A SIMULATION STUDY OF FUNCTIONAL DENSITY-BASED INVERSE REGRESSION

Noslen Hernández[1] *, Rolando J. Biscay[2,3], Nathalie Villa-Vialaneix[4,5], Isneri Talavera[1]

[1] Advanced Technologies Application Center, CENATAV - Cuba
[2] Institute of Cybernetics, Mathematics and Physics - Cuba
[3] Departamento de Estadística de la Universisad de Valparaiso y CIMFAV - Chile
[4] Institut de Mathématiques de Toulouse, Université de Toulouse - France
[5] IUT de Perpignan, Département STID, Carcassonne - France

ABSTRACT
In this paper a new nonparametric functional regression method is introduced for predicting a scalar random variable Y on the basis of a functional random variable X. The prediction has the form of a weighted average of the training data y_i, where the weights are determined by the conditional probability density of X given $Y = y_i$, which is assumed to be Gaussian. In this way such a conditional probability density is incorporated as a key information into the estimator. Contrary to some previous approaches, no assumption about the dimensionality of $E(X|Y = y)$ or about the distribution of X is required. The new proposal is computationally simple and easy to implement. Its performance is assessed through a simulation study.

RESUMEN
En este artículo se introduce un nuevo método de regresión funcional no paramétrico para predecir una variable aleatoria Y de valores reales, sobre la base de una variable aleatoria funcional X. Las predicciones se construyen mediante una promediación ponderada de los datos de entrenamiento y_i, donde las ponderaciones están determinadas por la densidad de probabilidad condicional de X dado $Y = y_i$, la cual se supone Gaussiana. De este modo, dicha densidad condicional es incorporada como información clave en el estimador que se propone. Contrariamente a otros enfoques existentes, no se requieren supuestos restrictivos sobre la dimensión de $E(X|Y = y)$ o la distribución de X. La nueva propuesta es computacionalmente simple y fácil de implementar. Su comportamiento es evaluado a través de un estudio de simulación.

KEY WORDS: Functional data analysis; Inverse regression; functional density estimation; Nonparametric regression

MSC code: 62G07, 62G08.

1. INTRODUCCIÓN

The fast development of instrumental analysis equipment and modern measurement devises provides huge amounts of data as high-resolution digitized functions. As a consequence, Functional Data Analysis (FDA)

*Ave. 7ma No.21812 e/ 218 y 222, Playa, C. Habana, Cuba, C.P 12200, nhernandez@cenatav.co.cu

has become a growing research field [24, 9]. In the FDA setting, each individual is treated as a single entity described by a continuous real-valued function rather than by a finite-dimensional vector: functional data (FD) are then supposed to have values in an infinite dimensional space, typically a Hilbert space $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle$.

Specifically, in functional regression analysis one intends to predict a random scalar variable (response variable) Y from a functional random variable (predictor) X (e.g., with values in $\mathcal{X} = L_2([a, b])$). That is, the goal is to estimate the regression function $\gamma(x) = \mathbb{E}(Y|X = x)$, where $\mathbb{E}(Y|X)$ denotes the conditional expectation, on the basis of a sample $(x_i, y_i)_{i=1,\dots,n}$ of independent realizations of (X, Y). Equivalently, the aim is to fit the functional regression model

$$Y = \gamma(X) + \epsilon,$$

where ϵ is a real random variable independent of X. Several approaches have been proposed for this problem, which can be classified into three main families:

a) *Linear regression methods.* Earlier work were focused on linear regression models where the regression function γ takes the linear form

$$\gamma(x) = c + \langle \beta, x \rangle,$$

where $c \in \mathbb{R}$ and $\beta \in \mathcal{X}$ are unknown parameters. A review of approaches for estimating this model can be found in [24]; see also, [14, 19, 3, 4, 2, 5, 23].

b) *Nonparametric regression methods.* A drawback of linear regression methods is that they can not deal with nonlinear dependencies between the predictor and the response variables. To overcome this, a number of nonparametric approaches has been proposed. We review the most important approaches developed in the past few years in that area.

The first approach, introduced in [9], is the use of functional kernel regression estimators:

$$\widehat{\gamma}(x) = \frac{\sum_{i=1}^{n} K(d(x_i, x)/h) y_i}{\sum_{i=1}^{n} K(d(x_i, x)/h)},$$

where $h > 0$ is the kernel bandwidth, d is a semi-metric on $\mathcal{X}$ and $K : R_+ \to R$ is a suitable kernel function. This kind of estimators allows for great flexibility in fitting nonlinear models. However, their consistency properties have been demonstrated only for restricted classes of kernel functions such as those of types I and II in [9]; furthermore, the data-driven selection of the kernel bandwidth h is a difficult problem, especially for this high dimensional setting.

Another class of nonparametric regression estimators are the Functional Neural Networks proposed by Rossi et. al. [25]. In particular, the single hidden layer perceptron is defined by

$$\widehat{\gamma}(x) = \sum_{j=1}^{q} \widehat{a}_j T(\widehat{u}_j + \widehat{l}_j(x)),$$

where $T : R \to R$ is a given activation function, $(l_j)_j$ are unknown linear functionals to be estimated (e.g., $l_j(x) = \langle w_j, x \rangle$ with $w_j \in \mathcal{X}$) and $(a_j)_j$, $(u_j)_j \subset \mathbb{R}$ are unknown parameters that also have to be estimated. Functional perceptrons have the universal approximation property that makes it possible to represent a wide variety of nonlinear functionals. But notice that they depend on a quite

large number of parameters $(w_j)_j$, $(a_j)_j$, $(u_j)_j$, which greatly increases with the number of neurons, q, and that their estimation by the optimization of the least square error leads to local minima issues. Also, the number of neurons has to be tuned which is a computationally difficult task. In the same spirit, functional versions of the radial basis functions neural networks have been introduced in [26].

Also, the general framework of function approximation in Reproducing Kernel Hilbert Spaces (RKHS) has been used [22, 15] to introduce functional regression estimators, which have the general form:

$$\widehat{\gamma}(x) = \sum_{i=1}^{n} \widehat{a}_i K(x_i, x),$$

where K is a reproducing kernel on $\mathcal{X}$, and $(a_i)_i \in \mathbb{R}$. In this framework can be included functional versions of support vector regression [15] and radial basis functions [22]. The latter has the specific form:

$$\widehat{\gamma}(x) = \sum_{i=1}^{m} \widehat{a}_i \phi(d(x, c_i)),$$

where ϕ is the adopted radial basis function, $c_1, ..., c_m \in \mathcal{X}$ are given centers, d is a distance defined on $\mathcal{X}$, and $(a_i)_i$ are unknown parameters.

An important advantage of the RKHS approach is that the resulting estimator is linear with respect to the unknown parameters $(a_i)_i$: their estimation by least squares optimization thus reduces to solve an algebraic linear problem. However, contrary to standard RKHS methods for approximating multivariate functions (e.g., standard multivariate splines and radial basis functions such as thin-plate splines), in the FDA setting the smoothness properties of $\widehat{\gamma}$ as a functional on $\mathcal{X}$ have not been yet defined. Hence, the selection of suitable reproducing kernels and radial basis functions is still an open issue.

Finally, more recently, k-nearest neighbors regression has been generalized to functional data [16, 1]. This approach leads to the following regression function:

$$\hat{\gamma}(x) = \frac{1}{k} \sum_{i=1}^{k} y_{(i,x)}$$

where $y_{(i,x)}$ is the value of Y for the *i-th* nearest neighbors of x among $(x_i)_{i=1,...,n}$. The consistency, as well as a rate of convergence, is given for this approach, depending on regularity conditions on γ.

c) *Functional Inverse Regression (FIR) methods.* More recently, an alternative methodology has been introduced that can be regarded as a compromise between too restrictive parametric methods (such as linear regression) and nonparametric ones (such as kernel methods). This is called Functional Inverse Regression (FIR or FSIR) [6, 12, 11], and constitutes a functional version of the Sliced Inverse Regression (SIR) previously proposed for multivariate regression models [17]. The FIR approach assumes that the following model holds:

$$Y = g(\langle \beta_1, X \rangle, ..., \langle \beta_d, X \rangle) + \epsilon,$$

where d is the so-called effective dimension and $g : \mathbb{R}^d \to \mathbb{R}$ is an unknown function. Under some additional assumptions (which are guaranteed if X has an elliptic distribution, e.g., a Gaussian distribution in the Hilbert space $\mathcal{X}$) the directions $(\beta_j)_j$ can be estimated from the spectral decomposition

of the covariance operators $\mathbb{V}(X)$ and $\mathbb{V}(\mathbb{E}(X|Y))$. The latter involves to fit the mean function of the "inverse" model

$$X = \mu(Y) + e, \tag{1.1}$$

where e is a random element in $\mathcal{X}$ with zero mean, not correlated with Y. More specifically, to estimate $(\beta_j)_j$ and g, the following steps are carried out:

1. Obtain an estimate of $\mu(Y) = \mathbb{E}(X|Y)$ by regressing each component of X versus Y through any univariate nonparametric regression method.
2. Obtain estimates of the covariance operators $\Gamma = \mathbb{V}(X)$ and $\Gamma_e = \mathbb{V}(\mathbb{E}(X|Y))$ on the basis of the results of step (1), and carry out the spectral decomposition of the operator $\Gamma^{-1/2}\Gamma_e\Gamma^{-1/2}$. $(\beta_j)_j$ are estimated as the eigenfunctions corresponding to the d first greatest eigenvalues.
3. Finally, estimate g through a d-variate nonparametric regression method.

In this semi-parametric approach, the dimension, d, is an hyper-parameter of the model. Several methods have been proposed to find a good d, such as the one proposed in [10].

In this paper, a new functional regression method to estimate $\gamma(X)$ is introduced that also relies on regarding the inverse regression model (1.1). Its main practical motivation arises from calibration problems in Chemometrics, specifically in spectroscopy, where some chemical variable Y (e.g., concentration) needs to be predicted from a digitized function X (e.g., an spectrum). In this setting, said "inverse" model represents the physical data generation process in which the output spectrum X is determined by the input chemical concentration Y, and e is a functional random perturbation mainly due to the measurement procedure. Though Y and X could have unknown complex probability distributions, it is a common assumption that the perturbation e follows a Gaussian distribution P_0, and so that the conditional distribution $P(\cdot/y)$ of X given $Y = y$ is a Gaussian distribution on $\mathcal{X}$ with mean $\mu(y)$. The relation between the regression function and the conditional density coming up from the definition of conditional expectation (see, e.g., [7, 27] in a nonparametric setting, and more specifically [18]) suggests the following estimate of $\gamma(x)$:

$$\widehat{\gamma}(x) = \frac{\sum_{i=1}^{n} \widehat{f}(x/y_i)\, y_i}{\widehat{f}_X(x)},$$

where $\widehat{f}(x/y)$ is an estimate of the density $f(x/y)$ of $P(\cdot/y)$ with respect to the measure P_0. This regression estimate will be refereed to as functional Density-Based Inverse Regression (DBIR). If X was a scalar variable, this would reduce to the approach for univariate calibration proposed in [18]. It requires more specific assumptions about the distribution of the perturbation e in the inverse model (1.1) (e.g., Gaussian distribution) but it has a number of appealing features in comparison with other approaches:

- Likewise other nonparametric approaches, it allows one to capture *nonlinear regression functions*.
- Likewise the FIR approach it requires to estimate the mean function μ of the inverse model and some covariance operator (specifically, $\mathbb{V}(e)$). μ should be estimated through a nonparametric regression method but, contrary to the classical nonparametric functional regression, this function is defined on $\mathbb{R}$ (and not in a infinite dimensional Hilbert space) and the estimation task is thus much easier, as well as the issue of tuning the hyperparameters (if so) in the chosen nonparametric regression method. Moreover, no other parameter have to be estimated. In particular, unlike functional kernel

regression, this approach does not require the selection of the bandwidth for a kernel on an infinite dimensional space. Also, it does not involve a large number of parameters related in a non quadratic way to the functional to optimize, contrary to, e.g., the case of functional perceptrons. DBIR is thus *computationally very easy to use*.

- Finally, unlike the FIR approach, *few assumptions are required*: in particular, γ does not need to be a function of a finite number d of projections nor X has to follow an elliptical distribution (or any other given distribution). Also notice that DBIR does not requires the additional multivariate nonparametric fitting step (c) aforementioned.

This paper is organized as follows. Section 2. defines the functional Density-Based Inverse Regression (DBIR) estimator. Section 3. carries out a simulation study in order to asses the feasibility and performance of the DBIR method. Finally, some conclusions are given in Section 4..

2. FUNCTIONAL DENSITY-BASED INVERSE REGRESSION (DBIR)

2.1 Definition of DBIR in an abstract setting

Let (X, Y) be a pair of random variables taking values in $\mathcal{X} \times \mathbb{R}$ where $(\mathcal{X}, \langle ., . \rangle)$ is the space of square integrable functions from $\mathbb{R}$ to $\mathbb{R}$ ($\mathcal{X} = L^2([a, b])$). Suppose also that n i.i.d. realizations of (X, Y) are given, denoted by $(x_i, y_i)_{i=1,\ldots,n}$. The goal is to build, from $(x_i, y_i)_i$, a predictor of the value of Y corresponding to a future observed value of X. This problem is usually addressed through the estimation of the regression function $\gamma(x) = \mathbb{E}(Y|X = x)$.
For this, the functional Density-Based Inverse Regression (DBIR) approach fits the inverse regression model:

$$X = F(Y) + \epsilon, \tag{2.2}$$

where ϵ is a random process (perturbation or noise) with zero mean, independent of Y, and $y \to F(y)$ is a function from $\mathbb{R}$ into $\mathcal{X}$. As was stated in Section 1., this is the commonly assumed data generating model in calibration problems [21].
Additionally, the following assumptions are made:

1. it exists a probability measure P_0 on $\mathcal{X}$ (not depending on y) such that the conditional probability measure of X given $Y = y$, say $P(\cdot / y)$, has a density $f(\cdot / y)$ with respect to P_0; i.e.,

$$P(A / y) = \int_A f(x / y) P_0(dx)$$

   for any measurable set A in $\mathcal{X}$;

2. it is assumed that Y is a continuous random variable, i.e., its distribution has a density $f_Y(y)$ (with respect to the Lebesgue measure on $\mathbb{R}$).

Under these assumptions, the regression function can be written as

$$\gamma(x) = \frac{\int_{\mathbb{R}} f(x / y) f_Y(y) y dy}{f_X(x)},$$

where

$$f_X(x) = \int_{\mathbb{R}} f(x/y)\, f_Y(y)\, dy.$$

Hence, given an estimate $\widehat{f}(x/y)$ of $f(x/y)$, the following (plug-in) estimate of $\gamma(x)$ can be constructed from the previous equation:

$$\widehat{\gamma}(x) = \frac{\sum_{i=1}^{n} \widehat{f}(x/y_i)\, y_i}{\widehat{f}_X(x)}, \tag{2.3}$$

where

$$\widehat{f}_X(x) = \sum_{i=1}^{n} \widehat{f}(x/y_i).$$

2.2 SPECIFICATION IN THE GAUSSIAN CASE

The general estimator given in Equation (2.3) will be here specified for the case where, for each $y \in \mathbb{R}$, $P(\cdot/y)$ is a Gaussian measure on $\mathcal{X} = \mathcal{L}_2[0,1]$. $P(\cdot/y)$ is then determined by its corresponding mean function $\mu(y) \in \mathcal{X}$ (which is then equal to $F(y)$ according to Equation (2.2)) and a covariance operator Γ (not depending on y), which is a symmetric and positive Hilbert-Schmidt operator on the space $\mathcal{X}$. Thus, there exists an eigenvalue decomposition of Γ, $(\varphi_j, \lambda_j)_{j\geq 1}$ such that $(\lambda_j)_j$ is a decreasing sequence of positive real numbers, $(\varphi_j)_j$ take values in $\mathcal{X}$ and $\Gamma = \sum_j \lambda_j \varphi_j \otimes \varphi_j$ where $\varphi_j \otimes \varphi_j(h) = \langle \varphi_j, h\rangle \varphi_j$ for any $h \in \mathcal{X}$.
Denote by P_0 the Gaussian measure on $\mathcal{X}$ with zero mean and covariance operator Γ. Assume that the following usual regularity condition holds: for each $y \in \mathbb{R}$,

$$\sum_{j=1}^{\infty} \frac{\mu_j^2(y)}{\lambda_j} < \infty,$$

where

$$\mu_j(y) = \langle \mu(y), \varphi_j \rangle.$$

Then, $P(\cdot/y)$ and P_0 are equivalent Gaussian measures, and the density $f(\cdot/y)$ has the explicit form:

$$f(x/y) = \exp\left\{ \sum_{j=1}^{\infty} \frac{\mu_j(y)}{\lambda_j} \left(x_j - \frac{\mu_j(y)}{2} \right) \right\},$$

where $x_j = \langle x, \varphi_j \rangle$ for all $j \geq 1$. This leads to the following algorithm to estimate $f(x/y)$:

1. Obtain an estimate $\widehat{\mu}(\cdot)(t)$ of the function $y \to \mu(y)(t)$ for each $t \in \mathbb{R}$. This may be carried out trough any standard nonparametric method for univariate regression based on the data set $(y_i, x_i(t))_{i=1,\ldots,n}$, e.g., a smoothing kernel method:

$$\hat{\mu}(y) = \frac{\sum_{i=1}^{n} K\left(\frac{y_i - y}{h}\right) x_i}{\sum_{i=1}^{n} K\left(\frac{y_i - y}{h}\right)} \tag{2.4}$$

   as proposed in [13] (note that, in this case, the bandwidth h has a common value for all t).

2. Obtain estimates $(\widehat{\varphi}_j, \widehat{\lambda}_j)_j$ of the eigenfunctions and eigenvalues $(\varphi_j, \lambda_j)_j$ of the covariance Γ on the basis of the empirical covariance of the residuals $\hat{e}_i = x_i - \widehat{\mu}(y_i)$, $i = 1, ..., n$

$$\hat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n} (\hat{e}_i - \bar{e}) \otimes (\hat{e}_i - \bar{e})$$

with $\bar{e} = \frac{1}{n}\sum_{i=1}^{n} e_i$. Only the first p eigenvalues and eigenfunctions are retained, where $p = p(n)$ is a given integer, smaller than n.

3. Estimate $f(x/y)$ by

$$\widehat{f}(x/y) = \exp\left\{\sum_{j=1}^{p} \frac{\widehat{\mu}_j(y)}{\widehat{\lambda}_j}\left(\widehat{x}_j - \frac{\widehat{\mu}_j(y)}{2}\right)\right\}, \tag{2.5}$$

where $\widehat{\mu}_j(y) = \langle\widehat{\mu}(y), \widehat{\varphi}_j\rangle$ and $\widehat{x}_j = \langle x, \widehat{\varphi}_j\rangle$.

Finally, substituting (2.5) into (2.3) leads to an estimate $\widehat{\gamma}(x)$ of $\gamma(x)$, which will be referred to as the DBIR estimator. It can be proved that this estimator is consistent ($\lim_{n\to\infty} \hat{\gamma}(x) =^{\mathbb{P}} \gamma(x)$) under mild regularity assumptions.

3. A SIMULATION STUDY

In this section, the feasibility and the performance of the nonparametric functional regression method described in Section 2. is discussed through a simulation study. All the simulations were done using Matlab [20]. The algorithm for the DBIR method was also implemented in Matlab. The FDA functions [8], developed by Ramsay et. al. were used for some operations with functional data. The DBIR code is available upon request.

3.1 Data generation

The data were simulated in the following way: values for the real random variable Y were drawn from a uniform distribution in the interval $[0, 10]$. Then, X was generated by 2 different models or settings:

M1 $X = Yv_1 + 2Yv_2 + 3Yv_5 + 4Yv_{10} + e$,

M2 $X = \sin(Y)v_1 + \log(Y+1)v_5 + e$,

where $(v_i)_{i\geq 1}$ is the trigonometric basis of $\mathcal{X} = \mathcal{L}^2([0,1])$ (i.e., $v_{2k-1} = \sqrt{2}\cos(2\pi kt)$, and $v_{2k} = \sqrt{2}\sin(2\pi kt)$), and e a Gaussian process independent of Y with zero mean and covariance operator $\Gamma_e = \sum_{j\geq 1}\frac{1}{j}v_j \otimes v_j$. More precisely, e was simulated by using a truncation of Γ_e, $\Gamma_e(s,t) \simeq \sum_{j=1}^{q}\frac{1}{j}v_j(t)v_j(s)$ by setting $q = 500$. From these two designs, training and a test samples were simulated with sizes $n_L = 300$ and $n_T = 200$ respectively.

Figures 1 and 2 give examples of realizations of X obtained under the first and the second model, respectively, for three different values of y. The underlying (non noisy) functions, $F(y)$ is also represented on them. In the case of model **M2**, the simulated data have a high level of noise which makes the estimation a hard task.

To apply the DBIR method, the discretized functions X were approximated by a continuous function using a functional basis expansion. Specifically, the data were approximated using 128 B-spline basis functions of order 4: Figures 1 and 2 show the comparison between the raw functions and their B-spline approximation.

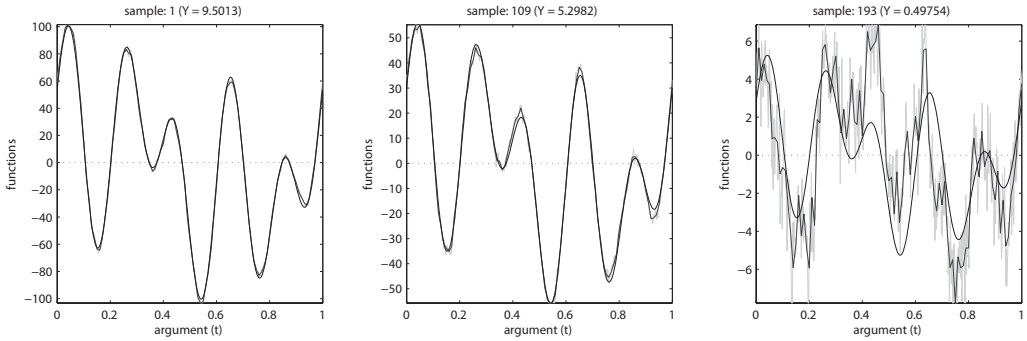

Figure 1: Model **M1**. True function $F(y)$ (smooth continuous line), simulated data X (gray rough line) and approximation of X using B-splines (rough black line) for three different values of y.

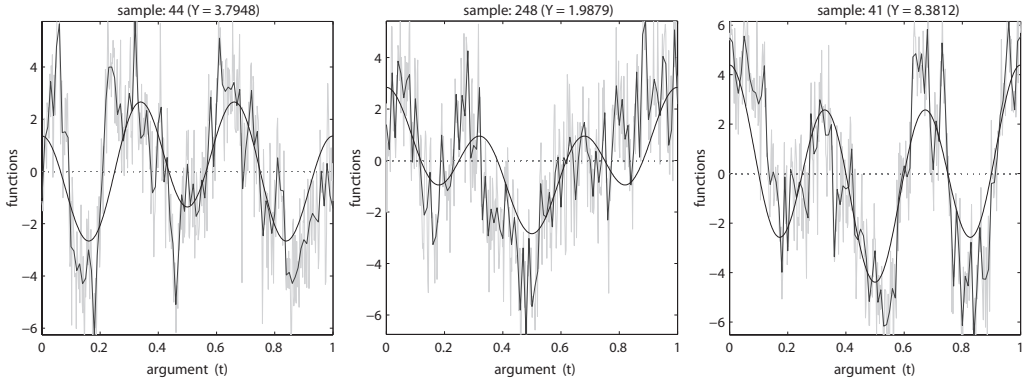

Figure 2: Model **M2**. True function $F(y)$ (smooth continuous line), simulated data X (gray rough line) and approximation of X using B-splines (rough black line) for three different values of y.

3.2 Simulation results

The conditional mean $\mu(y)$ was estimated by a kernel smoothing (such as in Equation (2.4)). Two hyper-parameters were to be tuned in this approach: the bandwidth parameter for the estimation of $\mu(y)$ and the number, p, of eigenfunctions involved in Equation (2.5). These two parameters were selected by a 10-fold cross-validation minimizing the mean square error (MSE) criterion on the training sampling.

3.2.1 Linear case: M1

This section gives the results obtained in the first simulated model, which presents a linear relation between the real explanatory variable and functional response variables in the inverse model of Equation (1.1). From Figure 1, it can be noticed that the level of noise in the data is greater for small values of Y.

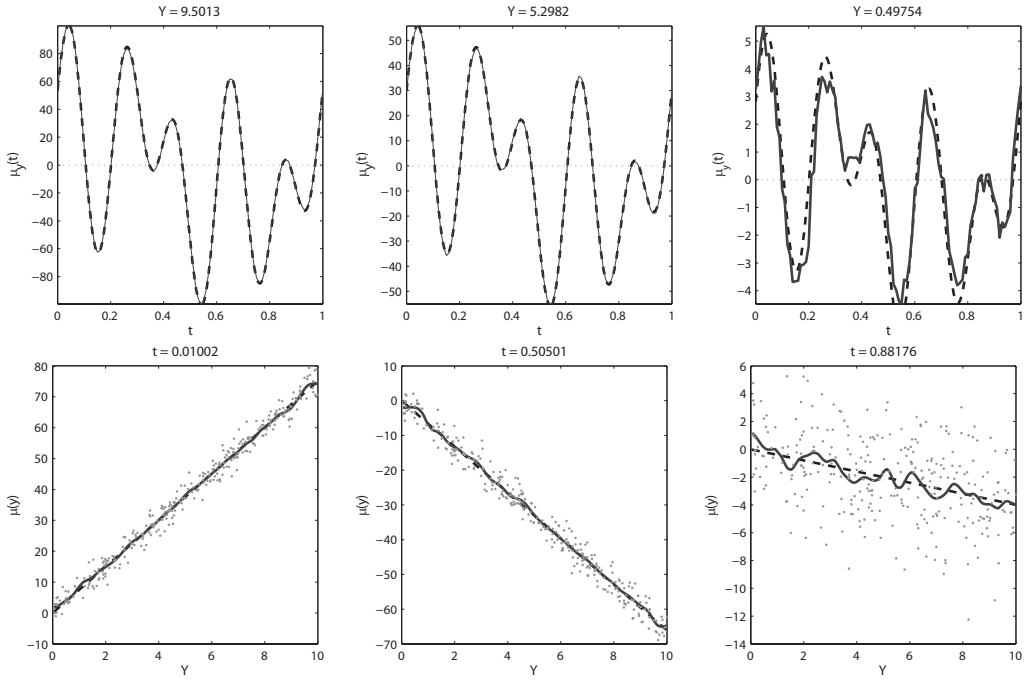

Figure 3: Model **M1**. Top: True (discontinuous lines) and estimated (continuous lines) values of $F(y)$ for different y values. Bottom: true and estimated values of $F(\cdot)(t)$ for different values of t (bottom). The dots (bottom) represents the simulated data $(x_i(t))_i$ in the training set.

The estimation of the conditional mean using a kernel estimator (Equation (2.4)) is shown in Figure 3. A comparison between true values and estimated values of $F(y)(t)$ are given for various values of y (top) and for various values of t (bottom). The linearity of the inverse model of Equation (1.1) is illustrated by the linear $F(y)$ in the bottom part of this figure. In general, the estimates are good but, in some cases (e.g., bottom right) the level of noise appears to be too high and the true mean (as a function of y) it is not as well estimated as in the other cases.
Figure 4 shows the estimated eigendecomposition of the empirical covariance of residuals r (Section 2.2,

step 2) and the predicted values of Y on the training and test sets by using the DBIR estimator. More specifically, the comparison between the true and the estimated eigenfunctions are shown in Figure 4 (a-c) and the comparison between the true and the estimated eigenvalues in Figure 4 (d). The results show very good predictions in both the training and test sets (Figure 4 (e-f)).

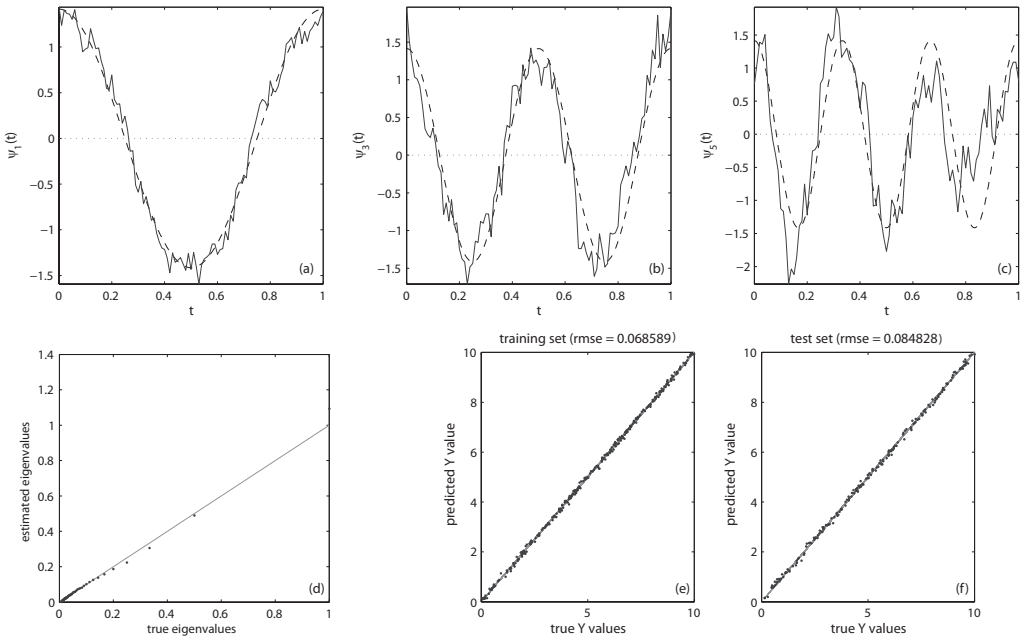

Figure 4: Model **M1.** (a-c): True (dashed line) and estimated eigenfunctions (continuous line); (d): estimated vs. true eigenvalues and (e-f): predicted values for Y vs. the true ones for training and test sets.

3.2.2 Nonlinear case: M2

For this nonlinear model **M2**, Figure 2 showed that the level of noise is much higher than for **M1**. In the same way as in the previous section, Figure 5 compares the true $F(y)(t)$ to its estimated values for various values of y (top) and for various values of t (bottom). The results are very satisfactory given the fact that the data have a high level of noise (which clearly appears in the bottom of this figure).
Figure 6 shows the results of the steps 2-3 of the estimation scheme: the estimated eigendecomposition of r is compared to the true one, and the predicted value for Y are compared to the true ones, both on training and test sets. The estimation of the eigendecomposition is also very satisfactory despite the high level of noise, and the comparison between training and test sets shows that the method does not overfit the data.

4. CONCLUSION

A new functional nonparametric regression approach has been introduced motivated by the calibration problems in Chemometrics. The new method, named functional Density-Based Inverse Regression (DBIR)

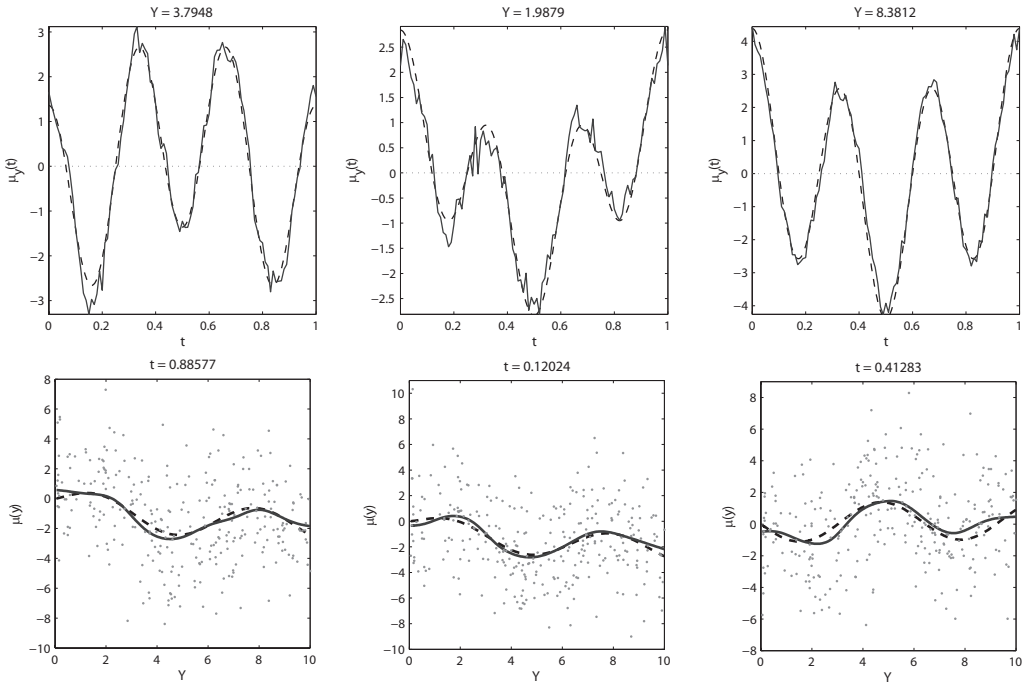

Figure 5: Model **M2**. Top: True values (discontinuous lines) and estimates (continuous lines) of $F(y)$ for various values of y. Bottom: true values and estimates of $F(\cdot)(t)$ for various values of t (bottom). The dots (bottom) are the simulated data $(x_i(t))_i$ in the training set.

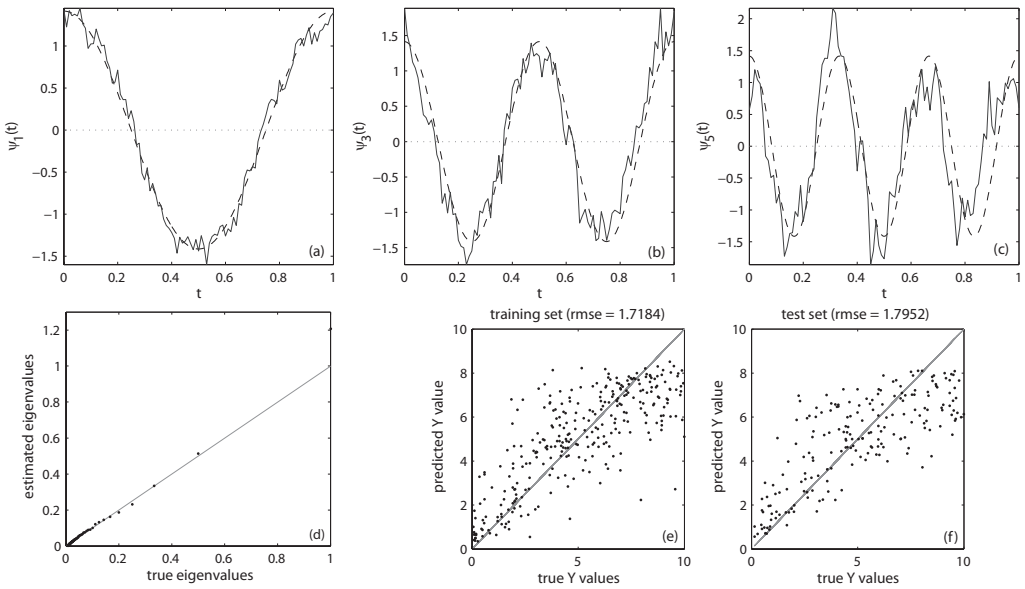

Figure 6: Model **M2**: (a-c): True (dashed line) and estimated eigenfunctions (continuous line); (d): estimated vs. true eigenvalues and (e-f): predicted values for Y vs. the true ones for training and test sets.

was fully described for the sample space $\mathcal{X} = L_2([a, b])$ under a Gaussian assumption for the conditional law $P(\cdot/Y)$ but it can be extended to other sample spaces and distribution families. Two appealing features of the new method are its rather mild model assumptions and its computational simplicity. The simulation study of DBIR has shown that it performs well for both linear and nonlinear models. Thus, DBIR can be considered as a promising functional regression methods, particularly appealing for calibration problems.

RECEIVED JULY 2010
REVISED NOVEMBER 2010

References

[1] Biau, G., Cérou, F., and Guyader, A. (2010): Rates of convergence of the functional k-nearest neighbor estimate **IEEE Transactions on Information Theory**, 56:2034–2040.

[2] Cardot, H., Crambes, C., Kneip, A., and Sarda, P. (2007): Smoothing spline estimators in functional linear regression with errors in variables **Computional Statistics and Data Analysis**, 51:4832–4848.

[3] Cardot, H., Ferraty, F., and Sarda, P. (1999): Functional linear model **Statistics and Probability Letter**, 45:11–22.

[4] Cardot, H., Ferraty, F., and Sarda, P. (2003): Spline estimators for the functional linear model **Statistica Sinica**, 13:571–591.

[5] Crambes, C., Kneip, A., and Sarda, P. (2008): Smoothing splines estimators for functional linear regression **The Annals of Statistics**, 37:35–72.

[6] Dauxois, J., Ferré, L., and Yao, A. (2001): Un modèle semi-paramétrique pour variable aléatoire hilbertienne **Compte Rendu de l'Académie des Sciences de Paris, Série I (Mathématique)**, 327(I):6947–952.

[7] DiNardo, J. and Tobias, J. L. (2001): Nonparametric density and regression estimation **The Journal of Economic Perspectives**, 15:11–28.

[8] FDA functions. Available at **ftp://ego.psych.mcgill.ca/pub/ramsay/FDAfuns**.

[9] Ferraty, F. and Vieu, P. (2006): **Nonparametric Functional Data Analysis: Theory and Practice (Springer Series in Statistics)** Springer-Verlag, New York.

[10] Ferré, L. (1998): Determining the dimension in sliced inverse regression and related methods **Journal of the American Statistical Association**, 93:132–140.

[11] Ferré, L. and Villa, N. (2006): Multi-layer perceptron with functional inputs : an inverse regression approach **Scandinavian Journal of Statistics**, 33:807–823.

[12] Ferré, L. and Yao, A. (2003): Functional sliced inverse regression analysis **Statistics**, 37:475–488.

[13] Ferré, L. and Yao, A. (2005): Smoothed functional inverse regression **Statistica Sinica**, 15:665–683.

[14] Hastie, T. and Mallows, C. (1993): A discussion of a statistical view of some chemometrics regression tools by i. e. frank and j. h. friedman **Technometrics**, 35:140–143.

[15] Hernández, N., Biscay, R. J., and Talavera, I. (2007): Support vector regression methods for functional data **Lecture Notes in Computer Science**, 4756:564–573.

[16] Laloë, T. (2008): A k-nearest neighbor approach for functional regression **Statistics and Probability Letters**, 78:1189–1193.

[17] Li, K. (1991): Sliced inverse regression for dimension reduction **Journal of the American Statistical Association**, 86:316–327.

[18] Lwin, T. and Maritz, J. S. (1980): A note on the problem of statistical calibration **Journal of the Royal Statistical Society. Series C**, 29:135–141.

[19] Marx, B. D. and Eilers, P. H. (1999): Generalized linear regression on sampled signals and curves: a P-spline approach **Technometrics**, 41:1–13.

[20] Matlab R2009a. The mathworks **Inc., Natick, MA**.

[21] Osborne, C. (1991): Statistical calibration: A review **International Statistical Review**, 59:309–336.

[22] Preda, C. (2007): Regression models for functional data by reproducing kernel Hilbert space methods **Journal of Statistical Planning and Inference**, 137:829–840.

[23] Preda, C. and Saporta, G. (2005): PLS regression on stochastic processes **Computional Statistics and Data Analysis**, 48:149–158.

[24] Ramsay, J. and Silverman, B. (2005): **Functional Data Analysis** Springer, New York, second edition.

[25] Rossi, F. and Conan-Guez, B. (2005): Functional multi-layer perceptron: a nonlinear tool for functional data anlysis **Neural Networks**, 18:45–60.

[26] Rossi, F., Delannay, N., Conan-Guez, B., and Verleysen, M. (2005): Representation of functional data in neural networks **Neurocomputing**, 64:183–210.

[27] Samb, R. (2010): Nonparametric kernel estimation of the probability density function of regression errors using estimated residuals **arXiv:1010.0439v1 [math.ST]**.