Interval Analysis and Optimization Applied to Parameter Estimation under Uncertainty

J. D. Gallego-Posada* M. E. Puerta-Yepes †

Abstract

We present a methodology through exemplification to perform parameter estimation subject to possible factors of uncertainty. The underlying optimization problem is posed in the framework of the theory of interval-valued optimization. The implementation of numerical procedures required to achieve efficient solutions implied the use of the $\ell 1$ norm instead of usual $\ell 2$ regression. Finally, an implementation using real data was performed, demonstrating the ability of interval analysis to encapsulate uncertainty while facing non-trivial parameter estimation problems.

1 Introduction

The majority of mathematical models developed to represent various problems depend on sets of parameters whose values are generally determined based on experimental measurements or inferred according to data obtained through observation of the studied phenomenon.

Given a set of experimental data $\{y_i\}$ (inaccurate because of the influence of randomness and uncertainty), we would like to estimate the values of those parameters, so that the model attains an acceptable fitting level to the observed data.

In the usual estimation techniques, a vector of real numbers is obtained (a single point estimation for each of the parameters involved), which causes that the probability that the result of the estimation happens to be the actual value of the parameter is fairly low and, therefore, is very sensitive to perturbations [1].

Nevertheless, the use of interval-valued analysis allows us to provide regions in which, given a determined confidence level, we can ensure the enclosing of the solution, thus reducing the impact of random disturbances and computational errors in the final result.

The foundations of interval analysis were established by Moore [1] in his PhD dissertation. Based on that work and further developments provided by Skelboe [2], Hansen [3] and Stroem [4], among others, several authors (Ratschek and Voller [5], Bhurjee and Panda [6]) have studied the potential of interval valued techniques in the field of optimization. On the other hand, authors like Ichida and Fujii [7] or Karmakar and Bhunia[8] have

*Mathematical Engineering student, jgalle29@eafit.edu.co, Universidad EAFIT, Medellín, Colombia.

†PhD. in Mathematical Sciences, Research Group in Functional Analysis and Applications, mpuerta@eafit.edu.co, Universidad EAFIT, Medellín, Colombia.

[1]**Key words:** optimization, interval-valued analysis, parameter estimation and inverse problems

developed significant research related to the implementation of multiobjective techniques including interval valued optimization. However, there is a lack of efforts in the use of these procedures in the context of parameter estimation which is one of the principal aims of this research work.

Thus, the uncertainty in those values due to measurement errors, rounding in calculations, computational representability and even possibly because of a lack of knowledge of qualitative information associated with those parameters, pose interval analysis and, particularly, interval-valued optimization as a very useful tool, leading to the development and establishment of robust techniques when facing the mentioned difficulties, obtaining more reliable and rigorous results in a mathematical perspective.

A methodology to perform parameter estimation under any kind of uncertainty is presented. The underlying optimization problem is posed in the framework of the theory of interval-valued optimization. Numerical results show that estimations obtained allow us to describe successfully the behaviour of the modelled phenomenons by enclosing adequately the uncertainty and sensitivity of the model, overcoming possible difficulties originated in errors generated while measuring or modelling, even in cases with low information availability.

In particular, the first numerical example allowed us to compare the efficacy of the optimization algorithms using an adaptation of the $\ell 1$ norm in the interval system using the Hausdorff metric, instead of usual $\ell 2$ regression, related to the Euclidean metric space. Information related to normed spaces can be found in [9] and [10].

This paper is organized as follows. In Section 2 we introduce the main concepts of interval arithmetic and analysis, in order to present a concise approach to this theory for the reader. In Section 3, we review some theoretical results related to interval-valued optimization. Finally, Section 4 presents the numerical results obtained for theoretical and real applications of interval optimization techniques.

2 Interval Arithmetic and Analysis

With the aim of presenting a self contained article, this section describes in general terms the notation, arithmetic and analytic components and some theorems related to interval computation.

Let $\mathscr{K}_c(\mathbb{R})$ denote the set of all non-empty, compact and convex subsets of $\mathbb{R}$. Let A, $B \in \mathscr{K}_c(\mathbb{R})$ and let $\odot \in \{+, -, \cdot, /\}$ be a binary operation on $\mathbb{R}$, e.g., addition, subtraction, multiplication and division.

Let us denote by $\mathscr{I}$ the class of all closed and bounded intervals in $\mathbb{R}$. For an interval A we adopt the notation $A = \left[a^L, a^U\right]$, where a^L and a^U mean the infimum and supremum of A, respectively. Throughout this paper, upper case non italic-letters represent intervals and lower case non italic-letters represent real number. Whenever the context may not be clear, proper comments shall be made.

2.1 Set Operations

Since, intervals are essentially sets of real numbers, it is often useful to define operations related to their behaviour as sets.

Definition 2.1. *Let A and $B \in \mathscr{I}$. The intersection of two intervals A and B is empty if either $b^U < a^L$ or $a^U < b^L$. Otherwise,*

$$A \cap B := \{x : x \in A \wedge x \in B\} = \left[max\left\{a^L, b^L\right\}, min\left\{a^U, b^U\right\}\right]$$

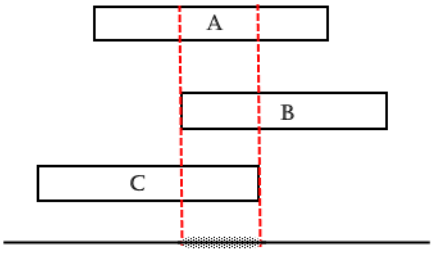

Figure 1: Intersection of intervals.

In particular, intersection plays a key role in interval analysis. If we have two intervals containing a result of interest, regardless of how they were obtained, then the intersection, which may be narrower (see Figure 1), also contains the result. Further set operations between intervals are described in [1] and [11].

2.2 Interval Properties

Recalling the notation $A = \left[a^L, a^U\right]$, we can describe some of the main properties of an interval in term of its endpoints, as can be seen in Figure 2.

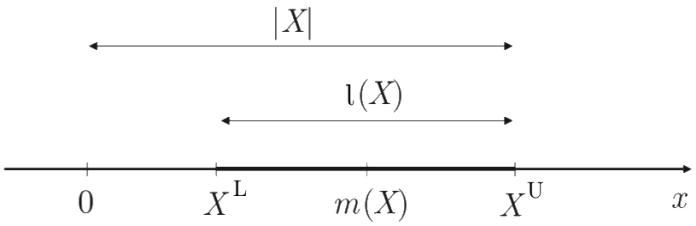

Figure 2: Geometric properties of an interval.

Definition 2.2. *Let $A \in \mathscr{I}$. We define the length of A as*

$$l(A) := a^U - a^L$$

Definition 2.3. *Let $A \in \mathscr{I}$. The absolute value of A is defined by*

$$|A| := max\left\{\left|a^L\right|, \left|a^U\right|\right\}$$

Definition 2.4. *Let $A \in \mathscr{I}$. The midpoint of A is given by*

$$m(A) := \frac{a^L + a^U}{2}$$

2.3 Arithmetic Operations

Let A, $B \in \mathscr{I}$. Let $\odot \in \{+, -, \cdot, /\}$ be a binary operation on $\mathbb{R}$, e.g., addition, subtraction, multiplication and division. The operation $A \odot B$ is defined by:

$$A \odot B := \{a \odot b : a \in A,\, b \in B\}$$

For instance, for the case of the sum of two intervals, the set $A + B$ can be explicitly described in terms of the endpoints of such interval:

$$A + B = \left[a^L + b^L,\, a^U + b^U\right]$$

The *Hukuhara difference*, $\ominus$, is a special kind of subtraction between two intervals [12], particularly important for the definition of differentiation of interval-valued functions.

$$A \ominus B = \left[a^L - b^L,\, a^U - b^U\right] \text{ provided } a^L - b^L \leq a^U - b^U$$

A complete list of endpoint formulas for the usual arithmetic operations and more details on the topic of interval computation can be found in [1] and [11].

There is a close relation in the mathematical structure of $\mathscr{I}$ and $\mathbb{R}$: for example, $\mathscr{I}$ is commutative and associative under addition and multiplication, the cancellation law for addition holds and there exist identity elements for such operations. However, we caution that, for $A \in \mathscr{I}$, $-A$ is not in general an additive inverse for A, since

$$A - A = [a^L, a^U] + [-a^U, -a^L] = \left[a^L - a^U,\, a^U - a^L\right] = l(A) \cdot [-1, 1]$$

Similarly, it can be proved that $A/A = [1, 1]$ only of $l(A) = 0$. Therefore, we do not have additive or multiplicative inverses except for degenerate intervals. However, we always have the inclusions $0 \in A - A$ and $1 \in A/A$ [1].

Example 2.1. *Consider the interval $A = [1, 3]$. It is straightforward to prove that $A - A = [-2, 2] = 2 \cdot [-1, 1] \ni 0$ and $A/A = [1/3, 1] \ni 1$.*

Because of this lack of inverse elements under addition, $\mathscr{I}$ can not constitute a vector space by itself. However, the work from Radstroem [13] develops the theory of an extension set via equivalence relations in which a commutative semigroup in which the law of cancellation holds, as is indeed true in $\mathscr{I}$, can be embedded in a vector space N where the product λA for $\lambda \geq 0$ coincides with the one given on $\mathscr{I}$.

This allows us to ensure the existence of the family $\mathscr{I}^m$, which represents the set of m-dimensional vectors with entries in $\mathscr{I}$, as well as the natural extension to sets of interval-valued matrices.

2.4 Interval Analysis

In order to build an analytical structure on $\mathscr{I}$ it is fundamental to define the notion of a metric between intervals. Hausdorff [14] proposed a metric between subsets X and Y of a metric space E, given by:

$$d_H(X,Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y}\|x-y\|, \sup_{y\in Y}\inf_{x\in X}\|x-y\|\right\}$$

where $\|\cdot\|$ is the distance defined on the metric space $(E, \|\cdot\|)$.

In particular, the Hausdorff norm induces a metric for the interval system that can be expressed in terms of the midpoints of the selected intervals $A = [a^L, a^U]$, $B = [b^L, b^U]$, as follows:

$$d_H(A,B) = \max\left\{|a^L - b^L|, |a^U - b^U|\right\}$$

Given the metric space $(\mathscr{I}, d_H(\cdot))$, it is possible to define the usual notions of convergence and limit in this space.

Definition 2.5. *Let $\{A_n\}$ and $A \in \mathscr{I}$. We say that the sequence of intervals $\{A_n\}$ converges to A, denoted by $\lim_{n\to\infty} A_n = A$, if, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that, for $n \geq N$, we have $d_H(A_n, A) < \epsilon$.*

Lemma 2.1. *Convergence in $\mathscr{I}$ can be reduced to the usual convergence in $\mathbb{R}$, that is,*

$$\lim_{n\to\infty} A_n = A \text{ if and only if } a_n^L \to a^L \wedge a_n^U \to a^U$$

2.4.1 Interval-valued Functions

Definition 2.6. *The function $f : \mathbb{R}^n \to \mathscr{I}$ defined on an Euclidean space $\mathbb{R}^n$ is called an interval-valued function. This function can also be written as $f(\mathbb{x}) = [f^L(\mathbb{x}), f^U(\mathbb{x})]$, where f^L and f^U are real-valued functions defined on $\mathbb{R}^n$ and satisfy $f^L(\mathbb{x}) \leq f^U(\mathbb{x})$ for every $\mathbb{x} \in \mathbb{R}^n$.*

This definition of interval-valued functions allows us to handle a wide range of situations in which we can describe a function of this nature in terms of two real valued functions, but also we could present a function whose parameters are intervals in $\mathscr{I}$.

Definition 2.7. *For $\mathbb{c} \in \mathbb{R}^n$, we write $\lim_{\mathbb{x}\to\mathbb{c}} f(\mathbb{x}) = A$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that, for $\|\mathbb{x} - \mathbb{c}\| < \delta$, we have $d_H(f(\mathbb{x}), A) < \epsilon$. In this case, we say A is the limit of f when $\mathbb{x}$ tends to $\mathbb{c}$.*

Theorem 2.1. *Let f be an interval-valued function defined on $\mathbb{R}^n$ and $A = [a^L, a^U] \in \mathscr{I}$. Then $\lim_{\mathbb{x}\to\mathbb{c}} f(\mathbb{x}) = A$ if and only if $\lim_{\mathbb{x}\to\mathbb{c}} f^L(\mathbb{x}) = a^L$ and $\lim_{\mathbb{x}\to\mathbb{c}} f^U(\mathbb{x}) = a^U$.*

Definition 2.8. *Let f be an interval-valued function defined on $\mathbb{R}^n$. We say that f is continuous at $\mathbb{c} \in \mathbb{R}^n$ if $\lim_{\mathbb{x}\to\mathbb{c}} f(\mathbb{x}) = f(\mathbb{c})$.*

Further information regarding interval-valued functions, their properties, and proofs of related theorems can be found in [15].

3 Karush-Kuhn-Tucker Optimality Conditions

3.1 Order Relations

Since, in the context of optimization its is necessary to compare the images of interval-valued functions to be minimized (maximized), a corresponding group of partial order relations can be defined for intervals. These relations are illustrated in Figures 3 - 5.

Definition 3.1. *Let* $A = [a^L, a^U]$ *and* $B = [b^L, b^U] \in \mathscr{I}$. *It is possible to express* A *as a function of its centre and width, as* $A = \langle m(A), w(A) \rangle$.

- $A \preceq_{LU} B$ *if and only if* $a^L \leq b^L$ *and* $a^U \leq b^U$
- $A \preceq_{UC} B$ *if and only if* $a^U \leq b^U$ *and* $m(A) \leq m(B)$
- $A \preceq_{CW} B$ *if and only if* $m(A) \leq m(B)$ *and* $w(A) \leq w(B)$

Figure 3: Order relations — $\preceq_{LU}$.

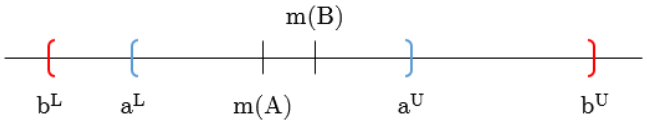

Figure 4: Order relations — $\preceq_{UC}$.

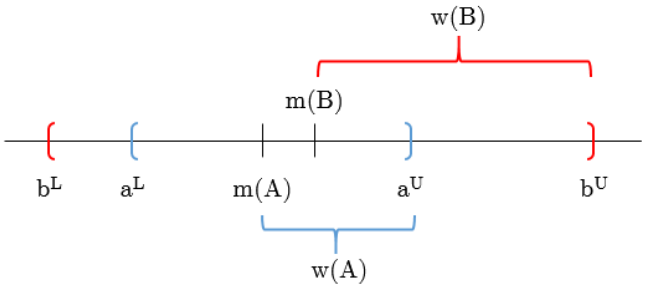

Figure 5: Order relations — $\preceq_{CW}$.

In the minimization case, the $\preceq_{CW}$ order is particularly interesting because it allows us to obtain the interval with the minimum midpoint and also the smaller width, which in this case can be seen as the variance or uncertainty in the measurement or evaluation of the function.

3.2 Differentiability

Differentiation is one of the key concepts in optimization. because the qualities of the derivatives of a function provide very relevant information related to minima and extrema values of this function.

The extension of differentiability to interval-valued function is based on the Hukuhara difference presented in Section 2.3 and, under certain conditions, keeps a very strong relation to usual real-valued functions derivatives, establishing a connection between Definitions 3.2 and 3.3.

Definition 3.2. *Let X be an open set in $\mathbb{R}$. An interval-valued function $f : X \to \mathscr{I}$ with $f(x) = [f^L(x), f^U(x)]$ is called weakly differentiable at $x_0 \in X$ if the real valued functions f^L and f^U are differentiable at x_0 (in the usual sense).*

Definition 3.3. *Let X be an open set in $\mathbb{R}$. We say $f : X \to \mathscr{I}$ is H-differentiable (strongly differentiable) at $x_0 \in X$ if there exists $A(x_0) \in \mathscr{I}$ such that*

$$\lim_{h \to 0^+} \frac{f(x_0 + h) \ominus f(x_0)}{h} \text{ and } \lim_{h \to 0^+} \frac{f(x_0) \ominus f(x_0 - h)}{h}$$

both exist and are equal at $A(x_0)$. Then $A(x_0)$ is the H-derivative of f at x_0.

The following theorem presents the conditions in which a function is (or not) *H-differentiable* at a point.

Theorem 3.1. *(Wu [15]) Let X be an open set in $\mathbb{R}$ and $f : X \to \mathscr{I}$ an interval-valued function defined on X. Suppose that f is weakly differentiable at $x_0 \in X$ with derivatives $(f^L)'(x_0) = \hat{a^L}(x_0)$ and $(f^U)'(x_0) = \hat{a^U}(x_0)$. Then,*

1. *If $f^L(x_0 + h) - f^L(x_0) \leq f^U(x_0 + h) - f^U(x_0)$ and $f^L(x_0) - f^L(x_0 - h) \leq f^U(x_0) - f^U(x_0 - h)$ for every $h > 0$, then f is H-differentiable at x_0 with H-derivative $A(x_0) = [\hat{a^L}(x_0), \hat{a^U}(x_0)]$.*

2. *If $\hat{a^U}(x_0) > \hat{a^L}(x_0)$, then f is H-non-differentiable at x_0.*

Based on Theorem 3.1, the next corollary is an immediate consequence given the definition of $\hat{a^L}(x_0)$ and $\hat{a^U}(x_0)$, and summarizes the concept of differentiability of interval-valued functions.

Corollary 3.1. *Let X be an open set in $\mathbb{R}$. If an interval-valued function $f : X \to \mathscr{I}$ is H-differentiable (strongly differentiable) at $x_0 \in X$ with H-derivative $A(x_0) = [a^L(x_0), a^U(x_0)]$, then f is weakly differentiable at x_0. Furthermore, we have $(f^L)'(x_0) = a^L(x_0)$ and $(f^U)'(x_0) = a^U(x_0)$.*

3.3 KKT Optimality Conditions

In optimization, the Karush-Kuhn-Tucker (KKT) conditions are first order necessary conditions for a solution in non-linear programming to be optimal, provided that some regularity conditions are satisfied [16]. The conditions derived in [15] allow us to determine the optimality of a feasible solution of an interval valued optimization problem.

Definition 3.4. *Let $f(\mathbb{x}) = [f^L(\mathbb{x}), f^U(\mathbb{x})]$ be an interval-valued function defined on a convex set $X \subseteq \mathbb{R}^n$. We say that f is LU-convex at $\mathbb{x}^* \in X$ if*

$$f(\lambda\mathbb{x}^* + (1-\lambda)\mathbb{x}) \prec_{LU} \lambda f(\mathbb{x}^*) + (1-\lambda) f(\mathbb{x})$$

for each $\lambda \in (0,1)$ and each $\mathbb{x} \in X$. It is possible to define UC-convexity in a similar fashion.

Remark 3.1. *Let X be a convex subset of $\mathbb{R}^n$ and f be an interval-valued function defined on X. The function f is LU-convex at $\mathbb{x}^* \in X$ if and only if f^L and f^U are convex (in the usual sense) at $\mathbb{x}^* \in X$.*

Definition 3.5. *Let $\mathbb{x}^*$ be a feasible solution, i.e., $\mathbb{x}^* \in X$. We say that $\mathbb{x}^*$ is a type-I solution [type-II solution] of an interval-valued optimization problem if there exists no $\overline{\mathbb{x}} \in X$ such that $f(\overline{\mathbb{x}}) \prec_{LU} f(\mathbb{x}^*)$ [or $f(\overline{\mathbb{x}}) \prec_{CW} f(\mathbb{x}^*)$].*

Theorem 3.2. *(Wu [15]) Consider the **interval valued** optimization problem (IVOP) given by*

$$min\ f(\mathbb{x}) = [f^L(x_1, ..., x_n), f^U(x_1, ..., x_n)] = [f^L(\mathbb{x}), f^U(\mathbb{x})]$$

$$subject\ to\ \ g_i(\mathbb{x}) \leq 0$$

where the real-valued constraint functions $g_i : \mathbb{R}^n \to \mathbb{R}$ are convex on $\mathbb{R}^n$ for $i = 1, ..., m$.

*Suppose that the **interval-valued** objective function $f : \mathbb{R}^n \to \mathscr{I}$ is **LU**-convex and **weakly** continuously differentiable at $\mathbb{x}^* \in \mathbb{R}^n$. If there exist (Lagrange) multipliers $0 < \lambda^L, \lambda^U \in R$ and $0 \leq \mu_i \in \mathbb{R}$, $i = 1, ..., m$, such that*

1. *$\lambda^L \nabla f^L(\mathbb{x}^*) + \lambda^U \nabla f^U(\mathbb{x}^*) + \sum_{i=1}^m \mu_i \nabla g_i(\mathbb{x}^*) = \mathbb{0}$ and*
2. *$\mu_i g_i(\mathbb{x}^*) = 0$ for all $i = 1, ..., m$*

*then $\mathbb{x}^*$ is a **type-I and type-II**, i.e. optimal under the selected order relation, solution of (IVOP) problem.*

Remark 3.2. *Note the existent resemblance between Theorem 3.2 an the usual KKT conditions for real-valued optimization, which is indeed equivalent to the last theorem removing the bold-face conditions.*

Remark 3.3. *The selected Lagrange multipliers $\lambda^L, \lambda^U \in R$ act as coefficients of a weighted sum applied to the components of the objective function f, thus reducing the dimensionality of the problem.*

4 Parameter Estimation using Interval-Valued Optimization

4.1 Generalized Polynomial Fitting

Let $c_i = [c_i^L, c_i^U] \in \mathscr{I}$ for $i \in \mathbb{N}$. We say $p(x)$ is an interval-valued polynomial if it can be expressed in the form

$$p(x) = \sum_{i=0}^{n} c_i \cdot x^i = \sum_{i=0}^{n} [c_i^L, c_i^U] \cdot x^i$$

Consider a set of observations $y_i = [y_i^L, y_i^U] \in \mathscr{I}$ for $i \in \{1, ..., m\}$. We can model this phenomenon using a n degree polynomial in a matrix form as follows:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbb{Y} = \mathbb{V}\mathbb{C} + \mathbb{E}$$

where $\mathbb{V}$ is called the Vandermonde matrix.

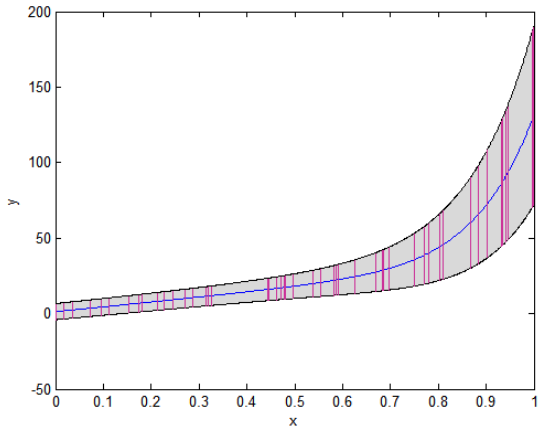

Figure 6: Interval-valued polynomial graphic.

A polynomial of degree 10 with random valued interval coefficients was generated and a random sample, purple intervals, was extracted, see Figure 6. With this information, the goal was to estimate the original values of the coefficients that generated this behaviour.

As it is usual in parameter estimation, and even more in polynomial fitting, because of its desirable computational speed and simplicity, our first attempt to achieve that goal was to perform an ordinary least square (OLS) estimation, in which the coefficients of a *real-valued polynomial* given *real observations* can be estimated by:

$$\mathbb{Y} = \mathbb{V}\mathbb{C} \quad \Rightarrow \quad \widehat{\mathbb{C}} = \left(\mathbb{V}^T\mathbb{V}\right)^{-1}\mathbb{V}^T\mathbb{Y}$$

This result can be obtain by minimizing the ℓ_2-norm of the residuals between the midpoints of fitted model and the real measurements.

$$\min \sum_{i=1}^{m} [m(y_i) - m(\widehat{y_i})]^2$$

Nevertheless this results provides good results in term of an unbiased estimation of the midpoints of the coefficients of the polynomial, there can exist very high overestimations in the length of the interval (see Figure 7) thus not allowing the identification of the real sensitivity of each parameter in the model.

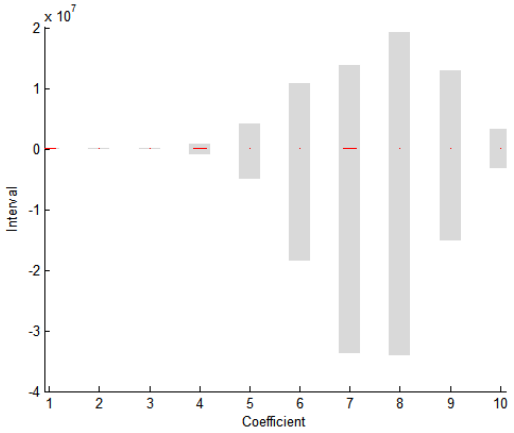

Figure 7: Parameter estimation result using OLS.

In Figure 7, the red bars represent the real values of the parameters, while the gray bars illustrate the value of the estimated coefficient. A good estimation should, of course, show little discrepancies between the gray and red bars for every coefficient, achieving an enclosure of the parameter, thus being unbiased in midpoint and length.

With the aim of reducing the mentioned difficulties, an heuristic differential evolution algorithm was implemented. Differential Evolution (DE) was developed originally by Price [17] while trying to solve a Chebychev polynomial fitting problem proposed by Storn. A complete description of DE can be found in [17].

The fundamental idea behind DE is the generation of vector through linear combinations of elements in the current population. Then, given an individual in the population, this element is compared with the trial vector generated via linear combination and, if the trial element has a better performance in terms of the objective function to be minimized (maximized), a crossing in the components of the original vector is made under a defined probability level.

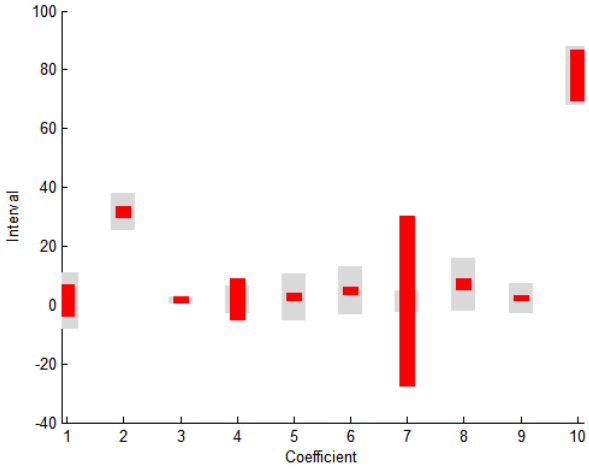

Figure 8: Parameter estimation result using Differential Evolution.

The estimations obtained using the heuristic search are presented in Figure 8. There is an improvement in the quality of the estimation according to the magnitude of the estimations in relation with the real values of the parameters. In most of the estimations there is an almost unbiased estimation in the midpoint of the intervals. However, some of the coefficients are underestimated or overestimated in the length of the interval, as can be seen in coefficients 7 and 8. Additionally, given the nature of the heuristic, the quality of the estimations is not very uniform and in some cases, the search does not converge to adequate values of the parameters.

As another alternative, we used the software implementation CVX for convex optimization developed in [18] [19]. To avoid the overestimation of the interval length, the metric induced in $\mathscr{I}$ by the ℓ_1 norm was used to measure the residuals, which can be expressed in terms of the Hausdorff distance in $\mathscr{I}$. In this way the optimization problem can be stated as follows

$$\min \sum_{i=1}^{m} d_H\left(y_i, \widehat{y}_i\right)$$

Given the definition of this metric, at an optimal point we can ensure an unbiased estimation of both endpoints, which is equivalent to a lack of midpoint-length bias.

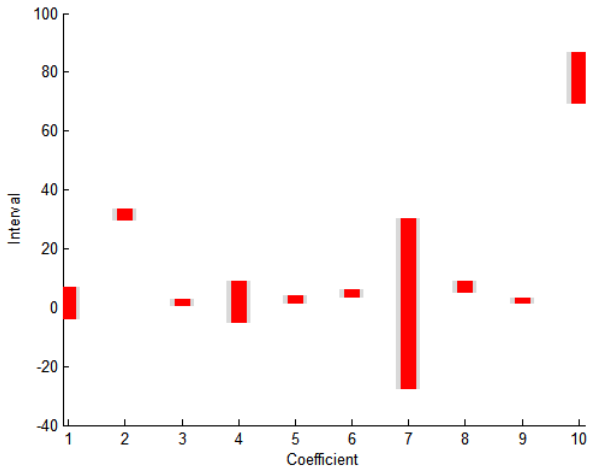

Figure 9: Parameter estimation result using CVX .

The results of this methodology are presented in Figure 9. As can be seen, the estimations agree successfully with the real values of the parameters, with errors of magnitude 10^{-9} in relation to the theoretical endpoints. Thus, allowing us to model closely the observed behaviour using an interval-valued polynomial model and enclosing the possible uncertainty in the problem.

4.2 Weierstrass Function

In order to evaluate the feasibility of an estimation of the parameters of a model using real data, the used techniques were tested using data sampled from a Weierstrass function,

which is an example of a pathological real-valued function on the real line [20], given by

$$f(x) = \sum_{n=0}^{\infty} a^n \cos\left(b^n \pi x\right)$$

In this case, the objective was to extract information about the value of the coefficient a, in this case treated as an interval, based on a set of measurements for $x \in [0, 1]$, using the ℓ_1 induced metric in $\mathscr{I}$ to define the optimization problem, as shown previously. To favour the clarity of the graphic, the estimated behaviour of the model and a subset of the measurements is presented in the interval [0, 0.125].

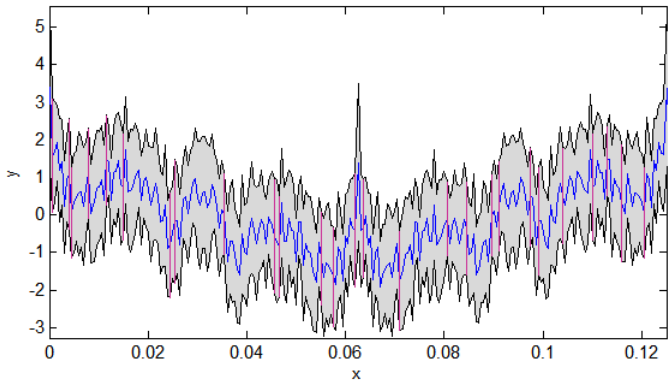

Figure 10: Weierstrass function model.

As we can see in Figure 10, the estimated coefficient for this model is able to handle the *chaotic* and *noisy* behaviour of this function, and also the extreme sensitivity that exists in this parameter, which generates changes of a large magnitude and form of the images of the function when small changes are applied to the value of a, for whom the estimation was very exact.

4.3 Fourier Series Applied to the Modelling of Spectral Power Densities

Using hydrophones, measurements of the spectral power density of the sound signals generated by vessels were performed in order to develop a characterization of such crafts. In total 36 measurements were performed, however 12 of those were discarded due to factors that generated changes in behaviour of the spectrum, for example, changes in the speed of the boat and its engines. The 24 accepted measurements are presented in Figure 11, where the horizontal axis represents the frequency in Hz and the vertical axis the spectral power density in dB/Hz. Because of confidentiality issues, the source of this data cannot be specified.

In order to describe this behaviour a Fourier series model was proposed. A Fourier series is a way to represent a wave-like function as the sum of simple sine waves, decomposing the signal into the sum of a (possibly infinite) set of simple oscillating functions, namely sines and cosines, as follows:

$$f(x) = a_0 + \sum_{i=1}^{n} a_i \cos\left(iwx\right) + b_i \sin\left(iwx\right)$$

where a_0 models a constant (intercept) term in the data and is associated with the $i = 0$ cosine term, w is the fundamental frequency of the signal, n is the number of terms (harmonics) in the series. In this case, several order models were estimated, however, an 8 order model followed the trends observed in the measurements in a much more adequate way, especially in the initial and final parts of the data.

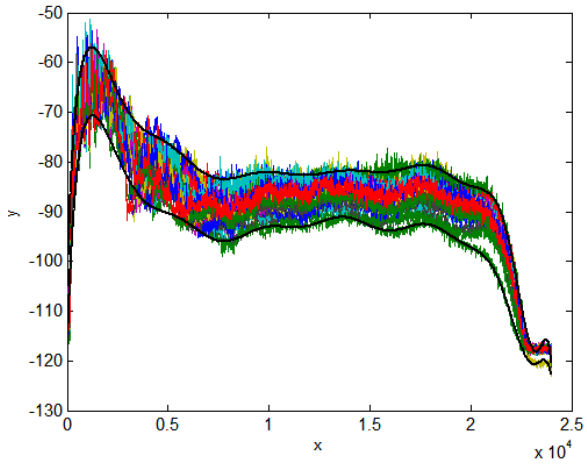

Figure 11: Real measurements - Level as a function of frequency.

Based on the complete set of measurements, the upper and lower bounds in each instant were extracted and, with this (*reduced*) information, the coefficients for the model of each of the bounding trajectories were estimated, obtaining the fits presented in Figures 12 and 13.

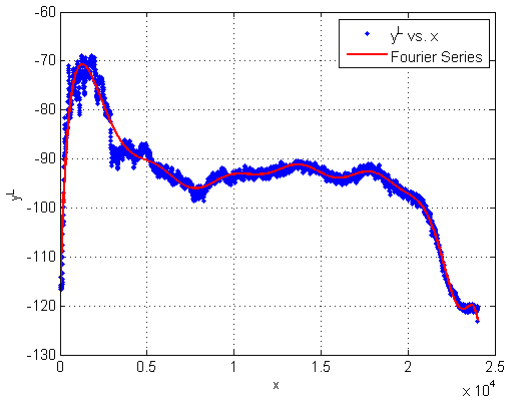

Figure 12: Fitted model lower bound.

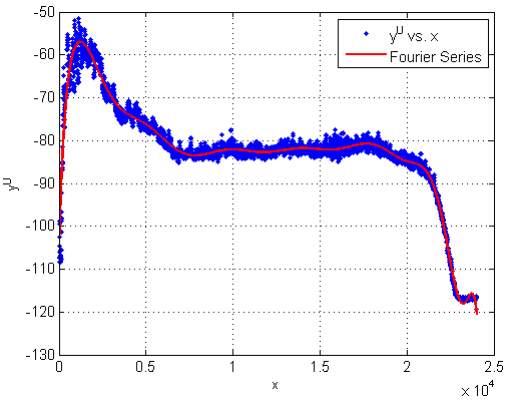

Figure 13: Fitted model upper bound.

Using these estimations an interval-valued function was proposed to enclose the volatility of the measurements using Fourier series to describe the lower and upper functions, i.e. $f : \mathbb{R} \to \mathscr{I}$, given by $f(x) = \left[f^L(x), f^U(x)\right]$, where the bounding functions can be expressed by:

$$f^L(x) = a_0^L + \sum_{i=1}^{n} a_i^L \cos\left(iw^L x\right) + b_i^L \sin\left(iw^L x\right)$$

$$f^U(x) = a_0^U + \sum_{i=1}^{n} a_i^U \cos\left(iw^U x\right) + b_i^U \sin\left(iw^U x\right)$$

An interval-valued plot generated from this model is presented in Figure 14. As can be seen in Figure 11, in which the black bold line represent the estimated bounding functions of the model, who is able to describe a significant proportion of the variance present in the measurements performed.

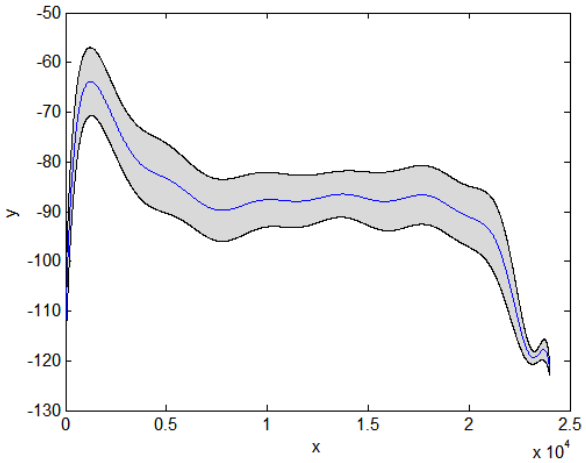

Figure 14: Interval-valued plot of the estimated Fourier series model.

It is remarkable that the effectively used information was related only to the maximum-minimum measurements in every instant. Therefore, the quality of the estimations obtained could have been equally as good as the presented in a situation of scarce available information. It is also possible to perceive a reduction in noise of the signals provided by the model, which is important in order to approximate the local behaviour of the phenomenon.

References

[1] R. E. Moore, "Interval arithmetic and automatic error analysis in digital computing," Ph.D. dissertation, Stanford, CA, USA, 1963, aAI6304614.

[2] S. Skelboe, "Computation of rational interval functions," *BIT*, vol. 14, no. 1, pp. 87–95, 1974.

[3] E. Hansen, "Global optimization using interval analysis: the multi-dimensional case," *Numerical Mathematics*, vol. 34, no. 3, pp. 247–270, 1980.

[4] T. Stroem, "Strict estimation of the maximum of a function of one variable," *BIT*, vol. 11, no. 2, pp. 199–211, 1971.

[5] H. Ratschek and R. Voller, "What can interval analysis do for global optimization?" *Journal of Global Optimization*, vol. 1, no. 2, pp. 111–130, 1991.

[6] A. Bhurjee and G. Panda, "Efficient solution of interval optimization problem," *Mathematical Methods of Operations Research*, vol. 76, no. 3, pp. 273–288, 2012.

[7] K. Ichida and Y. Fujii, "An interval arithmetic method for global optimization," *Computing*, vol. 23, no. 1, pp. 85–97, 1979.

[8] S. Karmakar and A. K. Bhunia, "An alternative optimization technique for interval objective constrained optimization problems via multiobjective programming," *Journal of the Egyptian Mathematical Society*, vol. 22, no. 2, pp. 292–303, 2014.

[9] H. Heuser, *Lehrbuch der Analysis*. Teubner, 1981.

[10] H. L. Royden, *Real analysis*, 4th ed. Pearson Education, 2010.

[11] G. Alefeld and J. Herzberger, *Introduction to interval computations*. Academic Press, 1983.

[12] M. Hukuhara, "Intégration des applications mesurables dont la valeur est un compact convexe," *Funkcialaj Ekvacioj*, no. 10, pp. 205–223, 1967.

[13] H. Radstroem, "An embedding theorem for spaces of convex sets," *Proceedings of the American Mathematical Society*, vol. 3, no. 1, pp. 165–169, 1952.

[14] F. Hausdorff, *Grundzuege der Mengenlehre*. Leipzig: Veit and Company, 1914.

[15] H.-C. Wu, "The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function," *European Journal of Operational Research*, vol. 176, no. 1, pp. 46 – 59, 2007.

[16] M. S. Bazaraa and C. M. Shetty, *Nonlinear programming*. Wiley, 1979.

[17] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[18] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," http://cvxr.com/cvx, Aug. 2012.

[19] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.

[20] G. H. Hardy, "Weierstrass's non-differentiable function," *Transactions of the American Mathematical Society*, vol. 17, no. 3, p. 301, 1916.