To cite this chapter:

Houben, M., Ceulemans, E., & Kuppens, P. (in press). Modeling intensive longitudinal data. In A. G. C. Wright & M. N. Hallquist (Eds.), *The handbook of research methods in clinical psychology*. New York, NY: Cambridge University Press.

Chapter 24: Modeling Intensive Longitudinal Data

Handbook of Research Methods in Clinical Psychology

Marlies Houben, Eva Ceulemans, Peter Kuppens

Faculty of Psychology and Individual Differences, KU Leuven, Tiensestraat 102, 3000 Leuven, Belgium

Research in clinical psychology is largely concerned with the investigation of the presence and severity of a range of different maladaptive feeling states, cognitions and behaviors, which we will refer to as psychopathological symptoms. Typical ways that have been used to assess such symptoms are clinical interviews or retrospective self-reported questionnaires, inquiring how people typically felt and behaved in a certain period of time or in a certain situation. For example, according to the Structured Clinical Interview for DSM- disorders (First, Williams, Karg, & Spitzer, 2015)and the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), respectively, depressed mood is typically assessed using the following items: “In the last month, has there been a period in which you felt down or depressed for the largest part of the day”, and “During the past week, I felt depressed”.

However, research increasingly finds that the way these symptoms behave over time — not their mere presence or absence — is a crucial feature that needs to be studied for a better understanding of the phenomenology and mechanisms underlying psychopathology (Myin-Germeys et al., 2009). Indeed, many psychopathological symptoms are not stable or constant over time; rather, they fluctuate and change at a moment-to-moment level, often in response to changes in the environment. Moreover, several symptoms or diagnostic criteria are inherently dynamic, such as emotional instability in borderline personality disorder (American Psychiatric Association, 2013; Ebner-Priemer et al., 2015; Houben et al., 2016), or persistent depressed mood for a depressive episode (American Psychiatric

Association, 2013), implying the importance of studying psychopathology from a dynamical viewpoint.

In order to examine the ups and downs of symptoms over time, a data collection method that allows for repeated moment-to-moment assessments of these symptoms in daily life or in the lab is required. In daily life, ambulatory assessment methods (AA) are a very promising approach in which self-reported symptoms, behaviors, or physiological processes are assessed in real-time using computerized devices (i.e. smartphones, ambulatory blood pressure devices etc.), while participants undertake their normal daily activities (Trull & Ebner-Priemer, 2013; see also Chapter XX). Similarly, repeated moment-to-moment assessments can also be collected in the lab, for example, in response to standardized stimuli that are presented to participants on a computer screen (Koval, Pe, Meers, & Kuppens, 2013), or during conversations with social interaction partners (Hollenstein & Lewis, 2006; Kuppens, Allen, & Sheeber, 2010). Repeated assessments can involve self-reported data, observational data, physiological parameters, and so forth.

These data collection approaches typically result in intensive longitudinal data, consisting of many repeated measurements from single participants, that are typically collected over many different time points during the day and potentially also over several consecutive days. These long strings of data contain rich dynamic information, as they allow us to explore the frequency and duration of symptoms over time, as well as dynamic properties of symptoms that are obscured by more traditional data collection methods. Relatedly, they allow us to examine the relationship among symptoms within-person. Rather than focusing on how symptoms co-occur on a person-level (e.g. people with higher levels of depressed mood also tend to experience fatigue), intensive longitudinal data allow us to examine how symptoms co-occur or predict one another in the moment within an individual person (e.g. at times where a participant experiences more depressed mood, this co-occurs with higher levels of fatigue), and how these s within-person relations differ between people, as a function of person-level features, such as diagnostic status or gender.

In this chapter, we present an overview of some major dynamical properties of single or multiple symptoms (and their interrelations) that can be studied with intensive longitudinal data, and how they can be calculated or modeled statistically. We focus primarily on properties that characterize time series as a whole, thereby implicitly assuming that these dynamic features remain unchanged throughout the time series (i.e., the models we present

here assume stationary processes). However, it is important to note that these properties can change abruptly or more gradually over time under certain conditions, such as when symptoms become less unstable over time due to treatment progress, or become more or less correlated with one another after the occurrence of a significant event. For simplicity, the majority of methods and models discussed in this chapter do not address such changes in dynamics, and assume stationarity. However, in the discussion section, some models and techniques will be briefly discussed that can be used to investigate such changes in dynamics.

We will first describe methods and models to examine the dynamic properties of a single symptom. Next, we will present methods and models to explore the dynamic relationships between two or more symptoms. For each approach, we will provide information on how to calculate simple indices on a more descriptive level, as well as how to model the dynamic features using more complex models. See Table 1 for an overview of the different dynamic properties that can be examined, involving one symptom (i.e. univariate approaches), and two or more symptoms (multivariate approaches), which indices or models to use, and what they exactly reflect. Last, in the discussion section, we will end with some additional considerations that should be taken into account when dealing with intensive longitudinal data.

Univariate approaches: the dynamics of one symptom over time

Using intensive longitudinal data, we can explore how a single symptom, for example, depressed mood, behaves over time. In doing so, one can focus on features related to the central tendency of the symptom, the spread and the dynamics over time.

Central tendency. On a descriptive and most simple level, several indices can be calculated to summarize the central tendency of the time series of one symptom, such as depressed mood. The average or median level of depressed mood of an individual reflects the typical or average moment-to-moment level of depressed mood for that person. Similarly, one can examine the frequency with which a certain symptom, such as depressed mood or the occurrence of non-suicidal self-injurious behavior, occurs throughout the time series. For symptoms that are assessed on a binary scale (such as non-suicidal self-injury) or after first dichotomizing symptoms that are assessed on a continuous scale at a clinically relevant cut-off score, one can calculate the proportion of time points in which a symptom

was present for each person. These indices provide information about overall or average intensity and frequency of occurrence of a symptom across a certain period of time.

When data from multiple individuals are available, one can examine between-person patterns in average symptom intensity and frequency using multilevel regression models (Hox, Moerbeek, & Schoot, 2010; Nezlek & Plesko, 2001; John B. Nezlek, 2008). If time series data of an individual are short, these models often provide more accurate estimates of the parameters of interest, because data from multiple people are combined in one model (Bulteel, Mestdagh, Tuerlinckx, & Ceulemans (2018). In case of repeated assessments nested within persons, the average or overall intensity of a symptom, such as depressed mood, can be modeled with a two-level linear regression model that includes an intercept that is allowed to differ between people (i.e., a so-called random intercept), and no other predictors (unless predictors are entered person-mean centered; See also chapter XX). These intercepts reflect the average depressed mood across all time points for each person. One can next investigate how these average levels of depressed mood differ as a function of certain person-specific variables, such as diagnostic status, by adding this variable as a predictor to the person level of the model. For example, Trull et al., (2008) used such an approach to examine differences between persons with a diagnosis of borderline personality disorder and major depressive disorder, and found that both groups tend to experience similar levels of overall momentary positive and negative affect.

In a similar way, the frequency or likelihood of occurrence of symptoms can be modeled with a multilevel logistic regression model, in which the log-odds for a symptom to occur at the moment-to-moment level is modeled with a random intercept, again allowing for individual differences between people, that can again be modeled in function of person-specific features. For instance, Houben et al., (2016) used this approach to study the occurrence of emotional switches between positive and negative emotional states in individuals who were diagnosed with borderline personality disorder versus healthy controls. However, no differences were found, which means that healthy participants and persons with a borderline personality disorder were equally likely to jump from a positive emotional state to a negative emotional state (or vice versa) on consecutive time points.

Spread measures. Moving beyond summaries of central tendency using means, medians and frequencies, the examination of the dispersion of these repeated assessments of symptoms can provide additional interesting insights. Dispersion scores reflect how much

individuals tend to deviate from their own overall or typical intensity of a symptom over time reflecting the degree of variability in the intensity of a symptom over time. On a descriptive level, the within-person standard deviation of the repeated scores per person provides the simplest index of variability. This measure reflects the degree of fluctuation or variability around the mean intensity level of a symptom. Similarly, and in combination with the median as a summary measure, one can compute the interquartile range for each person, indicating the range of the middle 50% of scores for each person. .

In a regression context and for data of multiple individuals, variability of a symptom over time can also be modeled using multilevel models. In case of repeated assessments nested within persons, the error variance at the moment-to-moment level would capture the variability of a symptom within-person over time. For instance, such an approach was used by Peeters, Berkhof, Delespaul, Rottenberg, and Nicolson (2006), in which variance at both moment-to-moment and day levels was compared between participants with major depressive disorder and healthy controls. Results showed larger moment-to-moment variability in negative affect in the depressed group. However, note that models are needed that allow for heterogeneity of variance across levels of a covariate (for example diagnostic status, thereby allowing variance estimates to differ between diagnostic groups), or across subjects (i.e. within-person variance estimates that are allowed to vary between people). This is not a standard assumption of multilevel models, as usually this within-person variance is assumed to be the same (i.e. homogeneous) across subjects. Hedeker, Mermelstein, and Demirtas (2012) describe such a model in which both the mean and within-person variance of repeated assessments are allowed to differ between subjects.

Additionally, one can also examine the proportion of variance in the independent variable at each level. Estimating empty multilevel models with only intercepts, the estimated variance at the moment-to-moment level and the person-level will typically be reported in the output. Next, one can compute the proportion of variance in the dependent variable at each level by dividing the variance at each level of the model by the total variance. In this context, an intraclass correlation coefficient (ICC) is an often used metric, calculated as the ratio of between-person variance and total variance. This reflects the proportion of variance located at the between-person level. The proportion of within-person variance is 1-ICC. This provides insight into where variability in the intensity of a symptom is

located: is the observed variance in symptom scores mainly driven by differences between people, or by differences from moment to moment within individuals?

Although these measures of spread provide some interesting information, one disadvantage is that they do not take into account *when* the changes in the intensity of a symptom occurred. Two participants could exhibit high levels of variability in the intensity of their depressed mood despite different temporal patterns. For example, one person could experience an increase in depressed mood followed by a decrease, following an inverted U-shape, while a second person could experience frequent ups and downs. Nevertheless, across the full time series, both participants may deviate from their mean levels to a similar extent, resulting in a similar overall variance. To capture the time course of symptoms, the dimension of time has to be taken into account.

Dynamics. When we are interested in how emotions behave over time, we can go beyond the dispersion measures and focus on different aspects of the dynamics of a symptom over time. As such, a researcher can examine how unstable or abrupt symptoms change over time, what the duration of symptoms is, how self-predictive symptoms are over time, or we can explore the full time-course of a symptom.

Instability. A first interesting temporal aspect is the degree to which symptoms abruptly go up and down over time, i.e. the degree of instability. One way of examining the ups and downs of symptoms, while taking into account the temporal dependency of the changes, is to calculate the mean squared successive difference (MSSD) or mean absolute successive difference (MASD) between consecutive assessments of a symptom. With this approach, we calculate the successive difference in the intensity of a symptom. Next, we square these differences (implying that more weight is given to larger changes). Another approach involves taking the absolute value of these differences, which is less influenced by extreme changes. This is done to remove signs, because we want to capture the magnitude of changes, irrespective of the direction of the change. Next, the average is taken of these successive (squared or absolute) differences, yielding an index for instability. The square root of the MSSD (RMSSD) can also be computed, to obtain an instability index that is in the original unit of the symptom scores. Because the MSSD and related indices quantify the average magnitude of change within a smaller time interval, rather than across the entire time series as is the case for dispersion measures, these measures are better able to capture frequent ups and downs in the intensity of symptoms over time. Indeed, large abrupt

changes that occur within short time intervals will yield larger MSSD values, while dispersion measures (such as the within-person standard deviation) can result in similar values both if changes occur orderly (i.e. the intensity slowly increases and subsequently decreases), or more abruptly (i.e. the intensity of a symptom rapidly shifts up and down). To illustrate, Figure 1 shows time series data of two fictive subjects. Both time series consist of exactly the same data points, therefore yielding the same mean and standard deviation across the time series data for both subjects. For subject 1, data points are ordered in such a way that intensity gradually increases and decreases, resulting in a low MSSD index. On the other hand, for subject 2, the intensity rapidly shifts up and down from time point to time point, resulting in a high MSSD index. This illustrates that, in contrast to measures of spread (reflecting overall variability), the MSSD indices capture time-ordered instability in symptoms. This approach has for instance been used by Thompson et al., (2012), who found significantly higher levels of emotional instability in negative affect but not positive affect in persons with a diagnosis of major depressive disorder compared to healthy controls.

Next to computing these metrics for each person on a descriptive level, we can again model MSSD and related metrics with multilevel models for multiple people at the same time. In case of repeated assessments of depressed mood, nested within persons, we can estimate two-level linear regression models in which the squared successive difference or the absolute successive difference of a symptom between any two time-points is used as the outcome, and modeled with a random intercept at the moment-to-moment level. This random intercept will again allow for individual differences between people in the average estimate, which can again be modeled as a function of person-specific characteristics, such as diagnostic status at the person-level of the model. Such an approach was used by Koval, Ogrinz, et al., (2013) for instance, in which they showed that emotional instability in positive affect, as measured with the MSSD of emotional states over time, was related to lower heart-rate variability, a physiological indicator of emotion regulation capacity. More detailed information about the MSSD measure and how it can be modeled can also be found in Jahng, Wood, & Trull (2008).

Of note with measures of instability is that a correction for differences in time intervals between consecutive measurements is sometimes necessary. In case the length of time intervals differs largely between consecutive measurements, one can adjust each

successive change for the actual time that elapsed, as was done by Thompson et al. (2012) and Trull et al. (2008). However, even if time intervals between consecutive measurements (within a day) are relatively equal, one should consider removing or correcting for overnight changes, as changes in a symptom between the evening and the next morning might not be meaningful.

Duration. Duration is another dynamic aspect of symptoms over time. One might be interested to know how long a symptom, such as depressed mood, persisted at or above a clinically meaningful severity level. To obtain this information on a descriptive level, a count of the number of time points, or a sum of the continuous time, that a symptom was consecutively rated or observed at or above a pre-specified intensity level would provide this information.

Survival analyses can be used to model the expected duration until a certain event happens, such as the termination of an ongoing process. Hence they can be used to model the duration of symptom, such as the time until depression levels decrease below a clinical cut-off score. An example of the application of the use of discrete-time survival analyses to model duration can be found in Verduyn, Delvaux, Van Coillie, Tuerlinckx, and Van Mechelen, (2009) where they used these models to estimate the duration of an emotional episode. Such an analysis provides several statistics related to symptom duration such as the hazard rate, which reflects the conditional probability that a process, such as an ongoing depressive episode, that has not yet ended at the beginning of a certain time interval will end during that interval. This hazard rate can next be modeled as a function of person-level predictors. More information can also be found in a study by Sbarra (2006).

Autocorrelation. Intensive longitudinal data allow for the examination of autocorrelations or autoregressive effects of symptoms over time. The autocorrelation captures how self-predictive a symptom is over time. A high autocorrelation reflects high self-predictability meaning that the current intensity of a symptom is highly predictive of the intensity of the symptom at a following time point, and thus it strongly carries over from one moment to the next. A high autocorrelation thus implies that a symptom is lingering, spiraling off and getting stuck, showing relatively little homeostatic recovery to normative states. It reflects strong resistance to change, or resistance to outside influences: independent of what is happening in the environment, symptoms are self-perpetuating. In the clinical and non-clinical literature, high levels of self-predictability are often referred to

as ‘inertia’, named after the concept of inertia in physics, which refers to the slowness and resistance of an object to change its current state of motion. To illustrate, depression has been consistently linked to high levels of emotional inertia, especially for negative affect (Houben, Van Den Noortgate, & Kuppens, 2015), showing that depression is characterized by patterns of negative emotions that have become stuck, and resistant to chance. The negative emotions have spiraled off, without homeostatic return to baseline. Moreover, research has consistently shown that persons with a diagnosis of borderline personality disorder exhibit large within-person variance in their emotional experiences. Additionally, some indications were found for stronger autoregressive effects in their emotional states (Ebner-Priemer et al., 2015; Houben et al., 2015). These findings imply that persons with borderline personality disorder tend to deviate more strongly from their emotional baseline levels, and subsequently tend to linger, and show slower return to baseline, indicating lower emotional recovery. On a descriptive level, an autocorrelation measure for each person can be calculated by taking the bivariate correlation of a symptom, with a time-lagged version of itself (i.e. the variable is shifted down with one or more time points), reflecting how strongly the intensity of a symptom is related to the intensity of the same symptom at the previous time point. Using a more advanced modeling approach for data from multiple individuals, we can use multilevel autoregressive models (multilevel AR model), where a symptom is predicted by a random intercept, and a time-lagged version of itself at the moment-to-moment level. Making this autoregressive slope a random effect facilitates capturing individual differences between people in the strength of this autoregressive effect, which can further be modeled as a function of person-specific variables. Similar approaches have been used in many studies, in which inertia of, for example mood, was typically related to low levels of well-being and higher levels of psychopathology (e.g. Kuppens et al., 2010; Thompson et al., 2012). More information regarding the modelling of inertia can also be found in Jahng et al. (2008).

Often, a lag of 1 is used in such AR models (i.e. first order AR models). However, note that depending on the time scale on which a process of interest is likely to occur and the sample frequency that was used in a study, a different number of lags might be required. A way to determine the correct order of the AR model (.i.e. the number of lags needed) is to check for the absence of serial dependencies in het residuals. Different methods are

available to do so (for the AR model see e.gg. Box & Jenkins; for the VAR model see e.g. Brandt & Williams, 2007).

Commonly expected values for such autoregressive effects in psychological data typically range between 0 and 1 (Rovine, & Walls 2006), with 0 reflecting no self-predictive effects, and 1 very high self-predictive effects. However, sometimes, also negative autoregressive effects can be found. For example, processes following a sine curve, in which low intensity at one point in time predicts high intensity at another time point, can result in negative autocorrelations, when captured at the correct time scale. Examples of such processes might be diurnal hormone levels, mood patterns in those with rapid cycling bipolar disorder, or food intake.

Note that measures of spread (i.e. within-person variance), instability (i.e. MSSD), and autocorrelation are not independent, but are statistically related. More information about these relations can be found in Jahng, Wood, and Trull (2008).

Time course. Last, for several research questions, it can be interesting to examine the overall course of a symptom over time. For example, how does the intensity of someone's depressed mood progress during the day? Such a question would involve the investigation of diurnal patterns of symptoms. However, the course of a symptom might also be examined over a longer period of time, such as a year. On a descriptive level, the simplest way to examine the time course of a symptom is to plot repeated measurements of a symptom over time, and visually inspect the time course. Such a visual inspection can provide insight into possible time trends in the intensity of a symptom over time. For example, does depressed mood increase or decrease in intensity across the course of a typical day, and how steep is the increase or decrease? On a more complex level, time courses can also be modeled in different ways. First, when there are data from multiple individuals, one can examine how symptom severity changes as a function of time, by using multilevel models, and adding time as a predictor of symptom severity at the moment-to-moment level. As such, we can examine whether symptom severity such as depressed mood increases or decreases linearly with time, or whether a quadratic relationship with time can be observed, which reflects a U-shaped or inverted U-shaped time course of symptom intensity over time. For example, this approach was used in Trull et al., (2008), in which time trends in positive and negative affect over days and within days in persons with borderline personality disorder and depression were examined and compared, and in Peeters, Berkhof, Delespaul, Rottenberg,

and Nicolson (2006), where they examined linear and quadratic diurnal patterns in affect, in persons with and without depression.

More broadly, growth curves are a highly flexible class of models that can be used to model growth and change in time-related phenomena and symptoms, and can accommodate a variety of shapes and structures of change. More detailed information can be found in McArdle and Nesselroade (2003).

The time course of symptoms can take many different shapes, however not always following linear or quadratic functions. As such, some people can experience depressed mood very intensely at the start of the day, and then experience a slow decrease, while other could well experiences the reverse pattern or experience an intense onset at the start of the day, with a second recrudescence later that day. Therefore, more advanced models have been developed that allow the modeling of more complex features of the time course of the intensity of symptoms, or intensity profiles. For example, the K-spectral centroid clustering method (Heylen, Van Mechelen, Verduyn, & Ceulemans, 2016; Yang & Leskovec, 2011) was developed to model different shapes (such as different steepness of onset and skewness) and different amplitudes (height of the profile) of such intensity profiles over time. With this method, first, based on all available data, reference or example profiles are determined, each of which captures a distinct profile shape that can be observed in the data based on shape and amplitude. An example of a typical shape is the early blooming shape, with a steep onset and peak in the intensity of a symptom at the beginning, followed by a slow return to baseline. More examples can be found in Heylen et al., 2016 and Heylen, Verduyn, Van Mechelen, and Ceulemans, 2015) Next, each observed intensity profile is assigned to one of the determined reference profiles, and receives an amplitude coefficient which indicates the extent to which the reference profile needs to be inflated or deflated to optimally approximate the observed intensity profile. As such, for each person can be determined which of several possible intensity profile shapes best describes the time course of their symptoms over time. Next, assignment of profile shape, but also duration and amplitude can be predicted by person-levels variables, such as diagnostic status. This approach is illustrated in a study by Heylen, Verduyn, Van Mechelen, and Ceulemans (2015) where they explored different shapes of intensity profiles of episodes of experienced anger. They also examined how these different shapes were related to emotion regulation use. First, they identified two types of anger episodes in the data: early-blooming episodes of

anger, during which experienced anger lasts relatively short and reaches a peak in intensity shortly after the start of the episode, and late-blooming episodes of anger, which last longer and reach their peak relatively later in the episode. Next, they found that early-blooming episodes of anger were associated with adaptive emotion regulation strategies, such as cognitive reappraisal, while late-blooming episodes of anger were linked to maladaptive strategies, such as rumination. Moreover, emotion regulation strategy use was related to the amplitude (reflecting overall intensity) and duration of anger episodes, with adaptive strategies being linked to shorter duration and lower amplitudes. In case several intensity profiles per person are collected, for example one for each day or profiles of multiple symptoms (meaning that intensity profiles are nested within person), a hierarchical (i.e. multilevel) extension of this model can be used to examine difference within the same person regarding the shape of intensity profiles he or she experiences, and how these differences related to other variables, such as overall well-being that day (Heylen et al., 2016).

Multivariate approaches: the dynamics of multiple symptoms

The first part of the chapter focused on how to model several dynamic features of a single symptom. Researchers or clinicians are often interested in how several symptoms behave over time, and more importantly, how they mutually co-occur and influence each other. Several options exist to explore such dynamic properties of multiple symptoms. We first propose some ways to examine concurrent relationships (i.e. relations between variables, measured at the same time) between two or more symptoms, and next discuss some options to investigate possible prospective relationships further. We will end with some more advanced techniques that allow for simultaneously modelling several of these properties and relations.

Concurrent relationships. Some research questions involve the examination of how two or more symptoms are related or co-occur in the moment. For example, does depressed mood typically co-occur with fatigue on a moment-to-moment level, and is the association positive or negative? On a descriptive level, such questions can be answered by computing indices of co-occurrence for each individual. First, a within-person correlation per person quantifies the strength and the direction of the moment-to-moment relationship between two symptoms for each individual. In case of data from multiple people, one can similarly center the variable within-person (i.e., center each person's time-series on their mean) and

next compute a normal correlation coefficient for all data points pooled across individuals. In the case of more than 2 variables, the intra-class correlations (ICC) can be used. Although different types of ICC's exist, the most often used ICC in this research context quantifies consistency between a group of related variables across different time points (Shrout & Fleiss, 1979), which is equivalent to a cronbach's alpha. A high ICC indicates a high correlation between several variables across different measurements points. For example, the ICC has been extensively used in research on emotion differentiation, which reflects the degree to which people label their emotions in a differentiated and specific way, rather than in global ways. In this field, ICC is used as a measure of emotion differentiation, where a high ICC is assumed to reflect low emotion differentiation (i.e., emotions are rated as rising and falling together), as emotions covary strongly with each other across the different measurement points, and therefore people do not strongly distinguish between these emotions (Erbas, Ceulemans, Pe, Koval, & Kuppens, 2014; Tomko et al., 2015).

Using a modeling approach for data from multiple individuals, the relationship between two or more symptoms can be examined using multilevel regression models, in which the intensity of one symptom is predicted by one or more other symptoms at the moment-to-moment level. Such analyses can provide insight into how strongly and in which direction (positively or negatively) one symptom, such as depressed mood, is related to one or more other symptoms, such as fatigue and concentration problems, in the moment. Using random slopes, individual differences in these relationships can be modeled as a function of other person-specific variables, such as diagnostic status. An example of such models can be found in Hepp, Carpenter, Lane, and Trull (2016), who examined how several core features of borderline personality disorder, such as negative affectivity, impulsivity, and interpersonal problems are affected by close social contact on a momentary level, and how personality traits further impact these relationships.

Another approach to examine concurrent relationships between several symptoms across time is to apply principal component analysis (PCA) to the data of each individual. This dimension reduction approach summarizes the variables (i.e., symptoms) to one or more components, that explain as much of the variance across the time points as possible (Cattell, 1952; Jones & Nesselroade, 1990) and are a linear combination of the symptoms. As in standard principal component analysis or factor analysis of cross-sectional data, the

interpretation of the components is based on the loadings of the variables on the components. For example, if depressed mood and other mood related symptoms would load highly on the same component, that component would likely reflect affect-related disturbances. Symptoms that load strongly on the same component typically covary significantly over time. Each time point has a score on each of the components, which can be plotted against time to visualize the fluctuations over time. Extensions of this approach have been proposed that allow to analyze symptom data from multiple individuals simultaneously, such as (multilevel) simultaneous component analysis; Timmerman, 2006; Timmerman & Kiers, 2003; Ceulemans, Wilderjans, Kiers, & Timmerman, 2016. These models allow to extract one set of components for all persons (i.e., the loadings are fixed across persons), and derive component scores for each time point of each individual. These component scores allow to examine differences between persons in within-person variance (e.g. one person might show more variability in affect-related disturbances over time than others). Moreover, one can inspect differences between people in within-person component correlations: the extracted components might co-occur more strongly over time for some persons and be almost independent for others.

Lagged or prospective relationships. One major strength of intensive longitudinal data is that it allows us to look beyond concurrent relationships between two or more symptoms, enabling researchers to examine prospective relationships between symptoms. That is, detailed analyses can reveal how symptom intensity, such as momentary fatigue, is related to other symptoms, such as depressed mood, at the next point in time, providing information about the temporal directionality of the relationship. On a descriptive level, an index for such prospective associations is the partial lagged correlation coefficient. This is computed by taking the correlation of a variable, such as depressed mood, with a second variable, such as fatigue, after being lagged one time point. Such a correlation quantifies the strength and the direction of the association between current depressed mood, and preceding fatigue. However, to preclude that the obtained association is driven by concurrent, rather than prospective relationships, the partial correlation is preferred, in which is additionally controlled for the autocorrelation effect (i.e. the lagged version) of the first variable (i.e. depressed mood in our example), and possibly also the concurrent version of the second variable (i.e. fatigue in our example).

On a modeling level, using data from multiple persons, prospective associations can be modeled with multilevel vector-autoregressive (VAR) models. Such models are extensions of the AR-models described earlier in which, next to the lagged version of the dependent variable (i.e. autoregressive effects), additional predictors measured at the previous time point are also added (i.e. cross-regressive effects). Using the same example, examining how fatigue is associated with depressed mood at the following time point, multilevel VAR-models can be used where depressed mood is predicted by fatigue at the previous time point (i.e. lagged with one time point), also adding depressed mood at the previous time point as a predictor. Possibly current fatigue can also be added as a predictor on the moment-to-moment level. This allows us to examine to what degree current depressed mood is predicted by or associated with fatigue at the previous time point, above and beyond depressed mood at the previous time point, and current fatigue. Using random slopes, individual differences in these prospective associations can be modeled, also as a function of person-level features, such as diagnostic status. For example, Houben and colleagues (2017) used VAR-models to examine the effect of the occurrence of non-suicidal self-injury on the intensity of subsequent positive and negative affect, correcting for preceding positive and negative affect, and found that the occurrence of NSSI is related to high levels of negative affect and low levels of positive affect at the following time point

Note that next to the associations between symptoms, also context information can be included in such models, allowing for the investigation of symptom reactivity to contextual variables. As such, estimating models in which depressed mood is predicted by depressed mood at the previous time point, and a context variable, such as the occurrence of a negative event or a social encounter, allows us to examine depressed mood reactivity on event occurrence. Such models have been used in research examining mood reactivity in psychopathology (e.g. Bylsma, Taylor-Clift, & Rottenberg, 2011).

Advanced techniques. Last, next to examining concurrent or prospective relationships between two or more variables, more advanced techniques exist that allow for the modelling of several types of relationships or more complex relations between two or more symptoms simultaneously. While we will not provide an exhaustive list of different models that exist, we will give a short overview of some possibilities.

Descriptive techniques. On a descriptive level, the state-space approach (extensively used by Hollenstein and colleagues; e.g. Hollenstein & Lewis, 2006; Lougheed & Hollenstein, 2016) provides a method to visually inspect how two categorical variables co-evolve over time and to derive descriptive indices regarding the dynamics of two variables simultaneously. Typically using specialized software, GridWare 1.1 (Lamey, Hollenstein, Lewis, & Granic, 2004), these methods have been used to inspect and describe the trajectory of dyadic emotions (i.e. combinations of emotional states from two individuals) of parent-child pairs during social interactions. These different dyadic emotional states are typically depicted in a grid, with the emotional state of one person on the x-axis, and the emotional state of the second person in the y-axis. Each combination represents a possible dyadic emotional state. Next, the trajectory of these dyadic states can be inspected, and indices can be calculated to quantify the dynamics of these states. As such, a dispersion measure can be computed that reflects the range of different dyadic states that are endorsed across the grid. This dispersion measure then reflects whether all behavior occurred in one cell versus whether behavior was equally distributed across the grid. Second, a measure of transitions can be computed, which captures the number of changes or transitions between different cells on the grid. In the context of dyadic emotion states, both measures are assumed to reflect greater flexibility, and have been shown to be associated with better psychosocial adjustment (Hollenstein & Lewis, 2006; Lougheed & Hollenstein, 2016).

Modeling approaches. Using a modeling approach, different options exist to examine the dynamics of two or more symptoms and complex interrelations between the symptoms. (Multilevel) vector-autoregressive (VAR) models can be used, in which, next to the lagged version of the dependent variable, a range of different predictors measured at the previous time point are also added. As such, a VAR model can be used to determine to what degree a symptom is predicted by itself, and by a range of other additional symptoms, measured at the previous time point. Next, these VAR-models can form the basis for network models (Bringmann et al., 2016), in which the unique and direct effects between a set of symptoms are visualized. Coefficients derived from such VAR-models, in which each time a different symptom serves at outcome, are used to determine the strength of the connections or ‘edges’ between the different symptoms or ‘nodes’ in such a network model (for some cautionary notes, however, see Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016). Additionally,

several features can be derived to describe the dynamical interplay between the symptoms or nodes in the network model, such as the density of the overall network, which indicates how strongly the symptoms in the network are interconnected (Pe et al., 2015), or node centrality, referring the importance of a single node in the network (Bringmann et al., 2013).

To examine the interrelationships between several symptoms within persons, Group Iterative Multiple Model Estimation (GIMME) is also a promising approach in which, next to several lagged relationships, as is the case for VAR-models, also contemporaneous relationships within-person over time can be examined. It is a structural equation method (SEM)-based method for identifying group-, subgroup-, and individual-level relations within time series data of several variables. Unlike the multilevel VAR model, GIMME estimates person-specific models for each individual, as well as searching for common features across individuals. This takes a different approach to dealing with heterogeneity of effects across individuals. This method can be applied, using the GIMME R package. Initially developed for fMRI data, it has recently been proven to work well with intensive repeated behavioral measures (e.g., daily diary data; Lane et al., in press; Wright et al., in press). More information can be found in Lane & Gates (2017).

Another approach to simultaneously model several dynamic features of two separate symptoms is the Bayesian (and Hierarchical) Ornstein–Uhlenbeck Model (BOUM; Kuppens, Oravecz, & Tuerlinckx, 2010; Oravecz, Tuerlinckx, & Vandekerckhove, 2011). It is based on the Ornstein–Uhlenbeck (OU) model, with an hierarchical (i.e., multilevel) extension. This implies that it can take into account individual differences between people. Moreover, it is estimated in a Bayesian framework. Using the BOUM model, three different dynamic parameters of two separate variables or dimensions can be estimated simultaneously: (1) it estimated a homebase for each dimension, which reflects the average or normative state for a variable, (2) it estimates the amount of variability around this homebase over time, and (3) it estimates the attractor strength, which reflects regulatory forces that pull a variable back towards the homebase. Because these processes are estimated for two variables at the same time, covariation between the two dimensions can be considered. As the model also allows for individual differences in the estimated parameters, these parameters can be regressed on covariates, to explain these individual differences. Moreover, this model takes time into account as a continuous variable, meaning that it takes into account the actual

time that elapsed between consecutive ratings. Therefore, it is ideal for measurements taken at possibly irregular time points. The BOUM model has been used in several studies (Ebner-Priemer et al., 2015; Santangelo et al., 2016) on affective dysregulation in borderline personality disorder (BPD), as defined by Linehan (1993). According to Linehan (1993), this dysregulation is reflected in heightened sensitivity to emotional stimuli, strong emotional reactivity, and a slow return to baseline. These three parameters can be mapped, one-on-one, onto the homebase, variability, and attractor strength parameters that are estimated in BOUM. Therefore, the BOUM model offered a way to estimate all three parameters underlying emotional dysregulation in BPD simultaneously, and to examine possible differences between persons with a diagnosis of BPD and healthy controls in these parameters, by adding diagnostic status as a covariate in the model. For detailed statistical information about the BOUM model and custom-made statistical software to run the analyses, see Oravecz et al. (2011) and Oravecz, Tuerlinckx, & Vandekerckhove (2012).

Discussion

The behaviors, thoughts, and feelings that indicate psychopathology are often not of a static nature, but rather change and fluctuate over time in response to changes in the environment and daily life situations. Therefore, clinical psychology research can benefit from focusing on how psychopathological features behave over time, as it can provide new perspectives and insights concerning the phenomenology and mechanisms underlying psychopathology.

The collection of intensive longitudinal data, consisting of many repeated measurements from single participants, allows for the investigation of several dynamical properties of single or multiple symptoms (and their interrelations). In this chapter, we presented an overview of some major dynamic properties that can be studied with intensive longitudinal data, and how they can be calculated or modeled statistically. First, we focused on several univariate approaches, allowing the examination of one single symptom or feature over time. Next, we discussed some methods and models to further examine the dynamic relationships between two or more symptoms. The diagram shown in Figure 2 provides a further guide on which method or model to use for which kind of data and/or research questions.

To end, some general issues that should be taken into account when dealing with intensive longitudinal data. First, in this chapter, we described several methods and models to examine dynamic features of symptoms and the interrelations between symptoms. However, as noted in the introduction, most methods and models discussed implicitly assume that these dynamic features of symptoms or the interrelations between symptoms remained unchanged throughout the time series. This is of course not always the case, as properties or relations between symptoms can change or shift abruptly, for example after the occurrence of a significant event, or gradually, for example during and after treatment. If it is known where the change occurred, for example because a participant initiated a treatment, single case study designs can be used to further explore these changes. For more information about such models, see for example Manolov and Onghena (2017). If it is not known where and how many changes occurred, changes can be detected using change point analyses (for a comparison of several techniques, see for instance, Cabrieto, Tuerlinckx, Kuppens, Grassmann, and Ceulemans, 2017)). Such a technique was also used in Wichers, Groot, Psychosystems, ESM Group, and EWS Group (2016) to detect a sudden change in depressive symptoms in data from one participant that had a history of multiple episodes of major depression; for a further analysis of this data, see Cabrieto, Tuerlinckx, Kuppens, Hunyadi, and Ceulemans, 2018).When changes in dynamic features occur more gradually, one can consider detrending the data first (Jahng et al., 2008), or applying methods that can deal with gradual changes in dynamic features such as time-varying AR and VAR models (Bringmann et al., 2017; Bringmann, Ferrer, Hamaker, Borsboom, & Tuerlinckx, 2018)

Second, when examining dynamic features of symptoms time plays an essential role. However, note that for some models and indices (such as the MSSD, AR and VAR models) time is considered as a discrete variable, meaning that time is seen as a categorical variable consisting of mere consecutive time points. Other the other hand, time can also be considered as a continuous variable, where the actual elapsed time is taken into account. Examples of such models that are suited for time series data are the BOUM model described above (Oravecz et al., 2011), and the time-varying structural equation modeling (that comes with a R package ctsem; Driver, Oud, & Voelkle, 2017). Such models would be more optimal when large variability exist in the time intervals between consecutive measurements.

Third, the more advanced models might offer elegant ways of modeling complex dynamic features and complex interrelations between symptoms. Still note that very complex models are not always the best choice, and more simple indices of dynamic features and interrelations between symptoms are often to be preferred in several situations, for example if the length of the time series data is limited (Bulteel, et al., 2018; Krone, Albers, & Timmerman, 2016; Liu, 2017).

Last, dynamic features of symptoms can be typically correlated with average levels of the symptom. Therefore, often correction for average levels of a symptom should be taken into account when examining dynamic features, for example see a method proposed by Mestdagh, Pe, Pestman, Verdonck, Kuppens, and Tuerlinckx (in press).To conclude, intensive longitudinal data allow us to examine how clinical symptoms and features change over time and dynamically influence each other. As a consequence, it offers us a new approach for studying psychopathology: it enables us to look beyond static traits, and model processes that change over time. It allows us to look beyond group effects, and focus on person-specific effects that unfold within individuals over time.

As such, it is a promising approach to further explore phenomenology and mechanisms underlying psychopathology, and model change in response to treatment.

References

American Psychiatric Association (Ed.). (2013). Diagnostic and statistical manual of mental disorders: DSM-5 (5th ed). Washington, D.C: American Psychiatric Association.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., & Ljung, G.M. (2016). Time Series Analysis: Forecasting and Control. New Jersey: John Wiley & Sons, Inc.

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

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling Nonstationary Emotion Dynamics in Dyads using a Time-Varying Vector-Autoregressive Model. Multivariate Behavioral Research. https://doi.org/10.1080/00273171.2018.1439722

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. Psychological Methods, 22(3), 409–425. https://doi.org/10.1037/met0000085

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., … Kuppens, P. (2016). Assessing Temporal Emotion Dynamics Using Networks. Assessment, 23(4), 425–435. https://doi.org/10.1177/1073191116645909

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., … Tuerlinckx, F. (2013). A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. PLOS ONE, 8(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Bulteel, K.*, Mestdagh, M.*, Tuerlinckx, F., & Ceulemans, E. (in press). VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. Psychological Methods.

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Using Raw VAR Regression Coefficients to Build Networks can be Misleading. Multivariate Behavioral Research, 51(2–3), 330–344. https://doi.org/10.1080/00273171.2016.1150151

Bylsma, L. M., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional reactivity to daily events in major and minor depression. Journal of Abnormal Psychology, 120(1), 155–167. https://doi.org/10.1037/a0021662

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Grassmann, M., & Ceulemans, E. (2017). Detecting correlation changes in multivariate time series: A comparison of four non-parametric change point detection methods. Behavior Research Methods, 49(3), 988–1005. https://doi.org/10.3758/s13428-016-0754-9

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., & Ceulemans, E. (2018). Testing for the Presence of Correlation Changes in a Multivariate Time Series: A Permutation Based Approach. Scientific Reports, 8(1), 769. https://doi.org/10.1038/s41598-017-19067-2

Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem.

Ebner-Priemer, U. W., Houben, M., Santangelo, P., Kleindienst, N., Tuerlinckx, F., Oravecz, Z., … Kuppens, P. (2015). Unraveling affective dysregulation in borderline personality disorder: A theoretical model and empirical evidence. Journal of Abnormal Psychology, 124(1), 186–198. https://doi.org/10.1037/abn0000021

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

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2012). Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. Statistics in Medicine, 31(27), 3328–3336. https://doi.org/10.1002/sim.5338

Hepp, J., Carpenter, R. W., Lane, S. P., & Trull, T. J. (2016). Momentary symptoms of Borderline Personality Disorder as a product of trait personality and social context. Personality Disorders, 7(4), 384–393. https://doi.org/10.1037/per0000175

Heylen, J., Van Mechelen, I., Verduyn, P., & Ceulemans, E. (2016). KSC-N: Clustering of Hierarchical Time Profile Data. Psychometrika, 81(2), 411–433. https://doi.org/10.1007/s11336-014-9433-x

Heylen, J., Verduyn, P., Van Mechelen, I., & Ceulemans, E. (2015). Variability in anger intensity profiles: Structure and predictive basis. Cognition and Emotion, 29(1), 168–177. https://doi.org/10.1080/02699931.2014.896783

Hollenstein, T., & Lewis, M. D. (2006). A state space analysis of emotion and flexibility in parent-child interactions. Emotion (Washington, D.C.), 6(4), 656–662. https://doi.org/10.1037/1528-3542.6.4.656

Houben, M., Claes, L., Vansteelandt, K., Berens, A., Sleuwaegen, E., & Kuppens, P. (2017). The emotion regulation function of nonsuicidal self-injury: A momentary assessment study in inpatients with borderline personality disorder features. Journal of Abnormal Psychology, 126(1), 89.

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. Psychological Bulletin, 141(4), 901–930. https://doi.org/10.1037/a0038822

Houben, M., Vansteelandt, K., Claes, L., Sienaert, P., Berens, A., Sleuwaegen, E., & Kuppens, P. (2016). Emotional switching in borderline personality disorder: A daily life study. Personality Disorders: Theory, Research, and Treatment, 7(1), 50–60. https://doi.org/10.1037/per0000126

Hox, J. J., Moerbeek, M., & Schoot, R. van de. (2010). Multilevel Analysis: Techniques and Applications, Second Edition. Routledge.

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

Koval, P., Ogrinz, B., Kuppens, P., Bergh, O. V. den, Tuerlinckx, F., & Sütterlin, S. (2013). Affective Instability in Daily Life Is Predicted by Resting Heart Rate Variability. PLOS ONE, 8(11), e81536. https://doi.org/10.1371/journal.pone.0081536

Koval, P., Pe, M. L., Meers, K., & Kuppens, P. (2013). Affect dynamics in relation to depressive symptoms: Variable, unstable or inert? Emotion, 13(6), 1132–1141. https://doi.org/10.1037/a0033579

Krone, T., Albers, C. J., & Timmerman, M. E. (2016). Comparison of Estimation Procedures for Multilevel AR(1) Models. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.00486

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

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. Journal of Personality and Social Psychology, 99(6), 1042–1060. https://doi.org/10.1037/a0020962

Lane, S. T., & Gates, K. M. (2017). Automated Selection of Robust Individual-Level Structural Equation Models for Time Series Data. Structural Equation Modeling: A Multidisciplinary Journal, 0(0), 1–15. https://doi.org/10.1080/10705511.2017.1309978

Lane, S.T., Gates, K.M., Pike, H.K., Beltz, A.M., & Wright, A.G.C. (in press). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods.*

Linehan, M. (1993). Cognitive-behavioral Treatment of Borderline Personality Disorder. Guilford Press.

Liu, S. (2017). Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels. The British Journal of Mathematical and Statistical Psychology. https://doi.org/10.1111/bmsp.12096

Lougheed, J. P., & Hollenstein, T. (2016). Socioemotional flexibility in mother-daughter dyads: Riding the emotional rollercoaster across positive and negative contexts. Emotion (Washington, D.C.), 16(5), 620–633. https://doi.org/10.1037/emo0000155

Manolov, R., & Onghena, P. (2017). Analyzing Data From Single-Case Alternating Treatments Designs. Psychological Methods. https://doi.org/10.1037/met0000133

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. Psychological Medicine, 39(09), 1533. https://doi.org/10.1017/S0033291708004947

Nezlek, J. B. (2008). An Introduction to Multilevel Modeling for Social and Personality Psychology. Social and Personality Psychology Compass, 2(2), 842–860. https://doi.org/10.1111/j.1751-9004.2007.00059.x

Nezlek, J. B., & Plesko, R. M. (2001). Day-to-Day Relationships among Self-Concept Clarity, Self-Esteem, Daily Events, and Mood. Personality and Social Psychology Bulletin, 27(2), 201–211. https://doi.org/10.1177/0146167201272006

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

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

Peeters, F., Berkhof, J., Delespaul, P., Rottenberg, J., & Nicolson, N. A. (2006). Diurnal mood variation in major depressive disorder. Emotion, 6(3), 383–391. https://doi.org/10.1037/1528-3542.6.3.383

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement, 1(3), 385–401. https://doi.org/10.1177/014662167700100306

Rovine, M., and Walls, T. (2006). "A multilevel autoregressive model to describe interindividual differences in the stability of a process," in Models for Intensive Longitudinal Data, eds J. Schafer and T. Walls (New York, NY: Oxford), 124–147

Santangelo, P. S., Limberger, M. F., Stiglmayr, C., Houben, M., Coosemans, J., Verleysen, G., … Ebner-Priemer, U. W. (2016). Analyzing subcomponents of affective dysregulation in borderline personality disorder in comparison to other clinical groups using multiple e-diary datasets. Borderline Personality Disorder and Emotion Dysregulation, 3(1), 5. https://doi.org/10.1186/s40479-016-0039-z

Sbarra, D. A. (2006). Predicting the onset of emotional recovery following nonmarital relationship dissolution: survival analyses of sadness and anger. Personality & Social Psychology Bulletin, 32(3), 298–312. https://doi.org/10.1177/0146167205280913

Thompson, R. J., Mata, J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Gotlib, I. H. (2012). The everyday emotional experience of adults with major depressive disorder: Examining emotional instability, inertia, and reactivity. Journal of Abnormal Psychology, 121(4), 819–829. https://doi.org/10.1037/a0027978

Timmerman, M. E. (2006). Multilevel component analysis. British Journal of Mathematical and Statistical Psychology, 59(2), 301–320. https://doi.org/10.1348/000711005X67599

Timmerman, M. E., & Kiers, H. A. L. (2003). Four simultaneous component models for the analysis of multivariate time series from more than one subject to model intraindividual and interindividual differences. Psychometrika, 68(1), 105–121. https://doi.org/10.1007/bf02296656

Tomko, R. L., Lane, S. P., Pronove, L. M., Treloar, H. R., Brown, W. C., Solhan, M. B., … Trull, T. J. (2015). Undifferentiated negative affect and impulsivity in borderline personality and depressive disorders: A momentary perspective. Journal of Abnormal Psychology, 124(3), 740.

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. Annu Rev Clin Psychol, 9. https://doi.org/10.1146/annurev-clinpsy-050212-185510

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology,* 117. https://doi.org/10.1037/a0012532

Verduyn, P., Delvaux, E., Van Coillie, H., Tuerlinckx, F., & Van Mechelen, I. (2009). Predicting the duration of emotional experience: two experience sampling studies. Emotion (Washington, D.C.), 9(1), 83–91. https://doi.org/10.1037/a0014610

Wichers, M., Groot, P. C., & Psychosystems, ESM Group, EWS Group. (2016). Critical Slowing Down as a Personalized Early Warning Signal for Depression. Psychotherapy and Psychosomatics, 85(2), 114–116. https://doi.org/10.1159/000441458

Wright, A.G.C., Gates, K.M., Arizmendi, C., Lane, S.T., Woods, W.C., & Edershile, E.A. (in press). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment.*

Yang, J., & Leskovec, J. (2011). Patterns of temporal variation in online media. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 177–186). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1935863

Table 1.

Overview of some major dynamical properties of single or multiple symptoms (and their interrelations) and how they can be calculated or modeled.

Univariate approaches: examining one variable			
Feature		How to calculate/model	interpretation
Central tendency			
	descriptive	Mean or median per person.	Typical moment-to-moment intensity level of a symptom.
		Count or proportion of time point a symptom was present per person.	Frequency of a symptom.
	model	(Multilevel) linear regression model with (random) intercept.	Average symptom intensity across all time points.
		(Multilevel) Logistic regression model.	Likelihood of occurrence of a symptom.
Spread			
	descriptive	Within-person standard deviation or variance, within-person interquartile range.	Variability; degree of overall deviation or dispersion around typical intensity levels of a symptom.
	model	Error variance at the moment-to-moment level in (multilevel) regression models.	Degree of variability of a symptom within-person over time.
Dynamics			
Instability	descriptive	Mean square successive difference and related metrics per person	Average magnitude of changes in intensity of a symptom from one time point to the next.
	model	Squared or absolute successive difference modeled using a (random) intercept in (multilevel) regression models.	Average magnitude of changes in intensity of a symptom from one time point to the next.

Duration	descriptive	Count of (successive) time (points) a symptom is present.	Duration of a symptom.
	model	Survival analyses.	Expected duration until the termination of a symptom.
Autocorrelation	descriptive	Within-person autocorrelation.	Inertia, self-predictability of a symptom over time.
	model	(Multilevel) AR models.	Inertia, self-predictability of a symptom over time.
Time course	descriptive	Visual plots.	Visual inspection of overall time trends in a symptom per person.
	model	(Multilevel) regression models with time as a predictor.	Testing time effects in the intensity of symptoms over time.
		Growth curve models.	Modeling of growth or change of time-related symptoms.
		K-spectral centroid clustering method.	Match the time course of a symptom with reference intensity profiles, that differ in shape and amplitude.

Multivariate approaches: examining the relation between two (or more) variables

Feature		How to calculate/model	interpretation
Concurrent Relationships			
	descriptive	Within-person correlation.	Strength and direction of the association between two symptoms assessed at a moment-to-moment level for each person.
		Intra class correlation per person	consistency between a group of related variables across different time points for each person.
	model	(Multilevel) regression models with more 1 or more predictors at the moment-to-moment level.	Strength and direction of the association between two symptoms on a moment-to-moment level.

		PCA and (multilevel) simultaneous component analysis	Summarizes a set of symptoms by extracting one or more components that explain as much of the variance across the time points as possible.
Lagged Relationships			
	descriptive	Partial lagged correlation coefficient per person.	Strength and direction of the association between two symptoms measured at consecutive time points for each person.
	model	(Multilevel) VAR-models with one lagged predictor.	Strength and direction of the association between two symptoms measured at consecutive moments in time; symptom reactivity to context at the moment-to-moment level.
Advanced Techniques			
	descriptive	State-space analysis.	visual inspection of how two categorical variables co-evolve over time; descriptive indices regarding the dynamics of two categorical variables at the moment-to-moment level simultaneously.
	model	(Multilevel) VAR models with multiple predictors in a network.	Network models showing unique and direct effects between a set of symptoms.
		GIMME.	Approach to examine several lagged and contemporaneous relationships within-person over time.
		BOUM.	Estimation of three different dynamic parameters of two separate continuous variables, simultaneously: homebase, the variability around this homebase, and the attractor strength.

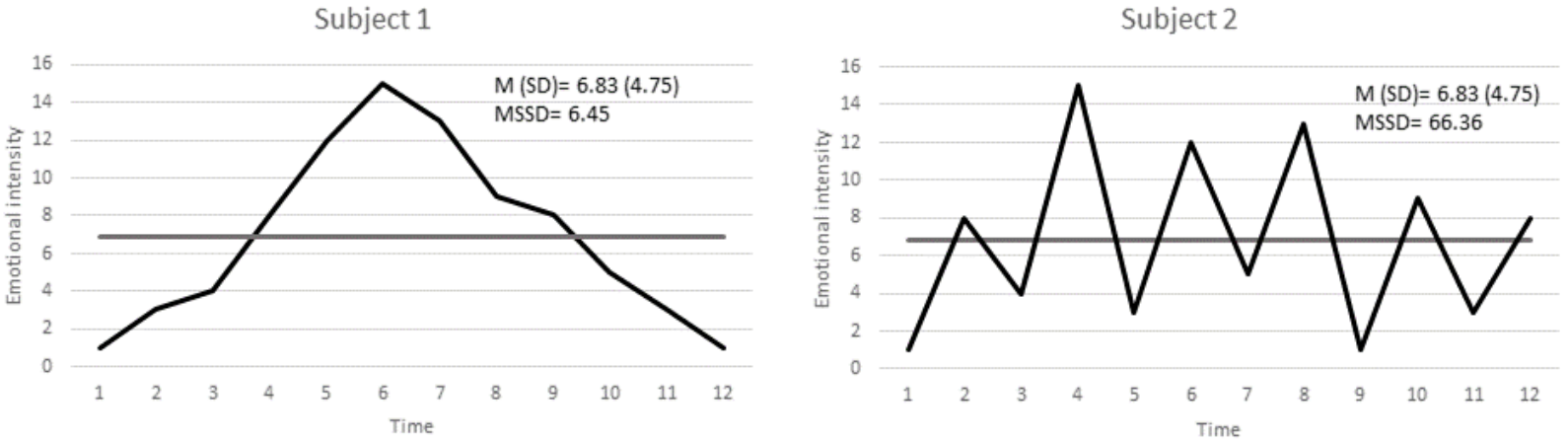
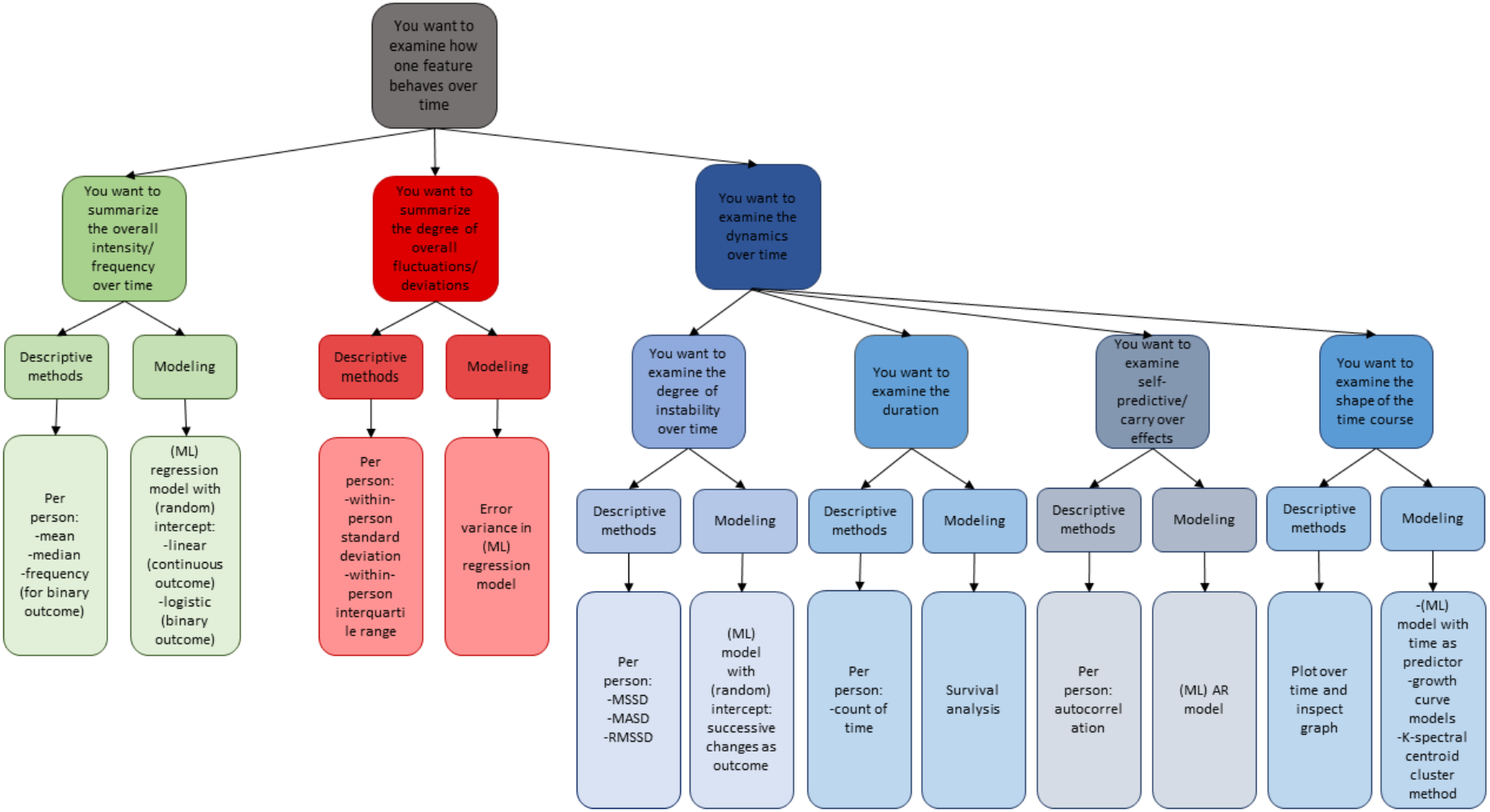


Figure 1. Graphs showing simulated time-series data for two fictive subjects. Both time series data strings consist of the same data points. Therefore, they have the same mean and standard deviation, showing the same degree of deviation or fluctuation from the average intensity level. However for subject 2, changes in intensity occur more frequent and more abrupt (reflecting larger instability), which is captured by a higher MSSD index. In contrast, in the time-series data of subject 1, changes are more gradual, resulting in a low MSSD index.


You want to examine how one feature behaves over time
You want to summarize the overall intensity/ frequency over time
You want to summarize the degree of overall fluctuations/ deviations
You want to examine the dynamics over time
Descriptive methods
Modeling
Descriptive methods
Modeling
You want to examine the degree of instability over time
You want to examine the duration
You want to examine self-predictive/ carry over effects
You want to examine the shape of the time course
Per person: -mean -median -frequency (for binary outcome)
(ML) regression model with (random) intercept: -linear (continuous outcome) -logistic (binary outcome)
Per person: -within-person standard deviation -within-person interquartile range
Error variance in (ML) regression model
Descriptive methods
Modeling
Descriptive methods
Modeling
Descriptive methods
Modeling
Descriptive methods
Modeling
Per person: -MSSD -MASD -RMSSD
(ML) model with (random) intercept: successive changes as outcome
Per person: -count of time
Survival analysis
Per person: autocorrelation
(ML) AR model
Plot over time and inspect graph
-(ML) model with time as predictor -growth curve models -K-spectral centroid cluster method

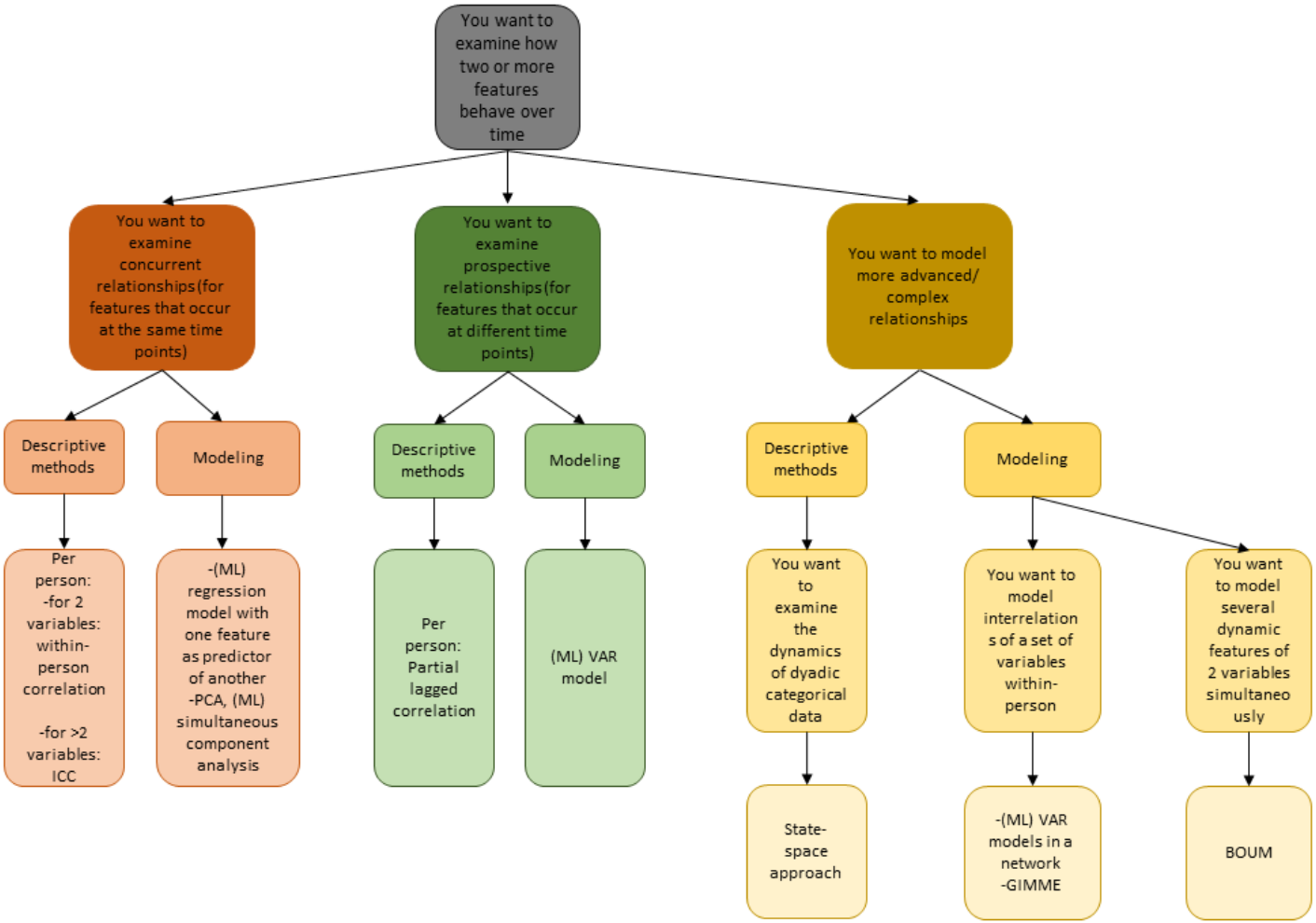


Figure 2. Diagram showing which models to use for several types of research questions, related to one variable (above) or multiple variables (below).