

Matt Rosebraugh
Mohammed El-Komy
22S: 138 Final Project Report
12/8/2008
Instructor: Dr. Cowles

Bayesian Hierarchical Models for Cadralazine Population Pharmacokinetics: Effect of Residual Error Structure

Introduction:

Population pharmacokinetic analysis is often done when population parameter information needs to be obtained but only sparse data exist. For the data set consider y_{it} , where $i=1, \dots, n$ and $t=t_1, \dots, t_n$ where y_{it} denotes the concentration of drug in the blood of individual i at time t . The drug is administered either intravenously (IV) or orally at time $t=0$. Following an IV dose concentration is assumed to reach a peak instantaneously. Oral drug administration causes the drug plasma concentration to rise slowly because absorption into the blood stream is occurring.

The data y_{it} is assumed to follow a normal or log normal distribution and a model is often used for the data that takes the form (2)

$$y_{it} = C^{(i)}(t, \theta) + \varepsilon_{it}$$

Where, $C^{(i)}(t, \theta)$ is a function of the predictor t , and of a vector θ with individual parameters. The ε_{it} is a random error term that can represent deviation from the assumed model. When the data come from an individual study, the parameters can be estimated separately for each individual. In population studies, the data are often insufficient to allow this, and therefore a non-linear mixed-effect modeling approach is used. A common model is assumed for all individuals and the inter-individual variation is accounted for by the assumption of a separate random parameter for each individual (2).

Among the various parametric and non-parametric approaches for population analysis, the non-linear mixed effects modeling with first order conditional estimation (NONMEM-FOCE) and the Bayesian approaches have been the most popular for the last decade. The NONMEM-FOCE approach is called the iterative 2-stage hierarchical maximum likelihood method, the first stage being the individual likelihood of the data $L(y_{it} | \theta, \sigma)$, given a set of model parameters (θ) and a residual error coefficient (σ) for

that individual, and the second stage being the likelihood of the set of model parameters, given some knowledge or assumptions of the distribution. Parameters among the population $h(\theta|\mu,\Omega)$, parameterized by the population mean (μ) and inter-subject variance (Ω) (3). The Bayesian approach is called the iterative 3-stage MCMC method. Stages 1 and 2 are similar to those characterized by random effects modeling with NONMEM and the third stage in this case refers to the prior (4). On the other hand, stages 1 and 2 in the Bayesian approach do not maximize the likelihood. They collect a series of possible μ 's, Ω 's and σ 's with a frequency that is based on their likelihood of explaining the data (3).

In this project we fitted three hierarchical non-linear mixed effects Bayesian models, with different residual error structures, to a real data set of the antihypertensive drug: Cadralazine. In detail, the fitted models were: a 1- compartment pharmacokinetic model with additive error, a 1- compartment pharmacokinetic model with proportional error and a 1- compartment pharmacokinetic model with mixed error.

Materials and Methods:

Dataset:

The data set consists of Cadralazine plasma concentration-time profiles from 10 cardiac failure patients. The patients were monitored for approximately 24 hours following an I.V. bolus administration of 30mg of Cadralazine (5). The dataset is displayed in Appendix I.

Bayesian Analysis:

The data was analyzed using WinBUGS differential interface (WBdiff) according to an open one compartment pharmacokinetic model parameterized in terms of clearance and volume of distribution and a 3 stage hierarchical model (4) (see below). The WBdiff code for the fitted model is displayed in appendix II.

Stage 1- Model for the data

$$y_{ij} = C(\theta_i, t_{ij}, D_i) + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \tau / v_{ij}), \tau = \sigma^{-2}$$

Where

y_{ij} is the j th observation for the i th subject

$C(\theta_i, t_{ij}, D_i)$ is the expected value of the data from the model

θ_i is a vector of individual pharmacokinetic parameters for the i th subject

t_{ij} is the sampling time of the j th observation for the i th subject

D_i is the drug dose for the i th subject

ε_{ij} is the residual error

τ / v_{ij} is the precision

v_{ij} is the residual error structure

$v_{ij} = 1$ for additive error

$v_{ij} = C_{ij}^2$ for proportional error

$v_{ij} = 1 + C_{ij}^2$ for mixed error

$C(\theta_i, t_{ij}, D_i)$ represents the analytical solution for the *one compartment first order ordinary differential equation*:

$$dC(\theta_i, t_{ij}, D_i)/dt_{ij} = -\left(Cl_i/V_i\right) \cdot C_{ij} \quad , C_{ij}(0) = \frac{D_i}{V_i}$$

Stage 2- Model for between subject variability

$$\theta_i \sim \text{MVN}(\mu, \Omega^{-1})$$

Where

μ is a vector of mean population pharmacokinetic parameters

Ω is a variance-covariance matrix of between subject random effects.

Stage 3- Model for the priors

$$\tau \sim \text{Gamma}(a, b)$$

$$\mu \sim N(\bar{\mu}, \Sigma^{-1})$$

$$\Omega^{-1} \sim W(R, \rho)$$

Where

$\bar{\mu}$ is a vector of prior population mean values of parameters

Σ^{-1} is the prior precision matrix

W represents wishart distribution with parameters R and ρ

The prior distributions of the parameters were obtained from literature (see Bayesian prior construction section) and three MCMC chains of 20,000 iterations were run for each model from different initial values for each chain (see Bayesian initial estimates section). Convergence diagnostics such as metropolis acceptance rate plot, history plots, autocorrelation plots and Gelman-Rubin diagnostic plots were collected for each model. In order to compare goodness of fit of the three error models, individual plots for each subject's observed and model predicted drug concentration verses time were obtained. Additionally, the residual plots as well as the deviance information

criterion were obtained after 25,000 iterations for each chain for each model. Summary statistics of posterior distributions of each model's parameters were depicted after discarding the first 8,000 iterations.

Bayesian Prior Construction:

Four previous studies (6-9) on Cadralazine pharmacokinetics in healthy subjects were used to obtain priors on the parameters of interest: clearance (CL) and volume of distribution (V). Appendix III section A shows the number of subjects included in each study and the means and variances of the parameters. In this project we assumed that the parameter values drawn from the literature follow a normal distribution. In order to force positive values on the Bayesian estimates of the parameters, the subject specific vector of parameters was assumed to be drawn from a log-normal distribution such that:

$$\bar{\theta}_i = \log(\theta_i) = [\log(CL_i), \log(V_i)]^T$$

Therefore, in constructing proper priors, the following procedure was followed:

1. Calculation of weighted means of normally distributed parameters

$$E(\theta_K) = \frac{\sum_{i=1}^n N_i \mu_{\theta_{k,i}}}{\sum_{i=1}^n N_i}$$

$$\text{Var}(\theta_K) = \frac{\sum_{i=1}^n N_i \sigma_{\theta_{k,i}}^2}{\sum_{i=1}^n N_i}$$

$$K=1,2 \quad i=1,2,3,4 \quad \theta_1=CL \quad \theta_2=V$$

Where N_i denotes the number of subjects in the i th study and $\mu_{\theta_{k,i}}$ denotes the mean and variance of the parameter θ_k in the i th study.

2. Log-transformation of normal distribution means and variances

$$E_{\log}(\theta_K) = \log[E(\theta_K)] - \frac{1}{2} \log \left[1 + \left(\frac{\sqrt{\text{var}(\theta_K)}}{E(\theta_K)} \right)^2 \right]$$

$$\text{var}_{\log}(\theta_K) = \log \left[1 + \left(\frac{\sqrt{\text{var}(\theta_K)}}{E(\theta_K)} \right)^2 \right]$$

$$K=1, 2 \quad \Theta_1=CL \quad \Theta_2=V$$

It is worth mentioning that the calculated log-normal variances were extremely low (<0.05). When we deal with population pharmacokinetic studies with significantly different random effects it is expected to have lower precisions than in individual studies. Accordingly, the obtained variances were empirically multiplied by 10.

3. Construction of population mean multivariate normal distribution parameters.

a. Prior mean $\bar{\mu}$:

$$\bar{\mu} = [E_{\log}(CL), E_{\log}(V)]^T$$

b. Prior precision matrix Σ^{-1}

$$\Sigma = \begin{bmatrix} \text{var}_{\log}(CL) & 0 \\ 0 & \text{var}_{\log}(V) \end{bmatrix}$$

$$\Sigma^{-1} = \text{Inverse}(\Sigma)$$

4. Construction of population variance-covariance matrix wishart distribution parameters.

a. The degrees of freedom (ρ)

ρ represents prior sample size. In this project ρ was approximated by the total number of subjects that provide information about the parameters (128 subjects) divided by the number of studies (4 studies).

b. The (R) parameter

$$R = \begin{bmatrix} \text{var}_{\log}(CL) \cdot (\rho - 2 - 1) & 0 \\ 0 & \text{var}_{\log}(V) \cdot (\rho - 2 - 1) \end{bmatrix}$$

Where 2 represents the number of parameters.

5. Residual error prior

Residual error represents the experimental error as well as model misspecification. Therefore, it is difficult to assign an informative prior on it. As a result, we selected a

non-informative uniform Gamma prior with a mean of 10 to represent our knowledge about the residual error.

The numerical values of the fore mentioned priors are given in appendix II section B.

Bayesian Initial Values Selection:

Three sets of initial values were used in the analysis. The first set is centered around the prior, the second is over dispersed (50% higher than the first set) and the third is under dispersed (50% lower than the first set). The numerical values of the fore mentioned initial values are given in appendix III section C.

Frequentist Analysis:

The frequentist maximum likelihood analysis of the Cadralazine data set was conducted using the first order conditional estimation (FOCE) method in NONMEM VI. The fitted pharmacokinetic model (Advan 1- Trans 2) was an open one compartment model parameterized in terms of clearance and volume of distribution. In common pharmacokinetic theory a one compartment model is also known as a single exponential disposition model. The identical model was run with the data using three different error structures- additive, proportional and mixed error. The initial estimates of the parameters used in NONMEM were identical to the initial estimates used in the Bayesian analysis. The control stream file (specifies exact model for NONMEM) is provided in appendix IV.

Results and Discussion:

The three models converged successfully in both WBdiff and NONMEM. Figure 1 shows that metropolis acceptance rates reach the desired 0.2 to 0.4 level within the first 500 iterations. The history plots in figure 2 show quick and good mixing for all parameters. Gelman-rubin diagnostic plots in figure 3 show that the pooled and within chains converged to stability between 5000 and 6000 iterations, so it was the decision to collect the statistical summary of the parameters starting from 8000 iterations. The ratio chain became close to 1 at around 4000 iterations. Almost no auto correlation, especially

at later stages, was observed for the proportional and mixed error models except for the parameter $\mu[2]$ (figure 4). However, the autocorrelation for this parameter reaches less than 0.2 during MC sampling. The summary statistics for all parameters were very similar for the three error models except for the residual error which was about 4 times lower in the additive error model compared to the other error models.

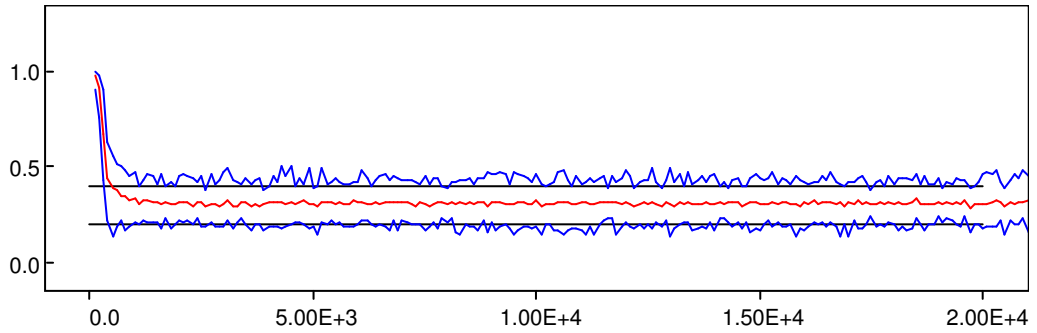
In checking goodness of fit, the first plots to look at would be plots of observed and individual predicted concentration versus time for all subjects (figure 5). It could be seen that additive error model gave the best fit, followed by mixed error model and finally comes the proportional error model. This notice was further confirmed by DIC values (table 1). The residual plots (figure 6), which could be looked at as an indicator for bias in predictions, show an overall trend for uniform distribution of points along the scatter line (i.e. good homoscedasticity). However, it is obvious that the homoscedasticity decreases in the order: additive, mixed and proportional error models respectively.

A comparison between the results obtained from the frequentist analysis and from the Bayesian analysis is given in table 4. The values of the theta parameters (the means of clearance and volume of distribution) are very similar for the two approaches. However, the variance covariance matrix as well as the residual error were much smaller in the case of the frequentist analysis.

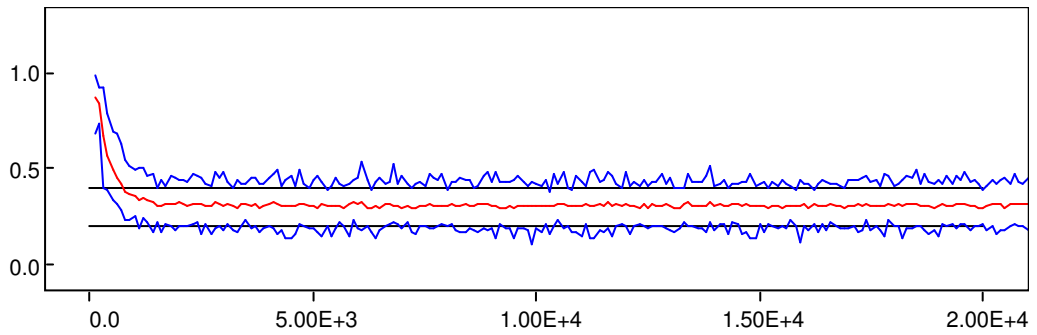
Further investigation into the results shows that the posterior means of clearance and volume of distribution were much lower than the prior means (2.82 vs. 12.5 for CL and 13 vs. 56 for V). An explanation for these results can be offered from a statistical point of view and a physiological point of view. Statistically, the posterior distribution depends on how much power is provided by the likelihood and the prior. It is clear from the results that the data was more influencing in posterior estimates. Physiologically, the prior parameters of Cadralazine were determined in subjects with healthy hearts while the data set used for analysis was obtained from cardiac failure patients. Congestive heart failure is associated with hypoperfusion to various organs including the sites of drug clearance such as the liver and kidney. Accordingly, the main changes in drug pharmacokinetics seen in congestive heart failure are a reduction in the volume of distribution and impairment of clearance as shown in the literature (10).

Figure 1. Metropolis acceptance rates for the Bayesian analysis:

(A) 1-CP with Additive Error Model



(B) 1-CP with Proportional Error Model



(C) 1-CP with Mixed Error Model

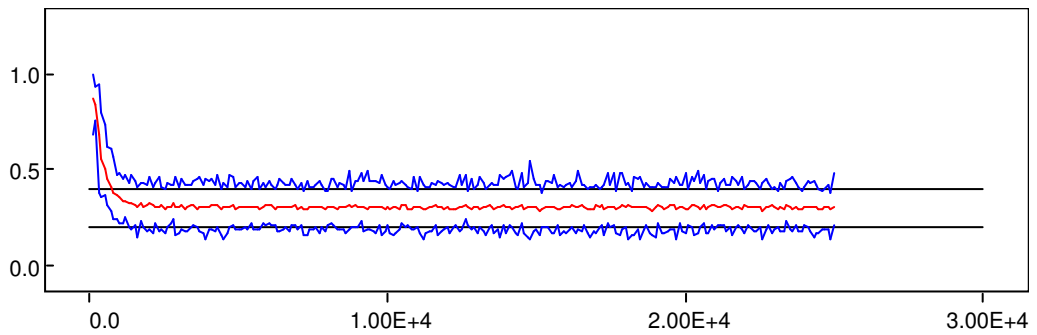
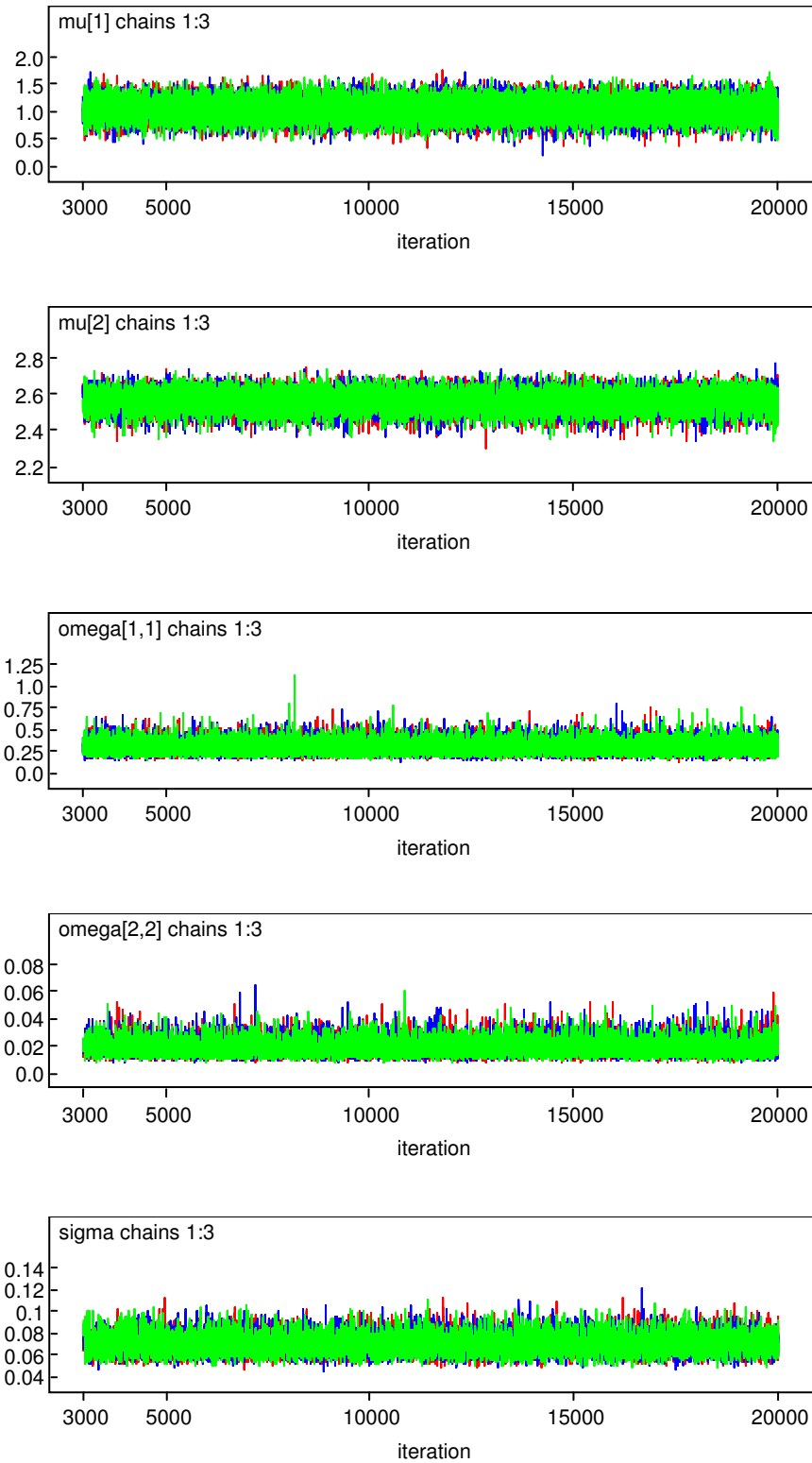
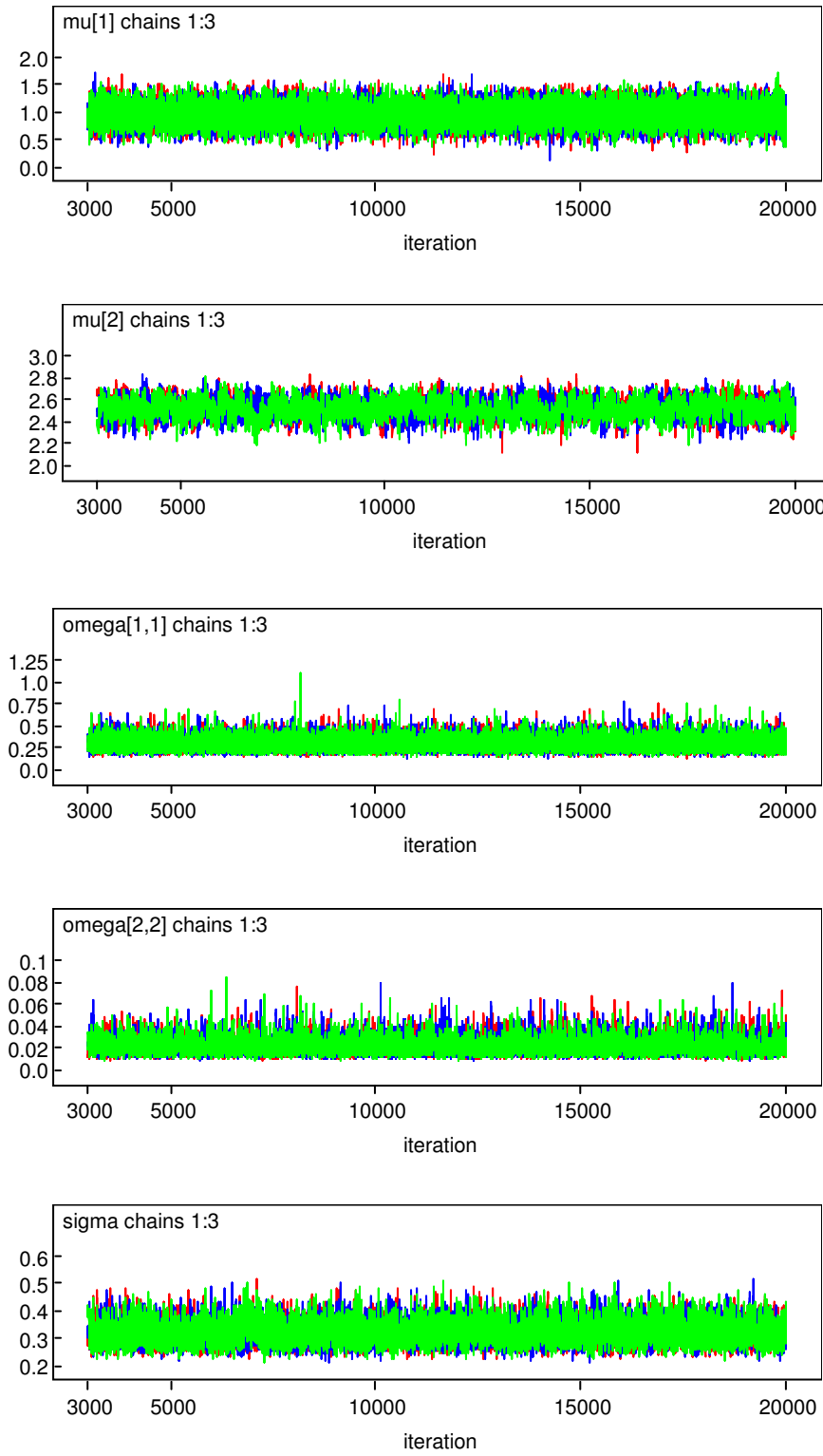


Figure 2. History Plots for Parameters μ , Ω (1,1) , Ω (2,2) , σ^2 :

(A) 1-CP with Additive Error Model



(B) 1-CP with Proportional Error Model



(C) 1-CP with Mixed Error Model

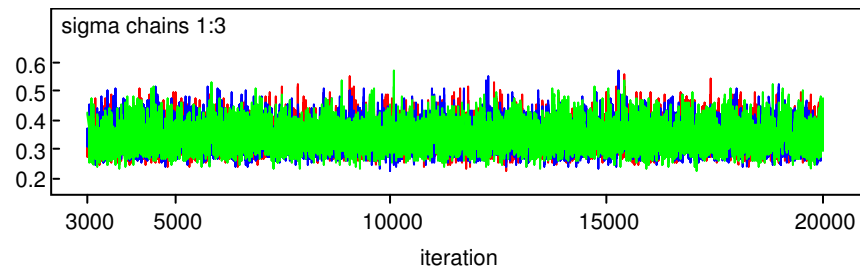
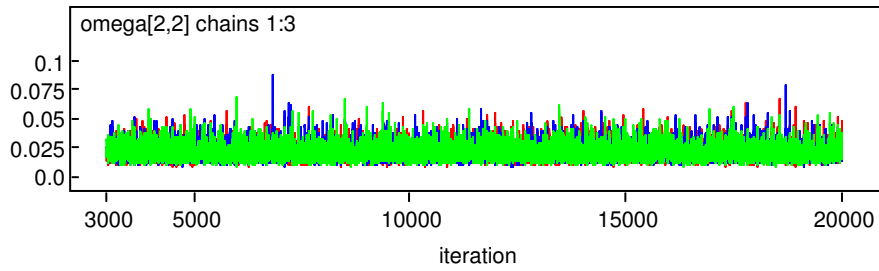
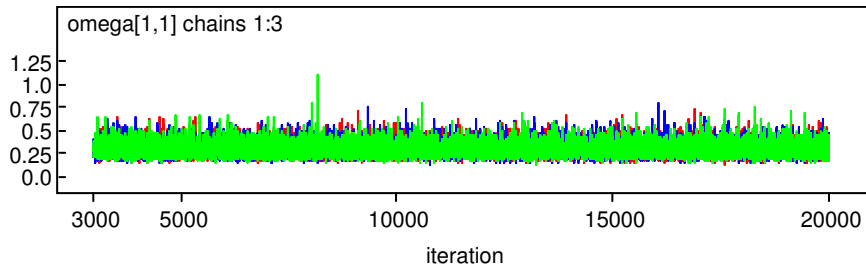
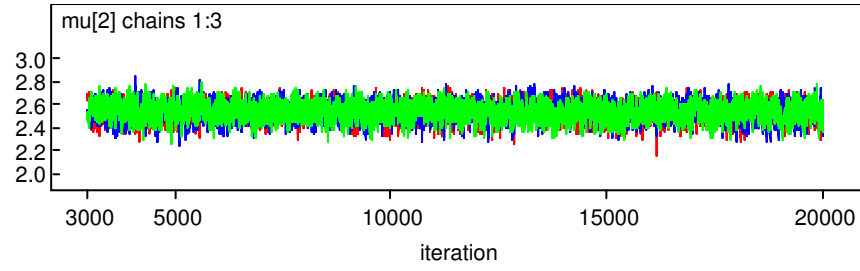
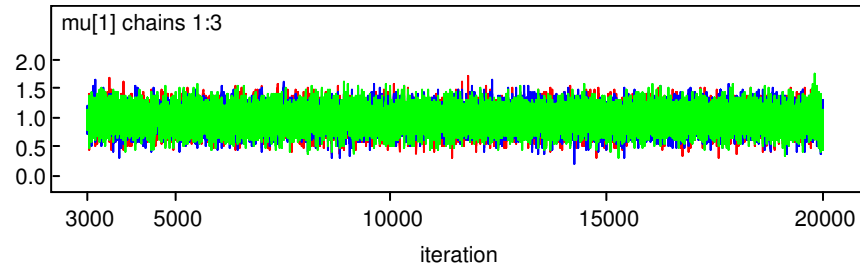
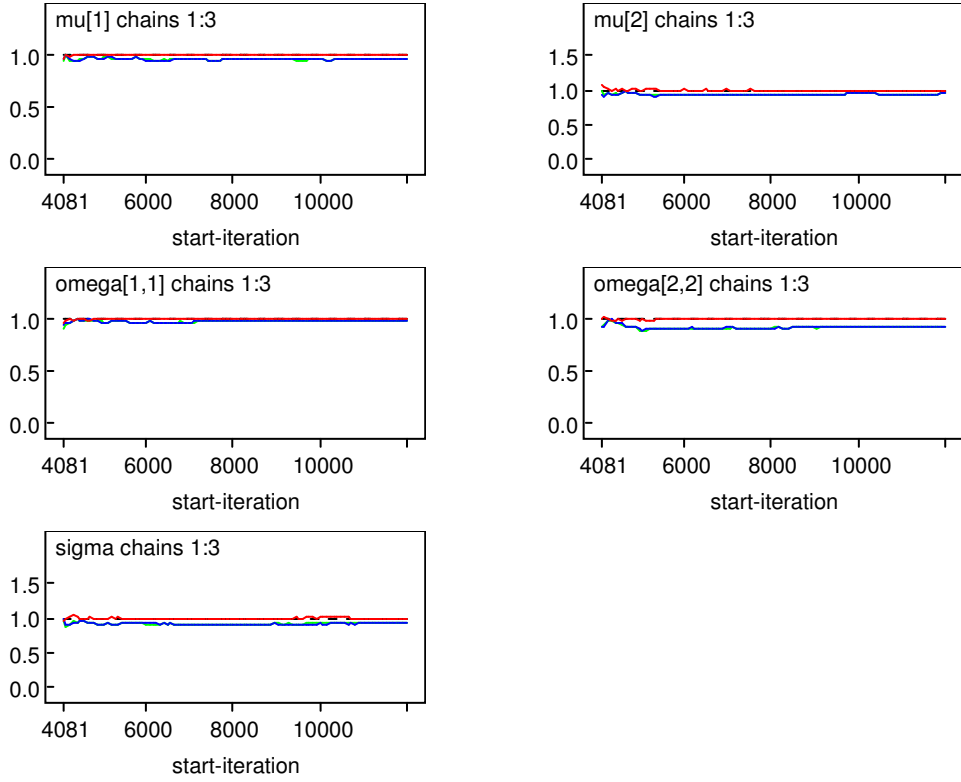
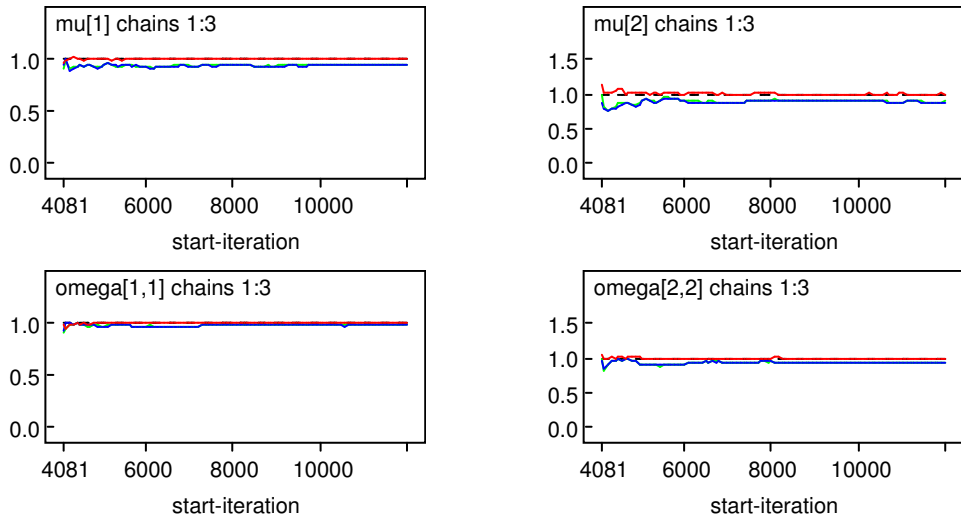


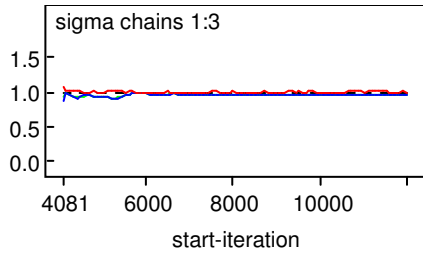
Figure 3. Gelman-Rubin Diagnostic Plots for Parameters μ , $\Omega (1,1)$, $\Omega (2,2)$, σ^2 :

(A) 1-CP with Additive Error Model



(B) 1-CP with Proportional Error Model





(C) 1-CP with Mixed Error Model

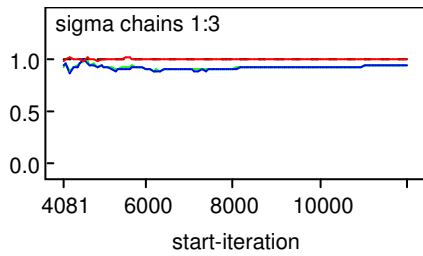
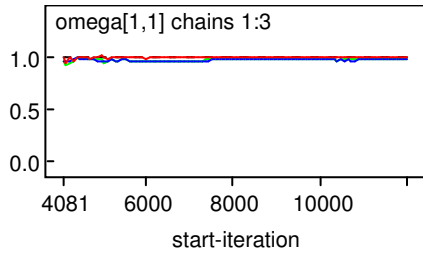
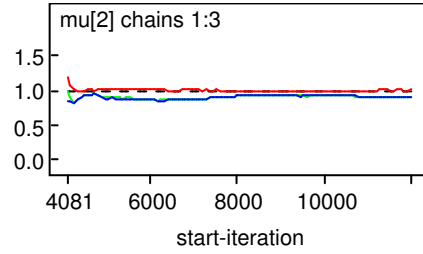
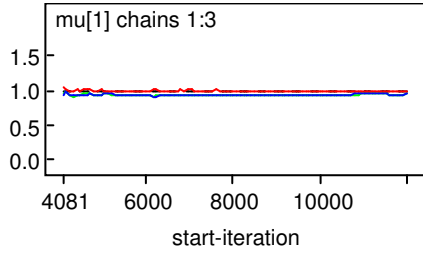
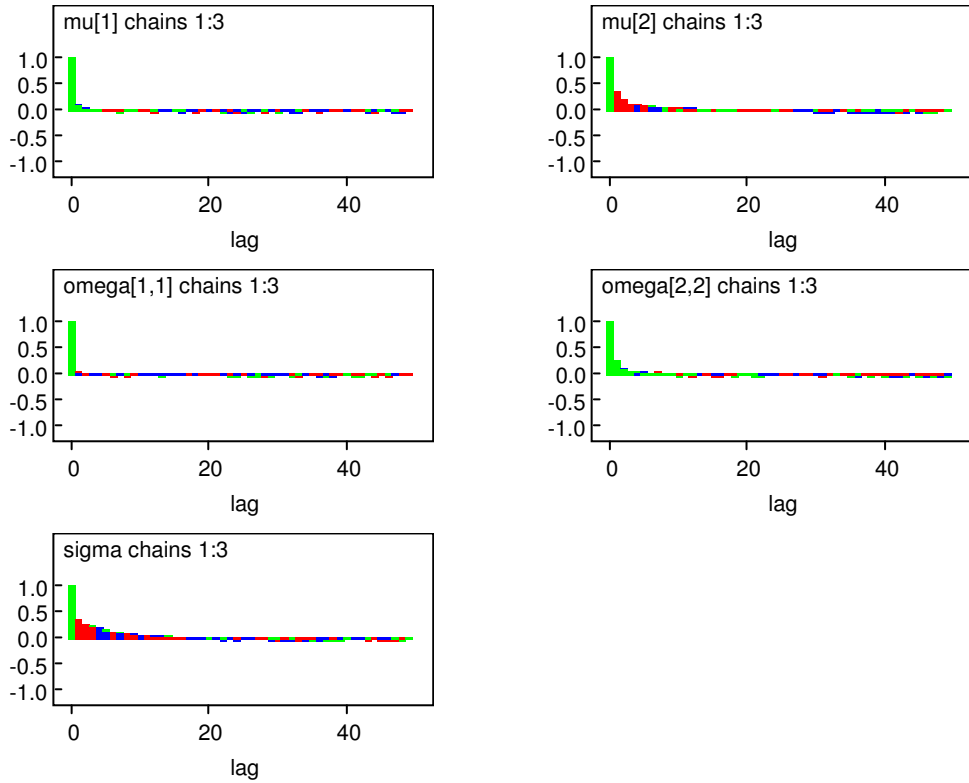
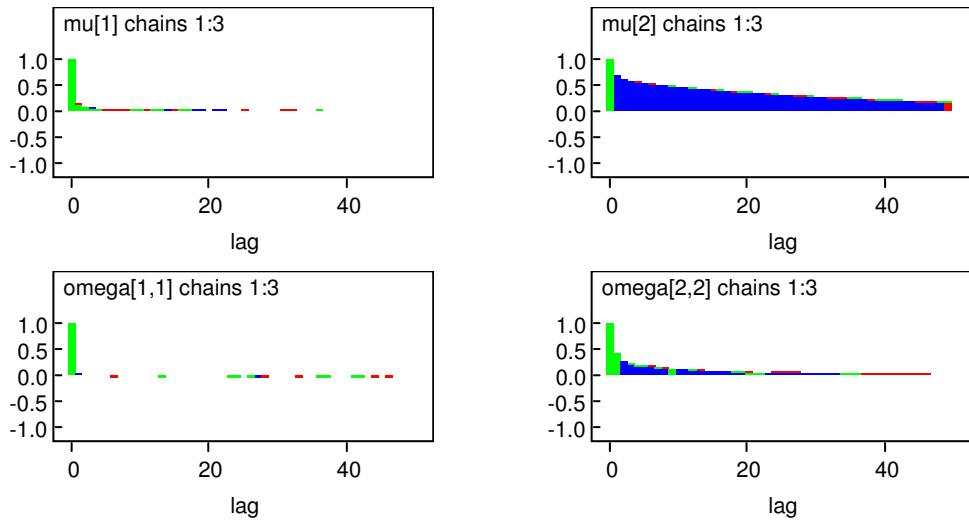


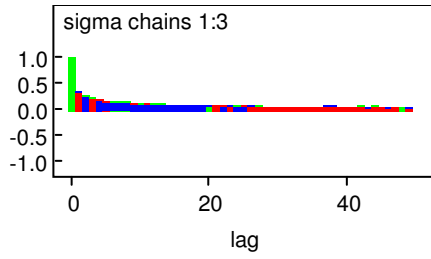
Figure 4. Autocorrelation Plots for Parameters μ , $\Omega (1,1)$, $\Omega (2,2)$, σ^2 :

(A) 1-CP with Additive Error Model



(B) 1-CP with Proportional Error Model





(C) 1-CP with Mixed Error Model

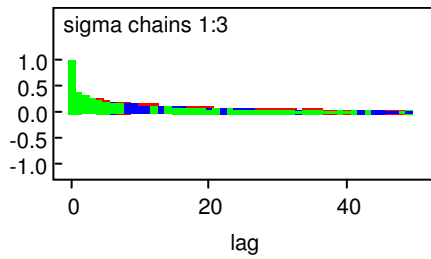
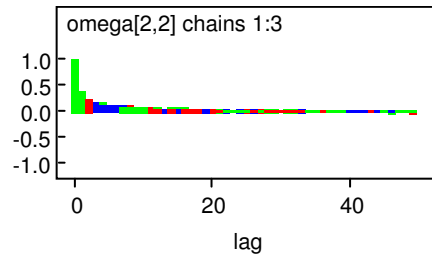
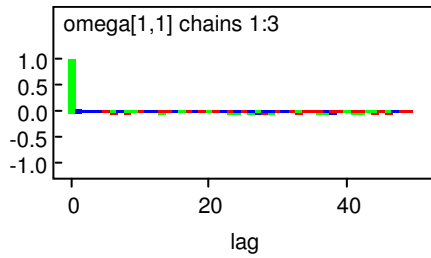
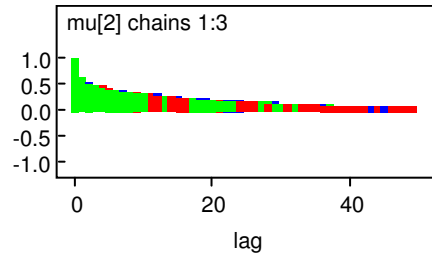
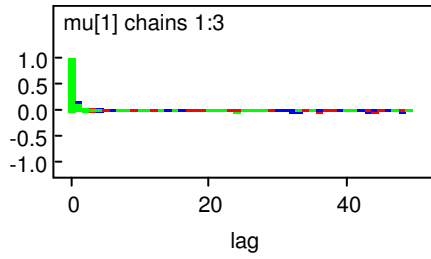
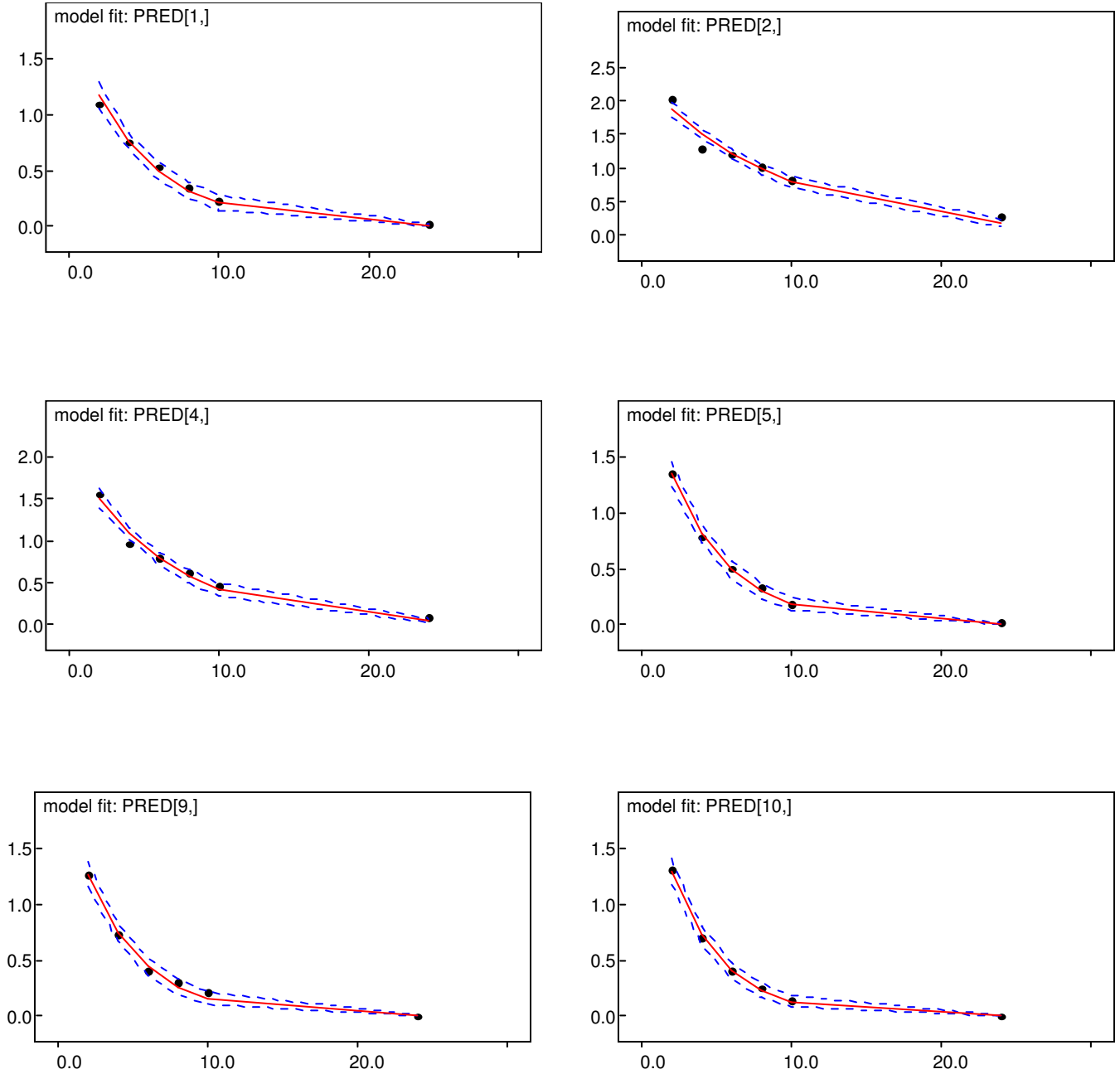
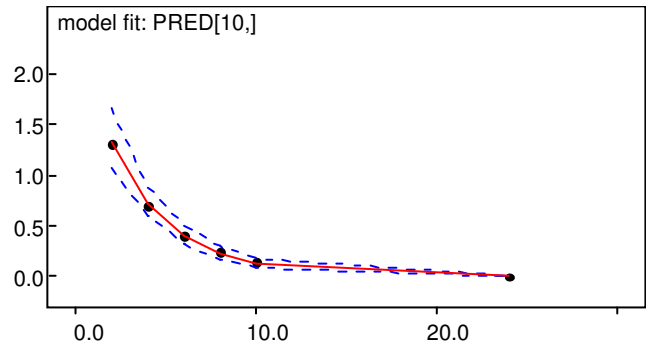
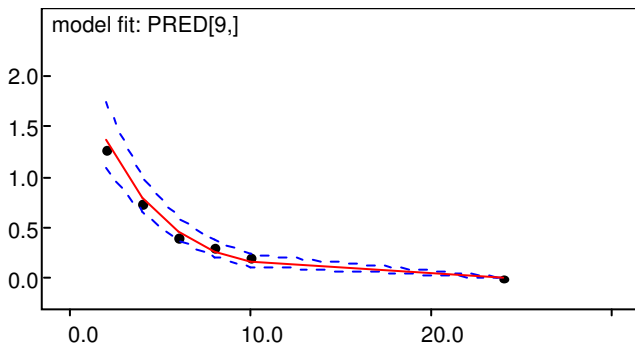
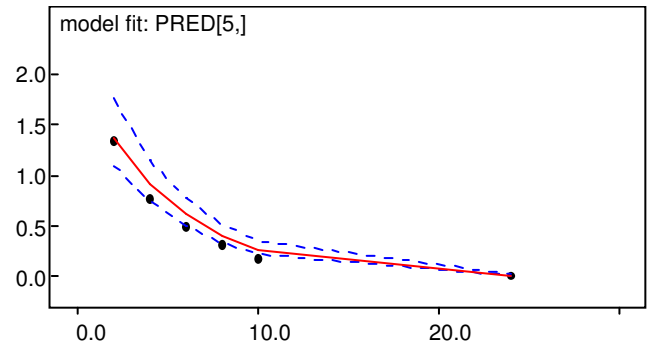
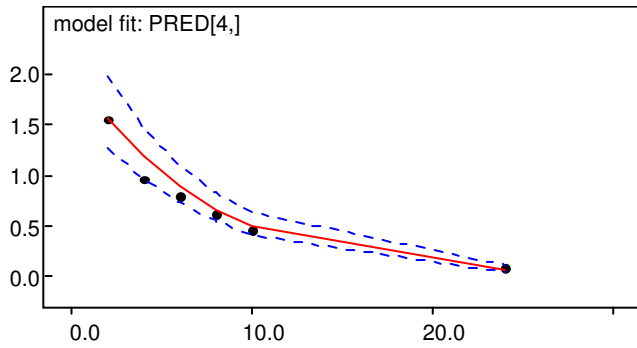
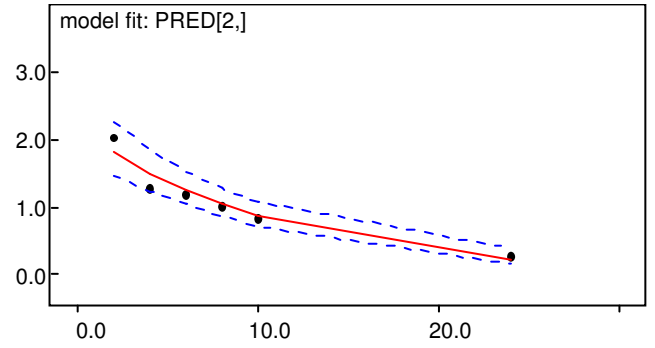
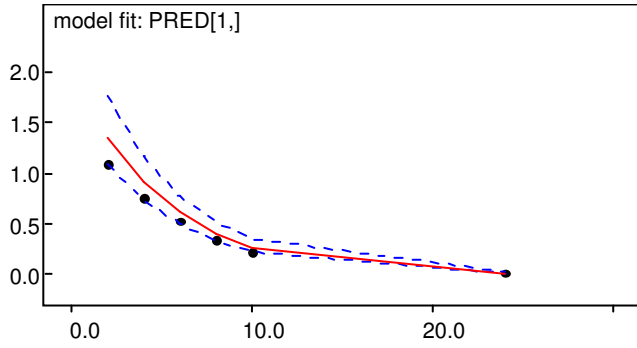


Figure 5. Fit of Bayesian models to Individual Subjects Data

(A) 1-CP with Additive Error Model



(B)1-CP with Proportional Error Model



(C)1-CP with Mixed Error Model

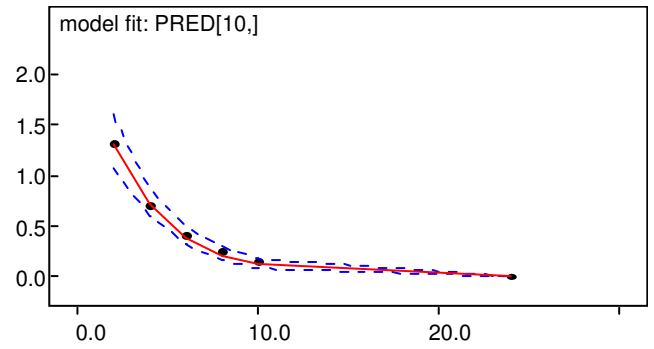
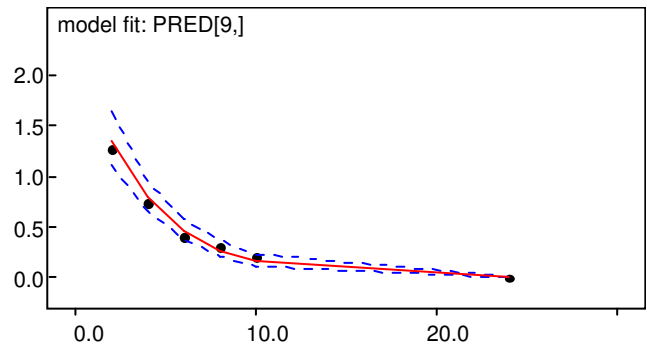
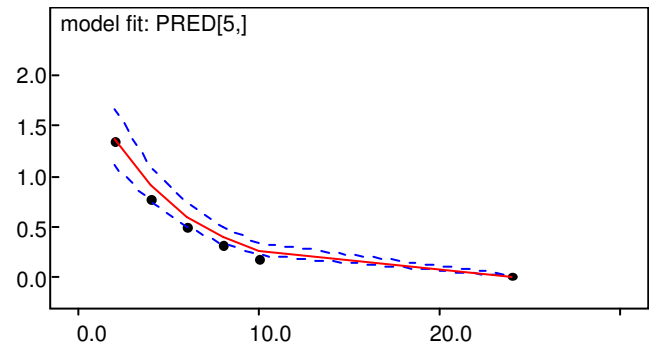
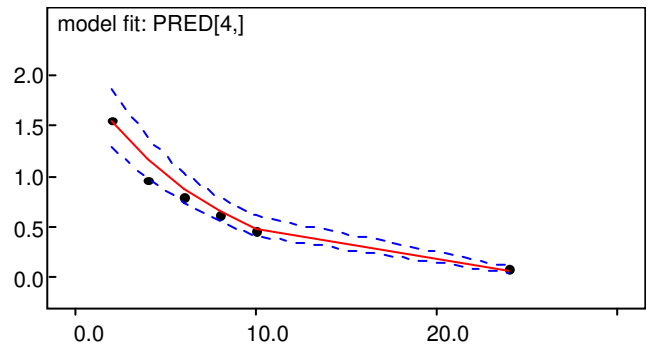
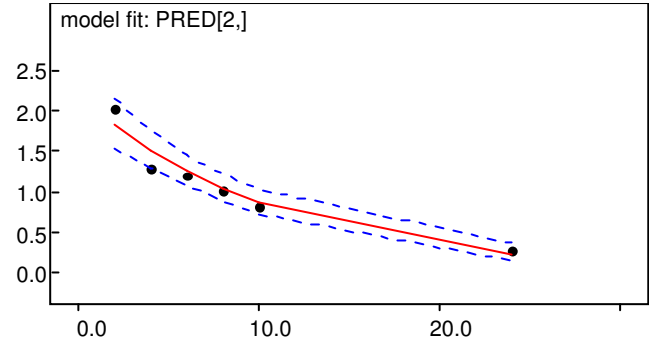
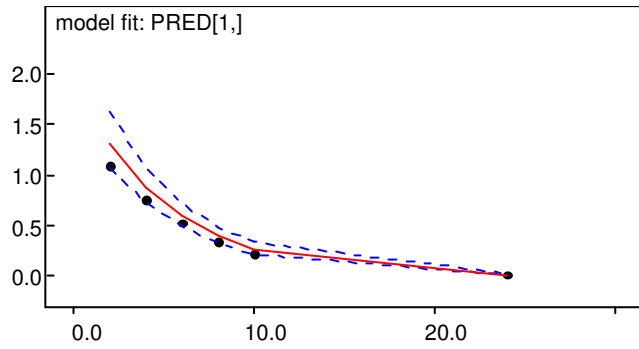
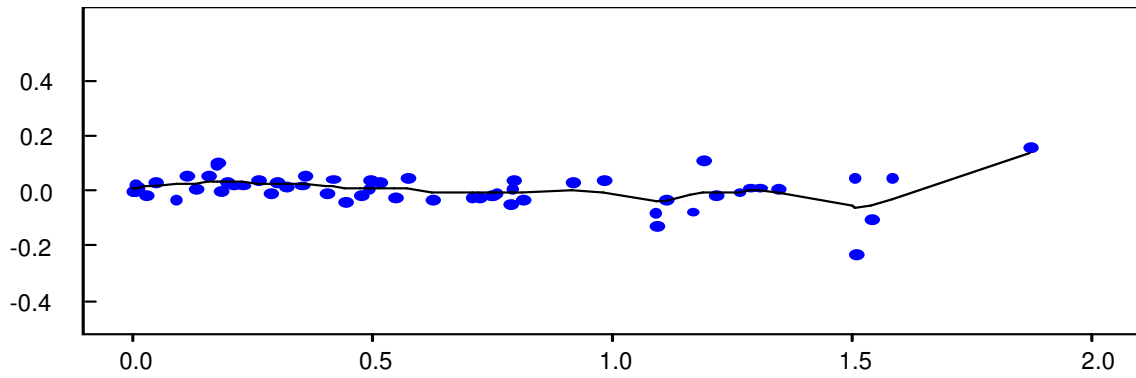
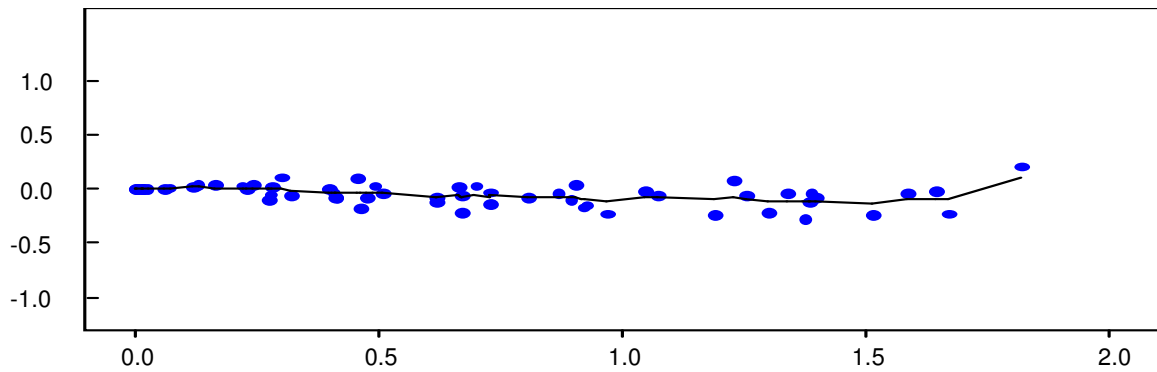


Figure 6. Residual Plots:

(A) 1-CP with Additive Error Model



(B) 1-CP with Proportional Error Model



(C) 1-CP with Mixed Error Model

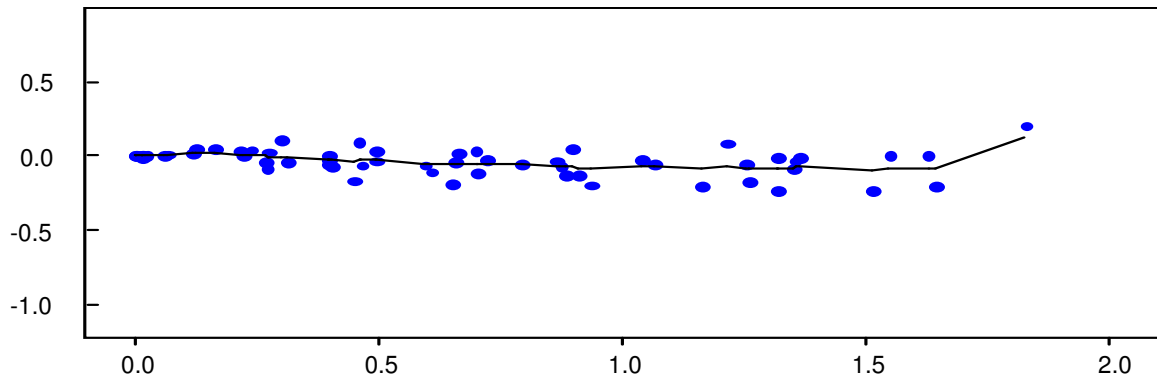


Table 1. DIC Parameters:

(A) 1-CP with Additive Error Model

	Dbar	Dhat	pD	DIC
OBS	-151.258	-170.295	19.037	-132.221
total	-151.258	-170.295	19.037	-132.221

(B) 1-CP with Proportional Error Model

	Dbar	Dhat	pD	DIC
OBS	-92.509	-105.777	13.267	-79.242
total	-92.509	-105.777	13.267	-79.242

(C) 1-CP with Mixed Error Model

	Dbar	Dhat	pD	DIC
OBS	-112.431	-127.27	14.839	-97.592
total	-112.431	-127.27	14.839	-97.592

Table 2. Statistical Summary of Parameters μ , Ω (1,1) , Ω (2,2) , σ^2 :

(A) 1-CP with Additive Error Model

node	mean	sd	MC error	2.50%	median	97.50%	start	sample
mu[1]	1.038	0.1717	0.001037	0.7003	1.038	1.374	8000	36003
mu[2]	2.565	0.05095	4.95E-04	2.459	2.567	2.659	8000	36003
omega[1,1]	0.2904	0.06845	3.31E-04	0.186	0.2807	0.4507	8000	36003
omega[2,2]	0.01846	0.005113	4.04E-05	0.01108	0.01758	0.03082	8000	36003
sigma	0.06931	0.007986	9.00E-05	0.05571	0.06858	0.08703	8000	36003

(D) 1-CP with Proportional Error Model

node	mean	sd	MC error	2.50%	median	97.50%	start	sample
mu[1]	0.9547	0.1761	0.001764	0.61	0.9539	1.303	8000	36003
mu[2]	2.514	0.08073	0.002362	2.353	2.515	2.67	8000	36003
omega[1,1]	0.2877	0.06801	3.58E-04	0.1841	0.2781	0.4468	8000	36003
omega[2,2]	0.02126	0.006686	1.02E-04	0.01194	0.02005	0.03771	8000	36003
sigma	0.3204	0.03697	5.91E-04	0.257	0.3171	0.4021	8000	36003

(E) 1-CP with Mixed Error Model

node	mean	sd	MC error	2.50%	median	97.50%	start	sample
mu[1]	0.9699	0.1737	0.001351	0.628	0.9703	1.312	8000	36003
mu[2]	2.536	0.06989	0.001652	2.392	2.538	2.667	8000	36003
omega[1,1]	0.2882	0.06805	3.43E-04	0.1844	0.2786	0.4465	8000	36003
omega[2,2]	0.02045	0.006148	7.98E-05	0.01167	0.01936	0.03551	8000	36003
sigma	0.3413	0.04122	7.02E-04	0.2721	0.3374	0.434	8000	36003

Table 3. NONMEM Minimum Value of Objective Function :

Model	Objective Function
1-CP with Additive Error Model	-236.863
1-CP with Proportional Error Model	-196.536
1-CP with Mixed Error Model	-282.524

Table 4. Bayesian versus Frequentist Estimates of Parameters:

Model	Parameter	Bayesian (mean \pm s.d)	Frequentist (mean \pm s.e)
1-CP with Additive Error Model	Theta[1]	2.82 \pm 0.1717	2.87 \pm 0.318
	Theta[2]	13 \pm 0.05095	14.5 \pm 0.559
	Omega[1,1]	0.2904 \pm 0.06845	0.12 \pm 0.0483
	Omega[2,2]	0.01846 \pm 0.005113	0.00644 \pm 0.00585
	Sigma	0.06931 \pm 0.007986	0.00417 \pm 0.00188
1-CP with Proportional Error Model	Theta[1]	2.598 \pm 0.1761	3.08
	Theta[2]	12.354 \pm 0.08073	18.4
	Omega[1,1]	0.2877 \pm 0.06801	0.116
	Omega[2,2]	0.02126 \pm 0.006686	0.00079
	Sigma	0.3204 \pm 0.03697	0.123
1-CP with Mixed Error Model	Theta[1]	2.638 \pm 0.1737	2.84 \pm 0.341
	Theta[2]	12.629 \pm 0.06989	15.3 \pm 0.637
	Omega[1,1]	0.2882 \pm 0.06805	0.128 \pm 0.0554
	Omega[2,2]	0.02045 \pm 0.006148	0.00194 \pm 0.00509
	Sigma	0.3413 \pm 0.04122	$\epsilon_1 = 0.00675\pm 0.00178$ $\epsilon_2 = 0.00037\pm 0.000187$

References:

1. Heikkila HJ: New Models for Pharmacokinetic Data based on a Generalized Weibull Distribution. *J Biopharm Stat.* 9:89-107, 1999
2. Lindstrom MJ, Bates DM: Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics* 46:673-687, 1990.
3. Bauer RJ, Guzys, Ng C: A Survey of Population Analysis Methods and Software for Complex Pharmacokinetic and Pharmacodynamic Models with Examples. *AAPS* 9:60-83, 2007.
4. Dufull S.B., Kirkpatrick CMJ, Green B, Holford NHG : Analysis of population pharmacokinetic data using NONMEM and WinBUGS, *J. Biopharm Stat.* 15:53-73, 2005.
5. Wakefield JC, Smith AFM, Racine-Poon A, Gelfand AE: Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Applied Stat.* 43:201-221, 1994.
6. Haglund K, Dahlquist R, Emilsson H, Englund G: Cadralazine Pharmacokinetics- A Pilot Study. *Eur J. Clin. Pharmacokinet.* 35:571-572, 1988.
7. Lecaillon JB, Dubois JP, Darragon T, Motolese M, Racine A, Ducret F, Grouberman D, Cordonnier D, Chanard J, Golorioso M: Pharmacokinetics of Cadralazine in a Large Group of Hypertensive Patients Chronically Treated with Cadralazine. Advantages over a Conventional Study in a Small Group of Patients. *Ther Drug Monit.* 13:103-108, 1991.
8. Hauffe SA, Dubois JP, Imhof PR: Human Pharmacokinetics of Cadralazine: A New Vasodilator. *Eur J Drug Metab. Pharmacokinet.* 10:217-223, 1985
9. Leonetti G, Parini J, Visconti M, Gradnik R: Pharmacokinetics of Cadralazine in Hypertensive Patients. *Eur J Drug Metab Pharmacokinet.* 13:295-300, 1988.
10. Shammas FV, Dickstein K: Clinical Pharmacokinetics in Heart Failure. An Updated Review. *Clin Pharmacokinet.* 15:94-113, 1988.

Appendices:**Appendix I. CADRALAZINE Pharmacokinetic Data Set**

<i>Patient</i>	<i>Results for the following times in hours after administration:</i>							
	<i>2</i>	<i>4</i>	<i>6</i>	<i>8</i>	<i>10</i>	<i>24</i>	<i>28</i>	<i>32</i>
1	1.09	0.75	0.53	0.34	0.23	0.02		
2	2.03	1.28	1.20	1.02	0.83	0.28		
3	1.44	1.30	0.95	0.68	0.52	0.06		
4	1.55	0.96	0.80	0.62	0.46	0.08		
5	1.35	0.78	0.50	0.33	0.18	0.02		
6	1.08	0.59	0.37	0.23	0.17	0.00		
7	1.32	0.74	0.46	0.28	0.27	0.03	0.02	0.00
8	1.63	1.01	0.73	0.55	0.41	0.01	0.06	0.02
9	1.26	0.73	0.40	0.30	0.21	0.00		
10	1.30	0.70	0.40	0.25	0.14	0.00		

Appendix II. WBdiff Code for Bayesian Model
model {

```

for (i in 1:n.ind) {
  for (j in 1:n.grid) {
    OBS[i, j] ~ dnorm(PRED[i, j], tau[i,j])

    #Additive Error Model
    tau[i,j]<-ta

    #Proportional Error Model
    tau[i,j]<-ta*PRED.sq.inv[i,j]
    PRED.sq[i,j]<-pow(PRED[i,j],2)
    PRED.sq.inv[i,j]<-1/(PRED.sq[i,j])

    #Mixed Error Model
    tau[i,j]<-ta*(1+PRED.sq.inv[i,j])
    PRED.sq[i,j]<-pow(PRED[i,j],2)
    PRED.sq.inv[i,j]<-1/(PRED.sq[i,j])

    RESID[i,j]<-OBS[i,j] - PRED[i,j]

    PRED[i, j] <- c.language[i,j,1]
  }
  theta[i, 1:p] ~ dnorm(mu[1:p], omega.inv[1:p, 1:p])

  CL[i] <- exp(theta[i, 1])

  V[i] <- exp(theta[i, 2])

  c.language[i, 1:n.grid, 1:dim] <- ode.block(inits[i, 1, 1:dim], grid[1:n.grid],
    D(A[i, 1:dim], t), origins[i, 1:n.block], tol)

  D(A[i, 1], t) <- -(CL[i]/V[i])*A[i,1]
  inits[i, 1, 1] <- D[i]/V[i]
  inits[i, 1, 2] <- 0
  origins[i, 1] <- 0
}
ta ~ dgamma(tau.a, tau.b)
sigma <- 1 / sqrt(ta)
mu[1:q] ~ dnorm(mu.prior.mean[1:q], mu.prior.precision[1:q, 1:q])
mu.prior.precision[1:q, 1:q] <- inverse(mu.prior.variance[1:q, 1:q])
omega.inv[1:p, 1:p] ~ dwish(omega.inv.matrix[1:p, 1:p], omega.inv.dof)
omega[1:p, 1:p] <- inverse(omega.inv[1:p, 1:p])
}

```

Appendix III.**Section. A** Cadralazine Clearance and Volume of Distribution from Literature

Study	No. of subjets	Clearance	Volume of Distribution	Notes
Ref(6)	5	$\mu = 19.788$ $SD = 15.77$	$\mu = 38.5$ $SD = 18.17$	CL used is renal CL at 0-8 hours V was calculated as CL times half life divided by 0.693
Ref(7)	100	$\mu = 11.58$ $SD = 15$	$\mu = 57.6$ $SD = 15$	No SD was provided. Since the CL and V variations in healthy subjects are affected by weight. The SD of the parameters was approximated by the SD of the subjects weights. The medians of parameters were used because the means were not provided
Ref(8)	17	$\mu = 16.42$ $SD = 1.82$	$\mu = 56$ $SD = 0.85$	V was calculated as CL times half life divided by 0.693
Ref(9)	6	$\mu = 10.85$ $SD = 2.74$	$\mu = 46.7$ $SD = 13.6$	V was calculated as CL times half life divided by 0.693

Section. B Bayesian Priors

$$\boldsymbol{\mu} \sim \mathbf{N} \left\{ \bar{\boldsymbol{\mu}} = [0.92772, 1.740201], \Sigma = \begin{bmatrix} 0.33902 & 0 \\ 0 & 0.01812 \end{bmatrix} \right\}$$

$$\boldsymbol{\Omega}^{-1} \sim \mathbf{W} \left\{ R = \begin{bmatrix} 9.83158 & 0 \\ 0 & 0.52548 \end{bmatrix}, \rho = 32 \right\}$$

$$\boldsymbol{\tau} \sim \mathbf{G} \{0.0001, 0.001\}$$

Section. C Bayesian Initial Values*Centered Initial Values:*

$$\boldsymbol{\mu} = [0.93, 1.74]$$

$$\boldsymbol{\Omega}^{-1} = \begin{bmatrix} 9.83 & 0 \\ 0 & 0.525 \end{bmatrix}$$

$$\boldsymbol{\tau} = 0.1$$

$$\boldsymbol{\theta} = \mathbf{2} \times \mathbf{10} = \begin{bmatrix} 0.93 & 1.74 \\ \vdots & \vdots \\ 0.93 & 1.74 \end{bmatrix}$$

Overdispersed (High) Initial Values:

$$\boldsymbol{\mu} = [1.395, 2.61]$$

$$\boldsymbol{\Omega}^{-1} = \begin{bmatrix} 14.75 & 0 \\ 0 & 0.788 \end{bmatrix}$$

$$\boldsymbol{\tau} = 0.15$$

$$\boldsymbol{\theta} = \mathbf{2} \times \mathbf{10} = \begin{bmatrix} 1.395 & 2.61 \\ \vdots & \vdots \\ 1.395 & 2.61 \end{bmatrix}$$

Overdispersed (Low) Initial Values:

$$\boldsymbol{\mu} = [0.465, 0.87]$$

$$\boldsymbol{\Omega}^{-1} = \begin{bmatrix} 4.9 & 0 \\ 0 & 0.263 \end{bmatrix}$$

$$\boldsymbol{\tau} = 0.05$$

$$\boldsymbol{\theta} = \mathbf{2} \times \mathbf{10} = \begin{bmatrix} 0.465 & 0.87 \\ \vdots & \vdots \\ 0.465 & 0.87 \end{bmatrix}$$

Appendix IV. NONMEM VI Control Stream File for Model Fitting

```

$PROBLEM          CAD DATA BAYESIAN PROJECT VERSION 1
$DATA            CadNM.TXT          IGNORE=#
$INPUT           ID   AMT TIME DV   MDV
$SUBROUTINES     ADVAN1   TRANS2
$PK
  CL = THETA(1)*EXP(ETA(1))
  V = THETA(2)*EXP(ETA(2))
  S1 = V
$ERROR
  IPRED = F
; Additive Error Model
  Y = F+EPS(1)
; Proportional Error Model
; Y = F+F*EPS(1)
; Mixed Error Model
; Y = F+F*EPS(1)+EPS(2)
$THETA
  (0.0, 12.51)      ; THETA(1)
  (0.0, 56.14)     ; THETA(2)
$OMEGA
  12.704
  12.54
$SIGMA
  0.1
$ESTIMATION
  METHOD = CONDITIONAL
  INTERACTION
  SLOW
  MAXEVAL=10000      PRINT=5
$COVARIANCE
  SLOW
$TABLE
  ID   TIME DV   IPRED      ETA1 CL   V
  NOPRINT ONEHEADER FILE = prob1_table.txt

```