Nonparametric Regression Estimation in Models with Weak Error's Structure

RICARDO FRAIMAN AND GONZALO PÉREZ IRIBARREN

Centro de Matemática, Universidad de la República, Montevideo, Uruguay

Communicated by the Editors

In this paper we propose nonparametric estimates of the regression function and its derivative when it is only assumed a weak error's structure. We study their local and global asymptotic behaviour when we observe dependent trajectories.


1. INTRODUCTION

A classical statistical problem is the estimation of a regression function $g(x)$. Typically, $g(x)$ is assumed to belong to a parametric family $g(x, \theta)$ and parameters are estimated via least square methods. However, these parametric models are often motivated more in methodological methods than in an intrinsic structure of the real phenomena. That is one of the reasons why nonparametric techniques seem more natural in order to obtain evidence which allows us to understand the structure of the problem.

In this paper we consider the problem of estimating the function $g(x)$ that verifies a general nonparametric regression model,

$$ Y(x) = g(x) + e(x), \quad E(e(x)) = 0 $$

from a discrete set of observations of the process $Y(\cdot)$, defined on a probability space $(\Omega, \mathcal{A}, P)$, at the points $\{x_i/1 \leq i \leq n\}$.

This problem has been considered by several authors; Rosenblatt [19], Priestley and Chao [16], Benedetti [11], Clark [3], Gasser and Müller [7], Gasser et al. [8], Härdle and Luckhaus [11], and Georgiev [9, 10] are some references.

However, it is quite often required that the process $e(x)$ verifies...
\( E(e(x_i) e(x_j)) = 0 \) for \( x_i \neq x_j \), or to impose a more restrictive condition, namely, that the "errors" given by the sequence \( \{e(x_i), i \geq 1\} \) be independent. This case corresponds mainly to measurement errors, and it cannot reasonably be applied to other situations as, for example, the case of some growth curves models where the observable response of each individual may be better modeled as a sampling path \( Y(x, \omega), \omega \in \Omega \) of a process \( Y(x) \) with expected value \( g(x) \). Moreover, square mean continuity of the observable process \( Y(x) \) entails that the variables \( e(x) \) and \( e(x') \) are strongly correlated if \( x' \) is near to \( x \). For instance, in hydrology, many phenomena may be represented by a sequence of continuous response curves \( \{Y^{(j)}(t, \omega) | j \geq 1\} \), where \( t \) represents the time elapsed from the beginning of a certain year, \( j \) indicates the corresponding year, and \( e^{(j)}(t, \omega) \) is the measurement of the deviation from the annual mean curve \( g(t) \). In some biological phenomena as the growth of individual (or populations) \( Y^{(j)}(x, \omega) \) will be the growth curve of the \( j \)-individual, \( e^{(j)}(x, \omega) \) the measurement of the deviation from the mean growth \( g(x) \) of the response of the \( j \)-individual, and \( \{x_i | 1 \leq i \leq n\} \) the points where measurements are taken. Moreover, in practice sometimes the observed responses in different \( j \)-units are also correlated and may be represented by a sequence of responses curve \( \{Y^{(j)}(\cdot), j \geq 1\} \) with an intrinsic dependence structure, such as mixing conditions. This situation is clear in the first example. In the second example \( m \)-dependence may also appear if we are dealing with some individuals with the same progenitors.

We would like to add a comment on the concepts of "weak error's structure." In this context we mean by "weak error's structure" to be "far from independence." More precisely, we want to assume as little as possible about the internal correlation structure of the process although we will have no problem assuming some regularity of trajectories when necessary. We think that for these kind of applications even the independent increments hypothesis must be avoided if possible.

As in Hart and Wehrly [12], we consider a nonparametric approach to regression problems with repeated measurements, in which a random sample of \( m \) experimental units of a response variable is available at the points \( x_1, \ldots, x_n \), of a controlled variable. For this model Hart and Wehrly [12] studied the asymptotic square mean error of a kernel estimator, assuming that the \( e^{(j)}(x_i) \) are zero mean random variables satisfying \( \text{cov}(e^{(j)}(x_j), e^{(k)}(x_h)) = \sigma^2 \rho(x_j - x_h) \) for \( j = k \) and zero elsewhere (for a smooth correlation function \( \rho \)), and obtained results concerning the optimum choice of a bandwidth. (See also Möcks et al. [15] and Raz et al. [17]). For parametric models there is also an extensive literature on the analysis of repeated measurements. See, for instance, a bibliography by Diggle, Donnelly, and Kirby (CSIRO Division of Mathematics and Statistics, Report No. ACT 85/18, 1985).
We will denote by \( Y^{(j)}(x_i) \) the \( j \)th response at the point \( x_i \), where \( 1 \leq j \leq m \), \( 1 \leq i \leq n \), and \( m \) is the number of experimental units. In order to allow dependence we will assume that the sequence \( \{e^{(j)}(\cdot), j \geq 1\} \) is an (non-necessarily stationary) \( \alpha \)-mixing process.

In Section 2, we propose nonparametric estimates of the function \( g(x) \) (population mean response) and of its derivative \( g'(x) \) based on locally weighted averages. For the derivative, instead of the derivative of the estimate of \( g(x) \), i.e., the derivative of a "smooth" version of \( g(x) \), we consider a smooth version of a numerical derivative of \( g(x) \), that is, a locally weighted average of the increments rate of the observed process. In this way we include nearest neighbor weights (which are not differentiable) and we also obtain—in general—a non-degenerate asymptotic joint distribution of the estimators of \( g \) and \( g' \).

Section 3 is divided into three parts. In the first one we show the consistency of the proposed estimates; in the second one we obtain their asymptotic distribution at a given point \( x \) and the asymptotic finite dimensional distributions. In the third one we give the asymptotic distribution of the corresponding processes on \( C[0, 1] \) (the space of continuous functions on the interval \([0, 1]\)) related to each one of the considered estimates. In Section 4 we give a real data example where we compare the salinity of the Rio de la Plata (a great estuary between Uruguay and Argentina) in two 6-year periods 26 years apart.

Similar results for the regression curve (about the consistency and the asymptotic finite dimensional distributions) may be obtained with straightforward modifications if we consider the model

\[
Y(x) = g(x) + e(x) + e_1(x)
\]

and it is assumed that \( e_1(x) \) is a white noise independent of \( e(x) \). The white noise \( e_1(x) \) may correspond, for instance, to measurement errors if we assume that we do not observe continuous trajectories.

Finally, it is shown in the Appendix that the more usual weight functions, based on kernel and nearest neighbor methods, verify the required set of design assumptions.

2. SOME NONPARAMETRIC ESTIMATES FOR \( g(x) \) AND \( g'(x) \)

Let is consider the model

\[
Y(x) = g(x) + e(x),
\]

(2.1)
where \( \{e(x) : x \in [0, 1]\} \) is a zero mean stochastic process defined on a probability space \((\Omega, \mathcal{A}, P)\) which—in general—will take values on the space \(C[0, 1]\), and \( g : [0, 1] \to \mathbb{R} \) is a continuous function. \( Y(x) \) will represent the individual response (sampling path) which will be observable at a discrete set of points \( x_i = x_m, 1 \leq i \leq n \), belonging to the unit interval. These measurements will be denoted by \( Y^{(j)}(x_1), \ldots, Y^{(j)}(x_n), 1 \leq j \leq m, 1 \leq i \leq n \), where the index \( j \) will correspond to the \( j \)th response.

We are interested in estimating the value of the function \( g(\cdot) \) at a given \( x \in [0, 1] \) and its derivative \( g'(\cdot) \) (when this last one does exist).

The classic growth curve model takes just \( m = 1 \), but makes strong assumptions on the structure of the errors process \( e(x) \) such as the independence of \( e(x_1), \ldots, e(x_n) \). As noted in the introduction there are many situations where this is an unrealistic assumption. For instance, mean square continuous sampling paths imply a strong correlation between \( e(x) \) and \( e(x') \) when \( x' \) is near to \( x \). Our main interest is to avoid as far as possible any assumptions on the error's structure.

In order to allow some dependence between individuals we will establish the following assumptions:

**H1.** The sequence \( \{e^{(j)}(x), j \geq 1\} \) is an \( \alpha \)-mixing sequence on \((\Omega, \mathcal{A}, P)\) with the same distribution as \( e(x) \), i.e., Rosenblatt [18], there exists a nonincreasing sequence of nonnegative numbers \( \{\alpha(j), j \geq 1\} \) with \( \lim_{j \to \infty} \alpha(j) = 0 \) such that for any integer \( j \), \( |P(AB) - P(A)P(B)| \leq \alpha(j) \) for all \( k \geq 1, A \in M_k, B \in M_{k+1} \), where \( M_k \) is the \( \sigma \)-field generated by \( \{e^{(j)}(x)/u, j \leq k\} \).

A nonparametric estimate of \( g(x) \) can be obtained as a local average of the response variables \( \{Y^{(j)}(x_i), 1 \leq i \leq n, 1 \leq j \leq m\} \). More precisely, for each \( x \in [0, 1] \), let \( w_{ni}(x) = w_{ni}(x, x_1, \ldots, x_n), 1 \leq i \leq n \), be a measurable weight function verifying the following design assumption:

**H2.**

(i) \( 0 \leq w_{ni}(x) \) for \( 1 \leq i \leq n \), and \( n \in \mathbb{N} \).

(ii) \( \sum_{i=1}^{n} w_{ni}(x) = 1 \)

(iii) \( \lim_{n} \sup_{0 \leq x \leq 1} \sum_{i=1}^{n} w_{ni}(x) I_A = 0 \) for all \( \delta > 0 \), where \( A = A(n, x, \delta) = \{i/|x_i - x| > \delta, 1 \leq i \leq n\} \) and \( I_A \) denotes the indicator function of the set \( A \).

**Remark 2.1.** (a) A useful way of rewriting assumption H2 is by introducing an artificial sequence of discrete random variables \( W_n(x) \) with probability function \( P \) defined by \( P(W_n(x) = x_i) = w_{ni}(x) \). Then (iii) is the same as \( \lim_{n} \sup_{x} P(|W_n(x) - x| > \delta) = 0 \) for all \( \delta > 0 \).

(b) (i) and (ii) can be weakened as in Georgiev [9, 10], as will be shown in the Appendix.
The estimator of the function $g(x)$ is defined by

$$g_{n,m}(x) = m^{-1} \sum_{j=1}^{m} \sum_{i=1}^{n} w_{ni}(x) Y^{(j)}(x_i) = \sum_{i=1}^{n} w_{ni}(x) \bar{Y}(x_i), \quad (2.2)$$

where $\bar{Y}(\cdot)$ denotes the average process $m^{-1} \sum_{j=1}^{m} Y^{(j)}(\cdot)$.

In order to estimate the velocity $g'(x)$ instead of the derivative of $g_{n,m}(x)$ we will use

$$Dg_{n,m}(x) = m^{-1} \sum_{j=1}^{m} \sum_{i=1}^{n} w_{ni}(x) (Y^{(j)}(x_{i+1}) - Y^{(j)}(x_i))/(x_{i+1} - x_i)$$

$$= \sum_{i=1}^{n} w_{ni}(x) (\bar{Y}(x_{i+1}) - \bar{Y}(x_i))/(x_{i+1} - x_i), \quad (2.3)$$

where $x_{n+1} = 1$. That is, instead of a derivative of a smooth version of $g(x)$, we use a smooth version of a "numerical derivative" of $g(x)$.

### 3. Asymptotic Results

#### A. Consistency.

Let $\{Y(x) : x \in [0, 1]\}$ be defined as in (2.1), and let $g_{n,m}$ and $Dg_{n,m}$ be the estimates considered in (2.2) and (2.3), respectively.

**Lemma 3.1.** Under H2 we have that:

(i) $\lim_{n} \sup_{0 \leq x \leq 1} |E(g_{n,m}(x)) - g(x)| = 0$.

(ii) If in addition $g'(x)$ is continuous on $[0, 1]$, then $\lim_{n} \sup_{0 \leq x \leq 1} |E(Dg_{n,m}(x)) - g'(x)| = 0$.

**Proposition 3.1.** If $\{e^{(j)}(\cdot), j \geq 1\}$ is a stationary sequence, under H1 and H2 we have that

(i) $P(\lim_{n} \lim_{m} g_{n,m}(x) = g(x)) = 1$.

(ii) If in addition the process $e(x)$ has continuous paths a.s. then $P(\lim_{m} \lim_{n} g_{n,m}(x) = g(x)) = 1$.

(iii) If $g'(x)$ is continuous at $x$ we have that $P(\lim_{m} \lim_{n} Dg_{n,m}(x) = g'(x)) = 1$.

(iv) If the derivative process $e'(x)$ is continuous with zero mean and $g'(x)$ exists then $P(\lim_{m} \lim_{n} Dg_{n,m}(x) = g'(x)) = 1$.

Proofs of Lemma 3.1 and Proposition 3.1 can be found in Fraiman and Pérez Iribarren [6].
The following theorem establishes the uniform mean square convergence of these estimates, with the only requirement on \( m \) and \( n \) that \( \min(n, m) \to \infty \).

**Theorem 3.2.** Assume that \( H_1, H_2, \sup_x E(|e(x)|)^{2+\theta} \leq K < \infty \) for some \( \theta > 0 \), and \( \sum_{k=0}^{\infty} \alpha(k)^{\theta/(2+\theta)} < \infty \) hold. Then

\[
\lim_{\min(n, m) \to \infty} \sup_x E((g_{n,m}(x) - g(x))^2) = 0.
\]

**Proof.** Denote \( V^{(j)}(x) = V_n^{(j)}(x) = \sum_{i=1}^{m} w_{ni}(x) e^{(j)}(x_i) \). Then

\[
E((g_{n,m}(x) - g(x))^2) = E\left( \left( m^{-1} \sum_{j=1}^{m} V^{(j)}(x) \right)^2 \right) + \left( m^{-1} \sum_{j=1}^{m} \sum_{i=1}^{n} w_{ni}(x)(g(x_i) - g(x))^2. \right. \quad (3.1)
\]

By Lemma 3.1 the second term on the right-hand side in (3.1) converges to zero uniformly in \( m \) and \( x \), as \( n \to \infty \). Since \( E(V^{(j)}(x)) = 0 \) for all \( x \in [0, 1] \) and

\[
E(|V^{(j)}(x)|^{2+\theta}) \leq E\left( \sum_{i=1}^{n} w_{ni}(x) |e^{(j)}(x_i)|^{2+\theta} \right) \leq K \sum_{i=1}^{n} w_{ni}(x) = K
\]

by the corollary of Lemma 2.1 of Davydov [4] (and the corresponding remark on it), with \( p = q = (2+\theta) \), \( E((m^{-1} \sum_{j=1}^{m} V^{(j)}(x))^2) \) can be majorized by

\[
m^{-2} \sum_{j=1}^{m} \sum_{h=1}^{m} C(\alpha(|j-h|))^\theta/(2+\theta) \leq 2Cm^{-1} \sum_{h=1}^{m} (\alpha(h))^\theta/(2+\theta)
\]

which converges to zero uniformly in \( n \) and \( x \) as \( m \to \infty \).

**Remark 3.1.** In the independent case, we will just need \( H_2 \) and \( \sup_x E(e^2(x)) < \infty \), and the proof is easier. This will be the case quite often in what follows.

For the derivative of \( g(x) \) we have the following result:

**Theorem 3.3.** Assume \( H_1, H_2, E(\sup_x |e'(x)|^{2+\rho}) \leq L < \infty \) for some \( \rho > 0 \) and \( \sum_{k=0}^{\infty} \alpha(k)^{\rho/(2+\rho)} < \infty \). If \( g(x) \) has a continuous derivative \( g'(x) \) for all \( x \in [0, 1] \), we have

\[
\lim_{\min(n, m) \to \infty} \sup_x E((Dg_{n,m}(x) - g'(x))^2) = 0.
\]
A proof can be found in Fraiman and Pérez Iribarren [6].

**Theorem 3.4.** Assume $H_1$, $H_2$, and that the mixing coefficients are geometric. Then we have that

(i) If there exists $L > 0$ such that $|\rho(x)| < L$ a.s., then $P(\lim_{n,m} g_{n,m}(x) = g(x)) = 1$.

(ii) If $e'(x)$ is bounded, then $P(\lim_{n,m} g_{n,m}(x) = g'(x)) = 1$.

**Proof.** (i) Since $\sum_{i=1}^n w_i(x)(g(x_i) - g(x))$ converges to zero as $n \to \infty$ it suffices to show that $m^{-1} \sum_{j=1}^m V^{(j)} \to 0$ a.s. as $m \to \infty$, with $V^{(j)}$ defined in the proof of Theorem 3.2.

We have that $E(W_j/m) = 0$, $|V_j|/m \leq L/m \leq 1$, for $m$ large enough and $E(|V_j|^3/m^3)^{1/3} < L/m$. Thus, we can apply a Bernstein type exponential inequality for $\alpha$-mixing processes due to Doukhan, León, and Portal [5, Theorem 6] to obtain

$$P\left(m^{-1} \left| \sum_{j=1}^m V^{(j)} \right| > \eta \right) \leq 3C_1 \exp(-C_2\eta^{1/2}m^{1/4})$$

for any $\eta > 0$, where $C_1$ and $C_2$ are positive constants which depend only on the mixing coefficients, which concludes the proof.

(ii) As $\sum_{i=1}^n w_i(x)[(g(x_{i+1}) - g(x_i))/(x_{i+1} - x_i) - g'(x)]$ converges to zero as $n \to \infty$, the proof follows in a similar way as in (i), applying the same exponential inequality to $U_n^{(j)} = U_n^{(j)}(x) = \sum_{i=1}^n w_i(x)(e^{(j)}(x_{i+1}) - e^{(j)}(x_i))/(x_{i+1} - x_i)$ which is a zero mean, bounded geometrically $\alpha$-mixing sequence.

B. Asymptotic Distribution for a Fixed $x$

In order to obtain the asymptotic distribution of the estimates, it will be necessary to relate $n$ and $m$ by considering $n = n[m]$. The asymptotic bias will depend on the asymptotic behaviour of $S_{lm}(x) = m^{1/2} \sum_{j=1}^{n[m]} w_m(x)(g(x_j) - g(x))$ for which we will need some considerations on the design and on the relationship between $n$ and $m$.

If $W_n(x)$ is the artificial sequence of random variables defined on Remark 2.1(a) and we denote by $\pi_n(x) = \bar{E}((W_n(x) - x)^2)$, the mean square error with respect to $\bar{P}$, it is easy to verify that $\pi_n = \sup_x \pi_n(x)$ converges to zero as $n \to \infty$. Let $H_3$ and $H_4$ be the following assumptions:

**H3.** The function $g$ verifies a Lipschitz condition of order one.

**H4.** There exist $\theta > 0$, $K > 0$, and $0 < a < 1$ such that:

(i) $\sup_x E(|e(x)|^{2+\theta}) < K < \infty$.

(ii) The mixing coefficients verify $\sum_{k=0}^{\infty} \lambda(k)^{(1-a)\theta/(2+\theta)} < \infty$. \

**Lemma 3.2.** Under $H_2$ and $H_3$ and for any sequence $n = n(m)$ such that there exists a sequence $\delta_n$ for which $m^{1/2}\delta_n \to 0$ and $m^{1/2}n_n/\delta_n^2 \to 0$ as $m \to \infty$, we have $\sup_x S_{\text{lm}}(x) \to 0$ as $m \to \infty$.

**Proof.**

$$|S_{\text{lm}}(x)| \leq m^{1/2} \sum_{i \in \mathcal{N}} w_{ni}(x) |g(x_i) - g(x)| + m^{1/2} \sum_{i \in \mathcal{A}} w_{ni}(x) |g(x_i) - g(x)|$$

$$\leq m^{1/2} M \delta_n + 2m^{1/2} \max_x |g(x)| \pi_n/\delta_n^2,$$

where $M$ is determined by $H_3$ and $A = A_n = A(n, x, \delta_n)$. Therefore $\sup_x |S_{\text{lm}}(x)| \to 0$.

**Remark 3.2.** Note that we can always choose $n = n(m)$ such that $\lim_{m \to \infty} m^{1/2}n_n = 0$. Moreover, we show in the Appendix that the usual kernel and nearest neighbor weights verify the assumptions of Lemma 3.2. In particular for a large class of kernel weights sufficient conditions are that $\lim \inf_{m \to \infty} nh_n > 0$ and $\lim_{m \to \infty} mh_n = 0$, where $h_n$ denotes the kernel bandwidth. From now on $n = n(m)$ will be chosen to satisfy the conditions of Lemma 3.2.

**Theorem 3.5.** Assume $H_1$ to $H_4$ and that for each $\delta > 0$ there exists $\eta = \eta(\delta)$ verifying $\sup_{|h| \leq \delta} E((e(x_0 + h) - e(x_0))^2) \leq \eta(\delta)$, where $\eta(\delta) \to 0$ as $\delta \to 0$. Then $m^{1/2}(g_{nm}(x_0) - g(x_0)) \to^\omega Z_1$, where $\to^\omega$ stands for weak convergence. $Z_1$ has a normal distribution with zero mean and variance $\sigma^2 = \sigma^2(x_0) = \lim_{m \to \infty} E((\sum_{j=1}^m e(j)(x_0))^2)/m$, assuming that this limit exists and $\sigma^2 > 0$. In the stationary case, $\sigma^2 = E(e(1)(x_0)^2) + 2 \sum_{j=1}^\infty E(e(j)(x_0) e(j+1)(x_0))$.

**Proof.** For each $x_0 \in [0, 1]$ we have that

$$m^{1/2}(g_{nm}(x_0) - g(x_0)) = m^{-1/2} \sum_{j=1}^m \sum_{i=1}^{n[m]} w_{ni}(x_0) e(j)(x_i)$$

$$+ m^{-1/2} \sum_{j=1}^m \sum_{i=1}^{n[m]} w_{ni}(x_0)(g(x_i) - g(x_0))$$

(3.3)

and Lemma 3.2 provides the limit of the second term in the right-hand side member of (3.3). On the other hand,

$$m^{-1/2} \sum_{j=1}^m \sum_{i=1}^{n[m]} w_{ni}(x_0) e(j)(x_i)$$

$$= m^{-1/2} \sum_{j=1}^m T(j)(x_0) + m^{-1/2} \sum_{j=1}^m e(j)(x_0)$$

(3.4)
with $T^{(j)}(x_0) = \sum_{i=1}^m w_{ni}(x_0) e^{(j)}(x_i) - e^{(j)}(x_0)$. As $E(T^{(j)}(x_0)) = 0$, and $m^{-1} E((\sum_{j=1}^m T^{(j)}(x_0))^2) \leq m^{-1} \sum_{j=1}^m \sum_{k=1}^m |E(T^{(j)}(x_0) T^{(k)}(x_0))|$, we will majorize $|E(T^{(j)}(x_0) T^{(k)}(x_0))|$ in order to prove that $m^{-1} \sum_{j=1}^m T^{(j)}(x_0)$ converges to zero in probability. For each $\delta > 0$ we have

$$|E(T^{(j)}(x_0) T^{(k)}(x_0))| \leq \sum_{(i,j) \in B} w_{ni}(x_0) w_{nj}(x_0) \min(\eta(\delta), C\alpha(|j-k|)^{\theta/(2+\theta)})$$

$$+ \sum_{(i,j) \in B^c} w_{ni}(x_0) w_{nj}(x_0) C\alpha(|j-k|)^{\theta/(2+\theta)}$$

$$\leq \min(\eta(\delta), C\alpha(|j-k|)^{\theta/(2+\theta)})$$

$$+ 2\lambda(\delta) C\alpha(|j-k|)^{\theta/(2+\theta)}$$

$$\leq (\eta(\delta)^a C^{-1} - \alpha((j_k)^{1-a})^{\theta/(2+\theta) + 2\lambda(\delta) C\alpha(|j-k|)^{\theta/(2+\theta)}})$$

where $B = \{(i, h)/i \in A(n, x_0, \delta), h \in A^c(n, x_0, \delta)\}$ with $A$ defined in H2(iii), $C$ given by Corollary 2.1 of Davydov [4], $\lambda(\delta)$ obtained from H2(iii), and $a, \theta$ given in assumption H4. Therefore,

$$m^{-1} \sum_{j=1}^m \sum_{k=1}^m |E(T^{(j)} T^{(k)})|^2$$

$$\leq 2C^{-1} - \eta(\delta)^a \sum_{j=0}^{\infty} (\alpha(j))^{1-a})^{\theta/(2+\theta)}$$

$$+ 4\lambda(\delta) C \sum_{k=0}^{\infty} (\alpha(k))^{\theta/(2+\theta)}$$

which can be made arbitrarily small if we first choose $\delta$ such that $\eta(\delta)$ is small, and for this $\delta$ we choose $m$ large enough so that $n = n(m)$ makes $\lambda(\delta)$ arbitrarily small.

The conclusion of Theorem 3.5 finally follows from Corollary 1 of Herrndorf [13], a Central Limit Theorem for $\alpha$-mixing sequences.

**Theorem 3.6.** Under the assumptions of Theorem 3.5, if $g'$ verifies a Lipschitz condition of order one, $E(e'(x_0)) = 0$, $\sup_x E(|e'(x)|^{2+\theta}) < \infty$, and $\sup_{|h| \leq \delta} E((e'(x_0 + h) - e'(x_0))^2) \leq \eta$ we have that $Z_m(x_0) = m^{-1/2}((g_{nm}(x_0) - g(x_0)), (Dg_{nm}(x_0) - Dg_{nm}(x_0))') \rightarrow^w Z$, where $Z$ is a normal random vector with zero mean and covariance matrix $A = ((\alpha_{ij}))$, $1 \leq i, j \leq 2$, $\alpha_{11}$ is given in Theorem 3.5. $\lambda_{22} = \lim_m E((\sum_{j=1}^m e^{(j)}(x_0))^2)/m$, $\lambda_{12} = \lim_m E((\sum_{j=1}^m e^{(j)}(x_0))(\sum_{j=1}^m e^{(j)}(x_0)))/m$, if all the limits exist and $\lambda_{11} > 0$, $\lambda_{22} > 0$. In the stationary case, $\lambda_{11}$ is given in Theorem 3.5, $\lambda_{22} =$
\[ E((e'(1)(x_0))^2) + 2\sum_{j=1}^{\infty} E(e'(1)(x_0) e'(j+1)(x_0)), \text{ and } \lambda_{12} = E(e'(1)(x_0) e'(1)(x_0)) + \sum_{j=1}^{\infty} (E(e'(1)(x_0) e'(j+1)(x_0)) + E(e'(1)(x_0) e'(j+1)(x_0))). \]

**Theorem 3.7.** If the assumptions of Theorem 3.6 holds for some fixed real numbers \( t_1, t_2, \ldots, t_p, t_i \in [0, 1], \) \( 1 \leq i \leq p, \) we have that

\[
m^{1/2}(g_{nm}(t_1) - g(t_1), \ldots, g_{nm}(t_p) - g(t_p),
\]

\[
Dg_{nm}(t_1) - g'(t_1), \ldots, Dg_{nm}(t_p) - g'(t_p))' \xrightarrow{\omega} Z,
\]

where \( Z \) is a normal random vector with zero mean and covariance matrix \( \Sigma \) with \( p \)-dimensional submatrices \( \Sigma_{ij}, 1 \leq i, j \leq 2, \) determined in a similar way to those of the preceding theorem.

A proof of Theorems 3.6 and 3.7 can be found in Fraiman and Pérez Iribarren [6].

C. **Asymptotic Distribution on** \( C[0,1] \) **of the Processes** \( \{ g_{nm}(x), x \in [0, 1] \} \) **and** \( \{ Dg_{nm}(x), x \in [0, 1] \}. \)

Let us now assume that for each \( n \in \mathbb{N} \) the weight functions \( w_{ni}(x), 1 \leq i \leq n, \) are continuous functions of \( x \) and the sequence \( \{ e^{(j)}(x), x \in [0, 1], j \geq 1 \} \) has continuous paths. Therefore \( \{ g_{nm}(x), x \in [0, 1] \} \) is a random element on the space \( C[0,1] \) of real-valued continuous functions on the interval \([0,1].\) On these conditions we are able to show the following result.

**Theorem 3.8.** Assume \( H_1 \) to \( H_3 \) and suppose that the following additional conditions are fulfilled:

(i) There exist \( \theta > 0 \) and \( 0 < a < 1 \) such that \( E(|e(0)|^{(2 + \theta)}) < \infty \) and \( \sum_{k=0}^{\infty} x(k)^{(1-a)\theta/(2+\theta)} < \infty. \)

(ii) There exist real constants \( K_1 > 0 \) and \( a_1 > (a\beta)^{-1} \) such that \( E(\sup_{|x-x'| \leq \delta} |e(x) - e(x')|^{2 + \theta}) \leq \eta(\delta) = K_1 \delta^{a_1}, \) where \( a \) is given in (i) and \( \beta = 2/(2 + \theta). \)

(iii) It can be chosen \( n = n(m) \to \infty \) such that \( m \delta_n \to 0 \) and \( m \pi_n/\delta_n^2 \to 0 \) as \( m \to \infty \) for some sequence \( \{ \delta_n \} \) with \( \lim_n \delta_n = 0, \) and \( b_1 = \min(2, a_1 \beta). \)

Then \( m^{1/2}(g_{nm}(x) - g(x)) \to^\omega \gamma(x), \) where \( \gamma(x) \) is a gaussian process with \( p \)-dimensional distributions with zero mean and covariance matrix \( \Sigma_{11} \) given by Theorem 3.7.

Note that assumption (ii) excludes some well-known processes such as the Brownian motion. However, in this case, the independent increment
structure can be used in order to obtain consistent estimators (see, for instance, Ibragimov et al. [14]).

Proof. Define

\[ S_{1m}(x) = m^{1/2} \sum_{i=1}^{n[m]} w_{ni}(x)(g(x_i) - g(x)), \]

\[ S_{2m}(x) = m^{-1/2} \sum_{j=1}^{m} e^{(j)}(x), \]

\[ S_{3m}(x) = m^{-1/2} \sum_{j=1}^{m} \sum_{i=1}^{n[m]} w_{ni}(x)(e^{(j)}(x_i) e^{(j)}(x)). \]

Since (iii) implies the design assumption of Lemma 3.2, \( \sup_x |S_{1m}(x)| \) converges to zero as \( m \to \infty \). Thus, the proof will be complete if we show that (a) \( S_{2m}(x) \to^w \gamma(x) \) and (b) \( \sup_x |S_{3m}(x)| \) converges to zero in probability as \( m \to \infty \).

(a) From Theorem 3.7 we know that the \( p \)-dimensional distributions of \( S_{2m}(x) \) are asymptotically normally distributed. On the other hand, we have that

\[
E \left( m^{-1/2} \left[ \sum_{j=1}^{m} e^{(j)}(t_2) - e^{(j)}(t_1) \right]^2 \right) \\
= m^{-1} \sum_{j=1}^{m} \sum_{k=1}^{m} E((e^{(j)}(t_2) - e^{(j)}(t_1))(e^{(k)}(t_2) - e^{(k)}(t_1))) \\
\leq m^{-1} \sum_{j=1}^{m} \sum_{k=1}^{m} \min(K_1^{\beta} |t_2 - t_1|^{\alpha_1 \beta}, Cx^{\theta/(2 + \theta)}(|j - k|)) \\
\leq m^{-1} K_1^{\alpha \beta} |t_2 - t_1|^{\alpha_1 \beta} \sum_{j=1}^{m} \sum_{k=1}^{m} \{ Cx^{\theta/(2 + \theta)}(|j - k|) \}^{1 - \alpha} \\
\leq C_1 |t_2 - t_1|^{\alpha_1 \beta},
\]

where \( C_1 = 2K_1^{\alpha \beta} C^{(1 - \alpha)} \sum_{k=0}^{\infty} x^k(k) \) and \( h = (1 - \alpha) \theta/(2 + \theta) \). Therefore (a) holds since Theorem 12.3 of Billingsley [2] implies that the sequence \( S_{2m}(x) \) is tight.

(b) Define \( V^{(j)} = \sup_x \sum_{i=1}^{n[m]} w_{ni}(x) |e^{(j)}(x_i) - e^{(j)}(x)| \). Then we have that

\[
E(\sup_x |S_{3m}(x)|^2) \leq m^{-1} \sum_{j=1}^{m} \sum_{k=1}^{m} E(V^{(j)}V^{(k)}) \tag{3.6}
\]
with $E(V^{(j)}V^{(k)})$ defined by

$$E(V^{(j)}V^{(k)}) = E\left( \sup_{x,y} \sum_{i=1}^{n} \sum_{h=1}^{n} w_{ni}(x) w_{nh}(y) \times |e^{(j)}(x_i) - e^{(j)}(x)| |e^{(k)}(x_h) - e^{(k)}(y)| \right). \quad (3.7)$$

Let $B^c = \{(i, h)/i \in A(n, x, \delta_n)\} \cup \{(i, h)/h \in A(n, y, \delta_n)\}$ with $A(n, x, \delta)$ defined in H2. Then we have that

$$E\left( \sup_{x,y} \sum_{(i, h) \in B^c} w_{ni}(x) w_{nh}(y) |e^{(j)}(x_i) - e^{(j)}(x)| |e^{(k)}(x_h) - e^{(k)}(y)| \right) \leq 2E\left( \sup_{x} \sum_{i \in A} w_{ni}(x) |e^{(j)}(x_i) - e^{(j)}(x)| \times \sup_{y} \sum_{h=1}^{n} w_{nh}(y) |e^{(k)}(x_h) - e^{(k)}(y)| \right)$$

$$\leq 2 \sup_{x} \sum_{i \in A} w_{ni}(x) \sup_{y} \sum_{h=1}^{n} w_{nh}(y) \times E(\sup_{x} |e^{(j)}(x_i) - e^{(j)}(x)| \sup_{y} |e^{(k)}(x_h) - e^{(k)}(y)|)$$

$$\leq 2 \sup_{x} \sum_{i \in A} w_{ni}(x) \left[C x^{\theta/2 + \theta}(|j - k|) + E^2(\sup_{y, h} |e(x_h) - e(y)|) \right] \leq 2 \lambda_n \left[C x^{\theta/2 + \theta}(|j - k|) + \eta(1)^{\theta/2} \right],$$

where $\lambda_n = \lambda(\delta_n) = \sup_x \sum_{i \in A} w_{ni}(x)$. On the other hand (ii) implies that

$$E\left( \sup_{x,y} \sum_{(i, h) \in B} w_{ni}(x) w_{nh}(y) |e^{(j)}(x_i) - e^{(j)}(x)| |e^{(k)}(x_h) - e^{(k)}(y)| \right) \leq E\left( \sup_{D} |e^{(j)}(x_i) - e^{(j)}(x)| |e^{(k)}(x_h) - e^{(k)}(y)| < \eta(\delta_n)^{\theta},

where $D = \{(x, y, x_i, x_h)/|x_i - x| \leq \delta_n, \ |x_h - y| \leq \delta_n, \ 1 \leq i, h \leq n\}$. Finally (3.6) and (3.7) imply that

$$E(\sup_{x} |S_{3m}(x)|^2) \leq m^{-1} \sum_{j=1}^{m} \sum_{k=1}^{m} (2 \lambda_n [C x^{\theta/2 + \theta}(|j - k|) + \eta(1)^{\theta/2}] + \eta(\delta_n)^{\theta})$$

$$\leq 4C \lambda_n \left[ \sum_{k=0}^{\infty} \alpha^{\theta/2 + \theta}(k) + m \eta(1)^{\theta/2} \right] + mn(\delta_n)^{\theta}$$
and the design assumption (iii) entails that $m\eta(\delta_n)^{b} = K'_{b}m\delta_{n}^{a+b} \to 0$ and $m\delta_{n}^{a+b} \to 0$ as $m \to \infty$, which concludes the proof.

Remark 3.3. In the case of independence the proof is quite simpler and the regularity conditions on Theorem 3.8 can be reduced to $E(|e(x) - e(x')|^{2}) \leq K_{1}|x - x'|^{a_{1}}$ with $a_{1} > 0$ and $\lim_{\delta \to 0} E(\sup_{|x - x'| \leq \delta} |e(x) - e(x')|^{2}) = 0$.

The design assumptions in this case are just those required in Lemma 3.2.

Remark 3.4. If, for instance, $w_{ni}(x) = C_{0}x^{(1-x)^{n-i}}$, $x_{i} = x_{i,n} = i/n$, $0 \leq i \leq n$, we have that $\pi_{n} \leq 1/(4n)$ and, choosing $n = m^{3}$, the sequence $\delta_{n} = n^{-\frac{1}{4}}$ satisfies (iii) if $a_{1} > \frac{5}{3}$ and $\theta = \frac{1}{3}$.

Theorem 3.9. Assume $H1$, $H2$, and the following conditions:

(i) The function $g(x)$ is continuously differentiable with derivative $g'(x)$ and $e(x)$ has a derivative for $e'(x)$ each $x$ with continuous path process $e'(x)$.

(ii) There exist $\theta_{1} > 0$ and $0 < a < 1$ such that $E(|e'(0)|^{2+\theta_{1}}) < \infty$ and $\sum_{j=0}^{\infty} a_{j}(1-a_{j}(1/2+\theta_{1})) < \infty$.

(iii) There exist $K_{2} > 0$ and $a_{2} > (a\beta)^{-1}$ such that $E(\sup_{|x - x'| \leq \delta} |e'(x) - e'(x')|^{2+\theta_{1}}) \leq K_{2}\delta^{a_{2}}$ with a given in (ii) and $\beta = 2/(2+\theta_{1})$.

(iv) It can be chosen that $n = n(m) \to \infty$ such that $m\delta_{n}^{b_{2}} \to 0$ and $m\pi_{n}/\delta_{n}^{2} \to 0$ as $m \to \infty$ for some sequence $\{\delta_{n}\}$ with $\lim_{m} \delta_{n} = 0$, and $b_{2} = \min(2, a_{2}\beta)$.

Then $m^{1/2}(Dg_{nm}(x) - g'(x)) \to^{m} \gamma_{1}(x)$, where $\gamma_{1}(x)$ is a Gaussian process with $p$-dimensional distributions with zero mean and covariance matrix $\Sigma_{22}$ given by Theorem 3.7.

Proof. The proof can be obtained by substituting $e'(x)$ for $e(x)$ in the proof of Theorem 3.8 and using that $\sup_{x} m_{1/2} \sum_{j=1}^{n} n_{i} w_{ni}(x) \{(g(x_{i+1}) - g(x_{i}))/n_{i+1} - x_{i}) - g'(x)\} \to 0$ as $m \to \infty$, with $x_{n+1} = 1$, i.e., an analogous result to Lemma 3.2.

4. A REAL DATA EXAMPLE

In this section we will describe briefly a real data example. Monthly measurements of the salinity at a fixed point in the Río de la Plata, near the city of Montevideo, 34°56' latitude S and 56°09' longitude W, were obtained by averages of daily data on each month. Two 6-year periods of
time were available: data from 1955–1960 and from 1981–1986. The two series of 72-month data are shown in Fig. 1.

As pointed in the Introduction we consider the model

\[ Y_{j}^{(I)}(t_i) = g_{i}(t_i) + e_{j}^{(I)}(t_i) \quad \text{for} \quad j = 1, \ldots, 6, I = 0, 1, \]

and \( t_i \in \{1, 2, \ldots, 12\} \). When \( I = 0 \) we are considering data from the first series—corresponding to the period 1955–1960—while we use the index \( I = 1 \) for the second series. The value \( j \) corresponds to observations taken \( (j - 1) \) years after the beginning of each series and \( t_i \) indicates which month of the year we are considering. For example, \( Y_{0}^{(3)}(5) \) stands for the observation at May 1957.

Finally in Fig. 2 a scatterplot of the smooth estimation defined by (2.2)
is given. Nearest neighbor weights with \( k = 3 \) have been considered, and the graph is obtained by linear interpolation.

The last graph indicates that the values of salinity in the last period are considerably lower than in the first one. That was one of the hypotheses suggested by F. Gascue and G. Manzzetta from the Oceanographic Center of the Navy at Montevideo. Some testing is considered later on.

**APPENDIX**

In this section we will show that assumption H2 can be slightly weakened in such a way that all the previous results still hold and that the more usual weight functions verify the design assumptions.

(a) H2 can be replaced by H2', where

\[
\text{H2':}\quad (i) \quad \sup_x \sum_{i=1}^n |w_{n_i}(x)| I_N \to 0, \text{ where } N = N_n = \{i/w_{n_i}(x) < 0\}.
\]

(ii) \( \lim_n \sum_{i=1}^n w_{n_i}(x) = 1 \) uniformly for \( x \in [0, 1] \).

(iii) \( \lim_n \sup_x \sum_{i=1}^n w_{n_i}(x) I_{N_n \cap A(n, \delta)} = 0. \)

All the results in Section 3A are valid replacing H2 by H2' with straightforward modifications. In Section 3B, Lemma 3.2 requires an additional condition on the behaviour of \( \beta_n = \sup_x \sum_{i=1}^n |w_{n_i}(x)| I_N \). It suffices to choose the sequences \( n = n(m) \) and \( \delta_n \) verifying \( m^{1/2} \beta_n \to 0 \), \( m^{1/2} \delta_n \to 0 \), and \( m^{1/2} \pi_n / \delta_n^2 \to 0 \) as \( m \to \infty \) with \( \pi_n^* = \sup_x \sum_{i=1}^n w_{n_i}(x) (x_i - x)^2 \). As in Section 3A, straightforward modifications provide the proofs of the corresponding results in Section 3B. Slight modifications on (iii) and (iv) are also required for the proofs of Theorem 3.8 and Theorem 3.9, respectively, which we omit.

(b) Finally we will show that the design assumptions are verified for nearest neighbor and kernel weights.

(b1) Uniform k-nearest neighbor weights. Let \( H_n(x) \) be the distance to the k-nearest neighbor to \( x \) between \( x_1, \ldots, x_n \), and \( k = k_n \) a sequence of positive integers. We define \( w_{n_i}(x) = k_n^{-1} \) if \( |x_i - x| \leq H_n(x) \) and 0 otherwise. Let \( d_n = \max_i |x_{i+1} - x_i| \). Then H2 is verified if \( k_n d_n \to 0 \) as \( n \to \infty \). In order to show that the assumptions of Lemma 3.2 are verified, since \( \pi_n \leq k_n^{-1} d_n^2 (1^2 + 2^2 + \cdots + k_n^2) \leq k_n^2 d_n^2 \), the conditions \( m^{1/2} \delta_n \to 0 \) and \( m^{1/2} k_n^2 d_n^2 / \delta_n^2 \to 0 \) for \( m \to \infty \) are sufficient. For instance, if \( d_n = (n - 1)^{-1} \), we may choose \( m = n^2 \), \( \delta_n = n^{-1/2} \) and \( k_n = o(n^{3/8}) \).

(b2) Nearest neighbor weights. Let \( c_{n1} \geq c_{n2} \geq \cdots \geq c_{nn} \geq 0 \), \( \sum_{i=1}^n c_{ni} = 1 \) and define \( w_{n_i}(x) = c_{nR_i} \) where \( R' = R_i(x) \) is the rank of \( |x_i - x| \) in the vector \( (|x_1 - x|, \ldots, |x_n - x|) \) as in Stone [20].

Then H2 is verified if \( \sum_{j \geq \lceil \delta / d_n \rceil + 1} c_{nj} \to 0 \) as \( n \to \infty \) with \( d_n \) as in \( b_1 \). It
is easy to see that we can choose \( n = n(m) \) and \( \delta_n \) such that the conditions in Lemma 3.2 are fulfilled.

Note that Theorems 3.8 and 3.9 do no hold in this case, since we have required the weight function to be continuous.

\((b_3)\) Kernel weights. Let \( K: \mathbb{R} \to \mathbb{R} \) be a continuous function verifying that \( I_1 \leq K(t) \leq I_2 \) for some \( 0 < a < b \), \( 0 < I_1 \leq I_2 \), and \( w_{nj}(x) = K((x-x_j)/h_n)/\sum_{i=1}^{n} K((x-x_i)/h_n) \), where \( h_n \) is a sequence of nonnegative real numbers, \( \lim_n h_n = 0 \). If \( d_n < 2ah_n \) the weights are well defined and \( H2(i) \) and \( (ii) \) hold. As \( h_n \to 0 \), \( \delta/h_n > b \) for \( n \) large enough and \( H2(iii) \) holds. Since \( \sum_{i=1}^{n} (x-x_i)^2 w_{ni}(x) \leq b^2 h_n^2 \), we can choose \( \pi_n = b^2 h_n^2 \) and therefore there exist \( n = n(m) \) such that \( \delta_n \to 0 \), \( m^{1/2} \delta_n \to 0 \), and \( m^{1/2} \pi_n/\delta_n^2 \to 0 \) as \( m \to \infty \). Therefore, it is easy to see that \( \lim_{m \to \infty} mh_n > 0 \) and \( \lim_{m \to \infty} mh_n = 0 \) are sufficient conditions for the asymptotic results. These conditions seems to be quite sharp and shows how the repeated measurements allows us to weaken the usual conditions for the independent case. For instance, \( \delta_n = h_n^{1/2} \) and \( h_n = o(m^{-1}) \) are sufficient conditions for Lemma 3.2.

Finally, for Theorems 3.8 and 3.9 we can choose \( n = n(m) \) for \((iii)\) and \((iv)\), respectively. For example, if \( a_1 \) in \((ii)\) of Theorem 3.8 is equal to 2, \( \theta = \frac{1}{2} \), and \( d_n = (n-1)^{-1} \), \( h_n = n^{-1/2} \) and \( n = m^3 \) give a possible choice.

**Acknowledgment**

We thank an anonymous referee for helpful suggestions.

**References**


