Analysis of Longitudinal Data using Mixed Models[a]

Kouros Owzar

March 20, 2001

OUTLINE:

This presentation aims to address the following topics:

1. To introduce and motivate the basic concept behind linear and non-linear mixed models to model longitudinal data using a number of examples [a]
2. To briefly address an analog Bayesian approach.
3. To give an overview of the available software available to fit these kinds of models.
4. To provide a list of representative literature on these types of models.

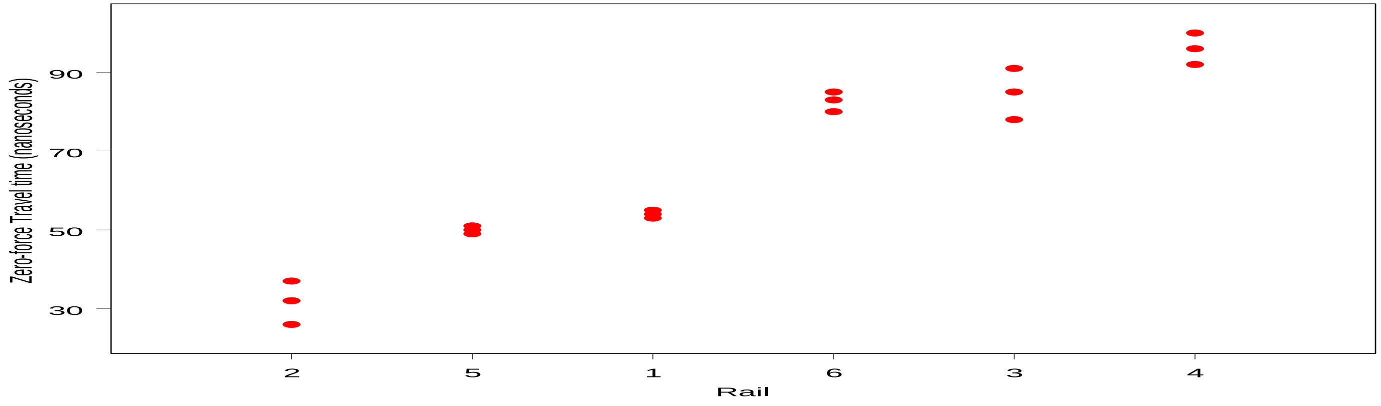
Discussions on the theoretical aspects are largely omitted (perhaps best left for a second presentation).

---

[a] All examples are taken from Pinheiro and Bates (2000)

Example I: Rail Data

1. Six randomly chosen rails were each tested three times for stress.
2. The test measured the length of time needed for a certain wave to travel from one end of the rail to the other.
3. The average travel time (call it β) for a typical rail is of interest.
4. There is apparent variability between the rails and within each rail there is variability between the three measurements.



Example I: Rails (cont.)

Let y_{ij} denote the j-th measurement for the i-th rail. The ϵ_{ij}'s are iid $N(0, \sigma_\epsilon^2)$ measurement errors.

Consider the following three models:

I. $y_{ij} = \beta + \epsilon_{ij}$ (β fixed parameter)

II. $y_{ij} = \beta_i + \epsilon_{ij}$ (β_i fixed parameters)

III. $y_{ij} = \beta + b_i + \epsilon_{ij}$ (β fixed parameter and b_i are $N(0, \sigma_b^2)$)

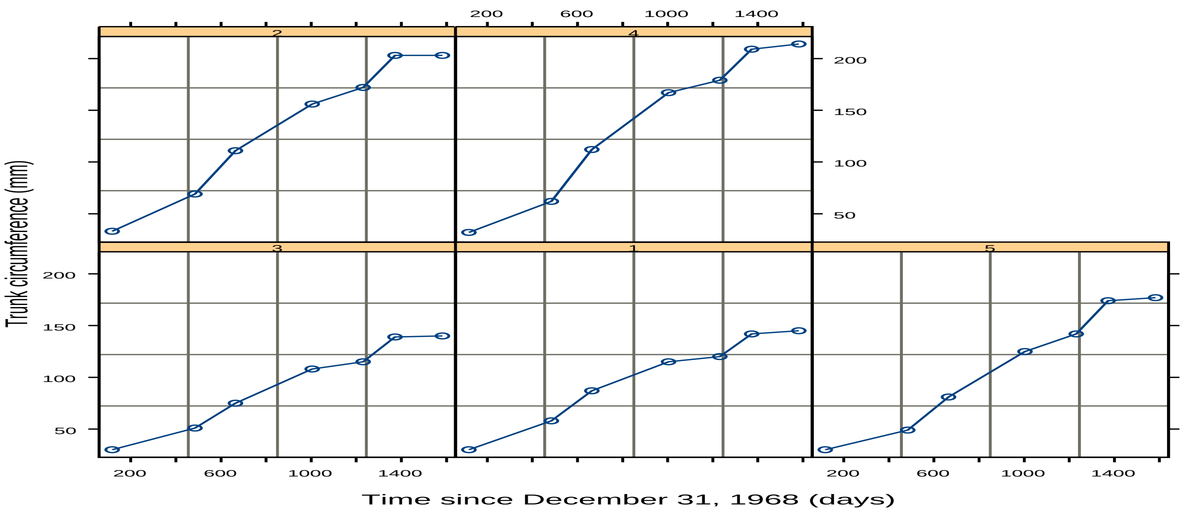
Note that:

1. Model (I) ignores the grouping structure.
2. Model (II) estimates individual rail travel times not the average travel time for a typical rail.
3. For Model (III) $\hat{\beta} = 66.5, \hat{\sigma}_b = 24.8$ and $\hat{\sigma}_\epsilon = 4.02$.

Example II: Orange Trees

1. The trunk circumference of 5 trees were measured over time.
2. The growth curve seems to be linear over time.
3. The variability between the growth curves (lines) could be explained using the following model:

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) + \epsilon_{ij}$$



Example I (Likelihood Function):

Write $\theta = \sqrt{\sigma_\epsilon^2/\sigma_b^2}$ (often called the precision factor). The likelihood function is given by

$$
\begin{aligned}
L(\beta, \theta, \sigma_\epsilon^2 | Y) &= \prod_{i=1}^{6} f(y_i|\beta, \sigma_\epsilon^2, \theta) = \prod_{i=1}^{6} \int f(y_i|\beta, \sigma_\epsilon^2, b_i) f(b_i|\sigma_\epsilon^2, \theta) db_i \\
&= \prod_{i=1}^{6} \int \left\{ \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{3}{2}}} \exp\left[-\frac{1}{2\sigma_\epsilon^2} \sum_{j=1}^{3} (y_{ij} - \beta - b_i)^2 \right] \right. \\
&\qquad \left. \times \frac{\theta}{(2\pi\sigma_\epsilon^2)^{\frac{1}{2}}} \exp\left[-\frac{\theta^2 b_i^2}{2\sigma_\epsilon^2} \right] \right\} db_i
\end{aligned}
$$

$$
= \prod_{i=1}^{6} \int \left\{ \frac{1}{2\pi\sigma_\epsilon^2} \exp\left[-\frac{1}{2\sigma_\epsilon^2} \left(\sum_{j=1}^{3} (y_{ij} - \beta - b_i)^2 + \theta^2 b_i^2 \right) \right] \right\} db_i
$$

General Linear Mixed Effects Model

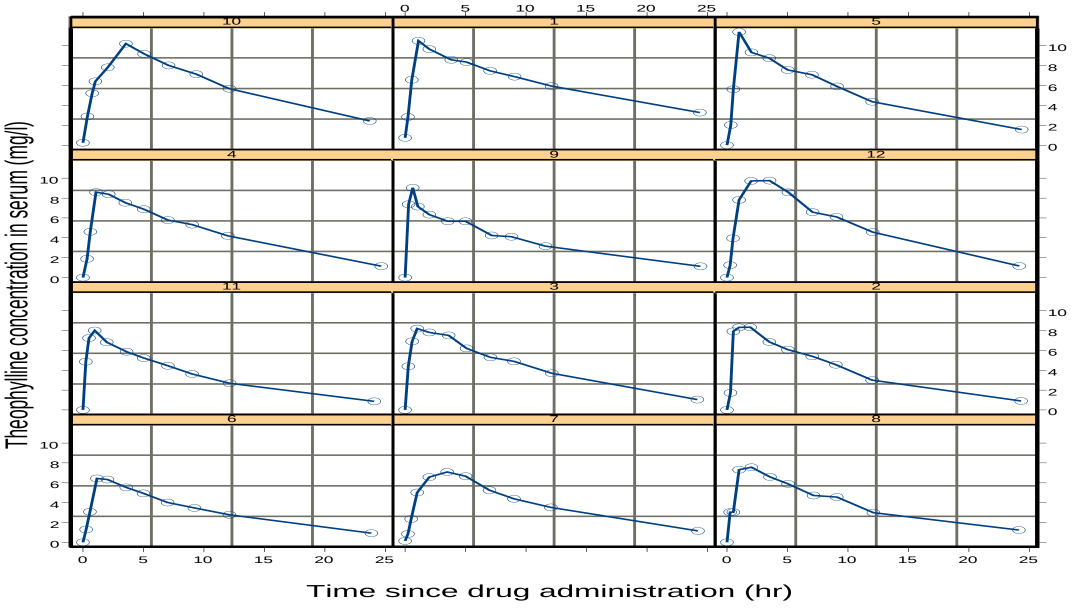
$$Y_i = X_i\beta + Z_i b_i + \epsilon_i, \tag{1}$$

where

1. X_i is the design matrix for the i-th subject;
2. β is the vector of parameters;
3. b_i is the random effects vector for the i-th subject;
4. ϵ_i are the measurement errors for the i-th subject;
5. $b_1, \ldots, b_n$ are iid $N_q(0, \Sigma_b)$
6. $\epsilon_1, \ldots, \epsilon_n$ are iid $N_{n_i}(0, \Sigma_\epsilon)$

Example III: Theophylline Data

1. One dose of theophylline (an anti-asthmatic drug) was administered orally to 12 subjects.
2. Following the initial administration the serum concentrations of the compound were measured 11 times in the following 25 hours.



Example III: Theophylline Data (cont):

The one compartment model with first order absorption and elimination

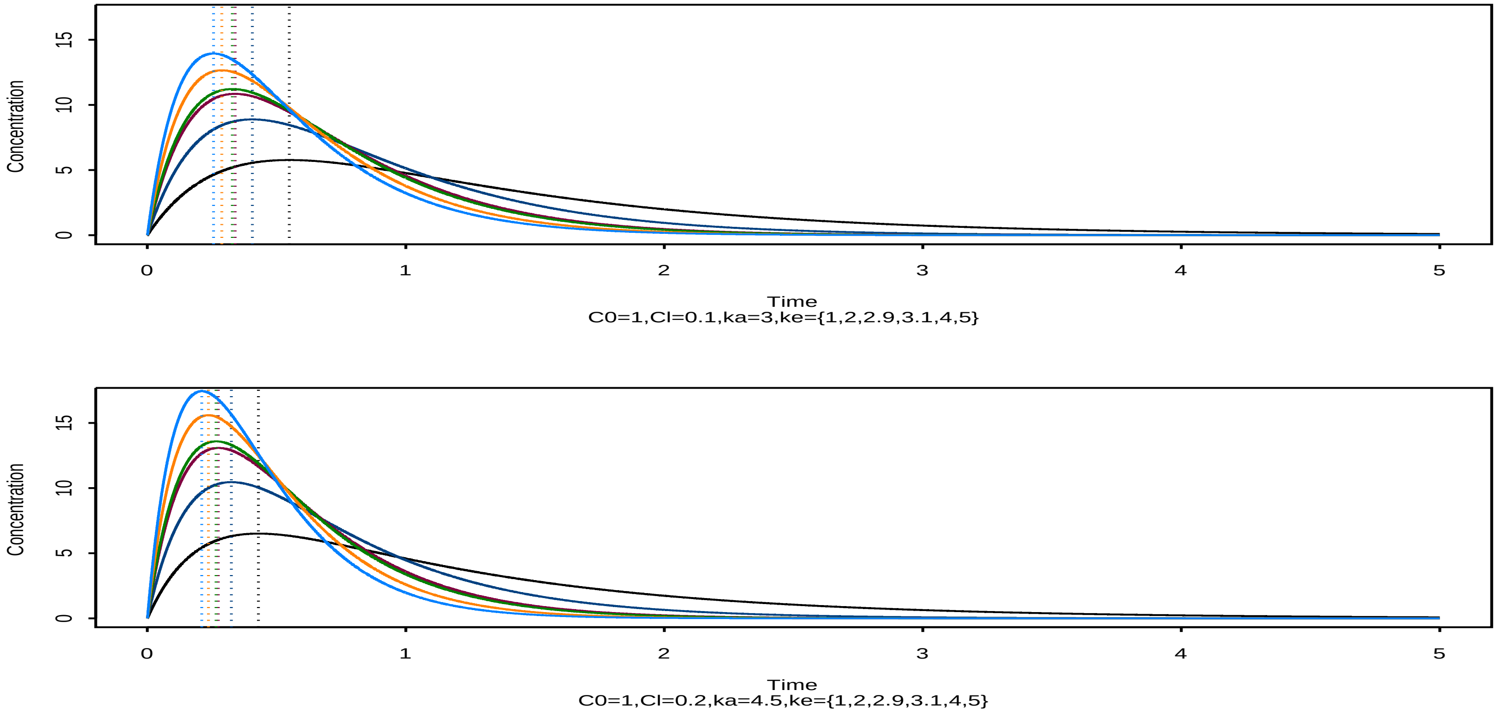
$$C(t) = \frac{C_0 k_e k_a}{Cl(k_a - k_e)} \{\exp(-k_e t) - \exp(-k_a t)\},$$

where

1. $C(t)$ is the concentration of the drug at time t;
2. C_0 is the amount of the drug administered;
3. k_a is rate of absorption;
4. k_e is rate of elimination;
5. Cl is rate of clearance.

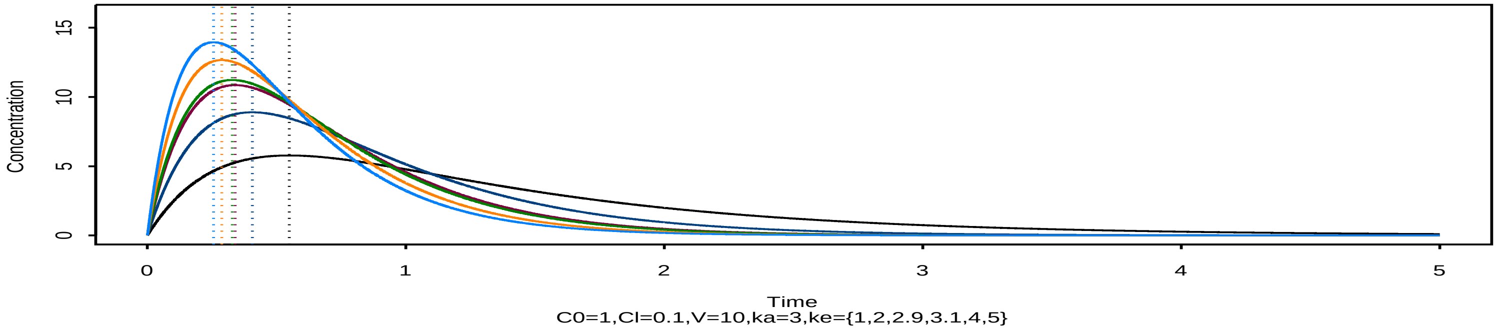
Example III: Theophylline Data (cont):

1. The time at which the concentration attains its maximum $\text{tmax} = (\log(\mathrm{k_a}) - \log(\mathrm{k_e}))/(\mathrm{k_a} - \mathrm{k_e})$ does not depend on Cl, C_0.

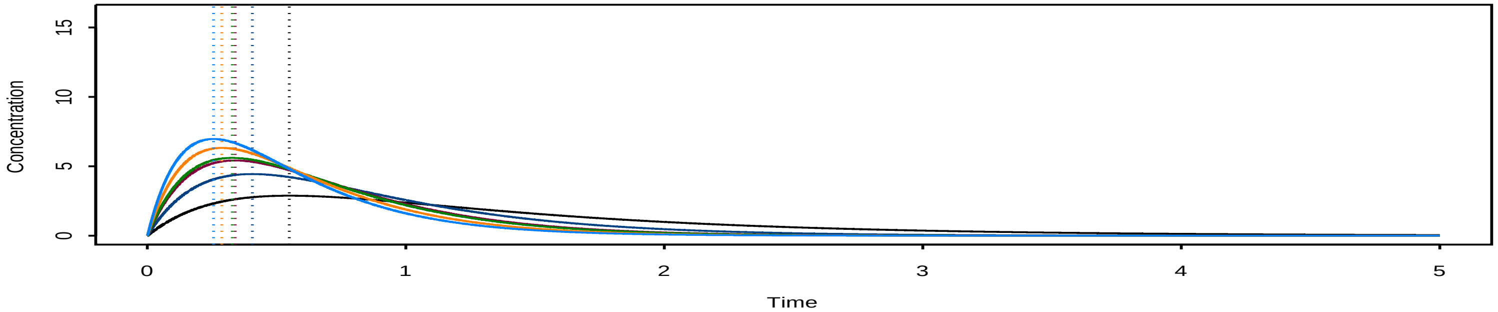


Example III: Theophylline Data (cont):

1. The Area Under the concentration Curve $AUC = C_0/Cl$ does not depend on k_a, k_e.



C0=1,Cl=0.1,V=10,ka=3,ke={1,2,2.9,3.1,4,5}



C0=1,Cl=0.2,V=5,ka=3,ke={1,2,2.9,3.1,4,5}

Example III: Theophylline Data (cont):

The observed concentration for the i-th subject at time t_j is given by:

$$C_i(t_j) = \frac{C_0 \beta_{ei} \beta_{ai}}{Cl(\beta_{ai} - \beta_{ei})} \{\exp(-k_e t) - \exp(-k_a t)\} + e_{ij},$$

where

$$\beta_i = \begin{pmatrix} \beta_{ai} \\ \beta_{ei} \\ \beta_{Cli} \end{pmatrix} = \begin{pmatrix} k_a \\ k_e \\ Cl \end{pmatrix} + \begin{pmatrix} k_{ai} \\ k_{ei} \\ Cl_i \end{pmatrix},$$

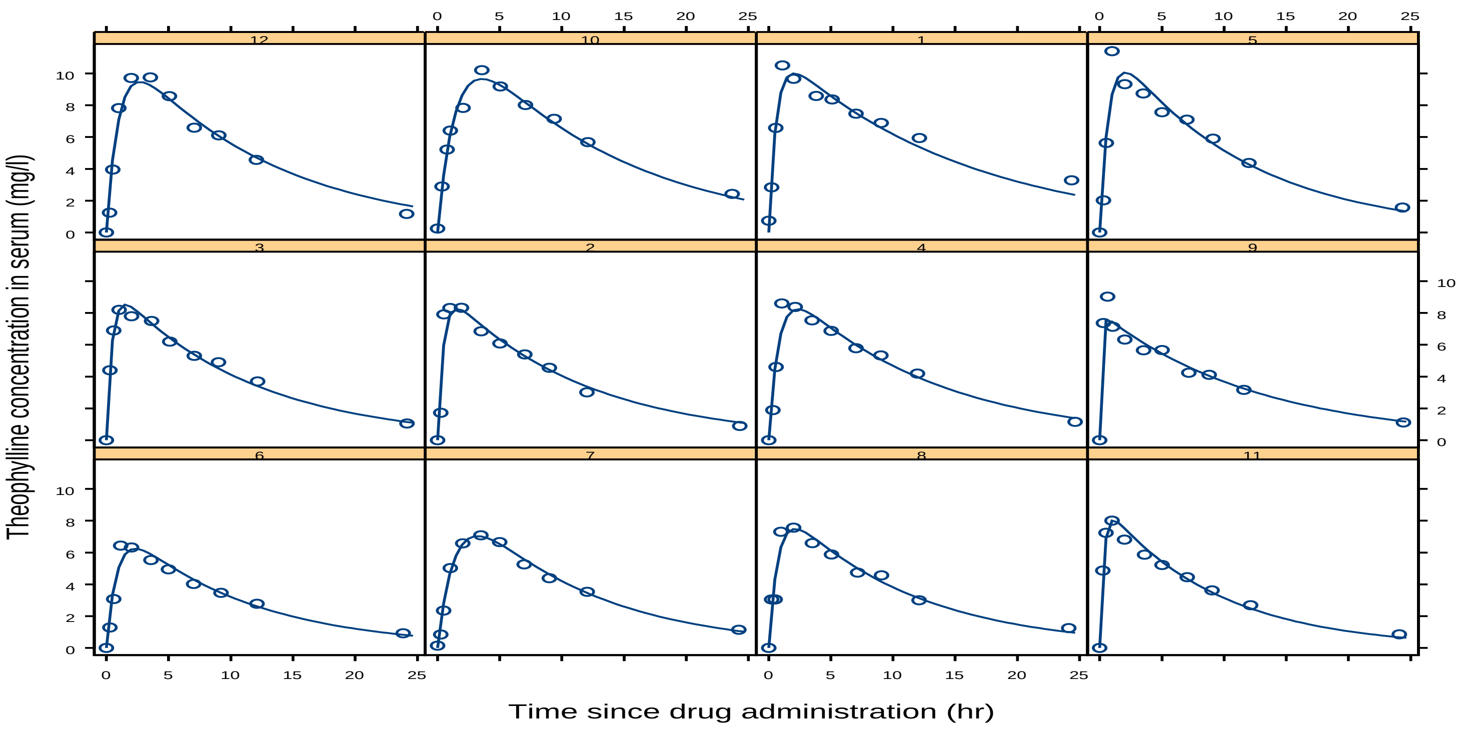
$$b_i = \begin{pmatrix} k_{ai} \\ k_{ei} \\ Cl_i \end{pmatrix} =_d N_3(0, \Sigma_b),$$

$$e_{ij} \text{ iid } \mathrm{N}(0, \sigma_\epsilon^2)$$

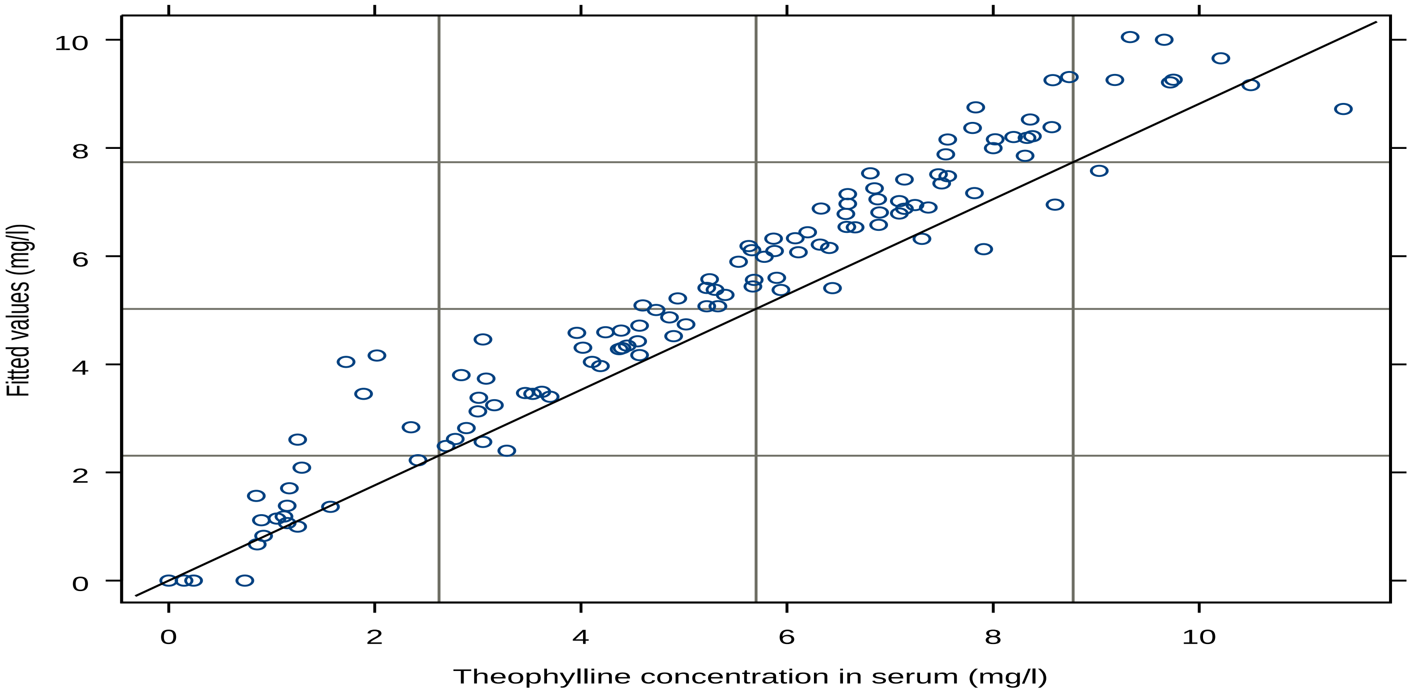
EXAMPLE III: NLME OUTPUT

```
> summary(mod1)
Nonlinear mixed-effects model fit by maximum likelihood
  Model: conc ~ SSfol(Dose, Time, lke, lka, lcl)
 Data: Theoph
       AIC      BIC    logLik
  366.6412 395.4693 -173.3206

Random effects:
 Formula: list(lke ~ 1, lka ~ 1, lcl ~ 1)
 Level: Subject
 Structure: General positive-definite
            StdDev   Corr
     lke 0.1310247 lke    lka
     lka 0.6376679  0.012
     lcl 0.2511744  0.995 -0.089
Residual 0.6818392

Fixed effects: lke + lka + lcl ~ 1
        Value Std.Error  DF   t-value p-value
lke -2.432670 0.0630209 118 -38.60097  <.0001
lka  0.451371 0.1962117 118   2.30043  0.0232
lcl -3.214451 0.0805949 118 -39.88403  <.0001
 Correlation:
       lke    lka
lka -0.142 lcl  0.854 -0.131

Number of Observations: 132 Number of Groups: 12
```

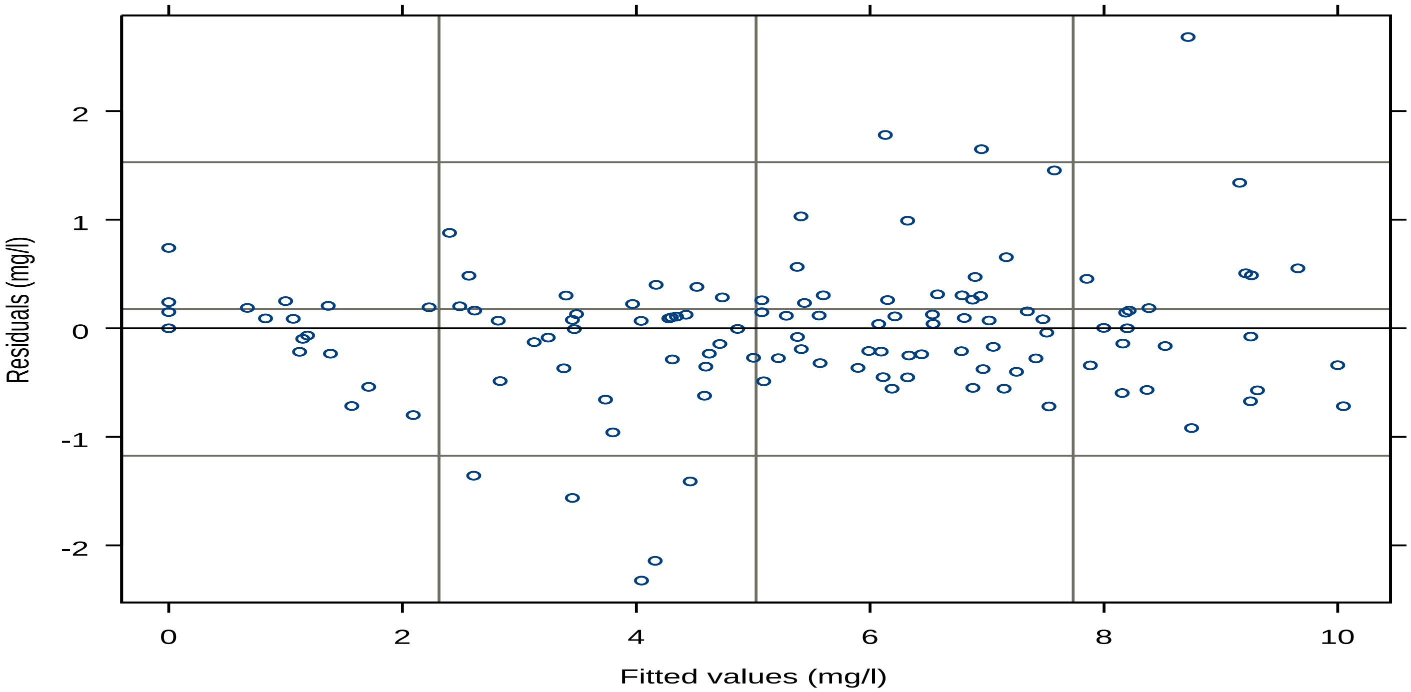
EXAMPLE III: Individual Fits



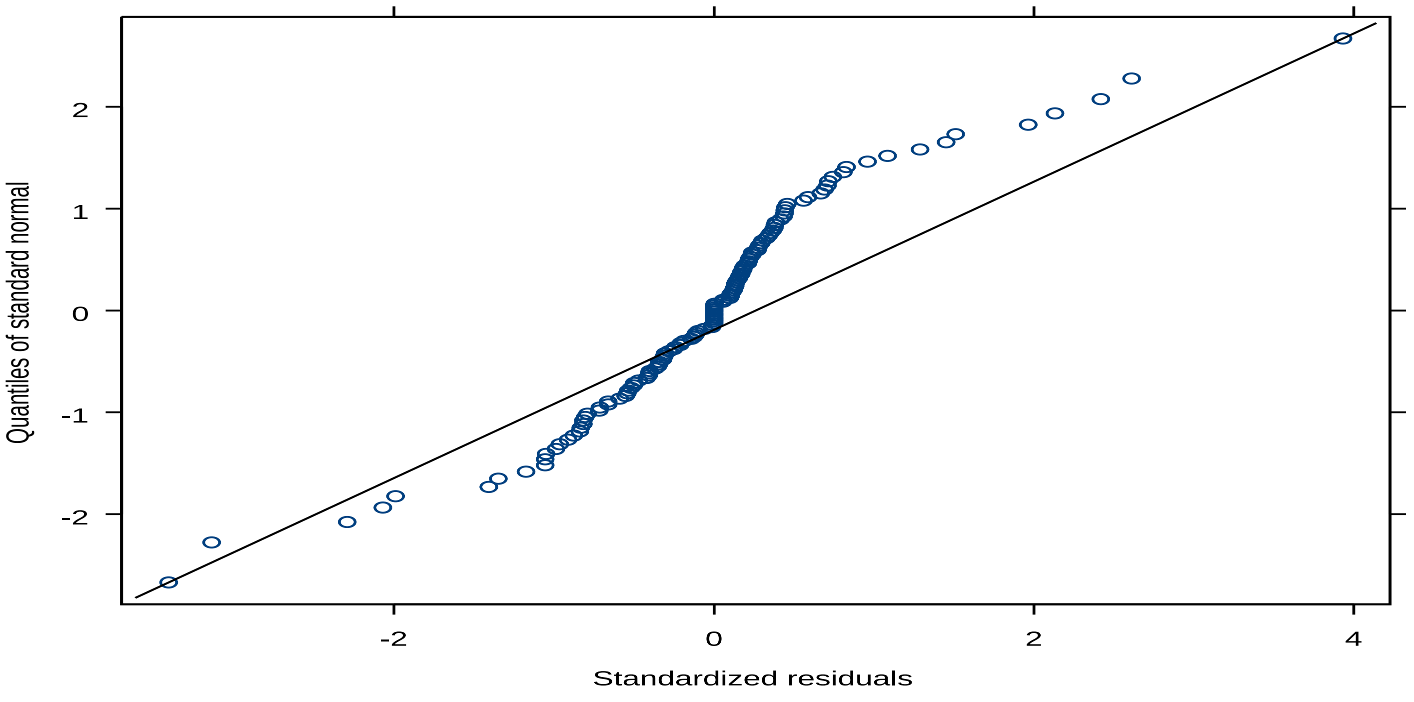
EXAMPLE III: Fitted vs. Observed Values



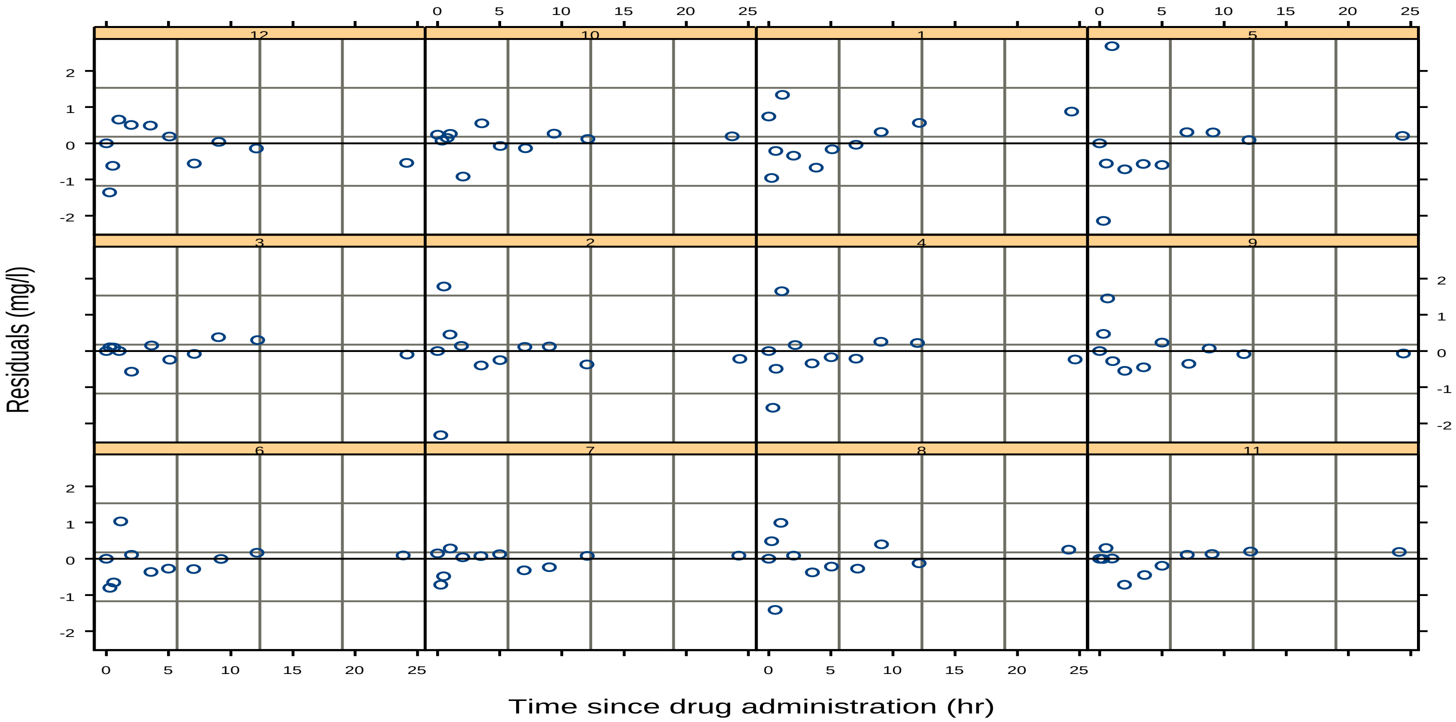
EXAMPLE III: Residuals vs. Fitted



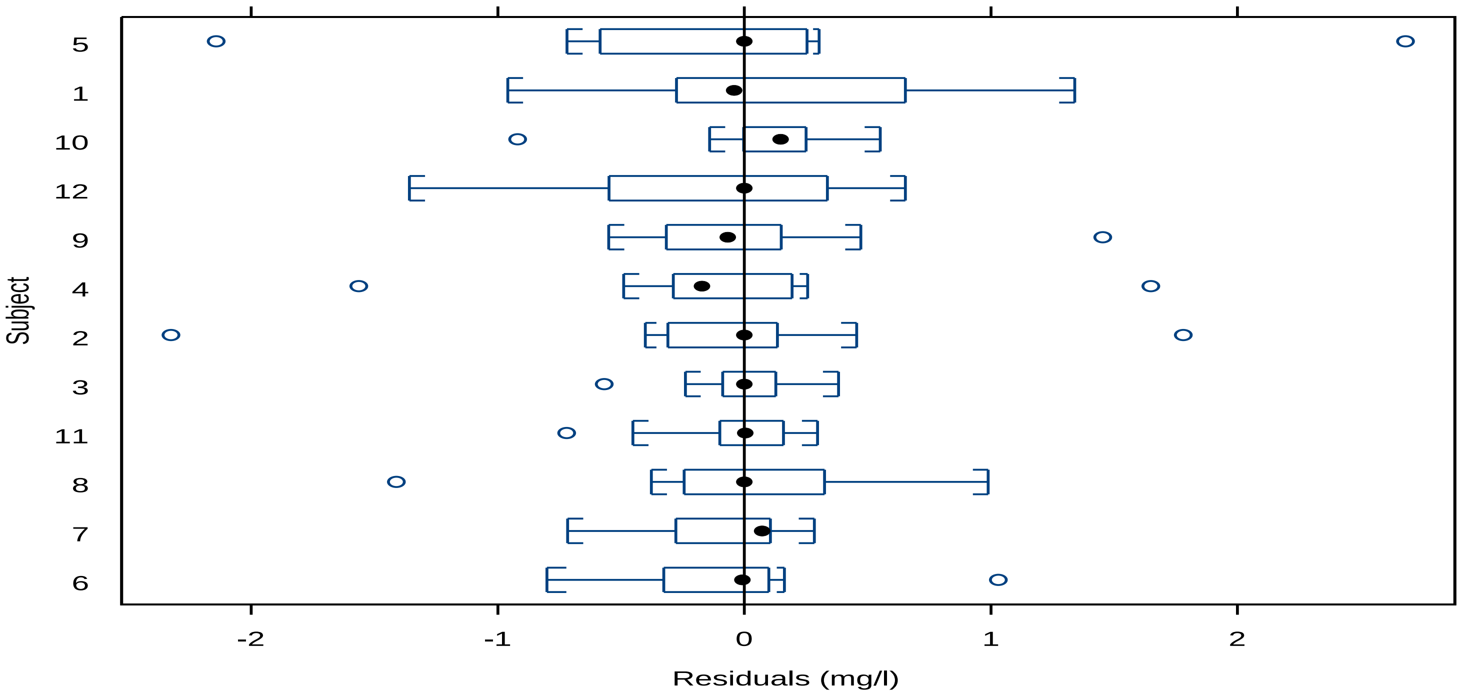
EXAMPLE III: QQ-PLOT of Residuals



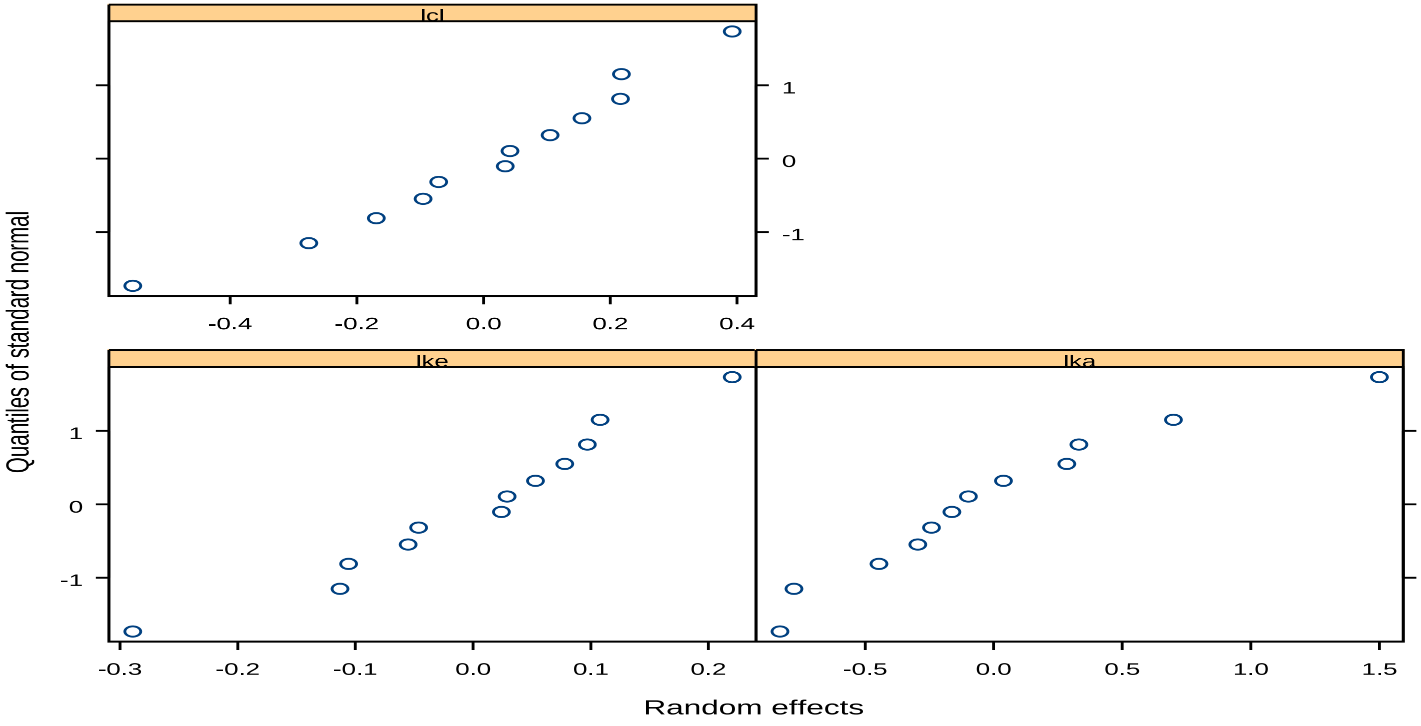
EXAMPLE III: Residuals vs. Time



EXAMPLE III: Boxplots of Individual Residuals



EXAMPLE III: QQ-PLOTS of Random Effects



EXAMPLE III: Reducing the Model

```
> mod11
Nonlinear mixed-effects model fit by maximum likelihood
  Model: conc ~ SSfol(Dose, Time, lke, lka, lcl)
  Data: Theoph
  Log-likelihood: -177.0231
  Fixed: lke + lka + lcl ~ 1
       lke       lka       lcl
 -2.454671 0.4655485 -3.227213

Random effects:
 Formula: list(lcl ~ 1, lka ~ 1)
 Level: Subject
 Structure: General positive-definite
            StdDev   Corr
     lcl 0.1669147 lcl
     lka 0.6438041 -0.002
Residual 0.7092300

Number of Observations: 132
Number of Groups: 12


> anova(mod11,mod1)
      Model df      AIC      BIC    logLik   Test L.Ratio p-value
mod11     1  7 368.0461 388.2257 -177.0231
 mod1     2 10 366.6412 395.4693 -173.3206 1 vs 2 7.40489  0.0601
>
```

Bayesian Hierarchical Model (cont):

1. Bayesian approaches to fitting mixed models have gained increased popularity (mostly due to improved computing power)
2. The idea behind this method is to simulate the model parameters from the posterior distribution.
3. In the Rails example, the model parameters are β, σ_b^2 and σ_ϵ^2.
4. Simulate N realizations $\{\beta^{*i}, \sigma_b^{*i}, \sigma_\epsilon^{*i}\}_i^N$ from the posterior distribution

$$f(\beta, \sigma_b^2, \sigma_\epsilon^2 | y_1, \ldots, y_n) \tag{2}$$

Bayesian Hierarchical Model:

In the Bayesian paradigm the model is not really mixed as all the parameters are random (no fixed effects). The repeated measurement models are expressed using the following hierarchy:

$$\text{STAGE1}: \qquad (Y_i|\beta_i, \sigma_\epsilon) =_d N_3(X_i\beta_i, \sigma_\epsilon I_3).$$

$$\text{STAGE2}: \qquad \beta_i = \beta + b_i \quad , \quad b_i =_d N_2(0, \Sigma_b)$$

$$\text{STAGE3}: \qquad \beta =_d N_2(\beta^\#, \Sigma_\beta^\#) \quad , \quad \Sigma_b^{-1} =_d \text{Wishart}\left[(\rho^\# \Sigma_{\text{b}}^\#)^{-1}, \rho^\#\right]$$

$$\sigma_\epsilon^{-2} =_d \text{Gamma}\left[\frac{\nu^\#}{2}, \frac{\nu^\# \tau^\#}{2}\right],$$

where Wishart[.,.] and Gamma[.,.] denote the Wishart and Gamma distributions respectively. The parameters $\beta^\#, \Sigma_\beta^\#, \rho^\#, \Sigma_b^\#, \nu^\#$ and $\tau^\#$ are assumed to be known.

Software:

1. The lme and nlme packages are freely available for fitting linear and non-linear mixed effects models in S-Plus and R.
2. SAS offers PROC MIXED and PROC NLMIXED for fitting linear mixed effects models and non-linear mixed effects models
3. The lme and nlme packages have built in diagnostic facilities.
4. Unlike nlme, PROC NLMIXED does not provide facilities for imposing structures on the covariance matrices Σ_ϵ and Σ_b.
5. PROC MIXED is very efficient for large scale simulations.
6. WinBUGS (Bayesian Analysis Using Gibbs Sampler) is freely available to fit Bayesian Hierarchical models. Its accompanying documentation provides several examples (including the modeling of the Orange data).

Comments and Remarks:

1. The procedures for fitting non-linear models require the assignment of good starting values. This is by no means a trivial task.

2. With (n)lme one can specify correlation structures on Σ_e as well as Σ_b

3. With (n)lme one can model unequal residual variances (e.g., as a function of time).

4. Graphical methods are typically used to determine if one should use impose these structures.

Literature and Acknowledgments:

The following books provide comprehensive information about Gaussian Mixed models. They will also give references to the original papers.

1. Davidian And Giltinan (1995), *"Nonlinear Models for Repeated Measurement Data"*;
2. Littell, Milliken, Stroup and Wolfinger (1996), *"SAS System for Mixed Models"*;
3. Vonesh and Chinchilli (1997), *"Linear and Nonlinear Models for Analysis of Repeated Measures"*;
4. Pinheiro and Bates (2000), *"Mixed-Effects Models in S and S-PLUS"*.

All of the examples discussed are taken from Pinheiro And Bates (2000). The material on Bayesian Hierarchical models is taken from Davidian and Giltinan (1995).