INSTITUT DE
STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION
PAPER

0715

A NON-PARAMETRIC BAYESIAN METHOD TO SMOOTH PET TIME-ACTIVITY-CURVES

JULLION, A., LAMBERT, Ph. and F. VANDENHENDE

A non-parametric Bayesian method to smooth PET Time-Activity-Curves

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

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

[2] *Méthodes quantitatives en sciences sociales, Université de Liège, Belgium*

[3]*ClinBAY, Belgium*

July 30, 2007

Abstract

Positron Emission Tomography (PET) is an imaging technique in which a radionuclide is introduced into a molecule of potential biological relevance (to form what is called a tracer) and administered to a patient. The regional evolution of the uptake of the tracer over time is called a Time-Activity-Curve (TAC) and is used to derive some clinical measures that give information about the process under study. One of these measures is the Distribution Volume (DV) which can be estimated by several methods, notably the Graphical Analysis Method (GA). It has been shown that

*Correspondence to: Philippe Lambert, Université de Liège, Méthodes quantitatives en sciences sociales, boulevard du Rectorat, 7 4000 Liège 1 (Belgium). E-mail: p.lambert@ulg.ac.be Phone: +32 4 366.59.90 Fax: +32 4 366.47.51

using GA method on noisy TAC leads to a systematic underestimation of the Distribution Volume (Hsu *et al.*, 1997; Slifstein and Laruelle, 1999).

We propose a method that allows to smooth the Time-Activity-Curves in a non-parametric way by using Bayesian P-splines. This method may be used in all the cases, whatever the compartmental model that might underly the observed data. Simulations have shown that this method gives an unbiased estimation of the true TAC, whatever the level of noise.

We show that smoothing TAC with the non-parametric method before computing the Distribution Volume allows to reduce considerably the bias. Logan *et al.* (2001) proposes to smooth data before computing the Distribution Volume by using the Generalized Linear Least Squares (GLLS) method if a one-tissue compartment model is considered and by applying the GLLS to the data in two parts for a two-tissues compartment model : one set of parameters is estimated from times 0 to T_1 and a second set from T_1 to the end time where T_1 has to be chosen from data. This method provides good results but, if we are not sure about the compartmental model suitable for the data or, if we want to avoid the choice of T_1, using the Bayesian nonparameteric model is a good alternative.

1 Introduction

Positron Emission Tomography (PET) is an imaging technique that allows the diagnosis of many diseases, particularly cancer and mental diseases. During a PET study, a radionuclide is synthetically introduced into a molecule of potential biological relevance and administered to a patient. The subsequent uptake of the radiotracer is measured over time and used to obtain information about the physiological process of interest. The evolution of the radiotracer uptake with time in a given region is called a Time-Activity-Curve (TAC).

Some important clinical measures are derived from these TAC, using different estimation methods. For instance, the Graphical Analysis Method allows to esti-

mate the Distribution Volume; the ratio method may be used to get an estimate of the receptor occupancy (van Warde, 2000). The kinetic measures obtained with these methods may be badly estimated if the TAC are noisy. It has been notably shown that using the Graphical Analysis Method on noisy TAC leads to a systematic underestimation of the Distribution Volume (Hsu *et al.*, 1997; Slifstein and Laruelle, 1999). That's why, prior to any quantitative analysis, it may be necessary to smooth the time-course of radioactivity measured in PET scans.

In this paper, we propose a method that allows to smooth the Time-Activity-Curves in a non-parametric way by using Bayesian P-splines. This method may be used in all the cases, whatever the compartmental model that might underly the observed data: it provides a satisfactory and unbiased fit of the radioactivity uptakes.

The plan of the paper is as follows. In Section 2, we present the non-parametric Bayesian method. In Section 3, we show results of some simulations where the proposed method is applied on data generated from different compartmental models with several levels of noise. In Section 4, we present an illustration of the method in the case of Distribution Volume estimation. We end this paper with a discussion of key results.

2 Method

In this section, we present the Bayesian model that allows to smooth Time-Activity-Curves: it is based on P-splines techniques.

2.1 Basic Bayesian P-splines model

Penalized B-splines (named P-splines) are a convenient tool to obtain a non parametric fit to a curve (Eilers and Marx, 1996). The idea is to take a linear

combination of a B-splines basis noted $b(x)$. A B-spline of degree q consists of $q+1$ polynomial pieces, each of degree q, which join at q inner knots of the experimental domain. Figure (1-a) presents a B-splines basis of degree 2 with 20 equidistant knots between 0 and 1.

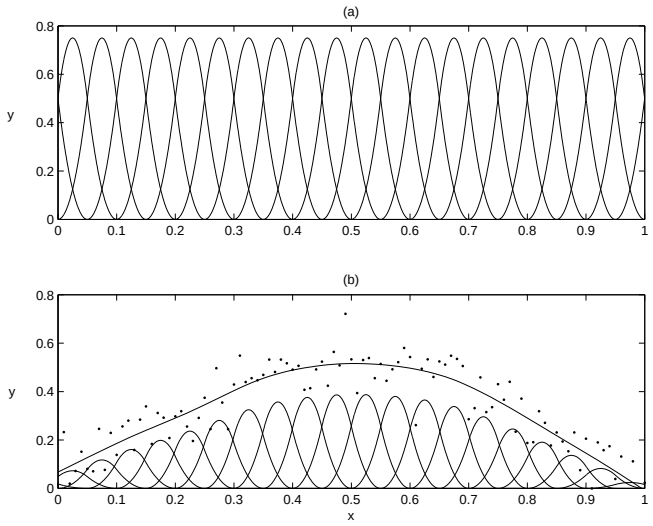

Figure 1: (a): B-splines basis of degree 2 with 20 knots. (b): Example of a fitted curve using the B-splines basis.

The fitted curve to the m data points (x_i, y_i) on Figure(1-b), $\hat{y}(x)$, is obtained by taking the linear combination $b(x)'\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is the estimated vector of B-splines coefficients obtained by minimizing $S = \sum_{i=1}^{m}\{y_i - b(x_i)'\boldsymbol{\theta}\}^2$.

The smoothness of the curve highly depends on the number of knots. To avoid this influence, Eilers and Marx (1996) propose to take a large number of equidistant knots and to add to the least squares equation, a penalty on finite differences of the coefficients of adjacent B-splines : $S = \sum_{i=1}^{m}\{y_i - b(x_i)'\boldsymbol{\theta}\}^2 + \lambda\sum_k(\Delta^r\theta_k)^2$ where Δ is the first order difference operator. In matricial terms, the penalty can be written as: $\lambda\boldsymbol{\theta}'P\boldsymbol{\theta}$ where $P = D'D$ and D is the rth order difference matrix. Parameter λ represents the weight that we give to the penalty. If λ is large, we give a large weight to the penalty, which results in a very smoothed

curve.

In a Bayesian setting, the penalty translates into a prior distribution for the rth order differences of successive B-splines parameters, $\boldsymbol{\theta}_j$, yielding for a conditional normal response Y_x :

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

where λ is the roughness penalty parameter and $\mathcal{G}\left(a,\ b\right)$ stands for a gamma distribution with mean a/b and variance a/b^2. The hyperparameters a and b are usually chosen to have a large variance conjugate prior distribution by taking for instance a equal to 1 and b equal to a small quantity (10^{-5}, say).

2.1.1 Two extentions to the basic Bayesian P-splines model

Jullion and Lambert (2007) have shown that the fit can be influenced by the choice of the hyperparameters a and b in some specific circumstances. They propose to use a more robust prior for the roughness penalty parameter by using the following reparametrisation : $a = \nu/2$ and $b = \delta\nu/2$. A prior distribution is added on δ:

$$\begin{aligned}
(\lambda|\delta) &\sim \mathcal{G}\left(0.5\ \nu,\ 0.5\ \delta\nu\right) \\
\delta &\sim \mathcal{G}\left(a_\delta,\ b_\delta\right)
\end{aligned}$$

The same authors show that more flexibility is obtained by using adaptive penalties : instead of having a global penalty λ, they propose to have a penalty parameter $\lambda \times \lambda^{(k)}$ for each rth-order difference between successive components of $\boldsymbol{\theta}$. The $\lambda^{(k)}$ are obtained sequentially to ensure a progressive evolution of the

penalty parameters with x :

$$\begin{aligned} p(\boldsymbol{\theta}|\lambda,\Lambda) &\propto \exp\left[-0.5\ \lambda \sum_{k=r+1}^{K}\left(\prod_{l=r+1}^{k}\lambda_l\right)(\Delta^r\theta_k)^2\right] \\ &= \exp\left[-0.5\ \lambda \sum_{k=r+1}^{K}\lambda^{(k)}(\Delta^r\theta_k)^2\right] \\ &= \exp\left[-0.5\ \lambda\ \boldsymbol{\theta}' D'\Lambda D\boldsymbol{\theta}\right] \\ \lambda_k &\sim \mathcal{G}\left(\omega,\ \omega\right) \quad \text{when} \quad k > r+1 \ \ ; \ \ \lambda_{r+1} = 1 \end{aligned}$$

where Λ is the matrix having the $\lambda^{(k)}$'s on its diagonal. The conditional posterior distributions are given in Appendix 1.

In Figure 2, we have generated 20 observations from $y_x = \mu_x \exp(\epsilon_x)$ with $\epsilon_x \sim \mathcal{N}\left(0,\ 0.01\right)$ and $\mu_x = \frac{A\ k_a}{k_a - k_e}[\exp(-k_e x) - \exp(-k_a x)]$ with $A = 3.74,\ k_e = 0.78,\ k_a = 50$. The thick solid line is the true curve. The thin solid line is estimated with the basic Bayesian P-splines model while the dashed line is estimated with the model having the adaptive penalties and the robust prior. The basic Bayesian P-splines model does not provide a satisfactory fit. The fit is markedly improved when using the two proposed extensions. For more details, we refer to Jullion and Lambert (2007).

3 Simulations

In this part, we shall study the quality of the fit obtained with the proposed method. Data coming from a one-tissue and a two-tissues compartment model were generated. The one-tissue compartment model is described by (Acton *et al.*, 1999):

$$\frac{dC_{tot}(t)}{dt} = K_1 C_p(t) - k_2 C_{tot}(t)$$

where $C_{tot}(t)$ is the concentration in tissue and $C_p(t)$ is a measured plasma input function, with $K_1 = 0.7\ (mLmin^{-1}mL^{-1})$ and $k_2 = 0.01\ (min^{-1})$. The equations

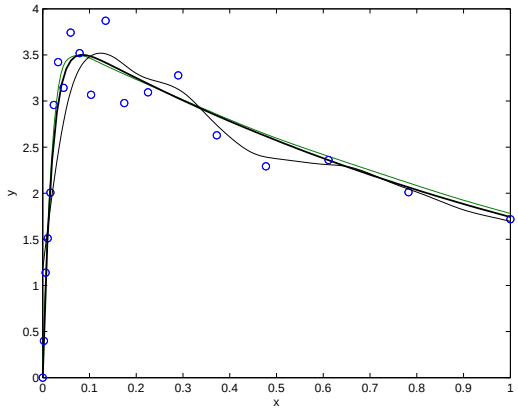

Figure 2: TAC estimation with the basic P-splines model (thin solid curve) and with the extended model (dashed line). The thick solid line is the true curve.

of the two-tissues compartment model are the following (Acton *et al.*, 1999):

$$\begin{aligned} \frac{dC_1(t)}{dt} &= K_1 C_p(t) - (k_2 + k_3) C_1(t) + k_4 C_2(t) \\ \frac{dC_2(t)}{dt} &= k_3 C_1(t) - k_4 C_2(t) \\ C_{tot}(t) &= C_1(t) + C_2(t) \end{aligned}$$

where $C_{tot}(t)$ is the measured concentration in the region of interest. $C_1(t)$ and $C_2(t)$ are the concentrations in the non-specific and specific compartments respectively with K_1 =0.7 ($mLmin^{-1}mL^{-1}$), k_2=0.05 (min^{-1}), k_3 =0.125 (min^{-1}) and k_4 = 0.3 (min^{-1}). Noise-free time activity data were generated at 27 time points for 90 minutes (with frame duration ranging from 30 seconds to 5 minutes). Then, normally distributed noise with mean 0 are added to these data. The chosen standard deviation is the same as in Ichise *et al.* (2002):

$$SD(t_i) = SF\sqrt{\frac{\exp(\lambda t_i) C_{tot}(t_i)}{\Delta t_i}} \tag{1}$$

such that

$$C(t_i) = C_{tot}(t_i) + \epsilon_i$$

with $\epsilon_i \sim N(0, SD(t_i)^2)$. SF is the scale factor that controls the level of noise, $C_{tot}(t_i)$ is the noise-free simulated radioactivity, Δt_i is the scan duration and λ is the radioisotope decay constant (fixed to $\log(2)/20$). The values of SF were selected such that the mean percent noise (f) contained in the noisy data, computed as the ratio of the mean SD to the mean tissue activity, is 4.49%, 8.99% and 15.73% for the one-tissue compartment model, and 4.52%, 9.05% and 18.1% for the two-tissues compartment model. Five hundreds noisy data sets are generated in each case.

Figure 3 and 4 show a generated dataset corresponding to the one- and two-tissues compartment model respectively, for each level of noise. Circles are generated with low noise, stars with medium noise and dots with high noise. The solid line represents the true curve.

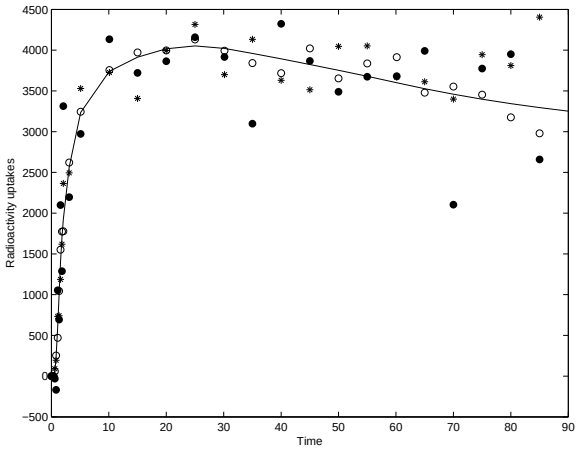

Figure 3: Example of generated datasets with the one-tissue compartment model. Circle are generated with low noise, stars with medium noise and dots with high noise. The solid line represents the true curve.

Figure 5 compares the estimated curves to the true curve for the one-tissue compartment model. The circles are the means of the 500 estimated uptakes obtained with the Bayesian method. The solid line is the true curve. The 90% credibility sets for the estimated concentration at each time-point are shown.

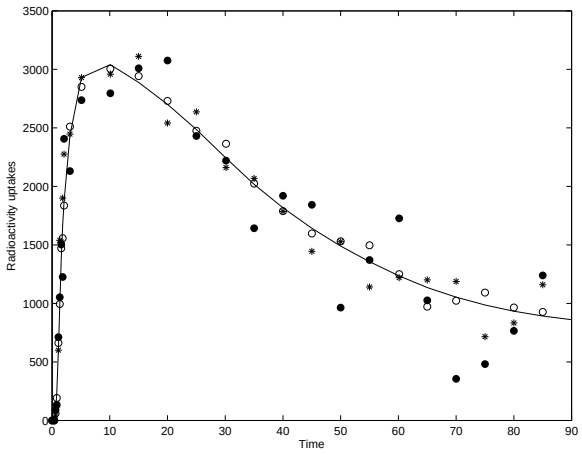

Figure 4: Example of generated datasets with the two-tissues compartment model. Circle are generated with low noise, stars with medium noise and dots with high noise. The solid line represents the true curve.

Figure 6 shows the same for the two-tissues compartment model.

We can see on Figures 5-6 that the proposed method provides virtually unbiased estimations of the true Time-Activity-Curve, whatever the underlying compartmental model and level of noise.

4 Estimation of the distribution volume (DV)

We now apply the Bayesian P-splines smoothing method for Distribution Volume (DV) estimation. Several strategies have been proposed in the literature to estimate this parameter. One commonly used method relies on compartmental kinetic models in combination with the Nonlinear least squares estimation method: it provides accurate estimates. However, due to its iterative aspects, it can be computationally intensive (Motulsky and Ransnas, 1987). Furthermore, convergence problems may arise. Linear least squares methods avoid these drawbacks by relying on some form of linearization of the compartment model equations which does not require an iterative search in the parameter space Blomqvist (1984).

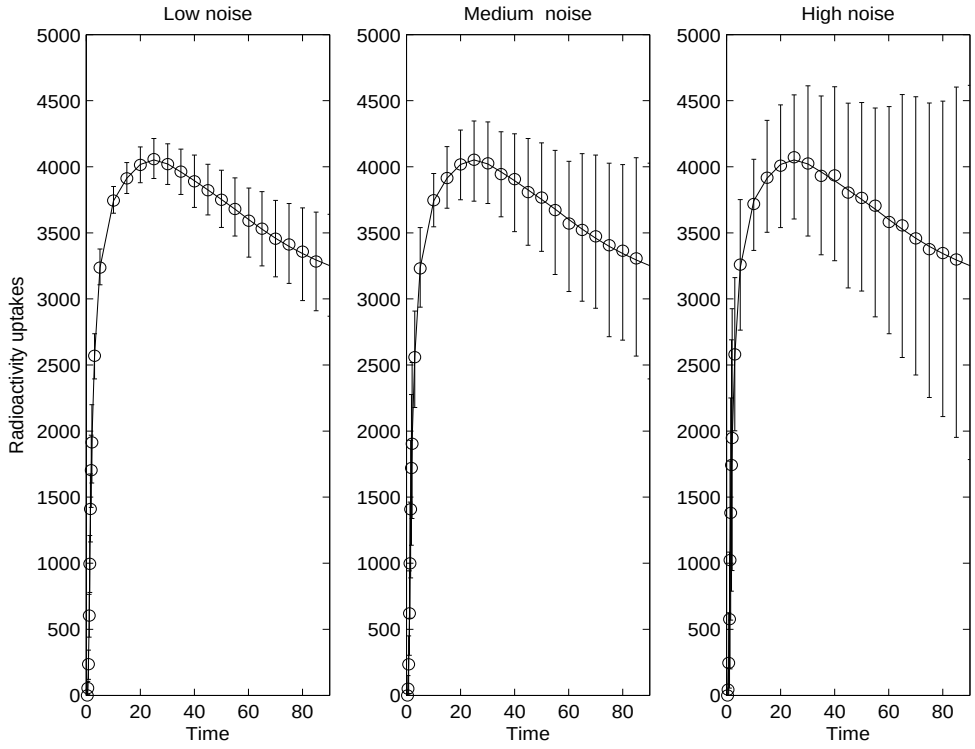

Figure 5: Results of the simulations for the one-compartment model. The true curve is the solid line. The circles are the means of the 500 estimated uptakes obtained with the Bayesian method. The 90% credibility sets for the estimated concentration at each time-point are shown

Logan *et al.* (1990) provide a simplified method, named *the Graphical Analysis Method* (GA) where the set of linear equations is associated to a linear plot. This method is simple to implement, does not face convergence problems and does not make any assumptions about the compartmental configuration of the underlying data. However, when the Time-Activity Curves are noisy, it has been shown that the slope of the linear plot, which is DV in the case of reversibly binding ligands, is underestimated (Hsu *et al.*, 1997; Slifstein and Laruelle, 1999).

Several approaches have been studied recently to reduce the bias in DV estimation in the presence of noisy data. Ichise *et al.* (2002) have proposed three

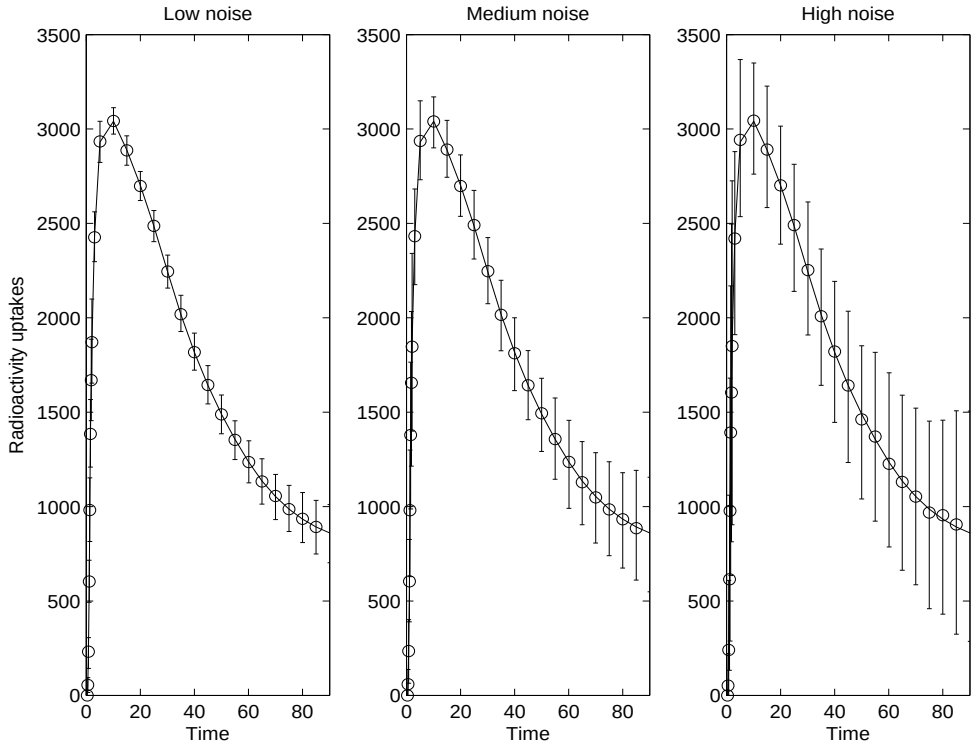

Figure 6: Results of the simulations for the two-compartments model. The true curve is the solid line. The circles are the means of the 500 estimated uptakes obtained with the Bayesian method. The 90% credibility sets for the estimated concentration at each time-point are shown

alternative noniterative linear methods to improve DV estimation. The first one is based on the total least squares estimation method while the two others rely on multilinear analyses based on mathematical arrangement of GA equations. An alternative approach is provided by Logan *et al.* (2001). Their idea is first, to smooth TAC with a parametric model and then apply the GA method to the smoothed data. If a one-tissue compartment model is considered, they apply the Generalized Linear Least Squares (GLLS) method (Feng *et al.*, 1996) to the data. For a two-tissues compartment model, they apply the GLLS to the data in two parts : one set of parameters is estimated from times 0 to T_1 and a second set

from T_1 to the end time.

We propose to estimate the Distribution Volume with the Graphical Analysis Method of Logan *et al.* (1990), but prior to this computation, we first smooth the Time-Activity-Curves with the Bayesian non-parametric method. We shall make some simulations to show that this procedure does not lead to an underestimation of the Distribution Volume.

4.1 Graphical Analysis Method

We briefly review the Graphical Analysis Method of Logan *et al.* (1990). This method, applied on PET data acquired with a reversible radiotracer, relies on the following equation :

$$\frac{\int_0^{t_i} C_1(t)dt}{C_1(t_i)} = DV \frac{\int_0^{t_i} C_p(t)dt}{C_1(t_i)} + b$$

where the t_i's are the midframe scanning times, $C_1(t)$ is the radioactivity concentration at time t in the target region of interest, $C_p(t)$ is the metabolite corrected plasma concentration at time t, DV is the Distribution Volume and b is the intercept which becomes constant when $t_i > t^*$ (where t^* is the equilibrium time).

4.2 Material and methods

We consider a two-tissues compartment model to simulate the data, using a measured plasma input function $C_p(t)$. We perform two simulations. For Simulation 1, the two-tissues compartment model parameters used are K_1 =0.0613 $(mLmin^{-1}mL^{-1})$, k_2 =0.0776 (min^{-1}), k_3=0.0734 (min^{-1}), k_4 =0.0135 (min^{-1}) and $DV = 4.56$ while for Simulation 2, $K_1 = 0.8542$ $(mLmin^{-1}mL^{-1})$, k_2 =0.0785 (min^{-1}), $k_3 = 0.0502$ (min^{-1}), $k_4 = 0.0227$ (min^{-1}) and $DV = 33.05$. Using these parameter values and the input function, noise-free time activity data are generated at 27 time points for 90 minutes (with framing time increasing from 30 seconds to 5 minutes). Then, normally distributed noise with mean 0 and stan-

dard deviation computed according to Equation (1) were added to these data (Ichise *et al.*, 2002). λ is fixed to $\log(2)/109.8$ for Simulation 1 and to $\log(2)/20.4$ for Simulation 2. Different ranges of SF values are selected : 0.35-3.00 for Simulation 1, and 0.75-7.00 for Simulation 2.

We generate 500 noisy data sets for each value of SF. The mean percent noise (f) contained in the noisy data ranges from 2.09 to 17.90% for Simulation 1 and from 2.07 to 19.34% for Simulation 2. To compute DV with the Logan method, a value for t^* must be selected : it was set to 20 min for both simulations. This t^* value was chosen graphically from preliminary GA of the noise-free data sets.

4.3 Results

Tables 1 and 2 summarize the results of the simulations. They give the median of the relative bias in the computation of DV i.e. $(DV_{estimated} - DV_{true})/DV_{true}$ and the Variation Coefficient (VC) computed as the standard deviation divided by the mean. As expected, Graphical Analysis applied to the noisy data underestimates DV as the noise increases. If we first smooth the data with the Bayesian method before computing DV, we can see that the bias is markedly reduced. However, we can note that, when the noise is more important, the method tends to have a larger VC than GA.

Figures 7 and 8 give, for the two simulations, the boxplots of the logarithm of the squared error ($SE = (DV_{true} - DV_{estimated,m})^2, m = 1, ..., M$, with M, the number of simulations). Boxplots for the Bayesian method are on the left. Those for the GA method are on the right. The mean SE (MSE) and the variability are smaller with the Bayesian method than with GA. In both simulations, the MSE tends to be smaller with the Bayesian method : it is particularly marked when noise (SF) is large.

SF		0.35	0.7	1.25	2	2.5	3
f		2.09%	4.18%	7.46%	11.93%	14.91%	17.90%
GA	Bias	-0.7%	-2.5%	-6.6%	-15.5%	-21.7%	-28%
Bayes	Bias	-0.11%	-0.15%	-0.16%	-1.25%	-1.57%	-1.96%
GA	VC	0.0173	0.0354	0.0567	0.0983	0.1067	0.1185
Bayes	VC	0.0168	0.0337	0.0596	0.0940	0.1070	0.1445

Table 1: Simulation 1: median of the Relative bias and Variation Coefficient for DV with the GA applied on noisy data and with the GA applied to data smoothed with the Bayesian method.

SF		0.75	1.5	3.5	5	7
f		2.07%	4.15%	9.67%	13.82%	19.34%
GA	Bias	-0.49%	-2.19%	-9%	-14.72%	-21.52%
Bayes	Bias	-0.47%	-0.65%	-0.66%	-0.68%	-0.9%
GA	VC	0.0145	0.0296	0.0685	0.0863	0.1133
Bayes	VC	0.0154	0.0294	0.0684	0.0994	0.1483

Table 2: Simulation 2: median of the Relative bias and Variation Coefficient for DV with the GA applied on noisy data and with the GA applied to data smoothed with the Bayesian method.

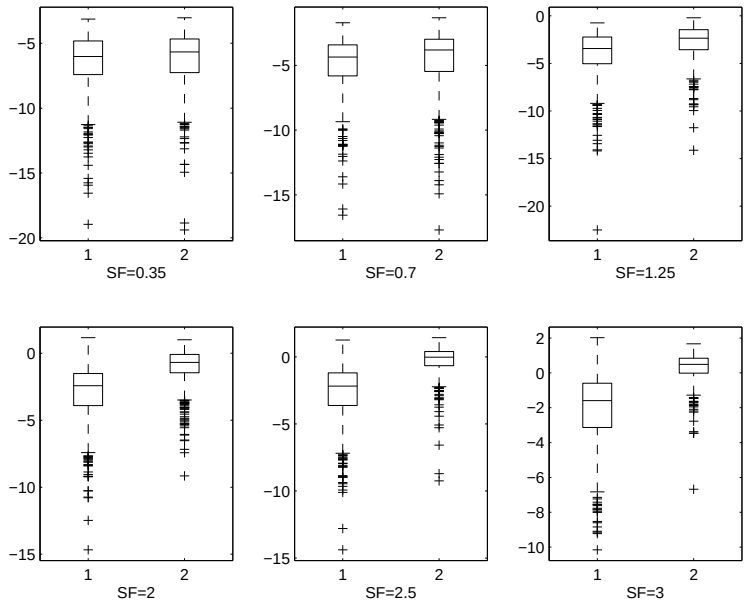

Figure 7: Boxplots of the logarithm of the squared error for Simulation 1 for the Bayesian method (on the left) and the GA method (on the right).

5 Discussion

We have proposed a method that allows to smooth Time-Activity-Curves without any prior assumptions on the behaviour of the data. Simulations have shown that this method gives an unbiased estimation of the true Time-Activity-Curve, for both one- and two-tissues compartment models, whatever the level of noise.

Smoothing TAC before computing the Distribution Volume improves the estimation. We have provided an example where smoothing noisy TAC with the non-parametric method before computing the Distribution Volume considerably reduces the bias in DV estimation. Logan already proposed to first smooth data before computing the Distribution Volume (Logan *et al.*, 2001). They advised to use the Generalized Linear Least Squares (GLLS) method if a one-tissue compartment model is considered and to apply the GLLS to the data in two parts for a two-tissues compartment model : one set of parameters is estimated from times

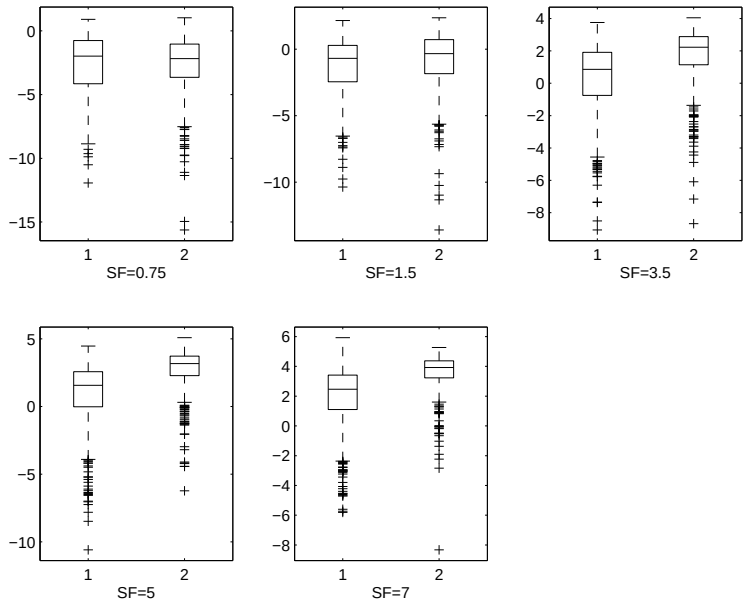

Figure 8: Boxplots of the logarithm of the squared error for Simulation 2 for the Bayesian method (on the left) and the GA method (on the right).

0 to T_1 and a second set from T_1 to the end time where T_1 has to be chosen from data. This method has been shown to provide good results but, if we are not sure about the compartmental model suitable for the data or, if we want to avoid the choice of T_1, using the Bayesian nonparameteric model is a good alternative.

References

Acton, P. D., Kushner, S. A., Kung, M. P., Mozley, P. D., Plössl, K., and Kung, H. F. (1999). Simplified reference region model for the kinetic analysis of [^{99mTc}]trodat-1 binding to dopamine transporters in nonhuman primates using single-photon emission tomography. *European Journal of Nuclear Medicine*, **26**, 518–526.

Blomqvist, G. (1984). On the construction of functional maps in positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism*, **17**, 437–447.

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

Feng, D., Huang, S.-C., Wang, Z., and Ho, D. (1996). An unbiased parametric imaging algorithm for nonuniformly sampled biomedical system parameter estimation. *IEEE Trans Med Imag*, **12**, 182–188.

Hsu, H., Alpert, N., Christian, B., Bonab, A., Morris, E., and Fishman, A. (1997). Noise properties of a graphical assay of receptor binding. *Journal of Nuclear Medicine*, **38**, 204P.

Ichise, M., Toyama, H., Innis, R. B., and Carson, R. E. (2002). Strategies to improve neuroceptor parameter estimation by linear regression analysis. *Journal of Cerebral Blood Flow and Metabolism*, **22**, 1271–1281.

Jullion, A. and Lambert, P. (2007). Robust specification of the roughness penalty prior distribution in spatially adaptative bayesian p-splines models. *Journal of Computational Statistics and Data Analysis*, **51**, 2542–2258.

Logan, J., Fowler, J. S., Volkow, N. D., Wolf, A., Dewey, S., Schyler, D., MacGregor, R., Hitzemmann, R., Bendriem, B., and Gatley, S. (1990). Graphical

analysis of reversible radioligand binding from time-activity measurements applied to [N $-^{11}$ C $-$ methyl] $-$ (.)$-$cocaine PET studies in human subjects. *Journal of Cerebral Blood Flow and Metabolism*, **10**, 740–747.

Logan, J., Fowler, J. S., Volkow, N. D., Ding, Y. S., Wang, G.-J., and Alexoff, D. L. (2001). A strategy for removing the bias in the graphical analysis method. *Journal of Cerebral Blood Flow and Metabolism*, **21**, 307–320.

Motulsky, H. J. and Ransnas, L. A. (1987). Fitting curves to data using nonlinear regression : a practical and nonmathematical review. *FASEB*, **1**, 365–374.

Slifstein, M. and Laruelle, M. (1999). Statistical bias in the logan graphical method. *Journal of Nuclear Medicine*, **40**, 35P.

van Warde, A. (2000). Measuring receptor occupancy with PET. *Current Pharmaceutical Design*, **6**, 1593–1610.

Appendix 1

The conditional posterior distributions are:

$$
\begin{aligned}
(\boldsymbol{\theta}|\tau,\lambda,\delta,\boldsymbol{\lambda};\boldsymbol{y}) &\sim \mathcal{N}\left(\tau\ \Sigma_\theta B'\boldsymbol{y},\ \Sigma_\theta\right)\\
(\tau|\text{rest};\boldsymbol{y}) \equiv (\tau|\boldsymbol{\theta};\boldsymbol{y}) &\sim \mathcal{G}\left(0.5\ n,\ 0.5\ (y-B\boldsymbol{\theta})'(y-B\boldsymbol{\theta})\right)\\
(\lambda_l|\text{rest};\boldsymbol{y}) \equiv (\lambda_l|\boldsymbol{\theta},\lambda,\boldsymbol{\lambda}_{-l};\boldsymbol{y}) &\overset{l>r+1}{\sim} \mathcal{G}\left(\omega+\frac{K-l+1}{2},\ \omega+\frac{\lambda}{2}\sum_{k=l}^{K}\frac{\lambda^{(k)}}{\lambda_l}(\Delta^r\theta_k)^2\right)\\
(\lambda|\text{rest};\boldsymbol{y}) \equiv (\lambda|\boldsymbol{\theta},\delta,\boldsymbol{\lambda};\boldsymbol{y}) &\sim \mathcal{G}\left(0.5\ \nu+0.5\ \rho(P),\ 0.5\ \delta\nu+0.5\ \boldsymbol{\theta}'D'\Lambda D\boldsymbol{\theta}\right)\\
(\delta|\tau,\lambda,\boldsymbol{\lambda};\boldsymbol{y}) &\sim \mathcal{G}\left(a_\delta+0.5\ \nu,\ b_\delta+0.5\ \nu\lambda\right)
\end{aligned}
$$

where

$$\Sigma_\theta^{-1} = \tau B'R^{-1}B + \lambda D'\Lambda D$$