Made available by Hasselt University Library in https://documentserver.uhasselt.be

Application of Penalized Splines in Analyzing Neuronal Data

Peer-reviewed author version

MARINGWA, John; FAES, Christel; GEYS, Helena; MOLENBERGHS, Geert;
Cadarso-Suarez, Carmen; Pardo-Vazquez, Jose L.; Leboran, Victor & Acuna, Carlos
(2009) Application of Penalized Splines in Analyzing Neuronal Data. In:
BIOMETRICAL JOURNAL, 51(1). p. 203-216.

DOI: 10.1002/bimj.200810501
Handle: http://hdl.handle.net/1942/9512

Application of Penalized Smoothing Splines in Analyzing Neuronal Data

John T. Maringwa *[1], **Christel Faes** [1], **Helena Geys** [2], **Geert Molenberghs** [1], **Carmen Cadarso-Suárez** [3], **José L. Pardo-Vázquez** [4], **Víctor Leborán** [4], and **Carlos Acuña** [4]

[1] Center for Statistics, Hasselt University, BE 3590 Diepenbeek, Belgium

[2] Johnson and Johnson Pharmaceutical Research and Development, Beerse, Belgium

[3] Dept of Statistics and Operations Research, University of Santiago de Compostela, Spain

[4] Departamento de Fisiologia, Facultad de Medicina, Complejo Hospitalario Universitario Universidad de Santiago de Compostela, Spain

Received 17 February 2008

* Corresponding author: e-mail: john.maringwa@uhasselt.be, Phone: +00 32 11 268285, Fax: +00 32 11 268299

Copyright line will be provided by the publisher

Summary

Neuron experiments produce high-dimensional data structures. Therefore, application of smoothing techniques in the analysis of neuronal data from electrophysiological experiments has received considerable attention of late. We investigate the use of penalized smoothing splines in the analysis of neuronal data. This is first illustrated when interested in the temporal trend of a single neuron. An approach to investigate the maximal firing rate, based on the penalized spline model is proposed. Determination of the time of maximal firing rate is based on non-linear optimization of the objective function with the corresponding confidence intervals constructed based on the first-order derivative function. To distinguish between the curves from different experimental conditions in a moment-by-moment sense, bias adjusted simulation-based simultaneous confidence bands leading to global inference in the time domain are constructed. The bands are an extension of the approach proposed by Ruppert *et al.* (2003). These methods are in a second step extended towards the analysis of a population of neurons via a marginal or population-averaged model.

Key words: Marginal model; Maximal firing rate, Neuronal data analysis; Penalized splines; Simultaneous confidence bands, Smoothing.

1 Introduction

Neurons carry information by means of electrical signals (action potentials or spikes) which are transmitted across synapses. The spike is a pulse signal of about 1 ms duration and of the same amplitude and it constitutes the relevant signal for the interactions between neurons. In electrophysiological experiments, these spikes are recorded by microelectrodes inserted in the brain as they occur in the time course, time stamps. It is assumed that the number of spikes per time unit, the spike rate, produced by single neurons is a relevant parameter for the coding in the brain. Single-unit activity is irregular, both within and across trials; hence to obtain the regularity of the response, trials are repeated several times. Furthermore, to assess that the behavior of single neuron activity is present at population level, the statistical analysis should be extended to the population of neurons with the same properties (Kass *et al* ˙, 2005).

The data considered here involve the electrical activity in 20 different neurons recorded in the ventral premotor cortex (VPM) while a monkey performs a continuous discrimination task (CD task). In this

Copyright line will be provided by the publisher

task, the monkey reports a decision, based on the comparison of the orientation of two visual stimuli shown sequentially, separated by a delay. The main determinant of the neuron's discharge was whether the second stimulus (test) was to the left or right of the first (reference). Several trials are performed, classified according to orientation of the second stimulus with reference to the first (left or right) and the degree of difficulty (easy, difficult). For each trial, within the experimental period, i.e., 1000 ms to 2500 ms, time points where electrical activity was noted are recorded. Full details of the experiment are given in Section 2. Usually, data from such electrophysiological experiments is summarized using a raster plot, displaying the complete set of spikes for each of the trials (Kass *et al.*, 2005). Also the peristimulus histogram (Gerstein and Kiang, 1960) can be used to summarize the overall activity and evolution in time by counting spikes in intervals of a certain width. Several ways to smooth instantaneous firing rates have been studied in the literature. An overview of the application of smoothing techniques in neuronal data can be found in Kass *et al.* (2003). Cadarso-Suárez *et al.* (2006) and Roca-Pardinas *et al.* (2006), for example, employ a flexible modeling technique based on the logistic Generalized Additive Model (GAM) with local linear kernel smoothers. Faes *et al.* (2007) apply a flexible method based on natural cubic splines to model synchrony in neuronal firing. Other recent techniques in this context include the Bayesian adaptive regression splines (DiMatteo *et al.*, 2001; Behseta *et al.* (2005); Behseta and Kass, 2005). Flexible regression-based techniques come out favorable since they enjoy the flexibility of capturing the temporal evolution without the restriction of parametric modeling as well as the possibility to include covariate or factor information. We revisit this aspect in Section 3 where the models discussed happen to share similar properties.

The number of spikes accumulated over the different trials can be assumed to come from an inhomogeneous Poisson counting process (Cadarso-Suárez *et al.*, 2006; Ventura *et al.*, 2002; Kass *et al.*, 2005), and our interest lies in estimating the instantaneous firing rate, denoted by $\lambda(t)$. Estimating the instantaneous firing rate is necessary, especially in our situation where one of the main research goals is to determine the time trend and the time of maximal firing rate. Capturing the temporal structure with a parametric function may

Copyright line will be provided by the publisher

prove difficult or unsatisfactory. For example, Ventura *et al.* (2002) define a piecewise parametric function to describe the mean of the intensity function. One of the problems they face is that, for some neurons, the proposed model does not conform to the observed pattern. An attractive alternative is to model the time evolution by a flexible semiparametric function estimated through use of penalized smoothing splines (Eilers and Marx, 1996; Verbyla *et al.*, 1999; Ruppert *et al.*, 2003), fitted in the mixed model framework. The representation of the penalized splines as a mixed model is very appealing, since this connection presents an opportunity for ordinary software packages for mixed models, such as, for example S-plus, R, or SAS, to be used to fit the penalized spline model, and to select the amount of smoothing automatically via maximum likelihood (ML) or restricted maximum likelihood (REML). Molenberghs and Verbeke (2005) present an example together with SAS code for analyzing an ordinal outcome in a clinical-trial setting. Other advantages are the unified framework for inference and the flexibility with which the models can be extended.

One of the main objectives of the study is to summarize certain characteristics of the time evolution of activity in the neurons from a population (all neurons) point of view. The important characteristics under consideration include time at which the maximum firing rate is observed. In addition, confidence intervals on the time of maximum firing rate are also required. We consider modeling the data from two perspectives, namely, single-neuron analysis and a population-averaged approach. Modeling of the time evolution in the two perspectives mentioned above essentially follows the same route as explained in Section 3.1.1.

The remainder of this paper is organized as follows. In Section, 2 details of the experiment are given. Section 3 focuses on the methodology applied herein. In particular, penalized splines methodology, and construction of simultaneous confidence bands for comparing curves in time are discussed. In Section 4 the methodology is applied to the study motivating this work.

Copyright line will be provided by the publisher

2 The Experiment

One monkey was trained to discriminate between different line orientations (stimuli) (Vazquez *et al.*, 2000). The stimuli (reference and test) consisted of stationary bright line segments presented on a monitor screen in front of the monkey. Reference stimuli were presented with three different orientations $(85.5^o, 90^o, 94.5^o)$. Eight test stimuli per reference were presented rotated clockwise or counter-clockwise to the reference line in steps of 1.5^o. Two bright circles were laterally displayed to the right and to the left of the center of the screen.

A trial was initiated when the monkey fixated a small line centred on the screen. Then, when the fixation line disappeared, the two stimuli, reference and test, each of 500 ms duration, appeared in sequence, separated by a fixed inter-stimulus interval (ISI, 1000 ms). At the end of the second stimulus, the subject had to make a saccadic eye movement towards one of the two circles to indicate whether the orientation of the second stimulus was clockwise (right) or counter-clockwise (left) to the reference stimulus. Monkeys were rewarded for correct discriminations (see Table 1). Once trained, extra-cellular single unit activity was recorded in the ventral premotor cortex (VPM).

The activity of some VPM neurons was modulated depending on the monkeys' decision about the orientation of the test to the reference stimuli. We were interested in (1) determining the maximum peak activity when the monkey correctly decided that the test stimuli were to the right and to the left of the reference line, and (2) comparing the neural response between behaviorally relevant conditions.

The data collected were summarized across the different trials in the form of spike counts per time unit. The period of analysis was between 1000–2500 ms. This time period occupies the last 500 ms of the ISI and the 500 ms of the comparison/decision period (2000–2500 ms), this being the relevant period for this analysis of the correct decisions to the left and to the right. Table 1 gives a summary of the events in each trial.

Table 1 ABOUT HERE

Copyright line will be provided by the publisher

3 Methods

3.1 Single Neuron Analysis

Let us introduce the methodology, first in the context of a single neuron. It should be noted that extension of the model to several neurons, in matrix notation, simply involves stacking together matrices corresponding to the different neurons. To reduce the computational burden, the time scale is subdivided into 20ms bins and each bin represented by the median time point of that bin.

3.1.1 Penalized Splines with Radial Basis

Let $y_t (t = 1, \ldots, T)$ represent the total count of activities recorded in bin t, aggregated over all the trials, x_t the median time point of bin t and $\kappa_1, \ldots, \kappa_K$ be a set of knots in the range of x_t. To flexibly model the response $\boldsymbol{Y} = (y_1, \ldots, y_T)^{'}$, consider the model,

$$h(E[\boldsymbol{Y}]) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_k \boldsymbol{u} \ , \tag{1}$$

where $h(\cdot)$ is an appropriate link function, such as the log link, and $\boldsymbol{X}$ and $\boldsymbol{Z}_k$ are design matrices of the form

$$\boldsymbol{X} = \left[1 \ x_t \ \ldots \ x_t^{m-1}\right]_{1 \le t \le T}, \quad \boldsymbol{Z}_k = \left[|x_t - \kappa_k|^{2m-1}\right]_{1 \le t \le T, \quad 1 \le k \le K}.$$

The parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{u}$ are fixed and random effects, respectively. This defines the radial basis spline function of degree $2m - 1$. When using a large number of knots, this spline function has a lot of flexibility. To overcome the problem of overfitting the data, the parameters $\boldsymbol{u}$ are restricted by assuming that (Ruppert *et al.*, 2003)

$$\text{Cov}(\boldsymbol{u}) = \sigma_u^2 (\boldsymbol{\Omega}_k)^{-1/2} (\boldsymbol{\Omega}_k^{-1/2})^T, \quad \text{with} \quad \boldsymbol{\Omega}_k = \left[\left|\kappa_k - \kappa_k^{'}\right|^{2m-1}\right]_{1 \le k, k' \le K}.$$

To fit the model using standard mixed model software, the transformation $\boldsymbol{Z} = \boldsymbol{Z}_k \boldsymbol{\Omega}_k{}^{-1/2}$ is applied, resulting in an equivalent model

$$h(E[\boldsymbol{Y}]) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u}, \quad \text{where} \quad \text{Cov}(\boldsymbol{u}) = \boldsymbol{G} = \sigma_u^2 \boldsymbol{I}. \tag{2}$$

Copyright line will be provided by the publisher

Estimation of the parameters in the model above are obtained via use of the SAS procedure GLIMMIX, which uses pseudolikelihood estimation techniques (Wolfinger and O'Connell, 1993). The knots in this case are selected as equally-spaced quantiles of the time variable (Ruppert, 2002; Ngo and Wand, 2004). Broadly speaking, the idea is to have a large number of knots and then constrain their influence to yield a smooth fit.

Since interest is in a comparison of the time trend among different experimental conditions (e.g., left versus right, and easy versus difficult), a complex model is assumed which considers different smooth functions in each of the experimental conditions. Such a model can be said to have factor by factor by curve interactions (Ruppert *et al.*, 2003). More specifically, consider the two-factor degree of difficulty ($d = 0$ if easy and $d = 1$ otherwise) combined with orientation ($z = 0$ if left and 1 otherwise), then placing this covariate information in the matrix $\mathbf{X}$, the penalized spline model of degree 3 considered herein takes the form:

$$\begin{aligned}\log(\lambda_t|\mathbf{X}) = \;& \beta_0 + \beta_1 x_t + \textstyle\sum_{k=1}^{K} b_k|x_t - \kappa_k|^3 + d\alpha_d + d\left(\sum_{k=1}^{K} v_k^d|x_t - \kappa_k|^3\right) + z\gamma_z \\ & + z\left(\textstyle\sum_{k=1}^{K} \omega_k^z|x_t - \kappa_k|^3\right) + dz\delta_{dz} + dz\left(\sum_{k=1}^{K} \phi_k^{dz}|x_t - \kappa_k|)^3\right),\end{aligned} \tag{3}$$

where b_k, v_k^d, ω_k^z, and ϕ_k^{dz} are random effects for smoothing, assumed to be N(0, σ_s^2). The parameters α_d, v_k^d correspond to orientation, γ_z, ω_k^z correspond to difficulty and δ_{dz}, ϕ_k^{dz} to the interaction between the two factors. Using appropriate design matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$, this model can be written as (2) and fitted as described in Section 3.1.1. Note that, although we fit different functions under the varying experimental conditions, i.e., independent random effects for smoothing by condition, the amount of smoothing is assumed the same, hence only one variance component has to be estimated. A more general model, smoothing the different experimental conditions with varying smoothing parameters, can be considered as well, in which case one may be inclined to test the need for varying the smoothing parameter by condition. The model considered can easily be extended or reduced in several ways. For example, one can assume a constant shift in the curves, if that is a reasonable assumption.

For comparison of curves from different experimental conditions, we propose use of bias adjusted simultaneous confidence bands around the fitted curves. Construction of such intervals requires the use of the

Copyright line will be provided by the publisher

variance-covariance matrix (see Ruppert *et al.*, 2003; SAS Institute Inc., 2004):

$$\boldsymbol{V} = \text{Cov}\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{u}} - \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^{'}\boldsymbol{S}^{-1}\boldsymbol{X} & \boldsymbol{X}^{'}\boldsymbol{S}^{-1}\boldsymbol{Z} \\ \boldsymbol{Z}^{'}\boldsymbol{S}^{-1}\boldsymbol{X} & \boldsymbol{Z}^{'}\boldsymbol{S}^{-1}\boldsymbol{Z} + \boldsymbol{G}^{-1} \end{bmatrix}^{-}.$$

Here, the matrix $\boldsymbol{S}$ is the conditional variance of the pseudo-data (see e.g., Molenberghs and Verbeke, 2005), generated during the fitting process. Let $\boldsymbol{P}$ denote the pseudo-data, then

$$\boldsymbol{S} = \text{var}[\boldsymbol{P}|\boldsymbol{u}] = \Delta^{-1}\boldsymbol{A}^{\frac{1}{2}}\boldsymbol{R}\boldsymbol{A}^{\frac{1}{2}}\Delta^{-1}.$$

From (1), one can write $E[\boldsymbol{Y}] = h^{-1}(\boldsymbol{X\beta} + \boldsymbol{Z}_k\boldsymbol{u}) = h^{-1}(\boldsymbol{\eta})$ and subsequently define $\Delta = \frac{\partial h^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}$. The matrix $\boldsymbol{A}$ is diagonal, containing variances of the response as a function of the mean and $\boldsymbol{R}$ is a residual variance matrix, dependent on user specifications.

Let $\boldsymbol{g} = (x_1, \ldots, x_T)^{'}$ be a set of values for which a simultaneous confidence band for $\boldsymbol{f} = (f_{x_1}, \ldots, f_{x_T})^{'}$ is required. It can be assumed that, approximately,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\boldsymbol{u}} - \boldsymbol{u} \end{bmatrix} \sim \text{N}\left(\boldsymbol{0}, \boldsymbol{V}\right). \tag{4}$$

Simultaneous confidence bands for $\boldsymbol{f}$ can then be obtained as

$$\left[\hat{f}_{x_t} \pm h_{(1-\alpha)}\widehat{\text{stdev}}\{\hat{f}(x_t) - f(x_t)\}\right]_{1 \le t \le T},$$

where $h_{(1-\alpha)}$ is the $1-\alpha$ quantile of (see Ruppert *et al.*, 2003)

$$\max_{1 \le x_t \le T}\left|\frac{\left(\boldsymbol{C_g}\begin{bmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\boldsymbol{u}} - \boldsymbol{u} \end{bmatrix}\right)_{x_t}}{\widehat{\text{stdev}}\{\hat{f}(x_t) - f(x_t)\}}\right|, \tag{5}$$

with $\boldsymbol{C_g} = [\boldsymbol{C}\ \boldsymbol{Z}]$, see Section 3.1.1. Essentially one simulates from (4) and computes (5) for N times and the value with rank $(1-\alpha)N$ becomes $h_{(1-\alpha)}$. As a result, $h_{1-\alpha}/z_\alpha$ approximates how wide the simultaneous confidence bands are, compared to their pointwise counterparts. The construction of confidence

Copyright line will be provided by the publisher

bands can be performed on the scale of the linear predictor and then transformed to intervals for the mean of the firing rate.

3.2 Derivation of the Time of Maximal Firing Rate and its Confidence Interval

One of the goals of the research is to detect the time at which the maximal firing rate occurs. To that effect, non-linear optimization techniques are implemented. After estimating the instantaneous firing rate, optimization and specifically, the conjugate gradient method (Gill, Murray and Wright, 1981; Fletcher, 1987) is applied. The optimization of the firing rate function λ_t is implemented in the SAS procedure IML using the NLPCG subroutine. The maximum firing rate is an immediate by-product of the maximization procedure. As is commonly encountered with non-unimodal optimization problems, the optimization algorithm implemented converges towards local rather than global optima. The smallest local minimum of an objective function is called the global minimum, and the largest local maximum of an objective function is called the global maximum. It is therefore not unusual that the algorithm occasionally fails to obtain the global optimum. Therefore, several starting values within the time range of interest are used. The objective function is evaluated over all these possible candidates, and the time that gives the maximum probability is taken as the time resulting in the maximal firing probability. This approach can be considered along the lines of a general concept for looking for features such as peaks, often referred to in literature as bump hunting (e.g., Heckman, 1992). Related approaches also include tests for monotonicity of regression functions (Gijbels *et al.*, 2000).

To determine a confidence interval for the time of maximal firing, we make use of the first-order derivative of the objective function. Similar approaches have been used in the literature, for example, Ganguli and Wand (2007) who apply tests for feature significance using the significant zero crossings methodology (SiZer) due to Chaudhuri and Marron (1999). Harezlak, Naumova and Laid (2007) employ bootstrap techniques to construct a test for bump hunting with penalized spline regression methodology. While these authors also make use of the first derivative of the objective function in the construction of their test, the

Copyright line will be provided by the publisher

main focus here is to determine the time corresponding to the maximal firing rate accompanied by a confidence interval.

Denote the derivative of $h(E[\boldsymbol{Y}])$ with respect to time by $h^{'}(E[\boldsymbol{Y}])$ and define the following matrices:

$$\boldsymbol{X}^d = [0\ 1]_{1 \leq t \leq T}, \quad \boldsymbol{Z}^d = \boldsymbol{Z}_k^* \boldsymbol{\Omega}_k^{-1/2}, \text{ and } \boldsymbol{Z}_k^* = [3(x_t - \kappa_k)|x_t - \kappa_k|]_{1 \leq t \leq T,\ 1 \leq k \leq K}.$$

It then follows from Section 3.1.1 that

$$h^{'}(E[\boldsymbol{Y}]) = \boldsymbol{X}^d \boldsymbol{\beta} + \boldsymbol{Z}^d \boldsymbol{u}. \tag{6}$$

Note that the model we consider assumes that $m = 2$ in (1). In general,

$$\boldsymbol{X}^d = \left[0\ 1\ 2x_t \ldots (m-1)x_t^{(m-1)}\right]_{1 \leq t \leq T} \quad \text{and} \quad \boldsymbol{Z}_k^* = \left[(2m-1)(x_t - \kappa_k)|x_t - \kappa_k|^{2m-3}\right]_{1 \leq t \leq T,\ 1 \leq k \leq K}.$$

The derivative function should be zero at the time corresponding to the maximum firing rate. To construct a confidence interval for time of maximal firing, a confidence interval for the derivative function is constructed. Defining $g(\lambda_t) = h^{'}(E[y_t])$, the variance function at a particular time point t is (Ruppert *et al.*, 2003)

$$\text{var}\{(\hat{g}(\lambda_t) - g(\lambda_t))\} \simeq \boldsymbol{C}_t^T \boldsymbol{V} \boldsymbol{C}_t\ ,$$

where $\boldsymbol{C}_t = \left[\boldsymbol{X}_t^d\ \boldsymbol{Z}_t^d\right]$. Construction of simultaneous confidence now follows the discussion in Section 3.1.1, with an appropriate adjustment to $\boldsymbol{C}_{\boldsymbol{g}}$ using $\boldsymbol{X}^d$ and $\boldsymbol{Z}^d$. Confidence limits for the time of maximal firing rate are then taken as the points where the obtained confidence limits for the derivative function cross the zero line. Note that the first-order derivative function at the scale of the link function is readily obtained by substituting the parameters by their estimates obtained from the model. Due to the monotonicity property of the link function, confidence intervals constructed for $h^{'}(\cdot)$ therefore suffice in this situation.

3.3 Population-averaged Model: Combining Information from Different Neurons

Our discussion until this far has focused on data from a particular neuron. However, it is the goal to combine the information from different neurons, resulting in a so-called population based analysis. We propose

Copyright line will be provided by the publisher

the use of a marginal or population-averaged model (Molenberghs and Verbeke, 2005). The approach used in this paper, which is implemented in the SAS procedure GLIMMIX, is based on linearization of the outcome variable (see e.g., Molenberghs and Verbeke, 2005), resulting in the application of weighted least squares (McCullagh and Nelder, 1989). In particular, the neurons are considered as the different subjects, accounting for the correlation of the measurements from each neuron through specification of a particular correlation structure, for example, an unstructured correlation or compound symmetry. The more general case of the generalized linear mixed model (GLMM) is of course a useful paradigm. Thus, one can consider a random-effects model, where neuron-specific random effects are used to account for the association. Such models are also handled by the procedure GLIMMIX. The marginal average evolution can then be obtained by averaging the conditional means over the random effects (Molenberghs and Verbeke, 2005), essentially integrating over them. Here, the random effects implied are neuron-specific effects, e.g., random intercepts, and not the random coefficients for smoothing. The former approach, which directly results in a population-averaged fit, is preferred in this situation.

Information from different experimental conditions from the different neurons is therefore combined to obtain condition-specific population-averaged profiles from which aspects of interest will be calculated. Since for each neuron the number of trials per experimental condition varies, we fit the model with the number of trials as an offset variable.

4 Application to the Data

4.1 Single Neuron Analysis

This section applies the penalized splines methodology on single neuron data, with the main focus on the population-averaged model being the subject of Section 4.2. Figure 1 displays data from a particular neuron, selected from the 20 neurons performing the CD task.

Figure 1 ABOUT HERE

Copyright line will be provided by the publisher

The graph shows the observed firing rates from the different experimental conditions, the corresponding raster plot from 176 trials as well as the fitted curves for each experimental condition, obtained using the penalized spline model. From the graphs, one observes that for this neuron, there is increased activity around 2000 ms into the experiment. The maximum firing rate occurs in this period. The more relevant comparison of the left and right decisions, for a fixed level of difficulty, is done using pointwise confidence intervals and simultaneous confidence bands in Figure 2. Other than the small section between about 1750 and 2000 ms, which appears to show a difference, the pointwise confidence intervals to suggest no significant differences between the left and right decisions in this case.

Figure 2 ABOUT HERE

Using the simultaneous confidence bands, which are expected to be wider than the pointwise counterparts, the apparent difference mentioned above is absorbed and the hypothesis of no difference between left and right is upheld. Figure 3 shows the first-order derivative function corresponding to the fitted profiles in Figure 1, together with their 95% pointwise confidence intervals and simultaneous bands. One can then approximate the confidence intervals for the time of maximal firing rate as suggested by the shaded areas in Figure 3. We return to the use of the derivative function and its confidence interval in the following section, where a marginal model that combines information from different neurons is considered.

Figure 3 ABOUT HERE

To give an overview on individual neurons, a similar analysis has been performed on each of the other neurons. There appears to be relatively large variability between neurons (see also Figure 4). A comparison of left and right orientations based on confidence intervals for each neuron separately is performed and plots given in Figures 7 and 8 in the Appendix. The results indicate that although in most neurons, differences between left and right occur in the region 1500-2000 ms, for some neurons, differences occur elsewhere. Moreover, the maximal firing rates in the sections showing differences are highly variable, suggesting that

Copyright line will be provided by the publisher

different neurons have different peaks. As such, in a population analysis, it may be difficult to detect differences between left and right with differences in different places tending to cancel other.

4.2 Overall Average Profile

In this section, focus is put on the marginal or population-averaged model. Essentially, all data from the different neurons have been combined to produce condition-specific profiles. Figure 4 shows individual neuron profiles for each of the experimental conditions wherein within-neuron variability appears substantial.

The model fitted assumes the same linear trend in the different experimental conditions albeit with independent random effects for smoothing each experimental condition, as in (3).

Figure 4 ABOUT HERE

This effectively produces different curves for different experimental conditions. Although the random effects are different, a single smoothing parameter is used, implying similar amount of smoothing in all experimental conditions. First, an independence model with an overdispersion parameter is considered.

Figure 5 ABOUT HERE

Other types of correlation structures, for example the compound-symmetry structure and AR(1) process, can be specified. We present results based on the compound-symmetry structure, under which the model was relatively easier to converge. Note that an unstructured variance-covariance matrix would yield a computationally prohibitive number of parameters and therefore may not be a good choice.

The fitted profiles in each of the experimental conditions are given in Figure 4 (right panel). The plot shows that curves from the same decision (left or right) look rather similar, suggesting no differences between levels of difficulty for a fixed decision.

Figure 6 ABOUT HERE

Copyright line will be provided by the publisher

From Figure 4, one can observe increased firing activity for decision to the left in the time period 1500 to 2000 ms. However, for decision to the right, no clear peak is evident, rather an overall increase is apparent between approximately 1500 and 2250 ms. The plot suggests that the time of maximal firing occurs earlier for decisions to the left compared to the right with the maximal firing rate being higher for decisions to the left.

Table 2 ABOUT HERE

Based on pointwise confidence intervals, apart from a small section between 1500 and 2000 ms, no differences between left and right are apparent in either of the levels of difficulty. Again, as one might expect, this apparent difference disappears as one considers simultaneous confidence bands. The simultaneous confidence bands used were found to be about 1.54 times wider than their pointwise counterparts. Note that the simultaneous confidence bands allow us to reach overall conclusions regarding differences or equality between the curves under comparison.

Figure 6 shows the first-order derivative and its 90% pointwise and simultaneous confidence intervals, together with an indication of the time of maximal firing. Since we already know the time of maximal firing, one would expect that its confidence interval may be deduced from the confidence interval of the derivative function, as exemplified by the shaded region in Figure 6. Note that in some experimental conditions, such as, for example, decisions to the right (see Figure 4), the time of maximal firing rate is not clear and may occur in a relatively wide region. In such instances, the upper and lower limits of the interval do not cross the zero line, leading to open-ended intervals. Here, we illustrate use of the proposed methodology using 90% confidence intervals. Such an interval can be interpreted as an interval such that in an indefinite repeat of similar experiments, 90% of the calculated confidence intervals for the maximal firing time will contain the true value of the time of maximal firing time. Table 2 displays the maximal firing times for each of the experimental conditions and the corresponding 90% pointwise and simultaneous confidence bands. The results in Table 2 suggest that for a fixed decision, there are no drastic differences between the levels of difficulty, neither in terms of the maximal firing rate nor the time of its occurrence. A similar conclusion

Copyright line will be provided by the publisher

may be drawn for the comparison between the left and right-oriented decisions. It is important to note the relatively wide confidence intervals/bands, especially for decision to the right. It is clear from the fitted curves in Figure 4 that for this condition, the maximum or peak may not be clearly determined, hence the wide confidence intervals on the time of maximal firing rate. The observed wide confidence intervals could be attributed to the large variability among the neurons, as mentioned before.

5 Discussion

We have considered an application of a flexible modeling technique, penalized smoothing splines in smoothing neuronal data. This approach is convenient, it can be applied by means of widely available commercial software for mixed models. The models can also be fitted in the Bayesian framework and consequently WinBugs, a publicly available free software package can be used. The models we have discussed are so general that they can be extended in several ways, exactly as required by the researcher. In particular, several possible scenarios depicting the evolution of curves in different experimental conditions, can be assumed. Differences or similarities can be assumed in the linear part of the models, the non-parametric part, or in both parts of the model. Extension of these models may also include variation of levels of smoothing in the different experimental conditions.

Focus has been on detecting the time of maximal firing rate and the maximal firing rate in a population of neurons subjected to different experimental conditions. Moreover, we were also interested in comparing the temporal evolution across the experimental conditions, the comparison between the left- and right-oriented decisions being the main focus. The model we focused on is of a marginal or population-averaged type, wherein correlations of observations from independent subjects, neurons in this case, is specified. Different types of correlation structures can be used in this context. However, with the number of time points encountered in such electrophysiological experiments, some structures like the unstructured correlation are simply computationally infeasible. As a result, less computationally demanding structures, for example the compound-symmetry or the simple structure, can be used. To compare the curves in a moment

Copyright line will be provided by the publisher

by moment sense (Cardárso-Suarez *et al.*, 2006), simultaneous confidence around the fitted instantaneous firing rates are constructed. This effectively solves the problem of testing for a difference at multiple time points and therefore allows global conclusions in the time domain.

For the time of maximal firing one can fit the model and obtain the time corresponding to the maximum firing rate. However, the time of maximal firing rate does not necessarily have to be one of the design points, and therefore we implemented an optimization procedure based on the penalized spline model fitted and confidence intervals on the time of maximal firing determined via the first order derivative function. It is clear that a number of properties of our proposed approach need to be investigated. For example, it is useful to assess the impact of the form of the underlying function, since it makes a difference whether it is constant, exhibits a single and small maximum, or features two local maxima. Also, the mean square error of the estimator as well as the coverage probability of the confidence intervals need to be investigated. In addition, a comparison with alternative estimators is worth undertaking. Also, it is worth exploring further to what extent results depend on the choices made for the grid on the time variable. This is topic of further research.

The analysis performed here suggests no significant difference between the experimental conditions under consideration, both in terms of temporal evolution and the occurrence of the time of maximal firing rate. Three possible causes for the result in the population analysis may be anticipated: (1) there is temporal variation between neurons and this jittering provokes the lack of significance; (2) the different heights of discharge rate at single neuron level damped the differences at population level. If this were the case, perhaps normalizing the firing rates could solve it; (3) the firing rate maximum peaks for left and right are almost of the same height but occur at different times; as a consequence there is no statistical difference between the two peaks. In our situation the data may be considered as heterogeneous in some sense. This means that events occur at different times so peaks tend to cancel the possible differences. It would be interesting to compare the results with a population of neurons known to be homogeneous or in the same 'phase'.

Copyright line will be provided by the publisher

It is important to mention that, with the methodology considered here, inference and specifically hypotheses testing may be compromised due to violation of existing theory. For example, the likelihood ratio test in the presence of penalized splines is discussed by Ruppert *et al.*, 2003 and Crainiceanu *et al.*, 2004).

Acknowledgements

The first four authors gratefully acknowledge the financial support from the IAP research network nr. P5/24 of the Belgian Government (Belgian Science Policy) and from the Institute for the Promotion of Innovation by Science and Technology (IWT) in Flanders, Belgium.

The research of Carlos Acuña was supported by the Ministerio de Educacion y Ciencia (MEC) grant BFU2006-06657. Work of José L. Pardo-Vázquez and Víctor Leborán was supported by pre-doctoral MEC fellowships.

The research of Christel Faes, Geert Molenberghs, and Carmen Cadarso-Suárez was partially supported by the Ministerio de Educacion y Ciencia (MEC) grant MTM2005-00818.

Copyright line will be provided by the publisher

References

Behseta, S. and Kass, R.E. (2005) Testing equality of two functions using BARS. *Statistics in Medicine*, **24**, 3523–3534.

Behseta, S., Wallstrom, G.L., and Kass, R.E. (2005) Hierachical models for assessing variability among functions. *Biometrika*, **92**, 419–434.

Cleveland, W.S. and Grosse, E. (1991) Computational methods for local regression. *Statistics and Computing*, **1**, 47–62.

Cadarso-Suárez, C., Roca-Pardinas, J., Molenberghs, G., Faes, C., Nacher, V., Ojeda, S., and Acuna, C. (2006) Flexible modeling of neuronal firing rates across different experimental conditions: an application to neural activity in the prefrontal cortex during a discrimination task. *Journal of the Royal Statistical Society, Series C*, **55**, 431–447.

Chaudhuri P., Marron, J.S. (1999) SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, **94**, 807–823.

Crainiceanu, C.M., Ruppert, D., Claeskens, G., and Wand, M.P. (2005) Exact likelihood ratio tests for penalized splines. *Biometrika*, **92**, 91–103.

DiMatteo, I., Genovese, C.R., and Kass, R.E. (2001) Bayesian curve-fitting with free-knot splines. *Biometrika*, **88**, 1055–1071.

Durbán, M., Harezlak, J., Wand, M.P., and Carroll R.J. (2005) Simple fitting of subject-specific curves for longitudinal data. *Statistics in Medicine*, **24**, 1153–1168.

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

Copyright line will be provided by the publisher

Faes, C., Geys, H., Molenberghs, G., Aerts, M., Cadarso-Suárez, C., and Acuna, C., and Cano, M. (2007) A flexible method to measure synchrony in neuronal firing. *Journal of the American Statistical Association, 00,000-000*

Ganguli, B., Wand, M.P. (2007) *Statistical Computing*, **17**, 179-192.

Gerstein G.L., and Kiang N.Y.S (1960) An approach to the quantitative analysis of electrophysiological data from single neurons. *Biophysical Journal*, **1**, 15-28.

Gijbels, I., Hall, P., Jones, M.C., Koch, I. (2000) Tests for monotonicity of a regression mean with guaranteed level. *Biometrika*, **87**(3), 663–673.

Heckman, N.E. (1992) Bump hunting in regression analysis. *Statistics and Probability Letters*, **14**, 141–152.

Kass, R.E., Ventura, V., and Brown, E.N. (2005) Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, **94**, 8–25.

Kass, R.E., Ventura, V., and Cai, C. (2003) Statistical smoothing of neuronal data. *Network: Computation in Neural Systems*, **14**, 5–15.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models.* London: Chapman & Hall.

Molenberghs, G. and Verbeke, G. (2005) *Models for Discrete Longitudinal Data.* New York: Springer.

Ngo, L. and Wand, M.P. (2004) Smoothing with mixed model software. *Journal of Statistical Software*, **9**, 1–56.

Roca-Pardinas, J., Cadarso-Suárez, C., Nacher, V., and Acuña, C. (2006) Boostrap-based methods for testing factor-by-curve interactions in generalized additive models: assessing prefrontal cortex neural activity related to decision-making. *Statistics in Medicine*, **25**, 2483-2501.

Copyright line will be provided by the publisher

Ruppert, D. (2002) Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics*, **11**, 735–757.

Ruppert D., Wand M.P., and Carroll R.J. (2003) *Semiparametric Regression*. Cambridge University Press.

SAS Institute Inc. (2004) *The GLIMMIX Procedure (Experimental).* Cary, NC: SAS Institute Inc.

Ventura, V., Carta, R., Kass, R.E, Gettner, S.N., and Olson, C.R. (2002) Statistical analysis of temporal evolution in single-neuron firing rates. *Biostatistics*, **1**, 1–20.

Verbeke, G. and Molenberghs, G. (2002) *Linear Mixed Models for Longitudinal data.* New York: Springer.

Verbyla, A.P, Cullis, B.R., Kenward, M.G., and Welman, S.J. (1999) The analysis of designed experiments and longitudinal data by using smoothing splines. *Applied Statistics*, **48**, 269–311.

Vazquez, P., Cano, M., and Acuna, C. (2000) Discrimination of the line orientation in humans and monkeys. *Journal of Neurophysiology*, **83**, 2639–2648.

Wolfinger, R. and O'Connell, M. (1993) Generalized linear mixed models: a pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*, **4**, 233-243.

Appendix

Figures 7 and 8 show the comparison between left and right decisions in either of the two levels of difficulty for neuron-specific data.

Figure 7 ABOUT HERE

Figure 8 ABOUT HERE

Copyright line will be provided by the publisher

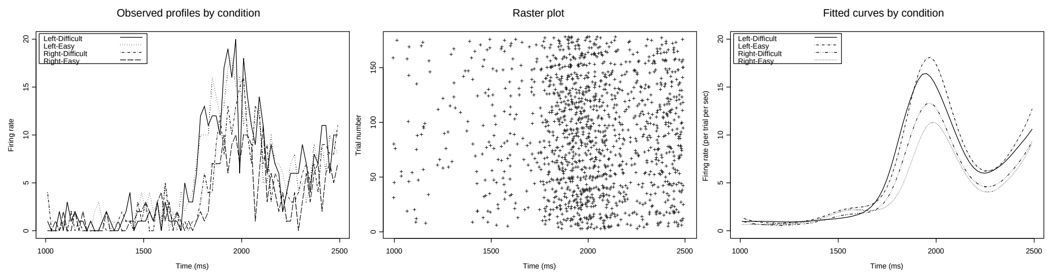

Fig. 1 *Observed firing rates by experimental condition (left) raster plot for that particular neuron (center) and fitted profiles by experimental condition (right).*

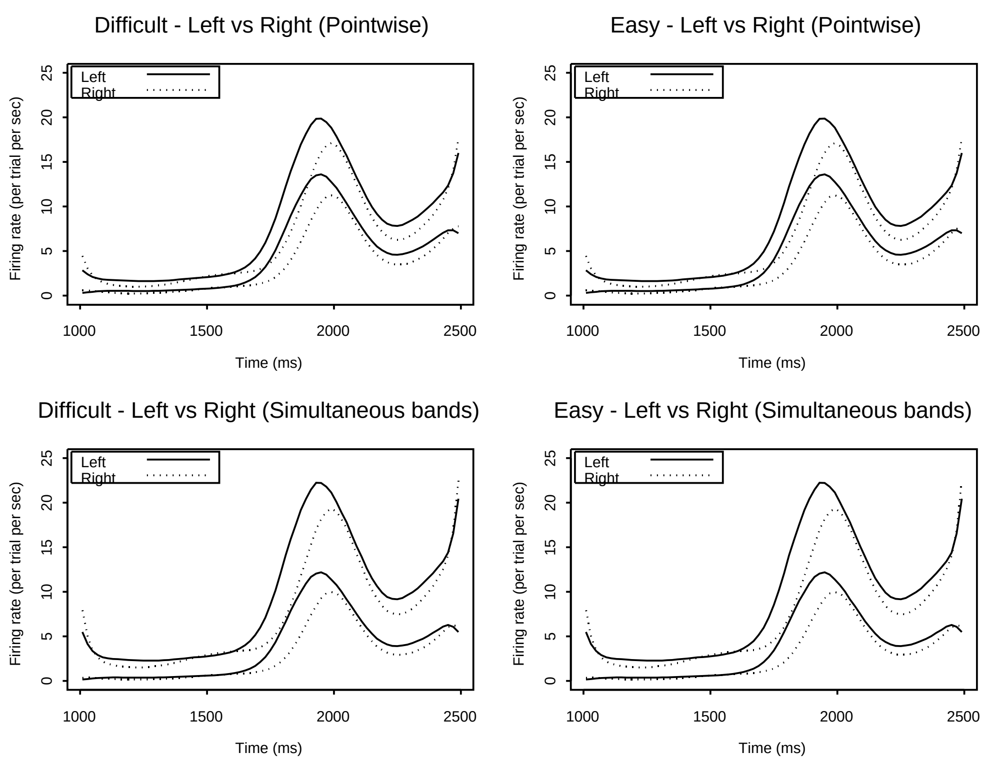

Fig. 2 *Pointwise confidence intervals and simultaneous bands comparing the left and right orientation for a fixed level of difficulty.*

Copyright line will be provided by the publisher

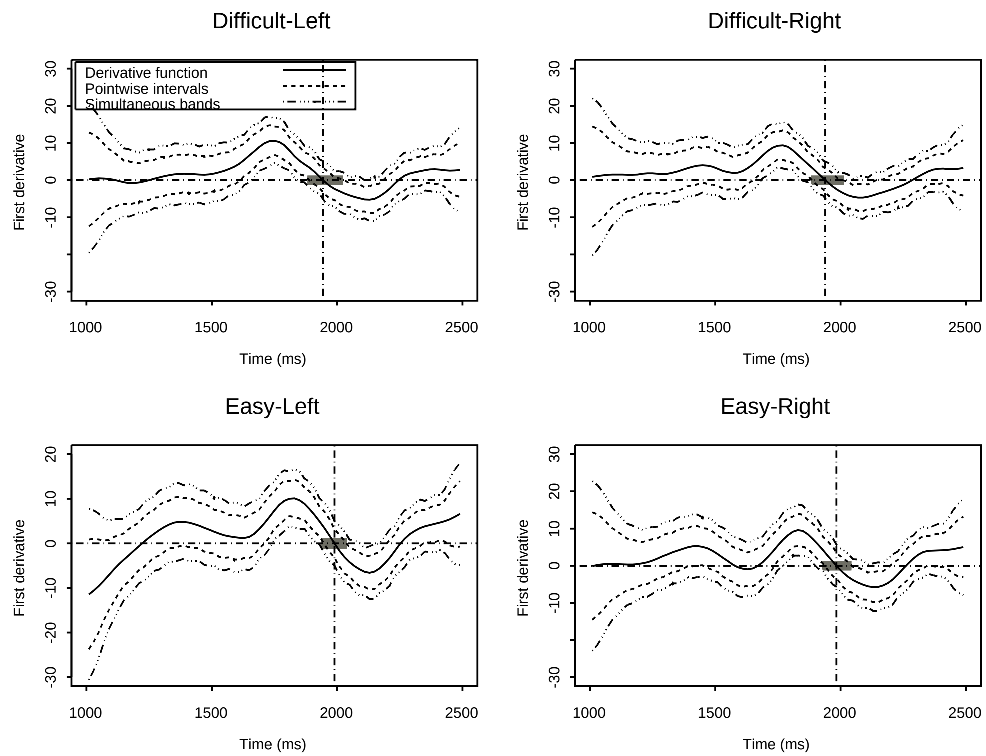

Fig. 3 *First-order derivative functions (continuous line) for each condition together with corresponding 95% pointwise and simultaneous confidence intervals. The vertical dashed line indicates the time of maximal firing rate and the shaded regions indicate its corresponding confidence interval, approximated by the pointwise intervals.*

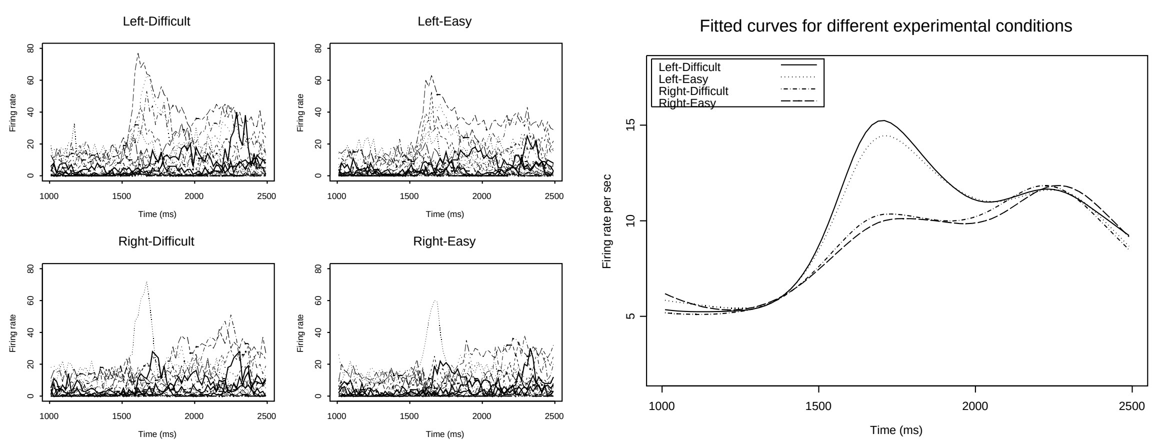

Fig. 4 *Observed firing rates for all neurons in the different experimental and fitted curves in each condition.*

Copyright line will be provided by the publisher

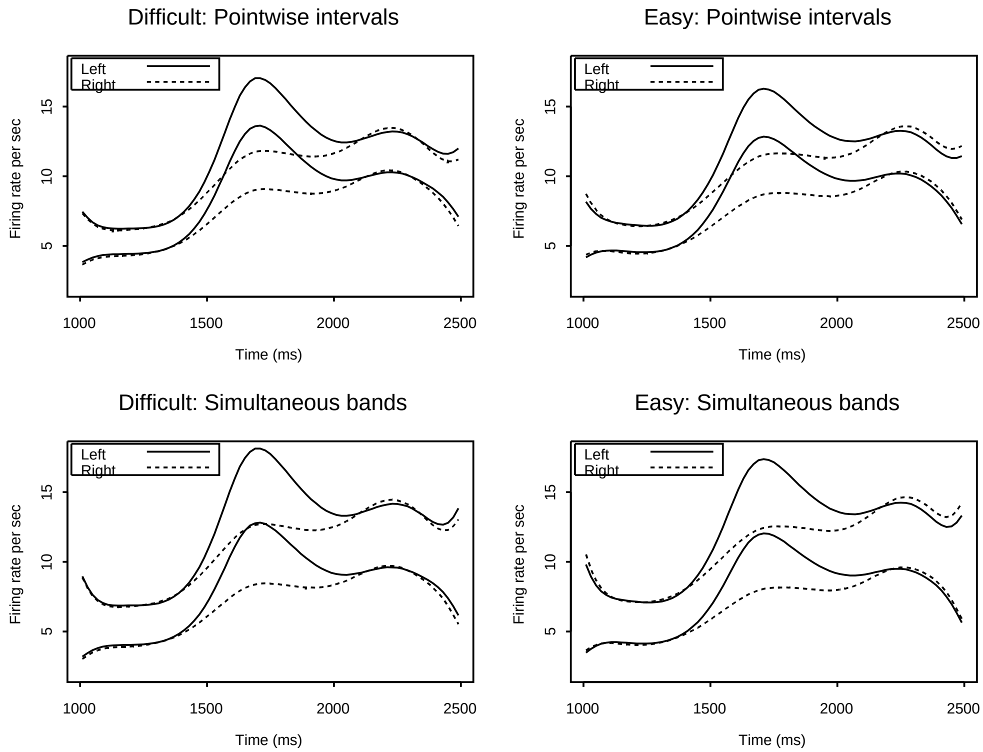

Fig. 5 *Overall comparison of left and right orientations for a fixed level of difficulty using 95% pointwise and simultaneous confidence bands.*

Table 1 *The sequence of events over time in each trial, starting from -500 to 4500 ms.*

Time period	Event
-500-0ms	Control period
0-500ms	Presentation of first (reference) stimulus
500-1500ms	Interstimulus interval (ISI) or delay period
1500-2000ms	Presentation of second (test) stimulus (comparison/decision period)
2000ms +	Subject makes a saccadic eye movement towards
	one of the 2 circles for reward

Copyright line will be provided by the publisher

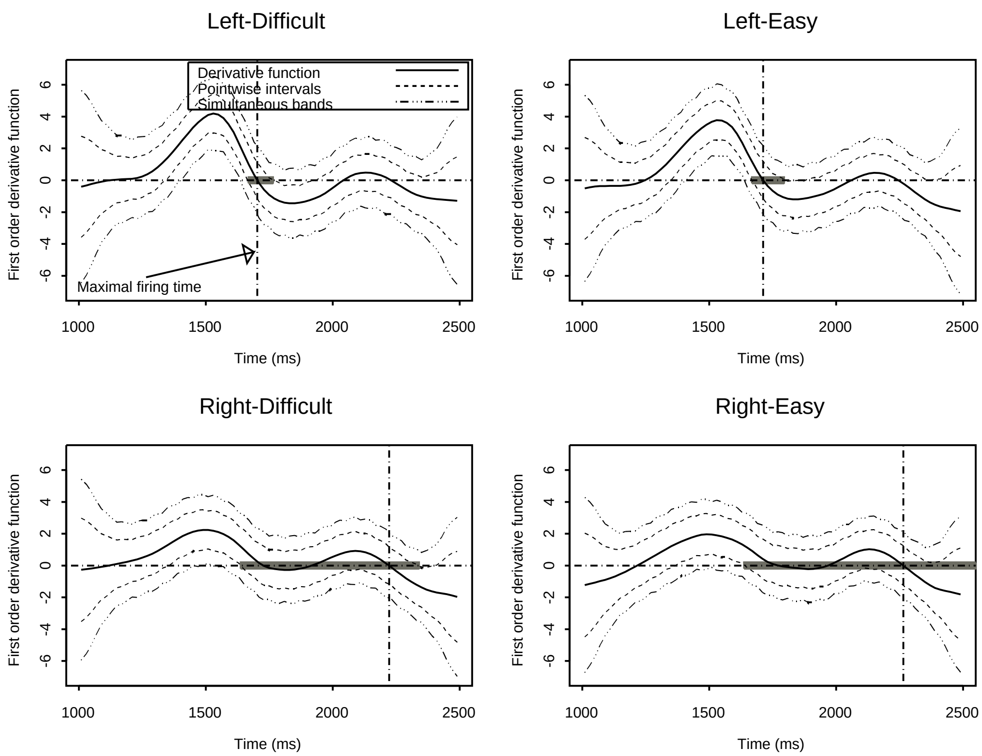

Fig. 6 *First-order derivative functions (continuous line) for each condition together with corresponding 90% pointwise and simultaneous confidence bands. The vertical dashed line indicates the time of maximal firing rate and the shaded region indicates its corresponding pointwise confidence interval.*

Copyright line will be provided by the publisher

Table 2 *Maximal firing times and corresponding 90% pointwise and simultaneous confidence bands. Also given are the maximal firing times (in milli seconds) and maximal firing rates for the different experimental conditions.*

			Approximate 90% confidence intervals				
			Pointwise		Simultaneous		
		T_{max}(ms)	Lower	Upper	Lower	Upper	Firing rate (spikes per sec)
Left	Difficult	1703	1668	1765	1635	∞	15
	Easy	1712	1668	1795	1630	∞	14
Right	Difficult	2222	1637	2340	1550	∞	12
	Easy	2264	1637	∞	$-\infty$	∞	12

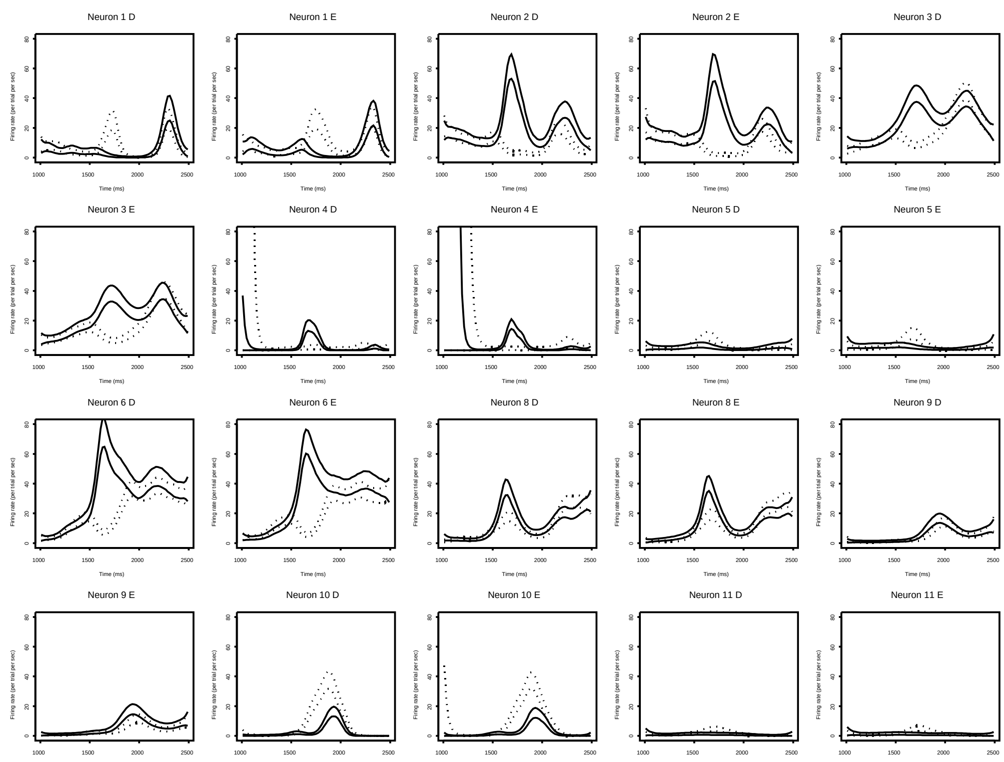

Fig. 7 *Confidence intervals comparing Left vs Right in individual neuron data: 1 (D represents "difficult" and E represents "Easy").*

Copyright line will be provided by the publisher

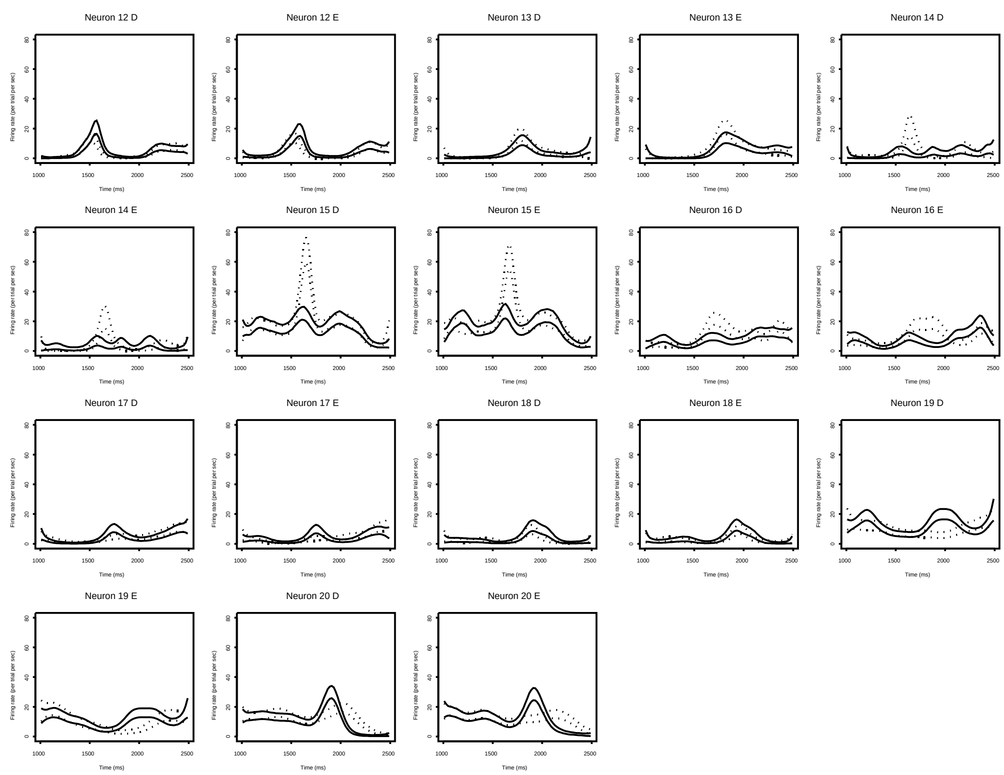

Fig. 8 *Confidence intervals comparing Left vs Right in individual neuron data: 2.*