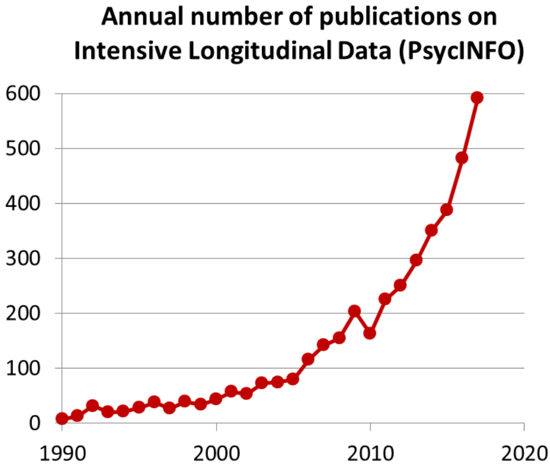
Times are changing



- Adapted from Hamaker & Wichers (2017).

Today's Workshop: Hosts

Utrecht University's Dynamic Modeling Lab

Oisín Ryan

- Causal analysis
- Differential Equation Models
- Theory formation
- website

Noémi Schuurman

- Multilevel modeling (DSEM)
- Bayesian stats
- Measurement
- website

Today's Workshop: Schedule

- 09.00-10.00 Lecture: Discrete-Time Modeling
- 10.00 – 10.45 Lab: Discrete-Time Modeling (n=1)
- 10.45 – 11.45 Lecture: Continuous-Time Modeling
- 11.45 – 12.30 Lab: Continuous-Time Modeling (n=1)
- 12.30 – 13.00 Plenary Discussion

Discrete Time Modeling in Psychology

N. K. Schuurman

Utrecht University
EAM

July 2021

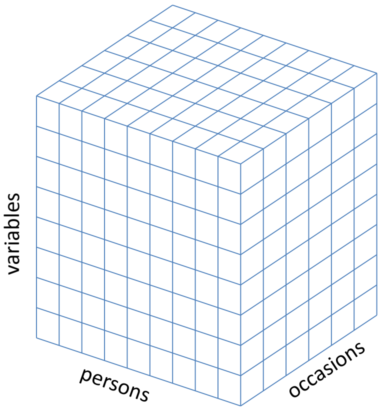
Overview

- Intensive Longitudinal Data
- Single Subject Univariate Autoregressive Modeling
- Single Subject Multivariate (Vector) Autoregressive Modeling
- Some Advanced Issues

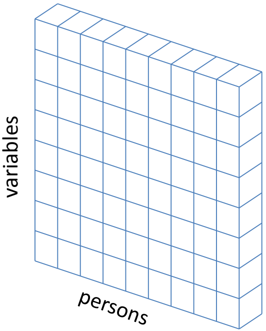
Overview

- **Intensive Longitudinal Data**
- Single Subject Univariate Autoregressive Modeling
- Single Subject Multivariate (Vector) Autoregressive Modeling
- Some Advanced Issues

Cattell's data box

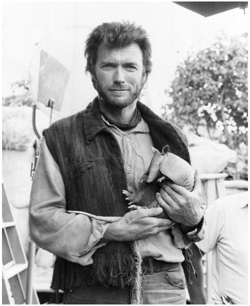


Cross-sectional research: N is large, T=1

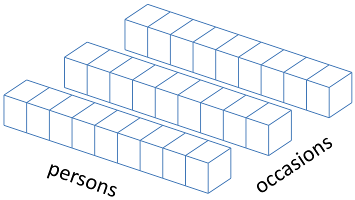


Cross-sectional research: N is large, T=1



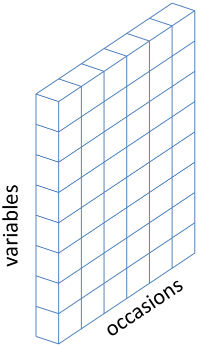


Panel research: N is large, T is small

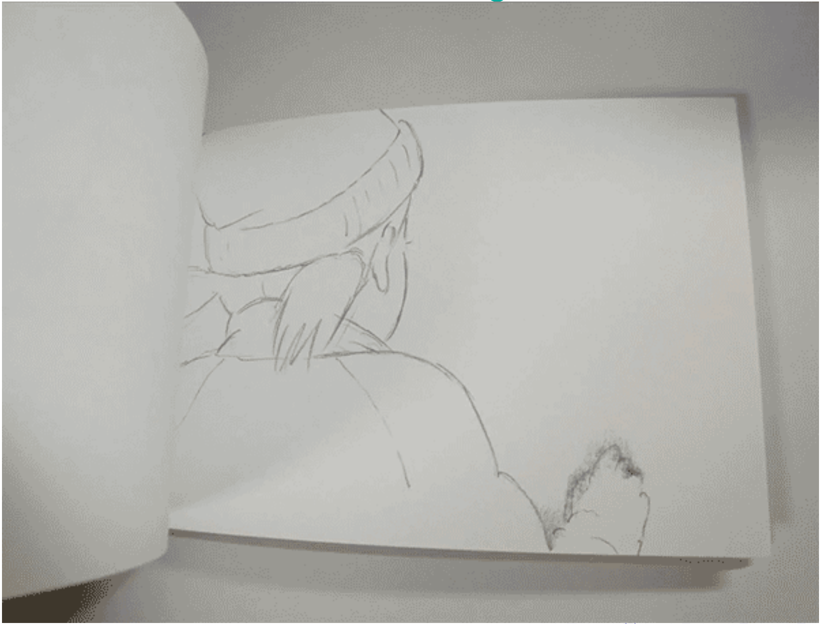


Panel research: N is large, T is small

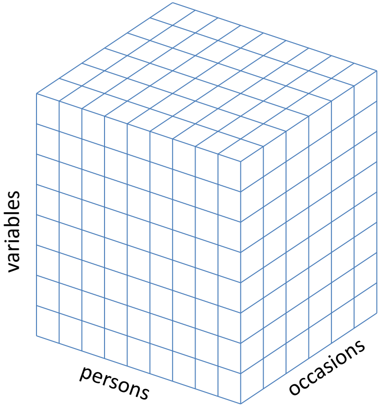
Time series data: N=1 and T is large



Time series data: N=1 and T is large



Intensive Longitudinal Data



Characteristics of these kind of data

Data structure:

- one or more measurements per day
- typically for multiple days
- sometimes multiple waves (i.e., Nesselroade's measurement-burst design)

Characteristics of these kind of data

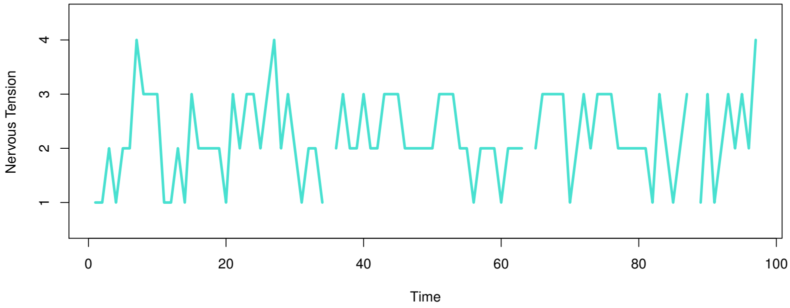
Data structure:

- one or more measurements per day
- typically for multiple days
- sometimes multiple waves (i.e., Nesselroade's measurement-burst design)

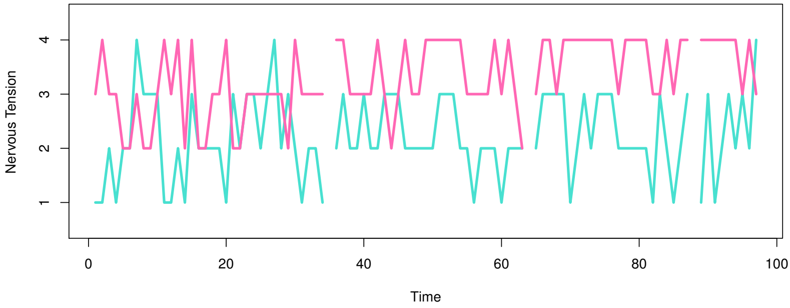
Advantages of ESM, EMA and AA

- no recall bias
- high ecological validity
- physiological measures over a large time span
- monitoring of symptoms and behavior, with new possibilities for feedback and intervention (e-Health and m-Health)
- window into the dynamics of processes

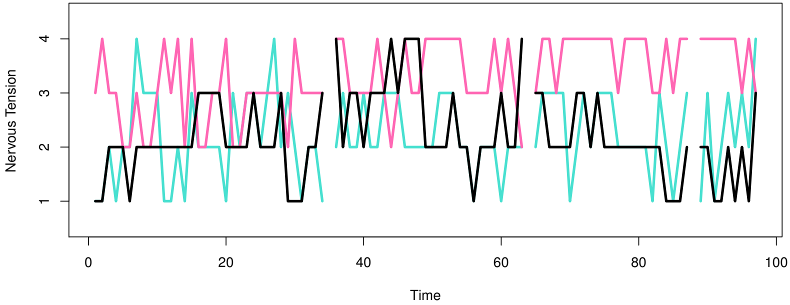
Time Series



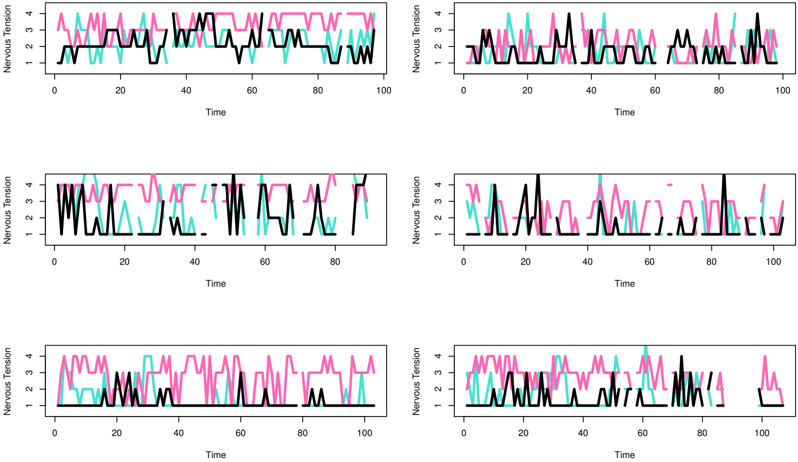
Multivariate Time Series



Multivariate Time Series



Intensive Longitudinal Data



ILD research in psychology

Different forms of intensive longitudinal data:

- daily diary (DD); self-report end-of-day
- experience sampling method (ESM); self-report of subjective experience
- event-based measurements; self-report after a particular event
- observational measurements; expert rater
- physiological measurements

ILD research in psychology

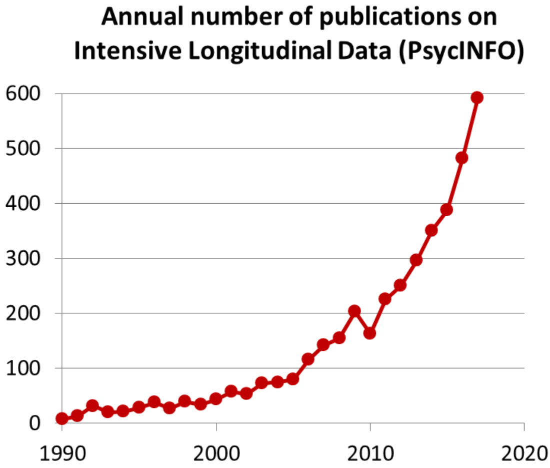
Different forms of intensive longitudinal data:

- daily diary (DD); self-report end-of-day
- experience sampling method (ESM); self-report of subjective experience
- event-based measurements; self-report after a particular event
- observational measurements; expert rater
- physiological measurements

Critiques of ILD research:

- within-person fluctuations are just **noise**
- results are **not generalizable**
- no one has these data

Times are changing



- Adapted from Hamaker & Wichers (2017).

Collecting Intensive Longitudinal Data

Ambulatory Assessment or Ecological Momentary Assessment

Experience Sampling, Daily diary, Tracking apps...See work by Timothy Trull and Ulrich Ebner-PriemerSociety of Ambulatory AssessmentLifedata, Ethica, Movisens, Expimetrics, ...

Collecting Daily Diary Data

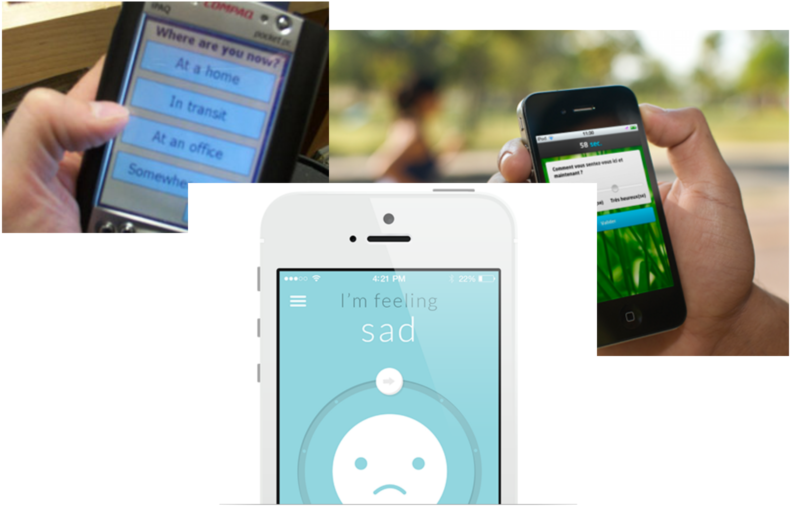
usually once at the end of the a day

Collecting Daily Diary Data

usually once at the end of the a day

Collecting Experience Sampling Data

Alert people randomly throughout the day



Tamlin Conner: https://www.youtube.com/watch?v=nQBBVp9vBIQ

Collection: Monitoring or Tracking Technology

Collection: Monitoring or Tracking Technology

Collection: Monitoring or Tracking Technology

Collection: Ambulatory/Ecological Momentary Assessment

Advantages

- limited recall bias
- high ecological validity
- allows for consistent monotoring, with new possibilities for feedback and intervention
- window into the dynamics of processes

How to Analyze This Stuff?

- Fairly young methodological area
- Not part of basic curriculum
- Huge development
- Already many options: discrete or continuous variables, latent variables, linear models, nonlinear models, and so on (Hamaker et al. 2015).

Overview

- **Intensive Longitudinal Data**
- Single Subject Univariate Autoregressive Modeling
- Single Subject Multivariate (Vector) Autoregressive Modeling
- Some Advanced Issues

What is time series analysis?

Time series analysis is a class of techniques that is used in econometrics, seismology, meteorology, control engineering, and signal processing.

What is time series analysis?

Time series analysis is a class of techniques that is used in econometrics, seismology, meteorology, control engineering, and signal processing.

Main characteristics:

- N=1 technique
- T is large (say $>$50)
- concerned with *trends*, *cycles* and *autocorrelation structure* (i.e., serial dependency)
- goal: forecasting ($\neq$ prediction)

Lags

Y
y_1
y_2
y_3
y_4
y_5
y_6
y_7
y_8
$\dots$
y_T

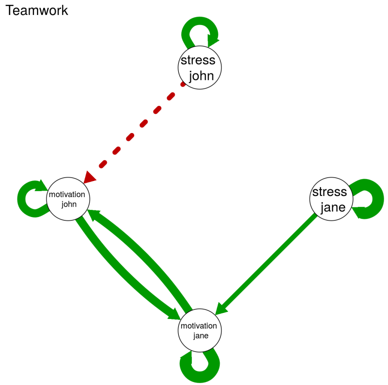
Lags

Y	Y at lag 1
y_1	
y_2	y_1
y_3	y_2
y_4	y_3
y_5	y_4
y_6	y_5
y_7	y_6
y_8	y_7
$\dots$	$\dots$
y_T	y_{T-1}
	y_T

Lags

Y	Y at lag 1	Y at lag 2
y_1		
y_2	y_1	
y_3	y_2	y_1
y_4	y_3	y_2
y_5	y_4	y_3
y_6	y_5	y_4
y_7	y_6	y_5
y_8	y_7	y_6
$\dots$	$\dots$	$\dots$
y_T	y_{T-1}	y_{T-2}
	y_T	y_{T-1}
		y_T

Simple models: Autoregressive Modeling



Why?

- Simple model (linear regression relationships, continuous variables)
- Appealing interpretation
- Basis for or related to many other dynamic models
- Can use coefficients to make pretty dynamic networks
- ***Hence, popular***

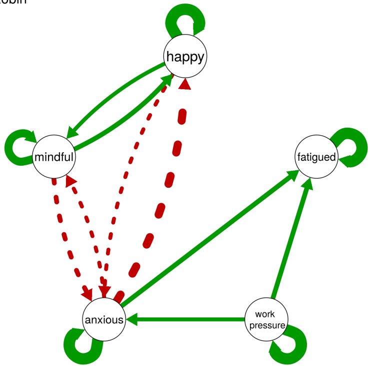
Intermezzo: Dynamic Networks/Intraindividual Networks

- Visualize how psychological variables are associated with themselves, and each other over time
- Conceptual models, or based on statistical estimates from (intensive longitudinal) data
- Currently, such statistical estimates are typically based on Vector Autoregressive Models

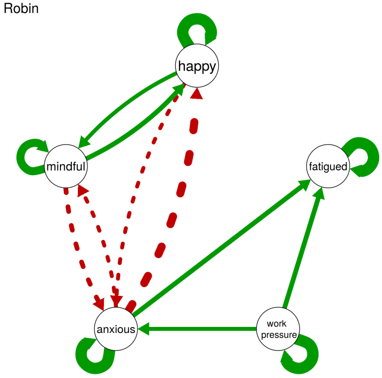
Read more: Borsboom (2017), Bringmann et al (2013), Cramer et al (2010).

Intermezzo: Dynamic Networks/Intraindividual Networks

Robin



Intermezzo: Dynamic Networks/Intraindividual Networks

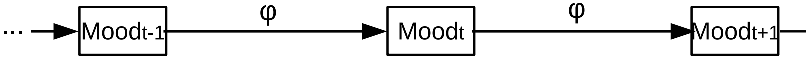


Autoregressive Modeling: The Basic Idea

> “The best predictor of future behavior is past behavior”

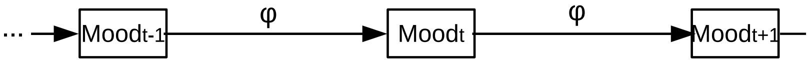
The N=1 Univariate Model (AR Model)

- Model for the time series of a specific person (N=1, T=many)
- Variable is regressed on itself at (a) previous occasion(s)
- AR(1) model: on the nearest previous occasion

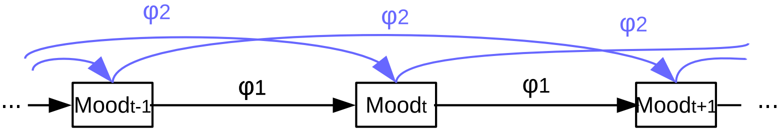


The N=1 Univariate Model (AR Model)

- AR(1) model: on the nearest previous occasion



- AR(2) model: on the nearest previous occasion, and the occasion before that



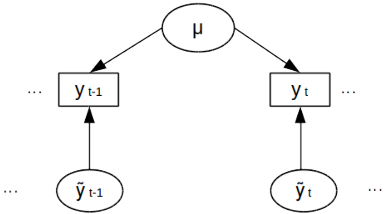
- AR(3) model: on the nearest previous occasion, and the occasion before that, and the one before that
- etc

The N=1 AR(1) Model

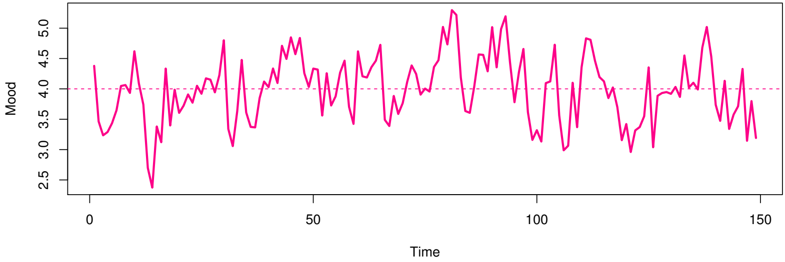
$$\cdots \rightarrow \boxed{Mood_{t-1}} \xrightarrow{\varphi} \boxed{Mood_{t}} \xrightarrow{\varphi} \boxed{Mood_{t+1}} -$$

Mood t	Mood t-1
5	.
3	5
3	3
4	3
2	4
3	2
1	3
1	1
2	1
.	2

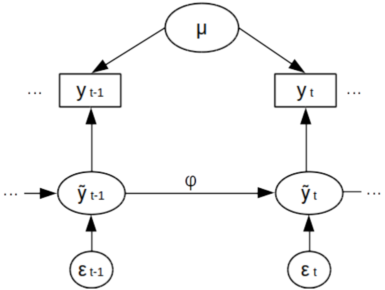
The N=1 AR(1) Model: Delving Deeper



$$y_t = \mu + \tilde{y}_t$$

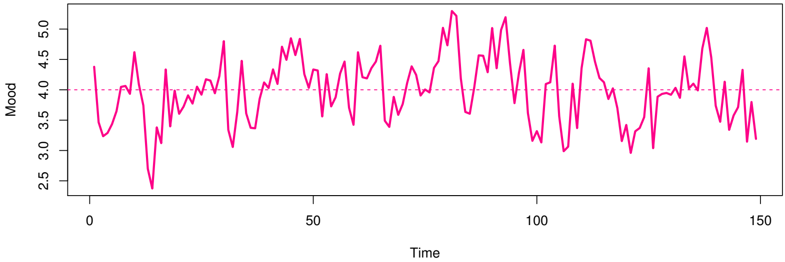


The N=1 AR(1) Model: Delving Deeper



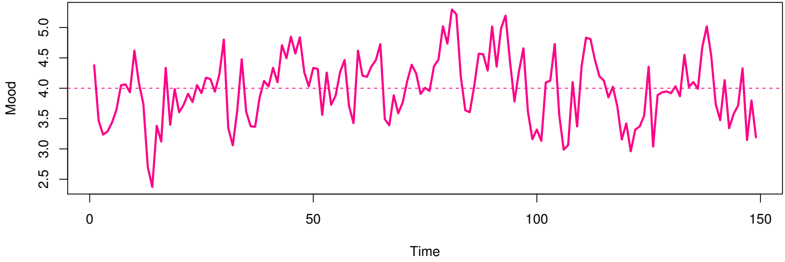
$$y_t = \mu + \tilde{y}_t$$
$$\tilde{y}_t = \phi \tilde{y}_{t-1} + \epsilon_t$$

$$\epsilon_t \sim Normal\left(0, \sigma^2\right)$$



The N=1 AR(1) Model: Delving Deeper

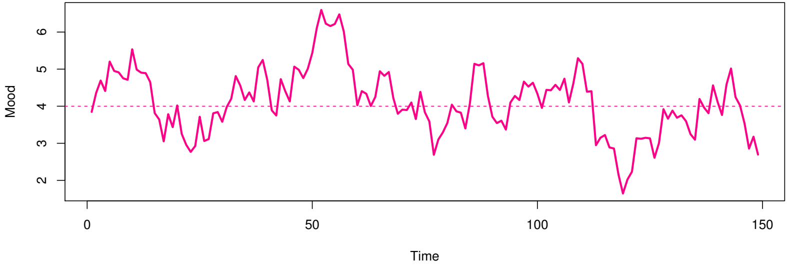
- In the AR(1) model ϕ lies between -1 and 1



AR(1) with ϕ = .5

The N=1 AR(1) Model: Delving Deeper

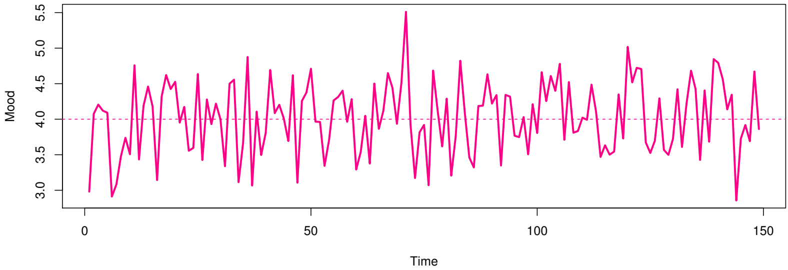
- In the AR(1) model ϕ lies between -1 and 1



AR(1) with $\phi = .8$

The N=1 AR(1) Model: Delving Deeper

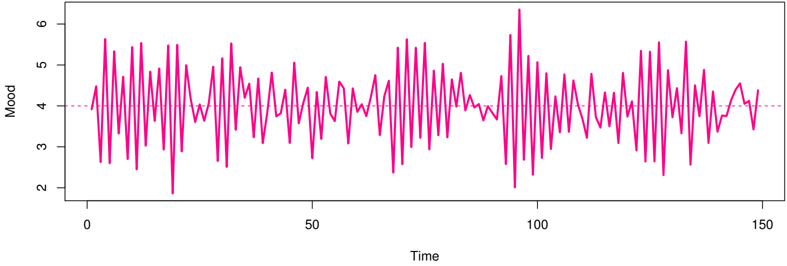
- In the AR(1) model ϕ lies between -1 and 1



AR(1) with $\phi = 0$

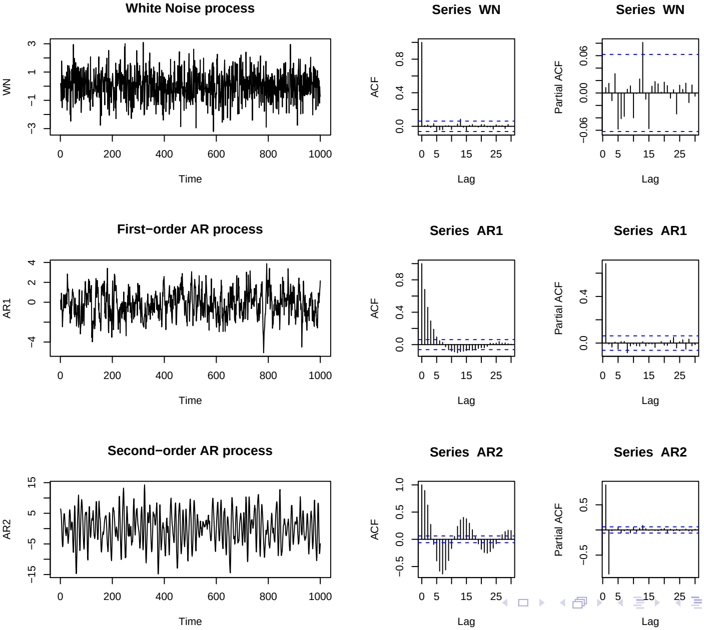
The N=1 AR(1) Model: Delving Deeper

- In the AR(1) model ϕ lies between -1 and 1



AR(1) with ϕ = -.8

Sequence, ACF and PACF



The N=1 AR(1) Model: Psychological Practice?

- The autoregressive effect as resilience
- emotional inertia positively related with psychological maladjustment (Kuppens et al. 2011)
- emotional inertia positively related with rumination and depression severity (Koval, 2012)
- emotional inertia predicts the onset of depressive disorder in adolescence (Kuppens et al. 2015)

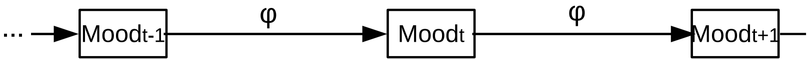
The N=1 AR(1) Model: Software?

	N=1	**multilevel**
uni-variate	- any regression software - arima in R - State Space Modeling software - Openmx - Bayesian modeling software (Including WinBUGS, STAN, JAGS and Mplus v8!)	
some-what multi-variate		
multi-variate		

The N=1 AR(1) Model: Software?

	N=1	multilevel
uni-variate	- any regression software - arima in R - State Space Modeling software - Openmx - Bayesian modeling software (Including WinBUGS, STAN, JAGS and Mplus v8!)	
some-what multi-variate		
multi-variate		

The N=1 AR(1) Model: Missings



Mood t	Mood t-1
5	.
3	5
3	3
4	3
2	4
3	2
1	3
1	1
2	1
.	2

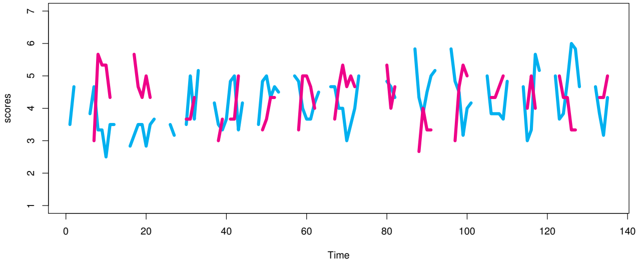
Overview

- **Intensive Longitudinal Data**
- Single Subject Univariate Autoregressive Modeling
- Single Subject Multivariate (Vector) Autoregressive Modeling
- Some Advanced Issues

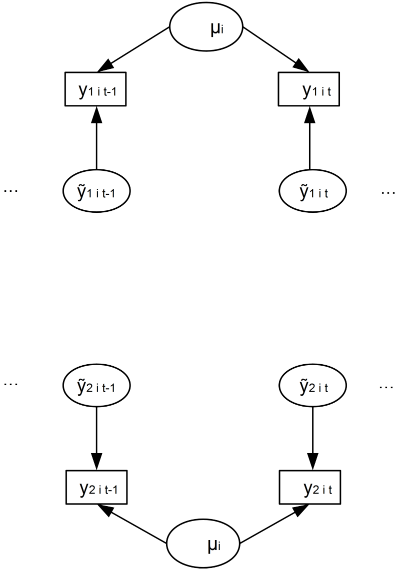
VAR modeling: Example

Competence and Exhaustion of people diagnosed with burnout

- Experience Sampling study by Sonnenschein et al. (2006)
- 54 persons diagnosed with burnout
- On average 80 repeated measures for exhaustion and 40 for feeling competent

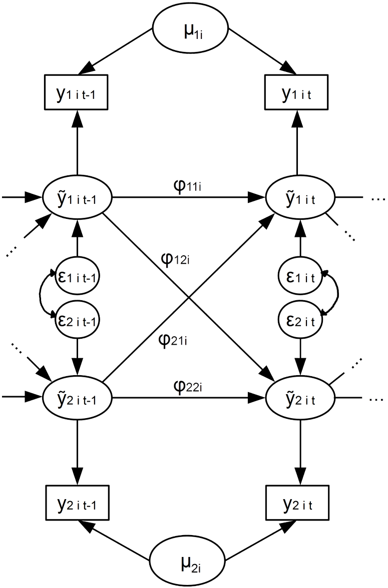


Bivariate autoregressive model



$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix}$$

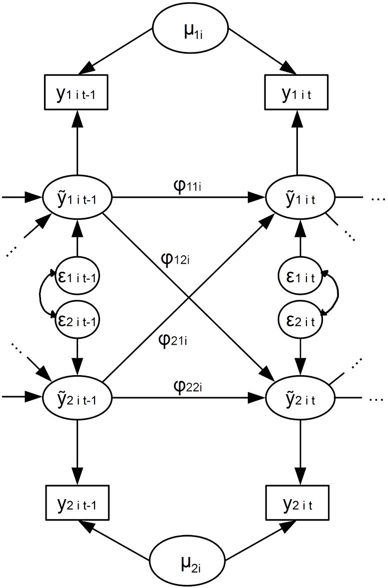
Bivariate Vector Autoregressive Model



$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} \tilde{y}_{1t-1} \\ \tilde{y}_{2t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

Bivariate Vector Autoregressive Model

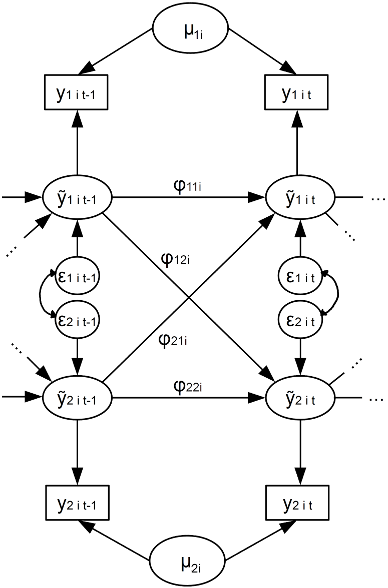


$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} \tilde{y}_{1t-1} \\ \tilde{y}_{2t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

$$\begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \sim MvN \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right)$$

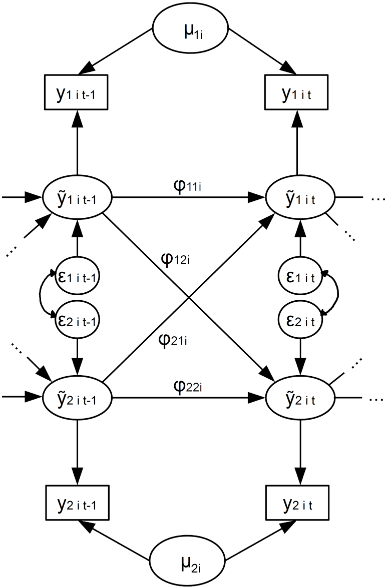
Bivariate Vector Autoregressive model



$$y_t = \mu + \tilde{y}_t$$
$$\tilde{y}_t = \Phi \tilde{y}_{t-1} + \epsilon_t$$

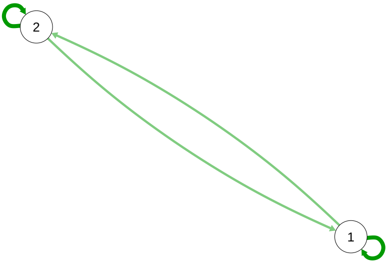
$$\epsilon_t \sim MvN\left(0, \Sigma\right)$$

Bivariate autoregressive model

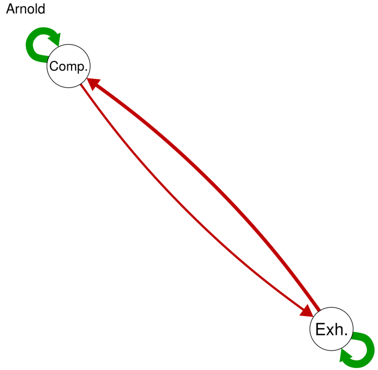


$$y_t = \mu + \tilde{y}_t$$
$$\tilde{y}_t = \Phi \tilde{y}_{t-1} + \epsilon_t$$

$$\epsilon_t \sim MvN(0, \Sigma)$$

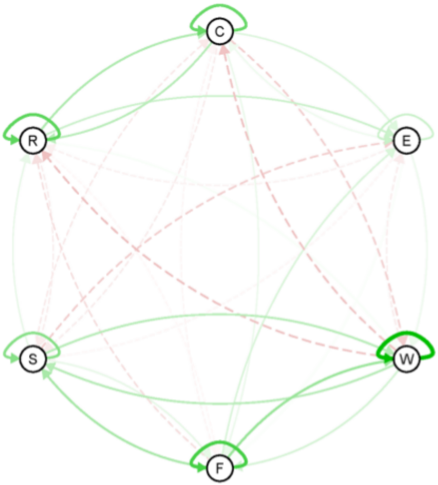


Vector Autoregressive Modeling: Multiple Variables



Based on results from Schuurman et al. 2016

Dynamic Network Examples



C=Cheerful; E=Event; W=Worried; F=Fear; S=Sad; R=Relaxed.
Image from Bringmann et al. (2013)

The N=1 VAR(1) Model: Software?

	N=1	multilevel
uni-variate	- any regression software - arima in R - State Space Modeling software - Openmx - Bayesian modeling software	
some-what multi-variate	- any regression software - VARS package in R - State Space Modeling Software - Bayesian software	
multi-variate	- State Space Modeling Software (mkfm6; Ox; fkf, dlm, KFAS, and MARSS in R) - Bayesian software (Winbugs, Openbugs, JAGS, STAN, Mplus v8)	

Intermezzo on Bayesian analysis

Bayesian analysis is based on combining the **density of the data** with a **prior distribution** for the unknown parameters, to get a **posterior distribution** of these parameters.

Intermezzo on Bayesian analysis

Bayesian analysis is based on combining the **density of the data** with a **prior distribution** for the unknown parameters, to get a **posterior distribution** of these parameters.

Posterior distribution of θ

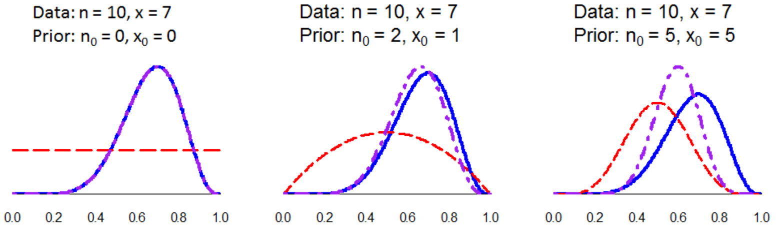
$$p(\theta|y) = \frac{f(y|\theta)p(\theta)}{f(y)}$$

where:

- $f(y|\theta)$ be the **density of the data** y given the parameters θ (also referred to as the likelihood)
- $p(\theta)$ be the **prior distribution** of the parameter(s) θ, which the user needs to specify
- $\int f(y,\theta)d\theta = f(y)$ is the **marginal density**, which can be ignored (because it is a constant)

Intermezzo on Bayesian analysis

Density (blue), prior (red), and posterior (purple):



When the prior is flat (no information), the posterior is identical to the likelihood.

If you have prior knowledge, you can add this to the equation by specifying a prior that reflects this.

For each to be estimated parameter, a prior needs to be specified. In the lab we'll aim to specify uninformative priors.

Intermezzo on Bayesian analysis: Convergence

Bayesian analysis is (often) based on using an **MCMC algorithm** which iteratively **samples** the parameters from their conditional posteriors.

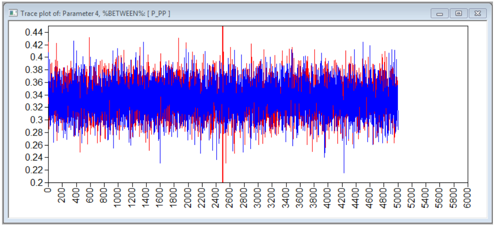
We have to check whether the analysis has **converged** (or: whether there are signs it did **not** converge).

Tools we use for this are:

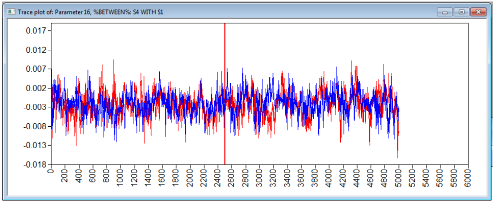
- Multiple chains; multiple runs of the analysis with different starting values.
- These chains should end up at approximately the same estimates.
- Burnin: Part of the iterations (before convergence) are discarded, leaving only 'converged' samples.
- Plots of the chains (fat hairy caterpillars), density plots (should look smooth and normal-ish), gelman rubin statistic: should be very close to 1.

Intermezzo on Bayesian analysis: Trace plots

This looks good (lazy, fat caterpillar):



This looks less good but not really bad; just needs more samples:



Overview

- Intensive Longitudinal Data
- Single Subject Univariate Autoregressive Modeling
- Single Subject Multivariate (Vector) Autoregressive Modeling
- Some Advanced Issues

Advanced Issues

Extensions to Multiple Subjects

- Multilevel time series & Dynamic SEM
- Clustering approaches (e.g., GIMME by Gates & Molenaar)

Advanced Issues

Extensions to Multiple Subjects

- Multilevel time series & Dynamic SEM
- Clustering approaches (e.g., GIMME by Gates & Molenaar)

General Modeling Issues/Assumptions

- Linear vs Non-linear models
- Categorical models (markov models)
- Models with other distributional assumptions
- Absence of Measurement Error
- Variable selection/model selection

Advanced Issues

Extensions to Multiple Subjects

- Multilevel time series & Dynamic SEM
- Clustering approaches (e.g., GIMME by Gates & Molenaar)

General Modeling Issues/Assumptions

- Linear vs Non-linear models
- Categorical models (markov models)
- Models with other distributional assumptions
- Absence of Measurement Error
- Variable selection/model selection

Assumptions/issues related to Dynamics

- Stationarity
- Equidistant measurements
- Mediation, Interventions and Causality
- Modeling processes on that take place at different time scales

Advanced Issues

Extensions to Multiple Subjects

- Multilevel time series & Dynamic SEM (Schuurman et al. 2016; Asparouhov, Hamaker & Muthen,2018).
- Clustering approaches (e.g., GIMME by Gates & Molenaar)

General Modeling Issues/Assumptions

- Linear vs Non-linear models
- Categorical models (markov models)
- Models with other distributional assumptions
- Absence of Measurement Error
- Variable selection/model selection

Assumptions/issues related to Dynamics

- Stationarity
- Equidistant measurements
- Mediation, Interventions and Causality
- Modeling processes on that take place at different time scales

Going Multilevel: Software

	N=1	multilevel
uni-variate	- arima in R - State Space Modeling software - Openmx - Bayesian modeling software - Mplusv8	- any multilevel software - MLvar package in R - Bayesian modeling software - Mplusv8
some-what multi-variate	- VARS package in R - State Space Modeling Software - Openmx - Bayesian modeling software - Mplusv8	- any multilevel software - MLVar package in R - Bayesian modeling software - Mplusv8
multi-variate	- State Space Modeling Software (mkfm6; Ox; fkf, dlm, KFAS, and MARSS in R) - Bayesian software (Winbugs, Openbugs, JAGS, STAN) - Mplusv8	- Bayesian software (Winbugs, Openbugs, JAGS, STAN) - Mplusv8

DSEM in Mplus v8

- Designed for continuous, normal variables
- N=1 or multilevel (all parameters can be allowed to vary across persons
- Explicit separation of within/between (so a multilevel context)
- Similar to the State Space modeling framework (but even more general!).
- Allows for specifying many different time series models, including classic AR, ARMA, ARIMA models
- Allows for adding predictors or outcome variables on between level and the within level in one step
- Can deal with categorical variables via a probit link function (I believe dynamic IRT models are possible)
- Bayesian estimation

DSEM Software

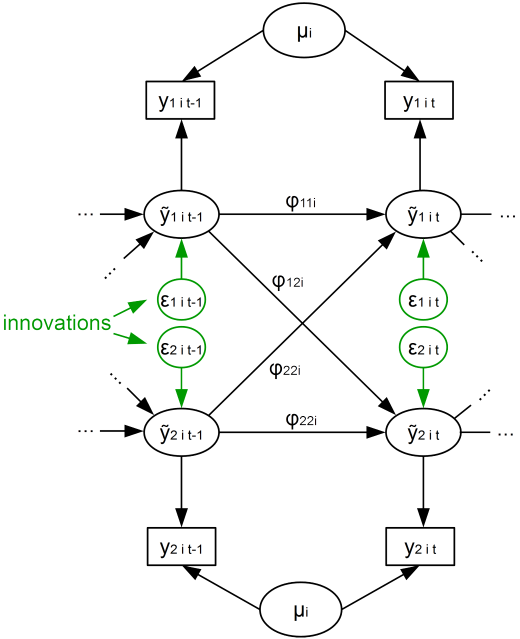
Mplus v8

- Specifically developed for DSEM
- tailored to DSEM specific issues, time saving features
- fast, stable
- less flexible
- Not free (aside from student version), not open source
- Support from Mplus
- Probably more user friendly

Bugs, Stan, Jags

- Not specifically developed for DSEM, very general
- dealing with specific DSEM issues requires (much) more work
- less fast, can be less stable (depending on your implementation)
- more flexible
- Free, open source
- Tips/advice everywhere, but you are basically on your own
- Probably less user friendly

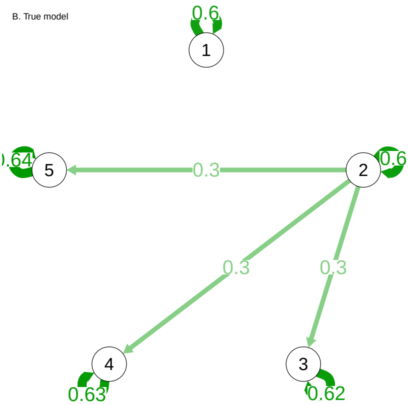
Innovations =/= Measurement errors

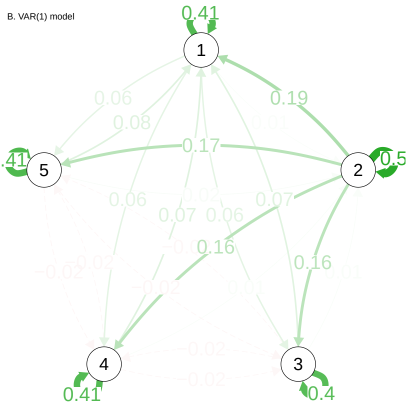


$$y_{it} = \mu_i + \tilde{y}_{it}$$
$$\tilde{y}_{it} = \Phi_i \tilde{y}_{it-1} + \epsilon_{it}$$

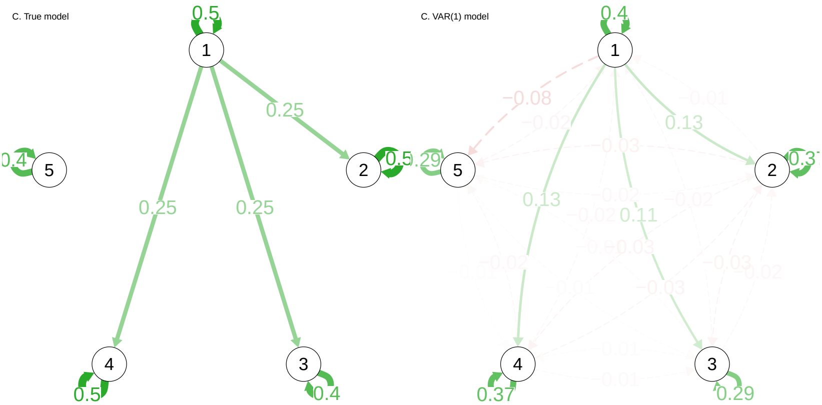
$$\epsilon_{it} \sim MvN\left(0, \Sigma_i\right)$$

Disregarding Measurement Error...

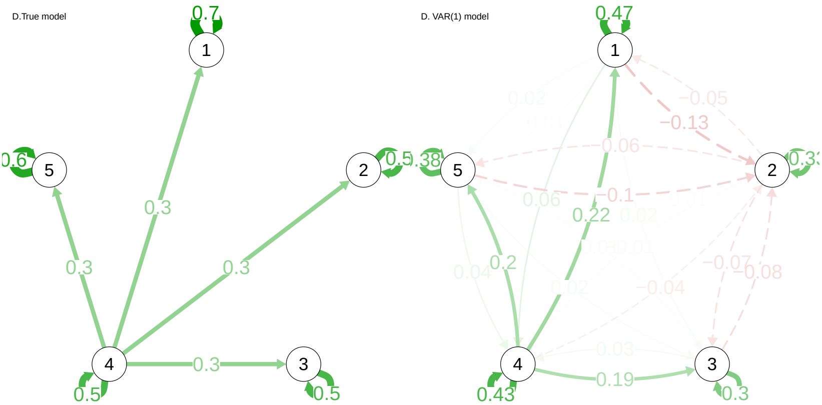




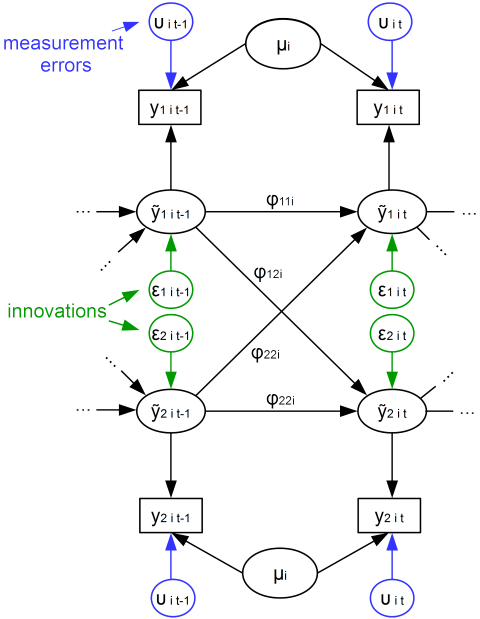
Disregarding Measurement Error...



Disregarding Measurement Error...



Innovations =/= Measurement errors



$$y_{it} = \mu_i + \tilde{y}_{it} + \upsilon_{it}$$
$$\tilde{y}_{it} = \Phi_i \tilde{y}_{it-1} + \epsilon_{it}$$

$$\upsilon_{it} \sim MvN\left(0, \Omega_i\right)$$

$$\epsilon_{it} \sim MvN\left(0, \Sigma_i\right)$$

Note: Multilevel approaches often disregard interindividual differences in residual (co)variances

Reasons to assume **individual differences** for these variances:

- individuals may differ with respect to the **variability in exposure** to external factors
- individuals may differ with respect to their **reactivity** to external influences (see reward experience and stress sensitivity research)

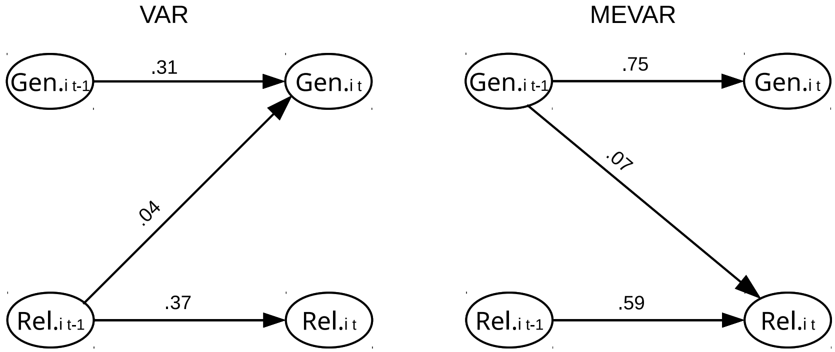
Empirical Example: General PA and Relationship PA

(Multilevel) VAR modeling with ME: Example

Positive affect of women in a heterosexual relationship

- Data from study by Ferrer, Steele, and Hsieh (2012)
- 190 women filled out a diary every evening
- about 60 to 90 repeated measures on daily General Positive Affect and Relationship Positive Affect

Empirical Example: General PA and Relationship PA



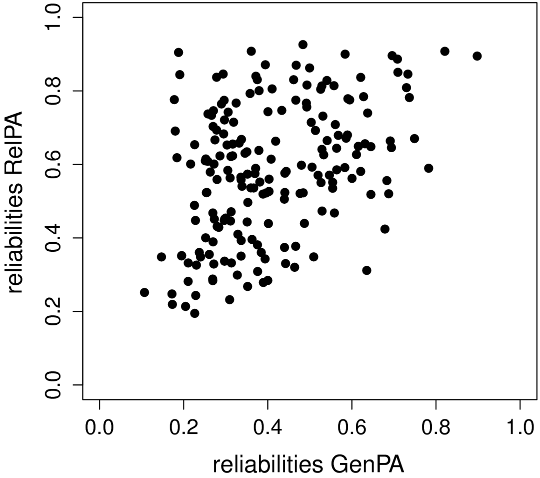
mean ϕ_{geni}: .31 (.28, .34)
mean ϕ_{reli}: .37 (.34, .40)
mean $\phi_{gen->reli}$: .04 (.02, .07)
mean $\phi_{rel->geni}$: .02 (.00, .04)

mean ϕ_{geni}:.75 (.69, .80)
mean ϕ_{reli}: .59 (.53, .64)
mean $\phi_{gen->reli}$: -.03 (-.07, .00)
mean $\phi_{rel->geni}$: .07 (.02, .13)

Person-specific reliabilities

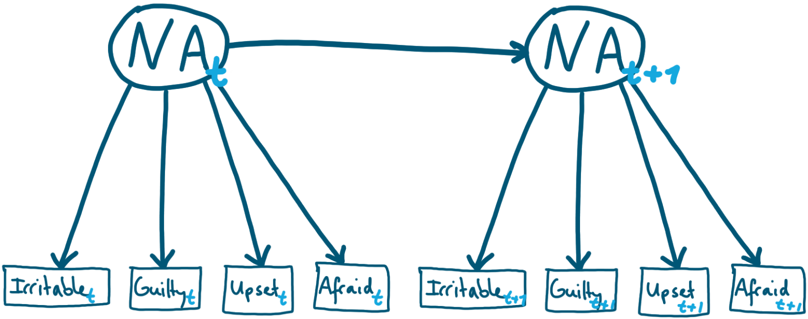
- Unique measurement error variances per person (and variable) also implies unique reliabilities!
- For each person: Calculate the proportion of that person's total variance and the part of the variance which is not due to measurement errors

Person-specific reliabilities

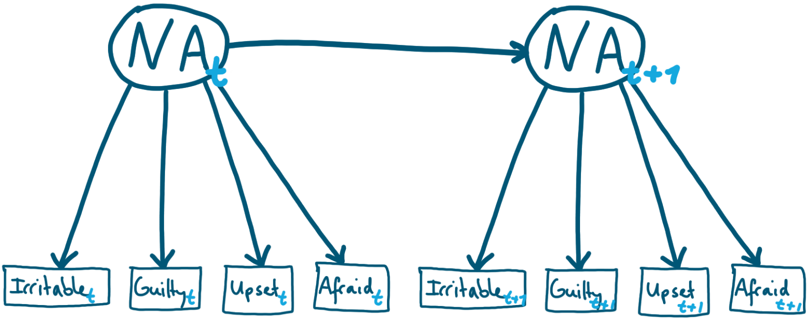


Read more:
Schuurman, Grasman & Hamaker (2015), Schuurman & Hamaker (2020).

Factor modeling as a tool for filtering out ME

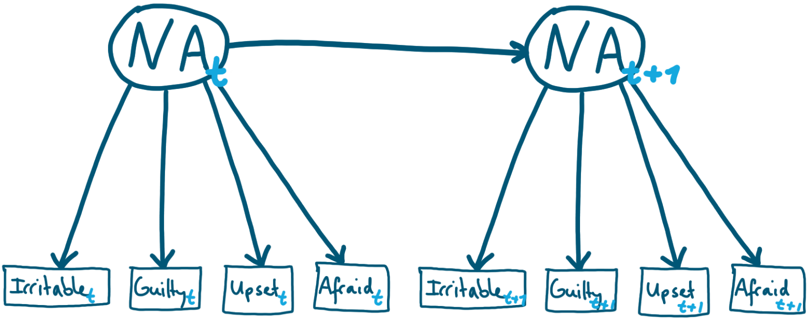


Factor modeling as a tool for filtering out ME



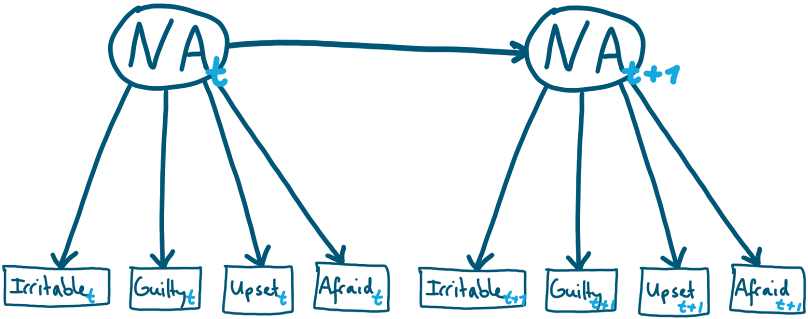
- Are they really exchangable, parallel items, that measure the same thing?

Factor modeling as a tool for filtering out ME



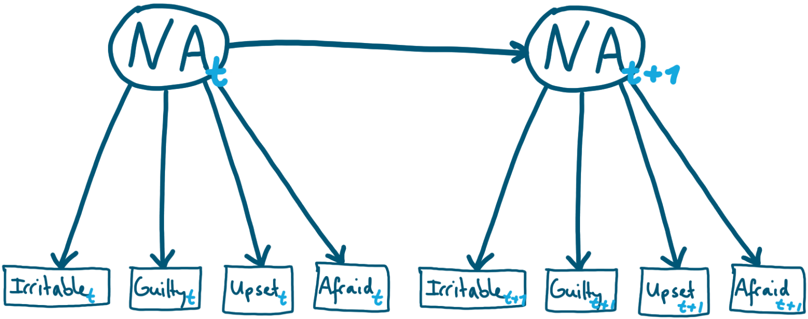
- Are they really exchangable, parallel items, that measure the same thing?
- Do we really expect these items to all usually increase and decrease together at each occasion?

Factor modeling as a tool for filtering out ME



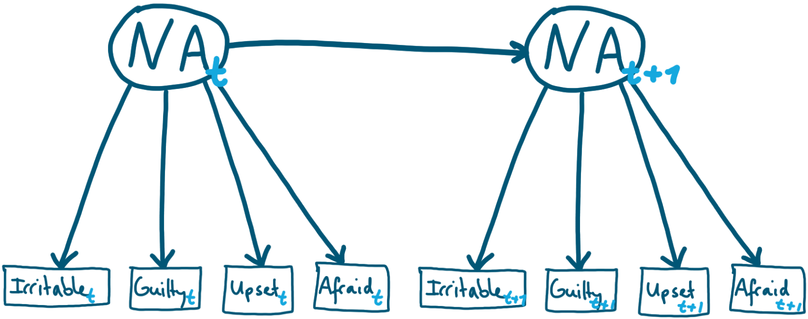
- Are they really exchangable, parallel items, that measure the same thing?
- Do we really expect these items to all usually increase and decrease together at each occasion?
- Is that the case for all persons?

Factor modeling as a tool for filtering out ME



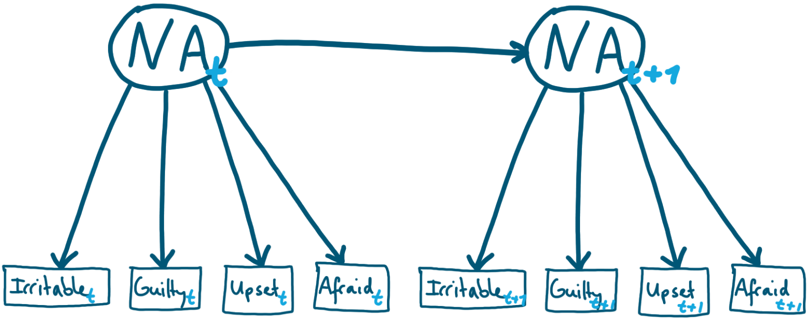
- Could we intervene directly on the latent variable for any particular person?

Factor modeling as a tool for filtering out ME



- Could we intervene directly on the latent variable for any particular person?
- Is the latent variable something real that acts within a person?

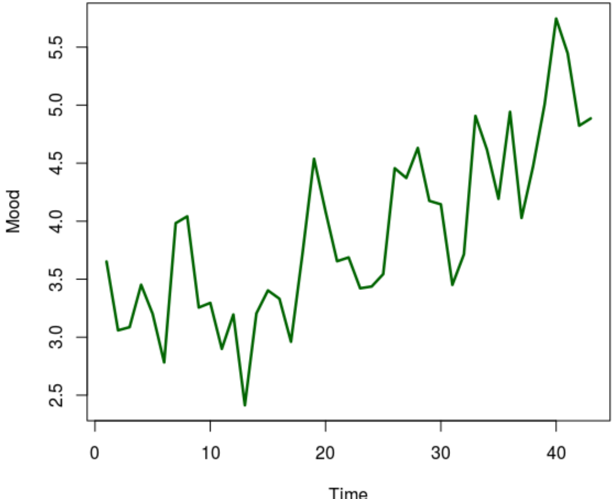
Factor modeling as a tool for filtering out ME



- Could we intervene directly on the latent variable for any particular person?
- Is the latent variable something real that acts within a person?
- How do changes in the latent variable result exactly in the observed scores?

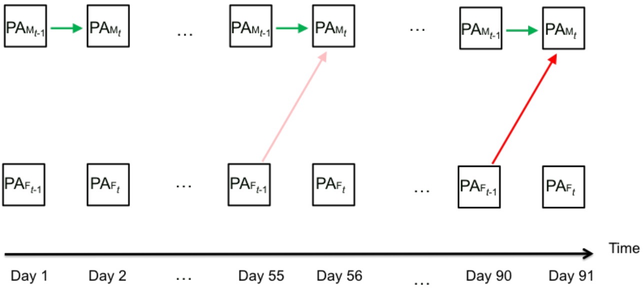
Stationarity Assumption

Parameters must not change over time (means, regression coefficients, variances, and so on).



Stationarity Assumption

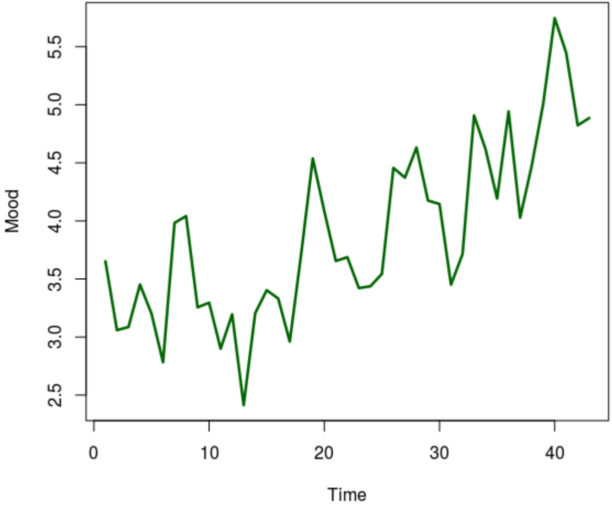
Time Varying VAR Read more: Bringmann, Hamaker, Vigo, Aubert, Borsboom, & Tuerlinckx (2016; only n=1)



More sudden changes?: Regime switching models, change point analysis, Threshold-AR models,... Read more: de Haan-Rietdijk et al. (2016), Hamaker, Grasman & Kamphuis (2016).

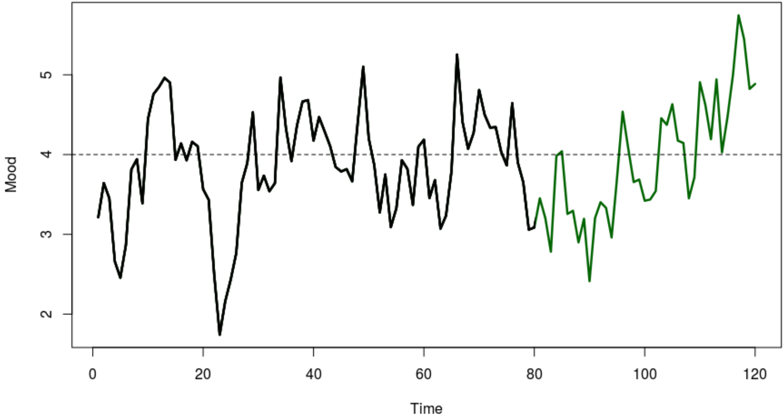
Stationarity Assumption

Trend...?

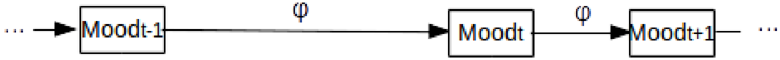


Stationarity Assumption

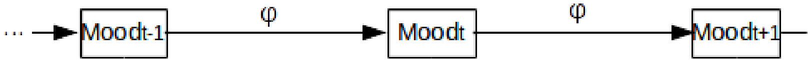
Trend...? No! Autoregressive process.



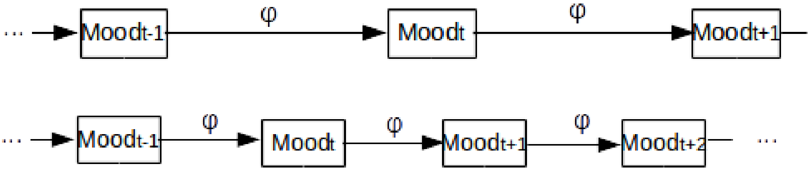
Equal Spacing Between Measurements



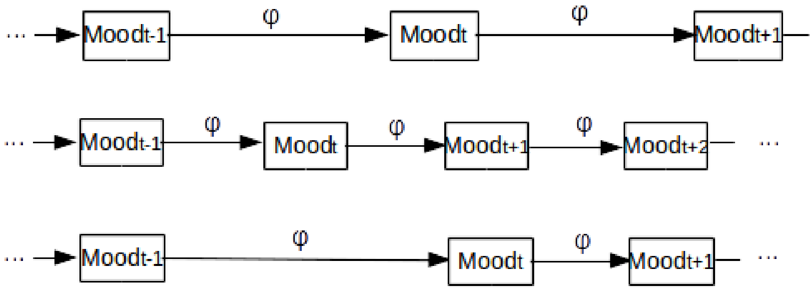
Equal Spacing Between Measurements



Equal Spacing Between Measurements



Equal Spacing Between Measurements



Discrete Time vs Continuous Time

- Ad hoc solution: add in missing observations to equally space measurements (TINTERVAL feature in Mplus)
- Continuous time models can directly take the length of the time intervals into account
- More in part 2 of this workshop!