Mplus 8: Dynamic SEM

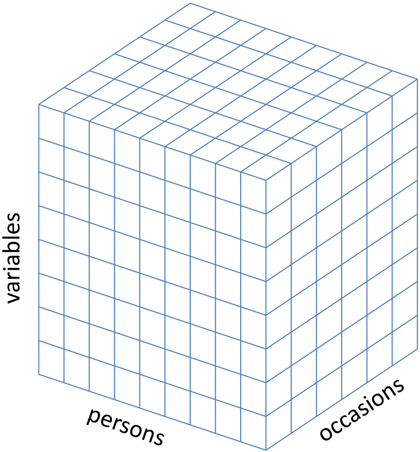
Time series analysis and state-space modeling

Ellen L. Hamaker
Utrecht University

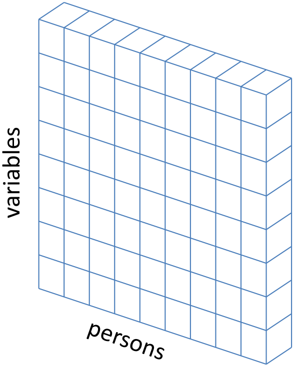
Tihomir Asparouhov & Bengt Muthén
Muthén & Muthén

May 23, 2016

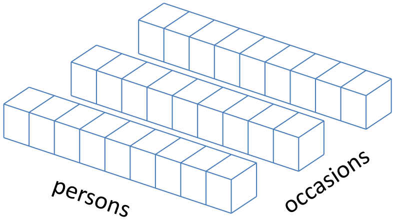
Cattell's data box



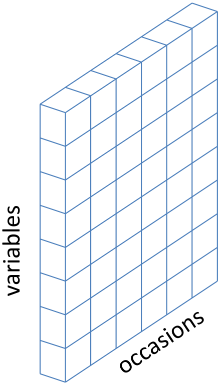
Cross-sectional research: A single snapshot



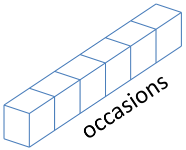
Panel research: A few snapshots



Time series data: Looking at the movie



Time series data: Looking at the movie



Outline

- **Why time series analysis?**
- Autocorrelation
- ARMA models
- Stationarity
- The state-space model
- Kalman filter and parameter estimation
- Diverse models in state-space format
- Miscellaneous

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)

TSA in the social and medical sciences

In **sociology**:

- quarterly unemployment numbers
- effect of alcohol consumption per capita on criminal violence rates
- effect of suicide news on suicide rates

In **medical research**:

- effect of safety warnings on antidepressants use
- effects of pain control strategies
- effect of 9/11 attacks on weekly psychiatric patient admissions

In **psychology**:

- network of symptoms in depressive patient
- effect of feedback on academic performance
- effect of an intervention on the relationship between stress and affect

Intensive longitudinal data

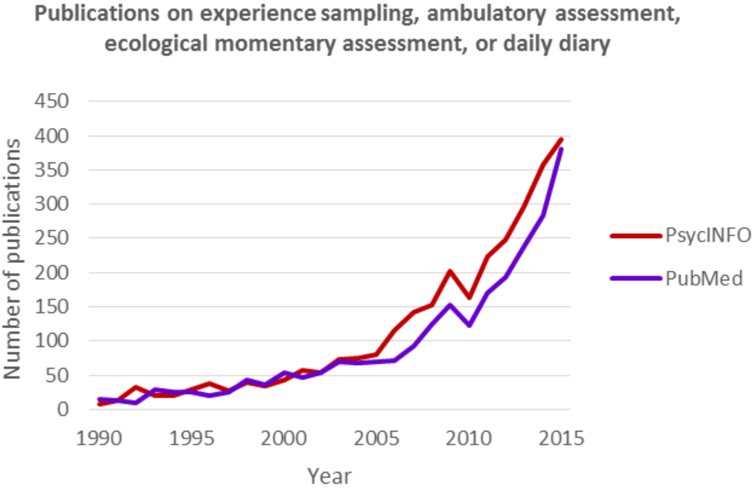
Intensive longitudinal data are gathered using:

- daily diary with end-of-day-measurements (self-report)
- experience sampling method (self-report)
- ecological momentary assessment (self-report)
- ambulatory assessment (including physiological variables)
- event contingency (self-report)
- observational measurements (expert rater)

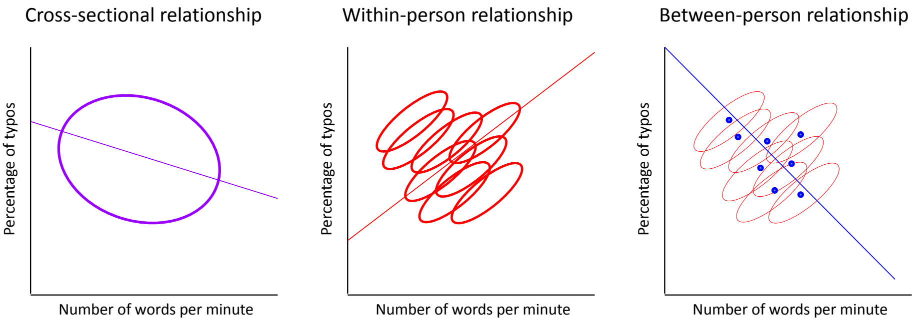
For more info on methodology, check out:

- Tamlin Conner (e.g., her seminar with Joshua Smyth on YouTube)
- Society for Ambulatory Assessment
- Trull and Ebner-Priemer (2013)

It's a revolution!

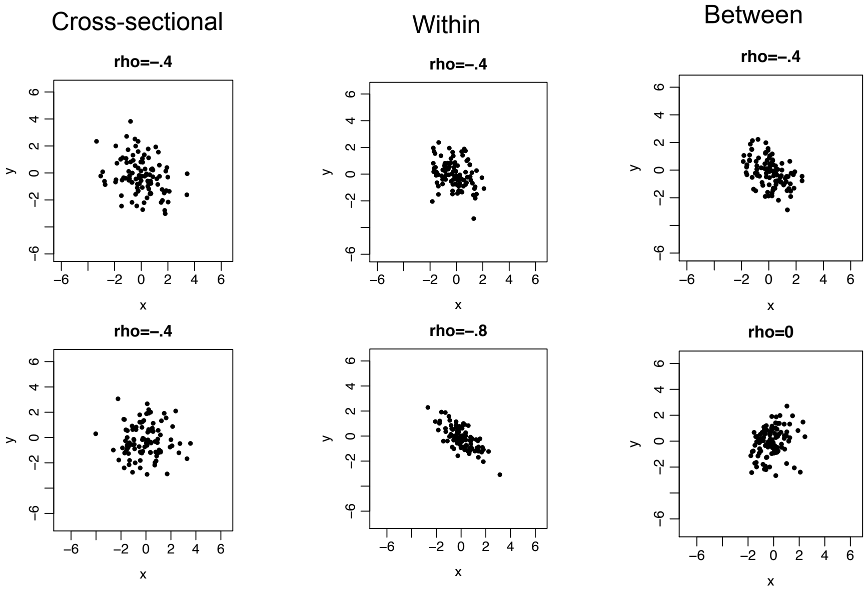


A fundamental problem in a nutshell



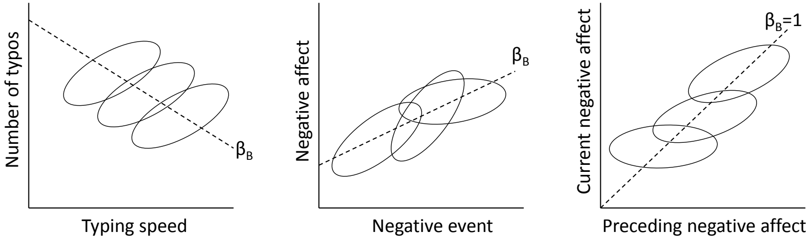
Taken from Hamaker (2012).

Three perspectives on data



Taken from Hamaker (2012).

Interindividual differences in intraindividual variation



Taken from Hamaker and Grasman (2014).

Cross-sectional correlations: A blend

Schmitz (2000):

$$r_{cs} = \eta^2 r_b + (1 - \eta^2) r_w$$

where

- r_{cs} is the cross-sectional correlation
- r_b is the between-person correlation
- r_w is the within-person correlation
- η^2 is the proportion of between-person variance of the total variability

Consequences:

- cross-sectional and panel research may result in an "**uninterpretable blend**" of within-person and between-person relationships (cf. Raudenbush and Bryk, 2002)
- in N=1 time series analysis there is **only within-person** variance

Outline

- Why time series analysis?
- **Autocorrelation**
- ARMA models
- Stationarity
- The state-space model
- Kalman filter and parameter estimation
- Diverse models in state-space format
- Miscellaneous

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

Autocorrelation function (ACF)

The ACF and the PACF can be used as **diagnostic tools** to determine the nature of the underlying process.

Variance (or: auto-covariance at lag 0):

$$\gamma_0 = \frac{1}{T} \sum_{t=1}^{T} \big(y_t - \bar{y}_t\big)^2$$

Auto-covariance at lag k:

$$\gamma_k = \frac{1}{T-k} \sum_{t=k+1}^{T} \big(y_t - \bar{y}_t\big)\big(y_{t-k} - \bar{y}_t\big)$$

Autocorrelation at lag k:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

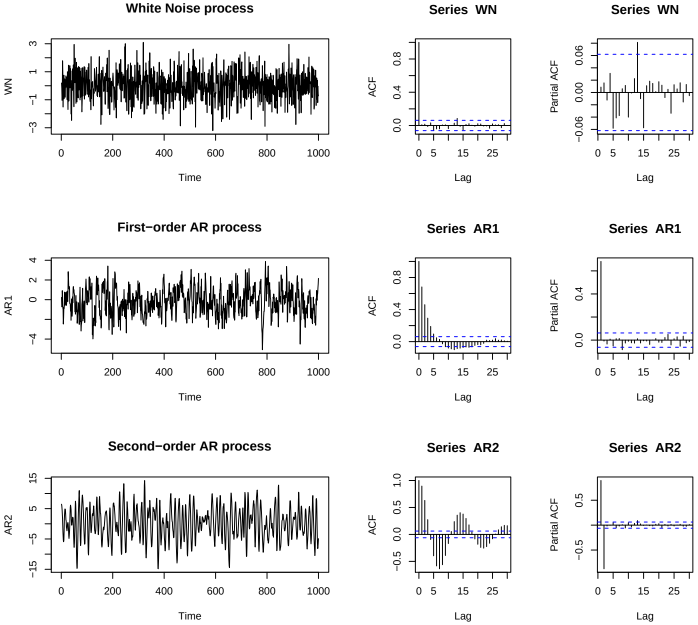
Partial autocorrelation function (PACF)

Partial autocorrelation at lag k is the correlation between y_t and y_{t-k} after **removing the effect of the intermediate observations** (i.e., y_{t-1} to y_{t-k+1}).

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
$\dots$	$\dots$	$\dots$
y_T	y_{T-1}	y_{T-2}
	y_T	y_{T-1}
		y_T

For instance: Is there a relationship between yesterday's positive affect and tomorrow's positive affect above and beyond their relationship to today's positive affect?

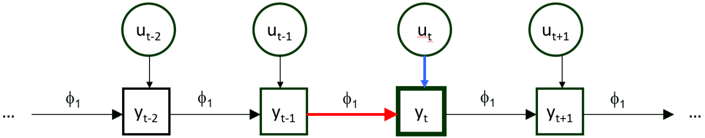
Sequence, ACF and PACF



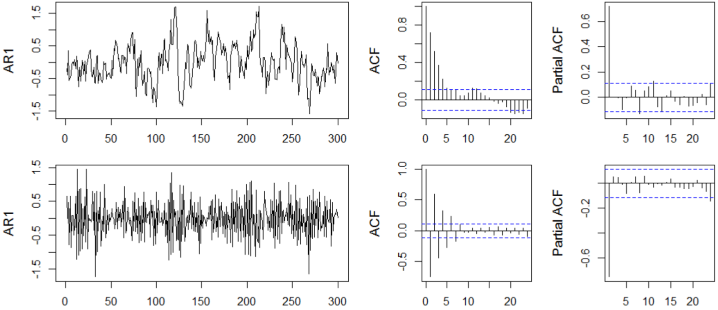
Outline

- Why time series analysis?
- Autocorrelation
- **ARMA models**
- Stationarity
- The state-space model
- Kalman filter and parameter estimation
- Diverse models in state-space format
- Miscellaneous

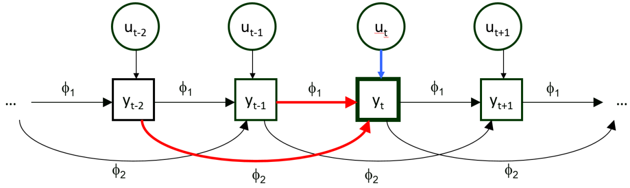
AR(1): $y_t = \phi_1 y_{t-1} + u_t$



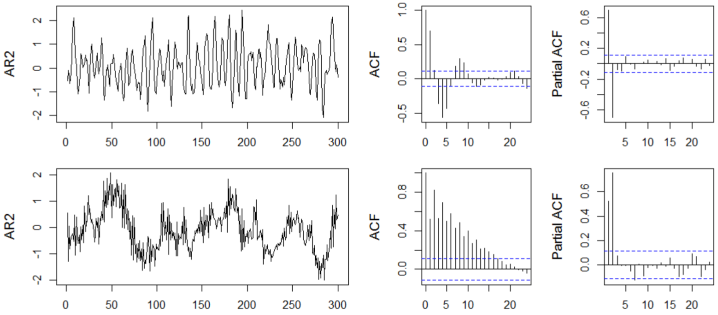
Example with $\phi_1 = 0.7$ and $\phi_1 = -0.7$:



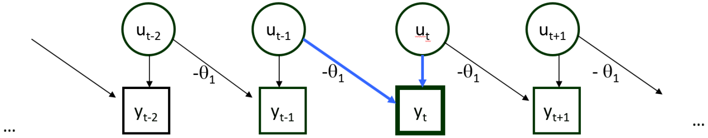
AR(2): $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$



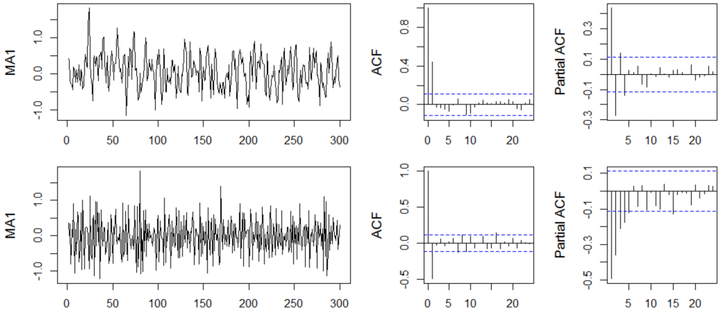
Example with $\phi_1 = 1.2$ and $\phi_2 = -0.7$ and with $\phi_1 = 0.2$ and $\phi_2 = 0.7$:



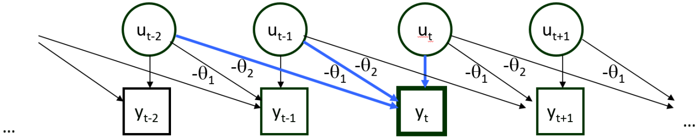
MA(1): $y_t = u_t - \theta_1 u_{t-1}$



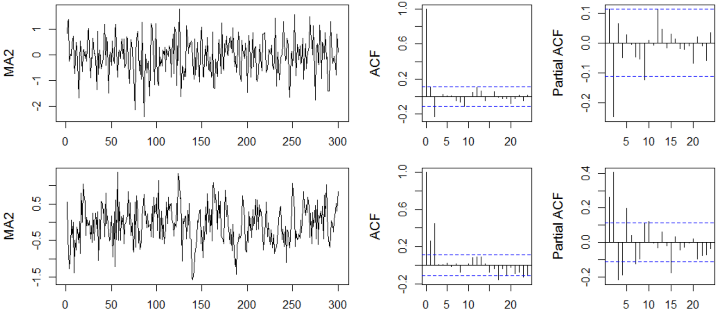
Examples with $\theta_1 = 0.7$ and with $\theta_1 = -0.7$:



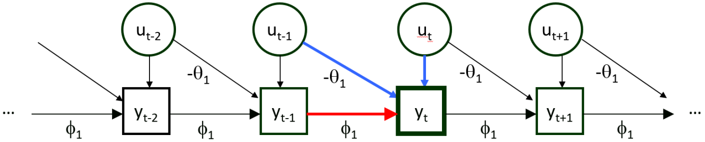
MA(2): $y_t = u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2}$



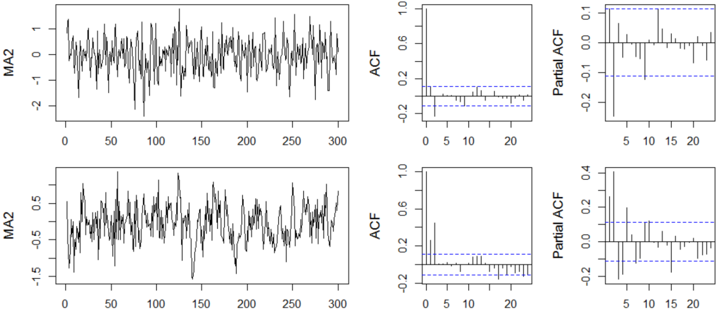
Examples with $\theta_1 = 1.2$ and $\theta_2 = -0.7$, and with $\theta_1 = 0.2$ and $\theta_2 = 0.7$:



ARMA(1,1): $y_t = \phi y_{t-1} + u_t - \theta_1 u_{t-1}$



Example with $\phi_1 = .8$ and $\theta_1 = 0.8$, and with $\phi_1 = -0.8$ and $\theta_1 = -0.8$:



Pure AR, pure MA, or an ARMA(p, q)?

In general:

- an AR(p) can **always** be written as an MA(∞)
- an MA(q) can **always** be written as an AR(∞)

Other (rather unexpected) results found by **Granger and Morris (1976)**:

- AR(1) + WN $\rightarrow$ ARMA(1,1)
- AR(1) + AR(1) $\rightarrow$ ARMA(2,1)
- MA(1) + WN $\rightarrow$ MA(1)

You may consider:

- interpretation (social sciences)
- forecasting (econometrics)
- parsimony

Outline

- Why time series analysis?
- Autocorrelation
- ARMA models
- **Stationarity**
- The state-space model
- Kalman filter and parameter estimation
- Diverse models in state-space format
- Miscellaneous

Stationarity

Stationarity is an important concept in time series analysis:

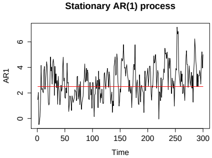
- is based on using **backshift operators** and the **unit root circle** (as all introductory texts on time series analysis do!)
- implies that **all moments** (i.e., means, variances, covariances, lagged covariances, etc.) are **independent of time**

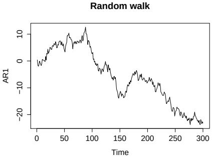
For instance:

- mean is constant over time
- γ_k **depends on the lag** k, **not on** t (i.e., the occasion itself)

Two typical **examples of nonstationary processes**:

- trends over time (including cycles?)
- random walk: $y_t = y_t + e_t$





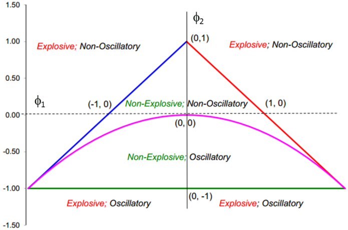
Stationarity of an AR(p)

For an AR(1) to be stationary, $|\phi| < 1$.

For an AR(2) to be stationary we need:

- $\phi_2 - \phi_1 < 1$
- $\phi_2 + \phi_1 < 1$
- $|\phi_2| < 1$

which leads to the following triangle:



(Check out: http://freakonometrics.hypotheses.org/12081)

Outline

- Why time series analysis?
- Autocorrelation
- ARMA models
- Stationarity
- **The state-space model**
- Kalman filter and parameter estimation
- Diverse models in state-space format
- Miscellaneous

Rocket science

State space model with **known parameters**:

- Kalman **filter** predicts the future state (e.g., the location of your space rocket), based on current and previous observations (on-line procedure)
- Kalman **smoother** predicts the state based on previous, current and future observations (off-line procedure)

Often, the **parameter values are NOT known**.

Then, certain **by-products** of the Kalman filter/smoother can be used in a **likelihood function** (see later).

The basic framework

Measurement equation

$$y_t = c_t + Z_t a_t + e_t \qquad \text{with} \qquad e_t \sim MN(0, GG_t)$$

- c_t is the vector with intercepts in the measurement equation
- Z_t is the matrix with factor loadings
- GG_t is the covariance matrix of the measurement errors

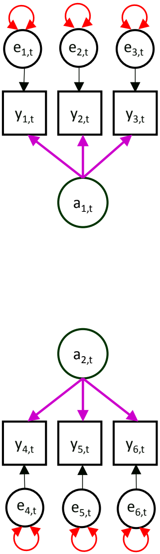
Transition equation

$$a_t = d_t + T_t a_{t-1} + u_t \qquad \text{with} \qquad u_t \sim MN(0, HH_t)$$

- d_t is the vector with intercepts in the transition equation
- T_t is the matrix with cross- and auto-regressive coefficients
- HH_t is the covariance matrix of the dynamic errors

In a **more basic version** these model matrices are **fixed over time**.

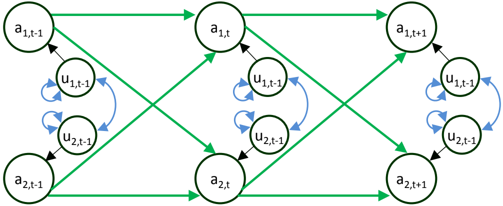
Measurement equation: regressing y_t on a_t



$$y_t = c + \mathbf{Z} a_t + e_t$$
$$e_t \sim MN(0, \mathbf{GG})$$

Transition equation: Regressing a_t on a_{t-1}

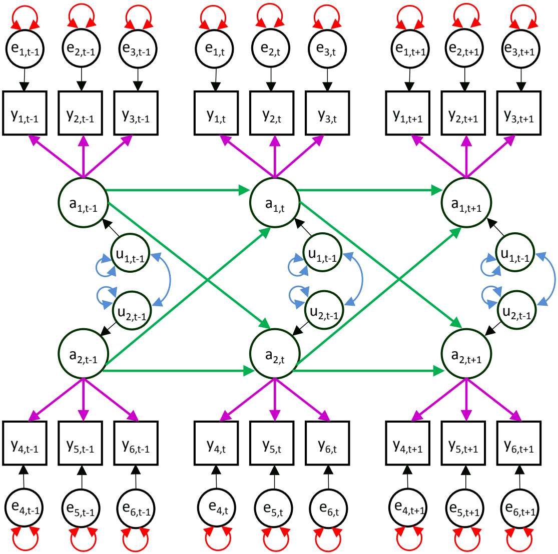
$$a_t = d + \mathbf{T} a_{t-1} + u_t$$
$$u_t \sim MN(0, \mathbf{HH})$$



State-space model = Latent VAR(1) model

$$y_t = c + \mathbf{Z} a_t + e_t$$
$$e_t \sim MN(0, \mathbf{GG})$$

$$a_t = d + \mathbf{T} a_{t-1} + u_t$$
$$u_t \sim MN(0, \mathbf{HH})$$



Outline

- Why time series analysis?
- Autocorrelation
- ARMA models
- Stationarity
- The state-space model
- **Kalman filter and parameter estimation**
- Diverse univariate models in state-space format
- Miscellaneous

State-space model versus SEM

Two ways in which SEM can be use to do TSA:

Toeplitz method, based on making lagged variables

$$\begin{array}{cc} y_1 & \\ y_2 & y_1 \\ y_3 & y_2 \\ y_4 & y_3 \\ \dots & \\ y_T & y_{T-1} \end{array}$$

Advantage: easy

Disadvantage: violates assumption of independent cases (=rows); no true ML estimates (and wrong fit measures)

(cf. Hamaker, Dolan & Molenaar, 2002)

Raw maximum likelihood estimation, based on N=1

$$y_1 \quad y_2 \quad y_3 \quad \dots \quad y_T$$

Advantage: gives ML estimates

Disadvantage: requires inversion of (at least) a $T \times T$ matrix (computationally troublesome)

(cf. Hamaker, Dolan and Molenaar, 2003)

See Chow, Ho, Hamaker and Dolan (2010) for further comparison of state-space modeling and SEM.

Kalman filter for parameter estimation

The Kalman filter can be used to **predict future states** when the **parameters are known**.

In practice, the parameter values are often **unknown**.

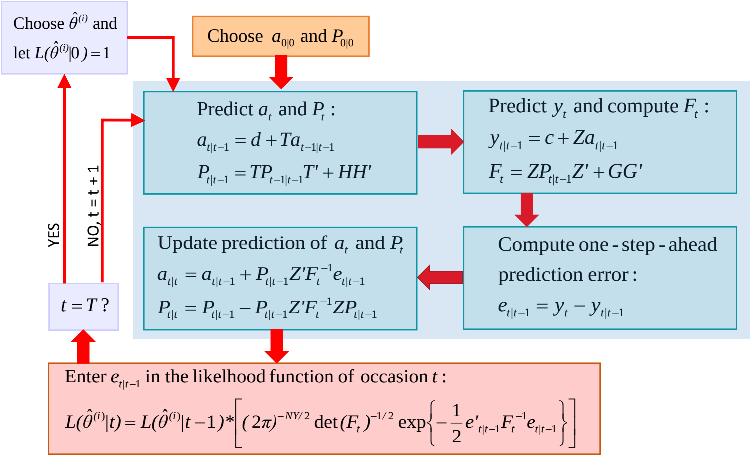
In that case, **by-products** of the Kalman filter can be used to **estimate the parameters**:

- the **one-step-ahead-prediction error** $e_{t|t-1} = y_t - y_{t|t-1}$
- the **covariance matrix** of $e_{t|t-1}$ (i.e., F_t)

These are **plugged into a likelihood function**, which is then **optimized** with respect to the unknown parameters.

Hence, for **each set** of possible parameter values, the **entire Kalman filter** is run from $t = 1$ to $t = T$.

Kalman filter for parameter estimation



Outline

- Why time series analysis?
- Autocorrelation
- ARMA models
- Stationarity
- The state-space model
- Kalman filter and parameter estimation
- **Diverse models in state-space format**
- Miscellaneous

Just a latent vector AR(1) model?

At first sight the state-space model **seems to be** just a latent VAR(1) model.

However, it is actually a **very flexible framework** for all sorts of time series models:

- all ARIMA models
- multivariate extensions
- dynamic factor analysis

Extensions may consist of:

- predictors (e.g., time, intervention, weather conditions) in the measurement and/or transition equation
- time-varying parameters
- regime switches (through combination with a hidden Markov process)

AR(1) in state-space format

Measurement equation:

$$y_t = c + a_t$$

- c is a vector containing the unknown mean
- Z is a 1 by 1 matrix containing 1
- GG is a zero matrix

Transition equation:

$$a_t = Ta_{t-1} + u_t$$

- d is a zero vector
- T is a 1 by 1 matrix containing the autoregressive parameter
- HH is a 1 by 1 covariance matrix containing the variance of the innovations

AR(1) with measurement error

Measurement equation:

$$y_t = c + a_t + e_t$$

- c is a vector containing the unknown mean
- Z is a 1 by 1 matrix containing 1
- GG is a 1 by 1 covariance matrix with the variance of the measurement error

Transition equation:

$$a_t = Ta_{t-1} + u_t$$

- d is a zero vector
- T is a 1 by 1 matrix containing the autoregressive parameter
- HH is a 1 by 1 covariance matrix containing the variance of the innovations

AR(2) in state-space format

Measurement equation:

$$y_t = c + \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} a_t \\ a_{t-1} \end{bmatrix} = c + a_t$$

where GG is a zero matrix.

Transition equation:

$$\begin{bmatrix} a_t \\ a_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{t-1} \\ a_{t-2} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \end{bmatrix} = \begin{bmatrix} \phi_1 a_{t-1} + \phi_2 a_{t-2} + u_t \\ a_{t-1} \end{bmatrix}$$

- d is a zero vector
- HH is a 2 by 2 covariance matrix containing only the variance of the innovations (element 1,1)

Bivariate VAR(1) in state-space format

The measurement equation:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}$$

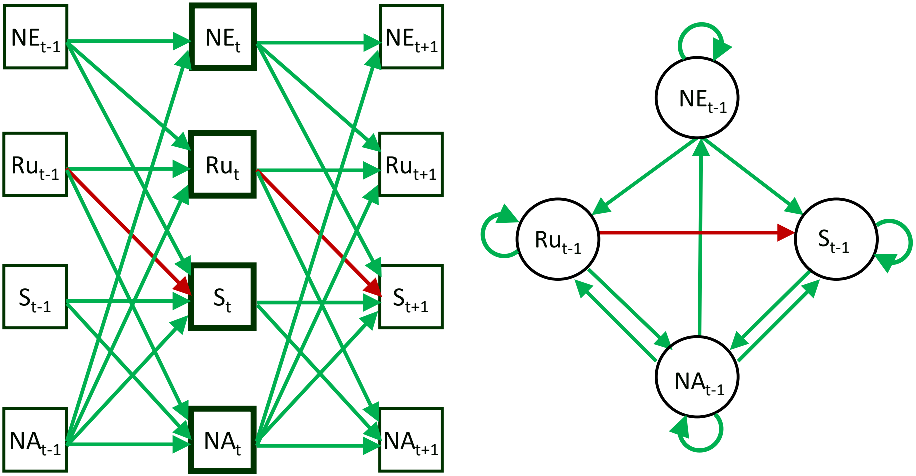
where GG is a zero matrix.

The transition equation:

$$\begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} a_{1,t-1} + \phi_{21} a_{2,t-1} + u_{1,t} \\ \phi_{22} a_{2,t-1} + \phi_{12} a_{1,t-1} + u_{2,t} \end{bmatrix}$$

- d is a zero vector
- HH is a 2 by 2 covariance matrix of $u_{1,t}$ and $u_{2,t}$

Graphical representations



Applications of VAR models

VAR models are of interest, because

- they allow you to study **Granger causality**: Can you predict Y from X, after controlling for previous levels of Y?
- they allow you to determine which variable is "**causally dominant**" when there are reciprocal effects
- they can be interpreted as **networks** (alternative to latent variable approach)

Some interesting **replicated VAR applications**

- Schmitz and Skinner (1994): Perceived control, effort and academic performance
- Rosmalen et al. (2012): Depression and physical activity
- Snippe et al. (2014): Mindfulness, repetitive thinking and depressive symptoms
- Van Gils et al. (2014): Stress and functional somatic symptoms

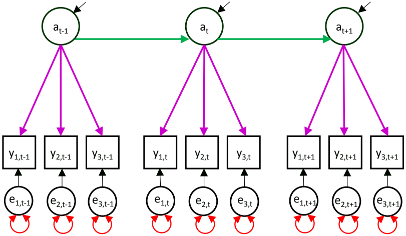
In all these studies they find **important differences across individuals**.

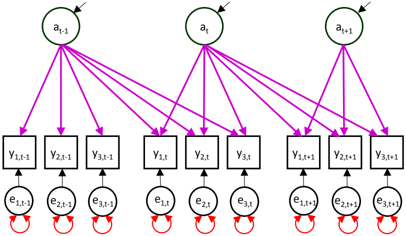
Dynamic factor model

Dynamic factor analysis is used for time series data consisting of **multiple indicators** of an underlying construct.

There are **two popular versions**:

- at the latent level there is a **VARMA model**; the factor loadings only appear **at lag 0**
- at the latent level there is **white noise**; the factor loadings appear at **different lags** (e.g., EEG data)





Outline

- Why time series analysis?
- Autocorrelation
- ARMA models
- Stationarity
- The state-space model
- Kalman filter and parameter estimation
- Diverse univariate models in state-space format
- **Miscellaneous**

Covariance matrix of the series

For a **univariate AR(1)**, we have: $\sigma_y^2 = \frac{\sigma_u^2}{1-\phi^2}$.

Similarly, for a **(latent) VAR model** we can express the **covariance matrix** of y_t in terms of

- **lagged regression parameters** Φ
- **covariance matrix of the innovations** Γ (i.e., HH in the state-space model)

Specifically (from Kim and Nelson, 1999):

$$\Sigma_y = mat\Big[\big(I - \Phi \otimes \Phi\big)^{-1} vec(\Gamma)\Big]$$

where

- $vec()$ implies you put all the **matrix elements in a vector**
- $mat()$ implies you place all the **vector elements in a square matrix**

Model fit

Despite the **similar appearance**, state-space modeling and SEM are **not the same**: For a time series there is **no saturated model** against which we can test other models.

We can **compare our model to other models**, including the white noise model (independence model), using

- log likelihood ratio test (for nested models)
- AIC, BIC, DIC, etc. (for all models)

Fit may be **less interesting** to econometricians and meteorologists: Their primary interest is **forecasting**.

To conclude

- time series analysis is a large class of diverse techniques to analyze N=1 data
- ARMA models are only a small (but basic) part of this
- time series models may be extended with cycles or trends over time
- in psychology we typically have N>1; there are different ways of handling this

References and suggested readings

- Chow, S.-M., Ho, M-H. R., Hamaker, E. L., & Dolan, C. V.(2010). Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques. *Structural Equation Modeling: A Multidisciplinary Journal, 17*, 303-332.
- Granger & Morris (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society, 139*, 246-257.
- Hamaker (2012). Why researchers should think within-person: A paradigmatic rationale. In Mehl & Conner (Eds.), *Handbook of research methods for studying daily life.* (pp. 43-61). New York, NY: The Guilford Press.
- Hamaker, E. L., & Dolan, C. V. (2009). Idiographic data analysis: Quantitative methods - from simple to advanced. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra and N. Chaudhary (Eds). *Dynamic Process Methodology in the Social and Developmental Sciences*, 191-216. New York: Springer-Verlag.
- Hamaker, E. L. and Dolan, C. V. and Molenaar, P. C. M. (2003). ARMA-based SEM when the number of time points T exceeds the number of cases N: Raw data maximum likelihood. Structural Equation Modeling, 10, 352-379
- Hamaker, E. L. and Dolan, C. V. and Molenaar, P. C. M. (2002). On the nature of SEM estimates of ARMA parameters. Structural Equation Modeling, 9, 347-368.
- Hamaker, & Grasman (2014). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology, 5*, 1492. doi:10.3389/fpsyg.2014.01492
- Hamilton, J. D. (1994). *Time series analysis.* Princeton, NJ: Princeton University Press.

References and suggested readings

- Kim, C-J, and Nelson, C. R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* Cambridge, MA: The MIT Press.
- Raudenbush S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods (Second Edition)*. Thousand Oaks, CA: Sage Publications.
- Rosmalen, Wenting, Roest, de Jonge & Bos (2012). Reveaing causal heterogeneity using time series analysis of ambulatory assessments: Application to the association between depression and physical activity after myocardial infarction. *Psychosomatic Medicine, 74*, 377-389.
- Schmitz (2000). Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen. *Psychologische Rundschau, 51*, 83-92.
- Schmitz & Skinner (1994). Perceived control, effort, and academic performance: Interindividual, intraindividual, and multivariate time-series analyses. *Journal of Personality and Social Psychology, 64*, 1010-1028.
- Snippe, Bos, van der Ploeg, Sanderman, Fleer & Schroevers (2014). Time-series analysis of daily changed in mindfulness, repetitive thinking, and depressive symptoms during mindfulness-based treatment. *Mindfulness,* doi:10.1007/s12671-014-0354-7.
- van Gils, Burton, Bos, Janssens, Schoevers, & Rosmalen (2014). Individual variation in temporal relationships between stress and functional somatic symptoms. *Journal of Psychosomatic Research, 77*(1), 34-39.

Mplus 8: Dynamic SEM

Applications

Ellen L. Hamaker
Utrecht University

Tihomir Asparouhov & Bengt Muthén
Muthén & Muthén

May 23, 2016

Intensive longitudinal data

Two approaches we can take when T is large and N>1:

1. **Top-down approach** (i.e., **dynamic multilevel modeling**):
   - use time series models as level 1
   - allow for quantitative individual differences in model dynamics at level 2
   - can be used with relative small T (say 20), but requires at least moderate N (say >30)

2. **Bottom-up approach** (i.e., **replicated time series analysis**)
   - use time series models to model N=1 data
   - allow for quantitative and qualitative differences between persons
   - can be used with small N (say 2), but requires relative large T (say >50)

Alternative approach: **pooled time series analysis** (requires N*T>50).

Outline

1. Top-down approach:
   - **Univariate multilevel AR(1) model**
   - Multiple indicator multilevel AR(1) model
   - Multilevel VAR(1) model
2. Bottom-up approach:
   - Comparison of linear models and regime-switching models
3. Discussion

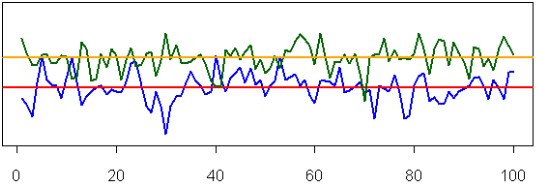
Univariate multilevel AR(1) model: Random mean

Centering part:

$$PA_{it} = \mu_i + PA^*_{it}$$

where

- μ_i is the individual's **mean** (i.e., baseline, trait, equilibrium) of positive affect
- PA^*_{it} is the **within-person centered** (cluster-mean centered) score



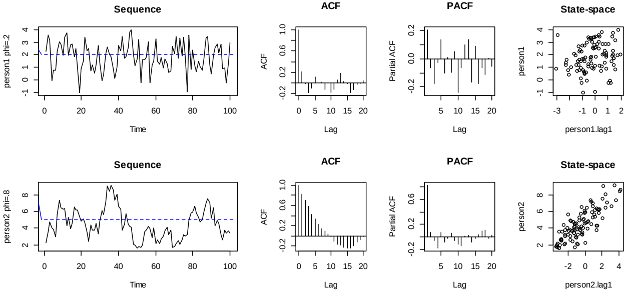
Univariate multilevel AR(1) model: Random inertia

Autoregressive part:

$$PA^*_{it} = \phi_i PA^*_{i,t-1} + \zeta_{it}$$

where

- ϕ_i is the **autoregressive parameter** (i.e., inertia, carry-over, or regulatory weakness)
- ζ_{it} is the **innovation** (residual, disturbance, dynamic error) (with $\zeta_{it} \sim N(0, \sigma^2_\zeta)$)



Univariate multilevel AR(1) model: Level 1

Putting these together we can write:

Level 1: Random mean and inertia

$$PA_{it} = \mu_i + \phi_i PA^*_{i,t-1} + \zeta_{it}$$

where $\zeta_{it} \sim N(0, \sigma^2)$.

Level 2:

$$\mu_i = \mu + v_{0i}$$
$$\phi_i = \phi + v_{1i}$$

where

$$\begin{bmatrix} v_{0i} \\ v_{1i} \end{bmatrix} \sim MN \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{11} & \\ \psi_{21} & \psi_{22} \end{bmatrix} \right]$$

Intermezzo: Centering level 1 predictors?

There are three ways in which we can include level 1 predictors:

- non-centered (**NC**)
- grand mean centered (**GMC**)
- cluster mean centered (**CMC**)

NC and GMC are **equivalent** (i.e., alternative parametrizations).

CMC is **equivalent under some circumstances** (i.e., no random slopes, and predictor means included as level 2 predictor of random intercept), but not always.

Converging consensus: The slope from NC/GMC can be an "**uninterpretable blend**" of the within and between relationship (Raudenbush & Bryck, 2002).

Intermezzo: Centering the lagged predictor?

Hamaker and Grasman (2015) compared four ways of centering the **lagged predictor** in a multilevel AR(1) model:

- NC: no centering
- CMC($\bar{y}_{.i}$): cluster mean centering using the sample mean
- CMC($\hat{\mu}_i$): cluster mean centering using the multilevel estimate
- CMC(μ_i): cluster mean centering using the true mean

Table 4 | Bias and coverage rates for fixed autoregressive parameter ϕ in multilevel autoregressive model under diverse scenarios.

AR parameter	Sample size		Bias				$CR_{0.95}$			
	N	T	NC	C($\bar{y}_{.i}$)	C($\hat{\mu}_i$)	C(μ_i)	NC	C($\bar{y}_{.i}$)	C($\hat{\mu}_i$)	C(μ_i)
$\phi_i \sim N(0.3, 0.1)$	20	20	0.002	−0.072	−0.069	−0.068	0.928	0.762	0.785	0.787
		50	0.000	−0.027	−0.027	−0.026	0.940	0.900	0.901	0.898
		100	0.000	−0.013	−0.013	−0.013	0.932	0.932	0.932	0.932
	50	20	0.005	−0.071	−0.069	−0.067	0.893	0.480	0.512	0.518
		50	0.001	−0.027	−0.026	−0.026	0.936	0.800	0.804	0.805
		100	0.000	−0.013	−0.013	−0.013	0.946	0.902	0.902	0.903
	100	20	0.006	−0.070	−0.068	−0.066	0.892	0.196	0.227	0.242
		50	0.001	−0.027	−0.027	−0.027	0.930	0.623	0.630	0.637
		100	0.000	−0.013	−0.013	−0.013	0.930	0.851	0.854	0.851

Intermezzo: Centering the lagged predictor?

Conclusion (from Hamaker & Grasman, 2015):

- **CMC leads to a downward bias** in the estimation of the AR parameter
- **CMC is better** when interest is in a **level 2 predictor of the AR parameter**

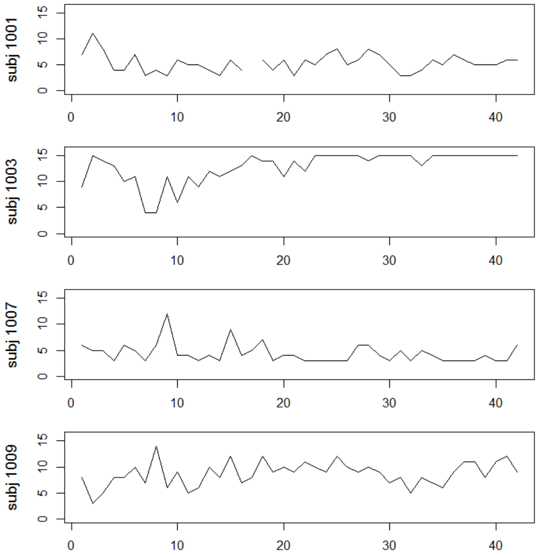
Note that when N=1, the OLS estimate of the AR parameter is known to be biased (e.g., Marriott & Pope, 1954).

BUT: CMC in Mplus is not associated with this bias (nor is it in WinBUGS, see Jongerling et al., 2015), probably because the **same (individual) parameter** is used as the intercept and for CMC of the lagged predictor.

NOTE: CMC is the default in Mplus when creating lagged variables.

Daily diary data on positive affect (PA)

Data: 89 females measured for 42 days (see Jongerling, Laurenceau & Hamaker, 2015).



Input: Create an observed lagged variable

```
TITLE: Multilevel AR(1) with random mean

DATA: file is fem.dat;

VARIABLE:
names=subj couple day   dhappy
dexcited    denerget    denthusi    PA;
cluster=subj;
useobs are
(subj .ne. 1003) .and.
(subj .ne. 1107) .and.
(subj .ne. 1223) .and.
(subj .ne. 1233) .and.
(subj .ne. 1249) .and.
(subj .ne. 1327) .and.
(subj .ne. 1425);

MISSING = all(999);

USEVAR are PA;

LAGVAR = PA(1);  ! CREATE AN OBSERVED LAGGED VARIABLE
```

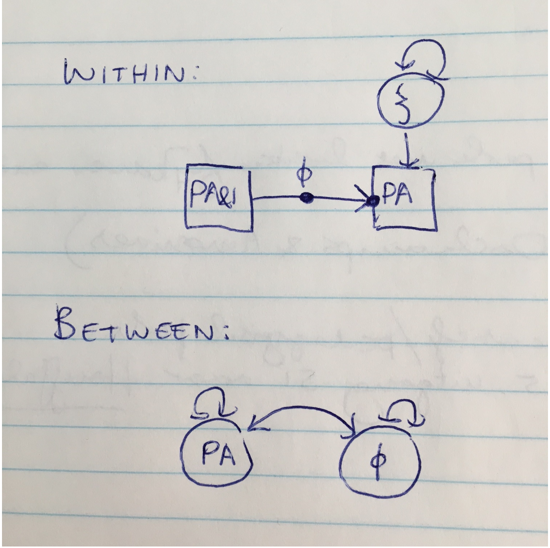
NOTE: Using **LAGVAR = PA(1);** gives a lagged variable based on lagging the observed variable PA by one.

Input: Random AR parameter and random mean

```
ANALYSIS:   TYPE IS TWOLEVEL random;
            estimator=bayes;
            fbiter=10000;
            bseed = 7487;
            proc = 2;

MODEL:

  %WITHIN%
  phi | PA on PA&1; ! AUTOREGRESSION IS RANDOM

  %BETWEEN%
  PA with phi;      ! CORRELATED RANDOM MEAN AND AR
```

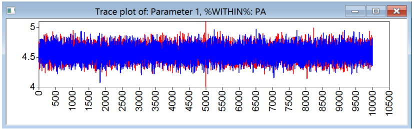
NOTE: The lagged variable (created by **LAGVAR = PA(1);**) is referred to as **PA&1**.

Path diagram of the multilevel AR(1) model

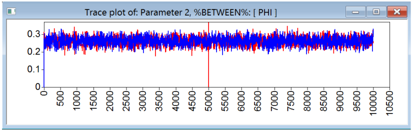


Results: Trace plots (10,000 iterations)

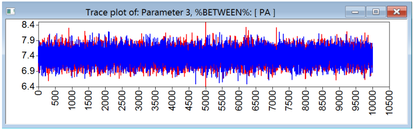
Level 1 residual variance:



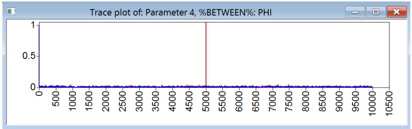
AR parameter:



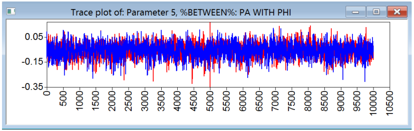
Average mean:



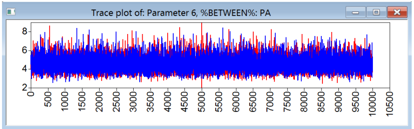
Variance of AR parameter:



Cov. mean and AR parameter:



Variance of mean:



Results: Parameter estimates

MODEL RESULTS

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Within Level						
Residual Variances						
PA	4.563	0.109	0.000	4.357	4.784	*
Between Level						
PA WITH						
PHI	-0.053	0.049	0.129	-0.152	0.039	
Means						
PA	7.393	0.231	0.000	6.933	7.842	*
PHI	0.263	0.021	0.000	0.221	0.304	*
Variances						
PA	4.470	0.752	0.000	3.316	6.260	*
PHI	0.010	0.005	0.000	0.002	0.022	*

Testing whether a random effect is significant is problematic; instead we can compare two models (with and without a random effect).

Input: Fixed AR parameter and random mean

```
ANALYSIS:   TYPE IS TWOLEVEL random;
            estimator=bayes;
            fbiter=10000;
            bseed = 6186;

MODEL:

  %WITHIN%
  PA on PA&1 (phi); ! AUTOREGRESSION

  %BETWEEN%
  PA;               ! RANDOM MEAN

OUTPUT: TECH8 TECH1;

PLOT: TYPE = PLOT2;
```

In this model there is no random AR parameter; only a random mean.

Random AR parameter?

Warning: Make sure the DIC is **stable** (this may take *many more iterations* than apparent from trace plots).

To ensure the DIC is stable, run the model at least **twice with a different seed**: This should give the same DIC and pD.

Here we compare the model with a fixed AR parameter (ϕ) to a model with a random AR parameter (ϕ_i).

Model	DIC	pD
ϕ	16501	192
ϕ_i	16498	216

Only slight preference for model with random AR parameter.

Literature on inertia

Affective inertia has been **empirically related to**

- neuroticism (+) and agreeableness (-) (Suls, Green & Hillis, 1998)
- concurrent depression (+) (Kuppens, Allen & Sheeber, 2010, *Psychological Science*)
- future depression (+) (Kuppens, Sheeber, Yap, Whittle, Simmons & Allen, 2012)
- rumination (+) (Koval, Kuppens, Allen & Sheeber, 2012)
- self-esteem (-) (Houben, Van den Noortgate & Kuppens, 20150)
- life-satisfaction (-) (Houben et al., 2015)
- PA (-) and NA (+) (Houben et al., 2015)

Note that inertia in positive affects seems also maladaptive.

Autoregressive parameter in **daily drinking behavior** has been positively related to being female (Rovine & Walls, 2006); however, the **average** was close to **zero**.

Extension 1: Random innovation variance

Level 1: Random mean, inertia, and innovation variance

$$PA_{ti} = \mu_i + \phi_i PA^*_{t-1,i} + \sigma_i \zeta_{ti}$$

where $\zeta_{ti} \sim N(0, 1)$.

Level 2:

$$\mu_i = \mu + v_{0i}$$
$$\phi_i = \phi + v_{1i}$$
$$\sigma_i = \sigma + v_{2i}$$

where

$$\begin{bmatrix} v_{0i} \\ v_{1i} \\ v_{2i} \end{bmatrix} \sim MN \left[\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \right]$$

Why random innovation variance? Statistical

For N=1 we have: $y_t = \mu + \phi(y_{t-1} - \mu) + \zeta_t$, such that:

$$\begin{aligned} Var(y_t) = E\Big[\Big\{y_t - \mu\Big\}^2\Big] &= E\Big[\Big\{\mu + \phi(y_{t-1} - \mu) + \zeta_t - \mu\Big\}^2\Big] \\ &= E\Big[\Big\{\phi(y_{t-1} - \mu) + \zeta_t\Big\}^2\Big] \\ &= \phi^2 E\Big[\Big\{y_{t-1} - \mu\Big\}^2\Big] + \sigma^2 \end{aligned}$$

where $E\Big[\Big\{y_t - \mu\Big\}^2\Big] = E\Big[\Big\{y_{t-1} - \mu\Big\}^2\Big] = \sigma_y^2$

$$\begin{aligned} \sigma_y^2 &= \phi^2 \sigma_y^2 + \sigma^2 \\ \sigma_y^2 - \phi^2 \sigma_y^2 &= \sigma^2 \\ (1 - \phi^2)\sigma_y^2 &= \sigma^2 \\ \sigma_y^2 &= \frac{\sigma^2}{1 - \phi^2} \end{aligned}$$

Hence, **individual differences** in σ_y^2 can come from individual differences in ϕ **and/or** σ^2.

Why random innovation variance? Substantive

Level 1: Random mean, inertia, and innovation variance

$$PA_{ti} = \mu_i + \phi_i PA^*_{t-1,i} + \sigma_i \zeta_{ti}$$

where $\zeta_{ti} \sim N(0, 1)$.

Substantive interpretation of random innovation variance:

- individual differences in exposure
- individual differences in reactivity

Level 1: Reactivity to Positive Events (PE)

$$PA_{ti} = \mu_i + \phi_i PA^*_{t-1,i} + \beta_i PE^*_{ti} + \zeta_{ti}$$

Some results for stress sensitivity and reward experience:

- Suls et al. (1998)
- Wichers: relationship with depression and effect of therapy

Extension 2: Measurement error

Level 1: Measurement equation

$$PA_{it} = \mu_i + \eta_{it} + \epsilon_{it}$$

where

- μ_i is the individual's mean
- η_{it} is the individual's true score at occasion t
- ϵ_{it} is the individual's measurement error at occasion t (could also consider individual differences in its variance)

Level 1: Transition equation

$$\eta_{it} = \phi_i \eta_{i,t-1} + \sigma_i \zeta_{it}$$

where $\zeta_{it} \sim N(0,1)$.

Some thoughts about measurement error in a multilevel AR(1) model:

- advantage: separate signal from noise
- advantage: reliability per person
- disadvantage: AR-effects in error end up in signal
- disadvantage: not identified when $\phi = 0$

Outline

1. Top-down approach:
   - Univariate multilevel AR(1) model
   - **Multiple indicator multilevel AR(1) model**
   - Multilevel VAR(1) model

2. Bottom-up approach:
   - Comparison of linear models and regime-switching models

3. Discussion

Multiple indicator AR(1) model for PA

We have three indicators: excited (EXC), energetic (ENE), and enthusiastic (ENT).

Level 1: Within-person factor model

$$\begin{bmatrix} EXC_{it} \\ ENE_{it} \\ ENT_{it} \end{bmatrix} = \begin{bmatrix} \mu_{EXC,i} \\ \mu_{ENE,i} \\ \mu_{ENT,i} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda_{2W} \\ \lambda_{3W} \end{bmatrix} PAW_{it} + \begin{bmatrix} \epsilon_{EXC,it} \\ \epsilon_{ENE,it} \\ \epsilon_{ENT,it} \end{bmatrix}$$

where

- μ's are the individual's means
- λ's are the within-person factor loadings
- PAW_{it} is the individual's latent score at occasion t
- ϵ's are the individual's measurement errors at occasion t

Multiple indicator AR(1) model for PA

Note that PAW_{it} has a mean of zero for each person (hence no within-person means here).

Level 1: Within-person latent AR(1)

$$PAW_{it} = \phi_i PAW_{i,t-1} + \sigma_i \zeta_{it}$$

where

- ϕ_i is the individual's autoregressive parameter
- $\sigma_i \zeta_{it}$ is the individual's innovation at occasion t (with var(ζ)=1)

Multiple indicator AR(1) model for PA

Level 2: Between-person factor model

$$\begin{bmatrix} \mu_{EXC,i} \\ \mu_{ENE,i} \\ \mu_{ENT,i} \end{bmatrix} = \begin{bmatrix} \mu_{EXC} \\ \mu_{ENE} \\ \mu_{ENT} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda_{2B} \\ \lambda_{3B} \end{bmatrix} PAB_i + \begin{bmatrix} \epsilon_{EXC,i} \\ \epsilon_{ENE,i} \\ \epsilon_{ENT,i} \end{bmatrix}$$

Level 2: Fixed and random effects

$$PAB_i = v_{0i}$$
$$\phi_i = \phi + v_{1i}$$
$$\zeta_i = \zeta + v_{2i}$$

where

$$\begin{bmatrix} v_{0i} \\ v_{1i} \\ v_{2i} \end{bmatrix} \sim MN \left[\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \right]$$

Input: Multiple indicator AR(1) model

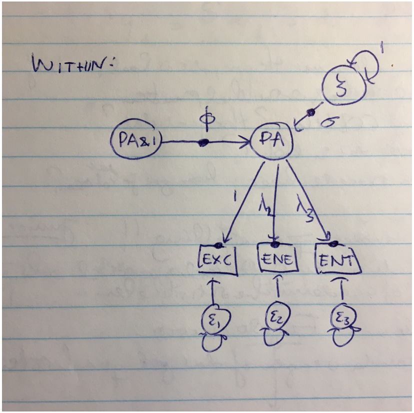
Allowing for:

- random means
- random autoregression
- random innovation SD

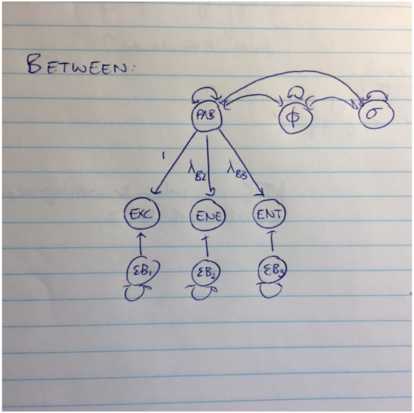
```
MODEL:
  %WITHIN%
  PA BY excited energet enthusi (&1);! FACTOR MODEL AND LAGGED LATENT VARIABLE
  PA@0;                              ! FIX THE RESIDUAL TO ZERO
  zeta BY;                           ! CREATE AN INNOVATION TERM
  PA with zeta@0;                    ! FIX COVARIANCE BETWEEN PA AND ZETA TO ZERO
  zeta@1;                            ! FIX VARIANCE OF THIS TERM TO 1
  sigma | PA on zeta;                ! ALLOW FOR A RANDOM LOADING: INDIVIDUAL SD OF THE INNOVATION
  phi | PA on PA&1;                  ! AUTOREGRESSION IS RANDOM

  %BETWEEN%
  PAB BY excited energet enthusi;    ! FACTOR MODEL
  PAB with sigma;                    ! ALLOW FOR CORRELATED RANDOM EFFECTS
  PAB with phi;                      ! ALLOW FOR CORRELATED RANDOM EFFECTS
  phi with sigma;                    ! ALLOW FOR CORRELATED RANDOM EFFECTS
  [phi*0.2]; phi*0.03;
  [sigma*1.2]; sigma*0.1;
```

Path diagram

Within level:



Between level:



Results: Parameter estimates (within)

MODEL RESULTS

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Within Level						
PA BY						
EXCITED	1.000	0.000	0.000	1.000	1.000	
ENERGET	0.953	0.029	0.000	0.898	1.012	*
ENTHUSI	1.049	0.029	0.000	0.993	1.108	*
PA WITH						
ZETA	0.000	0.000	1.000	0.000	0.000	
Variances						
ZETA	1.000	0.000	0.000	1.000	1.000	
Residual Variances						
EXCITED	0.431	0.014	0.000	0.404	0.459	*
ENERGET	0.323	0.012	0.000	0.300	0.346	*
ENTHUSI	0.318	0.012	0.000	0.294	0.343	*
PA	0.001	0.000	0.000	0.001	0.001	

Remember: $Var(PA_i) = \frac{\sigma_i^2}{1-\phi_i^2}$

Results: Parameter estimates (between)

```
PAB      BY
   EXCITED          1.000       0.000      0.000       1.000       1.000
   ENERGET          1.069       0.069      0.000       0.945       1.218      *
   ENTHUSI          1.035       0.067      0.000       0.915       1.178      *

PAB      WITH
   SIGMA            0.038       0.022      0.032      -0.002       0.085
   PHI             -0.033       0.022      0.056      -0.080       0.008

PHI      WITH
   SIGMA           -0.025       0.009      0.000      -0.046      -0.010      *

Means
   SIGMA            0.562       0.031      0.000       0.502       0.623      *
   PHI              0.393       0.029      0.000       0.336       0.450      *

Intercepts
   EXCITED          2.404       0.082      0.000       2.242       2.565      *
   ENERGET          2.513       0.083      0.000       2.349       2.676      *
   ENTHUSI          2.470       0.081      0.000       2.311       2.629      *

Variances
   PAB              0.470       0.095      0.000       0.321       0.692      *
   SIGMA            0.059       0.012      0.000       0.041       0.087      *
   PHI              0.025       0.011      0.000       0.009       0.051      *

Residual Variances
   EXCITED          0.086       0.019      0.000       0.056       0.130      *
   ENERGET          0.037       0.014      0.000       0.012       0.069      *
   ENTHUSI          0.035       0.013      0.000       0.011       0.064      *
```

NOTE: Means are the fixed effects, variances are the random effects.

Factorial invariance across levels

Are the **factor loadings** for PA **identical across levels**?

```
Within Level

 PA       BY
    EXCITED            1.000      0.000      0.000      1.000      1.000
    ENERGET            0.953      0.029      0.000      0.898      1.012      *
    ENTHUSI            1.049      0.029      0.000      0.993      1.108      *

Between Level

 PAB      BY
    EXCITED            1.000      0.000      0.000      1.000      1.000
    ENERGET            1.069      0.069      0.000      0.945      1.218      *
    ENTHUSI            1.035      0.067      0.000      0.915      1.178      *
```

If $\lambda_w = \lambda_b$, this implies that within-person, state-like fluctuations are **situated on the same underlying dimension** as stable between-person, trait-like differences.

DICs using 500,000 iterations

	$\lambda_w \neq \lambda_b$	$\lambda_w = \lambda_b$
	22355	22364
	22349	22358
	22353	22360
Average:	22352	22361

Outline

1. Top-down approach:
   - Univariate multilevel AR(1) model
   - Multiple indicator multilevel AR(1) model
   - **Multilevel VAR(1) model**

2. Bottom-up approach:
   - Comparison of linear models and regime-switching models

3. Discussion

Multilevel VAR(1) model

In a vector autoregressive (VAR) model, a vector is regressed on preceding versions of itself.

VAR(1):

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{\Phi}\boldsymbol{y}_{t-1} + \boldsymbol{\zeta}_t \qquad \text{with} \quad \boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{\Phi})^{-1}\boldsymbol{c}$$

Alternative expression of a VAR(1):

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\zeta}_t$$

When considering a multilevel extension, we want to allow for individual differences in:

- $\boldsymbol{\mu}$: the trait scores of individuals
- $\boldsymbol{\Phi}$: the inertias and cross-lagged relationships

NOTE: We write $\boldsymbol{y}^*_{t-1} = \boldsymbol{y}_{t-1} - \boldsymbol{\mu}$.

Example of a multilevel VAR(1) model

We make use of bivariate data from Emilio Ferrer: Positive Affect and Rumination (see Schuurman, Grasman & Hamaker, 2016).

Six days of ESM data with N=129 and T about 45.

Within level:

$$\begin{aligned}
\begin{bmatrix} PA_{it} \\ RU_{it} \end{bmatrix} &= \begin{bmatrix} \mu_{PA,i} \\ \mu_{RU,i} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} PA^*_{it-1} \\ RU^*_{it-1} \end{bmatrix} + \begin{bmatrix} \zeta_{PA,it} \\ \zeta_{RU,it} \end{bmatrix} \\
&= \begin{bmatrix} \mu_{PA,i} + \phi_{11} PA^*_{it-1} + \phi_{12} RU^*_{it-1} + \zeta_{PA,it} \\ \mu_{RU,i} + \phi_{21} PA^*_{it-1} + \phi_{22} RU^*_{it-1} + \zeta_{RU,it} \end{bmatrix}
\end{aligned}$$

Model specification

```
MODEL:

    %WITHIN%
    E1 BY PA@1 (&1);
    PA@0.01;
    E2 BY pieker@1(&1);
    pieker@0.01;
    E1 with E2;
    E1;
    E2;
    phi11 | E1 on E1&1;
    phi22 | E2 on E2&1;
    phi12 | E1 on E2&1;
    phi21 | E2 on E1&1;
```

At the between level the means and lagged effects are all allowed to correlate.

Results within level

```
Within Level

 E1       BY
    PA                 1.000      0.000      0.000      1.000      1.000

 E2       BY
    PIEKER             1.000      0.000      0.000      1.000      1.000

 E1       WITH
    E2                 0.496      0.047      0.000      0.413      0.593      *

 Residual Variances
    PA                 0.010      0.000      0.000      0.010      0.010
    PIEKER             0.010      0.000      0.000      0.010      0.010
    E1                 1.961      0.046      0.000      1.890      2.063      *
    E2                 2.640      0.062      0.000      2.518      2.759      *
```

Note that the measurement error variances **fixed at 0.01** are **negligibly small** compared to the total variances.

Results between level

```
Between Level

 PA       WITH
    PHI11          0.025      0.007      0.000      0.014      0.040      *
    PHI12          0.015      0.008      0.035      0.000      0.034
    PHI21         -0.033      0.011      0.000     -0.052     -0.010      *
    PHI22         -0.028      0.011      0.000     -0.056     -0.009      *

 PIEKER   WITH
    PHI11         -0.028      0.008      0.000     -0.046     -0.015      *
    PHI12         -0.012      0.010      0.110     -0.034      0.006
    PHI21          0.045      0.013      0.000      0.020      0.074      *
    PHI22          0.067      0.015      0.000      0.041      0.103      *

 PHI11    WITH
    PHI12         -0.002      0.002      0.040     -0.006      0.000
    PHI21         -0.006      0.002      0.000     -0.010     -0.003      *
    PHI22         -0.002      0.002      0.070     -0.005      0.001

 PHI12    WITH
    PHI21          0.000      0.001      0.435     -0.003      0.003
    PHI22         -0.004      0.003      0.055     -0.010      0.001

 PHI21    WITH
    PHI22         -0.002      0.003      0.280     -0.008      0.003

 PA       WITH
    PIEKER        -0.070      0.048      0.085     -0.169      0.018
```

Results between level (continued)

Means						
PA	2.244	0.063	0.000	2.117	2.357	*
PIEKER	1.752	0.069	0.000	1.599	1.872	*
PHI11	0.620	0.008	0.000	0.605	0.635	*
PHI22	0.356	0.017	0.000	0.318	0.392	*
PHI12	0.140	0.011	0.000	0.117	0.160	*
PHI21	0.265	0.014	0.000	0.236	0.292	*
Variances						
PA	0.382	0.055	0.000	0.291	0.496	*
PIEKER	0.624	0.092	0.000	0.446	0.811	*
PHI11	0.006	0.001	0.000	0.003	0.009	*
PHI22	0.020	0.004	0.000	0.013	0.030	*
PHI12	0.006	0.002	0.000	0.003	0.010	*
PHI21	0.014	0.003	0.000	0.008	0.022	*

Means are the fixed effects; **variances** are for the random effects.

Standardizing the cross-lagged parameters

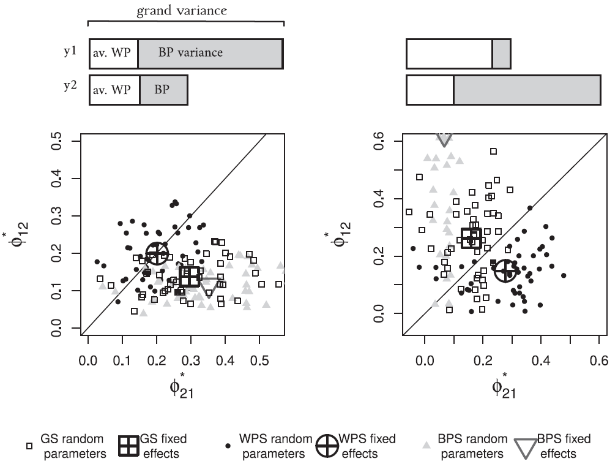
Schuurman et al. (2016) presents three forms of **standardization in multilevel models**:

- total variance (i.e., grand standardization)
- between-person variance (i.e., between standardization)
- average within-person variance
- within-person variance (i.e., within standardization)

Conclusion: last form is most meaningful, as it **parallels standardizing when N=1**.

Standardized fixed effect should be the **average standardized within-person effect**.

Does it make a difference?



From Schuurman et al. (2016)

Networks based on multilevel VAR models

Borsboom has used the idea of **networks as an alternative to latent variables** (in the context of psychopathology).

Dynamical networks are often based on a VAR(1) model.

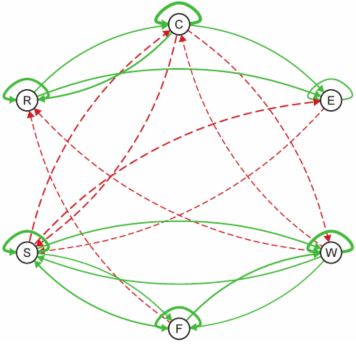
Bringmann et al. (2013) analyzed the lagged relationships between the following variables:

- cheerful (C)
- pleasant event (E)
- worry (W)
- fearful (F)
- sad (S)
- relaxed (R)

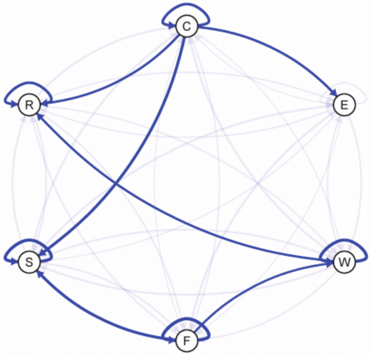
NOTE: They performed **separate multilevel regression analyses** on each of these variables, using all (lagged) variables as predictors.

Results at the population level

Average (fixed effects) network

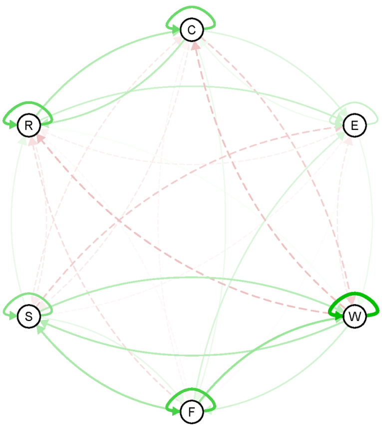


Individual differences network



C=cheerful; E=pleasant event; W=worry; F=fearful; S=sad; and R=relaxed; red solid lines represent positive relationships; green dashed lines represent negative relationship. From Bringmann et al. (2013)

Results at the individual level (2 individuals)



C=cheerful; E=pleasant event; W=worry; F=fearful; S=sad; and R=relaxed
From Bringmann et al. (2013)

Outline

1. Top-down approach:
   - Univariate multilevel AR(1) model
   - Multiple indicator multilevel AR(1) model
   - Multilevel VAR(1) model

2. Bottom-up approach:
   - **Comparison of linear models and regime-switching models**

3. Discussion

Bottom-up: Replicated time series analysis

Characteristics of TSA include:

- N=1
- T is large
- observations are ordered (in time)

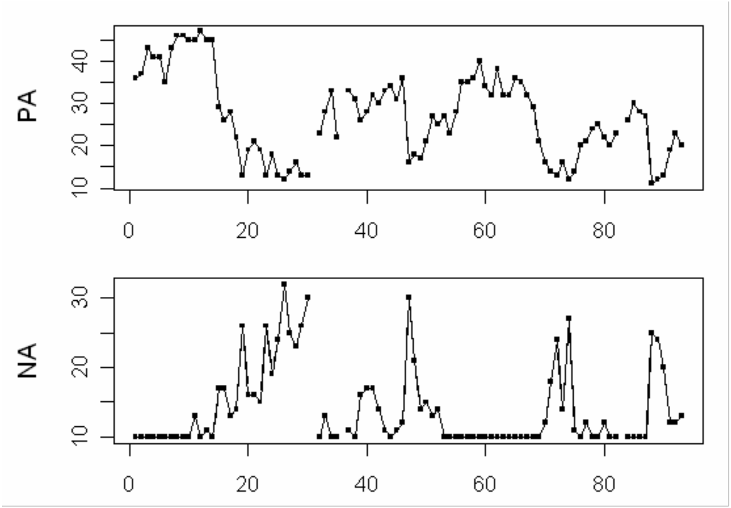
Goals of TSA include:

- prediction and forecasting: weather, currency, earthquakes, epidemic
- signal estimation (Kalman filter): e.g. to control your spacecraft
- identify the nature of the process

Example considered here is based on Hamaker, Grasman and Kamphuis (2016).

Bipolar disorder (BD)

Bipolar disorder is characterized by severe changes in affect and activity: Bipolar patients suffer from **manic** and **depressed episodes**.



BAS dysregulation in BD

BAS may play a crucial role:

- **active BAS**: expecting reward; difficulty inhibiting behavior when approaching a goal; hope
- **inactive BAS**: not expecting reward; difficulty to be motivated; despair

Two forms of BAS dysregulation:

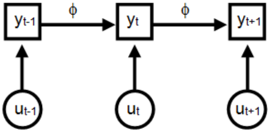
Slow return to baseline

Switches between distinct states

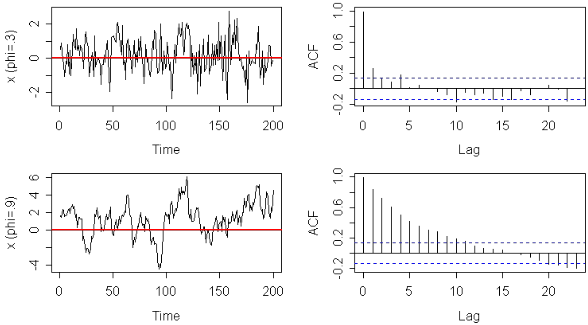
Slow-return-to-baseline model 1: AR(1)

AR(1)

$$y_t = c + \phi y_{t-1} + u_t$$



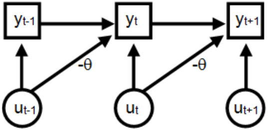
Carry-over: In the AR(1) model today's mood is influenced by yesterday's mood, and the higher ϕ, the more yesterday's mood carries over to today's mood.



Slow-return-to-baseline model 2: ARIMA(0,1,1)

ARIMA(0,1,1)

$$y_t = y_{t-1} - \theta u_{t-1} + u_t$$



Balancing preservation and adaption: The closer θ is to 1, the stronger preservation is; if θ is zero, the system fully adapts to perturbations.

$$\mathrm{E}[y_t | y_{t-1}] = y_{t-1} - \theta e_{t-1}$$

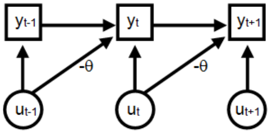
$$= \mathrm{E}[y_{t-1} | y_{t-2}] + e_{t-1} - \theta e_{t-1}$$

The parameter θ is considered to indicate the balance between **preservation** and **adaption**.

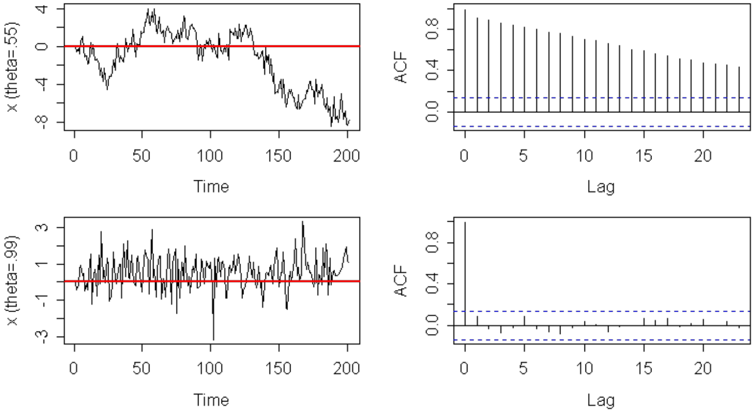
Slow-return-to-baseline model 2: ARIMA(0,1,1)

ARIMA(0,1,1)

$$y_t = y_{t-1} - \theta u_{t-1} + u_t$$



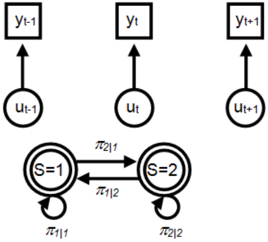
Balancing preservation and adaption: The closer θ is to 1, the stronger preservation is; if θ is zero, the system fully adapts to perturbations.



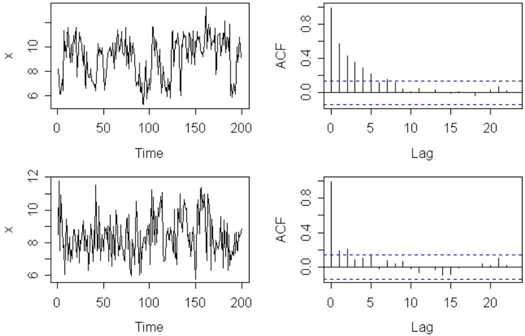
Regime-switching model 1: HM model

HMM

$$y_t = \mu_{S_t} + \sigma_{S_t} u_t$$



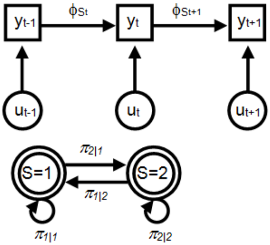
Switching: In the HMM model the system switches between two different WN processes (different means and variances). For each state, there is a probability to stay in it ($\pi_{1|1}$ and $\pi_{2|2}$) and a probabilities to switch ($\pi_{1|2}$ and $\pi_{2|1}$) .



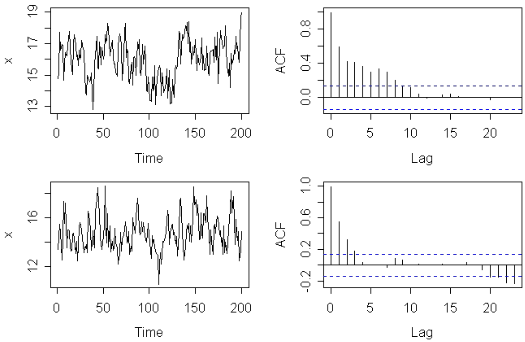
Regime-switching model 2: MSAR(1) model

MSAR(1)

$$y_t = c_{S_t} + \phi_{S_t} y_{t-1} + \sigma_{S_t} u_t$$



Switching with carry-over: The MSAR model is characterized by switches between two different AR(1) processes (different constant c, AR parameter ϕ and variance). Switches are smoother than in the HMM, due to the carry-over.



VAR(1) model and results

```
model:
    y1 with y2;
    y1 y2 on y1&1 y2&1;
```

Note we make use of observed lagged variables y1&1 and y2&1.

MODEL RESULTS

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Y1 ON						
Y1&1	0.881	0.079	0.000	0.717	1.042	*
Y2&1	0.041	0.140	0.379	-0.234	0.312	
Y2 ON						
Y1&1	-0.101	0.072	0.066	-0.246	0.037	
Y2&1	0.476	0.124	0.000	0.236	0.709	*
Y1 WITH						
Y2	-15.438	3.366	0.000	-23.886	-10.165	*
Intercepts						
Y1	2.439	3.873	0.242	-4.875	10.487	
Y2	9.931	3.443	0.004	3.565	17.121	*
Residual Variances						
Y1	26.481	4.307	0.000	19.982	36.577	*
Y2	22.405	3.674	0.000	17.120	31.582	*

VARIMA(0,1,1) model

```
model:
    e1 with e2;
    y1-y2@0.5; [y1-y2@0];
    e1 by y1@1 (&1);
    e2 by y2@1 (&1);
    y1 on y1&1@1 e1&1;
    y2 on y2&1@1 e2&1;
```

where:

- e1 by y1@1; defines e1 as the innovation of the process y1
- e1 by (&1); defines a lagged version of e1 (i.e., innovation at previous time point)
- y1 on y1&1@1; defines the I(1) part (random walk)
- y1 on e1&1; defines the MA(1) part (moving average process)

and:

- y1@0.5; sets the measurement error variance to a negligible small number
- and [y1@0]; sets the mean of the process to zero (because it is a unit root process; mean is not identified)

VARIMA(0,1,1) results

```
MODEL RESULTS

                                Posterior  One-Tailed         95% C.I.
                    Estimate       S.D.      P-Value   Lower 2.5%  Upper 2.5%  Significance

 E1       BY
    Y1                 1.000      0.000      0.000       1.000       1.000

 E2       BY
    Y2                 1.000      0.000      0.000       1.000       1.000

 Y1         ON
    E1&1              -0.200      0.098      0.025      -0.384      -0.000      *

 Y2         ON
    E2&1              -0.483      0.098      0.000      -0.658      -0.271      *

 Y1         ON
    Y1&1               1.000      0.000      0.000       1.000       1.000

 Y2         ON
    Y2&1               1.000      0.000      0.000       1.000       1.000

 E1       WITH
    E2               -17.078      3.639      0.000     -25.295     -11.432      *

 Intercepts
    Y1                 0.000      0.000      1.000       0.000       0.000
    Y2                 0.000      0.000      1.000       0.000       0.000

 Variances
    E1                27.380      4.583      0.000      20.380      37.988      *
    E2                24.409      3.974      0.000      18.196      33.781      *

 Residual Variances
    Y1                 0.500      0.000      0.000       0.500       0.500
    Y2                 0.500      0.000      0.000       0.500       0.500
```

HMM model

```
model:
    %overall%
    C on C&1;
    y1 with y2; y1-y2; [y1-y2];

model c:

     %C#1%

     y1 WITH y2*-0.12152 (v3);

     [ y1*2.02322 ];
     [ y2*1.66623 ];

     y1*0.40301 (v1);
     y2*0.27785 (v2);

     %C#2%

     y1 WITH y2*-0.12661 (w3);

     [ y1*2.05252 ];
     [ y2*1.61515 ];

     y1*0.40550 (w1);
     y2*0.20074 (w2);

model prior:
v1~IW(2,2);
v2~IW(2,2);
v3~IW(0,2);

w1~IW(2,2);
w2~IW(2,2);
w3~IW(0,2);
```

The overall model part:

- C ON C&1; specifies hidden Markov model
- y1 with y2; ensures the variables are allowed to correlate

Rest is used for specifying starting values and priors

HMM results

MODEL RESULTS

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Latent Class Pattern 1 1						
Y1 WITH						
Y2	-29.667	8.942	0.000	-52.482	-16.570	*
Means						
Y1	20.767	1.241	0.000	18.314	23.326	*
Y2	17.659	0.962	0.000	15.725	19.515	*
Variances						
Y1	59.325	14.126	0.000	39.518	94.774	*
Y2	37.273	8.495	0.000	25.278	58.170	*
Latent Class Pattern 1 2						
Y1 WITH						
Y2	0.176	0.394	0.317	-0.618	0.936	
Means						
Y1	33.508	1.283	0.000	30.991	35.930	*
Y2	10.044	0.055	0.000	9.949	10.157	*
Variances						
Y1	57.985	14.079	0.000	39.033	93.425	*
Y2	0.092	0.032	0.000	0.044	0.167	*
Categorical Latent Variables						
C#1 ON						
C&1#1	0.819	0.061	0.000	0.682	0.921	*
C&1#2	0.192	0.061	0.000	0.087	0.327	*
Class Proportions						
Class 1	0.409	0.031	0.000	0.341	0.460	
Class 2	0.091	0.031	0.000	0.039	0.158	
Class 3	0.096	0.031	0.000	0.044	0.164	
Class 4	0.404	0.031	0.000	0.336	0.456	

MSVAR(1) model

```
model:
    %overall%
    C on C&1;
    y1 with y2; y1-y2; [y1-y2];
     y1 y2 on y1&1 y2&1;

MODEL C:
     %C#1%
     y1 y2 on y1&1 y2&1;

     [ y1*20.76743 ] (1);
     [ y2*17.65870 ] (2);

     y1*59.32514 (v1);
     y2*37.27272 (v2);
     y1 WITH y2 (v3);

     %C#2%
     y1 y2 on y1&1 y2&1;

     [ y1*33.50785 ] (6);
     [ y2*10.04370 ] (7);

     y1*57.98539 (w1);
     y2*0.09211 (w2);
     y1 WITH y2*0 (w3);


model prior:
v1~IW(2,2);
v2~IW(2,2);
v3~IW(0,2);

w1~IW(2,2);
w2~IW(2,2);
w3~IW(0,2);
```

The overall model part:

- C ON C&1; specifies hidden Markov model
- y1 y2 on y1&1 y2&1; specifies a VAR(1) model
- y1 with y2; ensures the innovations are allowed to correlate

Rest is used for starting values and priors

MSVAR(1) results

MODEL RESULTS

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Latent Class Pattern 1 1						
Y1 ON						
Y1&1	0.814	0.131	0.000	0.543	1.053	*
Y2&1	0.133	0.182	0.219	-0.220	0.494	
Y2 ON						
Y1&1	-0.096	0.126	0.224	-0.338	0.159	
Y2&1	0.370	0.184	0.026	-0.001	0.732	
Y1 WITH						
Y2	-21.215	5.869	0.000	-35.638	-12.993	*
Intercepts						
Y1	1.008	5.500	0.428	-9.502	11.979	
Y2	13.713	5.439	0.007	2.799	24.266	*
Residual Variances						
Y1	27.773	6.527	0.000	18.603	43.894	*
Y2	29.673	6.904	0.000	20.094	46.646	*
Latent Class Pattern 1 2						
Y1 ON						
Y1&1	0.836	0.091	0.000	0.649	1.009	*
Y2&1	0.063	0.276	0.404	-0.477	0.611	
Y2 ON						
Y1&1	0.001	0.006	0.394	-0.010	0.013	
Y2&1	0.054	0.020	0.011	0.014	0.091	*
Y1 WITH						
Y2	-0.001	0.192	0.499	-0.368	0.407	
Intercepts						
Y1	5.076	5.182	0.155	-5.268	15.452	
Y2	9.401	0.341	0.000	8.728	10.097	*
Residual Variances						
Y1	17.086	4.395	0.000	11.082	27.990	*
Y2	0.063	0.024	0.000	0.038	0.130	*

MSVAR(1) results

```
Categorical Latent Variables

 C#1       ON
    C&1#1              0.807       0.064       0.000       0.663       0.914       *
    C&1#2              0.215       0.064       0.000       0.107       0.355       *

Class Proportions

    Class 1            0.404       0.032       0.000       0.332       0.457
    Class 2            0.096       0.032       0.000       0.043       0.168
    Class 3            0.107       0.032       0.000       0.054       0.177
    Class 4            0.393       0.032       0.000       0.322       0.446
```

Outline

1. Top-down approach:
   - Univariate multilevel AR(1) model
   - Multiple indicator multilevel AR(1) model
   - Multilevel VAR(1) model

2. Bottom-up approach:
   - Comparison of linear models and regime-switching models

3. **Discussion**

Some other issues to consider

- data may be irregularly spaced (e.g., ESM data), which should be taken into account when estimating lagged effects
- time is treated as discrete here, but it might be more appropriate to consider it as continuous (Deboeck & Preacher, 2015; Voelkle et al., 2012)
- there may be trends and cycles present which should (or not?) be accounted for (Liu & West, 2015; Wang & Maxwell, 2015)
- random factor loadings (allowing for idiographic loadings)
- level 2 predictors for the individual differences in dynamics
- time-varying parameters
- multilevel extension of the regime-switching models
- fit measure that allows for all models to be compared...

References and suggested readings

- Bringmann, Vissers, Wichers, Geschwind, Kuppens, Peeters, Borsboom & Tuerlinckx (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*, e60188, 1-13.
- Deboeck & Preacher (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling, 23*, 61-75.
- Geschwind, Peeters, Drukker, van Os & Wichers (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 79*, 618-628.
- De Haan-Rietdijk, Gottman, Bergeman & Hamaker (2014). Get over it! A multilevel threshold autoregressive model for state-dependent affect regulation. *Psychometika*. doi: 10.1007/s11336-014-9417-x
- Hamaker (2012). Why researchers should think within-person: A paradigmatic rationale. In Mehl & Conner (Eds.), *Handbook of research methods for studying daily life.* (pp. 43-61). New York, NY: The Guilford Press.
- Hamaker & Grasman (2014). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology, 5*, 1492. doi:10.3389/fpsyg.2014.01492
- Hamaker, Grasman & Kamphuis (2016). Modeling BAS dysregulation in Bipolar Disorder: Illustrating the potential of time series analysis. *Assessment*.

References and suggested readings

- Houben, Van den Noortgate & Kuppens, (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin, 141*, 901-930.
- Jongerling, Laurenceay & Hamaker (2015). A Multilevel AR(1) Model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research, 50*, 334-349.
- Koval, Kuppens, Allen & Sheeber (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition & Emotion, 26*, 1412-1427.
- Kuppens, Allen & Sheeber (2010). Emotional inertia and psychological maladjustment. *Psychological Science, 21*, 984-991.
- Kuppens, Sheeber, Yap, Whittle, Simmons & Allen (2012). Emotional inertia prospectively predicts the onset of depressive Multilevel AR(1) model 33 disorder in adolescence. *Emotion, 12*, 283-289.
- Liu & West (2015). Weekly cycles in daily report data: An overlooked issue. *Journal of Personality*. doi: 10.1111/jopy.12182
- Marriott & Pope(1954). Bias in the estimation of autocorrelations. *Biometrika, 41*, 390âĂŞ402. doi:10.1093/biomet/41.3-4.390
- Rovine & Walls (2006). A multilevel autoregressive model to describe interindividual differences in the stability of a process. In Schafer & Walls (Eds.), *Models for intensive longitudinal data* (pp. 124-147). New York, NY: Oxford.

References and suggested readings

- Raudenbush S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods (Second Edition)*. Thousand Oaks, CA: Sage Publications.
- Schuurman, Ferrer, de Boer-Sonnenschein & Hamaker (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods.*
- Schuurman, Houtveen, & Hamaker (2015). Incorporating measurement error in n=1 psychological autoregressive modeling. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.01038
- Suls, Green & Hillis (1998). Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin, 24*, 127-136.
- Voelkle, Oud, Davidov & Schmidt (2012). An SEM approach to continuous time modeling of panel data: relating authoritarianism and anomia. *Psychological Methods, 17*, 176-192. doi: 10.1037/a0027543
- Wang, Hamaker & Bergeman (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods, 17*, 567-581.
- Wang & Maxwell (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*, 63-83. http://dx.doi.org/10.1037/met0000030
- Wichers, Barge-Schaapman, Nicolson, Peeters, de Vries, Mengelers & van Os (2009). Reduced stress-sensitivity or increased reward experience: The psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology, 34*, 923-931.

Thank you

e.l.hamaker@uu.nl