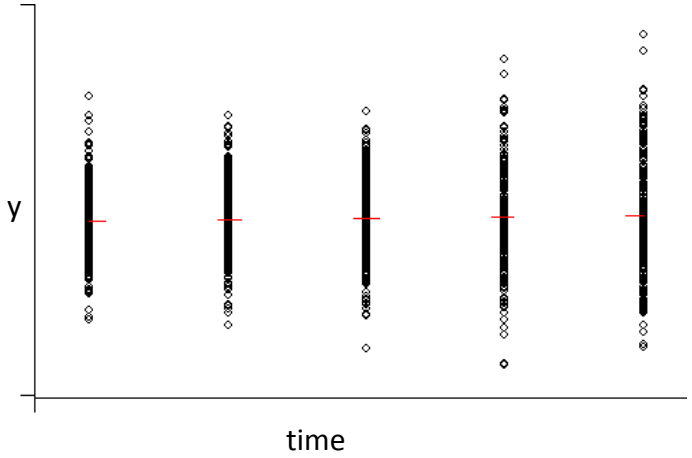
Fitting decision trees to multilevel and longitudinal data

Marjolein Fokkema, PhD
Maaike Jorink, MSc
Leiden University

Recursive partitioning of linear models

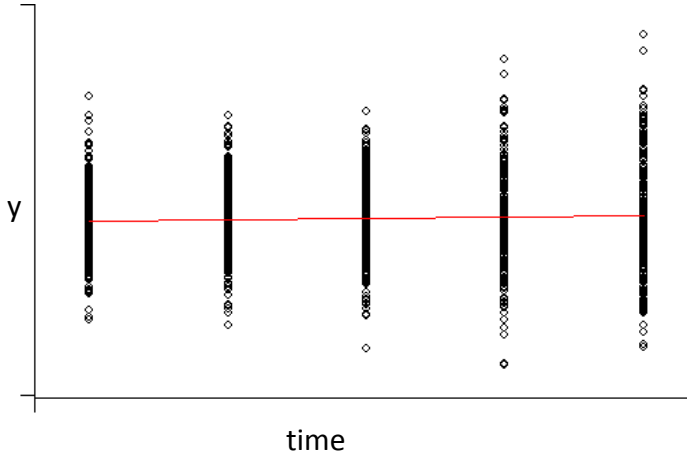
(very basic) LM: $\hat{y}_i = x_i^T \beta$

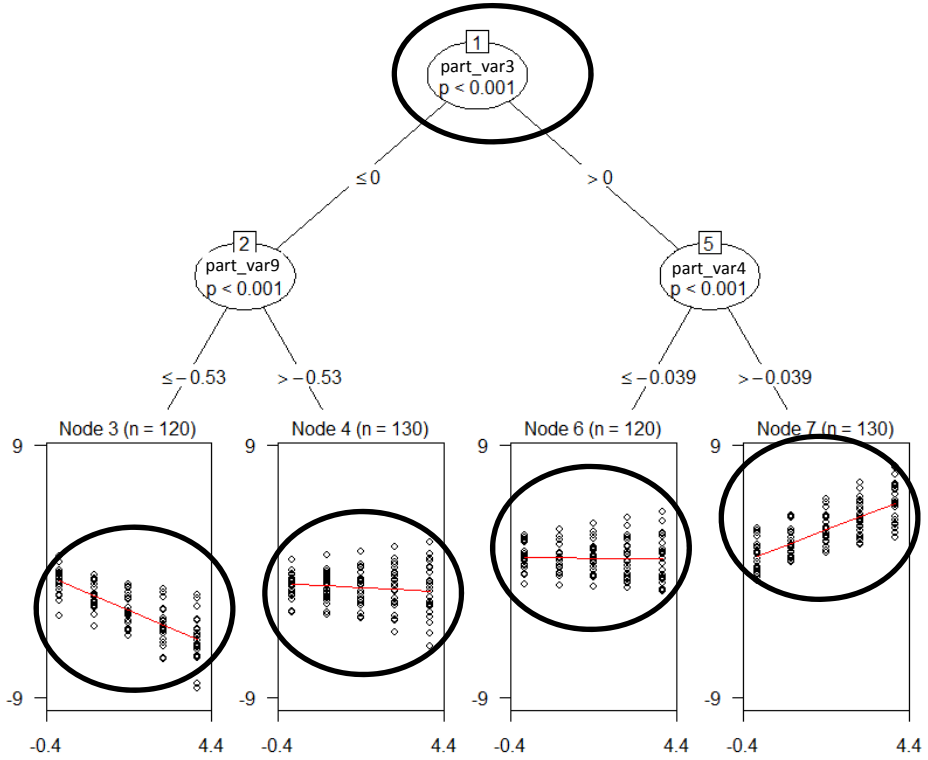


Recursive partitioning of linear models

LM: $\hat{y}_i = x_i^T \beta$

LM tree: $\hat{y}_i = x_i^T \beta_j$ (Zeileis, Hothorn & Hornik, 2006)



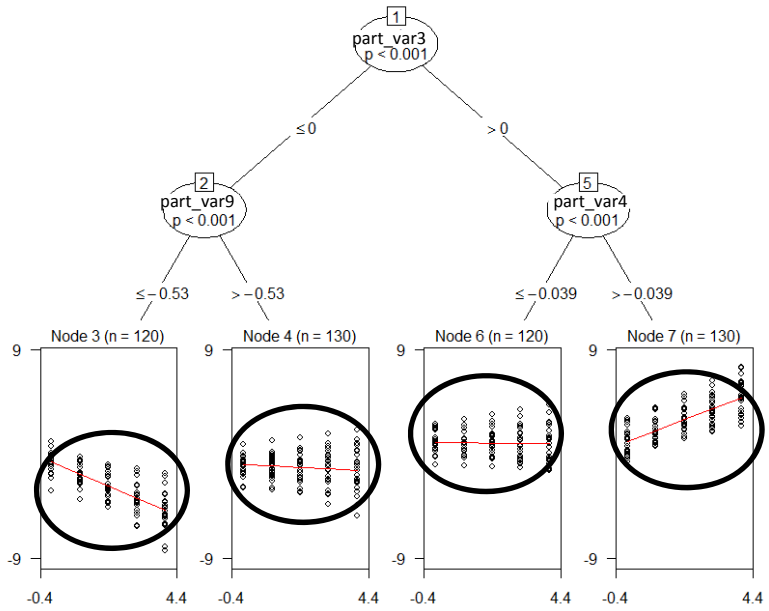


Recursive partitioning of linear models

LM: $\hat{y}_i = x_i^T \beta$

LM tree: $\hat{y}_i = x_i^T \beta_j$ (Zeileis, Hothorn & Hornik, 2006)

LMM tree: $\hat{y}_i = x_i^T \beta_j + z_i^T b$ (Fokkema et al., 2018)



LMM trees

Estimate LMM tree: $\hat{y}_i = x_i^T \beta_j + z_i^T b$

0) Initialize estimation assuming σ_b = 0

1) LM tree, given current random effects predictions

2) Random effects, given current LM tree

3) Iterate between steps 1 and 2 until convergence

Fokkema et al. (2018): Works well in clustered, non-longitudinal data:

- σ_b not very large
- partitioning variables measured on observation level

How about growth curve models?

Recursive partitioning methods for longitudinal data

Method (R package)	Support for partitioning of LGCMs	Type of mixed-effects models	Estimation of random effects parameters	Possible levels of partitioning variables
REEMtree (REEMtree)	No	-	Global	Observation and/or cluster
LMM tree (glmertree)	Yes	Linear	Global	Observation and/or cluster
SEM tree (semtree)	Yes	Linear	Local	Cluster
longRPart (longRpart)	Yes	Linear	Local	Cluster
longRPart2 (longRpart2)	Yes	Linear and non-linear	Local	Cluster

Recursive partitioning methods for longitudinal data

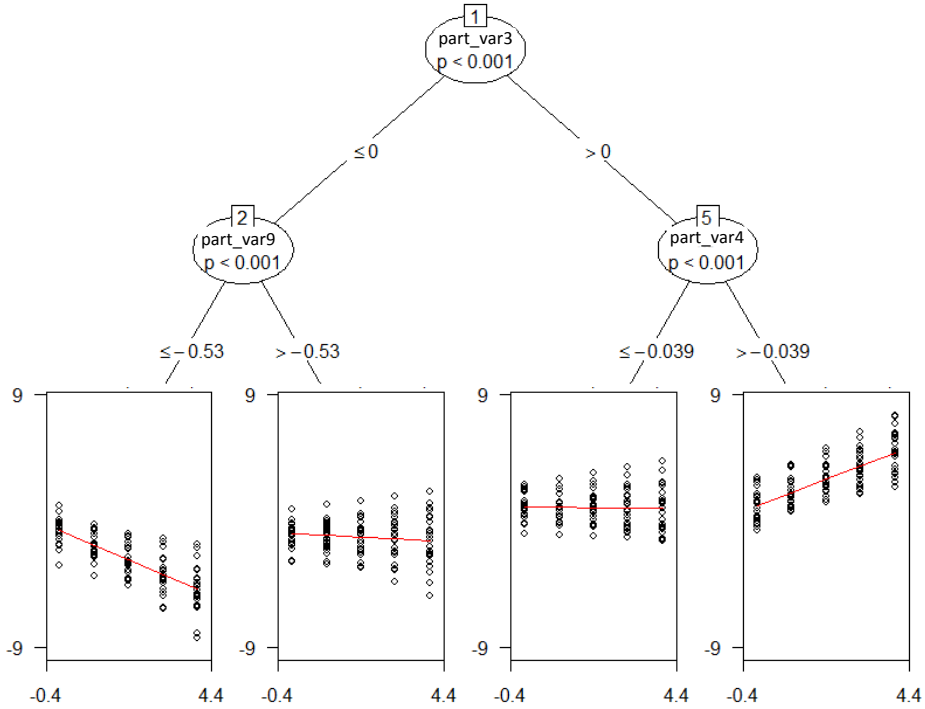
Method (R package)	Support for partitioning of LGCMs	Type of mixed-effects models	Estimation of random effects parameters	Possible levels of partitioning variables
REEMtree (REEMtree)	No	-	Global	Observation and/or cluster
LMM tree (glmertree)	Yes	Linear	Global	Observation and/or cluster
SEM tree (semtree)	Yes	Linear	Local	Cluster
longRPart (longRpart)	Yes	Linear	Local	Cluster
longRPart2 (longRpart2)	Yes	Linear and non-linear	Local	Cluster

Research questions

When partitioning linear GCMs:

1. Better to estimate random effects locally (SEM trees) or globally (LMM trees)?
2. How to best estimate LMM trees:
   - Initialize with tree or random effects?
   - Variable selection tests at observation or cluster level?
3. How complex should random-effects specification be?
   - Random intercept only, or random intercept + slope?

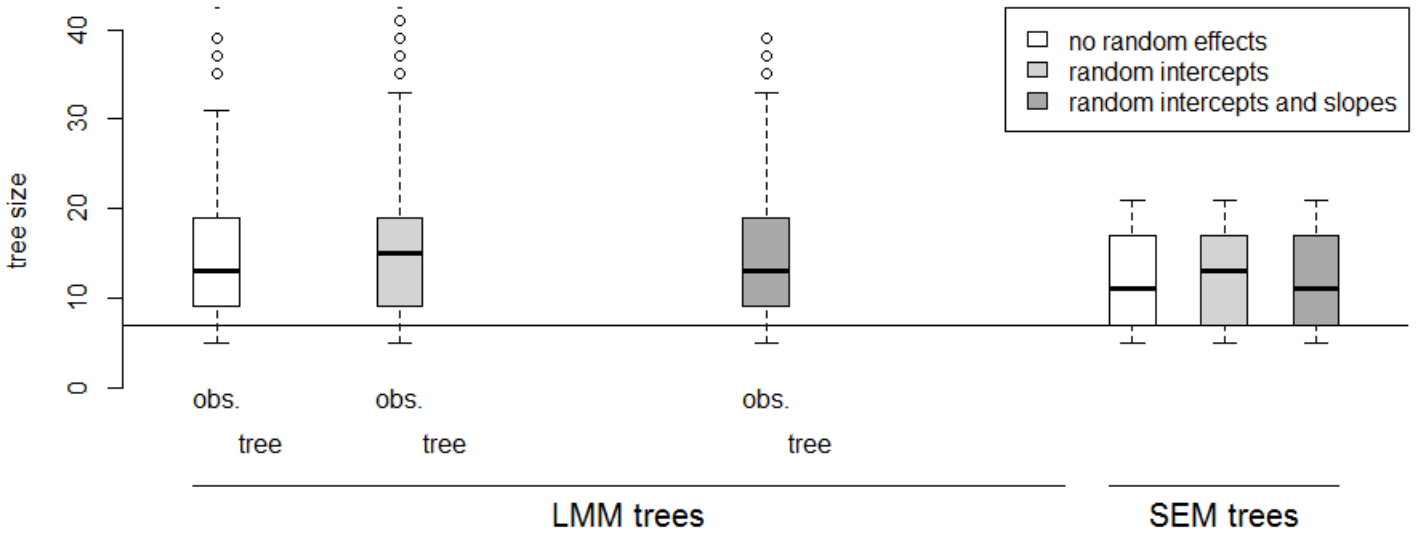
Simulation: Design



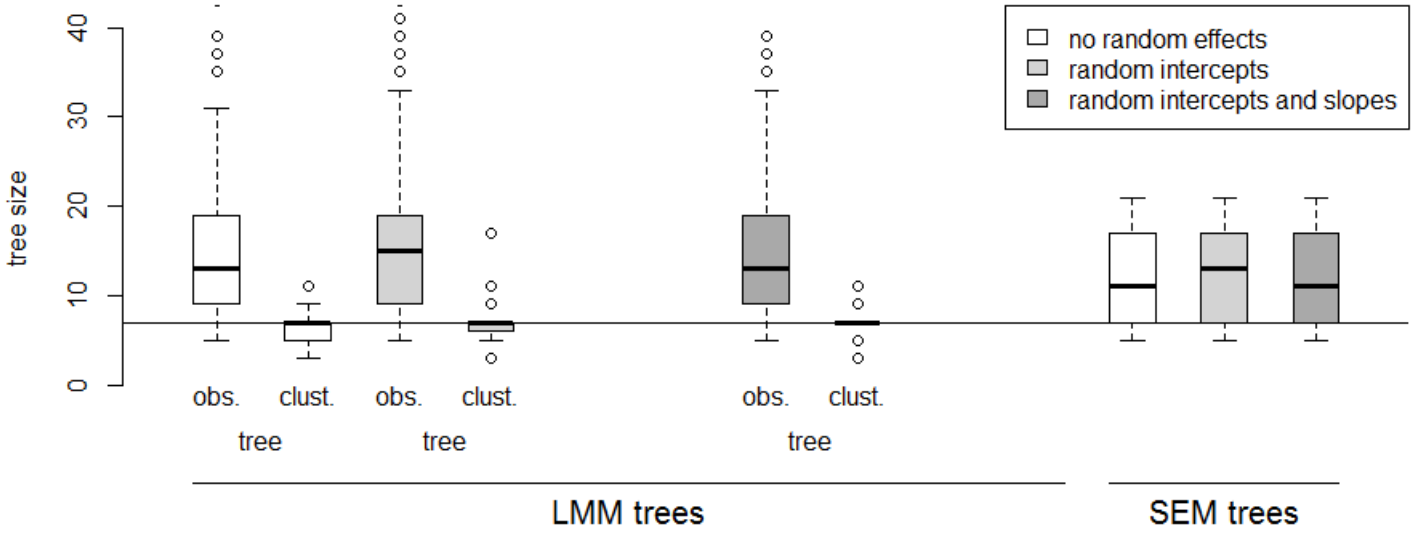
Parameters varied:

- # of subjects (clusters)
- # of noise variables
- variance of random effects
- …

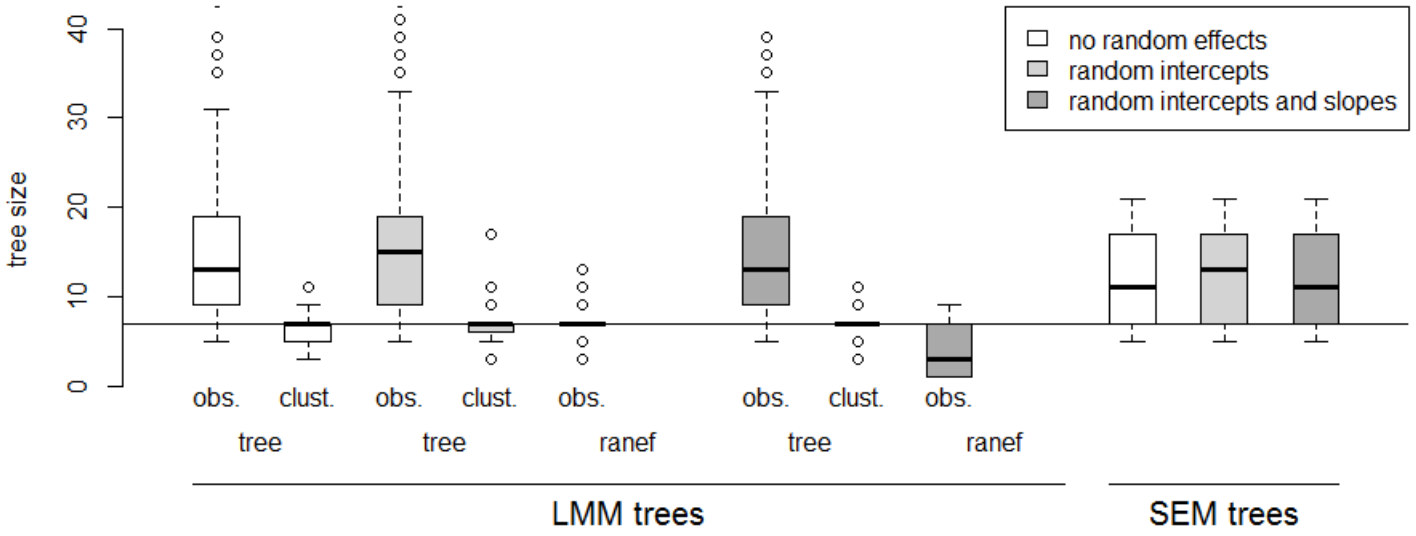
Simulation: Results



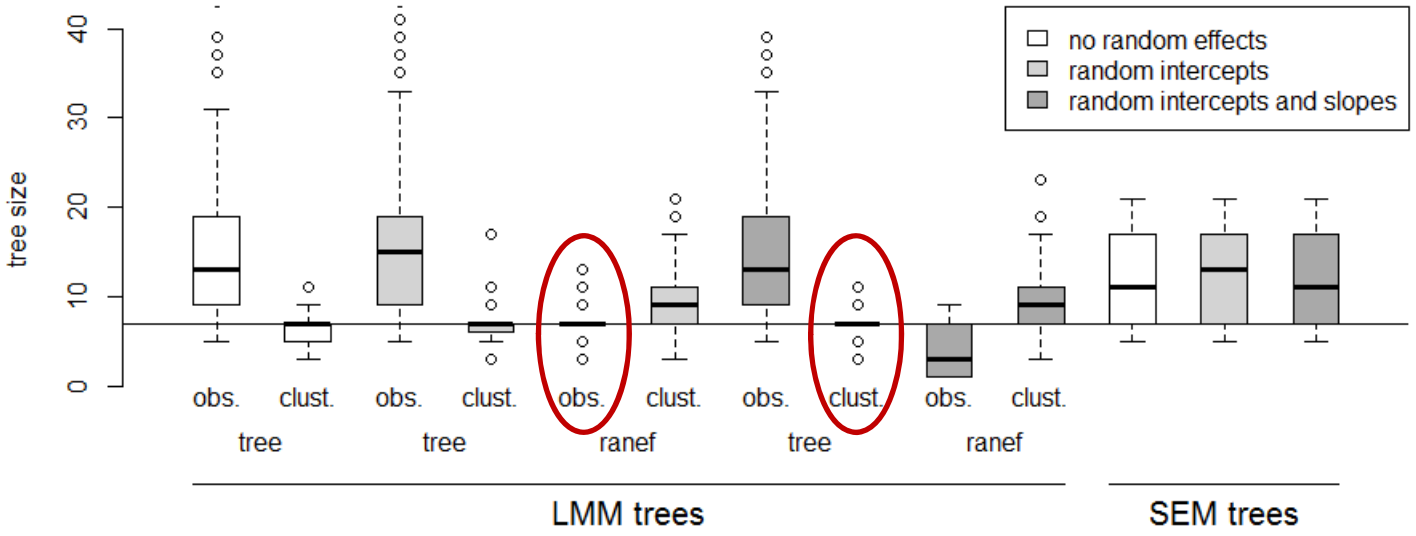
Simulation: Results



Simulation: Results



Simulation: Results



Application:
Early Childhood Longitudinal Study

Repeated assessments of reading (N ≈ 6,500; ages 5 – 12)

Response variable: theta scores (IRT)

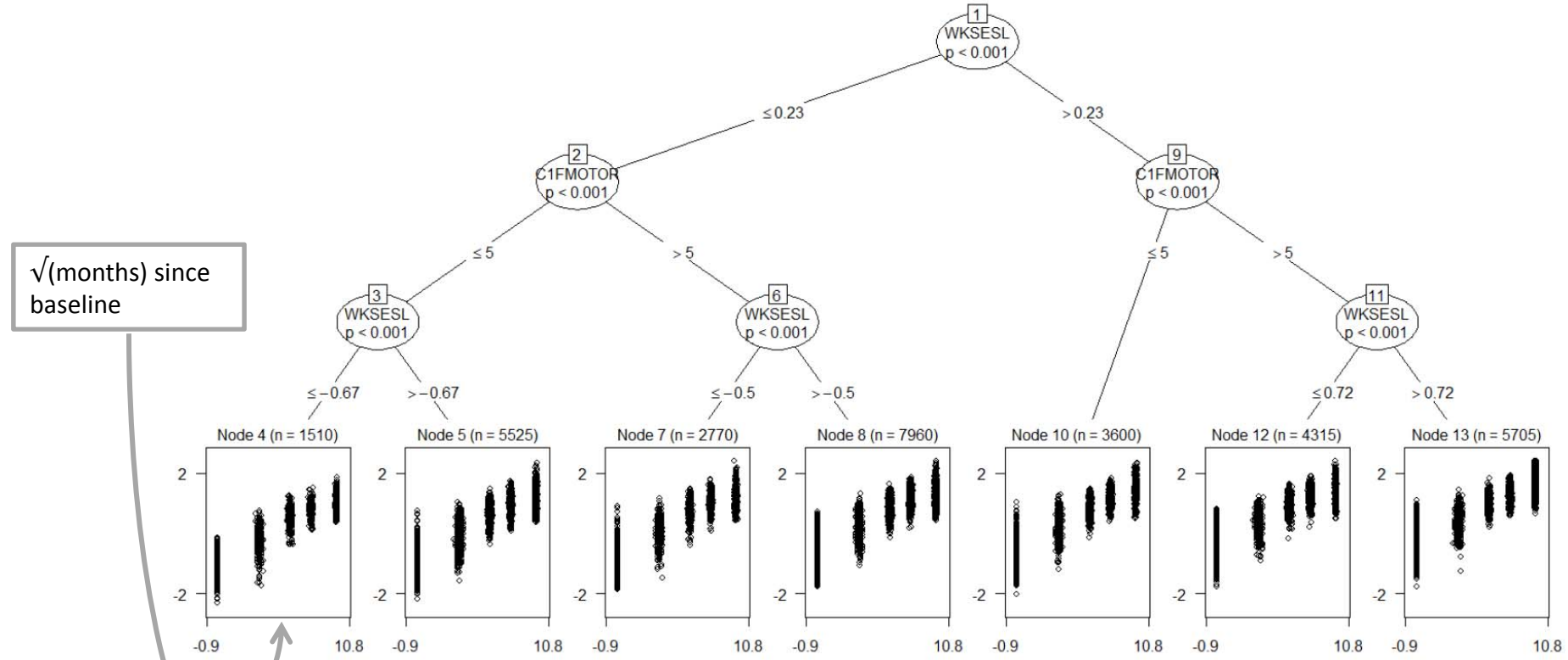
8 potential partitioning variables:

- gender, SES, motor skills, psycho-social functioning…

Current aim: Compare tree size and prediction error (10 fold CV) for different LMM tree specifications

Application: Results

First splits very similar between different settings:



Fixed intercept	-1.00	-0.78	-0.66	-0.51	-0.56	-0.41	-0.32
Fixed slope	0.22	0.22	0.21	0.21	0.22	0.21	0.21

Application: Results

Model	Random intercept only		Random intercept + slope	
	MSE (SE)	Mean no. of nodes (SD)	MSE (SE)	Mean no. of nodes (SD)
Observation-level tests				
Tree initialization	.157 (.002)	304.4 (31.71)	.164 (.002)	333.4 (27.93)
Random-effects initialization	.157 (.002)	304.4 (31.71)	.117 (.002)	19.6 (2.12)
Cluster-level tests				
Tree initialization	.145 (.002)	104.2 (18.38)	.149 (.002)	107.4 (16.38)
Random-effects initialization	.145 (.002)	104.2 (18.38)	.150 (.002)	245.8 (23.04)

Preliminary conclusion

LMM trees can perform well in recovering subgroups in GCMs

- But should use **either** cluster-level tests or random-effects initialization (not both)
- Which is best?
  - Different conclusions between simulation study and real data
  - Likely depends on:
    - ICC / variance of random effects in data
    - Complexity of random-effects specification
    - …. ?

Thank you!

References

Brandmaier, A.M., von Oertzen, T., McArdle, J.J., & Lindenberger, U. (2013). Structural equation model trees. *Psychological Methods*, *18*(1), 71.

R package: https://CRAN.R-project.org/package=semtree

Fokkema, M., Smits, N., Zeileis, A., Hothorn, T. & Kelderman, H. (2018). Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. *Behavior Research Methods 50*(5), 2016-2034. DOI: 10.3758/s13428-017-0971-x.

R package: https://CRAN.R-project.org/package=glmertree

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651-674.

R package: https://CRAN.R-project.org/package=partykit