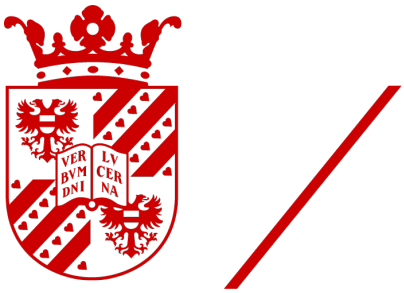
Improving on the approaches in Pharmacokinetical Modelling

Investigating approaches in pharmacokinetic modelling

Cornelis Scipio Leeuwenhart Rook

December 2022



rijksuniversiteit groningen

Rijksuniversiteit Groningen
Supervisors: W.P. Krijnen and Marco Grzegorczyk

Abstract

When it comes to creating a model in pharmacokinetics, the use of parametric, mixed effect models is often employed. This is a reliable approach when prior research has been performed to shine a light on the behavior of a drug. However, this approach is not necessarily optimal because of the assumptions it makes and requires about the behavior of a drug. Especially when the exact behavior of a drug is relatively unknown, the use of semi-parametric or non-parametric models might prove to be a useful tool to still build models that can be used to predict the behavior of a drug, albeit at the cost of not providing an explicit model formula. This thesis will showcase the use, pros and cons of the non-linear mixed effect models and generalized additive mixed effect models by performing model building and analysis on three separate data sets containing measurements in participants in pharmacokinetic studies. The numerical stability of the methods is also considered for each model. The drugs in question in the data sets are remifentanil and theophylline. The third data set contains measurements of blood flow in participants in a study on high-flux hemodialyzers. The goal of this thesis is to widen the scope of the models used in pharmacokinetic research, to explore the limits of such models and when to use them. The models were built and analysed in the statistical software tool R-studio.

Contents

1 Introduction

In the world of pharmacokinetic modeling, there are several problems and challenges that can arise. Some of these problems are practical, such as not being able to draw a measurement from a patient at every planned time interval. But other problems are more specifically related to the research of pharmacokinetcs. For instance, when investigating a new type of drug, it can be hard to apply a suitable type of model if you do not know the exact physiological behavior or reaction of the body to the drug. It can be the case that a one-compartment model is suitable, or a two-compartment model. What (non-)linear function would be appropriate to fit to the data? These are some of the troubles that one can walk into when trying to fit a model to pharmacokinetic data. For this reason, this thesis will investigate and compare two major model-fitting methods that can be used for different reasons and advantages to analyze pharmacokinetic data: generalized additive models and (non-)linear mixed effect models.

When trying to fit a model to data, one can choose out of a couple of regression models. On the one hand, there are non-parametric models, and on the other hand, there are parametric models. The first describes methods that do not make any prior assumptions about the model or parameters. Hence, the relationship between variables is assessed through non-parametric methods. However, this does not mean that findings such as residuals cannot be treated with parametric assumptions. The main advantage of models that are generated in this way, is that they have the advantage of being very flexible, as they are not subjected to specific parametrizations that may or may not closely approximate the data. However, the drawback of this flexibility is that it to some extent diminshes the degree to which one can interpret the meaning of the model. Since the result of a non-parametric model did not make use of any prior assumptions, the model also leaves the user with relatively little physical implications that can be derived from the model, meaning that the relationship between the dependent and the independent variables is not expressed as, for instance, some explicit formula. At this point it becomes clear that interpretation is the main drawback of this method and this is why we also consider a parametric method.

Contrary to a non-parametric method, a parametric method does allow for assumptions about the distribution and possible parametrizations of the data. Since a parametrization already expresses relationships between the dependent and independent variables, the user is able to gauge how well the respective parametrization can approximate the data. This means that there is already an interpretation of the model available by definition. The obvious downside of this approach is that a suitable parametrization might not be available and is hard to find. For example, this can be the case when the data shows irregular patterns that are hard to describe with a single formula. What's more, one does not always know if the used parametrization is the best one. However, another advantage that a parametric approach has compared to a non-parametric approach, is that it often requires less computational effort to generate the model. Hence, these parametric models hold appealing qualities, but are also limited to the quality and appropriateness of the parametrization used.

With regards to pharmacokinetic modelling, the above presents multiple reasons as to why to use certain approaches in modelling. When there is uncertainty about what explicit relationship may or may not exist between the dependent and the independent variables, a non-parametric method may be better suited for the job. However, even if a suitable parametric representation is available, would this assumed relationship outperform a non-parametric approach? Although it might seem reasonable to assume that a parametric approach would generally be better if within reach, it might not be the case by definition. For that reason, we would require methods that allow us to compare these different methods.

Generally speaking, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are widely used to compare (nested) models that are of similar structure and use similar likelihood estimation methods such as maximum likelihood estimation (MLE) and restricted maximum likelihood estimation (REML). Information criterion methods such as AIC and BIC reward goodness of fit and penalize the amount of parameters used in various ways to comprise a score. This will be elaborated on more in the method section. However, there exists some discussion between scientists that believe that such criterions can be used liberally to compare different

models and those that believe these criterions should be limited to nested models (Spiegelhalter et al., 2002). The root cause of this discussion comes from the fact that the comparison between nested models is reliable because the same terms that are neglected in the approximation of the error term $y - f(x)$ —Where f is some function of x and an approximation of y— will cancel (Wood, 2017). For that reason, the comparison between different types of models is not ideal, but generally accepted within the field of modelling.

In short, the goal of this paper is to unveil the qualities and purpose of both parametric and non-parametric models. Most notably, investigating when the two approaches differ in result and quality and why will be one of the main focuses. What conditions or circumstances will cause one type of candidate pharmacokinetic model to perform better than some other type? As the models will be discussed and compared, arguments will also be made in favor of either or both methods based on the performance and properties displayed. Ultimately, we hope that the information shown will bring attention and insight to others such that they might widen their horizon when it comes to modelling pharmacokinetic data.

2 Methods

In order to uncover potential discrepancies in the performance of the parametric and non-parametric approaches, three different data sets have been used to apply the various models to. These methods were applied in the statistical software program R. The data sets used are: Theophylline, from the 'datasets' package and Dialyzer and Remifentanil from the 'nlme' package. The data set on theophylline is a data frame consisting of 132 rows containing information on 12 subjects that participated in the study by Dr. Robert Upton of the kinetics of the anti-asthmatic drug. The data itself is taken from the report by Boeckmann et al. (1994). The columns contain information on the respective dose administered, weight, elapsed time in hours and current concentration of the drug in the blood stream of the participant after oral administration of the drug.

The Dialyzer data set was taken from research performed by Vonesh and Carter (1992) investigating the characteristics of in vivo ultrafiltration for hemodialyzers. Meaning that it measured the behavior of a dialyzer that filters the blood that are placed in the body. The data set holds 140 rows with measures about 20 patients. Next to a subject identifier, the data frame holds information about the bovine blood flow rate which comes as a factor with two levels, the transmembrane pressure and the ultrafiltration rate of the dialyzer.

Lastly, the Remifential data set is a larger data set containing 2107 rows and 12 columns, taken from a research on the influence of age and gender of the pharmacokinetics of remifentanil by Minto et al. (1997) The data frame holds information on 65 participants and how the remifentanil concentrations in their blood changed over time. Several other measurements about the patient are also included in the columns, such as age, sex, height, weight, body surface area and lean body mass, in addition to the infusion rate and amount of remifentanil administered in the current time interval.

The reason that we opted for these sets of pharmacokinetic data is that each data set has something that makes them fundamentally different from the others. Theophylline holds a fixed number of only seven data points per subject, whereas Remifentanil holds a much larger, varying amount of data points per subject. This difference is of interest in this study in order to investigate the performance of both models when fed both few and many data points. Secondly, Dialyzer is of a different nature in that it also holds a grouping factor and an interesting potential random effect structure, as will be shown later.

Now that the data sets have been discussed, it is time to introduce the methods used to model the data. All methods and analysis has been carried out within the statistical modelling software 'R-studio' using packages available on the repository (CRAN). As mentioned in the introduction, we will use two different approaches to modelling pharmacokinetic data. For the non-parametric method, we used GAM, short for generalized additive models, from the mgcv package and we used (n)lme (non-linear mixed effects models) from the nlme package for

the parametric method.

2.1 Generalized additive models

Introducing the generalized additive model as formulated in Wood (2017), the general idea of the method is to model the dependent variable by generating multiple curves that are nonzero for some (partially overlapping) interval along the axis of independent variables and add these curves together. The sum of these curves equals the function that approximates the data. In the case of multiple covariates that need to be incorporated into the model, this amounts to a general formal formulation of the form:

$$y_i = a + f_1(x_i) + f_2(v_i) + ... + f_k(z_i) + \epsilon_i$$

where y_i is the response variable, a is some intercept parameter, f, also called a smooth, is a function of some covariate that is generated from a sum of multiple underlying functions such as we just described. The ϵ_i term is a noise parameter that follows a $N(0, \sigma^2)$ distribution. Each f_j can be represented in a form equivalent to:

$$f_1(x) = \sum_{j=1}^{k_1} b_j(x)\delta_j$$

where δ_j are unknown coefficients and the b_j are piecewise linear functions based on some spacing of know along the axis of the input variable. Now, in order to find the functions that best approximation to the "true" function that is generating the data, the most conventional method is to minimize the difference

$$\|y - X\beta\|^2$$

However, in order to not just basically perform linear interpolation, a penalty term is introduced. This penalty term will prevent the overfitting to the noise that is typical of interpolation methods by penalizing the wiggliness of the curve. Consequently, penalizing the wiggliness of the curve more and more by increasing λ means that the curve will become more and more like a straight line. This term is implemented as follows

$$\|y - X\beta\|^2 + \lambda S$$

Where S is some expression that encapsulates the wiggliness of the curve in some form. Most intuitively, this could be performed using the second derivatives of the functions described in $X\beta$, as is the case in cubic smoothing splines. This would look like

$$\sum_{i=1}^{n}\{y_i - g(x_i)^2\} + \lambda \int g''(x)^2 dx$$

It should be noted however, that this is not the default definition used in the GAM function, as the default type of smoothing splines are thin plate smoothing splines. The main difference of this method is that it provides a general solution to fitting a smooth function of more than one independent variables. This is done by providing a different penalty expression defined as

$$S_{md} = \int_{R^d} \sum_{v_1+...+v_d} \frac{m!}{v_1!...v_d!}(\frac{\partial^m f}{\partial x_1^{v_1}...\partial x_d^{v_d}})^2 dx_1...dx_d$$

Note that x is a d-dimensional vector. Here, m is to be chosen, causing all of the combinations of the v-terms to be determined as well. Another advantage is that thin plate regression splines do not require any particular ordering of nodes.

A challenge that this method introduces, is its flexibility. Namely, within the mgcv package that provides the usage of this these types of models, a wide variety of options available to give certain characteristics to certain smooths in order to make an appropriate model. Apart from the fact that there are multiple types of smooths, specifying and adjusting a smooth means evaluating the dimension of a smooth, the order of the penalty term, point constraints that the smooth should satisfy, weights of certain measures and overall desired or tolerated smoothness of the curve. It is not always required to specify non-default values for most of these settings, but there certainly are cases in which the offered flexibility allows one to significantly improve the fit of a smooth and we will also draw from these resources in our models.

2.2 Non-linear mixed effect models

Secondly, we have the non-linear mixed effect models. These are often thought of as hierarchical models and therefore its initial formulation is slightly different from that of the generalized additive model. We speak of a mixed model because of the inclusion of both fixed and random effects. The first can be though of as population-level effects. The latter are the effects associated with individual experimental units drawn at random from a population (Pinheiro & Bates, 2000) . In this case an observation of data point within a group is modeled by

$$y_{ij} = f(\phi_{ij}, v_{ij}) + \epsilon_{ij}, \qquad I = 1, \ldots, M \qquad j = 1, \ldots, n_i$$

Where j is the number of the observation and i is the group number, M is the number of groups and n_i is the number of observations present in the respective group. Here, ϕ_{ij} is a parameter vector that is specific to each group i and v_{ij} is a covariate vector. Similar to earlier, the ϵ_{ij} term is an error term, but this time it is normally distributed differently per group. (Pinheiro & Bates, 2000) Because the model is non-linear, f is non-linear with respect to some element of the parameter vector ϕ_{ij}. Subsequently, this parameter is specified in such a way that it holds information about both fixed, population-level effects and random, subject-level effects. That is,

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i \qquad b_{ij} \sim \mathcal{N}(0, \Psi)$$

where β and b_i are the vectors representing the fixed and random effects by group respectively. The matrices A and B then naturally specify the (non-linear) ways the fixed and random effects are to be put together in order to represent the desired parameter vector ϕ_{ij} for the model.

2.3 Mixed effect models and additive models

From the general definitions of generalized additive models and non-linear mixed effects models, it is clear how fixed and random effects are incorporated in a nlme model. However, it is not immediately obvious how fixed and random effects can be represented in the context of a GAM. The analogy between the mixed effect formulation and the smooths in an additive model comes from the penalty matrix S as discussed in Silverman (1985) This matrix penalizes the wiggliness of the smooth along with its scalar, λ. Now, in the context of fixed and random effect, a fixed effect is reflected by a smooth that is not penalized. This is represented in the addition of a separate column vector containing the parameters that are fixed effects. Consequently, the random effects are the smooths that did undergo penalization and hence adopt a certain wiggliness based on the level of penalization. When we fit the smooths for each subject, or subgroups, we introduce the hierarchical structure similar to that of a nlme model. Moreover, the addition of each penalized smooth on top of the main, fixed effect smooth, represents the estimated correction on the fixed effect fit, using only the participant's personal measures and parameters.

More formally, when we suspect that random effects are present, we might formulate this belief in a way that these random effects behave in a random way, such a as a random variable following a certain probability distribution.

Let us introduce some smooth $y_i = f(x_i) + \epsilon_i$, such that it be can defined as $y \sim \mathcal{N}(\mathbf{X}\beta, \mathbf{I}\sigma^2)$ for some parameter vector β and where

$$\beta \sim \mathcal{N}(\mathbf{0}, \mathbf{S}^-/\lambda)$$

Here, $\mathbf{S}^-$ is a psuedoinverse of the penalty matrix S (Wood, 2017). Then, by taking the log and some other minor operations, we find that the estimator of this parameter vector can be defined as

$$\hat{\beta} = \arg\min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|/\sigma^2 + \lambda\beta^T\mathbf{S}\beta$$

Assuming positive definiteness of $\mathbf{S}$, $\hat{\beta}$ can be obtained by computing the derivative of the previous expression with respect to β, setting it to zero and solving for β. The steps are shown below. The σ^2 term has been removed through multiplying by it and adopting it in the λ.

$$\frac{\partial}{\partial\beta}\{\|\mathbf{y} - \mathbf{X}\beta\| + \lambda\beta^T\mathbf{S}\beta\}$$

$$= \frac{\partial}{\partial\beta}\{(\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta) + \lambda\beta^T\mathbf{S}\beta\}$$

$$= \frac{\partial}{\partial\beta}\{(\mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{X}^T\mathbf{y} + \beta^T(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{S})\beta\}$$

Setting the entire expression equal to zero and performing the differentiation yields:

$$= -2\mathbf{X}^T\mathbf{y} + 2(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{S})\beta$$

Note that the differentiation of the quadratic term on the right simply results in $2(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{S})\beta$ because $\mathbf{X}^T\mathbf{X}$ is symmetric, but S is too from the assumption of positive definiteness. Lastly, the latter result simplifies to

$$\beta = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{S})^{-1}\mathbf{X}^T\mathbf{y}$$

If we now substitute $\hat{\beta}$ for β in our expression for the distribution of y, we see that we arrive at a model structure that is very similar to that of the random effects ϕ_{ij} specified in the previous section. In this way, we are also able to generate random effects for our non-parametric additive models.

In practice, a when we want to produce a GAM in R, we specify the dependent variable and its relation to the independent variable with a model formula, like in most conventional regression methods. Additionally, the `gam()` function allows us to also use the `s()` function that introduces the ability to specify a smooth based on the input variables and parameters. A typical call to construct a GAM will look like the following

```
gam(y ~ s(x), data = D)
```

for dependent variable `y`, independent variable `x` and some data frame `D` containing these variables as columns. Now, in order to produce a smooth that is treated as a random effect, we can perform the similar call

```
gam(y ~ s(x, bs = "re"), data = D)
```

These random effects can be interpreted as random slopes that are added to the model. If a situation requires that a random effect is not merely a random slope, but rather a random smooth, we can use a different type of smooth. These are called factor smooths as they apply a smooth for each level of the specified random effect factor

and can be specified as follows

```
gam(y ~ s(fac, x, bs = "fs"), data = D) ,
```

where `fac` stands for the factor and x for the variable that we want a smooth of for each factor level.

However, we should note that the input for such a random effect is limited to being an independent variable or some interaction effect between independent variables. For this reason, we will also make use of the `gamm()` function, which actually uses the nlme package to fit random effects. For simple random effects, `gam()` is the preferred function as it has faster performance and is more numerically robust. On the other hand, the implementation of the nlme package for fitting random effects also allows the user to exploit the same random effect inputs as in a nlme function call, allowing for much more intricate random effect structures.

2.4 Assessing and comparing models

Before we start to compare methods, we need to realise that the two approaches we are considering do not use the same assumptions and methods. Therefore an optimal formulation of one model will not by definition produce an optimal model for the other, if we try to mimic the structure of the first model as much as possible in the other model. Secondly, we need to establish how we will arrive at an optimal model. It should be noted that the interpretation of optimal in our case mostly comes down to having adjusted the model such that it accounts for the structure of the data (e.g. its distribution and heteroscedasticity) and that the coefficients and parameters that reside in the model are significant and make for a parsimonious model. Generally speaking, the approach to realise the aforementioned definition of an optimal model amounts to starting with a nlme model with fixed and random effects present for all of the coefficients and then performing analysis and adjustments to improve the model. In order to compare models, a conventional ANOVA will be performed in order to assess the quality and improvement between the previous model and the new model.

In the ANOVA, F-tests are performed to compute the ratio of the between- and within-group variability of two or more models. Subsequently, the p-value for this ratio is computed and reflects how significant this ratio is. Additionally, the `anova` command in R additionally presents AIC and BIC scores that indicate how well the models fit the data relative to the amount parameters used and their log-likelihoods, as discussed earlier. Note that an F-test only works for nested models and hence the `avnova` command is mostly useful for comparison within a model type, rather than between model types. Therefore, when comparing non-nested or different types of models, no p-value is presented, but one can still draw inferences from the AIC and BIC values. The use of ANOVA tables within the process of analysing and adjusting the models will be demonstrated in the results section.

In addition to assessing the quality of a model on its own, we also consider the properties of the convergence and fit of the model. This will be done by inspecting the Hessian of the log-likelihood of the parameters and the quantile plots of the model fits respectively. The first of these represents the matrix of second derivatives of the log-likelihood of the combination of parameter, given the parameter vector used in the model. To better understand this, we can formalise this into

$$\mathbf{H}_{ij} = \frac{\partial^2}{\partial\theta_i \partial\theta_j} \log p(Y|\theta)\bigg|_{\theta=\theta^*}$$

meaning that we assess the second derivative of the log-likelihood of some probability density function p, for some random variable Y at some value of parameter vector θ^*. The entries of the matrix h_{ij} are then defined as the double partial derivative of the log-likelihood function with respect to the i-th and then j-th element of the parameter vector. In practice, this θ^* is the instance of the parameter vector at which the parameters provide the optimal fit to the given model structure.

Now, in order to use interpret the values of this matrix, it is useful to make an analogy to calculus, where the second derivative describes the acceleration of the curve. This property is then often used to derive properties at local optima and minima of the curve, which are obtained when the first derivative is zero. Namely, when the curve arrives at such a point, the first derivative, which equals zero, does not provide any additional information as to if the point is a local minimum or maximum. In our case, we know beforehand if we converge at some minimum or maximum based on if we choose to minimize or maximize the positive or negative log-likelihood respectively. However, the second derivative does show us the rate at which the model converges at around the optimal value, which is given by the optimal parameter vector θ^*.

The way this applies to our models is that we can look at the definiteness of this Hessian. If the Hessian is positive or negative definite for minimizing or maximizing the log-likelihood respectively at the optimal vector parameter solution, we have a converging model. In practice this condition is met by definition for any model that did not throw any convergence errors. However, we can still say something about the quality of this convergence if based on the properties of the Hessian. The main property of the Hessian is that it describes the rate at which the objective function, which is the log-likelihood function in our case, changes based on a difference in the input parameter vector θ^*. This amount of change is typically described by the condition number, κ, which is the ratio between the largest and smallest eigenvalues of $\mathbf{H}$ in the case of a symmetric matrix. Note that this is the case because here $\mathbf{H}$ is defined as the inverse of the variance-covariance matrix, which is symmetric by construction. Then, if κ is large it means that small differences in the θ^* will produce relative large differences in the output of the log-likelihood, meaning that the model is probably not very reliable. The opposite also holds: a κ closer to 1 (as the condition number is a ratio) implies that the log-likelihood behave more consistently. Note that generally speaking, a relatively small condition number does not imply convergence, but rather shows that the output will not vary as wildly compared to an output that has a greater underlying condition number when θ^* changes. Therefore, models with lower condition numbers will be preferable as they are more computationally stable.

Most importantly, it can tell us something about how distinct the given solution is and thus how certain we are of this model being appropriate for our observed values. By distinct we mean how spread out the locations of the minima and optima are if we were to fit our model to new or replicated data. The closer these minima or optima are to each other, the more distinct the value of the θ^* and the less scattered the values of the vector are between each fit to similar data. Hence, in order to assess the quality of convergence of our models.

Secondly, we also look at the quantile plots produced by the models. In these plots, dots are plotted representing the standardized residuals of the fit. These dots are plotted against a straight line with a slope of one. If these dots line up near this line, it indicates that the residuals are distributed approximately normal. Therefore, the quantile plots will inform us on how well the model manages to be in or around the center of how the response variables are plotted.

3 Results

3.1 Remifentanil

3.1.1 Nlme

The first models we will consider are the models we built for the Remifentanil data set. This data set was of great interest because of the large amount of data it provided because of the amount of observations per participant, but also the amount participants. Let us begin with the general model with fixed and random effects for all parameters by calling the following command in R.

```
RemiMod.nlme0 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject, data = RemiT)
```

where 'conc' (short for concentration) is the response variable, SSfol is a so called 'self-starting function' given by

$$y(x) = \frac{Dose * exp(lKe) * exp(lKa)}{exp(lCl)[exp(lKa) - exp(lKe)]}\{exp[-exp(lKe)x] - exp[-exp(lKa)x]\}$$

that feeds initial values of the parameters to the non-linear mixed effect model lKe, lKa and lCl. The latter parameters are shorthand for the log value of the extermination rate, absorption rate and clearance rate of the substance administered and x is stands for time. However, when running this line, R throws an error and warnings about functions not achieving convergence. Because increasing the allowed number of iterations did not solve this problem, we decided to look at the correlation structure of the variables by inspecting the `pairs()` of the parameters for a simpler model generated using the `nlsList()` function. This revealed the following plot

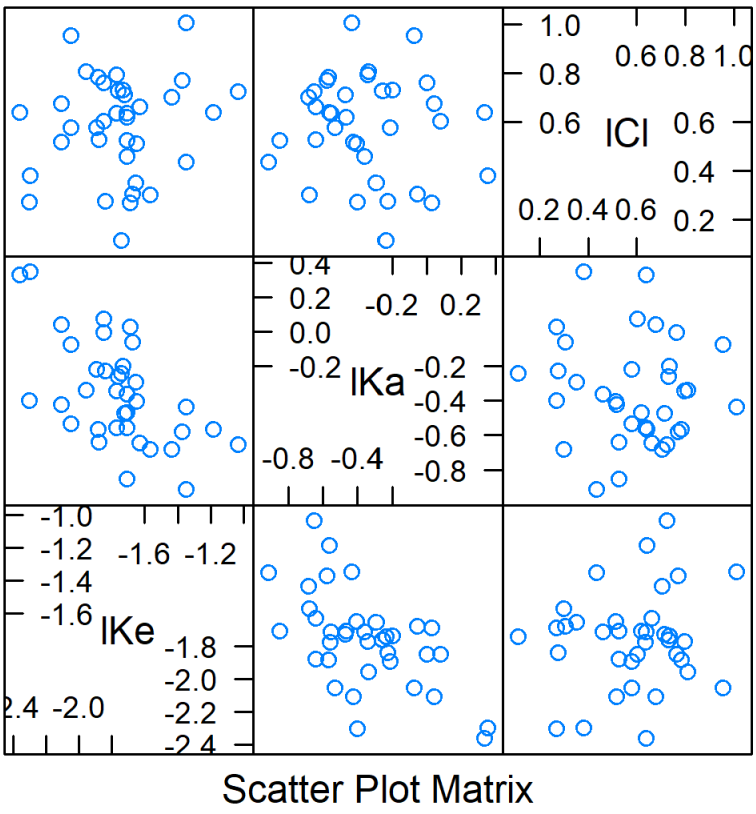


Figure 1: Scatter plot of the extermination rate, absorption rate and clearance.

In this plot it can be seen that the variables lKe and lKa are possibly negatively correlated as there is some diagonal structure to those particular plots. This means that our model is probably over-parameterized and that as a consequence, some of our parameters have become too dependent of each other. Therefore, we reformulate our model to

```
RemiMod.nlme01 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject,
random = pdDiag(list(lKe + lKa + lCl ~ 1)), data = RemiT) .
```

The additional specification of our random variables in this way forces the random-effects variance-covariance structure to be diagonal in order to remove the correlation between our lKe and lKa variables. This resulted in the function call converging and producing our first nlme model. The summary output can be seen below. Throughout

	RemiMod.nlme0
lKe_f	-1.85^{***}
	(0.04)
lKa_f	-0.44^{***}
	(0.07)
lCl_f	0.54^{***}
	(0.03)
lKe_σ	0.20^{***}
	(0.00)
lKa_σ	0.16^{***}
	(0.00)
lCl_σ	0.19^{***}
	(0.00)
N	1133.00
No. groups	36.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe	lKa
lKa	-0.398	
lCl	0.110	0.229

Table 1: Summary table and correlation matrix of RemiMod.nlme0

this thesis, values of fixed and standard deviations of random effects for coefficients are indicated by subscripts of f and σ respectively.

This output shows that the model at hand is already performs in a sound way as we can see from a couple of indicators. First of all, we see that the p-values for all of the fixed effects are significant. Secondly none of values of the fixed and random effects have values that are small to a degree that they are redundant compared to the general scale of the other values. Lastly, we also see that the correlation between our parameters is never close to -1 or 1, meaning that all of the parameters in the model are not (virtually) linearly dependent numerically speaking.

It would be acceptable to stop at this point, as the model does not show any glaring issues or obvious improvements. However, in order to see if we can do better we resort to literature on modelling the pharmacokinetics of Remifentanil. For instance Kim et al. (2017) showed that the weight of the participant was related to the clearance observed. We included this relationship in the fixed effect structure to see if this would lead to a better model as displayed below. We follow this up with an anova test to see if the new model is significantly better than our initial model.

```
RemiMod.nlme01 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject,
fixed = list(lKe + lKa ~ 1, lCl ~ LBM), random = pdDiag(list(lKe + lKa + lCl ~ 1)),
data = RemiT) .
```

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
RemiMod.nlme0	7	7866.368	7901.597	-3926.184		
RemiMod.nlme01	8	7854.403	7894.664	-3919.201	13.966	$1*10^{-4}$

Table 2: Anova table of models RemiMod.nlme0 and RemiMod.nlme01

From the anova test, we can see that the new model has a siginficantly better AIC and BIC with a p-value of around 0.0002. However, in the summary table of RemiMod.nlme01 in table 3 we can see that the addition of the interaction fixed effect between clearance and lean body mass (LBM) caused the intercept of the clearance to be no longer significant in the model. Therefore, we remove the intercept and perform another anova test to see if the

model has improved again under this adjustment.

	RemiMod.nlme01
lKe_f	-1.85^{***}
	(0.04)
lKa_f	-0.44^{***}
	(0.07)
lCl.(Intercept)_f	-0.07
	(0.15)
lCl.LBM_f	0.01^{***}
	(0.00)
lKe_σ	0.20^{***}
	(0.00)
lKa_σ	0.17^{***}
	(0.00)
$\text{lCl.(Intercept)}_\sigma$	0.15^{***}
	(0.00)
N	1133.00
No. groups	36.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe	lKa	lCl.(I)
lKa	-0.395		
lCl.(I)	0.035	-0.065	
LCL.LBM	-0.010	0.013	-0.982

Table 3: Summary table and correlation matrix of RemiMod.nlme01

```
RemiMod.nlme02 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject,
fixed = list(lKe + lKa ~ 1, lCl ~ 0 + lCl), random = pdDiag(list(lKe + lKa + lCl ~ 1)),
data = RemiT) .
```

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
RemiMod.nlme01	8	7854.403	7894.664	-3919.201		
RemiMod.nlme02	7	7852.839	7887.839	-3919.305	6.935	0.0085

Table 4: Anova table of models RemiMod.nlme01 and RemiMod.nlme02

Again, the new model supposedly has a significantly better fit in terms of AIC-score. As hoped, we have created a significantly better model that contains only significant coefficients, as displayed in table 4.

Additionally, as stated by Kim et al. (2017), it has been known for some time now that age is also a significant covariate in the pharmacokinetics of Remifentanil. Apparently, the age of a participant is most often found to be related to the elimination rate and/or the clearance. In our sample, the addition of age in the random effect structure did not lead to significant improvement of the model and this was also the case when adding it to the fixed effect component of the clearance. However, we did find strong support for participant age being an important covariate for the fixed effects corresponding to the elimination rate parameter when exploring how to include age into our model optimally by the following command.

	RemiMod.nlme02
lKe_f	-1.85^{***}
	(0.04)
lKa_f	-0.44^{***}
	(0.07)
lCl.LBM_f	0.01^{***}
	(0.00)
lKe_σ	0.20^{***}
	(0.00)
lKa_σ	0.17^{***}
	(0.00)
$\text{lCl.(Intercept)}_\sigma$	0.16^{***}
	(0.00)
N	1133.00
No. groups	36.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe	lKa
lKa	-0.395	
LCL.LBM	0.129	-0.266

Table 5: Summary table and correlation matrix of RemiMod.nlme02

```
RemiMod.nlme03 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject,
fixed = list(lKe ~ Age, lKa ~ 1, lCl ~ 0 + lCl), random = pdDiag(list(lKe + lKa + lCl
~ 1)), data = RemiT)
```

We also see from the summary in table 6 that again we have constructed a model that holds only significant coefficients.

	RemiMod.nlme03
lKe.(Intercept)_f	-2.07^{***}
	(0.09)
lKe.Age_f	0.01^{**}
	(0.00)
lKa_f	-0.45^{***}
	(0.07)
lCl.LBM_f	0.01^{***}
	(0.00)
$\text{lKe.(Intercept)}_\sigma$	0.18^{***}
	(0.00)
lKa_σ	0.20^{***}
	(0.00)
$\text{lCl.(Intercept)}_\sigma$	0.15^{***}
	(0.00)
N	1133.00
No. groups	36.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe.(I)	lKe.Ag	lKa
lKe.(I)	-0.902		
lKa	-0.060	-0.131	
LCL.LBM	0.012	0.054	-0.262

Table 6: Summary table and correlation matrix of RemiMod.nlme03

What's more, when comparing the models we see that we have again produced a significantly better model.

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
RemiMod.nlme02	7	7852.839	7887.839	-3919.305		
RemiMod.nlme03	8	7848.644	7887.905	-3916.822	6.966	0.083

Table 7: Anova table of models RemiMod.nlme02 and RemiMod.nlme03

Now that we have incorporated several influences from the literature and did not find any other worthwhile adjustments to the model, we revisit our initial issue with the random effects variance-covariance structure that lead us to using the `pdDiag` function. As we saw earlier, the absorption and elimination rate seemed to be strongly correlated for their random effects. From this standpoint, it might be worthwhile to remove one or the other to see if this significantly influences the model in some way. From the results, we saw that the removal of the elimination rate did the model no good, increasing the AIC up to 7898 with a p-value of near zero. However, the exclusion of the absorption random effect parameter did give some interesting results. For our new model, specified below, we also got the following ANOVA results

```
RemiMod.nlme04 <- nlme(conc ~ SSfol(Dose, Time, lKe, lKa, lCl) | Subject,
fixed = list(lKe ~ Age, lKa ~ 1, lCl ~ 0 + lCl), random = pdDiag(list(lKe + lCl ~ 1)),
data = RemiT)
```

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
RemiMod.nlme03	8	7848.644	7887.905	-3916.822		
RemiMod.nlme04	7	67847.980	7883.209	-3916.990	2.336	0.126

Table 8: Anova table of models RemiMod.nlme03 and RemiMod.nlme04

As we had suspected, the removal of the absorption parameter did not seem to significantly change the quality of the model by neither the AIC or BIC metrics. Hence, we cannot say that one model is strictly better than the other. However, because of the principle of parsimony, we favor the latter model, as it has a simpler random effect structure.

However, as we inspect the plot for the fitted values verses the response values in figure 2, it becomes clear that the residuals are seem heteroscedastic. For that reason, the model was updated again in order to account for the varying residuals. This was done by modelling the variance in the model as a power function plus some constant, according to the formulation of the `varConstPower` function from the nlme package as shown below.

```
RemiMod.nlme05 <- update(RemiMod.nlme0, fixed = list(lKe ~ 1, lKa ~ 1, lCl ~ 1), random
= pdDiag(list(lKe ~ 1, lCl ~ 1 )), groups = ~ Subject, weights = varConstPower(fixed
= list(power = .71), form = ~ fitted(.)))
```

Note that the fixed effect structure was also reduced to the basic structure as it appeared that the last addition to the model revealed that our prior structure did not yield a model that was significantly better than one without this complex fixed effect structure. Consequently, the results shown in the anova table reveal that the introduction of this variance structure improves the model significantly.

Lastly, the plots shown in figure 3 show the fit of the RemiMod.nlme05 model. As one can see, the shape of the fitted values roughly corresponds to the way the values observed values are spread out. However, the fitted values more often than not tend to be slightly shifted or delayed compared the observed values.

	RemiMod.nlme05
lKe_f	-2.03^{***}
	(0.02)
lKa_f	-0.38^{***}
	(0.06)
lCl_f	0.56^{***}
	(0.03)
lKe_σ	0.12^{***}
	(0.00)
lCl_σ	0.18^{***}
	(0.00)
N	1133.00
No. groups	36.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe	lKa
lKe	-0.344	
lKa	0.013	-0.146

Table 9: Summary table and correlation matrix of Remi.nlme05

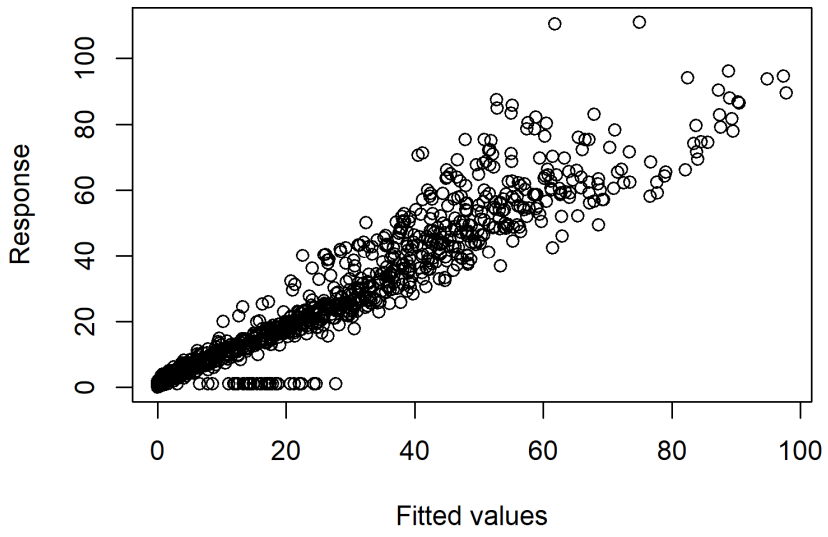


Figure 2: Plot of the fitted values versus the response values for RemiMod.nlme04. Note that the spread of the nodes seems to widen as the value of the fitted value increases

Model	df	AIC	BIC	LogLik.
RemiMod.nlme04	7	7847.980	7883.209	-3916.990
RemiMod.nlme05	7	6385.288	6420.516	-3185.644

Table 10: Anova table of models RemiMod.nlme04 and RemiMod.nlme05

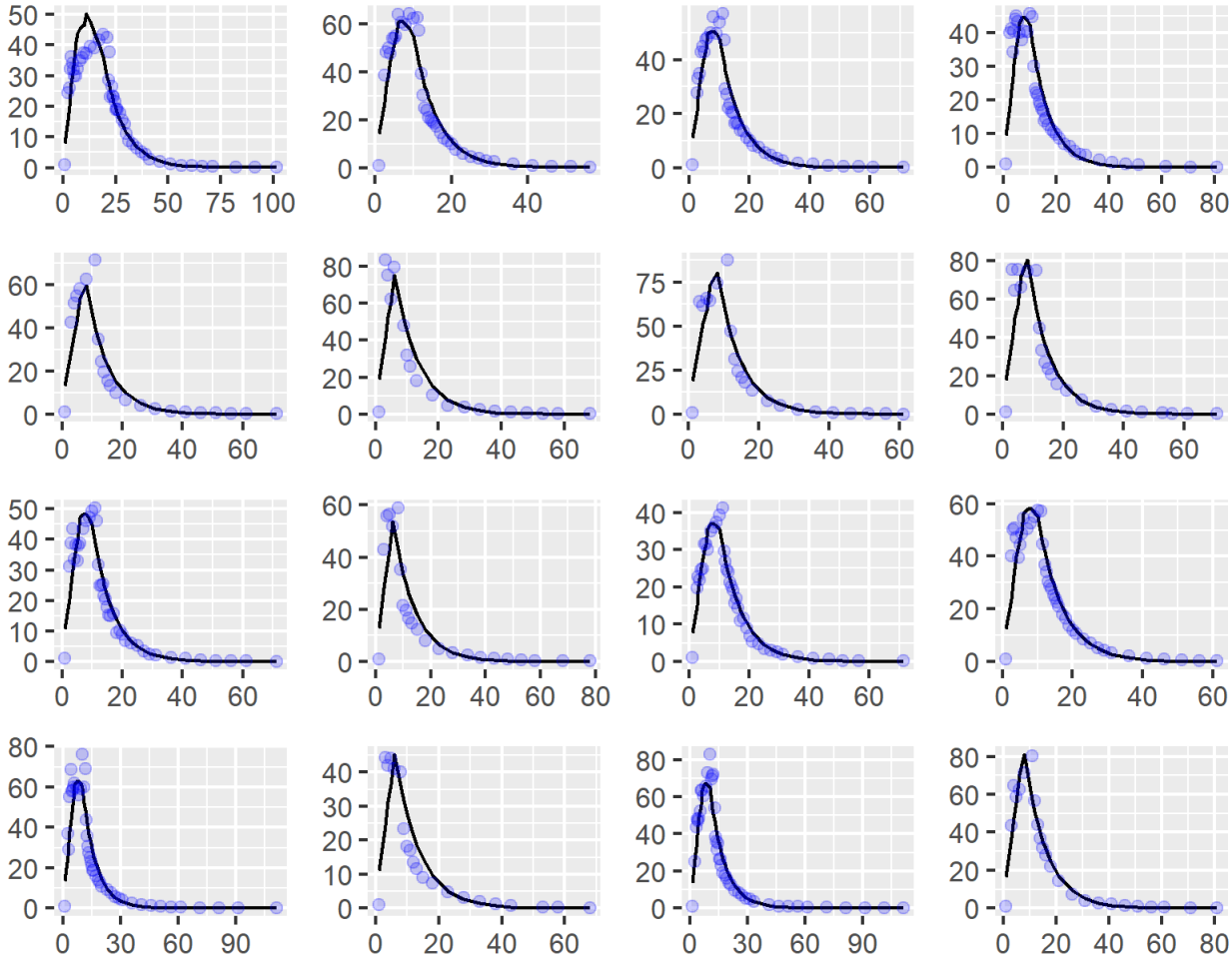


Figure 3: Plot of the RemiMod.nlme05 model for 16 random participants from the remifentanil study

3.1.2 GAM

Now, for the formulation of a proper GAM model we started with evaluating the various smooths that we could implement in the model and how define these. In the light of the theory on the analogy between smooths and random effects, it makes sense to think a model constructed out of a smooth that represents the general population effect of the drug and one or more smooths that represent the present per subject effects with respect to the fixed effect. Because the concentration of remifentanil was measured over time, it makes sense to specify the fixed effect smooth as a function of time. Subsequently, we need to assess what type of random effects we need to consider. Based on the knowledge we have from the previous models, a good starting point would be to incorporate random effects of both lean body mass (LBM), age and the dose administered. Although the dose was not explicitly used in our function call for our nlme models, it was present in the parametrization introduced by the self starting function we used from the nlme package. These covariates will serve as a starting point for formulating our model. For our initial model, we start by specifying our fixed effect model, which will be generated from:

```
RemiMod.gam1 <- gam(conc ~ s(Time, k = 15, pc = 1), data = RemiT, method = "REML", family
= Gamma(link = "log"), weights = c(4, rep(1,nrow(RemiT) -1)), select = TRUE)
```

In the specification of the smooth, `k` represents the dimension of the basis to use for constructing the smooth. The term `pc` is a so called point constraint that force the smooth to pass trough a certain point. This constraint

was added since the measures of the concentration would often jump from 0 to the a much higher value that was measured at the first time interval, causing the starting point to behave like an outlier in some cases where the difference between the first and second measurement we relatively big. Therefore, introducing the point constraint caused the the fit to pass through the first measurement in most cases (only most because a smooth is not equal to the fit). In addition we also increased the weight of the first measurement to put even more emphasis on the fit going through the first observed value. We also allowed the model to fully penalize the smoothing terms if necessary by stating `select = TRUE`. In other words, we allow the model to remove smooth terms from the model if they are too wiggly. We added this setting because the fits tended to become wiggly for some participants where there were large clusters of observations. Lastly, the gamma distribution was used to model the values of the fit. The reason for this was that we required the fit to be nonzero because a negative concentration would not make sense.

Now that we have introduced the fixed effect smooth, we can start adding random effect smooths, or factor smooths, rather. The approach will be similar in terms of how we assess the model. However, the main difference is that we have to find the best model from a pool of potential models and we have to find a way to find the best model. As mentioned earlier, we do have some prior knowledge of what covariates probably are present in the random effect structure. Additionally, it is intuitive to add a factor smooth dependent on time as a random effect, so that we have both a fixed and random effect smooth of time. Secondly, because it is hard to decide in which direction to take the model from there, we opt to compare this model to a relatively full model in which we add a multitude of fixed effects and random effects such as random intercepts, slopes, smooths. Then by comparing the residual variances and AIC scores, we can see if our basic fixed and random effect model needs some of the components of the fuller model. If this is the case, we can trim down the fuller model by checking for non-significant smooths and removing these. Hence, for our basic model, we define

```
RemiMod.gam2 <- gam(conc ~ s(Time, k = 15, pc = 1) + s(Subject, Time, k = 10, bs = "fs"),
data = RemiT, method = "REML", family = Gamma(link = "log"), weights = c(4, rep(1,nrow(RemiT)
-1)), select = TRUE)
```

This will be the basic model that we are going to compare to the fuller model. The reduction of the basis dimension of the factor smooth to 10 is because it appeared that larger values would cause the model to have more coefficients than data for future models.

From here, we will introduce our fuller model. We should mention that we have checked for the presence of some fixed and random effects, or even linear relationships beforehand. Namely, there appeared to be no significant interaction effects that were not between the subject and some other variable. We also left out the random intercept smooth of the dose along with the random slopes for the lean body mass and dose, because they greatly reduced the quality of the model compared to all other smooths. Only the variables of lean body mass en dose would produce significant linear effects for the model, but these would yield either virtually equal or worse AIC scores. Taking all of the above into considerations, our fuller model used will be given by:

```
RemiMod.gam.F <- gam(conc ~ s(Time, k = 15, pc = 1) + s(Age) + s(LBM)
+ s(Subject, bs = "re", k = 10) + s(Age, bs = "re", k = 10) + s(LBM, bs = "re", k=10)
+ s(Subject, Age, bs = "re", k = 10) + s(Subject, Time, k = 10, bs = "fs"), data = RemiT,
method = "REML", family = Gamma(link = "log"), weights = c(4, rep(1,nrow(RemiT) -1)))
```

We see here that we have included fixed effect smooths for age and lean body mass and that we have included random intercepts for each subject, age, and lean body mass. Additionally we have also added the random slope

	edf	F	p-value
s(Time)	$1.314 * 10^1$	186.204	$< 2 * 10^-16$
s(Age)	4.366	15.939	$< 2 * 10^-16$
s(LBM)	2.855	25.303	$< 2 * 10^-16$
s(Subject)	2.035	0.308	0.0150
s(Age)	1.39510^{-15}	0	0.9660
s(LBM)	$-6.068 * 1-^{-16}$	0	0.0127
s(Subject, Age)	$2.206 * 10^{-3}$	0	0.0919
s(Subject, Time)	$3.773 * 10^1$	2.445	$< 2 * 10^-16$

Table 12: Summary table of the smooths of the RemiMod.gam.F model

for age. When we look at the AIC for both models, we see that the AIC value for the full model is about 17 points lower than the AIC value for the basic model. Meaning that something is going right in the full model.

Model	df	AIC
RemiMod.gam2	58.76	3086.483
RemiMod.gam.F	53.84	3069.704

Table 11: Comparison of the degrees of freedom and the AIC values between the models RemiMod.gam2 and RemiMod.gam.F

When we inspect the summary of the this model in table 12, we see that the random effect smooths are not significant.

Removing the random effect smooths resulted in a model with only significant coefficients. However, upon further testing, it turned out that removing some of the significant smooths did not affect the AIC score, nor the residual deviance much. This way, upon continuing to iteratively include and exclude components from the full model, we found a new model, RemiMod.gam3, that has very similar properties to the full model, but uses fewer terms. This model was given by the following line.

```
RemiMod.gam3 <- gam(conc ~ s(Time, k = 15, pc = 1) + s(Subject, bs = "fs", k = 10) +
s(Age) + s(LBM), data = RemiT, family = Gamma(link = "log"), weights = c(4, rep(1,nrow(RemiT)
-1)))
```

Although a traditional anova F test cannot provide an accurate quantification of the p-value of the ratio of the F-values, we can draw information from the residual deviance results. Apparently, the residual deviance equals 83.079 for RemiMod.gam3 and 83.000 for RemiMod.gam.F. This shows that the RemiMod.gam3 model performs virtually the same with respect to this metric compared to the full model, even though it uses fewer terms. On the other hand, the AIC score for RemiMod.gam3 appears to be slightly greater than the full model. What's more, the AIC-scores are still worse than that of the original model with the regular fixed and random effect structure, represented by RemiMod.gam2.

Model	df	AIC
RemiMod.gam2	68.61	3038.232
RemiMod.gam3	69.02	3042.280
RemiMod.gam.F	53.84	3039.49

Table 13: Comparison of the degrees of freedom and the AIC values between the GAM models

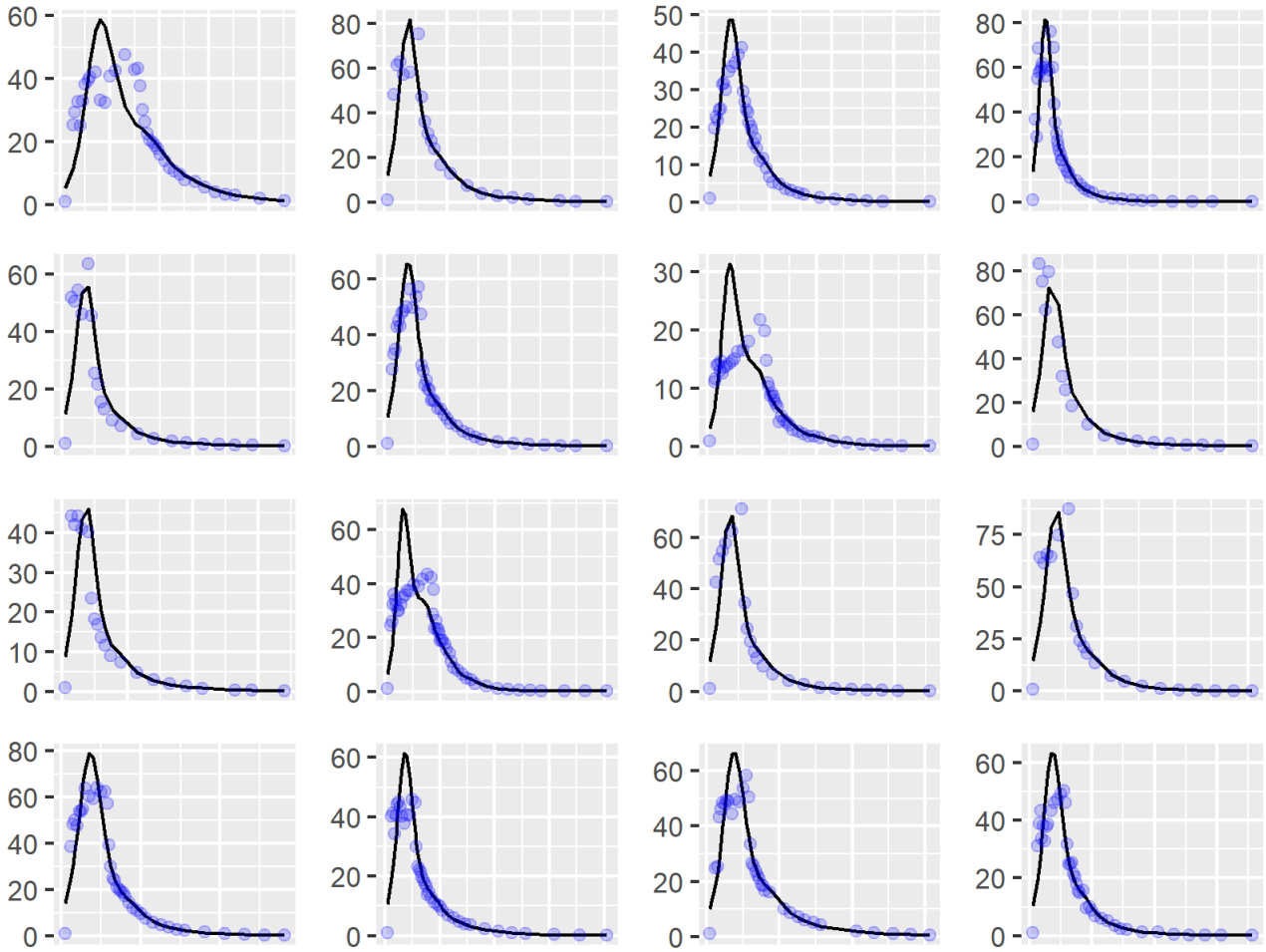


Figure 4: Plots of the fitted values of the RemiMod.gam2 model for 16 participants

Now that we have obtained a satisfactory model, we look at the plots to see how well it performs from a visual perspective. In figure 4, the fitted values look decent for some of the curves, but seem to overshoot the observed values more often than not. The most obvious explanation for this result is that the random effects do not adjust the fixed effect strongly enough for every subject. For that reason, we also attempted to formulate an additional model that does not assume a mixed model structure in order to see if this would produce a better result. In this model, a smooth is fitted directly per subject with entirely separate smooth coefficients per subject, rather than first generating a general population effect as a starting point for the model, as we can see in the model formula below.

```
RemiMod.gam0 <- gam(conc ~ s(Time, by = Subject, k = 20, pc = 1), gamma = .3, data =
RemiT, method = "REML", family = Gamma(link = "log"), weights = c(4, rep(1,nrow(RemiT)-1)),
select = TRUE)
```

Because we move away from the mixed model formulation here, we also leave behind some of the benefits this model. The most important aspect is that because we no longer consider population level effects, we do also not consider the inclusion of other explanatory variables to generate our fit. We are able to drop these explanatory variables in this case because we only fit smooth of time directly and independently for each subject. Therefore, the fixed effect plus some other explanatory variable-based random effects is no longer of any additional value and we are left with a model that provides less explanation or interpretation, at the trade-off of more precise fits. In

the plot below it is clearly visible that this method produces more suitable models.

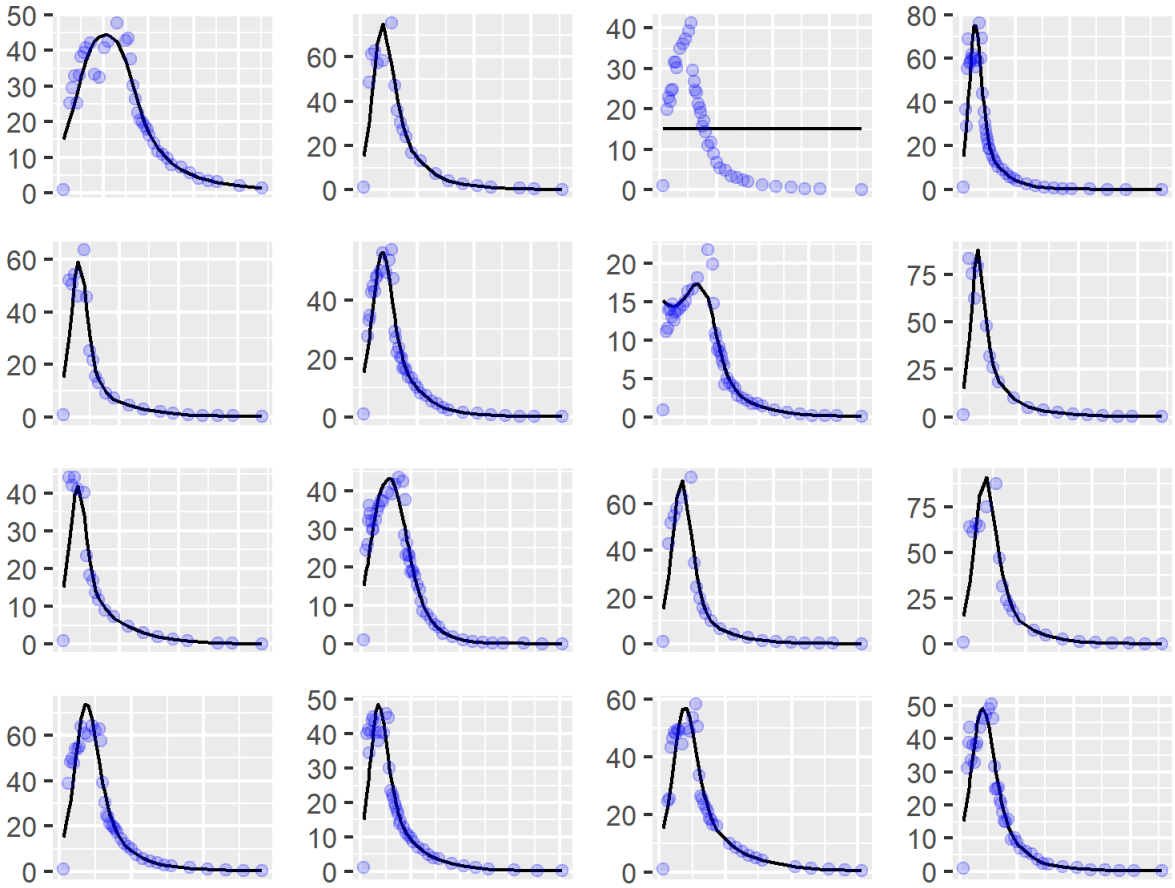


Figure 5: Plots of the per subject GAM of the same random 16 participants of the remifentanil study

Aside from the third graph showing a straight line for the fit (for some unknown reason), the other graphs display fairly nice fits. The seventh graph appears to suffer from a cluster of values that causes the fit to be dominated by this cluster. The result is that the model skips the first value as if it is some outlier and does not match the peak displayed in the measurements. Other than this, the fits behave well generally speaking. However, although this model provides good fits, the quality of these fits was paid for in a less explicit result that specifies nothing about the relationship between the dependent variable and any independent variable other than time.

3.1.3 Analysis of remifentanil models

Now that have formulated models for both the nlme and GAM methods, we will investigate some of they key properties that tell us something about the quality of a fit in a broader sense. First off, we will compare the quantile plots produced from the best model we were able to generate from both methods. From the quantile plot of model RemiMod.nlme05 below, we see that the extreme values of the residuals occur more often than expected compared to a Gaussian distribution. This means that largest residuals, both positive and negative occur too often. In combination with the plot on the right in figure 6, it becomes evident that there are multiple outliers in the observed or response values that are corroborated with smaller fitted values. On the other hand, we can also see that there is a cluster of fitted values that are structurally larger than the observed values in the zero to forty range, which partially explains the tail in the left of the quantile plot in figure 6.

The first phenomenon of response values being greater than some of the fitted values can also be observed to

some degree in figure 3, where in most plots the peaks of the fitted line do not quite attain most of the observed values nor the shape or curvature of this peak.

Still, despite the use of the constant plus power function for modelling the variance in the RemiMod.nlme05 model, the heteroscedasticity seems present to some degree, albeit is less than before. This observed trend could also be partially due to the parametrization not being fit to model such sharp peaks or various kinds of curvatures of these peaks in the data, which would cause this plot to have this shape for most power functions that would model the variance.

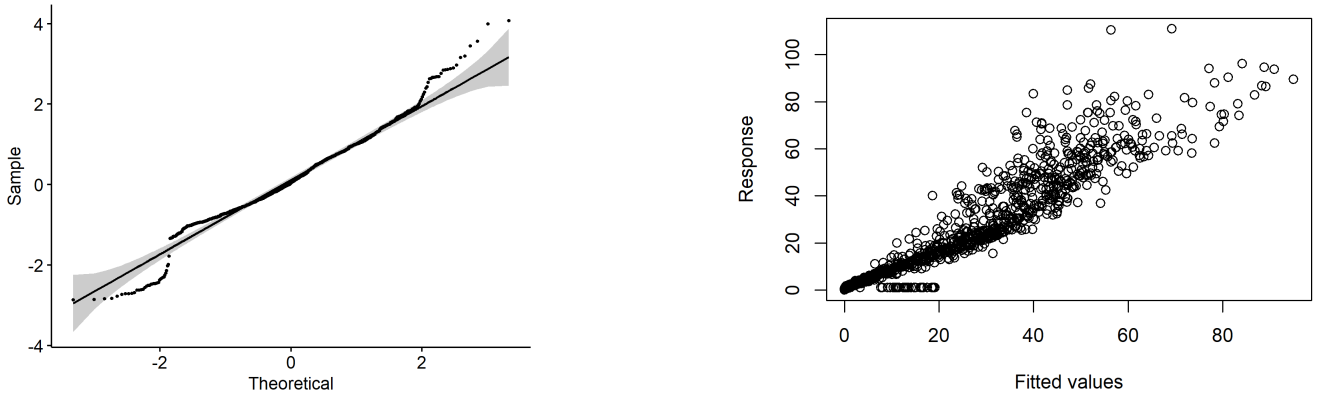


Figure 6: Left: quantile plot of the RemiMod.nlme05 model. Right: scatter plot of the fitted versus the residual values

Secondly, we have the quantile plots of the additive models. The first plot in figure 7, obtained from the `gam.check` function applied to the RemiMod.gam2 model, again shows that the deviance residuals are probably distributed in a non Gaussian fashion, as is supported by a Shapiro-test returning a p-value smaller than $2.2 * 10^{-16}$, meaning that the p-value is essentially zero. We can also see this from the plot in the bottom let corner, where the residuals are too heavily stacked around zero for the distribution to be normal. Moreover, the plot on the lower right also shows that the model struggles more with fitting a model to the observed values the greater the values get. This is also immediately clear from figure 9, where we see multiple peaks in the plots not quite matching the trend of the observed values.

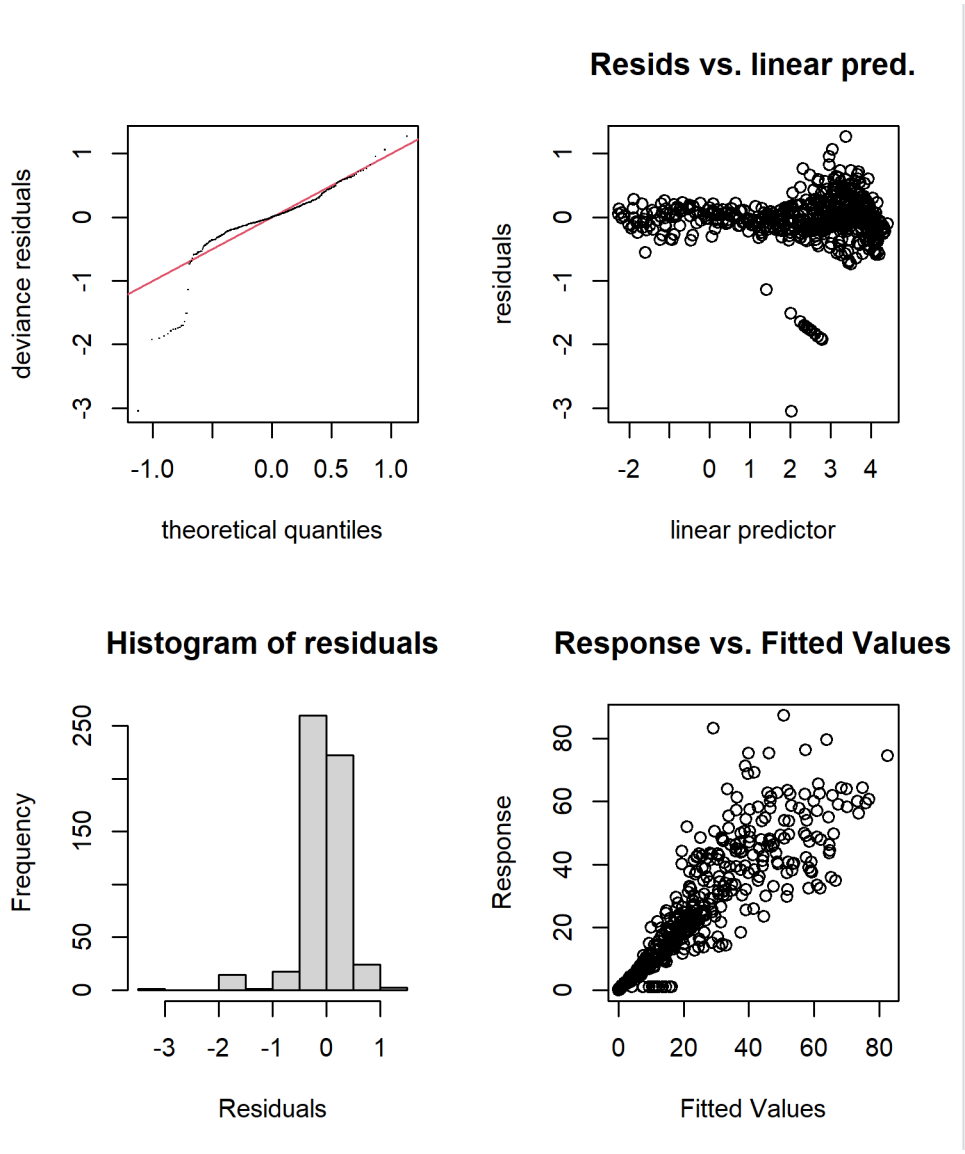


Figure 7: Output of the `gam.check()` function RemiMod.gam2.

If we compare the results from our mixed effect additive model, RemiMod.gam2, to our additive model that fits separately for each subject, RemiMod.gam0, we see that are some stark differences. Perhaps the most interesting difference between the results can be found in the plots displaying the response versus the fitted values (lower right in figure 7 and right in figure 8). Namely, the spread of the scatter is a lot wider and, moreover, more symmetrical for RemiMod.gam2 than for RemiMod.gam0. This pattern can also seen from the bottom left plot of figure 7, which shows that there are only slightly more negative residuals than positive residuals. Now, when we look at the right plot of figure 8, we see that the scatter is less symmetrical, with the spread showing a clear trend of response values being greater than the fitted values most of the time. This means that the residuals tend to be mostly positive and tend to be greater as the fitted values also become larger.

Still, this shows us that the fit is a fair bit more precise overall. However, we should mention that the quantile plot in figure 8 by no means displays normal behavior unfortunately, as can be seen from the confidence interval band in the left plot. We should note that for the plots below in figure 8, we have removed the results of the plot for which model produced a constant value, as well as some the fitted values corresponding to the first observation, as these residuals are not a fair representation of the overall performance of the model for all of the participants.

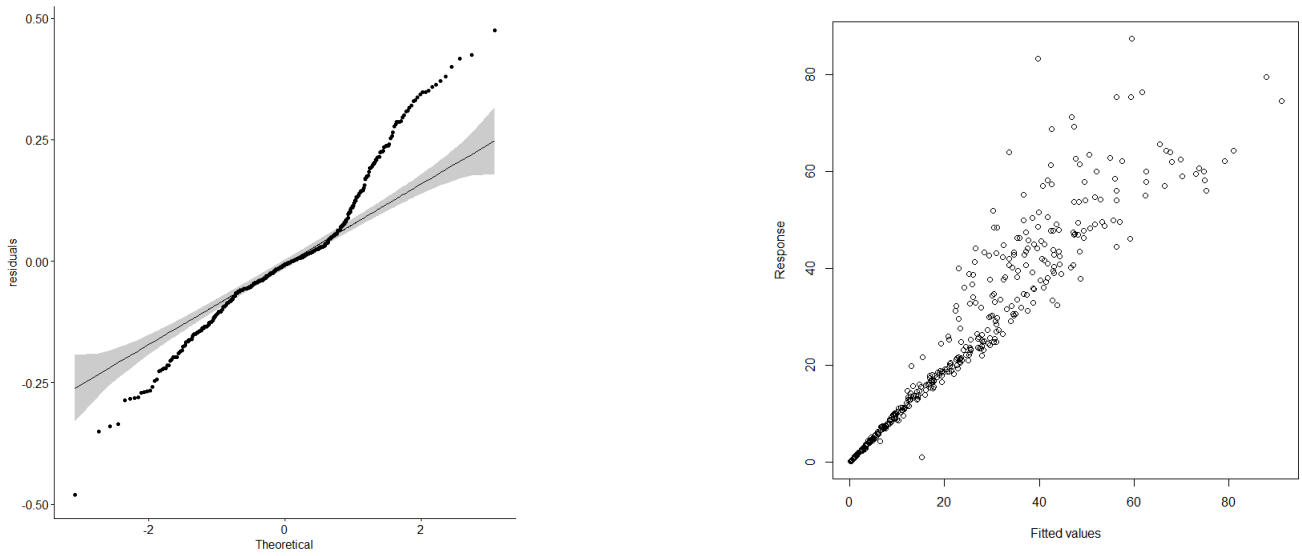


Figure 8: Left: Quantile plot of the residuals of the RemiMod.gam0 model. The band in this plot displays the 95% confidence interval for normality. Right: a fitted values vs Response plot.

So far we have assessed metrics such as residuals, AIC-values, correlation coefficients to judge whether a model has desired properties on its own, but also with respect to other models. Additionally, we have investigated behavior of residuals for our selected models to see what model structurally produces the most desirable type of fit (i.e. how are residuals of a model distributed generally). However, we still have to consider the condition number, κ, to see which models provide us with the most stable results. The results are printed in table 14 below.

Model	AIC	κ	R
RemiMod.nlme05	6385	34121.6	0.06
RemiMod.gam2	3038	583532.7	1
RemiMod.gam0	3426	137678.1	0.24

Table 14: Condition numbers of the Remifentanil models and the ratio R between the κ of the respective model and the largest κ of these models.

It can be seen from the table that the condition numbers are some differences between the model types we have constructed, with the smallest value being approximately twenty times smaller than the largest κ. This result was expected for the the non-linear mixed effect model and it is reassuring that was indeed the case. The reason that parametric models tend to be more consistent in their performance is because of the fact that the model function is included in the assumptions of the model and hence the model does not have to generate some arbitrary function to approximate the data, unlike non-parametric methods. Following up on the additive models, the difference is again very noticeable and expected because the RemiMod.gam2 model uses more smoothers on more variables and is therefore naturally more dependent on the input parameter vector θ^*. Before we conclude more about the performance of the model types as a whole, we will first consider the performance of our models on other data sets.

	fm4Theo.nlme
lKe_f	-2.45^{***}
	(0.05)
lKa_f	0.43^{*}
	(0.20)
lCl_f	-3.23^{***}
	(0.06)
lKa_σ	0.64^{***}
	(0.00)
lCl_σ	0.17^{***}
	(0.00)
N	132.00
No. groups	12.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	lKe	lKa
lKe	-0.344	
lKa	0.013	-0.146

Table 15: Summary table and correlation matrix of fm4Theo.nlme

3.2 Theophylline

3.2.1 Nlme

The following models that will be constructed are based on the theophylline data set. This sample has also been analyzed by Pinheiro and Bates (2000) and we will refer to this book for the exploratory analysis performed on the data. Additionally, the model selection steps were also presented in their book and we will therefore skip to the final theophylline model proposed. First of all, as general practice we generate a model using a self-starting function within the `nlsList` function and use this as the input for our basic nlme model.

```
fm1Theo.lis <- nlsList(conc   SSfol(Dose, Time, lKe, lKa, lCl) | Subject , data = Theoph)
fm1Theo.nlme <- nlme(fm1Theo.lis)
```

Again, this initial nlme model uses the `nlsList` input to determine starting values and the fixed and random effect structures. By default it assumes the most basic fixed and random effect R-studio model formula for each parameter (i.e. lKe + lKa + lCl $\sim$ 1 for both the fixed and random effects). According to Pinheiro and Bates (2000), the following model fits the data well.

```
fm4Theo.nlme <- update(fm3Theo.nlme, random = pdDiag(lKa   1, lCl   1)
  , weights = varConstPower(const = 1, power = 0.1))
```

From this model function we see that the random effect structure has been diagonalized and that the lKe parameter has been removed. These adjustments were made because there was high correlation between parameters and because the lKe parameter contributed virtually nothing to the random effects. Secondly, the `varConstPower` argument has been added in order to account for the heteroscedasticity in the standardized residuals. The summary of the final model below in figure 9 shows no worrying signs and hence we will use the model for our final analysis. The fit of the model is also shown in figure 9.

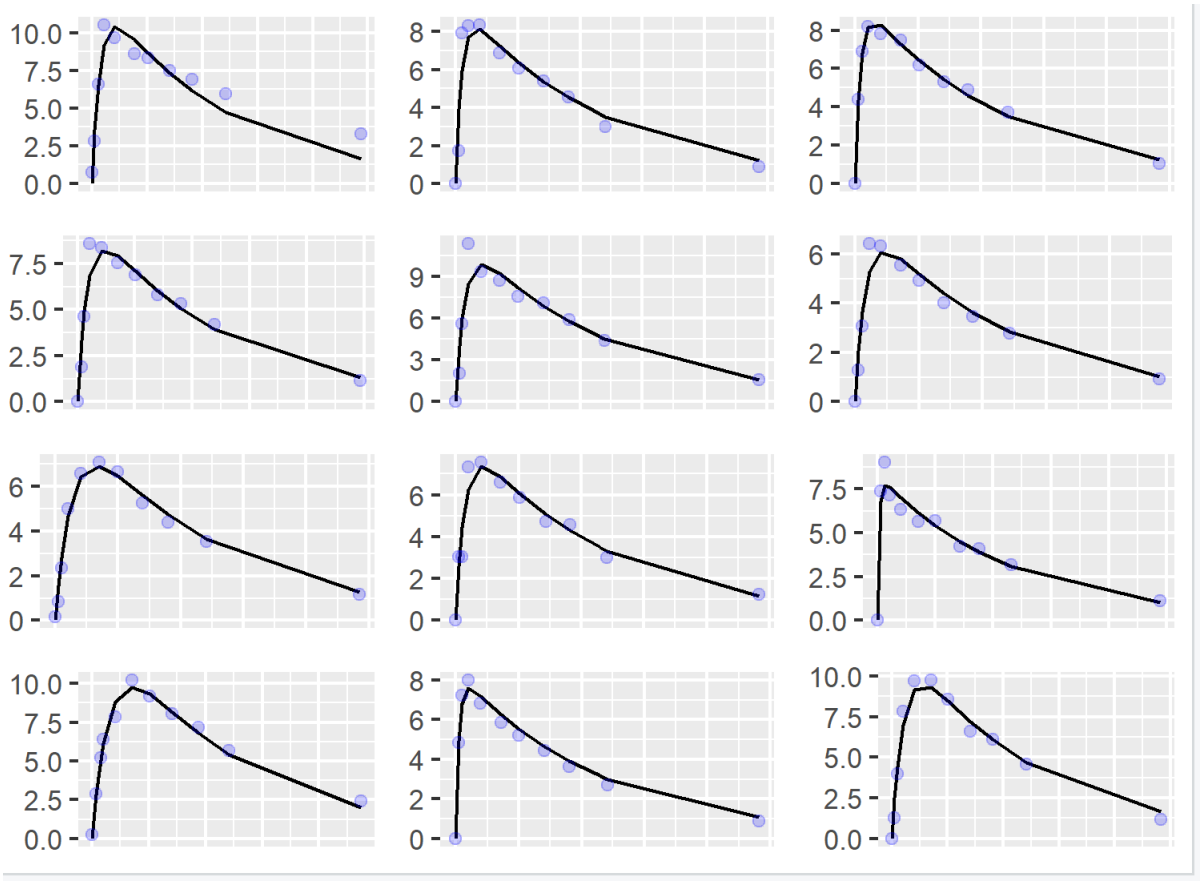


Figure 9: Plots of the fm4Theo.nlme model fits for each subject

3.2.2 GAM

For the construction of the additive models theophylline sample, we started with the same basic fixed plus random effect (random smooth) model structure shown below.

```
TheoMod.gam1 <- gam(conc   s(Time) + s(Time, Subject, bs = "fs"), data = Theo, method
= "REML" , gamma = .6 )
```

The gamma parameter was also reduced in order to reduce the smoothness of the fitted curve, to better represent the data. This did not change the overall patterns for each individual curve, but increased the overall fit quite well without overfitting. Subsequently, additional smooths were introduced into the model. However, upon adding these smooths, it appeared that the model had often become an overdetermined system because the model would require more parameters than the number of observed values. In the cases where the extra smooth did not introduce too many additional parameters, the contribution to the fit was not significant. Therefore, the best suited model for this data set appeared to simply the initial TheoMod.gam1 model. The summary and plots of the fit are displayed below.

	TheoMod.gam1
s(Time)	8.73***
s(Time,Subject)	22.22***
N	132.00
No. groups	12.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 16: summary of TheoMod.gam1

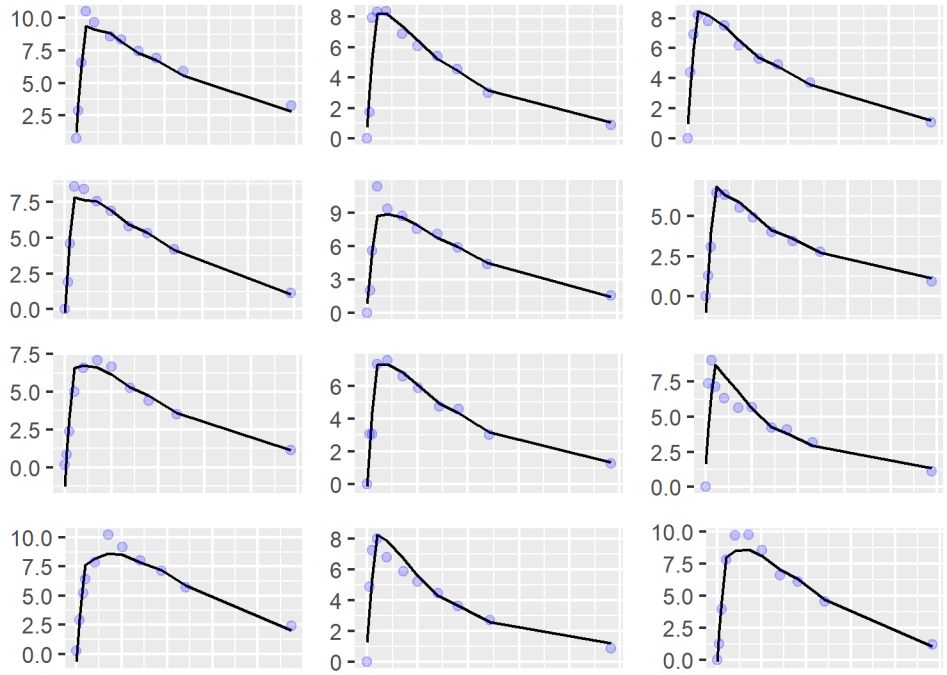


Figure 10: Plots of the TheoMod.gam model fits for each subject

3.2.3 Analysis of theophylline models

With the models being formulated, we again continue with the analysis of the quality of the model. Starting with the nlme model, in figure 11 we see that the the spread of the residuals in the left plot is approximately normal, except for the extreme values. Consequently, we also see in the right plot that the fitted values display relatively consistent behavior with respect to their observed values. All of this indicates that we have a suitable model at hand and that the model functions appropriately.

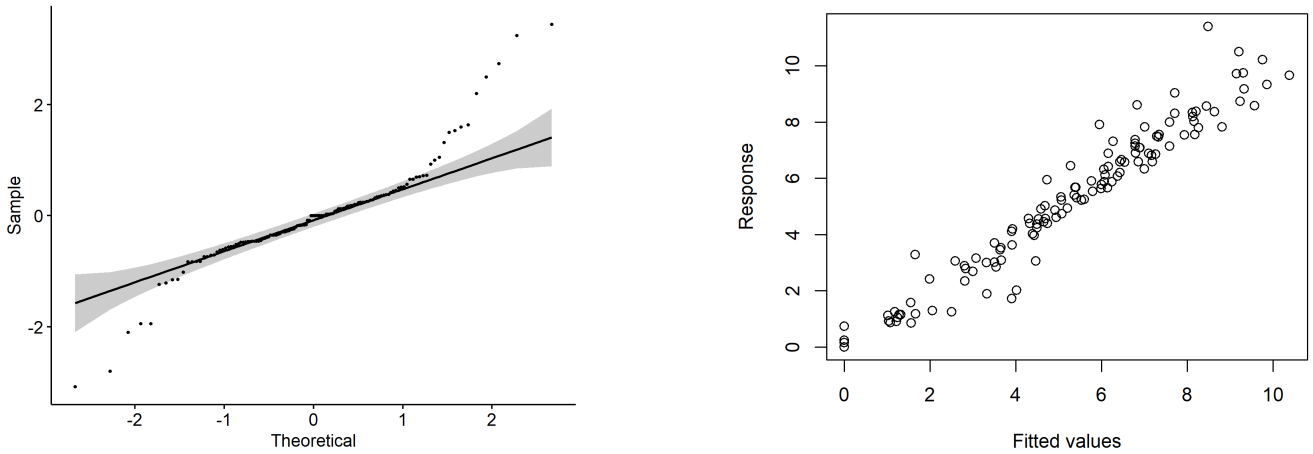


Figure 11: Left: quantile plot of the fm4Theo.nlme model. Right: scatter plot of the fitted versus the residual values

Secondly, we have the plots for the generalized additive model in figure 12. The quantile plot on the left shows us that the residuals are spread out mostly in a normal fashion. Most notably, there appear to be a few too many extreme residuals on the positive side, as indicated by the tail going upwards for the larger positive residuals. Additionally, we see in the plot on the right that the fitted values versus the response values show a clear diagonal trend. However, contrary to the plot in figure 11, we see that the fitted values do not quite reach the value 10 and up, contrary to the response values. This means that for the larger response values, the model structurally fits values to the peak that are generally smaller than the peak. This effect can also be seen in figure 10, most notably in plots (from left to right) 1, 4, 5, 10 en 12.

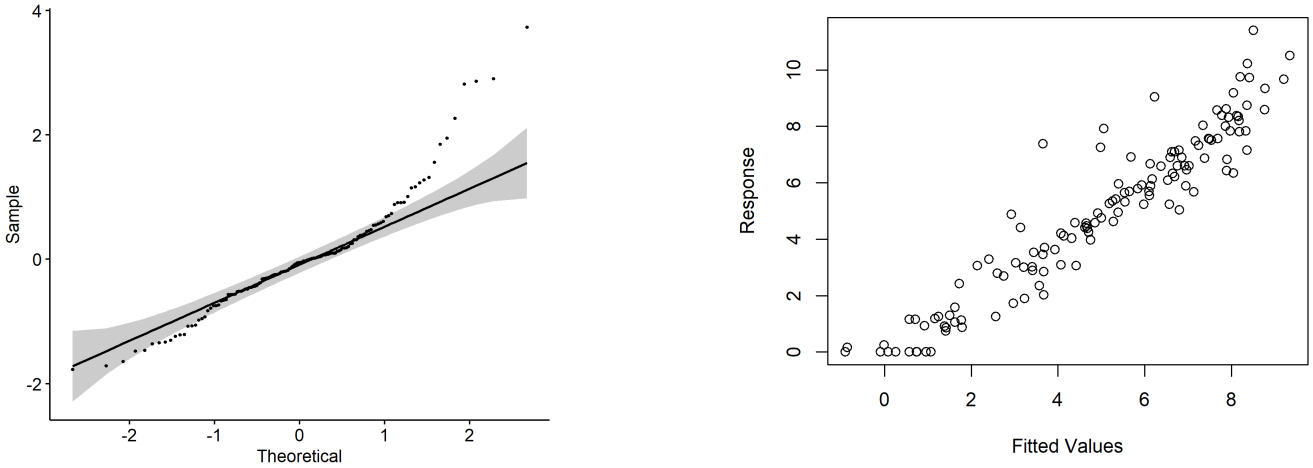


Figure 12: Left: quantile plot of the TheoMod.gam1 model. Right: scatter plot of the fitted versus the residual values

Lastly, we also have the condition numbers. For this sample, we see that the two models yield wildly different values for κ. The main takeaway from this is that the additive model is a lot more sensitive to changes in the data and hence to changes in the parameter vector θ^*, as discussed in the theory section.

Model	AIC	κ	R
fm4Theo.nlme	351	226.7416	$3.1*10^{-4}$
TheoMod.gam1	425	731631	1

Table 17: Condition numbers of the Theophylline models and the ratio R between the κ of the respective model and the largest κ of these models.

3.3 Dialyzer

3.3.1 Nlme

For our last sample, we have the dataset of Vonesh and Carter (1992) which was described in the theory section. Again, we start with building the nlme model first. A suitable model was demonstrated and proposed by Pinheiro and Bates (2000), but in order to demonstrate the potential (mis)use of the nlme model, we opt for using a new parametrization that also generally fits the spread of the data. The introduced parametrization is of course still a function of pressure, but belongs to a different family of functions compared to the parametrization used in Bates and Phinheiro (2000). The family used in our model is given by the parametrization that satisfies the following:

$$f = b_1 * p^{b_2} * e^{-p*b_3}$$

where, b_1, b_2 and b_3 are just some parameters and p represents the independent variable of pressure. In this case, these parameters are not chosen in such a way that they represent a specific physiological effect —although this could have been the case like in the self-starting functions from the nlme package—, but rather as an arbitrary collection of parameters that happen to fit the spread of the data. Now, for the nlme model, we start with creating the non-linear least squares model to use as our input for our `nlme` command.

```
DiaMod.lis <- nlsList(rate ~ b1*Pressure^(b2)*exp(-Pressure*b3)|Subject , data = Dia2
, start = c(b1 = 60, b2 = 1.5, b3 = 1))
```

From the pairs plot in figure 13, we see that there is some correlation between the parameters b_1 and b_3, but it does not seem to be too worrying from this point. After feeding this model to our nlme model and inspecting the summary in table 18, it appears that the standard deviation for the random effect of the b_2 parameter is essentially negligible given its order of magnitude, as shown below.

	DiaMod.nlme0
$b1_f$	71.87***
	(3.01)
$b2_f$	1.55***
	(0.05)
$b3_f$	0.69***
	(0.04)
$b1_\sigma$	8.55***
	(0.00)
$b2_\sigma$	0.00***
	(0.00)
$b3_\sigma$	0.09***
	(0.00)
N	140.00
No. groups	20.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	b1	b2
b1	0.637	
b2	0.856	0.798

Table 18: Summary table and correlation matrix of DiaMod.nlme0

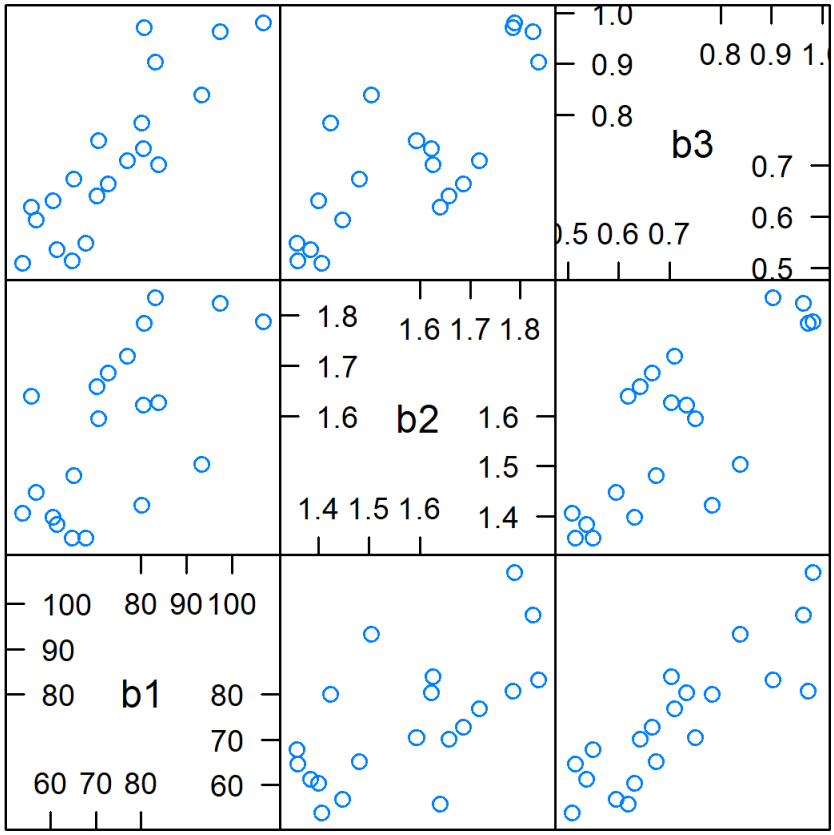


Figure 13: Plots of every possible pair of coefficients, for every coefficient in the DiaMod.lis model.

Therefore, we decide to drop it as a random effect to see if it improves the model, as shown below

```
DiaMod.nlme1 <- nlme(DiaMod.lis, random = list(b1 ~ 1, b3 ~ 1)
```

	DiaMod.nlme1
$b1_f$	71.88***
	(3.01)
$b2_f$	1.55***
	(0.05)
$b3_f$	0.69***
	(0.04)
$b1_\sigma$	8.55***
	(0.00)
$b3_\sigma$	0.09***
	(0.00)
N	140.00
No. groups	20.00

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

	b1	b2
b2	0.637	
b3	0.856	0.798

Table 19: Summary table and correlation matrix of DiaMod.nlme1

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
DiaMod.nlme0	10	792.1557	821.5721	-386.0779		
DiaMod.nlme1	7	786.1505	806.7420	-386.0753	0.0052	0.9999

Table 20: Anova table of the semi-parametric DiaMod.nlme0 and non-parametric DiaMod.nlme1

```
DiaMod.nlme3 <- nlme(DiaMod.lis , random = list(b1 ~ 1, b3 ~ 1) , weights = varPower(form
= ~ Pressure + I(Pressure)^2) )
```

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
DiaMod.nlme1	7	768.1505	806.7420	-338.7195		
DiaMod.nlme3	8	786.1995	809.7327	-385.0998	1.9510	0.1625

Table 21: Anova table of the semi-parametric DiaMod.gamm1$lme and non-parametric DiaMod.gamm2$lme

As can be seen in table 21, the similar values in AIC and the higher value of BIC for the DiaMod.nlme3, and non-significant p-value, show that the model is not significantly better and that the more parsimonious model, DiaMod.nlme1, is to be preferred. What's more, figure 14 shows that the fit of the final model shows a very plausible and nice fit.

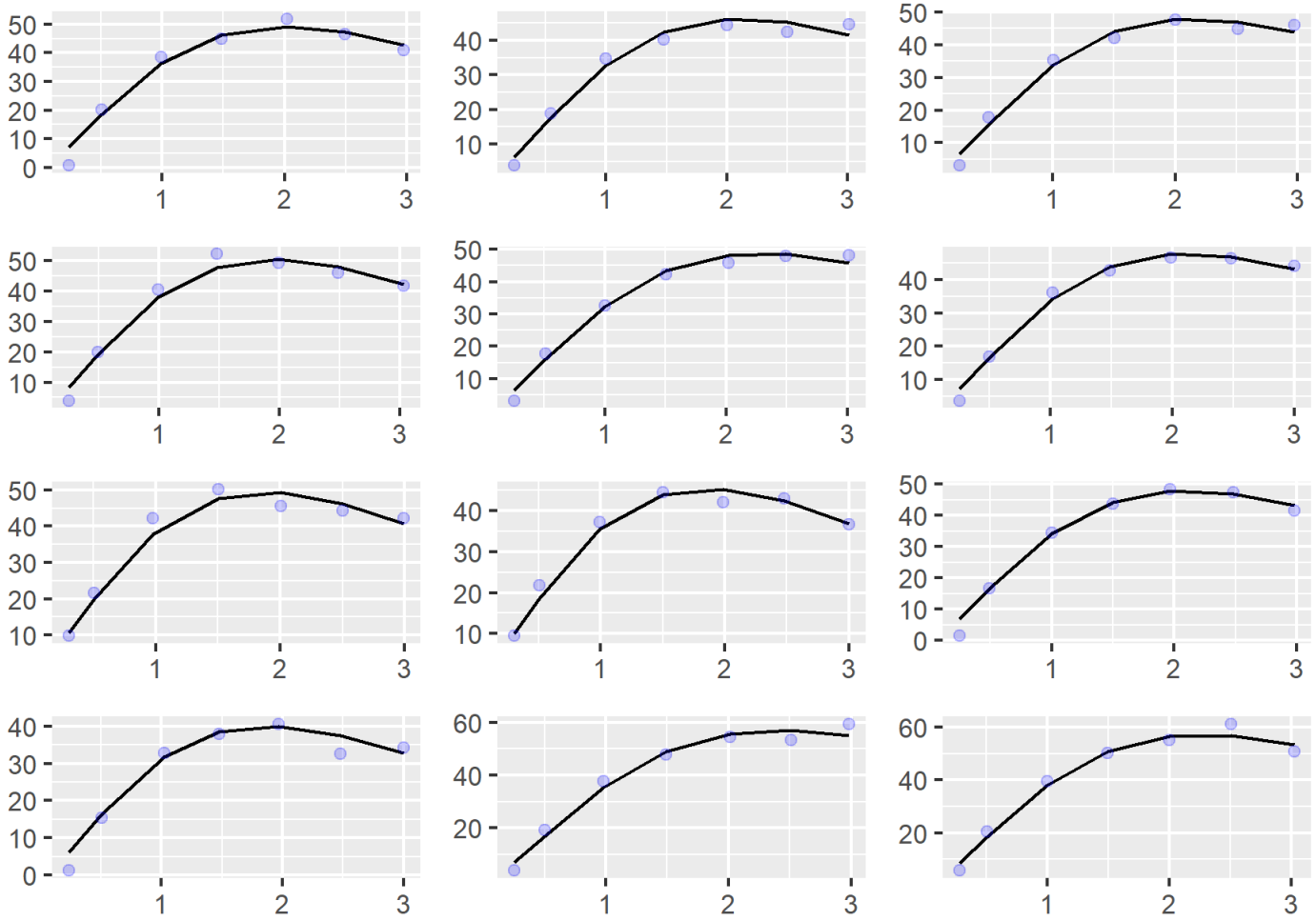


Figure 14: Plot of the DiaMod.nlme1 model fits for 12 of the subjects in the dialyzer data set with the pressure on the horizontal axis and the rate of the bloodflow in the dialyzer on the vertical axis.

3.3.2 GAM

In our generalized additive model, the `gamm` function was used in order to exploit the functionalities of the nlme package. Based on the work from Krijnen and Trapman (2022), the model below was obtained as the best model found in this research.

```
DiaMod.gamm1 <- gamm(rate ~ s(pressure) , random = list(Subject = ~ pressure +
I(pressure^2)) , data = Dia , method = "REML" , weights = varPower(form = ~ pressure)
```

This model equals the final model proposed by Krijnen and Trapman with the addition of the variance power function on the weights of the variance. This yielded a significantly better model. Consequently, the inputs for the random effects were checked after this addition to the model and the steps taken by Krijnen and Trapman were reiterated in order to see if these steps were still similarly valid, which they were. However, it should be noted that because functionalities of the nlme package were used through the `gamm` function, certain structures on the random effects and weights were specified, meaning that these components were parameterized. Therefore, this model is no longer in the realm of strictly non-parametric models, but rather belongs to the class of semi-parametric models. In the summary table and plot below we can see that the semi-parametric model shows no immediate issues and that the fit seems to be decent.

	DiaMod.gamm1$lme
X(Intercept)$_f$	37.38***
	(1.02)
Xs(pressure)Fx1$_f$	31.87***
	(5.18)
(Intercept)$_\sigma$	2.16
Pressure$_\sigma$	5.56
Pressure$^2_\sigma$	1.67
N	140.00
No. groups	20.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

	Intercept	Pressure
Pressure	-0.603	
Pressure^2	0.117	-0.727

Table 22: Summary table and correlation matrix of DiaMod.gamm1$lme

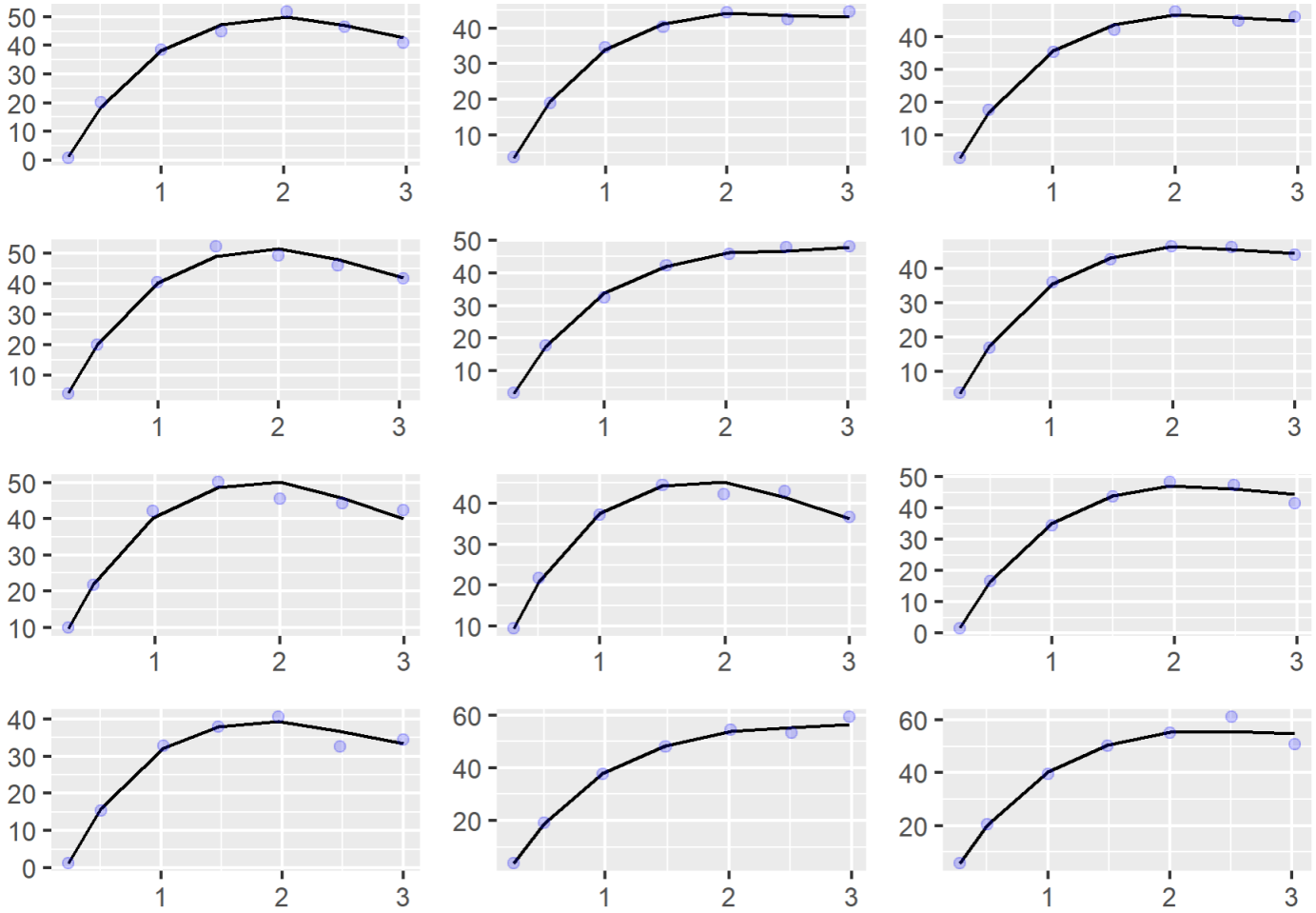


Figure 15: Plot of the model fits of DiaMod.gamm1 for 12 subjects.

In order to further complement the list of types of models investigated for this sample, a fully non-parametric model was also created in order to compare the performance of all of these models. To this end, a model was introduced similar to RemiMod.gam2 and TheoMod.gam1, where the fixed effect was represented by a regular smooth of the main independent variable and the random effect represented by a factor smooth of the same independent variable. This led to the following model, summary and plot. In the summary we see that again, there are no immediate glaring problems with the model and that the fits seem to be good. Additionally, the fit also seems suitable.

	DiaMod.gamm2$lme
X(Intercept)$_f$	37.39***
	(0.95)
Xs(pressure)Fx1$_f$	29.87***
	(5.35)
Xr.1$_\sigma$	34.07***
	(0.00)
Xr.2$_\sigma$	34.96
N	140.00
No. groups	20.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 23: Summary table of DiaMod.gamm2$lme

```
DiaMod.gamm2 <- gamm(rate   s(pressure) + s(pressure, Subject, bs = "fs") , method =
"REML" , select = TRUE , data = Dia )
```

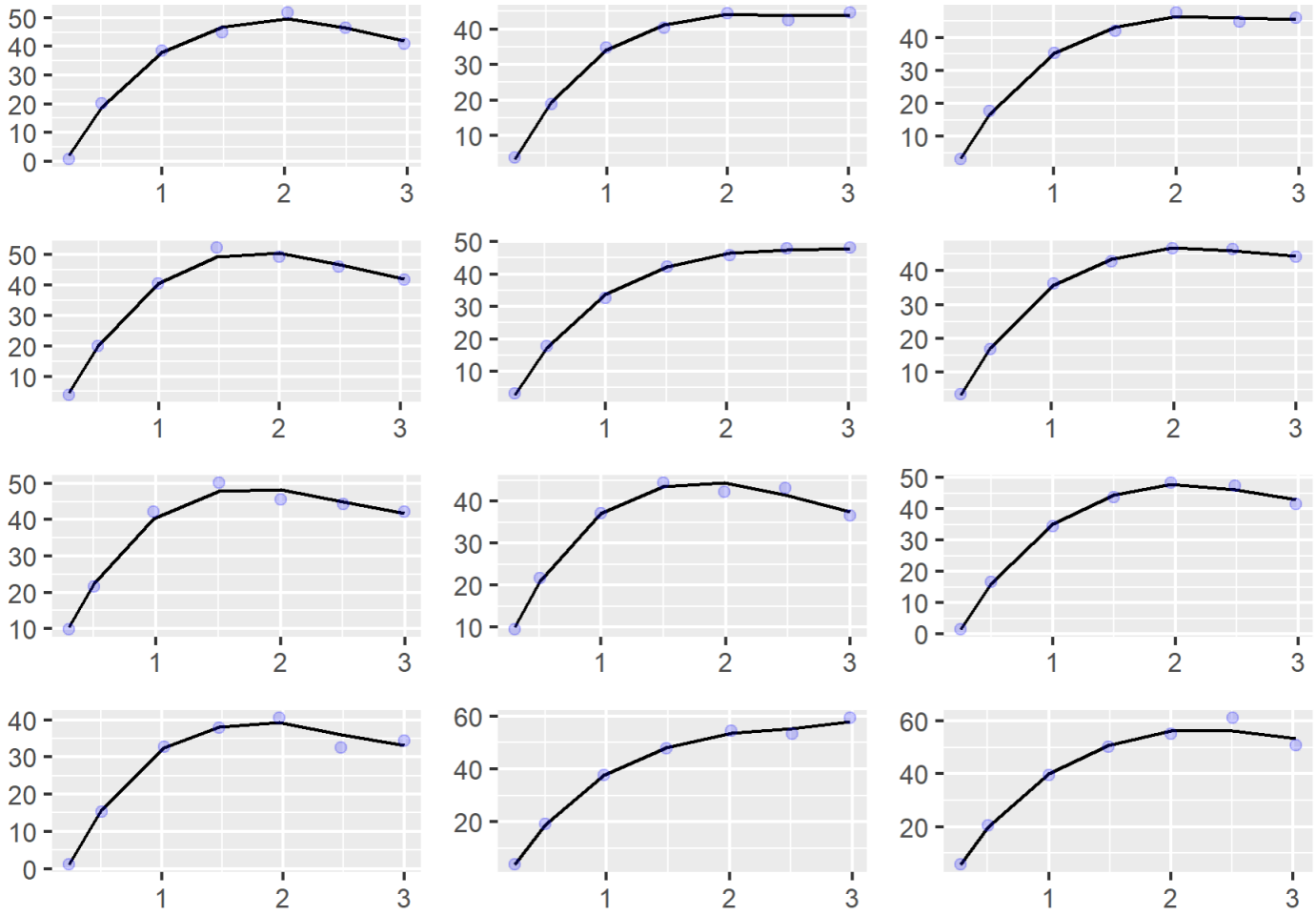


Figure 16: Plot of the model fits of DiaMod.gamm2 for 12 subjects.

Now, because we have used the `gamm` for both models, we can inspect the ANOVA table for these two models and compare them directly. As can be seen in table 24, the semi-parametric model performs significantly better,

showing the power of mixing parametric and non-parametric methods.

Model	df	AIC	BIC	LogLik.	L.Ratio	p-value
DiaMod.gamm1$lme	11	699.439	731.6388	-338.7195		
DiaMod.gamm2$lme	7	759.673	780.1638	-372.8365	68.23395	$< .0001$

Table 24: Anova table of the semi-parametric DiaMod.gamm1$lme and non-parametric DiaMod.gamm2$lme

3.3.3 Analysis of dialyzer models

For the comparison of the dialyzer model, the same diagnostic plots were performed in order to assess further assess and compare the model fits in terms of normality, skewness of the residuals and numerical stability of the methods. First of all, the nlme model for the dialyzer data does not show any major defects in the model as its residuals follow a normal distribution and does not suffer from heteroscedasticity. These things can be clearly observed in figure 17 and show that DiaMod.nlme1 indeed is a suitable model for the data.

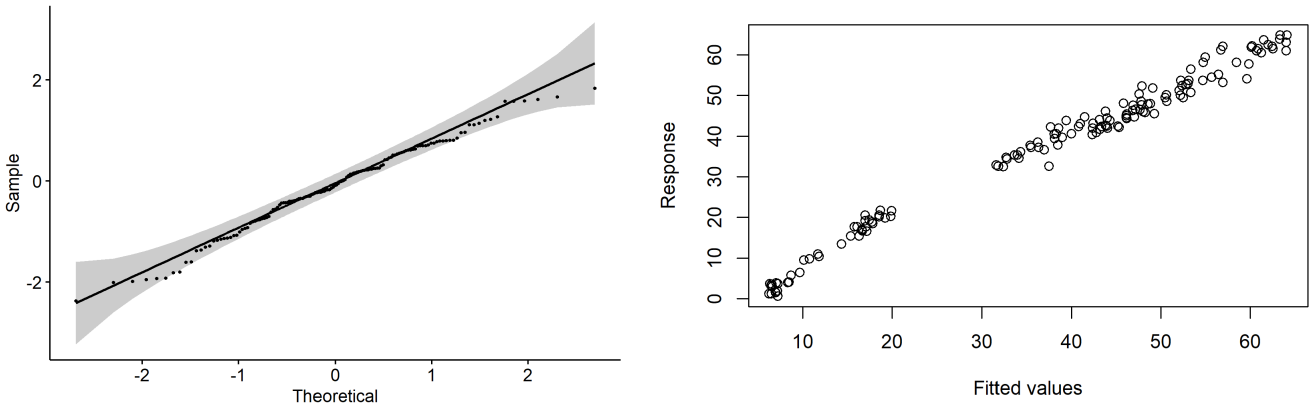


Figure 17: Left: quantile plot of the DiaMod.nlme1 model. Right: scatter plot of the fitted versus the residual values

Additionally, we have both the non-parametric and the semi-parametric to analyse. For these models we see a slight disparity in the for the quantile plots, while the fitted values versus the response values plots are quite similar. Namely, the quantile plot for the semi-parametric model, DiaMod.gamm1, seems to behave better in terms of normality of the residuals compared to the non-parametric plot. This can be inferred from the fact that the quantile plot in figure 18 shows more points outside of the confidence interval band for the more extreme values of the residuals. Even though this is a common phenomenon, it still shows that DiaMod.gamm1 has slightly more desirable properties. On the other hand, the patterns shown in the right plots for both figures 32 and 33 seem nearly identical and show no heteroscedasticity. However, both models are still acceptable as the slightly lacking normality only holds for few extreme values of the residuals and is hence not too worrying usually.

Secondly, we have the condition numbers of these three respective models. However, because of the way the mixed effect model is realised through the `gamm` function implies that there are two model components to each model. The gam component represents only the fixed effect estimated by the mgcv package machinery and the lme component represents the model fixed and random effects estimated by the nlme package machinery. Therefore, the gam components were omitted from the analysis because they have no estimated random effects. Interestingly enough, table 3 shows that the semi-parametric model has the highest condition number, implying that this method is the least robust numerically with regard to the estimated log-likelihood. Moreover, the most robust method would be the non-parametric method, which seems counter-intuitive but can be understood better given that the number of

parameters in this model is close to the number of parameters in the nlme model, judging from the the `apVar` output from both models. The parametric and non-parametric models hold four and five parameters respectively, which partially explains why the condition numbers could be similar. In contrast, the semi-parametric model uses nine parameters, which also explains why the condition number is roughly double that of both of the other models. For these reasons, given the diagnostic plots in figures 31 to 33 and the condition numbers, it seems that the parametric method is the best functioning method in terms of model properties. Lastly, the semi and non-parametric methods seem similar in performance as the semi parametric model performs better based on the quantile plots in figures 29 and 30, but performs worse in terms of numerical stability and vice versa for the non-parametric method.

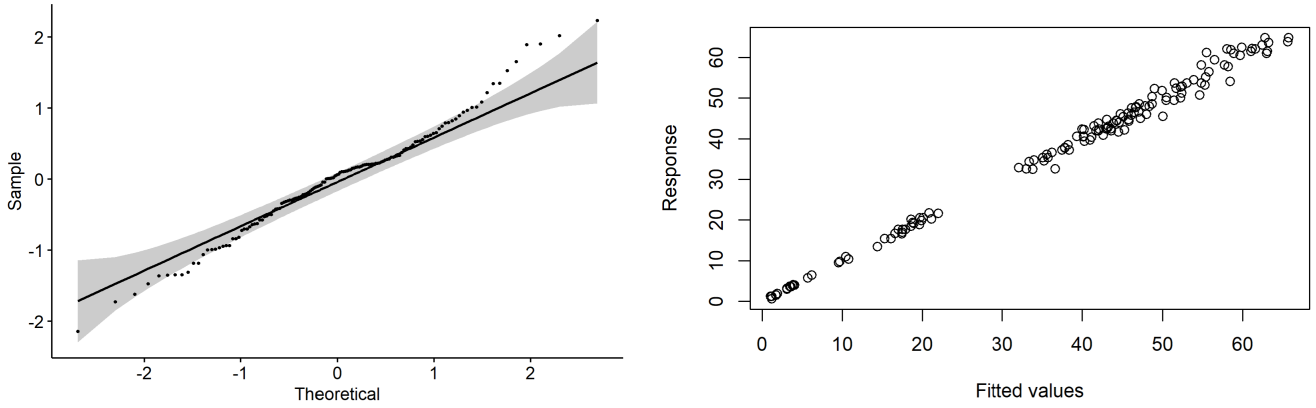


Figure 18: Left: quantile plot of the DiaMod.gamm1 model. Right: scatter plot of the fitted versus the residual values

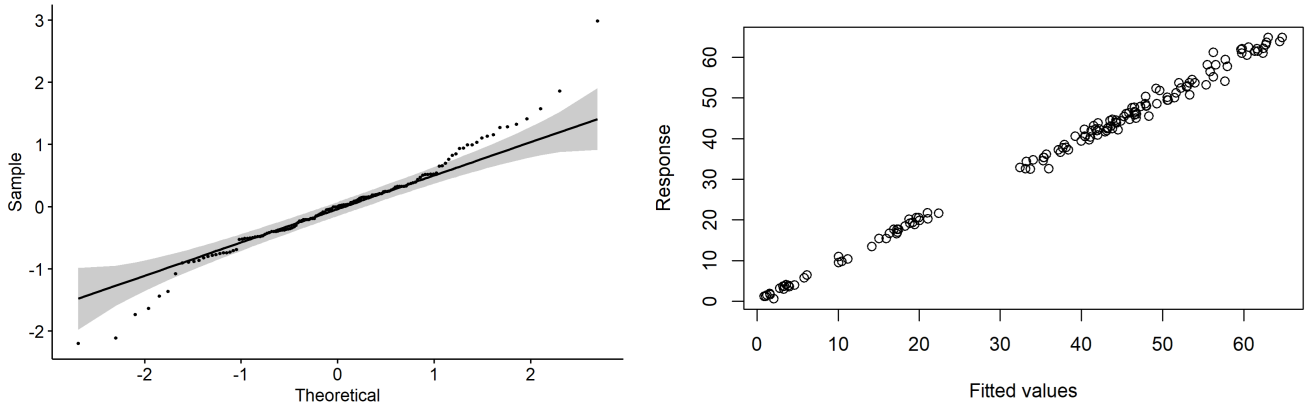


Figure 19: Left: quantile plot of the DiaMod.gamm2 model. Right: scatter plot of the fitted versus the residual values

Model	AIC	κ	R
DiaMod.nlme1	768	71.72192	0.67
DiaMod.gamm1$lme	699	108.7929	1
DiaMod.gamm2$lme	759	48.34529	0.44

Table 25: Condition numbers of the Dialyzer models and the ratio R between the κ of the respective model and the largest κ of these models.

4 discussion

Throughout this thesis, varying results have been found. Most notably, the parametric models were not absolutely the most stable models, although this was nearly the case. As for the parametric model for the dialyzer sample, the chosen parametrization of the rate as a function of pressure also influences the quality of the model. For the particular parametrization used in this paper, the result was satisfactory, but comes with the problem that we are stuck with the used parametrization and that it is not possible to show in this moment that is parametrization was the optimal one. As can be seen in the final chapters of Bates & Pinheiro (2000), the dialyzer model built shows a similar fit, but uses a different parametrization based on the self starting functions provided by the nlme package. Consequently, other questions also arise about what parametrization to choose and on what basis is one supposed to make this choice. It can be argued that one should use a parametrization based on findings from literature on the respective topic or drug and this seems reasonable. However, such parametrizations are still dependent on particular papers with their own respective confounds and limitations, meaning that the parametrization would become reduced to the best possible option that the literature so far has been able to unveil. When standing on this crossroad, the appeal of non-parametric or semi-parametric methods become more evident. The ability to have the parametrization be implicitly determined by the empirically obtained measures, allows one to perform analysis and prediction based on what is measured, rather than what was expected.

On the other hand, the beauty of the non-parametric or semi-parametric methods is also its necessary weakness. The fact that there is no explicit relationship assumed between the dependent variable, independent variables and other parameters, means that it becomes harder make derivations from the model for specific variables simply because this relationship is not explicitly assumed nor determined in the process. Therefore, the appeal for parametric methods remains and has their own place in science and with great success. When strong evidence for certain parametrizations is available, these parametrizations allow for prediction with the ability to explicitly quantify the relationship between variables and parameters and the dependent variable. This trade-off was also clearly visible in the case of remifentanil sample. The RemiMod.nlme05 model had some trouble with matching the peaks present in the observed values for multiple Subjects. The most likely cause of this is that the parametrization used was not well suited to model the shapes of these peaks given the available parameters. Therefore, such a finding implies that some different parametrization might be required to more accurately model the data, such as a parametrization based on a multicompartmental differential equation system (Eleveld et al., 2017).

Secondly, the models built for the theophylline data set seem strikingly similar, both in their respective quantile plots and fitted versus response value plots. This means that both models do a similar job at approximating the observed values. However, table 17 also shows that the condition number for the additive model is several orders of magnitude larger than its nlme counterpart. This is most likely explained by the fact that the data contains a small collection of data points per subject in combination with the fact that the spread of the data points is more difficult to model because of the steep peak early on in the measurements. This makes balancing the wiggliness penalty and minimizing residuals more difficult because following the spread of the observed values would come at the cost of a sharp increase in wiggliness in the model. Therefore, the additive model is more prone to changes in the data because these might give the model more reason to decide more strongly between wiggliness and fit. As for the similarities, these came from the distribution of the residual being close to normal for both methods. Additionally, the locations of outliers relative to the fitted values appeared to be similar too, showing that the models generated very similar fits. All in all, both methods yielded similar results in terms of fit, but the nlme model came with better qualities as it will be less sensitive to changes in the data and while performing similarly regarding the normality of the residuals between the models.

It would be interesting to further investigate and formalise the properties of a data set with the goal of assessing how well suited an additive model would be for the current spread of the data. This could be formalised by using metrics such as amount of data points, the distance between these points and some metric on the first derivative of

the lines obtained using linear interpolation for instance. This way one could determine a priori if the application of an additive model could prove to be problematic and it would help in choosing a suitable modelling method. In the context of the theophylline data set, this could mean that a generalized additive model would not be recommended by such a tool, because the model would vary relatively much for small changes in the data, meaning that the model would behave less predictably when predicting for different theophylline data sets.

However, at the core of these models lie fundamentally different model functions, ranging from sums of of polynomials to explicit parametrizations. For that reason, it is challenging at a theoretical level to soundly compare the quality of two such different models. Most notably, because the models are not nested by definition, comparison between model types through regular anova analysis and AIC scores is meaningless. This was the main motivation to choose to look at model fits through checking the normality of the residuals, but also through investigating the observerd patterns of these residuals for each fitted value and their respective observed value, sometimes indicating heteroscedasticity to some degree at times for instance. In order to improve on the current insights on parameters and variables, an additional analysis through the `brms` from the rstan package. This package provides tools to perform Bayesian regression model analysis through the use of Markov Chain Monte Carlo simulations. This Bayesian approach comes with the advantage of being able to alos investigate the stability of the solution through assessing the distribution of all kinds of variables, parameters and statistic, allowing one to say additional meaningful things about how appropriate a certain parametrization or additive model formula, given the data. Therefore, the integration of Bayesian regression could provide more insights on the properties of multiple parts of the model, allowing one to give a more detailed comparison between model, especially when the models behave relatively similarly based on a frequentist approach such as for the theophylline models.

For the Dialyzer data set, a relatively thorough model building process was followed, similar to that of the remifentanil models. In addition to the parametric and non-parametric models, a semi-parametric model was also introduced based on some of the properties of the data that were known from prior research by Pinheiro and Bates (2000) and Krijnen and Trapman (2022). All of the models seemed to perform quite well also having relatively similar condition numbers and quantile plots. When we compare this to our findings of the previous samples, we see that the high condition numbers are found when the curvature of a fit needs to be high and/or the amount of observations is low. However, in this sample we found that the condition numbers were small, even though there was only a limited amount of observations per participant. Therefore, it makes sense to suspect that the shape of the fit plays a considerable role in the stability of the parameter vector θ^*. This consideration highlights the limits of the empiricist nature of the semi- and non-parametric methods. Even though it might be more objective and can fit any curve, it might still struggle based on challenges hidden in the shape of the data points, as it can only work with what is observed.

However, as we touched on before in the discussion, the parametrization used for the DiaMod.nlme1 model is known to not be the only viable parametrization and it begs the question how much value the parametrization adds, given that our parametrization is not based on a physiological phenomenon or other measured variables. For that reason, one could claim that the non-parametric or semi-parametric models are preferable as they provide models that are similar in quality and performance, but are more objective in the sense that no explicit parametrization was assumed beforehand. A nice addition to this thesis would be to formally explore more of the variations of possible parametrizations to see which would perform well and why. Looking back, the results from the models on this sample displayed the considerable value of the non-parametric methods given the right conditions, especially when a suitable parametrization based in physiology is unknown or not forthcoming.

One limitation of this study would be that the approach taken for the remifentanil study does not translate very well to the models built for the other data sets. The main difference, namely, is that no factor smooth was used for RemiMod.gam0, but rather a smooth that was fitted for each subject. This meant that no real mixed effect model was implemented after all and that the implications from this model are harder to compare to the implications from

the models that did include both fixed and random effects. Therefore, the contribution of this model to the aim of this thesis is limited.

Additionally, the method of obtaining RemiMod.nlme3 could also have been more refined. Although, the method did yield a decent fit and provided a starting model that was clearly improved upon, a more formal approach could prove to be worthwhile. For instance, using a function suitable for generalized additive models similar to `step.AIC` function from the MASS package could be nice attempt to more structurally exhaust the list of possible combinations of smooths. The fact that there were so many potential smooths in this model, given the elaborate data set, meant that a lot more combinations of smooths had to be explored and with good reason, as the best model found used more smooths than just a basic fixed effect smooth and a factor smooth for the random effects. In addition, the large sample size of this data set was also relatively problematic because it caused some of the run times to become quite extensive, taking sometimes up to over 30 minutes to generate a generalized additive model. In order to combat this problem, a model was generated using a random subset of 16 subjects in order to reduce the amount of required computations. Unfortunately, this means that every iteration of a model could be slightly different because of the different subjects that were taken from the pool each time. To compensate for this, the models were run multiple times to check if different properties would arise. Fortunately, a subset of 16 subjects was large enough to create very similar results for each model. A possible solution to this problem could have been to use the `bam` function from the mgcv package because it is designed to handle larger data sets. All in all, there is room to improve on the performed analysis by expanding on the exploration of the best set of smooths and on adopting different functions to decrease computational workload.

5 conclusion

Multiple models with different types were employed for three different types of data sets, each model performing with some degree of succes. Between these models, the results from the ones that were used for the remifentanil sample were perhaps the most interesting as the models showed varying abilities to fit a model to the data. For instance RemiMod.gam2, had some decent qualities with respect to its residuals, but seemed to perform worse when the observed values were not spread out nicely enough. This showcased the robustness of the parametric methods, which would more or less force the graph of the fitted values to follow a certain shape, suffering less from an inconvenient spread of the data for some subjects.

On the other hand, the limits of mixed effect models was also displayed through the RemiMod.gam0 model. Because this model used a different fit per subject, essentially having random effects explaining everything for each subject, showed that such an approach also bears fruit when model fitting is the focus. However, this approach has even less room for interpretation than the mixed effect model because the common denominator of the fixed effects is no longer there as a point of reference. Regardless of the nice properties of mixed models, we saw that not every type of mixed model came with desirable properties. For instance, the normality of the

The models for the Theophylline sample had more similar performances. The main difference between the two models is that the shape of the plot was slightly different in the peaks with the nlme model generating more round or smooth peaks, whereas the peaks of the gam model fits often were more sharp and seemed less natural becuase of the lack of smoothness. Therefore, we know that a gam model is not the best candidate for a pharmacokinetic data set if there is a scarce amount of observed values in a shape that is not very smooth.

Lastly, the dialyzer sample was analyzed. Interestingly enough, the scarcity of available data points per subject was not as much of an issue for the stability of the additive model. This probably comes from the nice and smooth shape the observed values form as it is the main difference between this sample and the theophylline sample. Additionally, the models were very similar for all three types of models used in many regards. First of all, the residuals for all three models closely resembled a Gaussian distribution. Most notably, however, was that all the

condition numbers were relatively close and in the same order of magnitude, the model with the largest κ being the semi-parametric model and the lowest being the fully parametric model. This was still somewhat surprising because the parametric model usually displays better robustness. However, a meta analysis or dedicated research on the exact conditions of under what conditions additive models perform well should be further investigated before we can confidently state this as a fact.

All in all, this research has given insights in the qualities and limitations of various applications and implementations of mixed modelling. We have seen under what circumstances additive models fail compared to parametric nlme models, but also when non-parametric additive models provide advantages over parametric non-linear models in the mixed model context. Furthermore, we hope that this research has shown the merits of utilizing non-parametric models and how they can be implemented.

References

Boeckmann, A. J., Sheiner, L. B., & Beal, S. L. (1994). *Nonmem users guide: Part v.* NONMEM Project Group, University of California, San Francisco.

Eleveld, D. J., Proost, J. H., Vereecke, H., Absalom, A. R., Olofsen, E., Vuyk, J., & Struys, M. M. R. F. (2017). An allometric model of remifentanil pharmacokinetics and pharmacodynamics. *Anesthesiology*, *126*(6), 1005-1018. doi: 10.1097/ALN.0000000000001634

Kim, T. K., Obara, S., Egan, T. D., Minto, C. F., Colla, L. L., Drover, D. R., ... the Remifentanil Pharmacokinetics in Obesity Investigators (2017). Disposition of remifentanil in obesity: A new pharmacokinetic model incorporating the influence of body mass. *Anesthesiology*, *126*(6), 1019–1032. doi: https://doi.org/10.1097/ALN.0000000000001635

Krijnen, W. P., & Trapman, J. P. (2022). *Mixed models with gamm.*

Minto, C. F., Schnider, T. W., & Shafer, S. L. (1997). Pharmacokinetics and pharmacodynamics of remifentanil: Ii. model application. *Anesthesiology*, *86*(1), 24-33. doi: 10.1097/00000542-199701000-00005

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-plus (ser. statistics and computing).* Springer.

Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society. Series B (Methodological)*, *47*(1), 1-52.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583-639. doi: https://doi.org/10.1111/1467-9868.00353

Vonesh, E. F., & Carter, R. L. (1992). Mixed-effects nonlinear regression for unbalanced repeated measures. *Biometrics*, *48*(1), 1-18.

Wood, S. N. (2017). *Generalized additive models : an introduction with r* (2nd ed.). Boca Raton : CRC Press/Taylor Francis Group.