Evaluation of software tools for Bayesian estimation on population models: an update based on current software versions

Università degli studi di Pavia

R. Bartolucci, S.M. Lavezzi, E.M. Tosca, N. Melillo, S. Grandoni, E. Borella, L. Pasotti, G. De Nicolao, P. Magni

Department of Electrical, Computer and Biomedical Engineering, University of Pavia, via Ferrata 5, Pavia, Italy

BACKGROUND. Bayesian modelling based on Markov Chain Monte Carlo (MCMC) methods is acknowledged as a useful instrument in pharmacometrics. This work provides, after 3 years, an updated picture of a previous study [1], in which the performances of several software tools performing Bayesian estimation in a population context were compared in terms of efficiency and reliability of estimates, using as case studies an algebraic model and an ordinary differential equation (ODE) model.

METHODS.

WinBUGS 1.4.3 (with BlackBox Component Builder 1.5 and WBDiff interface)
RUNS ON: Windows.

Stan 2.17
RUNS ON: Windows, Linux, Mac OSX.

JAGS 4.3.0
RUNS ON: Windows, Linux, Mac OSX.

NONMEM 7.4.1 (with BAYES and NUTS)
RUNS ON: Windows, Linux, Mac OSX.

The study was conducted on a Windows 10 ASUS desktop PC, with Intel Core i5 3.30Ghz 4 cores and 8GB RAM.
The R *coda* package was used to analyze Markov chains.

For each model and tool, the number of iterations in the burn-in and stationary phases was computed based on the **Raftery algorithm** [2].

The tools capability to obtain uncorrelated samples was evaluated through **K,** i.e. the lag between two independent samples in the generated chain.

The **Effective Sample Size per execution time** unit (ESS/T) was calculated as an efficiency index.

$$\frac{ESS}{T} = \frac{n}{1 + 2\sum_{k=1}^{\infty}\rho_k} \cdot \frac{1}{T}$$

n: original sample size
ρ_k: autocorrelation function at lag K
T: execution time

ALGEBRAIC MODEL

Poisson count model, describing a clinical trial of an anticonvulsant therapy [3].
Covariates: treatment (Trt), 8-week baseline seizure counts ($Base$), age (Age), indicator variable for the 4th visit (V_4).
Random effects: Inter-Individual (b_{1j}) and Inter-Occasion (b_{jk}) variability.

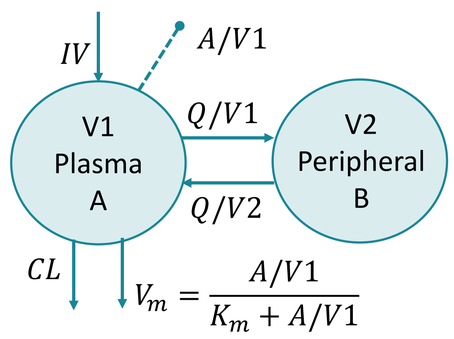
Patient	y_1	y_2	y_3	y_4	Trt	Base	Age
1	5	3	3	3	0	11	31
2	3	5	3	3	0	11	30
3	2	4	0	5	0	6	25
…							
59	1	4	3	2	1	12	37

$$y_{jk} \sim Poisson(m_{jk})$$
$$b_{1j} \sim N(0, \sigma_{b1}^2)$$
$$b_{jk} \sim N(0, \sigma_b^2)$$

$$\log(m_{jk}) = \alpha_0 + \alpha_{Base} \cdot \log(Base_j/4) + \alpha_{Trt} \cdot Trt_j + \alpha_{Age} \cdot \log(Age_j) + \alpha_{BT} \cdot Trt_j \cdot \log(Base_j/4) + \alpha_{V4} \cdot V4_k + b1_j + b_{jk}$$

ODE MODEL

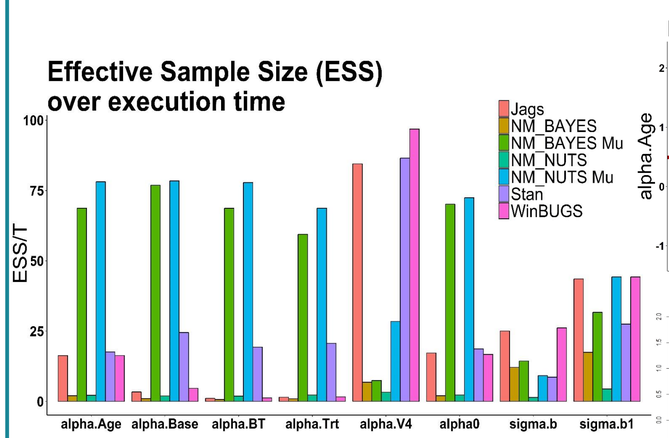
2-compartment PK ODE model with linear and non-linear elimination for a Phase I study of a monoclonal antibody for epilepsy [4,5].
Data: generated via Simulx using the reported parameter values.
Prior distributions: defined according to [4].
Error model: Additional + Proportional.

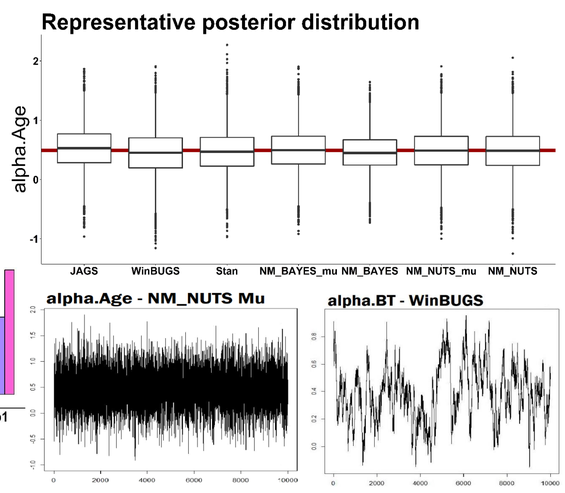


CL_{pop}=0.0045 [L/kg/day]	σ^2_{CL}=0.0495
$V1_{pop}$=0.0535 [L/kg]	σ^2_{v1}=0.025
$V2_{pop}$=0.036 [L/kg]	σ^2_{v2}=0.073
Q=0.0139 [L/kg/day]	σ^2_{add}=0.0001
Km=0.173 [mg/L]	σ^2_{prop}=0.0107
Vm=0.0037 [mg/kg/day]	

$$\log(CL_j) = \log(CL_{pop}) + \eta_{CL,j} \quad \eta_{CL,j} \sim N(0, \sigma^2_{CL})$$
$$\log(V1_j) = \log(V1_{pop}) + \eta_{V1,j} \quad \eta_{V1,j} \sim N(0, \sigma^2_{V1})$$
$$\log(V2_j) = \log(V2_{pop}) + \eta_{V2,j} \quad \eta_{V2,j} \sim N(0, \sigma^2_{V2})$$

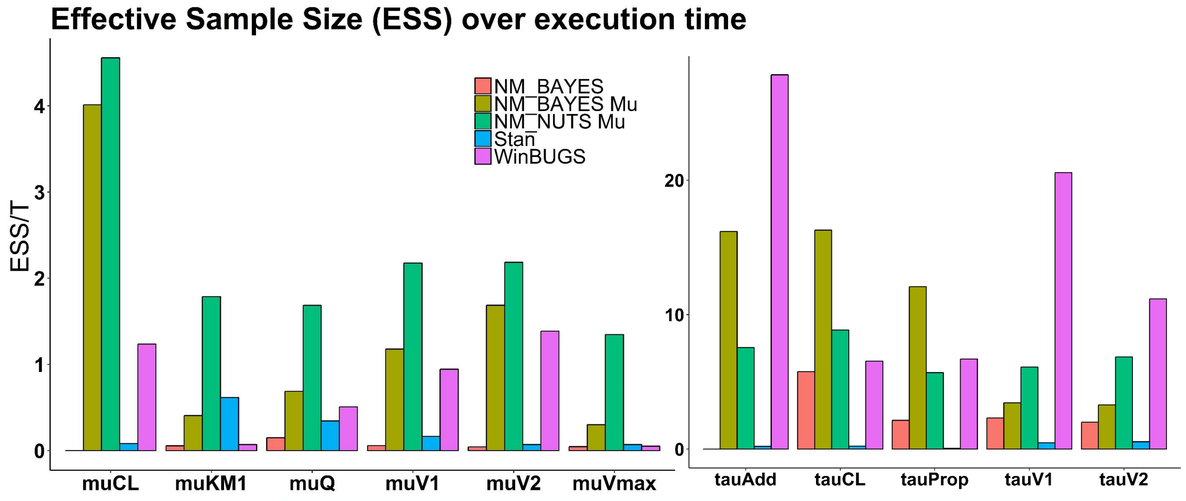
RESULTS. ALGEBRAIC MODEL



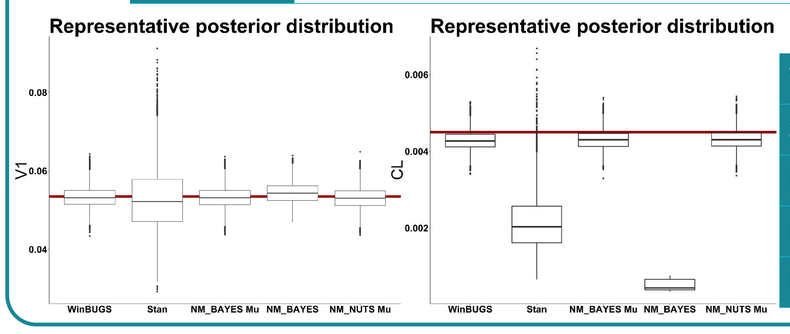


	WinBUGS	Jags	Stan	NM_BAYES Mu	NM_BAYES	NM_NUTS Mu	NM_NUTS
Time [s]	97.5	108	28.2	33.1	241.7	38.4	339.8

RESULTS. ODE MODEL



	WinBUGS	Stan	NM_BAYES Mu	NM_BAYES	NM_NUTS Mu
Time [min]	19.9	63.6	9.9	287.8	7.9



	Kmax	Kmin
WinBUGS	205	1.03
Stan	6.73	0.93
NM_BAYES Mu	34.4	1.1
NM_BAYES	1838.9	6.26
NM_NUTS Mu	3.63	0.62

	WinBUGS	Jags	Stan	NM_BAYES Mu	NM_BAYES	NM_NUTS Mu	NM_NUTS
Algebraic	✓	✓	✓✓	✓✓	✓	✓✓	✓
ODE	✓	✗	✓	✓✓	✓	✓✓	✗

CONCLUSIONS.

Algebraic model: the posterior distributions of all the tools were similar to the expected ones.
With count data, NONMEM required the objective function to be written explicitly, resulting in a less user-friendly model encoding.
As for ESS/T, NONMEM NUTS and BAYES methods with mu referencing showed better performance with respect to the other tools.
Compared to BAYES, NUTS slightly improved both the efficiency and the estimation results.
ODE model: no tool was able to recover the expected posterior distributions [4] for all model parameters.
In terms of ESS/T, the best performances were obtained with NONMEM NUTS and BAYES methods with mu referencing for fixed effects, whereas WinBUGS showed higher ESS/T for random effects.
Improvements: the NUTS algorithm used in Stan has been successfully implemented as a new feature in NONMEM 7.4.1. In this version, more flexibility has also been given to users in terms of prior distribution choices.
Differently from the previous study, Stan was able to finish the estimation process, even if the estimated posterior distributions are biased and highly skewed.

REFERENCES.

[1] E. Borella et al. Evaluation of software tools for Bayesian estimation on population models with count and continuous data. Proceedings of the 2015 PAGE meeting, June 2-5 Hersonissos, Greece
[2] A.E. Raftery and S.M. Lewis. Implementing MCMC. Markov Chain Monte Carlo in Practice (1996), London: Chapman and Hall, pp. 115-130
[3] http://www.openbugs.net/Examples/Epil.html
[4] F. Strimenopoulou et al. Bayesian non-linear PK modelling applied to dose escalation studies using WinBUGS. Proceedings of the Bayes 2012 meeting, May 9-11, Basel, Switzerland
[5] R. Lledo-Garcia et al. Dose escalation studies for mAb: prior distributions selection and software comparison. Proceedings of the PAGE meeting (2012), June 5-8, Venice, Italy