doi.org/10.26398/IJAS.0029-003

ANAEROBIC THRESHOLD AND KERNEL ESTIMATORS

Marco Maioli[1]

Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia - Modena, Italy

Abstract *The transition threshold from aerobic to anaerobic metabolism of athletes practising endurance sports is identified with the deflection point in the graph of cardiac frequency versus velocity. A parametric statistical background was provided (Calderoni et al., 1990) by assuming a transition from linear to parabolic regression. By interpreting the threshold as a jump-point for the first derivative of a continuous regression curve, the use of one-sided kernel estimators can be a useful explorative treatment. Here Müller's theory of such kernels is summarized and simulations are discussed for this type of application.*

Keywords: *Anaerobic threshold, Transition point, Kernel estimators.*

1. INTRODUCTION

The search of the deflection point of cardiac frequency of athletes as the velocity increases has a considerable interest since such a point coincides with the anaerobic threshold, i.e. the metabolic level at which the increasing of lactic acid becomes perceptible during a protracted intense physical effort. Indeed, beyond this threshold an anaerobic mechanism is added to the (aerobic) mechanism of production of ATP (adenosinetriphosphate), which is an important energy compound in metabolism. Hence, for an athlete in an endurance sport, the anaerobic threshold corresponds to the greatest speed which can be kept without a rapid exhaustion of energies via an aerobic process. Now, the direct methods to single out this threshold are certainly accurate, but they involve a laborious procedure. An indirect method was first proposed (Conconi et al., 1982) and is based on the almost linear increase of the cardiac frequency in a wide initial velocity interval (which coincides with the aerobic zone), which is followed by a deflection when the lactic acid starts to accumulate (anaerobic zone). Afterwards (Calderoni at al., 1990) a methodological framework was given to that procedure by a model which consists of two polynomial regression curves, where the transition from linear to parabolic regression corresponds to the unknown change-point from aerobic to anaerobic

[1] e-mail: marco.maioli@unimore.it

(all the moments vanish except the ones of order ν and k). The sequence of bandwidths $b = b(n)$ is required to satisfy

$$b \to 0, \ \ nb^{2\nu+1} \to \infty, \ \ \limsup_{n\to\infty} \ nb^{2(k+\nu)+1} < \infty. \tag{5}$$

Thus a choice of $b(n)$ allows a corresponding balance of accuracy and precision as $n \to \infty$, by virtue of asymptotic behaviours (Gasser and Müller, 1979) of the type

$$E[\hat{g}_n(t) - g(t)] = O[b(n)^k], \quad E[(\hat{g}_n(t) - g(t))^2] = \sigma^2 O(\frac{1}{nb(n)}). \tag{6}$$

In practice, for each t, b determines the interval of data $[t-b, t+b]$ involved in fitting.

A development of that theory (Müller, 1992) concerns change-points, i.e. discontinuity points which are detectable in the regression curve or in its derivatives and describes sudden and lasting changes appearing in various contexts of experimental sciences.

Starting from the above model, let us suppose that a change-point τ, $0 < \tau < 1$, exists for the derivative $g^{(\nu)}$, so that

$$g^{(\nu)}(t) = f^{(\nu)}(t) + \Delta_\nu \cdot I_{[\tau,1]}(t), \ \ \Delta_\nu > 0, \ \ 0 \le t \le 1, \tag{7}$$

where $f \in C^k([0,1])$ and

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Δ_ν is the (negative or positive) jump in the discontinuity point τ for the ν-th derivative of g. If $\Delta_\nu = 0$ g is continuous in τ.

By defining the right and left limits of $g^{(\nu)}$

$$g_+^{(\nu)}(\tau) = \lim_{t\to\tau^+} g^{(\nu)}(t), \quad g_-^{(\nu)}(\tau) = \lim_{t\to\tau^-} g^{(\nu)}(t) \tag{9}$$

and assuming $g^{(\nu)}(\tau) = g_+^{(\nu)}(\tau)$, one can notice that

$$\Delta_\nu = g_+^{(\nu)}(\tau) - g_-^{(\nu)}(\tau). \tag{10}$$

The idea of the method consists of introducing one-sided kernel estimators of the regression function, and looking for the point where the maximal difference of the (all the moments vanish except the ones of order ν and k). The sequence of bandwidths $b = b(n)$ is required to satisfy

$$b \to 0, \;\; nb^{2\nu+1} \to \infty, \;\; \limsup_{n\to\infty} \; nb^{2(k+\nu)+1} < \infty. \tag{5}$$

Thus a choice of $b(n)$ allows a corresponding balance of accuracy and precision as $n \to \infty$, by virtue of asymptotic behaviours (Gasser and Müller, 1979) of the type

$$E[\hat{g}_n(t) - g(t)] = O[b(n)^k], \quad E[(\hat{g}_n(t) - g(t))^2] = \sigma^2 O(\frac{1}{nb(n)}). \tag{6}$$

In practice, for each t, b determines the interval of data $[t-b, t+b]$ involved in fitting.

A development of that theory (Müller, 1992) concerns change-points, i.e. discontinuity points which are detectable in the regression curve or in its derivatives and describes sudden and lasting changes appearing in various contexts of experimental sciences.

Starting from the above model, let us suppose that a change-point τ, $0 < \tau < 1$, exists for the derivative $g^{(\nu)}$, so that

$$g^{(\nu)}(t) = f^{(\nu)}(t) + \Delta_\nu \cdot I_{[\tau,1]}(t), \;\; \Delta_\nu > 0, \;\; 0 \le t \le 1, \tag{7}$$

where $f \in C^k([0,1])$ and

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Δ_ν is the (negative or positive) jump in the discontinuity point τ for the ν-th derivative of g. If $\Delta_\nu = 0$ g is continuous in τ.

By defining the right and left limits of $g^{(\nu)}$

$$g_+^{(\nu)}(\tau) = \lim_{t\to\tau^+} g^{(\nu)}(t), \quad g_-^{(\nu)}(\tau) = \lim_{t\to\tau^-} g^{(\nu)}(t) \tag{9}$$

and assuming $g^{(\nu)}(\tau) = g_+^{(\nu)}(\tau)$, one can notice that

$$\Delta_\nu = g_+^{(\nu)}(\tau) - g_-^{(\nu)}(\tau). \tag{10}$$

The idea of the method consists of introducing one-sided kernel estimators of the regression function, and looking for the point where the maximal difference of the two estimates is attained. Let $K_+^{(\nu)}$ and $K_-^{(\nu)}$ be one-sided kernel functions, that is with compact supports $[-1,0]$ and $[0,1]$ respectively. Let us define the one-sided kernel estimators of the $\nu-$th derivative $g^{(\nu)}$:

$$\hat{g}_{\pm}^{(\nu)} = \frac{1}{b^{\nu+1}} \sum_{i=1}^{n} y_i \int_{s_{i-1}}^{s_i} K_{\pm}^{(\nu)}(\frac{t-u}{b})du$$

Then the inference on the change-points is based on the difference between right and left estimates:

$$\hat{\Delta}^{(\nu)}(t) = \hat{g}_+^{(\nu)}(t) - \hat{g}_-^{(\nu)}(t). \tag{11}$$

The point $\hat{\tau}$ on which the maximum of such difference (considered in absolute value) is attained, will be a suitable estimate of the change-point. Actually a closed subinterval $Q \subset (0,1)$ containing τ must be chosen to exclude the boundary, where a more complex situation occurs (Müller, 1991). So the estimator is defined

$$\hat{\tau} = inf\{\rho \in Q : \ |\hat{\Delta}^{(\nu)}(\rho)| = \sup_{x \in Q} |\hat{\Delta}^{(\nu)}(x)| \ \}. \tag{12}$$

As for the jump of the ν-th derivative, its estimator is:

$$\hat{\Delta}^{(\nu)}(\hat{\tau}) = \hat{g}_+^{(\nu)}(\hat{\tau}) - \hat{g}_-^{(\nu)}(\hat{\tau}). \tag{13}$$

The following hypotheses for one-sided kernels are assumed:

$$K_+^{(\nu)} \in C^{\mu}([-1,0]) \cap \mathscr{K}_{\nu,k}([-1,0]), \tag{14}$$

$$K_+^{(\nu+j)}(-1) = K_+^{(\nu+j)}(0) = 0, \ \ 0 \leq j < \mu, \tag{15}$$

$$K_-^{(\nu)} \in C^{\mu}([0,1]) \cap \mathscr{K}_{\nu,k}([0,1]), \tag{16}$$

$$K_-^{(\nu+j)}(0) = K_-^{(\nu+j)}(1) = 0, \ \ 0 \leq j < \mu. \tag{17}$$

Here, denoting either $[0,1]$ or $[-1,0]$ by $[a,b]$, $\mathscr{K}_{j,k}([a,b])$ is the set of continuous functions f such that $supp(f) = [a,b]$ and

$$\int f(x)x^i dx = \left\{ \begin{array}{ll} (-1)^j j!, & i = j \\ 0, & 0 \leq i < k,\ i \neq \nu \\ \neq 0 & i = k \end{array} \right\}. \tag{18}$$

The parameter $\mu = 0, 1, 2, \ldots$ expresses regularity, i.e. the order of differentiability inherited by the estimator $\hat{g}^{(\nu)}$ from the kernel.

If $\nu = 0$ and $k = 2$, from the hypotheses it follows that

$$\int_{-1}^{0} K_+(x)dx = 1,\ \int_{-1}^{0} K_+(x)x\ dx = 0,\ \int_{-1}^{0} K_+(x)x^2\ dx \neq 0, \tag{19}$$

and

$$\int_{0}^{1} K_-(x)dx = 1,\ \int_{0}^{1} K_-(x)x\ dx = 0,\ \int_{0}^{1} K_-(x)x^2\ dx \neq 0. \tag{20}$$

If the first derivative of a regression curve has to be estimated ($\nu = 1$), fixing $k = 3$, the kernel functions must satisfy:

$$\int K'_\pm(x)dx = 0, \quad \int K'_\pm(x)xdx = -1,$$

$$\int K'_\pm(x)x^2 dx = 0, \quad \int K'_\pm(x)x^3 dx \neq 0, \tag{21}$$

and the kernel and its derivative must vanish at the endpoints. Thus, in the role of the one-sided kernels, polynomials can be chosen, which vanish either in 0 and 1 or in -1 and 0, with suitable coefficients; besides they can be multiplied by other polynomials to increase the order of regularity. Notice that K_+ acts on the right of t, while K_- acts on the left by definition of the convolution.

To write down examples of left one-sided kernels satisfying the above hypotheses, with $\nu = 0$ and $k = 2$, let us take the following $\mu-$dependent polynomials vanishing in 0 and 1:

$$K_{-,\mu}(x) = x^{\mu}(1-x)^{\mu}(\alpha_\circ + \alpha_1 x), \quad x \in [0,1]. \tag{22}$$

By the prescription on left one-sided kernels, the moment of order 0 is 1 and the moment of order 1 is 0. These two conditions give a system of two equations by which the coefficients $\alpha_\circ$ and α_1 are completely determined after fixing μ. Setting

$\mu = 0$ we obtain the kernel function $K_-(x) = 2(2-3x)$; setting $\mu = 1$ we obtain $K_-(x) = 12x(1-x)(3-5x)$, while setting $\mu = 2$ we get $K_-(x) = 60x^2(1-x)^2(4-7x)$. This last kernel, when derived, can be used to estimate the first derivative ($\nu = 1$) of the regression curve by taking its derivative and choosing $\mu = 1$ and $k = 3$.

The right one-sided kernels can be recovered from the left ones by remarking that

$$K_+^{(\nu)}(x) = (-1)^\nu K_-^{(\nu)}(-x) \tag{23}$$

as it follows from the preceding assumptions.

3. THE SEARCH OF DEFLECTION POINTS OF CARDIAC FREQUENCY AS A FUNCTION OF VELOCITY

The search of the deflection points of cardiac frequency, as the velocity increases, can be typical for many problems. Simulated data are discussed relative to a typical Conconi's test.

The athlete, after an usual warming of $15-30$ minutes, has to increase his speed without stopping. The increase must be gradually carried out, by predetermined fractions: then the achieved velocity must be kept constant during the rest of the fraction. The measurements are taken in this final part of the fraction.

The regression curve is assumed to be continuous and the change-point has to be detected as a discontinuity of the first derivative.

By definition the estimator kernel for the first derivative is

$$\hat{g}'_\pm(t) = \sum_{i=1}^{n} \frac{y_i}{b^2} \int_{s_{i-1}}^{s_i} K'_\pm\left(\frac{t-u}{b}\right) du, \tag{24}$$

By the change of variable $v = (t-u)/b$, whence $du = -b\,dv$, one gets

$$\hat{g}'_\pm(t) = \sum_{i=1}^{n} \frac{y_i}{b^2} \int_{\frac{t-s_{i-1}}{b}}^{\frac{t-s_i}{b}} (-b) \cdot K'_\pm(v) dv = \sum_{i=1}^{n} \frac{y_i}{b} \int_{\frac{t-s_i}{b}}^{\frac{t-s_{i-1}}{b}} K'_\pm(v) dv. \tag{25}$$

The left one-sided kernel estimator is calculated as a sum of the products of the data y_i times the integral of the kernel function considered in the interval

$$[0,1] \cap \left(\frac{t-s_i}{b}, \frac{t-s_{i-1}}{b}\right). \tag{26}$$

Indeed the left one-sided kernel function has compact support in $[0,1]$; an analogous subinterval of $[-1,0]$ supports integration of the right one-sided kernel.

The algorithm makes use of the fact that two intervals $[a,b]$, $[c,d]$ have non-empty intersection if and only if $a \leq d$ and $b \geq c$. The derivatives of the kernels

$$K_{-}(x) = 60x^2(1-x)^2(4-7x), \quad K_{+}(x) = K_{-}(-x), \quad \mu = 1 \tag{27}$$

$$H_{-}(x) = 280x^3(1-x)^3(5-9x), \quad H_{+}(x) = H_{-}(-x), \quad \mu = 2 \tag{28}$$

are chosen as kernel functions for $\hat{g}'_{\pm}$.

Now 100 values of velocities are taken from 9 to 20, and corresponding values of cardiac frequency (about from 148 to 188) are normally generated with different variances around the curve

$$y = \begin{cases} 86 + 6.5t, \quad 9 \leq t \leq 14.29 \\ 27.519 + 17.067t - 0.453t^2, \quad 14.29 \leq t \leq 20. \end{cases} \tag{29}$$

One of our aims is to find a suitable range of values of the bandwdith b: some values of b are employed, and the validation is provided by the simulations for which the resulting estimate $\hat{\tau}$ is near enough to 14.29.

The change point is detected as a minimum of the jump size $\hat{\Delta}'(x) = \hat{g}'_{+}(x) - \hat{g}'_{-}(x)$, since the "defaillance" of cardiac frequency implies $g'_{+}(x) < g'_{-}(x)$.

In figures 1, 2, 3 an example of generated data (with standard deviation $\sigma = 0.4$) and the graphs of $\hat{\Delta}'(x)$ (coming from kernel estimators (27) and (28), respectively, with bandwidth $b = 2$) appear. The resulting estimate was $\hat{\tau} = 14.44$.

An asymptotic $(1-\alpha)$-confidence interval for τ (Müller, 1992, 3.11 at page 744) is provided by

$$\tau = \hat{\tau} \pm b \left[\Phi^{-1}(1 - \frac{\alpha}{2})(\mu+\nu)! \frac{\hat{\sigma}}{|\hat{\Delta}^{(\nu)}(\hat{\tau})| K_{-}^{(\mu+\nu)}(0)} \right]^{\frac{1}{\mu+\nu}}$$

$$\times \left[2 \int K_{-}^{(\nu+1)}(v)^2 dv / (nb^{2\nu+1}) \right]^{\frac{1}{2\mu+2\nu}} \tag{30}$$

Here

$$\hat{\Delta}^{(\nu)}(\hat{\tau}) = \hat{g}_{+}^{(\nu)}(\hat{\tau}) - \hat{g}_{-}^{(\nu)}(\hat{\tau}) \tag{31}$$

is the jump size in the $\nu-$th derivative, while an estimate of the variance σ^2 can be given as

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} (-\frac{1}{\sqrt{2}} y_i + \frac{1}{\sqrt{2}} y_{i+1})^2 \tag{32}$$

This last expression is a particular case of formula (3.10) in the above theory (Müller, 1992, p. 744),

$$\hat{\sigma}^2 = \frac{1}{n-(m_1+m_2)} \sum_{i=m_1+1}^{n-m_2} \left(\sum_{j=-m_1}^{m_2} \omega_j y_{j+i} \right)^2,$$

$$\sum \omega_j = 0, \ \sum \omega_j^2 = 1, \ \ m_1, m_2 \geq 0, \ m_1 + m_2 \geq 1.$$

We have chosen $m_1 = 0, m_2 = 1$ and the weights ω_j so that $\omega_{0,1} = \mp \frac{1}{\sqrt{2}}$.

By computing in the above simulations, we found:

$$\hat{\tau} = 14.44, \ \hat{\Delta} = -3.73, \ \ \hat{\sigma} = 0.57,$$

$$\int K''_-(v)^2 \, dv = 25920, \ \ K_-^{\mu+\nu}(0) = K''_-(0) = 480 \ (\mu = 1, \nu = 1, \ b = 2)$$

and a 95% confidence interval is $[\tau_1, \tau_2] = [12.44, \ 16.44]$.

By varying the bandwidth b, in similar simulations with $\sigma = 0.4$, the resulting estimates $\hat{\tau}$, $\hat{\Delta}$ were:

b	$\hat{\tau}$	$\hat{\Delta}$	$\hat{\sigma}$
0.5	13.66	-16.49	0.57
1	15.77	-5.16	0.54
1.3	14.44	-3.67	0.56
1.5	14.33	-4.00	0.56
2	14.88	-2.66	0.55
2.5	14.77	-2.81	0.58
2.6	14.44	-4.69	0.53
3	10.44	-95.11	0.55
3.5	18.44	-154.30	0.51

Therefore, if $\sigma \leq 0.4$, the acceptable bandwidths are found from about 1.3 to 2.6.

Increasing the standard deviation σ in simulations the estimates can go away from the expected $\tau = 14.29$:

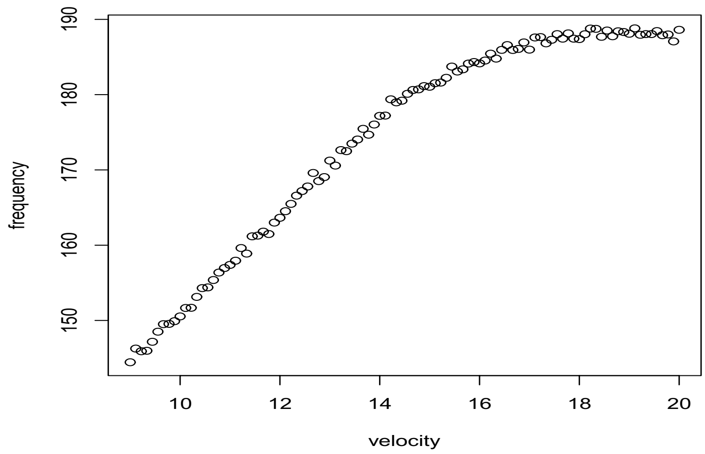

Figure 1: Simulation of line - parabola transition, with change point $t \simeq 14.29$

b	σ	$\hat{\tau}$	$\hat{\Delta}$	$\hat{\sigma}$
2	0.5	14.55	-4.60	0.70
2	1	17.00	-3.94	0.97
2	2	14.22	-7.23	2.18
2	3	17.11	-14.49	2.89
2.5	4	16.44	-15.45	4.16
2.5	5	16.77	-9.55	5.47
2.5	6	16.22	-11.01	6.50

As expected, a limitation of the method is the variability of the assigned data.

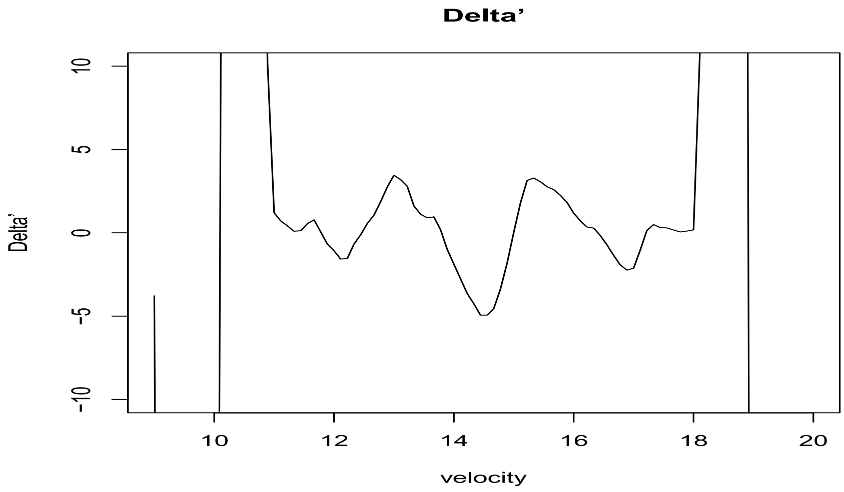

Figure 2: $\hat{\Delta}'$ **with** $K\_(x) = 60x^2(1 - x)^2(4 - 7x)$**,** μ**= 1.**

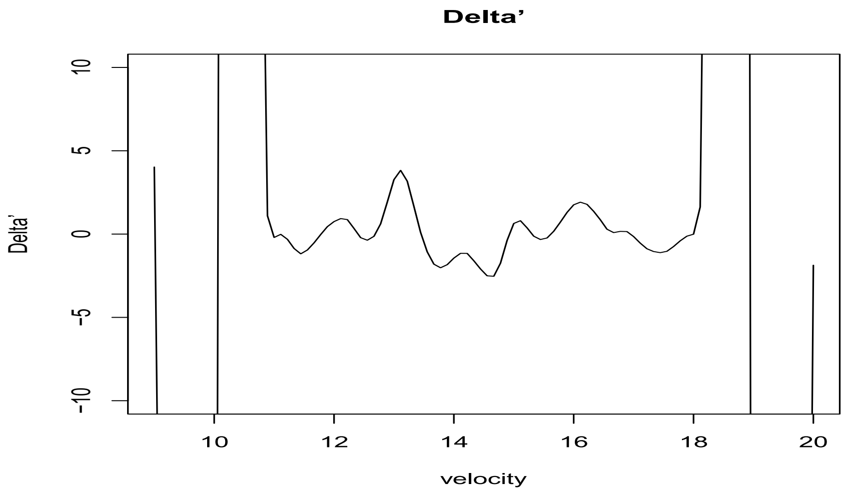

Figure 3: $\hat{\Delta}'$ **with** $K\_(x) = 280x^3(1 - x)^3(5 - 9x)$**,** μ**= 2.**

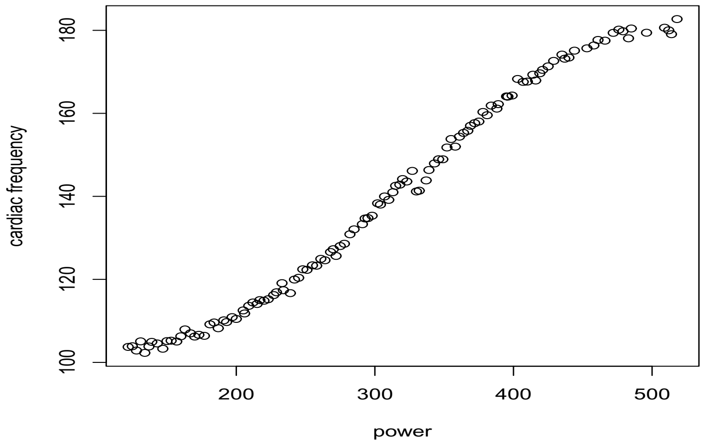

Figure 4: Simulation of transition from convex to concave parabola with discontinuity point $t \simeq 327$

Another transition which can appear in these problems is from convex to concave parabola, with evidence of discontinuity of $g(t)$ itself, rather than of $g'(t)$ (it can be verified when a parabolic regression restricted to the aerobic regime turns out to give better p-values than a simple line regression).

We consider as t an indirect measure of velocity, namely in watt, and a random perturbation of the curve

$$y = \begin{cases} 114.4 - 0.191t + 0.00088t^2, & 122 \leq t \leq 327 \\ -109.3 + 1.121t - 0.00108t^2, & 328 \leq t \leq 518 \end{cases} \tag{33}$$

with independent normal errors. An example of generated data is in Fig. 4. The search of a discontinuity point is done with kernels

$$K_-(x) = 12x(1-x)(3-5x),\ 0 < x < 1;\ K_+(x) = -12x(1+x)(3+5x),\ -1 < x < 0.$$

For $\sigma = 2$, $\sigma = 3$, acceptable estimates of the transition point are obtained with bandwidth about from 25 to 50:

b	$\hat{\tau}$	$\hat{\Delta}$
20	330	-0.573
30	327	-0.263
40	327	- 0.213
50	327	-0.175
55	327	- 0.166
60	173	-0.190

while further values of standard deviation in simulation make the estimates less satisfactory (as expected): for example, when $\sigma = 4$, a result is $\hat{\tau} = 291$, $\hat{\Delta} = -0.294$.

REFERENCES

Calderoni, G., Rogantin, M.P. and Piastra, G. (1990). Ricerca del punto di rottura di due regressioni per la determinazione del punto di deflessione della frequenza cardiaca al variare della velocità. In *Statistica Applicata.* 2: 207-223.

Conconi, F., Ferrari, M., Ziglio, P.G., Droghetti, P. and Codeca, L. (1982). Determination of the anaerobic threshold by a noninvasive field test in runners. In *Journal of Applied Physiology.* 52: 869-973

Gasser, T. and Müller, H.G. (1979). Kernel estimation of regression functions. In A. Dold and B. Eckmann, editors, *Smoothing techniques for curve estimation, Lecture Notes in Mathematics.* 757: 23-68.

Müller, H.G. (1991). Smooth optimum kernel estimators near endpoints. In *Biometrika*. 78: 521-530

Müller, H.G.(1992). Change points in non-parametric regression analysis. In *Annals of Statistics.* 20: 737-761.

Tosi, D. (1994). Nuclei stimatori unilaterali e modificati per la regressione non parametrica. *Thesis,* University of Modena.