Stochastic EM algorithms in population PKPD analyses

Pr France Mentré, INSERM U738, University Paris Diderot
Pr Marc Lavielle, INRIA, Universities Paris 5 & 11
Paris, France

Outline

1. Introduction
2. Stochastic EM algorithms
3. S-ADAPT & MONOLIX
4. Comparisons of S-ADAPT & MONOLIX to NONMEM
5. Examples of application of MONOLIX
6. Conclusion

1. Introduction

- Population PKPD analyses based on nonlinear mixed effects models (NLMEM)
- Results increasingly used in drug development
  - Parameter estimation
  - Model selection
  - Covariate testing
  - Predictions & Simulations

➢ Good estimation methods needed

- Focus here on Maximum Likelihood Estimation (MLE) parametric methods

The FO and FOCE methods

- First proposed methods for estimation of population parameters by maximum likelihood in NLMEM
- Problem in NLMEM: No closed form of the likelihood
- FO: First order linearisation of the model around random effects = 0
- FOCE: First order linearisation of the model around current estimates of random effects

Implemented in NONMEM, WinNonMix, nlme (R and Splus), Proc NLMIXED (SAS)

Limitations of FO and FOCE

- FO
  - assume that mean response = response for mean parameters
  - not true for nonlinear models!!
  - ➢ Bias if "not very small" inter-patient variability

- FOCE
  - not consistent for sparse designs
  - very sensitive to initial estimates:
  - ➢ Lot’s of run failed to converge, waste of time for modellers

- Both: Not real Maximum Likelihood Estimates (MLE)
  - good properties of MLE not demonstrated (LRT, standard errors from Fisher Information matrix, …)

Other approaches for computation of likelihood

- With approximation: linearisation using Laplace (NONMEM)
  - Similar problems of initial values than FOCE

Wolfinger (1993). Laplace's approximation for nonlinear mixed models. *Biometrika*, 80:791-5.

- Integration of the likelihood by Adaptative Gaussian Quadrature (Proc NLMIXED in SAS)
  - Limited to models with small number of random effects

Pinheiro & Bates (1995). Approximations to the Log-Likelihood function in the nonlinear mixed-effects model. *J Comput Graph Stat*, 1:12-35.

Guedj, Thiebaut & Commenges (2007). Maximum likelihood estimation in dynamical models of HIV. *Biometrics*, 63: 1198-1206

2. Stochastic EM algorithms

EM algorithm

- Developed for MLE in problems with missing data
- Two steps algorithm
  - E-step: expectation of the log-likelihood of the complete data
  - M-step: maximisation of the log-likelihood of the complete data
- Mixed-effects models
  - individual random-effects = missing data

Dempster, Laird & Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm, *JRSS B*, 1:1-38.
Lindstrom & Bates (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data, *JASA*, 83:1014-22

EM in NLMEM

- Problem in EM for NLMEM
  - no analytical solution for integral in E-step

1. Linearisation around current estimates of random effects (PPharm, ITS)
   - ➢ Similar problems for sparse design than FOCE

   Mentré & Gomeni (1995). A two-step algorithm for estimation on non-linear mixed-effects with an evaluation in population pharmacokinetics. *J Biopharm Stat*, 5:141-158.

2. Computation with high accuracy numerical techniques (PEM)
   - ➢ Can be very time consuming, not in available software

   Leary, Jelliffe, Schumitzky & Port (2004). Accurate Maximum Likelihood Estimation for Parametric Population Analysis. *PAGE*, 2004.

Stochastic EM in NLMEM

3. Full stochastic E-step
   - Can be very time consuming, not in available software

   Walker (1986). An EM algorithm for nonlinear mixed effects models, *Biometrics*, 52:934-3944.

4. MCPEM: Monte Carlo integration during the E step using importance sampling around current individual estimates

   Bauer & Guzy (2004). Monte Carlo Parametric Expectation Maximization Method for Analyzing Population PK/PD Data. In: D'Argenio DZ, ed. *Advanced Methods of PK and PD Systems Analysis*. pp: 135-163.

5. SAEM: Decomposition of E step in two steps
   – S-step: simulation of individual parameters using MCMC
   – SA-step: stochastic approximation of expected likelihood

   Delyon, Lavielle & Moulines (1999). Convergence of a stochastic approximation version of the EM procedure. *Ann Stat*, 27: 94-128.

SAEM in NLMEM

- Do not compute integral of E-step at each iteration
  - less time consuming than MCPEM
- Good statistical properties clearly demonstrated

  Delyon, Lavielle & Moulines (1999). Convergence of a stochastic approximation version of the EM procedure. *Ann Stat*, 27: 94-128.

  Kuhn, Lavielle (2004). Coupling a stochastic approximation version of EM with a MCMC procedure. *ESAIM Prob & Stat*, 8: 115-131.

  Kuhn & Lavielle (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Comput Stat Data Analysis,* 49: 1020-1038.

  Samson, Mentré, Lavielle (2007). The SAEM algorithm for group comparison tests in longitudinal data analysis base on nonlinear mixed-effects model. *Stat Med*, 26: 4860-4875.
- Addition of a Simulated Annealing algorithm to converge more quickly around the MLE
  - robust with respect to choice of initial estimates
  - fast

Recent extensions of the SAEM algorithm

- Correct handling of BQL data

Samson, Mentré & Lavielle (2006). Extension of the SAEM algorithm to left-censored data in nonlinear mixed effects models: application to HIV dynamic data. *Comput Stat Data Analysis,* 51: 1562-1574,

- Models defined by ODE or SDE

Donnet, Samson (2007). Estimation of parameters in incomplete data models defined by dynamical systems. *J Stat Plan Infer*, 137:2815-2831

Donnet, Samson (2007). Parametric inference for mixed models defined with stochastic differential equations. *ESAIM Prob & Stat*, 12: 196-218

- REML Estimation

Meza, Jaffrézic, Foulley (2007). REML estimation of variance parameters in nonlinez miwed effects models using the SAEMalgorithm. *Biometrical J*, 49:867_888

- Inter-occasion variability

Panhard, Samson (2008). Extension of the SAEM algorithm for nonlinear models with two levels of random effects. *Biostatistics* (accepted for publication)

- Binary data

Meza, Jaffrezic, Foulley (2008). Estimation in the probit normal model for binary outcomes using the SAEM (submitted)

3. S-ADAPT & MONOLIX

- Free software
- MCPEM algorithm in S-ADAPT (B. Bauer)
  - Fortran 95
  - Scripts
  - Other estimation approaches

http://bmsr.usc.edu/Software/Adapt/sadapt.html

(NB: also PDx-MCPEM, S.Guzy, sell by Icon)

- SAEM algorithm in MONOLIX (M. Lavielle)
  - MATLAB (stand alone version for models in library)
  - ODE models in MATLAB or C++
  - User Interface
  - Graphical outputs (GOF, VPC, ...)

http://software.monolix.org

Estimation & outputs: without linearisation

- Estimation of all components of variability (even small) and their standard errors
- Estimation of individual random effects
  - From simulated posterior/conditional distribution
  - Mean, Var and Mode without approximation
- Likelihood estimated by importance sampling
- Population and Individual residuals
  - From simulated marginal and posterior distributions
- No real importance of shrinkage

4. Comparisons of S-ADAPT & MONOLIX to NONMEM

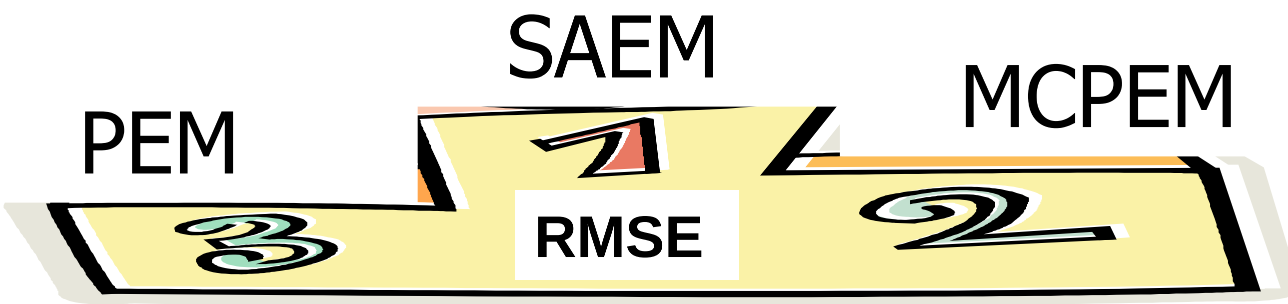
- Girard & Mentré PAGE 2005
  - 100 replicated simulated data sets: one PK and one PD
  - NONMEM V & VI, MCPEM, SAEM (blind evaluation)
- Bauer, Guzy & Chee AAPS J 2007
  - Four simulated examples: PK or PKPD
  - NONMEM VI, PDx-MCPEM, S-ADAPT, MONOLIX 1.1
- Bazzoli, Retout & Mentré PAGE 2007
  - 1000 replicated simulated data sets: one PKPD model
  - NONMEM V, MONOLIX 2.1
- Lavielle & Laveille Monolix website 2007
  - 150 simulated examples: PK (linear & nonlinear)
  - NONMEM V & VI, MONOLIX 2.2 & 2.3
- …

Main conclusions from comparisons

S-ADAPT (MCPEM) and MONOLIX (SAEM)

- Never failed to provide results whatever the models
  - in replicated simulated data sets, often NONMEM results on several data sets are missing
- Much faster than NONMEM FOCE for complex ODE models
- Better results (bias, RMSE) than NONMEM FOCE in sparse designs
- Applied successfully to real complex PKPD data sets (where NONMEM failed to converge)

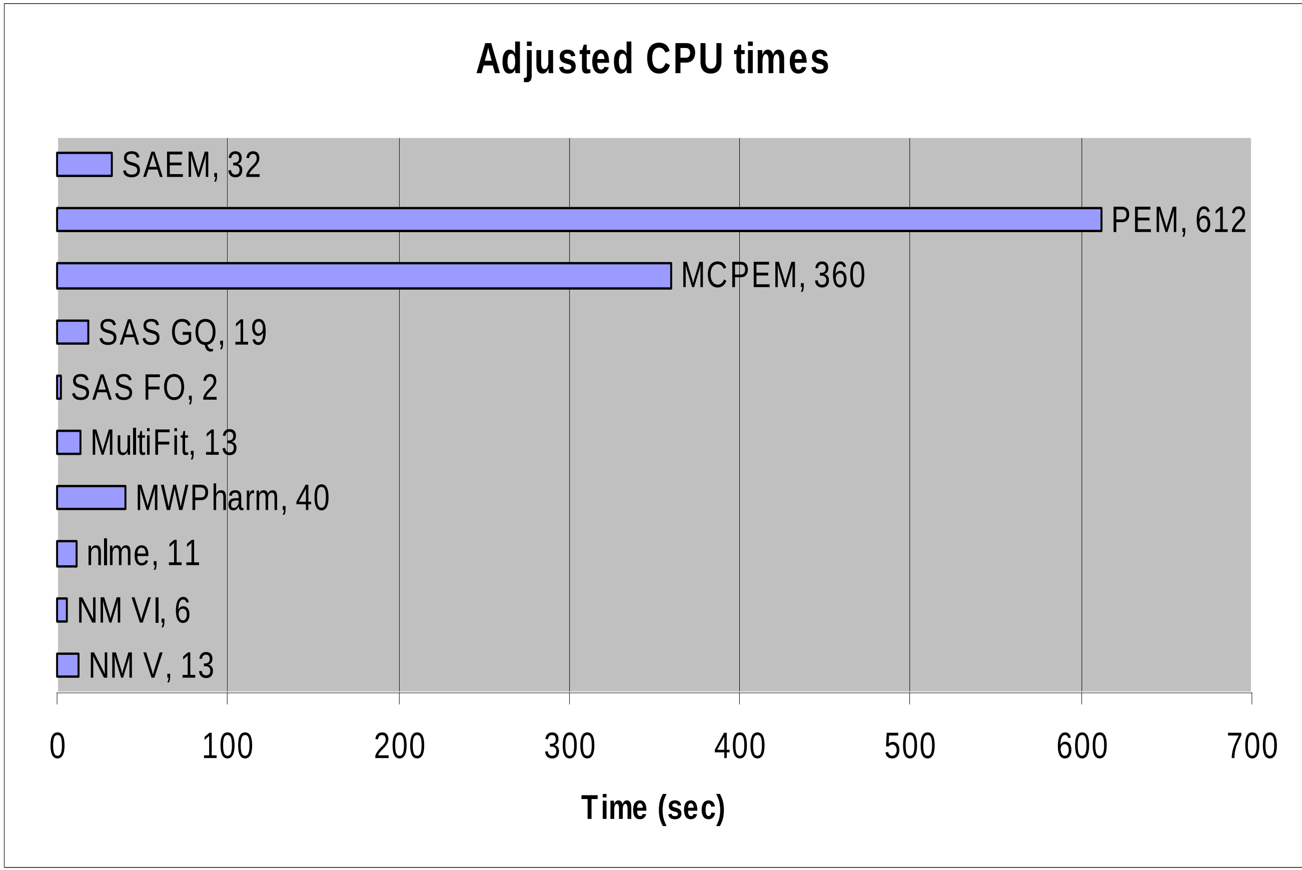
Simulated sparse design PK example
Classification of methods based on RMSE of the 10 population parameters (%)



	1	2	3	4	5	6	9	10
	SAEM	MC PEM	PEM	SAS AGQ	NM V	NM VI	nlme	SAS FO
Pts	89	78	73	70	64	59	44	17
Run	100	100	100	100	100*	49	88	100

Simulated sparse design PK example
Adjusted CPU times for one dataset

	GHz	CPU time for 1 run (sec)
SAEM[1]	1.6	20
PEM	1.7	360
MCPEM	1	360
SAS GQ	3.1	6
SAS FO	3.1	1
Nlme [2]	1.6	7
NM VI [3]	2.13	3
NM V [3]	2.13	6



1. A compiled version would probably be considerably faster than actual implementation in Matlab for PEM and SAEM
2. For a successful nls + nlme run
3. For a successful NONMEM run (EST & COV)

Evaluation of PK library in MONOLIX 2.3
CPU times

	Single Dose 120 subjects 2280 observations	Multiple Doses (7 doses) 120 subjects 5520 observations
1 compartment model IV bolus *linear elimination* *non linear elimination*	 7" 34"	 25" 1' 36"
2 compartments model 1st order oral absorption **lag time** *linear elimination* *non linear elimination*	 15" 2' 18"	 51" 6'52"

5. Examples with MONOLIX (v 2.3.1)

- PK and Viral dynamic model

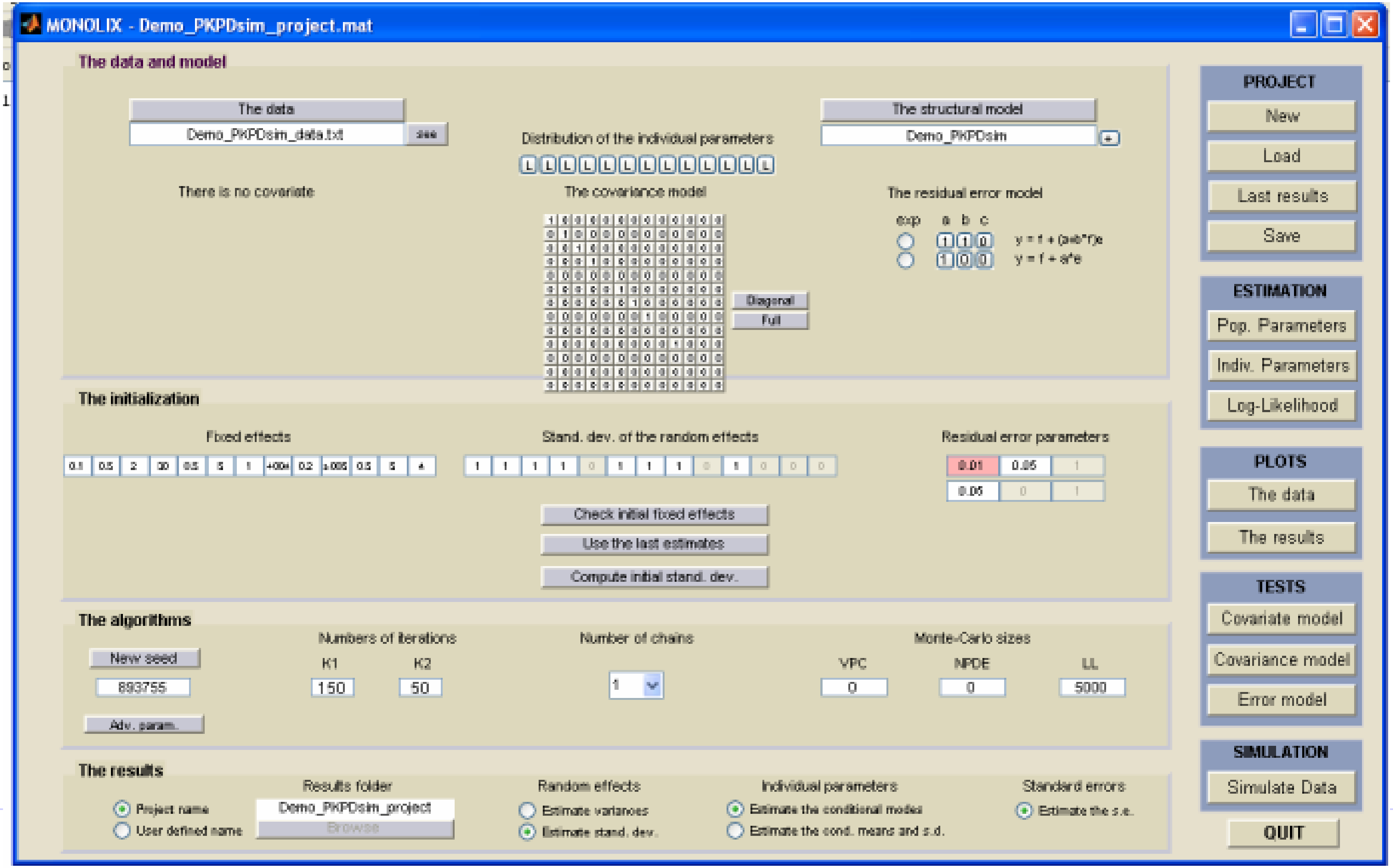
$$dC/dt = D_0 - kC$$
$$dCe/dt = ke_0(C - Ce)$$
$$dT/dt = s - d_1 T - \beta V T$$
$$dI/dt = \beta V T - d_2 I$$
$$dV/dt = (1 - INH)\,p\,I - cV$$

$$D_0 = \begin{cases} \dfrac{\text{Dose}}{\text{Vol} \times \text{Tk}_0} & \text{if } t - t_{Dose} < Tk_0 \\ 0 & \text{otherwise} \end{cases}$$

$$INH(t) = \frac{Ce(t)}{IC_{50} + Ce(t)}$$

- 12 parameters: $Tlag, Tk_0, Vol, Cl, ke_0, IC_{50}, \gamma, s, d_1, \beta, d_2, p, c$.
- Simulated data set: 100 patients
  - Oral dose every day for 5 days
  - 25 PK (D2 - D6) and 20 PD (D1 - D15) observations

MONOLIX interface



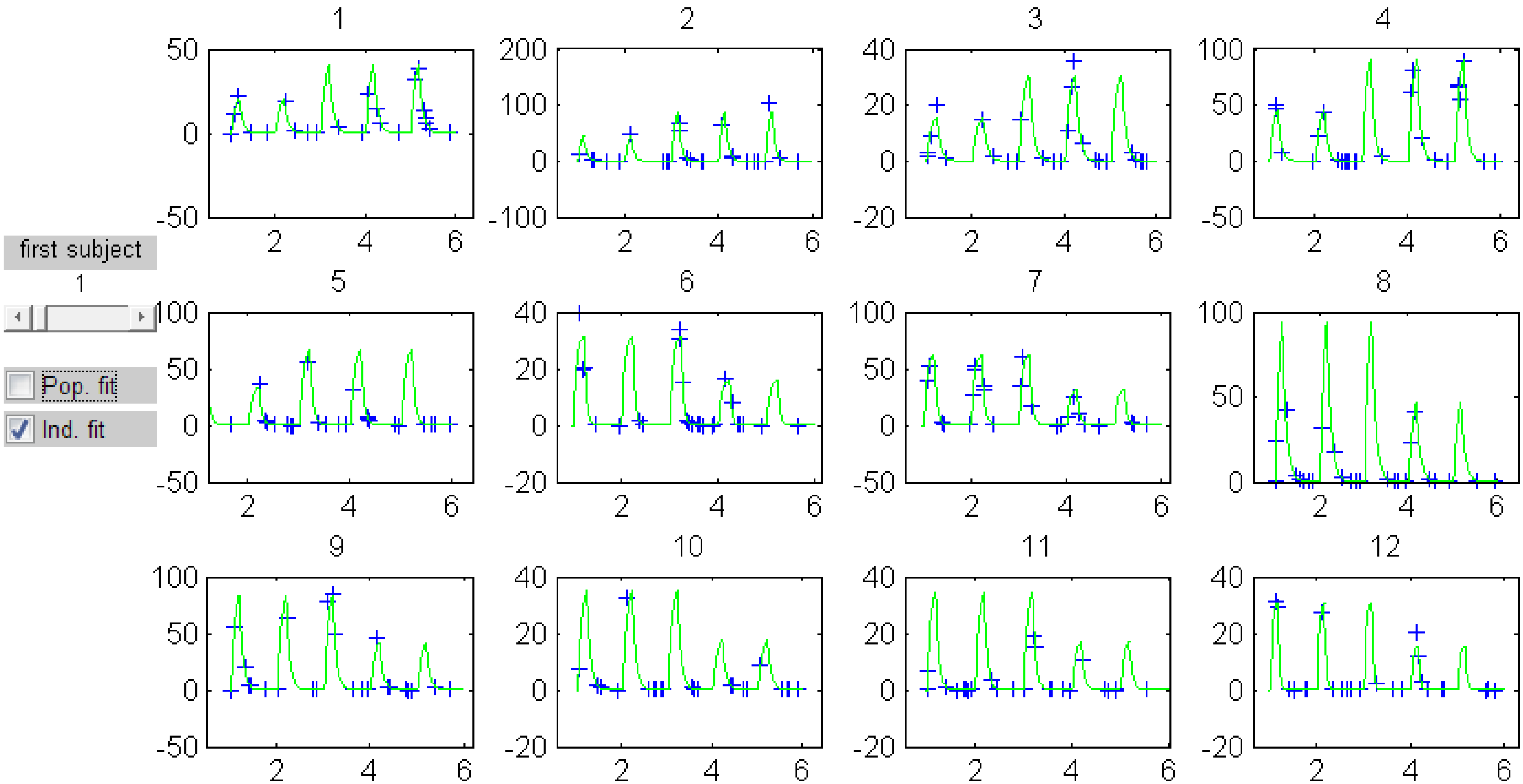
```
**************************************************************
*      March 05, 2008 at 13:16:54
*************************************************************
Estimation of the population parameters

                     parameter      s.e.
Tlag          :        0.0204     0.000612
Tk0           :         0.195      0.00372
V             :          0.95        0.018
Cl            :          13.6        0.266
ke0           :         0.758     0.000733
IC50          :          1.98       0.0436
gamma         :          1.96       0.0483
s             :     4.26e+004    1.14e+003
d1            :          0.26      0.00149
beta          :     2.49e-005    1.04e-006
d2            :          1.54      0.00294
p             :          10.6       0.0868
c             :          3.17       0.0126
```

CPU time is 885 seconds

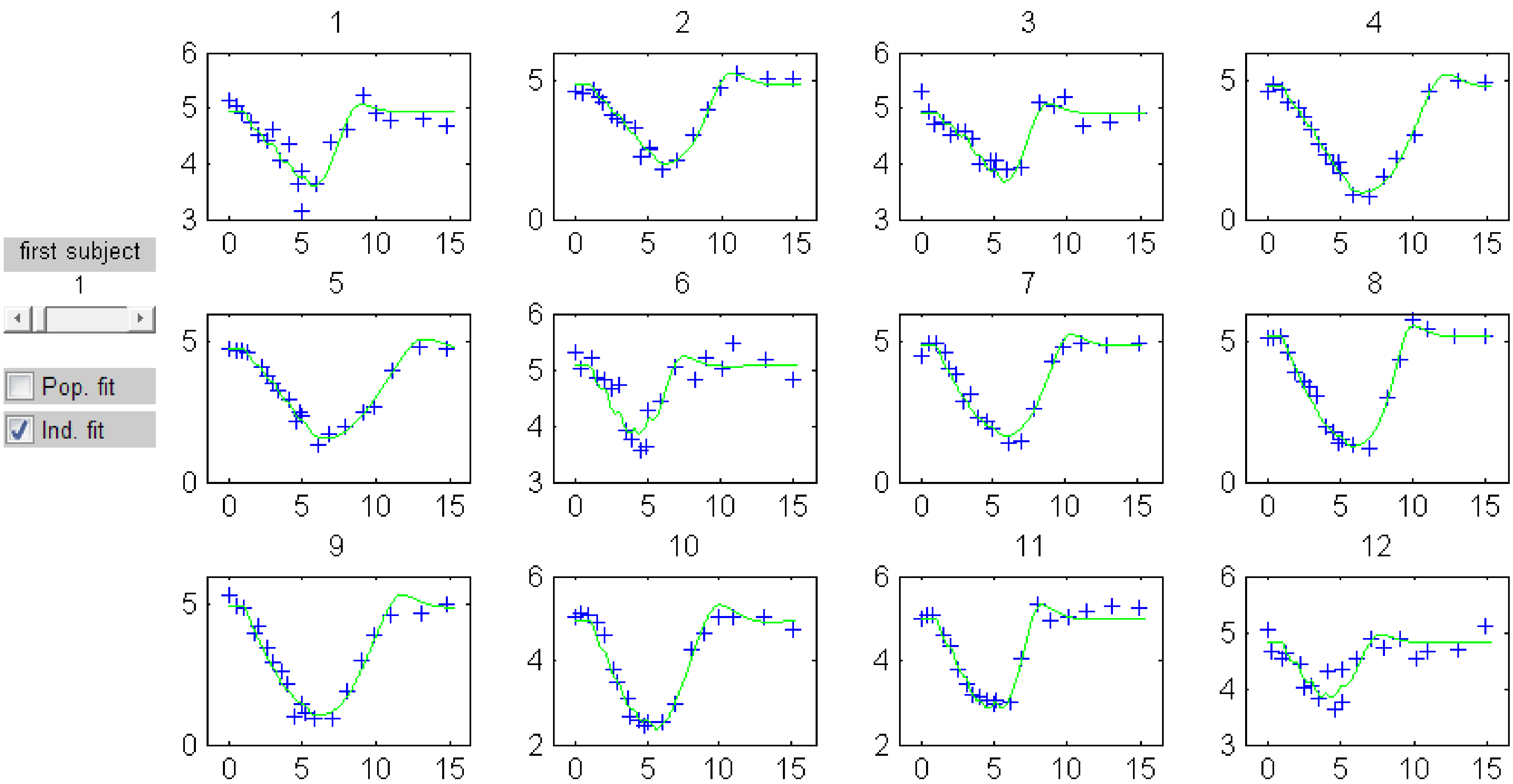
(≅ 15 min, Dual core 2.4 GHz)

		parameter	s.e.
omega_Tlag	:	0.194	0.0257
omega_Tk0	:	0.184	0.0138
omega_V	:	0.187	0.0136
omega_Cl	:	0.193	0.014
omega_ke0	:	0	0
omega_IC50	:	0.183	0.0169
omega_gamma	:	0.228	0.0184
omega_s	:	0.213	0.0218
omega_d1	:	0	0
omega_beta	:	0.315	0.0357
omega_d2	:	0	0
omega_p	:	0	0
omega_c	:	0	0
a_1	:	0.01	0
b_1	:	0.202	0.00391
c_1	:	1	0
a_2	:	0.215	0.00389
b_2	:	0	0
c_2	:	1	0

Some individual PK fits

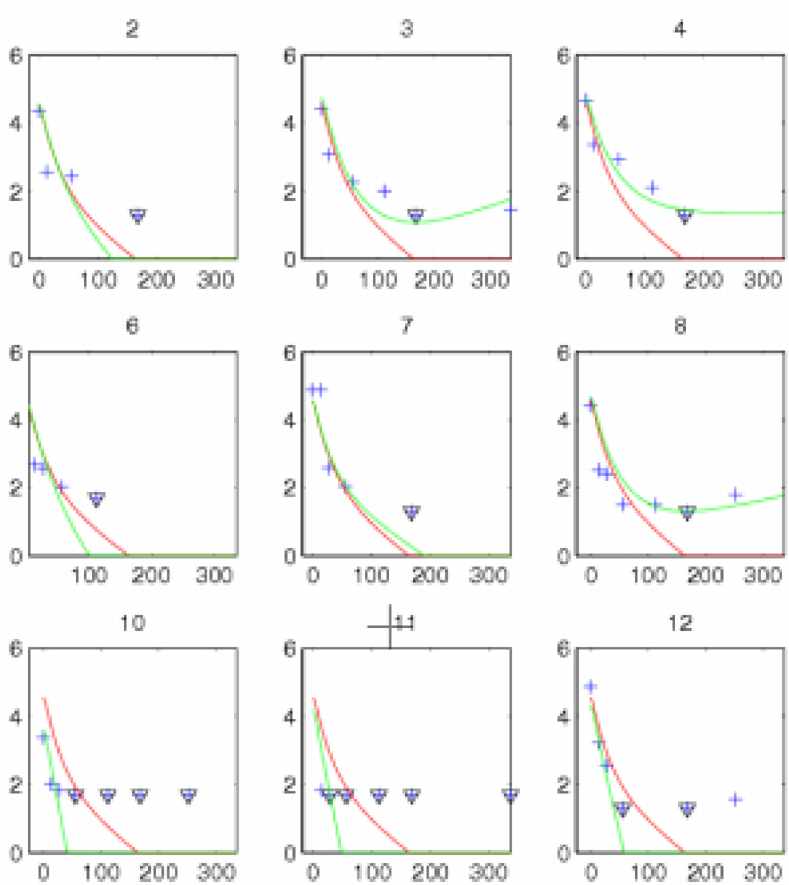


Some individual PD fits

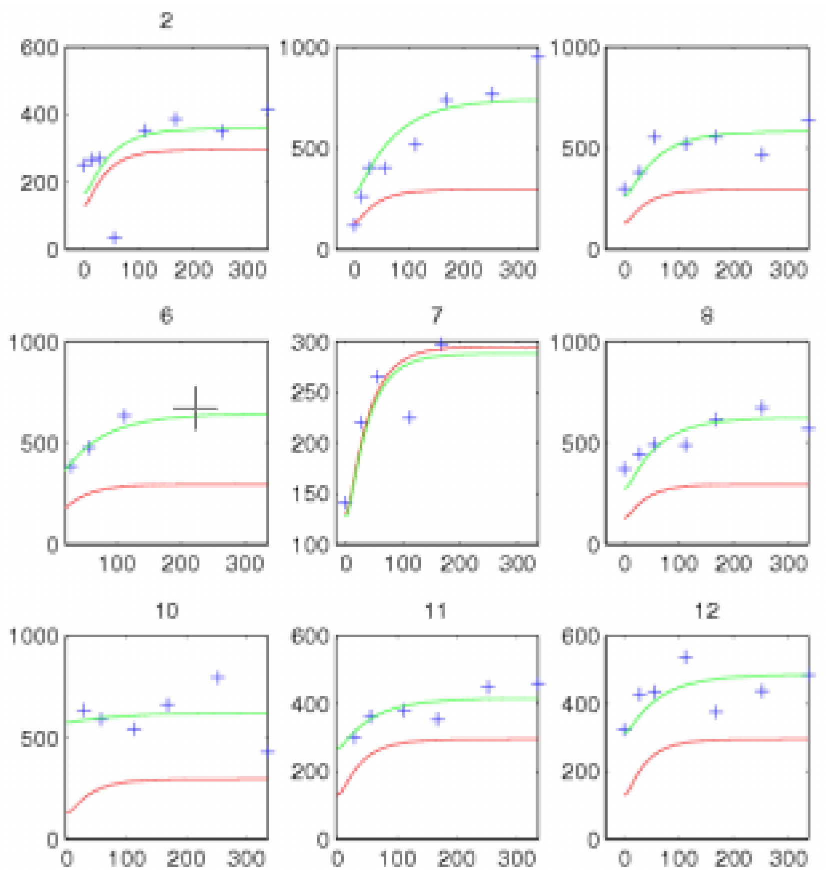


Preliminary results of HIV model on real data
100 HIV patients starting HAART followed one year
5 ODE, 12 parameters

Log Viral Load



CD4 cells



6. Conclusion

- NLMEM using MLE applied for increasingly complex dynamic models
- Drug companies mostly used NONMEM
  - FOCE developed 15 years ago: several drawbacks
- New method based on AGQ
  - limited to problems of small dimension
- New MLE methods based on stochastic EM developed by statisticians
  - fast, consistent, no linearization, …
  - SAEM cleverly used the iteration process

➢ *New software / algorithms should be used*

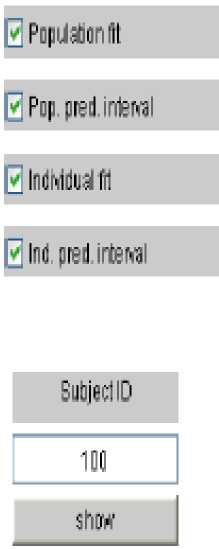
➢ *Extensions of S-ADAPT, MONOLIX … are ongoing*

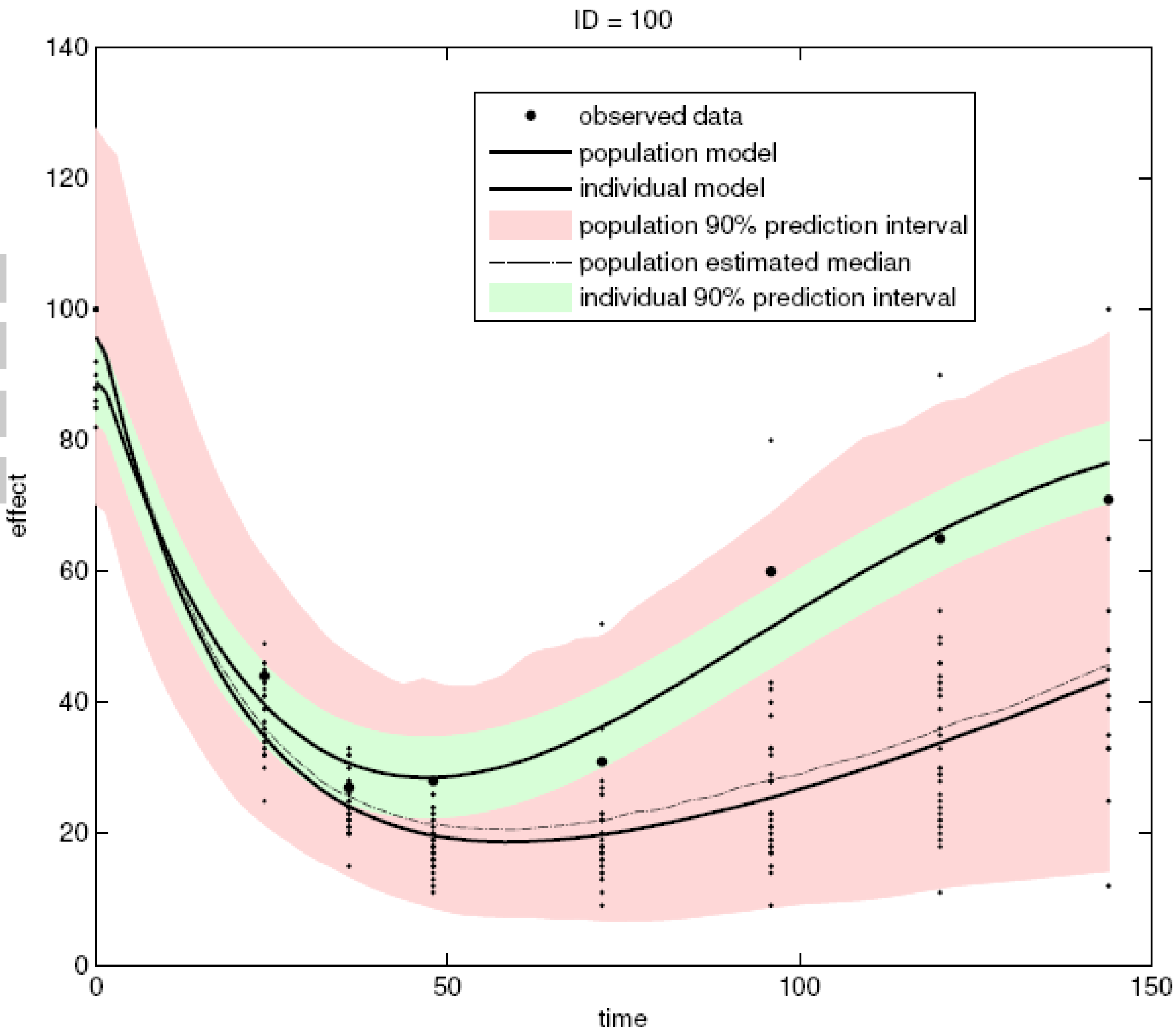
Backup slides

Monolix extensions

- Extensions of MONOLIX software (V2.4, June 2008)
  - Inter occasion variability
  - Other models in PK/PD library
  - Matlab / C
  - Interface for general ODE models
- Other Statistical extension
  - Mixture models
  - Missing covariates
  - Missing data and dropouts
  - Repeated discrete and count data
  - Strategy for model building
  - …

Example of VPC provided by MONOLIX
Warfarin PD data





MONOLIX interface