

Amsterdam Public Health



**Ecological Momentary Assessment
In Mental Health Research:**

**A practical introduction,
with examples in R**

Jeroen Ruwaard, Lisa Kooistra & Melissa Thong

Ecological Momentary Assessment in Mental Health Research

A Practical Introduction, with Examples in R (1st edition)

Jeroen Ruwaard

Lisa Kooistra

Melissa Thong

2018-11-16

Contents

Preface	vii
I Introduction	1
1 Introducing EMA	3
1.1 What is EMA?	3
1.2 Why EMA?	4
1.3 EMA Research Findings: A Birds-eye View	6
1.4 What is in this Manual?	8
2 Introducing R & RStudio	9
2.1 What are R and RStudio?	9
2.2 Why R?	9
2.3 Installing R & RStudio	10
2.4 Interacting with R through the RStudio Console	10
2.5 Writing R-scripts	11
2.6 Importing Your Data	11
2.7 Extending R with Packages	13
2.8 Getting Help	15
II EMA Methods	19
3 Study Design	21
3.1 What is the EMA Research Question?	21
3.2 Who are the Study Participants?	21
3.3 What are the Qualities of the EMA Measures?	22
3.4 What is the Sample Plan?	22
3.5 What is the Optimal Sample Size?	23
3.6 What are the Ethical Aspects?	25
4 Data Management	27
4.1 Using RStudio-projects	27
4.2 An Example Project Directory Structure	28
4.3 Data	29
4.4 Import Scripts	29
4.5 Cleaning	29
4.6 Pre-process	31
4.7 Combine	31
4.8 Reproducible Analyses	31
4.9 Documentation	32

4.10 Discussion	33
III EMA Outcomes	35
5 Mood	37
5.1 Unidimensional Mood Assessment	37
5.2 Bag-of-Items	38
5.3 Multi-dimensional Mood Assessment	39
6 Activity	43
6.1 Actigraphy	44
6.2 Geotracking	46
IV Analytic Approaches	49
7 Feature Extraction	51
7.1 Simulated EMA Time-series Example	51
7.2 Central Tendency and Variability	52
7.3 Modelling the Trend	53
7.4 Missing Values	54
7.5 Auto-correlation	55
7.6 Periodicity	56
7.7 Discussion	57
8 Mixed Modeling	61
8.1 The Mixed Model	61
8.2 Simulating Example Data	61
8.3 Fitting a Mixed Model in R	62
8.4 Adding Time as a Predictor	63
8.5 Adding a Two-Group Comparison	65
8.6 Next Steps	66
V A Case Study	67
9 Early Warning Signs of Depression	69
9.1 Critical Slowing Down	69
9.2 Plotting the Course of Depression	70
9.3 Mental state EMA Items	70
9.4 Running the DFA	71
9.5 Results	72
9.6 Discussion	73
VI EMA Catalogues	75
10 EMA Research within the APH Mental Health Consortium	77
10.1 Depression	78
10.2 Psychosomatic Symptoms	81
10.3 Suicidal Ideation	85
10.4 Psychotic symptoms	85
10.5 Sleep	87
10.6 Stress & Emotion	89

11 EMA Instruments Catalogue	91
11.1 EMA Platforms & Apps	91
11.2 Wearables	95
12 R Packages for EMA Research	97
12.1 Accelerometry	98
12.2 Autoregressive modeling	100
12.3 Data management & Visual Exploration	101
12.4 Mixed-effects Modeling	104
12.5 Power analysis	105
12.6 Symptom Network Analysis	106
12.7 Timeseries analysis	108
VII Closing Matters	111
Acknowledgements	113
Bibliography	115
Index	124

Preface

Given known limitations of retrospective self-report questionnaires, such as recall bias and poor generalizability of assessment results to real-life situations, mental health researchers increasingly adopt alternative assessment methods. One of the promising alternatives is Ecological Momentary Assessment (EMA), in which emotions and behaviors are repeatedly sampled in everyday life, through wearable electronic devices.

Repeated measurement can reveal important characteristics of the dynamics of phenomena, as illustrated by the cover of this book. With EMA, we can tap into mental health processes that were, up to very recently, unavailable to scientific research.

Conducting an EMA study, however, can be challenging. Researchers face a dazzling array of options related to the electronic wearables, outcomes selection, study design considerations, ethical and regulatory constraints, data management, statistical analysis, and study reporting. Although standards are emerging, clear guidelines for EMA research do not - at present - exist. EMA studies have unique characteristics that require specialist research skills, related to study design and statistical analysis. This manual was written to fill this gap.

The manual provides a practical introduction to EMA-research. It was written to aid beginning researchers of the Amsterdam School of Public Health (APH), who are looking for practical advice in conducting EMA studies. The manual provides an overview of EMA instruments, outcomes, methods and analytic techniques, guidelines for EMA-studies, and a catalogue of EMA research in the APH consortium.

This manual is available online at https://jruwaard.github.io/aph_ema_handbook/. Sources are available at https://github.com/jruwaard/aph_ema_handbook. Please post your comments and suggestions there, or via e-mail, through aph.ema@ggzingeest.nl.

This manual will be continuously updated. In citations, please include explicit build dates, as in:

Ruwaard, J., Kooistra, L. and Thong, M. (2018). *Ecological Momentary Assessment in Mental Health Research: A Practical Introduction, with Examples in R* (1st edition - build 2018-11-16). Amsterdam: APH Mental Health.

Part I

Introduction

Chapter 1

Introducing EMA

1.1 What is EMA?

Ecological Momentary Assessment (EMA) has many aliases. It is known as ‘Experience Sampling’ (Larson and Csikszentmihalyi, 1983), ‘Ambulatory Assessment’ (Ebner-Priemer and Trull, 2009) or ‘Ambulatory Self-reporting’ (Conner and Feldman Barrett, 2012), ‘Real-time Data Capturing’, the ‘Continuous Unified Electronic Diary Method’ (Ellis-Davies et al., 2012), and as the ‘Intensive-longitudinal Study Design’ (Bolger and Laurenceau, 2013). The different terms stress different aspects of EMA research. All, however, refer to research methods that involve the *repeated sampling of people’s current thoughts, emotions, behavior, physiological states, and context, in their natural environment, typically (but not necessarily) via electronic wearable devices* (Shiffman et al., 2008).

1.1.1 Active versus Passive EMA

In EMA research, we distinguish two forms of data collection: 1) ‘Active EMA’, with which *self-report* data are collected, and 2) ‘Passive EMA’, with which *observational* data are collected. Active EMA requires participants to consciously provide information, for example by rating their current mood in response to a question that is prompted on their smartphone (see Figure 1.1). In passive EMA, information is collected through wearables or log files without active involvement of participants, for example on heart-rate, activity, smartphone use or engagement on social media (see Figure 1.2). Studies may combine active and passive EMA. A study into sleep patterns, for example, may involve both a self-report sleep diary and an accelerometer sensor (Van der Meijden et al., 2016).

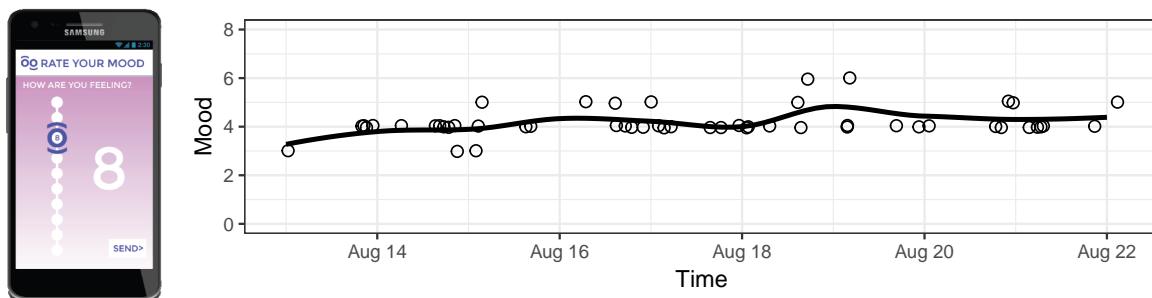


Figure 1.1: Active EMA: data are collected by prompting questions to participants, for instance by using an EMA app such as Moodbuster.



Figure 1.2: Passive EMA: data are collected automatically, for instance by a wearable device such as the GENEActiv accelerometer.

1.1.2 EMA Sampling

EMA is further typified by the way in which data sampling is triggered. EMA sampling may be triggered by a signal (signal-contingent sampling), an event (event-contingent sampling), or a combination of both (Conner and Lehman, 2012).

In *signal-contingent sampling*, participants respond to questions when they are prompted to do so by a signal (or ‘beeps’, as these signals are often called in the literature). In *event-contingent sampling*, study participants complete an assessment whenever a specific event occurs, such as a panic attack or alcohol consumption.

Signal-contingent sampling can follow a fixed or a random scheme. In a fixed scheme, participants are prompted at fixed time-points, for example at 9:30, 12:30, and 16:30. In a random scheme, prompts are sent at random time points, typically in pre-set intervals. For example, participants could be prompted to complete two assessments per day, one at a random time point between 10:00 and 14:00, and one at a random time point between 14:00 and 16:00. Using pre-set intervals ensures that participants do not receive several prompts within a limited time-frame (Piasecki et al., 2007). In addition, it ensures that participants are not bothered by prompts at inappropriate times (e.g., participants may not appreciate prompts after 22:00 and before 7:30).

With ‘event-contingent sampling’, the sample rate is determined by the occurrence rate of the event. One way to implement this is to simply ask study participants to complete a questionnaire whenever the event occurs. When active EMA is combined with passive EMA, it may also be possible to trigger event-based prompts automatically based on changes in passive data, for example by triggering an EMA questionnaire automatically whenever a significant change in activity level is detected (Smyth and Stone, 2003).

1.2 Why EMA?

1.2.1 To Minimize Recall Bias

In clinical research, self-report questionnaires are often used to assess the presence and severity of symptoms in the recent past. While useful, these retrospective self-reports are not without drawbacks, since they tap into the memory of respondents, which can be distorted (Shiffman et al., 2008; Moore et al., 2016). EMA circumvents this recall bias, by asking participants to rate their *current state*, rather than asking them to reflect on past experiences.

1.2.2 To Maximize Ecological Validity

A key feature of EMA is the collection of data in real-world environments, as participants go about their daily activities, as opposed to data collection in controlled labs or research settings (Shiffman et al., 2008). Thus, EMA data can be expected to result in research findings that have better ecological validity and better generalization to the subject's lived experience in real-world settings. Practical applications derived from EMA data are therefore expected to be more relevant to real-life situations.

1.2.3 To Advance Ideographic Research

In clinical research, a distinction is often made between *ideographic* and *nomothetic* methods (Allport, 1937). Idiographic methods are those that “aim to identify patterns of behavior within the person across a population of experiences or situations, and nomothetic methods[are] those that aim to identify patterns of behavior across a population of individuals, rather than for any given individual” (Conner et al., 2009). The difference is important. As is increasingly recognized, group-level findings do not necessarily generalize to the individual members of the group, as shown by Figure 1.3 (Hamaker, 2012).

In contrast to more qualitative idiographic methods, such as interviews and $N = 1$ case studies, EMA allows for the collection of large amounts of quantitative data on the individual level. Thus, EMA offers a quantitative method for idiographic research, measuring characteristics of (unique) individuals across time and context (Shiffman et al., 2008). This allows for a better understanding of the factors that account for the variability within and between individuals.

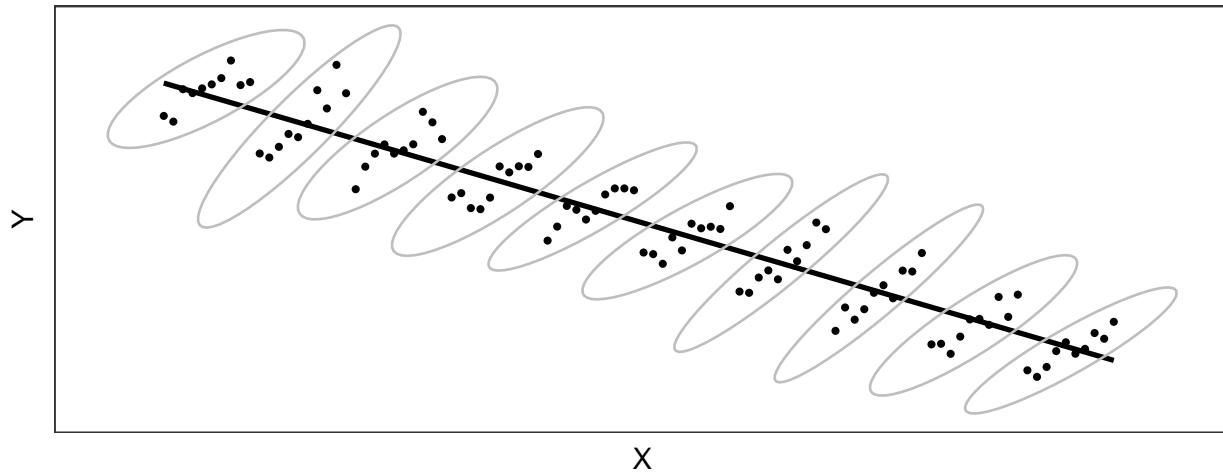


Figure 1.3: An illustration of how group-level and individual-level processes can differ dramatically: the relationship between x on y is negative in the group (as shown by the descending regression line), but positive for individuals (marked by ellipses).

1.2.4 To Understand the Dynamic Interplay between Symptoms

In the Network Theory of Psychopathology (Borsboom and Cramer, 2013; Borsboom, 2017), mental health disorders are conceptualized as networks of psychopathology symptoms, in which recurrent causal loops keep the network in a “disorder” state (e.g., sleeping problem \rightarrow fatigue \rightarrow rumination \rightarrow sleeping problem). Network Theory encourages the identification of patient-specific symptom networks, so that central symptoms can be targeted with personalized interventions, to break the self-sustaining loops. The identification of these networks requires repeated assessments of symptoms in real life (see, e.g., Bringmann et al., 2015), a task for which EMA is particularly well suited.

1.2.5 To Enable EMI

EMA enables Ecological Momentary Interventions (EMI): interventions that are provided to people during their everyday lives, in real time, and in their natural settings (Heron and Smyth, 2010). If dynamic disease processes can be adequately monitored in everyday-life through EMA, it should also be possible to intervene when EMA data reflect clear changes in these processes, in a way that is maximally effective given what is known about the individual. A key term in the last sentence, however, is “adequately”. We should be aware of the “garbage-in/garbage-out” principle. Before EMI’s can be considered, the psychometric properties of EMA measures should be demonstrated first.

1.3 EMA Research Findings: A Birds-eye View

The use of EMA in mental health research might seem novel but this methodology has a long track record. Already in the 1980’s, early pioneers used electronic devices to elicit responses from study participants to tap into (mental) health processes in everyday life (see, e.g., Csikszentmihalyi and Larson, 2014). Recent years, however, have witnessed a large increase in EMA research. Rapid technological developments, a marked interest in the individual, and a wide recognition of the need to study health-related processes in real-life situations have all contributed to this.

EMA-based research in mental health has produced an impressive trove of findings that have supported and, sometimes, challenged existing theories on behavior. EMA data, whether collected as self-report or via wearable device/sensor, have diagnostic, monitoring, management, or intervention applications (Patel et al., 2012; Aung et al., 2017; Evenson et al., 2015). Its feasibility for mental health research is evidenced by its use in observational studies and randomized controlled trials on a wide range of topics and populations. Below, we present a non-exhaustive summary of systematic reviews and meta-analyses of EMA research.

1.3.1 Active EMA

Mood disorders have been well-studied using active EMA methods (Wenze and Miller, 2010), with several reviews outlining recent findings in depression (Telford et al., 2012; Wichers et al., 2011; Burton et al., 2013), anxiety disorders (Walz et al., 2014), and depression/bipolar disorder (Aan het Rot et al., 2012). The potential of EMA for use among young populations showed promising results (Dubad et al., 2018). Innovations in mobile devices has improved the feasibility and popularity of ecological momentary interventions (EMIs) for anxiety and depression (Schueller et al., 2017). A systematic review and meta-analysis of EMIs reported small to medium effects on mental health (Versluis et al., 2016).

EMA appears to be particularly well-suited to examine the role of emotions in the development and maintenance of obesity and eating disorders (Engel et al., 2016). Meta-analytic results suggest that negative affect, rather than hunger, is associated with binge eating among individuals with eating disorders (Haedt-Matt and Keel, 2011; Haedt-Matt et al., 2012). EMA methods have also significantly contributed to the understanding of the processes that drive substance use, cessation, and relapse, often in contrast with theory-driven studies largely derived from global reports collected through retrospective questionnaires (Shiffman, 2009; Swendsen, 2016).

1.3.2 Passive EMA

Objective EMA data collected passively through bio-sensors, smart devices, or context/environmental (e.g. location) has been shown to be a feasible and promising method for the longitudinal monitoring of individuals with affective disorders (Dogan et al., 2017; Kirchner and Shiffman, 2016). The potential of passive sensing via smartphone for mental health research are outlined in two reviews, encompassing the assessment of health and well-being (Cornet and Holden, 2018), and the measuring, understanding, and intervening in mental illness and maintaining mental health (Aung et al., 2017). A systematic review and

meta-analysis on actigraphy reported diurnal variations in activity levels among individuals with depression (Burton et al., 2013). Compared with traditional self-reports, passive sensing was reported to be less intrusive and to result in more accurate data, while providing options for continuous monitoring and feedback.

1.3.3 Limitations

While available reviews suggest that EMA is a feasible research method that has the potential to make significant contributions to mental health research, one significant drawback of EMA highlighted in several reviews is the lack of high-quality studies. Other general limitations considered include:

- Unclear generalizability of research findings due to selective samples (Burton et al., 2013), and small sample sizes (Dogan et al., 2017)
- Unclear practice effects (Telford et al., 2012), and issues related to reactivity (Shiffman, 2009)
- Issues related to the feasibility and tolerability of prolonged and intense data periods of data collection (Wichers et al., 2011)
- Issues of privacy, ethics and informed consent (Cornet and Holden, 2018)
- Changing responsibilities of researchers and clinicians in EMA studies, e.g. in suicide ideation research (Wenze and Miller, 2010)
- Unresolved methodological issues (Dubad et al., 2018)

Table 1.1: A selection of reviews of EMA studies targeting various mental health conditions

Topic / Author (Year)	Summary
Anxiety disorders	
Walz et al. (2014)	Provides insights to the temporal variability of symptoms, and associations between daily affect, behaviors, and situational cues. Discusses the combination of EMA and ambulatory assessment of physiological variables in treatment evaluations.
Schueller et al. (2017)	Provides an overview of the distinction of EMIs from other types of treatment. Also discusses the considerations of conducting EMI research, such as design, deployment, and evaluation.
Eating disorders	
Engel et al. (2016)	An overview of studies on eating disorders, obesity, and bariatric surgery using EMA.
Mood disorders	
Aan het Rot et al. (2012)	Provides an overview of EMA studies on correlates of mood, treatment effects, residual symptoms of remitted patients, pediatric populations, MDD/BD specificity, and links with neuroscience,
Aung et al. (2017)	Provides a conceptual review of passive sensing techniques for measuring, understanding, and treatment of mental illness.
Burton et al. (2013)	Focuses on diurnal variations in activity levels among depressed individuals.
Telford et al. (2012)	Identifies six themes of EMA research in MDD: methodology, positive and negative affect, cortisol secretion, antidepressant treatment, work performance, and genetic risk factors.
Wenze and Miller (2010)	Provides an overview of EMA in mood disorder research comprising techniques used, types of population assessed, types of research questions, and a discussion of the potential of EMA in treatment settings.

Topic / Author (Year)	Summary
Wichers et al. (2011)	Provides an overview for the potential clinical application of EMA in the diagnostic and treatment of MDD.
Versluis et al. (2016)	Provides an overview of interventions (EMI) addressing anxiety, depression, and perceived stress on positive psychological outcomes.
Dubad et al. (2018)	Provides an overview of the feasibility and clinical impact of mood-monitoring applications targeting young populations (10-24 years old).
Substance-related disorders	
Shiffman (2009)	Review of processes that drive substance use, cessation, and relapse, sometimes in contrast with theory-driven studies that are largely derived from global reports collected through questionnaires.
Swendsen (2016)	Conceptual review of the use of mobile technologies for research on addiction and its treatment.
Mental health/ Well-being	
Cornet and Holden (2018)	Outlines the potential and challenges of passive sensing to detect status change and behavior change following feedback on behavior.
Kirchner and Shiffman (2016)	Provides an overview of ‘geographically explicit momentary assessment’ (GEMA) research to enrich EMA research in mental health and well-being.
Dogan et al. (2017)	Provides an overview of studies that combine subjective ratings with objective EMA-data collected using smartphone-based systems.

1.4 What is in this Manual?

This manual comprises six parts, in which various aspects of EMA research are discussed.

- Chapter 2, the second chapter of **Part I**, introduces R, the statistical program that is an indispensable tool for the EMA researcher.
- **Part II** focuses on EMA study design (Chapter 3), and EMA data management (Chapter 4).
- **Part III** details the momentary assessment of two specific outcomes: Mood (Chapter 5) and Activity (Chapter 6).
- **Part IV** introduces EMA data analysis techniques: Feature Extraction (Chapter 7), and Mixed Modelling (Chapter 8).
- In **part V**, the application of the preceding material is illustrated in a case study, in which EMA is used to detect early warning signs of depression (Chapter 9).
- **Part VI** provides three catalogues of EMA resources. Chapter 10 lists EMA research groups within APH, chapter 11 lists EMA instruments that were found to be in use among APH researchers, and Chapter 12 summarizes useful R extensions (packages) for EMA data analysis.

Chapter 2

Introducing R & RStudio

In this chapter, you will learn how to install and use two programs that are indispensable for the management and analysis of EMA data: R and RStudio.

2.1 What are R and RStudio?

R is a programming language and software environment for statistical computing and data visualization. RStudio is a powerful user interface to R. It has many useful features that greatly simplify R-work. We strongly advise you to adopt the R/RStudio-combo.

2.2 Why R?

R, you may have been told, is for data scientists, methodologists, and scientific programmers only. It has a steep learning curve. If you are trained in SPSS, it will take time to become as productive in R as in SPSS. Why then, should you invest in R?

- Unlike SPSS, R is free of charge. It does not eat up your budget. Why pay for something that you can get for free?
- R is cutting-edge. Methodological innovations first appear in R. Network analyses, for example, can be run in R, but not (yet) in SPSS. For some analyses, you need this alternative.
- Mastering R improves your connection to the statisticians in your team. They probably prefer R over SPSS. It is more efficient and less error-prone to all speak the same language.
- R is great for data-management. Clinical research, and especially EMA research, requires hundreds of operations on multiple raw data files. R excels at that. SPSS, frankly, does not. If you care about reproducible research (which you should), R can be a great help in putting it into practice.
- R can be used at different levels. If you want to be a basic user, that's fine. However, if you want to dive deeper, you will find that you can easily do so. You can study source code to understand a particular technique better. You can code new functions. R allows you to grow.
- R's user base is expanding every year. Chances are high that R will be the standard in your next workplace. R will look great on your CV.

You don't have to be a programmer or methodologist to use R. Yes, it takes time to unlock its full potential, but you should be able to run basic analyses in it within a week. This chapter will get you started.

2.3 Installing R & RStudio

Both R and RStudio are available, at no costs, for all major operating systems.

- Download R from the Comprehensive R Archive Network (CRAN), at <https://cran.r-project.org/bin/>
- Download RStudio from <http://rstudio.org>

Install R first, and RStudio second. If you install the programs in this order, RStudio will automatically find R on your computer.

If you installed R or RStudio previously, please update. This manual assumes you will be working with version 3.4.2 (or higher) of R, and version 1.1.414 (or higher) of RStudio.

2.4 Interacting with R through the RStudio Console

If you open RStudio, you will be presented with the interface shown in Figure 2.1. RStudio's main window is divided in four panes (sub-windows), which further contain several tabbed windows.

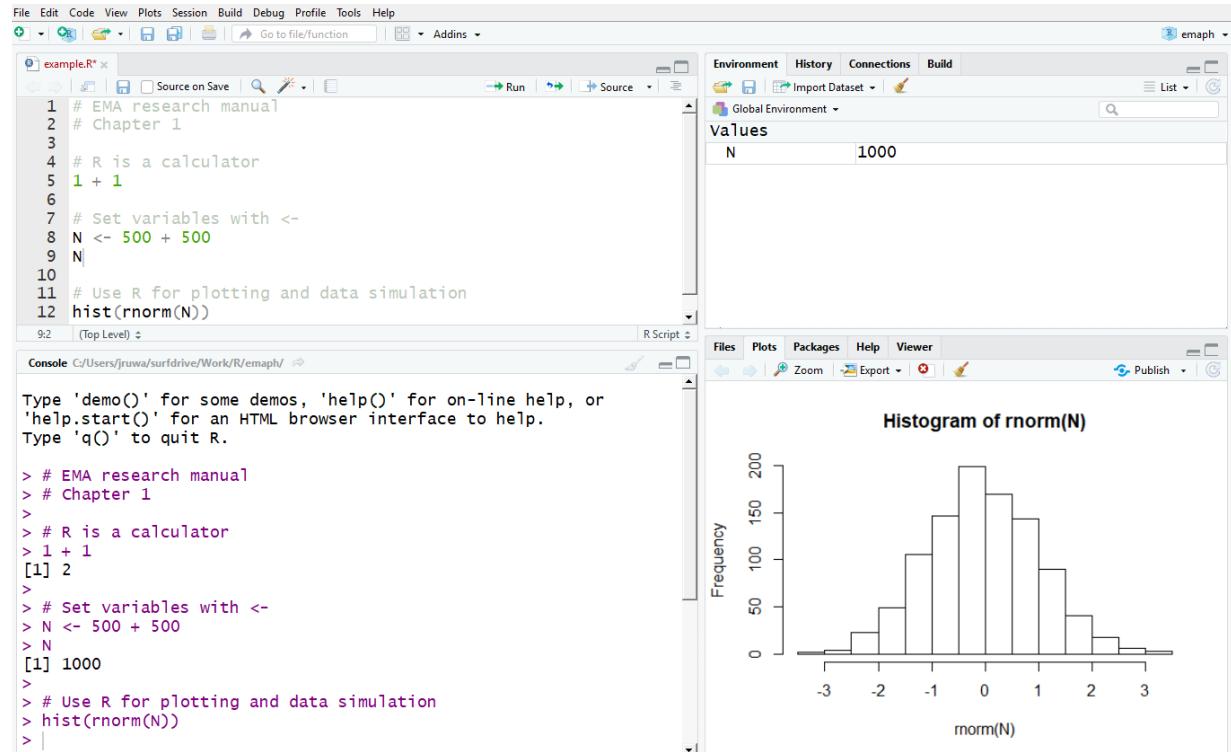


Figure 2.1: The RStudio Interface

Commands are sent to R in the bottom-left pane, named “Console”. To test this, move your cursor to the bottom line, immediately after the prompt sign (>). Next, type the statement below (note that # denotes a comment line; R ignores it, so there is no immediate need to type that). To execute, press **Enter**.

R will execute the command and return the answer back to the console.

```
# R is a calculator.
1 + 1
```

Results of calculations can be saved into variables, by making use of the assignment operator (`<-`). If you type the name of a variable, R returns its value.

```
# Use <- to declare and set a variable.
N <- 50 + 50
N
#> [1] 100
```

To understand why R is such a popular tool for statistical computing, consider the following command, which, in one line, 1) uses the variable `N`, just created, to 2) generate 100 random numbers from the normal distribution, and 3) plot a histogram of these numbers.

```
# Plot the histogram of a sample from the normal distribution.
hist(rnorm(N))
```

The plot appears in the bottom-right pane, as in Figure 2.1.

2.5 Writing R-scripts

Working in the console is a great way to interactively explore R and data, but what if you want to save a particularly useful chain of statements? For this, you can use a script file.

To create a script file, use the RStudio menu: `File > New File > R Script`. This will open a new tab in the top-left pane of RStudio, where you can edit the script.

- In the script window, type all statements that you have been entering in the console in the previous section.
- Next, select all lines in the script.
- Press `Ctrl+Enter` to run the script.

All commands in the script are executed. The commands are echoed in the console pane, and results are shown immediately, as was the case before, when you typed the commands in the console yourself.

Scripts can also be run line by line. Move the cursor to the line you want to run, and press `Ctrl+Enter`. The line is copied to the console and executed, and the cursor in the script will move to the next line, allowing you to walk through the script, step by step.

2.6 Importing Your Data

Something that confuses new RStudio users, who are more familiar with SPSS, is that it is not obvious how to import data into RStudio. In SPSS, the data are in plain sight. In R, you first have to import the data.

2.6.1 Using RStudio Menus to Import Data

One way to load data into R is to use RStudio's data import wizard. Follow the steps below to see how this works with data stored in a comma-separated-values (csv) format, a common data format to which many programs, including SPSS and Excel, can export data to.

- Download the example csv data file at <https://tinyurl.com/yczmjdat> (or create a csv-version of one of your own data files).
- In RStudio's menu, choose `File > Import Dataset > From Text (base)`.

- In the window that appears, click on **Browse** to locate the csv- file on your computer, and click **Import** in the next window (see Figure 2.2).

RStudio shows the data, in tabular view, in the top-left window, ready for analysis. You will also find a new entry in the **Environment**-tab in the top-right pane. When you click the small arrow, at the left of the name, you will see a brief summary of the contents of the data.



Figure 2.2: RStudio's CSV import wizard.

2.6.2 Using Functions to Import Data

While RStudio's Data import wizard is useful, you will probably use it less over time. Most likely, you will convert to using the more efficient R commands to import data. For example, it takes only a single line to download and import the example data.

```
# Import csv-data, from the internet.
ESMdata <- read.csv(url("https://tinyurl.com/yczmjdat"), row.names = NULL)
```

2.6.3 Accessing your Data

Since the data is now in the environment (under the name **ESMdata**), you can use it in other R commands. For example, to produce a more detailed summary of the first four columns of **ESMdata**, you type:

```
# Summarize data.
summary(ESMdata)

  dayno          beepno      mood_relaxed      mood_down
Min.   : 1.0   Min.   : 1.00   Min.   :1.000   Min.   :-3.0000
1st Qu.: 61.0  1st Qu.: 3.00   1st Qu.:4.000   1st Qu.: 0.0000
Median :252.0  Median : 5.00   Median :4.000   Median : 0.0000
Mean   :198.9  Mean   : 5.24   Mean   :4.173   Mean   : 0.1784
3rd Qu.:303.0  3rd Qu.: 8.00   3rd Qu.:5.000   3rd Qu.: 0.0000
Max.   :366.0  Max.   :10.00   Max.   :7.000   Max.   : 3.0000
                           NA's    :2

  mood_irritat
Min.   :1.000
1st Qu.:1.000
Median :2.000
```

```
Mean    : 2.241
3rd Qu.: 3.000
Max.    : 7.000
NA's    : 3
```

To inspect the first 6 lines of data, type:

```
# Show first 6 lines of a data frame.
head(ESMdata)

#>   dayno beepno mood_relaxed mood_down mood_irritat
#> 1    226     1          5        -1          1
#> 2    227     5          4         0          3
#> 3    227     6          4         0          2
#> 4    227     8          4         0          1
#> 5    227     9          4         0          2
#> 6    227    10          5         0          1
```

To view all rows of data in a spreadsheet (as in Figure 2.2), type:

```
# Show data as spreadsheet.
View(ESMdata)
```

To work with a specific variable in the data set, use \$, for instance, to print the first 20 numbers in the mood_relaxed variable, type:

```
# Access a single variable in a data frame.
head(ESMdata$mood_relaxed, n = 20)
```

This allows you to apply functions to specific variables. For example, to calculate the mean of scores in mood_relaxed, type:

```
# Calculate the mean of a variable.
mean(ESMdata$mood_relaxed)
#> [1] 4.173442
```

There are many ways in which you can summarize and manipulate your data. At this point, the important milestone is that you have imported and accessed data in R.

2.7 Extending R with Packages

R's attractiveness lies in the ease with which it can be extended with new functionality. Through so-called packages, which can be freely downloaded from the internet, specialized functions can be added to your work-space.

2.7.1 Installing R-packages from CRAN

Packages can be found at the CRAN website. To browse through the impressive list of available packages, see https://cran.r-project.org/web/packages/available_packages_by_name.html

If you find a package you like, you can install it via the RStudio menu system, choosing Tools > packages. But you can also use the console, via the `install.package` function.

A popular package, `tidyverse`, is used extensively in the examples of this manual. This package comprises a set of popular packages from the creators of RStudio, that greatly simplify working with R. So, while you are at it, install this package now.

```
# Install a package from CRAN.
install.package(tidyverse)
```

The `tidyverse` contains a package called `haven`, which allows you to read and write SPSS data files (.sav files). This is very convenient. You don't have to convert all your SPSS data to csv files. See `?read_spss` to learn how to import an SPSS-file (or use the data import wizard, by choosing **File > Import Dataset > From SPSS**, in RStudio's top-right pane).

2.7.2 Installing R-packages from GitHub

Not all packages are at CRAN. Many ‘unofficial’ packages are shared at a site called ‘GitHub’. This book’s companion R package `emaph`, for example, which contains specialized EMA functions data sets, is on GitHub. You need package `emaph` to run many examples in the book, so let’s install this package now.

GitHub packages can be installed via the `install_github` function, which is defined in a package called ‘`devtools`’. So, to install `emaph`, enter the following in the console:

```
# Install the GitHub 'emaph' package.
install.packages("devtools")
devtools::install_github("jruwaard/emaph")
```

2.7.3 Using Packages

To use packages, you have to tell R to load them, each session you want to work with them. You do this with the `library` function. For example, to use package `tidyverse` and `emaph`, type:

```
# Load packages.
library(tidyverse)
library(emaph)
```

Once loaded, you can use the functions and data sets of the packages. Package `emaph` provides data set `csd`, which contains the data from the ‘critical slowing down’-study (Kossakowski et al., 2017; Wichers et al., 2016), in which a patient recorded his mood, for 239 days (see also Chapter 9).

To plot the irritation levels of this patient in the first six days, using the `ggplot` function from package `ggplot2` (which is in `tidyverse`), type:

```
# Using ggplot to plot EMA time series.
ggplot(data = subset(csd, dayno <= 6),
       mapping = aes(x = beepno, y = mood_irritat)) +
  geom_point() + geom_step() +
  ylab("Irritation level") +
  scale_x_continuous(breaks = 1:10) +
  facet_wrap(~ dayno, nrow = 2)
```

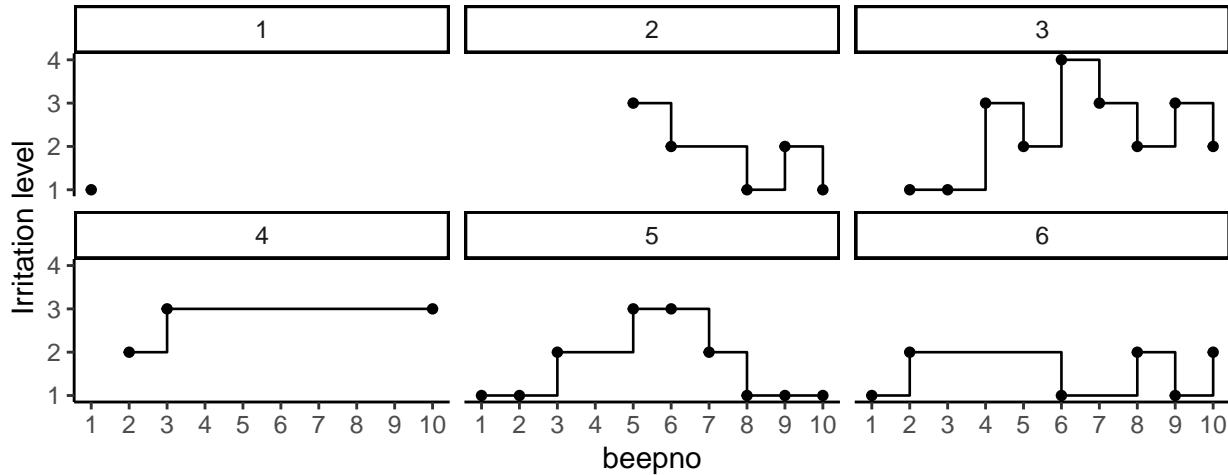


Figure 2.3: Irritation levels of a single patient, in the first six days of an EMA study. Missing values were most prominent at day 1, and irritation varied most at day 3.

2.8 Getting Help

R has no point-and-click menu's that you can browse through to select a statistical procedure. This is a problem for many new users. What if you want, for example, to generate random numbers from a distribution with a mean of 2 and standard deviation of 4? How to tell this to R?

2.8.1 Using ‘?’ to Consult the Documentation

The good thing is that you already know the name of the function to use, since we used it in the previous section: it is `rnorm`. To check the documentation of this function, type `?rnorm` in the console.

```
# Use '?' to find the documentation of a function.
?rnorm
```

This opens the documentation of the `rnorm` function in the Help-tab, in the bottom right pane, from which you learn that the `rnorm` function accepts `mean` and `sd` (standard deviation) as additional parameters, which are 0 and 1 default, respectively (which explains why `rnorm(100)` worked in the previous examples). So, to generate the required numbers, you type:

```
# Plot the histogram of a custom random sample.
hist(rnorm(1000, mean = 2, sd = 4))
```

All functions in R are documented, and this documentation is shown in RStudio's Help pane when you prepend `?` to the name of the function in the console.

2.8.2 Using RStudio's Global Documentation Index Search

What if you do not know the name of a function? Suppose you want to run a t-test for independent groups. Does R have a function for that?

At the top-right of the Help pane, RStudio has a search input field, which allows you to search through all documentation that is installed on your computer. The search field auto-completes your input. If you type a 't' in this field, you will be presented with a list of functions starting with a 't'. In this list, you find a likely

candidate: a function called `t.test`. From the documentation of this function (`?t.test`), you learn that, indeed, this is the function you were looking for.

```
# Run a t-test, on two simulated samples.

# generate two samples (N = 100 per group) from the normal distribution
A <- rnorm(100); B <- rnorm(100)

# the t-test should be non-significant
t.test(A, B)
#>
#> Welch Two Sample t-test
#>
#> data: A and B
#> t = -0.63201, df = 197.18, p-value = 0.5281
#> alternative hypothesis: true difference in means is not equal to 0
#> 95 percent confidence interval:
#> -0.3808144 0.1959674
#> sample estimates:
#> mean of x mean of y
#> -0.01695238 0.07547113
```

2.8.3 Learning from Examples

This book contains many R code snippets. By studying these examples, you will become more familiar with R.

Some examples will introduce R language constructs and functions that are unknown to you. Learn from these examples, by using `?` on each element that you do not understand.

2.8.4 Google

With Google, you will find many answers to your R questions. Googling for “t-test R”, for example, results in a rich set of online resources. Good resources are:

- RSeek (see <http://rseek.org/>)
- Stackoverflow: (see <https://stackoverflow.com/questions/tagged/r>)
- SearchR (see: <http://search.r-project.org/>)

2.8.5 Read Books

This book does not provide a comprehensive tutorial. There is no need for that, since excellent resources are readily available. A selection is presented below.

- Many mental health researchers own a copy of Andy Field’s popular book “Discovering Statistics Using IBM SPSS Statistics” (Field, 2013). For those, Field’s R-version of this book, “Discovering Statistics Using R” (Field et al., 2012) provides a familiar companion in making the transition to R. See <https://www.discoveringstatistics.com/>
- Free manuals can be found at the official CRAN site. The manuals are dry, but complete and authoritative, since the authors are members of the R core development team. See <https://cran.r-project.org/manuals.html> (or type `help.start()` in the console).

- While at CRAN, be sure to browse the ‘contributed documentation’-section. On this page, you will find many freely available manuals contributed by the R community. See <https://cran.r-project.org/other-docs.html>

2.8.6 Online Courses

- DataCamp, an online data science education platform, offers several interactive courses in R. See <http://www.datacamp.com>
- The Try-R course at the CodeSchool website provides an alternative to DataCamp. See: <http://tryr.codeschool.com/>
- The Quick-R website provides a concise introduction to R. See <https://www.statmethods.net/>

2.8.7 Learn R, in R

Package `swirl` contains a set of interactive courses that teach many aspects of the R language. See <http://swirlstats.com>

```
# Start the interactive swirl-course.  
install.packages("swirl")  
library("swirl")  
swirl()
```


Part II

EMA Methods

Chapter 3

Study Design

As with all scientific research, EMA studies start with mindful consideration of the study design. Issues that need to be considered are the research question(s), the hypotheses, the population of interest, and the nature of the comparison groups (Shiffman et al., 2008).

Ample information on general study design issues can be found elsewhere (see for example, the APH quality handbook, at <http://www.emgo.nl/kc/>). This chapter highlights key design aspects of EMA studies.

3.1 What is the EMA Research Question?

Given the plethora of new research options that emerged from the rapid development in EMA technologies, it can be tempting to dive straight into exploratory data collection, without giving much consideration to the theoretical background of the study. That, however, would be one pitfall of EMA research to avoid. Data mining is no substitute for theory. Asking participants to contribute data without a rationale is unethical. As in all scientific activities, defining the research question should be the first step.

Ask yourself what EMA could bring to your topic of interest. How is it different from traditional assessment methods? What questions does it allow you to address that you could not answer without it? For this, you could use any of the EMA advantages discussed in Chapter 1. Are you interested in real-life behavior, in differences between participants, in changes within participants over time, or in potential causal pathways between health-related variables? What relationships do you expect to find, and why? A solid theoretical background, and clearly formulated explicit research questions and hypotheses will help to make the right choices when you have to decide on the other aspects of the study design.

3.2 Who are the Study Participants?

Given the experimental nature of EMA, studies are often piloted in healthy or sub-clinical populations. This is a recommended first step, to test the experimental procedures and to avoid unnecessary burdening of vulnerable patient populations. You should be aware, though, that results obtained in non-patient populations do not necessarily generalize to patient populations. EMA mood ratings, for example, might be much more variable in patients compared to non-patients. Pilot studies should therefore also be conducted in the target population. It is also advisable to write a manual on how to operate the EMA device and spend time on briefing participants on what is expected of them during the study. Depending on your study topic and EMA method, de-briefing might also be necessary, along with instructions on how to return a wearable or de-install an EMA app.

3.3 What are the Qualities of the EMA Measures?

With the study hypotheses in place, theoretical constructs must be operationalized into quantitative measures. For this, you should consult the scientific literature on the reliability and validity of existing EMA measures (e.g., Moore et al., 2016; Van Rijssbergen et al., 2012).

In active EMA research, complex multi-dimensional constructs such as mood and anxiety are often measured using single-item questions, to reduce the assessment burden of participants, who are prompted several times per day. You should ask yourself (and the scientific literature) what the psychometric properties are of these measures. How do these EMA-measures compare to scores on traditional assessments (e.g., self-report questionnaires, clinical interviews)? What is known about the variability of items scores? And what is known about the measurement errors? Surprisingly, these last questions are often ignored in EMA research, even though it has been estimated that up to 30% to 50% of the observed variance in intensive repeated measures data can be attributable to measurement error (Schuurman et al., 2015).

In passive EMA research, it is important to be aware of the characteristics and limitations of the ‘data acquisition interfaces’ (Stone and Shiffman, 2002). For example, if you want to collect accelerometer data, a variety of options exist. You can collect these data via the built-in accelerometer of the smartphone of the participants, via cheap commercially available activity trackers, such as Fitbit, or via expensive wearable devices that are developed specifically for scientific research. Each option comes with specific advantages and disadvantages. Smartphone-based accelerometer data, for example, can be collected with relative ease, but these data can also be noisy and incomplete, since samples rates can often not be set and the precision of the built-in sensors varies considerably from device to device. Commercial accelerometers may have better precision and reliability (see, e.g., Evenson et al., 2015), but manufacturers often limit access to raw data and data pre-processing algorithms, making it difficult (or even impossible) to fully analyze outcomes. ‘Scientific wearables’ do offer this access, but often choose function over form (design). They can therefore draw attention, prompting unwanted questions to participants. Being aware of these issues when you plan the study, will help considerably in the analysis stage of your study.

3.4 What is the Sample Plan?

An important next step is to define the EMA data sample plan. Questions that need to be answered are:

- How many days will data collection last?
- On each day, how often are participants assessed?
- How and when are participants invited for assessment?

The questions above should be answered as detailed as possible to best serve the research question and the statistical power (see below). In practice, however, it is often necessary to balance research interests, respondent burden, and practical considerations, such as hardware limitations.

When determining the appropriate sample plan, start with mapping the expected fluctuation or patterns, based on available knowledge. For example, when an event is rare, it can be sufficient to ask participants to initiate EMA whenever the event occurs, or prompt them with an end-of-day diary. Adding more prompts in this scenario would not lead to more reliable data (Piasecki et al., 2007). Increasing the assessment frequency and study duration will allow for a more detailed assessment of the outcome of interest. It is tempting to collect often and for a long period of time. However, this may also increase respondent burden, which may, in turn, affect study adherence and accuracy. Measurement reactivity could also occur, where the EMA-induced enhanced focus on the outcome of interest causes participants to increase or decrease on this outcome (Hufford et al., 2002; Van Ballegooijen et al., 2016).

Issues related to hardware should also be considered. Electronic wearables have limited battery life and memory storage space. Memory space limitations of actigraphy watches may require participants to visit the research site. GPS-monitoring apps may have a negative impact on the battery life of the smartphone of the participants. These practical issues may also result in data loss and study drop-out.

Once all decisions related to the sampling plan are made, the procedure should be thoroughly tested. As a first step, it can be insightful to simulate the sample plan, as is done below, using the `sample_plan` function, which is part of package `emaph`:

```
# Simulating a signal-contingent sample plan.
library(emaph)
plan <- sample_plan(n_participants = 5,
                     n_days = 2,
                     times = c("09:00-11:00", "12:30", "17:00-19:00"),
                     plot = TRUE)
```

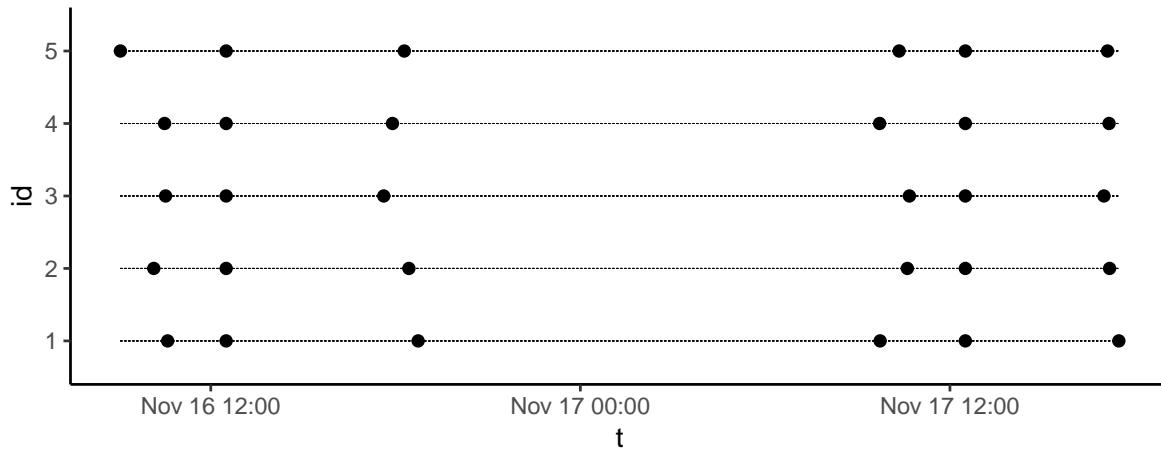


Figure 3.1: Generated two-day EMA sampling plan, for 5 participants

3.5 What is the Optimal Sample Size?

The power of a statistical test is the probability that it will detect an effect when this effect, in reality, exists. It is a function of the strength of the effect size, sample size, the significance level (α), and the statistical model. Determining the power of the experiment is an important step in the design of any study - EMA studies included. Both under-powered and overpowered studies are a waste of time and resources.

Conducting a power analysis can be easy or very difficult, depending on the complexity of the experimental design and the adopted statistical technique. For simple tests, such as the t-test and ANOVA, straightforward analytic solutions exist, which are implemented in readily available tools. In R, one of those tools is package `pwr` (which you can install, as you now know, via `install.packages('pwr')`).

For example, to use `pwr` to calculate the power of a t-test to detect a moderate effect size ($d = 0.5$), with $n = 30$ per group, and a (two-sided) significance level $\alpha = .05$, type:

```
# Power analysis of a t-test
# (analytical approach).

library(pwr)
pwr.t.test(d = 0.5,
            n = 30,
            sig.level = 0.05,
            type = "two.sample",
            alternative = "two.sided")
#>
```

```
#>      Two-sample t test power calculation
#>
#>      n = 30
#>      d = 0.5
#>      sig.level = 0.05
#>      power = 0.4778965
#>      alternative = two.sided
#>
#> NOTE: n is number in *each* group
```

The power is 48% - not even close to the generally adopted standard of 80%. More participants are needed to detect the hypothesized effect. Can you find the `n` for which power is 80%?

EMA study designs are often characterized by repeated measures, complex multi-level structures and the application of advanced statistical techniques. You may find that available power calculators are too limited to properly take key aspects of your design into account. If this happens, simulation techniques may help. If power is the probability that a test will detect an effect it exists, it can be determined by noting the proportion of times a statistical test reaches significance, if it is run, many times, on simulated data, in which the hypothesized effect is present. To illustrate how this works, we will calculate the power of the t-test again, through simulation:

```
# Power analysis of a t-test
# (simulation approach).

m1 = 0      # mean in group 1
m2 = 0.5    # mean in group 2
sd = 1       # sd (in both groups)
n = 30      # sample size, per group

# conduct the experiment many times
nsim <- 10000
p <- numeric(nsim)
for (i in 1:nsim) {

  data <- data.frame(
    outcome <- c(
      rnorm(n, m1, sd), # group 1 data
      rnorm(n, m2, sd)  # group 2 data
    ),
    group <- c(
      rep(1, n), # group 1 indicator
      rep(2, n)) # group 2 indicator
  )

  # save significance of test
  p[i] <- t.test(outcome ~ group, data)$p.value
}

# power
sum(p < 0.05) / nsim
#> [1] 0.4717
```

As can be seen, the simulation results are very close to the output of `pwr.t.test`. There was no immediate need to run this simulation. We already knew that the power was 48%. The example illustrates, however, that simulation is a valid option when power calculators are too limited. Simulating the right data, of course,

can be challenging as well, but you will find that R has packages that simplify data simulation. For example, `mvrnorm` in package MASS (Venables and Ripley, 2002) can be used to generate correlated data, and package `simstudy` (Goldfeld, 2018) can be used to generate complex longitudinal and hierarchical data.

3.6 What are the Ethical Aspects?

All clinical studies that involve human participants need to be evaluated by a Medical Research and Ethics Committee (MREC; Dutch: ‘METC’). Recently, the committees have also been tasked to determine whether a medical device is used and to evaluate the safety and quality of the device. Researchers are therefore required to add a section in the research protocol, explaining why the software/device is or is not a medical device. The official definition of a medical device (Medical Device Act, or ‘Wet Medische Hulpmiddelen’) is as follows:

“Any instrument, apparatus or appliance, any software or material or any other article that is used alone or in combination, including any accessory and the software required for its proper operation, that is intended by the manufacturer to be used specifically for diagnostic or therapeutic purposes, and is intended by the manufacturer to be used for human beings for the purpose of:
- diagnosis, prevention, monitoring, treatment or alleviation of disease - diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap - investigation, replacement or modification of the anatomy or of a physiological process - control of conception, and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its function by such means.”
(CCMO, 2018)

In short, software can be classified as a medical device when it collects patient-specific data and when it is specifically intended for one of the above-mentioned objectives. In practice, the definition of medical devices leaves a lot of room for confusion. Researchers often struggle with the question whether their assessment tools should be considered a medical device or not. For this purpose, flowcharts exist that help to determine whether an app or another software product should be classified as a medical device (see, e.g., Ekker and van Rest, 2013, and <http://cetool.nl/general/scanAid>). Figure 3.2 shows such a flow-chart.

While planning your EMA study, you should also be mindful of the rules and regulations that apply to data collection, storage and sharing. From May 2018 onward, the European Committee has enforced the General Data Protection Regulation (GDPR; in Dutch: “Algemene Verordening Gegevensbescherming - AVG”; see <https://gdpr-info.eu/>), which protects the data and privacy of EU citizens. Complying to the GDPR can be a complex and time-consuming process, depending on nature of your study. Do not hesitate to consult local experts or standard guidelines provided by your organization (see, e.g., <http://www.emgo.nl/kc/checklist-avg/>).

In EMA studies, an important GDPR-related issue concerns the Data Processing Agreement (DPA). When data processing is (partly) outsourced to a third party, a Data Processing Agreement (DPA) should be drafted, that specifies the agreements between the ‘controller’ and ‘processor’. In this context, a controller is the person or organization that determines the why and how of data collection (for example you as a researcher). The processor processes the data on behalf of the controller, for example by providing a data-collection service, or by storing raw study data on a cloud-server.

Aspects of data processing that need to be addressed in this agreement include the context, duration and termination of agreement, the nature of the data collected, the duration of data storage, security measures taken to prevent unauthorized access, and agreements on data ownership, data sharing, the handling of personal data breaches, and liability. Most organizations offer model agreements in which all relevant issues are addressed. Third parties may offer model agreements as well. If so, however, these agreements need to be checked for compliance to local regulations.



Figure 3.2: Flow-chart to determine whether study devices (including EMA apps) should be considered a medical device.

Chapter 4

Data Management

In a typical EMA study, lots of data are collected. Repeated self-reports, GPS-data, accelerometer data, background demographic data and traditional questionnaire data quickly add up to gigabytes of raw data. Without proper data management, the EMA researcher would drown in these data. Fortunately, R and RStudio are useful aids to prevent this from happening. R is very flexible in the management of multiple data files, and RStudio includes a handy feature, called “Projects”, with which data and analysis scripts can be stored in an orderly fashion.

4.1 Using RStudio-projects

RStudio projects can be opened by double-clicking existing projects, or by creating new projects from RStudio’s file menu. To create a new project, choose **File > New Project....** You will be asked to specify the project name and its disk location (as shown in 4.1, after which the project will open in a new window).

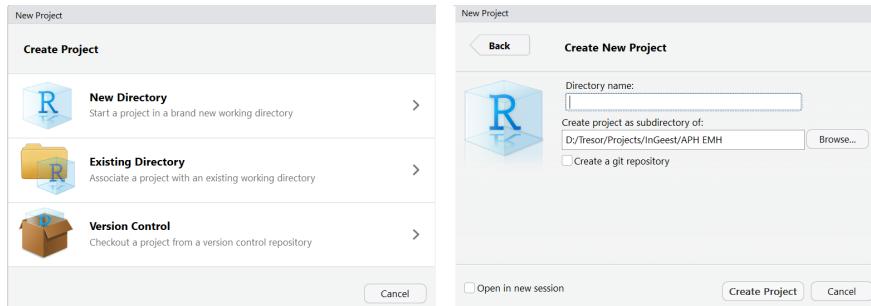


Figure 4.1: creating a project in RStudio

One of the advantages of using RStudio Projects is that projects set the working directory to the project directory location. You can verify this by asking R to print the current working directory, by typing the `getwd()` function in the console, while the project is open.

This may look like a trivial feature. It is, however, a great advantage, because it allows for the use of relative paths, which is very convenient. For example, to open a data source in a script, you no longer have to specify its full path (e.g., ‘D:/work/projects/my_ema_project/data/raw/ema_data.csv’). With relative paths, you can simply type ‘data/raw/ema_data.csv’. This saves typing, but, more importantly, it allows you to freely move your projects to other locations, without breaking the proper functioning of your scripts.

4.2 An Example Project Directory Structure

RStudio imposes no limitations to the contents of project directories. You are free to organize the project in the way you want. In clinical research, however, you are advised to choose a structure that aids you best to implement the following research guidelines:

- You should adhere strictly to a clear and logical directory structure, to ensure that co-workers and external auditors can quickly grasp and replicate your work;
- Raw data should be part of your project, so that results can be traced back to their source;
- Cleaned data should be separated from raw data, and data cleaning procedures should be explicit and replicable;
- Data and analyses should be clearly separated;
- All analyses should be explicit and replicable;
- Output of (published) analyses should be saved.

A directory structure based on these guidelines is listed in Figure 4.2. It shows the organization of a (hypothetical) project in which data were collected via:

1. an online survey system, to assess demographics and pre/post study depression severity (with the PHQ-9 questionnaire, Kroenke et al. (2009)),
2. an active EMA smartphone app, to assess day-to-day changes in mood, and
3. an accelerometer to assess activity levels.

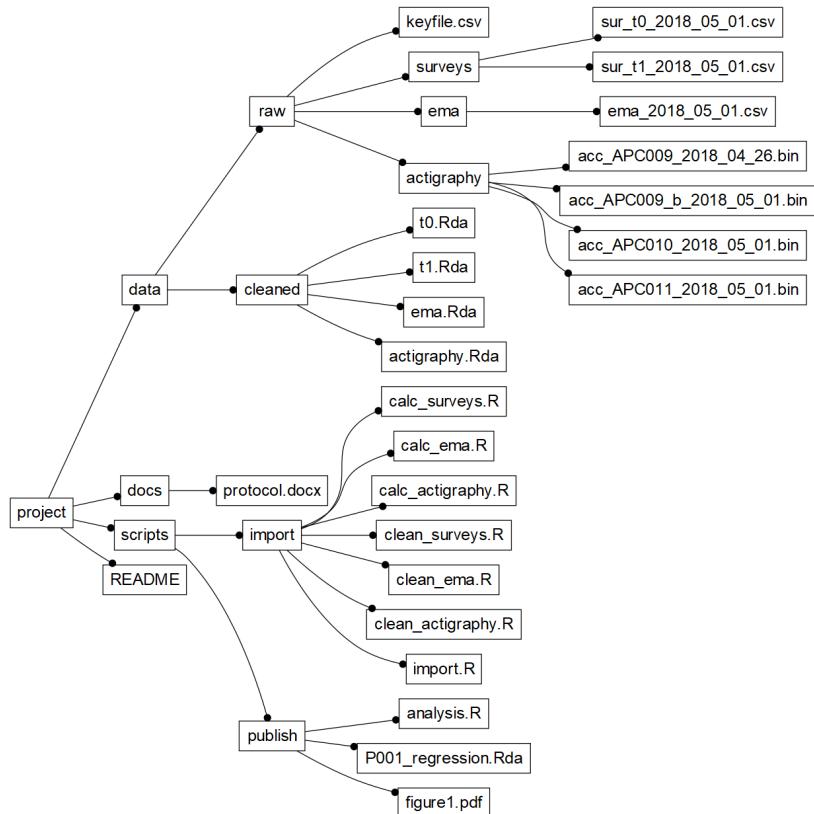


Figure 4.2: Example project directory structure

4.3 Data

Full replicability implies that raw data can be traced back to the source. Wherever possible, this should translate to the availability of raw study data in your project.

In the example project, unprocessed data files exported from the three data collection systems (i.e., survey data, EMA data, and actigraphy data) are stored in the 'data/raw'-directory, in separate sub-directories per data type.

In 'data/raw/survey', we find two files: 1) 'sur_t0_2018_05_01.csv', containing the results of demographic questionnaire and the PHQ-9 pre-test, and 2) and 'sur_t1_2018_05_01.csv', containing the results of the PHQ-9 post-test. In exports of such survey systems, data from all participants are typically stored in one file per assessment moment. Note how the export date is added to the files, to make sure that future updates are only used in the analysis when explicitly noted.

In 'data/raw/ema', we find one file: 'ema_2018_05_01.csv', containing the results of the EMA mood measurements of all participants, exported from the back-office of the EMA platform.

Finally, in 'data/raw/actigraphy', we see a series of '.bin' files: binary data files that were exported from, e.g., GENEActiv smart watches, that were worn by the participants. Actigraphy data are high-volume data: these files are typically large (500MB is no exception). By using the '.bin' format, in which data are compressed, disk space is saved (in uncompressed format, data in a single .bin file can amount up to 2 GB). Unlike the survey and EMA mood data, each data file in this directory contains data of one participant. In this case, there are even two files for one participant ('APC009'), perhaps because this participant started the study with a faulty device, which was swapped during the study.

In 'data/raw', a final very important file is 'key_file.csv'. This file is important because it ties all the data together. It contains the unique identifiers (IDs) that are assigned to the study participants in the various data collection systems. Ideally, of course, participants would be designated by a single ID in each data collection system. In practice, unfortunately, this is often not possible due to limitations of the systems used. As a result, researchers are forced to deal with the fact that participants are known under different IDs in each system. With a key-file, data can be tied together through scripts.

In the example key-file below, we find four columns. The first column defines the global study ID for each participant (i.e., the ID that the researcher intends to use as the "official" ID). The other three columns define how the participant is identified in each data collection system.

Table 4.1: Example Study Key-file

ID	Survey_ID	EMA_ID	Actigraphy_ID
P001	QM01221	192.A102.83A	APC009
P002	QM01228	192.B106.73X	APC010
P003	QM01230	192.B220.00N	APC011

4.4 Import Scripts

4.5 Cleaning

In 'scripts/import', we find the scripts with which the raw data are imported and cleaned, to produce ready-to-analyze data that are stored in 'data/cleaned'.

By making these scripts parts of your project, you ensure that you and others can always trace the decisions that were made to prepare the raw data for analysis. Raw data may be updated, for instance because more participants are recruited, or because new data exports are made from the data collection systems. You may

also detect errors in the import routines during the analyses, or peer reviewers may request information that can only be found by going back to the raw data. In all these cases, the availability of ready-to-run raw data import routines is crucial.

The code snippet below illustrates the kind of data transformations that you can expect to find in an import script. With a few lines of code, raw baseline questionnaire data are imported into R, participant id's that are specific to the data collection system are replaced by the proper global study id, variables of interest are selected, data ranges are checked (and corrected if needed), and data types are set (in accordance to the study code-book). Finally, the cleaned data set is saved to the 'data/cleaned' folder, ready for further processing in the final analyses.

```
# -----
# Clean data
# JR - 2018-10-16
# -----



# import raw baseline survey data (csv file) -----
t0 <- read.csv("data/raw/survey/sur_t0_2018_05_01.csv")

# inject study ID, from study key file -----
keys <- read.csv("data/raw/keyfile.csv")
t0 <- merge(t0, keys)

# only keep variables of interest -----
# (gender, age, phq item scores)
t0 <- t0[c("ID", "gender", "age", paste0("phq", 1:9)]



# turn gender into a proper factor -----
t0$gender <- factor(t0$gender, levels = c("M", "F"))



# replace out-of-range data with missing values (NA) -----
t0$age[t0$age > 100] <- NA
t0$age[t0$age < 5] <- NA



# replace out-of-range phq item data with missing values -----
t0[paste0("phq", 1:9)] <- lapply(
  t0[paste0("phq", 1:9)],
  function(x) {
    x[x < 0] <- NA
    x[x > 3] <- NA
  })
 



# save cleaned baseline data -----
save(t0, file = "data/cleaned/t0.Rda")
```

The code snippet above illustrates the importance of code documentation. You may struggle to immediately understand some of the code segments. For instance, you might not be familiar with the `paste0` function that is used to create the names of the variables that contain the PHQ-9 item scores. Fortunately, however, the comment (the lines starting with `#`) make it clear that the function is used to select individual PHQ-9 items. Comment your code. You will do yourself and your colleagues a big favor by making it much easier to quickly grasp the meaning of your code.

4.6 Pre-process

Once data are cleaned, you can enrich the data sets with variables that can be derived from the raw data, such as, e.g., survey scale and subscale scores, if needed), or actigraphy summary measures (see Chapter 6). The example below calculates the PHQ-9 sum scores from the item scores in the cleaned baseline data:

```
# -----
# Pre-process data
# JR - 2018-10-16
# -----  
  

# import cleaned baseline survey data -----
load("data/cleaned/baseline.Rda")
load("data/cleaned/posttest.Rda")  
  

# add PHQ-9 sum scores -----
baseline$phq9 <- rowSums(baseline[paste0("phq", 1:9)],  

                           na.rm = TRUE)
posttest$phq9 <- rowSums(posttest[paste0("phq", 1:9)],  

                           na.rm = TRUE)  
  

# re-save baseline data -----
save(baseline, file = "data/cleaned/baseline.Rda")
save(posttest, file = "data/cleaned/posttest.Rda")
```

4.7 Combine

While working on your project, you will probably want to re-run the cleaning and pre-processing scripts a lot, for instance in response to additional data coming in, or to fix bugs in the import routines. For this, it can be helpful to create one file in which all import scripts are executed in proper sequence. For this, you can use R's `source` command, with which scripts can be read and executed directly from disk:

```
# -----
# Import (clean & pre-process) all data files
# JR - 2018-10-16
# -----  
  

# clean -----
source("scripts/import/clean_surveys.R")
source("scripts/import/clean_ema.R")
source("scripts/import/clean_actigraphy.R")  
  

# pre-process-----
source("scripts/import/calc_surveys.R")
source("scripts/import/calc_ema.R")
source("scripts/import/calc_actigraphy.R")
```

4.8 Reproducible Analyses

When raw data stored, imported and cleaned, final analyses can be run. By basing these analyses on the cleaned data in 'data/cleaned', you ensure that these analyses can be fully reproduced from the raw study

data.

The code snippet below illustrates a typical analysis file: cleaned data are loaded into the R work environment, after which EMA scores of a single participant are selected, plotted and analyzed in a simple linear regression. Both the plot and the result of the regression are saved in the analysis directory. The plot is saved as a PDF-file (ready for submission to the journal), and the regression results are saved in a standard R data structure.

```
# -----
# N = 1 Analysis (P001)
# note: part of manuscript!
# JR - 2018-10-16
# -----

# import cleaned EMA mood study data -----
load("data/cleaned/ema.Rda")

# create and save Figure 1: EMA mood data, of participant P001
d <- subset(mood, ID == "P001")
pdf(file = "scripts/published/figure1.pdf")
plot(mood ~ time, d)
dev.off()

# run a regression model on P001 mood data
fm <- lm(mood ~ time, d)
summary(fm)

# save regression results
save(fm, file = "scripts/published/P001_regression.Rda")
```

R's ability to save the results of analyses to disk is yet another example of how R promotes accountability in clinical research. Suppose you used the regression analysis of 'P001' in a manuscript that you submitted for publication. Reviewers ask you to send residual plots, to convince them that the residual errors are normally distributed. When the regression results are saved to disk, the following three lines are all you need to satisfy their request:

```
# Revisiting regression results, for a visual regression residual check
load("scripts/published/P001_regression.Rda")

pdf(file = "scripts/publised/residual_plot.pdf")
plot(fm)
dev.off()
```

4.9 Documentation

4.9.1 Protocol

In “docs/protocol”, we find the “protocol.docx” file, which should contain a detailed description of the study. This file provides information on the background of the study, the methods, analysis techniques, defines the codebook of the study, and should include full descriptions of all surveys, assessment protocols, etc.. If needed, you can use this folder to store additional background material, such as PDFs of published psychometric studies.

4.9.2 README

Note, finally, the ‘README’ file in the root of the project directory. This file should contain a brief summary of the project, to quickly inform others (and your future self) of the context of the project and the contents of the project directory.

Element	Description
Title	Project title & Acronym .
Description	One-paragraph project description.
Author	Author name, e-mail affiliation.
Getting Started	Instructions on how to get the project up and running on a local machine for development.
Prerequisites	A listing of software required (i.e. R packages), and instructions on how to install this software.
Contents	A listing of project directories, with a brief description of their contents.
Restrictions	Notes about potential data access limitations.

4.10 Discussion

In this chapter, we aimed to show how adopting the RStudio Project can help you to implement key principles of EMA data management. To illustrate this, we discussed the project structure of a small-scale EMA study. No doubt, your project will differ from this example in many ways, forcing you to deviate from the example structure. The example may be too elaborate, for example, if your project only requires you to analyze a single data file. The structure is certainly too limited to support the requirements of a full PHD project (such as the one described, for example, in the APH quality handbook - see <http://www.emgo.nl/kc/folders-and-file-names/>). But as we mentioned earlier, RStudio Projects are flexible. It should be relatively straightforward to scale down or scale up the example that we discussed.

If you want to learn more about data management with R and RStudio, the book “*Reproducible Research with R and RStudio*” (Gandrud, 2015) would be a good place to start. You may also be informed by the data management techniques that are described in the first two chapters of the book “*Using R and RStudio for Data Management, Statistical Analysis, and Graphics*” (Horton and Kleinman, 2015).

Part III

EMA Outcomes

Chapter 5

Mood

Mood is a common outcome in EMA research (Myin-Germeys et al., 2016; Desmet et al., 2016). Having respondents rate their mood during the day allows researchers to assess mood fluctuation over time or reactivity to events and daily-stressors (Wenze and Miller, 2010). Often, it is studied in relation to depressive symptoms and mood disorders (Aan het Rot et al., 2012). In addition, mood can be linked to other variables, such as substance abuse (Kirchner and Shiffman, 2013; Serre et al., 2015), somatic health (Engel et al., 2016; Moore et al., 2016) or activity patterns (Dunton, 2017; Marszalek et al., 2014).

The definition of mood varies across studies. Usually the concept refers to a general affective state. Following this line of reasoning, a distinction can be made between mood states (e.g. irritable, cheerful, relaxed, etc.) and discrete emotions (e.g. happy, sad, anxious, etc.), where moods are thought to be less specific and more subjective, enduring and related to context (Beedie et al., 2005; Cranford et al., 2006; Desmet et al., 2016).

Depending on the study focus and research questions, mood measurement can be operationalized in several ways. Therefore, it is vital to consider the goal of measuring mood in your own study and to choose an operationalization that matches your hypothesis and theoretical framework. In this chapter, we will discuss the most commonly used options: 1) unidimensional mood assessment, 2) the ‘bag of items’ approach, and 3) dimensional models, namely the Circumplex model and Negative and Positive affect (NA/PA).

5.1 Unidimensional Mood Assessment

Perhaps the most seemingly straight-forward method to measure mood is to ask ‘face-valid’ unidimensional questions such as “*How is your mood right now*” (Van Ballegooijen et al., 2016) or “*How are you feeling right now*” (Van de Ven et al., 2017). Respondents usually rate these questions on a Visual Analogue Scale (VAS), aimed to indicate mood intensity. Typically, VAS scales will range from zero (low or worst mood) to 10 or 100 (good or best mood). Keep in mind that the middle of a VAS scale (e.g. 5 or 50) is generally considered a negative result, and only scores above 6 or 60 are considered acceptable or positive mood states (Groot, 2010). In order to address this issue, some researchers have proposed to use VAS-scales ranging from -1 to 1, with 0 as a neutral center. However, such a scale implies a mood state that ranges from negative to positive, rather than absent to present. Another alternative is to use Likert scales, where the scale center often reflects a neutral response.

Plotting data from an unidimensional item in a graph is an easy way to visually inspect within-subject change in general mood:

```
# Plotting data over time.  
library(ggplot2)  
library(emaph)
```

```
plotmood_down <- ggplot(csd, aes(x = date, y = as.numeric(mood_down))) +
  geom_smooth(method = "loess", span = .05, se = FALSE, colour="dodgerblue4") +
  geom_point(size = .3, alpha = .3, position = position_jitter(height = .1),
  colour="dodgerblue2") + scale_x_date() + scale_y_continuous(breaks = 1:7) +
  xlab("Time") + ylab("Mood")
print(plotmood_down)
```

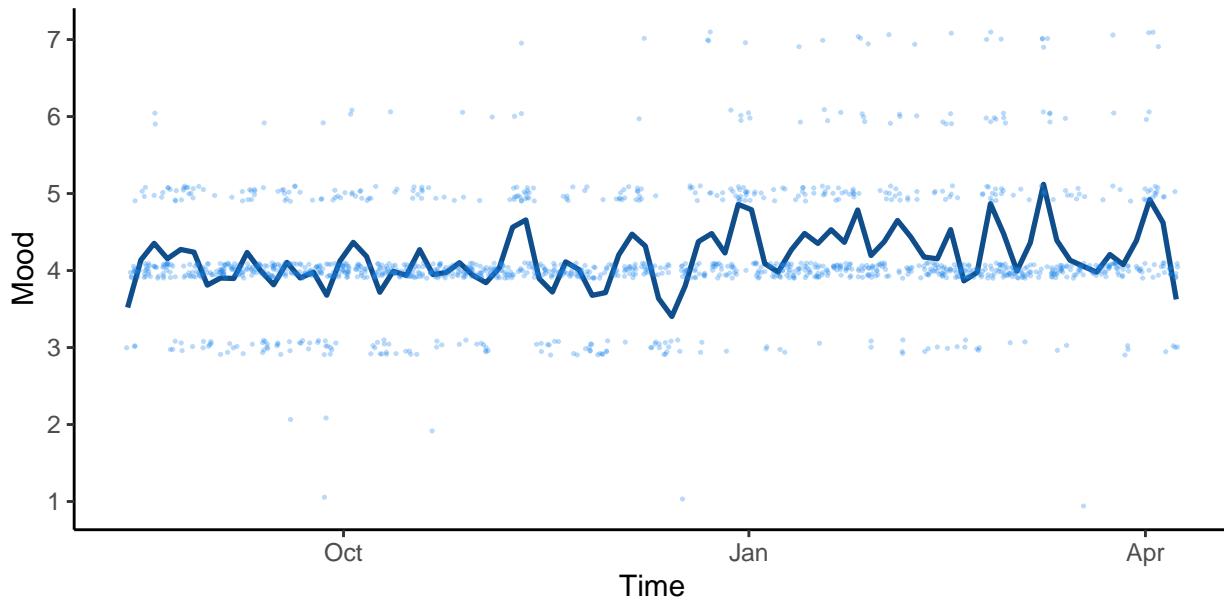


Figure 5.1: 34 weeks of mood data, from a single participant

5.1.1 Bipolar Unidimensional Items

Another option to assist respondents with the interpretation of one-item mood ratings, is to use a bipolar scale. These items place two opposing mood states at each end of the scale, for example by asking “Please rate your current mood on a scale of 0 to 100, on which 0 indicates happy, and 100 indicates sad” (Van Rijssbergen et al., 2014). This does assume that the opposing mood states, such as happy and sad, are mutually exclusive and thus cannot occur simultaneously. The bipolar-unidimensional method was shown to be able to predict time to relapse over 5.5 years in recurrently depressed out-patients, with 6.3% of variance in time to relapse explained. This percentage was comparable to that of the HAM-D (Van Rijssbergen et al., 2012). Also, the scale was able to detect relapse in patients with recurrent Major Depressive Disorder (based on SCID-I interview) at a cut-off score of 55, and outperformed the HAM-D and IDS-SR. However, 47% of patients indicated by the VAS scale did not fulfill formal criteria for relapse (false positives) (Van Rijssbergen et al., 2014).

5.2 Bag-of-Items

In order to make sure all constructs of interest are measured, you can also consider including a number of specific mood items in your EMA questionnaire, rather than one general uni-polar item or one bipolar item. For example, you can ask respondents *“How depressed are you feeling right now?”* and *“How anxious are you feeling right now?”* (Starr and Davila, 2012). This strategy often leads to a ‘bag-of-items’ approach, where single items from various sources, such as existing questionnaires, are combined into a new EMA

questionnaire. A benefit of this approach is that you can select items for which information on validity and test-retest reliability is available. A downside is that item scores can only be evaluated separately, rather than providing one overall indication of mood or well-being.

Combining data from multiple items, such as mood and loneliness, in one graph can provide respondents with insight in the interaction between the constructs. In R two variables can easily be plotted together:

```
# Plotting multiple variables in one graph.
library(ggplot2)
library(emaph)

combined <- plotmood_down +
  geom_point(data=csd, aes(date, as.numeric(mood_lonely)),
             size = .3, alpha = .3, position = position_jitter(height = .1),
             colour="indianred4") +
  geom_smooth(data=csd, aes(date, as.numeric(mood_lonely)),
              method = "loess", span = .05, se = FALSE, colour="indianred2")
print(combined)
```

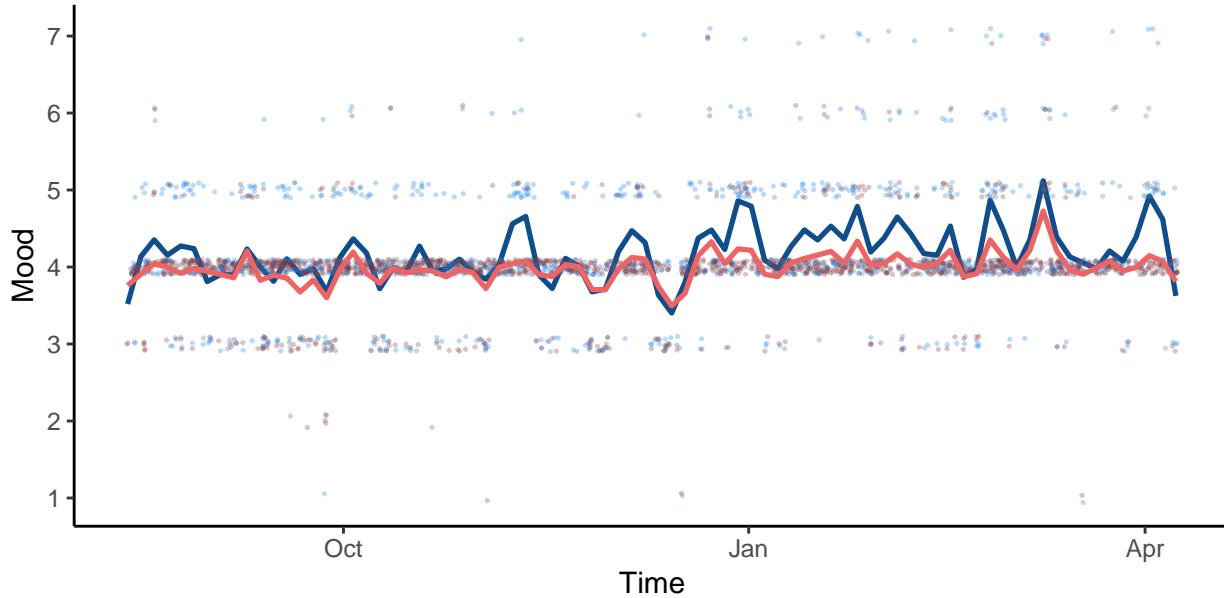


Figure 5.2: 34 weeks of combined mood data, from a single participant

5.3 Multi-dimensional Mood Assessment

Dimensional models assume that every affective state or emotion should be described by the combined score on (at least) two items, instead of just one. Over the past decades several multi-dimensional models have been specified (for an overview, see Sander and Scherer (2009)). In the context of EMA, researchers most often base their items on the Circumplex model (Russell, 1980) or the Negative/Positive affect (NA/PA) theory (Watson and Tellegen, 1985).

5.3.1 The Circumplex Model

The Circumplex Model of affect (Russell, 1980; Posner et al., 2005) argues that all mood states are a linear combination of two independent, bipolar scales: valence (ranging from unpleasant to pleasant) and arousal/activation (ranging from low to high arousal). Combining scores on these scales places the affective states in a circle on one of four quadrants (see Figure 5.3). States within one quadrant are believed to be positively correlated, while states in the opposing quadrant are thought to be negatively correlated.

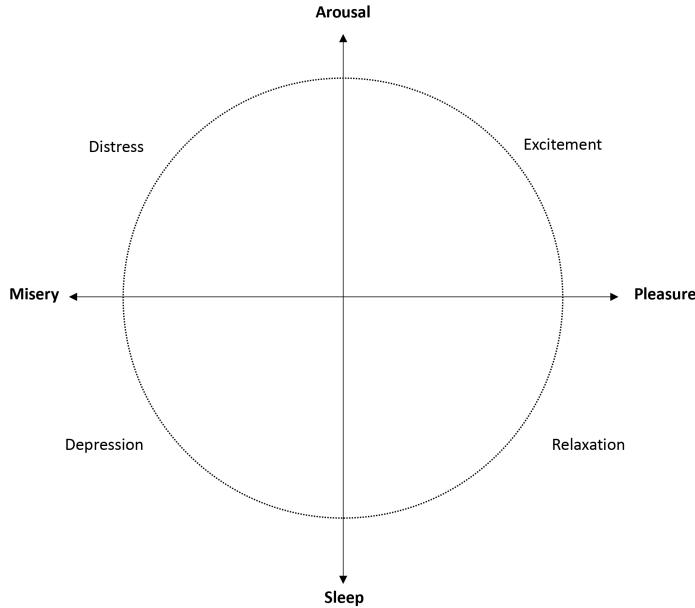


Figure 5.3: Russell's Circumplex model of affect.

There are several options to operationalize the Circumplex model in EMA research. For example, respondents can rate valence and arousal on two VAS scales. The most pragmatic approach is to report both scale scores separately (Asselbergs et al., 2016). Alternatively, scores can be combined to give insight into which of the four mood states (quadrants) respondents fall:

- Low aroused - unpleasant
- Low aroused - pleasant
- High aroused - unpleasant
- High aroused - pleasant

A downside of the Circumplex model is that the concepts of valence and arousal can be hard to convey to respondents, especially when translated to other languages, such as Dutch. One alternative is to adjust the scale ends, for example using “extreme sleepiness” and “extreme high energy” (Sharar et al., 2016). Another option is to use pictures or emoticons, rather than language. For this purpose, DeSmet and colleagues developed the pick-a-mood scale, which is a pictorial self-report scale (Desmet et al., 2016). The scale builds on the circumplex model and adds the third dimension “dominance” (level of experienced control over the mood state), rendering eight (instead of four) different mood states and one neutral state (see Figure 5.4).

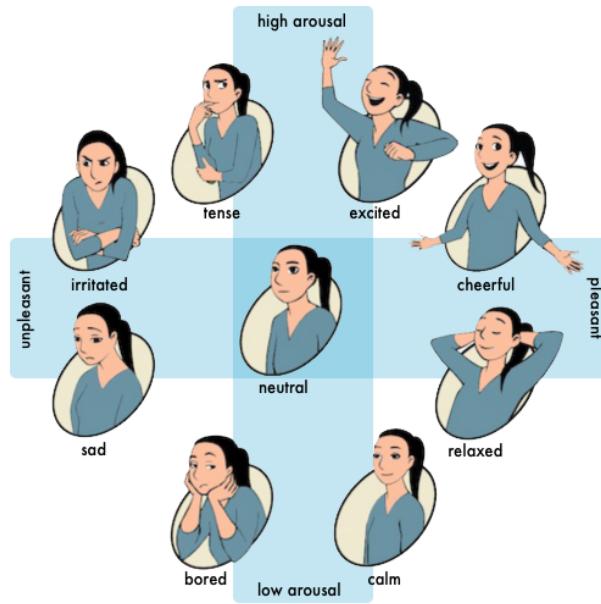


Figure 5.4: The Pick-A-Mood Circle.

5.3.2 Negative & Positive Affect

Watson and Tellegen (Watson and Tellegen, 1985) also specified the underlying theory of the Circumplex model, arguing that the diagonal quadrants represent Positive and Negative affect (PA/NA) and that these two terms are the main dimensions of affect (Watson and Clark, 1994). PA ranges from sadness to high energy, NA from calmness to distress (Watson et al., 1988). While bipolar-unidimensional assessment assumes that positive and negative affect are mutually exclusive, the PA/NA affect model assumes that these mood states can occur simultaneously. Watson and Clark (Watson and Clark, 1997) for example, showed a moderate correlation between the two constructs ($r = .32$). In order to measure Positive and Negative Affect, a designated Positive and Negative Affect Schedule was developed by Watson, Clark and Tellegen (Watson et al., 1988). Respondents are asked to indicate “*to what extent you feel this way right now*” on 20 affect items. The scale uses a 5-point Likert-scale, ranging from 1 (very slightly or not at all) to 5 (very) (Watson and Clark, 1994). Items include:

- Negative Affect (10): afraid, scared, nervous, jittery, irritable, hostile, guilty, ashamed, upset, distressed
- Positive Affect (10): active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong

There are several short-forms available. For example, Wichers and colleagues created a 10-item short-form of the PANAS for their EMA studies. The items were based on the PANAS and their own experience with EMA (Wichers et al., 2012). Using factor analyses, the following 7-point Likert items were chosen for the questionnaire:

- Negative affect (6): insecure, lonely, anxious, low, guilty, suspicious.
- Positive affect (4): cheerful, content, energetic, enthusiastic.

PA and NA were calculated as the average score across all items and weighted for their factor loadings (Wichers et al., 2012). In R, such a factor analysis can be executed as follows:

```
# Performing a factor analysis.
```

```
library(dplyr)
library(tibble)
```

```

items2 <- csd %>%
  select(c(
    "mood_relaxed", "mood_satisfi",
    "mood_enthus", "mood_cheerf",
    "mood_strong", "mood_down",
    "mood_lonely", "mood_anxious",
    "mood_guilty" )) %>%
  scale(.) %>%
  as.tibble(.) %>%
  mutate_all(funs(residuals(stats::arima(., order = c(1,0,0)))))

require(psych)

correlations <- cor(items2, use = "complete.obs")
model <- fa(items2,
            nfactors = 2,
            rotate = "oblimin",
            fm = "pa",
            scores = "regression")

colnames(model$loadings) <- c("PA", "NA")
psych::fa.diagram(model)

```

Factor Analysis

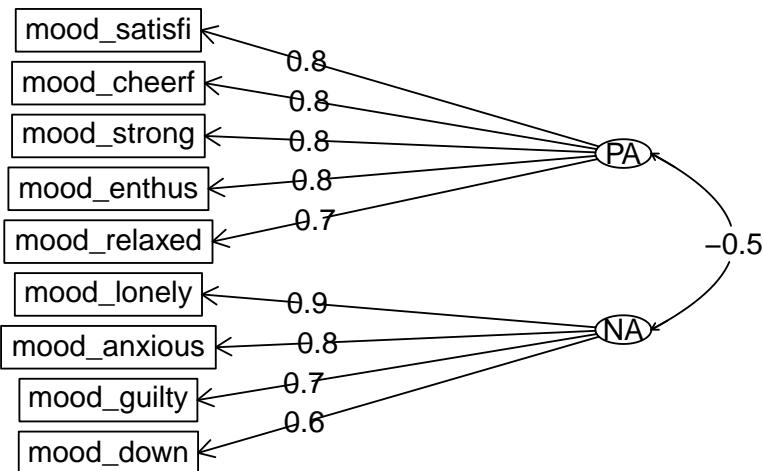


Figure 5.5: Factor analysis of scores of 9 EMA items, revealing two factors: Positive Affect (PA) and Negative Affect (NA).

Chapter 6

Activity

The objective study of human physical activity is one of the exciting opportunities created by passive EMA (Marszalek et al., 2014). Through technological advances in mobile sensing, we are now able to continuously monitor (in-)activity of participants in every-day life, with little to no participant burden.

While questions remain with regard to the validity, reliability and clinical utility of passive EMA of specific activities, such as (disturbed) sleep, sedentary behavior, and energy expenditure (see, e.g., Feehan et al., 2018; Gomersall et al., 2016), an increasing number of mental health studies are including activity tracking devices to better understand sleep habits, circadian rhythm disorders and depression (see, e.g., Saeb et al., 2015; Saunders et al., 2016; Tahmasian et al., 2013).

In this chapter, we will discuss two passive EMA methods to assess physical activity: actigraphy and geotracking. Of these two, actigraphy has been used most in human clinical research. However, due to the massive adoption of smartphones, researchers increasingly collect geolocation data as well, inspired perhaps by the elaborate geolocation data analysis techniques that have been developed in the past decades in wildlife telemetry research (Tomkiewicz et al., 2010).

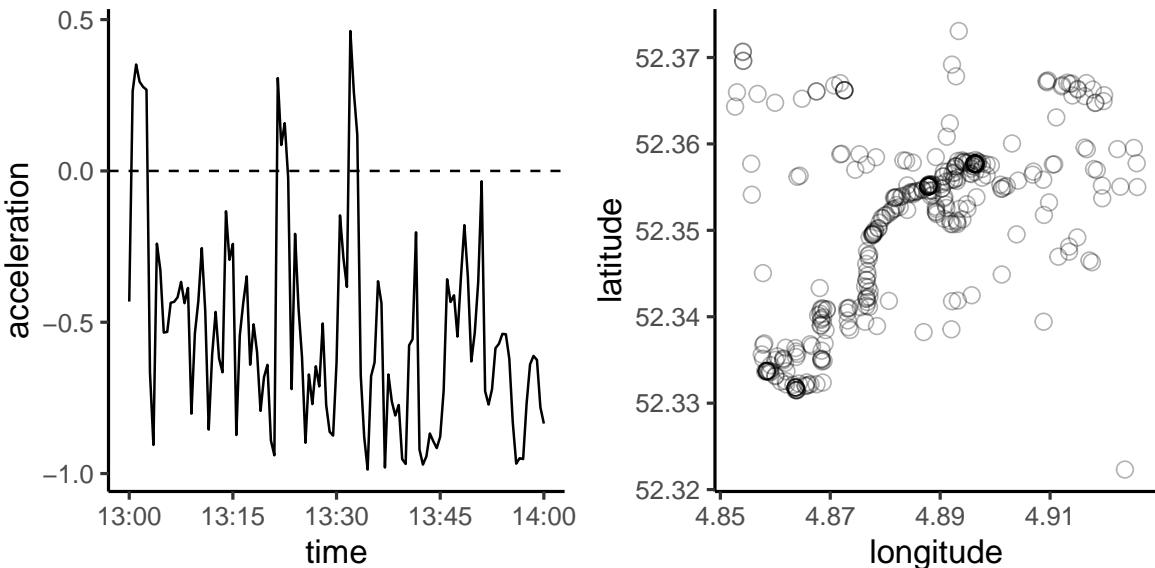


Figure 6.1: Actigraphy (left) and Geotracking (right): two methods for passive ecological momentary assessment of activity.

6.1 Actigraphy

Accelerometers are micro electro-mechanical systems (MEMS) that measure changes in acceleration forces (i.e., both static forces - earth's gravity - and dynamic forces - caused by movement), typically simultaneously on the vertical (Y), horizontal right-left (X) and horizontal front-back axis (Z). Through actigraphy, we study the frequency, duration, and intensity of physical activity. Figure 6.2 shows one hour of data collected from a wrist-worn GENEActiv accelerometer. As can be seen, three accelerometers (X, Y, Z) were simultaneously providing data. Data were sampled with a frequency of 30 Hertz (Hz; thirty measurements per second - which is common), but sub-sampled here to 0.1Hz (one measurement every 10 seconds), for practical reasons. If we would have plotted the data at 30Hz, the plot would have included 108.000 data points. At 0.1Hz, this reduces to 360 points.

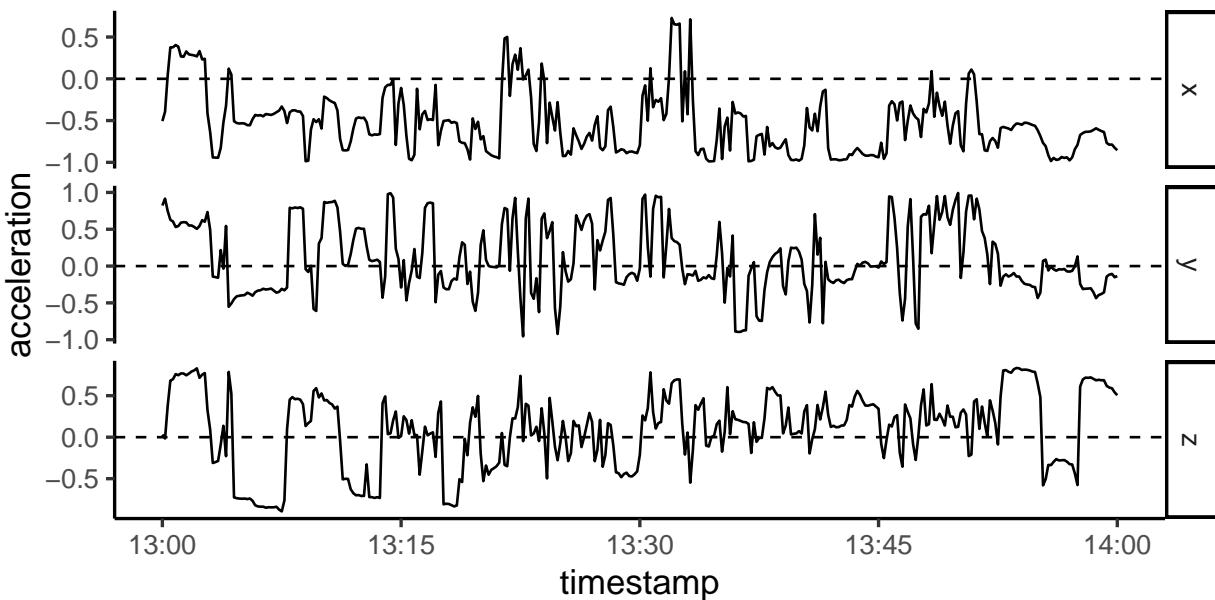


Figure 6.2: One hour of raw data collected with a wrist-worn GENEActiv accelerometer, sub-sampled to 10-second epochs (0.1 Hz)

Data shown are included in package `emaph`, and the R-code to reproduce the plot is listed below. Use this to familiarize yourself with actigraphy data. If you want to see how sub-sampling affects the number of points to plot, for example, you can set different values in the `round_date` function. For example, to get a point for each five seconds (0.2Hz), you would set the argument of this function to `5 seconds`.

```
# Plot one hour of emaph accelerometer data (of person 1).
library(dplyr)
library(ggplot2)
d <- subset(emaph::genea, timestamp > "2018-06-01 13:00" &
            timestamp < "2018-06-01 14:00" &
            id == 1)
d$timestamp <- lubridate::round_date(d$timestamp, "10 seconds")
d <- d %>% group_by(timestamp) %>% summarise_all(.funs = mean) %>%
  tidyr::gather(key = "sensor", value = "value", x, y, z)

ggplot(d, aes(timestamp, value)) + geom_line() +
  geom_hline(yintercept = 0, linetype = 2) + facet_grid(rows = vars(sensor) , scales = "free_y")
```

6.1.1 Data cleaning

Raw accelerometer data need to be cleaned before analyses can be run. Typical data import work-flows include re-calibration (to reduce systematic measurement error), the detection of non-wear periods (to ensure that non-informative data are removed), sub-sampling (reducing the sample rate to reduce analysis time) and filtering/aggregation (to smoothen the signal and reduce the impact of outliers, measurement error and occasional missing values). Study results can be highly dependent on these initial steps, which, unfortunately, are also complex and time-consuming. Specialized R-packages exist to help you with this (see, for example, package `GGIR` and `GENEARRead`, which are described in more detail in Chapter 12).

6.1.2 Feature Extraction

Properties of the signals that are of interest are highly dependent on the focus of the study. Highly detailed analysis of local peaks in the signal might be needed, for instance to reveal an association between activity and reported events. But analyses can also be more global, for instance when accelerometer data are used to study circadian rhythms in activity. Several approaches exist to combine the X, Y, Z measurements into a single meaningful metric. Two popular metrics are the ‘Signal Vector Magnitude’ (SVM) and the ‘Euclidean Norm Minus One’ (ENMO). Validation studies suggest that ENMO should be the preferred metric (Van Hees et al., 2014, 2015), although recent findings also suggest that alternative metrics should perhaps be considered when sedentary and light activities are of interest (Bai et al., 2016).

SVM and ENMO are closely related. SVM is the magnitude of the raw tri-axial signals (the Euclidean distance in the three-dimensional space), i.e. $SVM = \sqrt{x^2 + y^2 + z^2}$. ENMO is the corrected SVM: the vector magnitude remaining after removing one Earth Standard Gravitational unit ($1g = 9.81 \text{ m/s}^2$), i.e. $ENMO = SVM - 1$. The metrics can, in principle, be calculated for each $\{x, y, z\}$ -data point in the raw series. Typically, however, the metrics are calculated for time-windows (called *epochs*), in which case the mean can be used to characterize the overall activity in each epoch.

Figure 6.3 shows the development of ENMO over one day, as sampled by GENEActiv accelerometers that were worn by a young adult (top) and a middle-aged person (bottom). This figure is much easier to interpret than the plot of the raw x-y-z values in Figure 6.2. Activity levels over the day follow a similar pattern, but the activity levels in the two plots are strikingly different. Age appears to matter here: activity levels of the middle-aged person are consistently lower than those of the young adult.

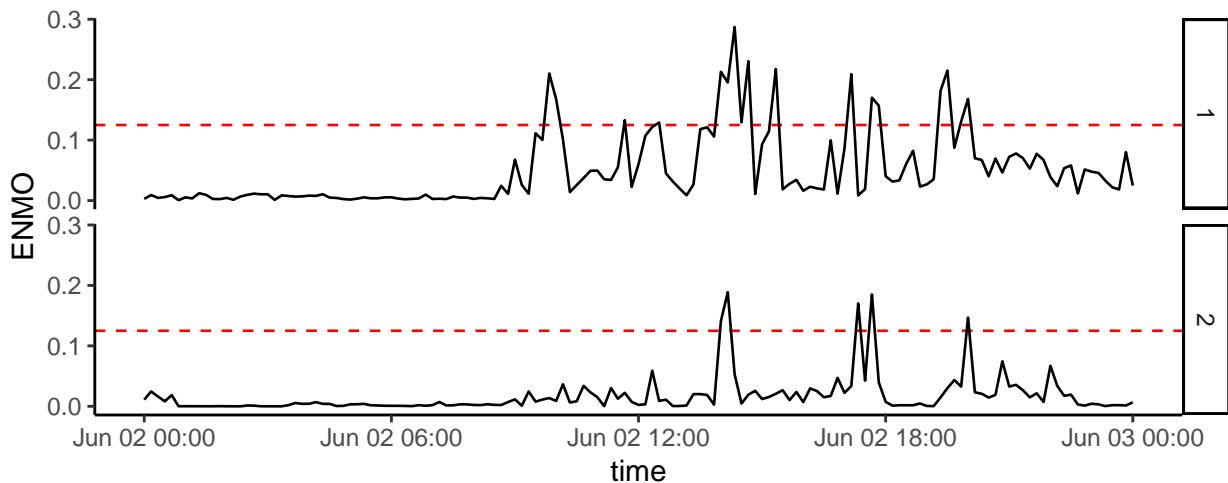


Figure 6.3: One day of data of the two persons in the GENEActiv data set of package ‘emaph’, summarised with ENMO, in 10-minute epochs

For SVM and ENMO, cut-off values for various activity classes have been determined (Da silva et al., 2014; Hildebrand et al., 2014; Kim et al., 2017; Rowlands et al., 2016). Although these cut-offs vary somewhat from study to study, a common ENMO cut-off for Moderate-to-Vigorous-Physical-Activity (MVPA) is 0.125g (125 milligravity units; Femke Lamers, personal communication, 15 november 2018). The dotted line in Figure 6.3 marks this cut-off.

With this cut-off, we can summarize the two series shown in Figure 6.3 by the number of times on which ENMO is higher than the MVPA cut-off. The daily MVPA-count for the young adult is 17. For the middle-aged person, this is 5: considerably lower.

You should be aware that the choice of the width of the epoch matters when MVPA-counts are calculated. By averaging values in each window, ENMO acts as a smoother, which may prevent you from the detection of short bursts of activity when the window is large. If we would have used a 5-second window to generate Figure 6.3, for example, the MVPA-counts would go up considerably for each person.

6.2 Geotracking

6.2.1 The Geographic Coordinate System

In the geographic coordinate system, each location on the earth is uniquely represented by two numbers: *Latitude* and *Longitude*. Latitude marks the north-south position of a point on the earth's surface, and longitude marks the east-west position (see Figure 6.4). The center of Amsterdam, for example, is {latitude: 52.37022; longitude: 4.89517}, which can be verified by punching these numbers in Google maps.

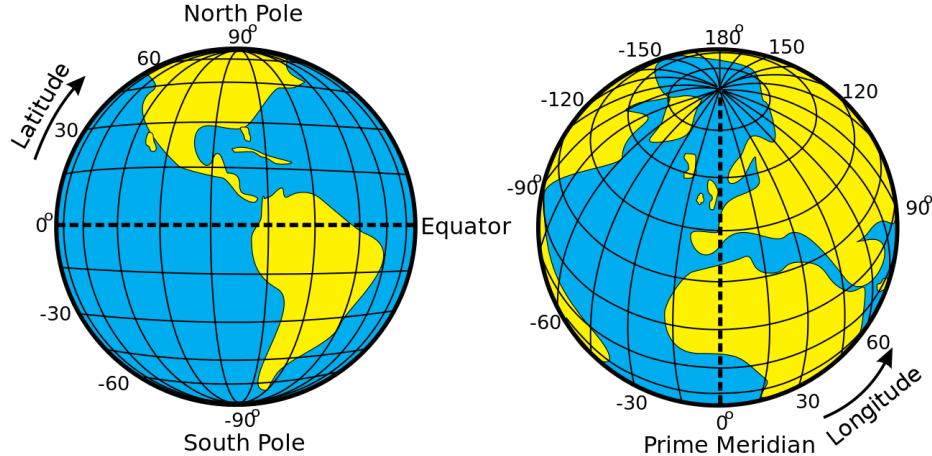


Figure 6.4: Latitude and Longtitude of the Earth (source: Wikipedia).

6.2.2 The Global Positioning System

The Global Positioning System (GPS) is a satellite-based radio-navigation system that provides geolocation and time information. With GPS-receivers, latitude and longitude can be determined, to track geographical locations and movement. Due to the increasing ease with which GPS-data can be collected via modern smartphones, recent years have witnessed a marked increase in the use of GPS-based activity measures in the study of mental health.

Figure 6.5 shows GPS-data of two people, collected over a period of four weeks, via the Google timeline smartphone app. Data can be found in the `emaph` package (see `?locations`).

```
# Plot four-week location history of emaph location data
library(ggplot2)
library(emaph)

d <- subset(locations,
             accuracy <= 50 &
               lon >= 4.80 & lon <= 5.00 &
               lat >= 52.25 & lat <= 52.50) %>%
  sample_n(4000)

ggplot(d, aes(lon, lat)) +
  geom_point(alpha = .2, shape = 21, size = 3) +
  xlab("longitude") + ylab("latitude") +
  facet_wrap(~ id)
```

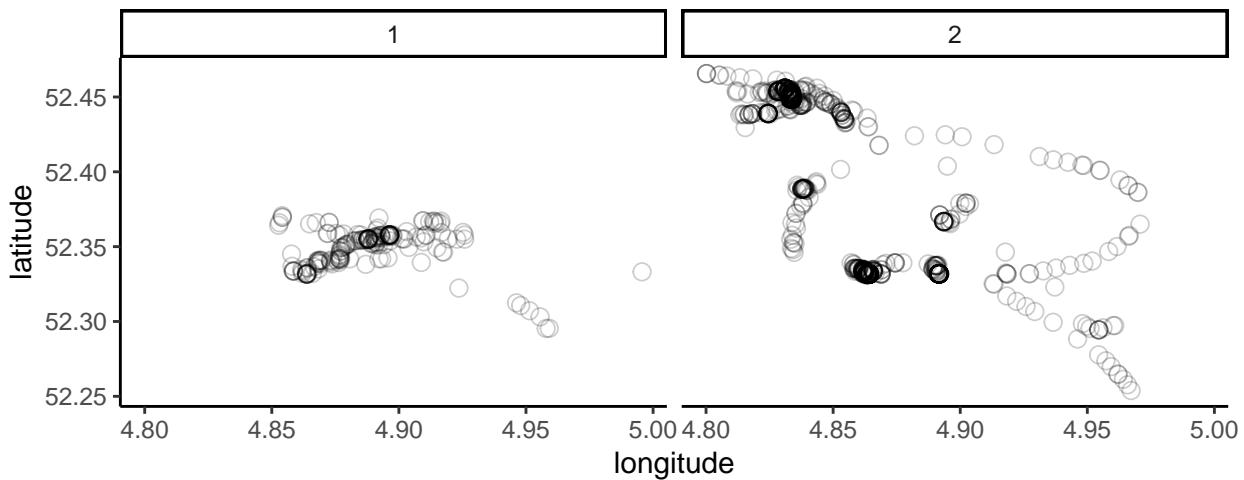


Figure 6.5: Four-week location history of two people, collected with Google Timeline.

Data-points are superposed, using transparent colors, to make a distinction between locations that were visited once (light areas) and places that were visited many times (darker areas). From the plot, we learn that these two people both lived and worked in the Amsterdam area (latitude and longitude are close to the coordinates of Amsterdam center). We also see that they shared a frequently visited location (they were co-workers, working in the same building). Locations of person 1 reveal that this person's home was probably in Amsterdam, while the locations of person 2 show that this person's home was probably located in an Amsterdam suburb. Commuting patterns (i.e., the recurrent traveling between the place of residence and place of work) are clearly visible.

It should be noted, though, that person 1 contributed much less data ($n = 722$) than person 2 ($n = 14031$). This can be explained by the different devices that were used by both: Person 1 used an iPhone (with standard GPS-settings) and Person 2 used a Sony Z1 Android (with high-precision GPS features enabled). This device-related variability in GPS sample rates and accuracy is one of the primary challenges of naturalistic EMA research and EMI applications.

The problem with the (in)accuracy of GPS-data is further illustrated by Figure 6.6, in which all data points are plotted that were registered by the smartphone of person 2 between 02:00 and 06:00. At those hours, the person was sleeping, in the bedroom of his house. He did not move. Yet, if we would take the GPS-data for granted, he regularly took a nightly random walk in the park. The red dot in the figure marks the median coordinate. This coordinate is very accurate: it marks the bedroom. All individual data points, however, fail

to identify this location.

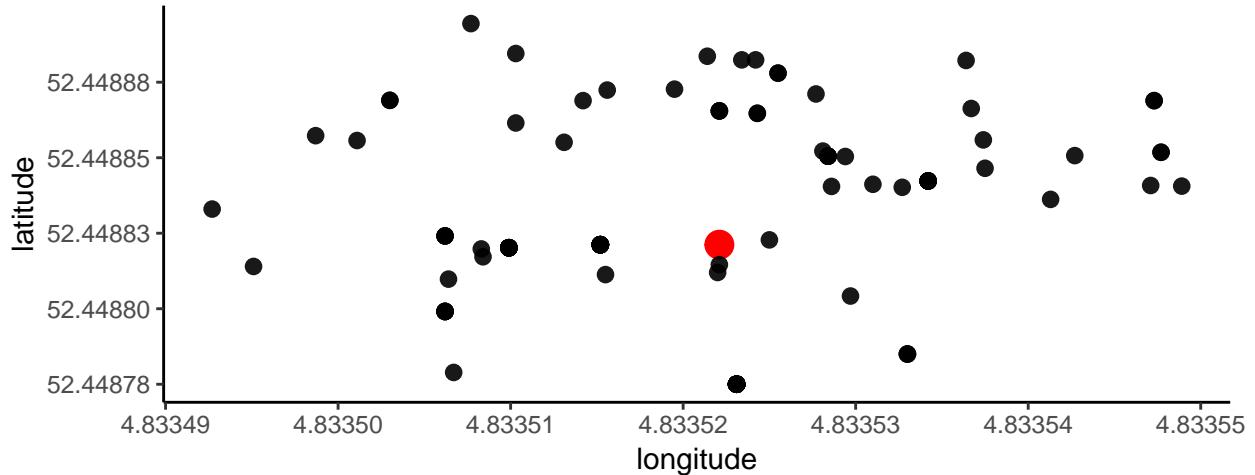


Figure 6.6: Nightly GPS-fluctuations, revealing inaccurate location measurements

6.2.3 GPS-based Activity Measures

Raw GPS-data reflect series of locations rather than activity per se. However, measures of activity can be extracted from these data.

Table 6.1 shows some of the measures that were derived from GPS data in a small ($n = 28$) study exploring the correlation between passive EMA data and depression, conducted by researchers of Northwestern University (Saeb et al., 2015). The researchers calculated total distance, location variance, the number of places visited by the participants during the study, the percentage of time spent at home (defined as a top 3 place which was most frequently visited between 24:00 and 6:00), and circadian movement - the consistency of location visits based on a 24-hour period. Circadian movement and location variance were found to be correlated with PHQ-9 scores in this study, but not - however - in a follow-up study, which included more participants (Saeb et al., 2017).

Table 6.1: Activity measures that can be derived from a GPS data set.

Name	Formula
Total distance between locations	$\sum(distance((lat_t, lon_t), (lat_{t-1}, lon_{t-1}))$
Location variance	$\log(\sigma_{lon}^2 + \sigma_{lat}^2)$
N Places	k-means(loc, lat)
Home Stay	time(cluster[home]) / time(clusters[j])
Circadian Movement	$\sum(psd(f_i)/(i1 - i2))$

Part IV

Analytic Approaches

Chapter 7

Feature Extraction

EMA data are streams: time-series that represent processes which may span hours, days, weeks or even months. To the beginning EMA researcher, it may not be immediately clear how to deal with these data. How, for example, can these series be summarized to study differences between individuals or changes within individuals over time? What ‘qualities’ or ‘features’ of a time-series should be considered? What are the available options? In this chapter, we will discuss some of the features that can be extracted from EMA time-series.

7.1 Simulated EMA Time-series Example

We will focus on a simulated three-week time-series of EMA mood responses of a single person, in which ratings were collected five times per day (using the `generate_features_dataset` function of package `emaph`). Figure 7.1 shows a plot of the simulated scores.

Stare at the figure for a moment. How would you characterize the development of scores over time? What features would you want to extract? How would you quantify these features?

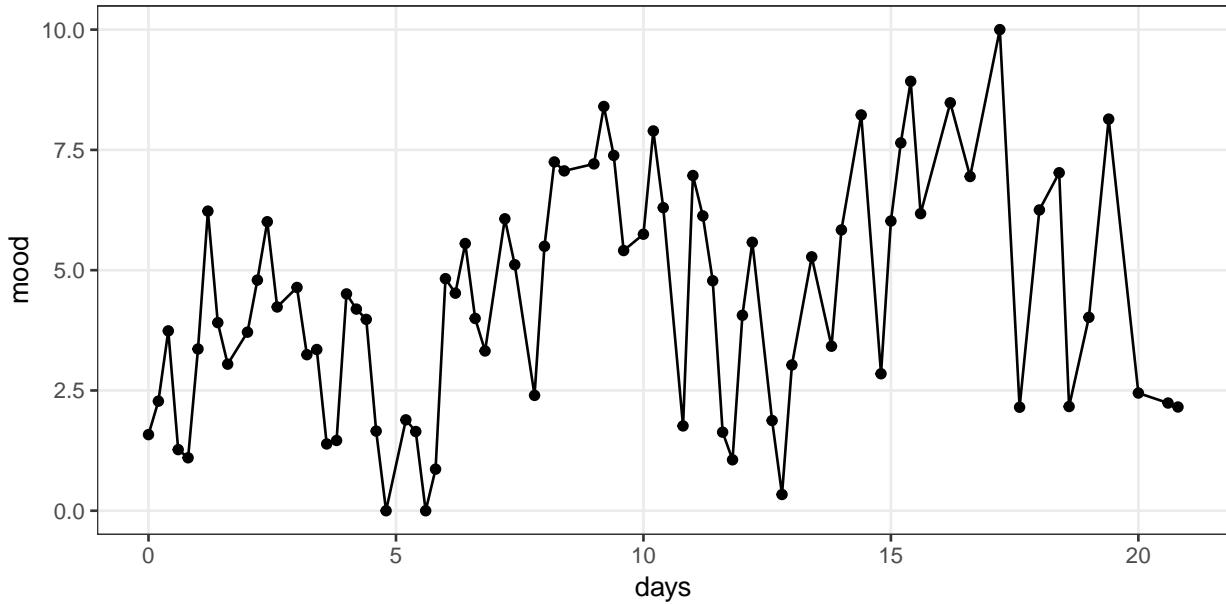


Figure 7.1: A simulated three-week time-series of EMA mood ratings.

7.2 Central Tendency and Variability

Features do not necessarily have to be complex. Familiar measures of central tendency of EMA time-series, such as the mode or the mean, can be useful predictors of traditional clinical assessments. Standard measures of the variability of EMA scores, such as the standard deviation (SD), have been shown to have diagnostic value (see, e.g., Bowen et al., 2006). You should not hesitate to consider these statistics when they help you to answer your research question.

You should check, though, whether the implicit assumptions behind the statistics are correct. When you use the mean and the standard deviation to summarize a time-series, you assume that the data-generating process behind the series is stable over time - in technical terms: that this process is *stationary* (Wikipedia, 2018). If this assumption is incorrect (i.e., if the process is unstable or non-stationary), the mean and SD are biased statistics.

Our simulated EMA time-series is non-stationary, as can be seen in Figure 7.2. We observe a trend: the mood of the person increases over time. EMA scores tend to be below the mean in the beginning, and above the mean at the end. The overall mean ($M = 4.4$) overestimates the EMA scores of the first week ($M = 3.3$), and underestimates the scores of the last week ($M = 5.5$). If you use the overall mean as a feature of the time-series in your analyses, you ignore the mood improvements over the three-week period. In clinical research, you would probably not want to miss this other important feature.

When a trend exists in a series, the overall standard deviation is overestimated, because of the extra variability that is introduced by the changes of the mean. This can be checked by comparing the SD of the full series to the SD of smaller parts. In our example, the overall SD is 2.4. If we calculate the average of the SD of scores of each day, however, the estimated SD is 1.8. If the additional variability introduced by the trend is cancelled out, our estimate of variability decreases considerably.

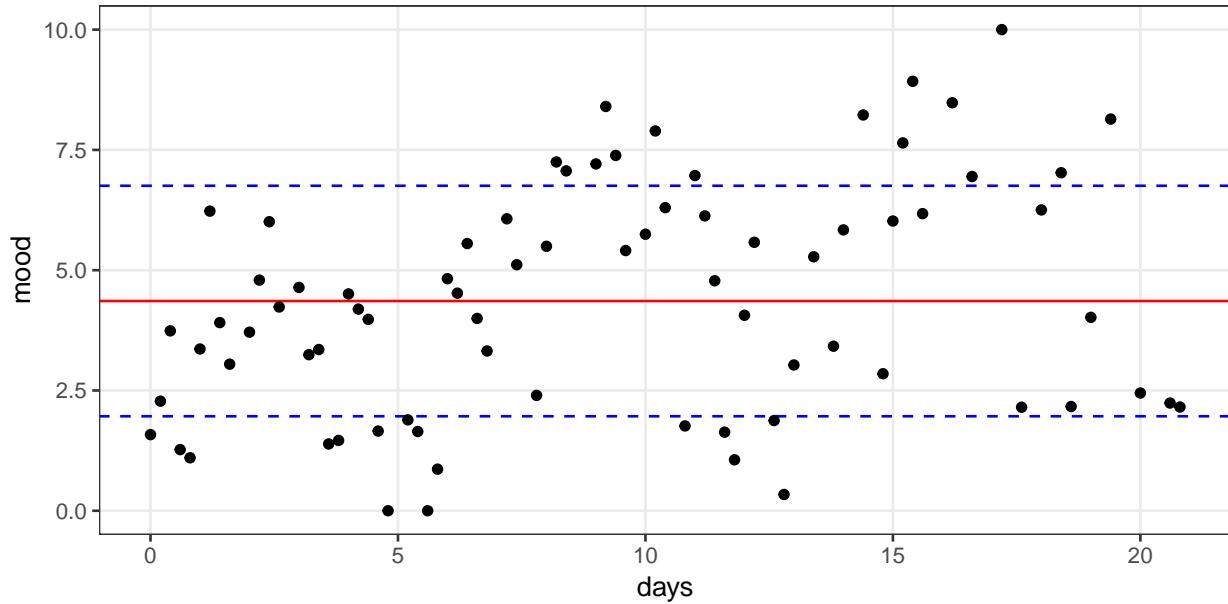


Figure 7.2: EMA time-series, with reference lines for the mean (red line) and the mean ± 1 standard deviation range (the area between the two blue lines). Both statistics are informative, but obviously do not do full justice to the variation in observations over time.

7.3 Modelling the Trend

Now that we know that the overall mean and SD are sub-optimal features of this time-series, the question is how we can do better. What other features can we extract? One option would be to characterize the series in simple regression terms, via an intercept and a slope. By doing so, we can economically describe the dynamics of the series with two numbers. As can be seen in Figure 7.3, the regression model provides a more realistic and informative summary. At the start, the estimated mean - the intercept - is 3.2. This mean increases approximately 0.9 points per week (the slope), to a final estimated mean of 5.9. Comparing Figure 7.2 and Figure 7.3, we see that Figure 7.3 is more successful in capturing the variation in the series.

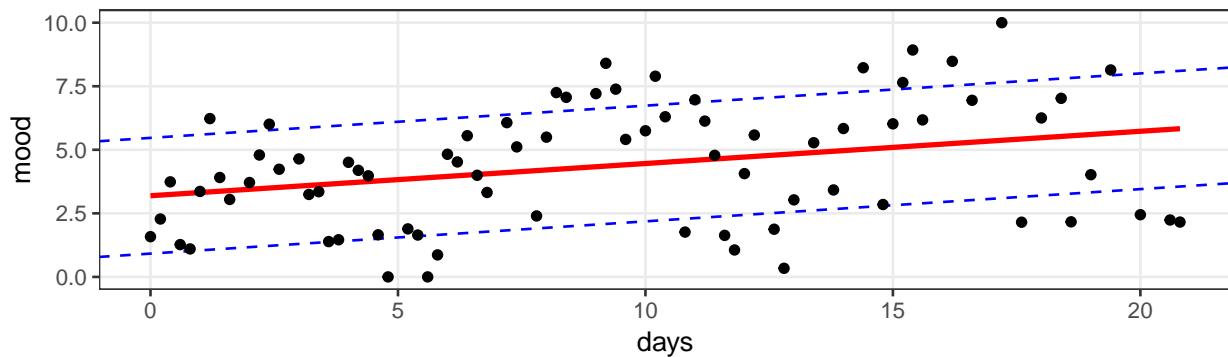


Figure 7.3: EMA time-series, with a regression reference line (red) and the residual error SD range around this line (the area between the two blue lines)

By using regression, we can also find a better estimate of the SD. Remember how the SD is conceptually

defined as the average distance between the scores and the mean. Since the mean is modeled by the regression line, we can estimate the SD by calculating the standard deviation of the residuals of the regression model that we used to retrieve the intercept and slope. For our series, this results in an estimated SD of 2.3. The new SD estimate is smaller than the overall SD, as expected. By accounting for the extra variability introduced by the trend, we get a more accurate estimate of the variability around the mean at each point in time. However, the new estimate is still higher than the daily estimate that we calculated above. Figure 7.4 provides a first indication of what is going on. In the figure, absolute regression residuals are plotted over time, with a regression line superposed. This reveals that residuals increase as time goes by. We discovered a new feature of the series: heteroscedasticity.

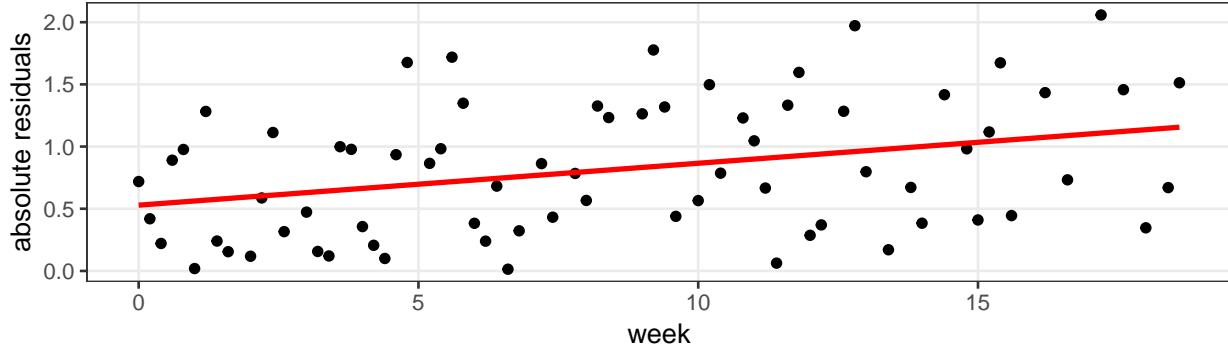


Figure 7.4: Plot of absolute regression residuals over time, revealing heteroscedasticity.

7.4 Missing Values

In EMA research, you should be prepared to deal with missing values. Especially with active EMA, in which participants have to consciously respond to prompted questions several times a day, non-response is inevitable. Missing values are typically considered to be a nuisance rather than a feature of EMA time-series. Non-response, however, is an important characteristic, that may even have clinical relevance. Missing values are also present in our simulated EMA time-series example, as you may have found out already when you inspected Figure 7.1. The points, denoting the observed responses, are denser in the beginning of the series. Over the course of the study period, the probability of missed ratings increases. This becomes immediately clear from Figure 7.5, in which the percentage of missed ratings is plotted per day. Every week, the probability of missed ratings increases approximately 16%.

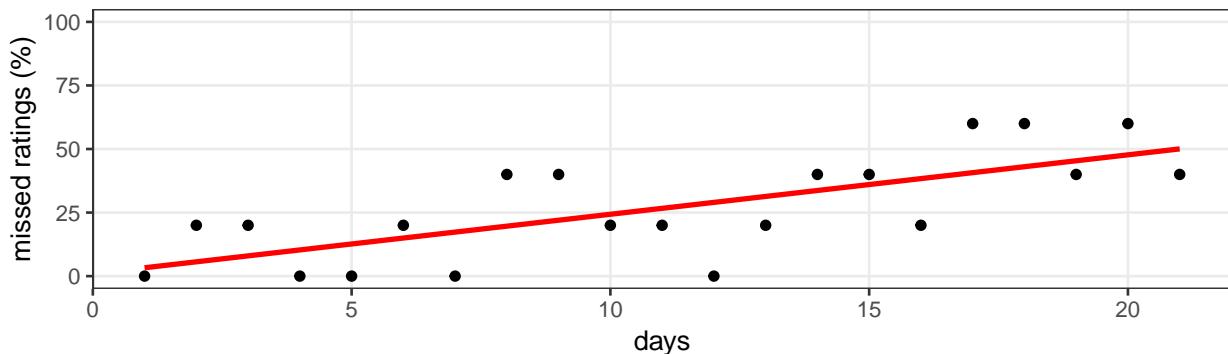


Figure 7.5: Percentage of missed mood ratings, per day, over the three-week study period, with a regression line (red) superposed, revealing a familiar trend in EMA data.

Now that we know about these missing values, the question arises how to deal with them? You can choose to ignore the missing values (assuming that the occurrence of missing values was driven by a completely random process). Alternatively, you could try to replace the missing values with plausible values (in which case it becomes important to think about the process that drives the missingness). If you decide to impute, there are several options. For example, you can replace missing values with the mean, the last-known value, an interpolated value or with a smoothing technique such as the Kalman-filter (see Hoogendoorn and Funk, 2017, for a discussion). Figure 7.6 illustrates the linear interpolation approach: missing values (marked by red dots) are replaced by values that lie between the non-missing values. Imagine that we would have used the mean to impute. Can you see why that would have disrupted local patterns in the series?

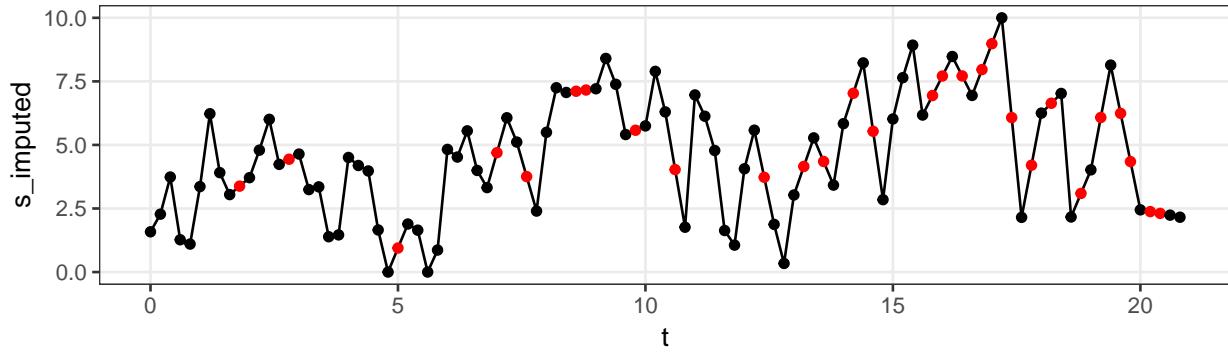


Figure 7.6: EMA time series, with missing values (red) imputed through interpolation.

7.5 Auto-correlation

When a series is correlated with delayed copies of itself, we say that it is auto-correlated. In repeated measurements of natural phenomena, this auto-correlation can often be found. The temperature of today is correlated with the temperature of recent days. Similarly, EMA mood ratings, taken at time t , are typically correlated with ratings taken at $t-1$, $t-2$, etc. Studying the auto-correlation of a time-series at different delays can be very revealing, as illustrated by the auto-correlation plot in Figure 7.7, in which the correlation of the EMA series with itself is plotted, for various delays (or ‘lags’ as these delays are called). First, the plot reveals that observations are correlated with previous values at lag 1: this series is auto-correlated. Second, there appears to be a pattern in the correlations at later lags: positive and negative correlations alternate, in a pattern that seems to reflect periodicity.

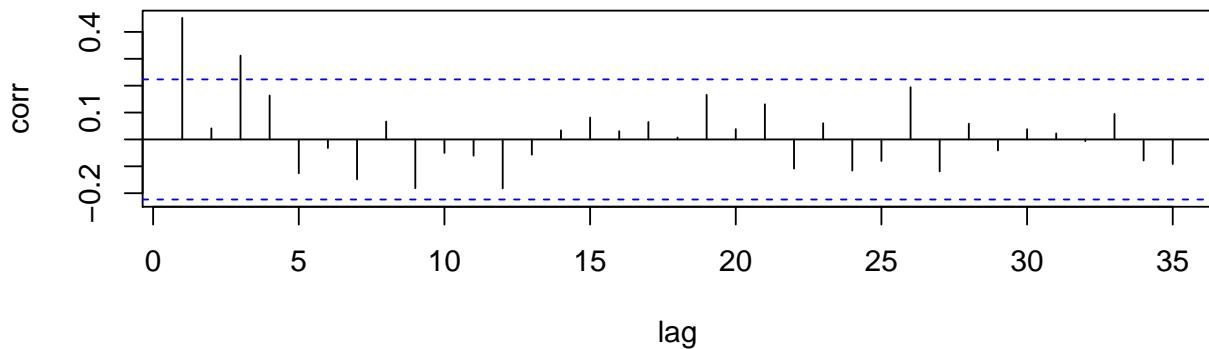


Figure 7.7: Autocorrelation plot of the EMA mood series, revealing periodicity.

7.5.1 Rolling Statistics

From the auto-correlation analysis, we learned that describing the EMA series as a simple trend might be too simplistic. The mean does not increase in a simple linear fashion. There seem to be periodic components in the series as well: the signal appears also to be characterized by a series of increases and decreases in the mean mood level. In the plot of the raw series, this is not immediately clear. However, if we “smooth” the series, by calculating the mean as it develops over time, the periodicity in the series is becoming more clear. Figure 7.8 illustrates this.

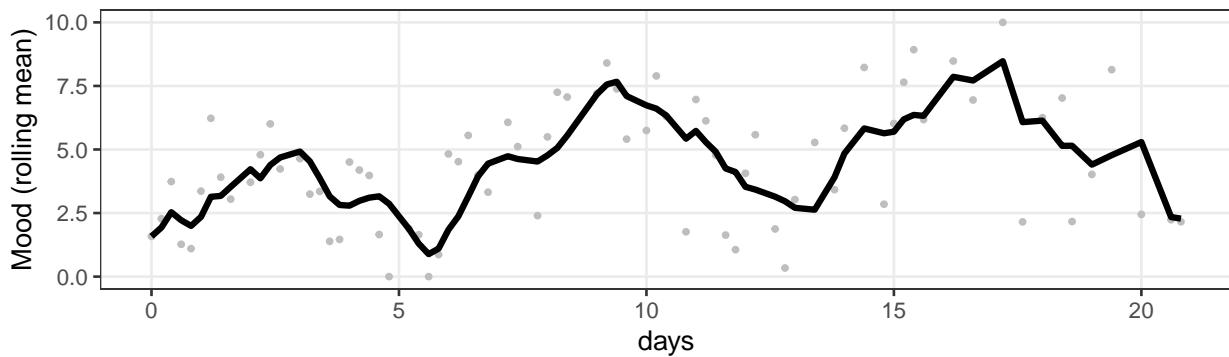


Figure 7.8: EMA mood series, with a rolling mean superposed.

7.6 Periodicity

Both the auto-correlation analysis and the moving average plot suggest periodicity in the EMA time series. Suppose we want to know more about this periodicity. Is there a way to quantify it? Yes, there is. With a technique called Fourier Analysis, the strength (‘power’) of the evidence for the presence of various frequencies can be quantified. Figure 7.9 shows what happens if we run a Fourier analysis of our time series (with R’s built-in function `spec.pgram`).

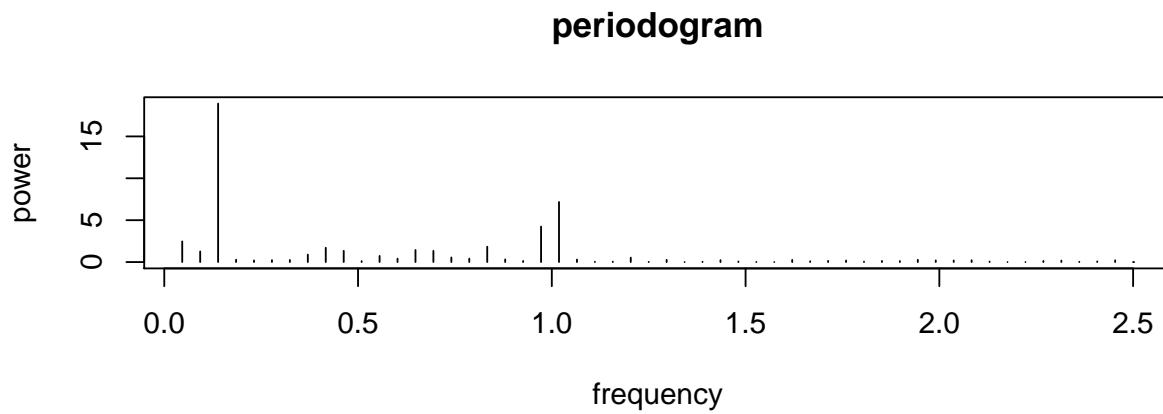


Figure 7.9: periodogram of the EMA time series, revealing a one-day and a one-week period.

There are two peaks in the figure: at a frequency of 1 (day) and at a frequency of 0.14 (about one week).

This series is *circadian*: mood ratings follow a daily and a weekly pattern. This is a common feature of EMA data, because of the circadian rhythms in our body and in our environment (our biological clocks, day-night cycles, seasons) (Doherty, 2018; Frank et al., 2000, 2007; Karatsoreos, 2014; Van Someren, 2000; Tahmasian et al., 2013).

7.7 Discussion

In this chapter, we used a number of statistical indicators to characterize a single time-series. By extracting these indicators (or “features”), we identified several regularities. We learned that the mean, variance and missingness increased over time, and we identified clear signs of circadian rhythmicity. You might have identified some of these features when you first inspected Figure 7.1. The increase in the mean was pretty easy to spot. However, the systematic increases in the variance and the missingness were less clear. Likewise, although you might have spotted the periodicity in the series, the circadian components were probably too subtle to spot for the naked eye.

The aim of the chapter was to illustrate the options available to the EMA researcher to quantitatively summarize EMA series. Other options, of course, exist. In chapter 9, for example, we will see how subtle changes in variation and auto-correlation in EMA mood scores can be summarized in a single dynamic statistical feature that may be predictive indicator of significant future state changes.

Finally, you might have been disappointed by the lack of code examples in this chapter. We did not include them, to avoid unnecessary distraction from the conceptual explanation. Instead, we provide a full code listing below, to allow you learn how to apply the feature extraction techniques yourself.

```
## (De)constructing a simulated EMA time-series

# libraries -----
library(ggplot2)
library(gridExtra)
library(psych)
library(zoo)
library(emaph)

# simulate the signal (using emaph function) -----
d <- generate_features_dataset(seed = 123)

# plot -----
e <- subset(d,!is.na(s))

# raw series
ggplot(e, aes(x = t, y = s)) +
  geom_line() +
  geom_point() +
  ylab("mood") + xlab("days") +
  ylim(0, 10) +
  theme_bw() + theme(panel.grid.minor = element_blank())

# plot mean and variance -----
ggplot(e, aes(x = t, y = s)) +
  geom_hline(yintercept = mean(e$s, na.rm = TRUE), color = "red") +
  geom_hline(
    yintercept = mean(e$s, na.rm = TRUE) + sd(e$s, na.rm = TRUE),
    color = "blue",
```

```

    linetype = 2
) +
geom_hline(
  yintercept = mean(e$s, na.rm = TRUE) - sd(e$s, na.rm = TRUE),
  color = "blue",
  linetype = 2
) +
#geom_line() +
geom_point() +
ylab("mood") + xlab("days") +
ylim(0, 10) +
theme_bw() + theme(panel.grid.minor = element_blank())

## The trends -----
## run a linear regression of mood ratings by time
fm = lm(s ~ t, e)

## mean
ggplot(e, aes(x = t, y = s)) +
  geom_smooth(method = "lm", color = "red", se = FALSE) +
  geom_abline(
    intercept = coef(fm)[1] + sd(resid(fm)),
    slope = coef(fm)[2],
    color = "blue",
    linetype = 2
) +
  geom_abline(
    intercept = coef(fm)[1] - sd(resid(fm)),
    slope = coef(fm)[2],
    color = "blue",
    linetype = 2
) +
  geom_point() +
  ylab("mood") + xlab("days") +
  ylim(0, 10) +
  theme_bw() + theme(panel.grid.minor = element_blank())

## residuals
e$residuals <- rstandard(fm)
e$week <- (e$t %% 1) + 1

ggplot(subset(e, week < 20), aes(x = t, y = abs(residuals))) +
  geom_point() +
  geom_smooth(method = "lm", color = "red", se = FALSE) +
  ylab("absolute residuals") + xlab("week") +
  theme_bw() + theme(panel.grid.minor = element_blank())

# Missing values -----

```

```

d$t_g <- cut(t, 21, labels = 1:21)
p <- prop.table(table(d$t_g, is.na(d$s)), 1)
p <- as.data.frame(p)
p <- subset(p, Var2 == TRUE)
fm2 <- lm(p$Freq ~ as.numeric(p$Var1))

ggplot(p, aes(x = as.numeric(Var1), y = Freq * 100)) +
  geom_smooth(method = "lm", color = "red", se = FALSE) +
  geom_point() +
  ylab("missed ratings (%)") + xlab("days") +
  ylim(0, 100) +
  theme_bw() + theme(panel.grid.minor = element_blank())

## imputation through interpolation
impute_interpolation <- function(x) {
  require(zoo)

  y = zoo(x)

  # fill initial and trailing NA
  if (is.na(y[1]))
    y[1] = y[which.max(!is.na(y))]
  if (is.na(y[length(y)]))
    y[length(y)] = y[max((!is.na(y)) * (1:length(y)))] 

  # interpolate
  y_ = as.numeric(na.approx(y))

  y_
}

d$s_imputed <- impute_interpolation(d$s)
ggplot(d, aes(x = t, y = s_imputed)) +
  geom_line(color = 1) +
  geom_point(aes(color = is.na(s))) +
  scale_color_manual(values = c("black", "red")) +
  guides(color = FALSE) +
  theme_bw() + theme(panel.grid.minor = element_blank())

# Autocorrelation -----
acf(e$s,
  type = "partial",
  lag.max = 35,
  main = "")

# Rolling mean -----
d$rolling_mean <-
  as.numeric(
    zoo::rollapply(

```

```

  as.numeric(d$s),
  width = 5,
  FUN = function(x)
    mean(x, na.rm = TRUE),
    fill = NA,
    align = "right",
    partial = TRUE
  )
)

e <- subset(d,!is.na(s))
ggplot(e, aes(x = t, y = rolling_mean)) +
  geom_point(aes(y = s), color = "grey", size = .8) +
  geom_line(size = 1.2) +
  ylab("Mood (rolling mean)") + xlab("days") +
  ylim(0, 10) +
  theme_bw() + theme(panel.grid.minor = element_blank())

# Frequency analysis -----
fa <-
spec.pgram(
  ts(impute_interpolation(d$s), frequency = n_measurements_per_day),
  detrend = TRUE,
  demean = TRUE,
  main = "periodogram",
  sub = "",
  xlab = "frequency",
  log = "no",
  ylab = "power",
  type = "h",
  spans = NULL,
  taper = 0.01,
  ci = FALSE
)

```

Chapter 8

Mixed Modeling

EMA data are time-series that are characterized by complex correlational structures, irregular sampling intervals, missing data, and substantive individual differences. Mixed models are well-suited to deal with these data. This chapter provides a brief introduction to conducting mixed model analysis of EMA data in R.

8.1 The Mixed Model

Mixed modeling can be understood as a regression technique in which separate regression functions are estimated for each cluster in the data set. In EMA data, these clusters are defined by the participants. Data from the same participant are expected to be correlated, and one way to honor this correlation is to conceptualize a separate regression for each participant. This idea, in the simplest regression model, can be expressed as:

$$Y_{ij} = \text{intercept}_i + \epsilon_{ij} \quad (8.1)$$

This models the expected value of the j-th measurement of participant i as the mean of all measurements of participant i, plus error. It defines a set of regression functions - one for each participant. The regression functions are, however, not independent. Mixed models divide the intercepts of the individual participant regression functions into two components: 1) the intercept of the group (intercept_g ; the mean intercept of all regression functions), and 2) a participant-specific component intercept_p (i.e., the difference between the intercept of the participant and the mean intercept), i.e.:

$$\text{intercept}_i = \text{intercept}_g + \text{intercept}_p \quad (8.2)$$

The group intercept is called the ‘fixed’ effect. If we would gather more data from new participants, we would expect to find approximately the same group intercept. The participant-specific component of the intercept is known as the ‘random’ effect. If we sample new data, we would expect a similar *variance* of the participant-specific intercept components around the group intercept. This “mixing” of fixed and random effects is what gives mixed modeling its name.

8.2 Simulating Example Data

To understand analysis techniques, it often helps to apply the technique to simulated data, in which parameters of interest are known. Here, we will use the `sim_ema` function from package `emaph`, to simulate EMA mood assessments of 100 participants, who rate their mood, three times per day, for one week. As you can learn from

the documentation of `sim_ema` (see `?sim_ema`), the function expects at least two arguments: the definition of a sample plan (see `?sample_plan`), and a specification of the data-generating model, in the form of a list defining fixed effects, the random effects, and residual variance (i.e, the error).

```
# Simulating ema data.
library(emaph)
plan <- sample_plan(n_participants = 100,
                     n_days = 7,
                     times = c("10:00-11:00",
                               "13:00-14:00",
                               "16:00-18:00"))

d1 <- sim_ema(plan,
               mm_par = list(fixed  = c(intercept = 5),
                             random = c(intercept = 1),
                             error = .5),
               lim = c(0, 10))
```

From the code, you learn that we set the mean mood (intercept_g) to 5, the variance around this mean - $\text{var}(\text{intercept}_p)$ - to 1, and the variance around these means within participants - the error - to .5. Figure 8.1 shows EMA mood ratings of the first 6 participants in the simulated data set, which we can use to check the simulation. As specified, mean mood ratings of the participants (the red lines) vary around 5 (the grey dashed line). So far, so good.

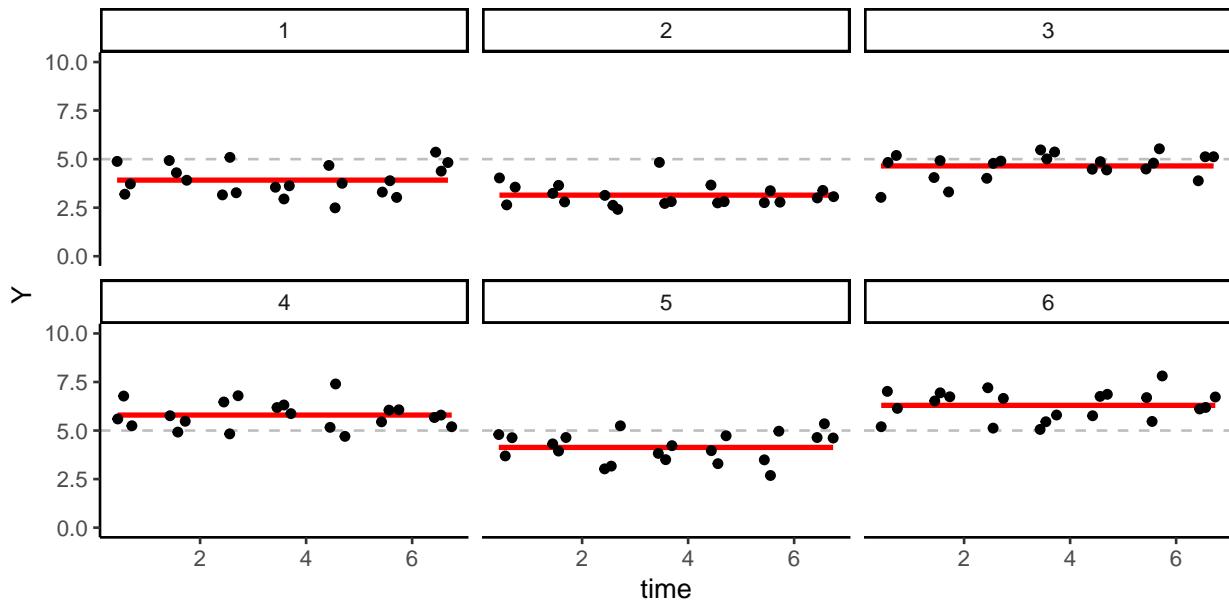


Figure 8.1: Simulated EMA data of Six Participants.

8.3 Fitting a Mixed Model in R

Now, let's fit a mixed model to the data, to see whether the simulation parameters are detected. For this, we will use the `lme` function, from package `nlme` (Pinheiro et al., 2018). The code snippet below shows how to do this.

The first argument of the `lme` function, `Y ~ 1`, specifies the fixed ‘effect’ (in this case: the mean intercept).

The second argument, `random = ~ 1 | id` specifies the random effect. In this model, intercepts are allowed to vary between participants. The fitted model is assigned to a variable (`fm`), which we will use later to study the fitted model.

```
# Fitting a mixed model with lme.
library(nlme)
fm <- lme(Y ~ 1, random = ~ 1 | id,
            data = d1)
```

We can now extract the fixed effects regression coefficients table, by calling the `summary` function on the fitted model. The estimated intercept should be around 5 (as this is a finite sample, we expect some deviation):

```
# Fixed effects.
summary(fm)$tTable
#>           Value Std.Error DF t-value p-value
#> (Intercept) 4.95     0.106 2000   46.7 1.28e-322
```

Random effects and residual variance are shown by the `VarCorr` function. Again, since we specified the data ourselves in this case, we know the ‘true’ value of these parameters: the random intercept variance should be around 1 and the residual error variance should be close to 0.5.

```
# Random effects.
VarCorr(fm)
#> id = pdLogChol(1)
#>          Variance StdDev
#> (Intercept) 1.097    1.048
#> Residual    0.502    0.709
```

It can be instructive to plot the predicted values of the model, to make clear how the model ‘thinks’. As shown by Figure 8.2, the model predicts a series of straight lines, one for each participant, that vary around 5 (the red line).

```
# Saving predicted values.
d1$predicted <- predict(fm)
```

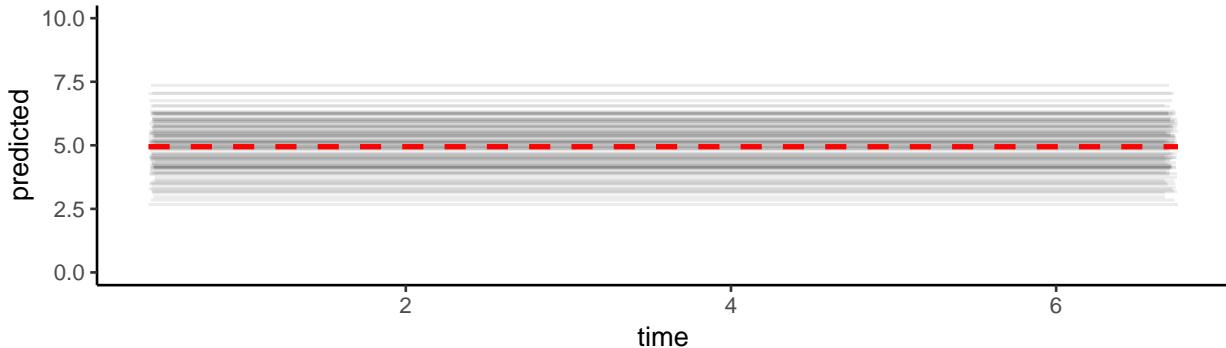


Figure 8.2: EMA ratings, of each participant in the simulated data set, as predicted by the intercept-only mixed linear model.

8.4 Adding Time as a Predictor

Now that we know how to fit a simple mixed model, we can consider a more complex scenario. In the first data set, participants’ mood ratings did not change over time. Scores varied around a stable mean during

the full week. Hence, there was no need to model a time effect. But suppose we would expect a systematic improvement of mood ratings over time, for instance in response to a mental health intervention?

Let's first call `sim_ema` again, with parameters that will result in data in which mood rating increase over the course of the week, 0.5 scale points per day. Let's also assume that individual participants will vary in the degree of mood improvement: the mean time effect will be 0.5, but this parameter is allowed to vary between participants, with a variance of 0.1.

```
# Simulating ema data (time effect).
d2 <- sim_ema(plan,
               mm_par = list(fixed = c(intercept = 5, time = 0.5),
                             random = c(intercept = 1, time = 0.1),
                             error = .5),
               lim = c(0, 10))
```

Figure 8.3 shows the data of the first six participants in the second data set. Both the intercept and the slope vary across the participants. Some participants improve more over time, and others improve less: the slope in this data set is a random effect.

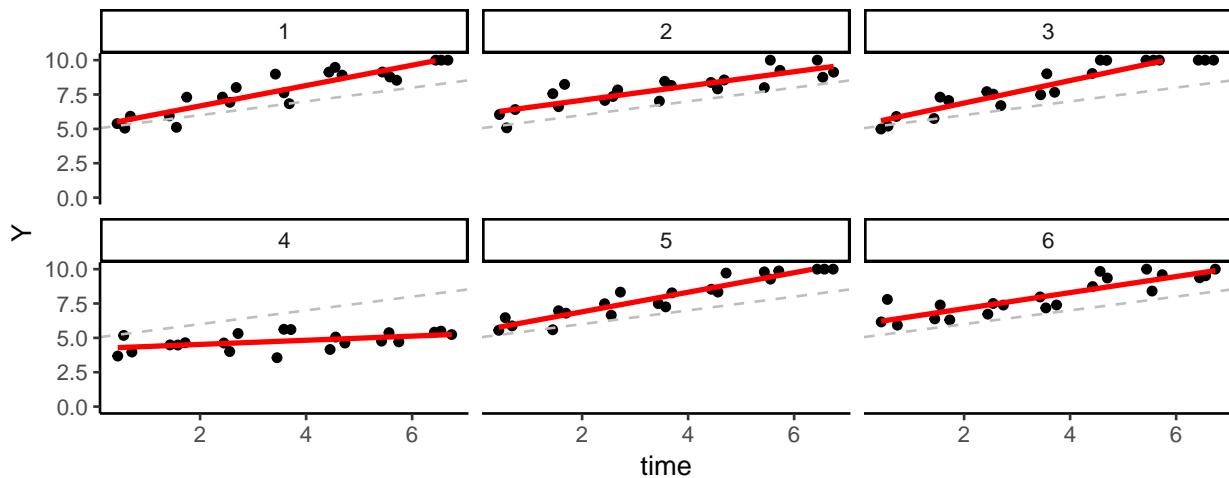


Figure 8.3: Simulated EMA data of Six Participants (Time-varying model).

To fit the extended mixed model, time can simply be added to both the fixed and random arguments of the ‘`lme`’ function. Fixed effects estimated of this model should be around 5 and 0.5, since that is how we specified the data. Calling `summary` on this function, we see that the fixed time effect is significant.

```
# A mixed model, with a random slope.
library(nlme)
fm <- lme(Y ~ 1 + time, random = ~ 1 + time | id,
            data = d2)
summary(fm)$tTable
#>           Value Std.Error   DF t-value p-value
#> (Intercept) 5.112    0.1141 1999    44.8 4.94e-304
#> time         0.467    0.0263 1999    17.7  1.52e-65
```

The random effects now have four components: the variance of the intercept, the variance of the slope, the residual error *and* the correlation between the random intercept and the random slope.

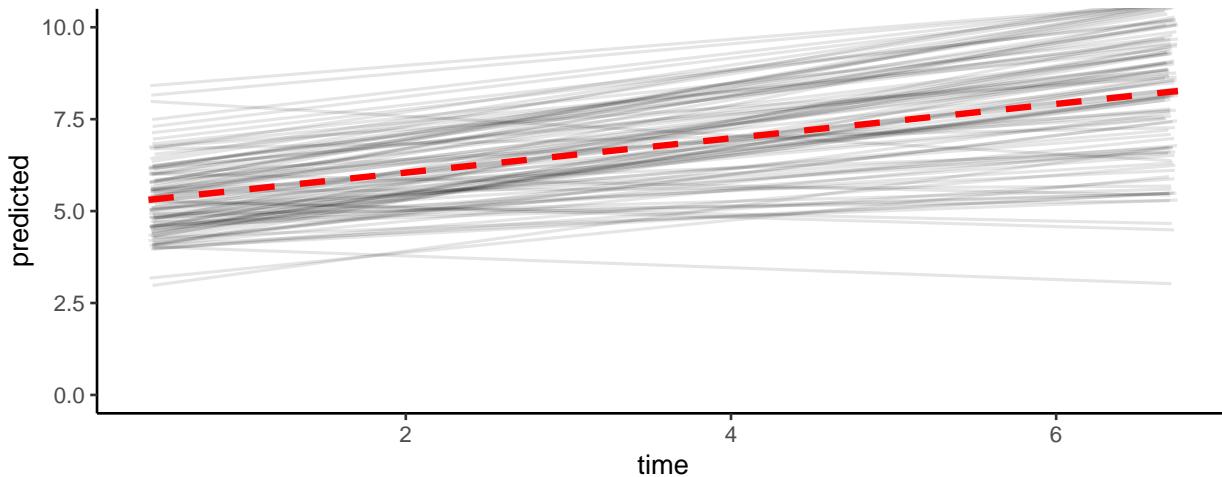
```
# Extracting random effects.
VarCorr(fm)
```

```
#> id = pdLogChol(1 + time)
#>          Variance StdDev Corr
#> (Intercept) 1.2071   1.099  (Intr)
#> time        0.0637   0.252 -0.231
#> Residual    0.4781   0.691
```

Model predictions clearly show how the mixed model estimated varying intercepts and slopes, that, on average, approximate the fixed effect regression formula $Y = 5 + 0.5 * \text{time}$ that was used to generate the data.

```
d2$predicted <- predict(fm)

ggplot(d2, aes(x = time, y = predicted, group = id)) +
  geom_line(alpha = .1, size = .6) +
  geom_smooth(aes(group = NULL), method = "lm", color = "red", size = 1.2, linetype = 2) +
  coord_cartesian(ylim = c(0, 10)) + theme_classic()
```



8.5 Adding a Two-Group Comparison

In data-set 1, mood ratings did not change during the week, while in data-set 2, the mood ratings increased. Suppose the two data-sets reflect the data that you collect in a two-group RCT, in which you compare the effects of a mental health intervention (data-set 2) against a waiting list condition (data-set 1). By combining the two data-sets, we can illustrate how to conduct a group comparison with `lme`.

Since the two data-sets are already available (in variables `d1` and `d2`), the new data set can be created with just three lines of code (below). In the first line, the `rbind` function is used to combine the rows of data-set 1 and 2 into a new variable: `d3`. The second line adds a group indicator to `d3`. The third line updates the IDs of the participants in the second group, to differentiate the participants in the second group from the participants in the first group.

```
# Two-group simulation.
d3 <- rbind(d1, d2)
d3$group <- factor(c(rep(0, nrow(d1)), rep(1, nrow(d2))), labels = c("control", "treatment"))
d3$id[d3$group == "treatment"] <- d3$id[d3$group == "treatment"] + 100
```

The effect of the intervention can be tested by adding a (fixed) ‘`time * group`’ interaction effect to the model. This effect, we know, is 0.5, and, as can be seen, this is what the model picks up:

```
# A mixed model, with two groups.
library(nlme)
fm <- lme(Y ~ 1 + time * group, random = ~ 1 + time | id,
            data = d3)
round(summary(fm)$tTable, 2)
#>           Value Std.Error   DF t-value p-value
#> (Intercept) 4.93     0.11 3998  43.98   0.00
#> time         0.00     0.02 3998    0.21   0.83
#> group:treatment 0.18     0.16 198    1.14   0.25
#> time:group:treatment 0.46     0.03 3998   16.71   0.00
```

In Figure 8.4 below, EMA mood ratings predicted by the fitted model show how the model detects 1) the fixed between-group effect, and 2) the variance in intercepts and slopes in both groups.

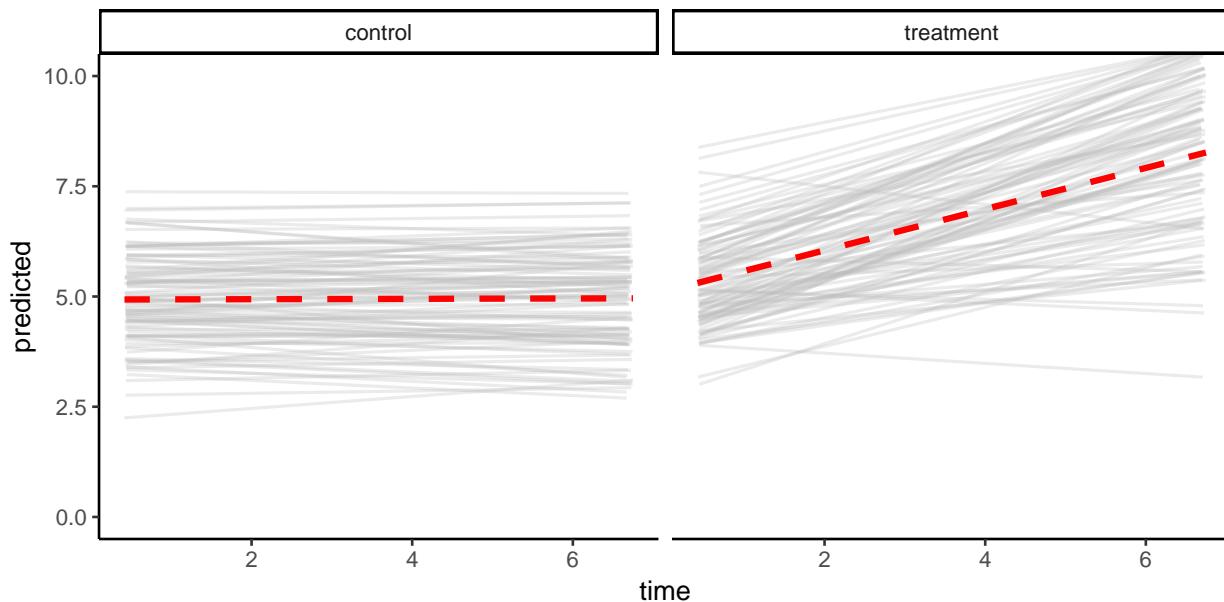


Figure 8.4: Predicted mood ratings

8.6 Next Steps

In this chapter, we introduced mixed model analysis of EMA data in R. To do so, we could only touch upon the theoretical foundations of mixed models, and we deliberately used simple examples with clean simulated data. Readers, who consider the application of mixed models, are strongly advised to study additional resources.

The authoritative reference for mixed effect modeling in R is a book by Pinheiro and Bates (2000). To fully appreciate this book, however, a strong background in mathematical statistics is required. Thorough introductions in the topic are further found in the work of Prof. dr. Jos Twisk (Twisk, 2006, 2013), who also teaches an applied mixed models course at the Department of Epidemiology and Biostatistics at the Vrije Universiteit Medical Center (see: <http://epidm.nl/en/courses/mixed-models/>).

Part V

A Case Study

Chapter 9

Early Warning Signs of Depression

One of the promises of EMA is that it might detect signs of mental health deterioration in an early stage. Subtle changes in time series of mood variables, for example, might signal a depression relapse. If we can detect these changes, preventive interventions can be triggered to avoid the relapse.

But what changes, exactly, should we look for? What are these early warning signs?

9.1 Critical Slowing Down

Critical Slowing Down (CSD) is a concept from dynamic systems theory. In dynamic systems, state transitions are preceded by a change in which the system reacts to disturbances. In a stable state, the system quickly recovers from disturbances. Prior to a transition to a new state, however, the system takes more and more time to recover back to its current state: the variance and auto-correlation of the system increases (Scheffer et al., 2009; Dakos et al., 2008).

In this chapter, we re-analyze data from a study that aimed to demonstrate CSD in EMA-data of a single patient with a history of major depression (Groot, 2010, Kossakowski et al. (2017); Wichers et al., 2016). The patient, a 57-year old male, used EMA to monitor himself during a 239-day single-case double-blind medication reduction trial. During the experiment, he experienced a relapse, and Wichers and colleagues showed how variance and auto-correlation in the EMA data increased, several weeks prior to this relapse. The transition appeared to be preceded by CSD.

We will try to reconstruct the finding, using an alternative analysis strategy. One of the limitations of the Wichers et al analysis is that auto-correlation was analyzed at lag 1 only (i.e., only the correlation between t and $t-1$ was considered). With another analysis technique, called ‘Detrended Fluctuation Analysis’, all lags can be considered.

To conduct the analysis, we need three R packages:

- Raw EMA data of this study were published in the public domain (Kossakowski et al., 2017). We included the data in the `emaph` package.
- To manipulate the raw data and reconstruct the plots of the article, we are going to use several functions from `tidyverse` packages.
- DFA is implemented in package `nonlinearTseries`, so we will need that as well.

```
# Required libraries for the CSD-study re-analysis.  
library(emaph)  
library(tidyverse)  
library(nonlinearTseries)
```

9.2 Plotting the Course of Depression

Let's take a look at the development of the depressive symptoms of the patient first. These were tapped with weekly assessments of the depression scale of the 'Symptom Checklist-90-Revised' (SCL-90-R; Derogatis, 1994), a well-established self-report questionnaire.

The code below reconstructs Figure 1 of the Wichers et al article. It plots the SCL90-R depression scale score over the time. Around day 120, the depression score increased considerable: the patient experienced a relapse.

```
# Plot depression score.
dep <- csd %>% select(dayno, scl90r_dep) %>%
  filter(!is.na(scl90r_dep)) %>% unique

# plot dep + change point (day 120 in our data)
ggplot(dep, aes(x = dayno, y = scl90r_dep, group = 1)) +
  geom_step() +
  xlab("Days") + ylab("SCL-90-R Depression") +
  theme_classic()
```

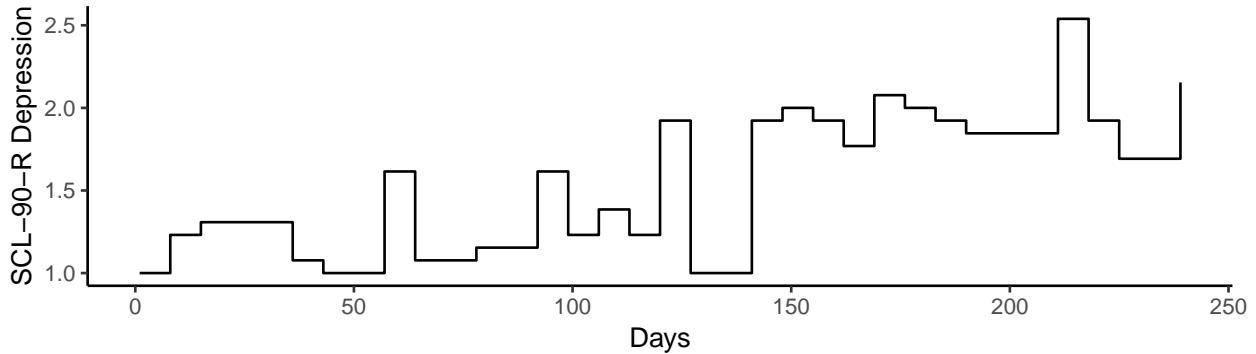


Figure 9.1: SCL-90 depression score, over the study period

9.3 Mental state EMA Items

Wichers and colleagues selected 13 items from the full EMA data set, which they grouped in 5 factors: positive affect (pa; 4 items), negative affect (na; 4 items), mental unrest (mu; 3 items), suspiciousness (su; 1 item), and worrying (wo; 1 item). From these factors, an overall mental state sum score can be calculated.

```
# Mood states calculation.

# positive affect
pa_items <- c("mood_enthus", "mood_cheerf",
             "mood_strong", "mood_satisfi")

csd$pa <- csd %>%
  select(pa_items) %>%
  rowMeans(., na.rm = TRUE)
csd$pa <- -csd$pa

# negative affect
na_items <- c("mood_lonely", "mood_anxious",
```

```

    "mood_guilty", "mood_doubt")

csd$na <- csd %>%
  select(na_items) %>%
  rowMeans(., na.rm = TRUE)

# mental unrest
mu_items <- c("mood_irritat", "pat_restl",
               "pat_agitate")
csd$mu <- csd %>%
  select(mu_items) %>%
  rowMeans(., na.rm = TRUE)

# 'single-item' states
csd$su <- csd$mood_suspic
csd$wo <- csd$pat_worry

# global mental state score
csd$ms <- rowSums(csd[c("pa", "na", "mu", "su", "wo")])

```

Rows, in which one or more of the factors have missing values, are removed from the analysis. In a full analysis, options to impute the missing values could and should be considered. However, since only 3 of the 1476 rows have missing item scores, not much is probably lost by running a simple complete-case analysis.

```

# Missing values removal.

# count number of items with missing items, per row
csd$nna <- csd %>%
  select(matches("mood_")) %>%
  is.na(.) %>% rowSums

# drop rows with missing values
csd <- csd %>% filter(nna == 0)

```

9.4 Running the DFA

We are now ready to run the DFA analysis. We split the full series in 31-day overlapping windows, in steps of 1 day (i.e., day 1-31, day 2-32, etc), calculate the DFA of each window and save that value so that we can compare it to the weekly depression assessments.

```

# Running the DFA.

# prepare result rows: one row for each day
d <- NULL
d <- csd %>%
  group_by(dayno) %>%
  summarise(ms = mean(ms))
d$ms_dfa = NA

# determine DFA, in a moving window of 31 days
window <- 31
for (i in seq(window, max(csd$dayno), 1)) {

```

```

# get the sliding window data
w <- subset(csd, dayno > (i - window) & dayno <= i)

# dfa: ms
dfa.analysis <- dfa(time.series = w$ms,
                      npoints = 30,
                      window.size.range = c(10, nrow(w)),
                      do.plot = FALSE)

fgn.estimation <- estimate(dfa.analysis,
                             do.plot = FALSE,
                             fit.col = "blue", fit.lwd = 2, fit.lty = 2,
                             main = "Fitting DFA to fGn")

d$ms_dfa[d$dayno == i] <- fgn.estimation
}

```

9.5 Results

We can now plot the DFA indicator over time, to see whether it peaks prior to the onset of the relapse. As can be seen, the DFA-indicator clearly peaks prior to the onset of the relapse, around day 110 (where the color changes to red). Interestingly, a second peak is reached at the end of the experiment, around day 239, indicating - perhaps - a recovery from the relapse.

```

# Plot DFA results.
ggplot(na.omit(d),
       aes(x = dayno, y = ms_dfa,
            colour = factor(dayno < 120),
            group = 1)) +
  geom_point() +
  geom_line() +
  ylab("dfa (60-day window)") +
  xlim(c(0, 250)) +
  guides(color = FALSE) +
  theme_classic()

```

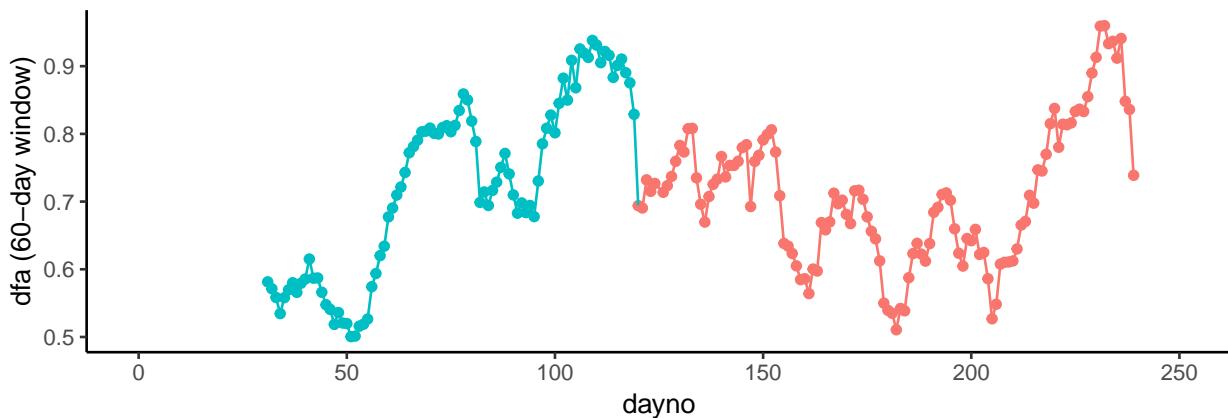


Figure 9.2: Results of the DFA analysis.

9.6 Discussion

Our re-analysis replicates the main finding of the Wichers et al article ((Wichers et al., 2016)): several weeks prior to a depression relapse, as predicted by complex systems theory, the variance and auto-correlation in EMA mood ratings increased.

Potential clinical applications, of course, are clear. If clinically relevant changes can be predicted algorithmically from EMA ratings, automated patient feedback systems could help to prevent pending deterioration or consolidate the path towards recovery.

One swallow, however, does not make summer. Yes, in this case, the DFA-indicator peaked prior to the relapse. This could be a mere coincidence. The predictive power of CSD needs to be confirmed in larger samples and prospective studies. Given the theoretical background, successful applications of CSD in other domains, and the present finding, the spending of resources on such studies seems warranted.

Important additional questions remain to be answered. When it predicts a state change, what is the expected delay between this prediction and the change? Does a critical DFA-value exist? Given that critical value, are false positive and false negative rates of this prediction acceptable? These are important questions that should be answered before any clinical application of DFA can be considered.

Re-analysis of data from completed clinical studies, in which EMA data were collected, may be a first step to further explore the value of CSD. One option, for example, would be to re-analyze data from the E-COMPARED study (Kleiboer et al., 2016). In this trial, patients, who were treated for major depression, completed weekly self-report questionnaires (the Patient Health Questionnaire; PHQ-8, Kroenke et al., 2009) and daily EMA mood ratings throughout treatment, which lasted up to 16-week. Since CSD is an indicator of *any* state change (i.e., irrespective of whether the change is clinically “good” or “bad”), theory would predict a (lagged) relationship between CSD (i.e., the DFA-score) and clinically significant changes in PHQ-scores (Jacobson and Truax, 1991).

Part VI

EMA Catalogues

Chapter 10

EMA Research within the APH Mental Health Consortium

This chapter summarizes recent EMA research projects within the APH Mental Health consortium, as a guide to other researchers looking for nearby EMA-expertise and research collaboration. An overview of identified projects is presented in Table 10.1. Detailed summaries are presented below.

Table 10.1: Overview of EMA research in the APH MH consortium.

Focus / PI	Project	APH MH Organization
Depression		
Bockting	Imagine Your Mood	AMC
	Stay Fine	
Penninx and Lamers	NESDA	VUmc / GGZ inGeest
	RADAR-CNS	VUmc / GGZ inGeest
Riper and Smit	ICT4Depression	VU / GGZ inGeest
	E-COMPARED	VU / GGZ inGeest
Psychosomatic symptoms		
Knoop	Chronic Fatigue study	AMC
Snoek	MERITS	VUmc
Sprangers	FAntasTIGUE	AMC
	IMPACT	AMC
Psychotic symptoms		
Van der Gaag	TemStem	VU
	VRETp	VU
Suicidal ideation		
Van Ballegooijen	CASPAR	VU
Sleep		
Van Someren		AMC / GGZ inGeest
Stress & Emotion		
De Geus	VU-AMS	VU

10.1 Depression

10.1.1 Bockting

Prof. dr. Claudi Bockting is affiliated with the AMC psychiatry department and the Rijksuniversiteit Groningen. Her group focuses on etiology, treatment and relapse prevention of (severe) depressive disorder and related common mental health disorders. In two studies of the group, EMA measures are included.

Aspect	Description
Project team	Claudi Bockting, PhD, in collaboration with Maaike Nauta, PhD (RuG), Yvonne Stikkelbroek, PhD (Universiteit Utrecht/ GGZ Oost Brabant); Gerard van Rijsbergen, PhD (GGZ Drenthe)
APH site	AMC
Project(s)	1) ‘Imagine Your Mood’ study (IYM; 2012-2017; Slofstra et al. (2017)): This was a three-arm RCT evaluating several relapse prevention strategies in remitted depressed patients. 2) STAY-FINE study (ongoing; http://stayfine.nl/): this RCT aims to evaluate a smartphone-based intervention (EMI) to prevent relapse in young people (12-23 yrs.) with remitted anxiety or depression. The study is funded by ZonMW (see http://tinyurl.com/yb2qr6ro).
EMA Measures	In the IYM RCT, EMA items included mood, positive affect (PA, 5 items), negative affect (NA, 9 items) and mental imagery (2 items). Participants were prompted 10 times per day, 3 days a week, over a period of 8 weeks. Prompts were sent randomly within fixed time-intervals. The VAS mood scale used (<i>‘Please rate your current mood on a scale of 0 to 100’</i> , on which 0 indicates ‘happy’, and 100 indicates ‘sad’), was shown to have a high positive predictive value without any false negatives at a cut-off score of 55 (Van Rijsbergen et al., 2014). Compared to the HAM-D17 and the IDS-SR, the VAS also was a better predictor of current relapse status, as measured by the SCID-1 interview (variance explained for VAS: 60%; HAM-D17: 49%; IDS-SR: 34%).
Platforms used	In the STAY-FINE RCT, The aim is to test whether EMA + EMI is more effective in preventing relapse in comparison to EMA alone. Structured clinical interviews will be conducted to assess the clinical status of participants. EMA will be used to monitor outcomes (for three weeks) at six assessment periods (between baseline and 36-month follow-up), in N = 212 participants. Polar watches will be used to collect passive activity measurement data. Prior to the main study, a feasibility study of the app will be conducted.
Contact	Psymate, Roqua (see also Chapter 11), ‘Imagine your mood’ app & STAY-FINE app (custom development), Polar watches (http://www.polar.com/nl). http://www.amc.nl/web/research-75/person-1/prof.-dr.-c.l.h.-bockting-phd.htm http://claudibockting.com http://stayfine.nl/

10.1.2 Penninx and Lamers

Prof.dr. Brenda Penninx and **Dr. Femke Lamers** are affiliated with the Department of Psychiatry, VU University Medical Center and GGZ inGeest. Both are involved in two EMA studies: an EMA study in the context of the Netherlands study of Depression and Anxiety (*NESDA*; www.nesda.nl), of which Penninx is primary investigator, and a large international EU H2020-funded collaborative research project, in which a variety of active and passive measures are used (the ‘Remote Assessment of Disease and Relapse - Central Nervous System’-project; *RADAR-CNS*; <http://www.radar-cns.org>).

10.1.2.1 NESDA EMA study

Aspect	Description
Project team	Femke Lamers, PhD, Brenda Penninx, PhD (VUmc), Harriette Riese, PhD (UMCG), Robert Schoevers, PhD (UMCG), the NESDA consortium
APH site	VUmc, GGZ inGeest, Amsterdam
Project(s)	The main aim of NESDA is to examine the long-term course and prognosis of anxiety and depression (Penninx et al., 2008). Within the 9-year follow-up of the NESDA study (2015 - 2017), NESDA participants (n = 384) were invited for an EMA diary/actigraphy study.
EMA measures	NESDA EMA study data were collected to explore the dynamic interplay between cognitions, emotions, behavior and environment in daily life of depressed versus non-depressed participants. EMA participants started within 31 days after the regular NESDA assessments (face-to-face interview and self-report measures). Participants could use their own smartphone (Android / iPhone) if they had sufficient data or access to WIFI for at least 80% of the two-week time period, and could borrow a phone if needed. To record the amount of physical activity during EMA monitoring, participants were also asked to wear an accelerometer (GENEActiv, 30Hz) on their non-dominant wrist, 24 hours a day, for two weeks. Recording started on the evening prior to the first EMA assessment and continued until the morning after the last assessment. Participants were prompted five times a day to complete a self-report questionnaire comprising ~30 items. Items (7-point Likert scale, ranging from ‘not [at all]’ to ‘very’) tapped 1) mood and cognition, 2) physical conditions, 3) social context, 4) sleep, 5) daily uplifts/hassles, 6) activities, and 7) substance use. Examples of items were: “ <i>I feel down</i> ”, “ <i>I feel cheerful</i> ” and “ <i>Where are you now</i> ” (e.g. at the neighbor’s house). In addition, there was an additional questionnaire inquiring about questionnaire burden and an open-ended question for general comments on circumstances that might have influenced answers. The items on Valance, Arousal and Current State have been used in a previous study; the Uncovering the Positive Potential of Emotional Reactivity (UPPER) study (Bennik, 2015). The other items are based on items from earlier EMA studies, such as the work by Mehl and colleagues (Mehl and Conner, 2012), van Os and colleagues (Wichers et al., 2012) and studies performed at the Interdisciplinary Center of Psychopathology and Emotion regulation (ICPE), such as the Mood and Movement in Daily life (MOOVD) study (Booij, 2015)
Platform used	RoQua (http://www.roqua.nl/), GENEActiv accelerometer. See also Chapter 11).
Contact Information	http://research.vumc.nl/en/persons/brenda-penninx http://research.vumc.nl/en/persons/femke-lamers

Aspect	Description
	http://www.nesda.nl/

10.1.2.2 RADAR-CNS

Aspect	Description
Project team	Femke Lamers, PhD; Brenda Penninx, PhD; Sonia Difrancesco, MSc, and the RADAR consortium.
APH site	VUmc, GGZ inGeest
Project	The European (EU H2020-IMI) <i>RADAR-CNS</i> project (Remote Assessment of Disease and Relapse - Central Nervous System; http://www.radar-cns.org/) aims to study the potential of wearable devices and smartphone technology to help prevent and treat depression, multiple sclerosis and epilepsy. The project, which started in 2016, was designed to examine how remote measurement technologies can monitor and improve quality of life and psychological well-being of patients. Within this project, the VUmc/GGZ inGeest research site focuses on depression (RADAR-MDD), in collaboration with King's College, London (KCL). Technical goals of the project are to 1) build an end-to-end open source system with generalized data aggregation capabilities, and 2) explore (technological aspects of) big data solutions. Clinical goals of the project are to 1) assess the feasibility of continuous monitoring of patients, and 2) predict disease onset or relapse with big data prevention and risk assessment approaches.
EMA Measures	Data are collected during a 2-year period, in which EMA will be activated every 6 weeks for 6 consecutive days. Active EMA outcomes include (variability in) sleep quality, activity, social interactions, mood, and stress as possible predictors of the clinical course of participants. Passive EMA measures include location and movement (GPS; actigraphy), skin temperature, galvanic skin conductance, heart rate (-variability), voice recognition, social interaction data (call/message logs), and smartphone app usage patterns.
Platform(s)	In the RADAR project, an comprehensive EMA platform is developed (the RADAR RMT application; see http://thehyve.nl/cases/radar-cns and http://github.com/RADAR-base); This platform focuses on classes of data rather than specific devices, to enhance modularity and adaptability, as new devices become available. The platform will be used to integrate and control data streams from an EMA smartphone app and several wearable device (including the Empatica E4 Wristband, Pebble 2 Smart-watch, Biovotion VSM, Faros 180, and Fitbit).
Contact	http://research.vumc.nl/en/persons/femke-lamers http://research.vumc.nl/en/persons/brenda-penninx http://www.radar-cns.org/

10.1.3 Riper and Smit

Prof. dr. Heleen Riper is affiliated with the department of clinical, neuro- and developmental psychology of the Vrije Universiteit and specialized mental health organization GGZ inGeest, Amsterdam. *Prof. dr. Jan Smit* is affiliated with the department of psychiatry of the Vrije Universiteit Medical Center (VUmc) and GGZ inGeest, Amsterdam. Riper and Smit have been driving forces behind the development of *Moodbuster*

(<http://www.ict4depression.eu/moodbuster/>), a research platform for the delivery of online and blended psychotherapy, which has built-in, integrated EMA functionality.

Aspect	Description
Project team	Heleen Riper, PhD, Jan Smit, PhD, Lise Kemmeren (GGz InGeest) and the Moodbuster/E-COMPARED consortium
APH site	Vrije Universiteit Amsterdam; GGZ inGeest (and other organisations)
Project(s)	Moodbuster was developed in the European FP7 project “ICT4Depression” (Warmerdam et al. (2012); http://www.ict4depression.eu/), and applied in the Horizon 2020 FP7 EU-project E-COMPARED [European COMPARative Effectiveness research on blended Depression treatment versus treatment-as-usual; Kleiboer et al. (2016); Kemmeren et al. (2016); http://www.e-compared.eu/], Moodbuster will also be used in the EU ImpleMentAll project (http://www.implementall.eu) and in several other clinical trials that are currently (October 2018) in preparation.
EMA measures	In the ICT4Depression project, the Moodbuster platform included a complex set of active and passive EMA measures, including physiological sensors, accelerometers, wearables to measure sympathetic nervous system responses (chest strap and glove), and an Android EMA smartphone app. For the E-COMPARED study, the Moodbuster website and the smartphone app were adapted to deliver (blended) treatment to patients with MDD, in 5 routine practice settings across Europe. A therapist portal was added, in order to allow therapists to monitor patients’ progress and send feedback messages. Smartphone-based EMA was used to asses sleep, mood, worrying, self-esteem, activities (2 items) and social contacts. Prompts were sent at a random time point between 10:00 and 22:00. At the beginning and during the final phase of treatment, patients received two additional prompts per day for one week. In the morning (around 10:00), sleep, worrying and self-esteem items were assessed. In the evening (around 22:00), these questions were repeated, along with the activity and social interaction items. Patients rated these EMA items on a visual analogue scale, ranging from 0 (low) to 10 (high). EMA data were used in several machine learning projects (Mikus et al., 2018; Rocha et al., 2018; Van Breda et al., 2018; Van de Ven et al., 2017).
Platform used	Moodbuster (http://www.moodbuster.eu ; see also Chapter 11)).
Contact	http://research.vu.nl/en/persons/heleen-riper http://research.vumc.nl/en/persons/jan-smit http://research.vumc.nl/en/persons/lise-kemmeren

10.2 Psychosomatic Symptoms

10.2.1 Knoop

Prof. dr. Hans Knoop is professor of evidence-based psychological and behavioral interventions for medical conditions and somatic symptoms at the department of Medical Psychology of the Amsterdam Medical Center (AMC). In 2015/2016, Knoop and colleagues ran an EMA study at the Expert Centre for Chronic Fatigue (ECCF; <http://nkcv.nl/>).

Aspect	Description
Project team	Margreet Worm-Smeitink, MSc; Hans Knoop, PhD, in collaboration with Judith Rosmalen, PhD, Anne van Gils, MSc, Rei Monden, PhD, and others

Aspect	Description
APH site	AMC
Project	The <i>CFS EMA study</i> was designed to investigate - through time-series analyses - whether there are differences in perpetuators of fatigue between individual patients.
**EMA measures	New patients attending the ECCF (n = 102) were asked to complete an e-diary, 5 times a day, for 2 weeks. The times were fixed in consultation with the participant, with a 3-hour break between each assessment. The e-diary assessed fatigue, pain, anxiety, depression, activity (physical, mental, social), patients' focus on fatigue, fatigue catastrophizing, self-efficacy, fear avoidance, and social incomprehension. Participants also wore an accelerometer (actigraphy) during the period when the self-reports were collected. The R 'auto-var' package was used to conduct network analyses. Results are forthcoming. Identified determinants of fatigue will be used to personalize treatment of CFS-patients.
Platform used	RoQua (http://www.roqua.nl ; see Chapter 11)). Participants received a link to the (web-based) e-diary via an SMS to their personal smartphone.
Contact	http: /www.amc.nl/web/research-75/publications/prof.-dr.-j.a.-knoop.htm http://nkcv.nl/onderzoek/expert-centre-chronic-fatigue/

10.2.2 Sprangers

Prof. dr. Mirjam Sprangers is professor at the Department of Medical Psychology, Academic Medical Center (AMC), University of Amsterdam. She coordinates a research line on Quality of Life (QoL) with focuses on patient-reported outcomes in somatic settings. Currently, she and her research group are involved in two projects that involve EMA: The *FAntasTIGUE study*, which focuses on fatigue in patients with chronic obstructive pulmonary disease and the NWO-funded *IMPACT project* (<http://www.impactonderzoek.nl/>) which targets patients undergoing a cardiac intervention.

10.2.2.1 FAntasTIGUE

Aspect	Description
Project team	Mirjam Sprangers, PhD, in collaboration with Martijn Spruit, PhD (PI); Yvonne Goërtz, MSc; Zjala Ebadi (from July 2018), MSc, Melissa Thong, PhD, Daisy Janssen, MD, PhD, Jeanette Peters, PhD, Jan Vercoulen, PhD, Chris Burtin, MSc, Yvonne Meertens-Kerris, Arnold Coors, Jean Muris, MD, PhD, Emiel Wouters, MD, PhD, Judith Prins, PhD, and Martijn Spruit, PhD
APH site	AMC (in collaboration with Ciro-Horn, Maastricht UMC, Radboud UMC, and Hasselt University)
Project goals	The <i>FAntasTIGUE</i> study examines fatigue in patients with chronic obstructive pulmonary disease (COPD; n = 400), by evaluating the course of fatigue, precipitating/perpetuating factors and hospitalization. A secondary aim is to identify diurnal fluctuations in fatigue, by using EMA in a sub-sample of participants (n = 40) (Goërtz et al., 2018).

Aspect	Description
EMA measures	The project comprises four data collection periods (baseline, and 4, 8, and 12-month follow-up). During data collection periods, patients are provided with iPods, on which an EMA application will be pre-installed. Participants will be prompted 8 times a day, at random moments between 7:30 and 22:30, for 5 consecutive days, to answer 19 items (including nine contextual items). Measured concepts include fatigue, relaxed feeling, breathlessness, agitation, uncertainty, irritation, satisfaction, anxiety, feeling energetic, and feeling mentally fit. Items are rated on a 7-point Likert-scale, ranging from 'Not at all' to 'Very much'. In addition, participants are asked to complete a morning questionnaire soon after they awaken to assess the quality of their sleep the previous night. Participants also complete an evening questionnaire, assessing the general perception of their day just before going to bed. During the EMA data collection period, patients will also be asked to wear an actigraph (activity monitor), 24 hours a day, for one week.
Platform used	Psymate (http://psymate.eu ; see Chapter 11).
Contact	http://www.amc.nl/web/research-75/publications/prof.-dr.-m.a.g.-sprangers-publications.htm

10.2.2.2 IMPACT

Aspect	Description
Project team	Mirjam Sprangers, PhD, with Iris Hartog, MSc; Justine Netjes (until 2017), MSc; Tom Oreel, MSc; Pythia Nieuwkerk, PhD; Michael Scherer-Rath, PhD; José Henriques, MD, PhD; Hanneke van Laarhoven, MD, PhD
APH site	AMC
Project	The <i>IMPACT</i> project (http://www.impactonderzoek.nl/) is an NWO-funded study targeting the Quality of Life (QoL) of patients with multiple chronic morbidities, specifically those undergoing a cardiac intervention. The project aims to improve the conceptualization of QoL and enhance the sensitivity and comprehensiveness of its measurement by taking the trait-state distinction, contextual factors, and response shift into account. Participants are cardiac patients with comorbidities who were scheduled for elective percutaneous coronary intervention (PCI) or elective coronary artery bypass graft (CABG) (N = 320). In a sub-sample of participants, QoL is monitored through EMA (n = 37), to monitor QoL states. The study has a longitudinal design, with three EMA data collection periods: 1) pre-treatment, 2) two weeks after treatment for PCI patients or 3 months post-treatment for CABG patients, and 3) six months post-treatment.

Aspect	Description
EMA measures	Participants are prompted to answer nine general and one evening questionnaire per day, for seven consecutive days. During the day, patients are beeped randomly between 7:30 and 22:30 to complete a 19-item questionnaire. Concepts measured include: positive mood (feeling energetic, relaxed feeling, cheerfulness, happy), negative mood (anxiety, sadness, irritation, worry), coronary artery disease symptoms (chest pain, tightness in chest, oppressive feeling on the chest), and general symptoms (tiredness, other types of pain, shortness of breath), and contextual items. Items are rated on a 7-point Likert-scale, ranging from 'Not at all' to 'Very much'. Patients are also asked to complete an evening questionnaire just before they go to bed. The evening questionnaire has, besides the general questionnaire, an additional set of questions from the EQ-5D, and a question on the overall health state of the day. This item is rated on a visual analogue scale from 0 (worst) to 100 (best). Data will be analyzed with vector auto-regressive models, using R.
Platform used	PsyMate (http://psymate.eu ; see Chapter 11)); Participants are provided with iPods with a pre-installed EMA application, or, if they prefer, can use their own device.
Contact	http://www.amc.nl/web/research-75/publications/prof.-dr.-m.a.g.-sprangers-publications.htm

10.2.3 Snoek

Prof. dr. Frank Snoek is Professor of Medical Psychology, specialized in psycho-social diabetology. He heads the Diabetes Psychology Research Group (<http://www.vumc.com/branch/diabetes-psychology/>) and is consulting clinical psychologist for the VUmc Diabetes Center <http://www.vumc.nl/afdelingen/diabetescentrum/>). Snoek and co-workers use EMA to study the relationship between blood glucose variability and wellbeing.

Aspect	Description
Project team	Frank J. Snoek, PhD, Maartje de Wit, PhD, Cat Racca, MSC, Linda T. Muijs, MSc
APH site	VUmc
Project	In the MERITS study ('Momentary assessment of patient Experiences in Real life of Insulin Glargine 300 in Type 1 diabetes'), Snoek and co-workers use EMA to explore whether a) blood glucose variability is associated with changes in wellbeing (mood / energy) during waking time, whether b) switching to U-300 results in less glucose variability and translates into improved mood over time within patients, and whether c) if individual differences (profiles) can be distinguished with regard to the (strength of the) association between glucose variability and changes in mood.
EMA measures	Adult patients (N = 70) with type 1 diabetes, will be (randomly) prompted to answer questions on mood (based on POMS questionnaire), diabetes distress, fear of hypoglycemia and sleep.
Platform used	Ilumivu (http://ilumivu.com , see also Chapter 11).
Contact	http://research.vumc.nl/en/persons/frank-snoek http://research.vumc.nl/en/persons/maartje-de-wit http://research.vumc.nl/en/persons/cati-racca http://research.vumc.nl/en/persons/linda-muijs

10.3 Suicidal Ideation

10.3.1 Van Ballegooijen

Dr. Wouter Van Ballegooijen is a post-doctoral researcher at the VU and GGZ inGeest, specialized in the use of information and communication technology in mental health care and suicide prevention. In the ‘Continuous Assessment for Suicide Prevention And Research’ study (*CASPAR*; Nuij et al., 2018) Van Ballegooijen and colleagues use EMA to study pathways to suicidal behavior.

Aspect	Description
Project team	Wouter van Ballegooijen, PhD; Chani Nuij, MSc.; Ad Kerkhof, PhD; Jan Smit, PhD; Heleen Riper, PhD, and others
APH site	VU, VUmc, GGZ inGeest
Project	The ‘Continuous Assessment for Suicide Prevention And Research’ (<i>CASPAR</i>) study ([Nuij et al. (2018)]) aims to evaluate the feasibility of mobile safety planning and daily mobile self-monitoring in routine care treatment for patients with major depression or dysthymia and suicide risk in mental health care. Feasibility will be operationalized in terms of uptake, usage, acceptability, usability and patient satisfaction. EMA data will be used to: (a) empirically validate hypothesized psychological processes and stages of suicide pathways, (b) identify individual pathways to suicidal behavior, and (c) profile (sub)types of suicidality. The project is expected to result in 1) an interactive smartphone-based safety plan that patients can access 24/7, 2) increased disease awareness of patients due to self-monitoring, and 3) input for the national and international field of mental health care by sharing our results and our data, ultimately contributing to more personalized interventions according to precision medicine principles, and more effective suicide prevention.
EMA measures	In the project, adult suicidal patients (N = 80) with major depression or dysthymia and suicide risk in mental health care will be prompted, three times a day, to answer eight self-report items (e.g. ‘I feel sad’) assessing mood, rumination, hopelessness, defeat, entrapment, burdensomeness, belongingness, impulsiveness, suicidal imagery and suicidal ideation. Items, which are based on existing questionnaires, such as the Patient Health Questionnaire (PHQ-9), are rated on a 7-point Likert-scale, ranging from ‘Not at all’ to ‘Very much’. Results are presented to patients in graphs, which patients encouraged to discuss with their clinicians. In addition to active EMA data, location data are also gathered to assess movement patterns and daily rhythms. Planned variables also include accelerometer data and smartphone usage patterns. These data are not visible to study participants.
Platform used	Illumivu (http://illumivu.com , see Chapter 11).
Contact	http://research.vu.nl/en/persons/wouter-van-ballegooijen http://research.vu.nl/en/persons/chani-nuij

10.4 Psychotic symptoms

10.4.1 Van der Gaag

Prof. dr. Mark van der Gaag is professor of Clinical Psychology at the VU University in Amsterdam, and the head of psychosis research at Parnassia, the Hague. He is specialized in the research and treatment of

psychosis. In two recent RCTs, described below, Van der Gaag and colleagues used EMA outcomes to collect in-the-moment data.

10.4.1.1 TemStem

Aspect	Description
Team	Mark van der Gaag, PhD, Alyssa Jongeneel, MSc, David van den Berg, PhD, Dorien Scheffers, MSc
APH site	Vrije Universiteit Amsterdam, Parnassia Psychiatric Institute
Project	The TemStem project focusses on people who suffer from hearing voices and are obstructed by them in their daily life. Study participants install an app that contains both an EMI and EMA function, which is designed to reduce distress and dysfunction caused by auditory verbal hallucinations (Jongeneel et al., 2018). Components of the app include 1) coping: addressing verbal working memory phonological loop with a language task, thereby blocking the hearing of voices, 2) positive reinforcement: decreasing self-reported negative self-esteem themes, 3) treatment: reducing emotional response to memories associated with voices by taxing the auditory working memory during recall of negative auditory memories (as in EMDR therapy). The TemStem study aims to explore the effect of the app on distress and dysfunction in an RCT, specifically with regard to the effect of TemStem on frequency and severity of AVH. Additional analyses will focus on the identification of working mechanisms (predictors and mediators of effects), and the usability of TemStem.
EMA measures	TemStem users are encouraged to fill-in nine self-report items on a daily basis. Items tap: 1) hearing voices: 6 items (e.g. “ <i>Today, the voices were disturbing</i> ”), 2) mood, 3) self-esteem, 4) the use of TemStem (“ <i>I used TemStem today</i> ”). Items, which are rated on a 7-point Likert-scale, are based on existing EMA questionnaires. Results are presented to users in separate graphs, to support users in gaining insight in the pattern of AVH over time, or after use of Temstem. Stored variables include scores of vividness of AVH pre and post use of TemStem, data on application use (duration), used function (e.g. ‘Silencing’ function which focuses on coping, or ‘Challenging’ function which is based on dual tasking), and how users feel when they hear voices. Users can choose to provide additional information on age, gender, which county in the Netherlands they are currently located, how they found the app, and why they want to use it
Platform used	The TemStem app was developed by Reframing Studio, in collaboration with Parnassia Psychiatric Institute and TU Delft. The app is available for IOS and Android.
Contact	http://research.vu.nl/en/persons/m-van-der-gaag http://research.vu.nl/en/persons/alyssa-jongeneel http://research.vu.nl/en/persons/dpg-van-den-berg

10.4.1.2 VRETp

Aspect	Description
Team	Mark van der Gaag, PhD, Roos Pot-Kolder, MSc, with others
APH site	Vrije Universiteit Amsterdam, VUmc, (in collaboration with UMC Groningen, Maastricht University)

Aspect	Description
Project	In the context of a trial exploring the effect of virtual reality exposure therapy on social participation in people with a psychotic disorder (VRETP; n = 116, Pot-Kolder et al., 2016), EMA data were collected to assess changes in social functioning and paranoid ideation.
EMA measures	EMA included momentary assessment of 1) Anxiety (one item, e.g. " <i>I feel anxious</i> "), 2) Perceived social threat (four items, e.g. " <i>In this company, I feel accepted</i> "), 3) Paranoia (three items, e.g. " <i>I feel suspicious</i> "), and 4) Time spent with others (max. three multiple choice items inquiring about type of company (nobody, family, non-family, etc.). These items were based on previous EMA work [e.g., Collip et al. (2011)]. Participants were prompted 10 times a day, during 6 days. Anxiety, threat and paranoia items were rated on a 7-point Likert scale, ranging from 1 ("not at all") to 7 ("very"). Reports had to be completed within 15 min of the beep. To be included in the analysis, participants had to complete diary entries for at least one-third of the beeps (i.e., a minimum of 20 measurements). All 116 participants completed EMA measurements at baseline (mean number of completed self-assessments 46.1, SD 13.3), 96 participants completed the post-treatment assessment sufficiently (43.1, SD 10.1), and 87 participants completed the follow-up (43.2, SD 11.1). The trial results suggest that the addition of VR-CBT to standard treatment can reduce paranoid ideation and momentary anxiety in patients with a psychotic disorder (Pot-Kolder et al., 2018).
Platform used	PsyMate (http://www.psymate.eu/ , see Chapter 11). Because the PsyMate application was not finished at the time of the trial, participants were provided with a small palmtop device for the duration of the study.
Contact	http://research.vu.nl/en/persons/m-van-der-gaag https://www.researchgate.net/profile/Roos_Pot-Kolder

10.5 Sleep

10.5.1 Van Someren

Prof. dr. Eus van Someren is affiliated with the department of sleep and cognition of the Netherlands Institute for Neuroscience (<http://herseinstituut.nl/>), the Department of Psychiatry of VUmc and GGZ inGeest, Amsterdam. His group focusses on the study of healthy and disturbed sleep, using neuro-imaging, actigraphy and EMA. Here, two examples of such studies are provided.

Aspect	Description
Team	Eus van Someren, PhD (department head), with Bart te Lindert, MSc, Wisse van der Meijden, MSc, Tessa Blanken, MSc, Michele Colombo, PhD, Kim Dekker, MSc, Jeanne Leerssen, MSc, Rick Wassing, MSc.
APH site	The Netherlands Institute for Neuroscience, VUmc, GGZ inGeest,

Aspect	Description
Project(s)	The Van Someren group is involved in a variety of projects in which sensory data are collected (see, e.g., Van Someren, 2011, Van Someren (2000)). Exemplary of current research activity is the work of PhD candidates <i>Wisse van der Meijden</i> and <i>Bart te Lindert</i> . Van der Meijden studies, among other things, the interplay between light, vigilance, and sleep. A recent study associated the post-illumination pupil response (PIPR) to blue light with multiple indices of sleep timing: a) a questionnaire on habitual lights-out time, sleep onset latency, and final wake-up time (Munich Chronotype Questionnaire, MTCQ (Roenneberg et al., 2003)); b) a one-week sleep diary on actual wake/sleep times; and c) actigraphy. Participants (adolescents and young adults, n = 71) with a later sleep timing had a stronger responsiveness to blue light. The mid-sleep timing estimates from sleep diaries and actigraphy shared 94.5% of their inter-individual variance (Van der Meijden et al., 2016). <i>Bart te Lindert</i> studies the effect of the environment on sleepiness, for example by measuring light, (skin) temperature, posture, and psychological variables. A recent study focused on the effect of light intensity on Liking, Wanting and mood in insomnia disorder (Te Lindert et al., 2018) The study combined active and passive EMA. EMA prompts were sent 8 times a day, for one week and were timed at quasi-random intervals between 8:00 and 22:00. In addition, participants provided input after waking up and before bedtime. EMA consisted of 22 items. Liking and Wanting (6 items each), focused on taste or smell, bodily sensation, watching or listening, interactions with other, physical activity or being busy, and receiving something, measured on a 0-100 VAS scale. Mood items were derived from the Daytime Insomnia Symptom Scale (DISS; Buysse et al., 2007), which was developed specifically for EMA. Items focused on positive mood (5 items) and negative mood (5 items), scored on a 0-100 VAS-scale. In addition, participants wore two light sensors (Dimesimeter) measuring wear-time and light. Participants received a designated Android smartphone for the duration of EMA. People with insomnia disorder (n = 17) had significantly lower subjective Liking and Wanting than matched controls without sleep complaints (n = 18). This was most apparent at low environmental light intensity. There were no overall differences between groups in Positive mood and Negative mood. Participants with insomnia did have a different diurnal profile of Positive mood and Negative mood (Te Lindert et al., 2018).
Platform(s) used	Philips Actiwatch Spectrum; Microelectromechanical accelerometer (Move II, Movisens GmbH). Sleep onset and final wake-up time were estimated using an algorithm that is implemented in Matlab (http://github.com/btlindert/actant-1). For EMA the MovisensXS platform was used (https://xs.movisens.com/ , see Chapter 11).
Contact	http://research.vumc.nl/en/persons/eus-jw-van-someren http://herseinstituut.nl/over-ons/de-organisatie/medewerkers/bart-te-lindert/ http://herseinstituut.nl/over-ons/de-organisatie/medewerkers/wisse-van-der-meijden/

10.6 Stress & Emotion

10.6.1 De Geus

Prof. dr. Eco de Geus is head of the department of Biological Psychology and co-director of the Netherlands Twin Registry at the VU University. De Geus has been the driving force behind the development of the VU University Ambulatory Monitoring System (VU-AMS; <http://www.vu-ams.nl/vu-ams/>), a non-invasive wearable device that is used for continuous ambulatory measurement of the autonomic nervous system.

Aspect	Description
Team	Eco de Geus, PhD; Gonneke Willemse, PhD; Martin Gevonden, PhD; Denise van der Mee, MSc, Mandy Tjew-A-Sin, MSc; Cor Stoof, MSc
APH site	VU
Project	VU-AMS was developed at the department of Biological Psychology of the Vrije Universiteit Amsterdam. The system is used worldwide by many research groups, to study stress and emotion in both laboratory and naturalistic settings (De Geus et al., 1995; De Geus and Van Doornen, 1996; Willemse et al., 1996). VU-AMS has been used in the study of ADHD, aggression, anxiety and depressive disorders, mental, social and work-related stress, circadian rhythms, hyperventilation, migraine, sleep, and in studies linking the autonomic nervous system to metabolic and immunological risk factors (see http://www.vu-ams.nl/research/publications/ and http://www.vu-ams.nl/research/phd-theses/). Current projects include: 1) Validation of a wristwatch-based technology, developed by Philips, to measure skin conductance responses in a laboratory (~2.5 h) and ambulatory (~22h, including the night) setting on a total of 100 subjects (van der Mee et al., ongoing, 2017 - 2021). The goal is to test whether wrist based (Philips) and palm based (VU-AMS) measured skin conductance responses accord, how these measures relate to the heart based measured pre-ejection period (PEP, VU-AMS) and whether it can be related to positive and negative affect (hourly diary prompts). The end goal for the wrist based technology is to detect sympathetic nervous system activity (measured as skin conductance responses) and present this information to the person's as an index of the current stress level, alongside a one-hour prediction of changes in stress level and cognitive functioning (Cognitive Zone Changes; http://www.ip.philips.com/licensing/program/121) 2) Ambulatory study on self-regulation among youth (Tjew A Sin et al., 2018-2019). This study is conducted at the schools as part of the NeuroLab project. A total of 50 students, will wear the VU-AMS device for approximately 22 hours. Autonomic nervous system activity is collected continuously over the course of a regular school day (including the night and morning thereafter) and will be related to multiple components of self-regulation, including emotion-regulation, cognitive functioning, impulsivity and inhibition

Aspect	Description
EMA measures	The VU-AMS device is a battery powered wearable that can record up to 48 hours of data (4GB storage). It measures the electrocardiogram, the impedance cardiogram, and skin conductance. With VU-AMS, the following outcomes can be collected: Heart Rate / Inter beat Interval (IBI), Heart Rate Variability (SDNN, RMSSD, IBI power spectrum: HF, LF), Respiratory Sinus Arrhythmia (RSA), Pre-Ejection Period (PEP), T-wave amplitude (TWA), Left Ventricular Ejection Time (LVET), Stroke Volume (SV) and Cardiac Output (CO), Respiration Rate (RR), Skin Conductance Level (SCL) and Skin Conductance Responses (SCRs), Movement (Hip-worn tri-axial accelerometer signals (g)). The freely available ‘Data Analysis and Management System’ (DAMS) package is used for extraction and processing of VU-AMS data (see http://www.vu-ams.nl/support/downloads/software/). The DAMS tool offers options for data inspection (visual inspection of raw data), automated detection of R-peaks in raw ECG signal and visual inspection of final IBI time series, event/diary-based data labeling, IBI spectral power calculation, automated scoring of RR, RSA, and PEP from the combined ECG.
Platform(s) used	VU-AMS / DAMS (see Chapter 11).
Contact	http://research.vu.nl/en/persons/jcn-de-geus http://research.vu.nl/en/persons/ahm-willemsen http://research.vu.nl/en/persons/martin-gevonden https://research.vu.nl/en/persons/denise-van-der-mee https://research.vu.nl/en/persons/mandy-tjew-a-sin

Chapter 11

EMA Instruments Catalogue

For conducting EMA studies, a variety of apps, online applications and wearable devices are on the market. In this chapter, we list a selection of instruments that we found to be in use in scientific studies within the APH consortium.

Table 11.1: EMA Instruments.

Name	Manufacturer	Description
EMA APPS		
Ilumivu	Ilumivu	Android / iOS EMA app
Moodbuster	ICT4D Consortium	Android EMA app
Movisens	MoviSens	Android EMA app
PsyMate	PsyMate	Android/iOS EMA app
RoQua	RoQua	SMS / Browser-based active EMA
Wearables		
GENEActiv	Activinsights	Wrist-worn accelerometer
VU-AMS	VU	Wearable for autonomic nervous system assessments

11.1 EMA Platforms & Apps

11.1.1 Ilumivu

Ilumivu (<https://ilumivu.com/>) is an American commercial company specialised in mobile EMA application development. It offers a cross-platform (Android and IOS) smartphone app (mEMA) that researchers can use to collect data from study participants.

Researchers can define the assessment plan of their study in an online back-office. The Ilumivu EMA toolbox library includes a rather complete set of survey elements, survey logic branching tools, and survey scheduling options. Passive EMA options include GPS tagging (of survey responses), Geo-fencing (triggering surveys at specific locations) and smartphone sensor logging (light and noise level, screen brightness, screen locked/unlocked, humidity, ambient temperature, barometric pressure, phone/SMS activity log, and basic device information). With the online back-office, researchers can invite participants, monitor study compliance and download data (in common data formats such as CSV).

A limited number of smart watch devices can optionally be linked to the Ilumivu app for further passive data collection. The company is open for custom development, when studies require features that are not included in the default app.

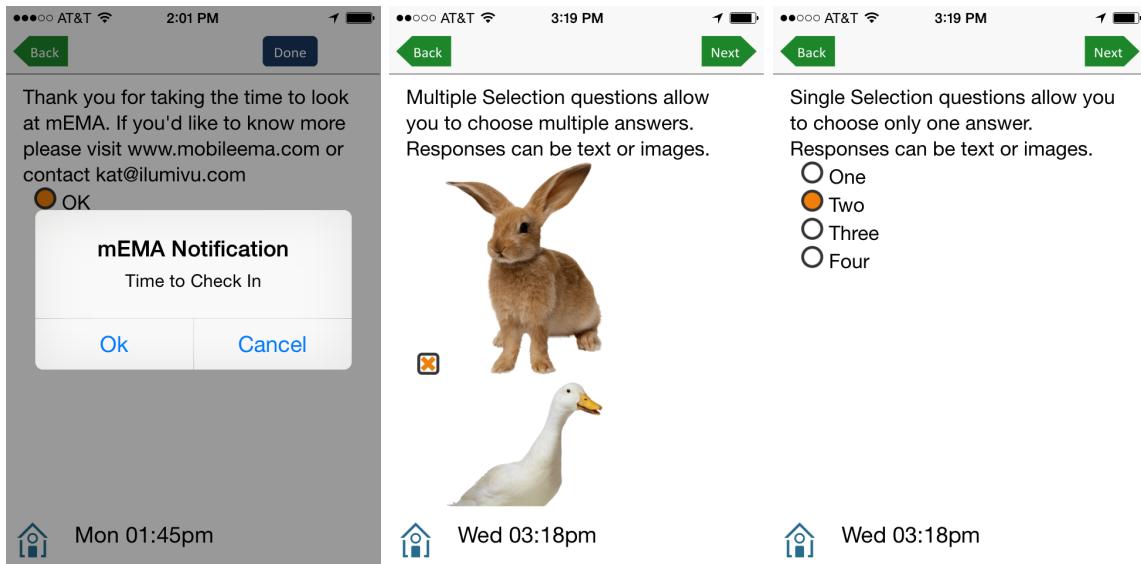


Figure 11.1: Ilumivu App Screenshots

Current (July 2018) subscription plans range from 3.375 dollar (basic features) to 12.000 dollar (all features), per year. Researchers, who are interested in purchasing a license of the service, are advised to consult the ‘Grant Writer’s Guide’ of Ilumivu (at <https://ilumivu.com/pricing/writing-a-grant/>). Ilumivu’s competitive advantage is that it has multi-platform support. It is, however, not the cheapest EMA product on the market. Being an American company, researchers should also consider EU regulations relating to personal data privacy protection, since regulations are stricter when personal data of EU citizens leave the EU.

11.1.2 Moodbuster

Moodbuster (<http://moodbuster.eu/>) is a web-based treatment platform with an integrated EMA app. The platform was developed by an international non-profit research consortium, including VU, VUmc and GGZ inGeest, in two major EU-funded research projects: ICT4Depression (see [ICT4Depression.eu](http://ict4depression.eu); Warmerdam et al., 2012) and E-COMPARED (see <http://e-compared.eu>; Kleiboer et al., 2016; Van de Ven et al., 2017).

In the E-COMPARED trial, Moodbuster was used, in five EU-countries, to test blended treatment of major depression. In this study, participants used the smartphone app to rate mood and various other depression-related variables, in the context of their treatment, over a period of up to 20 weeks.

E-COMPARED EMA data were used in several predictive machine learning studies (Mikus et al., 2018; Rocha et al., 2018). In addition, the EMA app was used in a satellite study designed to assess the effects of long-term EMA (Van Ballegooijen et al., 2016). Moodbuster will also be used in the EU ImpleMentAll project (<http://www.implementall.eu>) and in several other clinical trials that are in preparation. Currently, EMA assessment protocols are hard-coded in the Moodbuster app. New EMA assessments protocols can be implemented in collaboration with the Moodbuster development team. An online back-office is in development. Previous studies used an Android version of the EMA app. Cross-platform versions of the app are in development.

More information on Moodbuster can be obtained from prof. dr. Heleen Riper (h.riper@vu.nl).



Figure 11.2: MoodBuster App Screenshots

11.1.3 Movisens

Movisens (<http://www.movisens.com>) is a German company that is specialized in the development of hard- and software solutions for mobile sensing. The company sells small wearable devices that contain several high-precision sensors, including an accelerometer, gyroscope, barometer and thermometer. In addition, the company has developed an (Android) app, called MovisensXS, which can be used for active EMA research. The app can optionally be configured for smartphone logging (e.g., to log music that a study participant listens to). The wearable sensor can also be linked to the app, so that EMA questionnaires can be triggered based on targeted activity or energy expenditure patterns, such as extended periods of sedentary behavior.

Specialized software to import, pre-process and analyze raw sensor data is available for download. Like Ilumivu, researchers can define EMA sample schedules for their study in a web-based back-office (<https://xs.movisens.com>), using an online graphical editor. Once defined, participants can be invited to the study, through the back-office, to download the freely available Movisens App from Google Play store. The back-office also allows researchers to monitor study compliance and download data.

MoviSensXS EMA license costs vary from 500 to 10.000 euro's per year, depending on the required number of 'credits' which are linked to the number of EMA responses. Prospective users can test platform, without functional restrictions, with a free test account. An EMA test-study can thus be set up and started in less than a day. A current (November 2018) limitation of Movisens is the lack of an iOS version of the EMA app. Study participants who own an iPhone have to be excluded from studies, or will have to be provided with an Android phone.

11.1.4 PsyMate

The PsyMate app (<http://www.psymate.eu>) was developed by the Department of Psychiatry and Psychology at Maastricht University in the Netherlands to assess psychological problems in daily life. The app has been validated for use in depression, bipolar disorder, and psychosis, with new scales currently being developed for a range of diseases including Parkinson's disease, pain, cardiology, hypertension, diabetes and Irritable Bowel Syndrome. It was also used in a EU-funded project to study gene-environment interaction in schizophrenia (<http://www.eu-gei.eu/about-the-project/psymate>).

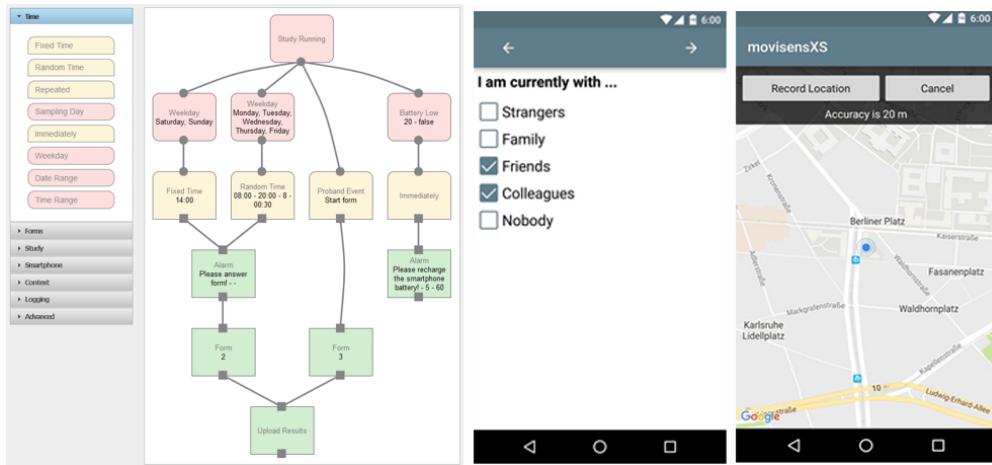


Figure 11.3: Movisens Sample scheme editor (left) and App Screenshots (right)

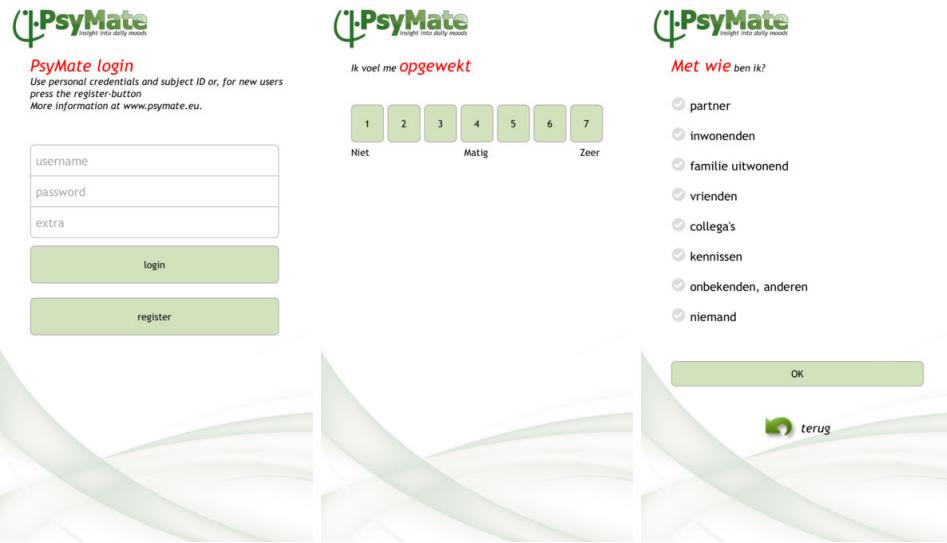


Figure 11.4: PsyMate App Screenshots

The app is free to download for iOS and Android devices on Apple and Google play stores. PsyMate is available in five languages (English, French, German, Dutch, Spanish). More translations are in preparation.

Uses of the app include self-monitoring of mood states, delivering professional support during treatment, or conducting EMA-research. The app can be customized to address specific client needs or research projects, with expertise from the PsyMate development team, which includes an active working group that meets regularly to discuss and advise new projects. Researchers have access to the raw data without having to go through the PsyMate development team. Communication from the PsyMate back office to researchers about updates and assistance with technical problems could be a point for consideration for using this platform.

11.1.5 RoQua

RoQua (<http://www.roqua.nl/>) is a web-based Routine Outcome Monitoring system, developed and maintained by a Dutch non-profit development and service organization that is funded by several northern GGZ organizations and the Department of Psychiatry, University Medical Center Groningen. RoQUA has



Figure 11.5: Screenshots of the participant feedback web-page of the 'HowNutsAreTheDutch' project, in which data is collected by the RoQua system

a sophisticated and user-friendly online back-office portal, with which researchers can define assessment protocols and invite study participants - through e-mail or SMS - to complete questionnaires online (on desktop or mobile devices).

By inviting study participants several times a day to complete a questionnaire via the standard web browser of their mobile phone, active EMA can be implemented. This approach was taken in several large-scale studies, including 'NESDA' (<http://nesda.nl>; see also Chapter 10) and 'HowNutsAreTheDutch' (<http://www.hoegekis.nl>; Van der Krieke et al., 2017, 2016). At present, RoQua does not support the collection of passive EMA data. However, preliminary results have been reported with a system called 'Physiqual' (Blaauw et al., 2016), with which active EMA data, collected with RoQUA, can be combined with wearable sensor data.

11.2 Wearables

11.2.1 GENEActiv

GENEActiv, sold by UK-based company Activinsights (<http://activinsights.com>), is a waterproof wrist-worn device with a high-precision, configurable 3-axial accelerometer (range: +/- 8g), an ambient light sensor, a (near-body) temperature sensor, and an event logger (a button that users can press to mark a targeted event). GENEActiv was developed to accurately assess human activity in scientific studies. The device has a storage capacity of 0.5GB of raw data. At 10Hz, the device can log activity up to one month. July 2018, the price for one unit was approximately 250 euro.

GENEActiv is used in a growing number of clinical studies to measure activity and sleep-wake cycles, in natural conditions, over longer periods of time. Dedicated R-packages to pre-process and analyze the raw data exist (see Chapter 12). Matlab users should consider the open source 'Actant - Activity Analysis Toolbox' (<https://github.com/maximosipov/actant>), by Osipov & Te Lindert, which includes functionality for visualization and analysis of behavioral and environmental timeseries, acquired using GENEActiv (and other) accelerometers. No accompanying app exists with which study participants can be provided feedback about their activity. This might negatively affect wear-time and study compliance in research participants, who are accustomed to consumer activity-sampling devices, such as Fitbit, where many options for real-time feedback exist. Further, researchers we spoke with noted that the event logger was hard to use in practice. This resulted in low compliance in field studies.



Figure 11.6: The GENEActiv Accelerometer

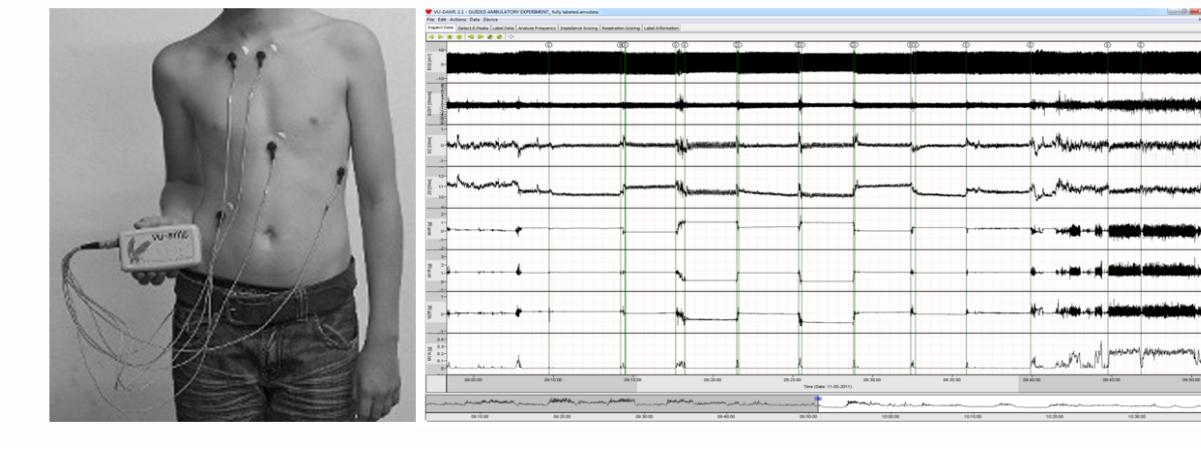


Figure 11.7: VU-AMS device (left) and VU-DAMS screenshot (right)

11.2.2 VU-AMS

The VU University Ambulatory Monitoring System (VU-AMS; <http://www.vu-ams.nl/>), developed by the department of Biological Psychology and the Technical Department (ITM) of the Faculty of Behavioral and Movement Sciences, allows ambulatory recording of autonomic and cardiovascular activity. VU-AMS measures heart rate, heart rate variability, Respiratory Sinus Arrhythmia, Pre-Ejection Period, Left Ventricular Ejection Time, Respiration Rate, Stroke Volume (SV) and Cardiac Output, Skin Conductance Level (SCL) and Skin Conductance Responses (SCRs) and Tri-Axial Accelerometry (of Body Movement). For the processing of VU-AMS data, a dedicated software suite called the 'Data Analysis and Management Software' (VU-DAMS) is available (for Windows and Mac). VU-AMS is used worldwide by many research groups, to study stress and emotion in both laboratory and naturalistic settings (De Geus et al., 1995; De Geus and Van Doornen, 1996; Willemsen et al., 1996).

Chapter 12

R Packages for EMA Research

Many R packages exist that can help you in the management and analysis of EMA data. In this chapter, a selection of these packages is discussed. For each, we provide a summary description, a code example code, and pointers to further documentation, to give you a head start in using the packages for your work.

Table 12.1: List of R packages that are useful in EMA research.

Category	Package	Description
Accelerometry	GENERead / GGIR	Import, pre-process and analyze raw accelerometer data.
	PhysicalActivity	Analyze Actigraph accelerometer data.
Data Management & Visual Exploration	dplyr	Data transformation.
	ggplot2	Create graphs.
	haven	Import and export SPSS data files.
Auto-regressive modeling	lubridate	Manipulate date and time variables.
	autovarCore	Automate the construction of vector autoregressive models.
Mixed-effects Modeling	lme4	Fit linear and nonlinear mixed-effects models. Fast alternative to package ‘nlme’.
	nlme	Fit linear and nonlinear mixed effects models. Pre-dates package <code>lme4</code> , but is still used because it provides advanced options to model correlational structures
Simulation	simr	Simulation-based power calculations for mixed models.
Symptom Networks	simstudy	Simulate complex study data.
	bootnet	Assess the stability of symptom networks.
Time series analysis	qgraph	Estimate and plot symptom networks.
	lomb	Calculate the Lomb-Scargle Periodogram for unevenly sampled time series.

12.1 Accelerometry

Accelerometer data need considerable pre-processing before final analyses can be run. Raw data have to be read in from a variety of brand-specific file formats, data have to re-calibrated on a per-device basis, non-wear periods have to be detected, and summarizing measures, such as activity counts and energy-expenditure measures, have to be calculated from imputed triangular (x, y, z) data, often in several time windows (i.e., epochs).

12.1.1 GENEAreaed / GGIR

Package *GGIR* (Van Hees et al., 2018) is a package to pre-process raw accelerometry data from three brands of wearables that are widely used in sleep and physical activity research: GENEActiv, ActiGraph and Axivity. GGIR includes package **GENEAreaed** (?), GENEAreaed, with which raw data can be imported into R for further processing, as illustrated below.

```
# Reading raw GENEActiv data.
library(GENEAreaed)
library(ggplot2)
library(tidyr)

dat <- read.bin(system.file("binfile/TESTfile.bin", package = "GENEAreaed"),
                 verbose = FALSE, downsample = 20)
#> Processing took: 0.178 secs .
#> Loaded 1560 records (Approx 0 MB of RAM)
#> 12-05-23 16:47:50.000 (Wed) to 12-05-23 16:53:01.799 (Wed)

d <- as.data.frame(dat$data.out)
d <- gather(d, key = "sensor", value = "value", -timestamp)
d$timestamp <- as.POSIXct(d$timestamp,
                           origin = "1970-01-01",
                           tz = "UTC")

ggplot(d, aes(x = timestamp, y = value)) +
  geom_line() +
  facet_wrap(~sensor, scales = "free_y")
```

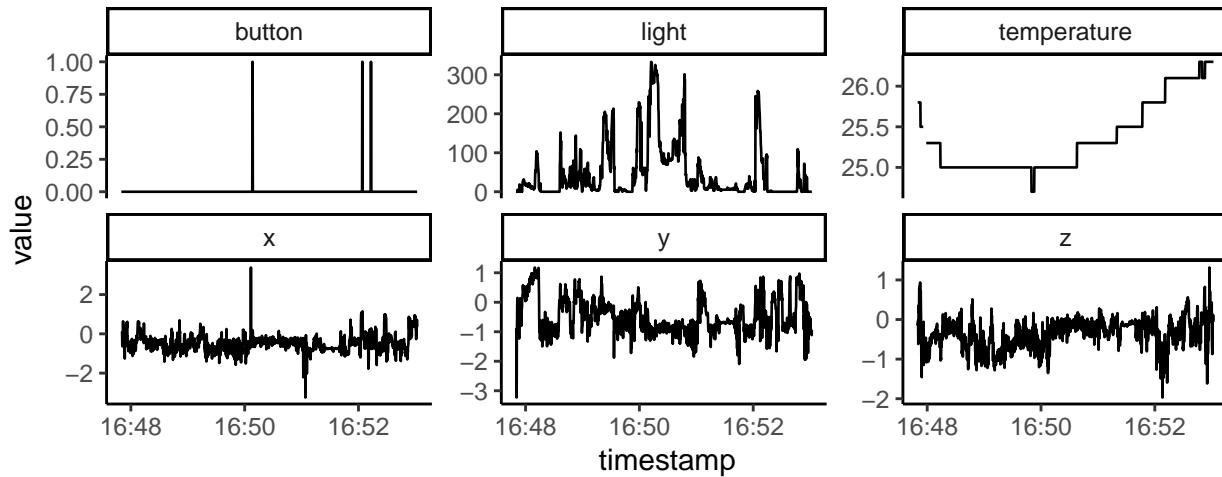


Figure 12.1: Raw sensor data of a GENEActiv wrist-worn tri-axial accelerometer (down-sampled from 100Hz to 5Hz).

12.1.2 PhysicalActivity

Package *PhysicalActivity* (Choi et al., 2018) provides an alternative to packages *GGIR* and *GENERead*, when ActiGraph data are available.

```
# Plotting activity counts.
library(PhysicalActivity)
library(ggplot2)

data(dataSec)

d <- dataCollapser(dataSec, TS = "TimeStamp", col = "counts", by = 300)

ggplot(d, aes(x = as.POSIXct(TimeStamp), y = counts)) +
  geom_line(size = .5, alpha = .5) +
  xlab("Time") + ylab("Activity Counts")
```

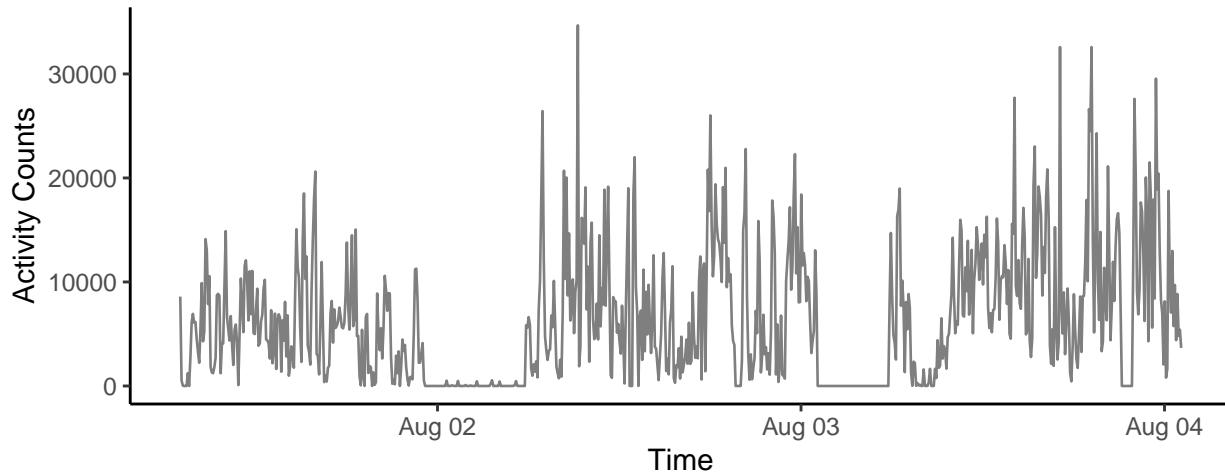


Figure 12.2: Activity Counts (5-minute windows), in a Three-day Accelerometer data set.

12.2 Autoregressive modeling

12.2.1 autovarCore

Vector auto-regressive (VAR) models can be used to detect lagged relationships between multiple time-series (see also Chapter 7). In VAR, each variable is modeled as a linear function of past values (lags) of itself and of present and past values of other variables. When EMA is used to capture multiple phenomena over time, VAR can provide insight in how these phenomena interact. One challenge in VAR modelling is that many alternative models potentially exist, Package *autovarCore* (Emerencia, 2018) was developed to help researchers to find the VAR model with the best fit to a given time-series data set.

In the (unrealistic) example below, function `autovar` is used to detect that changes in depression are positively related to past (lag 1) values of activity, in a simulated data set:

```
# Autovar analysis.
library(autovarCore)

# simulate data
N = 100
depression <- rnorm(N)
activity <- rnorm(N)
activity_lag1 <- c(NA, activity[1:(N -1)])

depression <- depression + 0.5 * activity_lag1
d <- data.frame(depression, activity)

models_found <- autovarCore::autovar(d, selected_column_names = c('activity', 'depression'))

# Show details for the best model found
summary(models_found[[1]]$varest$varresult$depression)
#>
#> Call:
#> lm(formula = y ~ -1 + ., data = datamat)
#>
```

```
#> Residuals:
#>      Min       1Q   Median      3Q     Max
#> -2.16684 -0.70663 -0.07416  0.55735  2.51048
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|)
#> activity.l1    0.50159   0.10708   4.684 9.25e-06 ***
#> depression.l1  0.14209   0.09166   1.550   0.124
#> const          0.09227   0.10229   0.902   0.369
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 1.015 on 96 degrees of freedom
#> Multiple R-squared:  0.2048, Adjusted R-squared:  0.1882
#> F-statistic: 12.36 on 2 and 96 DF,  p-value: 1.669e-05
```

`AutovarCore` is a simplified version of a more extensive package `autovar` (Emerencia, 2018), which was used in several publications (Van der Krieke et al., 2015; Emerencia et al., 2016). Further information can be found on <http://autovar.nl> and <http://autovarcore.nl>

12.3 Data management & Visual Exploration

The `tidyverse` is a collection of well-designed packages, authored by the team behind RStudio, that together add a consistent, modern, and efficient extension of base R functions. The `tidyverse` includes popular packages such as `ggplot2` (for plotting), `haven` (to read SPSS files), `dplyr` (for data manipulation), and many more (see: <http://tidyverse.org> for a full list).

12.3.1 dplyr

Package `dplyr` (Wickham et al., 2018) implements the ‘split-apply-combine’-strategy. With `dplyr`, elementary data manipulations can be chained (using the pipe operator `%>%`) to elegantly implement complex data transformations.

```
# Aggregate data by ID, through a 'pipe'.
require(dplyr)

d <- data.frame(
  id = factor(rep(1:5, each = 10)),
  score = rnorm(50)
)

b <- as_tibble(d) %>%
  group_by(id) %>%
  summarize(mean = mean(score))

knitr:::kable(b)
```

id	mean
1	0.3668117
2	0.1445131
3	0.4324790
4	-0.1027515
5	-0.3770372

A good introduction to `dplyr` can be found in the book ‘R for Data Science’ (Wickham and Grolemund, 2016), which can be freely accessed online (<http://r4ds.had.co.nz/>).

12.3.2 ggplot2

Package `ggplot2` (Wickham, 2016) provides a collection of high-level plotting commands with which graphs can be build up in layers. It is based on ‘The Grammar of Graphics’ (Wilkinson, 2006), an influential analysis of the structure of scientific graphs. Systematic introductions are available on the the tidyverse website, and in the book ‘ggplot2: Elegant Graphics for Data Analysis’ (Wickham, 2016).

The example below illustrates how graphs are layered. In the first step, a coordinate system is set up. In step 2, all time/scores points are plotted. In step 3, a smoothed line is fitted through these points. Finally, in step 4, the graph is split on a variable ID (a subject identifier), to show individual trajectories.

```
# Simple ggplot example.
library(ggplot2)

# simulate example data
d = data.frame(
  ID      = rep(1:4, each = 25),
  time   = rep(1:25, 4),
  score  = rnorm(100, 0, 2))

# step 1: initialise the coordinate system
g <- ggplot(d, aes(x = time, y = score)); g

# step 2: add scatterplot
g <- g + geom_point(); g

# step 3: add a smoothed line
g <- g + geom_smooth(method = "loess"); g

# step 4: split plot by subject ID
g + facet_wrap(~ ID)
```

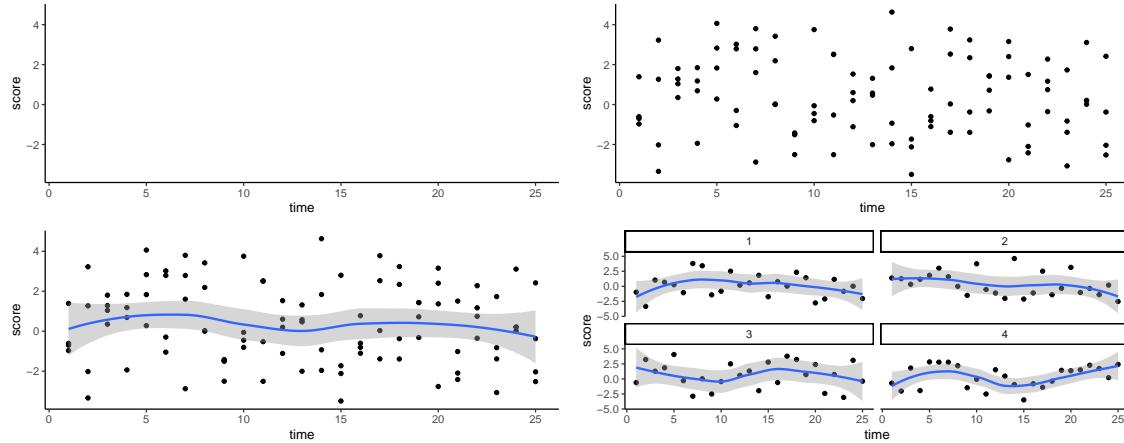


Figure 12.3: Plotting layers with ggplot2

12.3.3 haven

With package *haven* (Wickham and Miller, 2018), SPSS, STATA and SAS files can be read into R. A particular advantage is that variables definitions are retained (in attributes of the columns in the R `data.frames`).

```
# Read SPSS data
library(haven)

# read example data (included in package)
path <- system.file("examples", "iris.sav", package = "haven")
d <- read_sav(path)

# show attributes of variable
attributes(d$Species)
#> $format.spss
#> [1] "F8.0"
#>
#> $class
#> [1] "labelled"
#>
#> $labels
#>     setosa versicolor virginica
#>     1          2          3
```

12.3.4 lubridate

EMA data analyses frequently require manipulations of date-time variables. For this, package *lubridate* (Grolemund and Wickham, 2011), which provides many functions for common date and date time operations, can be very useful.

In the code snippet below, for example, the `round_date` function is used to calculate the ENMO value from raw tri-axial accelerometer data (see Chapter 6), in 15-minute epoch windows.

```
# Rounding datetime variables with lubridate.
library(emaph)
library(lubridate)
```

```

library(ggplot2)
library(dplyr)

d <- subset(genea,
            timestamp > "2018-06-02 12:00" &
            timestamp < "2018-06-02 18:00")

# round timestamp
d$epoch <- round_date(d$timestamp, "15 minutes")

# summarize x, y, z acceleration to ENMO
d <- d %>% group_by(id, epoch) %>%
  summarise(svm = sum(sqrt(x^2 + y^2 + z^2) - 1) / n())

```

id	epoch	svm
1	2018-06-02 12:00:00	0.0235192
1	2018-06-02 12:15:00	0.0486871
1	2018-06-02 12:30:00	0.0477664
1	2018-06-02 12:45:00	0.0128911
1	2018-06-02 13:00:00	0.0005558
1	2018-06-02 13:15:00	0.0027089

To learn more about handling dates and times with `lubridate`, Chapter 16 of the book ‘R for Data Science’ (Wickham and Grolemund, 2016) provides a good introduction.

12.4 Mixed-effects Modeling

Several R-packages for mixed-effects modelling exist. The most popular are package `nlme` (Pinheiro et al., 2018) and package `lme4` (Bates et al., 2015). Both packages are actively used, as both provide unique functionalities.

12.4.1 nlme

Package `nlme` is introduced in Chapter 8. It is an older package (in comparison to `lme4`), that is still used a lot because it provides options to model correlational structures that are not implemented (yet) in `lme4`.

```

# Fit a linear mixed model, with lme.
library(nlme)
fm <- lme(distance ~ age + Sex, data = Orthodont, random = ~ 1)

fixef(fm)
#> (Intercept)      age   SexFemale
#>  17.7067130  0.6601852 -2.3210227

```

12.4.2 lme4

Package `lme4` (Bates et al., 2015) is a faster reimplementation of the mixed-effects model. With large data sets and complex hierarchical models, this package should probably be preferred. As can be seen below, model specifications in `lmer` are different from model specifications in `lme`.

```

# Fit a linear mixed model, with lme.
library(lme4)

```

```
fm <- lmer(distance ~ age + Sex + (1 | Subject), data = nlme::Orthodont)

fixef(fm)
#> (Intercept)      age   SexFemale
#> 17.7067130  0.6601852 -2.3210227
```

12.5 Power analysis

12.5.1 simr

With package *simr* (Green and MacLeod, 2016), power of mixed-effects models can be determined via simulation. As illustrated below, the procedure requires the researcher to define the true parameters of a mixed model, and a single data set. Next, function `simPower` can be used to simulate new data sets and tests (of a specified parameter in the model), to determine the power of the test.

```
# Power analysis of a two-group repeated measures design
# (simulation approach).
library(simr)

# construct design matrix
time <- 1:24
subject <- 1:40
X <- expand.grid(time = time, subject = subject)
X$group <- c(rep(0, 24), rep(1, 24)) # group

# fixed intercept and slope
b <- c(2, -0.1, 0, -0.5)

# random intercept variance
V1 <- 0.5

# random intercept and slope variance-covariance matrix
V2 <- matrix(c(0.5, 0.05, 0.05, 0.1), 2)

# residual standard deviation
s <- 1

model1 <- makeLmer(y ~ time * group + (1 + time | subject),
                     fixef = b,
                     VarCorr = V2,
                     sigma = s,
                     data = X)

powerSim(model1,
          fixed("time:group", "lr"),
          nsim = 10, # set to high value for better results
          progress = FALSE)
#> Power for predictor 'time:group', (95% confidence interval):
#>      90.00% (55.50, 99.75)
#>
#> Test: Likelihood ratio
#>      Effect size for time:group is -0.50
```

```
#>
#> Based on 10 simulations, (0 warnings, 0 errors)
#> alpha = 0.05, nrow = 960
#>
#> Time elapsed: 0 h 0 m 5 s
```

12.6 Symptom Network Analysis

When EMA is used to tap various symptoms, network analysis can reveal the dynamic interplay between these symptoms (Borsboom and Cramer, 2013; Borsboom, 2017; Bringmann et al., 2015). Various packages exist to fit these networks in R. With these packages, it is relatively easy to fit a graphical network on multivariate data sets. If you are interested in conducting a network analysis, be sure to visit the Psycho-systems website, at <http://psychosystems.org>, or to register for the UvA network school course (see: <http://psychosystems.org/NetworkSchool>). A complete list of software related to network analysis from the Psycho-systems group is maintained at <http://psychosystems.org/software/>. Excellent tutorials can also be found at <http://psych-networks.com/tutorials/> and at <http://sachaepskamp.com/files/Cookbook.html>.

12.6.1 qgraph

Package *qgraph* (Epskamp et al., 2012) can be used to fit, visualize and analyze graphical networks.

In the example below, *qgraph* is used to fit a network on the ‘Critical Slowing Down’ (CSD) data set, which is included in package *emaph* (see Chapter 9).

```
# Fitting a network, with qgraph.

library(qgraph)
library(emaph)

# get mood_ scores from csd data set
d <- subset(csd,
             subset = phase == "exp: db: no change",
             select = grep("mood_", names(csd))[1:5])

# Fit and plot regularized partial correlation network
g <- qgraph(cor_auto(d, detectOrdinal = FALSE),
            graph = "glasso", sampleSize = nrow(d),
            nodeNames = names(d),
            label.scale = FALSE, label.cex = .8,
            legend = TRUE, legend.cex = .5,
            layout = "spring")
```

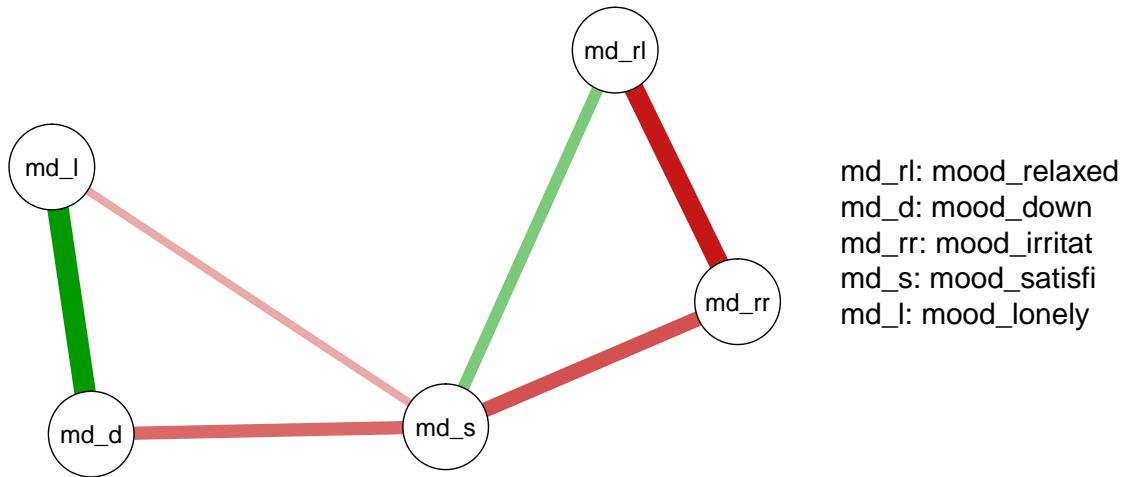
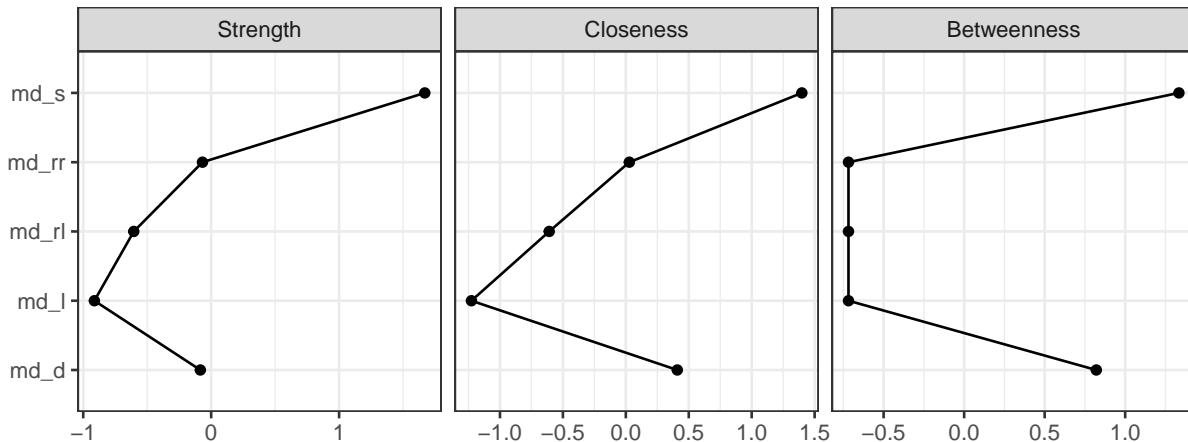


Figure 12.4: Network of mood items from CSD data set

Package `qgraph` also provides functions to analyze qualities of fitted networks, such as the centrality of nodes in the network. In the network plot above, node `md_s` appears to be a central node in the network. This is confirmed by calling `centralityPlot`:

```
centralityPlot(g)
```



12.6.2 bootnet

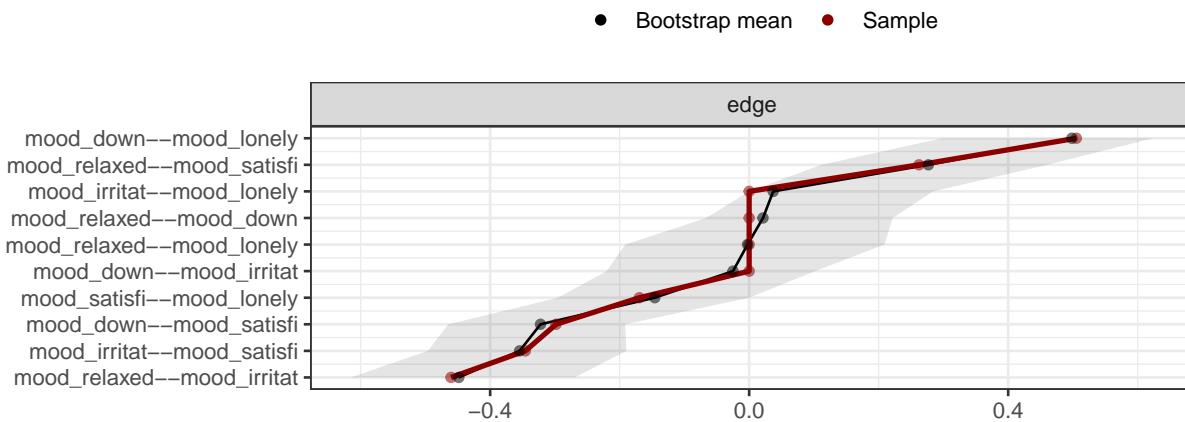
In the interpretation of fitted networks, it is important to take the stability of the network into account. Intuitively, networks that are fit on small sample data sets will be less stable than networks based on large data sets. One solution is to assess this is to fit a large number of networks on subsets of the original data, via bootstrapping. In stable networks, the variability of edge estimations and other characteristics will be small, while in unstable networks, the variance will be large. This idea is implemented in package `bootnet` (Epskamp et al., 2018).

Below, the stability of the network that was fit in the previous example is examined with `bootnet`: fifty networks are fit, based on fifty bootstrapped samples. In the results plot, the red line marks the strength of

the edges in the full sample, while grey confidence intervals mark the distribution of the edge weights in the bootstraps.

```
# A bootstrapped network.
library/bootnet

g <- estimateNetwork(d, default = "EBICglasso",
                      corArgs = list(detectOrdinal = FALSE))
results <- bootnet(g, nBoots = 50, verbose = FALSE)
plot(results, order = "mean")
```

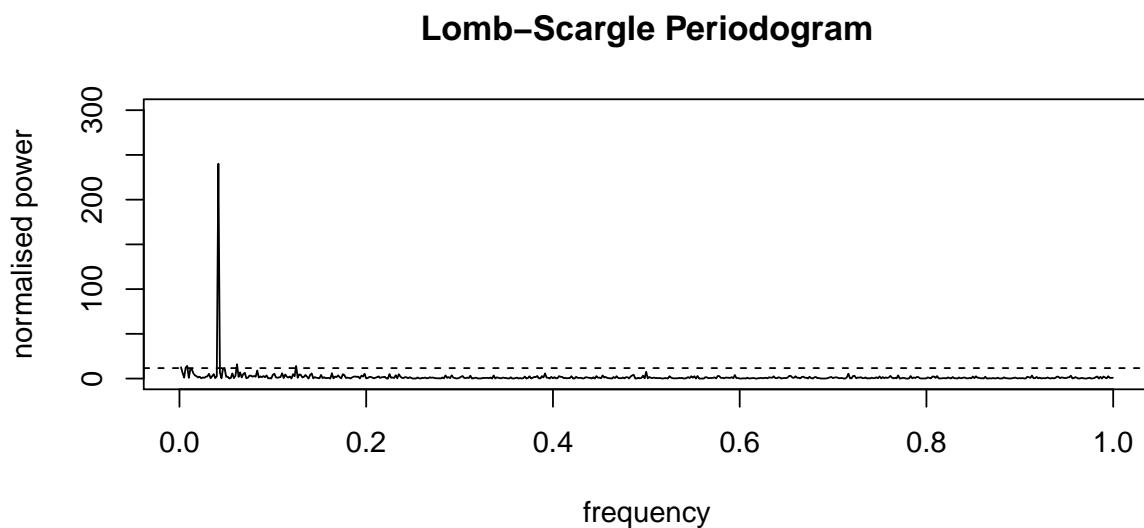


12.7 Timeseries analysis

12.7.1 lomb

Disturbances in circadian rhythms have been related to depressive symptoms (see, e.g., Saeb et al., 2015). With so-called periodograms, these circadian rhythms can be detected in EMA data (see Chapter 7). Standard analysis techniques, however, expect regular time-series, in which data are sampled at equidistant intervals. EMA data, typically, are not equidistant. One solution to this problem is to use the Lomb-Scargle periodogram procedure (Lomb, 1976), which can be applied to unevenly-sampled time-series as well. Package *lomb* (Ruf, 1999) implements this procedure.

```
# Calculating a Lomb-Scargle periodogram.  
data(ibex, package = "lomb")  
lomb::lsp(ibex[2:3])
```



Part VII

Closing Matters

Acknowledgements

We would like to thank the APH consortium, for giving us the opportunity to write this manual, by granting us one of the 2017 APH mental health research funds.

We also thank the APH researchers, to whom we spoke during the project, for sharing their knowledge, experience and insights. We are also indebted to the researchers who contributed data, scripts, and other materials that are presented in this manual.

We also like to thank the project advisors, Prof. Dr. Heleen Riper and Prof. Dr. Johannes Smit, for their support and valuable input.

This manual was written in RMarkdown with Bookdown (Xie, 2016). We thank Yihui Xie for this wonderful package.

Jeroen Ruwaard, Lisa Kooistra & Melissa Thong



Bibliography

- Aan het Rot, M., Hogenelst, K., and Schoevers, R. A. (2012). Mood disorders in everyday life: A systematic review of experience sampling and ecological momentary assessment studies. *Clinical psychology review*, 32(6):510–23.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. Holt and Company.
- Asselbergs, J., Ruwaard, J., Ejdys, M., Schrader, N., Sijbrandij, M., and Riper, H. (2016). Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: An explorative study. *Journal of medical Internet research*, 18(3):e72.
- Aung, M. H., Matthews, M., and Choudhury, T. (2017). Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies. *Depression and anxiety*, 34(7):603–609.
- Bai, J., Di, C., Xiao, L., Evenson, K. R., LaCroix, A. Z., Crainiceanu, C. M., and Buchner, D. M. (2016). An activity index for raw accelerometry data and its comparison with other activity metrics. *PLoS ONE*.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.
- Beedie, C., Terry, P., and Lane, A. (2005). Distinctions between emotion and mood. *Cognition & Emotion*, 19(6):847–878.
- Bennik, E. C. (2015). *Every dark cloud has a colored lining: The relation between positive and negative affect and reactivity to positive and negative events*. University of Groningen, Groningen.
- Blaauw, F. J., Schenk, H. M., Jeronimus, B. F., Van der Krieke, L., De Jonge, P., Aiello, M., and Emerencia, A. C. (2016). Let's get Physiqual – An intuitive and generic method to combine sensor technology with ecological momentary assessments. *Journal of Biomedical Informatics*.
- Bolger, N. and Laurenceau, J.-p. (2013). Introduction to intensive methods. In *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press, New York.
- Booij, S. (2015). *Dynamics of the human stress system in depression: A combined population- and person-based approach to assess long-term changes and daily life fluctuations*. University of Groningen, Groningen.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*.
- Borsboom, D. and Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9(1):91–121.
- Bowen, R., Baetz, M., Hawkes, J., and Bowen, A. (2006). Mood variability in anxiety disorders. *Journal of Affective Disorders*, 91(2-3):165–170.
- Bringmann, L. F., Lemmens, L. H., Huibers, M. J., Borsboom, D., and Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*.
- Burton, C., McKinstry, B., Tătar, A. S., Serrano-Blanco, A., Pagliari, C., and Wolters, M. (2013). Activity monitoring in patients with depression: A systematic review. *Journal of affective disorders*, 145(1):21–28.

- Buyssse, D. J., Thompson, W., Scott, J., Franzen, P. L., Germain, A., Hall, M., Moul, D. E., Nofzinger, E. A., and Kupfer, D. J. (2007). Daytime symptoms in primary insomnia: A prospective analysis using ecological momentary assessment. *Sleep Medicine*, 8(3):198–208.
- CCMO (2018). Research with a medical device. Retrieved May 25 2018, from <http://www.ccmo.nl/en/research-with-a-medical-device>.
- Choi, L., Beck, C., Liu, Z., Matthews, C. E., and Buchowski, M. S. (2018). *PhysicalActivity: Process Accelerometer Data for Physical Activity Measurement*. R package version 0.2-2.
- Collip, D., Oorschot, M., Thewissen, V., Van Os, J., Bentall, R., and Myin-Germeys, I. (2011). Social world interactions: how company connects to paranoia. *Psychological medicine*, 41(5):911–921.
- Conner, T. S. and Feldman Barrett, L. (2012). Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosomatic Medicine*, 74(4):327–337.
- Conner, T. S. and Lehman, B. J. (2012). Getting started: Launching a study in daily life. In Mehl, M. R. and Conner, T. S., editors, *Handbook of research methods for studying daily life.*, pages 89–107. Guilford Press, New York, NY, US.
- Conner, T. S., Tennen, H., Fleeson, W., and Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and personality psychology compass*, 3(3):292–313.
- Cornet, V. P. and Holden, R. J. (2018). Systematic review of smartphone-based passive sensing for health and wellbeing. *Journal of Biomedical Informatics*, 17(1):120–132.
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., and Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, 32(7):917–929.
- Csikszentmihalyi, M. and Larson, R. (2014). Validity and reliability of the experience-sampling method. In *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*, pages 35–54. Springer Netherlands.
- Da silva, I. C., Van Hees, V. T., Ramires, V. V., Knuth, A. G., Bielemann, R. M., Ekelund, U., Brage, S., and Hallal, P. C. (2014). Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *International Journal of Epidemiology*.
- Dakos, V., Scheffer, M., Van Nes, E. H., Brovkin, V., Petoukhov, V., and Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 105 VN -(38):14308–14312.
- De Geus, E. and Van Doornen, L. (1996). Ambulatory assessment of parasympathetic/sympathetic balance by impedance cardiography. In *Ambulatory Assessment. Computer assisted psychological and psychophysiological methods in monitoring and field studies*, pages 141–164. Hogrefe & Huber.
- De Geus, E., Willemse, G. H., Klaver, C. H., and Van Doornen, L. J. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology*.
- Derogatis, L. R. (1994). *Symptom Checklist-90-R (SCL-90-R): Administration, scoring, and procedures manual*. Minneapolis, MN: NCS Pearson.
- Desmet, P. M. A., Vastenburg, M. H., and Romero, N. (2016). Mood measurement with Pick-A-Mood: Review of current methods and design of a pictorial self-report scale. *Journal of Design Research*, 14(3):241–279.
- Dogan, E., Sander, C., Wagner, X., Hegerl, U., and Kohls, E. (2017). Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? systematic review. *Journal of medical Internet research*, 19(7).
- Doherty, A. (2018). Circadian rhythms and mental health: wearable sensing at scale. *The Lancet Psychiatry*.

- Dubad, M., Winsper, C., Meyer, C., Livanou, M., and Marwaha, S. (2018). A systematic review of the psychometric properties, usability and clinical impacts of mobile mood-monitoring applications in young people. *Psychological medicine*, 48(2):208–228.
- Dunton, G. F. (2017). Ecological momentary assessment in physical activity research. *Exercise and Sport Sciences Reviews*, 45(1):48–54.
- Ebner-Priemer, U. W. and Trull, T. J. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*, 21(4):463–475.
- Ekker, A. and van Rest, B. (2013). *Medische apps, is certificeren nodig? In 7 stappen naar een CE-markering voor uw app*. Nictiz, Den Haag.
- Ellis-Davies, K., Sakkalou, E., Fowler, N. C., Hilbrink, E. E., and Gattis, M. (2012). CUE: The continuous unified electronic diary method. *Behavior Research Methods*, 44(4):1063–1078.
- Emerencia, A. C. (2018). *autovarCore: Automated Vector Autoregression Models and Networks*. R package version 1.0-4.
- Emerencia, A. C., Van der Krieke, L., Bos, E. H., De Jonge, P., Petkov, N., and Aiello, M. (2016). Automating vector autoregression on electronic patient diary data. *IEEE Journal of Biomedical and Health Informatics*.
- Engel, S. G., Crosby, R. D., Thomas, G., Bond, D., Lavender, J. M., Mason, T., Steffen, K. J., Green, D. D., and Wonderlich, S. A. (2016). Ecological momentary assessment in eating disorder and obesity research: A review of the recent literature.
- Epskamp, S., Borsboom, D., and Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50(1):195–212.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4):1–18.
- Evenson, K. R., Goto, M. M., and Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1):159.
- Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., and Li, L. C. (2018). Accuracy of fitbit devices: Systematic review and narrative syntheses of quantitative data. *Journal of Medical Internet Research*.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications Ltd., 4th edition.
- Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications Ltd, 1 edition.
- Frank, E., Swartz, H. A., and Boland, E. (2007). Interpersonal and social rhythm therapy: An intervention addressing rhythm dysregulation in bipolar disorder. *Dialogues in Clinical Neuroscience*.
- Frank, E., Swartz, H. A., and Kupfer, D. J. (2000). Interpersonal and social rhythm therapy: Managing the chaos of bipolar disorder. *Biological Psychiatry*.
- Gandrud, C. (2015). *Reproducible Research with R and RStudio*. CRC Press / Taylor and Francis Group, Boca Raton.
- Goërtz, Y. M. J., Looijmans, M., Prins, J. B., Janssen, D. J. A., Thong, M. S. Y., Peters, J. B., Burtin, C., Meertens-Kerris, Y., Coors, A., Muris, J. W. M., Sprangers, M. A. G., Wouters, E. F. M., Vercoulen, J. H., and Spruit, M. A. (2018). Fatigue in patients with chronic obstructive pulmonary disease: Protocol of the Dutch multicentre, longitudinal, observational FAntasTIGUE study. *BMJ Open*, 8(4):e021745.
- Goldfeld, K. (2018). *simstudy: Simulation of Study Data*. R package version 0.1.9.

- Gomersall, S. R., Ng, N., Burton, N. W., Pavey, T. G., and Gilson (2016). Estimating physical activity and sedentary behavior in a free-living context: A pragmatic comparison of consumer-based activity trackers and actigraph accelerometry. *Journal of Medical Internet Research*.
- Green, P. and MacLeod, C. J. (2016). simr: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4):493–498.
- Grolemund, G. and Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3):1–25.
- Groot, P. C. (2010). Patients can diagnose too: How continuous self-assessment aids diagnosis of, and recovery from, depression. *Journal of Mental Health*, 19(4):352–362.
- Haedt-Matt, A. A. and Keel, P. K. (2011). Hunger and binge eating: A meta-analysis of studies using ecological momentary assessment. *International Journal of Eating Disorders*, 44(7):573–578.
- Haedt-Matt, A. A., Zalta, A. K., Forbush, K. T., and Keel, P. K. (2012). Experimental evidence that changes in mood cause changes in body dissatisfaction among undergraduate women. *Body image*, 9(2):216–220.
- Hamaker, E. (2012). Why researchers should think "within-person": A paradigmatic rationale. *Handbook of Research Methods for Studying Daily Life*, pages 43–61.
- Heron, K. E. and Smyth, J. M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British journal of health psychology*, 15(1):1–39.
- Hildebrand, M., Van Hees, V. T., Hansen, B. H., and Ekelund, U. (2014). Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Medicine and Science in Sports and Exercise*.
- Hoogendoorn, M. and Funk, B. (2017). *Machine Learning for the Quantified Self: On the Art of Learning from Sensory Data*. Cognitive Systems Monographs. Springer International Publishing.
- Horton, N. J. and Kleinman, K. (2015). *Using R and RStudio for data management, statistical analysis, and graphics*. Chapman and Hall/CRC.
- Hufford, M. R., Shields, A. L., Shiffman, S., Paty, J., and Balabanis, M. (2002). Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. *Psychology of Addictive Behaviors*, 16(3):205–211.
- Jacobson, N. S. and Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1):12–19.
- Jongeneel, A., Scheffers, D., Tromp, N., Nuij, C., Delespaul, P., Riper, H., Van Der Gaag, M., and Van den Berg, D. (2018). Reducing distress and improving social functioning in daily life in people with auditory verbal hallucinations: Study protocol for the 'temstem'randomised controlled trial. *BMJ open*, 8(3):e020537.
- Karatsoreos, I. N. (2014). Links between Circadian Rhythms and Psychiatric Disease. *Frontiers in Behavioral Neuroscience*.
- Kemmeren, L., Van Schaik, D., Riper, H., Kleiboer, A., Bosmans, J. E., and Smit, J. H. (2016). Effectiveness of blended depression treatment for adults in specialised mental healthcare: study protocol for a randomised controlled trial. *BMC psychiatry*, 16(1):113.
- Kim, Y., White, T., Wijndaele, K., Sharp, S. J., Wareham, N. J., and Brage, S. (2017). Adiposity and grip strength as long-Term predictors of objectively measured physical activity in 93 015 adults: The UK Biobank study. *International Journal of Obesity*.
- Kirchner, T. R. and Shiffman, S. (2013). Ecological Momentary Assessment. In *The Wiley-Blackwell Handbook of Addiction Psychopharmacology*, pages 541–565. Wiley-Blackwell.

- Kirchner, T. R. and Shiffman, S. (2016). Spatio-temporal determinants of mental health and well-being: Advances in geographically-explicit ecological momentary assessment (gema). *Social psychiatry and psychiatric epidemiology*, 51(9):1211–1223.
- Kleiboer, A., Smit, J., Bosmans, J., Ruwaard, J., Andersson, G., Topooco, N., Berger, T., Krieger, T., Botella, C., Baños, R., Chevreul, K., Araya, R., Cerga-Pashoja, A., Cieślak, R., Rogala, A., Vis, C., Draisma, S., Schaik, A., Kemmeren, L., Ebert, D., Berking, M., Funk, B., Cuijpers, P., and Riper, H. (2016). European COMPARE Effective research on blended Depression treatment versus treatment-as-usual (E-COMPARED): Study protocol for a randomized controlled, non-inferiority trial in eight European countries. *Trials*, 17(1).
- Kossakowski, J. J., Groot, P. C., Haslbeck, J. M. B., Borsboom, D., and Wichers, M. (2017). Data from ‘critical slowing down as a personalized early warning signal for depression’. *Journal of Open Psychology Data*, 5(1).
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., and Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114(1-3):163–173.
- Larson, R. and Csikszentmihalyi, M. (1983). The experience sampling method. In Reis, H. T., editor, *Naturalistic approaches to studying social interaction*, pages 41–56. Jossey-Bass Inc Pub, San Francisco.
- Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*.
- Marszalek, J., Morgulec-Adamowicz, N., Rutkowska, I., Kosmol, A., Marszalek, J., Morgulec-Adamowicz, N., Rutkowska, I., and Kosmol, A. (2014). Using ecological momentary assessment to evaluate current physical activity. *BioMed Research International*, 2014,:e915172.
- Mehl, M. R. and Conner, T. S., editors (2012). *Handbook of research methods for studying daily life*. Guilford Press, New York, NY, US.
- Mikus, A., Hoogendoorn, M., Rocha, A., Gama, J., Ruwaard, J., and Riper, H. (2018). Predicting short term mood developments among depressed patients using adherence and ecological momentary assessment data. *Internet interventions*, 12:105–110.
- Moore, R. C., Depp, C. A., Wetherell, J. L., and Lenze, E. J. (2016). Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of Psychiatric Research*, 75:116–123.
- Myin-Germeys, I., Klippe, A., Steinhart, H., and Reininghaus, U. (2016). Ecological momentary interventions in psychiatry. *Current opinion in psychiatry*, 29(4):258–63.
- Nuij, C., Van Ballegooijen, W., Ruwaard, J., de Beurs, D., Mokkenstorm, J., van Duijn, E., de Winter, R. F., O’Connor, R. C., Smit, J. H., Riper, H., and Kerkhof, A. (2018). Smartphone-based safety planning and self-monitoring for suicidal patients: Rationale and study protocol of the CASPAR (Continuous Assessment for Suicide Prevention And Research) study. *Internet Interventions*, 13:16–23.
- Patel, S., Park, H., Bonato, P., Chan, L., and Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation*, 9(1):21.
- Penninx, B. W. J. H., Beekman, A. T. F., Smit, J. H., Zitman, F. G., Nolen, W. A., Spinhover, P., Cuijpers, P., de Jong, P. J., Van Marwijk, H. W. J., Assendelft, W. J. J., Van der Meer, K., Verhaak, P., Wensing, M., de Graaf, R., Hoogendijk, W. J., Ormel, J., Van Dyck, R., and NESDA Research Consortium (2008). The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives and methods. *International journal of methods in psychiatric research*, 17(3):121–40.
- Piasecki, T. M., Hufford, M. R., Solhan, M., and Trull, T. J. (2007). Assessing clients in their natural environments with electronic diaries: Rationale, benefits, limitations, and barriers. *Psychological Assessment*, 19(1):25–43.

- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2018). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-137.
- Pinheiro, J. and Bates, D. M. (2000). *Mixed effects models in S and S-Plus*. Springer Verlag.
- Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3):715–734.
- Pot-Kolder, R. M., Geraets, C. N., Veling, W., Van Beilen, M., Staring, A. B., Gijsman, H. J., Delespaul, P. A., and Van der Gaag, M. (2018). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: a single-blind randomised controlled trial. *The Lancet Psychiatry*, 5(3):217–226.
- Pot-Kolder, R. M., Veling, W., Geraets, C., and Van der Gaag, M. (2016). Effect of virtual reality exposure therapy on social participation in people with a psychotic disorder (VRETp): Study protocol for a randomized controlled trial. *Trials*, 17(1):25.
- Rocha, A., Camacho, R., Ruwaard, J., and Riper, H. (2018). Using multi-relational data mining to discriminate blended therapy efficiency on patients based on log data. *Internet interventions*, 12:176–180.
- Roenneberg, T., Wirz-Justice, A., and Merrow, M. (2003). Life between clocks: Daily temporal patterns of human chronotypes. *Journal of biological rhythms*, 18(1):80–90.
- Rowlands, A. V., Yates, T., Davies, M., Khunti, K., and Edwardson, C. L. (2016). Raw Accelerometer Data Analysis with GGIR R-package: Does Accelerometer Brand Matter? *Medicine and Science in Sports and Exercise*.
- Ruf, T. (1999). The Lomb-Scargle periodogram in biological rhythm research: Analysis of incomplete and unequally spaced time-series. *Biological Rhythm Research*, 30:178–201.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161.
- Saeb, S., Lattie, E. G., Kording, K. P., and Mohr, D. C. (2017). Mobile Phone Detection of Semantic Location and Its Relationship to Depression and Anxiety. *JMIR mHealth and uHealth*.
- Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., and Mohr, D. C. (2015). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research*, 17(7).
- Sander, D. and Scherer, K. (2009). *Oxford companion to emotion and the affective sciences*. Oxford University Press.
- Saunders, K., Palmius, N., de Vos, M., Bilderbeck, A., Geddes, J., and Goodwin, G. (2016). Depression detection in bipolar disorder using geolocation data. *Bipolar Disorders*.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., Van Nes, E. H., Rietkerk, M., and Sugihara, G. (2009). Early-warning signals for critical transitions.
- Schueller, S. M., Aguilera, A., and Mohr, D. C. (2017). Ecological momentary interventions for depression and anxiety. *Depression and anxiety*, 34(6):540–545.
- Schuurman, N. K., Houtveen, J. H., and Hamaker, E. (2015). Incorporating measurement error in n=1 psychological autoregressive modeling. *Frontiers in psychology*, 6:1038.
- Serre, F., Fatseas, M., Swendsen, J., and Auriacombe, M. (2015). Ecological momentary assessment in the investigation of craving and substance use in daily life: A systematic review.
- Sharar, S. R., Alamdari, A., Hoffer, C., Hoffman, H. G., Jensen, M. P., and Patterson, D. R. (2016). Circumplex model of affect: A measure of pleasure and arousal during virtual reality distraction analgesia. *Games for health journal*, 5(3):197–202.

- Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*, 21(4):486–497.
- Shiffman, S., Stone, A. A., Hufford, M. R., Rev, A., Psychol, C., Shiffman, S., Stone, A. A., and Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, 4(1):1–32.
- Slofstra, C., Klein, N. S., Nauta, M. H., Wicher, M., Batalas, N., and Bockting, C. L. (2017). Imagine your mood: Study design and protocol of a randomized controlled micro-trial using app-based experience sampling methodology to explore processes of change during relapse prevention interventions for recurrent depression. *Contemporary clinical trials communications*, 7:172–178.
- Smyth, J. M. and Stone, A. A. (2003). Ecological momentary assessment research in behavioral medicine. *Journal of Happiness Studies*, 4(1):35–52.
- Starr, L. R. and Davila, J. (2012). Temporal patterns of anxious and depressed mood in generalized anxiety disorder: A daily diary study. *Behaviour Research and Therapy*.
- Stone, A. A. and Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine*, 24(3):236–243.
- Swendsen, J. (2016). Contributions of mobile technologies to addiction research. *Dialogues in clinical neuroscience*, 18(2):213.
- Tahmasian, M., Khazaie, H., Golshani, S., and Avis, K. T. (2013). Clinical application of actigraphy in psychotic disorders: A systematic review. *Current Psychiatry Reports*.
- Te Lindert, B. H., Itzhacki, J., Van der Meijden, W. P., Kringselbach, M. L., Mendoza, J., and Van Someren, E. (2018). Bright environmental light ameliorates deficient subjective ‘liking’ in insomnia: An experience sampling study. *Sleep*, 41(4):1–13.
- Telford, C., McCarthy-Jones, S., Corcoran, R., and Rowse, G. (2012). Experience sampling methodology studies of depression: The state of the art. *Psychological Medicine*, 42(6):1119–1129.
- Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., and Bates, K. K. (2010). Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1550):2163–2176.
- Twisk, J. (2006). *Applied multilevel analysis: a practical guide*. Cambridge University Press.
- Twisk, J. (2013). *Applied longitudinal data analysis for epidemiology: A practical guide (2nd Revised edition)*. Cambridge University Press.
- Van Ballegooijen, W., Ruwaard, J., Karyotaki, E., Ebert, D. D., Smit, J. H., and Riper, H. (2016). Reactivity to smartphone-based ecological momentary assessment of depressive symptoms (MoodMonitor): Protocol of a randomised controlled trial. *BMC Psychiatry*, 16(1):359.
- Van Breda, W., Bremer, V., Becker, D., Hoogendoorn, M., Funk, B., Ruwaard, J., and Riper, H. (2018). Predicting therapy success for treatment as usual and blended treatment in the domain of depression. *Internet Interventions*.
- Van de Ven, P., O’Brien, H., Henriques, R., Klein, M., Msetfi, R., Nelson, J., Rocha, A., Ruwaard, J., O’Sullivan, D., and Riper, H. (2017). ULTEMAT: A mobile framework for smart ecological momentary assessments and interventions. *Internet Interventions*, 9:74–81.
- Van der Krieke, L., Blaauw, F. J., Emerencia, A. C., Schenk, H. M., Slaets, J. P., Bos, E. H., De Jonge, P., and Jeronimus, B. F. (2017). Temporal dynamics of health and well-being: A crowdsourcing approach to momentary assessments and automated generation of personalized feedback. *Psychosomatic Medicine*.
- Van der Krieke, L., Emerencia, A. C., Bos, H. E., Rosmalen, G. J., Riese, H., Aiello, M., Sytema, S., and de Jonge, P. (2015). Ecological momentary assessments and automated time series analysis to promote tailored health care: A proof-of-principle study. *JMIR Research Protocols*, 4(3):e100.

- Van der Krieke, L., Jeronimus, B. F., Blaauw, F. J., Wanders, R. B., Emerencia, A. C., Schenk, H. M., Vos, S. D., Snippe, E., Wichers, M., Wigman, J. T., Bos, E. H., Wardenaar, K. J., and De Jonge, P. (2016). HowNutsAreTheDutch (HoeGekIsNL): A crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*.
- Van der Meijden, W. P., Van Someren, J. L., Te Lindert, B. H., Bruijel, J., Van Oosterhout, F., Coppens, J. E., Kalsbeek, A., Cajochen, C., Bourgin, P., and Van Someren, E. (2016). Individual differences in sleep timing relate to melanopsin-based phototransduction in healthy adolescents and young adults. *Sleep*, 39(6):1305–1310.
- Van Hees, V. T., Fang, Z., Langford, J., Assah, F., Mohammad, A., da Silva, I. C., Trenell, M. I., White, T., Wareham, N. J., and Brage, S. (2014). Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: An evaluation on four continents. *Journal of Applied Physiology*, 117(7):738–744.
- Van Hees, V. T., Fang, Z., Zhao, J. H., Heywood, J., Mirkes, E., Sabia, S., and Migueles, J. H. (2018). *GGIR: Raw Accelerometer Data Analysis*. R package version 1.5-24.
- Van Hees, V. T., Sabia, S., Anderson, K. N., Denton, S. J., Oliver, J., Catt, M., Abell, J. G., Kivimaki, M., Trenell, M. I., and Singh-Manoux, A. (2015). A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS One*, 10(11).
- Van Rijsbergen, G. D., Bockting, C. L. H., Berking, M., Koeter, M. W. J., and Schene, A. H. (2012). Can a one-item mood scale do the trick? Predicting relapse over 5-years in recurrent depression. *Plos One*, 7(10):1–5.
- Van Rijsbergen, G. D., Burger, H., Hollon, S. D., Elgersma, H. J., Kok, G. D., Dekker, J., de Jong, P. J., and Bockting, C. L. (2014). How do you feel? Detection of recurrent Major Depressive Disorder using a single-item screening tool. *Psychiatry Research*, 220(1-2):287–293.
- Van Someren, E. (2000). Circadian rhythms and sleep in human aging. *Chronobiology International*.
- Van Someren, E. (2011). Chapter 4 - Actigraphic monitoring of sleep and circadian rhythms. In Montagna, P. and Chokroverty, S., editors, *Sleep Disorders Part I*, volume 98 of *Handbook of Clinical Neurology*, pages 55 – 63. Elsevier.
- Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.
- Versluis, A., Verkuil, B., Spinhoven, P., Van der Ploeg, M. M., and Brosschot, J. F. (2016). Changing mental health and positive psychological well-being using ecological momentary interventions: A systematic review and meta-analysis. *Journal of medical Internet research*, 18(6).
- Walz, L. C., Nauta, M. H., and Aan het Rot, M. (2014). Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: A systematic review. *Journal of Anxiety Disorders*, 28(8):925–937.
- Warmerdam, L., Riper, H., Klein, M. C., Van de Ven, P., Rocha, A., Henriques, M. R., Tousset, E., Silva, H., Andersson, G., and Cuijpers, P. (2012). Innovative ict solutions to improve treatment outcomes for depression: the ICT4Depression project. *Annual Review of Cybertherapy and Telemedicine*, 181(1):339–343.
- Watson, D. and Clark, L. A. (1994). *The PANAS-X : Manual for the Positive and Negative Affect Schedule - expanded Form*. University of Iowa.
- Watson, D. and Clark, L. A. (1997). Measurement and mismeasurement of mood: Recurrent and emergent issues.
- Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect : The PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070.

- Watson, D. and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin*, 98(2):219.
- Wenze, S. J. and Miller, I. W. (2010). Use of ecological momentary assessment in mood disorders research. *Clinical psychology review*, 30(6):794–804.
- Wichers, M., Groot, P. C., Psychosystems, ESM, and EWS (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, 85(2):114–116.
- Wichers, M., Peeters, F., Rutten, B. P. F., Jacobs, N., Derom, C., Thiery, E., Delespaul, P., and Van Os, J. (2012). A time-lagged momentary assessment study on daily life physical activity and affect. *Health Psychology*, 31(2):135–144.
- Wichers, M., Simons, C. J. P., Kramer, I. M. A., Hartmann, J. A., Lothmann, C., Myin-Germeys, I., Van Bemmel, A. L., Peeters, F., Delespaul, P., and Van Os, J. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica*, 124(4):262–272.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.
- Wickham, H., François, R., Henry, L., and Müller, K. (2018). *dplyr: A Grammar of Data Manipulation*. R package version 0.7.6.
- Wickham, H. and Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. " O'Reilly Media, Inc. ".
- Wickham, H. and Miller, E. (2018). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. R package version 1.1.2.
- Wikipedia (2018). Stationary process — Wikipedia, The free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Stationary%20process&oldid=866157327>. [Online; accessed 30-October-2018].
- Wilkinson, L. (2006). *The grammar of graphics*. Springer Science & Business Media.
- Willemse, G. H., De Geus, E., Klaver, C. H., Van Doornen, L. J., and Carroll, D. (1996). Ambulatory monitoring of the impedance cardiogram. *Psychophysiology*.
- Xie, Y. (2016). *bookdown: Authoring Books and Technical Documents with R Markdown*. Chapman and Hall/CRC, Boca Raton, Florida. ISBN 978-1138700109.

Index

- Accelerometer, 44
Actant - Activity Analysis Toolbox, 95
Actigraphy, 44
Auto-correlation, 55
Autonomic and cardiovascular activity, 89, 95

CASPAR, 85
Chronic Fatigue, 81
Circadian rhythms, 56
CRAN, 13

DAMS, 89, 95
Datamanagement, 27
 Project directory structure, 28
 Raw data, 29
Datetime variables, 103

E-COMPARED, 80
Ecological Momentary Intervention, 5
Ecological validity, 4
EMA, 3
 Active EMA, 3, 6, 22
 Observational EMA, *see* Passive EMA
 Passive EMA, 3, 6, 22
 Self-report EMA, *see* Active EMA
EMA research
 Activity patterns, 37
 Anxiety disorders, 7
 Eating disorders, 7
 Mental health/ Well-being, 7
 Mood disorders, 7, 37
 Somatic health, 37
 Substance-related disorders, 7, 37
EMA synonyms, 3
 Ambulatory Assessment, 3
 Ambulatory Self-reporting, 3
 Continuous Unified Electronic Diary Method, 3
 Experience Sampling, 3
 Intensive-longitudinal Study Design, 3
 Real-time Data Capturing, 3
EMI, *see* Ecological Momentary Intervention
ENMO, 45

FAntasTIQUE, 82
Fixed effects, 61

GENEActiv, 95
Geographic Coordinate system, 46
Geotracking, 46
GitHub, 14
GPS, 46

ICT4Depression, 80
Ideographic research, 5
Ilumivu, 84, 85, 91
Imagine your Mood, 78
IMPACT, 82, 83
Intra-individual process, *see* Ideographic research

Latitude, 46
Limitations, 6
Lomb-Scargle periodogram, 108
Longitude, 46

Measurement error, 22
MERITS study, 84
Methods, 21
 Ethical considerations, 25
 Power analysis, 23
 Study design, 21
Missing values, 54
mixed effects, 104
Mixed Modeling, 61
Mood, 37
Mood assessment, 37
 Bag-of-Items, 38
 Bipolar unidimensional, 38
 Circumplex model, 39
 Multi-dimensional, 39
 PA/NA, 41
 Pick-a-mood scale, 39
 Unidimensional, 37
Mood definitions, 37
 Emotions, 37
 Mood states, 37
Moodbuster, 80, 92
Movisens, 92
MVPA, 45

- NESDA, 79
Network analysis, 105
Nomothetic research, 5
- Packages
 autovar, 100
 autovarCore, 100
 bootnet, 107
 dplyr, 101
 emaph, 14
 GENEAread, 98
 GGIR, 98
 ggplot2, 102
 haven, 103
 lme4, 104
 lomb, 108
 lubridate, 103
 nlme, 62, 104
 nonlinearTseries, 69
 PhysicalActivity, 99
 qgraph, 106
 simr, 104
 swirl, 17
 tidyverse, 13, 101
- Positive and Negative Affect, 41
 PANAS, *see* Mood assessment
- PsyMate, 78, 82, 83, 86, 93
- R and RStudio, 9
 Console, 10
 Data access, 12
 Data import, 11
 Help, 15
 Installing, 10
 Packages, 14
 Scripts, 11
 Why R, 9
- R packages, 13
RADAR-CNS, 79, 80
Random effects, 61
Recall bias, 4
RoQua, 78, 79, 81, 94
RStudio Projects, 27
- Sample plan, *see* Sampling
Sampling, 22
 Event-contingent, 4
 Signal-contingent, 4
- Simulating data, 24, 51, 57, 61
Sleep, 87
SPSS, 103
Stationarity, 52
StayFine, 78
SVM, 45
Symptom networks, 105
- TemStem, 85, 86
Unevenly-sampled timeseries, 108
- Vector autoregressive models, 100
VRETP, 85, 86
VU-AMS, 89, 95