Journal of Psychosomatic Research 137 (2020) 110191

Contents lists available at ScienceDirect

Journal of Psychosomatic Research

journal homepage: www.elsevier.com/locate/jpsychores

Review article

Time series analysis of intensive longitudinal data in psychosomatic research: A methodological overview

Sigert Ariens*, Eva Ceulemans, Janne K. Adolf

KU Leuven, Quantitative Psychology and Individual Differences, Faculty of Psychology and Educational Sciences, Leuven 3000, Belgium

ARTICLE INFO

Keywords:
Time series analysis
Intensive longitudinal data
Vector autoregressive modeling

ABSTRACT

Time series analysis of intensive longitudinal data provides the psychological literature with a powerful tool for assessing how psychological processes evolve through time. Recent applications in the field of psychosomatic research have provided insights into the dynamical nature of the relationship between somatic symptoms, physiological measures, and emotional states. These promising results highlight the intrinsic value of employing time series analysis, although application comes with some important challenges. This paper aims to present an approachable, non-technical overview of the state of the art on these challenges and the solutions that have been proposed, with emphasis on application towards psychosomatic hypotheses. Specifically, we elaborate on issues related to measurement intervals, the number and nature of the variables used in the analysis, modeling stable and changing processes, concurrent relationships, and extending time series analysis to incorporate the data of multiple individuals. We also briefly discuss some general modeling issues, such as lag-specification, sample size and time series length, and the role of measurement errors. We hope to arm applied researchers with an overview from which to select appropriate techniques from the ever growing variety of time series analysis approaches.

1. Introduction

Although psychological and somatic variables are often closely related within individuals (e.g. [8,88,160]), reviews of the literature almost exclusively discuss between-person research [133,148]. This discrepancy is problematic since findings between individuals, such as correlations computed across individuals, are generally not informative of how phenomena evolve over time within individuals (e.g. [68,112,114]). Moreover, recent research assessing these individual-level processes have highlighted that their precise nature can display considerable heterogeneity over individual cases [61,92,164,181].

In the domain of emotion research, the between versus within individuals debate has prompted a paradigm shift from between-subjects analysis at single, or limited numbers of time points, to investigating dynamical relationships between affective processes by employing intensive longitudinal sampling schemes, where dozens to thousands of data points are collected over time, within individuals (e.g. [102]). When subsequently analyzing this intensive longitudinal data (ILD) with statistical tools sensitive towards the serial dependence assumed to underlie the time series of data, a rich description of the temporal unfolding of emotional processes becomes available.

The field of psychosomatic research also appears particularly well-suited to embrace this novel paradigm and its tools: Many hypotheses in the domain imply some form of temporal dynamics, which could be assessed rigorously when appropriate study designs and statistical methodologies are used. For example, Ghiggia et al. [65] hypothesize that the prevalence of psychological distress in patients suffering from fibromyalgia is the result of a vicious cycle, or dynamical relationship, between psychological symptoms and psychosomatic manifestations of chronic pain. von Leupoldt et al. [172] call for an assessment of the relationship between asthma exacerbations and affective states in naturalistic settings. Researchers have also emphasized the need to assess the time course of intervention and treatment strategies [177], an example of which is the temporal profile of central sensitization processes in relation to different treatment and intervention strategies for chronic pain [110]. Here, a dynamic paradigm would provide concrete ways to assess patient-tailored approaches to therapy and intervention (e.g. [134]), often called for by researchers in the field.

As many others (e.g. [100,134,163]), we believe that time is ripe for committing to a truly dynamic psychosomatic research agenda. However, adequately applying the dynamic paradigm and its methods requires informed decisions on challenges related to data analysis. Therefore, this paper offers an approachable overview of five core challenges and the solutions that have been proposed in the

* Corresponding author.
E-mail address: sigert.ariens@kuleuven.be (S. Ariens).

https://doi.org/10.1016/j.jpsychores.2020.110191
Received 24 January 2020; Received in revised form 13 June 2020; Accepted 5 July 2020
0022-3999/ © 2020 Elsevier Inc. All rights reserved.

methodological literature.

The structure of this paper is as follows. In the first section, properties of ILD will be discussed. The second section will serve as a didactic introduction in the dominant modeling techniques for time series analysis (TSA) of ILD. In the third section, we elaborate on five identified topics, the proposed solutions, and provide directions for implementation. Specifically, we will discuss issues related to measurement intervals, the variables used in analysis, modeling stable and changing processes, concurrent relationships, and extending TSA to incorporate the data of multiple individuals. The discussion consists of a summary of the five discussed methodological problems and solutions, and an elaboration on some additional issues regarding model specification and data requirements.

2. Ingredients of the dynamic paradigm

2.1. Intensive longitudinal data

ILD consist of a large amount of repeated measures within individuals over time. They are most often obtained by means of repeated self-reports in ecologically valid contexts, often termed experience sampling method (ESM), ecological momentary assessment (EMA), or real time data collection (RTDC) [38,55,118,147,153,157]. However, ILD can also be collected in experimental studies (e.g. [120]), and may also involve other types of data than self-reports, such as physiological [117] or observational [80] data. The distinction between ILD and classical longitudinal designs is a matter of degree. Classical longitudinal designs have a limited number of measurement occasions, providing only a rough assessment of time-dependent processes [115], while ILD emphasizes a large number of repeated measures per individual [176], allowing dynamical relationships to be investigated in far more detail. Finally, the choice of which variables are included in an ILD study should primarily rest on theory and hypotheses about the processes under investigation.

As with more traditional designs, study design and data-analytic considerations are often closely related. This paper is primarily intended to elucidate data-analytic issues when conducting ILD studies, but we refer readers to Bolger et al. [16], Burton et al. [29], Collins [42], Ebner-Priemer and Trull [55], Piccirillo et al. [128], and Trull and Ebner-Priemer [128,158] for overviews of and guidelines for design-related issues. To illustrate how closely design and analysis are related we discuss a concrete example. Chronic pain in individuals suffering from spinal chord injury has been found to remain relatively stable over time [88]. Over minutes however, the level of experienced chronic pain may fluctuate around a stable mean in tandem with fluctuating attention towards nociceptive stimuli [100]. Investigating whether the fluctuations are centered around a relatively stable mean, or display trends or cycles, requires models which translate these theoretical assumptions, and to study such rapidly fluctuating processes demands measurement on a fine temporal scale. While ESM designs may be well suited to study the time course of chronic pain following spinal chord injury over days or weeks, experimental designs may be excel at revealing the dynamics between chronic pain and attention, allowing the inclusion of neuro-imaging techniques and explicit attentional manipulations. Other variables included in the study can be considered stable, fixed effects, such as the site at which pain is experienced, or dynamic, such as attention or psychological covariates such as pain catastrophizing.

2.2. Time series analysis

Many time series methods have in common the assumption that multiple variables display (mutual) dependence through time. That is, a variable X_1 at time t can to some extent be predicted by that same variable at time $t-1$, or by another variable X_2 at time $t-1$. For example, rumination at time $t-1$ may be predictive of itself and future social anxiety at time t. In turn, also social anxiety may affect itself, and rumination over time. Often, such relationships are of prime interest from a theoretical perspective, yet many researchers opt to apply classical analysis approaches which discount the very temporal dependence which endows ILD with its appeal [178].

By far the most prevalent methods used to capture these temporal dependencies are those belonging to the family of *vector autoregressive models* (VAR; see for instance [75,108], for more technical detail). VAR models do not only match our above example on rumination and anxiety, but also form the basis of many other methods. The key idea is the following: By including lagged versions of one or multiple variables (i.e., the values of the variables at previous time points) as predictors in a linear regression model, classical linear regression becomes applicable to time-dependent processes. Whereas autoregressive (AR) models are univariate, describing only one process over time, multivariate VAR models pertain to multiple interacting processes.

An AR model is given by:

$$y_t = \delta + \sum_{k=1}^{K} \phi_k y_{t-k} + \varepsilon_t \quad (1)$$

where t denotes time, and proceeds in one-unit steps, assuming values of 1, 2, …, T, and implying that intervals between subsequent time points are equal The variable y_t at time t then depends on the K previous values y_{t-1}, y_{t-2}, …, y_{t-K} via the AR effects ϕ_1, ϕ_2, …, ϕ_K. These parameters thus indicate how temporally dependent a process state is on K previous states. If we take social anxiety as an example, we would model the current state of an individuals' social anxiety as a function of K previous states of social anxiety. K thus reflects the number of lags, also called the order of the AR process, and the associated model is denoted an AR(K) model. In the case of a single lag, we get an AR(1) model, which can be written as:

$$y_t = \delta + \phi y_{t-1} + \varepsilon_t \quad (2)$$

Next to the temporal dependencies, an AR process has some additional defining properties: It fluctuates around its mean value, which depends on the intercept term δ and the AR parameter(s). The fluctuations originate from the residual or error term ε_t. The errors are assumed to be uncorrelated over time, independent of previous process states, and are usually assumed to be normally distributed, with a mean of 0, and variance σ_ε^2. The amount of overall fluctuations characterizing the process is a result of this residual variation, and the AR effects.

In the time series literature, the residual term is called innovation, reflecting the notion that ε_t provides a random perturbation or shock to the system at every time t. In the psychological literature, innovations are often assumed to reflect a collection of internal or external events influencing the system at a given time. In practice, these could be all sorts of situational factors such as various social interactions or stressful events, but also thoughts and appraisals [68]. Perturbations due to the innovations are then carried forward in the process of interest, to the extent the AR relationships are different from zero. Put simply, the innovations push the process away from its mean, and the size of an AR parameter indicates how quickly the process returns to its mean. The larger the effect size, the longer it will take for the effect of the innovation to disappear, and the slower the process will revert to its mean. Panel D of Fig. 1 displays such a response to a single, controlled innovation, also called impulse (hence, *impulse response*). Comparing the blue trajectories in the left column to the orange ones in the right column signaling the impulse responses of two distinct processes, we see a clear difference. Process y_1's mean reversion (in blue) is much slower than that of process y_2 (in orange). Of course, we will never observe such a clean and complete mean reversion in real data, since innovations are not isolated and affect the process continuously and randomly. It is a theoretical pattern that reveals differences in AR dynamics and is useful for understanding the behavior of VAR processes. The property of having a lasting effect distinguishes innovations from

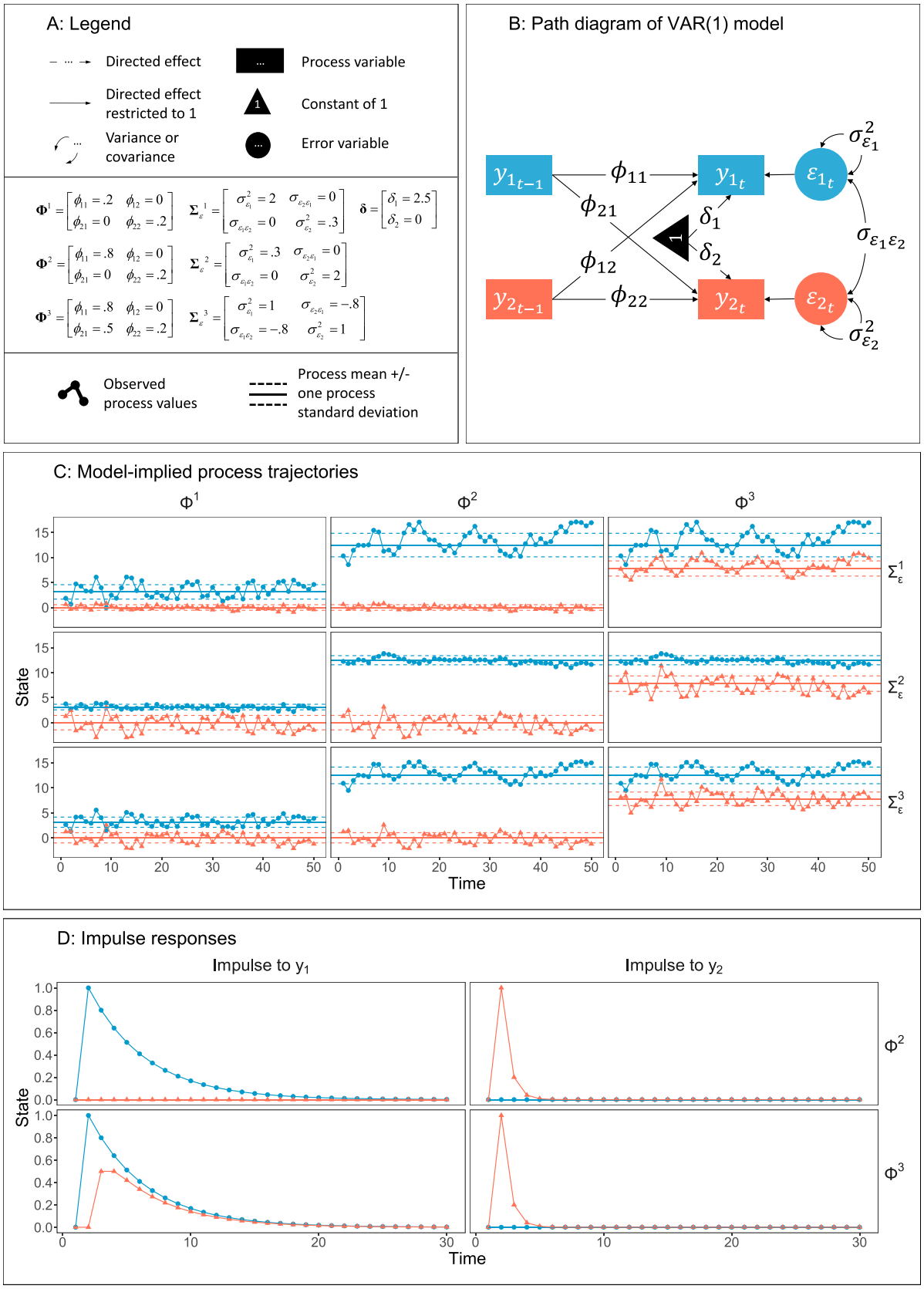

Fig. 1. Setup and behavior of a bivariate VAR(1) model. Panel A holds the legend to read the figure, Panel B displays a general path diagram for a VAR(1) model, Panel C displays simulated time series data generated from different model variants, and Panel D shows the model-implied impulse response functions. Here, the hypothetical scenario of a single innovation of value 1 affecting both processes at time $t = 2$ can be observed, driving the processes away from their respective means.

conventional error terms such as measurement error, the impact of which is assumed to be restricted to individual time points [146].

At this point, we briefly mention AR moving average (ARMA; for a more comprehensive overview of these classes of models, see [74,87,137,156]) as an extension of AR models. ARMA models regress a process variable not only on lagged versions of itself, but also on lagged versions of the innovation term. We will come back to them in later parts of the paper. An in-depth overview concerning the application of ARMA models to psychological data is given by Velicer and Molenaar [166].

Extending the univariate AR model to the multivariate case is possible by using VAR. Rather than single variables, VAR models feature a *set* of variables that are regressed on themselves and each other over time, yielding AR and crossregressive (CR) effects. VAR models thus allow us to not only assess the temporal dependence within a certain construct, such as social anxiety, but also between constructs. A first-order VAR model with two process variables, social anxiety and rumination, will include two CR effects, capturing the dependency between social anxiety at time $t-1$ and rumination at time t, and vice versa.

In terms of equations, such a bivariate VAR(1) model, linking two processes with a single lag, is:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t}_{\mathbf{y}_t} = \underbrace{\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}}_{\boldsymbol{\delta}} + \underbrace{\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}}_{\boldsymbol{\Phi}} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1}}_{\mathbf{y}_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}_t}_{\boldsymbol{\varepsilon}_t} \tag{3}$$

with

$$\underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}_t}_{\boldsymbol{\varepsilon}_t} \sim N\left(\underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\boldsymbol{\mu}_\varepsilon}, \underbrace{\begin{bmatrix} \sigma^2_{\varepsilon_1} & \sigma_{\varepsilon_1\varepsilon_2} \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2} \end{bmatrix}}_{\boldsymbol{\Sigma}_\varepsilon} \right)$$

The process variables, intercepts, and innovations are contained in 2×1 vectors $\mathbf{y}_t$ and $\mathbf{y}_{t-1}$, $\boldsymbol{\delta}$, and $\boldsymbol{\varepsilon}_t$ respectively. The AR and CR effects reside in the diagonals and off-diagonals of the 2×2 transition matrix $\boldsymbol{\Phi}$ respectively. The innovations affecting each process are assumed to be drawn from a multivariate normal distribution, with zero-mean vector and covariance matrix $\boldsymbol{\Sigma}_\varepsilon$. If the off-diagonal elements of the innovation covariance matrix, $\sigma_{\varepsilon 1 \varepsilon 2}$ and $\sigma_{\varepsilon 2 \varepsilon 1}$, take on a non-zero value, innovations are correlated, reflecting a contemporaneous relationship between the two processes.

Fig. 1 illustrates the bivariate VAR(1) model and its behavior in more detail. It contains a path diagram of the model (Panel B), and displays - for different parameter values - model-implied process trajectories (Panel C), and impulse response functions, i.e., how the VAR process would react to a single controlled innovation, also called impulse, given to either of the two processes (Panel D). Note that in the cases of no CR effects and no innovation covariances ($\boldsymbol{\Phi}^1$ or $\boldsymbol{\Phi}^2$ in combination with $\boldsymbol{\Sigma}^1_\varepsilon$ or $\boldsymbol{\Sigma}^2_\varepsilon$, see Fig. 1), the model can actually be thought of as two separate AR(1) models.

The different trajectories in Panel C demonstrate how combinations of different parameter values imply (and would thus fit) different processes. For instance, as the AR effect of process y_1 increases (left versus center column, corresponding to transition matrices $\boldsymbol{\Phi}^1$ and $\boldsymbol{\Phi}^2$), the corresponding time series in blue becomes more structured over time, which manifests in a somewhat smoother trajectory with longer departures from the process mean. Introducing the CR effect ϕ_{21} leads to process y_2 at time t to be dependent on process y_1 at time $t-1$ (center versus right column), with the resulting trajectory in orange demonstrating clear cross-variable temporal dependence in that process y_2 follows process y_1. The magnitude of the innovation variance (co-)determines how strongly the processes fluctuate around their mean levels, as can be seen directly by comparing the two process trajectories either within the top or the middle row of the left column. Here, the AR effects are the same across processes, but the innovation variances and thus the overall fluctuations are not. In addition, it can be seen that AR and CR effects not only contribute to the temporal pattern, but also to the amount of overall fluctuations. That is, in the top row, process y_1 has the same innovation variance between the left and the middle column, but different AR effects, leading to different amounts of overall fluctuation. Likewise, also in the top row, process y_2 has the same innovation variance between the middle and the right column, but different CR effects and amounts of overall fluctuations. When innovation covariances are introduced, the time series display contemporaneous relationships. This can best be seen in the absence of CR effects (bottom row, left column). In this case the correlation is rather strong and negative, resulting in the two process trajectories almost being mirror images of each other. When both CR and contemporaneous effects are included between processes (bottom row, right column), temporal and instantaneous relationships are blended together and become thus harder to see in the resulting trajectory. As we have already explained for the univariate model, the impulse response functions in Panel D illustrate differences in dynamics more clearly. The impulse leads to a sudden increase to a level of one in the process it is administered to. Following that, we see mean-reverting values, with the return-rate being inversely related to the strength of the AR effect (left versus right column). In addition, if present, CR effects can be seen to take effect (left column, bottom row). Note that it takes some time until that effect unfolds and the impulse to process process y_1 fully reaches process y_2.

An important assumption that underlies the above-described behavior of the AR(1) and VAR(1) models is that of stationarity. It implies that the modeled processes have time-invariant properties (i.e., stable means, (co-)variances, and autocovariances). An important condition for VAR(1) models to be stationary is that the temporal dependencies as captured by the AR and CR effects are not too strong, so that the perturbations due to the innovations "die out" over time, rather than accumulate (cf. [75], p. 54). For an AR(1) model, for instance, this condition is given by $|\phi| < 1$. For a VAR(1) model, this condition becomes a bit more complicated in that all eigenvalues of the transition matrix need to be smaller than one in absolute value [108]. Another, more obvious requirement for stationarity are time-invariant model parameters. We will later consider non-stationary models, in which some of these requirements are relaxed.

VAR models can be implemented and estimated in a variety of modeling frameworks, ranging from ordinary multiple regression modeling, as we have indicated above, to complex multivariate modeling frameworks such as structural equation modeling (SEM; [70,171]) and state space modeling (SSM; [36,77,84]). These different modeling frameworks also allow for simulation-based Bayesian estimation (e.g., [6,53,99]). The different modeling and estimation frameworks of course all come with their own pros and cons (e.g., the possibility to evoke common and unique latent variables in SEM and SSM, the flexibility with which distributional assumptions can be imposed during Bayesian estimation), and we will in part revisit them in the following.

3. Problems and solutions in time series analysis of intense longitudinal data

The VAR model has a number of important assumptions and implications which often do not hold for specific data sets, or do not correspond to the hypotheses of interest. Fortunately, many extensions have been proposed to accommodate researchers in handling these challenges. We will primarily focus on five specific issues and associated solutions throughout the body of the text.

Firstly, conventional VAR models yield results that are dependent on the time interval between ILD measurements, and furthermore assume that the process under investigation is measured at equidistant intervals. One must therefore ensure that time interval information is taken into account when comparing VAR modeling results from multiple studies. Moreover, within studies, the time intervals between observations should in principle be equal across observations, although this is often difficult or impossible to impose in practice. The issue here is thus

how to handle between- and within-studies differences in measurement intervals. *Secondly*, in many cases a relatively large number of variables are included in a study, and preferably jointly modeled to reveal dynamical relationships. A critical advantage of VAR models is the ability to model the joint dynamics of multiple variables, yet issues of multicollinearity and overfitting are an inevitable consequence in practical application. *Thirdly*, VAR models are appropriate when modeling stationary processes. Change, however, is often of critical interest in psychological research. In psychosomatic applications for example, effective intervention strategies imply some form of change in the trajectories of the processes of interest. We will discuss different options for studying such changes in processes. *A fourth issue* pertains to the emphasis on lagged relationships in VAR models. Researchers may also have specific hypotheses about contemporaneous relationships between variables, yet taking these into account using traditional VAR frameworks is not straightforward. *Finally*, while TSA has traditionally been employed to investigate processes within single subjects, between-subjects analyses are often of key interest to psychological researchers. Multiple extensions for between-subjects analysis have been developed, each imposing different assumptions concerning the nature of these between-subject differences.

3.1. Handling varying measurement intervals between and within studies

ILD can be collected at different time scales, and hence with different time intervals between measurements. The sampling frequency of time series data should foremost depend on the frequency at which the underlying process is assumed to change (e.g. [16,73,159,169]). For example, processes changing slowly over time, such as chronic pain, can be sampled at relatively large intervals. Others, such as mood swings and affective states, presumably vary more frequently and should be measured at shorter intervals. But additional theoretical and pragmatic considerations also influence the sampling scheme. For instance, participants of an ESM study need to sleep, necessitating night breaks. Also, intervals are often randomized, and hence vary in length during the day, to avoid expectancy effects. The decision about a sampling scheme is thus a complex one, to be made during the design phase of an ILD study. We come back to it here, because it has profound implications for data analysis, in the sense that both measurement intervals varying between studies and measurement intervals varying within studies can affect VAR modeling results, and can thus require data analytic responses.

We first consider the case of varying measurement intervals between studies. This becomes an issue if the same or similar phenomena are assessed across multiple studies, and modeling results are to be compared. As we mentioned above, VAR model parameters are specific to the time intervals over which they are estimated, and deviations in sampling schemes between studies compromise the comparability of effects. The reason for this is that VAR models are of recursive nature, predicting present from past states, making the modeled relationships inherently dependent on the time intervals over which they are estimated [124]. This is especially important in the case of AR and CR relationships, as these are often of prime interest: While AR effects normally show an exponential decay with increasing interval length, CR effects are close to zero if the interval is close to zero, reach their maximum or minimum (or their maxima and minima, in case of more then two processes) at some larger interval, and then diminish again. The latter pattern reflects that it takes time for one variable to exert its influence on another (e.g., [47]). In consequence, comparing the dynamics of one VAR process assessed at different time intervals naively (i.e., without taking time interval-information into account), can lead to misleading results. See Dormann and Griffin [51], Kuiper and Ryan [101], Oud and Delsing [124], Voelkle et al. [168] for more detailed illustrations. Note, again, that the time interval-dependency of VAR models is not some sort of anomaly, but simply an inherent feature of such model structures. Since it is (in principle) known, it can be accounted for.

Before we elaborate on proposed solutions, we turn to measurement intervals varying within a given study. Also here, the specificity of AR and CR parameters to the interval over which they are estimated holds. A core assumption of (conventional) VAR models discussed so far is that data constituting the time series are sampled at equidistant intervals, and neglecting this assumption will bias parameter estimates (e.g., [45]). Note that this is also the case if measurement intervals are varied randomly around some average interval, since the dependence of the VAR parameters on the time interval is non-linear implying that measurement interval-based effect size differences do not simply average out.

Finally, psychological processes are typically not considered to be discrete in nature, for instance they do not cease to unfold in the absence of measurement [12]. On a theoretical level, thinking about a process in terms of its dynamics and mapping theory to model can thus be less obvious in discrete time, where model interpretation is bound to a specific time interval.

3.1.1. Proposed solutions

As mentioned above, there are well justified reasons for allowing variability in ILD sampling schemes, and therefore, it is no surprise that recent developments in dynamic modeling have sought to address this issue.

3.1.1.1. Continuous-time modeling. A general approach involves dynamic models that are formulated in continuous time (e.g., [13,35,122–124,136,150,170]. Unlike conventional discrete-time models that formalize change between discrete time points, continuous-time models formalize change in the process(es) of interest as instantaneous change at any given time point. In more technical terms, continuous-time models feature differential equations, as compared to difference equations which underlie discrete-time models.

As such, continuous-time models are independent of a particular sampling interval [15,136,170]. Consequently, they readily handle unequal sampling intervals within and between studies, and corresponding discrete-time effects can be derived for any interval of interest. In addition, they offer a conceptually sound account of psychosomatic processes as continuously ongoing and smoothly changing [13,46,53,91]. This not only allows for a direct mapping between (most) theory and model, but also opens up new ways of thinking about prominent problems such as effects of interventions [54], mediation [47], or optimal sampling rates [2].

Continuous-time approaches are becoming more and more popular, resulting in increasing numbers of accessible didactic treatments (e.g., [136,168]) and software packages aimed at facilitating their use (e.g., [52,123]). It should however not be forgotten, that estimation of such models can be computationally intense to problematic, especially if more complex model variants are fit [168]. Also, such models are less straightforward than their discrete time counterparts on a technical level.

3.1.1.2. Discrete-time approaches. In addition to continuous-time modeling, there are also alternative strategies which can be employed while remaining in the discrete-time modeling framework.

Which strategy is appropriate depends properties of the measurement schedule used. For instance, if measurements are taken multiple times a day across multiple days, measurements occur in a blocked pattern, and intervals may be approximately invariant within days, but different between days (e.g., due to night breaks). A sometimes employed strategy is then to not regress the first measurement of a day on the last measurement of the previous day. This is typically done in classical regression implementations of (V)AR models by setting the lagged predictor variable(s) to missing for those data rows that pertain to the first occasion of each day [44,128]. Because missing predictor

values are not handled in classical regression modeling, these data rows are simply excluded when estimating the model.

If measurements do not occur in such regular blocks, or if one is hesitant to omit a considerable amount of data from model estimation, one can for instance artificially refine the time scale by including missing values in the data, with the result that time intervals between observations become approximately equal up to a certain level of precision. On the modeling side, this then requires the introduction of so-called phantom variables and the use of estimation techniques that can handle missing values (e.g., [3,6]). Phantom variables are latent variables without observed indicators [132], which represent the introduced missing data for a particular time point. Again, the applicability of this solution depends on the measurement pattern. If measurement intervals show large variations, the percentage of missing values in a data set, and consequently the number of phantom variables, can rise considerably, rendering such an approach infeasible [45,125].

In such a case of a disproportionate amount of missing data, or if phantom variables and missing values cannot be handled by the preferred estimation framework, then the imputation of missing values, for instance through multiple imputation techniques [161], can be an option. Software packages such as Amelia II [81] are specifically tailored to the imputation of missing data for TSA (see also [89,105]). Single imputation techniques exist as well, but are generally not recommended because they suffer from various shortcomings (for overviews, see [121,138]). For further reading on the topic of missing data in ILD, we refer to Black et al. [11], Ji et al. [89], and Velicer and Colby [165].

3.2. Handling multiple possibly correlated variables

A single process can be indicated by a single, or multiple observed variables, and in addition multiple processes may mutually interact over time to form an intricate dynamic relationship. As mentioned previously, the set of variables included in analysis depends primarily on theory and hypotheses about the processes under study, but the interpretability of model solutions may become cumbersome as the number of variables included increases.

Specifically, including many variables may lead to three problems. Firstly, interpreting VAR coefficients becomes problematic when large numbers of variables are employed, because the AR and CR effects (and regression weights in general) pertain to the unique direct effects of specific variables on top of all the other variables, and do not account for shared effects [27,41]. Secondly, if variables are highly correlated over time, one may run into multicollinearity issues, which lead to unstable VAR estimates (i.e., the bouncing beta problem in regression analysis, [41,93]). Thirdly, VAR models include a considerable number of parameters, and this number exponentially increases with each additional variable included in the model (i.e., in case of J variables, J^2 lagged effects have to be estimated). Such large numbers of parameters render VAR estimation vulnerable to overfitting, leading to poor generalisability of model results [7,111], as recently demonstrated by means of cross-validation (Bulteel, Tuerlinckx, et al., 2018). Importantly, insufficient numbers of observations can lead cross-validation techniques to prefer an AR model even when the process is actually generated from a VAR model, illustrating how prone VAR models are to overfitting, and emphasizing the importance of measuring the processes at a sufficiently large number of time points for highly parameterized models [25].

3.2.1. Proposed solutions

All three problems can be more or less solved by restricting the VAR model, so that less parameters have to be freely estimated and a sparser model is obtained. To achieve this, one can either collapse the variables into a smaller number of principal components or common factors and estimate the VAR model on the component or factor level, which vastly reduces the dimensionality of the VAR problem, or use techniques that limit complexity by forcing some estimates to zero. The former strategy uses dimension reduction methods, the latter uses regularization methods.

3.2.1.1. Dimension reduction techniques. To collapse the variables into components or factors one can either use a confirmatory approach or an exploratory one. So far, the confirmatory setting has received the most attention. In this setting, it is decided beforehand, usually based on theory, which variables are indicators of the same latent construct and therefore can be taken together in a factor. To this end, so-called dynamic factor analysis (DFA, [131]) and related approaches have been put forward. These translate the typically cross-sectional (confirmatory) factor analysis framework to the case of time series data and combine it with VAR modeling. Associated software has become available (e.g., [6,52,84]). Moreover, some of these model combinations come with fit indices that allow to assess whether the hypothesized structure in the variables seems to hold up.

In the exploratory setting, where the variables are collapsed in a data-driven way, approaches have been developed that build on principal component analysis [90]. Here, PC-VAR is a recent and intuitive proposition, that consists of two steps: First, perform principal component analysis on the variables and rotate them to simple structure, and in a second step, perform VAR modeling on the rotated component scores [28]. This technique has been demonstrated to outperform lasso-based VAR, a typical regularization method, with regard to predictive accuracy [28]. In addition, several exploratory dynamic factor analysis approaches [24,115] have been proposed.

3.2.1.2. Regularization methods. Regularization methods solve the above mentioned problems by penalizing properties of the estimated VAR parameters. A tuning or regularization parameter indicates how strongly this penalty is imposed when fitting the model. The most popular penalty in VAR analysis is the lasso penalty [82] which equals the sum of the absolute values of the regression weights [56,177]. The higher the value of the tuning parameter, the more estimates will be forced to zero, yielding a sparser model. Whereas this approach helps to minimize overfitting, it does not fully solve the interpretation and collinearity issues. Specifically, if two very strongly correlated variables are included, the lasso penalty will force one of the associated effects to zero. Yet, as demonstrated by Bulteel, Tuerlinckx, et al. [28] and discussed in general by Bondell and Reich [18], Tibshirani [154], Zou and Hastie [182], which of both parameters is set to zero is largely arbitrary.

3.3. Modeling changing processes

A core assumption of conventional VAR models, and other models using AR structures to quantify temporal effects (e.g. DFA), is that the processes under investigation are stationary. That means that their means, (co-)variances and auto-covariances do not change over time, thus, the process displays stable characteristics. However, change - also in how phenomena evolve - is often of key interest in psychopathology and psychology in general (e.g., [112,142,162]). How do individuals, for example, develop a depressive episode, how do they achieve remission, how do treatments affect the dynamics of the system? To answer these questions, statistical approaches are needed that make it possible to assess time-dependent processes, whose characteristics change. In such cases, many of the models discussed so far can be modified to capture systemic changes directly, and inference can be made about these changing processes.

3.3.1. Proposed solutions

The models discussed in this section thus deal with non-stationary time series, where change rather than stability of process characteristics is of interest. We present techniques as falling into three categories: Models that capture non-stationarities via time-varying parameters,

integrated models, and data-driven approaches. This last category can be of interest for more explorative research questions, or for the pre-processing of time series data if (certain forms of) non-stationarity are not of substantive interest but should be corrected for. In addition, we offer some distinctions between models in how they handle non-stationarities, or assume these non-stationarities to arise. We chose for these distinctions in order to serve as a practical guide for researchers trying to select from the diverse range of TSA models for non-stationary data, but do not assume that this taxonomy is substantive, nor exhaustive. Specifically, we will remark on, firstly, the degree to which models treat the sources of non-stationarity as unknown versus known, or in other words, the degree to which they incorporate uncertainty about such sources, and, secondly, whether changes are incorporated as smooth and continuous or as discrete and abrupt (cf. [3]).

Typically non-stationary models will include more parameters then their stationary counterparts to capture the more complex temporal patterns in the data. A noteworthy consequence is that this increased model complexity usually implies that more data is needed for estimation relative to stationary processes.

3.3.1.1. VAR models with time-varying parameters. In psychological applications where changing processes are of interest, VAR models with time-varying parameters are widely used. A first popular modeling option is given by so-called regime-switching models (e.g., [71,76,95,123]). As the name suggests, such models assume that the process of interest switches abruptly back and forth between a few regimes, during which it is governed by distinct parameters and hence displays distinct dynamics. Regime-switching may occur if people transition into and out of depressive episodes, or if they switch between different cognitive strategies when solving a problem [72]. The switching itself is usually governed by a Markov chain, a stochastic AR-type process that captures temporal dependencies among discrete-valued variables, in this case among the regimes. Since the parameter changes are produced by a stochastic process, whose parameters - next to the regime-specific parameters - are unknown and estimated from the data, regime-switching models can handle quite some uncertainty about non-stationarity. As a concrete example, Schneider et al. [141] investigated the temporal course of chronic pain in patients with rheumatic diseases, assuming that a person-specific regime-dependent intercept term could capture periods of low versus high pain ratings. Besides providing a more fine-grained assessment of the temporal course of chronic pain, dynamic measures such as the rapidity of switching between regimes and the difference in the magnitudes of the intercepts between the two regimes were associated with clinically relevant outcomes, beyond what could be explained by average pain levels.

Alternative modeling approaches, also allowing parameter changes to arise from unknown stochastic processes, assume continuous-valued AR-type parameter processes (e.g., [37,116]). As such, they can thus capture gradually - rather than suddenly - changing processes, which might be appropriate to formalize stress-dependent affect regulation [37] or long-term adjustments to major environmental changes [14]. In addition, there are also approaches that incorporate smooth parameter changes drawing upon techniques from nonparameteric statistics [22,78]. In some cases, it may be difficult to decide a priori whether one expects changes in parameters to occur in a gradual or abrupt fashion, or one may expect both types of changes to occur in the data. Albers and Bringmann [4] recently proposed a method capable of capturing both smooth and abrupt changes by combining time-varying AR modeling with change point detection.

Quite a different approach to time-varying parameters is taken by fixed moderated VAR and related models [3,23], which imply that we are more certain about parameter changes. Specifically, these models require that we know or hypothesize that specific time-varying covariates are associated with parameter changes in a VAR model. The parameter changes, and in particular their timing, are then determined via the observed changes in the covariates, while the amount of change is freely estimated. This applies to situations where an external event is likely to trigger a change in dynamics, such as major life events during and ESM protocol, the administration of medicine or treatments, or the experience of certain stimuli during an experimental procedure. Different shapes of change (e.g., gradual, abrupt, trending) can then flexibly be incorporated by using covariates of according format. Fixed moderated models are obviously more restrictive than the initially presented ones, but, if appropriate, they require less data for accurate model estimation. Note that it is of course also possible to include observed covariates into, for example, regime-switching models in order to aid the identification of parameter changes, but unlike in fixed moderated VAR, covariates are not needed for model estimation.

Finally, threshold AR models might be seen as bridging the gap between regime-switching models and VAR models with fixed (discrete) moderators, as both could be conceptualized as instances of this model family (see [155]).

3.3.1.2. Integrated models. We briefly introduced ARMA models in section 2, and discuss them in more detail here. When non-stationarities in the form of mean trends or cyclically recurring patterns are apparent in the data, an additional model parameter, *d*, is added to an ARMA model describing the order of differencing (difference scores between observations are used in the model, rather than the original scores) needed to result in a stationary series (e.g., [74]). The resulting AR *integrated* moving average (ARIMA) model thus, like an ARMA model, indexes short range AR and/or MA dependencies through the strength of these respective parameters.

A more general variant is the AR fractionally integrated moving average (ARFIMA) model, which is identical to the ARIMA model save for the fact that the differencing parameter is allowed to take on non-integer values. The result is a very slow decrease in serial dependence over time, which is why the magnitude of this parameter is often interpreted to reflect long memory, or long term dependence (LTD, see [156,174,175]).

In experimental settings, ARFIMA models have been employed to quantify long term serial dependencies arising in cognitive tasks [174]. LTD in physiological time series such as gait variability [48,83], heartbeat rhythm [126], and respiration [127] have been considered indicative of healthy functioning, with deviations of these dynamics considered to be indicative of pathology. Different types of ARFIMA specifications correspond to different types of serial dependencies in the data and can thus be used to test hypotheses, with information criteria measures typically being employed for model selection (e.g., [156]). ARFIMA can be extended to multivariate cases, termed fractionally integrated vector autoregression (FIVAR) [39,50], although to the best of our knowledge this has not been applied to psychological data thusfar.

Despite the promising and innovative contribution that the study of LTD in human behavior could have, there is much discussion on the extent to which it can be distinguished from multiple short-memory processes [174], and why specifically these patterns have indicative properties is a topic of debate [58,175]. In addition, ARFIMA models generally require a large number of time points for adequate estimation (see [40,152,156]).

3.3.1.3. Data-driven approaches. As indicated above, researchers might find themselves in a situation where they want to explore non-stationarities in the data without having to specify and fit a full time series model. Also, if non-stationarities are present but not of primary relevance to the research question, it might be more convenient or feasible to correct for them a priori to the actual analysis. In such cases, data-driven approaches that pinpoint the presence of non-stationarities can be useful.

One example is a recently developed method to screen time series for abrupt changes by performing non-parametric multivariate kernel

change point detection (KCP-RS). Building on work by Arlot et al. [5], this method was presented and validated by Cabrieto and colleagues [31,33,34] and comes with an associated R-package [32]. The method can be used to signal and locate abrupt, long-lasting changes in means, variances, (auto-)correlations, or any other summary statistic of interest, although some changes (such as changes in the mean) appear more easily detectable than others (such as changes related to (auto-) covariances). As such, KCP-RS can be used for exploratory purposes (e.g., to find early warning signs of a relapse into depressions; [30]), or to assess deviations from stationarity assumptions. For example, if evidence is found for a mean change, this implies that the time series can be split into multiple phases that are characterized by different means. Researchers can use these findings to detrend the data, run separate analyses for the different phases, or to inform a fixed moderated analysis.

3.4. The issue of concurrent relationships

As mentioned above, VAR techniques estimate unique direct effects of the lagged variables. They do however not specify explicitly how the different variables under study impact one another concurrently. Such information would be highly informative from a clinical point of view. To illustrate the issue, imagine running a classical VAR model on a bivariate time series encompassing social anxiety and rumination. If no CR relationships are found, one may be tempted to conclude that any intervention targeting one of the variables would leave the other variable unaffected. Indeed, although these variables may display some innovation covariance, this interrelation is usually understood as a common response to external perturbations and not as an "inherent" direct relationship between the variables. The question here is whether or how the latter type of interpretation can be made, since this is often desirable in psychology. In this section, we aim to briefly summarize what has already been said about this admittedly difficult and technical topic, which however has major interpretational ramifications.

At this point, it is useful to mention that VAR modeling was initially developed in econometrics. In this field, researchers are usually not interested in interpreting the VAR regression weights, but rather aim to forecast future observations by means of obtained VAR solutions. However, when interpretation is of central interest, a distinction has been made between two types of VAR models. These are statistically equivalent and would thus yield equal forecasts, but differ in how they handle contemporaneous relations. The first type is the VAR version that we have introduced in section 2.2 (and worked with throughout); this version is also called the reduced form [66,113]. The second type is called structural VAR (SVAR). SVAR includes contemporaneous unique direct effects among the variables. As a consequence, the innovations are not allowed to covary anymore. For a mean-centered time series the SVAR model of order 1 is formulated as follows (SVAR, [149]):

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_t + \mathbf{B}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (4)$$

Where **B** again holds estimates of AR and CR effects and the covariance matrix of the innovations is now diagonal. The additional regression weight matrix **A** captures contemporaneous relationships. Importantly, for each pair of variables only one of these concurrent regression weights is allowed differ from zero, which indicates in principle which of both variables drives the other.

Including contemporaneous effects this way may sound appealing. However, it comes with the drawback, in that SVAR solutions are not unique from a statistical point of view (for a non-technical elaboration and an in depth discussion of SVAR methodology, see [66,94], respectively). Any single (reduced-form) VAR model corresponds to multiple possible SVAR models, and each of the latter may imply that very different dynamics underlie the data. Indeed, these possible SVAR models will often not only differ in the estimates of contemporaneous relationships, but also in the estimates of lagged effects, rendering the interpretation of VAR regression weights problematic [113]. To obtain a unique solution, theoretical restrictions on the possible contemporaneous relationships are needed, indicating which of the corresponding regression weights can be set to zero.

3.4.1. Proposed solutions

The statistical equivalence of VAR and multiple SVAR solutions thus poses a significant problem for straightforward interpretation of VAR coefficients as reflecting dynamics of a psychologically meaningful system. To handle this issue, techniques have been proposed recently, which either aim to attenuate the problem, or to at least expose the possible statistically equivalent solutions, allowing a researcher to make informed decisions on grounds of theory-based hypotheses about the regression weights. One such technique is the GIMME framework [64]. It estimates SVAR models directly, rather than indirectly via the reduced form VAR model (see section 3.5 for more information on how GIMME also allows to model multiple time series) and allows to output all statistically equivalent solutions [10]. A second proposition is hybrid VAR, which, as indicated by its name, tries to find a compromise between VAR and SVAR by allowing some of the innovation covariances to differ from zero while simultaneously including the contemporaneous relationships [113].

3.5. Integrating multiple individuals

In their original formulations, the methods discussed here are applicable to data from single individual cases. Often, however, data from multiple individuals are gathered, and the aim is to learn not only about intra-individual processes and dynamics, but also about how they behave over individuals. Such situations render some sort of integration of data and/or modeling results across cases imperative. Which integration strategy one selects depends on the kind and degree of unifying assumptions one deems appropriate to make across individuals (cf. [69]). This in turn depends on substantive as well as statistical considerations.

From a substantive perspective, one might expect intra-individual processes to behave highly individual-specific or highly uniform across individuals. For instance, the relationship between the experience of emotional and physical pain seems to be highly heterogeneous across individual chronic pain patients [61]. On the other hand, some low-level psychological processes such as those related to visual perception may be rather homogeneous across individuals (i.e., the finding that visual orientation judgements are biased towards previously seen orientations, Fischer and Whitney [59]). Of course, in between these two extreme positions lies a whole continuum (e.g., [1,167], and it may well be that the to be described phenomena are situated somewhere on this continuum.

From a statistical perspective, a main question is whether modeling solutions gain or lose accuracy if one integrates across individuals. Here, arguments revolve around trading different sources of inaccuracy (bias versus variance), which can be of different importance depending on the goal of an analysis (for instance, explanation or prediction, see [179]). Whereas unifying assumptions across individuals might be overly restrictive and hence bias individual modeling solutions and inferences based on them, they might also prevent imprecise modeling solutions, under-powered tests and imprecise predictions if the amount of data per individual is limited [28,130]. The challenge thus lies in weighing between the group and the individual.

3.5.1. Proposed solutions

3.5.1.1. Modeling individuals (in)dependently. There exist multiple ways in which data and/or modeling solutions form multiple individuals can be combined. Matching the endpoints of the above construed continuum are the extreme positions (cf. [69]) of either modeling the data of each person independently of others, and possibly subsequently drawing comparisons, or modeling the pooled data in a way that ignores potential inter-individual differences completely, for instance

by treating different individuals as replicates of each other and fitting one VAR model to their data.

Whereas the latter strategy is acknowledged as inappropriate in light of plausible and actual differences between individuals in intra-individual processes and dynamics (e.g., [60,68,112,114,139]), the first strategy allows for a maximally unconstrained analysis of such differences. It can therefore be viewed as enabling "unbiased descriptions" of inter-individual differences and commonalities in intra-individual phenomena ([106], p.300), which might, on the one hand, lead to "informed aggregations" across individuals and the articulation of "lawful relationships" ([119], p.211). On the other hand, it can also capture highly individual-specific phenomena that do not permit (more than very abstract) generalizations between individuals [130]. Clearly, however, fitting a model for each person separately requires sufficient amounts of data within individuals to produce accurate results. If this requirement is not met, or if a process is assumed to unfold in a somewhat homogeneous way in multiple individuals, modeling data from multiple individuals as to some extent pooled is an option. Here, one can distinguish approaches that incorporate inter-individual differences as continuous and quantitative and approaches that group individuals into homogeneous subgroups that differ in more categorical or qualitative ways.

3.5.1.2. Modeling individual differences as quantitative. A prototypical quantitative integration strategy, also among VAR models, is to use hierarchical or multi-level or mixed effects model extensions (e.g., [23,124]). These comprise an average intra-individual process model (via so-called fixed effects) and an inter-individual model capturing individual deviations from the average structure via continuous distributions (via so-called random effects). Hence, the same model structure applies to all individuals, with individuals differing gradually in the strengths of the model's parameters. Traditional model formulations and implementations usually rely on normal distributions to describe individual differences, but some flexibility has been gained, also with respect to which parameters can be included as random effects, especially due to the rise of simulation-based estimation in the Bayesian modeling framework (e.g., [6,53,144]).

Another approach drawing upon multiple individuals to estimate the parameters of a VAR-type model is given by so-called group iterative multiple model estimation (GIMME; [63,64]). This recently developed method also represents individual differences as quantitative, but not subject to distributional assumptions. Instead, a data-driven strategy is employed. In a first iterative group-level step, the procedure decides which dynamic effects should be estimated for everyone based on modification indices (i.e., the expected improvement in model fit) associated with each effect for each person. Effects that are significant for a majority of the individuals are selected. Next, it proceeds with an iterative individual-level step, in which estimates of the selected effects are obtained for each individual separately and one or more of the other effects can be added until the time series of that individual is optimally fitted. GIMME thus yields estimates that are allowed to differ across persons, but the order in which effects are added to the person-specific models is based on group-level information. In addition, it features not the conventional VAR model, but the structural model variant, including contemporaneous relationships among the modeled processes directly. GIMME approaches have seen recent successes in psychosomatic applications. Bouwmans et al. [19] assessed dynamic relationships between endogenous melatonin fluctuations, affect, and fatigue in depressed individuals, using GIMME to accommodate large heterogeneity of these relationships across individuals. Despite the variability over individuals, some theoretically intriguing generalizations could be made, such as the finding that participants who did not display significant relationships with melatonin had more severe depression indices (for another example of GIMME application, see [92]).

3.5.1.3. Modeling individual differences as qualitative. Instead of continuous, quantitative differences, cluster-based approaches assume discrete differences between individuals, which are often interpretable as qualitative or categorical. Individuals are thereby grouped according to their full profile of AR and CR effects, such that cases that belong to the same cluster have similar intra-personal dynamics, whereas cases that are assigned to different clusters differ in at least some dynamic aspects. To our knowledge, four types of clustering approaches have been proposed so far.

Firstly, one can use a two-step approach, where a (V)AR model is initially fitted to the data of each individual separately. Subsequently, one computes dissimilarities (e.g., using the Euclidean distance) between the effect profiles for each pair of individuals and feeds these dissimilarities to a hierarchical clustering technique [180].

Secondly, Bulteel et al. [26] proposed an extension of the well-known K-means approach, which they called clusterwise VAR, as the method clusters on the VAR estimates. This method simultaneously partitions the individuals into clusters and fits a VAR model within each cluster.

Thirdly, Ernst et al. [57] built on mixture modeling and developed a probabilistic version of clusterwise VAR. One benefit of this method is that it yields an estimate of how probable the assignment of an individual to a cluster is, rather than this assignment being either true or false.

Finally, a subgrouping version of GIMME, called S-GIMME, has been presented that allows to subgroup individuals as well [62,104]. To this end, a subgroup-level step is added in between the group-level and individual-level steps as described in the previous section, implying that specific effects are added for subgroups of persons. An important distinction between S-GIMME and the previous approaches, is that within each obtained subgroup or cluster, the effects are ultimately still estimated per individual, implying that estimates can still differ within each subgroup.

4. Discussion

The benefits of employing intensive within-subjects paradigms are clear, and the complex relationships between mind and body appears well suited for analysis from a dynamical perspective. It is also clear, however, that the methodological challenges associated with employing these promising techniques are far from trivial, and have direct implications for the accuracy, interpretability, replicability, and generalizability of findings.

We would like to emphasize that most if not all scientific results ultimately depend on a vast range of researcher choices, from hypothesis to design to analysis, and the study of ILD is no different in this respect. Throughout this manuscript, the emphasis on VAR-type models is driven primarily by the wide applicability and prevalence of this modeling family in applied research. There are however many ways in which time-dependent processes can be modeled and analyzed. For instance, where we focus primarily on VAR-type models, typically analyzed in the time domain, Hamaker et al. [69] also discuss analysis in the frequency domain (e.g. [75]), which is often applicable to physiological signals (e.g. [129]).

However even within VAR frameworks, the wealth of decisions researchers are faced with can lead to highly heterogeneous results, as exemplified by a recent preprint where 12 leading research teams were asked to analyze a single data set [9]. Although most teams used VAR-type models, large variability was found between groups in all steps of the process. To give a few examples, teams differed in variable selection steps such as clustering, how they dealt with missing data, how they selected model order, and whether or not they included contemporaneous relationships. Ultimately, the researchers conclusions were as varied as the steps they took in the data-analytic process. However, the researchers also had very little information about the data they were analyzing, e.g. a lack of clinical context and relatively ambiguous research question, highlighting the need for both applied

researchers and analysts to be engaged throughout the research process. We again emphasize that data-analytic decisions must always be made in interaction with design and research question, and transparency throughout is key. Preregistration has been suggested as a general way in which all researcher decisions can be transparently reported, and guidelines for preregistration of ILD designs are being developed (a recent article discusses preregistration for ESM designs: [96]).

In the following, we briefly recapitulate the most important data-analytic issues discussed throughout the manuscript, and highlight three additional concerns we have not explicitly mentioned in the other sections. Specifically, we discuss the importance of the lag, or order of the model employed, data requirements such as sample size and time series length, and the role of measurement error in the discussed frameworks. We close with a note on causal inference in the dynamic paradigm.

4.1. Varying measurement intervals

Most conventional VAR models operate in discrete time and thus assume that intervals between observations are equal. Also, the parameter values recovered from such models are only specific to the time interval over which they are estimated. Within a study, parameter estimates can be biased when differences in time intervals are neglected. In addition, if time interval differences between studies are ignored, cross-study comparisons can confound actual differences between effects with time-interval induced ones. Conflicting findings may therefore pertain to differences regarding methodology rather than substance. The problem can be mitigated to some extent by employing various strategies, such as refining the time scale in combination with estimation techniques that accommodate latent variables and missing data, or multiple imputation of missing data. Alternatively, continuous-time methods offer an appealing solution for the issue of varying measurement intervals, since the modeled effects are not specific to a particular interval, but such approaches are paired with their own difficulties on both conceptual and statistical levels. However, the discussed procedures work optimally under circumstances where data is missing (completely) at random. In the worst case when data are not missing at random, a situation to be expected in psychological research, some process states are more likely to be missing than others. For instance, when asked to report one's momentary affect, one might be less inclined to respond on those occasions where one feels especially negative. Different types of missing data and their implications are discussed by Rubin [135] (see also [121,138]).

4.2. Including multiple possibly covarying variables

We have highlighted that VAR modeling is essentially a linear regression technique, and thus the problems associated with multivariate linear regression are applicable to the VAR framework too. Multicollinearity issues may be apparent in many practical applications, and can result in unstable VAR estimates. In addition, the fact that VAR models are intrinsically highly parameterized can quickly lead to issues concerning overfitting. Techniques aimed at limiting this complexity such as those employing dimension reduction or regularization could provide relatively straightforward solutions for these issues, and overfitting of models due to excessive parameterization can be assessed by employing cross-validation techniques.

4.3. The issue of concurrent effects

The VAR methodology has seen extensive use and development in econometrics, where economic predictions or forecasts rather than the structure of the economic system are typically of interest. This distinction may be more problematic for psychological research, where the underlying psychological dynamics and effects of variables on each other are often of key interest. VAR models do not explicitly model contemporaneous relationships between variables. To handle this issue, the structural VAR model has been proposed, but using this model will typically not result in a unique solution, in that any single VAR model is statistically equivalent to multiple structural VAR models. This conundrum can be solved by invoking theory to select the most plausible solution. However, given the fact that the dynamic paradigm in psychosomatic research is in its infancy, theoretical advances are also necessary in this regard.

4.4. Changing processes

TSA approaches should in principle directly translate the assumptions of the researcher. Assumptions on process change can be translated by employing models capable of capturing which features of the dynamical system change (e.g. changes in mean or variance over time), whether these changes occur abruptly, or as smooth transitions, and whether the sources of these changes are known or unknown. If little is known about the time-course of a process, tools have been developed to screen time series for temporal changes, and tests of stationarity can be applied. For psychosomatic applications, change may be of key interest in understanding how pathological mind-body relationships form and perpetuate. Theory should aim to make hypotheses about these temporal profiles concrete, allowing one to determine which methodological approaches are most appropriate.

4.5. Integrating multiple individuals

Regarding the integration of multiple individuals, significant recent methodological advances have been made, in that multiple approaches have been developed with different perspectives on the nature of inter-individual differences between individual-level processes. These perspectives can be situated on a continuum ranging from those assuming that each individual constitutes a unique system, to those assuming that all individual systems are equal with regard to the processes under investigation. The type of method(s) employed should ideally depend on the assumptions researchers have about the processes under investigation. The research already conducted in the psychosomatic domain appears to imply that heterogeneity may be a common issue. If little is known about how a specific psychosomatic process behaves across individuals, models that make less strong unifying assumptions at the population level should be preferred, because such models are less likely to impose group-level structures on individual-level solutions. At the same time, integrating over individuals can be beneficial for the accuracy of model solutions [25,107]. Aside from these statistical considerations, interpretability is also a concern. Interpreting and communicating individual differences is less straightforward when they are based on many independent individual-level model solutions than if a model summarizes them in terms of a normal distribution with a certain mean and variance, for instance. Choices regarding the integration of multiple individuals will thus always be a balancing act between these different aspects.

4.6. The importance of the lag or model order employed

Typical psychological applications of VAR models have focused on first-order relationships, where $\mathbf{y}_t$ is predicted from $\mathbf{y}_{t-1}$. There are no general reasons for excluding higher-order TSA models. The selection of the appropriate lag structure required to model the data is thus an issue which pertains to all models discussed throughout the paper, and incorrect specification will lead to biased parameter estimates [21]. Indeed, any inference made from model results is specific to the employed lag. The question thus boils down to how to select the lag structure which best fits the data. Including additional lags increases the parameterisation of the VAR model, which means that criteria used to select the most appropriate model order must be sensitive towards model complexity. Solutions therefore typically involve comparing

information criteria, such as the Akaike or Bayesian information criterion (AIC and BIC), or a combination of criteria [67,79], which are compared between models of differing order, such that an optimal selection can be made. In small samples specifically (e.g. $T = 40$), Hacker and Hatemi-J [67] recommend usage of either the corrected AIC (AICC [85]) or BIC. As another alternative, Jacobson et al. [86] recently developed the differential time-varying effect model (DTVEM), which employs a combination of exploratory modeling and confirmatory model selection to gauge the optimal lag structure. Although specifically tailored to accommodate multi-subject ILD, the framework currently imposes the same lag structure over all individuals, and can become computationally demanding in some cases. However, the technique offers an accessible tool for exploratory analysis.

4.7. Sample size and time series length

We have throughout this paper refrained from making concrete recommendations concerning sample size requirements, since no unitary answer to the question of 'how much data is sufficient' exists. The only general statement one can make in this regard is noting that more data is always better for purposes of statistical estimation and inference (e.g., [49,151]), and that this applies equally so to the methods discussed throughout this paper. Nevertheless, some general guidelines do exist. Researchers investigating non-stationary processes, processes with high dimensionality, or processes with large lag orders, should be especially considerate to include sufficient observations within individuals, or as we have illustrated above, think of appropriate ways to integrate information over individual cases (e.g., [25,107]).

However, collecting many data per participant may come at a price. Measurement reliability and validity may depend on participant burden, and excessive sampling schemes may introduce unwanted confounds [16]. In addition, people change. Assumptions of stationarity may become less plausible over long periods of time. To find a balance between data quantity and quality, power analyses can be informative since they allow one to gauge how many observations are needed to statistically detect a hypothesized effect if it exists at the population level.

How to perform power analyses in an ILD context is relatively understudied however, due to a number of difficulties. Time series length recommendations will differ depending on the type of models one wishes to employ, on which parameters one intends to make inferences about, and the magnitude of the the corresponding population-level effect sizes. For example, under circumstances of stationarity, correct model specification (a simple AR(1) model), and testing significance of a single AR parameter for a single individual, a power larger than .80 was achieved in relatively short series ($T = 50$) only when the magnitude of the AR coefficient generating the process was large enough (approximately 0.7, see [98]). Studies investigating effects of small sizes are thus quickly under-powered if a limited number of time points is used.

This case is perhaps oddly specific for the applications substantive researchers have in mind, as more complicated data structures where multilevel models are desired, are perhaps more commonplace. In these cases, power calculations must also take into account the number of individuals in addition to the number of time points. Power analyses can then in principle allow one to gauge trade-offs between the number of participants and the number of time points per participant allowing one to take into account issues such as participant burden [20,107,173]. A general solution is given by simulation-based procedures [17], which have the flexibility to accommodate in principle any design and assumed model. In essence, one uses a hypothesized model, for instance a multilevel or individual VAR model, to generate many data sets. These many data sets are then analyzed, and power can be defined simply as the proportion of generated data sets for which the effect of interest is significant. Despite its generality, the simulation-based approach is only as accurate as the prior information it is based on. If very little is known about plausible values for some effects, it may be difficult to construct a plausible model from which to simulate representative data. In addition, data will typically be simulated with perfectly equidistant intervals and without missing data, so prudent researchers should interpret the sample size and series length recommendations resulting from such analyses with due caution. We refer readers to Bolger et al. [17] for a more in-depth tutorial, and to Lafit et al. [103] who offer a straightforward shiny app which can accommodate the needs of many typically used multilevel models in ILD.

4.8. The role of measurement error

Innovations play an important role in VAR models and extensions thereof. They affect the process over time and therefore are part of the dynamics of the system. Measurement error in contrast is only assumed to arise from imperfect measurement of a construct at time *t*, lacking the temporal informative content of innovations. In conventional VAR models, measurement error is not taken into account, implying that each construct in question is measured perfectly. Many may justly consider this to be a problematic assumption, especially in psychological applications. In fact, neglecting measurement error has been found to lead to underestimations of AR parameters in AR(1) models [146]. These authors also note that what is considered measurement error, and what is considered innovation, can vary depending on the sampling interval. If large versus small intervals are used to assess the same process, effects that did carry over to subsequent observations in the small interval case, and are thus considered innovations, could fail to carry over in the large interval case, and could thus be considered measurement error.

Disentangling process error and measurement error effects can be done in a variety of ways, but these have in common that the difference is explicitly incorporated into the model. SSM and SEM frameworks employ this distinction, include measurement errors as latent variables in the measurement model, and process innovations as latent variables in the process model. In multilevel extensions of VAR models, a well-known problem is that variability between individuals must be disentangled from variability within individuals [43], but Schuurman and Hamaker [145] emphasize that variability within individuals will also consist of variability due to the process and variability due to measurement. These authors propose a multilevel model that allows one to specify whether measurement error should be common to all individuals, or whether measurement error is also person-specific.

4.9. Causal inference

A final remark we wish to make concerns causal inference. Multiple authors have made a case that adopting a dynamic multivariate approach to psychological and psychosomatic phenomena may aid the identification of causal effects (e.g., [12,54,143,168]). The presumed assets are threefold. Firstly, intra-individual change is formalized, and especially so in terms of interacting processes, instead of as a function of time. Secondly, the introduction of time and temporal order allows one to disentangle (statistical) antecedents and consequences, and to account for confounding variables within time points. Thirdly, the dynamic paradigm provides the opportunity to control for associations between variables across individuals as important confounding factors.

However, causal inference - which itself is not a unitary concept and might be approached from different theoretical standpoints with different techniques - remains a complicated problem even in the context of a dynamic research paradigm. For instance, since observational data often comprise only a limited set of variables, potentially measured with errors, information for causal inference is likely insufficient. Also, that the targeted phenomena are often poorly understood does not exactly help in formalizing causal mechanisms. Proposals to further improve causal inference under the dynamic paradigm mainly draw upon the idea of bringing the targeted phenomena partly under

experimental and/or statistical control (e.g. [97,109,120,140]).

5. Conclusion

In summary, if we wish to safeguard the promising possibilities associated with the dynamic paradigm, methodological issues must be taken seriously. Concrete steps can be taken to ensure the accuracy, interpretability, replicability, and generalizability of findings. While it may not always be possible to meet each modeling assumption from a pragmatic point of view, studies should nevertheless be transparent in reporting details associated with the sampling scheme, included variables, and reasons for employing certain modeling techniques. In this way one may more easily disentangle methodological issues from conflicting theoretical findings.

Declaration of competing interest

The authors declare the absence of any financial, intellectual, or other conflicts of interest which may have biased any aspect of this manuscript.

Acknowledgements

This work was supported by a research fellowship from the German Research Foundation (DFG) awarded to Janne Adolf, and by research grants from the Fund for Scientific Research-Flanders (FWO, Project No. G.074319N) and from the Research Council of KU Leuven (C14/19/054) awarded to Eva Ceulemans.

References

[1] J.K. Adolf, E.I. Fried, Ergodicity is sufficient but not necessary for group-to-individual generalizability, Proc. Natl. Acad. Sci. 116 (14) (2019) 6540–6541.
[2] J.K. Adolf, T. Loossens, F. Tuerlinckx, E. Ceulemans, Optimal Fixed Sampling Rates for Reliable Continuous-Time First-Order Autoregressive Modeling, (2019), https://doi.org/10.31234/osf.io/5cbfw.
[3] J.K. Adolf, M.C. Voelkle, A. Brose, F. Schmiedek, Capturing context-related change in emotional dynamics via fixed moderated time series analysis, Multivar. Behav. Res. 52 (4) (2017) 499–531.
[4] C. Albers, L.F. Bringmann, Inspecting Gradual and Abrupt Changes in Emotion Dynamics with the Time-Varying Change Point Autoregressive Model, (2020).
[5] S. Arlot, A. Celisse, Z. Harchaoui, Kernel Change-Point Detection, (2012) arXiv preprint arXiv:1202.3878, 6.
[6] T. Asparouhov, E.L. Hamaker, B. Muthén, Dynamic structural equation models, Struct. Equ. Model. Multidiscip. J. 25 (3) (2018) 359–388.
[7] M.A. Babyak, What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models, Psychosom. Med. 66 (3) (2004) 411–421.
[8] R. Balon, Mood, anxiety, and physical illness: body and mind, or mind and body? Depress. Anxiety 23 (6) (2006) 377–387.
[9] J.A. Bastiaansen, Y.K. Kunkels, F. Blaauw, S.M. Boker, E. Ceulemans, M. Chen, S.-M. Chow, P. de Jonge, A.C. Emerencia, S. Epskamp, et al., Time to Get Personal? The Impact of researchers' Choices on the Selection of Treatment Targets Using the Experience Sampling Methodology, (2019), https://doi.org/10.31234/osf.io/c8vp7.
[10] A.M. Beltz, P.C. Molenaar, Dealing with multiple solutions in structural vector autoregressive models, Multivar. Behav. Res. 51 (2–3) (2016) 357–373.
[11] A.C. Black, O. Harel, G. Matthews, Techniques for analyzing intensive longitudinal data with missing values, in: R. Mehl, T.S. Conner (Eds.), Handbook of Research Methods for Studying Daily Life, The Guilford Press, 2012, pp. 339–365.
[12] S.M. Boker, Consequences of continuity: the hunt for intrinsic properties within parameters of dynamics in psychological processes, Multivar. Behav. Res. 37 (3) (2002) 405–422.
[13] S.M. Boker, Dynamical Systems and Differential Equation Models of Change, in: H. Cooper, P.M. Camic, D.L. Long, A.T. Panter, D. Rindskopf, K.J. Sher (Eds.), APA Handbook of Research Methods in Psychology, Vol 3: Data Analysis and Research Publication, American Psychological Association, Washington, 2012, pp. 323–333 Retrieved December 6, 2018, from http://content.apa.org/books/13621-016.
[14] S.M. Boker, Adaptive equilibrium regulation: a balancing act in two timescales, J. Person-Orient. Res. 1 (1–2) (2015) 99–109.
[15] S.M. Boker, S.S. Tiberio, R.G. Moulder, Robustness of time delay embedding to sampling interval misspecification, in: K. van Montfort, J.H.L. Oud, M.C. Voelkle (Eds.), Continuous Time Modeling in the Behavioral and Related Sciences, Springer International Publishing, 2018, pp. 239–258, , https://doi.org/10.1007/978-3-319-77219-6_10.
[16] N. Bolger, A. Davis, E. Rafaeli, Diary methods: capturing life as it is lived, Annu. Rev. Psychol. 54 (1) (2003) 579–616.
[17] N. Bolger, G. Stadler, J.-P. Laurenceau, Power analysis for intensive longitudinal studies, in: M.R. Mehl, T. Conner (Eds.), Handbook of Research Methods for Studying Daily Life, The Guilford Press, 2012, pp. 285–301.
[18] H.D. Bondell, B.J. Reich, Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR, Biometrics 64 (1) (2008) 115–123.
[19] M.E. Bouwmans, A.M. Beltz, E.H. Bos, A.J. Oldehinkel, P. de Jonge, P.C. Molenaar, The person-specific interplay of melatonin, affect, and fatigue in the context of sleep and depression, Personal. Individ. Differ. 123 (2018) 163–170.
[20] A.M. Brandmaier, T. von Oertzen, P. Ghisletta, C. Hertzog, U. Lindenberger, LIFESPAN: A tool for the computer-aided design of longitudinal studies, Front. Psychol. 6 (2015) 272.
[21] P.A. Braun, S. Mittnik, Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions, J. Econ. 59 (3) (1993) 319–341.
[22] L.F. Bringmann, E.L. Hamaker, D.E. Vigo, A. Aubert, D. Borsboom, F. Tuerlinckx, Changing dynamics: Time-varying autoregressive models using generalized additive modeling, Psychol. Methods 22 (3) (2017) 409–425.
[23] L.F. Bringmann, N. Vissers, M. Wichers, N. Geschwind, P. Kuppens, F. Peeters, D. Borsboom, F. Tuerlinckx, A network approach to psychopathology: New insights into clinical longitudinal data, PLoS One 8 (4) (2013) e60188.
[24] M.W. Browne, J.R. Nesselroade, Representing psychological processes with dynamic factor models: Some promising uses and extensions of autoregressive moving average time series models, in: A. Maydeu-Olivares, J. McArdle (Eds.), *Multivariate applications book series. Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 415–452), Lawrence Erlbaum Associates Publishers, 2005.
[25] K. Bulteel, M. Mestdagh, F. Tuerlinckx, E. Ceulemans, VAR (1) based models do not always outpredict AR (1) models in typical psychological applications, Psychol. Methods 23 (4) (2018) 740–756.
[26] K. Bulteel, F. Tuerlinckx, A. Brose, E. Ceulemans, Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics, Front. Psychol. 7 (2016) 1540.
[27] K. Bulteel, F. Tuerlinckx, A. Brose, E. Ceulemans, Using raw var. regression coefficients to build networks can be misleading, Multivar. Behav. Res. 51 (2–3) (2016) 330–344.
[28] K. Bulteel, F. Tuerlinckx, A. Brose, E. Ceulemans, Improved insight into and prediction of network dynamics by combining var. and dimension reduction, Multivar. Behav. Res. 53 (6) (2018) 853–875.
[29] C. Burton, D. Weller, M. Sharpe, Are electronic diaries useful for symptoms research? a systematic review, J. Psychosom. Res. 62 (5) (2007) 553–561.
[30] J. Cabrieto, J. Adolf, F. Tuerlinckx, P. Kuppens, E. Ceulemans, An objective, comprehensive and flexible statistical framework for detecting early warning signs of mental health problems, Psychother. Psychosom. 88 (2019) 184–186.
[31] J. Cabrieto, J.K. Adolf, F. Tuerlinckx, P. Kuppens, E. Ceulemans, Detecting long-lived autodependency changes in a multivariate system via change point detection and regime switching models, Sci. Rep. 8 (1) (2018) 15,637.
[32] J. Cabrieto, K. Meers, Kcprs-package: Kcp on the running statistics, (2019).
[33] J. Cabrieto, F. Tuerlinckx, P. Kuppens, B. Hunyadi, E. Ceulemans, Testing for the presence of correlation changes in a multivariate time series: A permutation based approach, Sci. Rep. 8 (1) (2018) 769.
[34] J. Cabrieto, F. Tuerlinckx, P. Kuppens, F.H. Wilhelm, M. Liedlgruber, E. Ceulemans, Capturing correlation changes by applying kernel change point detection on the running correlations, Inf. Sci. 447 (2018) 117–139.
[35] S.-M. Chow, Practical tools and guidelines for exploring and fitting linear and nonlinear dynamical systems models, Multivar. Behav. Res. 54 (5) (2019) 690–718.
[36] S.-M. Chow, M.-H.R. Ho, E.L. Hamaker, C.V. Dolan, Equivalence and differences between structural equation modeling and state-space modeling techniques, Struct. Equ. Model. 17 (2) (2010) 303–332.
[37] S.-M. Chow, J. Zu, K. Shifren, G. Zhang, Dynamic factor analysis models with time-varying parameters, Multivar. Behav. Res. 46 (2) (2011) 303–339.
[38] C.A. Chun, The expression of posttraumatic stress symptoms in daily life: A review of experience sampling methodology and daily diary studies, J. Psychopathol. Behav. Assess. 38 (3) (2016) 406–420.
[39] C.-F. Chung, Calculating and analyzing impulse responses for the vector arfima model, Econ. Lett. 71 (1) (2001) 17–25.
[40] C.-F. Chung, R.T. Baillie, Small sample bias in conditional sum-of-squares estimators of fractionally integrated ARMA models, Empir. Econ. 18 (4) (1993) 791–806.
[41] J. Cohen, P. Cohen, S.G. West, L.S. Aiken, Applied multiple regression/correlation analysis for the behavioral sciences, Lawrence Erlbaum Associates Publishers, 2013.
[42] L.M. Collins, Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model, Annu. Rev. Psychol. 57 (2006) 505–528.
[43] P.J. Curran, D.J. Bauer, The disaggregation of within-person and between-person effects in longitudinal models of change, Annu. Rev. Psychol. 62 (2011) 583–619.
[44] S. de Haan-Rietdijk, P. Kuppens, E.L. Hamaker, What's in a day? a guide to decomposing the variance in intensive longitudinal data, Front. Psychol. 7 (2016) 891.
[45] S. de Haan-Rietdijk, M.C. Voelkle, L. Keijsers, E.L. Hamaker, Discrete-vs. continuoustime modeling of unequally spaced experience sampling method data, Front. Psychol. 8 (2017) 1849.
[46] P.R. Deboeck, J. Nicholson, C. Kouros, T.D. Little, J. Garber, Integrating developmental theory and methodology: Using derivatives to articulate change

theories, models, and inferences, Appl. Dev. Sci. 19 (4) (2015) 217–231.
[47] P.R. Deboeck, K.J. Preacher, No need to be discrete: A method for continuous time mediation analysis, Struct. Equ. Model. Multidiscip. J. 23 (1) (2016) 61–75.
[48] D. Delignières, K. Torre, Fractal dynamics of human gait: A reassessment of the 1996 data of Hausdorff et al, J. Appl. Physiol. 106 (4) (2009) 1272–1279.
[49] L. Deng, M. Yang, K.M. Marcoulides, Structural equation modeling with many variables: A systematic review of issues and developments, Front. Psychol. 9 (2018) 580.
[50] H.X. Do, R.D. Brooks, S. Treepongkaruna, Generalized impulse response analysis in a fractionally integrated vector autoregressive model, Econ. Lett. 118 (3) (2013) 462–465.
[51] C. Dormann, M.A. Griffin, Optimal time lags in panel studies, Psychol. Methods 20 (4) (2015) 489.
[52] C.C. Driver, J.H. Oud, M.C. Voelkle, Continuous time structural equation modeling with r package ctsem, J. Stat. Softw. 77 (5) (2017) 1–35.
[53] C.C. Driver, M.C. Voelkle, Hierarchical Bayesian continuous time dynamic modeling, Psychol. Methods 23 (4) (2018) 774–799.
[54] C.C. Driver, M.C. Voelkle, Understanding the time course of interventions with continuous time dynamic models, in: K. van Montfort, J.H.L. Oud, M.C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 79–109), Springer International Publishing, 2018.
[55] U.W. Ebner-Priemer, T.J. Trull, Ecological momentary assessment of mood disorders and mood dysregulation, Psychol. Assess. 21 (4) (2009) 463.
[56] S. Epskamp, E.I. Fried, A tutorial on regularized partial correlation networks, Psychol. Methods 23 (4) (2018) 617–634.
[57] A.F. Ernst, M.E. Timmerman, B.F. Jeronimus, C.J. Albers, Insight into individual differences in emotion dynamics with clustering, Assessment. (2019), https://doi.org/10.1177/1073191119873714.
[58] S. Farrell, E.-J. Wagenmakers, R. Ratcliff, 1/f noise in human cognition: Is it ubiquitous, and what does it mean? Psychon. Bull. Rev. 13 (4) (2006) 737–741.
[59] J. Fischer, D. Whitney, Serial dependence in visual perception, Nat. Neurosci. 17 (5) (2014) 738–743.
[60] A.J. Fisher, J.D. Medaglia, B.F. Jeronimus, Lack of group-to-individual generalizability is a threat to human subjects research, Proc. Natl. Acad. Sci. 115 (27) (2018) E6106–E6115.
[61] M. Frumkin, S. Haroutounian, T. Rodebaugh, Examining emotional pain among individuals with chronic physical pain: Nomothetic and idiographic approaches, (2020) Manuscript under review.
[62] K.M. Gates, S.T. Lane, E. Varangis, K. Giovanello, K. Guiskewicz, Unsupervised classification during time-series model building, Multivar. Behav. Res. 52 (2) (2017) 129–148.
[63] K.M. Gates, P.C. Molenaar, Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples, NeuroImage 63 (1) (2012) 310–319.
[64] K.M. Gates, P.C. Molenaar, S.P. Iyer, J.T. Nigg, D.A. Fair, Organizing heterogeneous samples using community detection of gimme-derived resting state functional networks, PLoS One 9 (3) (2014) e91322.
[65] A. Ghiggia, R. Torta, V. Tesio, M. Di Tella, A. Romeo, F. Colonna, G.C. Geminiani, E. Fusaro, A. Batticciotto, L. Castelli, Psychosomatic syndromes in fibromyalgia, Clin. Exp. Rheumatol. 105 (3) (2017) 106–111.
[66] Gottschalk, J. (2001). *An introduction into the svar methodology: Identification, interpretation and limitations of svar models* (tech. Rep.). Kiel working paper.
[67] S.R. Hacker, A. Hatemi-J, Optimal lag-length choice in stable and unstable var. models under situations of homoscedasticity and ARCH, J. Appl. Stat. 35 (6) (2008) 601–615.
[68] E.L. Hamaker, Why researchers should think "within-person": A paradigmatic rationale, in: M.R. Mehl, T. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61), The Guilford Press, 2012.
[69] E.L. Hamaker, E. Ceulemans, R.P.P.P. Grasman, F. Tuerlinckx, Modeling affect dynamics: State of the art and future challenges, Emot. Rev. 7 (4) (2015) 316–322.
[70] E.L. Hamaker, C.V. Dolan, P.C. Molenaar, Arma-based sem when the number of time points t exceeds the number of cases n: Raw data maximum likelihood, Struct. Equ. Model. 10 (3) (2003) 352–379.
[71] E.L. Hamaker, R.P.P.P. Grasman, Regime switching state-space models applied to psychological processes: Handling missing data and making inferences, Psychometrika 77 (2) (2012) 400–422.
[72] E.L. Hamaker, R.P.P.P. Grasman, J.H. Kamphuis, Regime-switching models to study psychological processes, in: P.C. Molenaar, K.M. Newell (Eds.), *Individual pathways of change: Statistical models for analyzing learning and development* (pp. 155–168), American Psychological Association, 2010.
[73] E.L. Hamaker, M. Wichers, No time like the present: Discovering the hidden dynamics in intensive longitudinal data, Curr. Dir. Psychol. Sci. 26 (1) (2017) 10–15.
[74] E.L. Hamaker, C.V. Dolan, Idiographic data analysis: Quantitative methods—from simple to advanced, in: J. Valsiner, P.C.M. Molenaar, M.C. Lyra, N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 191–216), 2009 Springer US. https://doi.org/10.1007/978-0-387-95922-1_9.
[75] Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). Princeton University Press.
[76] J.D. Hamilton, Regime switching models, in: S.N. Durlauf, L.E. Blume (Eds.), *Macroeconometrics and time series analysis* (pp. 202–209), Palgrave Macmillan, 2009.
[77] A.C. Harvey, Forecasting, structural time series models and the kalman filter, Cambridge University Press, 1989.
[78] Haslbeck, J., Bringmann, L. F., & Waldorp, L. J. (2017). How to estimate time-varying vector autoregressive models? a comparison of two methods. *arXiv preprint arXiv:1711.05204*.
[79] A. Hatemi-j, A new method to choose optimal lag order in stable and unstable var. models, Appl. Econ. Lett. 10 (3) (2003) 135–137.
[80] C.R. Henrie, R. Bodily, K.C. Manwaring, C.R. Graham, Exploring intensive longitudinal measures of student engagement in blended learning, Int. Rev. Res. Open Dist. Learn. 16 (3) (2015) 131–155.
[81] J. Honaker, G. King, M. Blackwell, et al., Amelia ii: A program for missing data, J. Stat. Softw. 45 (7) (2011) 1–47.
[82] N.-J. Hsu, H.-L. Hung, Y.-M. Chang, Subset selection for vector autoregressive processes using lasso, Comput. Stat. Data Anal. 52 (7) (2008) 3645–3657.
[83] N. Hunt, D. McGrath, N. Stergiou, The influence of auditory-motor coupling on fractal dynamics in human gait, Sci. Rep. 4 (2014) 5879.
[84] M.D. Hunter, State space modeling in an open source, modular, structural equation modeling environment, Struct. Equ. Model. Multidiscip. J. 25 (2) (2018) 307–324.
[85] C.M. Hurvich, C.-L. Tsai, A corrected akaike information criterion for vector autoregressive model selection, J. Time Ser. Anal. 14 (3) (1993) 271–279.
[86] N.C. Jacobson, S.-M. Chow, M.G. Newman, The differential time-varying effect model (dtvem): A tool for diagnosing and modeling time lags in intensive longitudinal data, Behav. Res. Methods 51 (1) (2019) 295–315.
[87] A.T. Jebb, L. Tay, W. Wang, Q. Huang, Time series analysis for psychological research: Examining and forecasting change, Front. Psychol. 6 (2015) 727.
[88] M. Jensen, A. Hoffman, D. Cardenas, Chronic pain in individuals with spinal cord injury: A survey and longitudinal study, Spinal Cord 43 (12) (2005) 704–712.
[89] L. Ji, S.-M. Chow, A.C. Schermerhorn, N.C. Jacobson, E.M. Cummings, Handling missing data in the modeling of intensive longitudinal data, Struct. Equ. Model. Multidiscip. J. 25 (5) (2018) 715–736.
[90] I.T. Jolliffe, N.T. Trendafilov, M. Uddin, A modified principal component technique based on the lasso, J. Comput. Graph. Stat. 12 (3) (2003) 531–547.
[91] J. Karch, *A machine learning perspective on repeated measures* (Doctoral dissertation), Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, 2016.
[92] D. Kelly, A. Weigard, A. Beltz, How are you doing? the person-specificity of daily links between neuroticism and physical health, (2020) Manuscript under review.
[93] H.A. Kiers, A.K. Smilde, A comparison of various methods for multivariate regression with highly collinear variables, JISS 16 (2) (2007) 193–228.
[94] L. Kilian, H. Lütkepohl, Structural vector autoregressive analysis, Cambridge University Press, 2017.
[95] Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and gibbs-sampling approaches with applications* (Vol. 1). The MIT press.
[96] O. Kirtley, G. Lafit, R. Achterhof, A.P. Hiekkaranta, I. Myin-Germeys, Making the black box transparent: A template and tutorial for (pre-) registration of studies using experience sampling methods (esm), Advances and Methods and Practices in Psychological Science (2019).
[97] P. Koval, M.L. Pe, K. Meers, P. Kuppens, Affect dynamics in relation to depressive symptoms: Variable, unstable or inert? Emotion 13 (6) (2013) 1132–1141.
[98] T. Krone, C.J. Albers, M.E. Timmerman, A comparative simulation study of ar (1) estimators in short time series, Qual. Quant. 51 (1) (2017) 1–21.
[99] Krone, T., Albers, C., & Timmerman, M. (2016). Bayesian dynamic modeling to assess differential treatment effects on panic attack frequencies. Stat. Model., 16(5), 343–359.
[100] A. Kucyi, K.D. Davis, The dynamic pain connectome, Trends Neurosci. 38 (2) (2015) 86–95.
[101] R.M. Kuiper, O. Ryan, Drawing conclusions from cross-lagged relationships: Reconsidering the role of the time-interval, Struct. Equ. Model. Multidiscip. J. 25 (5) (2018) 809–823.
[102] P. Kuppens, It's about time: A special section on affect dynamics, Emot. Rev. 7 (4) (2015) 297–300.
[103] Lafit, G., Adolf, J. K., Dejonckheere, E., Viechtbauer, W., Myin-Germeys, I., & Ceulemans, E. (2020). *Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial to perform power analysis* [Retrieved from osf.io/vguey]. Retrieved from osf.io/vguey.
[104] S.T. Lane, K.M. Gates, H.K. Pike, A.M. Beltz, A.G. Wright, Uncovering general, shared, and unique temporal patterns in ambulatory assessment data, Psychol. Methods 24 (1) (2019) 54–69.
[105] Y. Li, L. Ji, Z. Oravecz, T.R. Brick, M.D. Hunter, S.-M. Chow, Dynr.mi: An r program for multiple imputation in dynamic modeling, World Acad. Sci. Eng. Technol. 13 (5) (2019) 302–311.
[106] U. Lindenberger, T. von Oertzen, Variability in cognitive aging: From taxonomy to theory, in: E. Bialystok, F.I.M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 297–314), Oxford University Press, 2006.
[107] S. Liu, Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels, Br. J. Math. Stat. Psychol. 70 (3) (2017) 480–498.
[108] H. Lütkepohl, New introduction to multiple time series analysis, Springer, Berlin, Heidelberg, 2015.
[109] C. McCall, L.K. Hildebrandt, R. Hartmann, B.M. Baczkowski, T. Singer, Introducing the wunderkammer as a tool for emotion research: Unconstrained gaze and movement patterns in three emotionally evocative virtual worlds, Comput. Hum. Behav. 59 (2016) 93–107.
[110] K. McGreevy, M.M. Bottros, S.N. Raja, Preventing chronic pain following acute pain: Risk factors, preventive strategies, and their efficacy, Eur. J. Pain Suppl. 5 (2) (2011) 365–376.
[111] D.M. McNeish, Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences, Multivar. Behav. Res. 50 (5) (2015) 471–484.
[112] P.C. Molenaar, A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever, Measurement 2 (4) (2004) 201–218.

[113] P.C. Molenaar, Equivalent dynamic models, Multivar. Behav. Res. 52 (2) (2017) 242–258.
[114] P.C. Molenaar, C.G. Campbell, The new person-specific paradigm in psychology, Curr. Dir. Psychol. Sci. 18 (2) (2009) 112–117.
[115] P.C. Molenaar, J.G. De Gooijer, B. Schmitz, Dynamic factor analysis of nonstationary multivariate time series, Psychometrika 57 (3) (1992) 333–349.
[116] P.C. Molenaar, K.O. Sinclair, M.J. Rovine, N. Ram, S.E. Corneal, Analyzing developmental processes on an individual level using nonstationary time series modeling, Dev. Psychol. 45 (1) (2009) 260–271.
[117] T. Nakamura, K. Kiyono, H. Wendt, P. Abry, Y. Yamamoto, Multiscale analysis of intensive longitudinal biomedical signals and its clinical applications, Proc. IEEE 104 (2) (2016) 242–261.
[118] A.M. Nehrkorn-Bailey, M.S. Reardon, J. Hicks Patrick, Some methodological and analytical issues related to real-time data capture studies, Transl. Issues Psychol. Sci. 4 (4) (2018) 349.
[119] J.R. Nesselroade, On an emerging third discipline of scientific psychology, in: P.C. Molenaar, K.M. Newell (Eds.), *Individual pathways of change: Statistical models for analyzing learning and development* (pp. 209–218), American Psychological Association, 2010.
[120] A.B. Neubauer, V. Lerche, A. Voss, Interindividual differences in the intraindividual association of competence and well-being: Combining experimental and intensive longitudinal designs, J. Pers. 86 (4) (2018) 698–713.
[121] D.A. Newman, Missing data: Five practical guidelines, Organ. Res. Methods 17 (4) (2014) 372–411.
[122] Z. Oravecz, F. Tuerlinckx, J. Vandekerckhove, A hierarchical ornstein–uhlenbeck model for continuous repeated measurement data, Psychometrika 74 (3) (2009) 395–418.
[123] L. Ou, M.D. Hunter, S.-M. Chow, What's for dynr: A package for linear and nonlinear dynamic modeling in r, (2017) Journal of Statistical Software.
[124] J.H. Oud, M.J. Delsing, Continuous time modeling of panel data by means of sem, in: K. van Montfort, J.H. Oud, A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244), Springer, 2010.
[125] J.H. Oud, M.C. Voelkle, Do missing values exist? incomplete data handling in crossnational longitudinal studies by means of continuous time modeling, Qual. Quant. 48 (6) (2014) 3271–3288.
[126] C.-K. Peng, J. Mietus, J. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, Long-range anticorrelations and non-gaussian behavior of the heartbeat, Phys. Rev. Lett. 70 (9) (1993) 1343.
[127] C.-K. Peng, J.E. Mietus, Y. Liu, C. Lee, J.M. Hausdorff, H.E. Stanley, A.L. Goldberger, L.A. Lipsitz, Quantifying fractal dynamics of human respiration: Age and gender effects, Ann. Biomed. Eng. 30 (5) (2002) 683–692.
[128] M.L. Piccirillo, E.D. Beck, T.L. Rodebaugh, A clinician's primer for idiographic research: Considerations and recommendations, Behav. Ther. 50 (5) (2019) 938–951.
[129] M.M. Platisa, Z. Nestorovic, S. Damjanovic, V. Gal, Linear and non-linear heart rate variability measures in chronic and acute phase of anorexia nervosa, Clin. Physiol. Funct. Imaging 26 (1) (2006) 54–60.
[130] N. Ram, M. Brinberg, A.L. Pincus, D.E. Conroy, The questionable ecological validity of ecological momentary assessment: Considerations for design and analysis, Res. Hum. Dev. 14 (3) (2017) 253–270.
[131] N. Ram, A. Brose, P.C. Molenaar, et al., Dynamic factor analysis: Modeling person-specific process, The Oxford Handbook of Quantitative Methods, 2 2013, pp. 441–457.
[132] D. Rindskopf, Using phantom and imaginary latent variables to parameterize constraints in linear structural models, Psychometrika 49 (1) (1984) 37–47.
[133] C. Rometsch-Ogioun El Sount, P. Windthorst, J. Denkinger, K. Ziser, C. Nikendei, D. Kindermann, J. Ringwald, V. Renner, S. Zipfel, F. Junne, Chronic pain in refugees with posttraumatic stress disorder (ptsd): A systematic review on patients' characteristics and specific interventions, J. Psychosom. Res. 118 (2019) 83–97.
[134] J.G. Rosmalen, A.M. Wenting, A.M. Roest, P. de Jonge, E.H. Bos, Revealing causal heterogeneity using time series analysis of ambulatory assessments: Application to the association between depression and physical activity after myocardial infarction, Psychosom. Med. 74 (4) (2012) 377–386.
[135] D.B. Rubin, Inference and missing data, Biometrika 63 (3) (1976) 581–592.
[136] O. Ryan, R.M. Kuiper, E.L. Hamaker, 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* (pp. 27–54), Springer International Publishing, 2018https://doi.org/10.1007/978-3-319-77219-6_2.
[137] J.D. Scargle, Studies in astronomical time series analysis: Modeling random processes in the time domain, Astrophys. J. Suppl. Ser. 45 (1981) 1–71.
[138] J.L. Schafer, J.W. Graham, Missing data: Our view of the state of the art, Psychol. Methods 7 (2) (2002) 147–177.
[139] Schmiedek, F., Lövdén, M., von Oertzen, T., & Lindenberger, U. (2019). *Within-person structures of daily cognitive performance cannot be inferred from between-person structures of cognitive abilities* (tech. Rep.). PeerJ Preprints.
[140] F. Schmiedek, A.B. Neubauer, Experiments in the wild: Introducing the within-person encouragement design, Multivar. Behav. Res. 55 (2) (2020) 256–276.
[141] S. Schneider, D.U. Junghaenel, M. Ono, A.A. Stone, Temporal dynamics of pain: An application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases, Pain 159 (7) (2018) 1346.
[142] U. Scholz, It's time to think about time in health psychology, Appl. Psychol. Health Well-Being 11 (2) (2019) 173–186.
[143] N.K. Schuurman, E. Ferrer, M. de Boer-Sonnenschein, E.L. Hamaker, How to compare cross-lagged associations in a multilevel autoregressive model, Psychol. Methods 21 (2) (2016) 206–221.
[144] N.K. Schuurman, R.P.P.P. Grasman, E. Hamaker, A comparison of inverse-wishart prior specifications for covariance matrices in multilevel autoregressive models, Multivar. Behav. Res. 51 (2–3) (2016) 185–206.
[145] N.K. Schuurman, E.L. Hamaker, Measurement error and person-specific reliability in multilevel autoregressive modeling, Psychol. Methods 24 (1) (2019) 70–91.
[146] N.K. Schuurman, J.H. Houtveen, E.L. Hamaker, Incorporating measurement error in $n = 1$ psychological autoregressive modeling, Front. Psychol. 6 (2015) 1038.
[147] N.C. Scollon, C.-K. Prieto, E. Diener, Experience sampling: Promises and pitfalls, strength and weaknesses, in: E. Diener (Ed.), *Assessing well-being: The collected works of ed diener* (pp. 157–180), 2009 Springer Netherlands. https://doi.org/10.1007/978-90-481-2354-4_8.
[148] J. Sharp, B. Keefe, Psychiatry in chronic pain: A review and update, Focus 7 (4) (2006) 213–580.
[149] C.A. Sims, Macroeconomics and reality, Econometrica 48 (1) (1980) 1–48.
[150] H. Singer, Sem modeling with singular moment matrices part ii: Ml-estimation of sampled stochastic differential equations, J. Math. Sociol. 36 (1) (2012) 22–43.
[151] A. Solanas, R. Manolov, V. Sierra, Lag-one autocorrelation in short series: Estimation and hypotheses testing, Psicológica 31 (2) (2010) 357–381.
[152] T. Stadnytska, J. Werner, Sample size and accuracy of estimation of the fractional differencing parameter, Methodology 2 (4) (2006) 135–141.
[153] A.A. Stone, J.E. Broderick, Real-time data collection for pain: Appraisal and current status, Pain Med. 8 (3) (2007) S85–S93.
[154] R. Tibshirani, Regression shrinkage and selection via the lasso: A retrospective, J. Royal Stat. Soc. Ser. B (Stat. Methodol.) 73 (3) (2011) 273–282.
[155] H. Tong, Threshold models in time series analysis—30 years on, Stat. Interface 4 (2) (2011) 107–118.
[156] K. Torre, D. Delignières, L. Lemoine, Detection of long-range dependence and estimation of fractal exponents through arfima modeling, Br. J. Math. Stat. Psychol. 60 (1) (2007) 85–106.
[157] T.J. Trull, U.W. Ebner-Priemer, Using experience sampling methods/ecological momentary assessment (esm/ema) in clinical assessment and clinical research: Introduction to the special section, Psychol. Assess. 21 (4) (2009) 457–462.
[158] T.J. Trull, U.W. Ebner-Priemer, Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices, J. Abnorm. Psychol. 129 (1) (2020) 56–63.
[159] T.J. Trull, S.P. Lane, P. Koval, U.W. Ebner-Priemer, Affective dynamics in psychopathology, Emot. Rev. 7 (4) (2015) 355–361.
[160] E.R. Tunks, J. Crook, R. Weir, Epidemiology of chronic pain with psychological comorbidity: Prevalence, risk, course, and prognosis, Can. J. Psychiatry 53 (4) (2008) 224–234.
[161] S. Van Buuren, Flexible Imputation of Missing Data. Chapman; Hall/CRC, (2018).
[162] van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Et al. (2014). Critical slowing down as early warning for the onset and termination of depression. Proc. Natl. Acad. Sci., 111(1), 87–92.
[163] L. van der Krieke, F.J. Blaauw, A.C. Emerencia, H.M. Schenk, J.P. Slaets, E.H. Bos, P. de Jonge, B.F. Jeronimus, Temporal dynamics of health and well-being: A crowdsourcing approach to momentary assessments and automated generation of personalized feedback, Psychosom. Med. 79 (2) (2017) 213–223.
[164] A. Van Giels, A. Emerencia, E. Bos, J. Rosmalen, Exploring interrelationships among worrying, anxiety, and somatic symptoms using time-series analysis, (2020) Manuscript under review.
[165] W.F. Velicer, S.M. Colby, A comparison of missing-data procedures for arima time-series analysis, Educ. Psychol. Meas. 65 (4) (2005) 596–615.
[166] Velicer, W. F., & Molenaar, P. C. (2012). Time series analysis for psychological research. In *Handbook of psychology, second edition.* American Cancer Society. https://onlinelibrary.wiley.com/doi/ abs/https://doi.org/10.1002/9781118133880.hop202022.
[167] M.C. Voelkle, A. Brose, F. Schmiedek, U. Lindenberger, Towards a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms, Multivar. Behav. Res. 49 (3) (2014) 193–213.
[168] M.C. Voelkle, C. Gische, C.C. Driver, U. Lindenberger, The role of time in the quest for understanding psychological mechanisms, Multivar. Behav. Res. 53 (6) (2018) 782–805.
[169] M.C. Voelkle, J.H. Oud, Continuous time modeling with individually varying time intervals for oscillating and non-oscillating processes, Br. J. Math. Stat. Psychol. 66 (1) (2013) 103–126.
[170] M.C. Voelkle, J.H. Oud, E. Davidov, P. Schmidt, An sem approach to continuous time modeling of panel data: Relating authoritarianism and anomia, Psychol. Methods 17 (2) (2012) 176–192.
[171] M.C. Voelkle, J.H. Oud, T. von Oertzen, U. Lindenberger, Maximum likelihood dynamic factor modeling for arbitrary n and t using sem, Struct. Equ. Model. Multidiscip. J. 19 (3) (2012) 329–350.
[172] A. von Leupoldt, F. Riedel, B. Dahme, The impact of emotions on the perception of dyspnea in pediatric asthma, Psychophysiology 43 (6) (2006) 641–644.
[173] T. von Oertzen, A.M. Brandmaier, Optimal study design with identical power: An application of power equivalence to latent growth curve models, Psychol. Aging 28 (2) (2013) 414–428.
[174] E.-J. Wagenmakers, S. Farrell, R. Ratcliff, Estimation and interpretation of $1/f\ \alpha$ noise in human cognition, Psychon. Bull. Rev. 11 (4) (2004) 579–615.
[175] E.-J. Wagenmakers, S. Farrell, R. Ratcliff, Human cognition and a pile of sand: A discussion on serial correlations and self-organized criticality, J. Exp. Psychol. Gen. 134 (1) (2005) 108.
[176] T.A. Walls, J.L. Schafer, Models for intensive longitudinal data, Oxford University Press, 2006.

[177] B. Wild, M. Eichler, H.-C. Friederich, M. Hartmann, S. Zipfel, W. Herzog, A graphical vector autoregressive modeling approach to the analysis of electronic diary data, BMC Med. Res. Methodol. 10 (1) (2010) 28.
[178] J. Windt, C.L. Ardern, T.J. Gabbett, K.M. Khan, C.E. Cook, B.C. Sporer, B.D. Zumbo, Getting the most out of intensive longitudinal data: A methodological review of workload–injury studies, BMJ Open 8 (10) (2018) e022626.
[179] T. Yarkoni, J. Westfall, Choosing prediction over explanation in psychology: Lessons from machine learning, Perspect. Psychol. Sci. 12 (6) (2017) 1100–1122.
[180] Y. Zheng, R.P. Wiebe, H.H. Cleveland, P.C. Molenaar, K.S. Harris, An idiographic examination of day-to-day patterns of substance use craving, negative affect, and tobacco use among young adults in recovery, Multivar. Behav. Res. 48 (2) (2013) 241–266.
[181] J. Zink, M. Nicolo, K. Imm, S. Ebrahimian, Q. Yu, K. Lee, K. Zapanta, J. Huh, G. Dunton, M.I. Goran, K. Page, C.M. Dieli-Conwright, B. Belcher, Interstitial glucose and subsequent affective and physical feeling states: A pilot study combining continuous glucose monitoring and ecological momentary assessment in adolescents, (2020) Manuscript under review.
[182] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. Royal Stat. Soc. Ser. B (Stat. Methodol.) 67 (2) (2005) 301–320.