

Comparisons of Split-linear Fitting of Wind Curves

Philippe C. Besse¹ and Nathalie Raimbault²

¹*Université Paul Sabatier* and ²*EADS AIRBUS*

Abstract: The detection of slope change points in wind curves depends on linear curve-fitting. Hall and Titterington's algorithm based on smoothing is adapted and compared to a Bayesian method of curve-fitting. After prior spline smoothing of the data, the algorithms are tested and the errors between the split-linear fitted wind and the real one are estimated. In our case, the adaptation of the edge-preserving smoothing algorithm gives the same good performance as automatic Bayesian curve-fitting based on a Monte Carlo Markov chain algorithm yet saves computation time.

Key words: Bayesian curve-fitting, MCMC, slope change detection, spline smoothing, split-linear fitting.

1. Introduction

This study is aimed at the improvement of the aircraft autopilot conception process. The autopilot allows landings in bad weather conditions and must guarantee passengers safety, touchdown comfort, and precision. We study in particular the influence of wind during automatic landing. We focus on the effect of the linear wind components in the last 30 seconds to show that they are a decisive factor in touchdown precision. This is achieved by comparing simulated landings with either a real wind or its piecewise linear approximation. This has led us to develop a method of split-linear fitting based on slope change detection adapted to our data. This method is similar to those proposed by Jones (1998) which aim at predicting the influence of discrete gusts on linear systems.

Slope change detection is associated with edge and peak detection. Two kinds of method can be adapted : a classic one, based on smoothing and a Bayesian one based on Monte Carlo Markov Chains. Hall and Titterington (1992) proposed an edge-preserving smoothing algorithm. It is based on edge detection by comparisons of three smoothings. The aim of this method is to compute, for each given point, three smoothed estimates of the function, based on the data to the right, to the left and on both sides of the point. Each discontinuity is associated with a local maximum of the difference between the three fits. Wu and Chu (1993) took this algorithm and modified it using kernel smoothing instead

of smoothing by linear combinations of observations. Loader (1996) showed convergence properties of this kind of change point estimate adapted to the search for one discontinuity in a function. Müller and Song (1997) improved this kind of algorithm to increase the convergence rate of one discontinuity estimate. Green (1995) and Denison *et al.* (1998) applied reversible Monte Carlo Markov Chains and Bayesian model determination to curve-fitting, including, in particular, piecewise linear polynomial fitting. The automatic Bayesian curve-fitting algorithm proposed by Denison *et al.* (inspired by Green's (1995) algorithm) determines the number and the position of the slope change points, knots, by the estimation of their joint posterior distribution.

Our aim here is to derive an efficient split-linear fitting algorithm. Unlike other studies, the quality of fit is not measured by the usual residual squared error. The criteria to be met are estimation of a parsimonious model by detecting only the significant slope changes in the curve and achievement of the same landing properties with the real wind and with the split-linear approximation. This is an important feature of this paper, which especially involves adaptations of the algorithms which consist in prior smoothing of the data and algorithm refinements.

A variation on Hall and Titterington's (1992) method is computed on samples of noise-free wind data measured during flight tests. The results are compared with those obtained with the automatic Bayesian curve-fitting algorithm. Denison *et al.* (1998) algorithm is taken here as benchmark but considered as too computational expensive for real application. The Hall and Titterington's method consists in computing and comparing three smoothings. The local extremes of the difference between the three smoothings are used as estimates of slope change points. Moreover kernel smoothing is here replaced by spline smoothing to improve the quality of the fitting as further detailed. The proposed variation is based on a single slope sign change detection by slope comparisons. These slope comparisons are combined with smoothing comparisons to localize other slope changes between the previously detected ones.

Tests and comparisons show that the performances of the two methods are similar. Differences between touchdown with real wind and touchdown with split-linear fit are estimated with a complex simulation program. This allows us to perform free-choice but time-costly landing simulations. Implementation is carried out using S-plus(1997).

The specific problem and the data transformations are explained in Section 2. The algorithms will be presented in Section 3. The numerical results and the comparison between the methods are explored further in Section 4. We will conclude by a brief discussion in Section 5.

2. Aeronautical Context and Data Transformations

We aim to decompose wind speed evolution into linear components to study later their own influence on the behaviour of an aircraft autopilot. The stage of the flight we are interested in is the final approach, specifically the last 30 seconds before touchdown when wind perturbations are critical because of the imminence of the landing. Split-linear fit of wind curves will help to confirm the idea that the linear components of the wind are the main factor towards unusual landings. Then it will be possible to discriminate between severe landing due to particular external conditions and bad tuning of the autopilot. Moreover, the split-linear fit of real flight winds will add to the wind databases usually used for landing simulations.

Only the longitudinal wind component $W_x(t)$ along the runway axis is considered. Landing quality is estimated by the distance in meters between the beginning of the runway and the touchdown on the runway axis (X_{TP}). Our sample of data consists of 59 winds measured during flight tests. As we focus on the latest 30 seconds before landing, each wind recording is a time serie of $N=240$ points sampled every $\frac{1}{8}$ th of a second.

Simulation software allows us to perform landing simulations. This is convenient since the criterion to be met is expressed in term of error between touchdown performance (X_{TP}) instead of the usual residual square error. This software is used to set the different values of the parameters and to evaluate the algorithms. The difficulty we have to cope with, is that it cannot be included in an optimisation or a cross validation program.

Direct applications of the following algorithms on raw data are not satisfactory. Successful computations have been carried out using smoothed wind curves. We aim to determine only the significant slope change points and avoid finding too many knots. Experiments with raw data lead to non-significant point detection when using Hall and Titterington's algorithm and to difficulties in setting up the prior distribution of the number of knots in the Bayesian algorithm. Moreover, the estimated errors between touchdowns with the split-linear fits of raw data are too large.

So a necessary first step consists in clearing winds of non significant information. That is, components which do not change the touchdown. It may be measurement noise as well as some frequency components which do not disturb the aircraft during the flight. To be more specific, high frequencies have little influence on our aircraft trajectory because of its inertia. Smoothing splines are used to get rid of the high frequencies and allows the slope change detection to be efficient. Note that Fourier methods are not applicable here since the data are neither stationary nor periodic.

Let $v(t_i)$ be the cubic spline fit of the velocity values $\{(t_i, w_i), i = 1, \dots, N\}$:

$$\sum_{i=1}^N (w_i - v(t_i))^2 + \lambda \int_a^b (v''(t))^2 dt \text{ where } w_i = W_x(t_i), \quad \forall i = 1, \dots, N$$

The smoothing parameter λ governs the trade-off between smoothness and goodness of fit and is usually estimated by cross validation. In our particular case, we aim at keeping only the significant components of the wind towards the landing. Therefore goodness of fit is estimated here by the error between the real X_{TP} touchdown with raw data and the one with smoothed data. This criterion replaces the usual residual error minimization. Empirical cross validation has led us to choose $\lambda = 10^{-5}$ as the largest smoothing parameter minimizing error mean and scatter (4% error maximum).

3. Split-linear Fitting Algorithms

The aim is to detect the slope change times t_j^* among the times $\{t_1, t_2, \dots, t_n\}$. The split-linear algorithms are detailed in this section. They are applied on smoothed wind data as previously explained in Section 2. Observations are denoted (t_i, w_i) whereas the values of the smoothed wind curves are denoted by v_i .

Since one matching condition is to detect only significant slope changes, each method implies the search for the optimum values of the different parameters which govern the precision.

3.1 Variation on Hall and Titterington algorithm

The features of this algorithm inspired by Hall and Titterington (1992) are as follows:

1. The smoothed wind curve is split into intervals I_k , $k = 1, \dots, K$.
2. In each interval, the slope a_k of the best straight line is estimated by ordinary least-squares.
3. The consecutive slopes a_k and a_{k+1} are compared.
If $a_k \cdot a_{k+1} < 0$, a point of slope sign change t_j^* is detected.
4. If the interval $[t_j^*, t_{j+1}^*]$ is larger than T_{min} , points of single slope change t_j^{**} are sought by the comparison of three smoothings further detailed.
5. Since every t_j^*, t_j^{**} has been detected, the split-linear fitting of the smoothed wind curves is carried out by linking the $v_j^* = v(t_j^*)$ or $v(t_j^{**})$.

Linear least-squares model (step 1. and 2.)

The size of the K intervals governs the trade-off between accuracy of the method and calculation time. In our case, tests lead us to select 4-points intervals. In each interval, the piecewise linear least-squares model is fit to the data :

$$v_i = a_k t_i + b_k, \quad \forall i \in [t_{4(k-1)+1}, t_{4k}] \quad \forall k = 1, \dots, K.$$

Smoothing comparison (step 4.)

The bandwidth parameter h has to be chosen for each smoothing. It determines how large is the neighbourhood of each point used to calculate the three smoothings.

For each point in $[1+h, N-h]$, left, right and centre smoothing are computed. Hall and Titterington propose smoothers such as particular linear combinations of the observations around each point. This smoothing has been tested and gave worse results on our data than spline smoothing. So spline smoothing has been chosen for left and right smoothing whereas centre smoothing is obtained by averaging.

Let $f_r(i)$ be the value of the right smoothing at the point i computed on (t_k, v_k) , $k \in [i, i+h]$. Let $f_l(i)$ be the value of the left smoothing at the point i computed on (t_k, v_k) , $k \in [i-h, i]$. Let $f_m(i)$ be the mean of the $\{v_k, k \in [i-h, i+h]\}$.

For the two spline smoothings, the smoothing parameter λ has to be chosen. Moreover, the bandwidth h has to be large enough to avoid side-effects of the smoothing. But since $2h$ points are lost for the slope change detection, a trade-off has to be sought.

These two parameters λ and h can be optimized by cross validation. In our case, we have to choose prior satisfactory empirical values and perform the choice further as regards the numerical results in Section 4.

The diagnosis is made by estimating for each point $i \in [1+h, N-h]$, the difference d_1 defined by $d_1(i) = |f_r(i) + f_l(i) - 2f_m(i)|$. If the point i is a slope change point then this difference will be a local maximum.

Note that since the data are smoothed before the algorithm is applied, $f_r(i) \simeq f_l(i)$, $\forall i \in [1+h, N-h]$. So $d_1(i)$ can be replaced arbitrarily by $d_2(i) = |f_r(i) - f_m(i)|$. The detection of the local maxima is made by the comparison of each point with its two neighbours. A point i is a local maximum if :

$$|f_r(i) - f_m(i)| > |f_r(j) - f_m(j)| \quad \text{for } j = i-1 \text{ and } i+1$$

Each local maximum of d_2 noted i is associated with a slope change point $t_j^{**} = t(i)$.

In brief, the parameters to set are : the number of intervals K , the optional T_{min} , the smoothing parameter λ and the bandwidth h .

3.2 Bayesian automatic curve-fitting

Recall here that we aim at identifying the slope change times t_j^* among the t_i . Therefore, the algorithm of Denison *et al.* is applied on our data so as to obtain a split-linear fit and only modified to obtain the final knots. The Bayesian basis of the algorithm is briefly described here according to Denison's *et al.* (1998) and Green's (1995) articles.

The piecewise polynomial model to fit is as follows:

$$v_i = f_{K,1}(t_i) + \epsilon_i = \sum_{n=0}^1 \beta_{n,0}(t_i - t_0^*)_+^n + \sum_{j=1}^K \beta_{1,j}(t_i - t_j^*)_+ + \epsilon_i, \quad i = 1, \dots, N. \quad (3.1)$$

The dimension K of the true model is unknown as well as $\theta^{(K)} = (t_1^*, \dots, t_K^*, \sigma^2)$ representing knots positions and variance of the model error. Inference about K and $\theta^{(K)}$ is based on the joint distribution of $(K, \theta^{(K)}, v)$. The best $(K, \theta^{(K)})$ are assumed to maximize the posterior distribution $p(K, \theta^{(K)}|v) = \frac{1}{Z} p(K) p(\theta^{(K)}|K) p(v|K, \theta^{(K)})$. The coefficients $\beta_{n,j}$ in (3.1) are then estimated by least-squares minimization.

Since an analytic or numerical expression of this complicated distribution is impossible, samples of the joint posterior distribution have to be simulated.

Denison *et al.* (1998) propose the design of three types of move (addition, deletion and movement of a knot) between the subspace Θ_K for θ_K which allows us to explore all the subspaces, searching for the (K, θ_K) which minimizes $MSE = \frac{1}{n} \sum_{i=1}^N (v_i - f_{K,1}(t_i))^2$. Each move is chosen with a specific probability which depends on the prior distribution of K .

The prior probability of the dimension K of the model is assumed to be a Poisson distribution with parameter λ_p :

$$p(K) = \frac{\lambda_p^K \exp(-\lambda_p)}{K!}, \quad K = 0, 1, 2, \dots$$

Two examples are taken here to illustrate the sensitivity to λ_p . Few values of λ_p are tested on the two raw and on the two smoothed winds. The final number of knots after 10000 iterations is plotted against the different values of λ_p . The two top panel of Figure 1 corresponds to the raw data and the other two to the smoothed data.

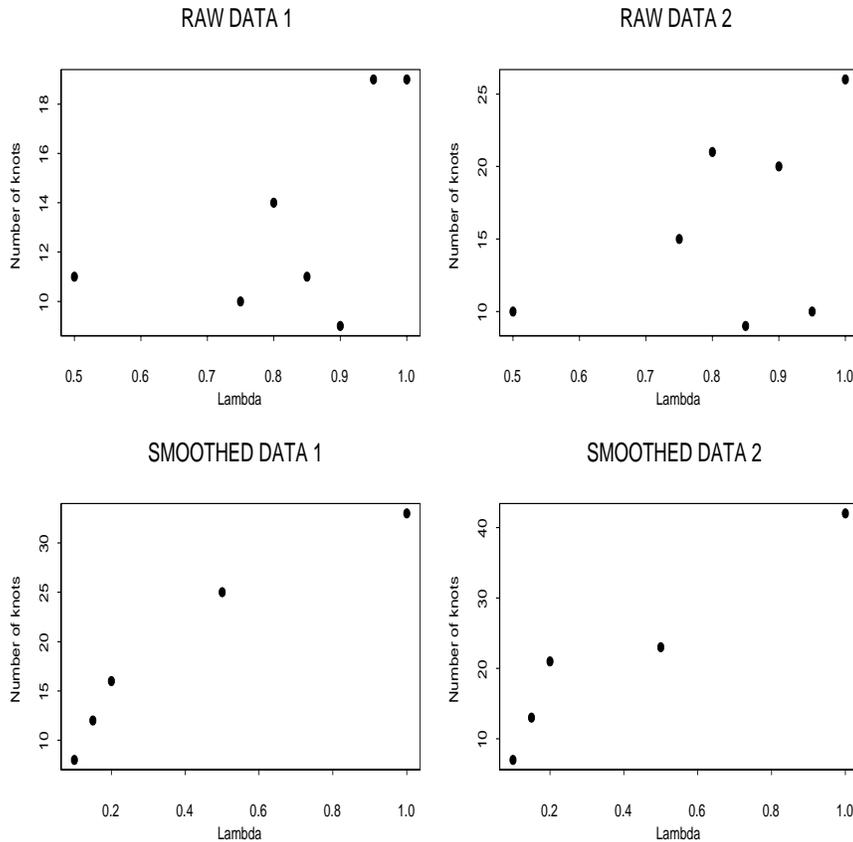


Figure 1: This Figure depicts the influence of λ_p on the final number of knots K obtained for two raw and smoothed winds. The difficulty to set λ_p for raw data leads us to use prior spline smoothing on our data.

These plots confirm that the Denison *et al.* (1998) algorithm gives better performance on “smooth” curves and show that the value of λ_p is impossible to set for raw data.

The algorithm of Denison *et al.* can be summarised in three steps as follows:

1. Initialisation : choose λ_p knots uniformly among the t_i and set k equal to the number of interior knots.
2. Choice of the type of move : generate u uniformly and compare it to b_k and d_k to choose a birth, death or move transition.
3. Iteration : repeat step 2 until the MSE no longer decreases.

In brief, the parameters to be set for this algorithm are the smoothing parameter λ and the parameter λ_p of the prior distribution of K .

Note that the quality of fit is measured by the MSE since, as already explained, it is impossible to replace the criterion by the one we focus on, that is difference between touchdown with the fitted wind curves and the real ones. However, verifications are carried out (see estimation of error in Section 4.) to ensure that the fits were satisfactory according to our criterion.

4. Practical Comparisons

In this section, both algorithms are applied on the wind data described in Section 2. The test sample is composed of all 59 available winds. Each algorithm is tested separately at first and then a numerical comparison is carried out. The variation on Hall and Titterington's algorithm is compared to the Bayesian approach. The Bayesian approach is here considered as a benchmark whose quality is well-known but computational costly.

Graphic comparisons are made between raw data and split-linear fitting to check the quality of the methods. Then touchdowns are simulated with the two kinds of wind, the errors are estimated and compared so as to choose the best algorithm.

The tests of the algorithm inspired by Hall and Titterington's one are carried out with empirically best values of the different parameters to be set. To carry out a fair comparison with the Denison's one, we have to focus on the final number of knots K (slope change point) as we are searching for a parsimonious linear-fits. The comparison protocol is as follows:

1. Choose the maximum number of knots allowed K_{max} (in our case, $K_{max} = 20$ since too close slope changes do not have any physical meaning).
2. For each wind, check that the number $K_{class} < K_{max}$ (K_{class} is the number of knots obtained by the classic algorithm).
3. For each wind, set λ_p (parameter of the Poisson distribution of the number of knots) such as $|K_{class} - K_{Bayesian}| \leq 3$ with $K_{Bayesian}$ number of knots given by the Bayesian linear curve-fitting algorithm.
4. Carry out the graphical comparisons for each wind.
5. Carry out the numerical comparisons for each wind by simulating the X_{TP} impact obtained with the two split-linear fits.

4.1 Graphic comparisons

Test of the variation on Hall and Titterington's algorithm

The parameters are set to the best empirical ones ($\lambda = 10^{-5}$ and $h = 15$ points). Further results in Section 4.2 will show that they also are the best as regards to error minimization.

Figure 2 et 3 illustrate the interest of the adaptation performed on the Hall and Titterington's algorithm.

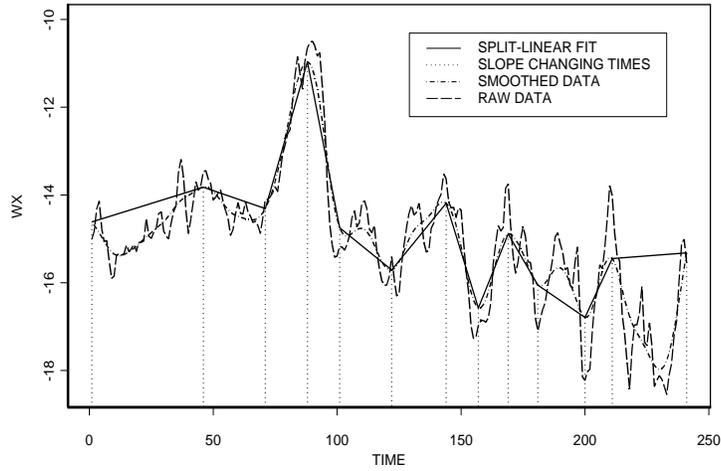


Figure 2: Example of an unsatisfactory graphic comparison between the raw data and the split-linear fitting obtained with the Hall and Titterington's algorithm. At least five slope changes are forgotten at times 15, 190 and between times 210 and 250.

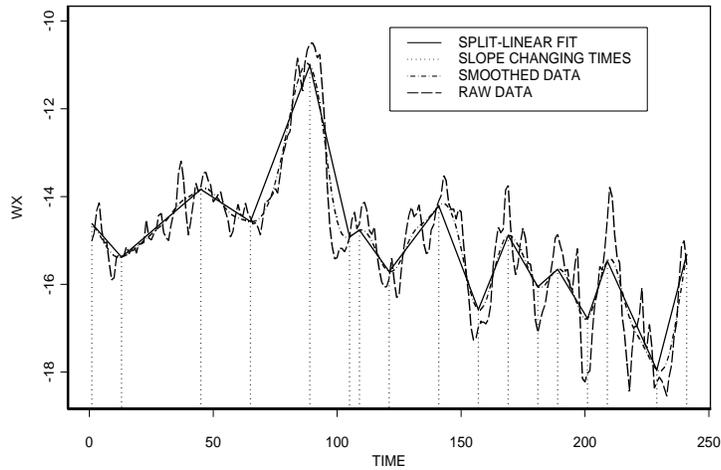


Figure 3: Example of a satisfactory graphic comparison between the raw data and the split-linear fitting obtained with the variation on Hall and Titterington's algorithm. This corresponds to Figure 2. Every slope sign changes are detected.

Test of the Bayesian automatic curve-fitting algorithm

The graphic example shown in Section 4.1 (Figure 3) is then compared to the results given by the Denison *et al.* algorithm on the same wind. The comparison protocol is followed. K_{class} is estimated then λ_p is chosen such that $|K_{class} - K_{Bayesian}| \leq 3$.

For the wind shown here as an example, $K_{class} = 15$, $K_{Bayesian} = 13$ for $\lambda_p = 0.15$.

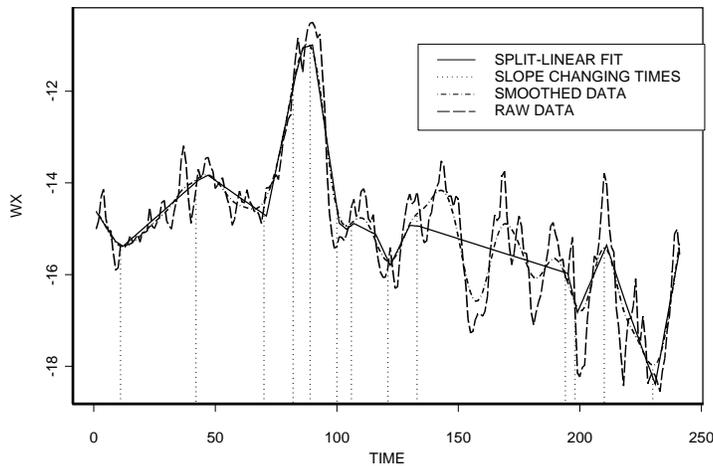


Figure 4: The raw data and the split-linear fit obtained on smoothed data with the Bayesian algorithm are plotted. This Figure is associated with Figure 3.

The graphical result seems to be less satisfactory than the one given by the first algorithm. But as our criterion to be met is mainly to obtain the same X_{TP} touchdown with the split linear fit and the raw wind as explained in Section 4.2, we carry out numerical comparisons.

4.2 Estimation of error

As previously said, the main criterion of satisfaction consists in the estimation of error. The usual residual error minimization is then replaced by the minimization of the difference between touchdown with raw data and the split-linear fit.

A subset of 10 winds is extracted. A split-linear fit is estimated with each algorithm using the same values of parameters as in Section 4.1. Then touchdown simulations are computed comparing the value of X_{TP} obtained with each raw wind and its three fits. The errors are estimated in percent of the X_{TP} calculated with the raw wind.

The numerical comparisons are carried out in the same way as the graphic comparisons. The comparison protocol is applied here. For each wind, K_{class} is estimated and λ_p is chosen such as $|K_{class} - K_{Bayesian}| \leq 3$.

Figure 5 depicts the boxplot of errors of the two algorithms. The performance of these two methods are similar.

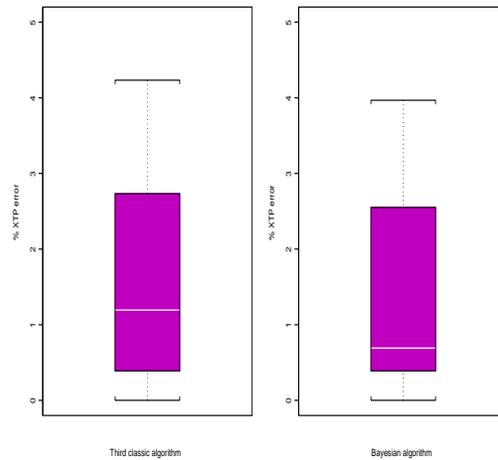


Figure 5: Boxplots of errors of the two algorithms estimated in percents of the X_{TP} calculated with the raw wind. The results are similar. In the two cases, the maximum error is less than 4.5%.

5. Discussion

The purpose was the split-linear fitting of time series of wind velocity before touchdown. Two algorithms have been discussed in this article and compared on real-life data issued from wind measurement during flight tests. We aimed at determining all the significant slope changes in the wind curves. The criterion to meet was the achievement of the same simulated touchdown with the raw wind and its linear fit. Two kinds of approach have been considered. The algorithm based on slope or smoothing comparisons (Hall and Titterington) has been adapted and compared to a Bayesian curve-fitting algorithm used as a benchmark. This involved particularly the choice of best values for the parameters of each method. For the first algorithm, the smoothing bandwidth and parameter

had to be set while the prior distribution of the number of slope changes was the main parameter to consider for the Bayesian algorithm. The main features of this article is the interest of a prior smoothing of the data and the setting protocol for all parameters. These settings have to be carried out empirically in our case because of the specificity of the satisfactory criterion to meet. The performance of these two methods are anyway satisfactory and similar. They both prove that only a piecewise approximation of the smoothed wind must be taken into account in further landing simulation processes. The choice between the two methods can only be made according to the difficulty to arrange the parameters and time calculation.

Acknowledgments

This research was carried on in collaboration between AEROSPATIALE-MATRA-AIRBUS, the ANRT and the Probability and Statistics laboratory of Paul Sabatier University in Toulouse, FRANCE. We are grateful to C.BES, J. IRVOAS from AEROSPATIALE-MATRA-AIRBUS for their suggestions. We thank D.DENISON for his help in the use of his algorithm and the referees for their helpful comments.

References

- Denison, D. G. T., Mallick B. K. and Smith, A. F. M. (1998). Automatic Bayesian curve-fitting. *J. R. Statistical Soc. B* **60**, 333-350.
- Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711-732.
- Hall, P. and Titterington, D.M. (1992). Edge-Preserving and Peak-Preserving Smoothing. *Technometrics* **34**, 429-440.
- Loader, C. R. (1996). Change point estimation using nonparametric regression. *Annals of Statistics* **24**, 1667-1678.
- Müller, H. G. and Song, K. S. (1997). Two-stage change-point estimators in smooth regression models. *Statistics and Probability Letters* **34**, 323-335.
- S-Plus (1997). *S-PLUS 4 Guide to Statistics*. Mathsoft.
- Wu, J.S. and Chu, C. K. (1993). Kernel-type estimators of jump points and values of a regression function. *Annals of Statistics* **21**, 1545-1566.

Received March 25, 2005; accepted December 6, 2005.

Philippe Besse Laboratoire de Statistique et Probabilites
UMR CNRS 583
Universite Paul Sabatier
31062 Toulouse cedex 9
France
philippe.besse@math.ups-tlse.fr

Nathalie Raimbault - Airbus France
31707 Blagnac cedec
France
nathalie.raimbault@airbus.com