

# Proposal of heuristic regression method applied in descriptive data analysis: case studies

Flávio A. Gomes, Alfredo de O. Assis, Márcio R. da C. Reis, Viviane M. Gomes, Sóstenes G. M. Oliveira, Wanderson R. H. de Araujo, Wesley P. Calixto

**Abstract**—The purpose of this paper is to use the hybridized optimization method in order to find mathematical structures for analysis of experimental data. The heuristic optimization method will be hybridized with deterministic optimization method in order to that structures found require not knowledge about data generated experimentally. Five case studies are proposed and discussed to validate the results. The proposed method has viable solution for the analysis of experimental data and extrapolation, with mathematical expression reduced.

**Index Terms** - regression, heuristic, modeling, optimization.

## I. INTRODUCTION

This paper is an extended version of our paper published in 2016 IEEE 16th International Conference on Environment and Electrical Engineering [1]. Traditionally, researches show the need to express the variable behavior through functions that represent experimental data. In several areas, regression methods are used to establish the relationship between variables, such as in the image processing [2], analysis of concrete structures [3] [4], extraction of tone of voice [5], health area [6] and waste flow forecasting [7].

To [8], the regression analysis consists in the study of the dependence between variables, verifying the relationship of the explanatory variables towards the dependent variable to perform forecasts and previews. This study is necessary due the existent lack of knowledge of the algebraic expression that rules the system being analyzed.

The absence of the function that describes the behavior of the system implies in simulations or experiments performing in order to define the outputs, every time the inputs are changed. Several times, this requires time and effort, which can make the process of study the system unpractical. The experiments (real or simulated) provide, as output, discrete data, however, in most cases, there is needed a function that describes the data in a continuous way [9].

Once the function that defines the system is found, many analyses can be performed, such as data prediction, which tries to obtain an output for a certain correspondent input beyond the predefined interval [10]. In case of forecasting of natural resources demand, the efficient use can be obtained based on the performed predictions.

In many situations, even simulations take a considerable amount of time, making the system analysis process difficult. In order to solve this problem, we use regression to replace part of the system by an expression that represents it, decreasing

the simulation time. In [1], a regression of collected data on a test bench of controlled rectifiers was performed.

Regression methods use techniques that seek flexibility and predictive capacity. Many studies base themselves on polynomials and trigonometric functions for approximating data. However, regressions by hybrid functions, polynomial and trigonometric, present themselves more representative than each of them apart, overcoming limitations as the periodicity for polynomial or prediction for trigonometric series [11]. Other methods are used for prediction and curve fitting, such as Artificial Neural Networks in [3] and [4], which got better results than quadratic regressions and additive models in [7], which are compared to *cubic smoothing splines*.

Researches about regression seek effective parametrization methods, in order to improve the curve fitting. In [12] is used *Darwinian Particle Swarm Optimization*, *P-Spline* method in [13], regularized algorithm of Levinson-Galerkin in [6], the least squares method to parametrize trigonometric series in [5] and in [14] has the solving of compound optimization criterion through weighted polynomial regression models.

The purpose of this work is to present a methodology to determine mathematical expressions that represent the systems with the least number of possible terms. The main contribution is to reduce the edge effect due to the reduced number of terms. Besides that, it contributes to the recognition of systems from the experimental data and also in assertive extrapolation at considerable intervals.

The proposed methodology is based on the generalization of the power and trigonometric series and the application of optimization methods. Section II presents the theoretical background, Section III brings the proposed methodology and the results achieved are presented in Section IV.

## II. BACKGROUND

According to [15], a bounded-input, bounded-output system (BIBO) is stable when it is limited in respect of the space's norm in which it is defined ( $L_2$ ,  $L_\infty$ ). Using the space:

$$L_2(\Omega) = \{f(t) \mid \|f(t)\|_2 < \infty\}, \quad (1)$$

The norm of (1) is defined by  $\|f(t)\|_2 = \sqrt{\int_{\Omega} |f(t)|^2 dt}$ , where  $\Omega$  is a subinterval in the real numbers and  $f(t)$  is a square-integrable function in  $\Omega$ . By analyzing the experimental data  $f_{ex}(x)$  of a BIBO system, we have according to [16] that the collected data are represented by:

$$f_{ex}(x) = f_{op}(x) + \epsilon, \quad (2)$$

Since  $f_{op}(x)$  represents the regression and  $\epsilon$  is the random additive error of the process that does not depend on "x" and satisfies the homoscedasticity criterion, which is, that the variance of  $\epsilon$  is constant. In this sense it is said that  $f_{op}(x)$  is the regression that represents the system if the mean square error (MSE) is as minimal as possible. Therefore, the following optimization problem is generated:

$$f_{op}(x) = \min_{x \in \Omega} \{ \|\epsilon\|_2^2 \}, \quad (3)$$

where  $f_{op}(x)$  depends on the used base for data interpolation.

For the representation of these events, there is a wide collection of interpolation and extrapolation theories, being the polynomial approximation of Weierstrass the main interpolation theorem. In this, it is shown that in the space of the continuous functions  $C_{[a,b]} \subset L_2[a,b]$  any function  $f \in C_{[a,b]}$ , where  $a, b \in \mathbf{R}$ , can be approximated by a polynomial function [17]. Extending its definition to the space of the analytic functions  $f \in C_{(-\infty, \infty)}$ , any function can be expressed as a power Series.

The standard methods vary from polynomial to trigonometric representations, using the base  $\beta_1$  for the power series or polynomial, given by (4), and the base  $\beta_2$  for the trigonometric series, given by (5).

$$\beta_1 = \{1, x, \dots, x^n, \dots\} \quad (4)$$

and

$$\beta_2 = \left\{ 1, \sin\left(\frac{\pi x}{p}\right), \sin\left(\frac{2 \cdot \pi x}{p}\right), \dots, \sin\left(\frac{n\pi x}{p}\right), \dots, \cos\left(\frac{\pi x}{p}\right), \cos\left(\frac{2 \cdot \pi x}{p}\right), \dots, \cos\left(\frac{n\pi x}{p}\right), \dots \right\} \quad (5)$$

The obtained approximations verify trends and represent data by means of functions [18]. Thus, the regression methods are chosen depending on the characteristics of the problem. The bases  $\beta_1$  and  $\beta_2$  have properties of representation in the space of continuous functions in the interval  $[a, b]$ . When there is some kind of frequent oscillation, the base  $\beta_1$  is insufficient to extrapolate the polynomial regression interval, since to represent the trigonometric frequencies, there is the need to transform the polynomial regression into a series. However, the extrapolation problem is also present in the base  $\beta_2$ , since it has limitations for data prediction for Non-periodic functions [11] [5].

### III. METODOLOGY

The proposed methodology will use hybridized optimization method (heuristic and deterministic) to determine parameters of predefined structures. Based on experimental data, the optimization process will return the mathematical expression that will represent the dynamics of the system, as Fig. 1.

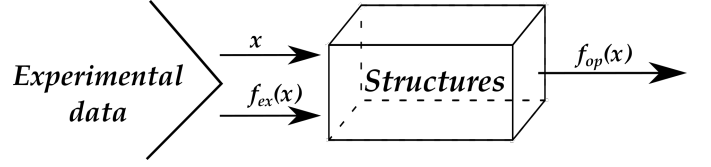


Figure 1. Flow of optimization process.

These structures, based on polynomial, trigonometric, and exponential functions, enable to represent a significant amount of curves. Regression will be performed by comparing the curve defined by the experimental data  $f_{ex}$  with the curve generated by structures, called optimized curve  $f_{op}$ . Structures that generalize the power and trigonometric series given by  $f_{op1}$ ,  $f_{op2}$  and  $f_{op3}$  will be proposed in order to meet the different curve profiles. These structures are presented in expressions (6), (7) and (8), respectively.

$$f_{op1} = a_0 + \sum_{i=1}^n a_i \cdot x^{b_i} \quad (6)$$

$$f_{op2} = a_0 + \sum_{i=1}^n a_i \cdot x^{b_i} \cdot \cos(c_i \cdot x + d_i) \quad (7)$$

$$f_{op3} = a_0 + \sum_{i=1}^n a_i \cdot x^{b_i} \cdot \cos(c_i \cdot x + d_i) \cdot \exp(e_i \cdot x) \quad (8)$$

where:  $a_0, a_i, b_i, c_i, d_i, e_i \in \mathbf{R}$ .

Unlike other methods [11] [14], the parameters of  $f_{op}$  will assume values belonging to the set of real numbers. Therefore, polynomials of the  $\beta_1$  base from the power series will be generalized to rational functions, well as trigonometric functions with fixed frequencies of the  $\beta_2$  base will be generalized to any real frequency. Thus, it will be possible to express experimental data with smaller structures, compared to other regression methods, maintaining the power of prediction.

Based on the characteristics of experimental curve  $f_{ex}$ , the proposed methodology will select the structure that have greater proximity between the optimized curve  $f_{op}$  and the experimental  $f_{ex}$ . Thus, the optimization process will be applied following the expression (3), but due to the fact of working with discrete signals of finite duration in the optimization process, the calculus of approximation error or evaluation function  $F_{aval}$  will be given by:

$$F_{aval} = \sqrt{\sum_{i=1}^n (f_{ex_i} - f_{op_i})^2} \quad (9)$$

where:  $n$  will be the number of  $f_{ex}$  points.

Before performing the regression, data set will be processed in order to select the characteristic intervals  $I_k$  to assist in the optimization process, that will express the orderly domain  $J$  of the  $f_{ex}$  curve in

$$J = I_1 \cup I_2 \cup \dots \cup I_k \quad (10)$$

where:  $k$  will be the number of intervals.

The first regression interval will be the one that contains the initial point of  $f_{ex}$  curve. The method will be applied successively by the union of subsequent intervals given by expression (10). In order to define the intervals, experimental curve will be divided into parts, based on inflection points and variation at the ordinates axis.

The inflection point is the main factor for choosing the structure and also the optimization method. This occurs because this concept is related to the change in the function's variation rate, being characterized by the point at which the derivative of the function changes from increasing to decreasing and vice versa.

This feature influences both at the choice of structure and the improvement of the optimization process. Due to the fact that structures with several inflection points tend to be more oscillatory, this parameter directly influences the choice of structure that best fits the data. If we analyze the optimization aspect, by dividing the interval based on inflection points reduces the possibility of stopping the process in some local optimum point. In this sense, the way of choosing the structures from the simple characteristic of the experimental data is defined.

The amount of inflection points will be the base parameter to define those intervals. If there are until 2 inflection points, it will mean that data set have no oscillatory characteristic. Therefore, data set will be divided into 10 equal parts and the intervals will be chosen based on variation at ordinate axis on these parts. The highest variation will be chosen as reference and the set of intervals ( $J$ ) of (10) will be compound by only those that will achieve variation higher than 30% in relation to the chosen reference.

If there are 3 or more inflection points, it will mean that data set presents sinuosity and its analysis will be based on these oscillations. Thus, the highest variation at ordinates axis for all set will be chosen as reference. The subinterval between first 3 inflection points will be chosen to check the higher variation at ordinates axis present in this subset. If this variation exceeds 5% of reference variation, then this subinterval will be selected as the set of interval ( $J$ ) of (10) for analysis in the optimization process. If this variation does not overcome that percentage, the subinterval will be grouped with other more relevant. The following inflection points will continue being analyzed in search of variations that meet this restriction.

These intervals will be passed for the optimization routine that hybridizes the heuristic methods, Genetic Algorithm, and deterministic, Nelder-Mead, in order to find the optimized parameters [19]. At the end, the result will be the values of structures parameters proposed and their respective evaluation functions  $F_{aval}$  of data set. The best result will selected and the parameter values will be replaced at the corresponding structure with the view to mount the function that describes the set of experimental data.

## IV. RESULTS

In order to generate the set of experimental data, known and used functions have been used to evaluate regression processes in mathematics and statistics. These functions do not represent physical systems and still present problems of mapping by both interpolating polynomials and extrapolations. These functions were used as case studies as well as data collected from a test bench of controlled rectifiers. This choice was done due to: i) the possibility to perform extrapolation of original set, ii) the approximation error with the results obtained at the initial simulation can be measured, and iii) the success of optimization process can be verified.

### A. Case Study 1

The generating function of experimental data chosen for this first case study was given by:

$$f_{ex} = \frac{1}{1+x^2} \quad (11)$$

This function was chosen because of presenting oscillation problem near the edges of interval analyzed using polynomial interpolation with polynomials of high order. This problem is known as Runge phenomenon like cited in [20]. In the expression (11),  $x$  assumes 1000 values in the range  $1 \leq x \leq 100$ . The smallest error was got by the structure that contains only polynomials derived from (6) and the eleven terms of final expression was given by:

$$\begin{aligned} f_{op} = & -6.08 \cdot 10^{-5} + 1.30 \cdot x^{-2.07} \\ & + 5.62 \cdot 10^{-5} \cdot x^{-1.52} - 0.81 \cdot x^{-2.89} \\ & - 9.54 \cdot 10^{-4} \cdot x^{-0.77} + 6.21 \cdot 10^{-4} \cdot x^{-0.40}. \end{aligned} \quad (12)$$

Fig. 2 illustrates experimental and optimized curves obtained with  $F_{aval} = 1.25 \cdot 10^{-2}$ . In the same figure there is a cut at the point 75 showing the difference between both curves with instantaneous error of about  $10^{-4}$ .

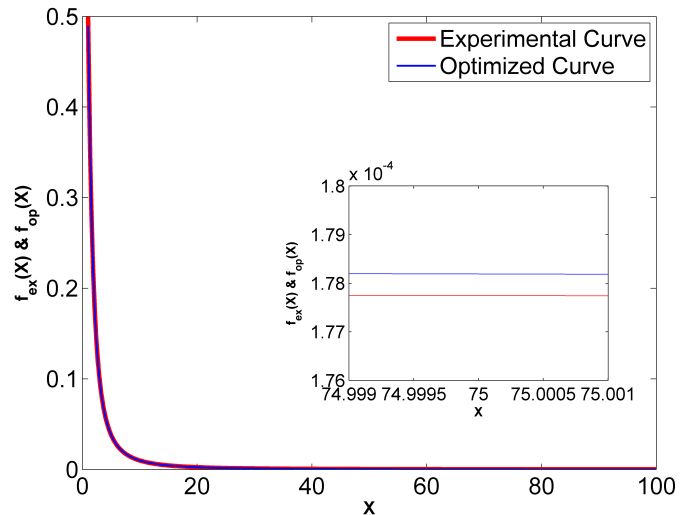


Figure 2. Case study 1.

### B. Case Study 2

For the second case study, the generating function of the chosen experimental data was given by:

$$f_{ex} = \sin(2 \cdot x + 3) \cdot \exp(-0.5 \cdot x) \quad (13)$$

This function was chosen because of presenting a difficult behavior to be mapped by the structures (6) and (7). It presents also different oscillations throughout data set analyzed. In the expression (13),  $x$  assumes 1000 values in the range  $0 \leq x \leq 40$ . The smallest error was got by the most complete structure that contains polynomials, cosine, and natural exponential derived from (8). The eleven terms found of final expression was given by:

$$\begin{aligned} f_{op} = & 2.18 \cdot 10^{-9} - 1.00 \cdot x^{4.27 \cdot 10^{-7}} \\ & \cdot \cos(1.99 \cdot x - 1.71) \cdot \exp(-0.50 \cdot x) \\ & - 4.61 \cdot 10^{-12} \cdot x^{1.12} \cdot \cos(-0.39 \cdot x + 1.48) \\ & \cdot \exp(0.14 \cdot x). \end{aligned} \quad (14)$$

Fig. 3 illustrates experimental and optimized curves obtained with  $F_{aval} = 0.14$ . In the same figure there is a cut at the point 30 showing the difference between both curves with instantaneous error of about  $10^{-8}$ .

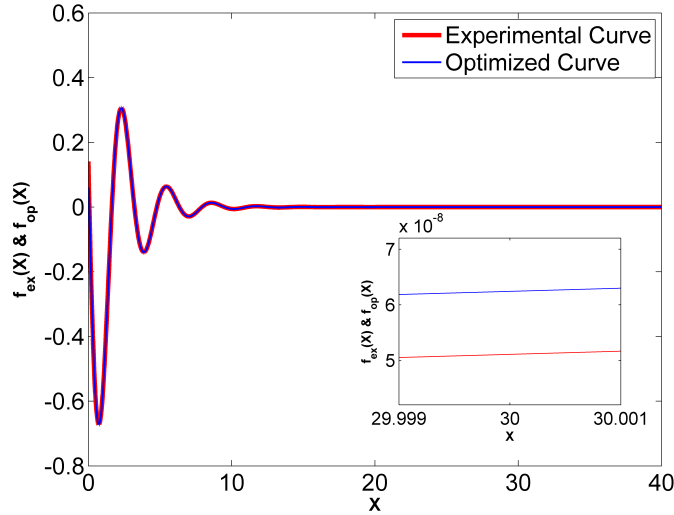


Figure 3. Case study 2.

### C. Case Study 3

The chosen generating function of the experimental data for this third case study was given by:

$$f_{ex} = x + \frac{1}{\tan(x)} \quad (15)$$

This function was chosen because it presents output data with negative values, increasing oscillation and also in order to compare with polynomial interpolation methods. In (15),  $x$  assumes 20 values in the interval  $1 \leq x \leq 20$ . The smallest error was obtained by the structure that has polynomials and

cosines (7) and the 25 terms of the final expression was given by (16).

$$\begin{aligned} f_{op} = & 19.16 + 44.95 \cdot x^{-0.11} \cdot \cos(6.31 \cdot x + 10.42) \\ & + 11.32 \cdot x^{-0.79} \cdot \cos(2.15 \cdot x + 3.10) \\ & + 2.96 \cdot 10^{-6} \cdot x^{4.73} \cdot \cos(2.17 \cdot x + 8.90) \\ & - 5.85 \cdot 10^{-6} \cdot x^{3.72} \cdot \cos(4.31 \cdot 10^{-3} \cdot x - 3.51 \cdot 10^{-3}) \\ & + 6.47 \cdot 10^{-3} \cdot x^{1.90} \cdot \cos(-0.22 \cdot x - 1.22) \\ & - 5.19 \cdot 10^{-5} \cdot x^{2.97} \cdot \cos(0.71 \cdot x + 19.87). \end{aligned} \quad (16)$$

Fig. 4 illustrates the experimental and optimized curves obtained with  $F_{aval} = 1.97 \cdot 10^{-1}$ . Within the same figure, there is a cut at the point  $x = 3$ , which illustrates the difference between both curves, with the order of the distance between them of approximately  $10^{-2}$ .

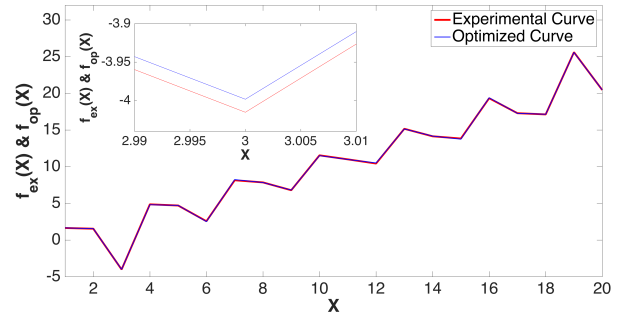


Figure 4. Case study 3.

Polynomial interpolations were also performed to the same generating function in (15) in order to compare the proposed method and this technique of curve fitting. Two polynomials were found, one being 20 degree in (17) and the other nine degree in (18).

$$\begin{aligned} f_{pol20} = & 3.01 \cdot 10^{-13} \cdot x^{19} - 6.29 \cdot 10^{-11} \cdot x^{18} \\ & + 6.10 \cdot 10^{-9} \cdot x^{17} - 3.65 \cdot 10^{-7} \cdot x^{16} \\ & + 1.51 \cdot 10^{-5} \cdot x^{15} - 4.57 \cdot 10^{-4} \cdot x^{14} \\ & + 1.05 \cdot 10^{-2} \cdot x^{13} - 1.87 \cdot 10^{-1} \cdot x^{12} \\ & + 2.61 \cdot x^{11} - 28.80 \cdot x^{10} + 2.52 \cdot 10^2 \cdot x^9 \\ & - 1.74 \cdot 10^3 \cdot x^8 + 9.43 \cdot 10^3 \cdot x^7 - 3.96 \cdot 10^4 \cdot x^6 \\ & + 1.27 \cdot 10^5 \cdot x^5 - 3.01 \cdot 10^5 \cdot x^4 + 5.07 \cdot 10^5 \cdot x^3 \\ & - 5.67 \cdot 10^5 \cdot x^2 + 3.73 \cdot 10^5 \cdot x - 1.06 \cdot 10^5. \end{aligned} \quad (17)$$

$$\begin{aligned} f_{pol9} = & -2.09 \cdot 10^{-7} \cdot x^9 + 1.84 \cdot 10^{-5} \cdot x^8 \\ & - 6.80 \cdot 10^{-4} \cdot x^7 + 1.38 \cdot 10^{-2} \cdot x^6 \\ & - 1.69 \cdot 10^{-1} \cdot x^5 + 1.30 \cdot x^4 - 6.25 \cdot x^3 \\ & + 18.52 \cdot x^2 - 29.4 \cdot x + 18.01. \end{aligned} \quad (18)$$

Fig. 5 illustrates the experimental and optimized curves by the proposed method and by the polynomials in (17) and (18).

The approximation error of the proposed method was  $F_{aval} = 1.97 \cdot 10^{-1}$ , whereas using the polynomial of 20 degree the error was  $F_{aval} = 2.03 \cdot 10^1$  and the polynomial of nine degree with error of  $F_{aval} = 4.67 \cdot 10^2$ .

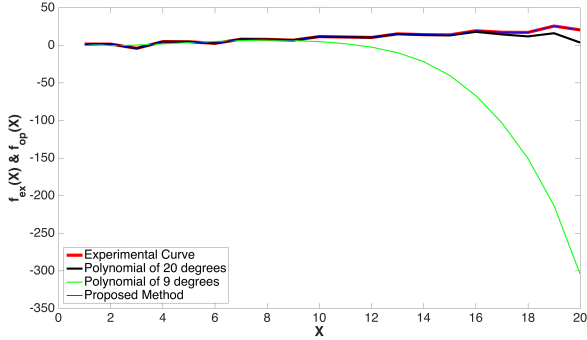


Figure 5. Proposed method and polynomial interpolation comparison.

#### D. Case Study 4

In this case study, the errors of extrapolations made for the previous case studies were calculated in order to verify the efficiency of the proposed method. In addition to reduction of terms of the expressions found, the extrapolations showed that the curve fitting captured the essence of the systems studied. The case study of section IV-A was extrapolated until point 300 in order to show the curve fitting after the original interval. Fig. 6 illustrates the experimental and optimized curves. The measured error for the new interval was  $F_{aval} = 1.13 \cdot 10^{-2}$  and within the same Fig. 6 there is a cut at the point  $x = 280$ , which illustrates the difference between both curves with the order of the distance between them being approximately  $10^{-5}$ .

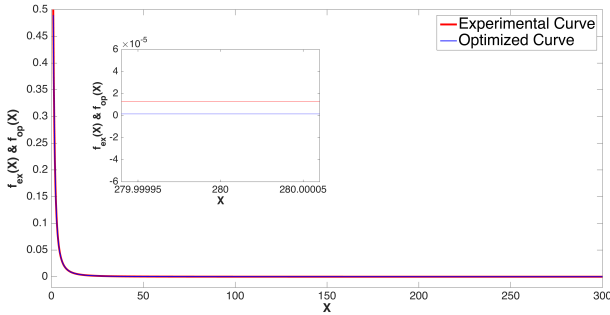


Figure 6. Extrapolation of the case study 1.

For the case study of section IV-B, the extrapolation was performed both before and after the initial interval. In Fig. 7, the explanatory variable  $x$  takes on values in the new interval  $-15 \leq x \leq 60$  and again, it can be noticed that (14) follows the behavior of the experimental data curve. The measured error for the new interval was  $F_{aval} = 2.36 \cdot 10^{-2}$  and within the same Fig. 7, there is a cut close to the point  $x = -11.84$ , which illustrates the difference between the two curves, being the order of distance between them approximately  $10^{-3}$ .

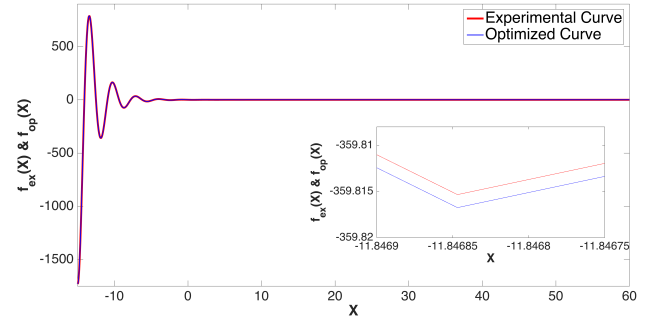


Figure 7. Extrapolation of the case study 2.

For the case study of section IV-C, the extrapolation was performed a little after the initial interval, since the approximation error of the curves by the methods becomes difficult to be perceived graphically. The nine degree polynomial in (18) was unable to adjust the curve in the original interval, remaining in the extrapolation process. The 20 degree polynomial in (17) obtained a suitable approximation in the analyzed interval and diverged abruptly when the extrapolation occurred shortly after the original interval due to the edge effect or Runge's phenomenon [20] which is noticed in polynomial interpolations.

In Fig. 8, there are presented the experimental and optimized curves by the proposed method and by the interpolating polynomials. The explanatory variable  $x$  assumes values in the new range  $5 \leq x \leq 21$  and again, it can be noted that (16) follows the behavior of the experimental data, whereas the interpolating polynomials lose their ability of approaching. For the new interval, the measured errors by using the proposed method and (17) and (18) were  $F_{aval} = 7.27 \cdot 10^{-1}$ ,  $F_{aval} = 6.68 \cdot 10^2$  and  $F_{aval} = 6.41 \cdot 10^2$ , respectively.

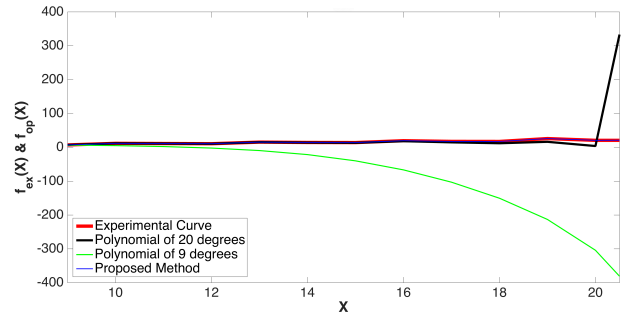


Figure 8. Extrapolation of the case study 3.

#### E. Case Study 5

At the fifth case study were analyzed data collected at a test bench for studies of controlled rectifiers. These rectifiers provide DC voltage of variable output as from a fixed AC voltage. Due to its ability to provide DC voltage continuously variable, the controlled rectifiers revolutionized the modern industrial control equipments. This converter was shown in Fig. 9.

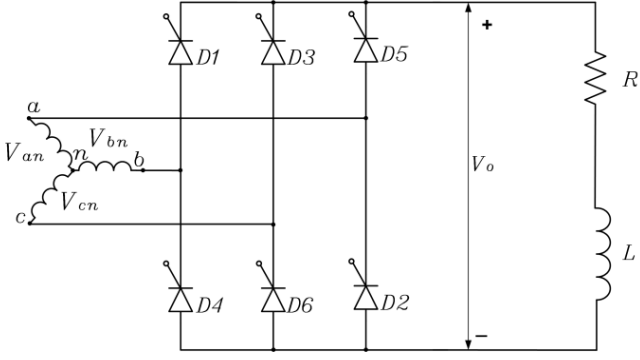


Figure 9. Power converter circuit with RL load.

In order to obtain the instantaneous value of voltage controlled output  $V_o$ , the literature has the solutions given by (19) according to [21].

$$V_o = \begin{cases} \beta\sqrt{2} V_{ab} & \text{if } \omega t \leq \frac{\pi}{6} \\ \beta \sin(\omega t + \frac{\pi}{6}) & \text{if } \frac{\pi}{6} + \alpha \leq \omega t \leq \frac{\pi}{2} + \alpha, \\ \beta \sin \omega t' & \text{if } \frac{\pi}{3} + \alpha \leq \omega t' \leq \frac{2\pi}{3} + \alpha, \end{cases} \quad (19)$$

where:  $\omega t' = \omega t + \frac{\pi}{6}$  and  $V_{ab}$  is the voltage (effective) of input line and  $\beta$  is the extinction angle of electric current described in [22].

A test bench has been developed for obtaining experimental data of the converter output voltage and the firing angles of keys. The collected data set was interpolated in order to also contain 1000 values, and then was applied the proposed method to obtain analytical expression that represent the voltage as a function just of the firing angle  $\alpha$ . The smallest error was obtained by the structure of polynomials and cosines derived from (7) and the 21 terms of found expression was given by (20):

$$\begin{aligned} f_{op} = & 263 + 40.9 \cdot x^{0.98} \cdot \cos(3.78 \cdot 10^{-4} \cdot x + 1.57) \\ & + 3.67 \cdot x^{0.16} \cdot \cos(8.20 \cdot 10^{-2} \cdot x + 3.01) \\ & - 2.84 \cdot 10^{-4} \cdot x^{2.62} \cdot \cos(-3.96 \cdot 10^{-2} \cdot x + 3.67) \\ & - 0.15 \cdot x^{0.93} \cdot \cos(0.10 \cdot x - 61.3) \\ & - 0.46 \cdot x^{3.57 \cdot 10^{-5}} \cdot \cos(0.22 \cdot x + 3.92 \cdot 10^{-2}). \end{aligned} \quad (20)$$

Fig. 10 presents the characteristic experimental curve of voltage of converter controlled three phase operating with load RL (resistor-inductor) and the optimized curve obtained. The approximation error found was  $F_{aval} = 37.6$ . The set of terms was analysed to identify the importance of each of them in the composition of encountered error. It was noticed that removing the last term in expression (20) the new value was  $F_{aval} = 42.7$ , that is, with 17 terms it still maintain an acceptable approximation error.

## V. CONCLUSION

This work presented the hybrid optimization method to be applied in the development of descriptive analysis data

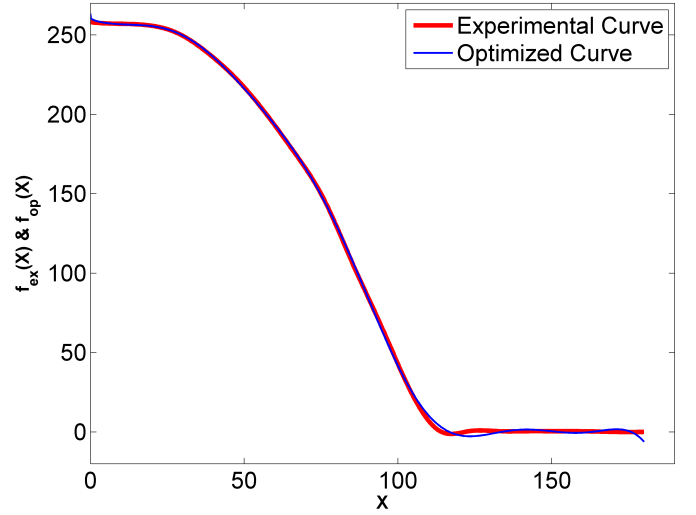


Figure 10. Case study 5.

structure. The study results indicate that the proposed method is able to formulate mathematical expressions, in the form of regression, allowing to explore the relationship between the dependent and independent or explanatory variables. The proposal finds values in the set of real numbers for the coefficients, exponents and frequency of structures that generalize the power and trigonometric series, in an attempt to minimize errors. This proposed method is able to find a continuous function expression that represents a set of experimental data described by a discrete function expression. Another advantage is the extrapolation performed in an assertive form at first and second case studies without observe problems like Runge phenomenon at the edges of analyzed sets. Researches are still being developed in order to compare the proposed method with the traditional methods of regression.

## ACKNOWLEDGMENT

The authors would like to thank Coordination for the Improvement of Higher Education Personnel (CAPES), the National Counsel of Technological and Scientific Development (CNPq) and Research Support Foundation of Goiás State (FAPEG) for financial support research and scholarships.

## REFERENCES

- [1] F. A. Gomes, V. M. Gomes, A. d. O. Assis, M. R. d. C. Reis, G. da Cruz, and W. P. Calixto, "Heuristic regression method for descriptive data analysis," 2016.
- [2] E. Garcia, R. Arora, and M. R. Gupta, "Optimized regression for efficient function evaluation," 2012.
- [3] W. H. Chien, L. Chen, C. C. Wei, H. H. Hsu, and T. S. Wang, "Modeling slump flow of high-performance concrete using a back-propagation network," 2010.
- [4] I.-C. Yeh, "Modeling slump flow of concrete using second-order regressions and artificial neural networks," 2007.
- [5] K. Steiglitz, G. Winham, and J. Petzinger, "Pitch extraction by trigonometric curve fitting," 1975.
- [6] T. Strohmmer, "A levinson-galerkin algorithm for regularized trigonometric approximation," 2000.
- [7] A. Antoniadis, I. Gijbels, and A. Verhasselt, "Variable selection in additive models using p-splines," 2012.

- [8] D. Gujarati and D. Porter, "Econometria básica - 5.ed.," 2011.
- [9] L. A. Aguirre, "Introdução à identificação de sistemas—técnicas lineares e não-lineares aplicadas a sistemas reais," 2004.
- [10] M. K. Goyal, "Monthly rainfall prediction using wavelet regression and neural network: an analysis of 1901–2002 data, assam, india," 2014.
- [11] R. L. Eubank and P. Speckman, "Curve fitting by polynomial-trigonometric regression," 1990.
- [12] M. S. Couceiro, D. Portugal, N. Gonçalves, R. Rocha, J. M. A. Luz, C. M. Figueiredo, and G. Dias, "A methodology for detection and estimation in the analysis of golf putting," 2013.
- [13] B. U. Park, E. Mammen, Y. K. Lee, and E. R. Lee, "Varying coefficient regression models: a review and new developments," 2015.
- [14] H. Dette, G. Haller, *et al.*, "Optimal designs for the identification of the order of a fourier regression," 1998.
- [15] P. T. Chen G, "Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems," 2000.
- [16] G. Broniatowski, Michel; Celant, "Interpolation and extrapolation optimal designs. 1, polynomial regression and approximation theory," 2016.
- [17] R. Glowinski, K. Atkinson, and W. Han, "Theoretical numerical analysis: A functional analysis framework," 2003.
- [18] H. A. Schilling and S. L. Harris, "Applied numerical methods for engineers using matlab," 1999.
- [19] W. P. Calixto, A. Paulo Coimbra, J. C. d. Mota, M. Wu, W. G. Silva, B. Alvarenga, L. d. C. Brito, A. J. Alves, E. G. Domingues, and D. P. Neto, "Troubleshooting in geoelectrical prospecting using real-coded genetic algorithm with chromosomal extrapolation," 2015.
- [20] L. Trefethen, *Approximation Theory and Approximation Practice*. Society for Industrial and Applied Mathematics, 2013.
- [21] M. Rashid, "Eletrônica de potência: circuitos, dispositivos e aplicações," 1999.
- [22] M. R. C. Reis, "Comparative analysis of optimization methods applied of tuning PI controller," 2014.