

Conditional non-parametric bootstrap for non-linear mixed effect models

Christelle Rodrigues¹, Vincent Jullien¹, Emmanuelle Comets^{2,3}★

¹ INSERM UMR1129, Paris, France ² INSERM, IAME, UMR 1137; University Paris Diderot, F-75018, Paris, France ³INSERM, CIC 1414; Univ Rennes-1, 35700 Rennes, France

Objective: propose and evaluate a new residual bootstrap using samples from the individual conditional distributions.

Introduction

- Estimation errors in non-linear mixed effect models (NLMEM)
 - generally obtained using the inverse of the Fisher information matrix
 - asymptotic approach valid with large number of subjects and of observations per subject
 - alternative approaches: likelihood profiling, bootstrap, SIR
- Bootstrap methods in NLMEM need to consider the hierarchical nature of the data [1]
- Evaluated in these models with contrasting results [2]
 - sparse and rich designs
 - good performance of case bootstrap
 - stratification issues with heterogenous designs
 - good performance of parametric bootstrap
 - in settings where the distribution of the residuals is correct
 - non-parametric bootstrap
 - requires corrections for shrinkage in the individual residuals
- Proposal: Conditional non-parametric bootstrap
 - samples from the conditional distribution are used to correct shrinkage in diagnostics [3]
 - idea: build bootstrap samples using residuals sampled from the individual conditional distributions

Methods

Statistical models

Model for observation y_{ij}

$$y_{ij} = f(\Psi_i, x_{ij}) + g(\Psi_i, \gamma, x_{ij})\epsilon_{ij}$$

- subject i ($i = 1, \dots, N$), with n_i observations $\mathbf{y}_i = \{y_{i1}, \dots, y_{in_i}\}$ at times t_{ij} , and covariates $\mathbf{z}_i$
- f : structural model (analytical expression, see example)
- $g = g(a, b, c)$: residual error model (one of: constant, proportional, combined, exponential) with parameters $\gamma = \{a, b, c\}$
- individual parameters Ψ_i
 - modelled parametrically as a function $\Psi_i = h(\phi_i, \mathbf{z}_i) = h(\mu(\mathbf{z}_i), \eta_i)$ of fixed effects μ and random effects η_i ($\eta_i \sim \mathcal{N}(0, \Omega)$)
 - in saemix, h can be the identity function (normal distribution for Ψ), the exponential function (log-normal distribution for Ψ), or the logit or probit transformations

Bootstrap methods

- Case bootstrap (CaseBoot)**: resample complete individuals from the dataset
- Parametric residual bootstrap (PBoot)**: resample within a theoretical distribution
- Non-parametric residual bootstrap (NPBoot)**: resample empirical residuals after centering and scaling

- Conditional non-parametric residual bootstrap (cNPBoot)**: resample residuals from individual conditional distributions (Figure 1)

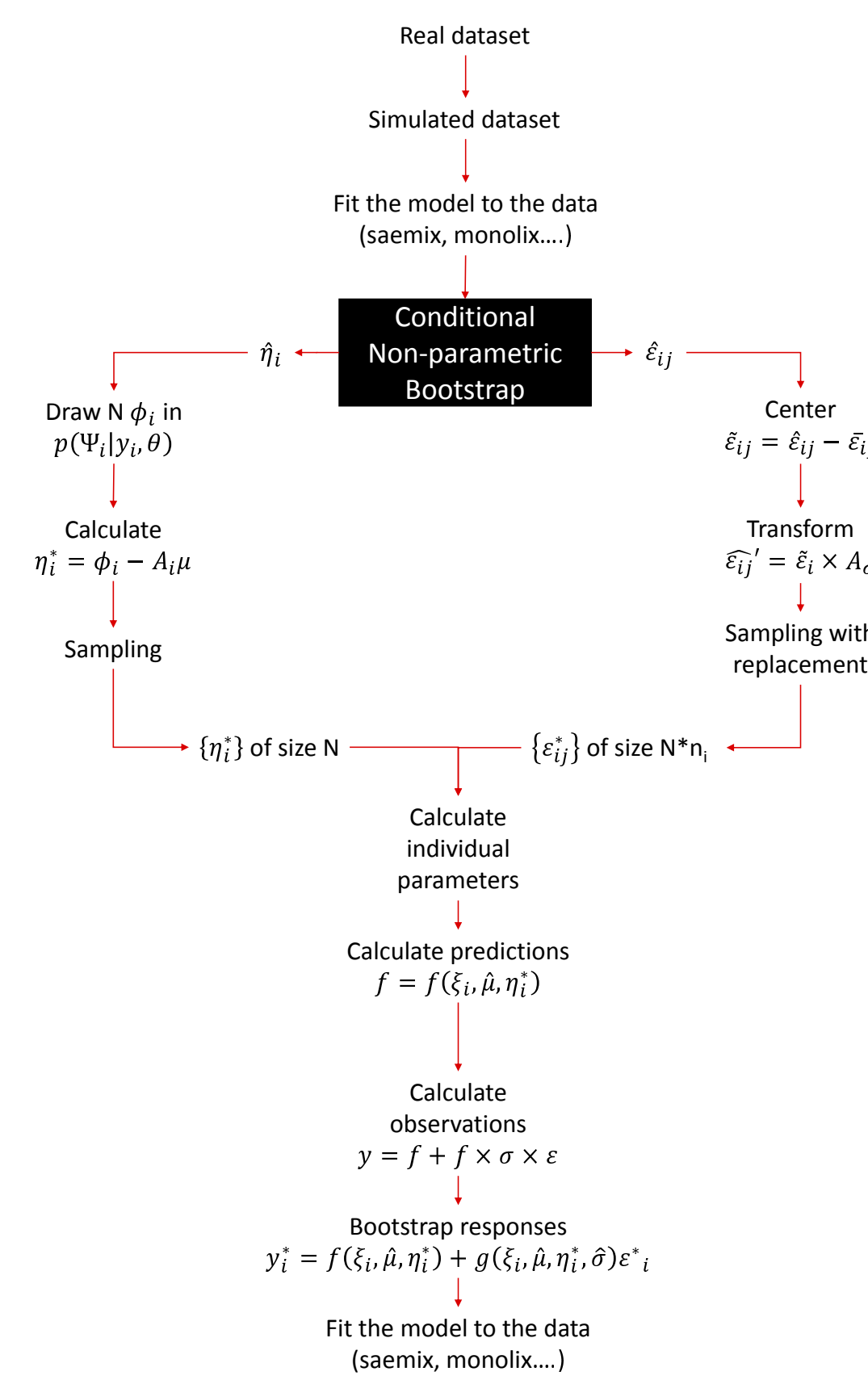


Figure 1: Algorithm for *Conditional non-parametric bootstrap*

Evaluation of conditional non-parametric bootstrap

Objective Compare the performance of conditional non-parametric bootstrap with other bootstrap approaches

Simulation study:

Models:

- structural model: one-compartment model with first-order absorption
- IIV: exponential model
- residual error model: proportional error
- parameter values from a PK analysis of valproic acid pharmacokinetics in epileptic children [4]

Parameter	Value	Parameter	Value (%)
k_a (h^{-1})	0.15	$\omega_{k_a}^2$	0
V (L)	5.66	ω_V^2	30
CL (L/h)	0.37	ω_{CL}^2	30
σ (-)	0.2		

Table 1: Parameter values used in the simulation.

Simulated datasets:

- for each scenario, 100 simulated datasets with $N=100$ subjects
 - 200 bootstrap samples for each bootstrap
 - 100 samples from the conditional distribution for each subject for **cNPBoot**
- two designs:
 - dose regimen: single dose and repeated dosing
 - rich design: 6 sampling times, which were obtained by optimal design using PFIM [5].

Evaluation of bootstrap methods:

- empirical SE obtained from K simulated datasets giving the "true" value observed across the simulations

$$SE_{\text{empirical}}(\hat{\theta}) = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (\hat{\theta}_k - \theta_0)^2} \quad (1)$$

- relative bias and RMSE: compared to the parameters used for the simulation

$$RBias(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{\theta}_b - \theta_0}{\theta_0} \times 100 \quad (2)$$

$$RBias(SE(\hat{\theta})) = \frac{1}{B} \sum_{b=1}^B \frac{SE_B(\hat{\theta}) - SE_{\text{empirical}}(\hat{\theta})}{SE_{\text{empirical}}(\hat{\theta})} \times 100 \quad (3)$$

- coverage rate of the 95% bootstrap CI for each estimated parameter: probability that CI contains the true value of the parameter
- compare with the asymptotic SE given by the software and obtained as the inverse of the Fisher information matrix
- good bootstrap: defined as a method without bias for parameter estimates and their corresponding SE (less than 5%), and ensuring a good coverage rate of the 95% CI (90% to 100%)

Simulations performed using the saemix library in R [6] and compared to the results obtained with the Rsmix library using the MixConnectors API [7].

Results

Single dose scenario

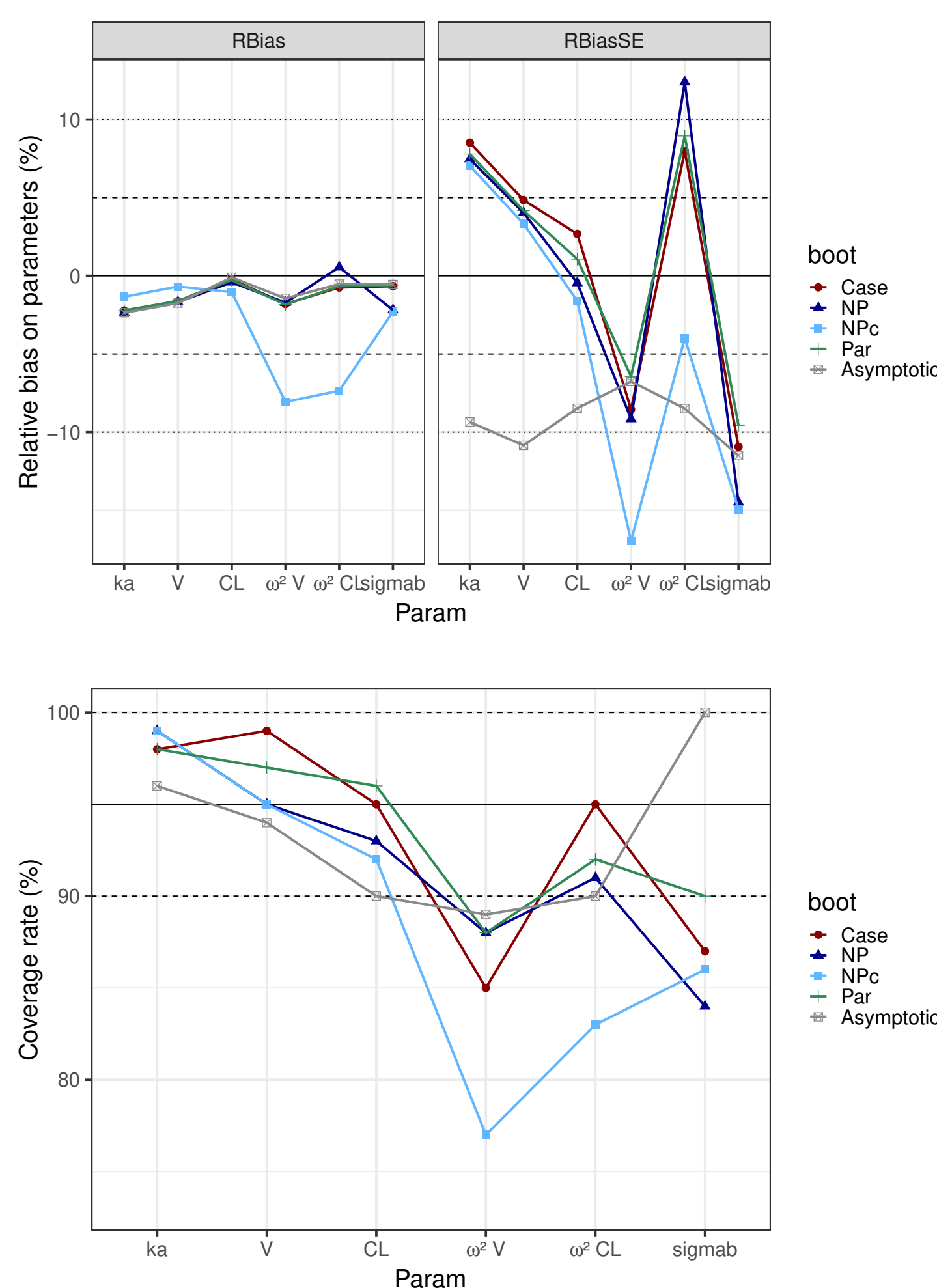


Figure 2: Single dose scenario - Top row: Relative bias on the parameters (left) and their estimation errors (right). Bottom: Coverage rate for the parameters.

- parameter estimates
 - no estimation bias with the asymptotic method
 - conditional non-parametric bootstrap underestimates the variabilities ω_{CL} and ω_V by about 8%
- SE estimates
 - biased estimations of the SE for at least some of the parameters for all bootstrap
 - although absolute differences in SE small in this example
 - more bias with conditional non-parametric bootstrap compared to the non-parametric bootstrap
- coverage rate
 - poor coverage rate for ω_V for all bootstraps
 - poor coverage rate for variance and residual error parameters with conditional non-parametric bootstrap

Method	ka	V	CL	ω_V	ω_{CL}	σ_b
CaseBoot	0.015	0.54	0.018	0.014	0.019	0.0070
NPBoot	0.014	0.54	0.017	0.014	0.020	0.0068
cNPBoot	0.014	0.53	0.017	0.013	0.017	0.0067
PBoot	0.014	0.54	0.017	0.015	0.019	0.0072
AsympSE	0.012	0.46	0.016	0.015	0.016	0.0070
Empirical	0.013	0.52	0.017	0.016	0.018	0.0079

Table 2: Bootstrap and asymptotic SE (AsymSE), compared to the empirical SE.

Steady-state scenario

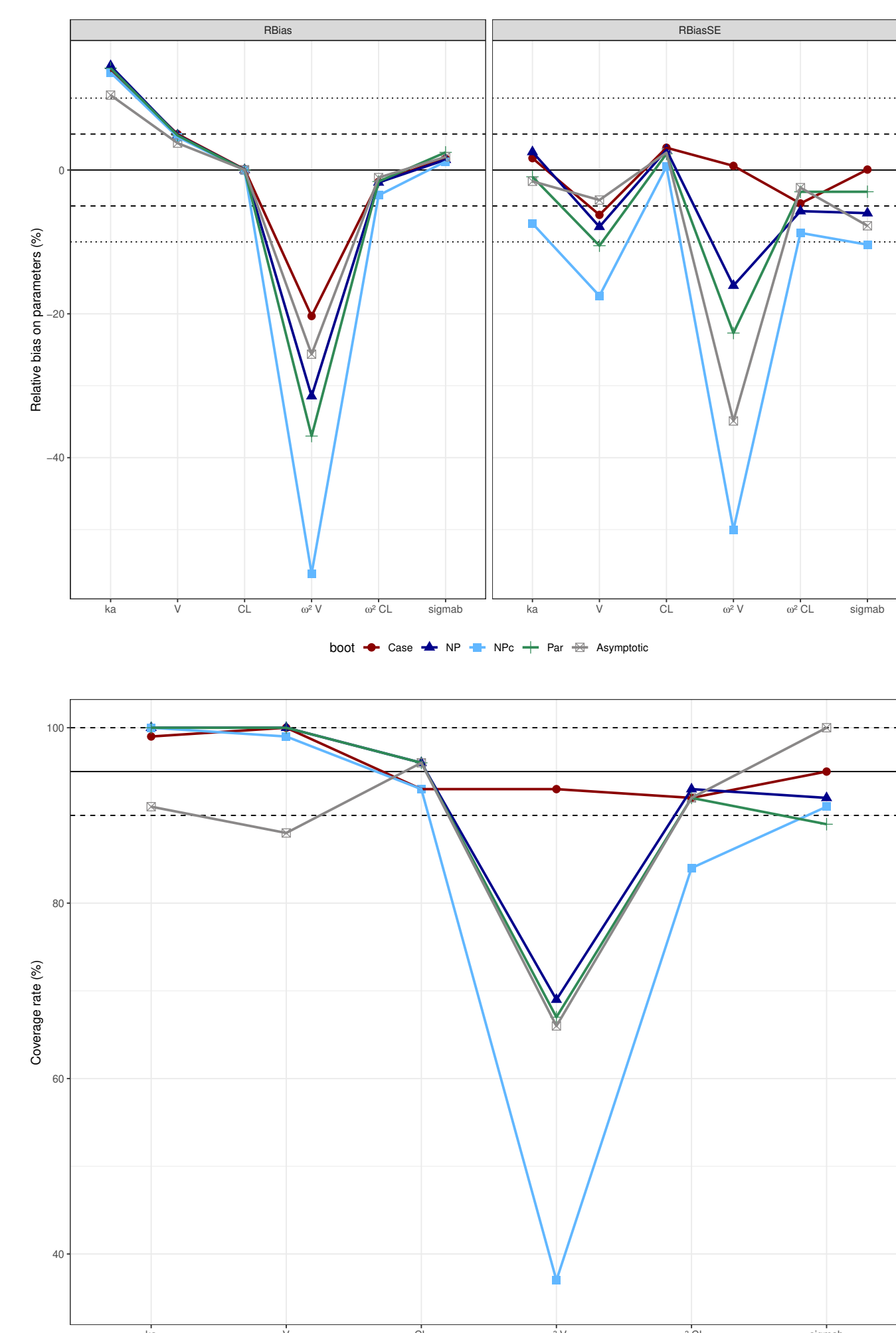


Figure 3: Steady state scenario - Top row: Relative bias on the parameters (left) and their estimation errors (right). Bottom: Coverage rate for the parameters.

- parameter estimates
 - biased estimates of k_a and ω_V
 - worst behaviour for conditional non-parametric bootstrap
 - similar figures when correction for flip-flop
- SE estimates
 - best performance for case bootstrap with only a 6% bias on $SE(V)$
 - good performance for asymptotic method except for the underestimation of SE on ω_V
 - again more bias with conditional non-parametric bootstrap compared to the non-parametric bootstrap
- coverage rate
 - lowest coverage rate for ω_V for all bootstraps
 - case bootstrap performs best
 - poor coverage rate for variance parameters with conditional non-parametric bootstrap

Conclusion

- Conditional non-parametric bootstrap
 - relies on samples from the conditional distribution in the hopes of correcting for shrinkage in a subject-specific way
 - but leads to increased bias on bootstrap estimates of both parameters and SE
- Simple evaluation scenario chosen for the simulations
 - none of the bootstraps improves over the asymptotic method
 - the bias in the estimates of the initial parameters, used as a basis for all bootstraps, is not corrected by any bootstrap

REFERENCES

- HT Thai, France Mentré, Nicholas H Holford, C Veyrat-Follet, and Emmanuelle Comets. A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models. *Pharmaceutical Statistics*, 12:129–40, 2013.
- HT Thai, France Mentré, Nicholas H Holford, C Veyrat-Follet, and Emmanuelle Comets. Evaluation of bootstrap methods for estimating uncertainty of parameters in nonlinear mixed-effects models: a simulation study in population pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 41:15–33, 2014.
- M Lavielle and B Ribba. Enhanced method for diagnosing pharmacometric models: random sampling from condi-

tional distributions. *Pharmaceutical Research*, 33:2979–88, 2016.

[4] C Rodrigues, S Chhun, C Chiron, O Dulac, E Rey, G Pons, and V Jullien. A population pharmacokinetic model taking into account protein binding for the sustained-release granule formulation of valproic acid in children with epilepsy. *Journal of Clinical Pharmacology*, 74:793–803, 2018.

[5] C Dumont, G Lestini, H Le Nagard, F Mentré, E Comets, TT Nguyen, and the PFIM Group. PFIM 4.0, an extended R program for design evaluation and optimization in nonlinear mixed-effect models. *Computer Methods and Programs Biomedicine*, 156:217–29, 2018.

[6] E Comets, A Lavenu, and M Lavielle. SAEMIX, an R version of the SAEM algorithm. *20th meeting of the Population Approach Group in Europe, Athens, Greece*, 2011. Abstr 2173.

[7] Marc Lavielle and Jonathan Chauvin. *Rsmix: R Speaks 'Monolix'*, 2018. R package version 1.1.0.

★ Presenting author

email: emmanuelle.comets@inserm.fr

