Drug Disease Model Resources

ddmore

Technical Report

Continuous Observation Models

With NMTRAN/MLXTRAN code

Maciej J Swat[1], Roberto Bizzotto[2], Marc Lavielle[3]

[1]*EMBL - European Bioinformatics Institute, Cambridge, UK*
[2]*National Research Council, Institute of Neuroscience, Padova, Italy*
[3]*Inria Saclay, France*

WORKING DRAFT VERSION

November 4, 2016

Contents

Chapter 1

Introduction

The purpose of this document is to discuss observation models used in nonlinear mixed effects (NLME) models for continuous data and supported in PharmML, [Swat et al., 2015]. It starts with a general classification of observation models and discusses aspects of the variability structure of residual error models. A short overview of the most common models follows which are then described in detail in chapter 2. Chapter 3 discusses then a number of non-standard models, among others from the [Keizer and Karlsson, 2013].

1.1 Classification of observation models

The observation models can be divided into three groups:

- **Structured** models (**S-type**)
  - **General model** – residual error, ϵ_{ij}, with a symmetric distribution with mean 0
    - for **transform-both-sides** data

$$\underbrace{u(y_{ij})}_{\substack{\text{Transformed}\\ \text{experimental}\\ \text{data}}} = \underbrace{u\big(f(x_{ij},\psi_i)\big)}_{\substack{\text{Transformed}\\ \text{model}\\ \text{prediction}}} + \underbrace{g(x_{ij},\psi_i,\xi_i)\,\epsilon_{ij}}_{\substack{\text{Residual}\\ \text{error}}} \tag{1.1}$$

    - for **untransformed** data – special case of the previous with $h \equiv identity$

$$\underbrace{y_{ij}}_{\substack{\text{Experimental}\\ \text{data}}} = \underbrace{f(x_{ij},\psi_i)}_{\substack{\text{Model}\\ \text{prediction}}} + \underbrace{g(x_{ij},\psi_i,\xi_i)\,\epsilon_{ij}}_{\substack{\text{Residual}\\ \text{error}}}$$

  - **Gaussian model** – special case of eq. (1.1) with $\epsilon_{ij} \sim \mathcal{N}(0,1)$ – $u(y_{ij})$ is normally distributed with mean $u(f(x_{ij},\psi_i))$ and the standard deviation $g(x_{ij},\psi_i,\xi_i)$.
- **Distribution** based models (**D-type**)

$$u(y_{ij}) \sim DistributionName(parameter1, parameter2, ...) \tag{1.2}$$

- **Equation** based models (**E-type**)

$$\underbrace{u(y_{ij})}_{\substack{\text{Transformed}\\ \text{experimental}\\ \text{data}}} = \underbrace{U\big(f(x_{ij},\psi_i),\xi_i,\epsilon_{1,ij},\epsilon_{2,ij},\dots\big)}_{\substack{\text{Transformed}\\ \text{model prediction}\\ \text{including residual errors}}} \tag{1.3}$$

Symbol description:

- $1 \le i \le N$
- $1 \le j \le n_i$
- y_{ij} – j^{th} observation for subject i

- f – structural model prediction
- x_{ij} – regression variables, e.g. time or concentration
- ψ_i – individual parameters
- ϵ_{ij} – normalised residual error
- g – standard deviation of the residual error
- ξ_i – parameters of the residual model.

1.2 Variability structure

In most cases the variability structure is reduced to one level, that of the observation, and its magnitude is equal among the entire population. However, in general, the magnitude of the residual error can vary between subjects or occasions. Essay errors are another variability source, the so called inter-replicates variability.

1.2.1 Inter-individual/occasion variability (IIV/IOV)

In analogy to the nested hierarchical structure for the variability of the individual parameters, see [Swat et al., 2015], the variability of residual error parameters inherits the same structure. This way one can capture the inter-individual and/or inter-occasion variability of the residual error parameters. The IIV on the residual error parameters, i.e. the varying magnitude of residual error between subjects, corresponds to the so called 'ETA-on-EPS' approach, see Figure 1.1.

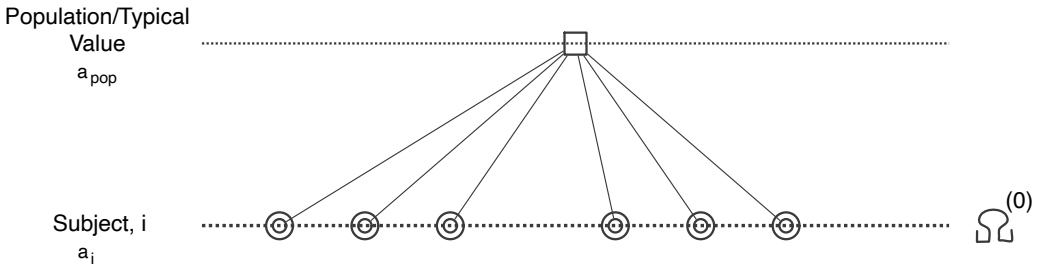

Figure 1.1: Inter-individual variability of the residual error parameter, a, visualised as basic nested hierarchy.

For example, if additive residual error model and log-normal distribution for a is assumed, then the parameter model reads

$$\log(a_i) \sim \mathcal{N}(\log(a_{pop}), \omega_a^2)$$

and the observation model reads

$$y_{ij} \sim \mathcal{N}(f_{ij}, a_i^2): \quad y_{ij} = f_{ij} + a_i\epsilon_{ij}, \quad \epsilon_{ij} \sim \mathcal{N}(0,1)$$

See also chapter 3 for IIV and IOV examples with NMTRAN and MLXTRAN code.

1.2.2 Inter-replicate variability (IRV)

This variability type is analog to the inter-occasion variability for parameters, see for example [Lavielle, 2014], and has been first considered and described in the literature by [Karlsson et al., 1995]. The replicate level in this residual error is described by an additional random effect and reads

$$y_{ijk} = f_{ij} + \epsilon_{ij} + \epsilon_{ijk}, \text{ with } \epsilon_{ij} \sim \mathcal{N}(0, \Sigma^{(0)}), \epsilon_{ijk} \sim \mathcal{N}(0, \Sigma^{(1)})$$

with the k index for the replicates and can therefore be visualised graphically in a very clear nested hierarchical structure, Figure 1.2. Each level can be fully described by a covariance matrix, Σ.

See also chapter 3 for IRV examples with NMTRAN and MLXTRAN code.

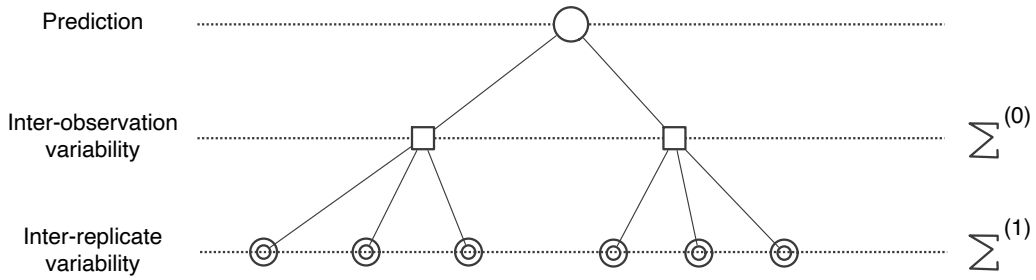

Figure 1.2: Inter-replicate variability in the residual error, visualised as nested hierarchy.

1.3 Selected observation model types

In the following we list a selection of commonly used observation/residual error models, based mainly on [Beal et al., 2006] and [Lavielle, 2014], and described in detail in the next chapters:

1. Constant/additive:

$$y_{ij} = f_{ij} + a\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

$$\text{OR} \quad y_{ij} = f_{ij} + \epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,\sigma^2)$$

2. Constant/additive for log-transformed data, aka *exponential*:

$$\log(y_{ij}) = \log(f_{ij}) + a\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1) \Longleftrightarrow y_{ij} = f_{ij}\exp(a\,\epsilon_{ij})$$

3. Constant/additive for logit-transformed data:

$$\log\Big(\frac{y_{ij}}{1-y_{ij}}\Big) = \log\Big(\frac{f_{ij}}{1-f_{ij}}\Big) + g\epsilon_{ij} \Longleftrightarrow y_{ij} = \frac{f_{ij}\exp(g\,\epsilon_{ij})}{1+f_{ij}\big(\exp(g\,\epsilon_{ij})-1\big)}$$

4. Constant/additive for extended-logit-transformed data:

$$\log\Big(\frac{y_{ij}-A}{B-y_{ij}}\Big) = \log\Big(\frac{f_{ij}-A}{B-f_{ij}}\Big) + g\,\epsilon_{ij} \Longleftrightarrow y_{ij} = A + (B-A)\frac{f_{ij}-A}{f_{ij}-A+(B-f_{ij})\exp(-g\,\epsilon_{ij})}$$

5. Constant/additive for Box-Cox transformed data:

$$\begin{cases} \big(y_{ij}^{\lambda}-1\big)/\lambda = \big(f_{ij}^{\lambda}-1\big)/\lambda + g\,\epsilon_{ij} & \text{for} \quad \lambda \neq 0 \\ \log(y_{ij}) = \log(f_{ij}) + g\,\epsilon_{ij} & \text{for} \quad \lambda = 0 \end{cases} \Longleftrightarrow \begin{cases} y_{ij} = \big(f_{ij}^{\lambda} + \lambda g\,\epsilon_{ij}\big)^{1/\lambda} & \text{for} \quad \lambda \neq 0 \\ y_{ij} = f_{ij}\exp(g\,\epsilon_{ij}) & \text{for} \quad \lambda = 0 \end{cases}$$

6. Proportional or constant coefficient of variation (CCV):

$$y_{ij} = f_{ij} + bf_{ij}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

$$\text{OR} \quad y_{ij} = f_{ij}(1+\epsilon_{ij}); \quad \epsilon_{ij} \sim N(0,\sigma^2)$$

7. Combined proportional 1:

$$y_{ij} = f_{ij} + (a + bf_{ij})\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

8. Combined proportional 2:

$$y_{ij} = f_{ij} + \sqrt{a^2 + b^2 f_{ij}^2}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

$$\text{OR} \quad y_{ij} = f_{ij} + a\,\epsilon_{1,ij} + bf_{ij}\,\epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,1); \quad \epsilon_{2,ij} \sim N(0,1)$$

$$\text{OR} \quad y_{ij} = f_{ij}(1 + \epsilon_{1,ij}) + \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0, \sigma_1^2); \quad \epsilon_{2,ij} \sim N(0, \sigma_2^2)$$

9. Power error model:

$$y_{ij} = f_{ij} + b\, f_{ij}^c\, \epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

$$\text{OR } y_{ij} = f_{ij} + f_{ij}^c\, \epsilon_{ij}; \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

10. Combined power error model 1:

$$y_{ij} = f_{ij} + (a + b f_{ij}^c)\, \epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1)$$

11. Combined power error model 2:

$$y_{ij} = f_{ij} + a\epsilon_{1,ij} + b f_{ij}^c \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,1); \quad \epsilon_{2,ij} \sim N(0,1)$$

$$\text{OR} \quad y_{ij} = f_{ij} + \epsilon_{1,ij} + f_{ij}^c \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0, \sigma_1^2); \quad \epsilon_{2,ij} \sim N(0, \sigma_2^2)$$

Models listed above are the most popular ones in use but the present PharmML structure allows for implementation of virtually any user-defined model. Chapter 2 describes in detail the standard error models listed above. Chapter 3 describes some of the lesser known residual error models described in [Keizer and Karlsson, 2013] and other sources.

Chapter 2

Gaussian observation models

It is by far the most popular type of observation models used. Table 2 lists the few common ones which will be described in details in this chapter.

Short name	Model name	Model formulation	Variance
const	constant (*aka* additive)	$y = f + a\epsilon$ $h(y) = h(f) + a\epsilon$	$var(y) = a^2$ $var(h(y)) = a^2$
prop	proportional	$y = f + bf\epsilon$ $h(y) = h(f) + bh(f)\epsilon$	$var(y) = b^2 f^2$ $var(h(y)) = b^2 h(f)^2$
comb1	combined 1 (*aka* constant + proportional 1)	$y = f + (a + bf)\epsilon$ $h(y) = h(f) + (a + b\, h(f))\epsilon$	$var(y) = (a + bf)^2$ $var(h(y)) = (a + bh(f))^2$
comb2	combined 2 (*aka* constant + proportional 2)	$y = f + a\epsilon_1 + bf\epsilon_2$ $h(y) = h(f) + a\epsilon_1 + bh(f)\epsilon_2$	$var(y) = a^2 + b^2 f^2$ $var(h(y)) = a^2 + b^2 h(f)^2$
power	power	$y = f + bf^c\epsilon$ $h(y) = h(f) + bh(f)^c\epsilon$	$var(y) = b^2 f^{2c}$ $var(h(y)) = b^2 h(f)^{2c}$
comb_power1	combined power 1	$y = f + (a + bf^c)\epsilon$ $h(y) = h(f) + (a + bh(f)^c)\epsilon$	$var(y) = (a + bf^c)^2$ $var(h(y)) = (a + bh(f)^c)^2$
comb_power2	combined power 2	$y = f + a\epsilon_1 + bf^c\epsilon_2$ $h(y) = h(f) + a\epsilon_1 + bh(f)^c\epsilon_2$	$var(y) = a^2 + b^2 f^{2c}$ $var(h(y)) = a^2 + b^2 h(f)^{2c}$

Table 2.1: List of most popular type of Gaussian observation models – both in their un-transformed and transformed forms with according variances.

2.1 Constant model – *aka* additive model

Transformation:

$$u \equiv identity, \quad i.e. \quad u(y_{ij}) = y_{ij}$$

Model definition:

$$y_{ij} = f_{ij} + a\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = a^2$$

Alternative model:

$$y_{ij} = f_{ij} + \epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,\sigma^2); \quad var(y_{ij}) = \sigma^2$$

NMTRAN

```
; default model
$ERROR
ADD = THETA(1)
Y = F + ADD*EPS(1)
$THETA 1; ADD
$SIGMA 1 FIXED

; alternative model
$ERROR
Y = F + EPS(1)
$THETA 1
$SIGMA 1
```

MLXTRAN

```
; default model
[LONGITUDINAL]
input = {..., a}

EQUATION:
F = ...

DEFINITION:
Y = {distribution=normal, prediction=F, sd=a}

; OR using a predefined error model:
Y = {distribution=normal, prediction=F, errorModel=constant(a)}
```

In the next sub-sections special cases of the constant model when applied various transformation types will be discussed.

2.1.1 Constant log-transformed model (aka *exponential* model)

Transformation:

$$u(y_{ij}) = \log(y_{ij})$$

Model definition:

$$\begin{aligned} \log(y_{ij}) &= \log(f_{ij}) + g\epsilon_{ij} \\ \iff y_{ij} &= f_{ij}\exp(g\,\epsilon_{ij}) \end{aligned}$$

with $\epsilon_{ij} \sim N(0,1)$.
Because of the last form this model is often called *exponential*.

NMTRAN

```
$ERROR
G = THETA(1)
Y = LOG(F) + G*EPS(1)
$THETA 1; G
$SIGMA 1 FIX
```

Note In NMTRAN the left hand side of the equation appears as not log-transformed. This is because the software assumes that the data is log-transformed. This assumption holds for other transformations as well as presented in next examples. Monolix doesn't make this assumption, expecting data on the natural scale.

MLXTRAN

```
[LONGITUDINAL]
input = {..., g}

EQUATION:
F = ...

DEFINITION:
Y = {distribution=logNormal, prediction=F, sd=g}
```

2.1.2 Constant logit-transformed

Transformation:

$$u(y_{ij}) = \log\Big(\frac{y_{ij}}{1-y_{ij}}\Big)$$

Model definition:

$$\begin{aligned}\log\Big(\frac{y_{ij}}{1-y_{ij}}\Big) &= \log\Big(\frac{f_{ij}}{1-f_{ij}}\Big) + g\epsilon_{ij} \\ \iff y_{ij} &= \frac{f_{ij}\exp(g\,\epsilon_{ij})}{1+f_{ij}\big(\exp(g\,\epsilon_{ij})-1\big)}\end{aligned}$$

with $\epsilon_{ij} \sim N(0,1)$.

NMTRAN

```
$ERROR
G = THETA(1)
Y = LOG(F/(1−F)) + G∗EPS(1)
$THETA 1; G
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., g}

EQUATION:
F = ...

DEFINITION:
Y = {distribution=logitNormal, prediction=F, sd=g}
```

Note see note on page 7 on differences between NONMEM and Monolix wrt treatment of transformed data

2.1.3 Constant extended-logit-transformed model

Transformation:

$$u(y_{ij}) = \log\Big(\frac{y_{ij}-A}{B-y_{ij}}\Big)$$

Model definition:

$$\begin{aligned}\log\Big(\frac{y_{ij}-A}{B-y_{ij}}\Big) &= \log\Big(\frac{f_{ij}-A}{B-f_{ij}}\Big) + g\,\epsilon_{ij} \\ \iff y_{ij} &= A + (B-A)\frac{f_{ij}-A}{f_{ij}-A+(B-f_{ij})\exp(-g\,\epsilon_{ij})}\end{aligned}$$

with $\epsilon_{ij} \sim N(0,1)$.

NMTRAN

```
$ERROR
A = THETA(1)
B = THETA(2)
G = THETA(3)
Y = LOG((F−A)/(B−F)) + G∗EPS(1)
$THETA 10; A
       20; B
       1; G
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., g}

EQUATION:
F = ...

DEFINITION:
Y = {distribution=logitNormal(a,b), prediction=F, sd=g}
```

Note see note on page 7 on differences between NONMEM and Monolix wrt treatment of transformed data

2.1.4 Constant Box-Cox transformed model

Transformation:

$$u(y_{ij}) = \begin{cases} \left(y_{ij}^{\lambda} - 1\right)/\lambda & \text{for} \quad \lambda \neq 0 \\ \log(y_{ij}) & \text{for} \quad \lambda = 0 \end{cases}$$

Model definition:

$$\begin{cases} \left(y_{ij}^{\lambda} - 1\right)/\lambda = \left(f_{ij}^{\lambda} - 1\right)/\lambda + g\,\epsilon_{ij} & \text{for} \quad \lambda \neq 0 \\ \log(y_{ij}) = \log(f_{ij}) + g\,\epsilon_{ij} & \text{for} \quad \lambda = 0 \end{cases}$$

$$\iff \begin{cases} y_{ij} = \left(f_{ij}^{\lambda} + \lambda g\,\epsilon_{ij}\right)^{1/\lambda} & \text{for} \quad \lambda \neq 0 \\ y_{ij} = f_{ij}\exp(g\,\epsilon_{ij}) & \text{for} \quad \lambda = 0 \end{cases}$$

with $\epsilon_{ij} \sim N(0,1)$.

NMTRAN

```
$ERROR
LAMBDA = THETA(1)
G = THETA(2)
IF(LAMBDA.NE.0)
Y = (F**LAMBDA - 1)/LAMBDA + G*EPS(1)
ELSE
Y = LOG(F) + G*EPS(1)
ENDIF
$THETA 5; LAMBDA
          1; G
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., lambda, g}

EQUATION:
F = ...

DEFINITION:
Y = {distribution=powerNormal(lambda), prediction=F, sd=g}
```

Note see note on page 7 on differences between NONMEM and Monolix wrt treatment of transformed data. CHECK powerNorm

2.2 Proportional or constant coefficient of variation (CCV) model

The coefficient of variation is defined as $CV(y_{ij}) = \frac{\sqrt{var(y_{ij})}}{E(y_{ij})}$.
Model definition:

$$y_{ij} = f_{ij} + b f_{ij}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = b^2 f_{ij}^2; \quad CV = b$$

Alternative model (e.g. part V in [Beal et al., 2006]):

$$y_{ij} = f_{ij}(1 + \epsilon_{ij}); \quad \epsilon_{ij} \sim N(0, \sigma^2); \quad var(y_{ij}) = \sigma^2 f_{ij}^2; \quad CV = \sigma$$

NMTRAN

```
; default model
$ERROR
PROP = THETA(1)
Y = F + PROP*F*EPS(1)
$THETA 1; PROP
$SIGMA 1 FIXED

; alternative model
$ERROR
Y = F*(1 + EPS(1))
$SIGMA 0.1
```

MLXTRAN

```
; default model
[LONGITUDINAL]
input = {..., b}

EQUATION:
F = ...
G = b*F

DEFINITION:
Y = {distribution=normal, prediction=F, sd=G}

; OR using a predefined error model:
Y = {distribution=normal, prediction=F, errorModel=proportional(b)}
```

2.3 Combined constant and proportional error model 1

Model definition:

$$y_{ij} = f_{ij} + (a + bf_{ij})\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0, 1); \quad var(y_{ij}) = (a + bf_{ij})^2$$

NMTRAN

```
$ERROR
ADD = THETA(1)
PROP = THETA(2)
Y = F + (ADD + PROP*F)*EPS(1)
$THETA 1; ADD
         1; PROP
$SIGMA 1 FIXED
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., a, b}

EQUATION:
F = ...
G = a + b*F

DEFINITION:
Y = {distribution=normal, prediction=F, sd=G}

; OR using a predefined error model:
Y = {distribution=normal, prediction=F, errorModel=combined1(a,b)}
```

2.4 Combined constant and proportional error model 2

All three representations are equivalent assuming uncorrelated $\epsilon_{1,ij}$ and $\epsilon_{2,ij}$ in the alternative models 1 & 2. Model definition:

$$y_{ij} = f_{ij} + \sqrt{a^2 + b^2 f_{ij}^2}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = a^2 + b^2 f_{ij}^2$$

Alternative model 1:

$$y_{ij} = f_{ij} + a\,\epsilon_{1,ij} + b f_{ij}\,\epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,1); \quad \epsilon_{2,ij} \sim N(0,1); \quad var(y_{ij}) = a^2 + b^2 f_{ij}^2$$

Alternative model 2:

$$y_{ij} = f_{ij}(1 + \epsilon_{1,ij}) + \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,\sigma_1^2); \quad \epsilon_{2,ij} \sim N(0,\sigma_2^2); \quad var(y_{ij}) = \sigma_1^2 f_{ij}^2 + \sigma_2^2$$

Note NMTRAN, in contrast to MLXTRAN, can handle multiple ϵ's. That's why for the latter only the first model definition can be used (see below).

NMTRAN

```
; default model
$ERROR
ADD = THETA(1)
PROP = THETA(2)
Y = F + SQRT(ADD**2 + PROP**2*F*F)*EPS(1)
$THETA 1; ADD
       1; PROP
$SIGMA 1 FIXED

; alternative model 1
$ERROR
ADD = THETA(1)
PROP = THETA(2)
Y = F + ADD*EPS(1) + PROP*F*EPS(2)
$THETA 1; ADD
       1; PROP
$SIGMA 1 FIXED
$SIGMA 1 FIXED

; alternative model 2
Y = F*(1 + EPS(1)) + EPS(2)
$SIGMA 1
$SIGMA 1
```

MLXTRAN

```
; default model
[LONGITUDINAL]
input = {..., a, b}

EQUATION:
F = ...
G = sqrt(a^2 + b^2*F^2)

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}

; OR using a predefined error model:
Y = {distribution=normal, prediction=F, errorModel=combined2(a,b)}
```

2.5 Power error model

Model definition:

$$y_{ij} = f_{ij} + b\,f_{ij}^c\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = b^2 f_{ij}^{2c}$$

Alternative model:

$$y_{ij} = f_{ij} + f_{ij}^{c}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,\sigma^2); \quad var(y_{ij}) = \sigma^2 f_{ij}^{2c}$$

NMTRAN

```
; default model
$ERROR
PROP = THETA(1)
POWER = THETA(2)
Y = F + (PROP*F**POWER)*EPS(1)
$THETA 1; PROP
       2; POWER
$SIGMA 1 FIXED

; alternative model
$ERROR
POWER = THETA(1)
Y = F + F**POWER*EPS(1)
$THETA 1; POWER
$SIGMA 1
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., b, c}

EQUATION:
F = ...
G = b*F^c

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}

; OR using a predefined error model
Y = {distribution=normal, prediction=F, error=proportionalc(b,c)}
```

2.6 Combined constant and power error model 1

Model definition:

$$y_{ij} = f_{ij} + (a + b f_{ij}^{c})\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = (a + b f_{ij}^{c})^2$$

NMTRAN

```
$ERROR
ADD = THETA(1)
PROP = THETA(2)
POWER = THETA(3)
Y = F + (ADD + PROP*F**POWER)*EPS(1)
$THETA 1; ADD
       1; PROP
       2; POWER
$SIGMA 1 FIXED
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., a, b, c}

EQUATION:
F = ...
G = a + b*F^c

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}

; OR using predefined error models
Y = {distribution=normal, prediction=F, error=combined1c(a,b,c)}
```

2.7 Combined constant and power error model 2

All three representations are equivalent assuming uncorrelated $\epsilon_{1,ij}$ and $\epsilon_{2,ij}$.
Model definition:

$$y_{ij} = f_{ij} + \sqrt{a^2 + b^2 f_{ij}^c}\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = a^2 + b^2 f_{ij}^{2c}$$

Alternative model 1:

$$y_{ij} = f_{ij} + a\epsilon_{1,ij} + b f_{ij}^c \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,1); \quad \epsilon_{2,ij} \sim N(0,1); \quad var(y_{ij}) = a^2 + b^2 f_{ij}^{2c}$$

Alternative model 2:

$$y_{ij} = f_{ij} + \epsilon_{1,ij} + f_{ij}^c \epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0,\sigma_1^2); \quad \epsilon_{2,ij} \sim N(0,\sigma_2^2); \quad var(y_{ij}) = \sigma_1^2 + \sigma_2^2 f_{ij}^{2c}$$

NMTRAN

```
; default model
$ERROR
ADD = THETA(1)
PROP = THETA(2)
POWER = THETA(3)
Y = F + SQRT(ADD**2 + PROP**2*F**POWER)*EPS(1)
$THETA 1; ADD
       1; PROP
       2; POWER
$SIGMA 1 FIXED

; alternative model 1
$ERROR
ADD = THETA(1)
PROP = THETA(2)
POWER = THETA(3)
Y = F + ADD*EPS(1) + PROP*F**POWER*EPS(2)
$THETA 1; ADD
       1; PROP
       2; POWER
$SIGMA 1 FIXED
$SIGMA 1 FIXED

; alternative model 2
$ERROR
POWER = THETA(1)
Y = F + EPS(1) + F**POWER*EPS(2)
$THETA 10
$SIGMA 1
$SIGMA 1
```

MLXTRAN

```
; default model
[LONGITUDINAL]
input = {..., a, b, c}

EQUATION:
F = ...
G = sqrt(a^2 + b^2*F^(2*c))

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}

; OR using predefined error models
Y = {distribution=normal, prediction=F, error=combined2c(a,b,c)}
```

Chapter 3

Other observation models

There is a large family of models which cannot be expressed as *structured* models. Here examples described in detail in this chapter.

- Non-standard combined log-transformed model
- BQL residual model
- Variability in residual error
  - IIV in residual error magnitude
  - IOV in residual error magnitude
  - IRV of the residual error
  - influence of regressors
- Joint residual error between multiple observations
- Two (or more) types of measurements error model

3.1 Non-standard combined log-transformed model

The following variance model was proposed specifically to get the same error structure for log-transformed data as in case of the 'combined constant and proportional model 2' for the untransformed data (see [NMusers] forum discussion).

Model definition:

$$\log(y_{ij}) = \log(f_{ij}) + \sqrt{a^2 + b^2/f_{ij}^2}\,\epsilon_{ij}; \quad \epsilon_{ij} \sim N(0,1); \quad var(y_{ij}) = a^2 + b^2/f_{ij}^2.$$

NONMEM

```
$ERROR
ADD = THETA(1)
PROP = THETA(2)
Y=LOG(F)+SQRT(PROP**2+ADD**2/F**2)*EPS(1)
$THETA 1; ADD
       1; PROP
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., a, b}

EQUATION:
F = ...
G = sqrt(a^2 + b^2/F^2)

DEFINITION:
Y = {distribution=logNormal, prediction=F, sd=G}
```

3.2 BQL residual model

This model was proposed by [Beal, 2001]: *Logarithmically transformed observations are used, and are modeled with the homogeneous variance model, when the distributions of the untransformed observations may be positively skewed, but their cv's seem to be constant. However, there may be observations whose pharmacokinetic predictions become theoretically small, but both their central tendency and variance seem to remain constant and above certain levels (assuming that the assay is accurate, this can only happen when the kinetics are misspecified), in which case another useful model for the logarithmically transformed observations is given by the formula below where M is an extra positively constrained parameter and ϵ_1 and ϵ_2 are random errors.*

Model definition:

$$\log(y_{ij}) = \log(f_{ij}+M) + \frac{f_{ij}}{f_{ij}+M}\epsilon_{1,ij} + \frac{M}{f_{ij}+M}\epsilon_{2,ij}; \quad \epsilon_{1,ij}, \epsilon_{2,ij} \sim N(0,1); \quad var(y_{ij}) = \frac{f_{ij}^2+M^2}{(f_{ij}+M)^2}$$

for uncorrelated $\epsilon_{1,ij}$ and $\epsilon_{2,ij}$.

Alternative model definition:

$$y_{ij} \sim \mathcal{N}\Big(\log(f_{ij}+M), \frac{f_{ij}^2+M^2}{(f_{ij}+M)^2}\Big)$$

NMTRAN (from `www.cognigencorp.com/nonmem/nm/99apr242002.html`):

```
$ERROR
M=THETA(1)
Y=LOG((F)+M)+(F*EPS(1))/(F+M)+(M*EPS(2))/(F+M)
$THETA (0,0.1); M
$SIGMA 1 FIX
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {...,M}

EQUATION:
F = ...
FM = F + M
G = sqrt(F^2 + M^2)/FM

DEFINITION:
Y = {distribution=logNormal, prediction = FM, sd=G}
```

3.3 Variability in residual error magnitude and parameters

3.3.1 IIV of the residual error magnitude – *aka* ETA-on-EPS

Model definition:
Let g_{ij} be the individual log-normally distributed standard deviation of the residual error

$$g_{ij} = g_{typ,ij}\exp(\eta_i); \quad \eta_i \sim \mathcal{N}(0,\omega_g)$$

with $g_{typ,ij}$ as the typical value, see section 1.2.1, e.g.

$$g_{typ,ij} = \sqrt{a^2 + b^2 f_{ij}^2}$$

The Gaussian observation model for untransformed data reads then

$$y_{ij} = f_{ij} + g_{ij}\epsilon_{ij}, \quad \epsilon_{ij} \sim \mathcal{N}(0,1); \quad var(y_{ij}) = g_{ij}^2$$

NMTRAN

```
$ERROR
ADD=THETA(1)
PROP=THETA(2)
IPRED = F
G = SQRT(ADD**2+PROP**2*IPRED**2) * EXP(ETA(1))
Y = IPRED + G*EPS(1)
$THETA 1; ADD
       1; PROP
$OMEGA 1
$SIGMA 1 FIX
```

MLXTRAN

```
[INDIVIDUAL]
input={..., omega_G}

DEFINITION:
eta_G = {distribution=normal, mean=0, sd=omega_G}

[LONGITUDINAL]
input={..., a, b, eta_G}

EQUATION:
F = ...
G = sqrt(a^2 + b^2*F^2)*eta_G

DEFINITION:
Y = {distribution=normal, prediction=F, sd=G}
```

3.3.2 IIV of the residual error parameters

Alternative IIV .... The models above illustrates modelling IIV of the residual error magnitude via the standard deviation function, G. An alternative is to assign IIV to the parameters of the residual error as shown in this example.

Model definition:

$$
\begin{aligned}
a_i &= a_{pop}\, \exp(\eta_{a,i}); \quad \eta_{a,i} \sim \mathcal{N}(0, \omega_a^2) \\
b_i &= b_{pop}\, \exp(\eta_{b,i}); \quad \eta_{b,i} \sim \mathcal{N}(0, \omega_b^2) \\
g_{ij} &= \sqrt{a_i^2 + b_i^2 f_{ij}^2} \\
y_{ij} &= f_{ij} + g_{ij}\epsilon_{ij}; \quad \epsilon_{ij} \sim \mathcal{N}(0,1); \quad var(y_{ij}) = g_{ij}^2
\end{aligned}
$$

NMTRAN

```
$ERROR
ADD_pop = THETA(1)
PROP_pop = THETA(2)
ADD = ADD_pop*ETA(1)
PROP = PROP_pop*ETA(2)
G = SQRT(ADD**2 + PROP**2*F**2)
Y = F + G*EPS(1)
$THETA 1; ADD_pop
       1; PROP_pop
$OMEGA 1
$OMEGA 1
$SIGMA 1 FIXED
```

MLXTRAN

```
[INDIVIDUAL]
input = {..., a_pop, b_pop, omega_a, omega_b}

DEFINITION:
a = {distribution=logNormal, prediction=a_pop,sd=omega_a}
b = {distribution=logNormal, prediction=b_pop,sd=omega_b}

[LONGITUDINAL]
input = {..., a, b}

EQUATION:
F = ...
G = sqrt(a^2+b^2*F^2)

DEFINITION
Y = {distribution=normal, mean=F, sd=G}
```

3.3.3 IOV of the residual error magnitude

Here the magnitude of the residual error varies with occasions, expressed with subscript k.

Model definition:

$$g_{typ,ij} = \sqrt{a^2 + b^2 f_{ij}^2}$$
$$g_{ijk} = g_{typ,ij} \exp(\eta_{ik})$$

with $\eta_{ik} \sim \mathcal{N}(0, \gamma^2)$.

The Gaussian observation model for untransformed data reads then

$$\begin{aligned} & y_{ijk} \sim \mathcal{N}(f_{ij}, g_{ijk}^2) \\ or \quad & y_{ijk} = f_{ij} + g_{ijk}\epsilon_{ij}; \quad \epsilon_{ij} \sim \mathcal{N}(0,1); \quad var(y_{ijk}) = g_{ijk}^2 \end{aligned}$$

NMTRAN (with two occasions indicated with values 1 or 2 in the OCC column)

```
$ERROR
ADD = THETA(1)
PROP = THETA(2)
IF(OCC.EQ.1) IOV = ETA(1)
IF(OCC.EQ.2) IOV = ETA(2)
G = SQRT(ADD**2+PROP**2*IPRED**2) * EXP(IOV)
Y = IPRED + G*EPS(1)

$THETA 1; ADD
       1; PROP
$OMEGA BLOCK (1) 0.5
$OMEGA BLOCK (1) SAME
$SIGMA 1 FIX
```

MLXTRAN
(with two occasions indicated with values 1 or 2 in the OCC column of type REG. Note that this cannot change the connectivity of records in the dataset in terms of time values.)

```
[INDIVIDUAL]
input={..., omega}

DEFINITION:
eta_G1 = {distribution=normal, mean=0,sd=omega}
eta_G2 = {distribution=normal, mean=0,sd=omega}

[LONGITUDINAL]
input={..., eta_G1, eta_G2, a, b}
regressor={OCC}

EQUATION:
F = ...
if OCC==1
   IOV=eta_G1
else
   IOV=eta_G2
end
G = sqrt(a^2+b^2*F^2) * exp(IOV)

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}
```

MLXTRAN

```
[INDIVIDUAL]
input={..., omega}

DEFINITION:
eta_G1 = {distribution=normal, mean=0,sd=omega}
eta_G2 = {distribution=normal, mean=0,sd=omega}

[LONGITUDINAL]
input={..., eta_G1, eta_G2, a, b}
regressor={OCC}

EQUATION:
F = ...
G_typ = sqrt(a^2 + b^2*F^2)
if OCC==1
   G=G_typ*exp(eta_G1)
else
   G=G_typ*exp(eta_G2)
end

DEFINITION:
Y = {distribution=normal, mean=F, sd=G}
```

Note IOV of the residual error magnitude can also be encoded via variability on its parameters, a and b, similarly to the model described in section 3.3.2.

3.3.4 IRV of the residual error magnitude

The background information was given in section 1.2. Here we cite from the first paper dealing with this model in the PK context.

[Karlsson et al., 1995] on inter-replicate variability: *By taking replicate measurements at a single time point it is possible to some extent to differentiate between sources of error. Variability between replicates arises from assay error and the error introduced by sample handling from the point in the handling chain where the replication is made. We will assume replication is made at the point of sampling. The difference between predicted and observed concentrations at the j^{th} time point now has two components, a consistent difference, ϵ_{ij}, between all replicates and the prediction, and replicate-specific differences, ϵ_{ijk}.*

Model definition:

$$y_{ijk} = f_{ij} + \epsilon_{ij} + \epsilon_{ijk}; \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma_1^2),\ \epsilon_{ijk} \sim \mathcal{N}(0, \sigma_2^2); \quad var(y_{ijk}) = \sigma_1^2 + \sigma_2^2$$

NMTRAN (with two replicates)

```
$ERROR
IF (REP.EQ.1) IRV = EPS(2)
IF (REP.EQ.2) IRV = EPS(3)
Y = IPRED + EPS(1) + IRV
$SIGMA 0.1 ; "Shared" error
$SIGMA BLOCK (1) 0.05
$SIGMA BLOCK (1) SAME
```

MLXTRAN (with two replicates, indicated by different values in a column of type YTYPE)

```
[LONGITUDINAL]
input = {..., a1, a2}

EQUATION:
F = ...
G = sqrt(a1^2 + a2^2)

DEFINITION:
Y1 = {distribution=normal, prediction=F, sd=G}
Y2 = {distribution=normal, prediction=F, sd=G}
```

Note IRV of the residual error magnitude can also be encoded via variability on its parameters, a1 and a2, similar to model described in section 3.3.2.

3.3.5 Residual error magnitude varying with time-varying covariates/regressors

Model definition example:

$$\begin{aligned} MAT &= 1/KA; \quad \text{mean absorption time} \\ FACT &= \begin{cases} 1 & \text{if} \quad time > MAT \\ c & \text{otherwise} \end{cases} \\ g &= \sqrt{a^2 + b^2 f^2} \times FACT \end{aligned}$$

NMTRAN

```
$ERROR
PROP = THETA(1)
ADD = THETA(2)
FACT = c
MAT = 1 / KA
IF(TIME.GT.MAT) FACT = 1
W = SQRT(ADD**2+PROP**2*IPRED*IPRED)*FACT
Y = IPRED + W*EPS(1)
$THETA 1; PROP
       1; ADD
$SIGMA 1 FIX
```

MLXTRAN

```
[LONGITUDINAL]
input = {..., KA, a, b, c}

EQUATION:
MAT = 1/KA
if t > MAT
       FACT = 1
else
       FACT = c
end
F = ...
G = sqrt(a^2 + b^2*F^2)*FACT

DEFINITION:
Y = {distribution=normal, prediction=F, sd=G}
```

3.4 Joint residual error between multiple observations

This models applies to joint parent/metabolite measurements, discussed here in the specific case of constant error model for both observations

Model definition:

$$y_{ij} = \begin{cases} f_{1,ij} + \epsilon_{1,ij} & \text{if} \quad \text{TYPE} = 1 \\ f_{2,ij} + \epsilon_{2,ij} & \text{if} \quad \text{TYPE} = 0 \end{cases} \qquad \text{with} \quad \begin{bmatrix} \epsilon_{1,ij} \\ \epsilon_{2,ij} \end{bmatrix} \in \mathcal{N}\left(0, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$$

where TYPE is the observation type.

NMTRAN

```
$ERROR
IF (TYPE.EQ.1) THEN
Y = IPRED1 + EPS(1) ; Observation type 1
ELSE
Y = IPRED2 + EPS(2) ; Observation type 2
ENDIF
$SIGMA BLOCK(2)
0.1
0.05 0.1
```

MLXTRAN (with observation types indicated by different values in a column of type YTYPE)

```
[LONGITUDINAL]
input = {..., a1, a2, r}

EQUATION:
F1 = ...
F2 = ...
G1 = a1
G2 = a2

DEFINITION:
Y1 = {distribution=normal, prediction=F1, sd=G1}
Y2 = {distribution=normal, prediction=F2, sd=G2}
correlation = {block={Y1,Y2} , corrcoef=r}
```

3.5 Two (or more) types of measurements error model

This model assumes that e.g. the drug concentration, is measured using different assays. This is expressed in the following example by the categorical variable ASY taking values 1 or 0, for the first or second assay respectively ([Beal et al., 2006]).

Model definition:

$$y_{ij} = f_{ij} + \text{ASY}_j \epsilon_{1,ij} + (1 - \text{ASY}_j)\epsilon_{2,ij}; \quad \epsilon_{1,ij} \sim N(0, \sigma_1^2); \quad \epsilon_{2,ij} \sim N(0, \sigma_2^2);$$

Alternative formulation:

$$y_{ij} = \left\{ \begin{array}{ll} f_{ij} + \epsilon_{1,ij} & \text{if} \quad \text{ASY}_j = 1 \\ f_{ij} + \epsilon_{2,ij} & \text{if} \quad \text{ASY}_j = 0 \end{array} \right. ; \quad \epsilon_{1,ij} \sim N(0, \sigma_1^2); \quad \epsilon_{2,ij} \sim N(0, \sigma_2^2);$$

In both formulations:

$$var(y_{ij}) = \left\{ \begin{array}{ll} \sigma_1^2 & \text{if} \quad \text{ASY}_j = 1 \\ \sigma_2^2 & \text{if} \quad \text{ASY}_j = 0 \end{array} \right.$$

NMTRAN

```
$ERROR
IPRED=F
IF (ASY.EQ.1) Y = IPRED + EPS(1)
IF (ASY.EQ.0) Y = IPRED + EPS(2)
$SIGMA 1
$SIGMA 1
```

MLXTRAN (with YTYPE for the column ASY)

```
[LONGITUDINAL]
input = {a1, a2, ...}

EQUATION:
F = ...
G1 = a1
G2 = a2

DEFINITION:
Y1 = {distribution=normal, mean=F, sd=G1}
Y2 = {distribution=normal, mean=F, sd=G2}
```

Bibliography

[Beal, 2001] Beal, S. L. (2001). Ways to fit a pk model with some data below the quantification limit. *Journal of Pharmacokinetics and Pharmacodynamics*, 28(5):481–504.

[Beal et al., 2006] Beal, S. L., B, S. L., and (Eds.), B. A. J. (2006). *NONMEM Users Guides, (1989-2006)*. Icon Development Solutions, Ellicott City, Maryland, USA.

[Karlsson et al., 1995] Karlsson, M. O., Beal, S. L., and Sheiner, L. B. (1995). Three new residual error models for population pk/pd analyses. *J Pharmacokinet Biopharm*, 23(6):651–72.

[Keizer and Karlsson, 2013] Keizer, R. and Karlsson, M. (2013). Stochastic models. Technical report, Uppsala Pharmacometrics Research Group, Uppsala.

[Lavielle, 2014] Lavielle, M. (2014). *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. Chapman & Hall/CRC Biostatistics Series.

[Swat et al., 2015] Swat, M. J., Wimalaratne, S. M., Kristensen, N. R., Yvon, F., Moodie, S., and Le Novère, N. (2015). Pharmacometrics Markup Language (PharmML), Language Specification for Version 0.6.