Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models.

Elodie L. Plan, Alan Maloney, France Mentré, Mats O. Karlsson, Julie Bertrand

▶ **To cite this version:**

HAL Id: inserm-00709829

https://inserm.hal.science/inserm-00709829

Submitted on 19 Jun 2012

TITLE

Performance Comparison of various Maximum Likelihood Nonlinear Mixed-effects Estimation Methods for Dose-Response Models

AUTHORS

Elodie L. Plan [1,2], Alan Maloney [1,3], France Mentré [4], Mats O. Karlsson [1], Julie Bertrand [4]

AFFILIATIONS

[1] Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

[2] Correspondence: Elodie Plan (e-mail: elodie.plan@farmbio.uu.se)

[3] Exprimo NV, Mechelen, Belgium

[4] UMR-S 738, INSERM and University Paris Diderot, Paris, F-75018, France

RUNNING HEAD

Maximum Likelihood Methods for Dose-Response Models

ABSTRACT

Estimation methods for nonlinear mixed-effects modelling have considerably improved over the last decades. Nowadays several algorithms implemented in different softwares are used. The present study aimed at comparing their performance for dose-response models.

Eight scenarios were considered using a sigmoid E_{max} model, with varying sigmoidicity factors and residual error models. 100 simulated datasets for each scenario were generated. 100 individuals with observations at 4 doses constituted the rich design and at 2 doses for the sparse design. Nine parametric approaches for maximum likelihood estimation were studied: FOCE in NONMEM and R, LAPLACE in NONMEM and SAS, adaptive Gaussian quadrature (AGQ) in SAS, and SAEM in NONMEM and MONOLIX (both SAEM approaches with default and modified settings). All approaches started first from initial estimates set to the true values, and second using altered values. Results were examined through relative root mean squared error (RRMSE) of the estimates.

With true initial conditions, full completion rate was obtained with all approaches except FOCE in R. Runtimes were shortest with FOCE and LAPLACE, and longest with AGQ. Under the rich design with true initial conditions, all approaches performed well except FOCE in R. When starting from altered initial conditions, AGQ, and then FOCE in NONMEM, LAPLACE in SAS, and SAEM in NONMEM and MONOLIX with tuned settings, consistently displayed lower RRMSE than the other approaches.

For standard dose-response models analyzed through mixed-effects models, differences could be identified in the performance of estimation methods available in current software.

KEYWORDS

MAXIMUM LIKELIHOOD ESTIMATION, FOCE, LAPLACE, ADAPTIVE GAUSSIAN QUADRATURE, SAEM

INTRODUCTION

Non-linear mixed-effects models (NLMEM) were introduced to the biomedical field about 30 years ago (1-3) and have substantially improved the information learned from preclinical and clinical trials. Within drug development, NLMEM were initially used for pharmacokinetic (PK) analyses (4), before being extended to pharmacokinetic-pharmacodynamic (PKPD) analyses (5), along with dose-response analyses. On top of the structural mathematical model fit to PK or/and PD observations, the statistical model components enable the modeller to characterize results obtained in a set of individuals with the same parametric model and, in addition, to estimate the interindividual variability (6), and to quantify the unexplained variability (7).

The estimation of the fixed effect and random effect parameters involve complex estimation methods. Maximum likelihood estimation (MLE) approaches constitute a large family of methods commonly used in NLMEM analyses (8). The non-linearity of the regression function in the random effects prevents a closed form solution to the integration over the random effects of the likelihood function (9), thus several algorithms have been developed for MLE. Gaussian assumptions for the distribution of the random effects are common among MLE methods, and form the group of parametric approaches (10).

Along with methodological developments, different software have emerged, the most commonly used one (11) being NONMEM (12). Estimation algorithms available were first restricted to First-Order (FO) and then First-Order Conditional Estimation (FOCE), which were subsequently implemented in Splus, R and WinNonMix. LAPLACE (13) then appeared in NONMEM, while SAS witnessed the addition of two macros MIXLIN and NLINMIX. A later procedure in SAS that represented a considerable improvement was NLMIXED, with FO and adaptive Gaussian quadrature (AGQ). Alternatives followed with stochastic expectation maximisation (EM) algorithms, and especially the SAEM algorithm (14) implemented in the MONOLIX (15) and the NONMEM (16) software.

Whilst the estimation algorithms use different statistical methods, all aim at producing reliable estimates of the model parameters. The complexity of the model and the approximations embedded in the algorithm could potentially lead to poor estimation

performance. This performance is measured through precision and accuracy. As the estimates may impact on clinical decisions and lead to biomedical conclusions, selecting an estimation method with lower bias and higher precision is desirable.

In the past, several studies comparing algorithms have been performed, stimulated by the introduction of new algorithms (17, 18), as a systematic comparison from a workgroup (19), in order to highlight practical applications (20), or as a complex-problem solving survey (21). However, apart from (17, 18), these investigations were not supported by a high number of simulations, but rather considered the analysis of only one simulated dataset (19, 21) or one real dataset (20).

Recently, large Monte Carlo simulation studies compared estimation methods performance for PD count (22, 23), categorical (24, 25), and repeated time-to-event (26) models, enlarging the challenge represented by the model type. Estimation methods compared over all these five investigations were LAPLACE in NONMEM, AGQ in SAS, SAEM in MONOLIX, SAEM in NONMEM and importance sampling in NONMEM. Nevertheless, rarely more than three approaches were compared within a study, although the panel of algorithms and software available to the modeller is now rich and diversified. A wider comparison has been performed for continuous PK data (27) and remained to be for dose-response analyses.

The objectives of this study were to measure and compare the estimation performance of FOCE in NONMEM and R, LAPLACE in NONMEM and SAS, adaptive Gaussian quadrature in SAS, and SAEM in NONMEM and MONOLIX for a set of dose-response scenarios.

METHODS

1. Statistical model

Let $d = d_1, \ldots, d_K$ be a set of ordered dose levels selected in a dose-response study and y_{ik} be the response of subject $i = 1, \ldots, N$ to the dose d_k. The dose-response is assumed to be adequately described by a function *f* such as:

$$y_{ik} = f(d_k, \phi_i) + \varepsilon_{ik} \qquad (1)$$

where ϕ_i is the p dimensional vector of the model individual parameters for subject i and ε_{ik} is the measurement error. ε_{ik} given ϕ_i are assumed to be independent and normally distributed with a zero mean and a variance σ_{ik}^2 which can be additive ($\sigma_{ik}^2 = \sigma^2$) or proportional ($\sigma_{tk}^2 = f(d_k,\phi_i)^2 \times \sigma^2$). f is a function than can be nonlinear with respect to the parameters ϕ.

ϕ_i depend on the fixed effect p-dimensional vector μ and the random effect q-dimensional vector η_i in the following manner when considering an exponential model to ensure positivity:

$$\phi_i = \mu \cdot e^{B \cdot \eta_i} \tag{2}$$

with the random effects following a Gaussian distribution with a zero mean and a variance matrix Ω of size (q×q), whose diagonal elements are variances ω^2. The (p×q)-matrix B allows some components of ϕ not to have a random part. Also, the exponential random effect model ensures the positivity of the model parameter.

Finally, let define the vector of all the model parameters as $\Psi = (\mu', \mathrm{Vech}(\Omega)', \sigma)$ where the operator Vech(.) creates a column vector from the matrix by stacking its lower diagonal elements below one another.

2. Likelihood function

The log-likelihood $L(y; \Psi)$ is the sum over the N subjects of the individual likelihoods, $L(y_i; \Psi)$:

$$L(y; \Psi) = \sum_{i=1}^{N} L_i(y_i; \Psi) \tag{3}$$

where the individual log-likelihood $L_i(y_i; \Psi)$ is defined as follows:

$$L_i(y_i; \Psi) = \log\left(\int p(y_i, \eta_i; \Psi) d\eta \right) = \log\left(\int p(y_i | \eta_i; \Psi) p(\eta_i; \Psi) d\eta \right) \tag{4}$$

with $p(y_i | \eta_i; \Psi)$ the conditional density of the observations given the individual random effects, $p(\eta_i; \Psi)$ the density of the individual random effects, and $p(y_i, \eta_i; \Psi)$ the likelihood of the 'complete' data which correspond to the observations plus the random effects, η_i.

3. Estimation algorithms

Estimation methods are briefly described here. More details may be obtained in the original articles.

3.1. First-Order Conditional Estimation (FOCE)

As initially described by Lindstrom and Bates (28), the algorithm approximates (4) by the log-likelihood of a linear mixed effect model. The η_i and updated estimates of μ are obtained by minimizing a penalized nonlinear least square (PNLS) objective function using the current estimates of Ω and σ. Then, the model function f is linearized using a first-order Taylor expansion around the current estimates of μ and the conditional mode of the η_i so that (4) can be approximated by the log-likelihood of a linear mixed effect (LME) model to estimate Ω and σ. The maximization is realized through a hybrid approach starting with a moderate number of EM iterations before switching to Newton-Raphson iterations. The approach alternates between PNLS and LME until a convergence criterion is met. They implemented their method in the nlme function of the R software (29).

In the NONMEM software, the conditional modes of the η_i are obtained by maximizing the empirical Bayes posterior density of η_i, $p(\eta_i|y_i ; \Psi)$, using the current estimates of vector Ψ:

$$p\left(\eta_i \middle| y_i ; \Psi\right) = \frac{p(y_i|\eta_i ; \Psi)p(\eta_i ; \Psi)}{\int p(y_i|\eta_i ; \Psi)p(\eta_i ; \Psi)d\eta} \qquad (5)$$

Also, (4) is approximated by a second order Taylor expansion of the integrand (also called Laplacian approximation) around the η_i ; however the Hessian is approximated by a function of the gradient vector to avoid the direct computation of second-order derivatives. For an additive residual error model, both the approximation by the linearization of the function f and the Laplacian approximation using an approximated Hessian have been shown to be equivalent asymptotically (9). However, this equivalence no longer holds in case of interaction between the η_i and the ε_{ik}, as in the proportional error model. A derivative-free quasi-Newton type minimization algorithm is used.

3.2. Laplacian approximation (LAPLACE)

The principle of this algorithm is to approximate (4) by a second order Taylor expansion of the integrand around the conditional mode of the η_i, which are obtained by maximizing the empirical Bayes posterior density of the η_i using the current estimates of vector Ψ.

In the NLMIXED procedure of the SAS software (30), this algorithm is implemented as a special case of the adaptive Gaussian quadrature algorithm (see below) where only one abscissa is defined at the conditional modes of the η_i with a corresponding weight equal to 1. Also, the η_i are also obtained by maximizing $p(\eta_i|y_i; \Psi)$ with a default dual quasi-Newton optimisation method.

3.3. Adaptive Gaussian Quadrature (AGQ)

The principle of this algorithm is to numerically compute (4) by a weighted average of $p(y_i|\eta_i; \Psi) \, p(\eta_i; \Psi)$ at predetermined abscissa for the random effects using a Gaussian kernel. Pinheiro and Bates (31) suggested using standard Gauss-Hermite abscissa and weights (32), with the abscissa centred around the conditional mode of the η_i and scaled by the Hessian matrix from the conditional mode estimation (33). The adaptive Gaussian approximation can be made arbitrarily accurate by increasing the number of abscissa.

3.4. Stochastic Approximation Expectation Maximization (SAEM)

SAEM is an extension of the EM algorithm where individual random effects are considered as missing data (34). It converges to maximum likelihood estimates by repeatedly alternating between the E and M steps. As the E step is often analytically intractable for nonlinear models, the E step in SAEM is replaced by a simulation step where the η_i are drawn by running several iterations of a Hastings-Metropolis algorithm using three different kernels successively (35). Then the expectation of the complete log-likelihood $Q(\Psi) = E(\log(p(y, \eta ; \Psi)))$ is computed according to a stochastic approximation:

$$Q_m(\Psi) = Q_{m-1}(\Psi) + \gamma_m \left(\log \left(p(y, \eta_m ; \Psi) - Q_{m-1}(\Psi) \right) \right) \qquad (6)$$

where γ_m is a decreasing sequence of positive numbers over the m = 1, …, M algorithm iterations with $\gamma_1 = 1$. The SAEM algorithm has been shown to converge to a maximum (local or global) of the likelihood of the observations under very general conditions (36).

4. Simulation and estimation study

This simulation study consisted, for each studied scenario, of 100 stochastic simulated datasets generated in NONMEM and subsequently analysed with the different studied approaches (i.e. implementation of the estimation algorithms in the various software).

4.1. Simulations

4.1.1. Design

The dataset structure mimicked a clinical trial including 100 individuals and investigating four dose levels: 0, 100, 300 and 1000 mg. A continuous PD outcome was recorded for each individual following two simulation designs: (i) the rich design counted four observations per individual, one at each dose level, whereas (ii) in the sparse design each individual was randomly allocated to only two of the four dose levels.

4.1.2. Base model

A dose-response model based on a sigmoid E_{max} function with a baseline (E_0) was constructed as in (7). The Hill factor (γ) is responsible for the sigmoidicity, i.e. the degree of non-linearity of the function shape.

$$E_i = E_{0_i} + \frac{E_{max_i} \cdot d^{\gamma}}{ED_{50_i}{}^{\gamma} + d^{\gamma}} \qquad (7)$$

Gaussian random components with normal zero-mean distribution were assumed for all individual parameters except for γ. A correlation in the variances of the random effects for E_{max} and ED_{50} was assumed. The residual error model was assumed to be additive or proportional (see 2.1). Selected parameters values are reported Table 1.

4.1.3. Scenarios

Eight simulation scenarios (s = 8) were derived, exploring (i) the two previously described simulation designs: rich (R) and sparse (S), (ii) three values of γ: 1, 2, and 3, and (iii) two error models: additive (A) and proportional (P). They were referred to as: R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P, and corresponded to eight sets of 100 simulated datasets to be analysed. Note that for the sparse design only sets with $\gamma = 3$, the most non-linear model, were evaluated.

4.2. Estimations

4.2.1. Initial conditions

The same model from which the simulated datasets were generated was used for estimation. Each dataset was analysed twice: (i) with true initial conditions, i.e. starting estimate values set to the original parameter values on which simulations were based, and (ii) with altered initial conditions: γ set to 1, the other fixed effects to two fold of their true value, and random effects to low numbers (Table 1). This procedure explored the robustness of the approaches.

4.2.2. Software settings

Estimation algorithms were mostly utilised with the default settings with which they are available in the different studied software. Changes from these defaults were listed Table 2 and reported below.

FOCE and LAPLACE in NONMEM 7.1.0 (FOCE_NM and LAP_NM) had the maximum number of iterations set to the highest possible value as done in common practice, and the option INTERACTION was added for the scenarios with a proportional error. FOCE in R 2.9.1 (FOCE_R) was using the nlme routine. LAPLACE and AGQ in SAS 9.2 (LAP_SAS and AGQ_SAS) were adaptive Gaussian quadrature respectively corresponding to a number of quadrature points (QPOINTS) of 1 and 9. Other settings listed in table 2 were adapted from the defaults (FTOL=1E-15.7 XTOL=0 TECH=QUANEW EBSTEPS=50 EBSUBSTEPS=20 EBSSFRAC=0.8 EBTOL=2.2E-12 INSTEP=1) in SAS. These settings were used previously (22) to improve robustness in the conditional modes calculations (the EB options) or to reduce the very high default convergence criteria (for FTOL and XTOL).

SAEM presents a number of settings the user is invited to modify, that can follow different terminologies depending on the software: NONMEM 7.1.0/MONOLIX 3.1. These include the numbers NBURN/K1 and NITER/K2 of iterations in the stochastic ($\gamma_k = 1$) and the cooling (decreasing γ_k) phases, respectively, as well as the number ISAMPLE/nmc of chains in the MCMC procedure. Stopping rules can also be defined for the two software for the stochastic phase, and also for the cooling phase in MONOLIX only. A simulated annealing version of SAEM during the first iterations can be set in NONMEM while it is automatically performed in MONOLIX. Moreover, ϕi can be defined as the log-transform of a Gaussian random vector to meet with constraints of positivity, which corresponds to mu-referencing in NONMEM and the default in MONOLIX. In light of these possibilities, SAEM was run with each software

9

twice: once with the default settings (SAEM_NM and SAEM_MLX), and a second time with modified settings (SAEM_NM_tun and SAEM_MLX_tun). SAEM_NM was run with the defaults NITER=1000, ISAMPLE=2 and IACCEPT=0.4, and with the number of iterations from the stochastic phase NBURN≤2000 being stopped with a convergence test for termination CTYPE=3 based on objective function, fixed effects, residual error and all random effect elements. SAEM_NM_tun had parameters linearly mu-referenced, decreased number of iterations in the two phases and increased number of individual samples. Concerning the convergence, it was stopped in the same manner as SAEM_NM, but instead of every 9999 iterations being submitted to the convergence test system, only every 25 were.. SAEM_MLX was run with setting the maximal number of iterations for the stochastic (K1≤500) and the cooling phase (K2≤200) using the following stopping rules: i) the stochastic phase is ended before K1 is reached if an iteration m is met where $p(y, \eta_m; \Psi_m) < p(y, \eta_{m-1K1}; \Psi_{m-1K1})$ with lK1=100 and ii) the cooling phase is ended before K2 is reached if an iteration m is met where the variances of the parameters, computed over a window of lK2 iterations, is reduced by a factor rK2 compared to their values at the end of the stochastic phase, with lK2=50 and rK2=0.1. SAEM_MLX_tun was tuned in the way that it had a number of iterations for the stochastic phase, K1=500 (i.e. not using the stopping rule for this phase), and increased individual samples, nmc=5.

Hence nine approaches (a = 9) were explored through the estimation of the simulated datasets: FOCE_NM, FOCE_R, LAP_NM LAP_SAS, AGQ_SAS, SAEM_NM, SAEM_NM_tun, SAEM_MLX, and SAEM_MLX_tun.

4.3. Computer power

FOCE, LAPLACE and SAEM run in NONMEM 7.1.0 were assisted with PsN 3.2.5 (37) on a Linux cluster node of 3.59 GHz with a G77 Fortran compiler. Estimations with FOCE in R were done on a 2.49 GHz CPU as well as some with SAEM in MONOLIX (others were on a 1.83 GHz), assisted by a Matlab version R2009b. All SAS runs (LAPLACE and AGQ) were performed on a 2.66 GHz computer using SAS 9.2 for Windows.

5. Performance comparison

5.1. Completion rates

The proportion of completed estimations, i.e. the number K of the 100 analysed datasets that produced parameter estimates with each approach was reported. Other computations were executed with these Z sets of results; however when less than 50 of the runs completed, statistical measures were not produced. Z, thereafter expressed as a percentage, was therefore assessing the stability of the different approaches, whereas results were given only when K ≥ 50.

5.2. Runtimes

Runtimes were recorded as the CPU time needed to estimate each of the 100 copies of a simulated scenario. Then the average was calculated. A correction was done with the clock rate of the processor in the computer on which runs were performed as in (8). Parallelization was not possible with the investigated approaches, so did not have to be accounted for.

$$NI_{s,a} = \frac{1}{K} \sum_{k=1}^{K} CPUt_{s,a,k} \cdot CPUf_{.,a} \qquad (8)$$

where $NI_{s,a}$ is the calculated number of instructions in billions for scenario s with approach a, $CPUt_{s,a,k}$ the real time in seconds recorded on a CPU to perform the corresponding k^{th} estimation, and $CPUf_{.,a}$ the frequency in GHz (equivalent to billion instructions per second) of the clock in the utilized CPU.

5.3. Accuracy and precision

Relative estimation errors (RER), relative bias (RBias), and root mean squared error (RMSE) were computed such that the accuracy and the precision of the estimation algorithms were evaluated for each of the 9 components (p) of the vector Ψ. The RER (%) are evaluated for each estimate and box-plot of RER(%) show both bias (mean) and imprecision (width). The RBias (%) describes the deviation of the mean over the estimated parameters from their true value. The relative RMSE (RRMSE %) summarize both the bias and the variability in estimates. The Standardized RRMSE (%) was constructed for each parameter and each approach as the RRMSE divided by the lowest RRMSE value obtained across all approaches for that parameter in (12).

$$RER_k(\Psi_{p,s,a}) = \left(\frac{\widehat{\Psi}^k_{p,s,a} - \Psi^*_{p,s,.}}{\Psi^*_{p,s,.}} \right) \cdot 100 \qquad (9)$$

$$RBias(\Psi_{p,s,a}) = \frac{1}{K} \sum_{k=1}^{K} RER_k(\Psi_{p,s,a}) \qquad (10)$$

$$RRMSE(\Psi_{p,s,a}) = \sqrt{\frac{1}{K} \sum_{k=1}^{K} RER_k(\Psi_{p,s,a})^2} \qquad (11)$$

$$\text{Standardized } RRMSE(\Psi_{p,s,a}) = \frac{RRMSE(\Psi_{p,s,a})}{\min_a\left(RRMSE(\Psi_{p,s,a})\right)} \qquad (12)$$

where $\widehat{\Psi}_p^k$ is the estimated p component for the kth data set and Ψ_p^* the true value.

For each scenario and each approach, mean standardized RRMSE across the 9 components of Ψ was computed as a global measure of the performance.

Computations were conducted in R 2.11.1.

RESULTS

1. Completion rates

100 % of the analyses started from true initial conditions completed with final estimates for all the approaches except FOCE_R (99, 62, 5, 69, 32, 2, 16, and 33 % for the R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P scenarios, respectively) (Figure 2). The same simulated datasets estimated with altered starting values gave completion rates of the same order with FOCE_R (98, 76, 16, 68, 8, 3, 5, and 10 % for the R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P scenarios, respectively), decreased ones with SAEM_NM (97, 91, 16, 74, 81, and 75 % for the R1A, R3A, R1P, R2P, R3P, and S3P scenarios, respectively) and SAEM_NM_tun (91 and 67 % for the R3A and S3A scenarios), and maximum completion (100 %) for all the other approaches. Therefore 133 sets of estimates were considered for further comparison statistics, 11 failing to meet the 50 % completion criterion.

2. Runtimes

Runtimes expressed as number of instructions (NI) ranged from 4 to 1614 billion instructions (BI), and are displayed for estimations starting from true initial conditions in Figure 2. FOCE_NM was the fastest approach (median NI = 7.2 BI and 9.6 BI, starting from

true and altered initial conditions, respectively), never taking longer than 15 BI, very closely followed by FOCE_R and LAP_SAS. LAP_NM was displaying equivalently short runtimes for the additive error models (median NI = 10.2 BI and 11.2 BI, starting from true and altered initial conditions, respectively), which were doubled (median NI = 22.7 BI and 27.3 BI, starting from true and altered initial conditions, respectively) for the proportional error models, the design having no noticeable impact. SAEM approaches with default settings were systematically slower than FOCE and LAPLACE, but it was faster in MONOLIX (median NI = 43.2 BI and 52.6 BI, starting from true and altered initial conditions, respectively) than in NONMEM (median NI = 147.7 BI and 287.8 BI, starting from true and altered initial conditions, respectively), by around 3 folds when the initial conditions were true and 6 folds when they were altered. The tuned version of the approach, SAEM_MLX_tun, took around 2.5 times longer (median NI = 117.6 BI) than the non-tuned version, whereas SAEM_NM_tun (median NI =79.9 BI) was almost 3 times faster than SAEM_NM and 1.5 times faster than SAEM_MLX_tun; both had very similar runtimes between true and altered initial conditions. The NI reached with AGQ_SAS was high (median NI = 674.8 BI and 864.1 BI, starting from true and altered initial conditions, respectively); it was consistently the slowest.

3. Accuracy and precision

Boxplots of RER for ED_{50} and $\omega^2(ED_{50})$ estimates are displayed on Figures 3a and 3b as they often are the main parameters of interest in dose-response studies. Standardized RRMSE star-plots with 9 radii for each of the elements of Ψ are represented in Figure 4; on a given radius, the closer to 1, the closer is the performance relative to the approach with the smallest RRMSE for the parameter of interest. For a global assessment across parameters, mean standardized RRMSE are illustrated in Figure 5.

3.1. True initial conditions

As displayed in Figure 3a, the parameter ED_{50} was globally accurately estimated under true conditions, but presented a lower precision for scenarios with $\gamma = 1$. The highest and most consistent biases were observed with FOCE_R, on the few scenarios for which metrics were produced due to poor completion rates. ED_{50} was better estimated with AGQ_SAS, LAP_NM and FOCE_NM on the sparse design than with the other tested approaches, which produced some bias, especially LAP_SAS (interquartile range excluding zero), and exhibited

13

imprecision (wide interquartile range and longer whiskers), especially the SAEM approaches (except SAEM_NM). For the parameter $\omega^2(ED_{50})$ (Figure 3b), estimates were slightly more biased, but essentially more imprecise, especially with $\gamma = 1$, and the additive error model. For the sparse design, most approaches exhibited a bias, except the four SAEM approaches, which appeared to provide more accurate but less precise estimates than the other approaches.

SAEM_NM obtained the lowest RRMSEs whatever the scenario and parameter (values available in appendix); as illustrated in Figure 4, when $\gamma > 1$ and the error model was additive, all approaches but SAEM_NM estimated large E_{max}, and when $\gamma = 3$ and the error model was proportional, all approaches but SAEM_NM estimated large ED_{50}.

Globally on the rich design, as represented Figure 5, all approaches had a mean standardized RRMSE below 1.5 for most of the scenarios with the exception of FOCE_R. Nevertheless, for scenario R3A, FOCE_NM and SAEM_MLX had it slightly above 1.5. On the sparse design, the LAPLACE methods, AGQ_SAS, and SAEM_NM had mean standardized RRMSEs below 1.5, whereas SAEM_MLX had it above 1.5 for both error models and SAEM_NM_tun and SAEM_MLX_tun for only the S3A scenario.

3.2. Altered initial conditions

On the rich design, most of the approaches estimated ED_{50} similarly as when starting from true values, as illustrated in Figure 3a. However, the results of the SAEM approaches changed compared to the previous initial conditions case and sometimes drastically for the versions with the default settings, even failing to reach 50 % of completion for SAEM_NM with scenario R1P. On the sparse design, most of the methods obtained biased estimates, with the exceptions of AGQ_SAS, SAEM_NM, and FOCE_NM, which gave the distributions of 100 estimated ED_{50} the most centred on the true value and tight. FOCE_R results could not be assessed, but the other approaches presented tailed distributions of estimated ED_{50}, with quartiles not including the true value for LAP_SAS with both scenarios models and for SAEM_MLX with S3A. As shown in Figure 3b, the bias and imprecision in the $\omega^2(ED_{50})$ estimates were increased by starting from altered initial conditions particularly for SAEM_NM, whereas SAEM_MLX_tun yielded the boxplot most centred on zero.

It can be observed in Figure 4 that FOCE_NM and AGQ_SAS obtained standardized RRMSEs below 1.5 on most scenarios and parameters. When the sparse design was adopted the SAEM approaches and the LAPLACE approaches obtained standardized RRMSEs above

1.5 on most parameters, but for the proportional error model scenario they were below 1.5 with SAEM_NM_tun. FOCE_R estimated most parameters with poor standardized RRMSE, but especially γ and σ.

On Figure 5, FOCE_NM, and AGQ_SAS are shown to have lowest mean standardized RRMSE whatever the scenario, with LAP_SAS and SAEM_MLX_tun having mean standardized RRMSE below 1.5 for all but one scenario (S3P and S3A respectively). FOCE_R obtained mean standardized RRMSE above 1.5 on all scenarios where its performance could be evaluated, whereas SAEM_NM, SAEM_MLX and LAP_NM also obtained elevated mean standardized RRMSE on at least half of the scenarios.

DISCUSSION

The present work provides a comparison in terms of speed, robustness, bias and precision of the most commonly used likelihood-based estimation approaches in nonlinear mixed effect modelling for the fitting of a dose-response model.

FOCE_R was shown to be the least robust approach with less than 50 % completion rate on 9 of the 16 combinations of scenarios and initial conditions settings investigated. All other approaches could be evaluated as they completed at least half of the data sets, with the exception of SAEM_NM in one situation. However the convergence criteria differed across estimation methods. In NONMEM, convergence of classical methods (FOCE and LAPLACE) is based only on the parameter estimation gradient, whereas it was set to be based on objective function, thetas, sigmas, and all omega elements for the SAEM methods. In MONOLIX, the automatic stopping rule for the stochastic phase is based on the complete log-likelihood. In SAS, convergence is primarily based on 6 key criteria, relating to the absolute and relative changes in the likelihood, gradients, and parameter values. The difficulty in defining convergence complicates these comparisons.

The convergence criteria used will affect runtimes, with less strict convergence criteria yielding shorter runtimes. However it is believed that the trends would remain the same, with the classical methods FOCE and LAPLACE being the fastest, and AGQ being the slowest. AGQ slow runtimes were due to the high number of quadrature points chosen (9 quadrature points across 3 random effects imply 729 (9^3) likelihood evaluations for each subject at each

iteration). Reducing this (e.g. to 3 quadrature points) would have significantly shortened the runtimes, and may have led to similar results (not inspected). Unsurprisingly, the estimation process speed was driven by the extent of the likelihood function simplification, with first-order linearization-based algorithms achieving the shortest run times. Within each iteration, the SAEM approaches are faster than the Gaussian quadrature-based method because they sample the integrand rather than fully integrating it, but many more iterations are needed with SAEM than with AGQ. Increasing the number of chains to the SAEM algorithm was additionally time-consuming in MONOLIX, whereas SAEM_NM_tun was overall faster than SAEM_NM due to the number of iterations being decreased.

Globally, the approximation based on a linearization of the model, but for FOCE_R, gave good results for the fixed effects (relative biases typically less than 3 %) when starting from the true conditions, with $\omega^2(ED_{50})$ and $Cov(E_{max},ED_{50})$ being least well estimated. As for their precision, it was decreasing in a similar extent using altered conditions and/or on a sparse design. The performance of adaptive Gaussian quadrature was high on all cases. The conclusions were less straightforward for the SAEM approaches. Indeed, SAEM_NM lacks a global search first step in order to refine the initial estimates; this could be appreciated with the results of the scenarios starting from altered values compared to SAEM_MLX. . However increasing the number of individual samples and linearly mu-referencing the parameters substantially improved the results. Mu-referencing appeared to yield more efficient behaviour of SAEM_NM_tun according to the implementation of the algorithm in NONMEM. SAEM_MLX performance with altered initial conditions comes from the fact that it is coupled with a simulated annealing algorithm slowing up the decrease in variance estimates during the first iterations allowing escape from the local maxima of the likelihood and convergence to a neighbourhood of the global maximum. However, the more reduced the information is in the data, the more iterations and the more chains are needed to be provided in order to improve the convergence. Of note, on the S3A scenario with altered initial conditions, which is a particularly challenging combination of error model, Hill parameter value and design, the SAEM_NM_tun performance was improved using a user-supplied Omega shrinking algorithm for fixed effects parameters without interindividual variability instead of the default gradient process (results not shown). A similar Omega shrinking approach is implemented in MONOLIX.

One noticeable aspect about the investigated approaches is the possibility for user-defined options. The main advantage is the opportunity for the modeller to adapt the search to their specific problem. This makes it necessary for the user to be educated to the different alternatives, and their need might change during the model building, or worst, their non-utilization might influence the model selection. Nevertheless, an implementation always entails default settings, chosen by the developer and enlightened by common usage. Hence the same estimation algorithm existing in distinct software represents a dissimilar approach not only because of the implementation, but also because of the defaults settings. For that reason, explored approaches were primarily run with the options set to the defaults and secondarily with settings changed or tuned, when possible.

As estimation approaches in NLMEM require the user to provide initial values for the parameters to estimate, it was decided to assess the impact of these values on their performances. For the sake of simplicity, only two scenarios were considered, with initial guesses respectively correct and reasonably altered. The real case scenario would probably lie in between both situations as the user would first explore the data at hand, as well as use prior knowledge on the compound to come up with reasonable guesses. Of note, low initial values for the variances may provide less power to the EM-like algorithms for exploring the parameter search space, however in MONOLIX the simulated annealing inflates initial values for the variances.

Models investigated in the present study were dose-response models, based on the most commonly used structure in the field, a sigmoid E_{max}. This model is fairly simple and contains a low number of parameters. The degree of nonlinearity is linked to the value of the Hill factor, which was varied across scenarios. Non-linearity is the major difficulty for ML estimation methods, for the reason mentioned earlier of no closed form solution for the integrand, whether the algorithm performs a linear approximation, a numerical integration or a stochastic approximation of the likelihood. Decreasing performance could hence be observed along the γ-increase with the additive error models, but not with the proportional error models, revealing other factors to take into account, such as the design. Models defined by ordinary differential equations represent also a challenge for estimation methods, and would perhaps result in conclusions of a different nature, but were not investigated in the present study.

17

Random effects are keys in the analysis of repeated data, allowing the modeller to quantify interindividual variability. The number of random effects that can be included in a model primarily depends on the amount of information generated under the chosen design, but also on the capacity of the algorithm to estimate them in addition to the fixed effects. The structure plays likewise a role, with considerations about the size of the variance-covariance matrices; therefore the studied structure included random effects on all parameters except one, plus one correlation.

Studies performing comparisons are bound to be limited by their tools. In the present work we used RMSE to sum-up information on both accuracy and precision which is a metric known to be sensitive to outliers. Yet, these choices provided us with the opportunity to present a readable comparison of 9 different estimation approaches across several combinations of true parameter values, error models and designs.

Drawbacks of FOCE_R experienced in this study had been described before (27). Nevertheless, previously reported (22, 24) poor performance of LAP_NM for skewed distributions was not as evident in this study, where LAP_NM mean standardized RRMSE was low for all scenarios. However parameters on which performance was the poorest were variance of random effects, which was the case here also. These studies and additional ones (23, 25) showed estimates were improved with the use of AGQ_SAS or SAEM_MLX_tun; these approaches gave good results here too. Another investigation (26) highlighted that for cases with low information content LAP_NM had problems that disappeared when SAEM_NM was used. Again, this was only retrieved for variances of random effects, but was accordingly the case for the sparse design scenarios S3A and S3P. The impact of initial conditions had not been explored before, and this study showed the lack of robustness of some otherwise accurate estimation methods. Notwithstanding, it is important to realize that none of the NONMEM nor MONOLIX methods has been tested before, as the sofware have been updated since previous publications (from versions NONMEM VI and MONOLIX 2.4, respectively). Another comparison (38) presenting EM methods as alternatives to gradient-based methods in terms of computation rates and runtimes was recently published (based on real data).

CONCLUSIONS

For standard dose-response models analyzed through mixed-effects models, differences could be identified in the performance of estimation methods available in current software. Along with the exploration of different settings, designs and initial conditions, the strength of the present investigation resides in the inclusion of a high number of estimation methods and software.

REFERENCES

1.	Sheiner BL, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm. 1981 Oct;9(5):635-51.

2.	Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm. 1980 Dec;8(6):553-71.

3.	Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm. 1983 Jun;11(3):303-19.

4.	Ette EI, Williams PJ. Population pharmacokinetics I: background, concepts, and models. The Annals of Pharmacotherapy. 2004 Oct;38(10):1702-6.

5.	Ette EI, Williams PJ. Population Pharmacokinetics II: Estimation Methods. The Annals of Pharmacotherapy. 2004;38(11):1907-15

6.	Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokin Pharmacodyn. 1977;5(5):445-79.

7.	Beal SL, Sheiner LB. Heteroscedastic Nonlinear Regression. Technometrics. 1988;30(3):327-38

8.	Dartois C, Brendel K, Comets E, Laffont CM, Laveille C, Tranchand B, et al. Overview of model-building strategies in population PK/PD analyses: 2002-2004 literature survey. Br J Clin Pharmacol. 2007 Nov;64(5):603-12.

9.	Wang Y. Derivation of various NONMEM estimation methods. J Pharmacokinet Pharmacodyn. 2007 Oct;34(5):575-93.

10.	Pillai GC, Mentre F, Steimer JL. Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science. J Pharmacokinet Pharmacodyn. 2005 Apr;32(2):161-83.

11.	Dartois C, Brendel K, Comets E, Laffont CM, Laveille C, Tranchand B, et al. Overview of model-building strategies in population PK/PD analyses: 2002–2004 literature survey. Br J Clin Pharmacol. 2007;64(5):603-12

12.	Beal S, Sheiner L. The NONMEM System. The American Statistician. 1980;34(2):118-9

13.	Wolfinger R. Laplace's approximation for nonlinear mixed models. Biometrika. 1993;80(4):791-5.

14.	Kuhn E, Lavielle M. Coupling a stochastic approximation version of EM with an MCMC procedure. ESAIM: P&S. 2004;8:115-31.

15.	Lavielle M. MONOLIX (MOdèles NOn LInéaires à effets miXtes). Orsay, France: MONOLIX

group; 2008.

16.	Beal S, Sheiner L, Boeckmann A, Bauer R. NONMEM User's Guides (1989-2009). Ellicott City, MD, USA2009.

17.	Steimer JL, Mallet A, Golmard JL, Boisvieux JF. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed-effect model. Drug Metab Rev. 1984;15(1-2):265-92.

18.	Kuhn E, Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. Computational Statistics & Data Analysis. 2005;49(4):1020-38.

19.	Roe DJ, Vonesh EF, Wolfinger RD, Mesnil F, Mallet A. Comparison of population pharmacokinetic modeling methods using simulated data: results from the population modeling workgroup. Statistics in Medicine. 1997;16(11):1241-62.

20.	Duffull SB, Kirkpatrick CM, Green B, Holford NH. Analysis of population pharmacokinetic data using NONMEM and WinBUGS. J Biopharm Stat. 2005;15(1):53-73.

21.	Bauer RJ, Guzy S, Ng C. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. AAPS J. 2007;9(1):E60-83.

22.	Plan EL, Maloney A, Troconiz IF, Karlsson MO. Performance in population models for count data, part I: maximum likelihood approximations. J Pharmacokinet Pharmacodyn. 2009 Aug;36(4):353-66.

23.	Savic R, Lavielle M. Performance in population models for count data, part II: a new SAEM algorithm. J Pharmacokinet Pharmacodyn. 2009 Aug;36(4):367-79.

24.	Jonsson S, Kjellsson MC, Karlsson MO. Estimating bias in population parameters for some models for repeated measures ordinal data using NONMEM and NLMIXED. J Pharmacokinet Pharmacodyn. 2004 Aug;31(4):299-320.

25.	Savic RM, Mentre F, Lavielle M. Implementation and Evaluation of the SAEM Algorithm for Longitudinal Ordered Categorical Data with an Illustration in Pharmacokinetics-Pharmacodynamics. AAPS J. 2010 Nov 11.

26.	Karlsson K, Plan, EL., Karlsson, MO. Performance of the LAPLACE method in repeated time-to-event modeling. In: American Conference of Pharmacometrics A, editor. ACoP; Mystic, US: American Conference of Pharmacometrics, ACoP; 2009.

27.	Girard P, Mentré, F. A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. In: Population Approach Group in Europe P, editor. Population Approach Group in Europe, PAGE; Pamplona, Spain. Pamplona, Spain: Population Approach Group in Europe, PAGE; 2005.

28.	Lindstrom ML, Bates DM. Nonlinear mixed effects models for repeated measures data. Biometrics. 1990 Sep;46(3):673-87.

29.	R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2008.

30.	SAS Institute Inc. SAS. 9.1.2 ed. Cary, NC, USA2004.

31.	Pinheiro JC, Bates DM. Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model. J Computational Graphical Statistics. 1995;4(1):12-35.

32.	Abramowitz M, Stegun IA. Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables: Dover Publications; 1965.

33.	Pinheiro JC, Bates DM, . Mixed-Effects Models in S and S-PLUS. London: Springer; 2000.

34.	Lindstrom MJ, Bates DM. Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data. J American Statistical Ass. 1988;83(4):1014-22.

35.	Lavielle M, Mentre F. Estimation of population pharmacokinetic parameters of saquinavir in HIV patients with the MONOLIX software. Journal of pharmacokinetics and pharmacodynamics. 2007 Apr;34(2):229-49.

36.	Delyon B, Lavielle M, Moulines E. Convergence of a stochastic approximation version of the EM algorithm. Annals of Statistics. 1999;27 (1):94-128.

37.	Lindbom L, Pihlgren P, Jonsson EN. PsN-Toolkit--a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. Comput Methods Programs Biomed. 2005;79(3):241-57.

38.	Chan P, Jacqmin P, Lavielle M, McFadyen L, Weatherley B. The use of the SAEM algorithm in MONOLIX software for estimation of population pharmacokinetic-pharmacodynamic-viral dynamics parameters of maraviroc in asymptomatic HIV subjects. J Pharmacokin Pharmacodyn.1-21.

Table 1: True initial conditions are the parameter values used for the simulation of 8 scenarios constructed with 3 different Hill factor (γ) values and 2 different residual error models: additive (A) and proportional (P). True and altered initial conditions were used for the estimation of the simulated datasets.

Parameters	E_0	E_{max}	ED_{50}	γ			$\omega^2(E_0)$	$\omega^2(E_{max})$	$Cov(E_{max}, ED_{50})$	$\omega^2(ED_{50})$	σ^2	
True initial conditions	5	30	500	1	2	3	0.090	0.490	0.245	0.490	A: 4	P: 0.010
Altered initial conditions	10	60	1000	1	1	1	0.100	0.100	0.010	0.100	A: 1	P: 0.0625

Table 2: Approach settings changed from defaults

Approach	Algorithm	Software	Settings
FOCE_NM	FOCE	NONMEM 7.1.0	A1, A2, A3: MAXEVALS=9999 P1, P2, P3: INTERACTION MAXEVALS=9999
FOCE_R	FOCE	R 2.9.1	-
LAP_NM	LAPLACE	NONMEM 7.1.0	A1, A2, A3: MAXEVALS=9999 P1, P2, P3: INTERACTION MAXEVALS=9999
LAP_SAS	LAPLACE	SAS 9.2	QPOINTS=1 FTOL=1E-8 XTOL=1E-8 TECH=QUANEW/DBLDOG EBSTEPS=300 EBSUBSTEPS=300 EBSSFRAC=0.2 EBTOL=1E-6 INSTEP=1E-1
AGQ_SAS	AGQ	SAS 9.2	QPOINTS=9 FTOL=1E-8 XTOL=1E-8 TECH=QUANEW/DBLDOG EBSTEPS=300 EBSUBSTEPS=300 EBSSFRAC=0.2 EBTOL=1E-6 INSTEP=1E-1
SAEM_NM	SAEM	NONMEM 7.1.0	INTERACTION CTYPE=3
SAEM_NM_tun	SAEM	NONMEM 7.1.0	INTERACTION CTYPE=3 NBURN=1000 NITER=200 ISAMPLE=5 IACCEPT=0.3 CINTERVAL=25 NOABORT
SAEM_MLX	SAEM	MONOLIX 3.1	-
SAEM_MLX_tun	SAEM	MONOLIX 3.1	K1=500 nmc=5

Figure 1: Individual response versus dose profiles from a typical dataset simulated using 6 of the 8 dose-response profiles: rich design, additional error model with Hill parameter = 1, 2 and 3: R1A, R2A, R3A and proportional error model with Hill parameter = 1, 2 and 3: R1P, R2P, R3P. On the x-axis are displayed the four doses considered.

Figure 2: Percentage of completion and number of instructions (in billions) obtained with the 9 investigated approaches for the true initial conditions. The barchart represents the median, and the arrows link the minimum to the maximum value of the range.

Figure 3: Relative estimation error (RER) for the parameter ED_{50} (3a) and its variance (3b), for the 8 scenarios R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P referring to 2 simulation designs (R for rich and S for sparse), 3 Hill factor values (1, 2, 3), and 2 residual error models (A for additive and P for proportional), with the estimation from true initial conditions and altered initial conditions. The boxplot represents the median (middle bar) and the interquartile range (box limits), with points for the mean (black) and the outliers (grey).

Figure 4: Standardized RRMSE of the 9 population parameters for the 8 scenarios R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P, and 2 initial conditions: true and altered. The following colour code was used: FOCE_NM = red, FOCE_R = pink, LAP_NM = orange, LAP_SAS = light green, AGQ_SAS = dark green, SAEM_NM = dark blue, SAEM_NM_tun = light blue, SAEM_MLX = dark violet, SAEM_MLX_tun = light violet.

Figure 5: Strip chart of the mean standardized RRMSE obtained with each approach for the 8 scenarios R1A, R2A, R3A, R1P, R2P, R3P, S3A, and S3P, and 2 initial conditions: true and altered, on a semi-log scale. The colour code used is described in the Figure 4 legend. The star symbol (*) represents the S3A estimate from SAEM_NM_tun that is above 45 units. The dashed line is drawn at the value 1.5 used for description purposes in the results section.

27

	R1A	γ	R2A		R3A		S3A

ED₅₀

E₀

E_max

ωE_max

Cov

ωE₀

ωED₅₀

σ

R1P

R2P

R3P

S3P

True initial conditions

R1A

R2A

R3A

S3A

R1P

R2P

R3P

S3P
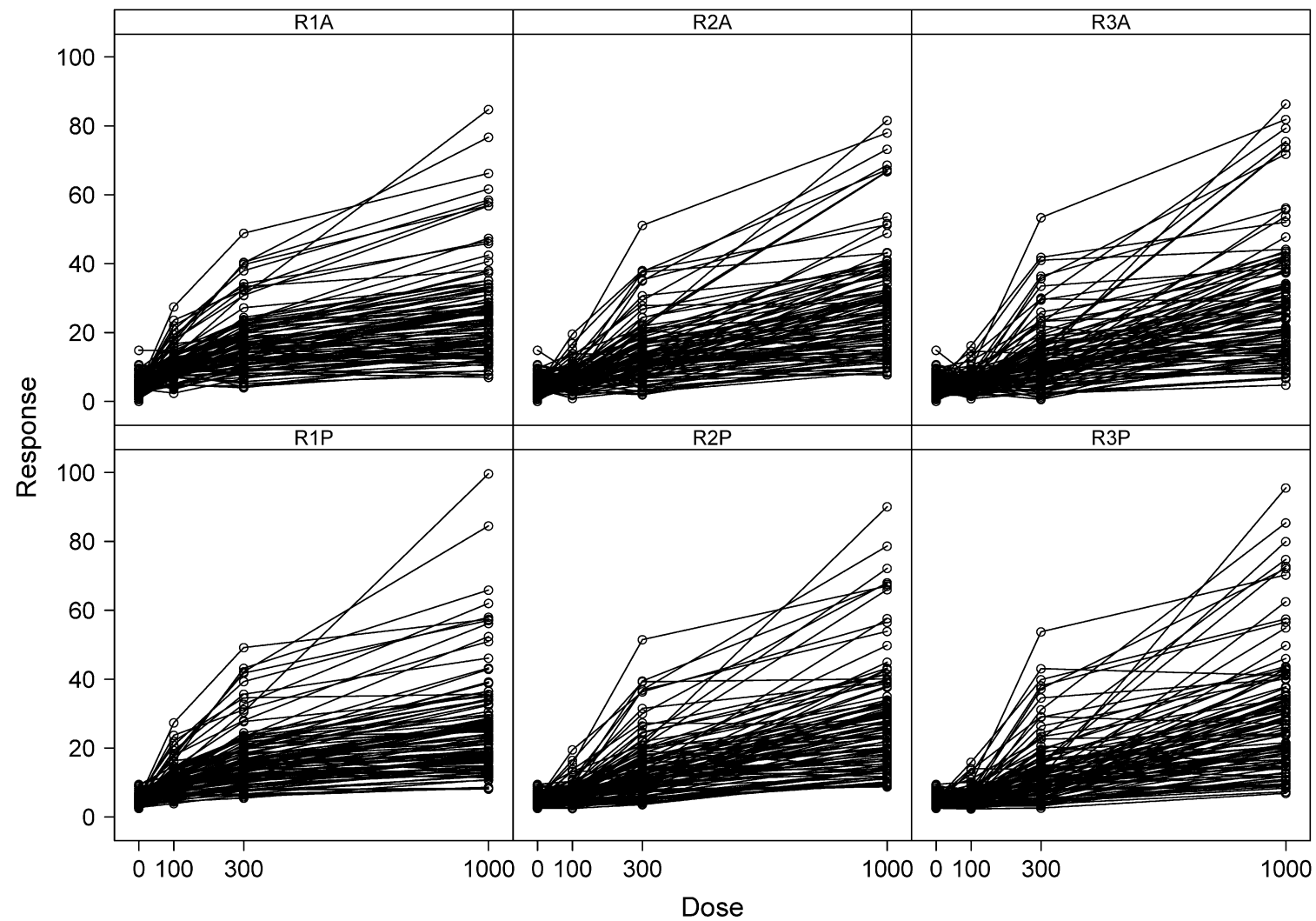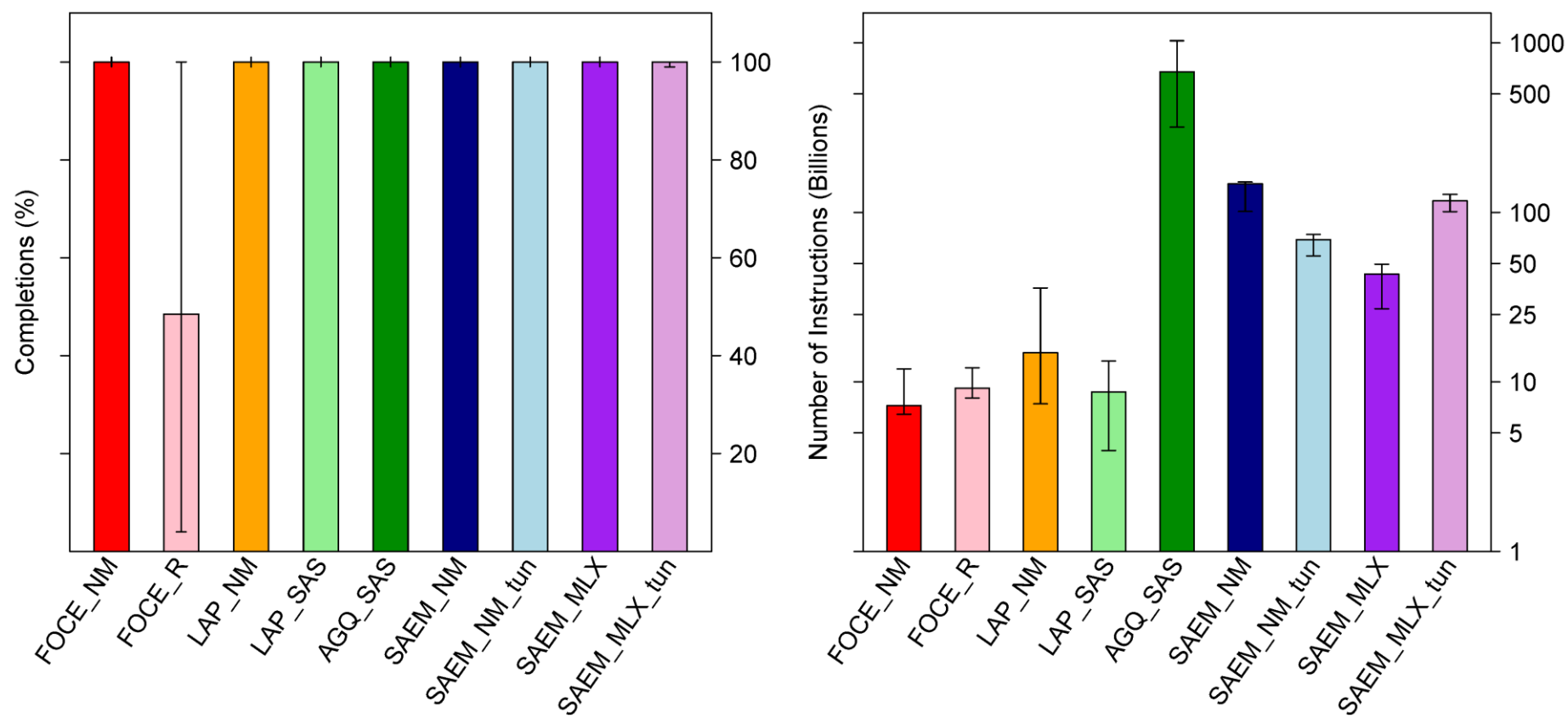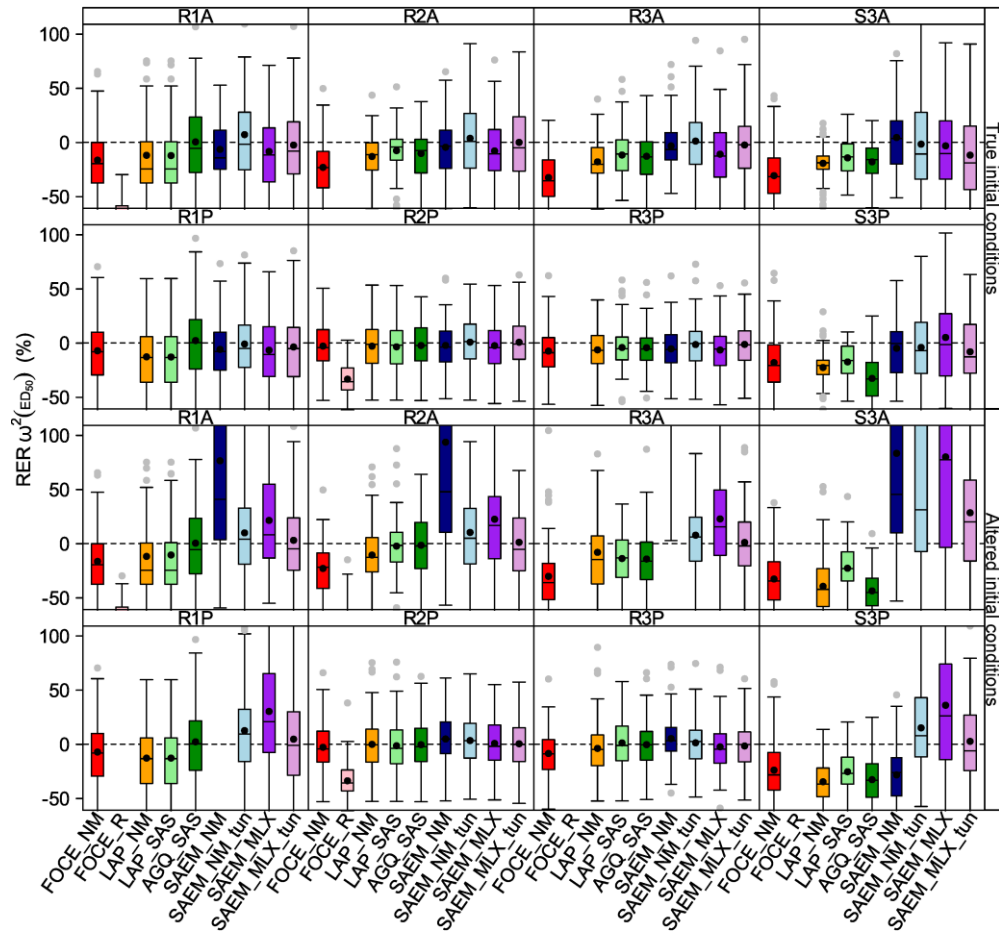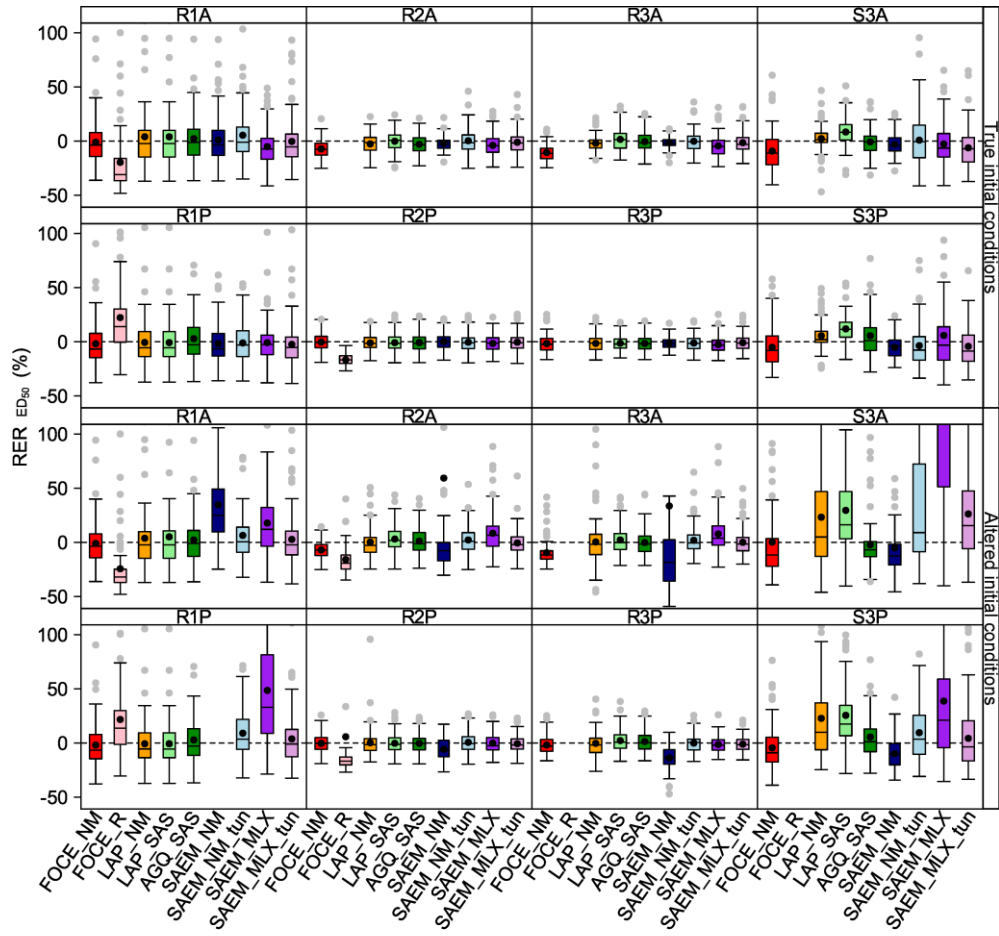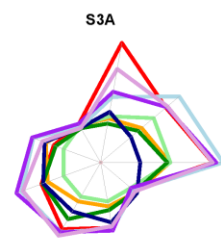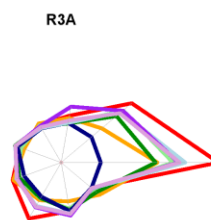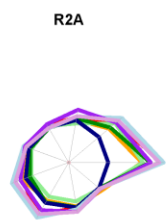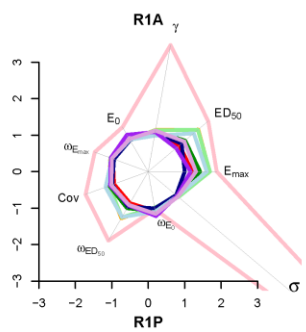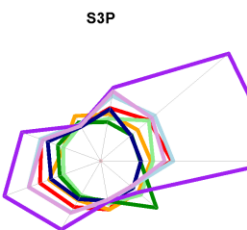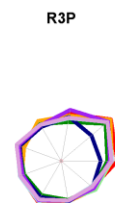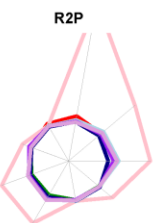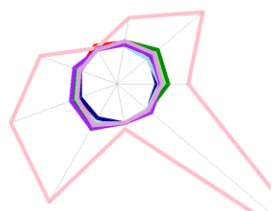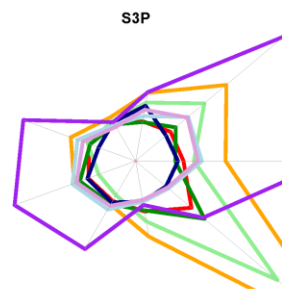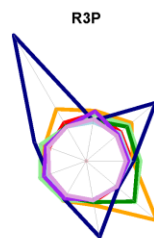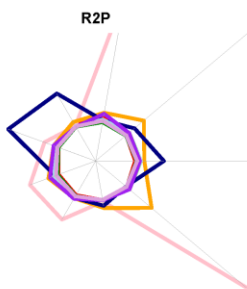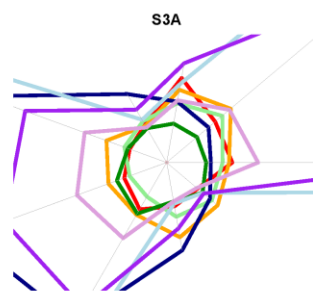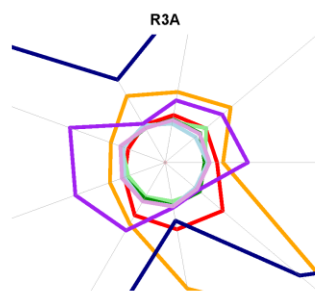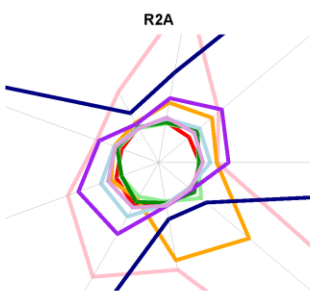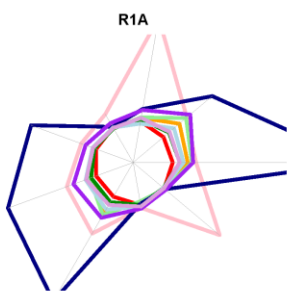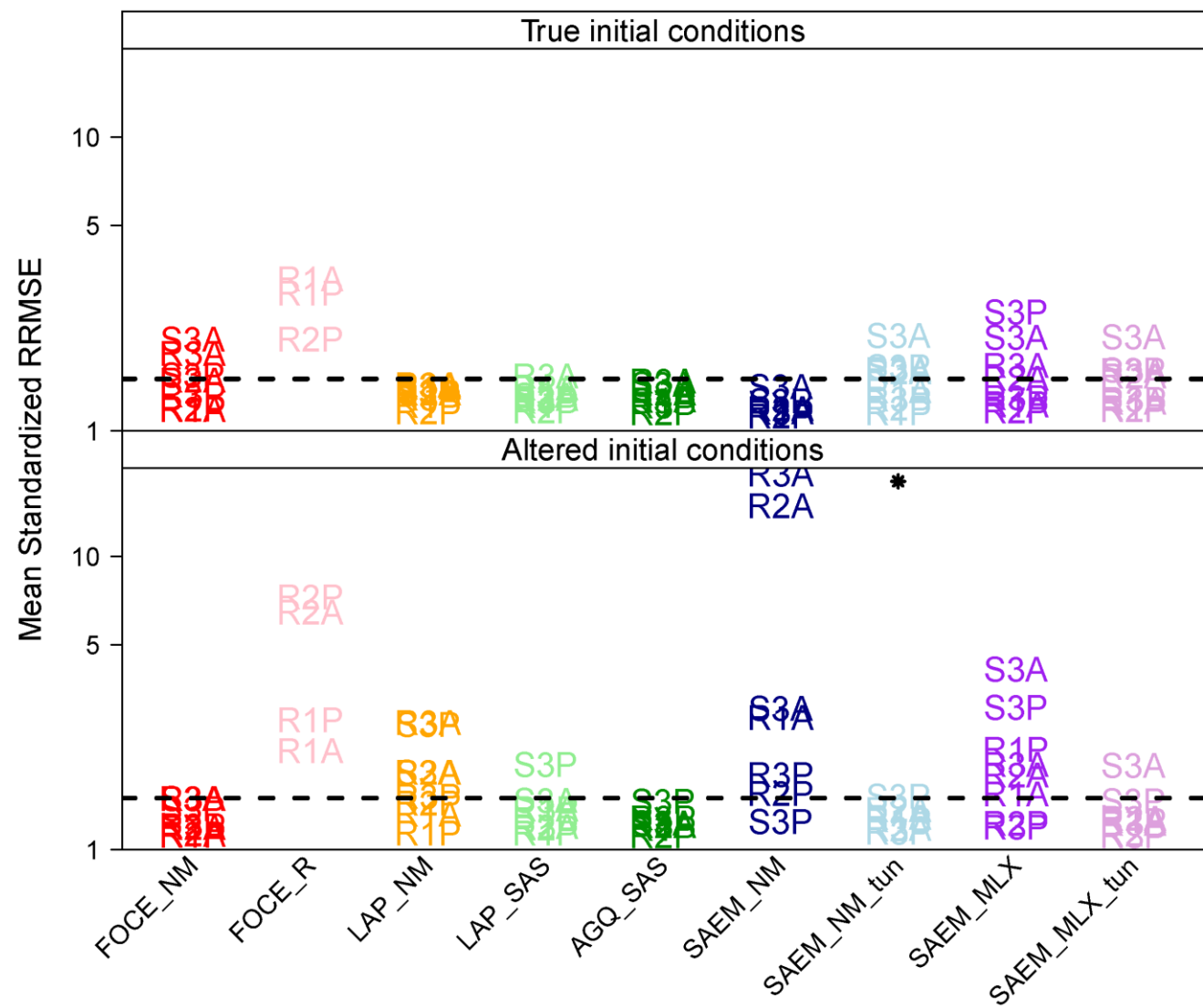
Altered initial conditions

Tables of relative bias (RBias) and relative RMSE (RRMSE) obtained with the 9 investigated approaches for the parameters of the explored scenarios (in %)

(* based on less than 50% convergence)

True initial conditions

Parameter		E_0		E_{max}		ED_{50}		γ		$\omega^2(E_0)$		$\omega^2(E_{max})$		$Cov(E_{max},ED_{50})$		$\omega^2(ED_{50})$		σ^2	
Approach	Scenario	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE
FOCE_NM	R1A	1.63	5.13	-0.57	10.99	0.03	20.15	-0.52	6.37	7.38	32.35	-2.49	18.16	-9.83	38.46	-15.46	33.45	-0.89	11.96
	R2A	1.74	4.81	-6.95	11.84	-6.17	10.28	0.11	5.62	6.00	29.21	-1.57	18.42	-24.69	39.95	-22.14	32.06	-2.89	12.29
	R3A	1.64	4.74	-13.71	16.53	-9.55	12.22	0.39	7.66	5.87	27.89	2.24	19.70	-42.16	51.76	-31.55	38.17	-2.09	12.53
	R1P	-1.48	3.26	-1.78	10.90	-0.89	19.69	-0.29	5.43	0.34	15.06	-0.99	16.78	-3.45	38.13	-6.24	32.32	-2.34	14.39
	R2P	-1.11	3.11	-1.29	8.86	0.70	8.08	-1.76	3.54	0.23	14.80	-1.30	16.97	-5.84	31.32	-1.90	20.92	-4.58	13.84
	R3P	-0.93	3.00	-4.05	10.75	-0.63	8.06	-1.77	3.95	0.08	14.34	-0.97	17.19	-14.72	35.46	-6.54	21.27	-4.52	13.61
	S3A	3.52	7.33	-17.30	22.97	-8.14	19.28	1.41	34.15	15.53	60.54	3.59	27.87	-32.83	57.49	-29.76	40.68	-3.55	22.93
	S3P	-0.22	3.74	-16.06	21.97	-4.18	17.85	-5.19	12.61	6.37	23.48	9.54	34.88	-19.34	59.31	-17.16	31.49	-18.03	40.40
FOCE_R	R1A	4.96	6.98	-10.74	17.10	-18.57	38.70	15.82	20.91	-0.67	30.75	-21.33	27.50	-66.19	70.34	-70.68	73.06	35.32	127.65
	R2A*	8.54	10.05	-14.65	18.03	-15.70	20.92	18.47	26.32	2.03	39.61	-28.22	32.73	-88.36	89.42	-78.93	82.41	154.19	413.29
	R3A*	-2.14	13.02	-1.00	12.60	-2.87	14.66	0.94	27.16	-34.89	57.34	-17.35	21.90	-74.52	78.89	-38.29	54.30	126.92	260.89

	R1P	-0.97	3.07	6.71	17.45	23.32	48.30	-6.23	8.60	-8.21	17.35	-30.02	34.28	-99.49	102.36	-97.68	97.73	94.80	97.36
	R2P	0.19	2.97	-13.86	15.57	-15.68	16.45	12.02	12.62	-2.85	14.74	-15.82	20.80	-51.26	54.71	-32.47	35.48	11.41	19.78
	R3P*	0.47	3.26	-16.19	17.98	-17.11	17.97	18.97	20.34	-1.32	14.55	-17.94	24.05	-68.42	70.72	-35.67	39.09	13.34	24.66
	S3A*	6.22	9.06	-5.38	30.51	-21.05	35.48	677.60	945.82	17.41	66.82	-28.44	40.55	-101.02	105.63	-59.24	67.61	2.73	32.73
	S3P*	1.01	4.99	-15.53	30.94	-33.27	42.94	731.17	1075.08	-2.95	23.51	-22.89	37.89	-103.73	109.53	-55.19	61.62	11.01	118.97
LAP_NM	**R1A**	-0.47	5.03	2.46	15.34	5.08	32.25	0.58	6.89	-0.02	29.61	-7.30	19.46	-15.72	44.51	-10.99	49.31	0.86	12.23
	R2A	-0.22	4.56	-1.18	9.30	-1.58	8.74	0.99	5.26	-0.04	28.31	-9.06	19.34	-21.65	39.61	-12.14	23.48	-1.53	13.48
	R3A	-1.08	4.92	-5.08	10.11	-1.00	7.03	-1.77	7.26	0.34	22.36	-5.17	17.62	-37.61	51.00	-17.02	25.62	1.71	15.72
	R1P	-0.26	2.93	0.53	11.56	0.33	20.85	0.95	5.52	-1.07	15.05	-5.69	17.40	-14.89	38.33	-11.84	33.56	-1.37	15.48
	R2P	-0.27	2.90	1.18	9.05	-0.11	7.56	0.46	3.18	-0.61	14.66	-4.07	17.29	-3.91	32.64	-1.98	20.57	-5.15	13.97
	R3P	-0.33	3.20	-0.27	9.74	-0.53	8.00	0.23	3.80	-1.01	14.34	-2.42	15.72	-10.16	34.28	-5.38	19.31	-4.44	14.57
	S3A	-1.04	7.11	-3.72	14.19	3.34	12.11	-9.11	12.55	3.68	40.75	0.62	32.50	-37.50	51.45	-18.37	23.76	3.69	25.13
	S3P	0.41	4.72	-1.77	15.57	6.52	13.60	-5.66	11.21	3.22	23.50	-0.33	21.80	-43.69	50.37	-21.66	26.83	8.86	50.11
LAP_SAS	**R1A**	-0.47	5.03	2.49	15.38	5.01	32.04	0.60	6.86	-0.25	29.45	-7.36	19.46	-16.00	43.75	-11.25	48.47	0.88	12.25
	R2A	-0.52	4.79	1.27	9.45	0.89	10.10	-0.25	4.79	-1.71	25.12	-8.15	17.78	-14.78	31.65	-6.67	21.06	-0.92	12.68
	R3A	-0.21	4.70	-0.28	11.75	2.61	10.70	-2.00	6.37	-1.06	26.56	-10.39	19.13	-32.57	45.96	-10.73	24.43	0.17	13.48
	R1P	-0.27	2.95	0.49	11.55	0.23	20.83	1.00	5.54	-1.10	15.06	-5.72	17.44	-15.00	38.47	-11.97	33.45	-1.37	15.48
	R2P	-0.18	2.94	1.19	9.23	0.17	8.32	0.42	3.16	-0.83	14.66	-5.33	17.30	-5.99	31.44	-2.59	19.83	-4.08	13.98
	R3P	-0.34	2.83	0.44	9.60	-0.40	7.23	0.26	3.61	-0.77	13.85	-1.35	17.76	-4.08	34.40	-3.23	19.50	-4.72	14.81
	S3A	-1.12	6.95	2.58	14.31	9.42	15.61	-10.39	13.19	1.41	36.17	1.55	21.97	-29.74	36.00	-13.36	20.88	4.71	21.49

Group	Row																		
AGQ_SAS	S3P	0.45	3.75	3.79	16.22	12.86	17.73	-6.97	10.11	2.09	20.64	4.91	22.05	-28.65	36.24	-16.53	23.39	7.00	53.67
	R1A	-0.55	5.00	2.40	12.94	3.26	23.81	0.16	6.66	1.44	32.20	-0.96	19.05	2.31	43.67	1.37	38.78	-0.68	11.70
	R2A	-0.46	4.51	-0.99	10.37	-1.95	9.02	1.11	5.43	0.03	28.85	-4.90	18.49	-13.55	33.20	-9.33	24.53	0.35	12.42
	R3A	0.12	4.51	-1.67	9.92	0.56	9.57	0.36	6.38	-3.13	27.35	-15.44	21.97	-34.94	44.47	-11.91	24.26	4.28	14.47
	R1P	-0.23	2.93	2.61	12.55	3.84	22.86	0.07	5.55	-0.32	14.99	-0.30	17.47	3.08	42.84	3.18	37.17	-1.31	14.21
	R2P	-0.17	2.94	1.52	9.17	0.24	8.31	0.30	3.12	-0.65	14.39	-2.48	17.24	-1.98	30.26	-1.35	19.53	-1.80	13.74
	R3P	-0.21	2.95	0.41	8.86	-0.56	7.21	0.54	3.26	-0.78	13.86	-3.44	17.98	-5.96	32.17	-3.40	17.93	-1.00	13.79
	S3A	-2.99	12.24	0.31	11.34	-2.22	9.27	0.26	4.05	-13.01	25.19	-32.48	40.06	-17.04	24.74	-6.46	19.36	-99.61	99.61
	S3P	-1.82	19.45	6.54	22.53	-4.63	12.58	0.35	3.94	-18.01	33.80	-65.20	75.15	-31.74	38.39	-3.42	23.98	57.63	90.29
SAEM_NM	R1A	-0.14	4.95	1.70	9.10	2.11	21.52	1.06	6.43	-0.63	30.16	-1.49	17.60	-6.80	40.47	-5.55	36.99	-1.93	11.32
	R2A	-0.33	4.59	-1.67	5.03	-1.03	6.56	1.36	5.22	0.75	29.06	-1.01	19.12	-5.00	36.72	-3.59	24.85	-2.87	12.32
	R3A	-0.40	4.32	-1.53	4.15	-0.40	5.24	1.26	6.48	0.66	28.06	-0.26	20.72	-3.68	39.52	-2.26	23.27	-2.48	12.40
	R1P	-0.60	3.00	0.19	9.71	-0.74	18.04	1.35	5.31	1.01	15.16	-1.43	16.51	-5.98	35.69	-4.96	28.34	-3.62	13.97
	R2P	-0.88	2.87	-1.80	7.54	1.09	7.51	0.72	3.10	1.13	14.78	-1.35	16.71	-3.48	31.37	-1.29	20.59	-5.67	14.12
	R3P	-1.11	2.74	-4.26	7.58	-0.45	4.92	1.40	3.70	1.01	14.40	-1.82	17.56	-9.51	33.60	-4.44	18.87	-5.36	13.52
	S3A	-1.91	6.66	-2.20	8.12	-2.02	8.93	0.67	14.45	24.83	55.69	9.07	33.25	13.59	55.81	5.40	29.88	-17.51	25.81
	S3P	-0.52	3.60	-8.09	12.46	-4.23	11.15	3.90	12.33	3.18	19.29	4.41	31.15	-7.71	48.44	-3.94	26.12	-32.42	43.16
SAEM_NM_tun	R1A	-0.11	5.02	3.77	14.61	6.45	29.38	-0.66	5.94	0.67	32.81	0.91	18.72	7.59	47.92	8.07	48.14	-0.63	11.79
	R2A	0.24	4.49	2.79	12.09	1.36	11.51	0.54	5.69	-0.51	30.20	0.77	21.68	6.77	48.50	4.68	33.93	-0.87	12.61
	R3A	0.08	4.36	2.78	13.04	0.87	10.40	0.55	6.50	-0.66	29.43	0.95	21.84	5.35	47.65	2.13	29.55	-0.38	12.66

	R1P	-0.20	2.97	0.97	10.48	-0.11	17.06	0.87	5.05	0.72	15.13	-0.43	17.11	-0.47	39.34	-0.01	32.32	-1.70	13.81
	R2P	-0.14	2.92	1.98	9.54	0.54	8.64	0.18	3.24	0.56	14.74	-0.88	17.26	1.83	32.56	1.72	21.70	-2.09	13.85
	R3P	-0.14	2.89	1.26	9.79	-0.17	7.50	0.61	3.58	0.46	14.16	-0.66	18.66	-0.18	35.97	-0.48	20.96	-2.16	11.79
	S3A	-0.89	6.76	2.64	25.12	2.04	23.22	1.46	18.73	11.57	61.74	4.01	39.87	1.06	76.93	-1.46	42.43	-5.56	23.73
	S3P	-0.29	3.84	--2.02	22.47	-2.62	19.92	4.44	15.50	1.27	19.29	3.02	35.38	-4.08	70.33	-3.11	35.04	-17.02	41.06
SAEM_MLX	R1A	-2.06	5.78	-0.59	10.56	-4.07	18.01	1.99	6.54	13.57	36.46	-3.50	19.29	-5.49	42.68	-7.45	35.85	-0.12	11.71
	R2A	-1.18	5.01	-1.02	10.45	-2.91	10.29	2.22	6.50	8.24	31.83	-4.07	21.02	-8.17	41.25	-6.95	29.33	0.14	12.78
	R3A	-1.02	4.60	-2.54	12.32	-3.34	10.68	4.06	9.14	8.29	30.44	-4.59	21.71	-14.26	46.70	-9.95	30.44	-0.29	12.55
	R1P	-0.23	2.92	0.67	10.69	0.16	19.35	0.85	5.17	-0.67	15.06	-2.60	19.17	-5.98	44.60	-5.68	36.92	3.36	15.48
	R2P	-0.17	2.92	0.83	8.58	-0.57	8.00	0.81	3.13	-0.66	14.35	-3.63	17.86	-3.16	32.41	-1.50	21.87	0.92	13.91
	R3P	-0.15	2.87	-0.96	9.34	-1.63	7.61	1.64	4.28	-0.87	14.10	-4.62	18.18	-9.95	33.47	-5.54	20.49	1.84	14.08
	S3A	-2.46	7.41	0.58	23.75	-1.69	19.48	4.08	20.19	29.33	63.33	7.75	40.63	6.29	81.81	-2.28	42.53	-8.02	21.48
	S3P	-0.36	3.73	11.02	55.84	6.86	47.15	2.69	17.51	-1.74	19.71	11.86	46.28	20.70	94.21	6.06	46.89	3.72	49.86
SAEM_MLX_tun	R1A	-0.82	5.11	1.29	11.91	0.73	23.18	0.96	6.77	2.44	34.61	-1.59	18.66	0.01	41.45	-1.54	35.32	-0.84	12.06
	R2A	-0.66	4.54	1.87	11.54	-0.07	10.81	0.79	5.70	1.56	31.83	-0.78	20.46	3.07	45.56	0.94	31.82	-0.81	12.68
	R3A	-0.67	4.46	1.44	12.44	-0.45	10.16	1.15	6.86	1.99	29.58	-0.69	21.79	0.48	47.55	-1.49	30.01	-0.84	12.49
	R1P	-0.23	2.93	0.18	10.99	-1.43	20.58	1.66	5.58	-0.27	14.96	-1.94	16.67	-3.16	38.18	-2.72	34.39	-1.70	14.33
	R2P	-0.21	2.93	2.15	9.24	0.56	8.31	0.24	3.08	-0.43	14.54	-0.98	17.02	3.00	33.31	1.64	22.25	-2.34	13.85
	R3P	-0.19	2.91	1.81	9.65	0.10	7.52	0.54	3.63	-0.57	14.05	-0.81	18.93	1.36	35.55	-0.36	20.16	-1.91	13.34
	S3A	-0.52	6.60	-4.31	23.01	-5.20	19.73	9.81	26.69	6.04	59.23	1.80	34.56	-10.17	75.57	-10.86	44.62	-3.18	24.42

| S3P | -0.26 | 3.69 | -3.07 | 19.79 | -3.23 | 18.85 | 5.21 | 16.58 | -1.08 | 18.82 | 0.08 | 34.03 | -10.32 | 70.22 | -7.14 | 35.41 | -8.14 | 40.60 |

Altered initial conditions

Parameter		E0		Emax		ED$_{50}$		γ		ω2(E0)		ω2(Emax)		Cov(Emax,ED$_{50}$)		ω2(ED$_{50}$)		σ^2	
Approach	Scenario	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE	RBias	RRMSE
FOCE_NM	R1A	1.63	5.13	-0.58	10.99	0.03	20.15	-0.52	6.37	7.38	32.34	-2.49	18.16	-9.83	38.45	-15.46	33.46	-0.89	11.96
	R2A	1.78	4.83	-6.87	11.41	-5.97	10.06	0.07	5.54	4.62	30.68	-2.01	18.91	-25.52	41.17	-21.97	31.49	-2.34	13.72
	R3A	1.96	4.79	-12.43	16.81	-8.55	13.84	-0.40	8.44	-9.43	44.90	1.54	20.76	-40.11	51.28	-29.23	41.32	5.06	23.16
	R1P	-1.48	3.26	-1.77	10.90	-0.88	19.69	-0.30	5.42	0.34	15.06	-0.99	16.78	-3.44	38.13	-6.23	32.32	-2.33	14.39
	R2P	-1.11	3.11	-1.30	8.86	0.69	8.42	-1.75	3.53	0.21	14.61	-1.38	16.60	-5.57	31.44	-1.86	21.42	-4.52	13.83
	R3P	-1.04	3.24	-4.77	11.04	-0.85	8.31	-1.58	4.07	0.17	14.64	-0.51	18.56	-16.44	39.32	-7.57	22.18	-3.88	13.24
	S3A	3.65	7.44	-9.31	42.25	1.22	42.62	-0.46	31.62	28.61	76.36	-4.71	32.37	-44.01	65.80	-31.73	42.56	-4.69	23.95
	S3P	0.09	3.91	-16.15	23.52	-3.44	19.55	-3.74	12.84	6.44	24.51	4.00	34.50	-32.25	63.83	-22.93	35.53	-2.81	73.12
FOCE_R	R1A	4.72	7.07	-12.82	17.49	-23.35	35.69	18.36	22.33	-4.00	34.44	-20.29	25.49	-64.71	68.87	-67.53	69.46	20.23	33.43
	R2A	3.88	9.61	-11.59	16.41	-14.86	19.91	11.95	22.78	-26.19	79.81	-26.75	31.55	-86.84	88.13	-80.61	83.50	187.97	397.65
	R3A*	-3.04	8.94	-1.31	21.84	-1.07	27.15	-7.75	26.73	-83.34	98.60	-37.80	49.80	-84.06	86.33	-77.98	86.15	691.81	1112.98
	R1P	-0.84	3.07	6.31	17.97	22.74	48.93	-5.83	8.60	-8.64	17.31	-30.72	34.97	-100.41	103.20	-97.42	97.50	94.51	97.42
	R2P	0.26	2.98	1.77	151.33	6.79	217.18	11.30	13.29	-3.50	15.14	-16.98	22.95	-51.65	55.73	-32.80	36.63	20.49	67.16
	R3P*	0.53	3.38	-19.12	20.54	-19.40	20.20	19.81	21.18	-1.95	21.04	-13.93	23.77	-66.40	68.55	-38.89	41.70	32.06	104.63
	S3A*	6.22	9.06	-5.38	30.51	-21.05	35.48	677.60	945.82	17.41	66.82	-28.44	40.55	-101.02	105.63	-59.24	67.61	2.73	32.73
	S3P*	1.01	4.99	-15.53	30.94	-33.27	42.94	731.17	1075.08	-2.95	23.51	-22.89	37.89	-103.73	109.53	-55.19	61.62	11.01	118.97

LAP_NM	R1A	-0.46	5.04	2.45	15.18	4.82	30.76	0.57	6.89	-0.05	29.57	-7.30	19.33	-15.94	43.35	-10.85	49.63	0.78	12.27
	R2A	1.87	5.44	0.20	16.14	1.23	17.72	0.12	8.12	-54.95	72.69	-14.45	22.82	-29.58	46.14	-9.63	27.02	22.51	36.47
	R3A	2.56	8.55	-4.27	18.76	1.40	22.35	-6.57	12.53	-73.81	84.79	-11.09	30.42	-45.75	69.37	-7.11	48.06	52.26	82.30
	R1P	-0.25	2.93	0.52	11.56	0.28	20.83	0.97	5.52	-1.10	15.06	-5.68	17.38	-14.95	38.40	-11.87	33.53	-1.37	15.48
	R2P	-0.09	3.31	2.20	10.90	1.20	13.27	0.09	3.83	-0.55	17.55	-5.45	18.85	-6.96	40.12	0.87	23.31	-0.35	24.88
	R3P	0.26	4.39	0.66	11.29	0.69	10.96	-0.74	4.33	-0.99	15.83	-5.65	20.82	-16.73	41.72	-2.88	23.77	4.32	30.93
	S3A	-1.48	7.22	5.96	40.16	24.23	56.27	-12.45	27.20	60.31	130.05	7.94	52.58	-65.45	90.29	-38.47	47.62	10.24	39.45
	S3P	1.60	4.72	12.50	44.25	23.79	50.97	-12.11	22.05	-9.65	36.82	-5.13	48.06	-62.17	82.82	-33.69	38.80	202.67	273.31
LAP_SAS	R1A	-0.45	5.03	2.91	16.34	6.13	35.18	0.50	6.98	-0.07	29.50	-7.17	19.65	-14.84	44.99	-9.53	51.43	0.83	12.26
	R2A	-0.01	4.68	3.73	11.01	4.11	12.92	-0.10	5.29	-3.70	28.94	-9.97	20.66	-13.22	36.58	-1.34	24.99	3.39	17.13
	R3A	-0.64	4.67	0.57	12.93	3.28	13.99	-2.80	7.23	0.72	26.14	-10.64	24.12	-32.81	46.54	-12.80	26.76	-1.45	13.99
	R1P	-0.26	2.93	0.53	11.56	0.32	20.84	0.96	5.52	-1.09	15.06	-5.67	17.37	-14.90	38.35	-11.84	33.56	-1.36	15.48
	R2P	-0.22	2.93	2.01	10.14	0.76	9.24	0.28	3.18	-0.79	14.54	-5.22	17.99	-3.41	35.32	-0.43	22.77	-3.89	15.05
	R3P	-0.11	2.90	4.60	13.16	3.23	10.31	-0.95	3.99	-0.10	14.52	-4.05	22.75	0.57	44.16	2.20	23.59	1.53	21.41
	S3A	-2.18	8.12	19.53	36.01	30.58	48.86	-17.24	22.16	46.03	94.45	9.77	36.94	-27.82	57.47	-21.78	31.10	13.51	34.97
	S3P	0.41	4.31	15.96	28.63	26.55	38.58	-15.40	18.89	-7.11	29.38	0.32	35.53	-32.54	49.92	-24.43	30.38	104.20	186.10
AGQ_SAS	R1A	-0.55	5.00	2.40	12.94	3.26	23.81	0.16	6.66	1.44	32.20	-0.96	19.05	2.31	43.67	1.37	38.78	-0.68	11.70
	R2A	-0.05	4.84	2.43	11.31	2.20	12.83	0.47	5.45	-1.71	30.06	-7.01	21.30	-9.10	35.94	-0.56	28.76	1.74	14.32
	R3A	0.02	4.40	-1.58	12.82	0.80	11.89	0.55	7.66	-0.94	28.96	-15.25	24.56	-37.08	51.40	-13.26	28.15	3.32	13.90
	R1P	-0.23	2.93	2.61	12.55	3.84	22.86	0.07	5.55	-0.32	14.99	-0.30	17.47	3.08	42.83	3.18	37.17	-1.31	14.21

SAEM_NM	**R2P**	-0.19	2.92	1.97	9.19	0.70	8.35	0.16	3.09	-0.46	14.63	-2.63	16.72	-0.33	32.26	0.52	21.91	-1.81	13.70
	R3P	-0.02	2.97	3.77	12.44	2.41	9.86	-0.23	3.50	-0.48	14.78	-6.10	19.27	-3.42	36.21	0.57	21.55	5.88	20.99
	S3A	-0.05	6.73	-8.40	25.45	-1.23	26.59	1.01	14.56	10.03	67.90	-17.26	33.84	-71.92	77.29	-42.69	46.51	8.89	27.75
	S3P	0.35	3.94	-1.82	19.45	6.54	22.53	-4.63	12.58	-3.42	23.98	-18.01	33.80	-65.20	75.15	-31.74	38.39	57.63	90.29
	R1A	-0.22	5.20	59.29	62.94	35.68	52.65	-4.67	8.26	8.01	31.20	38.20	50.00	77.94	129.89	77.44	135.55	-5.63	12.32
	R2A	-3.63	6.74	41.77	244.10	60.37	653.79	-6.84	12.23	20.14	42.11	49.78	140.05	142.50	367.92	94.60	239.33	-1.37	19.24
	R3A	-8.60	10.65	113.94	309.40	34.62	273.44	-28.84	32.14	6.30	39.11	436.02	683.44	962.02	1443.49	510.96	755.94	34.12	54.28
	R1P*	5.53	10.62	72.60	121.03	-15.84	37.16	-8.08	25.06	13.63	27.76	4589.61	18105.33	-4456.81	17451.48	1128.65	4210.09	14.02	51.81
	R2P	3.76	5.73	4.97	15.30	-4.74	10.73	-0.08	3.23	3.55	16.71	8.38	39.15	8.34	34.41	5.81	23.12	-4.50	13.47
	R3P	7.57	10.65	-0.72	11.42	-12.79	16.79	0.64	3.46	11.31	27.21	3.63	25.96	-0.49	36.41	6.17	22.75	-4.17	13.94
	S3A	-8.98	13.46	11.81	28.44	-4.16	36.74	-13.00	22.77	107.50	152.96	83.26	158.77	175.66	311.97	84.32	145.58	-26.92	33.45
	S3P	-21.77	4.18	-36.70	20.81	-31.61	17.01	-18.04	17.88	227.99	19.67	2690.59	27.43	-496.72	65.03	-12.60	36.78	-11.33	39.69
SAEM_NM_tun	**R1A**	-0.46	5.11	4.47	14.02	7.40	27.11	-1.14	6.01	1.77	33.44	1.65	19.95	10.81	49.32	10.87	46.02	-0.87	11.73
	R2A	-0.81	4.71	5.30	14.39	3.25	13.54	-0.86	5.94	2.51	31.05	3.82	22.66	17.14	55.94	11.22	39.52	-0.83	12.21
	R3A	-1.07	4.77	4.60	13.82	1.81	10.33	-1.33	6.92	3.17	29.27	2.90	23.49	12.92	51.19	6.56	30.41	-0.02	12.20
	R1P	-0.22	2.95	5.58	13.60	10.04	25.38	-1.58	5.77	0.85	15.11	3.08	19.25	13.69	49.33	13.57	41.51	-1.32	14.13
	R2P	-0.22	2.95	3.04	10.14	1.64	9.03	-0.25	3.29	0.70	14.69	0.50	17.65	6.37	35.07	4.36	23.47	-2.50	13.63
	R3P	-0.19	2.90	2.95	10.60	0.92	8.07	-0.09	3.35	0.57	14.13	1.18	19.15	5.87	37.92	2.28	21.45	-1.93	13.23
	S3A	-3.47	8.41	1857.98	6896.95	837.51	2788.61	-14.48	29.54	13.44	71.67	187.72	535.79	342.45	877.83	148.92	361.69	0.26	27.96
	S3P	-0.86	3.91	13.39	32.55	10.57	29.49	-3.68	14.50	2.52	19.80	12.15	43.05	28.56	85.23	16.08	43.03	-17.22	42.87

SAEM_MLX	**R1A**	-1.62	5.75	9.38	16.67	18.83	38.00	-3.93	8.11	7.94	34.90	2.95	23.34	22.21	61.89	22.37	54.02	0.80	12.58
	R2A	-1.91	5.26	11.00	19.46	9.42	20.98	-3.98	8.74	6.12	32.38	9.27	30.38	36.53	77.77	23.47	52.22	0.93	13.18
	R3A	-2.00	4.98	14.78	26.80	8.78	19.48	-6.25	11.04	4.15	30.33	21.59	53.35	53.40	112.68	23.77	53.14	2.89	13.26
	R1P	-0.26	2.96	19.56	26.88	49.55	75.31	-7.56	10.84	-1.55	15.67	7.46	26.07	33.37	77.62	31.24	67.52	11.48	27.02
	R2P	-0.18	2.91	2.30	9.91	0.86	9.21	0.27	3.75	-0.69	14.65	-0.91	18.86	3.29	38.72	1.73	23.64	0.67	15.04
	R3P	-0.11	2.90	1.16	10.20	-0.36	7.65	1.25	4.28	-0.80	13.94	-1.46	20.27	-0.72	40.47	-1.47	22.47	0.16	13.49
	S3A	-7.30	10.39	132.11	185.95	137.19	184.09	-34.04	37.15	70.51	114.85	80.32	123.52	165.84	251.64	81.00	132.09	6.50	28.62
	S3P	-1.10	4.18	47.54	95.87	39.74	83.43	-9.14	22.28	-1.16	21.37	37.18	83.16	79.36	162.95	36.93	78.15	25.72	88.55
SAEM_MLX_tun	**R1A**	-0.84	5.05	2.81	12.91	3.75	24.24	-0.07	6.89	2.26	33.38	-0.01	19.65	6.18	48.68	4.01	41.45	-0.77	11.87
	R2A	-0.55	4.67	2.36	12.33	0.47	11.80	0.69	6.11	1.10	32.10	-0.10	22.08	4.99	49.20	2.15	32.91	-0.72	12.65
	R3A	-0.71	4.41	3.43	14.40	1.04	11.73	0.17	7.55	1.42	29.51	1.71	25.12	7.33	54.26	2.04	30.42	-0.35	12.77
	R1P	-0.24	2.94	3.02	13.34	4.81	26.59	0.09	5.96	-0.29	15.01	0.55	17.69	6.19	45.83	5.64	41.55	-1.14	14.50
	R2P	-0.16	2.94	1.98	9.30	0.37	8.42	0.42	3.19	-0.48	14.53	-0.99	17.09	2.63	33.48	1.39	22.07	-2.11	13.68
	R3P	-0.16	2.90	1.62	9.57	-0.07	7.32	0.73	3.60	-0.61	14.02	-1.00	18.98	0.86	35.46	-0.64	20.44	-1.57	13.34
	S3A	-2.25	6.88	30.60	58.87	27.27	55.13	-11.26	23.25	10.74	67.76	29.30	71.63	61.71	139.04	29.45	68.79	-1.29	23.44
	S3P	-0.50	3.69	6.59	30.63	5.30	29.08	1.07	16.58	-1.34	18.92	3.12	39.15	6.11	79.49	3.60	39.80	-5.11	41.58