Philipps Universität Marburg

GIMME, GIMME more (network models): Multiverse analysis for dynamic network models

Björn Siepe[1] **Daniel W. Heck**[1]

May 26, 2025 – Society for Ambulatory Assessment

[1]Psychological Methods Lab, Department of Psychology, University of Marburg

Analytical Flexibility in Time Series Modeling

- Time series data often require **complex models**

Analytical Flexibility in Time Series Modeling

- Time series data often require **complex models**
- Many analytical choices: Model selection, preprocessing, model specification, interpretation...

Analytical Flexibility in Time Series Modeling

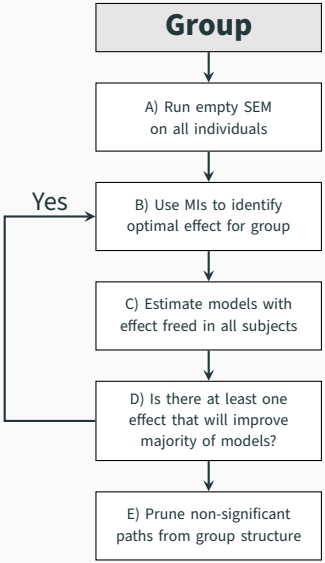
- Time series data often require **complex models**
- Many analytical choices: Model selection, preprocessing, model specification, interpretation...
- Often, only one set of choices is conducted and reported

Analytical Flexibility in Time Series Modeling

- Time series data often require **complex models**
- Many analytical choices: Model selection, preprocessing, model specification, interpretation...
- Often, only one set of choices is conducted and reported
- Robustness to **arbitrary choices** is underappreciated

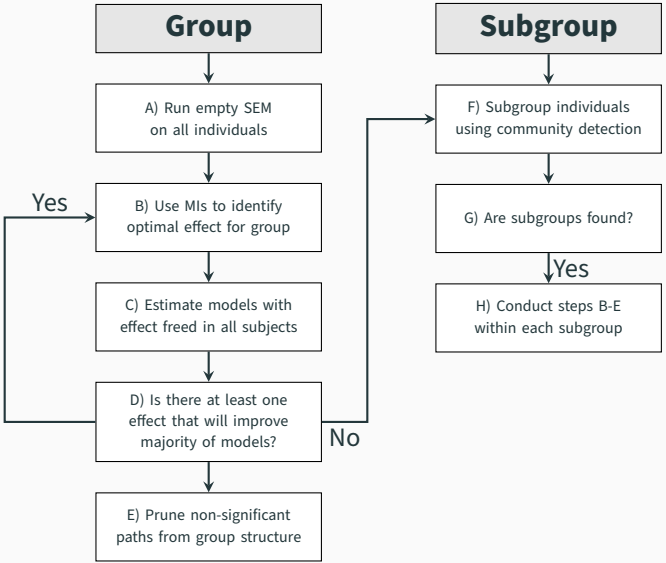
GIMME

GIMME



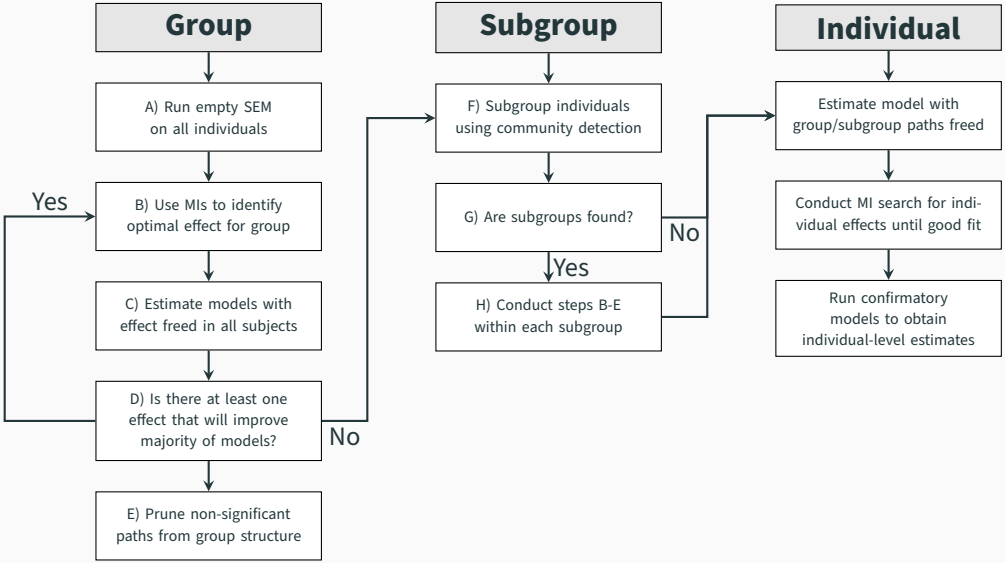
Adapted from Gates et al., 2017

GIMME



Adapted from Gates et al., 2017

GIMME



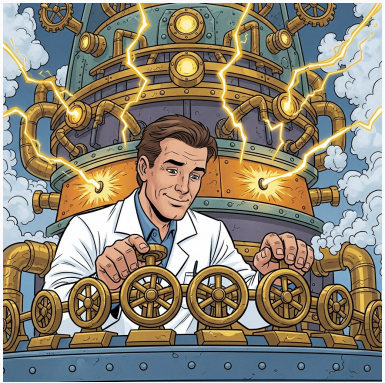
Adapted from Gates et al., 2017

This Study

- Investigate the impact of arbitrary alternative modeling decisions on previously published results
- Previous multiverse studies: Focus on measurement (Dejonckheere et al., 2018) or preprocessing (Weermeijer et al., 2022)

This Study

- Investigate the impact of arbitrary alternative modeling decisions on previously published results
- Previous multiverse studies: Focus on measurement (Dejonckheere et al., 2018) or preprocessing (Weermeijer et al., 2022)



Created with Gemini 2.5 Pro

Datasets

Datasets

Personality Dataset (Wright et al., 2019):

- $n = 94$ participants with personality disorder diagnosis, average $t = 91.48$ (1x/day)
- Items: Affect, interpersonal behavior, stress, and functioning (all sum scores), and daily functioning (5-point Likert)

Datasets

Personality Dataset (Wright et al., 2019):

- $n = 94$ participants with personality disorder diagnosis, average $t = 91.48$ (1x/day)
- Items: Affect, interpersonal behavior, stress, and functioning (all sum scores), and daily functioning (5-point Likert)

Emotion Dataset (Kullar et al., 2024):

- $n = 105$, average $t = 62.31$ ($SD = 8.11$)
- Items: Nine momentary emotions (Likert 7-point)

GIMME Variations

1. Group threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$
2. Subgroup threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$

GIMME Variations

1. Group threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$
2. Subgroup threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$

Five parameters refer to the fit indices used for model selection:

3. RMSEA cutoff $\in \{.03, .\mathbf{05}, .08\}$
4. SRMR cutoff $\in \{.03, .\mathbf{05}, .08\}$
5. NNFI cutoff $\in \{.90, .\mathbf{95}, .97\}$
6. CFI cutoff $\in \{.90, .\mathbf{95}, .97\}$
7. Fit measures satisfying the cutoffs $\in \{1, \mathbf{2}, 3\}$

GIMME Variations

1. Group threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$
2. Subgroup threshold $\in \{50\%, 60\%, \mathbf{75}\%, 80\%\}$

Five parameters refer to the fit indices used for model selection:

3. RMSEA cutoff $\in \{.03, .\mathbf{05}, .08\}$
4. SRMR cutoff $\in \{.03, .\mathbf{05}, .08\}$
5. NNFI cutoff $\in \{.90, .\mathbf{95}, .97\}$
6. CFI cutoff $\in \{.90, .\mathbf{95}, .97\}$
7. Fit measures satisfying the cutoffs $\in \{1, \mathbf{2}, 3\}$

Conducted a small simulation study showing the arbitrariness of these choices

Personality Results

- **Group**:

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups
- **Individual**

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups
- **Individual**
  - On average, $\sim$ 2 paths different in presence/absence from reference fit

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups
- **Individual**
  - On average, $\sim$ 2 paths different in presence/absence from reference fit
  - For those different effects: Absolute average difference of $\sim$ 0.13

Personality Results

- **Group**:
  - $\sim$ 92% of models with same group path
- **Subgroup**:
  - Same subgroups
- **Individual**
  - On average, $\sim$ 2 paths different in presence/absence from reference fit
  - For those different effects: Absolute average difference of $\sim$ 0.13
  - For 12 individuals, the most central node was identical to the reference model in less than one-third of all specifications

Emotion Results

- **Group**:

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit
  - Subgroup threshold of 50% instead of 51%: Change of 128 paths

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit
  - Subgroup threshold of 50% instead of 51%: Change of 128 paths
- **Individual**

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit
  - Subgroup threshold of 50% instead of 51%: Change of 128 paths
- **Individual**
  - On average, $\sim$ 9 paths different in absence/presence from reference fit

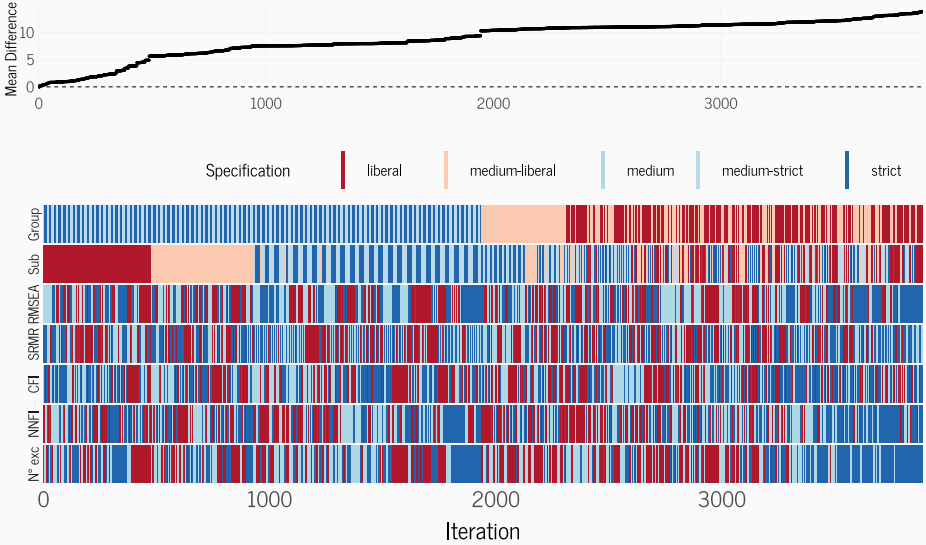
Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit
  - Subgroup threshold of 50% instead of 51%: Change of 128 paths
- **Individual**
  - On average, $\sim$ 9 paths different in absence/presence from reference fit
  - For those different effects: Absolute average difference of 0.2

Emotion Results

- **Group**:
  - In $\sim$ 50% specifications, four or five group effects (instead of one) were estimated
- **Subgroup**:
  - $\sim$ 50% of subgroup solutions very dissimilar to the reference fit
  - Subgroup threshold of 50% instead of 51%: Change of 128 paths
- **Individual**
  - On average, $\sim$ 9 paths different in absence/presence from reference fit
  - For those different effects: Absolute average difference of 0.2
  - For 30 individuals, the most central node was identical to the reference model in less than one-third of all specifications

Emotion Results



Shiny App

Specification Curve Analysis

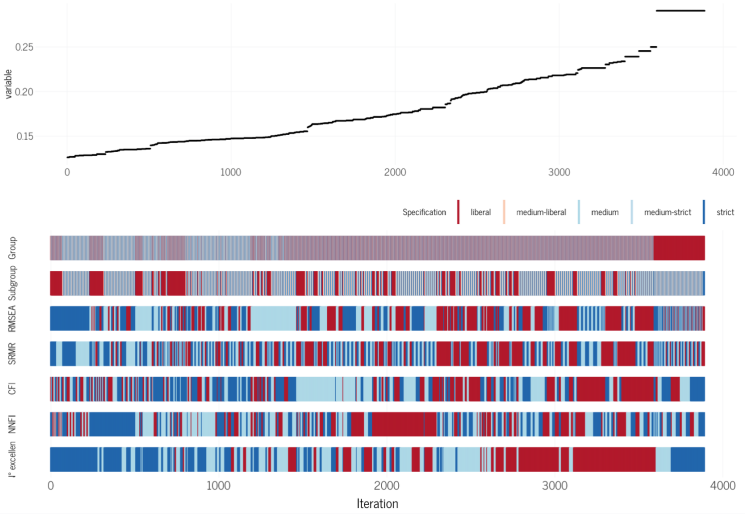
A Specification Curve Analysis (SCA) plot is a visual tool used to analyze the effects of a variable across different specifications.

Select a variable from the dropdown menu to explore its impact as ordered by its size across specifications.

Rendering of the plot can take some time. This plot is not interactive to speed up computation time.

Select Column:

Homogeneity



Shiny App for the paper "Network Multiverse" (Siepe & Heck, 2023). Find the source code on GitHub.

Takeaways

Applied

- Results largely robust for Personality dataset, less robust for Emotion dataset

Takeaways

Applied

- Results largely robust for Personality dataset, less robust for Emotion dataset
- Different algorithmic specifications can affect main conclusions, especially at individual level

Takeaways

Applied

- Results largely robust for Personality dataset, less robust for Emotion dataset
- Different algorithmic specifications can affect main conclusions, especially at individual level

Takeaways

Applied

- Results largely robust for Personality dataset, less robust for Emotion dataset
- Different algorithmic specifications can affect main conclusions, especially at individual level

Methodological

- Multiverse analyses have focused on preprocessing, but algorithmic decisions are also important

Takeaways

Applied

- Results largely robust for Personality dataset, less robust for Emotion dataset
- Different algorithmic specifications can affect main conclusions, especially at individual level

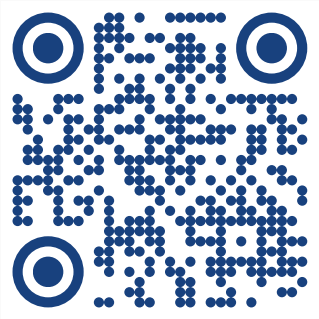
Methodological

- Multiverse analyses have focused on preprocessing, but algorithmic decisions are also important
- Relationship between simulation studies and multiverse analyses

Get In Touch

- @ bjoern.siepe@uni-marburg.de
- BlueSky: bsiepe
- https://bsiepe.github.io/

Paper & Slides



References i

Dejonckheere, E., Mestdagh, M., Houben, M., Erbas, Y., Pe, M., Koval, P., Brose, A., Bastian, B., & Kuppens, P. (2018).**The bipolarity of affect and depressive symptoms.** *Journal of Personality and Social Psychology*, *114*(2), 323–341.

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guiskewicz, K. (2017).**Unsupervised classification during time-series model building.** *Multivariate Behavioral Research*, *52*(2), 129–148. https://doi.org/10.1080/00273171.2016.1256187

References ii

Kullar, M., Carter, S., Hitchcock, C., Whittaker, S., Wright, A. G. C., & Dalgleish, T. (2024).**Patterns of emotion-network dynamics are orthogonal to mood disorder status: An experience sampling investigation..** *Emotion*, *Advance online publication*. https://doi.org/10.1037/emo0001245

Weermeijer, J., Lafit, G., Kiekens, G., Wampers, M., Eisele, G., Kasanova, Z., Vaessen, T., Kuppens, P., & Myin-Germeys, I. (2022).**Applying multiverse analysis to experience sampling data: Investigating whether preprocessing choices affect robustness of conclusions.** *Behavior Research Methods*, *54*(6), 2981–2992. https://doi.org/10.3758/s13428-021-01777-1

References iii

Wright, A. G. C., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019).**Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology.** *Psychological Assessment*, *31*(4), 502–515. https://doi.org/10.1037/pas0000617

Image Source

Britney Spears Cover:
https:
//lastfm.freetls.fastly.net/i/u/ar0/dff1377bbbf94c97cc8769d0ea072804.jpg