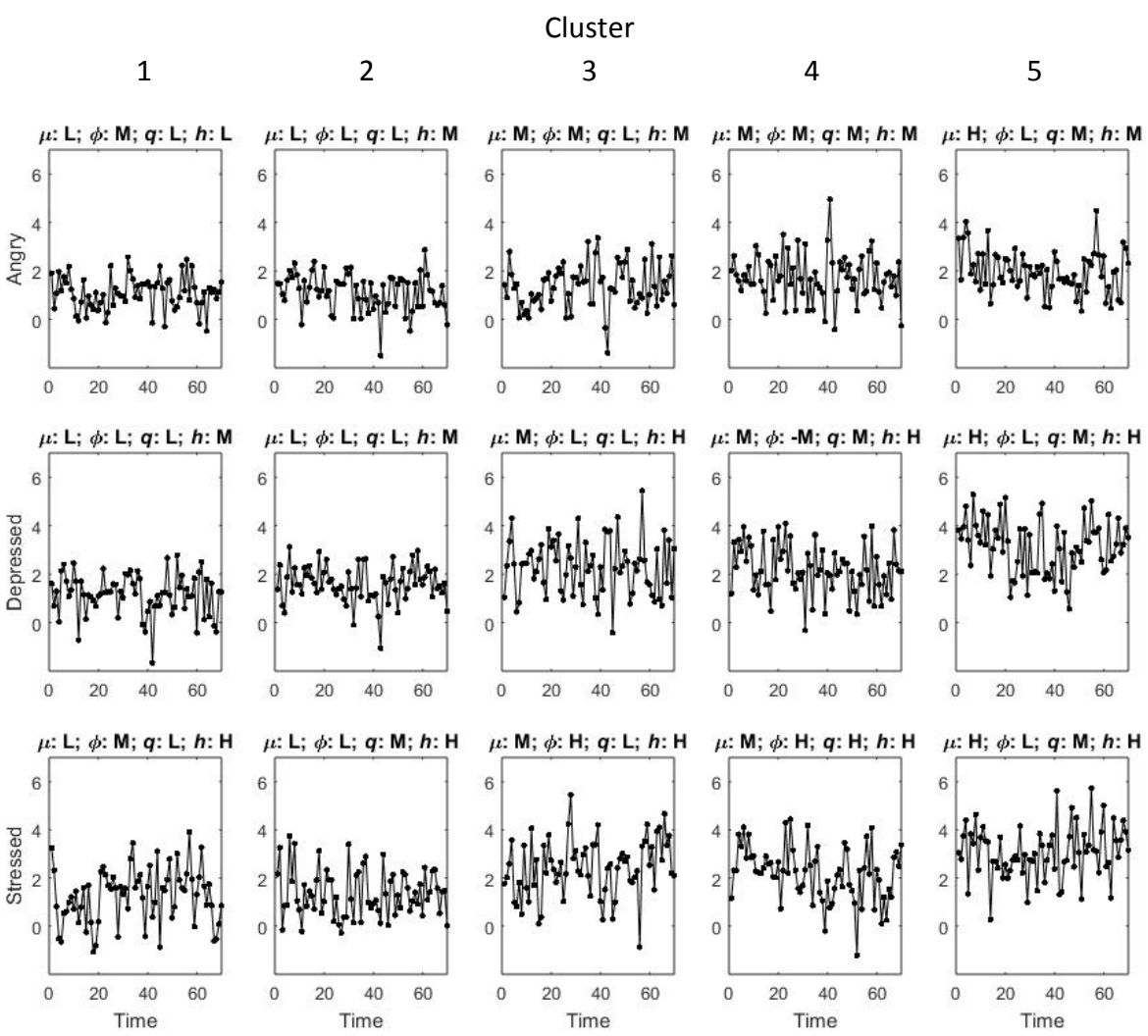

Figure 4. Typical time series for the emotions angry, depressed and stressed (in rows), for each of the 5 clusters (in columns) of the K-means analysis, using the parameter means per cluster as parameters of the VAR(1)-BDM.

μ indicates intensity, φ autoregression, q innovation variance, h white noise variance; Size of parameter (relative to other variables and clusters) is classified as Low (L), Medium (M), High (H).

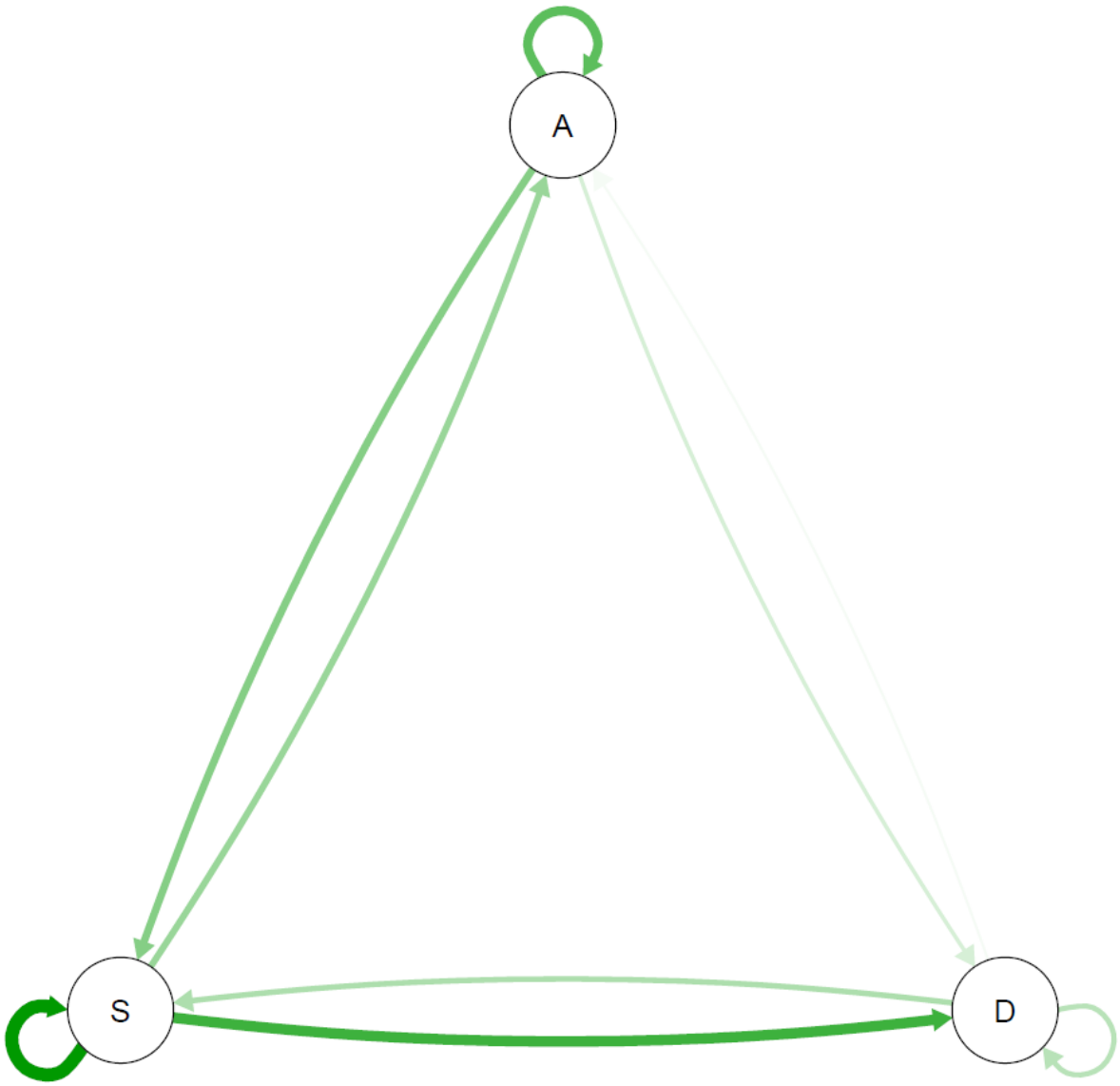

Figure 5. Network visualization of the auto- and cross-regressive effects for the average of the 50 individuals. A, D, and S stand for Angry, Depressed and Stressed, respectively. The thickness of the arrows is proportional to the size of the effect. The thickness of the curves is proportional to the value. Visualization constructed using the R-package qgraph (Epskamp, Cramer, Waldorp, Schittmann, Borsboom, 2012).

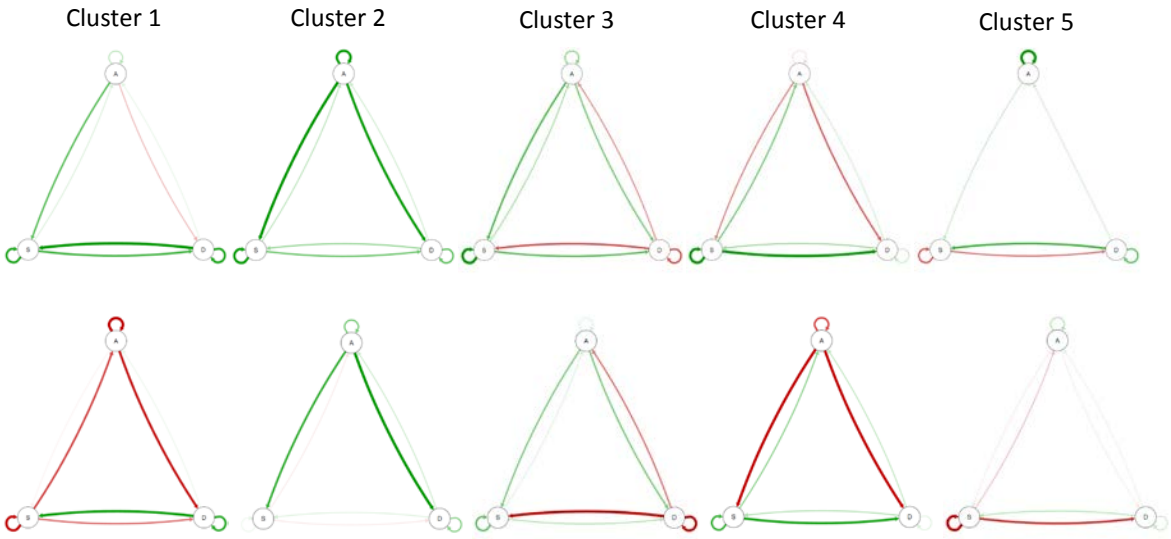

Figure 6. Network visualization (constructed using the qgraph package) for each of the five clusters. Top-row: visualization of the **Φ**-matrix per cluster; bottom-row: visualization of the deviation of $\mathbf{\Phi}_c$ (c = 1, ..., 5) from the overall average of **Φ**. A, D, and S stand for Angry, Depressed and Stressed, respectively. Green arrows indicate positive relations, red negative ones. Thickness of arrows is proportional to size of effects.

Appendix A

R code

```
require(rstan)
##take note: need to install Rtools first(see http://mc-stan.org/)
rstan_options(auto_write=TRUE)
options(mc.cores = 5)

#create data
data <- list()
data$N <- 1 #subjects
data$T <- 20 #maximum time points
data$I <- 2 #number of emotions
data$T_N <- array(data=20, dim=1) #time points per individual
#Missing value indicator
data$M<-array(data=c(1,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,0,
                     1,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,0),dim=c(N,T,I))
#Observed scores
data$Y<-array(data=c(1,3,2,5,3,5,5,4,3,3,1,1,2,6,6,5,5,4,1,6,1,1,1,1,1,2,3,3,2,1,1,1,5,6,3,5,2
,4,1,5),
              dim=c(N,T,I))


#call model file from working directory
Mod <- stan_model(file="Stancode_full_model.stan")

#sample model
fit <- sampling(
Mod
, data = data #The data list
, iter = 300 #This number should be large enough to reach convergence (i.e., 20.000 or such)
, chains = 3 #number of MCMC chains
, verbose = F
, refresh = 300 #update on every 300 iterations
, seed = 2016 #seed to replicate
)
```

Appendix B

Stan code

```
data{
  int <lower=2> T;      //maximum length of series across all individuals
  int <lower=1> I;      //number of emotions
  int <lower=1> N;      //number of individuals
  matrix [T,I] Y [N];   //observed scores in N matrices of T*I
  int M[N,T,I];         //m==1 if missing, m==0 is non-missing
  int T_N[N];           //length of individual time series
}

parameters{
  matrix <lower=-1, upper=1> [I,I] phi [N]; //cross-lag matrix, values restricted to (-1,1)
  corr_matrix[I] omega [N];                 //correlation matrix Q
  vector<lower=0>[I] tau [N];               //scale matrix Q
  row_vector [I] mu [N];                    //mean vector
  matrix [T, I] Z [N];                      //latent scores y*
  matrix <upper=0> [I,I] X [N];             //constant for prior of phi
  vector<lower=0>[I] lambda [N];            //vector matrix Q
}

transformed parameters{
  matrix [I,I] H[N];                        //covariance matrix of white noise
  matrix[I,I] Q[N];                         //covariance matrix of innovation
  for (n in 1:N){
  //quad_form=diag(tau)*omega*diag(tau)', diag(tau) is a diagonal matrix of vector tau
  H[n] <- quad_form_diag(omega[n],tau[n]);
  Q[n] <- diag_matrix(lambda[n]);
  }
}

model{
  for (n in 1:N){
    for (t in 2:T_N[n]){
    //system equation: estimate latent score theta
    Z[n,t] ~ multi_normal(Z[n,t-1]*phi[n], Q[n]);
        //link to vector Y_t of individual n if Y is observed for all emotions:
        if (sum(M[n,t])==0){
        //observation equation: estimate y (identity link function omitted)
        Y[n,t] ~ multi_normal(mu[n]+Z[n,t], H[n]);
        }
    }
  }
#prior
  for (n in 1:N){
  mu[n] ~ normal(0,2);
  tau[n] ~ cauchy(0,2.5);
  omega[n] ~ lkj_corr(2);
  lambda[n]~gamma(3,3);
    for (i in 1:I){
        for (j in 1:I){
        //symmetrized reference prior for phi
        increment_log_prob(X[n,i,j] - log(1-(phi[n,i,j]*phi[n,i,j]))/2);
        }
    }
 }
}
```