Methods

Methods

Multilevel AR(1) model

In this subchapter, I will describe the mathematical basis, assumptions and estimation procedures of the first-order multilevel autoregressive (MLAR(1)) model with random intercepts and random autoregressive effects, which is the focus of the simulation part of the thesis. While the notation for the model varies across different papers, the notation used by Lafit et al. (2020) will be adhered to throughout the thesis.

The MLAR model consists of two levels: the within-person Level 1 and the between-person Level 2. At level 1, described by Equation (1) (Lafit et al. 2020), each participant's first-order autoregressive process is modelled: The person-specific autoregressive parameter (inertia) γ_{1i} quantifies to what degree the process value esm_{it} of participant i at time t depends on the lagged process value $esm_{i,t-1}$. The person-specific intercept γ_{0i} represents the expected process value esm_{it} when the lagged variable $esm_{i,t-1}$ equals 0 (Jongerling, Laurenceau, and Hamaker 2015). Following the recommendations by (Hamaker and Grasman 2015), the predictor (lagged variable) was person-mean centered. The innovation ϵ_{it} (i.e., residuals, the part of the variance that is not explained by the lagged variable $esm_{i,t-1}$) is assumed to be independent and coming from a normal distribution with mean of 0 and variance σ_e^2 (Lafit et al., 2020). The model used in the present thesis assumes that the innovation variance is identical for all participants.

$$esm_{it} = \gamma_{0i} + \gamma_{1i} * esm_{i,t-1} + \epsilon_{it} \tag{1}$$

In the multilevel AR(1) model, the person-specific autoregressive effects 1i and the person-specific intercepts 0i are allowed to vary between participants. The Level 2 of the MLAR(1) model describes this between-person variability. The Level 2 is defined in Equation (2)). Each person-specific autoregressive effect γ_{1i} is a sum of a fixed effect β_{10} and a person-specific random effect ν_{1i}. The random effects ν_{1i} themselves come from a normal distribution with mean of 0 and variance $\sigma_{\nu 1}^2$ (Lafit et al., 2020). The same holds for the person-specific intercepts γ_{0i}: they are a sum of a fixed effect β_{00} and a random effect ν_{0i} that comes from $N(0,\, \sigma_{\nu 0}^2)$.

$$\begin{aligned} \gamma_{0i} &= \beta_{00} + \nu_{0i} \\ \gamma_{1i} &= \beta_{10} + \nu_{1i} \end{aligned} \tag{2}$$

Assumptions of the MLAR(1) model

In this part, the assumptions of the MLAR(1) model and the way they were taken into account in the present simulation study will be explained.

Stationarity. The MLAR(1) model is used to model stable processes in which no temporal trends are present. As such, it assumes weak stationarity: the (person-specific) process mean, innovation variance, and autoregressive parameter are assumed to not change through the time series (Rovine and Walls 2006). For this reason, the person-specific autoregressive effects γ_{1i} are assumed to be bounded by -1 and 1, as autoregressive effects larger than 1 (or lower than -1) cause a change in the process mean (Krone, Albers, and Timmerman 2016).

Exogeneity. TODO (see word file)

Equally spaced measurements. TODO (see word file again)

Estimation procedures for the MLAR(1) model

Initial conditions problem. TODO (see Word)

Multilevel AR(1) model and statistical power

Simulation study

The goal of the present exploratory simulation study is to assess the effects of four different patterns of missing data (data missing completely random, data missing in blocks, and two patterns of data missing dependent on process value) on estimation performance/bias, standard error and statistical power for the estimation of the fixed autoregressive effect in the MLAR(1) model.

Simulation procedure

The study followed the general principles of a Monte Carlo simulation procedure described by Lane & Hennes (2018).

Simulation conditions. Two simulation studies, Simulation A and Simulation B, were carried out to investigate the research questions. In Simulation A, no random autoregressive effects were simulated and estimated (i.e., each subject's time-series in the simulation had the same simulated autoregressive effect, and only fixed autoregressive effects were estimated). In Simulation B, random autoregressive effects were simulated and estimated (with the random effects variance set to either 0.05 or 0.1). Both random and fixed intercepts were estimated in Simulations A and B.

Simulation A followed a 4 × 2 × 3 × 4 × 3 factorial design (yielding 288 simulation conditions in total), and Simulation B followed a 4 × 2 × 2 × 4 × 2 × 2 design (256 conditions in total). Each of the conditions was simulated in 1,000 simulation runs. As such, 544,000 datasets were generated (and the same number of models was estimated) in this simulation study. The manipulated variables (as well as the other parameters that remained fixed throughout all simulation conditions) are listed in Table 1 and 2.

```
library(kableExtra)

r1 = c(
"Missingness pattern", "MCAR, block, extreme-onesided, extreme-twosided", "MCAR, block, extreme-onesided

df_manipulated = matrix(r1, ncol = 3, byrow = TRUE)

kable(df_manipulated, booktabs = TRUE, align = "ccc", caption = "Values of the manipulated parameters u
```

```
r2 = c("Fixed intercept", "0", "0","Variance of random intercepts", "3", "3", "Innovation variance", "3

df_stable = matrix(r2, ncol = 3, byrow = TRUE)

kable(df_stable, booktabs = TRUE, align = "ccc", caption = "Parameters used for the two simulation studi
```

Hamaker, Ellen L., and Raoul P. P. P. Grasman. 2015. "To Center or Not to Center? Investigating Inertia with a Multilevel Autoregressive Model." *Frontiers in Psychology* 5 (January). https://doi.org/10.3389/fpsyg.2014.01492.

Table 1: Values of the manipulated parameters used in the two simulation studies

Manipulated parameter	Simulation A	Simulation B
Missingness pattern	MCAR, block, extreme-onesided, extreme-twosided	MCAR, block, extreme-onesided, extreme-twosided
Simulated fixed AR effect	0.3, 0.5, 0.7	0.3, 0.7
Variance of random AR effects	-	0.05, 0.1
Compliance	0.4, 0.6, 0.8, 1	0.4, 0.6, 0.8, 1
Number of participants (N)	20, 50	20, 50
Timepoints per participant (T.obs)	20, 50, 100	50, 100

Table 2: Parameters used for the two simulation studies.

Simulation parameter	Simulation A	Simulation B
Fixed intercept	0	0
Variance of random intercepts	3	3
Innovation variance	3	3
Correlation between random intercepts and random slopes	0	0
Significance threshold	0.05	0.05
Simulation runs per condition	1000	1000

Jongerling, Joran, Jean-Philippe Laurenceau, and Ellen L. Hamaker. 2015. "A Multilevel AR(1) Model: Allowing for Inter-Individual Differences in Trait-Scores, Inertia, and Innovation Variance." *Multivariate Behavioral Research* 50 (3): 334–49. https://doi.org/10.1080/00273171.2014.1003772.

Krone, Tanja, Casper J. Albers, and Marieke E. Timmerman. 2016. "Comparison of Estimation Procedures for Multilevel AR(1) Models." *Frontiers in Psychology* 7 (April). https://doi.org/10.3389/fpsyg.2016.00486.

Lafit, Ginette, Janne Adolf, Egon Dejonckheere, Inez Myin-Germeys, Wolfgang Viechtbauer, and Eva Ceulemans. 2020. "Selection of the Number of Participants in Intensive Longitudinal Studies: A User-Friendly Shiny App and Tutorial to Perform Power Analysis in Multilevel Regression Models That Account for Temporal Dependencies." https://doi.org/10.31234/osf.io/dq6ky.

Lane, Sean P., and Erin P. Hennes. 2018. "Power Struggles: Estimating Sample Size for Multilevel Relationships Research." *Journal of Social and Personal Relationships* 35 (1): 7–31. https://doi.org/10.1177/0265407517710342.

Rovine, Michael J, and Theodore A Walls. 2006. "Multilevel Autoregressive Modeling of Interindividual Differences in the Stability of a Process." *Models for Intensive Longitudinal Data*, 124147.