

# Water quality assessment by means of HFNI valvometry and high-frequency data modeling

Mohamedou Sow · Gilles Durrieu · Laurent Briollais ·  
Pierre Ciret · Jean-Charles Massabuau

Received: 19 July 2010 / Accepted: 21 December 2010 / Published online: 13 January 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** The high-frequency measurements of valve activity in bivalves (e.g., valvometry) over a long period of time and in various environmental conditions allow a very accurate study of their behaviors as well as a global analysis of possible perturbations due to the environment. Valvometry uses the bivalve's ability to close its shell when exposed to a contaminant or other abnormal environmental conditions as an alarm to indicate possible perturbations in the environment. The modeling of such high-frequency serial valvometry data is statistically challenging, and here, a

nonparametric approach based on kernel estimation is proposed. This method has the advantage of summarizing complex data into a simple density profile obtained from each animal at every 24-h period to ultimately make inference about time effect and external conditions on this profile. The statistical properties of the estimator are presented. Through an application to a sample of 16 oysters living in the Bay of Arcachon (France), we demonstrate that this method can be used to first estimate the normal biological rhythms of permanently immersed oysters and second to detect perturbations of these rhythms due to changes in their environment. We anticipate that this approach could have an important contribution to the survey of aquatic systems.

**Keywords** Environmental monitoring · Nonparametric regression · High-frequency data · Valvometry · Water quality · Bivalves

---

M. Sow · G. Durrieu (✉) · P. Ciret · J.-C. Massabuau  
CNRS UMR 5805 EPOC, University of Bordeaux 1,  
Place du Dr Peyneau, 33120 Arcachon, France  
e-mail: gilles.durrieu@univ-ubs.fr,  
g.durrieu@epoc.u-bordeaux1.fr

M. Sow · L. Briollais  
Division of Epidemiology and Biostatistics,  
Samuel Lunenfeld Research Institute,  
Mount Sinai Hospital,  
Toronto, ON, M5G 1X5, Canada

*Present Address:*  
G. Durrieu  
LMAM, University of South Brittany,  
Campus de Tohannic, 56017 Vannes, France

## Introduction

Protection of the aquatic environment is a top priority for marine managers, policy makers, and the general public. Due to an increasing interest in the health of aquatic systems, there is a compelling need to using remote online sensors to instantly and widely distribute information on a daily basis, specifically in marine environment

(Kröger and Law 2005). Among these sensors, bio-indicators are increasingly used and are highly effective in their ability to reveal the presence of trace amounts (very low concentrations) of contaminants through accumulation in tissues (see the US Mussel Watch, <http://ccma.nos.noaa.gov/about/coast/nsandt/welcome.html>). The ability of mollusk bivalves to taste their environment is one of the possible ways to monitor water quality. Monitoring their opening and closing activities over time is yet another way to evaluate the behavior of the bivalves in reaction to their environmental exposure. The interest in investigating the bivalve's activities by recording their valve movements (e.g., valvometry) is not recent. Marceau was certainly pioneer in this method and performed the first recordings of molluscan valve movements (Marceau 1909). This technique has been explored in ecotoxicology for more than 20 years, first under laboratory conditions. The basic idea of valvometry is to use the bivalve's ability to close its shell when exposed to a contaminant as an alarm signal (Doherty et al. 1987; Shumway and Cucci 1987; Byrne et al. 1990; Kramer and Botterweg 1991; Borcharding 1992; Sluyts et al. 1996; Curtis et al. 2000; Kadar et al. 2001; Nagai et al. 2006). Nowadays, valvometers are available on the market and use the principle of electromagnetic induction (Sloff et al. 1983; Jenner et al. 1989) such as the Mossel Monitor (Kramer et al. 1989) or the Dreissena Monitor (Borcharding 1992). There has been a clear research interest in recent years to measuring the bivalves' behaviors directly in real conditions (Riisgard et al. 2006; Robson et al. 2007; Garcia-March et al. 2008). This is also our case where our technical innovations include the development of an online data system (no delayed analysis and graph production; near-real time data production), a remote control, a strong resistance under bad sea conditions, and no limit of distance between the sampling sites and the laboratory. Thus, our newly developed technique and analytical tools are able to monitor and analyze quantitatively the bivalve's behaviors very accurately. More specifically, we want to describe statistically the normal bivalve's activities under laboratory and/or field conditions in order to provide a reference from different seas, biotopes, and seasons

and then to further determine abnormal patterns. For this purpose, since 2003, we have developed a new approach using a noninvasive valvometric technique (high frequency, noninvasive (HFNI) valvometry) coupled with (a) the integration of the bivalve's respiratory physiology and ethology and (b) an in-depth statistical analysis that can handle very large amount of data, a key to fast online analysis of molluscan bivalve's behavior at a large scale. Three original key features have been introduced (Tran et al. 2003). First, a new valvometric system was built using lightweight electrodes (approximately 100 mg) linked by thin flexible wires to high-performance electronic units. This system allows the bivalves to be studied in their natural positions with minimal experimental constraints. Second, variations in the probability of spontaneous closures over the animals' biological rhythms were first taken into account under laboratory conditions by working during specific periods, for example, when the probability of spontaneous closing events was the lowest (Tran et al. 2003) and by controlling the animals' respiratory physiology and behavior (Tran et al. 2000, 2001; Massabuau 2001). Third, a statistical approach was developed, and it demonstrated its ability to use bivalves' closure response as a method for detecting trace amounts of metals (Tran et al. 2003, 2004, 2007; Fournier et al. 2004). It is noteworthy that from our results and observations, Liao et al. (2005, 2006) and Jou et al. (2009) developed mathematical models that describe the interaction between spontaneous valve daily rhythms and various valve response characteristics. In this study, we established two field stations, the first in the Bay of Arcachon, France, in February 2006 and the second one in the Havannah channel, New Caledonia, in August 2007. These field stations use the light electromagnetic and electronic principles as described in Chambon et al. (2007). Two major challenges we faced were (a) one data point is generated every 0.1 s at each site, 24 h a day, to optimize the description of the animals' behaviors and (b) the data sampled have a high variability. The objective of this paper is to apply and develop appropriate statistical tools to successfully summarize and analyze the high-frequency valvometry data. Although our statistical modeling was conducted

on the oysters from the Bay of Arcachon, our statistical approach was also validated on the giant clams from the site in New Caledonia. Here, we illustrate some applications using valvometry data on a group of 16 oysters from the Bay of Arcachon with valve activity measured over a period of 183 days. On a daily basis, the data and their statistical analysis are presented graphically and made available to public in the MolluSCAN Eye website [http://www.domino.u-bordeaux.fr/molluscan\\_eye](http://www.domino.u-bordeaux.fr/molluscan_eye).

## Materials and methods

### Experimental sites and animal species

The monitoring site is located in the Bay of Arcachon, France, at the Eyrac pier (latitude 44°40 N, longitude 1°10 W). Since February 2006, groups of 16 Pacific oysters, *Crassostrea gigas*, measuring from 8 to 10 cm in length are permanently installed on this site. These oysters are all from the same age group (1.5 years old) and come from the same local supplier. They also all grew in the Bay of Arcachon. They were placed in a traditional oyster farmer bag on a bench in the bottom and permanently immersed at 3 to 7 m deep in the water, depending on the tide activity. We present here a study covering the period from 1st April 2006 to 31 September 2006.

### Measurements of valve activity

The first electronic principle of monitoring was described by Tran et al. (2003) and further modified by Chambon et al. in 2007. The modification was necessary for field application in the ocean and was purely mechanical. It involved the hardening of the monitoring system to allow its constant immersion in seawater even in strong water current conditions and even during storms. Some information about these specific aspects can be found on [http://www.domino.u-bordeaux.fr/molluscan\\_eye](http://www.domino.u-bordeaux.fr/molluscan_eye). The main challenge was to ensure the complete autonomy of the system without in situ human intervention for at least one full year. In brief, each animal is equipped with two light coils (sensors),  $\approx 53$  mg each (unembedded), fixed on the edge of each valve. These coils mea-

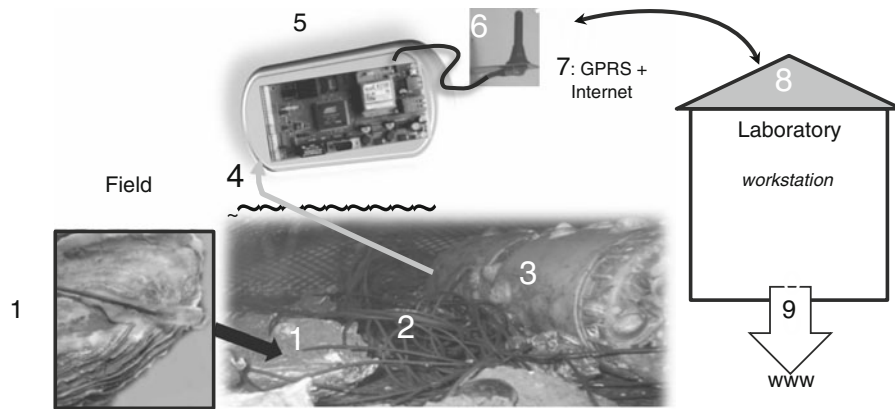
sured  $2.5 \times 2.5 \times 2$  mm and were coated with a resin-sealing before fixation on the valves. One of the coils emits a high-frequency, sinusoidal signal which is received by the other coil. In each group of animals, one measurement was received every 0.1 s (10 Hz) on one of the 16 animals. This means that each animal's behavior was measured every 1.6 s. Every day, a data set with 864,000 pairs of values (one distance value, one stamped time value) was generated representing a total of 1,728,000 data points. Each oyster's behavior was described by a total of 108,000 measurements per day. The strength of the electric field produced between the two coils decreases with the distance according to the transformation  $1/D$ , where  $D$  is the distance between the point of measurement and the center of the transmitting coil. The measured signal ( $D_{mv}$  in millivolts) is converted into distance in millimeters ( $D_{mm}$ ) using the following calibration model

$$\hat{D}_{mv} = \frac{151}{(\hat{D}_{mm})^{0.35}} - 1.48. \quad (1)$$

The model parameters were estimated using least-squares estimation, and the model was very significant ( $R^2 = 0.99$ ,  $P$  value  $< 0.001$ ).

### Data collection and data transmission

On the shore (i.e., by a pier or lighthouse), a second electronic unit handles the data acquisition and the programmed emission. The unit is equipped with a GSM/GPRS modem and uses Linux operating system for driving the first control unit, managing the data storage with a time stamp, accessing the internet, and transferring the data. A self-developed software module runs on mobile phone technology. After each 24-h period (or any other programmed period of time), the data collected are transmitted to a central workstation server located at the Marine Biological Station in Arcachon, France. The valve activity data are stored in a central database in a merged data set. Access to this database is provided via an Intranet or Internet connection (Fig. 1).



**Fig. 1** Synoptic representation of the system, from field to laboratory. 1 Oyster equipped with two electrodes, 2 connecting cables from oysters to first-level electronic card, 3 first-level electronic card in a waterproof case (immersed), 4 electrical connection between the first and the

second electronic card, 5 second electronic card (emersed), 6 GPRS antennae, 7 GPRS and internet connection, 8 Marine Biological Station of Arcachon, 9 daily update on Internet

#### Tide and diurnal/nocturnal data

Tidal data were collected including a measurement of the height of the water column every hour and the times of the low and high tides, provided by the hydrographic and oceanographic service of the marine (<http://www.shom.fr/>). To characterize the diurnal/nocturnal rhythms of the bivalve, we estimated the time of sunrise and sunset at the sampling site.

#### Estimators and models

We propose a kernel-based approach for the modeling and a statistical treatment of the high-frequency valvometry data. The main novelty is the use of a nonparametric regression, which is particularly advantageous in the context of environmental studies that often exhibit high variability. The principle of our approach is first to produce a density estimation of each animal's opening and closing valve activities over a 24-h period, second to make inference about these individual profiles using kernel regression, and third to correlate them with some measured environmental data. A graphical display is also provided to represent the daily and annual profiles of each animal, which can be easily read and interpreted by specialists and nonspecialists. Another possible choice of estimator is a local polynomial estimator

(Fan and Gijbels 1996), but Einmahl and Mason (2000) and Quian and Mammitzsch (2000) emphasized the advantage of kernel regression and the use of the Nadaraya–Watson approach for the variance estimator (Nadaraya 1964; Watson 1964). One of the strengths of this estimator consists in its automatic adaptation to designs where the local polynomial estimator may not perform reliably overall since its variance may fail to exist. Also, the Nadaraya–Watson estimator retains some optimality properties as exposed in Härdle and Marron (1985). Moreover, we recently applied successfully kernel estimation procedure to high-frequency data in genetics as described in Briollais et al. (2007) and Durrieu and Briollais (2009).

#### *Estimation of the probability density function of valve activities*

To determine a general profile of valve activities during a given period of time (24 h in general), we estimated the density of probability of the theoretical states of opening and closing which are unknown in our sample. There are a number of statistical problems for which it is necessary to obtain estimates of the density. For instance, it is of great interest to reveal features of the data such as skewness and multimodality which make possible the identification of different behaviors in

the whole population (a mixture of several distributions and clusters). A histogram-type estimator is the most commonly used estimator, but it could be strongly affected by the number of classes chosen. Therefore, we decided to use a nonparametric kernel density estimator that behaves better statistically (consistency, asymptotic normality, and numerical behavior). A kernel estimator of a function  $f$  is defined by Silverman (1986), Wand and Jones (1995):

$$\hat{f}_{h_n}(t) = \frac{1}{n h_n} \sum_{i=1}^n K\left(\frac{t - T_i}{h_n}\right) \tag{2}$$

where  $n$  is the sample size and  $h_n$  is the bandwidth (the smoothing parameter) satisfying (C1)  $h_n \rightarrow 0$  and  $n h_n \rightarrow \infty$ , as  $n \rightarrow \infty$ . The kernel  $K$  is a symmetric function satisfying (C2)  $\int_{-\infty}^{+\infty} K(t) dt = 1$ ,  $\int_{-\infty}^{+\infty} t K(t) dt = 0$ ,  $\int_{-\infty}^{+\infty} |t| K(t) dt < +\infty$ , and the distribution function  $f$  is twice differentiable (C3). We introduce the notation  $\int_{-\infty}^{+\infty} K^2(t) dt = \tau^2$  to simplify the following formulas, and here, we considered the Gaussian kernel

$$K(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right). \tag{3}$$

Under conditions (C1)–(C3), the bias and variance of the estimator  $\hat{f}_n(\cdot)$  can be approximated when  $n \rightarrow \infty$  by:

$$\text{bias}(\hat{f}_{h_n}(t)) = \frac{h_n^2}{2} f''(t) \int t^2 K(t) dt + o(h_n^2), \tag{4}$$

and

$$\text{var}(\hat{f}_{h_n}(t)) = \frac{1}{n h_n} f(t) \tau^2 + o\left(\frac{1}{n h_n}\right). \tag{5}$$

Under conditions (C1)–(C3), the bias tends to zero and the kernel density estimator  $\hat{f}_{h_n}$  is a consistent estimate of  $f$  asymptotically normal distributed as  $n \rightarrow \infty$ :

$$\sqrt{n h_n} (\hat{f}_{h_n}(t) - f(t)) \xrightarrow{D} N(0, \tau^2 f(t)). \tag{6}$$

Using Eq. 6, we derive the  $(1 - \alpha)\%$  confidence interval of  $f$  when  $n \rightarrow \infty$

$$\left[ \hat{f}_{h_n}(t) - z_{1-\alpha/2} \sqrt{\frac{\hat{f}_{h_n}(t) \tau^2}{n h_n}}, \hat{f}_{h_n}(t) + z_{1-\alpha/2} \sqrt{\frac{\hat{f}_{h_n}(t) \tau^2}{n h_n}} \right] \tag{7}$$

where  $z_{1-\alpha/2}$  is the  $(1 - \alpha/2)$ -quantile of the standard normal distribution.

### Nonparametric kernel regression model

The relationship between the distance between the two electrodes ( $Y$ ) and the time of the measurement ( $T$ ) is modeled using a nonparametric regression model given by:

$$Y_i = m(T_i) + \epsilon_i, \quad \text{for } i = 1, \dots, n \tag{8}$$

where  $n, m$ , and  $\epsilon$  denote the sample size (total number of couples of values), the regression function to be estimated, and the model error term, respectively. The stochastic distribution  $f$  of  $\epsilon$  is typically unknown and is unlikely to follow any familiar distribution such as a normal distribution and is independent of  $T$ . We therefore decided to use nonparametric statistics. Here, we consider the case of a fixed equidistant design. Also, the  $T_i$ 's have a uniform known design density  $f_T$ . However, the limiting bias and variance depend on unknown quantities which have to be estimated consistently in order to construct the asymptotic confidence intervals. The variation  $\epsilon_i$  is a random variable with a mean value equal to 0 and the stochastic distribution  $f$  of these random variations enable us to characterize the variation of the random variable  $Y$  around  $m$ :

$$m(t) = E(Y/T = t) = \frac{\int y f(y, t) dy}{f(t)}. \tag{9}$$

The regression function  $m$  thus depends on the joint density which is unknown. The joint density

was estimated using a multiplicative kernel and calculations show that

$$\int y \hat{f}(y, t) dy = \frac{1}{n} \sum_{i=1}^n K\left(\frac{t - T_i}{h_n}\right) Y_i. \tag{10}$$

Using the kernel estimator of the joint densities function, we deduced the regression estimator of  $m$  given by

$$\hat{m}_{h_n}(t) = \begin{cases} \frac{1}{n} \frac{\sum_{i=1}^n K\left(\frac{t - T_i}{h_n}\right) Y_i}{f_T(t)} & \text{if } f_T(t) \neq 0, \\ \frac{1}{n} \sum_{i=1}^n Y_i & \text{otherwise.} \end{cases} \tag{11}$$

Under conditions (C1)–(C3), the bias and variance of the estimator  $\hat{m}_{h_n}$  are given when  $n \rightarrow \infty$  by:

$$\text{bias}(\hat{m}_{h_n}(t)) = \frac{1}{2} h_n^2 m''(t) \int t^2 K(t) dt + o(h_n^2), \tag{12}$$

and

$$\text{Var}(\hat{m}_{h_n}(t)) = \frac{1}{n h_n} \sigma^2(t) \tau^2 + o\left(\frac{1}{n h_n}\right), \tag{13}$$

where  $\sigma^2(t)$  is the conditional variance  $\text{Var}(Y/T = t)$ . We chose  $h_n = n^{-\eta}$  and replaced  $\sigma^2(t)$  by its estimator

$$\hat{\sigma}^2(t) = \frac{1}{n f(t)} \sum_{i=1}^n K\left(\frac{t - T_i}{h_n}\right) (Y_i - \hat{m}_{h_n}(t))^2. \tag{14}$$

We added a new condition (C4) stating that  $m$  is twice differentiable,  $E(Y^2) < \infty$  and  $\eta \in [1/5, 1[$ . Under conditions (C1)–(C4),  $\hat{m}_{h_n}(t)$  is a consistent estimate of  $m(t)$  and is asymptotically normally distributed (Härdle 1990) when  $n \rightarrow \infty$ :

$$\sqrt{n h_n} (\hat{m}_{h_n}(t) - m(t)) \xrightarrow{D} N(B, \sigma^2(t) \tau^2), \tag{15}$$

where  $B$  denotes the bias of  $\hat{m}_{h_n}(t)$ . Our analyses show that the bias is negligible compared to the variance estimator (see “[Checking the model behavior](#)”). Then we can set the bias of  $\hat{m}_{h_n}(t)$  equal to zero. Equation 15 allows to compute an

$(1 - \alpha)\%$  confidence interval for  $m(t)$ , and it is given by

$$\left[ \hat{m}_{h_n}(t) - z_{1-\alpha/2} \sqrt{\frac{\sigma^2(t) \tau^2}{n h_n}}, \hat{m}_{h_n}(t) + z_{1-\alpha/2} \sqrt{\frac{\sigma^2 \tau^2}{n h_n}} \right] \tag{16}$$

where  $z_{1-\alpha/2}$  is the  $(1 - \alpha/2)$ -quantile of the standard normal distribution. From the expression of  $\hat{m}_{h_n}(t)$  (Eq. 11), we can estimate by derivation with respect to time the speed of opening and closing by using the derivative over time:

$$\frac{d\hat{m}_{h_n}(t)}{dt} = \frac{1}{n h_n} \frac{\sum_i K'\left(\frac{t - T_i}{h_n}\right) Y_i}{f_T(t)}. \tag{17}$$

### Bandwidth selection

Selecting a proper bandwidth parameter  $h_n$  is a critical step in estimating the density and the regression functions. Although in practice one can choose the bandwidth subjectively, this can lead to inaccurate estimation of the density and its variance. The choice of  $h_n$  is much more important for the behavior of the estimators than the choice of the kernel function  $K$ . Small values of  $h_n$  make the estimate look wiggly and show wrong features, whereas large values of  $h_n$  will lead to an estimate which is too smooth in the sense that is heavily biased and may not reveal important structural features, a bimodality for example. To optimize  $h_n$  for the density, we used a plug-in procedure developed by Sheather and Jones (1991). For the regression function estimator, we used a cross-validation (CV) criterion which leads to the score function

$$\text{CV}(h_n) = \sum_{i=1}^n (Y_i - \hat{m}_{(-i)}(T_i; h_n))^2 \tag{18}$$

with respect to  $h_n$  where  $\hat{m}_{(-i)}(T_i; h_n)$  is the regression function estimator without considering  $(T_i, Y_i)$ . This bandwidth is asymptotically optimal and yields the same speed of convergence than other techniques (Härdle 1991).

All our analyses were performed on a DELL workstation using both Bash Linux (Fedora 12) and mathematical codes written in *R* (Venables and Ripley 1999, <http://www.R-project.org/>).

## Results and discussion

### Analysis of the oysters' behaviors

The kernel density estimator (Eq. 2) provides an estimation of the distribution and probability of closing and opening states and is computed each day of our study period. Three examples are given in Fig. 2. When an oyster has several sequences of opening and closing states during the day, the distribution function is bimodal. In Fig. 2a, the density has two peaks, the first one corresponding to a closing state and the second one to an opening state. The probability of the closing state for this particular oyster is equal to 0.252, and the corresponding 95% confidence interval (Eq. 7) is [0.196, 0.323], so the total duration of the closing state during the studied day is 6.048 h and the corresponding 95% confidence interval is [4.704, 7.752] with  $\hat{h} = 0.42$ . In contrast, when the oyster remains completely closed or open during the day, the distribution function is unimodal. In Fig. 2b, it is observed that the oyster has spent more time open than closed (with an opening amplitude  $>4$  mm). The probability of closing for this oyster and this day is equal to 0.00208, and the corresponding 95% confidence interval is [0, 0.0206], so the total duration of closure is 0.0672 hours and the corresponding 95% confidence interval is [0, 0.494] with  $\hat{h} = 0.053$ . The closing probability for the oyster in Fig. 2c is 0.883, and the 95% confidence interval is [0.779, 0.986], so the total duration of closure is 21.16 h and the corresponding 95% confidence interval is [18.696, 23.664] with  $\hat{h} = 0.41$ . This means that this oyster spent more time closed than open during the day.

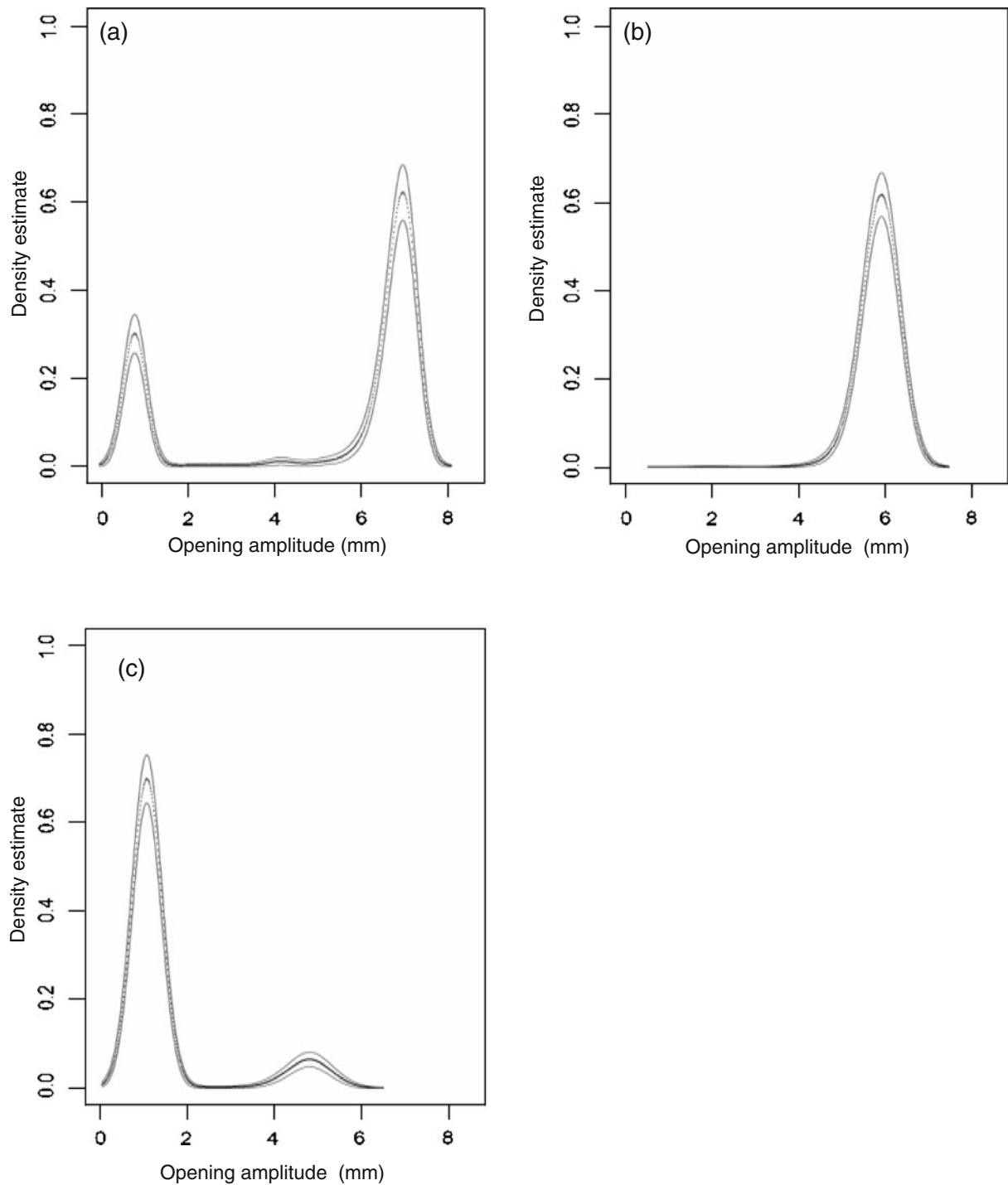
The application of the estimator in Eq. 11 gives an estimate of the function  $m$ , which models the opening amplitude as a function of time. Then, by using asymptotic normality and consistency properties of  $\hat{m}_{n_n}$ , we derived the following 95% confidence interval based on (Eq. 16). The bandwidth estimator has been determined by CV where the CV functions are convex and have a clear minimum. This bandwidth provides a good fit of our data (see Fig. 3).

Then, from the estimator in Eq. 11, we estimated the speed of each closing and opening

moves for each oyster (Eq. 17). In Fig. 4, we notice that the positive speeds correspond to the opening moves and the negative speeds to the closing moves of the valves. Table 1 gives the values corresponding to this activity. It shows that oyster number 1 (Fig. 3a) had three closing states and their total duration was 6.07 h. The second closing state is much quicker with a speed equals to 17.96 mm/s and opening speed equals to 2.4 mm/s. In Fig. 3b, we observe that the oyster had two micro-closures during the day that were shorter than 1 min. In Figs. 3c and 4c, this oyster has one opening move with speed equal to 7.93 mm/s and a total duration of closing state equal to 21.03 h. Therefore, the density estimation and the kernel regression both provide very useful information to establish the normal bivalve's activity for a given period of time and then can be used as a reference to determine abnormal responses to the environment (for example, due to water contamination), as illustrated below.

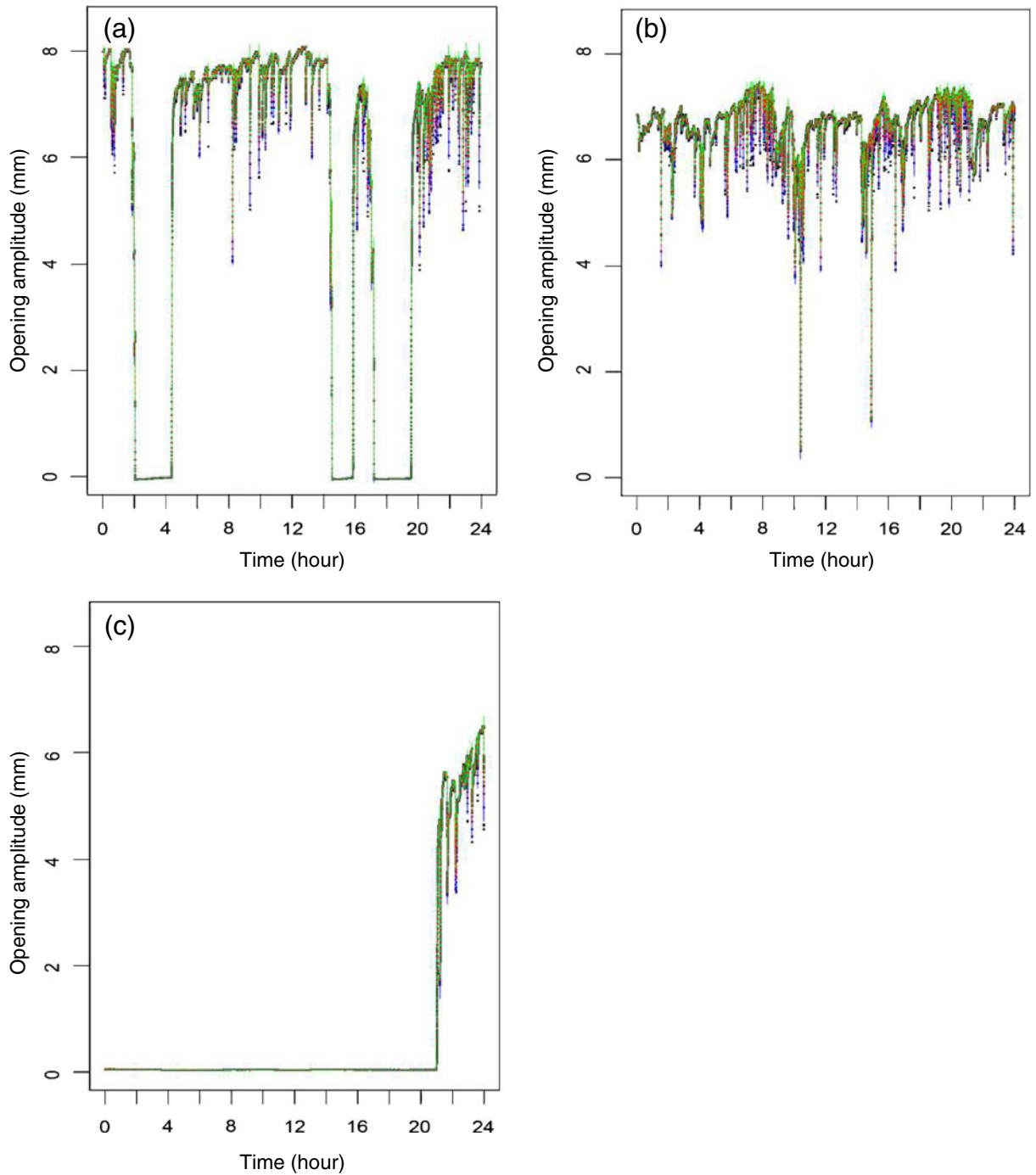
### Association of oysters' mean behavior with nycthemeral and tidal biological rhythms

Concerning the nycthemeral rhythms in the Bay of Arcachon, the difference between diurnal and nocturnal mean duration of closing was not significant in our study period (i.e., April to September 2006) with a Mann and Whitney  $U$  test giving a  $P$  value = 0.17. The mean duration of the closing state for an oyster was 5.06 h per day with a confidence interval of [1.98, 8.16] (mean  $\pm$  1 SE) decomposed into 2.63 h ([0.89, 4.39]) for the night and 2.43 h ([0.69, 4.17]) for the day. Using the model described in Eqs. 8 and 11, we can also model the biological rhythms of the permanently immersed bivalves. To extract a biological rhythm, we used a graphical representation of the period of closing estimates. The principle of this approach is represented in Fig. 5a, b, which shows a series of opening and closing states for each oyster during a 24-h period. The length of the closing period is proportional to the length of the solid line. Figure 5c shows this representation during a full month (May 2006), the  $x$ -axis represents the time in a 24-h time period, and the  $y$ -axis represents the number of days during this month. For each day, there are 16 lines which



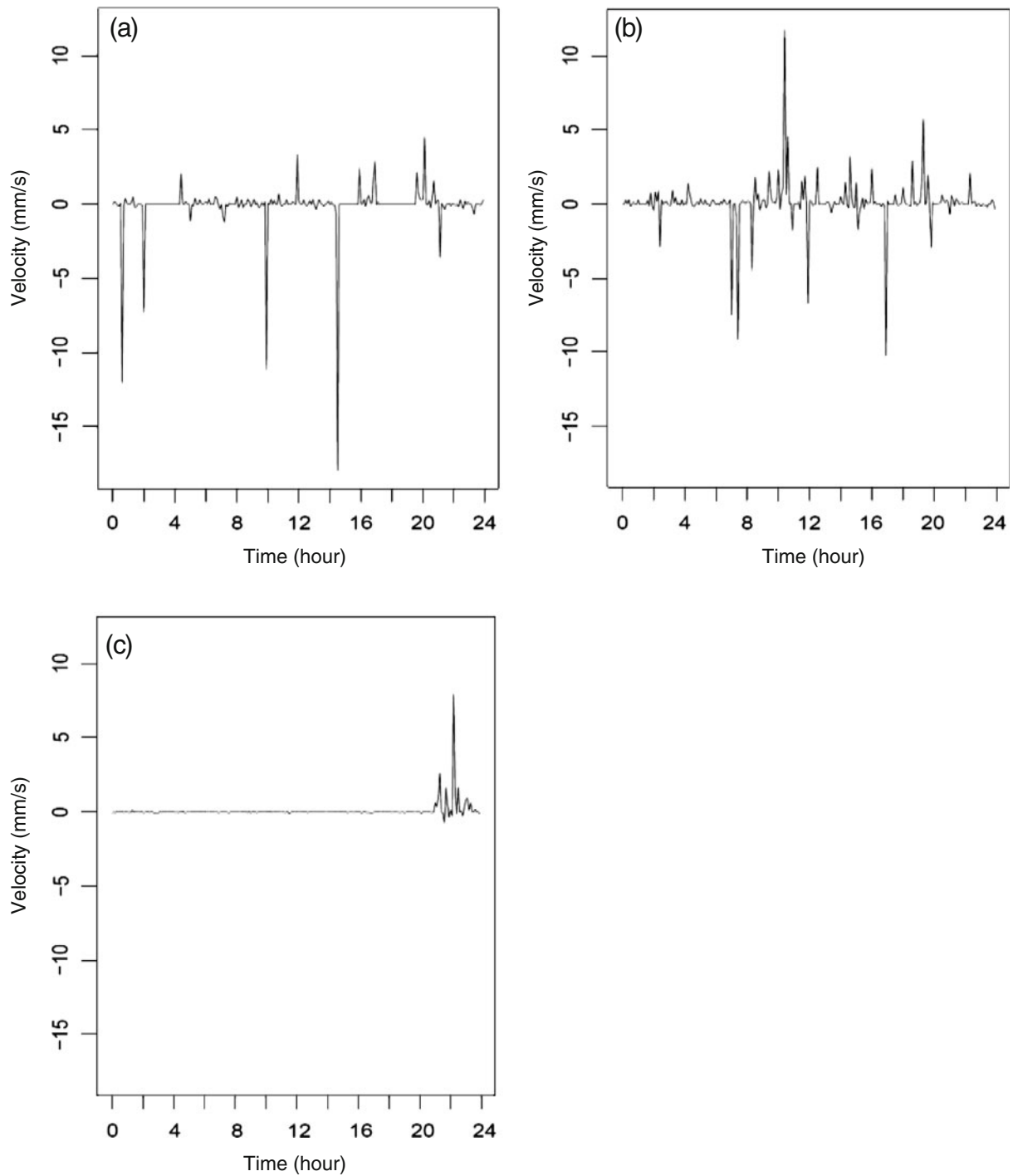
**Fig. 2** a–c Plot of kernel density estimator of opening amplitude (*dotted lines*) and 95% confidence bands for the underlying distribution function (*solid line*), for three different oysters (**a**), (**b**), and (**c**)





**Fig. 3 a–c** Relationship between the opening amplitude (millimeters) and the time of the experiment (in hour; represented by *points*), estimator of link regression estimator

(*dashed lines*), and 95% confidence interval of the link regression (*solid line*) for three different oysters (**a**), (**b**), and (**c**)



**Fig. 4** a–c Estimator of the opening and closing velocity (millimeters per second) according to time (in hour) for three different oysters (a), (b), and (c)

**Table 1** Example of estimated parameters for three different oysters (a), (b), and (c)

Oysters	Closing (h)	Opening (h)	Duration (h)	Sunrise (h)	Sunset (h)	Day/night	Tide amplitude (cm)	Closing velocity (mm/s)	Opening velocity (mm/s)
a	2.034	4.382	2.347	6.650	19.328	Night	424	7.29	2.07
a	14.522	15.869	1.346	6.650	19.328	Day	424	17.96	2.4
a	17.191	19.565	2.373	6.650	19.328	Day/night	424	0.011	2.15
b	10.388	10.403	0.014	6.650	19.328	Day	424	0.36	11.77
b	14.915	14.924	0.009	6.650	19.328	Day	424	0.24	1.44
c	0.001	21.026	21.026	6.650	19.328	Day/night	424	0.071	7.93

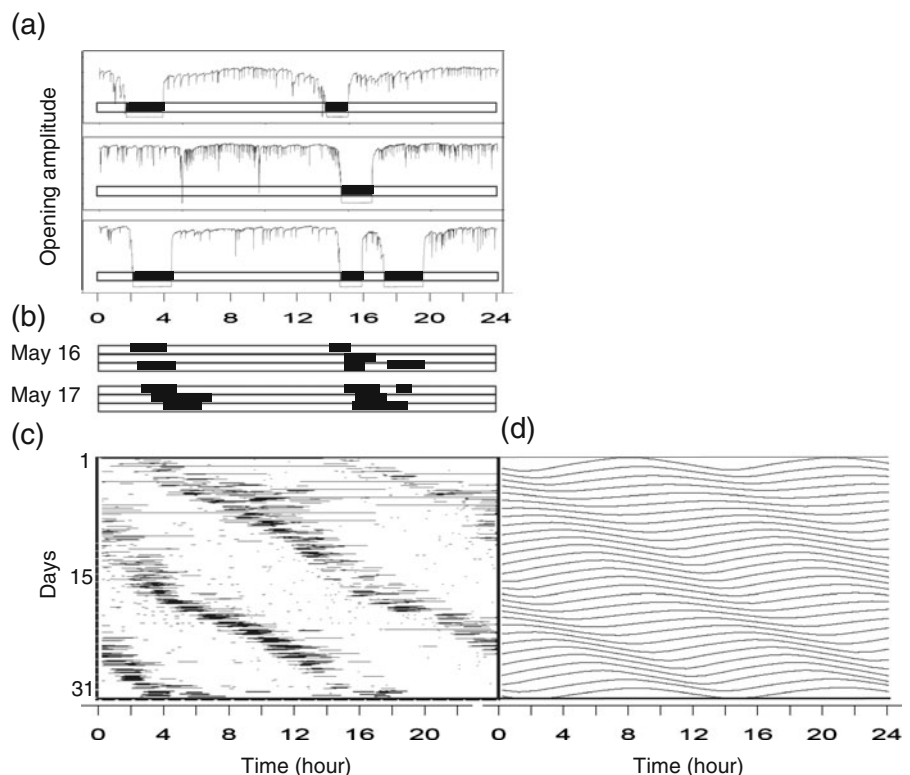
represent the periods of different time intervals of each oyster’s closing states (solid lines). The advantage of this representation is to give the time duration and number of closing and opening states for each of the 16 oysters for a given time period in one single graph. We also notice in Figs. 5c, d and 6 (April to August) that the closing and opening time periods do not occur randomly. In particular, this is very clear in Fig. 5d which depicts the change of the tide levels beginning at noon every day. More specifically, these figures show that the closing activity is highly correlated with the tidal amplitude and that the closing state

is synchronized with the low tide period. Based on these observations, we proposed a statistical method to evaluate the relationship between the oysters’ closing events and the tide amplitude. This is the first time that the bio-rhythms of the permanently immersed oysters is studied.

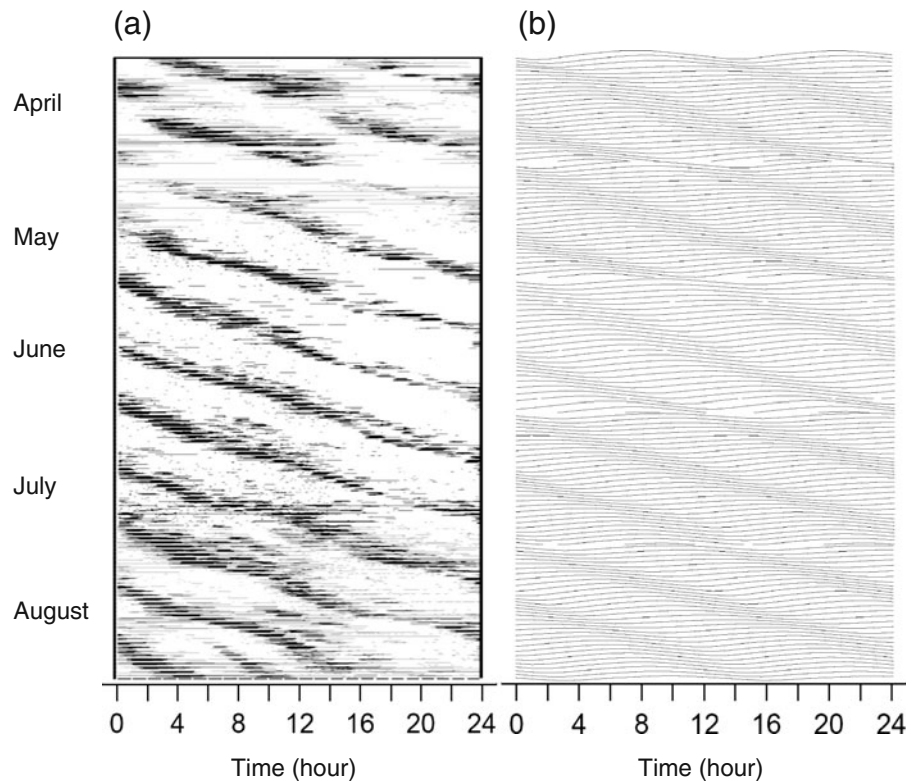
Correlation between tide amplitude and valve closing states

We determined the moment of the day where the oysters close their valves and correlated this event with the tide parameters in our entire study

**Fig. 5** Principle of biological rhythm related to tide (May 2006): **a** closing periods of three different oysters (filled and open bars denote respectively the closing and opening periods); **b** superposing of the three oysters closing/opening periods for 2 days (May 16 and 17); **c** superposing of the oyster closing/opening periods for 31 days (for each day, there are 16 lines which represent the periods of different time intervals of the closing/opening states for the 16 oysters); **d** representation of the tide levels per hour. The values of the x-axis represent the hours per day



**Fig. 6** Biological rhythm related to tide: **a** superposing of the oyster closing/opening periods from April 1 to August 31 following the same graph principle as in Fig. 5 (filled and open bars denote respectively the closing and opening periods), **b** the tide levels per hour from April to August 2006



period. For any particular day, there are 16 broken lines representing the time intervals of the closing states for the 16 oysters. Figure 5b represents for example this information for 3 oysters and during a 2-day period. Figure 5c compiles this information for the 16 oysters for a 1-month period. The boundaries of the closing state interval were calculated using a threshold value that takes into account the distance between the two valves. We denote the estimated vector of threshold values for the 16 oysters by

$$\hat{s} = (\hat{s}_1, \dots, \hat{s}_{16})$$

where for  $i = 1, \dots, 16$

$$\hat{s}_i = \min(\hat{Y}_i) + (\max(\hat{Y}_i) - \min(\hat{Y}_i)) \times 5\%. \quad (19)$$

Using the tide data (<http://www.shom.fr/>), we defined the interval  $C$  as the time period between two high tides occurring every lunar day. Using the vector of thresholds, we estimated for each day  $k$ , the vector  $\hat{\tau}$  corresponding to the time at which the first closing event occur for an oyster  $i$  in the interval  $C$ :

$$\hat{\tau}^k = (\hat{\tau}_1^k, \dots, \hat{\tau}_i^k, \dots, \hat{\tau}_{16}^k)$$

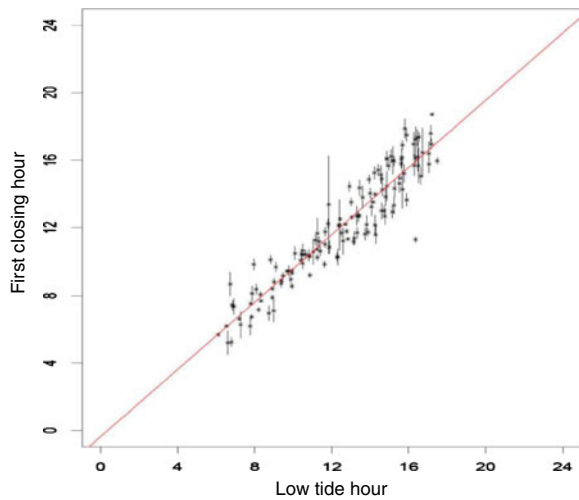
where for  $i = 1, \dots, 16$

$$(\hat{\tau}_i^k = \min \{t_j \in C^k | \hat{m}_{h_n}^k(t_j) \leq s_i^k\}. \quad (20)$$

Then, we estimated the mean of each oyster's time at the first closing period in the interval  $C$  denoted by  $(\bar{\tau}_i^k)$  and the standard error for the 16 oysters. In Fig. 7, we plotted  $(\bar{\tau}_i^k)$  as a function of time of the low tide for each day of the study period. We found evidence for a very significant linear relationship (least square estimators of the slope and the intercept being respectively equal to 0.995 and  $-0.36$ ;  $R^2 = 0.94$ ,  $P$  value  $< 0.001$ ). So, we have showed that on average, the oyster's first closing hour is significantly correlated with the time of low tide.

#### Detection of unusual behaviors

Animals' reactions to changes in their environment can be very variable. One way to assess unusual oysters' activity due to environmental changes is to estimate the proportion of oysters which close their valves at a moment of the day where they usually are open. For example, the

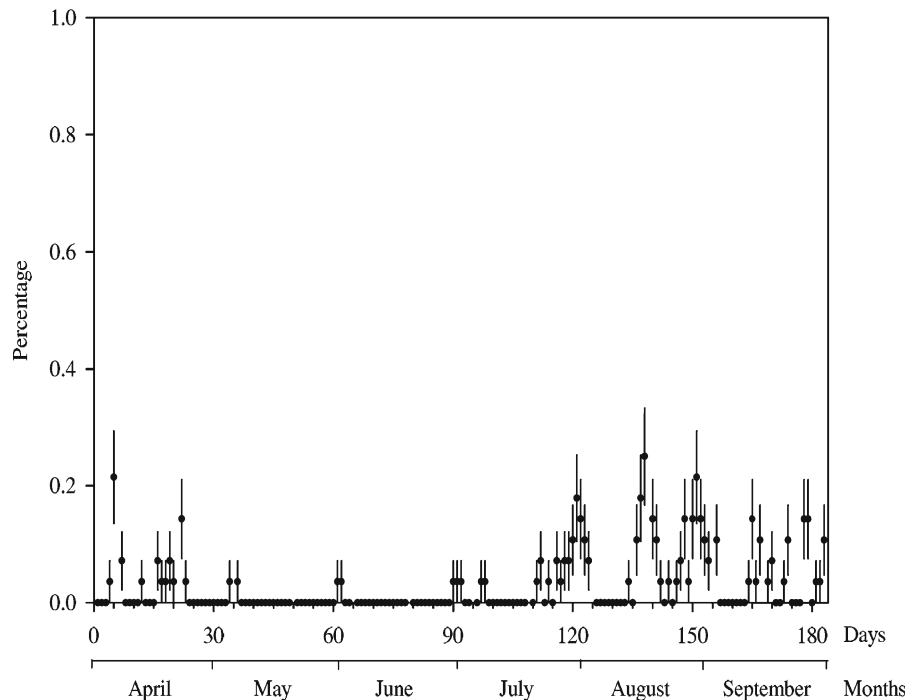


**Fig. 7** Relationship between the time of the low tide and the average of first closing period (16 oysters) for each day (April and September). Error bars denote the standard errors from the mean denoted by asterisk

proportion of oysters closing their valves during the high tide period  $\pm 1$  h should be very small. This proportion estimated for our study period was  $0.064 \pm 0.0068$  (mean  $\pm$  1SE) whereas in the low tide period  $\pm 1$  h, this proportion was equal

to  $0.45 \pm 0.015$  (mean  $\pm$  1SE), SE denoting the standard error. Figure 8 shows some results for the proportion of oysters closing their valves during the high tide period over our study period. We can easily detect the days when this rhythm was the most modified: April 5, 2006 ( $p = 41\%$ ), July 30, 2006 ( $p = 36\%$ ), August 16, 2006 ( $p = 51\%$ ), and August 29, 2006 ( $p = 42\%$ ). To better characterize the animals' behaviors during these periods, we defined two time periods, one corresponding to a reference activity (46 days between June 1 and July 15, 2006) and another one corresponding to the modified activity during the consumption ban (i.e., 46 days between August 1 and September 15, 2006; Ifremer 2006). The proportion of oysters closing their valves during the high tide interval  $\pm 1$  was significantly different between the two periods (Mann and Whitney  $U$  test,  $P$  value =  $5.508E - 06$ ). An estimation of the probability that this proportion be superior or inferior to a fixed threshold, for example 25%, was 0.048 and 0.29 for the normal and perturbed period, respectively. It suggests that changes in the bio-rhythms might be associated with field perturbations. The reasons for these changes remain to be determined although there is some

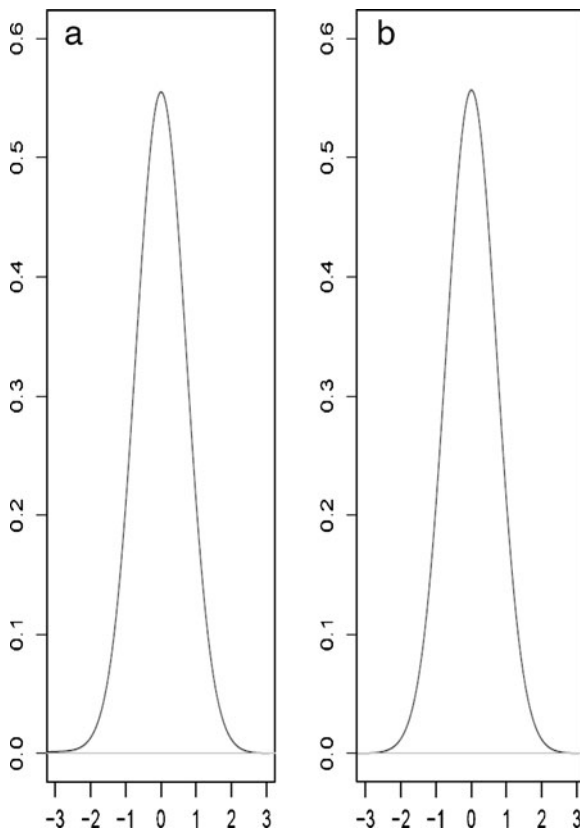
**Fig. 8** Percentage of oysters that closed their valves in the time interval high tide  $\pm 1$  h for each day (April to September 2006)



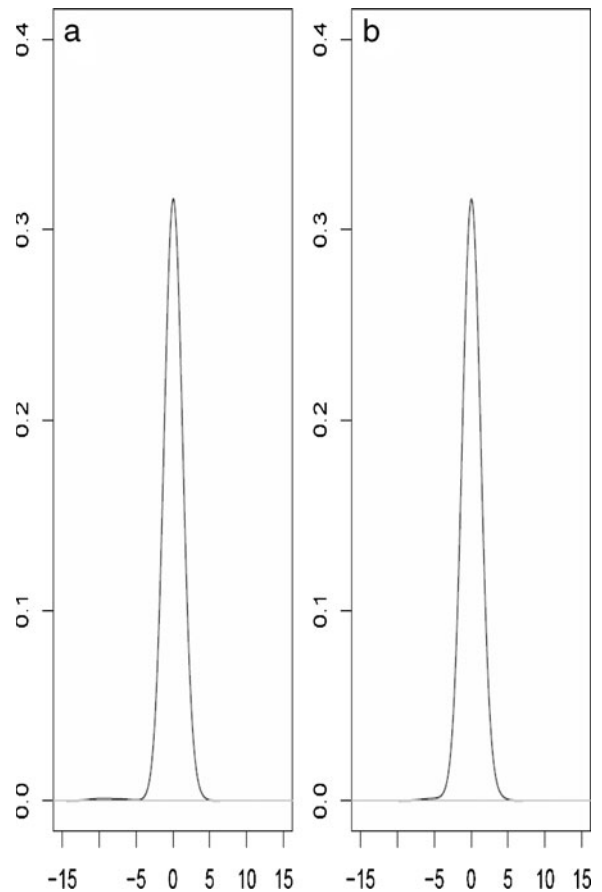
evidence that the oyster *C. gigas* can modified its normal behavior in presence of the harmful algae *Alexandrium minutum* (Tran et al. 2010). Similar results were also provided by Basti et al. (2009) in the manila clam *Ruditapes philippinarum*.

### Checking the model behavior

We numerically checked the asymptotic distribution of  $m(x)$  and its first derivative. We selected one oyster from our sample and chose two time points randomly between 0 and 24 h over a period of  $N = 450$  days. Figure 9 shows that the distribution of the bias of  $\hat{m}_{h_n}$  is normally distributed and centered around 0. The mean of the bias is 0.001 with a confidence interval of  $[-0.01, 0.014]$ , we employed a Mann and Whitney  $U$  test to compare the mean of the bias to 0, and the test yielded a



**Fig. 9** Convergence in distribution of estimator of the fixed design regression estimator on  $N = 450$  days at 1.5 h (a) and 10 h (b)



**Fig. 10** Convergence in distribution of fixed design derivative estimator on  $N = 450$  days at 1.5 h (a) and 10 h (b)

$P$  value = 0.887. Figure 10 also shows that the first derivative of  $\hat{m}_{h_n}$  is normally distributed.

### Conclusion

This work demonstrates the possibility to develop a general mathematical and statistical approach for analyzing high-frequency serial valvometry data in the context of an environmental study. The main interest of this work is to provide efficient analytical tools to synthesizing and analyzing the bivalves' behaviors directly exposed to their natural environment. We present an application to a group of 16 oysters living in the Bay of Arcachon, studied in 2006. Our statistical modeling based on density estimation and nonparametric kernel regression allows a significant reduction

in the calculations needed for the estimation of the oysters' activity over a long period of time and provides a very informative summary from a huge flow of data produced at high frequency. This methodology also enables us to make some statistical inference and correlate the serial valvometry data with serial environmental data in order to explain unusual behaviors of the animals studied. For example, we can identify several days in our study period at which the oysters exhibited abnormal activities, corresponding to periods of oyster banning from human consumption. We think that our statistical approach provides a very useful tool for measuring the oysters' behaviors and also assessing changes in these behaviors due to perturbations in their environment. Therefore, we anticipate that this approach could have an important contribution to the survey of aquatic systems.

**Acknowledgements** M. SOW was supported by a grant from the Aquitaine Region and University of Toronto. All experiments presented in this article complied with the current laws of France, where they were performed. The authors thank Lindsay Hiscop for the improvement of the narrative style of the manuscript.

## References

- Basti, L., Nagai, K., Shimasaki, Y., Oshima, Y., Honjo, T., & Segawa, S. (2009). Effects of the toxic dinoflagellate *Heterocapsa circularisquama* on the valve movement behaviour of the Manila clam *Ruditapes philippinarum*. *Aquaculture*, 291, 41–47.
- Borcherding, J. (1992). Another early warning system for the detection of toxic discharges in the aquatic environment based on valve movements of the freshwater mussel *Dreissena polymorpha*. In D. Neumann & H. A. Jenner (Eds.), *The zebra mussel Dreissena polymorpha: Ecology, biological monitoring and first applications in the water quality management* (pp. 127–146). New York: Gustav Fischer.
- Briollais, L., Durrieu, G., & Upathilake, R. (2007). Novel approach for genome scan meta-analysis of rheumatoid arthritis: A kernel-based estimation procedure. *BMC Genetics*, 1, 1–6.
- Byrne, R., Gnaiger, E., McMahon, R., & Dietz, T. (1990). Behavioral and metabolic responses to emersion and subsequent reimmersion in the freshwater bivalve *Corbicula fluminea*. *The Biological Bulletin*, 178, 251–259.
- Chambon, C., Legeay, A., Durrieu, G., Gonzalez, P., Ciret, P., & Massabuau, J. C. (2007). Influence of the parasite worm *Polydora sp.* on the behaviour of the oyster *Crassostrea gigas*: A study of the respiratory impact and associated oxidative stress. *Marine Biology*, 152, 329–338.
- Curtis, T. M., Williamson, R., & Depledge, M. H. (2000). Simultaneous, long-term monitoring of valve and cardiac activity in the blue mussel *Mytilus edulis* exposed to copper. *Marine Biology*, 136, 837–846.
- Doherty, F.G., Cherry, D.S., & Cairns, J. (1987). Valve closure responses of the Asiatic clam *Corbicula fluminea* exposed to cadmium and zinc. *Hydrobiologia*, 153, 159–167.
- Durrieu, G., & Briollais, L. (2009). Sequential design for microarray experiments. *Journal of the American Statistical Association*, 486, 650–660.
- Einmahl, U., & Mason, D.M. (2000). An empirical process approach to the uniform consistency of kernel type function estimators. *Journal of Theoretical Probability*, 13, 1–37.
- Fan J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman and Hall.
- Fournier, E., Tran, D., Denison, F., Massabuau, J. C., & Garnier-Laplace J. (2004). Valve closure response to uranium exposure for a freshwater bivalve *Corbicula fluminea*: Quantification of the influence of pH. *Environmental Toxicology and Chemistry*, 23, 1108–1111.
- Garcia-March, J. R., Sanchis Solsona, M. A., & Garcia-Carrascosa, A. M. (2008). Shell gaping behavior of *Pinna nobilis* L., 1758: Circadian and circalunar rhythms revealed by in situ monitoring. *Marine Biology*, 153, 689–698.
- Härdle, W. (1990). *Applied nonparametric regression*. Cambridge: Cambridge University Press.
- Härdle, W. (1991). *Smoothing techniques with implementation in S*. New York: Springer.
- Härdle, W., & Marron, J. S. (1985). Asymptotic nonequivalence of some bandwidth selectors in nonparametric regression. *Biometrika*, 72, 481–484.
- Ifremer (2006). *Bilan de la crise phycotoxines sur le Bassin d'Arcachon en 2006* (35 pp.). Direction des opérations/Laboratoires Environnement Ressources d'Arcachon.
- Jenner, H. A., Noppert, F., & Sikking, T. (1989). A new system for the detection of valve movement response of bivalves. *Kema Scientific and Technical Reports*, 7(2), 91–98.
- Jou, L. J., Chen, W. Y., & Liao, C. M. (2009). Online detection of waterborne bioavailable copper by valve daily rhythm in freshwater clam *Corbicula fluminea*. *Environmental Monitoring Assess.* doi:10.1007/s10661-008-0433-0.
- Kadar, E., Salanki, J., Jugdaohsingh, R., Powell, J. J., McCrohan, C. R., & White, K. N. (2001). Avoidance responses to aluminium in the freshwater bivalve *Anodonta cygnea*. *Aquatic Toxicology*, 55, 137–148.
- Kramer, K. J. M., & Botterweg, J. (1991). *Aquatic biological early warning systems: An overview*. In D. W. Jeffrey & B. Madden (Eds.), *Bioindicators and environmental management* (pp. 95–126). London: Academic.
- Kramer, K. J. M., Jenner, H. A., & De Zwart, D. (1989). The valve movement response of mussels: A tool in

- biological monitoring. *Hydrobiologia*, 188(189), 433–443.
- Kröger, S., & Law R. J. (2005). Sensing the sea. *TRENDS in Biotechnology*, 23, 250–256.
- Liao, C. M., Jou, L. J., & Chen, B. C. (2005). Risk-based approach to appraise valve closure in the clam *Corbicula fluminea* in response to waterborne metals. *Environmental Pollution*, 135, 41–52.
- Liao, C. M., Lin, C. M., Jou, L. J., & Chiang, K. C. (2006). Linking valve closure behavior and sodium transport mechanism in freshwater clam *Corbicula fluminea* in response to copper. *Environmental Pollution*, 147, 656–667.
- Marceau, F. (1909). Recherches sur la morphologie, et l'histologie, et la physiologie compares des muscles adducteurs des mollusques acphales. *Archives de Zoologie expérimentale et générale*, 2, 295–469.
- Massabuau, J. C. (2001). From a low blood- to low tissue-oxygenation strategy, an evolutionary theory. *Respiratory Physiology and Neurobiology*, 128, 249–262.
- Nadaraya, E.A. (1964). On estimating regression. *Theory of Probability and its Application*, 9, 141–142.
- Nagai, K., Honjo, T., Go, J., Yamashita, H., & Oh, S. J. (2006). Detecting the shellfish killer *Heterocapsa circularisquama* (Dinophyceae) by measuring the bivalve valve activity with Hall element sensor. *Aquaculture*, 255, 395–401.
- Quian, W., & Mammitzsch, V. (2000). Strong uniform convergence for the estimator of the regression function under  $\Phi$  mixing conditions. *Metrika*, 52, 45–61.
- Riisgard, H. U., Lassen, J., & Kittner, C. (2006). Valve-gape response times in mussels (*Mytilus edulis*): Effects of laboratory preceding feeding conditions and in situ tidally induced variation in phytoplankton biomass. *The Journal of Shellfish Research*, 25, 901–911.
- Robson, A., Wilson, R., & Garcia de Leaniz, C. (2007). Mussels flexing their muscles: A new method for quantifying bivalve behavior. *Marine Biology*, 151, 1195–1204.
- Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B*, 53, 683–690.
- Shumway, S. E., & Cucci, T. L. (1987). The effects of the toxic dinoflagellate *Protogonyaulax tamarensis* on the feeding and behavior of bivalve mollusks. *Aquatic Toxicology*, 10, 9–27.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.
- Sloff, W., De Zwart, D., & Marquenie, J. M. (1983). Detection limits of a biological monitoring system for chemical water pollution based on mussel activity. *Bulletin of Environmental Contamination and Toxicology*, 30(4), 400–405.
- Sluyts, H., Van Hoof, F., Cornet, A., & Paulussen, J. (1996). A dynamic new alarm system for use in biological early warning systems. *Environmental Toxicology and Chemistry*, 15, 1317–1323.
- Tran, D., Boudou, A., & Massabuau, J. C. (2000). Mechanism of oxygen consumption maintenance under varying levels of oxygenation in the freshwater clam *Corbicula fluminea*. *Canadian Journal of Zoology*, 78, 2027–2036.
- Tran, D., Boudou, A., & Massabuau, J. C. (2001). How water oxygenation level influences cadmium accumulation pattern in the Asiatic clam *Corbicula fluminea*: A laboratory and field study. *Environmental Toxicology and Chemistry*, 20(9), 2073–2080.
- Tran, D., Ciret, P., Ciutat, A., Durrieu, G., & Massabuau, J. C. (2003). Estimation of potential and limits of bivalve closure response to detect contaminants: application to cadmium. *Environmental Toxicology and Chemistry*, 22, 116–122.
- Tran, D., Fournier, E., Durrieu, G., & Massabuau, J. C. (2004). Copper detection in the Asiatic clam *Corbicula fluminea*: Optimum valve closure response. *Aquatic Toxicology*, 65(3), 317–327.
- Tran, D., Fournier, E., Durrieu, G., & Massabuau, J. C. (2007). *Corbicula fluminea* valve closure response as sensitivity threshold to inorganic mercury contamination. *Environmental Toxicology and Chemistry*, 26(7), 1545–1551.
- Tran, D., Haberkorn, H., Soudant, P., Ciret P., & Massabuau, J. C. (2010). Behavioral responses of *Crassostrea gigas* exposed to the harmful algae *Alexandrium minutum*. *Aquaculture*, 298, 338–345.
- Venables, W. N., & Ripley, B. D. (1999). *Modern applied statistics with S-PLUS*. New York: Springer.
- Wand, M. P., & Jones, M. C. (1995). *Kernel smoothing*. London: Chapman and Hall.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhya Series A*, 26, 359–372.