ACoP12 Poster #: STPM248

Parallel Hierarchical Gibbs-NUTS MCMC Algorithm for Nonlinear Mixed Effects Models in Pumas®

Manu Francis, Mohamed Tarek, Vijay Ivaturi
Pumas-AI Inc.

Objectives

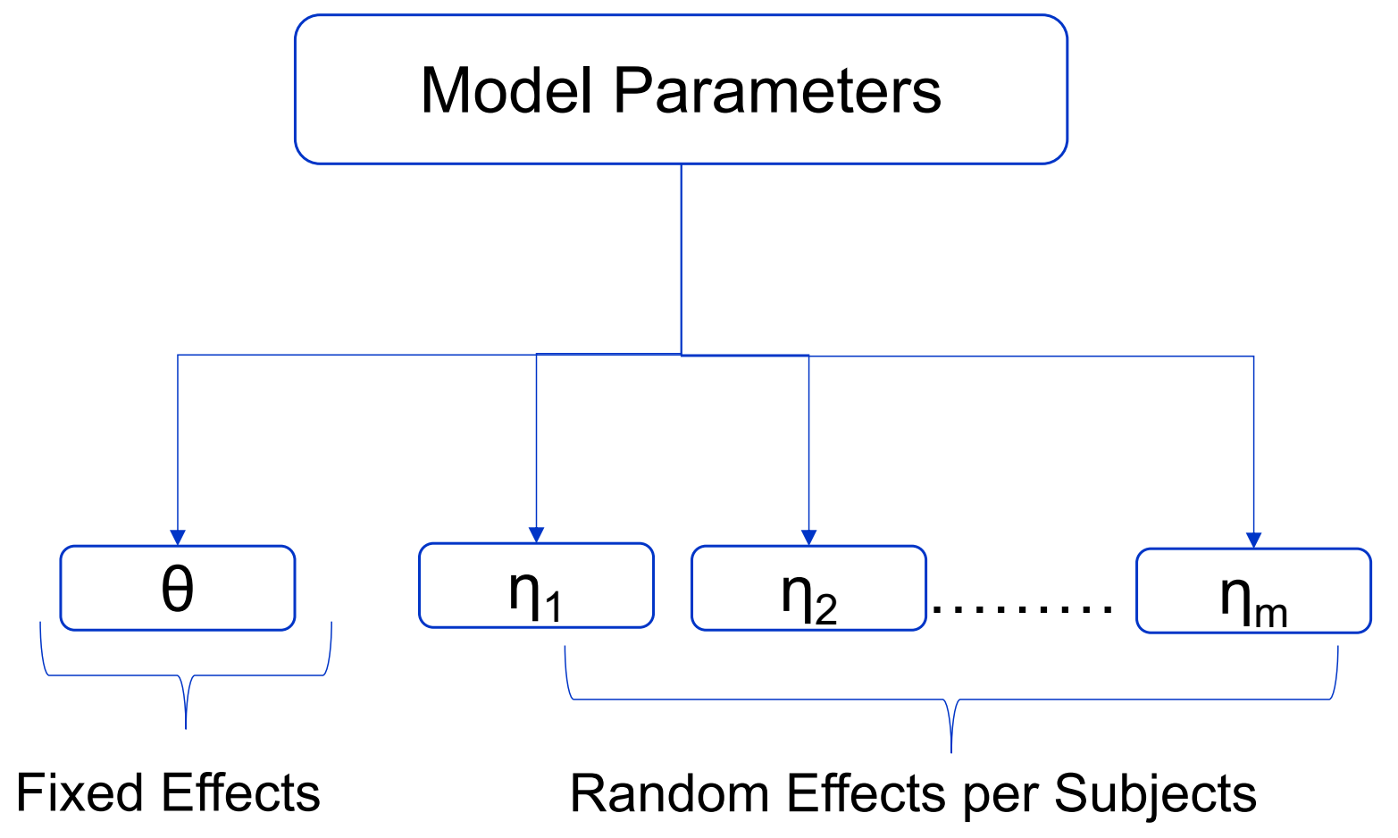
In pharmacokinetic and pharmacodynamic (PK/PD) modelling, Bayesian inference-based parameter estimation attempts to integrate uncertainty into predictions. Markov chain Monte Carlo samplers based on the No-U-Turn sampler (NUTS) algorithm accurately estimate the posterior distribution of parameters and random effects applied to the population [1]. The computation of the posterior distribution becomes computationally costly as the population and model parameters grow larger. Through using the conditional independence of random effects per subject and the hierarchical nature model, this work aims to reduce the sampling time of posterior distribution estimation of PK/PD model parameters using a full Bayesian inference.

Methods

This work explores Hierarchical Gibbs Samplers in parameter estimation for PK/PD models to minimise sampling time by using the computing machine's parallel processing capability [2].

Instead of diagonal adaptation per subject, this Gibbs sampling is implemented using NUTS adaptation with a dense momentum matrix. This implementation allows each subject in the population to have its own Hamiltonian, proposals, and step size for NUTS based Gibbs sampling. Parallel sampling of random effects parameters for each subject is possible since they are conditionally independent per subject in a population. Aside from that, this approach allows for fine-tuning of MCMC samplers for individual subjects and fixed effects.

This work is implemented as part of the Pumas®-NLME^{TM} [3] framework and will be made available for general usage. The multithreading capabilities of Julia programming language is used for parallelizing Bayesian inference step.



Model Parameters

θ

η_1 η_2 ……… η_m

Fixed Effects

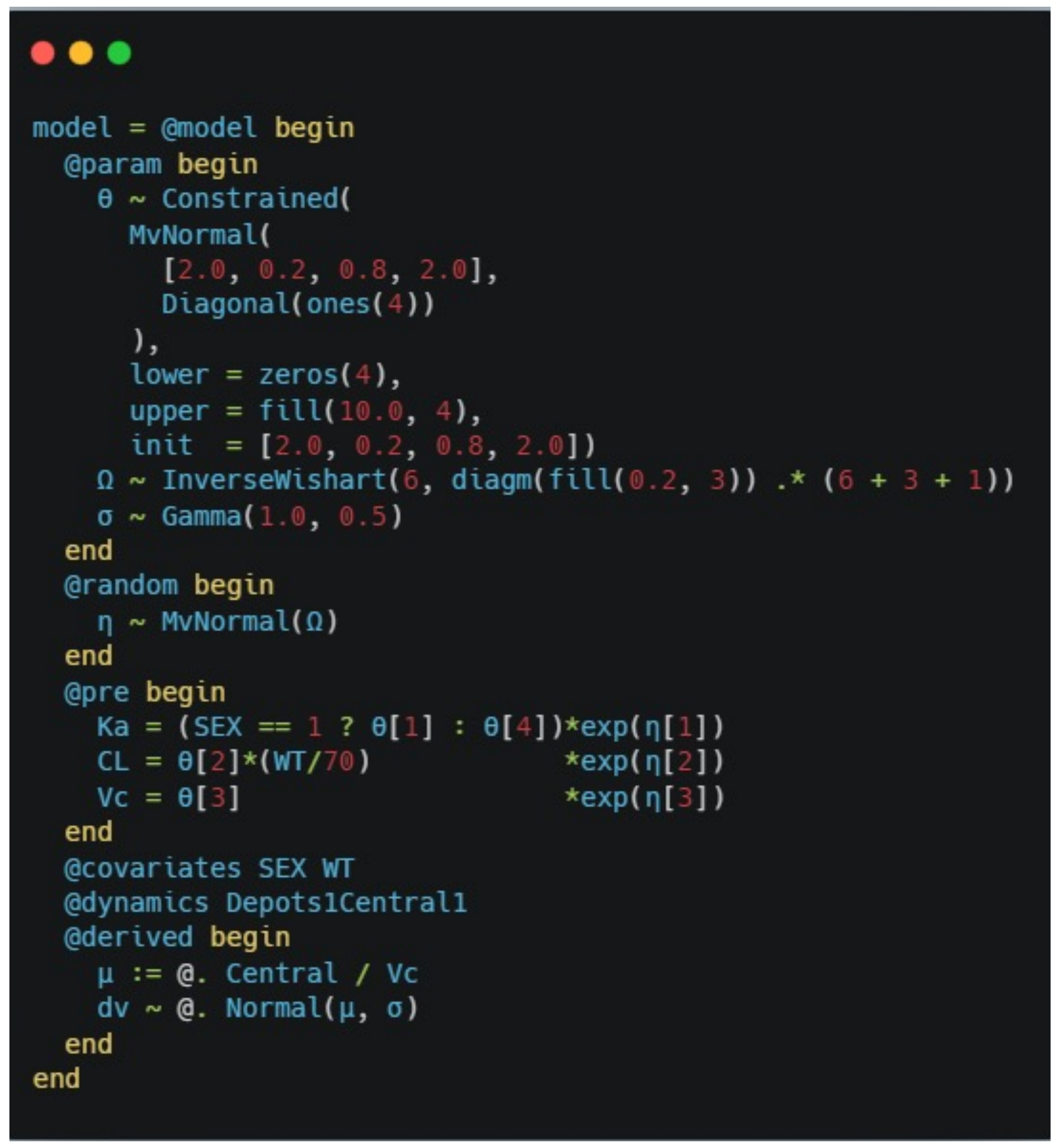
Random Effects per Subjects

Figure 1: Conditionally Independent NLME Model

```
Algorithm 1 Parallel Hierarchical Gibbs-NUTS MCMC Algorithm
  Initialize $X_1 = \{\theta_1^1, ...\theta_n^1, \eta_1^1, ..\eta_m^1\}$   ▷ $\theta$ : Fixed Effects, $\eta$ : Random Effects
  for i=1:T do                                           ▷ For every iteration
      Sample $\theta_1^i, ...\theta_n^i \sim p(\theta|\eta_1^{i-1}, ...\eta_m^{i-1})$
      Sample all following random effects in parallel
      Sample $\eta_1^i \sim p(\eta_1|\theta_1^i, ...\theta_n^i)$
      Sample $\eta_1^2 \sim p(\eta_2|\theta_1^i, ...\theta_n^i)$
      ⋮
      Sample $\eta_1^i \sim p(\eta_m|\theta_1^i, ...\theta_m^i)$
      New sample $X_i = \{\theta_1^i, ...\theta_n^i, \eta_1^i, ..\eta_m^i\}$
  end for
```

Results

Model 1



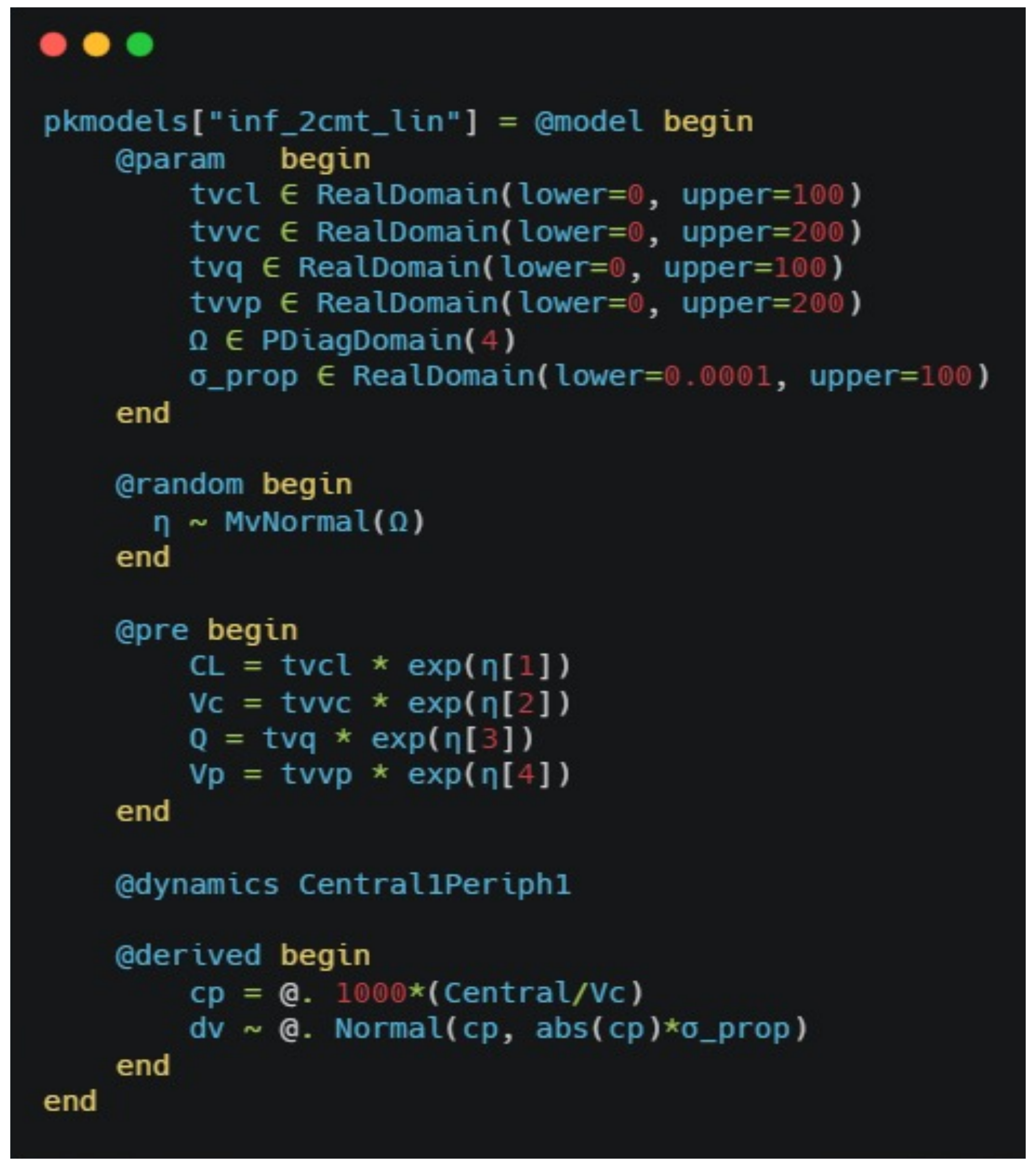
```
model = @model begin
  @param begin
    θ ~ Constrained(
      MvNormal(
        [2.0, 0.2, 0.8, 2.0],
        Diagonal(ones(4))
      ),
      lower = zeros(4),
      upper = fill(10.0, 4),
      init  = [2.0, 0.2, 0.8, 2.0])
    Ω ~ InverseWishart(6, diagm(fill(0.2, 3)) .* (6 + 3 + 1))
    σ ~ Gamma(1.0, 0.5)
  end
  @random begin
    η ~ MvNormal(Ω)
  end
  @pre begin
    Ka = (SEX == 1 ? θ[1] : θ[4])*exp(η[1])
    CL = θ[2]*(WT/70)             *exp(η[2])
    Vc = θ[3]                     *exp(η[3])
  end
  @covariates SEX WT
  @dynamics Depots1Central1
  @derived begin
    μ := @. Central / Vc
    dv ~ @. Normal(μ, σ)
  end
end
```

- This experiment is designed to discuss the effective sample size and sampling time
- Population size of 12 is considered for this experiment
- Total 3000 samples were generated for the analysis

Table 1: Effective sample size and sampling time for different algorithms

Sampler	Parameter	Mean	ESS	Sampling time
Hierarchical Gibbs [4 threads]	θ1	1.8054	119.8489	00:00:53
	θ2	0.0430	40.0543	
	θ3	0.4857	27.1227	
	σ	0.7021	900.2086	
Generalized NUTS	θ1	1.8440	2230.4643	00:04:18
	θ2	0.0419	1206.0245	
	θ3	0.4740	1516.1839	
	σ	0.7043	1347.3785	

Model 2



```
pkmodels["inf_2cmt_lin"] = @model begin
    @param   begin
        tvcl ∈ RealDomain(lower=0, upper=100)
        tvvc ∈ RealDomain(lower=0, upper=200)
        tvq ∈ RealDomain(lower=0, upper=100)
        tvvp ∈ RealDomain(lower=0, upper=200)
        Ω ∈ PDiagDomain(4)
        σ_prop ∈ RealDomain(lower=0.0001, upper=100)
    end

    @random begin
      η ~ MvNormal(Ω)
    end

    @pre begin
        CL = tvcl * exp(η[1])
        Vc = tvvc * exp(η[2])
        Q = tvq * exp(η[3])
        Vp = tvvp * exp(η[4])
    end

    @dynamics Central1Periph1

    @derived begin
        cp = @. 1000*(Central/Vc)
        dv ~ @. Normal(cp, abs(cp)*σ_prop)
    end
end
```

- This experiment is designed to show the improvement in sampling time ratio compared to effective sampling time ratio
- Different population size is considered
- Total 750 samples were generated for the analysis

Table 2: Effective sample size vs Sampling time for tvvc parameter.

Population Size	Sampling Time		Sampling Time Ratio	ESS		ESS Ratio
	NUTS	Gibbs		NUTS	Gibbs	
10	0:00:35	0:00:07	5	100.6928	34.7432	2.83
20	0:01:08	0:00:12	5.66	105.8525	37.3022	2.83
50	0:03:22	0:00:25	8.08	257.1940	113.5819	2.26
100	0:06:49	0:00:51	8.01	170.0646	60.7751	2.79
120	0:08:38	0:01:12	7.19	127.3898	40.8092	3.121

Conclusions and Discussions

- The parameter estimation of an NLME model with 11 parameters and 3 random effect per subject took 4 minutes 18 seconds using NUTS sampler; however the hierarchical Gibbs sampler with a single thread took only 53 seconds. But owing to high autocorrelation between samples, the effective sample size was lower in the hierarchical Gibbs sampler.
- The saving in sampling time ratio is greater than the effective sample size ratio in some examples. This means that the lower effective sample size can be compensated for with the lower sampling time of the hierarchical Gibbs sampler.
- Although the autocorrelation between samples is higher in the hierarchical Gibbs sampler for PK/PD parameter identification than in NUTS based sampling, the mean values of parameters are similar to NUTS based sampling.

References

1. Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.
2. Gelfand, A. E. (2000). Gibbs sampling. Journal of the American statistical Association, 95(452), 1300-1304.
3. Rackauckas, Chris, Yingbo Ma, Andreas Noack, Vaibhav Dixit, Patrick Kofod Mogensen, Simon Byrne, Shubham Maddhashiya, et al. "Accelerated Predictive Healthcare Analytics with Pumas, a High Performance Pharmaceutical Modeling and Simulation Platform." BioRxiv, November 30, 2020, 2020.11.28.402297. https://doi.org/10.1101/2020.11.28.402297
4. Neal, R. M. (2011). MCMC using Hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11), 2.