

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

Improved Insight into and Prediction of Network Dynamics by Combining VAR and Dimension Reduction

Kirsten Bulteel, Francis Tuerlinckx, Annette Brose & Eva Ceulemans

To cite this article: Kirsten Bulteel, Francis Tuerlinckx, Annette Brose & Eva Ceulemans (2018) Improved Insight into and Prediction of Network Dynamics by Combining VAR and Dimension Reduction, Multivariate Behavioral Research, 53:6, 853-875, DOI: 10.1080/00273171.2018.1516540

To link to this article: https://doi.org/10.1080/00273171.2018.1516540

View supplementary material

Published online: 19 Nov 2018.

Submit your article to this journal

Article views: 61

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hmbr20

MULTIVARIATE BEHAVIORAL RESEARCH
2018, VOL. 53, NO. 6, 853–875
https://doi.org/10.1080/00273171.2018.1516540

Routledge
Taylor & Francis Group

Check for updates

Improved Insight into and Prediction of Network Dynamics by Combining VAR and Dimension Reduction

Kirsten Bulteel[a], Francis Tuerlinckx[a], Annette Brose[a,b], and Eva Ceulemans[a]

[a]KU Leuven, University of Leuven; [b]Humboldt University Berlin

ABSTRACT

To understand within-person psychological processes, one may fit VAR(1) models (or continuous-time variants thereof) to multivariate time series and display the VAR(1) coefficients as a network. This approach has two major problems. First, the contemporaneous correlations between the variables will frequently be substantial, yielding multicollinearity issues. In addition, the shared effects of the variables are not included in the network. Consequently, VAR(1) networks can be hard to interpret. Second, crossvalidation results show that the highly parametrized VAR(1) model is prone to overfitting. In this article, we compare the pros and cons of two potential solutions to both problems. The first is to impose a lasso penalty on the VAR(1) coefficients, setting some of them to zero. The second, which has not yet been pursued in psychological network analysis, uses principal component VAR(1) (termed PC-VAR(1)). In this approach, the variables are first reduced to a few principal components, which are rotated toward simple structure; then VAR(1) analysis (or a continuous-time analog) is applied to the rotated components. Reanalyzing the data of a single participant of the COGITO study, we show that PC-VAR(1) has the better predictive performance and that networks based on PC-VAR(1) clearly represent both the lagged and the contemporaneous variable relations.

KEYWORDS

Network modeling; vector autoregressive modeling; multicollinearity; lasso; principal components; single-case

Psychological research increasingly focuses on how processes such as affect or stress regulation unfold at the within-person level. An important reason for this shift in focus is the availability of time series data for single individuals due to technological advances (e.g., smart phones for experience sampling studies; wearable devices for physiological measurements; see Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015). Time series data refer to many repeated measurements (often >50 measurements) of one or more variables (also termed as intensive longitudinal data; Bolger & Laurenceau, 2013; Walls & Schafer, 2006).

Although time series data are often gathered for multiple persons, we focus, in this article, on how to analyze the time series data of a single person. In addition and in line with most research on within-person psychological processes, we treat the data as being stationary (requiring the data to have constant mean values, (co)variances, and autocorrelation, without any cycles or trends; Liu & West, 2015). In addition, we will assume that the time intervals between the measurements are (roughly) equal although we will also indicate what to do if this assumption is violated.

A remarkable example of such time series data stems from the COGITO study (Schmiedek, Lövdén, & Lindenberger, 2010). A sample of 101 younger and 103 older adults was tested on approximately 100 days. The core data set of the COGITO study is composed of performance data on 12 cognitive tasks; other constructs were also measured such as task-related motivation (Brose, Schmiedek, Lövdén, Molenaar, & Lindenberger, 2010), positive and negative affect (Brose, Voelkle, Lövdén, Lindenberger, & Schmiedek, 2015), and physical health (Wolff, Brose, Lövdén, Tesch-Römer, Lindenberger, & Schmiedek, 2012). Given its intensive longitudinal design, the study can be used to gain insight into within-person dynamics (e.g., Brose, De Roover, Ceulemans, & Kuppens, 2015). Between-person analyses only particularize to the within-person level under very stringent and rather unlikely conditions, which makes it necessary to analyze within-person time series data to truly

CONTACT Eva Ceulemans eva.ceulemans@kuleuven.be Faculty of Psychology and Educational Sciences, KU Leuven, University of Leuven, Tiensestraat 102, Box 3713, Leuven 3000, Belgium.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hmbr.
Supplemental data for this article can be accessed here.
© 2018 Taylor & Francis Group, LLC

grasp within-person processes (Browne & Nesselroade, 2005; Molenaar, 2004; Molenaar & Campbell, 2009). In this manuscript, we will analyze data of a single, depressed participant of the COGITO study to shed light on the dynamics of her affective states as well as her depressive symptoms.

Despite the recent interest in and increasing availability of within-person time series data, analysis methods that model the dynamics present in these data are still under development (Hamaker et al., 2015). Among the proposed methods for stationary data are continuous-time modeling (Voelkle, Oud, Davidov, & Schmidt, 2012), dynamic factor analysis (Molenaar, 1985), and coupled latent differential equations (Boker, Neale, & Rausch, 2004). At this moment, however, the most popular method for representing stationary processes is to fit a lag-one vector autoregressive (VAR(1)) model (e.g., Bringmann et al., 2013, 2016; Pe et al., 2015; van der Krieke et al., 2015; Wichers, 2014; Wigman et al., 2015). This method is so popular because it is very easy to fit and interpret: it consists of a set of regression equations in which each variable is regressed on all variables (including itself) at the previous time point. Within the dynamic network approach (Bringmann, Lemmens, Huibers, Borsboom, & Tuerlinckx, 2015; Wichers, 2014), the resulting VAR(1) coefficients are displayed in a network figure in which the variables are the nodes and the VAR(1) coefficients constitute the directed links or edges between the nodes. The dynamic networks can be analyzed further by computing characteristics such as centrality measures. These measures aim to single out the most important nodes of the network (Newman, 2010, Chapter 7). This way, one tries to identify which variables strongly affect multiple other variables when activated, which is the important information when designing interventions (Borsboom & Cramer, 2013; Bringmann et al., 2015).

VAR(1)-based networks typically involve a relatively large number of variables leading to two problems: multicollinearity and an increased risk of overfitting (Babyak, 2004). Regarding multicollinearity, (subsets of) the variables can often be expected to be substantially correlated. As an example, in a study of the emotion networks of persons with and without major depressive disorder (MDD; American Psychiatric Association, 2013), participants were asked to rate the items 'happy' and 'excited' to measure their current level of positive affect (Pe et al., 2015). These items will be highly correlated across time. This high correlation is not a consequence of strong time–dynamical interaction, but rather of large semantic overlap (which is often unavoidable when researchers design questionnaires with good psychometric properties). Such substantial correlations and the resulting multicollinearity (which can be quantified in several ways, see e.g., Cohen, Cohen, West, & Aiken, 2003, Chapter 10) are problematic in regression analysis. Multicollinearity leads to highly unstable regression estimates (so-called bouncing betas), which are difficult to interpret (Cohen et al., 2003; Kiers & Smilde, 2007). Furthermore, regression coefficients capture only the unique direct effects of the variables, ignoring shared effects. Thus, in case of substantial correlations between the variables, the shared dynamics are not included in the VAR(1)-based networks.[1] As a consequence, computing network characteristics on the basis of these VAR(1) networks may be misleading, and may not help to identify which variables should be targeted first in an intervention (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016b). Furthermore, multicollinearity also adds to the overfitting problem, which we discuss next.

As shown by Bulteel, Mestdagh, Tuerlinckx, and Ceulemans (in press), the highly parametrized VAR(1) model (i.e., in case of J variables, there are J^2 lagged effects to estimate) is prone to overfitting. Overfitting occurs when the VAR(1) coefficients not only capture the lagged effects present in the data but also part of the error (Babyak, 2004; McNeish, 2015), because there is an insufficient amount of data compared to the number of VAR(1) parameters. Overfitting is problematic, because the VAR(1) coefficients do not properly describe the dynamic process of the individual, implying that it is not meaningful to use the associated network figures and characteristics for designing interventions. Overfitting can be prevented by evaluating the predictive performance of the model on a new data set, which was not used to compute the model estimates. In the absence of a second independent data set for that same individual, predictive accuracy can be assessed by means of K-fold crossvalidation (CV; Hastie, Tibshirani, & Friedman, 2009). In K-fold CV, the available data are divided into K equally sized subsets. Each of these subsets is consecutively considered as a test set for model testing, using the remaining subsets as the training set for model fitting. Finally, a predictive accuracy measure is computed between the observed and the predicted values of the test set and averaged over all the test sets.

[1] In Bulteel, Tuerlinckx, Brose and Ceulemans (2016b), we propose the use of relative importance metrics as a solution to include shared effects. As it does not resolve the other problems identified here, we do not consider this approach in this article.

Assessing the predictive accuracy ensures that we are not capitalizing on chance (Diaconis, 1985), which is a prerequisite for formulating sound scientific explanations (Breiman, 2001; Shmueli, 2010). We emphasize that we use CV in this article to examine whether we have sufficient data to estimate the VAR(1) coefficients, rather than whether or not this model is correct.

The multicollinearity issue and the risk of overfitting are well-known problems in the literature on regression analysis (e.g., Cohen et al., 2003). Two broad classes of statistical solutions have been put forward: penalization and dimension reduction. The aim of this article is to compare the pros and cons of both types of solutions, by assessing the extent to which they improve interpretability and predictive accuracy of VAR(1)-based network analysis. To this end, we will analyze a simulated toy data set and two COGITO data sets (affective states and depressive symptoms) of one COGITO participant.

Regarding penalized regression methods, we focus on the least absolute shrinkage and selection operator (lasso; Tibshirani, 1996) as this is the only penalty that has already been applied in the context of VAR(1)-based network modeling of within-person psychological processes (Epskamp, Waldorp, Mõttus, & Borsboom, 2018; Wild et al., 2010). The lasso shrinks the regression coefficients and even sets some of them to zero, yielding sparse and therefore easier to interpret networks (i.e., such networks have fewer edges). Furthermore, the shrinkage and variable selection also reduces the risk of overfitting (Babyak, 2004; McNeish, 2015). However, this lasso VAR(1) approach also has some disadvantages. First, sparseness may be an incorrect assumption because in reality the data may not be sparse. However, in such cases applying the lasso will yield a sparse network (Epskamp, Kruis, & Marsman, 2017). In addition, in case of groups of highly correlated variables, the lasso is known to arbitrarily select only some of them (Bondell & Reich, 2008; Tibshirani, 2011; Zou & Hastie, 2005) which makes proper interpretation of the VAR(1) coefficients problematic. Third, the shared effects of the variables are still not included in the network figure.

In contrast to the lasso, dimension reduction methods have not been rigorously adopted so far to improve VAR(1)-based network analysis. We will explore the benefits of this approach for studying within-person dynamics. More specifically, we propose to use principal components VAR(1) (or PC-VAR(1)), which extends the key principles of principal component regression (PCR; Jolliffe, 1982) to the VAR(1) case. PC-VAR(1) consists of two easy-to-conduct steps. First, the variables are reduced to a smaller set of components by means of principal component analysis (PCA; Jolliffe, 2002; Meredith & Millsap, 1985; Pearson, 1901). Applying a simple structure rotation to the obtained components provides a reduced dimensional representation of the contemporaneous correlation structure. The lagged relations, that is, the time dynamics, among the variables are shown in the second step by conducting a VAR(1) analysis on these rotated component scores. The obtained VAR(1) results can be visualized in a network figure, using the components as the nodes. These nodes are connected through directed arrows, yielding a representation of the crosslagged relationships among the components. It should be noted that these edges will reflect the dynamics almost completely as the shared effects of the components on one another can be expected to be small in most cases. In addition, the observed variables are also displayed as nodes, and are connected to the components on which they have their primary loadings, representing the contemporaneous correlations. It should be noted that the VAR(1) step can be replaced by its continuous-time extension, to handle violations of the assumption of equal measurement intervals. We will illustrate this alternative approach when analyzing the COGITO data. We focus on VAR(1) when introducing the dimension reduction approach, because the continuous-time extension is more difficult to fit and to interpret, requiring transformations of the parameter estimates (Ryan, Kuiper, & Hamaker, in press), and is therefore less likely to be applied by users.

We expect CV results and predictive accuracy to improve due to the dimension reduction step for two reasons: First, the VAR(1) analysis is performed on a small number of components rather than on all the observed variables. This substantially reduces the number of parameters that have to be estimated. Second, because the component scores are weighted sums of the original variables, the measurement error is reduced, which, in turn, will reduce overfitting. Our emphasis on this overfitting and CV perspective accounts for the reason why we utilized PCA in the dimension reduction step instead of exploratory factor analysis (EFA). To the best of our knowledge, it has not been clearly studied how to perform *K*-fold CV in EFA. Yet, as we will present in a later section, replacing PCA by EFA yields networks that are very similar in interpretation. It should be noted that the complete PC-VAR(1) method (both dimension reduction and regression) is closely related to a specific dynamic

factor analysis approach, that is, exploratory process factor analysis (EPFA). We will return to the comparison of PC-VAR(1) and EPFA in the discussion section.

The remainder of this article is organized as follows. First, we discuss the VAR(1)-based dynamic network approach, focusing on the issues identified above, multicollinearity, and overfitting, as well as the lasso and dimension reduction-based solutions. We illustrate the behavior of the different approaches initially by means of a simulated toy data set. Then, in the next section, we further demonstrate the differences among the methods by applying them to data from one of the participants of the COGITO study. We will analyze two time series data sets from this single participant, reflecting her affective states and her depressive symptoms. We conclude the article with a discussion of some limitations of the present approaches and provide directions for future research.

Methods

In this section, we discuss each of the methods introduced in the previous section and illustrate them by means of a toy example. This example data set is simulated in accordance with a PC-VAR(1) model. This way we get a first indication of how the previously proposed methods address this kind of data. The data set consists of six observed variables, each measured at $T = 500$ equidistant occasions. Recall that the standard VAR(1) approach builds on this equidistance, but that strong violations of this assumption can be handled through the use of continuous time approaches as will be discussed below. The generated scores on the six variables are based on three underlying components. The first three variables load on the first component, the fourth variable on the second component, and the last two variables on the third component. The scores of the time points on these components follow a VAR(1) process, thus yielding stationary data (for a formal definition, see Lütkepohl, 2005). The values of the loadings and VAR(1) coefficients for the resulting network are shown in Figure 1.[2] The loadings are indicated by the gray-directed arrows connecting components and variables, and the VAR(1) coefficients are displayed as green-directed arrows between the components. Randomness is introduced in the models through the dynamic errors or innovations, which were randomly sampled from a multivariate normal distribution with zero means and a covariance matrix with ones on the diagonal and 0.2 as off-diagonal elements (hence inducing contemporaneous correlations). To remove the effect of the starting values on the VAR(1) process, we generated 1500 observations, and discarded the first 1000. In addition, the final scores on the six variables were slightly perturbed by random noise, drawn from a standard normal distribution. The noise was rescaled such that the expected noise variance equals 5%. We use a low-noise percentage because the toy example serves a didactical purpose only. Complete results of the analysis of the toy example with 50% expected noise variance can be found in the Supplemental Materials. The conclusions are similar to the ones for the reported toy example with 5% expected noise variance. R code to generate the toy example data set and to apply the methods under study can be found in the Supplemental Material.

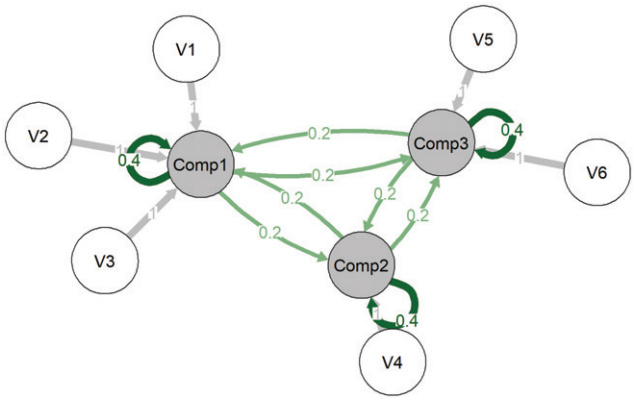


Figure 1. The parameters used to generate the toy example data set. The labels on the edges connecting a variable and a component display the component loadings. The labels on the edges connecting components show the crossregressive VAR(1) coefficients. The semicircular arrows represent the autoregressive VAR(1) effects.

Before analyzing the data, we standardized the scores on each variable. As we are interested only in the lagged covariance structure of the data, and not in the mean levels, the implied centering simplifies notation. In addition, scaling the centered variables ensures comparability of the VAR(1) coefficients (Bulteel et al., 2016b).

Var(1)

Model

In a VAR model of order 1, we model J time series of length T. For example, for our toy data, J equals 6 and $T = 500$. In a VAR(1), the $J \times 1$ vector of

[2]Similar graphs are drawn to visualize structural equation models although they are based on other conventions. In particular, manifest variables are displayed as squares, and latent variables as circles. Covariances and residuals are drawn as double-headed arrows and directed relations as single-headed arrows.

observations $\mathbf{y}_t$ at time point t is regressed on the $J \times 1$ vector of observations $\mathbf{y}_{t-1}$ at time point t–1:

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{y}_{t-1} + \mathbf{u}_t \quad (1)$$

where $\mathbf{\Phi}$ is a $J \times J$ matrix holding the VAR(1) coefficients, and the $J \times 1$ vector $\mathbf{u}_t$ contains the innovations of the variables at time point t (Hamilton, 1994; Lütkepohl, 2005).

The innovations represent what cannot be predicted based on the observations one time point earlier. They capture all internal and external events that occurred between time points t–1 and t and that influence the process under investigation (Hamaker, 2012). The innovations are assumed to follow a multivariate normal distribution with zero means and a covariance matrix $\mathbf{\Sigma}$. The innovations must be clearly distinguished from measurement error as their influence reaches beyond a particular point in time. The effect of the innovation at a particular occasion lingers on via the VAR(1) parameters (Schuurman, Houtveen, & Hamaker, 2015).

The diagonal elements of $\mathbf{\Phi}$ are the autoregressive effects, showing the effect of each variable on itself one time point later. The off-diagonal elements are the effects of one variable on another variable at the next measurement occasion (i.e., crossregressive or crosslagged effect). To ensure stationarity, the modulus of the absolute values of the eigenvalues of $\mathbf{\Phi}$ should be smaller than one.

In case of unequal measurements intervals, using the VAR(1) model can result in biased parameters (de Haan-Rietdijk, Voelkle, Keijsers, & Hamaker, 2017). A more appropriate model in this case is the continuous-time variant of the VAR(1) model (e.g., Oud & Delsing, 2010; Voelkle et al., 2012). In this continuous-time model, the exact duration of each time interval between two adjacent observations is taken into account when predicting the score at time point t based on that at time point t–1. In fact, the regular or discrete time VAR(1) model that we used can be viewed as a special case of the continuous-time model, in which time intervals are assumed to be equal between all observations. The parameters of the continuous-time VAR(1) model can always be transformed to regular VAR(1) coefficients for a specific time interval (Ryan et al., in press). The R package ctsem is available to estimate the parameters of such a model (Driver, Oud, & Voelkle, 2017).

Estimation

Different estimation procedures are available (see e.g., Hamilton, 1994; Lütkepohl, 2005). We use the least squares (LS) estimator. It should be noted that the covariance structure $\mathbf{\Sigma}$ is not directly estimated, but can be computed based on the regression residuals.

Crossvalidation

Standard CV techniques rely on the assumption that training and test sets are independent, an assumption that is violated for time series data. We use blocked CV (Snijders, 1988; as cited in Bergmeir & Benítez, 2012) following the recommendations of Bergmeir and Benítez (2012) and Bulteel et al. (in press). Blocked CV differs only from K-fold CV in its first step, in that the data are now divided into K equal-length subsamples of consecutive observations. One subsample is selected as the test set and the remaining subsamples as training set. The VAR(1) model is fitted to the training set and the resulting VAR(1) coefficients are used to predict the test set observations. These steps are repeated using each of the K subsamples once as test set. Then, we compute the mean squared error (MSE) between the actual observations and their predictions (each observation is used exactly once as part of a test set):

$$\mathrm{MSE}_{\mathrm{CV}} = \frac{1}{JT}\sum_{k=1}^{K} \|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2 \quad (2)$$

where K equals the number of test sets (typically, $K = 10$), $\mathbf{y}_k$ contains the $T_k \times J$ test set observations in the kth subsample, $\hat{\mathbf{y}}_k$ represents the predictions for these test set observations based on the training set, and with $\|.\|^2$ being the squared Euclidean norm.

Network

The VAR(1) model can be visualized in a network figure. The variables are the nodes and the edges are drawn based on the VAR(1) coefficients. The autoregressive effects are displayed as self-loops and the crossregressive effects as directed arrows connecting different nodes. We used the R package qgraph to plot the network (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). The positive VAR(1) weights are displayed as a green solid arrows, and negative weights as red dashed arrows. The edge width is relative to the size of the VAR(1) coefficient.[3] The thicker the edge, the larger the VAR(1) coefficient of this lagged effect. In addition, network characteristics can be computed. In this article, we will use the outstrength centrality measure as an example. The outstrength of a particular variable node is computed

[3]Similar graphs are drawn to visualize structural equation models although based on other conventions. In particular, manifest variables are displayed as squares, and latent variables as circles. Covariances and residuals are drawn as double-headed arrows, and directed relations as single-headed arrows.

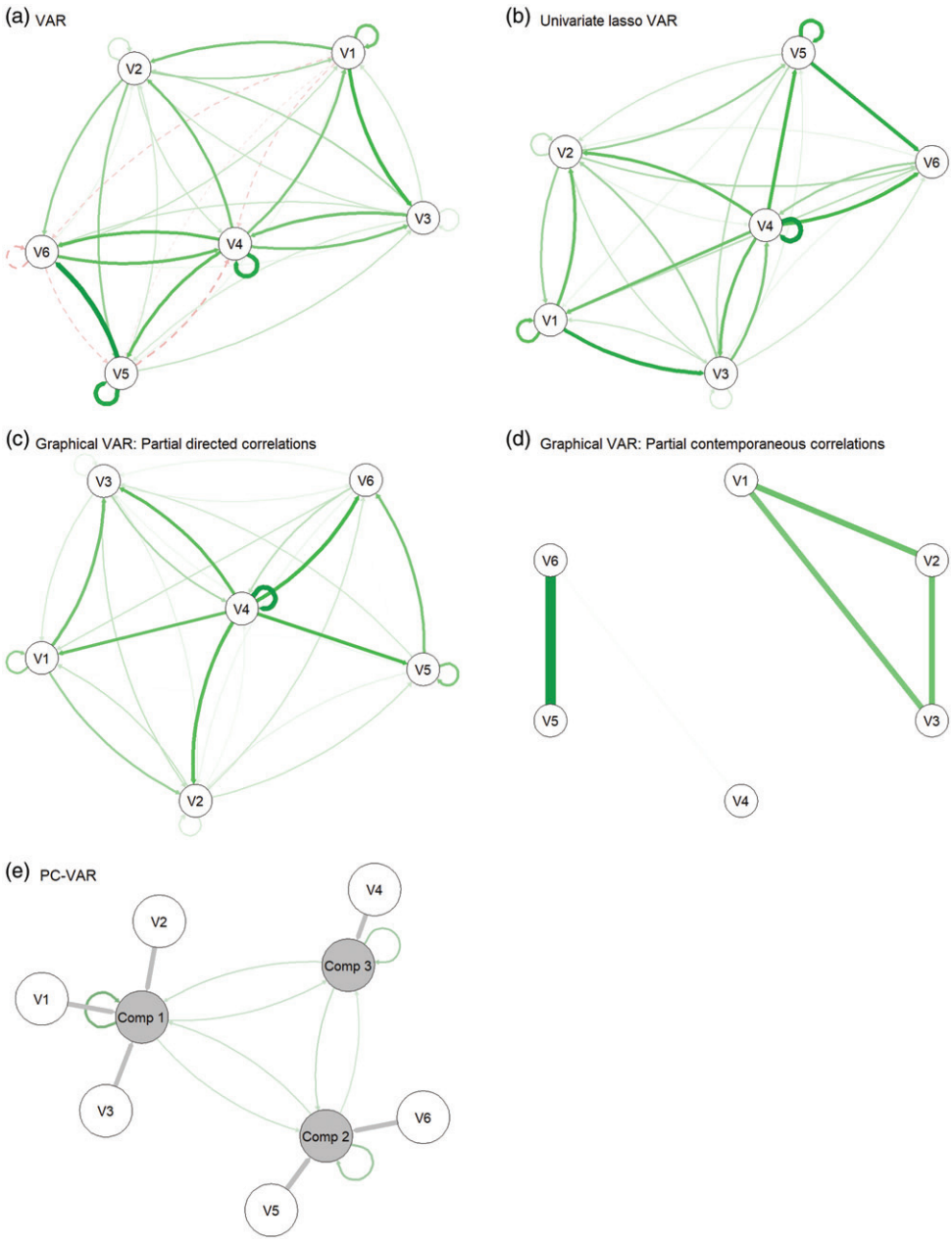


Figure 2. Network figures of the toy example depicting (a) the VAR(1) coefficients, (b) the univariate lasso VAR(1) coefficients, (c) the PDC, and (d) the PCC resulting from the graphical VAR(1) approach, and (e) the loadings and VAR(1) weights obtained by applying PC-VAR(1). In all panels, green solid edges indicate positive coefficients and red dashed edges negative coefficients. Additionally, the loadings are indicated by the gray-directed arrows connecting components and variables.

as the sum of the absolute values of the VAR(1) coefficients corresponding with the edges leaving from this node (Opsahl, Agneessens, & Skovetz, 2010). When the VAR(1) analysis is performed on standardized variables, the outstrength values within a network are directly comparable (Bulteel et al., 2016b).

Application to toy example

We applied the VAR(1) model to the six variables in the toy example. The resulting network figure is shown in Figure 2a. The six variables (labeled as 'V1'–'V6') are the nodes. At first glance, it is difficult to identify any structure in the variable dynamics due to the tangle of edges. In addition, it is notable that a few red dashed edges (negative links) are present which is inconsistent with the data-generating mechanism (Figure 1). Only the fourth and fifth variables have strong autoregressive effects as shown by the thick self-loops. We conjecture that the relatively weak autoregressive effect of the first three variables is due to the high correlations among these variables, which all have high loadings on the first component.

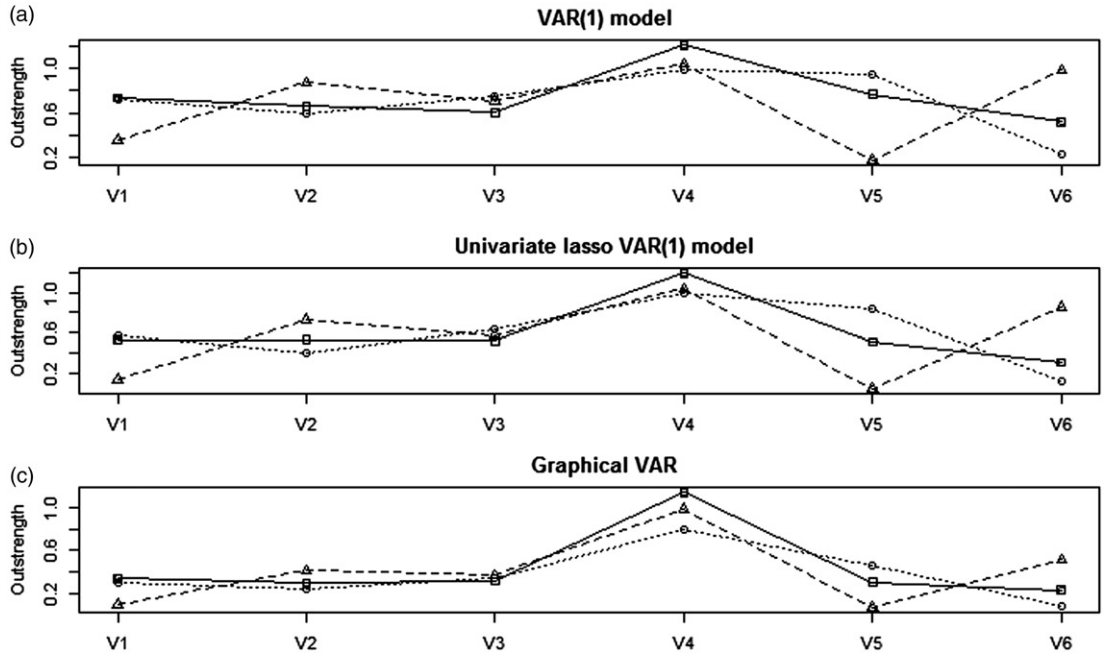


Figure 3. The outstrength values of the six variables of the toy example data sets. The values in panel (a) are based on the VAR(1)-based network, those in (b) on the univariate lasso VAR(1) model, and those in (c) on graphical VAR(1). The solid line presents the results for the first toy example data set, the dotted line for the second toy example, and the dashed line for the third toy example. The three toy example data sets were simulated using the same data generating mechanism.

Table 1. Regular VAR(1), univariate lasso VAR(1), and graphical VAR(1) results for three different data sets with 500 measurements (Toy examples 1–3), simulated using the same data generating mechanism (see main text and Figure 1).

		Criterion					
	Predictor	V1	V2	V3	V4	V5	V6
VAR(1)	V1				1/2/3	1	1
	V2					2	2
	V3					3	
	V4						
	V5	3	3	3	1/3		
	V6		2	2	2	1	1
Univariate lasso VAR(1)	V1				1/2/3	1	1
	V2					2	2
	V3					3	
	V4						
	V5	3	3	3	1/3		
	V6		2	2	2	1	1
Graphical VAR(1)	V1				1/2/3	2/3	1
	V2						2
	V3				3	3	
	V4						
	V5	2/3			1/3		
	V6		2	2	2		1

The number indicates for which of the three toy examples an unexpected VAR(1) coefficient was obtained, in that the value is smaller or equal to zero. Negative or zero values are inconsistent with the data generating process. The results differ from simulation to simulation precluding clear interpretation.

The strength of the self-loop of the fourth and fifth variables is identical to the ones of the second and third components used to generate the data. However, the autoregressive effect of the sixth variable, which also loads on the third component, is negative. The strongest crossregressive link (i.e., the thickest edge) is the one connecting the fifth variable with the sixth. Computing the outstrength, we see that the highest value is obtained for the fourth variable, followed by the fifth and the first variable (Figure 3a – solid line).

Due to multicollinearity, these VAR(1) weights are not stable. Table 1 summarizes which VAR(1) coefficients are negative for the toy data as well as for two additional data sets that were generated in the same way. As it can be seen, the results differ substantially across the three data sets although the data-generating mechanism is identical. This instability is also reflected in the outstrength plot in Figure 3a. For the two additional data sets, the fourth variable is again selected as the most central one according to the outstrength measure. In contrast, for the second data set, the fourth variable is followed by the fifth and the third (dotted line), whereas for the third data set, it is followed by the sixth and the second (dashed line). Furthermore, we computed the prediction error by means of blocked CV for the toy example data. The MSE averaged across the K-folds equals 0.5785 as can be seen in the right plot of Figure 4a. This CV-value acts as a benchmark for the predictive accuracy of the lasso VAR(1) models and the PC-VAR(1) approach.

Lasso VAR(1)

To address multicollinearity, a range of penalized regression approaches have been proposed, including

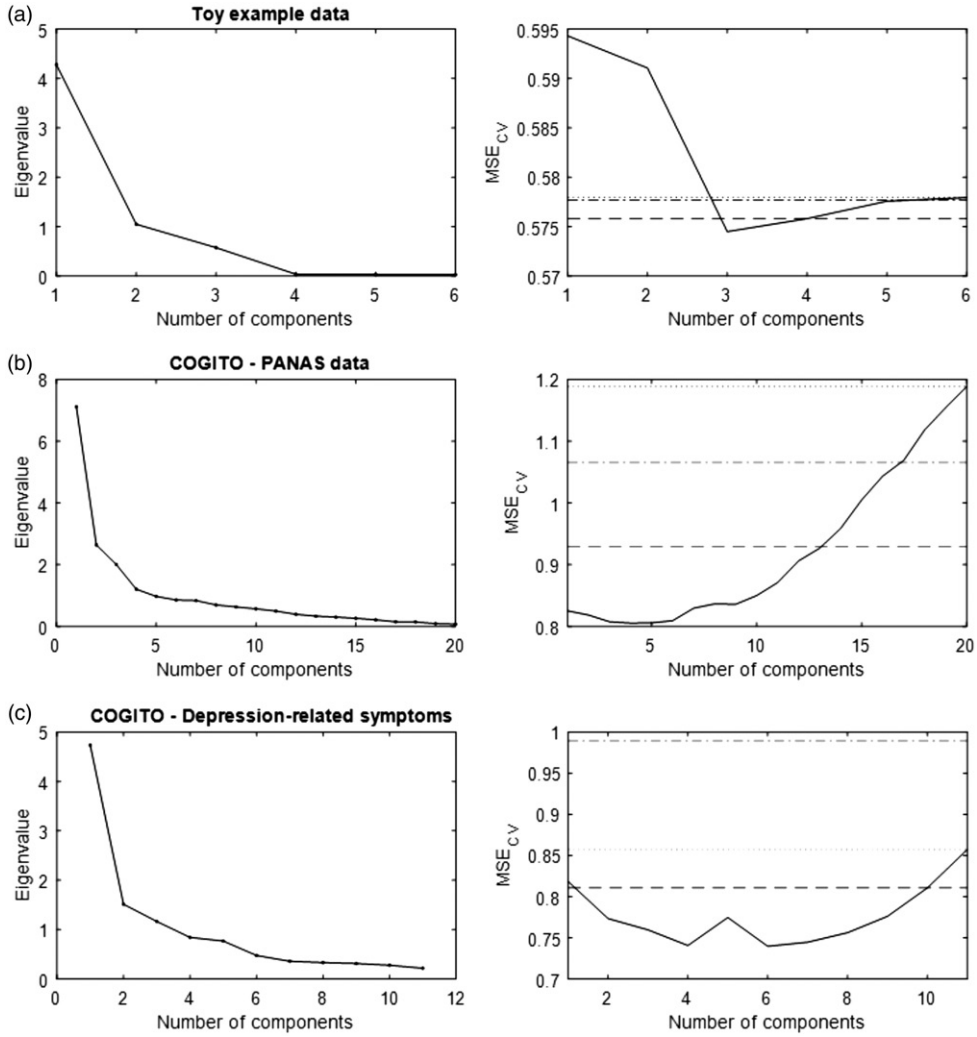


Figure 4. Scree and CV plots. (a) The plots for the toy example, (b) for the affective states data of the COGITO study, and (c) for the depression-related symptoms data of the COGITO study. In the CV plot, the dotted line depicts the MSE of the VAR(1) model, the dashed line depicts the MSE of the univariate lasso VAR(1) model, the dash–dot line depicts the MSE of the graphical VAR model, and the solid line depicts the MSE of the PC-VAR(1) approach. In the case of the PC-VAR(1) approach, MSE varies as a function of the number of components; the other MSE values are constant.

the lasso (Tibshirani, 1996), ridge regression (Hoerl & Kennard, 1970), and the elastic net (Zou & Hastie, 2005). In this manuscript, we focus on the lasso penalty because this approach has previously been adopted in network analysis (Epskamp et al., 2018; Wild et al., 2010).

In general, a lasso approach implies the same regression equations as the original model (in our case Equation (1)). A penalty is added to the LS loss function (or to the log likelihood function) when estimating the parameters to shrink some of them to zero. The penalty equals the sum of the absolute values of all or a subset of the parameters. The magnitude of the impact of the penalty can be varied by multiplying it by a tuning parameter. If the tuning parameter equals zero, the lasso VAR(1) model reduces to the standard VAR(1) model. The larger the tuning parameter, the more parameters will be forced to zero. In case of groups of strongly correlated predictors, the lasso tends to arbitrarily select one or a few predictors per group, which is clearly a disadvantage of this method (Bondell & Reich, 2008; Zou & Hastie, 2005). In this article, we discuss two variants of the lasso VAR(1) model: A univariate and a multivariate variant.

Univariate lasso

The standard VAR(1) model implies a separate regression equation for each criterion variable. The univariate lasso VAR(1) approach (Hsu, Hung, & Chang, 2008) tunes the lasso penalty for each of these regressions separately. The penalty term thus reflects the VAR(1) coefficients of that particular regression equation and the value of the tuning parameter λ can

differ across the regression equations (Rothman, Levina, & Zhu, 2010 termed this approach 'separate lasso'). To choose the optimal set of λs, we use blocked CV for each regression equation separately. As was the case for the standard VAR(1) model, the covariance structure of the innovations is computed afterward on the basis of the residuals. The network figure is built in the same way as for the standard VAR(1) model, but we expect fewer arrows because some of the VAR(1) coefficients are set to zero. Outstrength centrality measures can be computed in the same way as for the standard VAR(1) model.

To determine the predictive accuracy of this univariate lasso VAR(1) approach, blocked CV is applied to compute the MSE_{CV}. Given that we set the optimal λs by means of blocked CV as well, this implies that the blocked CV to determine the λ values for the training set is applied within the overall blocked CV procedure.

The obtained results for the toy example are shown in Figure 2b. The figure is similar to the standard VAR(1) picture. The counter-intuitive negative cross-lagged relations found with the standard VAR(1) model are set to zero in the univariate lasso VAR(1) model. In addition, the variables associated with the three underlying components in this simulated data set are partitioned into the three expected groups. In line with our expectations, at least one variable per group is not set to zero in each regression equation. However, which variable is selected is arbitrary as summarized in Table 1. This table shows which predictors were selected by the lasso approach for the toy data set and the two other data sets that were identically generated. Indeed, for each data set other VAR(1) weights are set to zero. Only the edge connecting the first variable to the fourth variable is systematically set to zero across the three data sets. The outstrength results in Figure 3b show that the outstrength ranking of the variables differs across the three data sets. The predictive accuracy of the univariate lasso VAR(1) model is shown in the right plot of Figure 4a. The MSE_{CV} equals 0.5763, which is slightly lower than the value for the regular VAR(1) model, suggesting that the univariate lasso VAR(1) model generalizes somewhat better to a new data set.

Multivariate lasso

The multivariate lasso VAR(1) (Abegaz & Wit, 2013; Rothman et al., 2010), or graphical VAR(1) model (Wild et al., 2010) further improves the predictive performance of the estimates by also imposing a penalty on the precision matrix $\boldsymbol{\Omega}$, the inverse covariance matrix of the innovations (i.e., $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$). Thus, two penalty terms are now added to the likelihood function: the sum of the absolute values of all VAR(1) coefficients, weighted with a tuning parameter λ_1, and the sum of the absolute values of the off-diagonal elements of the precision matrix $\boldsymbol{\Omega}$ weighed with a tuning parameter λ_2 (implying that there is no shrinkage of the variances of the variables). Even if the underlying precision matrix is not sparse, this method is still expected to outperform univariate VAR(1) lasso because the latter ignores the innovation covariances (Rothman et al., 2010). In contrast to the above methods, the graphical VAR(1) returns two sparse networks: One network is based on the lagged relations (similar to the previous methods), and one network is based on the contemporaneous relations (using the precision matrix). Specifically, the partial-directed correlations (PDC) and the partial contemporaneous correlations (PCC) values are computed and visualized. The PDC are obtained by rescaling the VAR(1) coefficients as follows:

$$\mathrm{PDC}(y_{i,t-1}, y_{jt}) = \frac{\Phi_{ij}}{\sqrt{\boldsymbol{\Sigma}_{jj}\boldsymbol{\Omega}_{ii} + \Phi_{ij}}}, \tag{3}$$

where the subscripts of the matrices indicate which element is selected. The first index refers to the row, and the second to the column. The PCC are computed as follows:

$$\mathrm{PCC}(y_{it}, y_{jt}) = -\frac{\boldsymbol{\Omega}_{ij}}{\sqrt{\boldsymbol{\Omega}_{ii}\boldsymbol{\Omega}_{jj}}}. \tag{4}$$

We estimate the graphical VAR(1) model by means of the R package graphicalVAR (Epskamp, 2016). In this package, the tuning parameters are selected on the basis of the extended Bayesian information criterion (Chen & Chen, 2008) using a grid search. The plot of the PDC and PCC values is automatically provided by the package. To obtain outstrength values that are comparable to those for the previous methods, we compute them on the basis of the PDC network. To compute the MSE_{CV} for the multivariate lasso, we use only the estimates of the lagged effects $\boldsymbol{\Phi}$ to predict the test set observations (when using the squared Euclidean norm as a measure of predictive accuracy, the contemporaneous relations do not influence the result). This also implies that MSE_{CV} is computed in the same way as for the standard VAR(1) model and the univariate lasso VAR(1) model.

The PDC network displaying the lagged effects for the toy example is shown in Figure 2c. In line with the standard VAR(1) and univariate lasso VAR(1)-based networks, the fourth variable is placed in the

middle of the PDC figure. In addition, the directed effects suggest that this variable is driving the other variables (the arrows starting at V4 and heading at the other variables). This suggestion is supported by the outstrength plot in Figure 3c (solid line). A possible explanation is that only one variable (i.e., V4) loads strongly on the second component, implying that the total effect of this component on the other variables does not have to be divided across multiple variables as it is the case for the other components. Again, which edges are set to zero seems to be arbitrary as is listed in Table 1. Only the edge connecting the lagged value of the first variable with the fourth variable is consistently estimated to be a zero value. In addition, the outstrength ranking of the variables differs across the three generated data sets (Figure 3c). As it can be seen in the right plot of Figure 4a, the MSE_{CV} for the multivariate lasso VAR(1) approach equals 0.5777, which indicates that the multivariate approach has a somewhat worse predictive performance than the univariate approach, but outperforms the standard VAR(1) model. The PCC network (Figure 2d), showing the contemporaneous partial coefficients, reveals the component structure in that the variables that load on the same component are strongly connected, whereas the components are clearly separated. It is noteworthy that this PCC network is stable across the three toy example data sets.

PC-VAR(1)

In the literature on multicollinearity in regression analysis, the use of dimension reduction techniques such as PCR offers an important alternative to the use of penalty functions (Kiers & Smilde, 2007). In a PCR analysis, the predictor variables are first reduced to a smaller number of components. In a second step, the criterion variables are regressed on these components, rather than on the predictor variables themselves (Jolliffe, 2002).

Given that in VAR(1) analyses the predictors are lagged versions of the criteria, we adapted the two PCR steps as follows, yielding PC-VAR(1) (for a similar approach, see Morana, 2012). In the first step, the variables are reduced to a smaller set of components to shed light on the contemporaneous correlation structure. Second, a VAR(1) analysis is performed using the components instead of the variables to capture the dynamics across time. As the second step is the same as the regular VAR(1) analysis explained above, we focus on the first step, the selection of the number of components, and possible alternative approaches in the following paragraphs.

Step one

PCA is applied to the predictor variables. In particular, the predictor scores (i.e., the within-person standardized $(T–1) \times J$ time series data) are decomposed as:

$$\mathbf{y}_{t-1} = \mathbf{B}\mathbf{f}_{t-1} + \mathbf{e}_{t-1} \tag{5}$$

where $\mathbf{B}$ $(J \times Q)$ is the loading matrix, $\mathbf{f}_{t-1}$ is a $Q \times 1$ vector containing the component scores of the observations at time point $t–1$ on the Q-retained components, and $\mathbf{e}_{t-1}$ is the residuals vector $(J \times 1)$ at time $t–1$ (Tsay, 2005). To identify the solution, the component scores are uncorrelated and have a variance of one. In this case, the loadings are equal to the correlation between the corresponding components and the variables.

In addition, we rotate the loadings $\mathbf{B}$ to simple structure, and recalculate the component scores accordingly:

$$\begin{aligned} \mathbf{B}_{\text{rot}} &= \mathbf{B}(\mathbf{M}')^{-1} \\ \mathbf{F}_{\text{rot}} &= \mathbf{F}(\mathbf{M}')^{-1} \end{aligned} \tag{6}$$

where $\mathbf{B}_{\text{rot}}$ and $\mathbf{F}_{\text{rot}}$ represent the loading matrix and component score matrix after rotation, and $\mathbf{M}$ is the $Q \times Q$ rotation matrix. The simplicity criterion that we maximize, produces a solution with two properties: (a) one high loading for each variable with the other loadings for each variable close to zero; and (b) components characterized by high loadings for a few variables and near zero loadings for the others (which is also termed a perfect cluster configuration; Browne, 2001). The goal is to yield easy-to-interpret components. In practice, the researcher can opt for an orthogonal or an oblique rotation. The advantage of an orthogonal rotation is that the resulting components, which will be subjected to a VAR(1) analysis in the next step, will be uncorrelated. The potential problem that the shared variance is not taken into account in a VAR(1) analysis (**Introduction** section), is no longer an issue (Bulteel et al., 2016b). The disadvantage is that an orthogonal rotation is often less realistic for psychological data because psychological constructs often covary or may even be causally related at the within-person level (e.g., Borsboom & Cramer, 2013). Oblique rotation therefore can provide more realistic results, but with the disadvantage that the shared variance between the components will not be reflected in the VAR(1) coefficients. In the following, we will maximize the varimax criterion for orthogonal rotation

(Kaiser, 1958), and the promax criterion for oblique rotation (Hendrickson & White, 1964).

Step two

A regular VAR(1) analysis is performed on the rotated component scores:

$$\mathbf{f}_t = \mathbf{\Phi}\mathbf{f}_{t-1} + \mathbf{u}_t \tag{7}$$

with $\mathbf{\Phi}$ being a $Q \times Q$ matrix containing the lagged associations between the components. It should be noted that the regular VAR(1) model is equivalent to the PC-VAR(1) approach when the number of components Q equals the number of variables J.

Selection of the number of components

A first approach is to inspect a scree plot (Cattell, 1966). In such a plot, the components are plotted on the x axis, and the corresponding eigenvalues on the y axis. To optimally balance fit and complexity, one searches for an elbow in the downward curve, indicating that additional components do not imply a substantial gain in information. The final number is set to the number right before the elbow. Often, multiple elbows are present in the scree plot, however. Because we aim to describe the dynamics of the initial set of variables, we recommend selecting a sufficient number of components in such cases to retain enough information in the extracted components. Ideally, one should retain the major as well as the minor components, and exclude noise components (i.e., the measurement error; Ceulemans & Kiers, 2009). Vervloet, Van Deun, Van den Noortgate, and Ceulemans (2016) demonstrated that one risks removing components necessary to explain the criterion variables properly when including only the major components. Therefore, we suggest looking at the last elbow in case of multiple elbows. The CHull criterion and software (Ceulemans & Kiers, 2006; Wilderjans, Ceulemans & Meers, 2013) can be used to compute the elbows in an automated way.

A second approach to decide on Q is to use CV, and compute the MSE_{CV} for the test set data. MSE_{CV} should be computed for each solution for the number of components ranging from 1 to the number of variables J. We then select the Q-value that minimizes this MSE value. To compute the MSE_{CV} for PC-VAR(1), we take into account that the component scores $\mathbf{F}_{t-1}$ are obtained as weighted sums of the observed variables: $\mathbf{F}_{t-1} = \mathbf{Y}_{t-1}\mathbf{W}\sqrt{T}$. The $J \times Q$ weight matrix $\mathbf{W}$ can be derived as follows:

$$\mathbf{W} = (\mathbf{Y}_{t-1})^{-1}\mathbf{U} \tag{8}$$

where $(\mathbf{Y}_{t-1})^{-1}$ is the Moore–Penrose pseudoinverse of the predictor scores matrix $\mathbf{Y}_{t-1}$, and $\mathbf{U}$ contains the Q left singular vectors that result from a singular value decomposition of $\mathbf{Y}_{t-1}$. Hence, when conducting CV, we compute the component scores of the test set using the same weight matrix and rotate them identically to the training data. Next, the predictions for the test set observations are calculated, by multiplying the rotated component scores by the VAR(1) coefficients matrix $\mathbf{\Phi}$ and the transpose of the loading matrix.

Other approaches to selecting the optimal number of components are possible, for instance, the use of parallel analysis (Crawford et al., 2010; Horn, 1965; see however the cautionary remarks by Saccenti & Timmerman, 2017; Vitale et al., 2017), CHull (Lorenzo-Seva, Timmerman, & Kiers, 2011), or both. Golino and Epskamp (2017) recently proposed exploratory graph analysis to estimate the number of components, which is based on a graphical lasso approach (equivalent to the PCC network presented above).

Alternative methods in both steps

In both steps of the PC-VAR(1) procedure, it is possible to use alternative methods. In the first step, one could use EFA instead of PCA to perform dimension reduction. As PCA decomposes the complete variance–covariance matrix, whereas EFA focuses on the common variance among the variables, discarding their unique variances, the loadings for PCA will be slightly higher compared to the loadings for EFA (Velicer & Jackson, 1990). However, given that the interpretation of the components and factors is in most cases very similar (i.e., relative differences between the loadings are comparable across both methods), we conjecture that the component scores and factor scores will typically be almost identical. We verified this for the toy example. Using the weighted least-squares method to compute the factor scores, the correlations between the component and the factor scores were >0.99. The subsequent VAR(1) analysis takes only these component or factor scores into account, not the loadings. Therefore, we expect the differences in obtained parameter estimates between both approaches will be minimal in practice. We present the PCA approach in this article because the application of CV is less straightforward in the context of EFA. The indeterminacy of factor scores presents a challenge as several potential methods can be used to estimate the factor scores (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Grice, 2001). In addition, it is unclear how the factor scores for new data, required to perform CV, should be computed.

In the second step of the PC-VAR(1) procedure, the regular VAR(1) analysis can also be replaced by alternative dynamic models or estimation procedures. As discussed in an earlier section, the continuous-time variant of PC-VAR(1) is useful in case of irregular measurement intervals. To estimate the continuous-time VAR(1) parameters, we can specify the loadings obtained from the PCA analysis as input parameters in the ctsem package.

Network

In the network representation of PC-VAR(1), we include the variables as well as the components as nodes in the network, similar to the representation in structural equation models (for the differences see Footnote 2). The component nodes are colored gray to distinguish them from the variable nodes. The component loadings, describing the contemporaneous correlation structure, are displayed as gray solid edges connecting the variables with the components on which they have the highest loading. The edges connecting the components describe the crosslagged relations which are displayed as in the standard VAR(1) network figure. Such a network can also be created with the R package qgraph. The network characteristics of a PC-VAR(1) network are computed at the level of the components. If a certain component has a high outstrength, this result suggests that one should consider intervening on all the variables that load strongly on that component.

Application to toy example

The PC-VAR(1) network of the toy example is shown in Figure 2e. To select the number of components, we inspected the scree plot (left plot in Figure 4a), and the CV plot (right plot in Figure 4a). In the scree plot, two elbows are clearly visible at two components and at four components. In line with the above recommendations, we focus on the last elbow selecting three components. In the CV plot, the MSE_{CV} is minimized when extracting three components (i.e., 0.5712). Hence, the scree and CV plots both point toward using three components. These three components explain 98% of the variance in the predictor variables. Furthermore, the right plot in Figure 4a shows that the prediction error of PC-VAR(1) is clearly lower than the MSE_{CV} values of the previous methods when three or more components are retained. This result implies that PC-VAR(1) has the better predictive performance. As noted above, the MSE_{CV} value for the VAR(1) model equals the value of the PC-VAR(1) when retaining all extracted components in the analysis. To facilitate the interpretation of the network figure, we plot only the highest loading for each variable. The underlying component structure is clearly revealed as the loadings closely reflect the ones used to generate the data. Also, the VAR(1) coefficients for the components resemble those used to generate the data. Summarizing the network figure, we see three groups of variables that each comprises a component. These three groups have similarly strong autoregressive and crossregressive effects as could be expected based on how the toy example data are simulated. For completeness, we also analyzed the two additional data sets simulated using the same data generating model (Table 1). The results (data not shown) were similar for the three data sets. We also analyzed the original data set used in the toy example using EFA instead of PCA in the first step of the procedure. As expected, we obtained highly similar results as shown in the Supplemental Material.

Applications

To further demonstrate the potential of the PC-VAR(1) approach, we applied the approaches under study to data of one young women participating in the COGITO study (Schmiedek et al., 2010). The goal of the COGITO study was to collect 100 daily observations from each participant. Data were not ever collected on Sundays or holidays and, for the majority of participants, also not on Saturdays. Participants also were not fully adherent, frequently missing testing days despite the best efforts of the researchers. As a result, the mean duration of the study was 162 days (range, 117–373) for younger adult participants and 150 days (range, 117–373) for older adult participants.

We selected the same individual as in Bulteel et al. (2016b). The data are from a young woman who was highly adherent and exceeded the cutoff score used to identify individuals at risk for clinical depression (Center for Epidemiological Studies Depression Scale [CES-D] score of 40, whereas the cutoff value in a representative adult German sample is 28; Hautzinger & Bailer, 1993). This participant had 78 daily observations that were exactly 1 day apart (measured on days t and $t+1$), 12 observations measured 2 days apart (days t and $t+2$), 6 observations measured 3 days apart, and 4 observations measured four or more days apart. To apply PC-VAR(1) analysis, the lag length needs to be the same between all measurements. Therefore, we used

Table 2. The loadings and the VAR(1) coefficients of the PC-VAR(1) solution with three components for the affective states from the COGITO data, after orthogonal and oblique rotation.

		Orthogonal rotation			Oblique rotation		
		Comp 1	Comp 2	Comp 3	Comp 1	Comp 2	Comp 3
Loadings	Distressed	0.79	−0.15	0.11	0.83	0.04	0.01
	Upset	0.12	0.09	0.89	0.01	0.04	0.90
	Guilty	0.74	−0.20	0.04	0.77	−0.01	−0.06
	Scared	0.82	−0.28	0.15	0.82	−0.10	0.04
	Hostile	0.13	−0.12	0.90	−0.03	−0.18	0.90
	Irritable	0.30	−0.02	0.80	0.20	−0.02	0.78
	Ashamed	0.69	−0.19	0.12	0.71	−0.03	0.03
	Nervous	0.86	−0.04	0.13	0.93	0.19	0.02
	Jittery	0.83	−0.07	0.14	0.89	0.14	0.04
	Afraid	0.85	−0.18	0.11	0.89	0.03	0.01
	Enthusiastic	−0.29	0.58	−0.30	−0.11	0.59	−0.26
	Excited	0.19	0.76	−0.16	0.46	0.9	−0.17
	Strong	−0.44	0.55	0.07	−0.35	0.48	0.14
	Interested	−0.18	0.67	−0.17	0.02	0.71	−0.13
	Proud	−0.62	0.28	−0.10	−0.60	0.14	−0.02
	Alert	−0.26	0.62	−0.07	−0.10	0.61	−0.02
	Inspired	−0.06	0.47	0.01	0.06	0.50	0.03
	Determined	−0.15	0.57	0.18	−0.03	0.57	0.22
	Attentive	−0.14	0.59	0.24	−0.03	0.58	0.28
	Active	−0.30	0.64	−0.04	−0.15	0.63	0.01
VAR(1) coefficients	Comp 1, lag 1	0.76	−0.21	0.07	0.81	−0.30	0.16
	Comp 2, lag 1	−0.12	0.44	0.02	−0.03	0.40	0.06
	Comp 3, lag 1	0.13	−0.04	0.14	0.05	−0.02	0.14

Note: Component loadings >0.50 in magnitude are shaded.

only the 78 observations that were exactly 1 day apart (i.e., retaining 78 pairs of observations). As this implies a substantial loss of information (i.e., discarding 22 observations), we also used the continuous-time variant of the regular VAR(1) model in step 2 of the PC-VAR(1) procedure to verify the findings. To be able to directly compare the results of both approaches, the continuous-time VAR(1) parameters were transformed into regular VAR(1) parameters with a time interval of 1 day.

The participant filled out questionnaires and worked on cognitive tasks in a lab. In this article, we focus on two different variable sets that correspond to two typical dynamic network applications. First, we analyze the affective states over the course of the study using the PANAS items (for a similar analysis, see e.g., Pe et al., 2015). Second, we focus on the dynamics of depression-related symptoms (as was done in Bringmann et al., 2015 and Bulteel et al., 2016b).

COGITO data set on affective states

Affect was measured using 20 items from the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988; the items are listed in Table 2) in the COGITO study. Ratings were provided on an 8-point scale. Previous research (Pe et al., 2015) indicated that MDD is characterized by dense emotion networks. For the selected individual who is potentially depressed, we investigated the specific constellation of such a network. We took into account that multicollinearity might be present as items such as scared and afraid or jittery and nervous can be expected to be highly correlated. Multicollinearity was indeed present as the absolute values of the pairwise within-person correlations between the variables range up to 0.86 ($M = 0.29$, $SD = 0.18$).

Comparing the predictive accuracy of the different methods as shown in Figure 4d, it can be seen that PC-VAR(1) has the best performance when the number of components is <12. The regular VAR(1) model has the worst predictive accuracy, followed by graphical VAR(1), and the univariate lasso VAR(1) model. Figure 5 shows the networks that result from applying the different methods under study.[4] With so many variables, the VAR(1) network in Figure 5a is neither insightful nor easy-to-read. Although many edges are dropped in the univariate lasso VAR(1) network in Figure 5b, it is still hard to interpret. Regarding the graphical VAR(1) results, the PDC network in Figure 5c is empty. In the PCC network in Figure 5d, some correlations show up between the negative emotions, but no clear grouping structure is revealed. We will not further interpret and discuss these networks,

[4]A lag length selection analysis (Brandt & Williams, 2007) was performed by means of the Bayesian information criterion (BIC; Schwarz, 1978). The information criterion was computed using the R package vars (Pfaff, 2008). For a VAR model of order 1, the BIC value was minimized (i.e., the values for a number of lags ranging up to 3 are −2.45, 6.03, and 12.08, respectively). The conclusions are the same after standardizing the data.

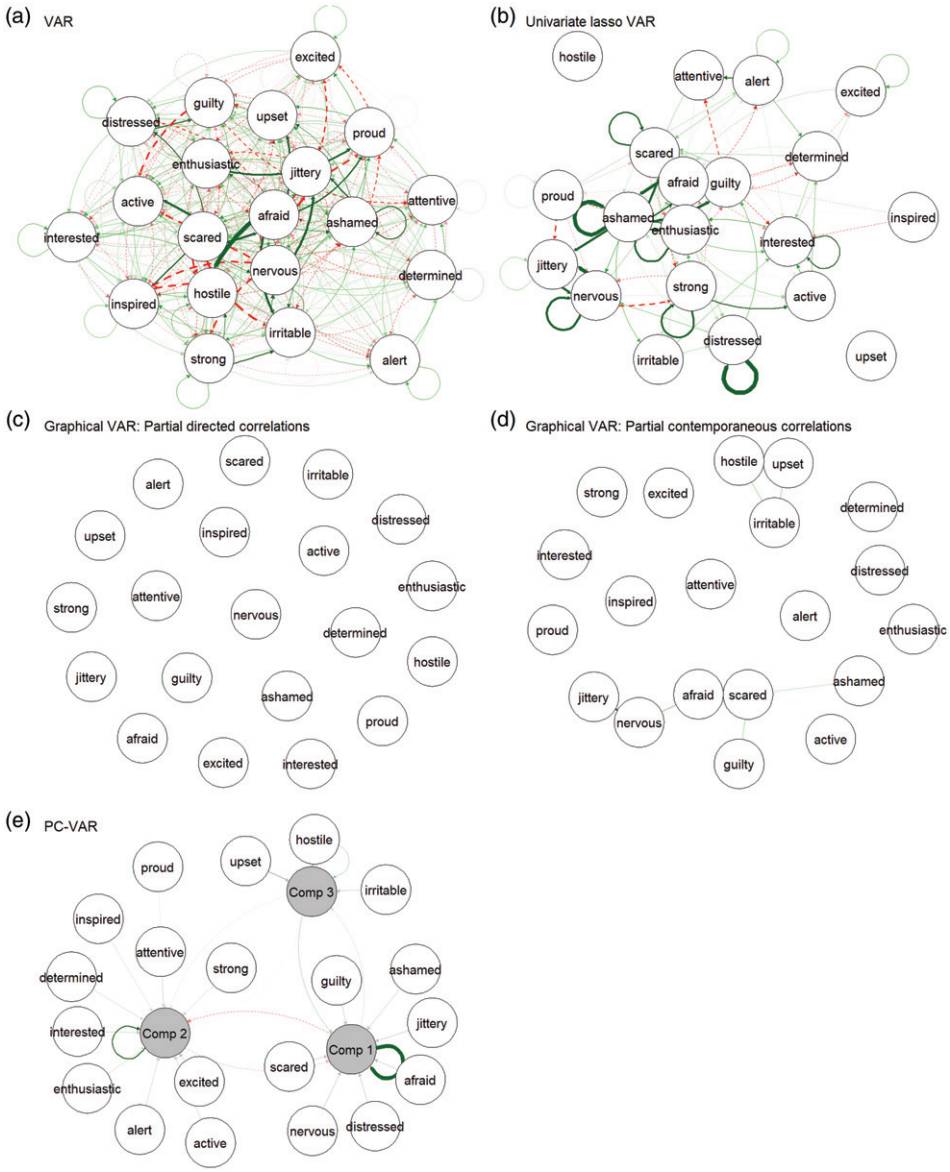


Figure 5. Network figures of the COGITO data set on affective states. (a) The network for the VAR(1) coefficients, (b) for the univariate lasso VAR(1) coefficients, (c) for the PDC from the graphical VAR(1) approach, (d) for the PCC from the graphical VAR(1) approach, and (e) the loadings and VAR(1) weights obtained by applying PC-VAR(1). In all panels, green solid edges indicate positive coefficients and red dashed edges negative coefficients. Additionally, the loadings are indicated by the gray-directed arrows connecting components and variables.

because they are based on the models with a relatively low predictive accuracy.

The results for PC-VAR(1) are shown in Figure 5e and summarized in Table 2. Based on the scree plot at the left of Figure 4b, we retain three components as the last clear elbow is situated at four components. In addition, the prediction error is also minimized for three components as shown in the right plot of Figure 4b (solid line). The first three components explain almost 59% of the data variance (i.e., 36, 13, and 10%, respectively). Table 2 shows that the results based on an orthogonal or an oblique rotation are comparable. As shown in Figure 5e (based on the orthogonal rotation) and listed in Table 2, the variables that have relatively high loadings on the first component are 'internalizing' negative affective states from the content domains of fear and nervousness. The second component contains positive affective states, and the third negative affective states related to the content domain of anger.

Regarding the lagged relations, the self-loop of the first component (i.e., 'internalizing' negativity) is the strongest, the self-loop of the second (i.e., positivity) is substantially smaller, and the self-loop of the third component (i.e., anger) is even lower. That is, affective states from the content domain of anger persist less

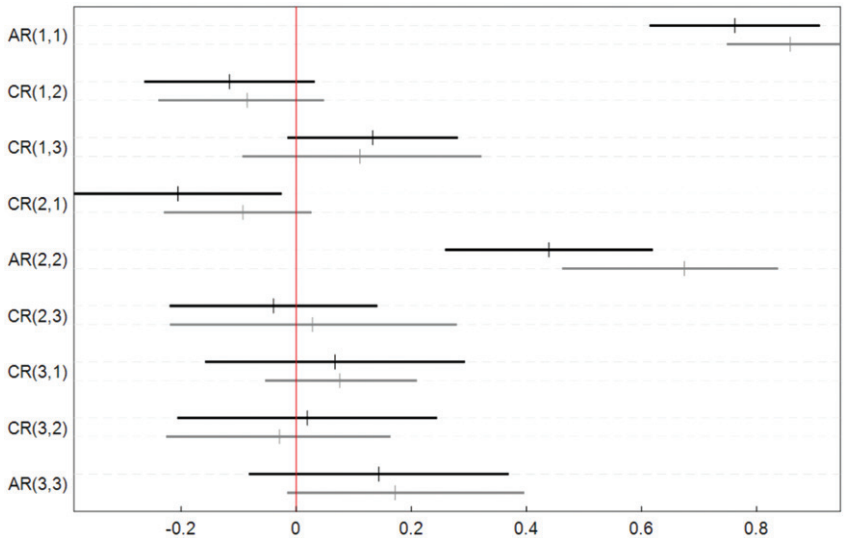


Figure 6. PC-VAR(1) coefficients for the COGITO data set on affective states. The autoregressive effects (AR) are the diagonal elements of the VAR(1) coefficient matrix, and the crossregressive effects (CR) are the off diagonal elements. The numbers between the brackets indicate the row and column of the element in the matrix, respectively. For each coefficient, the upper set of estimates (black) is obtained with the regular VAR(1) model, and the lower (gray) with the continuous-time variant. The point estimates are represented by a small vertical line (the long vertical reference line indicates a zero value) and the corresponding confidence intervals by the horizontal line.

across time than do more internal negative affective states. This seems plausible, given that anger is likely to be more tied to specific emotion-eliciting situations than nervousness (Gross, 2010).

The strongest and only significant crosslagged relation in the PC-VAR(1) network is the negative crosslagged effect of the first component on the second component. This implies that internalizing negative affect is not only pervasive across time; it may also reduce the experience of positive feelings in the future. This is perhaps the case because the individual keeps being absorbed by thoughts about the negative feelings and its causes (e.g., Brose, Schmiedek, Koval, & Kuppens, 2015). All other crossregressive effects are nonsignificant which might be due to low statistical power resulting from the relatively low number of observations (i.e., 78) and to the time scale of the COGITO study: Measurement intervals of 1 day might be too wide to reveal potentially causal influences among different components of affect if one considers the multitude of aspects that impact the ups and downs of daily mood and emotions (e.g., circadian rhythms, the weather, daily hassles, just to name a few). The network figure in this case mainly sheds light on the self-sustaining aspects of the within-person affect dynamics. Importantly, comparing these results to those shown in Figure 5a,b, our skepticism increases regarding the validity of the crosslagged effects found with the VAR(1) and the univariate lasso approach.

We also applied a PC-VAR(1) analysis with the continuous-time variant of the regular VAR(1) model in the second step. The results are shown in Figure 6. The autoregressive effects are even stronger in the continuous time model, compared to the regular discrete time VAR(1) model. The crossregressive coefficients, on the other hand, are somewhat closer to zero. Taking the confidence intervals into account, the interpretation of the resulting coefficients is identical, however, except for one crosslagged effect.

COGITO data set on depression-related symptoms

In this application, we focus on the same participant as in the previous section, but we now examine the dynamics of her depression-related symptoms (the symptoms are provided in Table 3). This data set was used before to illustrate the importance of using standardized VAR(1) weights (Bulteel et al., 2016b). The following 11 depression-related symptoms are included: Physical symptoms, rumination, feeling guilty, feeling unhappy, feeling downhearted, loss of activation, loss of interest, poor sleep quality, cognitive problems, loss of energy and restlessness. The absolute values of the pairwise within-person correlations for this data set range up to 0.64 ($M = 0.34$, $SD = 0.15$). More information on the data and tests of assumptions (including lag length specification analysis) can be found in Bulteel et al. (2016b).

Table 3. The loadings and the VAR(1) coefficients of the PC-VAR(1) solution with three components for the depression-related symptoms from the COGITO data, after orthogonal and oblique rotation.

		Orthogonal rotation			Oblique rotation		
		Comp 1	Comp 2	Comp 3	Comp 1	Comp 2	Comp 3
Loadings	PhysSymptoms	0.04	0.73	−0.03	−0.12	0.80	−0.16
	Rumination	0.12	−0.05	0.82	−0.03	−0.12	0.87
	Guilty	0.70	−0.06	0.45	0.72	−0.23	0.33
	Unhappy	0.88	−0.01	0.19	0.97	−0.19	0
	Down	0.63	0.09	0.55	0.58	−0.06	0.43
	Loss of act	0.56	0.60	−0.01	0.51	0.55	−0.23
	Loss of interest	0.70	0.30	0.03	0.72	0.20	−0.18
	SleepQual	0.18	0.44	0.62	−0.02	0.41	0.57
	CognProblems	0.43	0.57	0.34	0.29	0.52	0.18
	Loss of energy	−0.04	0.78	0.38	−0.31	0.84	0.30
	Restlessness	0.27	0.35	0.73	0.07	0.29	0.68
VAR(1) coefficients	Comp 1, lag 1	0.67	−0.05	0.14	0.69	−0.06	0.16
	Comp 2, lag 1	−0.14	0.60	0.01	−0.17	0.55	−0.05
	Comp 3, lag 1	0.14	0.29	0.59	0.20	0.33	0.62

Note: Component loadings >0.50 in magnitude are shaded.

The predictive accuracies of the models under study are shown in the right plot of Figure 4c. The PC-VAR(1) approach results overall in the lowest prediction error. The next best performance is achieved by the univariate lasso VAR(1) model, followed by the VAR(1) model, and finally graphical VAR(1).

The VAR(1) network is shown in Figure 7a. Feeling guilty and rumination have strong self-loops. Restlessness seems rather central, in that it has many links to other nodes such as loss of energy and cognitive problems. (A discussion of this network can also be found in Bulteel et al., 2016b). In the univariate lasso VAR(1) network in Figure 7b, many of the links are set to zero. Most self-loops are still present. The PDC network of the graphical VAR(1) approach in Figure 7c contains only the self-loop of guilty. In the PCC network in Figure 7d, unhappy and downhearted are strongly linked. Again, restlessness has many links. As the underlying models have a relatively poor predictive accuracy, we will not examine these network figures by means of network characteristics.

The scree plot (left plot in Figure 4c) does not give a clear indication of the number of components to select. Inspecting the CV plot (right plot in Figure 4c), it can be seen that the MSE_{CV} is the lowest for the PC-VAR(1) models with four and six components. However, inspecting the loading matrices of these PC-VAR(1) models, we see many crossloadings (i.e., high loadings of items on more than one component) which make the results hard to interpret. In addition, the network figures for the higher component solutions contain many negative edges connecting the components, which is counterintuitive. Therefore, we selected the three-component solution. The corresponding results are shown in Figure 7e and summarized in Table 3. The three components explain 67% of the variance of the predictor variables. The first component combines feeling guilty, feeling unhappy, feeling downhearted, loss of activation, and loss of interest (we label this component being absorbed by own negative affect). Physical symptoms, loss of activation, cognitive problems, and loss of energy form the second component (somatic component). Rumination, feeling downhearted, sleep quality, and restlessness load on the third component (agitation component). Together, the components capture different aspects of depression (DSM-V; American Psychiatric Association, 2013). The crossloadings may be indicative of the dense emotion networks in depressed people (Pe et al., 2015).

Again, we computed the regular and the continuous-time VAR(1) coefficients for the PC-VAR(1) approach. As shown in Figure 8, the differences between both methods are limited. The autoregressive coefficients for all components are relatively strong. That is, the different aspects of depressive symptomatology persist across time – which could be an additional indication that this person is at risk (van de Leemput et al., 2014). The strongest crossregressive link connects the third component (i.e., agitation component) to the second component (i.e., somatic component) at the next time point; this positive relation would be in line with the idea that depressive symptoms infect each other across time (Borsboom & Cramer, 2013). More specifically, more somatic problems seem to follow symptoms of restlessness and agitation in this person – an insight that could be relevant for interventions.

Discussion

In this article, we compared the lasso VAR(1) and PC-VAR(1) approaches by applying them to a toy example

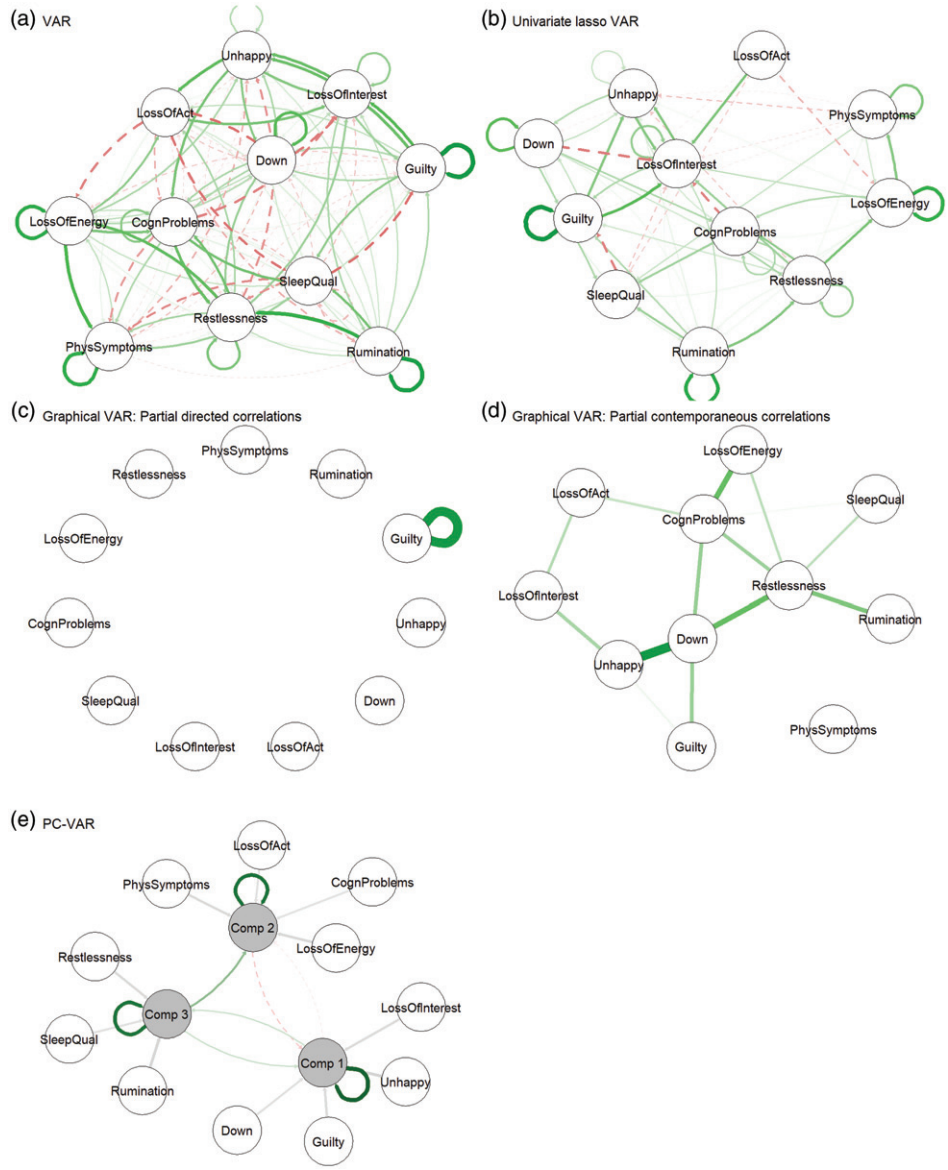


Figure 7. Network figures of the COGITO data set on depression-related symptoms. (a) The VAR(1) coefficients, (b) the univariate lasso VAR(1) coefficients, (c) the PDC from the graphical VAR(1) approach, (d) the PCC from the graphical VAR(1) approach, and (e) the loadings and VAR(1) weights obtained by applying PC-VAR(1). In all panels, green solid edges indicate positive coefficients and red dashed edges negative coefficients. Additionally, the loadings are indicated by the gray-directed arrows connecting components and variables.

as well as to data of a single participant in the COGITO study. We focused on the interpretability of the resulting solutions and on their predictive accuracy. Although augmenting the VAR(1) model with a lasso penalty had a superior predictive performance compared to the standard VAR(1) model, PC-VAR(1) always had the highest predictive accuracy. Regarding interpretability, the PC-VAR(1) yields a readable network figure, combining the contemporaneous correlation structure and lagged effects in a single figure. In contrast, the interpretability of the lasso-based networks is jeopardized by the arbitrary choices in case of groups of highly correlated variables (as it was shown for the toy example).

Limitations

The first limitation of this study is our exclusive focus on lasso regularization. Initially, we also included the elastic net (Zou & Hastie, 2005). This penalty seems attractive because it was developed, among others, to deal with groups of highly interrelated variables. It adds a group effect to the lasso: It includes or excludes groups of variables as a whole, and if such a group is retained in the model, it assigns similar regression weights to the group variables. However, the results were not satisfactory for the two COGITO data sets we considered. First, it was unclear whether two variables

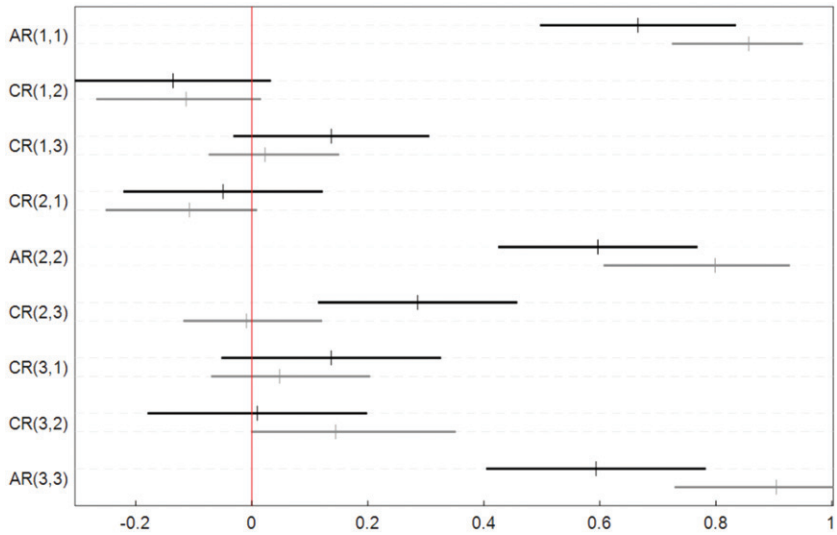


Figure 8. PC-VAR(1) coefficients for the COGITO data set on depression-related symptoms. The autoregressive effects (AR) are the diagonal elements of the VAR(1) coefficient matrix, and the crossregressive effects (CR) are the off diagonal elements. The numbers between the brackets indicate the row and column of the element in the matrix, respectively. For each coefficient, the upper set of estimates (black) is obtained with the regular VAR(1) model, and the lower (gray) with the continuous-time variant. The point estimates are represented by a small vertical line (the long vertical reference line indicates a zero value) and the corresponding confidence intervals by the horizontal line.

had similar VAR(1) coefficients because they formed a group, or because they had a comparable effect on the criterion variable. A second reason is that the network figures were still difficult to interpret as all edges were retained for the data sets shown here.

Second, in this article we focused on the data of a single person only. In cases in which time series data are collected from multiple persons, there are different possibilities for combining their data depending on the hypothesized similarities and differences between these persons. For instance, it can sometimes be reasonable to assume that the underlying components and/or VAR(1) coefficients are equal for all persons. In other cases, it can be more realistic to expect that subgroups of individuals exist and that each has its own specific underlying components and/or VAR(1) structure. Simultaneous component analysis (SCA; Ceulemans, Wilderjans, Kiers, & Timmerman, 2016; Kiers & Berge, 1994; Timmerman & Kiers, 2003; Van Deun, Smilde, van der Werf, Kiers, & Van Mechelen, 2009) can be used, to extract one shared set of components for all persons. When subgroups of persons exist with different component structures, an extension of clusterwise SCA would be useful (De Roover et al., 2012; De Roover, Ceulemans, Timmerman, Nezlek, & Onghena, 2013). To model categorical differences in VAR(1) coefficients, one can build on the previous work on clusterwise regression approaches (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016a; Wilderjans, Vande Gaer, Kiers, Van Mechelen, & Ceulemans, 2017).

A third limitation is that we assume that the loadings, and thus the contemporaneous correlations, remain constant over time when using PC-VAR(1). Of course, it is possible that this is not the case. The current PC-VAR(1) approach might be extended based on the principles behind switching PCA (De Roover, Timmerman, Van Diest, Onghena, & Ceulemans, 2014), in which the time series is divided into separate phases that are each characterized by a separate set of component loadings. In this regard, change point detection methods for signaling correlation change might be a useful screening tool to decide on an appropriate modeling strategy (Cabrieto et al., 2018; Cabrieto, Tuerlinckx, Kuppens, Hunyadi, & Ceulemans, 2018).

Directions for future research

We see two important directions for future research. First, we showed that the PC-VAR(1) model was capable of recovering the structure underlying our toy example. In future research, an extensive simulation study should be performed to thoroughly investigate the performance of the PC-VAR(1) model. Important factors to manipulate in such a study are the number of time points, the number of variables, the number of components, the strength and correlation of the

Table 4. The advantages and the disadvantages of the methods under study.

	Pros	Cons
Lasso VAR(1)	*Sparse network for a large number of variables	*Only unique direct effects represented *Arbitrary selection of edges *Measurement error not addressed
PC-VAR(1)	*Shared effects represented *Clear figure for a large number of variables *Measurement error removed	*Loss of information (due to a reduction of the variables to a smaller set of components) *Interpretability of the components not always straightforward

components, and the crosslagged structure (see e.g., Molenaar & Nesselroade, 2009). In addition, it would be interesting to vary the amount of measurement error to further study its influence on the PC-VAR(1) estimates of the loadings and the VAR(1) coefficients.

Second, further research is required to study EPFA, the factor analysis-based counterpart of PC-VAR(1) (Browne & Nesselroade, 2005).[5] In short, in this model, it is assumed that the measurements are manifestations of Q underlying factors which follow a VAR(1) process:

$$\begin{aligned} \mathbf{y_t} &= \mathbf{\Lambda f_t} + \mathbf{e_t} \\ \mathbf{f_t} &= \mathbf{A f_{t-1}} + \mathbf{u_t} \end{aligned} \tag{9}$$

where the $J \times Q$ matrix $\mathbf{\Lambda}$ holds the factor loadings, $\mathbf{f}_t$ $(Q \times 1)$ and $\mathbf{f}_{t-1}$ $(Q \times 1)$ contain the factor scores at time t and t–1, respectively, $\mathbf{e}_t$ $(J \times 1)$ contains the scores on the unique factors at time t, and the $Q \times Q$ matrix $\mathbf{A}$ holds the VAR(1) coefficients. To identify the model, different constraints are possible (e.g., on $\mathbf{\Lambda}$ or on the (co)variances of $\mathbf{f}_t$). This model is more complex to fit than PC-VAR(1) as it extracts a few latent factors, and simultaneously models, among others, their lagged relations. It is also not straightforward to perform CV in the context of this model. One practical reason is that currently available software to estimate the EPFA model (Browne & Zhang, 2005) does not provide the necessary output to apply CV. More specifically, it is unclear how the factor scores for the test observations should be computed. In general, it seems more straightforward to apply CV in PCA based than in FA-based models (e.g., Bro, Kjeldahl, Smilde, & Kiers, 2008). It should be noted, moreover, that it is common practice in FA to validate a model by applying a confirmatory FA on a second independent sample (e.g., Ferrer & Nesselroade, 2003). It would be interesting to compare the loadings, and the component and factor scores (on which the VAR(1) coefficients are based) resulting from both approaches. We conjecture that the loadings of PC-VAR(1) will be higher on average, and that differences between the component and factors scores will be rather small.

[5]We also applied EPFA to the three presented data sets. Results can be requested from the first author.

Conclusion

In sum, building PC-VAR(1)-based networks is a promising alternative to the standard VAR(1) or lasso VAR(1) approaches. Such dimension reduction approaches have the potential to handle multicollinearity, give insight into both the lagged effects and the contemporaneous correlation structure of the variables, and offer a clear overview of the dynamics even when the number of variables is large. In addition, they might be suited to deal with measurement error and generalize better to new data. A complete overview of the advantages and disadvantages is summarized in Table 4. Future research is required to investigate the performance of the PC-VAR(1) approach in more detail, but the application to a few data sets proves its potential, especially in the case of a larger number of variables.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. Dr. Tuerlinckx and Dr. Ceulemans report a grant from the Belgian Government.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by Grant GOA/15/003 from The Research Fund of KU Leuven and by Grant (IAP/P7/06) from The Interuniversity Attraction Poles programme financed by The Belgian government. Kirsten Bulteel is a doctoral research fellow with the Research Foundation – Flanders (FWO).

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The authors would like to thank Dr. West and the reviewers for their comments on prior versions of this article. The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions or the listed funding agencies is not intended and should not be inferred.

References

Abegaz, F., & Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics*, *14*(3), 586–599. doi:10.1093/biostatistics/kxt005.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Babyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, *66*(3), 411–421. doi:0033-3174/04/6603-0411.

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*, 192–213. doi:10.1016/j.ins.2011.12.028.

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

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Bondell, H. D., & Reich, B. J. (2008). Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR. *Biometrics*, *64*(1), 115–123. doi:10.1111/j.1541-0420.2007.00843.x.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. doi:10.1146/annurev-clinpsy-050212-185608.

Brandt, P. T., & Williams, J. T. (2007). *Multiple time series models*. Thousand Oaks, CA: Sage Publications.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, *16*(3), 199–231. doi:10.1214/ss/1009213726.

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*, *45*(4), 747–757. doi:10.1017/S0033291714001809

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., … Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, *23*(4), 425–435. doi:10.1177/1073191116645909

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. 1–13. doi:10.1371/journal.pone.0060188

Bro, R., Kjeldahl, K., Smilde, A. K., & Kiers, H. A. L. (2008). Cross-validation of component models: A critical look at current methods. *Analytical and Bioanalytical Chemistry*, *390*(5), 1241–1251. doi:10.1007/s00216-007-1790-1.

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

Brose, A., Schmiedek, F., Koval, P., & Kuppens, P. (2015). Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition and Emotion*, *29*(3), 527–538. doi:10.1080/02699931.2014.916252

Brose, A., Schmiedek, F., Lövdén, M., Molenaar, P. C. M., & Lindenberger, U. (2010). Adult age differences in covariation of motivation and working memory performance: Contrasting between-person and within-person findings. *Research in Human Development*, *7*(1), 61–78. doi:10.1080/15427600903578177

Brose, A., Voelkle, M. C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2015). Differences in the between-person and the within-person structures of affect are a matter of degree. *European Journal of Personality*, *29*(1), 55–71. doi:10.1002/per.1961

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150. doi:10.1207/S15327906MBR3601_05.

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of autoregressive moving average time series models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 415–452). Mahwah, NJ: Lawrence Erlbaum Associates.

Browne, M. W., & Zhang, G. (2005). DyFA: Dynamic factor analysis of lagged correlation matrices (Version 2.03) [Computer software and manual]. Retrieved from http://faculty.psy.ohio-state.edu/browne/software.php

Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (in press). VAR(1) based models do not outpredict AR(1) models in current psychological applications. *Psychological Methods*. doi: 10.1037/met0000178.

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016a). Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics. *Frontiers in Psychology*, *7*, 1–15, 1540. doi: 10.3389/fpsyg.2016.01540

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016b). Using raw VAR regression coefficients to build networks can be misleading. *Multivariate Behavioral*

Research, *51*(2–3), 330–344. doi:10.1080/00273171.2016.1150151

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., & Ceulemans, E. (2018). Testing for the presence of correlation changes in a multivariate time series: A permutation based approach. *Scientific Reports*, *8*(1), 1–20, 769. doi:10.1038/s41598-017-19067-2

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Wilhelm, F., Liedlgruber, M., & Ceulemans, E. (2018). Capturing correlation changes by applying kernel change point detection on the running correlations. *Information Sciences*, *447*, 117–139. doi:10.1016/j.ins.2018.03.010

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. doi:10.1207/s15327906mbr0102_10

Ceulemans, E., & Kiers, H. A. L. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 133–150. doi:10.1348/000711005X64817

Ceulemans, E., & Kiers, H. A. L. (2009). Discriminating between strong and weak structures in three-mode principal component analysis. *British Journal of Mathematical and Statistical Psychology*, *62*(3), 601–620. doi:10.1348/000711008X369474

Ceulemans, E., Wilderjans, T. F., Kiers, H. A. L., & Timmerman, M. E. (2016). Multilevel simultaneous component analysis: A computational shortcut and software package. *Behavior Research Methods*, *48*(3), 1008–1020. doi:10.3758/s13428-015-0626-8.

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771. doi:10.1093/biomet/asnO34.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York, NY: Routledge.

Crawford, A. V., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Dubravka, S., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*(6), 885–901. doi:10.1177/0013164410379332.

de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*, *8*, 1849, 19. doi:10.3389/fpsyg.2017.01849.

De Roover, K., Ceulemans, E., Timmerman, M. E., Nezlek, J. B., & Onghena, P. (2013). Modeling differences in the dimensionality of multiblock data by means of clusterwise simultaneous component analysis. *Psychometrika*, *78*(4), 648–668. doi:10.1007/S11336-013-9318-4

De Roover, K., Ceulemans, E., Timmerman, M. E., Vansteelandt, K., Stouten, J., & Onghena, P. (2012). Clusterwise simultaneous component analysis for analyzing structural differences in multivariate multiblock data. *Psychological Methods*, *17*(1), 100–119. doi:10.1037/a0025385

De Roover, K., Timmerman, M. E., Van Diest, I., Onghena, P., & Ceulemans, E. (2014). Switching principal component analysis for modeling means and covariance changes over time. *Psychological Methods*, *19*(1), 113–132. doi:10.1037/a0034525

Diaconis, P. (1985). Theories of data analysis. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Exploring data tables, trends, and shapes* (pp. 1–36). New York, NY: Wiley.

Driver, C. C., Oud, J. H. L., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, *77*(5), 1–35. doi:10.18637/jss.v077.i05.

Epskamp, S. (2016). graphicalVAR: Graphical VAR for Experience Sampling Data (Version 0.1.4) [Computer software]. Retrieved from https://CRAN.R-project.org/package=graphicalVAR

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18. doi:doi:10.18637/jss.v048.i04.

Epskamp, S., Kruis, J., & Marsman, M. (2017). Estimating psychopathological networks: Be careful what you wish for. *PLoS One*, *12*(6), e0179891. 1–13. doi:10.1371/journal.pone.0179891.

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, 1–28. doi:10.1080/00273171.2018.1454823.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. doi:10.1037/1082-989X.4.3.272.

Ferrer, E., & Nesselroade, J. R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion*, *3*(4), 344–360. doi:10.1037/1528-3542.3.4.344.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One*, *12*(6), e0174035. doi:10.1371/journal.pone.0174035.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*(4), 430–450. doi:10.1037/1082-989X.6.4.430.

Gross, J. J. (2010). The future's so bright, I gotta wear shades. *Emotion Review*, *2*(3), 212–216. doi:10.1177/1754073910361982.

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.). *Handbook of research methods for studying daily life* (pp. 43–61), New York, NY: Guilford.

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

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

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer.

Hautzinger, M., & Bailer, M. (1993). *Allgemeine Depressions Skala [General Depression Scale]*. Weinheim, Germany: Beltz.

Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, *17*(1), 65–70. doi:10.1111/j.2044-8317.1964.tb00244.x.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. doi:10.2307/1267351.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. doi:10.1007/BF02289447.

Hsu, N.-J., Hung, H.-L., & Chang, Y.-M. (2008). Subset selection for vector autoregressive processes using Lasso. *Computational Statistics & Data Analysis*, *52*(7), 3645–3657. doi:10.1016/j.csda.2007.12.004.

Jolliffe, I. T. (1982). A note on the use of principal components in regression. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *31*(3), 300–303. doi:10.2307/2348005.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). New York, NY: Springer.

Kaiser, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. doi:10.1007/BF02289233.

Kiers, H. A. L., & Berge, J. M. F. (1994). Hierarchical relations between methods for simultaneous component analysis and a technique for rotation to a simple simultaneous structure. *British Journal of Mathematical and Statistical Psychology*, *47*(1), 109–126. doi:10.1111/j.2044- 8317.1994.tb01027.x

Kiers, H. A. L., & Smilde, A. K. (2007). A comparison of various methods for multivariate regression with highly collinear variables. *Statistical Methods and Applications*, *16*(2), 193–228. doi:10.1007/s10260-006-0025-5.

Liu, Y., & West, S. G. (2016). Weekly cycles in daily report data: An overlooked issue. *Journal of Personality*, *84*(5), 560–579. doi:10.1111/jopy.12182

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340–364. doi:10.1080/00273171.2011.564527.

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

McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate Behavioral Research*, *50*(5), 471–483. doi:10.1080/00273171.2015.1036965.

Meredith, W., & Millsap, R. E. (1985). On component analyses. *Psychometrika*, *50*(4), 495–507. doi:10.1007/BF02296266.

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

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201–218. doi:10.1207/s15366359mea0204_1.

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

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

Morana, C. (2012). PC-VAR estimation of vector autoregressive models. *Open Journal of Statistics*, *2*(3), 251–259. doi:10.4236/ojs.2012.23030.

Newman, M. E. J. (2010). *Networks: An introduction*. New York, NY: Oxford University Press.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251. doi:10.1016/j.socnet.2010.03.006.

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

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*, *3*(2), 292–300. doi:10.1177/2167702614540645.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, *2*(11), 559–572. doi:10.1080/14786440109462720.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, *27*(4), 1–32. doi:10.18637/jss.v027.i04.

Rothman, A. J., Levina, E., & Zhu, J. (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics*, *19*(4), 947–962. doi:10.1198/jcgs.2010.09188.

Ryan, O., Kuiper, R. M., & Hamaker, E. L. (in press). A continuous time approach to intensive longitudinal data: What, why and how? In K. van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences*. New York, NY: Springer.

Saccenti, E., & Timmerman, M. E. (2017). Considering Horn's parallel analysis from a random matrix theory point of view. *Psychometrika*, *82*(1), 186–209. doi:10.1007/s11336-016-9515-z.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience*, *2*, 27. doi:10.3389/fnagi.2010.00027.

Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology*, *6*, 1038. doi:10.3389/fpsyg.2015.01038.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. doi:10.1214/aos/1176344136.

Shmueli, G. (2010). To explain or to predict?. *Statistical Science*, *25*(3), 289–310. doi:10.1214/10-STS330.

Snijders, T. A. B. (1988). On cross-validation for predictor evaluation in time series. In T. K. Dijkstra (Ed.), *On model uncertainty and its statistical implications* (pp. 56–69). New York, NY: Springer-Verlag.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*, 267–288. doi:10.1111/j.1467-9868.2011.00771.x.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical*

Society: Series B (Statistical Methodology), *73*(3), 273–282. doi:10.1111/j.1467-9868.2011.00771.x.

Timmerman, M. E., & Kiers, H. A. L. (2003). Four simultaneous component models for the analysis of multivariate time series from more than one subject to model intraindividual and interindividual differences. *Psychometrika*, *68*(1), 105–121. doi:10.1007/BF02296656

Tsay, R. S. (2005). *Analysis of financial time series*. Hoboken, NJ: John Wiley & Sons.

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., … Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(1), 87–92. doi:10.1073/pnas.1312114110.

van der Krieke, L., Emerencia, A. C., Bos, E. H., Rosmalen, J. G. M., Riese, H., Aiello, M., … de Jonge, P. (2015). Ecological momentary assessments and automated time series analysis to promote tailored health care: A proof-of-principle study. *JMIR Research Protocols*, *4*(3), e100. doi:10.2196/resprot.4000.

Van Deun, K., Smilde, A. K., van der Werf, M. J., Kiers, H. A. L., & Van Mechelen, I. (2009). A structured overview of simultaneous component based data integration. *BMC Bioinformatics*, *10*(1), 246. doi:10.1186/1471-2105-10-246

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*(1), 97–28. doi:10.1207/s15327906mbr2501_12.

Vervloet, M., Van Deun, K., Van den Noortgate, W., & Ceulemans, E. (2016). Model selection in principal covariates regression. *Chemometrics and Intelligent Laboratory Systems*, *151*, 26–33. doi:10.1016/j.chemolab.2015.12.004

Vitale, R., Westerhuis, J. A., Naes, T., Smilde, A. K., de Noord, O. E., & Ferrer, A. (2017). Selecting the number of factors in principal component analysis by permutation testing: Numerical and practical aspects. *Journal of Chemometrics*, *31*(12), e2937. 1–15. doi:10.1002/cem.2937.

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*(3), 384–192. doi:0.1037/a0027543.

Walls, T. A., & Schafer, J. L. (2006). *Models for intensive longitudinal data*. New York, NY: Oxford University Press.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. doi:10.1037/0022-3514.54.6.1063.

Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, *44*(07), 1349–1360. doi:10.1017/S0033291713001979.

Wigman, J. T. W., van Os, J., Borsboom, D., Wardenaar, K. J., Epskamp, S., Klippel, A., … Wichers, M. (2015). Exploring the underlying structure of mental disorders: Cross-diagnostic differences and similarities from a network perspective using a both top-down and a bottom-up approach. *Psychological Medicine*, *45*(11), 2375–2387. doi:10.1017/S0033291715000331.

Wild, B., Eichler, M., Friederich, H.-C., Hartmann, M., Zipfel, S., & Herzog, W. (2010). A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC Medical Research Methodology*, *10*(1), 28. doi:10.1186/1471-2288-10-28.

Wilderjans, T. F., Ceulemans, E., & Meers, K. (2013). CHull: A generic convex hull based model selection method. *Behavior Research Methods*, *45*(1), 1–15. doi:10.3758/s13428-012-0238-5

Wilderjans, T. F., Vande Gaer, E., Kiers, H. A. L., Van Mechelen, I., & Ceulemans, E. (2017). Principal covariates clusterwise regression (PCCR): Accounting for multicollinearity and population heterogeneity in hierarchically organized data. *Psychometrika*, *82*(1), 86–111. doi:10.1007/s11336-016-9522-0.

Wolff, J. K., Brose, A., Lövdén, M., Tesch-Römer, C., Lindenberger, U., & Schmiedek, F. (2012). Health is health is health? Age differences in intraindividual variability and in within-person versus between-person factor structures of self-reported health complaints. *Psychology and Aging*, *27*(4), 881–891. doi:10.1037/a0029125

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. doi:10.1111/j.1467-9868.2005.00503.x.