Network estimation from time series and panel data

Burger, J.; Hoekstra, R.H.A.; Mansueto, A.C.; Epskamp, S.

DOI
10.4324/9781003111238-13

Publication date
2022

Document Version
Final published version

Published in
Network Psychometrics with R

Chapter 10

Network Estimation from Time Series and Panel Data

Julian Burger[1,2,3], Ria H. A. Hoekstra[2], Alessandra C. Mansueto[2,3,4], & Sacha Epskamp[2,3]

1. University of Groningen, University Medical Center Groningen
2. University of Amsterdam, Department of Psychology
3. University of Amsterdam, Centre for Urban Mental Health
4. University of Amsterdam, Department of Communication Science

10.1 Introduction

The previous chapter introduced *time* into the sampling design of studies. With the addition of time, longitudinal analysis of multiple measures per person becomes possible. As the previous chapter discussed, this step is vital in separating within- from between-person relationships. Time adds a new level of complexity to the modeling frameworks we have used before: temporal dependencies between observations of the same person over time. This complicates the models—as many more parameters need to be estimated—but also comes with the substantial benefit of allowing researchers to study temporal effects—often termed 'dynamical relationships.' With the advent of modern data collection methodologies, such as electronic diaries or wearables, time series data have become a new data source to estimate networks. For example, in experience sampling method and ecological momentary assessment, a smartphone can be used to query a subject multiple times per day (also termed 'beeps') on a set of questions. Through these advances, the

Cite this chapter as:

Burger, J., Hoekstra, R. H. A., Mansueto, A. C., & Epskamp, S. (2022). Chapter 10. Network estimation from time series and panel data. In Isvoranu, A. M., Epskamp, S., Waldorp, L., J. & Borsboom, D. (Eds.). *Network psychometrics with R: A guide for behavioral and social scientists*. Routledge, Taylor & Francis Group.

use of dynamical network models has grown prominent. This chapter introduces the estimation of dynamical networks from longitudinal data. The chapter begins with a summary of the main modeling framework used, followed by explanations on how to estimate these models from $N = 1$ and $N > 1$ data sets. The chapter will then conclude with an overview of practical and methodological challenges in dynamical network analysis.

10.2 Graphical vector auto-regression

The main model we focus on in this chapter is the (lag-1) graphical vector auto-regression (GVAR; Epskamp, Waldorp, et al., 2018; Wild et al., 2010) model for continuous (and assumed normal) data introduced in Chapter 9 and further detailed in Technical Box 10.1. In this model, a person's responses on a certain measurement are modeled as a Gaussian graphical model (GGM; see Chapter 6) after conditioning on their responses in the previous measurement. Alternatively, the GVAR model can also be interpreted as a multivariate regression on the previous responses, with the residuals (then termed *innovations*) being modeled through a GGM. The GVAR model includes two network structures: a *temporal network* that encodes how well deviations from the person-wise mean in one variable at a certain measurement occasion predict deviations from the person-wise mean in the *next* measurement occasion,[1] and a *contemporaneous network* that encodes relationships between variables within the same measurement occasion and after controlling for temporal effects. These networks allow for a within-person interpretation and can be estimated per person or for the average person (fixed effects). In $N > 1$ data, relationships between the means can further be investigated to construct a GGM termed the *between-person network*.

Figure 10.1 shows an example of these three network structures, estimated from time series data collected by Fried et al. (2021).[2] The temporal network in the left panel of Figure 10.1 contains self-loops (auto-regressions) and edges between nodes (cross-lagged regressions). For example, the self-loop on 'difficulty relax' indicates that when a person had more difficulty relaxing than their average in one measurement occasion, that person likely still had a higher than average difficulty to relax in the next measurement occasion.[3] Edges between nodes indicate similar predictions but then for different nodes. For example, we can see that 'angry' and 'irritable' predict one another well over time. The contemporaneous network in the middle panel of Figure 10.1 indicates, for example, that a person who is currently more 'worried' than their average is likely also more 'nervous' than their average at the same time. Lastly, the between-subjects network in the right panel of Figure 10.1 shows, for example, that individuals who 'worry' a lot on

[1] Temporal connections can also be said to encode *Granger causality* (Eichler, 2007; Granger, 1969) as they encode temporal predictions. However, that does not mean the edges necessarily encode *causal* relationships. Ultimately, temporal edges are just partial correlations between a lagged (encoding the previous time point) and a non-lagged variable after controlling for all other lagged variables. As such, similar reservations to causal interpretations in the temporal network should be taken as discussed in Chapter 6.

[2] The data, including a detailed overview of the measures used, are available online at https://osf.io/mvdpe/.

[3] The inverse interpretations are also true: a positive edge also indicates that whenever people experienced *less* difficulty relaxing than their average in one measurement occasion, they likely also experienced *less* difficulty relaxing in the following measurement occasion.

average also tend to be individuals that are 'irritable' on average. The between-person network also shows an interesting negative relationship between feelings of hopelessness[4] and difficulty relaxing. As explained in Chapter 6, this could be due to a common effect structure. For example, perhaps both being a person who often feels hopeless and being a person who often has difficulty relaxing leads one to become a person who often worries.[5] Interestingly, this common effect structure can also be seen at the temporal level. Another explanation is Simpson's paradox, in which an effect becomes different when conditioning on a different level of the data (Hamaker, 2012; Kievit et al., 2013).

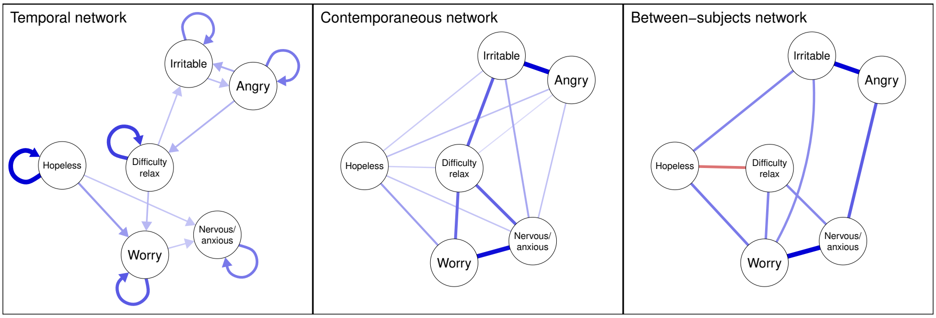

Figure 10.1. Example of a temporal (fixed effects), contemporaneous (fixed effects), and between-subjects network with six nodes based on data collected by Fried et al. (2021). The data consist of four measurements per day for 14 subsequent days, filled in by 80 undergraduate students. Networks were estimated using two-step multi-level estimation with the *mlVAR* package using correlated random effects. The network structure was obtained by thresholding edges at $\alpha = 0.05$ (using an AND-rule for the contemporaneous and between-person networks).

10.3 *N* = 1 estimation: personalized network models

The GVAR model can be estimated in various ways from time series data of a single subject ($N = 1$). Doing so will return a temporal and a contemporaneous network.[6] Estimation of the GVAR model from $N = 1$ data mostly follows the same principles as estimating GGMs from single measurement data: the GVAR model can be estimated through multivariate or univariate (nodewise) estimation, using frequentist and Bayesian estimation, and model selection can be performed through regularization or other model search strategies. This section will discuss estimation strategies for the GVAR, followed by model selection strategies.

[4]The actual measure used for this node was "I felt that I had nothing to look forward to."

[5]These data were collected during the first weeks of the 2020 COVID-19 pandemic outbreak. As such, any between-person effects could also be due to the general atmosphere of this time.

[6]It is not uncommon that only the temporal network obtained through a (G)VAR analysis is of interest. Not reporting the contemporaneous network, however, is not recommended, as (a) the temporal network is only half the statistical model, (b) the method might lack sensitivity, especially in temporal connections (which are often weak), and (c) relationships in the same window of measurement can be interesting as well (Epskamp, van Borkulo, et al., 2018).

The vector auto-regression (VAR) model is a model that assumes responses from longitudinal data of a single subject to be normally distributed after controlling for the previous measurement. Let $\boldsymbol{y}_{p,t}$ represent a set of responses from a subject p measured at time point t (for simplicity we do not denote random vectors different from realizations here). The VAR model can then be written as:

$$\boldsymbol{y}_{p,t} \mid \boldsymbol{y}_{p,t-1} \sim N\left(\boldsymbol{\mu}_p + \boldsymbol{B}_p\left(\boldsymbol{y}_{p,t-1} - \boldsymbol{\mu}_p\right), \boldsymbol{\Sigma}_p^{(\mathrm{C})}\right). \tag{10.1}$$

The $\boldsymbol{B}_p$ matrix encodes these temporal regression effects, with element β_{ijp} (row i, column j) encoding the temporal effect from variable j to variable i for subject p. The transpose of this matrix therefore encodes a *temporal network*.[a] The matrix $\boldsymbol{\Sigma}_p^{(\mathrm{C})}$ can be interpreted as the *contemporaneous* variance–covariance matrix: the variance–covariance structure after controlling for temporal dependencies. The subscript p indicates that these matrices can be modeled per person. In *graphical* vector auto-regression (GVAR; Epskamp, Waldorp, et al., 2018; Wild et al., 2010), we further model $\boldsymbol{\Sigma}_p^{(\mathrm{C})}$ through the use of a Gaussian graphical model (GGM; see Chapter 6):

$$\boldsymbol{\Sigma}_p^{(\mathrm{C})} = \boldsymbol{\Delta}_p^{(\mathrm{C})}\left(\boldsymbol{I} - \boldsymbol{\Omega}_p^{(\mathrm{C})}\right)^{-1}\boldsymbol{\Delta}_p^{(\mathrm{C})}.$$

The matrix $\boldsymbol{\Omega}_p^{(\mathrm{C})}$ encodes a *contemporaneous network*. Finally, in $N > 1$ data, the means can subsequently be treated as a random variable and also be modeled with a GGM:

$$\boldsymbol{\mu}_p \sim N\left(\boldsymbol{0}, \boldsymbol{\Sigma}^{(\mathrm{B})}\right); \quad \boldsymbol{\Sigma}^{(\mathrm{B})} = \boldsymbol{\Delta}^{(\mathrm{B})}\left(\boldsymbol{I} - \boldsymbol{\Omega}^{(\mathrm{B})}\right)^{-1}\boldsymbol{\Delta}^{(\mathrm{B})},$$

in which the matrix $\boldsymbol{\Omega}^{(\mathrm{B})}$ encodes the *between-persons network*.

[a] Unlike the contemporaneous and between-person networks, the temporal network is not standardized in this notation. Typically, all variables are standardized before analysis to make the $\boldsymbol{B}$ matrices interpretable (Bulteel et al., 2016). Alternatively, these coefficients can be standardized to *partial directed correlations* (Wild et al., 2010). Elements from $\boldsymbol{\Omega}_p^{(\mathrm{C})}$ are sometimes termed *partial contemporaneous correlations*, and elements from $\boldsymbol{\Omega}^{(\mathrm{B})}$ are sometimes termed *partial between-person correlations* (PBC).

Technical Box 10.1. Technical description of the graphical vector auto-regressive model.

Maximum likelihood estimation

Chapter 6 introduced maximum likelihood estimation as an approach to establish multivariate structures: parameters are found under which the data were most likely to occur. This process requires a *joint likelihood* expression of the entire data. For single measurement data, computing the joint likelihood is relatively straightforward: this quantity can be computed by multiplying (summing) the (log) likelihoods of every individual case—the rows in the data set. While proper maximum likelihood estimation of the GVAR model parameters from time series data is possible in principle as well, it becomes computationally much more challenging in practice compared to analyzing single measurement data (Ciraki, 2007). This is because an inherent property of time series data is that cases in the data set are no longer independent. As a result, the likelihood can no longer be formed easily. To this end, the covariance between every case needs to be modeled, resulting in a covariance matrix that contains a row/column for every variable of every case; in the case of 100 measures on 10 variables, the variance–covariance matrix modeled would be a 1,000 × 1,000 matrix, which is too large for most software packages to handle.

Data augmentation

To overcome the computational challenges of maximum likelihood estimation, GVAR estimation typically relies on a trick that involves augmenting the data (Hamaker et al., 2002; Lane et al., 2019). While this trick no longer results in 'true' maximum likelihood estimation, the resulting parameters and standard errors are comparable to true maximum likelihood estimates, especially at larger sample sizes. Figure 10.2 demonstrates how the data can be augmented: by making a copy of the data, shifting that copy by one row, and appending the shifted data set to the original data set. This way, each row t contains both responses at time t as well as the previous time $t - 1$. If the mean structure is not explicitly modeled, the data can also be centered. The variance–covariance matrix of this augmented data takes the form of a block Toeplitz matrix and can be modeled in the same manner as the variance–covariance matrix of single measurement data (see Technical Box 10.3 at the end of this chapter). Alternatively, regression models can be used on the

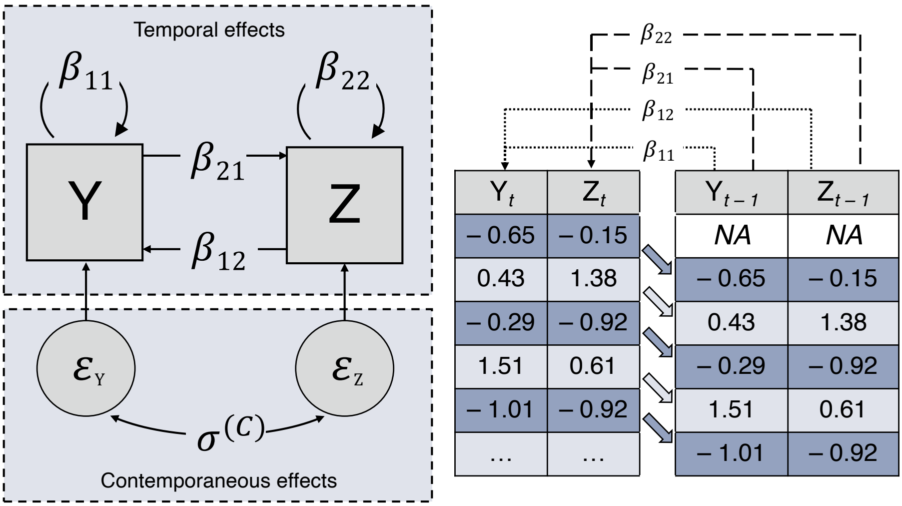

Figure 10.2. Example of a graphical vector auto-regression (GVAR) model (left) and data augmentation usually used to estimate model parameters (right). In the data augmentation, all variables (here Y and Z) are copied and shifted by one row (also termed *lagged*). Rows in the augmented data that cross a night or non-equal measurements can be removed before analysis. The temporal effects (β parameters) encode within-person prediction over time, and the contemporaneous effects (here $\sigma^{(C)}$) can be used to form a GGM encoding relationships in the same window of measurement. For interpretable parameters the data should also be within-person centered or the mean structure should be explicitly modeled.

augmented data: a multivariate regression model can be used with the set of responses at t as dependent variables and the set of responses at $t - 1$ as independent variables, resulting in the parameters of the temporal network. Subsequently, the residuals can be analyzed using the methods discussed in Chapter 6 to obtain the contemporaneous network. Another alternative is to use univariate estimation to obtain all temporal effects by regressing each variable on all lagged variables first, and subsequently, to obtain all contemporaneous effects by using univariate estimation tools on the residuals as discussed in Chapter 6.

Model selection

Model selection can be performed in similar manners as discussed in Chapter 7. For example, edges could be selected based on some threshold or through stepwise model selection search strategies, which has been implemented in the *psychonetrics* package, using mostly the same code as described in Chapter 6 and Chapter 7 except that the `gvar` model function is used instead of the `ggm` model function (Epskamp, 2020). A popular way in which the GVAR is estimated is through a regularization procedure closely related to the *EBICglasso* procedure used in GGM models. The multivariate regression with the covariance estimation (MRCE; Rothman et al., 2010) algorithm can be used to sequentially estimate a regularized temporal network (using LASSO regularization) and a contemporaneous network (using the GLASSO algorithm) until convergence. This algorithm utilizes two tuning parameters, one for the temporal coefficients and one for the contemporaneous coefficients, which can be selected using EBIC model selection (Abegaz & Wit, 2013). This algorithm has been implemented in the *graphicalVAR* package (Epskamp, 2021a) and the *SparseTSCGM* package (Abegaz & Wit, 2021). In Tutorial Box 10.1, we illustrate an example for estimating temporal and contemporaneous networks using *graphicalVAR* and *psychonetrics*.

Bayesian estimation

Another popular method for estimating VAR models (and, by extension, GVAR models) is through the use of multivariate Bayesian estimation by implementing the model in Bayesian sampling software such as JAGS (Plummer et al., 2003) or Stan (Carpenter et al., 2017). These software packages model a response vector as a function of the previous response vector by looping over the data when specifying the likelihood. To this end, the data need not be augmented as described above for other settings. Additional benefits of the Bayesian approach are that missing responses can easily be handled—even allowing for continuous time modeling (Ryan & Hamaker, 2021)—and that prior information could be used to improve estimation (Burger et al., 2021). The GVAR model has also been implemented in the BGGM package (Williams & Mulder, 2020).

Suppose a data frame `data` in R has the following form:

subject	day	beep	worry	relax	angry
1	1	1	2	1	2
1	1	2	2	2	2
1	1	3	3	1	1
1	2	1	2	1	1
1	2	2	2	3	1
⋮	⋮	⋮	⋮	⋮	⋮

We can estimate GVAR model parameters from this $N = 1$ data set using the R packages *graphicalVAR* (Epskamp, 2021a) for regularized estimation (Abegaz & Wit, 2013; Rothman et al., 2010) and *psychonetrics* (Epskamp, 2021b) for maximum likelihood estimation. The input to both packages is comparable, and requires information about the columns in the data to be stored first:

```
vars <- c("worry","relax","angry") # Variables used in the model
dayvar <- "day" # The day variable, only use with >1 assessment/day
beepvar <- "beep" # The beep variable
```

These objects correspond to argument names in the R packages. The `vars` argument specifies the variables used in the analysis, the optional `dayvar` argument specifies the days and is used to cut out pairs of measurements that cross a night[a], and the optional `beepvar`, corresponding to the measurement number within each day, can be used if the data contain missing measurements. Now, the *graphicalVAR* package can be used as follows:

```
library("graphicalVAR")
graphicalVAR(data, vars = vars, dayvar = dayvar, beepvar = beepvar)
```

The estimated network structures are stored as *partial directed correlations* (PDC) and *partial contemporaneous correlations* (PCC) for the temporal and contemporaneous networks respectively, and we can visualize the networks using the `plot` function. In *psychonetrics*, the `gvar` model function can be used:

```
library("psychonetrics"); library("dplyr")
gvar(data, vars = vars, dayvar = dayvar, beepvar = beepvar,
     estimator = "FIML") %>% runmodel
```

Optionally, further model search functions can be applied such as `prune` and `modelsearch`. The weights matrices can be obtained using the `getmatrix` function, with `omega_zeta` indicating the contemporaneous network and `PDC` the standardized temporal network.

[a]Importantly: do not use the `dayvar` argument with only one observation per day, as then all data will be removed!

Tutorial Box 10.1. Estimating $N = 1$ networks from time series data using *graphicalVAR* and *psychonetrics*.

10.4 *N* > 1 estimation: multi-level estimation

If intensive longitudinal data are available from multiple subjects, we might be interested in constructing a network of the average temporal and contemporaneous effects, the *fixed effects* network structures introduced in Chapter 9. A first, intuitive approach to estimating these fixed effects is to compute a network for each subject separately using the methods discussed above and subsequently calculate the averages for each parameter, as well as their variance–covariance structure. This approach of *pooling parameter estimates*, however, discards the nested structure of the data and relies on the—potentially underpowered—estimation of many (*N*) models separately. An alternative to estimating separate models per person is to estimate only one model for all observations. The most straightforward way to do this is to within-person center all variables[7], combine all within-person centered data sets, and estimate a single GVAR model (making sure that responses from one person are not regressed on responses from another person). This process has been automated in the `mlGraphicalVAR` function in the *graphicalVAR* package[8] and the `gvar` function in *psychonetrics* (if the `idvar` argument is used). These methods provide good estimates of the fixed effect structures, but still do not properly take nesting of data points into account. These methods also do not provide insight into the variability of parameters across the sample, as individual networks have to be estimated separately per person.

Multi-level modeling

We can actively incorporate the nested structure of our data by estimating a *multi-level* (G)VAR model (Bringmann et al., 2013; Epskamp, Waldorp, et al., 2018). The term multi-level refers to the data being organized in two levels: within-subject variance on level 1, and between-subject variance on level 2. In this approach, each parameter in the model (e.g., edge weights and means) is assumed to have a distribution over the population. Thus, in the estimation procedure, only these distributions (mean and variance of the parameter, and possibly the covariance between parameters) need to be estimated. The fixed effects can then be obtained from the centers of these distributions. Subsequently, the deviations from this center point—the *random effects*—can be sampled, which, together with the fixed effects, lead to estimates for the personal network models. Figure 10.3 shows how the parameter distributions of a multi-level VAR can inform fixed and random parameters in temporal networks.

Estimating GVAR models through multi-level estimation has four main benefits. First, a single analysis can be performed on the entire data set, leading to a well-powered analysis based on a large sample size, especially for the estimation of fixed effects. Second, the multi-level analysis provides not only insight into the fixed effects structure, but also into the heterogeneity around these fixed effects through the standard errors of the random

[7] For each variable removing for each person the mean of that person from the scores of the variable. This step is necessary to ensure that between-person effects are not included in the analysis.

[8] Unlike the name suggests the `mlGraphicalVAR` function does *not* perform multi-level estimation. The function merely computes a pooled GVAR over all combined within-person centered data sets and runs the `graphicalVAR` function per person separately for individual networks (Epskamp, Waldorp, et al., 2018).

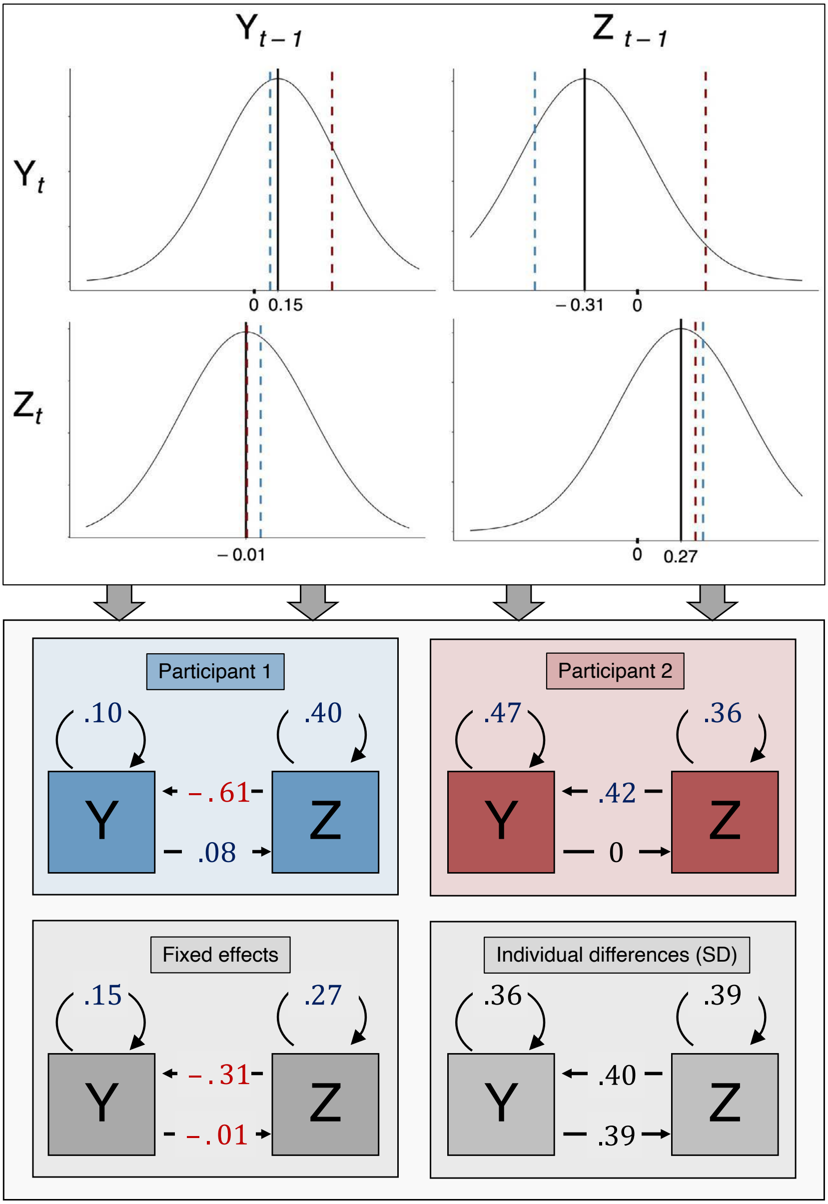

Figure 10.3. Multi-level model for the temporal effects of two variables *Y* and *Z*, with data simulated for $N = 150$. The distributions for the temporal effect parameters (top panel) are used to establish the random temporal effects for participant 1 and 2 (bottom panel, top row), as well as fixed temporal effects and the standard deviations of subjects on these effects (bottom panel, bottom row).

effects. Third, multi-level modeling can be used to separate within- and between-person variances, which also leads to estimates of the between-person structure. Finally, and perhaps most importantly, estimated individual network structures are typically closer to the fixed effects estimate compared to parameters that are estimated for that individual directly. This is termed *shrinkage*, as the estimates of different persons in the sample are 'shrunken' towards each other. In other words, individuals' effects are also informed by other individuals and result in estimates that lie close to each other. This can lead to better estimates of the personal network structures, requiring fewer observations per person than performing many $N = 1$ analyses separately.

Multi-level modeling also has some downsides. Assuming a (typical normal) distribution across the population on parameters entails that these parameters do not differ in *structure*, only in weight. For example, suppose that a temporal edge $A \to B$ is modeled with a normal distribution across the population with mean (fixed-effect) 0.2 and standard deviation (of the random-effect) 0.1. This means that we would expect roughly 95% of the population to have individual edge weights for $A \to B$ between 0 and 0.4, with the remaining 5% lower than 0 (negative edge weights) or higher than 0.4. The model, therefore, does not assume that any of these persons have an edge-weight of *exactly* 0, which would lead to the edge not being included in the temporal network. To this end, multi-level modeling does not estimate individual network structures, only individual network parameters. These parameters are also shrunken towards the fixed effect, which makes it questionable if the individual network structures genuinely allow for a within-person interpretation (after all, the networks were not estimated within the data of every person separately). Another prominent downside of multi-level modeling is that the models quickly become very complicated and computationally too challenging to be estimated. This is because a joint distribution over all parameters needs to be estimated. To this end, it is generally not possible to include many nodes in multi-level analyses. Multivariate estimation methods typically only allow for a few nodes to be included, which is why the main estimation method we will discuss uses univariate estimation. In univariate estimation, about six (with correlated random effects) to 20 (with uncorrelated/orthogonal random effects) can be included at most.

The two-step multi-level VAR algorithm

Univariate estimation—using sequential univariate multi-level models and combining the results—was first proposed for the multi-level VAR model in psychological literature by Bringmann et al. (2013). Epskamp, Deserno, et al. (2021) extended this approach for estimating GVAR models by separating within- and between-person effects (allowing for the estimation of between-person networks) and by estimating contemporaneous networks. This extension is termed *two-step multi-level GVAR* estimation (see Technical Box 10.2), and is implemented in the *mlVAR* package (Epskamp, Deserno, et al., 2021), further described in Tutorial Box 10.2. In step 1, the algorithm estimates the temporal and between-subjects networks by performing univariate multi-level modeling, predicting each variable from within-person centered lagged variables and personwise means. In step 2, the estimated residuals of the models run in step 1 are used in a new sequence of univariate multi-level models to estimate the contemporaneous effects. For separating within- and between-person variances, the algorithm makes use of within-person centering: using the sample means from every person separately to center variables. This requires decent estimates of the within-person means, meaning that several (at least about 20) measures have to be available per person.[9] To this end, two-step multi-level GVAR estimation can be used with $N > 1$ time series data, but not with panel data. The *mlVAR* package uses the *lme4* package (Bates et al., 2015) for multi-level estimation of all effects.

The *two-step multi-level* GVAR estimation algorithm, proposed by Epskamp, Waldorp, et al. (2018), is an algorithm for estimating multi-level GVAR models through a series of univariate multi-level regression analyses. First, the entire data set used is standardized to z-scores (subtracting the mean and dividing by the standard deviation). Let $z_{tp}^{(\text{worry})}$, $z_{tp}^{(\text{relax})}$, and $z_{tp}^{(\text{anger})}$ represent three variables answered by person p at measurement occasion t. To separate within- and between-person variances, we within-person center lagged variables as predictors (Hamaker & Grasman, 2014): $\tilde{z}_{t-1,p}^{(\dots)} = z_{t-1,p}^{(\dots)} - \bar{z}_{p}^{(\dots)}$, in which $\tilde{z}$ represents a within-person centered variable and $\bar{z}$ the person-wise mean. In step 1, for each variable a univariate multi-level regression is performed using that variable as a dependent variable and all within-person centered lagged variables together with the person-wise means as independent variables:

$$
\begin{aligned}
z_{t,p}^{(\text{worry})} = \beta_{0p} + \beta_{11p}^{(\text{T})} \cdot \tilde{z}_{t-1,p}^{(\text{worry})} + \beta_{12p}^{(\text{T})} \cdot \tilde{z}_{t-1,p}^{(\text{relax})} + \beta_{13p}^{(\text{T})} \cdot \tilde{z}_{t-1,p}^{(\text{anger})} \\
+ \beta_{12}^{(\text{B})} \cdot \bar{z}_{p}^{(\text{relax})} + \beta_{13}^{(\text{B})} \cdot \bar{z}_{p}^{(\text{anger})} + \varepsilon_{tp}^{(\text{worry})}.
\end{aligned}
$$

The $\beta^{(\text{T})}$ parameters form the individual temporal networks, and the $\beta^{(\text{B})}$ parameters can be used to form a GGM in the same way univariate regressions in univariate GGM estimation are averaged to partial correlation coefficients (Epskamp, Waldorp, et al., 2018). In the second step, the estimated residuals of the multi-level regression models in step 1 are used in a second round of univariate multi-level models:

$$
\hat{\varepsilon}_{tp}^{(\text{worry})} = \beta_{12p}^{(\text{C})} \cdot \hat{\varepsilon}_{tp}^{(\text{relax})} + \beta_{13p}^{(\text{C})} \cdot \hat{\varepsilon}_{tp}^{(\text{anger})} + \zeta_{tp}^{(\text{worry})}.
$$

The $\beta^{(\text{C})}$ parameters are subsequently used to form the contemporaneous networks.

Technical Box 10.2. Two-step multi-level graphical vector auto-regression.

[9] With too few observations per person, the estimated network structures will likely be biased. This bias is termed *Nickel's bias* (Jordan et al., 2020), and most notably leads to erroneous negative auto-regressions (self-loops in the temporal network).

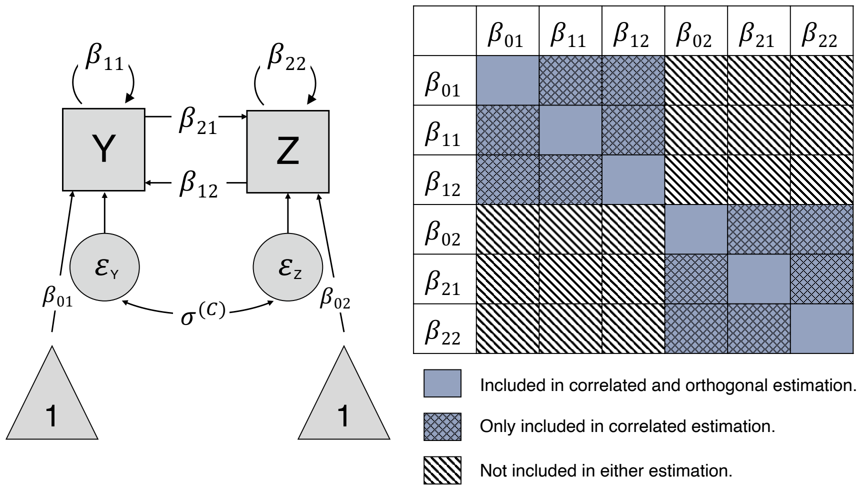

Figure 10.4. Parameter covariation included in *mlVAR* for the estimation of temporal effects. Choosing orthogonal estimation assumes that parameters are independent, whereas correlated estimation considers some (but not all) parameter correlations.

Parameter covariance. A challenging aspect of multi-level modeling is that often covariances between random effects need to be estimated as well. For example, it could be that people that have strong edges between some variables also tend to have strong edges between other variables (Pe et al., 2015). Estimating sequential univariate models, as is done in two-step multi-level GVAR estimation, provides a computationally efficient alternative to estimating the multi-level GVAR model because the univariate multi-level models do not include all parameters; the estimation routine does not have to include all potential covariances between random effects.[10] More specifically, univariate models only include the intercept and (incoming) edge-weights that are connected to the dependent variables. In *mlVAR*, the covariance between these included random effects can be estimated by using the arguments `temporal` and/or `contemporaneous`. The default (`"correlated"`) will include correlated random effects, which is feasible for up to about 8 to 10 nodes. For about 10 to 20 nodes, uncorrelated (`"orthogonal"`) random effects can be used, which introduces a limitation to the estimation procedure, as random effects can be assumed to be correlated. Figure 10.4 shows that estimating models with no correlated random effects will only assume some random effect covariances to be 0, not all (as not all random effects are included in each model).

[10]The upside of not having to estimate covariances between random effects also comes with the downside of not being able to investigate these covariances. This is why the between-person effects are estimated through level 2 predictors rather than by studying the random effects covariances between means.

We can estimate the multi-level VAR model from $N > 1$ time series data using the *mlVAR* package and from panel data using the *psychonetrics* package (fixed effect networks only). Both packages can handle data as structured in Tutorial Box 10.1. The *mlVAR* package uses the `mlVAR` function, which can be used as follows (assuming a column in the data containing information on the subject id is called `subject`):

```
library("mlVAR")
mlVAR_results <- mlVAR(data, vars = vars, idvar = "subject",
  temporal = "correlated", contemporaneous = "correlated")
```

Optionally the `dayvar` and `beepvar` arguments can be used which work similarly as in Tutorial Box 10.1. For non-correlated random effects, the `temporal` and `contemporaneous` arguments can be set to `orthogonal`. The plot method can be used to threshold and visualize the network. For example, the following command plots the temporal network with non-significant effects hidden:

```
plot(mlVAR_results, "temporal", nonsig = "hide")
```

Replacing `plot` for `getNet` will return the weights matrix instead. Networks showing the standard deviation of random effects can be obtained by setting `SD = TRUE` in `plot(...)` or `getNet(...)`.

In *psychonetrics* the model can be estimated using the `ml_gvar` function:

```
library("psychonetrics"); library("dplyr")
ml_gvar(data, vars = vars, idvar = "subject", standardize = "z") %>%
runmodel
```

Optionally, the `beepvar` argument can be used (the `dayvar` argument is not supported because this model is not designed for intensive time series), and further model search functions can be applied such as `prune` and `modelsearch`. Standardizing data is recommended to improve estimation. The contemporaneous network is stored as `omega_zeta_within` and the between persons network is stored as `omega_zeta_between`. If data are encoded in a wide format (variables encoded as different columns for each measurement), the `panelgvar` function can be used instead. Both `ml_gvar` and `panelgvar` are wrapper functions on the main `dlvm1` function used for panel data modeling, which allows for some more options (e.g., modeling between-person effects as a Cholesky decomposition, which can be useful if between-person networks are seemingly not estimated well).

Tutorial Box 10.2. Estimating multi-level GVAR models from $N > 1$ longitudinal data.

Multivariate estimation

Panel data. A considerable downside of the two-step multi-level GVAR estimation algorithm is that due to within-person centering with the person-wise sample means, a decent number of observations per person is required. To this end, it is not recommended to use this algorithm with less than about 20 observations, making it applicable to $N > 1$ time series data but not to panel data. One estimation method for estimating a GVAR model from panel data (also termed *panel GVAR*) has been proposed by Epskamp (2020) and is implemented in the *psychonetrics* package, further described in Tutorial Box 10.2.[11] This model is a multi-level GVAR model with only random intercepts/means. This means that it assumes the same network structure for every person but allows people to differ on their averages. The variance–covariance structure of these random means is used to model the between-person network. The implementation in *psychonetrics* is a full-information implementation, meaning that all covariances between every possible measurement are included in the model. This makes the model computationally challenging to use with many nodes or many time points. It is recommended not to use this model with more than 10 measurements and more than around 10 to 20 nodes.

Bayesian estimation. A final powerful method for multivariate multi-level (G)VAR estimation is Bayesian estimation through sampling procedures. In these frameworks, all effects can be random and are included in the same model. As such, these methods return all possible random effect correlations (e.g., also allowing for between-person networks to contain edges). While it is possible to implement this model manually in software such as JAGS (Plummer et al., 2003) or Stan (Carpenter et al., 2017), doing so is quite challenging, requiring many prior distribution choices and likely leading to long computations if more than a few nodes are modeled. The Mplus software includes a module on dynamic structural equation models from version 8 onwards, which simplifies this process (Asparouhov et al., 2018; McNeish & Hamaker, 2020). This framework accommodates the multi-level VAR model, and while modeling GGMs is not included, partial correlation coefficients can manually be obtained from posterior samples (Epskamp, Waldorp, et al., 2018). The main downside is that even though the implementation in Mplus is very powerful, the number of nodes that can realistically be included in the analyses is still quite limited (about six). Another downside is that Mplus is not open-source and, therefore, not free to use. A more detailed discussion on differences between Bayesian estimation and the two-step multi-level GVAR algorithm can be found in Epskamp, Waldorp, et al. (2018).

10.5 Challenges to GVAR estimation

In this section, we discuss some of the most prominent practical and methodological challenges that researchers may face when estimating GVAR models from data.

Power and feasibility

The required number of observations to estimate reliable networks from time series data of a single subject ($N = 1$) is at least comparable to the number of participants needed to estimate networks from single measurement data. In fact, the required number may even be higher, as the GVAR model includes a temporal network and is estimated from

[11] The panel GVAR model is implemented as a special case of a larger modeling framework that also includes latent variables. Epskamp (2020) termed this model the *panel-LVGVAR* model, and the *psychonetrics* package terms this model the `dlvm1` (dynamic latent variable model with lag-1) model. The panel GVAR can be obtained by representing each observed variable with a latent variable, setting all factor loadings to 1, and all residual variances to 0. This is done automatically in the *psychonetrics* package if no latent variable structure is assigned.

data with auto-correlated responses (reducing effective sample size). Collecting large time series for a single subject, however, is challenging in most fields of psychological research. Not only can it be burdensome for participants, but extending the measurement to long periods may also hinder the assumption of stationarity (see below). Furthermore, the number of time points needed depends on the estimated network structure and on the number of nodes included. Sparse and well-defined network structures containing a few strong edges can be retrieved with smaller samples than dense networks with fewer strong edges that stand out.

In a series of simulation studies, Epskamp, Waldorp, et al. (2018) showed that reliable estimation of sparse synthetic networks is possible with 100 time points and eight nodes (Epskamp, Waldorp, et al., 2018). However, a recent simulation study by Mansueto et al. (2021) used empirical networks as generating structures and instead found a relatively poor sensitivity (power to detect edges) with around 100 observations. A different generating network with six nodes led to better recovery of the global structure at 100 observations, but weaker edges were still not reliably retrieved. Such weak edges do not necessarily represent small and negligible effects; edges may be weak because of sampling bias or slight variations of the variables in time and may still be relevant for research or clinical purposes. Unfortunately, there is no way to know if the generating structure (assuming data were generated through a GVAR model) was dense or sparse and included strong or weak edges. As such, it is questionable if GVAR estimation is feasible from $N = 1$ data sets that may realistically be obtainable, and this will rely on certain assumptions (notably, that the generating model is sparse). It is advisable to consider that with about 50 to 100 time points, sensitivity likely is low, meaning that only a few edges may be discovered. The best solution to this problem, outside of aiming to collect more data, is to keep the model as small as possible. For $N = 1$ GVAR models, it is generally advisable to include as few nodes as possible (e.g., less than 10).

Heterogeneity

In addition to discovering individual network structures, researchers may also be interested in how much people differ in their network structure (heterogeneity). The detection of heterogeneity between GVAR models is directly related to the reliability of GVAR estimation. If it is not feasible to estimate reliable network structures, visually comparing network structures of individuals may lead to an illusionary sense of heterogeneity. Hoekstra et al. (2021) discuss that even if the generating structure is the same for two people, network structures estimated from their data may differ substantively. For example, suppose that the generating model contains 10 (true) edges, but sensitivity (power) is only 50%, meaning that we only expect to find 5 out of 10 edges in the network of one particular person. Suppose also that the chance of including an edge is the same for all 10 true edges. Then, there is only a 0.000016 probability that the *exact* same edges are detected in two people. As such, even though the generating structure is the same, we would expect to find different networks. This entails high sensitivity (and specificity) are needed to separate true from illusionary heterogeneity when estimating individual network models, and to this end, it is advisable to not interpret differences in personal network models as evidence for heterogeneity, especially when these networks are sparse.

If a large number of people are included in the data set, multi-level modeling can be used to gain insight in the heterogeneity of parameter values. When estimating a multi-level network using the *mlVAR* package, the standard deviations of random effects across the population on the temporal and contemporaneous network parameters are returned and can be visualized as networks (e.g., Figure 10.3). The width of the edges in this network shows the degree to which network parameters exhibit individual differences. Bringmann et al., 2013 recommend using a cut-off score of 0.10 for the edge weights. Alternatively, random effects can in principle be tested statistically by comparing a model with random effects to a fixed effects only model, although this may be hard in practice.[12]

Missing data

In intensive time series designs missing data are very common. Usually, time series data are characterized by wave missingness, where every item at a particular measurement point is missing (McLean et al., 2017; Schafer & Graham, 2002). Many factors can affect missing data, for example, measurement frequency and timing, length of the measurement period, physical activity, substance use, age, and gender (Jones et al., 2019; McLean et al., 2017; Ono et al., 2019; Rintala et al., 2019; Wen et al., 2017). Techniques based on imputation or maximum likelihood can be used to handle data missing completely at random (MCAR) and at random (MAR), while with data missing not at random (MNAR), these may yield biased estimates. For example, Kalman filter imputation (Hamaker & Grasman, 2012; Harvey, 1990) can be applied prior to network estimation with the R package *imputeTS* (Moritz & Bartz-Beielstein, 2017). Alternatively, full information maximum likelihood estimation is implemented in the R package *psychonetrics*, which estimates a model using only observed responses (Epskamp, Isvoranu, et al., 2021). Finally, Bayesian estimation methods are well capable of handling missing data (McNeish & Hamaker, 2020). Mansueto et al. (2021) propose that such methods for handling missing data may lead to promising avenues for reducing participant burden through the use of planned missingness and adaptive testing—only asking a subset of questions in each measurement (Graham et al., 2006).

Stationarity assumption

As any statistical model, (multi-level) GVAR estimation relies on a set of assumptions. A core assumption of the VAR model is *stationarity*. A stationary time series does not indicate changes over time in its defining characteristics, such as the means, variances, and network parameters. Violations of this assumption arise if there are trends in the time series, for example, seasonal or linear trends or changes in volatility. Deviations from the stationarity assumption are not implausible in psychological time series. For example, we might observe mean-shifts in symptoms following certain life events or obtain seasonal patterns for affect variables depending on the time of the year. To understand the dynamic process without such trends, a time series can be broken down into its constituent trend components through a method called *decomposition*, resulting in a trend component, a seasonal/cyclical component, and a residual/regular component. Another scenario

[12] A limited implementation for testing random effects of temporal coefficients is implemented in *mlVAR* in the `mlVARcompare` function.

introducing non-stationarity is the presence of a so-called *unit root*. A unit root is present if the auto-regressive parameter of a time series equals one and can be detected using the (Augmented-)Dickey-Fuller test (Dickey & Fuller, 1979). Figure 10.5 visualizes a stationary distribution, as well as different cases of non-stationarity.

There are several methods of handling non-stationarity. Generally, most of these methods aim to remove existing linear trends or seasonal components. For example, linear trends in the data can be accounted for by performing a regression on time and subsequently modeling the residuals as the time series adjusted for linear changes. Simulation studies showed that it is generally recommended to detrend present linear trends before estimating networks (Epskamp, van Borkulo, et al., 2018). In this study, detrending all versus only significant trends performed comparably, while not detrending led to lower specificity (especially in temporal networks) and lower sensitivity (especially in contemporaneous networks). However, in many situations changing means, variances, or network parameters are of central interest, and therefore removing them from the data through detrending would be detrimental. Instead, these changes over time can be explicitly modeled in time-varying network models, which are further discussed in Chapter 11.

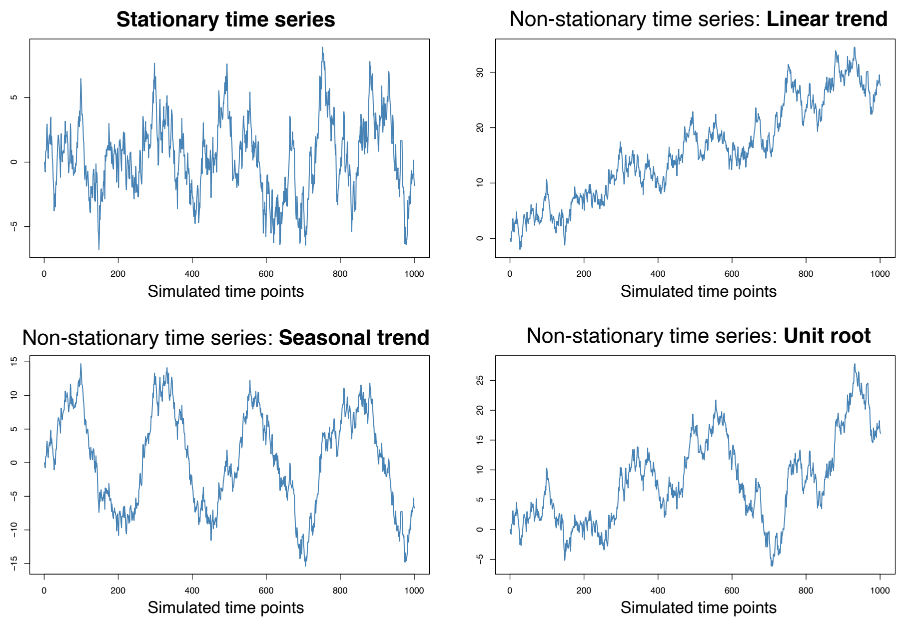

Figure 10.5. Simulated time series data under four conditions. Top left: A stationary time series with auto-regressive effect $\beta = 0.95$. Top right: Non-stationarity, due to a linear trend added to the time series used for the first plot. Bottom left: Non-stationarity due to a seasonal trend added to the time series used for the first plot. Bottom right: Non-stationarity due to the presence of a unit root, time series with auto-regressive effect $\beta = 1$.

Assumption of equidistant measures

The GVAR model establishes temporal dependencies in the form of lagged relationships. In doing so, it treats all lags of the same level as equally distant from one another. In other words, it is assumed that the time difference between any two subsequent assessment points is equal, an assumption referred to as *equidistant measures*. Two primary manners in which this assumption may be violated are (1) when there are multiple measures per day, because there will likely be a larger time difference between the last assessment of a day and the subsequent day's first assessment compared to time differences between other consecutive measurements, and (2) when a participant failed to fill in all measurements, and the data are not properly encoded with missing values on the measurements that were not filled in. These problems can adequately be handled using the software packages discussed in this chapter through the use of the `dayvar` argument (removes pairs of observations that cross a night) and `beepvar` arguments (removes pairs of responses that are not consecutive). Another way this assumption may be violated is (3) when measurement occasions are at random time intervals. While it should not be a big problem if time intervals are roughly equal (e.g., sometimes two hours and sometimes three hours), it may be problematic if time intervals show large differences (e.g., sometimes 10 minutes and sometimes four hours). In this setting, an alternative is to use continuous time modeling, further discussed by Ryan and Hamaker (2021).

Time scales

In the GVAR model, we aim to predict dynamics as lagged relationships, typically including only one time lag. Consequently, to interpret temporal effects, we need to make sure that the time scale chosen for our analysis matches the type of dynamics we want to investigate. For example, if we want to model a temporal effect A → B, we want to ensure that we also capture this effect by appropriately timing our assessment intervals. This, however, is not always possible or feasible; in many cases, we either do not know the true time scale our processes are operating at, or it is not feasible to measure at the desired frequency. A mismatch between true time scale and modeled lags can lead to problematic inferences in two situations: first, the true dynamic process can unfold *faster* than specified in our assessment (e.g., panic symptoms occur within seconds, but we measure every two hours). In this case, the effects will not be captured in the temporal prediction. Such fast effects might be found in the contemporaneous rather than the temporal effects (Epskamp, van Borkulo, et al., 2018). An alternative approach to modeling discrete time lags is to conceptualize dynamics on a continuous level, for example, using continuous structural equation modeling (Driver et al., 2017; Ryan & Hamaker, 2021) or differential equations. Second, the true dynamic process can unfold *slower* than specified in our assessment (e.g., investigating mood dynamics in relation to hormones, but we assess hormone levels every two hours). Such slower effects might be better understood using panel designs (Epskamp, 2020) because they require more distance between assessments.

Epskamp (2020) detail multivariate estimation of the GVAR model from $N = 1$ time series and panel data. To estimate a GVAR model, the data can first be augmented:

$$\boldsymbol{Y}^{(\text{aug})} = \begin{bmatrix} \boldsymbol{Y}^{(\text{lag})} & \boldsymbol{Y} \end{bmatrix},$$

in which $\boldsymbol{Y}^{(\text{aug})}$ represents the augmented data set, $\boldsymbol{Y}$ represents the original data (with measurement t on row t), and $\boldsymbol{Y}^{(\text{lag})}$ the original data set shifted by one row (measurement $t-1$ on row t). If needed, several rows of $\boldsymbol{Y}^{(\text{aug})}$ can be removed, especially when the pair of measurements $t-1$ and t feature a large gap in time, such as across a night. The variance–covariance matrix of $\boldsymbol{Y}^{(\text{aug})}$ takes the following form:

$$\boldsymbol{\Sigma} = \begin{matrix} t-1 \\ t \end{matrix} \overset{\begin{matrix} t-1 & t \end{matrix}}{\begin{bmatrix} \boldsymbol{\Sigma}^{*} & \boldsymbol{\Sigma}_1^{\top} \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0. \end{bmatrix}}$$

Also termed the Toeplitz variance–covariance matrix. The block $\boldsymbol{\Sigma}_0$ can be modeled with the following expression:

$$\text{Vec}\,(\boldsymbol{\Sigma}_0) = (\boldsymbol{I} \otimes \boldsymbol{I} - \boldsymbol{B} \otimes \boldsymbol{B})^{-1}\,\text{Vec}\left(\boldsymbol{\Sigma}^{(\text{C})}\right),$$

in which $\boldsymbol{\Sigma}^{(\text{C})}$ is the innovation variance–covariance matrix that can further be modeled as a GGM (see Technical Box 10.1). The lag-1 variance–covariance matrix can subsequently be modeled as:

$$\boldsymbol{\Sigma}_1 = \boldsymbol{B}\boldsymbol{\Sigma}_0.$$

Finally, with large samples we would expect $\boldsymbol{\Sigma}^{*} = \boldsymbol{\Sigma}_0$. In the *psychonetrics* package, however, $\boldsymbol{\Sigma}^{*}$ is modeled using a Cholesky decomposition instead:

$$\boldsymbol{\Sigma}^{*} = \boldsymbol{L}\boldsymbol{L}^{\top},$$

such that this block is always positive semi-definite and such that the stationary variance–covariance structure is not modeled twice. This Technical Box explains the main expressions used in maximum likelihood estimation of $N = 1$ GVAR models. The variant for panel data follows mostly the same steps, but creates a larger Toeplitz matrix with all waves of data and models between-person variance in addition to the within-person variances discussed here.

Technical Box 10.3. Toeplitz variance–covariance structure for the graphical VAR model.

10.6 Conclusion

Time series analysis is a fruitful field for constructing dynamical networks. The graphical vector auto-regression (GVAR) model separates longitudinal information in *contemporaneous*, *temporal*, and *between-persons* network structures. This chapter discussed how these networks could be estimated from time series analyses for single subjects and multiple subjects, using intensive longitudinal data collected via novel ambulatory assessment techniques or panel data. Current challenges in estimating time series networks span from a trade-off between power, stationarity, and feasibility to identifying appropriate time scales for lagged relationships.

While this chapter provides an introduction to network analysis from longitudinal data, the topic of longitudinal data analysis itself goes beyond the scope of this book. Indeed, entire textbooks could be written on this topic. Important to note is that the GVAR model, which was the focus of this chapter, is only one of several possible models. Another approach to constructing networks from time series data is the estimation of a structural VAR model (SVAR; Chen et al., 2011; Gates et al., 2010). In contrast to GVAR, the SVAR model uses *directed* effects for the contemporaneous network. Structural VARs can be estimated by transforming (G)VAR results (Lütkepohl, 2005) or through unified structural equation modeling (Beltz & Molenaar, 2016; Gates et al., 2010; Kim et al., 2007), and structure estimation is usually done through stepwise model search. In $N > 1$ data, 'group iterative multiple model estimation' GIMME; (GIMME; Gates & Molenaar, 2012) is an often-used method for estimating structural VAR models for multiple persons. In short, GIMME searches for qualitative similarity across people—using stepwise model search strategies through structural equation models—to find network structures that contain group-level (edges that are included for every person in a group) as well as person-specific temporal and contemporaneous effects. Other variants of network models estimated from time series data are time-varying VAR networks and VAR networks that include non-Gaussian variables. These will be introduced in Chapter 11.

10.7 Exercises

Conceptual

10.1. Explain why pairs of observations that cross a night have to be removed after augmenting the data.

10.2. Suppose you have $N = 1$ data of a person measured once per day on weekdays only for several months. Which argument in *graphicalVAR* or *psychonetrics* can be used to make sure you do not regress the responses from Mondays on the responses from Fridays?

10.3. How can *shrinkage* help in estimating a network for each subject in a data set?

10.4. Describe the difference between *fixed effects* and *random effects*.

10.5. Suppose a researcher has two time series data sets of patients measured in clinical practice, both with 75 measurements in total. The researcher estimates two GVAR models, and finds that these network models are different due to having different edges. Can the researcher conclude that there is *heterogeneity* between these people? Why (not)?

True or false

10.6. In the second step of two-step multi-level GVAR estimation the between-person network parameters are estimated.

10.7. Using multi-level estimation, for every person in the data set, a temporal and contemporaneous network can be estimated but not a between-persons networks.

10.8. It is generally recommended to remove trends (such as linear trends) prior to analyzing your $N = 1$ time series.

10.9. *Correlated* estimation in *mlVAR* includes all parameter correlations.

10.10. Contrary to graphical VAR, structural VAR estimation estimates directed contemporaneous effects.

Practical

For practical exercises in R, please navigate to the appropriate folder of this chapter, available on the online *Companion Website*.

References

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

Abegaz, F., & Wit, E. (2021). *SparseTSCGM: Sparse time series chain graphical models* [R package version 4.0]. https://CRAN.R-project.org/package=SparseTSCGM

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Beltz, A. M., & Molenaar, P. C. M. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate Behavioral Research*, *51*(2-3), 357–373.

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS One*, *8*(4), e60188.

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Using raw VAR regression coefficients to build networks can be misleading. *Multivariate Behavioral Research*, *51*(2-3), 330–344.

Burger, J., Epskamp, S., van der Veen, D. C., Dablander, F., Schoevers, R. A., Fried, E. I. I., & Riese, H. (2021). A clinical PREMISE for personalized models: Towards a formal integration of case formulations and statistical networks. *PsyArXiv*. https://doi.org/10.31234/osf.io/bdrs7

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32.

Chen, G., Glen, D. R., Saad, Z. S., Hamilton, J. P., Thomason, M. E., Gotlib, I. H., & Cox, R. W. (2011). Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis. *Computers in Biology and Medicine*, *41*(12), 1142–1155.

Ciraki, D. (2007). *Dynamic structural equation models: Estimation and interference.* (Doctoral dissertation). London School of Economics and Political Science (United Kingdom).

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427–431.

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).

Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *Journal of Econometrics*, *137*(2), 334–353.

Epskamp, S. (2021a). *graphicalVAR: Graphical VAR for experience sampling data* [R package version 0.3]. https://CRAN.R-project.org/package=graphicalVAR

Epskamp, S. (2021b). *psychonetrics: Structural equation modeling and confirmatory network analysis* [R package version 0.10]. https://CRAN.R-project.org/package=psychonetrics

Epskamp, S., Isvoranu, A. M., & Cheung, M. (2021). Meta-analytic Gaussian network aggregation. *Psychometrika*. https://doi.org/10.1007/s11336-021-09764-3

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A. M., Riese, H., & Cramer, A. O. J. (2018). Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections. *Clinical Psychological Science*, *6*(3), 416–427.

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480.

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231.

Epskamp, S., Deserno, M. K., & Bringmann, L. F. (2021). *Mlvar: Multi-level vector autoregression* [R package version 0.5]. https://CRAN.R-project.org/package=mlVAR

Fried, E. I., Papanikolaou, F., & Epskamp, S. (2021). Mental health and social contact during the COVID-19 pandemic: An ecological momentary assessment study. *Clinical Psychological Science*. https://doi.org/10.1177/21677026211017839

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319.

Gates, K. M., Molenaar, P. C., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage*, *50*(3), 1118–1125.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*(4), 323–343.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, *37*(3), 424–438.

Hamaker, E. L., & Grasman, R. P. P. P. (2014). To center or not to center? investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, *5*, 1492.

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). The Guilford Press.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2002). On the nature of SEM estimates of ARMA parameters. *Structural Equation Modeling*, *9*(3), 347–368.

Hamaker, E. L., & Grasman, R. P. P. P. (2012). Regime switching state-space models applied to psychological processes: Handling missing data and making inferences. *Psychometrika*, *77*(2), 400–422.

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

Hoekstra, R. H. A., Epskamp, S., & Borsboom, D. (2021). Heterogeneity in individual network analysis: Reality or illusion? *Manuscript in review*.

Jones, A., Remmerswaal, D., Verveer, I., Robinson, E., Franken, I. H., Wen, C. K. F., & Field, M. (2019). Compliance with ecological momentary assessment protocols in substance users: A meta-analysis. *Addiction*, *114*(4), 609–619.

Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology*, *76*(9), 1591–1612.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, 513.

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*, *28*(2), 85–93.

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*, *24*(1), 54–69.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Mansueto, A. C., Wiers, R., van Weert, J. C. M., Schouten, B. C., & Epskamp, S. (2021). Investigating the feasibility of idiographic network models. *Psychological Methods*. https://doi.org/10.1037/met0000466

McLean, D. C., Nakamura, J., & Csikszentmihalyi, M. (2017). Explaining system missing: Missing data and experience sampling method. *Social Psychological and Personality Science*, *8*(4), 434–441.

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in mplus. *Psychological Methods*, *25*(5), 610–635.

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *R Journal*, *9*(1), 207–218.

Ono, M., Schneider, S., Junghaenel, D. U., & Stone, A. A. (2019). What affects the completion of ecological momentary assessments in chronic pain research? an individual patient data meta-analysis. *Journal of Medical Internet Research*, *21*(2), e11398.

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., Mata, J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., et al. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science*, *3*(2), 292–300.

Plummer, M. et al. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, *124*(125.10), 1–10.

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2019). Response compliance and predictors thereof in studies using the experience sampling method. *Psychological Assessment*, *31*(2), 226–235.

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

Ryan, O., & Hamaker, E. L. (2021). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*. https://doi.org/10.1007/s11336-021-09767-0

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Wen, C. K. F., Schneider, S., Stone, A. A., & Spruijt-Metz, D. (2017). Compliance with mobile ecological momentary assessment protocols in children and adolescents: A systematic review and meta-analysis. *Journal of Medical Internet Research*, *19*(4), e132.

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.

Williams, D. R., & Mulder, J. (2020). BGGM: Bayesian Gaussian graphical models in R. *Journal of Open Source Software*, *5*(51), 2111.