Which Methods Do We Need for Intensive Longitudinal Data?

Bengt Muthén

Tihomir Asparouhov & Ellen Hamaker

bmuthen@statmodel.com
Mplus
www.statmodel.com

Keynote address at the Modern Modeling Methods Conference
UConn, May 24, 2016

Expert assistance from Noah Hastings is acknowledged
“Foreshadowing or guessing ahead is a literary device by which an author hints what is to come. Foreshadowing is a dramatic device in which an important plot-point is mentioned early in the story and will return in a more significant way. It is used to avoid disappointment. It is also sometimes used to arouse the reader”.

The Transitional Aspect of Snow (Foreshadowing)
Non-intensive longitudinal data:
- $T$ small (2 - 10) and $N$ large
- Modeling: Auto-regression and growth

Intensive longitudinal data:
- $T$ large (30-200) and $N$ smallish (even $N = 1$) but can be 1,000.
  - Often $T > N$
- Modeling: We shall see
Ecological momentary assessment (EMA): a research participant repeatedly reports on symptoms, affect, behavior, and cognitions close in time to experience and in the participants’ natural environment using smartphone, handheld computer, or GPS.

Experience sampling method (ESM): a research procedure for studying what people do, feel, and think during their daily lives.

Daily diary measurements

Burst of measurement

Ambulatory assessment (Trull & Ebner-Priemer, 2014. *Current Directions in Psychological Science*)
### Examples of Intensive Longitudinal Data Sets

<table>
<thead>
<tr>
<th>Data Sets</th>
<th>( N )</th>
<th>( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trull data in Jahng et al. (2008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mood:</td>
<td>84</td>
<td>76-186</td>
</tr>
<tr>
<td>Bergeman data in Wang et al. (2012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive and negative affect:</td>
<td>230</td>
<td>56</td>
</tr>
<tr>
<td>Shiffman data in Hedeker et al. (2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking urge:</td>
<td>515</td>
<td>34</td>
</tr>
<tr>
<td>Laurenceau data in Jongerling et al. (2015)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive affect:</td>
<td>96</td>
<td>42</td>
</tr>
</tbody>
</table>
Publications on Analysis of Intensive Longitudinal Data

Publications on experience sampling, ambulatory assessment, ecological momentary assessment, or daily diary

Number of publications vs. Year

- PsycINFO
- PubMed
Common Methods for Non-Intensive Longitudinal Data

$N$ large and $T$ small (2 - 10):

(1) Auto-Regression Modeling

Cross-lagged modeling:
Common Methods for Non-Intensive Longitudinal Data:
(2) Growth Modeling

Individual Curves

Time

y outcome

y1 y2 y3 y4 y5
i
s

0.0 0.5 1.0 1.5 2.0 2.5 3.0
0 5 10 15

Individual Curves
The Fashion Pendulum of Longitudinal Modeling

Auto-regression modeling

Growth modeling

Growth with AR

Auto-regression modeling

Growth modeling
Growth Modeling Adding Correlated Residuals
Growth Modeling Adding Auto-Regression (ALT Model)

\[
y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5
\]

Methods for Intensive Longitudinal Data 13/67
Growth Modeling Adding Auto-Regressive Residuals

Instead of $T = 5$, ILD has $T = 50$ or $100$: Too wide for the single-level wide approach (and often $T > N$)

Within-individual correlation across time is a nuisance in growth modeling but a focus of ILD analysis

Development over time typically not as simple for ILD data as using a few random slope growth factors: seasonality, cycles
Growth Modeling from a Two-Level Perspective

Within (level-1) Variation across time
Between (level-2) Variation across individual
Two-Level, Long Format Version

Within (level-1)
Variation across time

Between (level-2)
Variation across individual

y
w
i
s
Two-Level, Time-Series Version

- 3 key features: Random mean ($y$), random autoregression ($\phi$), random variance (not shown)

Within (level-1) Variation across time
Between (level-2) Variation across individual

Methods for Intensive Longitudinal Data 19/ 67
3 Key Features of Two-Level Time-Series Model: Inter-Individual Differences in Intra-Individual Characteristics

1. Random mean: Individual differences in level
2. Random autoregression: Inertia (resistance to change). Related to:
   - Neuroticism and agreeableness (Suls et al., 1998)
   - Depression (Kuppens et al., 2010)
   - Rumination, self-esteem, life satisfaction, pos. and neg. affect (gender)
3. Random variance: Innovation variance
   - Individual differences in reactivity (stress sensitivity) and exposure
Two-Level Time-Series Analysis
How Big do $N$ and $T$ Need to be?

Within (level-1) Variation across time

Between (level-2) Variation across individual

Bengt Muthén
MONTECARLO:  NAMES = y w z;
             NOBSERVATIONS = 5000;
             NREP = 1000;
             NCSIZES = 1;
             CSIZES = 100(50); ! N=100, T= 50
             LAGVARIABLES = y(1);
             BETWEEN = w z;

ANALYSIS:  TYPE = TWOLEVEL RANDOM;
             ESTIMATOR = BAYES;
             BITERATIONS = (1000);
             PROCESSORS = 2;
             ! MODEL MONTECARLO omitted

MODEL:  %WITHIN%
         phi  |  y ON y&1*0.5;
               y*1;
         %BETWEEN%
               ! w*1;
               y ON w*.3;
               y*0.09;
         phi ON w*.1;
         phi*.01; [phi*.3];
         z ON y*.5 phi*.7;
         z*0.0548;
E(φ) = 0.3, V(φ) = 0.02, 1000 replications. Ignorable bias, good coverage in all cases. Power results:

- \( N = 25, T = 50: \phi \) on \( w = 0.73, z \) on \( \phi = 0.15 \)
- \( N = 50, T = 50: \phi \) on \( w = 0.96, z \) on \( \phi = 0.32 \)
- \( N = 100, T = 50: \phi \) on \( w = 1.00, z \) on \( \phi = 0.58 \)
- \( N = 200, T = 50: \phi \) on \( w = 1.00, z \) on \( \phi = 0.87 \)
What About Other Methods for Intensive Longitudinal Data: Part 1. Multilevel Analysis

- **Papers/talks:**
  - Hedeker (2015): Keynote address at the 2015 M3 meeting

- **Data from MIXREGLS:**
  - Positive mood related to alone (tvc) and gender (tic), $N = 515$, $T = 34$ ($3 - 58$)
Two-Level Modeling of Hedeker Mood data: Hedeker’s Model (Left) versus Time-Series Model (Right)
The Transitional Aspect of Snow: Temperature and Snow Depth
Bivariate Time-Series Data with a Lagged Effect
- Implications for Sledding
Bengt’s grad school term project related to time-series analysis:
- Repeated measurements on respiratory problems of 7 dogs
- Fortran program for ML estimation with autoregressive and heteroscedastic residuals
Sources:
3 yearly measurements:

- Harvest index: Swedish grain harvest rated on a nine-point scale. Total crop failure scored 0; superabundant crop scored 9
- Fertility: Births per 1,000 female population
- Population rate: Birth rate - death rate
  - Birth rate: the total number of live births per 1,000 of a population in a year
  - Death rate: the number of deaths per 1,000 of the population in a year
Figure: Harvest (bottom) and Fertility (top)

Figure: Fertility (bottom) and Population Rate (top)
Irrespective of which party had gained control, or whether the King himself was on the throne, if the harvest was good, marriage and birth rates were high and death rates comparatively low, that is, the bulk of the of the population flourished.

On the contrary, when the harvest failed, marriage and birth rates declined and death devastated the land, bearing witness to need and privation and at times even to starvation. Whether the factories fared well or badly or whether the bank-rate rose or fell - all the things at this time, were scarcely more than ripples on the surface (Thomas, 1940: 82).
Theory for how Harvest Influences Fertility

Thomas (1940) cites a number of plausible mechanisms for this relationship.
First, in years following crop failure, marriage rates (and hence, fertility rates) drop.
Second, and more importantly, in years following a crop failure, young women who might otherwise bear children in Sweden are likely to emigrate (primarily to Finland and the United States during this period).
As a result of emigration, the average age of the female population rises dramatically in years following a crop failure and fertility drops accordingly.
How can you identify the a, b, c parameters from $N = 1$?
Rotating the Mediation Figure 90 Degrees Counter-Clock-Wise

- $f_t$
- $h_{t-1}$
- $p_t$
- $b$
- $c$
- $a$
Time-Series Mediation Model ($N = 1$):
Harvest, Fertility, and Population Rate, 1750 - 1850

\[ p_{t-1} \rightarrow p_t \rightarrow p_{t+1} \]
\[ f_{t-1} \rightarrow f_t \rightarrow f_{t+1} \]
\[ h_{t-2} \rightarrow h_{t-1} \rightarrow h_t \rightarrow \ldots \]
DATA:
  FILE = swedish_harvest_data.txt;
VARIABLE:
  NAME = year harvest fertil poprate;
  USEVARIABLES = harvest-poprate;
  MISSING = all (999);
LAGVARIABLES = harvest(1) fertil(1) poprate(1);
DEFINE:
  fertil = fertil/10;
ANALYSIS:
  ESTIMATOR = BAYES;
PROCESSORS = 2;
BITERATIONS = (10000);
MODEL:
  poprate ON fertil (b)
  harvest&1;
  fertil ON harvest&1 (a);
  \[ \text{!! auto-regressive part:} \]
  poprate ON poprate&1;
  fertil ON fertil&1;
  harvest ON harvest&1;
MODEL CONSTRAINT:
  NEW(indirect);
  indirect = a*b;
$N > 1$, Two-Level Time-Series Mediation Analysis: Random Effects, Temporal Order, Causality

Within (level-1) Variation across time
Between (level-2) Variation across individual

Within (level-1)
Variation across time
Between (level-2)
Variation across individual

Bengt Muthén

Methods for Intensive Longitudinal Data
Maxwell et al. (2007; 2011): Cross-sectional mediation analysis does not capture indirect/direct effects of longitudinal mediation processes.

Special 2011 issue of Multivariate Behavioral Research (editor S. West; comments by Reichardt, Shrout, Imai et al.)

What is missing? - Random effects: Variation across subjects in mean, variance, and auto-regression.
\[ N = 1: \text{Implications for ABAB designs} \]


- Susan Murphy: Just-In-Time Adaptive Interventions (JITAI) in which real-time, passively or actively collected, information on the patient (e.g., Ecological Momentary Assessments: EMA) is used to inform the real-time delivery of intervention options (e.g., recommendations, information and prompts)

- ARIMA impact assessment: transfer functions
$N = 1$ case versus $N > 1$ case

- Time-varying treatments (exposure), time-varying confounding
- Marginal structural models and inverse probability weighting:
  - Robins et al. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*
  - Vandecandeleare et al. (2016). Time-varying treatments in observational studies: Marginal structural models of the effects of early grade retention on math achievement. *Multivariate Behavioral Research*

- Granger causality (cross-lagged modeling)
For individual $i$ at the $j^{th}$ observation,

$$y_{ij} = \beta_0(t_{ij}) + \beta_1(t_{ij}) x_{ij} + \varepsilon_{ij},$$

where $\beta_0(t_{ij})$ and $\beta_1(t_{ij})$ are continuous functions of time using P-spline-based methods and varying the number of knots (Hastie & Tibshirani, 1990).

Based on the varying coefficient model of Hastie & Tibshirani (1993) in *Journal of the Royal Statistical Society, Series B*. 
Tan, Shiyko, Li, Li & Dierker (2012). A time varying effect model for intensive longitudinal data. *Psychological Methods*

- RCT of a smoking cessation intervention
- People asked to smoke on personal digital assistant (PDA) prompts - eliminating free will and increasing self-efficacy (confidence) for quitting
- Analysis of the regression of smoking abstinence self-efficacy ($y$) on momentary positive affect ($x$)
- Immediately prior to and following a quit attempt (QA)
- $N = 66$, $T = 25$ (1 – 117)
Analysis of the regression of smoking abstinence self-efficacy (y) on momentary positive affect (x) immediately prior to and following a quit attempt (QA)

- $N = 66$, $T = 25$ ($1 - 117$)
- Time-varying intercept and slope
A multilevel \((N > 1)\) time-series is analyzed but no time-series modeling features are included (individually-varying auto-regression coefficient \(\phi\), mean, and variance)

This is presumably difficult to do in combination with splines

An alternative for \(N = 1\) is presented in:


\[
Time - varying - AR(1) \ model : \ y_t = \beta_{0,t} + \beta_{1,t} y_{t-1} + \epsilon_t
\]

Bringman et al. (2016) discusses generalizations:

- Multivariate - challenging
- Multilevel - even more challenging
What’s Missing in Regular TVEM?
Dynamic Mediation Analysis

- Continuous-time modeling

Huang & Yuan (2016; online). Bayesian dynamic mediation analysis. *Psychological Methods*
  - Multilevel \((N > 1)\) and multivariate
  - Time-varying coefficients in a mediation model
  - Nonparametric penalized spline approach
  - Autoregressive modeling

What is missing here?
- Continuous-time modeling
CTMs applied in a mediation context would allow researchers to gain insight into how key effects vary as a function of lag.


- Voelkle & Oud (2013). Continuous-time modeling with individually-varying time intervals. *British Journal of Mathematical and Statistical Psychology*

- Voelkle & Oud (2015). Relating latent change score and continuous time models. *Structural Equation Modeling*

Continuous latent variables:
- Multilevel \((N > 1)\) cross-lagged analysis with factors measured by multiple indicators
- Cross-classified factor model (time crossed with subject)

Categorical latent variables:
- Transition modeling (Hidden Markov, regime switching, time-series LTA) with latent class variables
Multilevel Time-Series/Dynamic Factor Analysis

Bengt Muthén
Methods for Intensive Longitudinal Data

\[
\begin{align*}
\text{Within} & \quad \phi \\
\text{Between} & \quad f_{t-1} \quad f_t \\
\end{align*}
\]
**Time-Series Factor Analysis**

- **N = 1 methods (single-level modeling):**

- **N > 1 methods - two-level modeling with random effects:**
  - *Mplus Version 8*
Example: Affective Instability
In Ecological Momentary Assessment

- An example of the growing amount of EMA data
- 84 outpatient subjects: 46 meeting borderline personality disorder (BPD) and 38 meeting MDD or DYS
- Each individual is measured several times a day for 4 weeks for a total of about 100 assessments
- A mood factor for each individual is measured with 21 self-rated continuous items
- The research question is if the BPD group demonstrates more temporal negative mood instability than the MDD/DYS group
3-Factor EFA/CFA DAFS

- EFA suggests 3 factors (although time-series ESEM is needed):
  - Angry: Upset, Distressed, Angry, Irritable
  - Sad: Downhearted, Sad, Blue, Lonely
  - Afraid: Afraid, Frightened, Scared

- 3-factor EFA-within-CFA DAFS factor autocorrelation:
  - 0.536 (Angry), 0.578 (Sad), 0.623 (Afraid)
  - To which you could add random effects for the factor autocorrelations to see if they have different variability across subjects
How to standardize the coefficients?

Schuurman et al. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. Forthcoming in *Psychological Methods*
ANALYSIS:
TYPE = TWOLEVEL RANDOM;
ESTIMATOR = BAYES; PROCESSORS = 2; THIN = 5;
BITERATIONS = (10000);

MODEL:
%WITHIN%
  f1 BY jittery-scornful*0 (&1);
  f2 BY jittery-scornful*0 (&1);
  f3 BY jittery-scornful*0 (&1);
  f1 BY downhearted*1 sad*1 blue*1 lonely*1;
  f1 BY angry@0 afraid@0;
  f2 BY angry*1 irritable*1 hostile*1 scornful*1;
  f2 BY downhearted@0 afraid@0;
  f3 BY afraid*1 frightened*1 scared*1;
  f3 BY downhearted@0 angry@0;
  f1-f3@1;
  s1 | f1 ON f1&1;
  s2 | f2 ON f2&1;
  s3 | f3 ON f3&1;
  f1 ON f2&1 f3&1;
  f2 ON f1&1 f3&1;
  f3 ON f1&1 f2&1;
%BETWEEN%
  fb BY jittery-scornful*;
  fb s1-s3 ON group; fb@1;
Cross-classified time-series factor analysis

Cross-classification of time and subject gives flexible modeling with random intercepts and factor loadings:


- **Mplus Version 8**: Expanded to time-series version
  - Auto-regression for factor and factor indicators
  - Random factor loadings varying across subjects and random intercepts for measurement and factor intercepts varying across time
  - Can variation across time in random slopes be used to study trends?

Figure : Intercept

Figure : Slope
Latent Transition Analysis (LTA; Hidden Markov model; HMM):

\[ c_{t-1} \rightarrow na_{t-1} \rightarrow \rightarrow pa_{t-1} \rightarrow na_{t} \rightarrow \rightarrow pa_{t} \rightarrow \rightarrow c_{t} \]

LTA with autoregression (Markov switching autoreg. model; MSAR):

\[ c_{t-1} \rightarrow na_{t-1} \rightarrow \rightarrow na_{t} \rightarrow \rightarrow c_{t} \]

\[ c_{t-1} \rightarrow pa_{t-1} \rightarrow \rightarrow pa_{t} \rightarrow \rightarrow c_{t} \]

- Hamaker et al. (2016). Modeling BAS dysregulation in bipolar disorder: Illustrating the potential of time series analysis. *Assessment*
What’s the answer?
- All the above and more

For this we need the input from many researchers
To Learn More About Time-Series Analysis

- Classic but only $N = 1$:
  - Less hard: Shumway & Stoffer (2011)

- Accessible, applied writings with $N > 1$:
  - Chow et al. (2010). Equivalence and differences between SEM and state-space modeling. *Structural Equation Modeling*
  - More in the pipeline
Epilogue: The Transitional Aspect of Snow: Temperature and Snow Depth
- Implications for Sledding
“Foreshadowing or guessing ahead is a literary device by which an author hints what is to come. Foreshadowing is a dramatic device in which an important plot-point is mentioned early in the story and will return in a more significant way. It is used to avoid disappointment. It is also sometimes used to arouse the reader”.