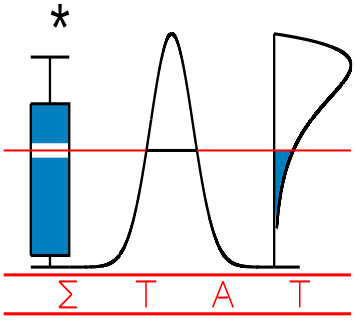
T E C H N I C A L
R E P O R T

0632

PHARMACOKINETIC PARAMETERS ESTIMATION USING ADAPTIVE BAYESIAN P-SPLINES MODELS

JULLION, A., LAMBERT, P., BECK, B., and F. VANDENHENDE



I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Pharmacokinetic parameters estimation using adaptive Bayesian P-splines models

Astrid Jullion[1],Philippe Lambert[1,2],Benoit Beck[3] and F. Vandenhende[3] *

[1]*Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgium*

[2]*Unité d'épidémiologie, biostatistique et méthodes opérationnelles, Faculté de Médecine, Université catholique de Louvain, Belgium*

[3]*Eli Lilly, Mont-Saint-guibert, Belgium*

September 19, 2006

ACKNOWLEDGEMENTS

Astrid Jullion thanks Eli Lilly for financial support through a patronage research grant and the UCL for a FSR research grant. Financial support from the IAP research network nr P5/24 of the Belgian State (Federal Office for Scientific, Technical and Cultural Affairs) is also gratefully acknowledged by Philippe Lambert. No relevant conflict of interest has to be declared.

*Correspondence to: Philippe Lambert, Université catholique de Louvain, Institut de Statistique, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve (Belgium). E-mail: lambert@stat.ucl.ac.be Phone: +32-10-47.28.01 Fax: +32-10-47.30.32

Contents

List of Figures

SUMMARY

In preclinical experiments, pharmacokinetic (PK) studies are designed to analyse the evolution of drug concentration in plasma over time i.e. the PK profile. Some PK parameters are estimated in order to summarize the complete drug's kinetic profile : area under the curve (AUC), maximal concentration (C_{max}), time at which the maximal concentration occurs (t_{max}) and half life time ($t_{1/2}$).

Several methods have been proposed to estimate these PK parameters. A first method relies on interpolating between observed concentrations. The interpolation method is often chosen linear. Another method relies on compartmental modelling. In this case, non linear methods are used to estimate parameters of a chosen compartmental model. Two problems can arise with this method. The first one is the difficulty to choose the suitable comparmental model given the small number of data available in preclinical experiment. Secondly, non linear methods may fail to converge with sparse data. Hence, there are some limitations in practice that prevent its use in preclinical PK studies.

In this paper, we propose a Bayesian nonparametric model based on P-splines. Simulations show that the proposed method provides better PK parameters estimations than the interpolation method, both in terms of bias and precision. We extend the basic model to a hierarchical one that treats the case where we have concentrations from different subjects. We are then able to get individual PK parameter estimations. Finally, with Bayesian methods, we can get easily some uncertainty measures by obtaining crediblity sets for each PK parameter.

1. INTRODUCTION

In drug discovery experiments, pharmacokinetic (PK) studies are designed to assess the systemic exposure of animals to a compound under investigation. This kind of studies attempts to analyse the evolution of drug concentration in plasma over time i.e. the PK profile. To do so, blood samples are collected at several time points after drug administration and some PK parameters are estimated in order to summarize the complete drug's kinetic profile. The usual parameters are: area under the curve (AUC), maximal concentration (C_{max}), time at which the maximal concentration occurs (t_{max}) and half life time ($t_{1/2}$). These estimated PK parameters are notably used for drug screening to rapidly take the decision of droping a compound or of keeping on working with promising ones.

There are several types of design in animals PK studies. When working with large animals, a series of blood samples may be taken from each individual such that the whole PK profile can be characterized on each animal. In this case, we talk about complete design. For small animals, due to ethical considerations, only a limited volume of blood can be collected from each subject. A classical approach is to reduce the sampling frequency per animal and to collect blood from different blocks of animals across timepoints [1]. The design is then said to be incomplete. In the extreme case where we only have one observation per animal, we talk about destructive design [2]. The scope of this work includes both complete and incomplete designs.

Several methods exist to estimate PK parameters. A first method, named below the "traditional method", is based on interpolating between observed concentrations. Typically, the interpolation is linear in the ascending phase and log-linear in the descending phase although other interpolation methods have been proposed in the literature like for example splines interpolation [3]. In an incomplete design, the traditional method consists of interpolating at each time point, the means of the observed concentrations and thus, it does not allow to have individual estimations of PK parameters for each animal.

Compartmental modelling approach is an alternative. One difficulty is to choose a suitable compartmental model given the limited number of samples available in preclinical investigations. The nonlinear estimation method often fails to converge with sparse data. So, even if this method is theoritically attractive, there are practical limitations that prevent its use in screening PK studies.

In this paper, we propose a Bayesian nonparametric method that improves the

PK parameters estimation compared to the interpolation method and that has not the convergence issues of the compartmental method. The idea is to fit individual PK profile, using penalized splines. With this method, we get quick results and we do not face the problem of model choice. Furthermore, some comparisons with the traditional method will show that our approach provides better estimations of the PK parameters, i.e. smaller bias and higher precision. Finally, with Bayesian methods, one can easily obtain measures of uncertainty for the estimated PK parameters using credibility sets.

The plan of the paper is as follows. In Section 2, we present the PK experiments as well as a brief reminder of the traditional method. Our proposal is presented in Section 3. Section 4 explains how to estimate PK parameters with the Bayesian non parametric method and how to get measures of uncertainty. Section 5 gives the results of some simulations which aim is to compare the performances of our approach with those of traditional methods. Some applications on real data are shown in Section 6. We end this paper with a discussion in Section 7.

2. PK EXPERIMENT

Pharmacokinetic studies are aimed at studying the absorption, distribution, metabolism and elimination of a pharmaceutical product. To do so, blood sample are collected at multiple times after dosing in a panel of animals.

2.1. Incomplete and complete designs

In an incomplete design, animals are sampled at one of possible subsets of predefined time points. Typically, the time points are assigned to blocks of animals in the following way. The population of animals (n) is divided into G groups ($g = 1, .., G$), containing each n_g individuals. The K sampling times t_j, $j = 1, .., K$ are distributed into these groups, so that k_g samples are taken from animals in group g, at times points $\{t_i : i \in I_g \subset \{1, ..., K\}\}$ with $\#I_g = k_g$. Sampling times are different between groups. The complete design is a particular case when $G = 1$ and $\#I_g = K$, i.e. each animal is sampled at every time point. In this situation, a PK profile may be straigthforwardly estimated for each subject.

2.2. Parameter definitions

The most familiar non-compartmental PK parameters are the AUC, C_{max}, t_{max}, $t_{1/2}$. The C_{max} represents the systemic concentration and t_{max} is the time needed

to achieve C_{max}. AUC is the area under the PK curve. It measures the extend of systemic exposure. The terminal half-life time, $t_{1/2}$, is the half-life time associated with the terminal phase, which is the final log-linear portion of the concentration versus time curve, for multicompartmental pharmacokinetics.

2.3. Traditional methods

We give here a brief reminder of the most simple and popular way to estimate the different PK parameters non compartimentally [4,5]. Denote by $C(t)$ the concentration at the observation time t and assume that the concentrations are observed at K different times. To estimate C_{max}, we use :

$$\hat{C}_{max} = \max\{C(t) : t \in \{t_1, .., t_K\}\}.$$

t_{max} is estimated by :

$$\hat{t}_{max} = \arg\max\{C(t) : t \in \{t_1, .., t_K\}\}.$$

Several investigators have considered different rules to estimate the AUC [6,7,8,9]. The simplest one is the trapezoidal rule given by :

$$\widehat{\mathrm{AUC}} = \sum_{i=1}^{K-1} 0.5(C(t_{i+1}) + C(t_i))(t_{i+1} - t_i).$$

Methods that use the log-trapezoidal rule instead of the trapezoidal one on the descending portion of the curve were found to improve results with respect to accuracy while not losing out on statistical precision [7,8,9]. With this method, we estimate the AUC as :

$$\begin{aligned} \widehat{\mathrm{AUC}} &= \sum_{i=1}^{j-1} 0.5(C(t_{i+1}) + C(t_i))(t_{i+1} - t_i) \\ &+ \sum_{i=j}^{K-1} (C(t_{i+1}) - C(t_i))(t_{i+1} - t_i)/\log(C(t_i)/C(t_{i+1})), \end{aligned}$$

where t_j is the first observation time in the descending portion of the curve.

To estimate $t_{1/2}$, there exist several methods, more or less sophisticated [10, 11,12]. In this paper, we shall consider a simple one where we first estimate the slope λ given by the 2 last observations on the time-log(concentration) scale and then, we compute :

$$\hat{t}_{1/2} = log(2)/\hat{\lambda}.$$

In the case of an incomplete design, everything stays the same except that the first step consists in computing the means of the concentrations at each time

point. It means that, while we can get individual estimations of PK parameters in a complete design, only mean PK parameters estimations are available in an incomplete one. With an incomplete design, we can get uncertainty measures for the AUC. Indeed, several efforts have been made to get confidence intervals for this PK parameter [3,2,13,14,15,16]. For the other parameters, obtaining uncertainty measures is still an issue.

3. BAYESIAN P-SPLINES MODEL

In this section, we present Bayesian models to obtain a nonparametric estimation of the individual PK profiles for complete and incomplete design. These models are based on P-splines techniques.

3.1. P-splines definition

To obtain a non parametric fit to a curve, we use penalized B-splines, also named P-splines by Eilers and Marx [17]. A B-spline of degree q consists of $q+1$ polynomial pieces, each of degree q. These polynomial pieces join at q inner knots of the experimental domain. Each B-spline is positive on a domain spanned by $q+2$ knots and it is zero everywhere else. Figure (1a) presents a B-splines basis of degree 2 with 20 equidistant knots between 0 and 1.

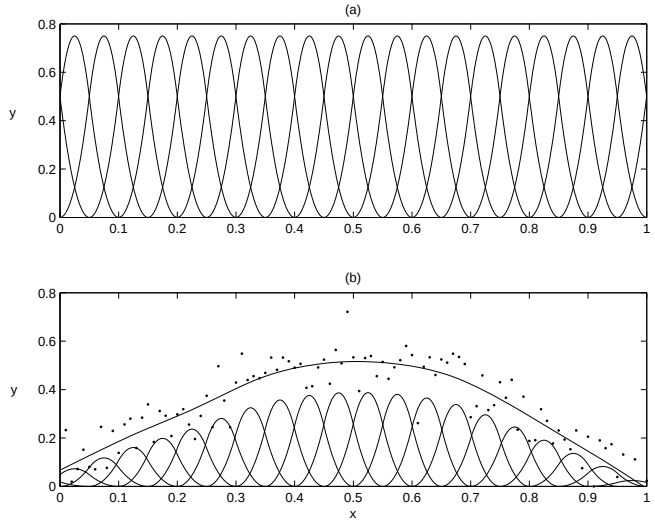


Figure 1: (a): B-splines basis of degree 2 with 20 knots. (b): Example of a fitted curve using the B-splines basis.

Let $\boldsymbol{b}(x)$ denote the B-spline basis at x for a given equidistant grid of knots. A

fitted curve $\hat{y}$ to data $\{(x_i, y_i)\}$ is a linear combination $\hat{y}(x) = \boldsymbol{b}(x)'\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is the estimated vector of B-splines coefficients. Figure(1b) presents an example of a fitted curve obtained with a linear combination of the B-splines basis presented in Figure(1a).

When m data points (x_i, y_i) are available, the least squares estimator $\hat{\boldsymbol{\theta}}$ minimizes the function :

$$S = \sum_{i=1}^{m} \{y_i - \boldsymbol{b}(x_i)'\boldsymbol{\theta}\}^2.$$

The parameters estimates are highly dependent on the number of knots and their location. The fitted curve will show more variation than is justified by the data if we let the number of knots be relatively large. To make the estimates less sensitive, Eilers and Marx propose to consider a large set of equidistant knots and to introduce a penalty term in the objective function :

$$S = \sum_{i=1}^{m} \{y_i - \boldsymbol{b}(x_i)'\boldsymbol{\theta}\}^2 + \lambda \boldsymbol{\theta}' P \boldsymbol{\theta},$$

where $P = D'D$ is the penalty matrix and D the rth-order difference matrix, yielding $\boldsymbol{\theta}' P \boldsymbol{\theta} = \sum_k (\Delta^r \theta_k)^2$ where Δ is the first-order difference operator. Thus, for $r = 2$, we have

$$D = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix}$$

By adding the term $\sum_k (\Delta^r \theta_k)^2$, we add a penalty on finite differences of the coefficients of adjacent B-splines. Parameter λ expresses the weight that we give to this penalty. If λ is high, the weight is large and we force the coefficients of adjacent B-splines to be close to each other. It will result in a very smooth fit. In contrary, if λ is small, the penalty plays a small role and the coefficients of adjacent B-splines are allowed to be highly different from each other. It will yield a wiggly fit. λ is usually selected using cross-validation or information criteria.

3.2. Basic Bayesian P-splines model

In terms of likelihood, the penalty appears as a term that we subtract from the log-likelihood $l(y; \theta)$. The penalized likelihood function has the following form :

$$l_{\text{pen}} = l(y; \boldsymbol{\theta}) - \frac{\lambda}{2} \boldsymbol{\theta}' P \boldsymbol{\theta}.$$

We get the same log-posterior in a Bayesian setting with the following model specification [18, 19]:

$$\begin{aligned}(Y_x|\boldsymbol{\theta},\tau) &\sim \mathcal{N}\left(\boldsymbol{b}(x)'\boldsymbol{\theta},\ \tau^{-1}\right)\\ p(\boldsymbol{\theta}|\tau_\lambda) &\propto \exp\left[-0.5\ \tau_\lambda\ \boldsymbol{\theta}' P\boldsymbol{\theta}\right].\end{aligned}$$

τ_λ is the roughness penalty parameter and plays the same role as λ in the frequentist setting. The penalty from the frequentist penalized likelihood approach translates, in a Bayesian setting, into a prior distribution for the rth order differences of successive B-splines parameters, θ_j.

3.3. Prior definition

There are some hyperparameters for which we have to propose prior distributions. For τ, the conditional precision of the vector response Y_x, it is common to take a noninformative prior:

$$p(\tau) \propto \tau^{-1}$$

The prior of the roughness penalty parameter τ_λ can be conveniently chosen to be the conditional conjugate prior

$$\tau_\lambda \sim \mathcal{G}(a,\ b),$$

where $\mathcal{G}(a,\ b)$ denotes a gamma distribution with mean a/b and variance a/b^2. Lang & Brezger [18] have recommended using a large variance by setting a equal to 1 and b equal to a small quantity, or $a = b$ equal to a small quantity.

3.4. Three extentions to the basic Bayesian P-splines model

In this part, we propose some extensions of the basic Bayesian P-splines model to improve its fit to animal PK data [19].

3.4.1. Preliminary remark

In preclinical PK studies, sampling schedules are often specified to have the two last times widely separated [15]. This suggests to specify the Bayesian model for log-transformed data, $\log(1 + time)$ and $\log(1 + concentration)$. This ensures positive values for the fitted concentrations and better handles the large interval between the two last observations.

3.4.2. First extension : robust prior

Jullion and Lambert [19] have emphasized the sensitivity of the Bayesian fit to the choice of the prior for τ_λ in some specific circumstances. In Figure 2, we have fitted the model of Section 3.2, on simulated sparse PK data with several gamma priors for τ_λ corresponding to different values for a and b. We can see the influence of these hyperparameters on the fit in such a setting. When $a = 1$, b=0.001, we even get a straigth line.

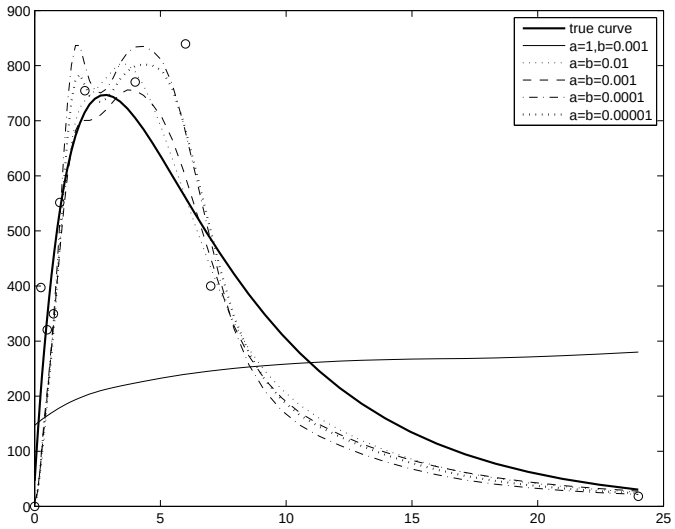


Figure 2: Influence of the prior distribution on the estimated profile.

To deal with this issue, Jullion and Lambert [19] propose to consider as prior distribution for τ_λ a weigthed sum of M Gamma distributions with different values for b. In this case, a is fixed to 1. This gives the prior :

$$\begin{aligned} (\tau_\lambda|\boldsymbol{p}) &\sim \sum_{m=1}^{M} p_m \mathcal{G}\left(a,\ b_m\right) \\ \boldsymbol{p} &\sim \mathcal{D}(\boldsymbol{u}) \end{aligned}$$

where $\{b_1, ..., b_M\}$ is a set of prespecified values. For instance, we may consider a grid of 33 values, logarithmically equally spaced between 10^{-5} to 10^3. $\mathcal{D}$ stands for the Dirichlet distribution, and $\boldsymbol{u'} = \{u_1, \ldots, u_M\}$ is a set of (small and equal) hyperprior parameters expressing our likely prior ignorance about the optimal choice for b. We use for instance the value 0.01 for each u_i.

3.4.3. Second extension : adaptive penalties

We provide even more flexibility to the model by allowing the roughness penalty to change in a progressive way along the x axis. Adapting penalties can be integrated into the previous model as follows:

$$\begin{aligned} p(\boldsymbol{\theta}|\tau_\lambda, \Lambda) &\propto \exp\left[-0.5\ \tau_\lambda\ \boldsymbol{\theta}' D' \Lambda D \boldsymbol{\theta}\right] \\ \lambda_k &\sim \mathcal{G}(\omega,\ \omega) \quad \text{when} \quad k > r+1 \ ; \ \lambda_{r+1} = 1 \end{aligned}$$

where

$$\begin{aligned} \Lambda &= \text{diag}\left(\lambda^{(r+1)}, \ldots, \lambda^{(K)}\right) \\ \lambda^{(k)} &= \prod_{l=r+1}^{k} \lambda_l \end{aligned}$$

Instead of having a single penalty τ_λ, we now have a penalty parameter $\tau_\lambda \lambda^{(k)}$ for each rth-order difference between successive components of $\boldsymbol{\theta}$. The penalty parameters are obtained sequentially by multiplying the previous one by a Gamma random variable with mean 1 and an (arbitrarily large) variance ω^{-1}. That construction yields a progressive evolution of the penalty parameters with x (see Appendix 1). For more details, we refer to [19].

3.4.4. Third extension : concavity condition

One could further constrain $\boldsymbol{\theta}$ through its prior. Indeed, a general information about a PK profile with oral dosing is its global shape. As Gibadli expresses in [20], kinetic profiles after oral administration of a drug, first show a continuous increase in drug concentration in the blood stream, then, after having reached a peak, drug concentration slowly decreases over time, following a negative exponential elimination curve. We can constraint the estimated profile to have this global shape by imposing a concavity condition on the fitted curve. In the Gibbs sampler (see Section 4), this would translate by a rejection of a $\boldsymbol{\theta}$ generated in the unconstrained specification if it does not meet the concavity condition.

In Figure 3, the thick solid line is the true PK profile. The dashed line is estimated with basic Bayesian P-splines model combined with the robust prior while the thin solid line is estimated with the model also having the adaptive penalties and the concavity condition. The fit is markedly improved.

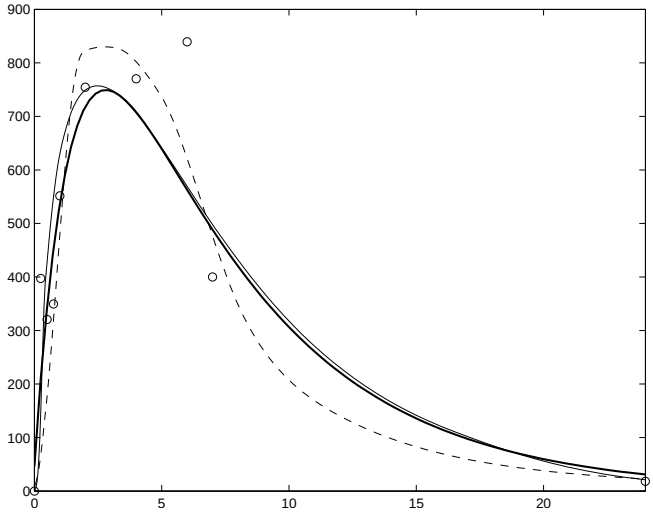


Figure 3: Profile estimation with the basic P-splines model (dashed curve) and with the extended model (solid line). The thick solid line is the true curve.

3.5. Hierarchical model

The above model allows the fit of an individual time curve. In this section, we adapt it to the case where we have several subjects such that we can get an estimate of the mean PK profile but also individual PK profiles for each subject. The following model is then suitable to incomplete designs. Considering that we have n_s subjects, we define the vector of parameters :

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_0', \quad \boldsymbol{\theta}_1', \quad .... \quad , \boldsymbol{\theta}_{n_s}']'$$

where the coefficient vector $\boldsymbol{\theta}_0$ yields the mean PK profile and the coefficient vectors $\boldsymbol{\theta}_j, j = 1, ..., n_s$ enable to correct $\boldsymbol{\theta}_0$, to get the spline parameters for subject j. We denote by K the number of B-splines in the basis. In order to force the above interpretation for θ_0, we specify for the first and last B-splines coefficients θ_{j1} and $\theta_{j,K}$, $j = 1, ..., n_s$, a normal prior distribution with mean 0 and a variance η^{-1} :

$$\begin{aligned} (\theta_{j1}|\eta) &\sim \mathcal{N}\left(0,\ \eta^{-1}\right) \quad \forall j = 1,...,n_s \\ (\theta_{jK}|\eta) &\sim \mathcal{N}\left(0,\ \eta^{-1}\right) \quad \forall j = 1,...,n_s \end{aligned}$$

By combining the previous equations with the usual smoothness prior:

$$p(\boldsymbol{\theta}_j|\tau_j) \quad \propto \quad \exp\left[-0.5\ \tau_j\ \boldsymbol{\theta_j}'D'D\boldsymbol{\theta_j}\right] \quad \forall j = 1,...,n_s,$$

we get the following prior distribution for $\boldsymbol{\theta}_j$:

$$(\boldsymbol{\theta}_j|\eta,\tau_j) \quad \sim \quad \mathcal{N}(\mathbf{0},\ \Sigma_j) \quad \forall j = 1,...,n_s$$

with

$$\Sigma_j^{-1} = \tau_j D'D + \eta \text{diag}(1,0,..,0,1)$$

The complete model specification is given in Appendix 2.

4. POSTERIOR AND PARAMETER ESTIMATION

In this section, we shall explain how to explore the posterior distribution using MCMC techniques and, from this, how to derive estimates and credibility sets for the PK parameters in the Bayesian P-splines model.

4.1. Exploring the posterior using MCMC

Markov Chain Monte Carlo (MCMC) technique is a powerful method to generate samples from posterior distributions in a Bayesian framework [21]. The Gibbs sampler [22] is a MCMC sampler by which each component is updated conditionally on the last available updates for the other components.

In our case, for each presented model, all the conditional posterior distributions can be identified. Thus, we can use the Gibbs sampler to generate random samples from the posterior distribution. We give here the conditional posterior distributions for the basic Bayesian P-splines model of Section 3.2. :

$$\begin{aligned}
(\boldsymbol{\theta}|\tau,\tau_\lambda;\boldsymbol{y}) &\sim \mathcal{N}(\tau\ \Sigma_\theta B'\boldsymbol{y},\ \Sigma_\theta) \\
(\tau|\text{other};\boldsymbol{y}) \equiv (\tau|\boldsymbol{\theta};\boldsymbol{y}) &\sim \mathcal{G}(0.5\ n,\ 0.5\ (y-B\boldsymbol{\theta})'(y-B\boldsymbol{\theta})) \\
(\tau_\lambda|\text{other};\boldsymbol{y}) \equiv (\tau_\lambda|\boldsymbol{\theta};\boldsymbol{y}) &\sim \mathcal{G}\left(a+0.5\ \rho(P),\ b+0.5\ \boldsymbol{\theta}'P\boldsymbol{\theta}\right)
\end{aligned}$$

where $\rho(P)$ is the rank of P and

$$B = [\boldsymbol{b}(x_1),\ldots,\boldsymbol{b}(x_n)]', \quad \text{and} \quad \Sigma_\theta^{-1} = \tau B'B + \tau_\lambda P$$

and 'other' generically denotes all the other parameters from the joint distribution.

The conditional posterior distributions for the extended P-splines model and for the hierarchical one are given in Appendix 3.

4.2. Estimation of PK parameters

At each iteration of the MCMC sampler, a vector $\boldsymbol{\theta}$ is generated yielding a chain of vectors $\{\boldsymbol{\theta}^{(1)},...,\boldsymbol{\theta}^{(M)}\}$ (see Section 4.1). For each $\boldsymbol{\theta}^{(m)}$, the predicted concentrations are calculated over a detailed grid of time points and PK parameters are

derived as follows. The estimate for C_{max} is the maximal predicted concentration on the detailed grid and the estimate for t_{max} is the time at which that maximum occurs. An estimate for the AUC can be computed using trapeze integration based on this grid. To obtain $t_{1/2}$, even if more sophisticated methods could be used, like for instance by using linear regression [10 11], we simply compute $t_{1/2}$ as $\log(2)/\lambda$ where λ is the slope formed by the predicted values at the two last time points.

PK parameters are so obtained for each generated $\boldsymbol{\theta}^{(m)}$, yielding a MCMC sample of size M for each PK parameter. The posterior median of these chains is used to estimate the PK parameters and 95% credibility sets are obtained by taking the 2.5 and 97.5 percentiles of the chains.

5. SIMULATIONS

In this part, we perform some simulations to compare the performances of the extended Bayesian P-splines model and of the traditional method for complete and incomplete designs.

For the complete design, we consider that we have one animal sampled at the following time points: 0, 0.25h, 0.5h, 0.75h, 1h, 2h, 4h, 6h, 7h, 24h, post dosing. The choice of the design is important when using traditional methods. Indeed, the estimated values for C_{max} and t_{max} highly depend on the choice of the observed time points.

At each time point t_i, we generate concentration y_i using :

$$y(t_i) = \mu(t_i)(1 + \sigma\epsilon_i) \quad (1)$$

$$\mu(t_i) = -1555\exp[-\exp(-0.45)t_i] + \{1600\exp[-\exp(-1.8)t_i]\}, \quad (2)$$

where $\epsilon \sim \mathcal{N}(0,\ 1)$. We consider 3 different values for σ : $\sigma = 0.1, 0.3$ and 0.9 which correspond to a low, medium and high level of noise in the data. We have generated 1000 simulations in each case.

For the incomplete case, we consider a design with two animals sampled at times (0.25h, 1h, 5h), two at times (0.5h, 2h, 7h) and two at times (0.75h, 3h, 24h). We use the same PK model as the one used in the complete design but with an extra inter-animal variability : we generate a different curve $\mu_j(t), (j = 1, ..., 6)$ for each subject by adding to each of the numerical values in Equation (2), a normal perturbation with standard deviation equal to 80,0.2,40,0.1 respectively. For each of the generated curves, we generate an observation at the three time points selected by the design as done previously.

Figures 4 and 5 summarize the results of the simulations. It gives, for each PK parameter, the boxplot of the relative bias for the incomplete and complete designs (for 3 different values of σ in the latter case); results based on the Bayesian model are given on the left part and the ones obtained with the traditional method are shown on the right.

For AUC, the relative bias is always smaller with the Bayesian method. The precision of this method is larger for the complete design and a little bit smaller for the incomplete design. For t_{max}, the Bayesian method provides better results, both in terms of bias and precision. For C_{max}, the precision is always larger with the Bayesian method. The relative bias of the Bayesian method is markedly smaller when the level of noise is high ($\sigma = 0.3$ and $\sigma = 0.9$). Finally, the relative bias for $t_{1/2}$ (see Figure 4 and a zoomed version in Figure 5) is comparable for the 2 methods but the precision is larger with the Bayesian one.

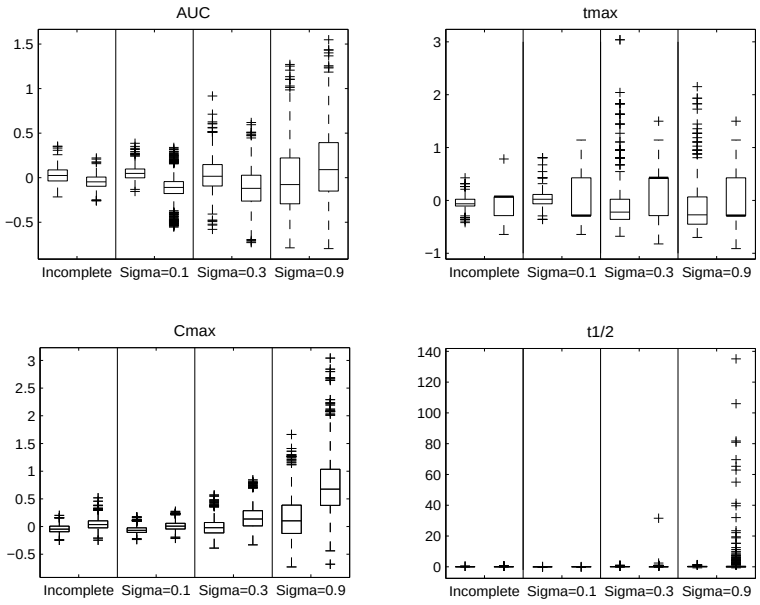


Figure 4: Simulation results : boxplots of the relative bias computed for each PK parameter, in the case of incomplete and complete designs (with 3 values for σ in the latter case). Boxplots for the Bayesian (resp. traditional) method are given on the left (resp. right) side of each case.

In the incomplete design case, we have the possibility to fit a separate PK profile for each subject with the hierarchical Bayesian model. To evaluate the quality of these estimations, we have reported in Figure 6, the boxplots of the relative bias, computed for each subject at each simulation. The model provides good estimates

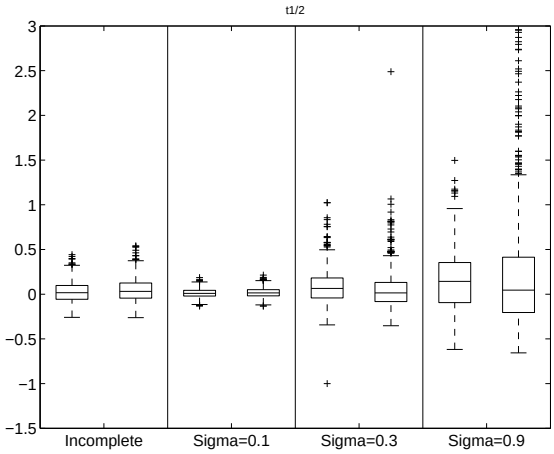


Figure 5: Simulation results : boxplots of the relative bias computed for $t_{1/2}$, in the case of incomplete and complete designs (with 3 values for σ in the latter case). Boxplots for the Bayesian (resp. traditional) method are given on the left (resp. right) side of each case. The scale has been truncated to [-1.5,3].

of individual PK parameters.

6. APPLICATION

6.1. Complete design

We present an application of the Bayesian model (see Section 3.4) on real data observed on a single subject (see Table 1).

Figure 7 shows the estimated PK profile with 95% crediblity sets for it. To obtain these curves, we estimate at each MCMC generation, the concentrations over a detailed grid of time. Then, we take the quantiles 2.5%, 50% and 97.5% of the MCMC chain of the concentration estimated at each time point. Table 2 reports the estimations of each PK parameter with 95% credibility set.

6.2. Incomplete design

Table 3 gives PK data for 6 rats observed in an incomplete design.

Figure 8 shows the fitted mean PK profile (thick solid line) and the individual PK profiles. The estimated PK parameters are given in Table 4 with 95% credibility sets. Thanks to the hierarchical model, we can get some information on the inter-animal variability.

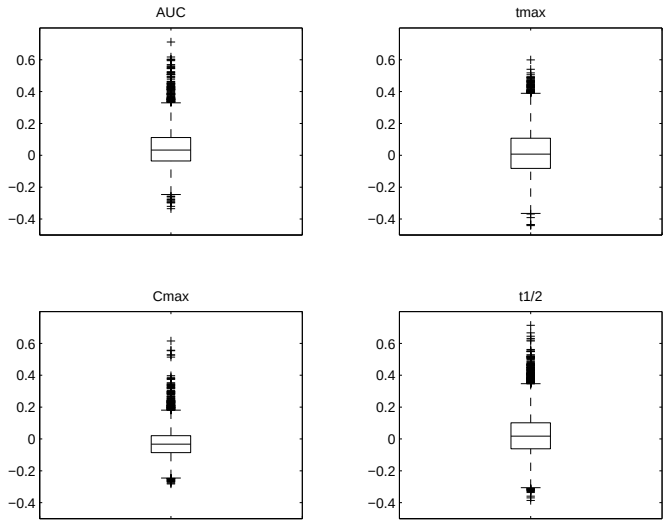


Figure 6: Simulation results : individual fits

7. DISCUSSION

We have presented a Bayesian method based on P-plines to estimate the PK profile and derive PK parameters. We have shown that the proposed Bayesian P-splines model is superior to the traditional non-compartmental methods to estimate PK parameters in complete and incomplete sampling designs. Individual predictions are also available in incomplete sparse designs.

Like the traditional non-compartmental method, the Bayesian estimation is fast to compute and does not require the assumptions of a compartmental model. The advantages of the Bayesian method over the traditional one are shown with the simulations where we get more accurate and more precise estimates for the PK parameters, both in the complete and incomplete designs. Furthermore, as we work in a Bayesian setting, we can get uncertainty measures through the credibility sets obtained using MCMC.

The presented hierarchical model offers the advantage to fit an individual profile for each subject even when the design is incomplete. For our datasets, we do not notice any influence of the choice of the hyperparameters in the prior distribution for η (see Appendix 3.2) on the fitted curves. However, the sensitivity of the results to such a choice has been reported with some hierarchical models [23]. If necessary, a mixture prior can be used (as with τ_λ in Section 3.4.2).

The hierarchical model can be extended to analyse repeated IV/oral dosing, dose proportionality studies, or to perform group comparisons for instance.

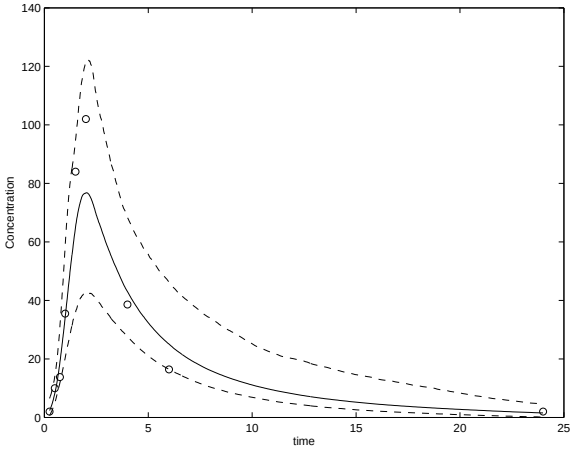


Figure 7: Estimated PK profile with credibility sets.

We have not disputed the added value of parametric models such as non linear mixed effect models to estimate PK profiles. These methods are performant provided that sufficient data are available to validate the underlying compartmental model and that the estimation algorithm converges in this nonlinear context. The Bayesian method proposed in this paper is a good alternative in drug discovery when one or both of these 2 conditions are not met.

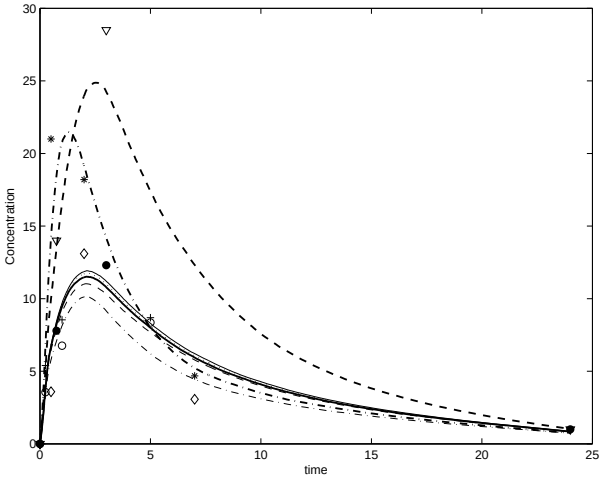


Figure 8: Application to real data for the incomplete case. The thick solid line is the mean profile, the thin dashed one is the profile for rat1 (circles), the thin dotted one for rat2 (plus), the thin dashed dotted one for rat3 (diamonds), the thick dashed-dotted one for rat4 (stars), the thick dashed one for rat 5 (triangles) and the thin solid one for rat 6 (full circles).

REFERENCES

1. Vandenhende, F., Comblain, M. and et al., M.-H. D. 'Construction of an optimal destructive sampling design for noncompartmental auc estimation', *Journal of Pharmacokinetics and Biopharmaceutics*, **27**, 191–212 (1999).

2. Nedelman, J. and Jia, X. 'An extension of Satterthwaite's approximation applied to pharmacokinetics', *Journal of Biopharmaceutical Statistics*, **8**, 317–328 (1998).

3. Bailer, A. 'Testing for the equality of area under the curves when using destructive measurement techniques', *Journal of Pharmacokinetics and Biopharmaceutics*, **16**, 303–309 (1988).

4. Westlake, W. 'Use of statistical methods in evaluation of in vivo performance of dosage forms', *Journal of Pharmaceutical Sciences*, **62**, 1579–1589 (1973).

5. Cochetto, D., Wargin, W. and Crow, J. 'Pitfalls and valid approaches to pharmacokinetic analysis of mean concentration data following intravenous administration', *Journal of Pharmacokinetics and biopharmaceutics*, **8**, 539–552 (1980).

6. Yeh, K. and Small, R. 'Pharmacokinetic evaluation of stable piecewise cubic polynomials as numerical integration functions', *Journal of Pharmacokinetics and Biopharmaceutics*, **17**, 721–739 (1989).

7. Piegorsch, A. B. W. 'Estimating integrals using quadrature methods with an application in pharmacokinetics', *Biometrics*, **46**, 1201–1211 (1990).

8. Purves, R. 'Bias and variance of extrapolated tails for area-under-the-curve (AUC) and area-under-the-moment-curve (aumc)', *Journal of Pharmacokinetics and Biopharmaceutics*, **20**, 211–226 (1992).

9. Yu, Z. and Tse, F. 'Pharmacokinetic evaluation of stable piecewise cubic polynomials as numerical integration functions', *Journal of Pharmacokinetics and Biopharmaceutics*, **17**, 721–739 (1989).

10. Lee, M., Poon, W.-Y. and Kingdon, H. 'A two-phase linear regression model for biologic half-life data', *Journal of Laboratory and Clinical Medicine*, **115**, 745–748 (1990).

11. Wright, J. and Boddy, A. 'All half-lifes are wrong but some half-lifes are useful', *Clinical Pharmacokinetics*, **40**, 237–244 (2001).

12. Kowalski, K. 'An algorithm for estimating the terminal half-life in pharmacokinetic studies', *Computer Methods and Programs in Biomedicine*, **42**, 119–126 (1994).

13. Yuan, J. 'Estimation of variance for auc in animal studies', *Journal of Pharmaceutical Sciences*, **82**, 761–763 (1993).

14. Tang-Liu, D.-S. and Burke, P. 'The effect of azone on ocular lebobunolol absorption: calculating the area under the curve and its standard error using tissue samplint compartments', *Pharmarceutical Research*, **5**, 238–241 (1988).

15. Nedelman, J. and Gibiansky, E. 'The variance of a better AUC estimator for sparse , destructive sampling in toxicokinetcs', *Journal of Pharmaceutical Sciences*, **85**, 884–886 (1996).

16. Yeh, C. 'Estimation and significant tests of area under the curve derived from incomplete blood sampling', *American Statistical Association Proceedings of the Biopharmaceutical Section*, pp. 74–81 (1990).

17. Eilers, P. H. C. and Marx, B. D. 'Flexible smoothing with B-splines and penalties (with discussion)', *Statistical Science*, **11**, 89–121 (1996).

18. Lang, S. and Brezger, A. 'Bayesian P-splines', *Journal of Computational and Graphical Statistics*, **13**, 183–212 (2004).

19. Jullion, A. and Lambert, P. 'Robust specification of the roughness penalty prior distribution in spatially adaptative bayesian p-splines models', *Journal of Computational Statistics and Data Analysis* (2006). (in press).

20. Gibaldi, M. and Perrier, D. *Pharmacokinetics*, New-York : Marcel Dekker, 1982.

21. Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. *Markov chain Monte Carlo in Practice*, Chapman and Hall/CRC, 1996.

22. Casella, G. and George, E. 'Explaining the gibbs sampler', *The American Statistician*, **46**, 167–174 (1992).

23. Antoniadis, A., Grégoire, G. and McKeague, I. W. 'Bayesian estimation in single-index models', *Statistica Sinica*, **14**, 1147–1164 (2004).

Times (in *hours*)	0	0.25	0.5	0.75	1	1.5	2	4	6	24
Concentrations (in mg/l)	0	2	10	13	35	84	102	38	16	2

Table 1: Data observed on a single subject

PK parameter	Median	95% Credibility set
t_{max}	2.03	[1.36,2.86]
C_{max}	78.57	[45.60,123.63]
AUC	399.08	[274.04,650.08]
$t_{1/2}$	4.21	[2.43,6.40]

Table 2: Estimated PK parameters with credibility sets

Time	0	0.25	0.5	0.75	1	2	3	5	7	24
Rat 1	0	5.4			6.76			8.69		
Rat 2	0	3.56			8.54			8.36		
Rat 3	0		3.59			13.1			3.07	
Rat 4	0		21			18.2			4.68	
Rat 5	0			14			28.5			1
Rat 6	0			7.79			12.3			1

Table 3: Observed concentrations (in mg/l) in 6 rats under an incomplete design

	t_{max}	C_{max}	AUC	$t_{1/2}$
Global	2.15 [1.03;3.40]	11.84 [8.23;18.29]	130.52 [101.33;182.91]	5.96 [3.76;9.22]
Rat1	2.19 [1.01;3.64]	11.37 [7.41;17.99]	127.20[94.06;181.24]	6.03 [3.74;9.61]
Rat2	2.15 [1.03;3.49]	12.12 [8.18;20.24]	132.59 [100.86;186.74]	5.92[3.78;9.36]
Rat3	2.05[0.99;3.26]	10.51 [6.96;18.3]	109.65 [79.58;154.51]	6.59 [3.32;13.08]
Rat4	1.36[0.82;2.51]	22.76 [10.94;39.41	149.81 [109.38;222.99]	5.93 [2.84;13.37]
Rat5	3.18 [2.05;4.61]	25.95 [14.52;44.42]	231.14[133.99;414.81]	4.69 [3.28;7.60]
Rat6	2.43 [1.46;5.06]	12.32 [8.67;19.20]	134.62 [100.24;197.28]	5.89 [3.74;9.32]

Table 4: Estimated PK parameters

APPENDIX 1

Model specification with the 2 extensions :

$$\begin{aligned}
(Y_x|\boldsymbol{\theta},\tau) &\sim \mathcal{N}\left(\boldsymbol{b}(x)'\boldsymbol{\theta},\ \tau^{-1}\right)\\
p(\tau) &\propto \tau^{-1}\\
p(\boldsymbol{\theta}|\tau_\lambda,\Lambda) &\propto \exp\left[-0.5\ \tau_\lambda\ \boldsymbol{\theta}'D'\Lambda D\boldsymbol{\theta}\right]\\
\lambda_k &\sim \mathcal{G}\left(\omega,\ \omega\right)\ \text{ when }\ k>r+1\ \ ;\ \ \lambda_{r+1}=1\\
(\tau_\lambda|\boldsymbol{p}) &\sim \sum_{m=1}^{M} p_m\mathcal{G}\left(a,\ b_m\right)\\
\boldsymbol{p} &\sim \mathcal{D}(\boldsymbol{u})
\end{aligned}$$

APPENDIX 2

The model specification for the hierarchical model is the following :

$$\begin{aligned}
(\boldsymbol{Y}|\boldsymbol{\theta},\tau) &\sim \mathcal{N}\left(\boldsymbol{M\theta},\ \tau^{-1}\boldsymbol{I}_n\right)\\
p(\boldsymbol{\theta}_0|\tau_0,\Lambda) &\propto \exp\left[-0.5\ \tau_0\ \boldsymbol{\theta_0}'D'\Lambda D\boldsymbol{\theta_0}\right]\\
(\boldsymbol{\theta_j}|\tau_n,\tau_\lambda) &\sim \mathcal{N}\left(\boldsymbol{0},\ \Sigma_j\right)\ \ \forall j=1,...,n_s\\
\lambda_k &\sim \mathcal{G}\left(\omega,\ \omega\right)\ \text{ when }\ k>r+1\ \ ;\ \ \lambda_{r+1}=1\\
p(\tau) &\propto \tau^{-1}\\
(\tau_0|\boldsymbol{p}_0) &\sim \sum_{m=1}^{M} p_{0,m}\mathcal{G}\left(a,\ b_m\right)\\
\boldsymbol{p}_0 &\sim \mathcal{D}(\boldsymbol{u})\\
(\tau_j|\boldsymbol{p}_j) &\sim \sum_{m=1}^{M} p_{j,m}\mathcal{G}\left(a,\ b_m\right)\ \ \forall j=1,...,n_s\\
\boldsymbol{p}_j &\sim \mathcal{D}(\boldsymbol{u})\ \ \forall j=1,...,n_s\\
\tau_n &\sim \mathcal{G}\left(a_n,\ b_n\right)
\end{aligned}$$

with

$$\Sigma_j^{-1}=\tau_j D'D+\tau_n\text{diag}(1,0,..,0,1)\ \ \forall j=1,...,n_s$$

We give here an example of matrix M for 6 subjects with 4 observations per subject. B is the B-splines basis for the 24 observations with 9 knots. The matrix

M is given by :

$$\boldsymbol{M} = \begin{bmatrix}
B(1,:) & B(1,:) & 0 & 0 & 0 & 0 & 0 \\
B(2,:) & B(2,:) & 0 & 0 & 0 & 0 & 0 \\
B(3,:) & B(3,:) & 0 & 0 & 0 & 0 & 0 \\
B(4,:) & B(4,:) & 0 & 0 & 0 & 0 & 0 \\
B(5,:) & 0 & B(5,:) & 0 & 0 & 0 & 0 \\
B(6,:) & 0 & B(6,:) & 0 & 0 & 0 & 0 \\
B(7,:) & 0 & B(7,:) & 0 & 0 & 0 & 0 \\
B(8,:) & 0 & B(8,:) & 0 & 0 & 0 & 0 \\
B(9,:) & 0 & 0 & B(9,:) & 0 & 0 & 0 \\
B(10,:) & 0 & 0 & B(10,:) & 0 & 0 & 0 \\
B(11,:) & 0 & 0 & B(11,:) & 0 & 0 & 0 \\
B(12,:) & 0 & 0 & B(12,:) & 0 & 0 & 0 \\
B(13,:) & 0 & 0 & 0 & B(13,:) & 0 & 0 \\
B(14,:) & 0 & 0 & 0 & B(14,:) & 0 & 0 \\
B(15,:) & 0 & 0 & 0 & B(15,:) & 0 & 0 \\
B(16,:) & 0 & 0 & 0 & B(16,:) & 0 & 0 \\
B(17,:) & 0 & 0 & 0 & 0 & B(17,:) & 0 \\
B(18,:) & 0 & 0 & 0 & 0 & B(18,:) & 0 \\
B(19,:) & 0 & 0 & 0 & 0 & B(19,:) & 0 \\
B(20,:) & 0 & 0 & 0 & 0 & B(20,:) & 0 \\
B(21,:) & 0 & 0 & 0 & 0 & & B(21,:) \\
B(22,:) & 0 & 0 & 0 & 0 & 0 & B(22,:) \\
B(23,:) & 0 & 0 & 0 & 0 & 0 & B(23,:) \\
B(24,:) & 0 & 0 & 0 & 0 & 0 & B(24,:)
\end{bmatrix}$$

We summarize the matrix M with this notation :

$$\boldsymbol{M} = [\boldsymbol{M}_0, \boldsymbol{M}_1, \boldsymbol{M}_2, ..., \boldsymbol{M}_{n_s}].$$

APPENDIX 3

1. EXTENDED P-SPLINES MODEL

The conditional posterior distributions are:

$$\begin{aligned}
(\boldsymbol{\theta}|\text{other}; \boldsymbol{y}) \equiv (\boldsymbol{\theta}|\tau, \tau_\lambda \boldsymbol{\lambda}; \boldsymbol{y}) \quad &\sim \quad \mathcal{N}\left(\tau\ \Sigma_\theta B' \boldsymbol{y},\ \Sigma_{\theta_g}\right) \\
(\tau|\text{other}; \boldsymbol{y}) \equiv (\tau|\boldsymbol{\theta}; \boldsymbol{y}) \quad &\sim \quad \mathcal{G}\left(0.5\ n,\ 0.5\ (y - B\boldsymbol{\theta})'(y - B\boldsymbol{\theta})\right)
\end{aligned}$$

$$
\begin{aligned}
(\lambda_l|\text{other};\boldsymbol{y}) \equiv (\lambda_l|\boldsymbol{\theta},\tau_\lambda,\boldsymbol{\lambda}_{-l};\boldsymbol{y}) &\overset{l>r+1}{\sim} \mathcal{G}\left(\omega+\frac{K-l+1}{2},\ \omega+\frac{\tau_\lambda}{2}\sum_{k=l}^{K}\frac{\lambda^{(k)}}{\lambda_l}(\Delta^r\theta_k)^2\right)\\
(\tau_\lambda|\text{other};\boldsymbol{y}) \equiv (\tau_\lambda|\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{p};\boldsymbol{y}) &\sim \sum_{m=1}^{M} p_m\mathcal{G}\left(a+0.5\ \rho(P),\ b_m+0.5\ \boldsymbol{\theta}'D'\Lambda D\boldsymbol{\theta}\right)\\
(\boldsymbol{p}|\text{other};\boldsymbol{y}) \equiv (\boldsymbol{p}|\tau_\lambda;\boldsymbol{y}) &\propto \sum_{m=1}^{M}\frac{c_m}{\sum_{k=1}^{M}c_k}\mathcal{D}(u_1,...,u_m+1,...u_M)
\end{aligned}
$$

where

$$
\begin{aligned}
c_m &= \exp(-\tau_\lambda b_m)b_m^a\frac{\sum_{k=1}^{M}u_k}{u_m}\\
\Sigma_{\theta_g}^{-1} &= \tau B'B+\tau_\lambda D'\Lambda D
\end{aligned}
$$

2. HIERARCHICAL MODEL

The conditional distributions are :

$$
\begin{aligned}
(\tau|\boldsymbol{\theta};\boldsymbol{y}) &\sim \mathcal{G}\left(0.5\ n,\ 0.5\ (\boldsymbol{y}-\boldsymbol{M\theta})'(\boldsymbol{y}-\boldsymbol{M\theta})\right)\\
(\tau_0|\boldsymbol{\theta}_0,\boldsymbol{p_0};\boldsymbol{y}) &\sim \sum_{m=1}^{M}p_{0,m}\mathcal{G}\left(a+0.5\ \rho(P),\ b_m+0.5\ \boldsymbol{\theta}_0'D'\Lambda D\boldsymbol{\theta}_0\right)\\
(\boldsymbol{p_0}|\tau_0;\boldsymbol{y}) &\propto \sum_{m=1}^{M}\frac{c_{0,m}}{\sum_{j=1}^{M}c_{0,j}}\mathcal{D}(u_1,...,u_m+1,...,u_M)\\
(\tau_j|\boldsymbol{\theta}_j,\boldsymbol{p_j};\boldsymbol{y}) &\sim \sum_{m=1}^{M}p_{j,m}\mathcal{G}\left(a+0.5\ \rho(P),\ b_m+0.5\ \boldsymbol{\theta}_j'D'D\boldsymbol{\theta}_j\right)\quad\forall j=1,...,n_s\\
(\boldsymbol{p_j}|\tau_j;\boldsymbol{y}) &\propto \sum_{m=1}^{M}\frac{c_{j,m}}{\sum_{k=1}^{M}c_{j,k}}\mathcal{D}(u_1,...,u_m+1,...,u_M)\\
(\boldsymbol{\theta}_0|\tau,\tau_0,\theta_j;\boldsymbol{y}) &\sim \mathcal{N}\left(\boldsymbol{\mu}_0,\ \boldsymbol{\Sigma_0}\right)\\
(\boldsymbol{\theta}_j|\tau,\tau_j,\theta_g;\boldsymbol{y}) &\sim \mathcal{N}\left(\boldsymbol{\mu}_j,\ \boldsymbol{\Sigma}_{p,j}\right)\quad\forall j=1,...,n_s\\
(\tau_n|\theta_{j1},\theta_{jK},\boldsymbol{y}) &\sim \mathcal{G}\left(a_n+n_s,\ b_n+0.5\sum_j(\theta_{j1}^2+\theta_{jK}^2)\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\mu_0} &= \tau\Sigma_0B'W\\
\boldsymbol{\mu_j} &= \tau\Sigma_{p,j}B'W_j\\
\Sigma_0^{-1} &= \tau B'B+\tau_0D'\Lambda D\\
\Sigma_{p,j}^{-1} &= \tau B'B+\Sigma_j\quad\forall j=1,...,n_s\\
W &= \boldsymbol{y}-\sum_{i=1}^{n_s}M_i\boldsymbol{\theta_i}\\
W_j &= \boldsymbol{y}-M_0\boldsymbol{\theta_0}-\sum_{i=1,i\neq j}^{n_s}M_i\boldsymbol{\theta_i}\quad\forall j=1,...,n_s
\end{aligned}
$$

$$
\begin{aligned}
c_{j,m} &= \exp(-\tau_j b_m) b_m^a \frac{\sum_{k=1}^{M} u_k}{u_m} \\
c_{0,m} &= \exp(-\tau_0 b_m) b_m^a \frac{\sum_{k=1}^{M} u_k}{u_m}
\end{aligned}
$$