

Article

A Novel Non-Isotonic Statistical Bivariate Regression Method—Application to Stratigraphic Data Modeling and Interpolation

Daniele Polucci ¹, Michele Marchetti ¹ and Simone Fiori ^{2,*} 

¹ School of Computer Science and Automation Engineering, Marches Polytechnic University, White Pebbles Rd., 60131 Ancona, Italy; S1077089@studenti.univpm.it (D.P.); S1077737@studenti.univpm.it (M.M.)

² Department of Information Engineering, Marches Polytechnic University, White Pebbles Rd., 60131 Ancona, Italy

* Correspondence: s.fiori@univpm.it

Received: 13 February 2020; Accepted: 9 March 2020; Published: 10 March 2020



Abstract: The present paper deals with nonlinear, non-monotonic data regression. This paper introduces an efficient algorithm to perform data transformation from non-monotonic to monotonic to be paired with a statistical bivariate regression method. The proposed algorithm is applied to a number of synthetic and real-world non-monotonic data sets to test its effectiveness. The proposed novel non-isotonic regression algorithm is also applied to a collection of data about strontium isotope stratigraphy and compared to a LOWESS regression tool.

Keywords: non-monotonic nonlinear data-fitting; statistical bivariate regression; non-isotonic regression; marine stratigraphy

1. Introduction

Every experiment or phenomenological study produces a set of data that, in a large number of instances, is monotonic (see, for example, the study [1] on nonlinear magnetostatic problems). When a data set is not non-monotonic, it is harder to obtain a model of the data and to infer the value of missing records by interpolation than for a monotonic data set. One solution is to infer a functional relationship between variables using regression analysis as illustrated, to cite a few, in the paper [2] on evolutionary algorithms, in the contribution [3] on autonomous agents, and in the contributions [4–6] which cover several practical aspects of regression analysis. Regression is a computation application of paramount importance as testified by the research paper [7] that illustrates an application to drowsiness estimation using electroencephalographic data, by the book [8] on statistical methods for engineers and scientists, by [9] that explores an improved power law for nonlinear least-squares fitting, in the papers [10–12] that exploit regression analysis in forecasting and prediction, by the research paper [13] that compares a number of linear and non-linear regression methods, in the paper [14] that uses support vector regression for the modeling and synthesis of antenna arrays, and by the contribution [15] that applies kernel Ridge regression to short-term wind speed forecasting.

The present research takes its moves from the isotonic statistical bivariate regression (SBR) method presented in [16] (which was successfully applied to estimate the glomerular filtration rate in kidneys). Previous comparative studies [5,16] have clearly shown how statistical regression implemented by look-up tables is much faster in execution than traditional techniques while ensuring the same modeling/regression performance.

Since isotonic regression is based on the assumption that the independent variable and the dependent variable are bound by a monotonic relationship, the statistical bivariate regression

method cannot be applied directly to data sets that are not monotonically increasing nor decreasing. We proposed in [5,6], as a remedy, to make use of a data-transformation technique referred to as *data monotonicization*. As a novel contribution to this research topic, in Section 2 of the present paper we propose a non-linear integral transformation which turns a bivariate data set into a modified set in which the relation between the dependent variable and the independent variable is monotonically increasing. Section 3 contains a summary of the SBR method. In Section 4, some results of numerical tests performed to assess the effectiveness of the proposed technique are illustrated and discussed.

A regression problem that motivated the present research endeavor is marine stratigraphy as tackled in [17]. Marine stratigraphy is at the heart of geology and deals with the study of marine deposits over ages of the Earth [18–20]. The principal aim of stratigraphy is to produce a time scale to date geological processes by arranging rocks in chronological order on the basis of their inorganic and organic characteristics [21]. Absolute radiometric dating is the base for investigating the gross speed of processes such as tectonic movements or organic evolution [21]. The stratigraphy data set analyzed through non-isotonic regression as well as the results of regression are explained in Section 5. Section 6 concludes the paper.

2. Proposed Transformation and Pseudo-Codes

The proposed paper is an extension of previous work of the third author on non-monotonic regression by means of transformation of a non-monotonic response to a monotonic one. This section introduces a non-linear integral transformation that is combined with the previously published statistical bivariate regression method in order to predict the response values for the same or different values of predictor.

2.1. A Non-Linear Integral Transformation

Assume that a smooth function $f : [a, b] \rightarrow \mathbb{R}$ is not monotonic: a non-linear transformation that makes it monotonically increasing is defined as

$$h(x) := r \int_a^x \exp(c f'(\xi)) d\xi + f(a), \quad (1)$$

where $r, c > 0$ are constants. The main reason why this non-linear integral transformation has been chosen is that $h'(x) = r \exp(c f'(x))$ is positive for every x , therefore the derivation-exponentiation-integration chain that such transformation is based on guarantees the resulting function h to be smooth and monotonically increasing even if the function f is not.

There exists an interesting connection between the proposed non-linear transformation and a family of functions suggested by Ramsay in [22], which are defined by the the second-order differential equation $\frac{d^2g}{dx^2} = w \frac{dg}{dx}$, where w is an unconstrained coefficient function and g is the sought unknown. The solution to this equation is

$$g(x) = c_0 + c_1 \int_0^x \exp\left(\int_0^\xi w(\tau) d\tau\right) d\xi, \quad (2)$$

where c_0 and c_1 are arbitrary constants. The family of solutions include the strictly monotonic twice-differentiable functions. (In fitting data, it might be useful to regularize g by penalizing the integral of w^2 since this is a measure of the relative curvature in f .) Clearly the non-linear transformation (1) is a member of the family (2) whenever f is a twice-differentiable function, in which case, setting $w := c f''$ makes such identification explicit.

An important feature of the transformation (1) is that it admits an exact inverse

$$f(x) = \frac{1}{c} \int_a^x \log\left(\frac{1}{r} h'(\xi)\right) d\xi + h(a). \quad (3)$$

We shall assume that a model f to be inferred is represented by a finite data set \mathbb{D} . A finite bivariate data set is a set of pairs

$$\mathbb{D} := \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}, \quad (4)$$

where x_i denotes a sample of the independent variable, y_i a sample of the dependent variable, and n denotes the total number of samples, where it is assumed that $y_i \sim f(x_i)$. Such assumption subsumes some statistical relationship between the variables (x, y) like, for instance, $y = f(x + \nu) + \mu$, where ν and μ denote measurement noises. Further, we assume that the x_i values are unique (i.e., the data set contains no repetitions) and that the pairs in \mathbb{D} are sorted in ascending order according to the x_i values.

The integral transformation (1) may be adapted to a finite data-set as follows:

$$z_i = \begin{cases} y_1, & \text{for } i = 1, \\ z_{i-1} + r \exp\left(c \frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right) (x_i - x_{i-1}), & \text{for } i > 1, \end{cases} \quad (5)$$

where the pairs (x_i, z_i) constitute the resulting monotonic dataset, namely

$$\mathbb{M} := \{(x_1, z_1), (x_2, z_2), (x_3, z_3), \dots, (x_n, z_n)\}. \quad (6)$$

The algorithm (5) is based on a Newton's divided differences approximation that affects its performances: it was chosen because it represents the simplest way to approximate a first order derivative. The inverse transformation is achieved by

$$y_i = \begin{cases} z_1, & \text{for } i = 1, \\ y_{i-1} + \frac{1}{c} (x_i - x_{i-1}) \log\left(\frac{1}{r} \frac{z_i - z_{i-1}}{x_i - x_{i-1}}\right), & \text{for } i > 1. \end{cases} \quad (7)$$

The first point of the transformed data set and the original data set coincide.

The purpose of the constants r and c is to control the range of the transformed data so that, for instance, they keep in the same range of the original data set. Values of the constant c are likely to be far lesser than unity in order to prevent the exponential to blow up.

A proposed pseudo-code to implement the monotonicity transformation is outlined in the Algorithm 1. In this pseudo-code, Line 1: the call to the function requires as arguments the x and y arrays that will be transformed and the constants r and c ; Lines 2–3: performs initialization; Lines 4–9: the monotonic transformation is computed; Line 10: the array z is returned as output of the function.

Algorithm 1 Pseudo-code to implement the (direct) transformation (5).

```

1: function MONO( $r, c, x, y$ )
2:    $z_1 \leftarrow y_1$ 
3:    $i \leftarrow 2$ 
4:   while  $i \leq \text{length}(x)$  do
5:      $\Delta x_i \leftarrow x_i - x_{i-1}$ 
6:      $\Delta y_i \leftarrow y_i - y_{i-1}$ 
7:      $z_i \leftarrow z_{i-1} + r \exp(c \Delta y_i / \Delta x_i) \Delta x_i$ 
8:      $i \leftarrow i + 1$ 
9:   end while
10:  return  $z$ 
11: end function

```

The Algorithm 2 shows a pseudo-code to implement the inverse of the transformation in Algorithm 1.

Algorithm 2 Pseudo-code to implement the (inverse) transformation (7).

```

1: function INVMONO( $r, c, x, z$ )
2:    $y_1 \leftarrow z_1$ 
3:    $i \leftarrow 2$ 
4:   while  $i \leq \text{length}(x)$  do
5:      $\Delta x_i \leftarrow x_i - x_{i-1}$ 
6:      $\Delta z_i \leftarrow z_i - z_{i-1}$ 
7:      $y_i \leftarrow y_{i-1} + \Delta x_i \log((\Delta z_i / \Delta x_i) / r) / c$ 
8:      $i \leftarrow i + 1$ 
9:   end while
10:  return  $y$ 
11: end function

```

In this pseudo-code, Line 1: the call to the function requires as arguments the x and z arrays that will be transformed back to non monotonic data and the constants r and c whose values are necessarily the same as in the direct processing; Lines 4–9: the inverse transformation is computed; Line 10: the array y is returned as output of the function.

2.2. Un-blended and Blended Methods

Once the SBR method is used on a monotonic data set, the inverse of the transformation is applied on interpolated data.

In this paper, a direct application of the INVMONO function to the model is called *un-blended method*. In addition, an alternative solution is proposed, which is referred to as *blended method*. This method consists in gathering the x - and q_x -values (respectively, the original data and the query point) into an enlarged data set and in gathering the y - and q_y -values (respectively original data and the algorithmic response to query points) into an enlarged data set, namely:

$$\{\hat{x}\} := \{x\} \cup \{q_x\} \text{ and } \{\hat{y}\} := \{y\} \cup \{q_y\}. \quad (8)$$

Applying the INVMONO function to the pair $(\{\hat{x}\}, \{\hat{y}\})$ results in a model that contains information from both the original data-points as well as the point inferred by the SBR procedure. Recovering the results of de-monotonization corresponding to the query points $(\{q_x\}, \{q_y\})$ results in the sought model.

We shall illustrate the proposed methodology by numerical experiments conducted on both synthetic and real-world data and via a comparison to another existing regression method.

3. Statistical Bivariate Regression

Although the monotonization/demonotonization stages are independent of the monotonic regression algorithm sandwiched inbetween, hence, in principle, every regression algorithm might be invoked, our aim is to make use only of computationally simple (rather elementary) numerical operations, which are incompatible with complex methods such as kernel-based smoothers (see, e.g., [23]), neural networks (see, e.g., [24]) or splines-based regressors (see, for example, [25]).

Statistical bivariate regression is a mathematical method to deduce the value of missing points between adjacent pairs of data points (x_i, y_i) and constitutes an improvement over isotonic regression. In the paper [16] it was presented an algorithm that estimates the cumulative distribution function (CDF) of the x -set, the inverse cumulative distribution function (INVCDF) of the y -set, and combines such estimations to obtain the sought model. This algorithm can process effectively only monotonic data as it cannot cope with non-monotonic relationships. A pseudo-code for the SBR procedure is shown in the Algorithm 3.

Algorithm 3 Pseudo-code to implement statistical bivariate regression.

```

1: function SBR( $x, y, q_x$ )
2:    $P_x \leftarrow \text{CDF}(x, q_x)$ ;
3:    $q_y \leftarrow \text{INVCDF}(y, P_x)$ 
4:   return  $q_y$ 
5: end function

```

In the pseudo-code, q_x denotes a set of query-points where the model is needed to be inferred, while the set q_y denotes the corresponding response. In other words, the set q_x contains values of the independent variable that were not observed, hence that do not belong to the x -set, and the procedure SBR infers the corresponding values of the dependent variable. For a detailed explanation of the underlying theory, interested readers might consult the published paper [16].

4. Numerical Experiments

This section discusses the results of several preliminary numerical tests. These tests were performed on synthetic as well as real-world data drawn from public repositories.

4.1. Specifications of Data Sets Used in the Experiments

Several data sets, each exhibiting different features, were used to tests the monotonicity transformation and the statistical bivariate regression method applied in combination. Figure 1 shows these data sets, which were borrowed from the articles [5,6].

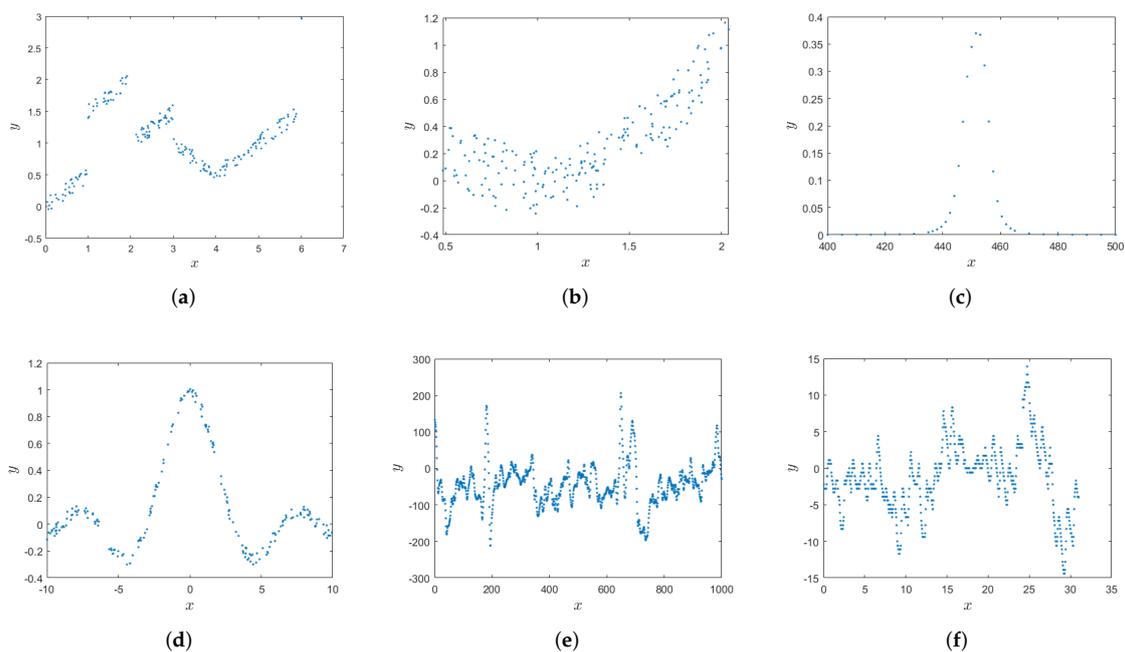


Figure 1. Graphical illustration of the three synthetic and three real-world data sets used in the numerical experiments: (a) Dataset 1; (b) Dataset 2; (c) Dataset 3; (d) Dataset 4; (e) Dataset 5; (f) Dataset 6.

The functional expressions of the synthetic data are explained in [5,6]. Some further specifications about these data are as follows:

- The Dataset 1, Dataset 2 and Dataset 4 were synthetically generated to exhibit specific features. The Dataset 1 was generated to exhibit a discontinuous dependency between the independent variable x and the dependent variable y as well as a moderate amount of noise. The Dataset 2

was designed on the basis of a quadratic dependency and large additive noise. The Dataset 4 was designed on the basis of a moderately noisy, oscillating (cardinal-sine-type) dependency.

- The Dataset 3 was downloaded from the repository described in [26] and is the result of a NIST study involving circular interference transmittance. It has been chosen because it contains only 35 records for a comparison with other large data sets.
- The Dataset 5 arises from an electrocardiogram (ECG) readout. The x variable represents a data-sample (or temporal) index, while the y variable represents an ECG voltage reading. This dataset contains 1000 sample pairs.
- The Dataset 6 is a real-world data set of temperature readings (in Celsius scale) taken every hour at the Logan Airport for the entire month of January 2011. This dataset contains 744 sample pairs.

The real-world data sets exhibit large variability in the variables' ranges. All numerical experiments were performed on a MATLAB[®] platform. In these preliminary tests, the value of r was set to 0.1.

4.2. Results of Monotonization

In this subsection, results of monotonic transformation are illustrated through numerical examples. Figure 2 shows results of monotonization applied to the Data sets 1 to 6. In these panels, blue dots denote data transformed according to the algorithm (5), while red dots denote results of the inverse transformation (7). These results were obtained by setting $c = 0.0001$.

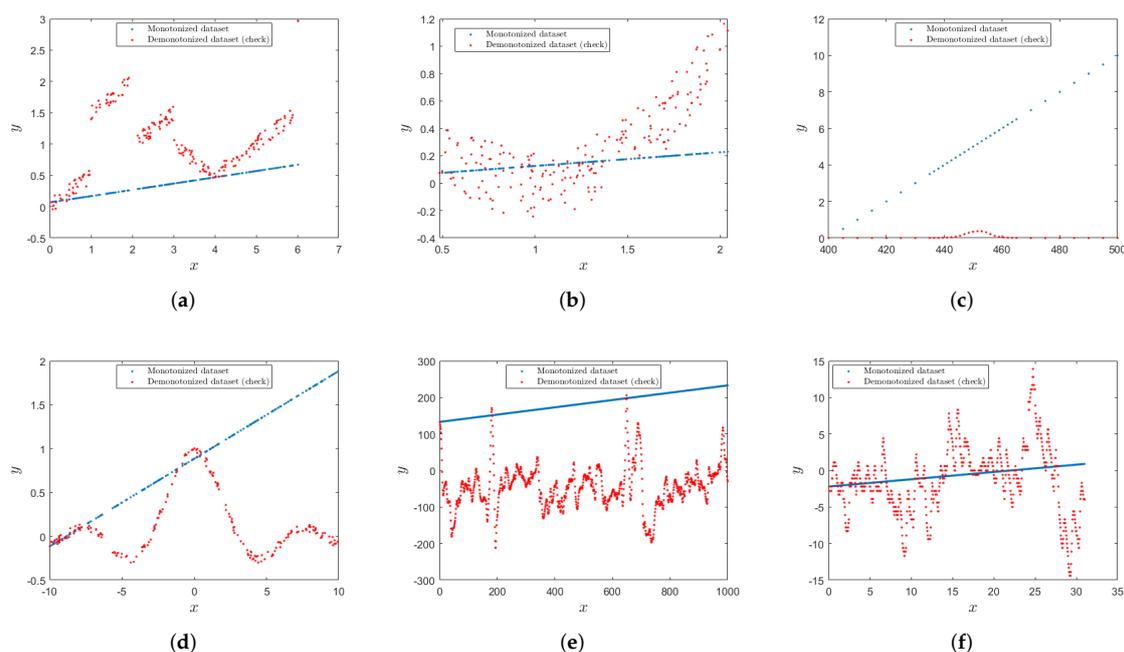


Figure 2. Results of monotonization/demononization applied to: (a) Data set 1, (b) Data set 2, (c) Data set 3, (d) Data set 4, (e) Data set 5, (f) Data set 6.

In all figures the red dots coincide with the original data-points, confirming that the monotonization/demononization cascade is an approximate identity.

4.3. Results of Data Regression

The blended and un-blended methods to achieve de-monotonization were compared. Three different values of the constant c were chosen, namely $c = 0.0001$, $c = 0.00001$ and $c = 0.000001$.

Figure 3 shows results obtained on the Data set 1. Even in the presence of data that present large jumps, the proposed regression procedure is able to fit the data satisfactorily.

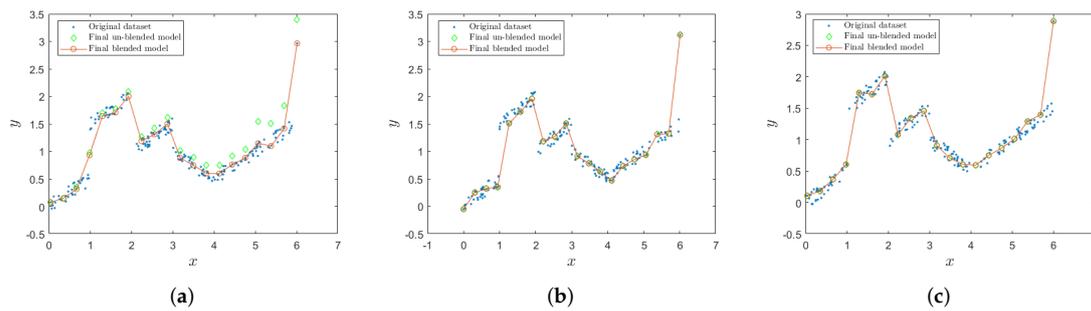


Figure 3. Results of non-isotonic modeling obtained on the Data set 1 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

Figure 4 shows results obtained on the Data set 2. Due to the presence of large noise components in the data, the proposed non-monotonic regression algorithm is unable to infer a consistent data model.

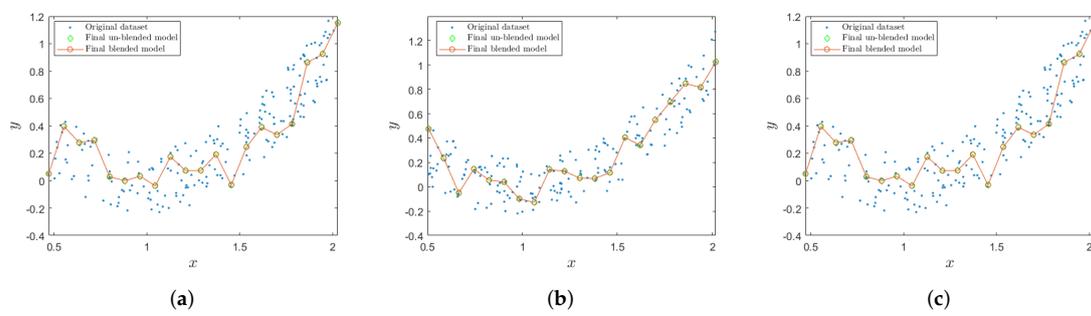


Figure 4. Results of non-isotonic modeling obtained on the Data set 2 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

Figure 5 shows results obtained on the Data set 3. Even in the presence of a very limited number of data-points in the training set, the proposed non-linear regression algorithm is able to infer a consistent model of the underlying process.

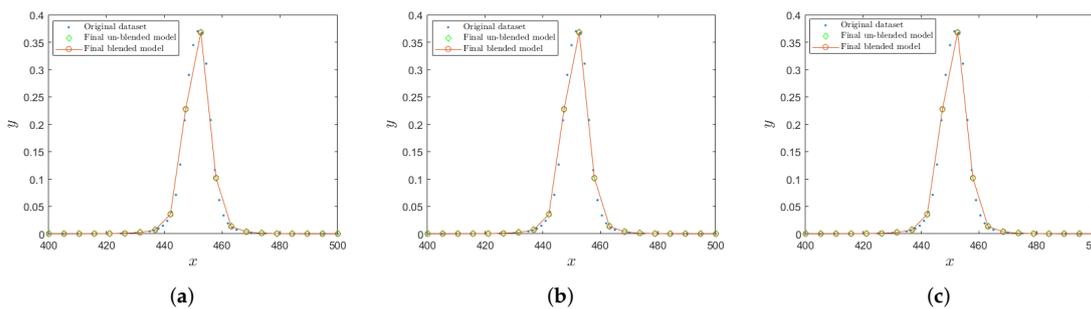


Figure 5. Results of non-isotonic modeling obtained on the Data set 3 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

Figure 6 shows results obtained on the Data set 4. On this noiseless and smooth data set, the proposed regression procedure performs satisfactorily.

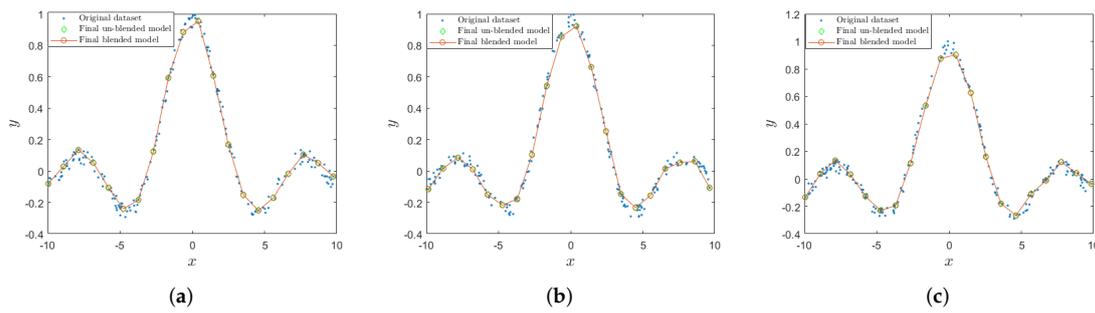


Figure 6. Results of non-isotonic modeling obtained on the Data set 4 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

Figure 7 shows results obtained on the Data set 5, while Figure 8 shows results obtained on the Data set 6. In both cases, the proposed regression procedure is able to capture some features of the underlying model, although it is ineffective in capturing fast changes in the data structure.

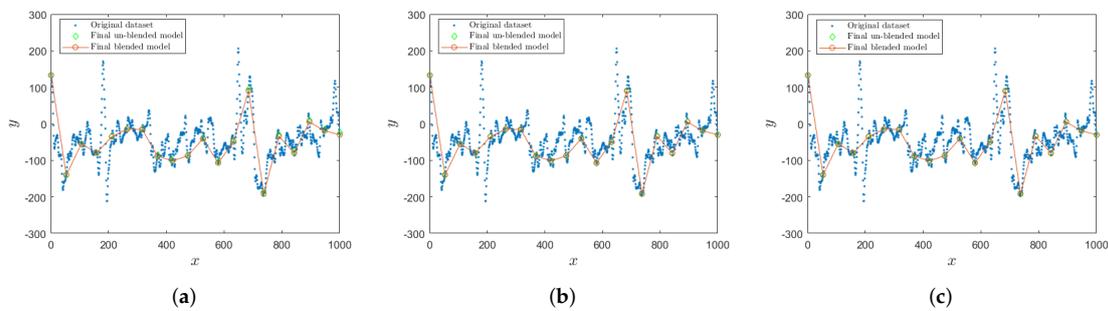


Figure 7. Results of non-isotonic modeling obtained on the Data set 5 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

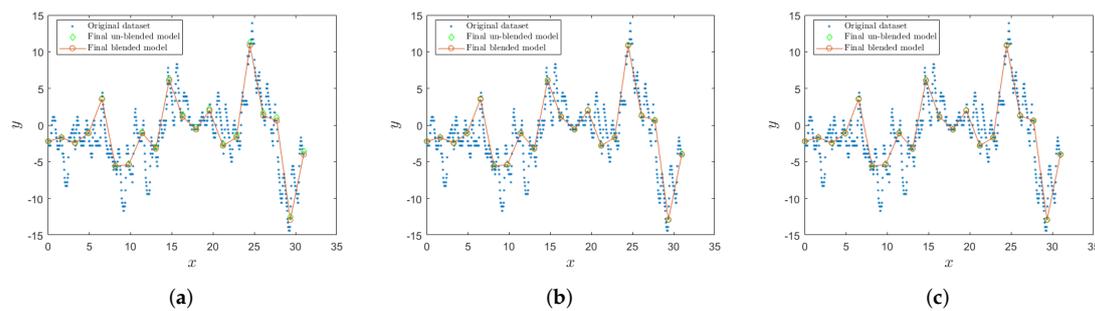


Figure 8. Results of non-isotonic modeling obtained on the Data set 6 with $r = 0.1$ and: (a) $c = 0.0001$, (b) $c = 0.00001$, (c) $c = 0.000001$.

The illustrated results were further evaluated by a mean-squared error (MSE) metric on 20 samples chosen randomly from each data set. The resulting figures are reported in Table 1.

Table 1. Evaluation of the blended and un-blended regression methods by a mean-squared error metric. Results pertaining to the parameter value $c = 0.000001$.

	Data Set 1	Data Set 2	Data Set 3	Data Set 4	Data Set 5	Data Set 6
Un-blended	9.31×10^{-1}	8.40×10^{-2}	5.91×10^1	6.02×10^{-3}	7.16×10^4	4.21×10^1
Blended	1.28×10^{-18}	2.75×10^{-18}	3.73×10^{-17}	1.75×10^{-18}	5.02×10^{-12}	1.69×10^{-15}

These numerical tests show, especially for Data sets 1, 2, 5 and 6, that the un-blended-based model deviates from the original one, while the blended model appears to be faithful to the actual model for sufficiently small values of the constant c .

5. Application to Strontium Isotope Stratigraphy

The present section deals with an application of the proposed non-isotonic regression method to a collection of strontium isotope stratigraphic data from a study by McArthur, Howarth and Bailey [17], where the authors refer to this collection with the name of “V3”. The purpose of the study by McArthur and coworkers was to compile a table that affords assigning numerical ages to sediments based on concentration ratios $^{87}\text{Sr}/^{86}\text{Sr}$ of radioactive strontium isotopes.

The V3 dataset comes in pairs of 3401 records of the type (x_i, y_i) , where the variable x denotes the age of a sediment, expressed in Ma (‘mega-annum’ corresponding to a period of 1 million years) and the variable y denotes the ratio $^{87}\text{Sr}/^{86}\text{Sr}$ of strontium isotopes [27]. The data set V3 is indeed incomplete, since 11 records are missing a y -value and one record presents a 0 value in the y attribute.

In order to apply the devised non-isotonic regression method to this data set, it was necessary to pre-process the data. The devised regression method cannot be applied in the presence of more records that present the same value in the x -attribute, therefore, as a first fix, all those values have been replaced with a pair (\bar{x}, \bar{y}) where the \bar{x} 's are unique and the \bar{y} 's denotes the mean value taken among all records whose x -attribute was repeated. Furthermore, incomplete records were removed from V3 to realize regression. After this pre-processing, the data set reduced to 3389 pairs.

Figure 9 shows results obtained on the whole dataset. The interpolation covers only the range 0–509 Ma as in the study by McArthur, Howarth and Bailey.

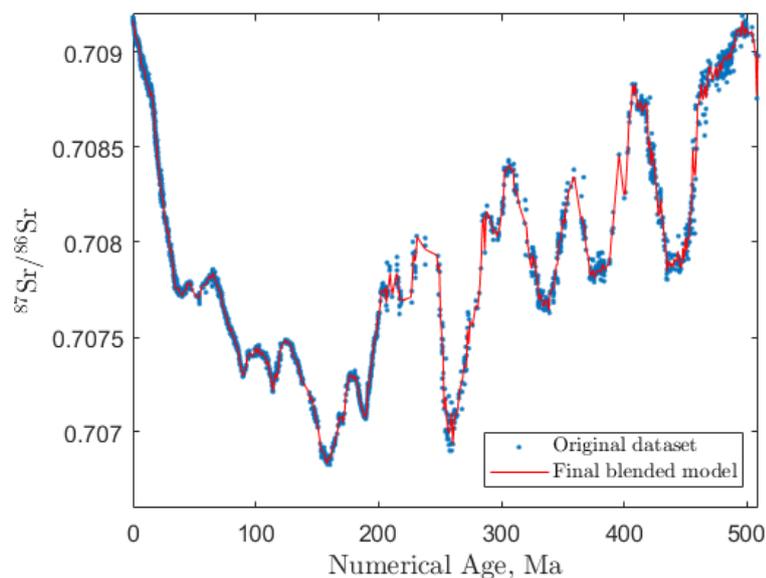


Figure 9. Result of non-isotonic modeling obtained on V3 with $r = 0.1$ and $c = 0.1$ (range 0–509 Ma).

In order to probe in more detail the result of regression, following [17] we have divided the interval 0–509 Ma into several sub-intervals (two of which are partially overlapping).

Figure 10 shows results obtained on the first half (0–210 Ma). In three out of four panels, the data-points are pretty dense, hence the statistical regression method, which is based on the probability distribution of the data, possesses enough information to infer a data model. In the panel corresponding to the interval 30–70 Ma, the data-points are less dense and the regression algorithm yields a coarse model of the relationship between the dependent and the independent variable.

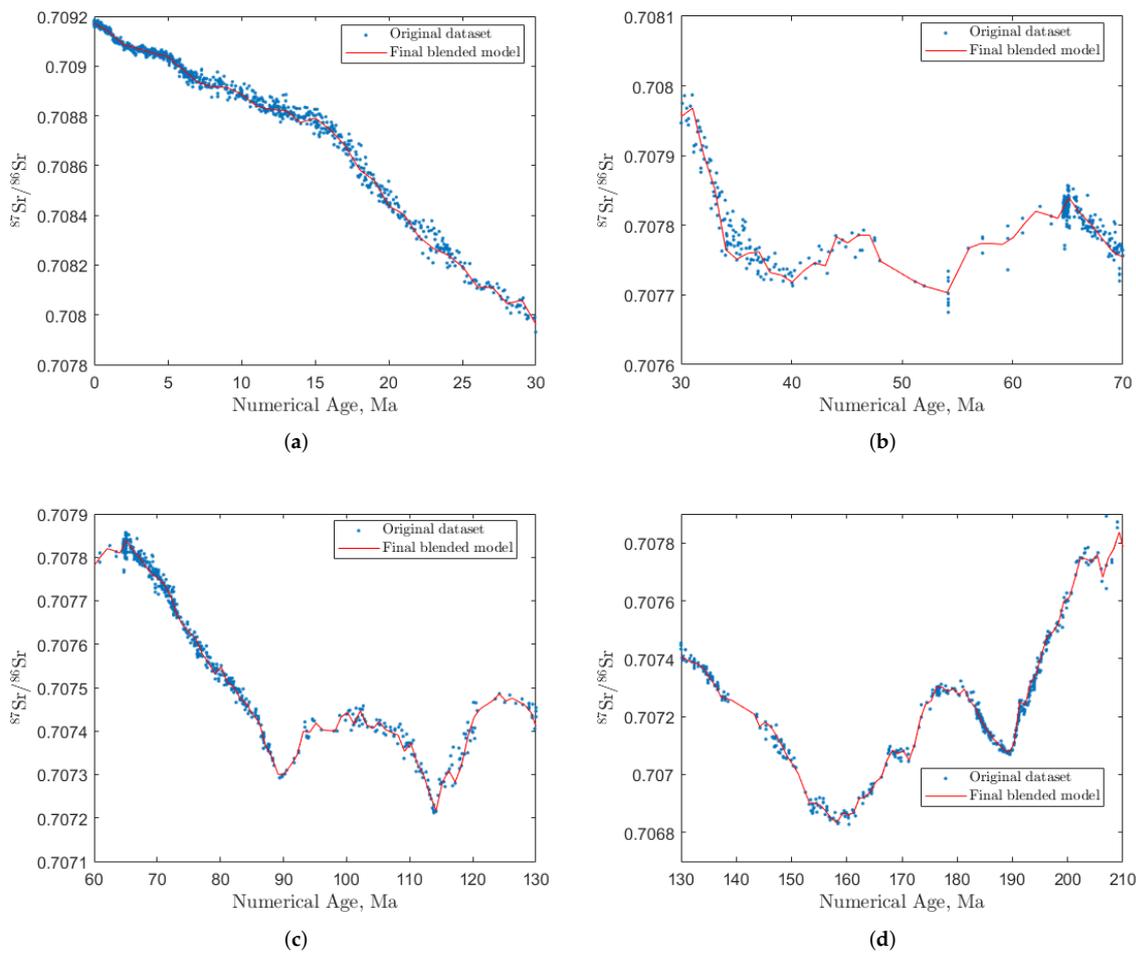


Figure 10. Results of non-isotonic modeling obtained on V3: First four subintervals, two of which are slightly overlapping as in [17]. Only incomplete records were omitted from the graphs.

Figure 11 shows results obtained on the second half (200–509 Ma). As already noted, wherever the data-points are scarce, the regression algorithm returns a coarse model. It is also interesting to observe how the regression algorithm ignores some data-points, treating them as outliers, as it happens for example in the interval 360–370 Ma.

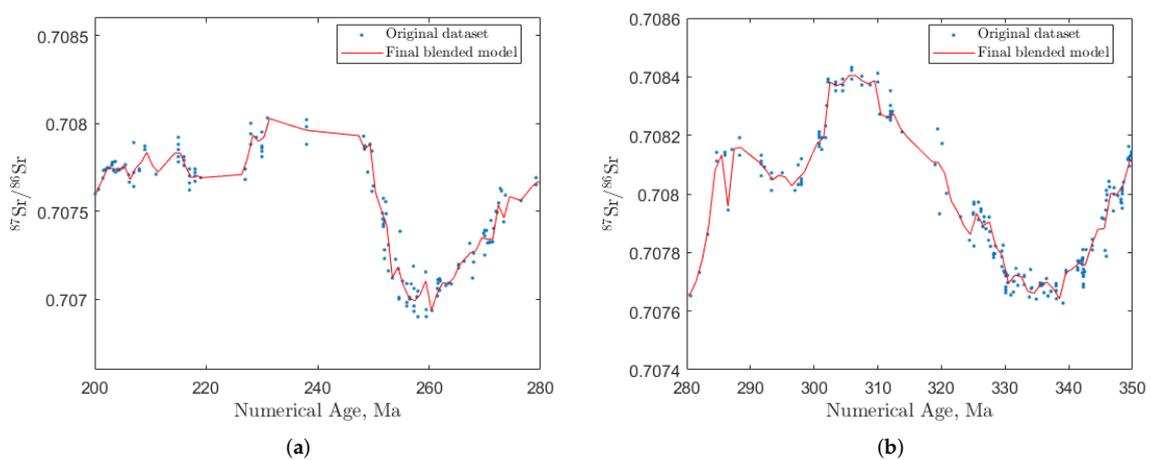


Figure 11. Cont.

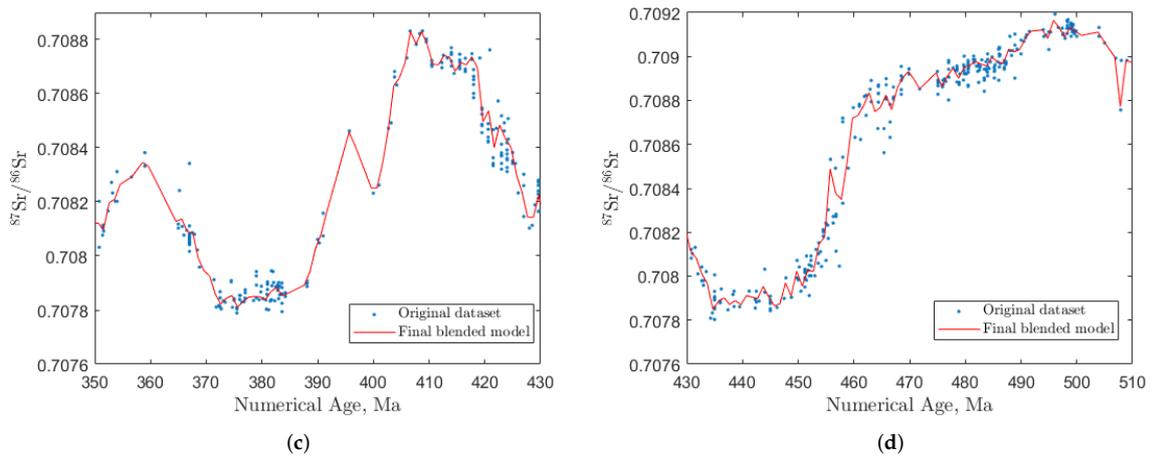


Figure 11. Results of non-isotonic modeling obtained on V3: Last four subintervals. The interval 200–210 Ma was repeated to get a clearer vision of the graphs, as in [17]. Only incomplete records were omitted from the graphs.

For comparison purposes, the non-isotonic regression method is contrasted with the LOWESS method on the strontium dataset. The LOWESS method is a nonparametric regression technique explained in an earlier paper by McArthur and Howarth [28].

The LOWESS fit is expressed by three graphs according to the mean of the model, its maximum and its minimum. Figure 12 compares estimations provided by the devised statistical regression method and by the LOWESS method. As it may be readily observed, the line output of the statistical regression method discussed in the present work agrees pretty well with the ‘Min LOWESS’ inference, except perhaps for a point around 220 Ma where statistical regression seems to adhere more closely to the data than the LOWESS prediction, for the interval 230–250 Ma, where the LOWESS method predicts some spikes, while our method predicts a flatland, and for a point around 400 Ma where the Min LOWESS curve looks pretty smooth, while the curve pertaining to our method presents a spike.

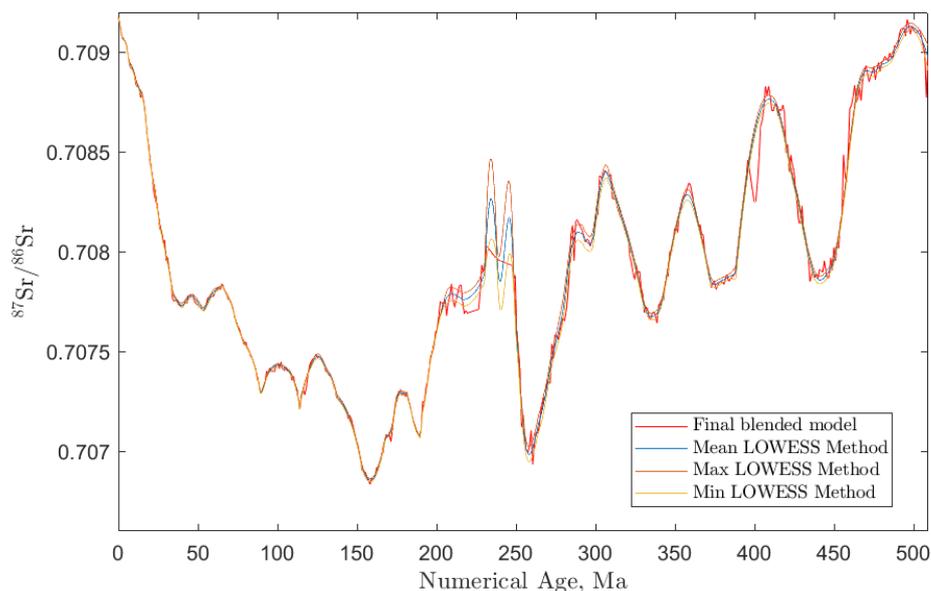


Figure 12. Comparison of the non-isotonic regression method and the three estimations gotten by the LOWESS method.

The ‘Min LOWESS’ and ‘Max LOWESS’ were contrasted with the ‘blended’ fit by the Diebold-Mariano test [29]. The p-value obtained by using absolute loss-differentials is $p_1 \approx 0.364$, which is far larger than the largest reference p-value 0.05, hence the ‘blended’ model is in excellent agreement with the ‘Min LOWESS’ and ‘Max LOWESS’ fits.

The proposed method certainly honors the data better than the LOWESS fit, in which spans were chosen that deliberately down-weight data that are known to be aberrant. In fact, it is known from geo-chemical reasoning that the curve of $^{87}\text{Sr}/^{86}\text{Sr}$ against time should not change sharply over time intervals of 1–2 Ma, which explains why the sharp inflections at 286 Ma in Figure 11b are smoothed out by LOWESS.

Since the V3 data table extends to more than 600 Ma its age span, we have also tried to extend the model using the records of all the numerical ages even if they were further than 509 Ma. Figure 13 represents results obtained by the non-isotonic regression method applied to the whole data set. These data appear not particularly well-tamed in the interval 500–600 Ma and are quite scarce, therefore the inferred regression line appears quite inconsistent.

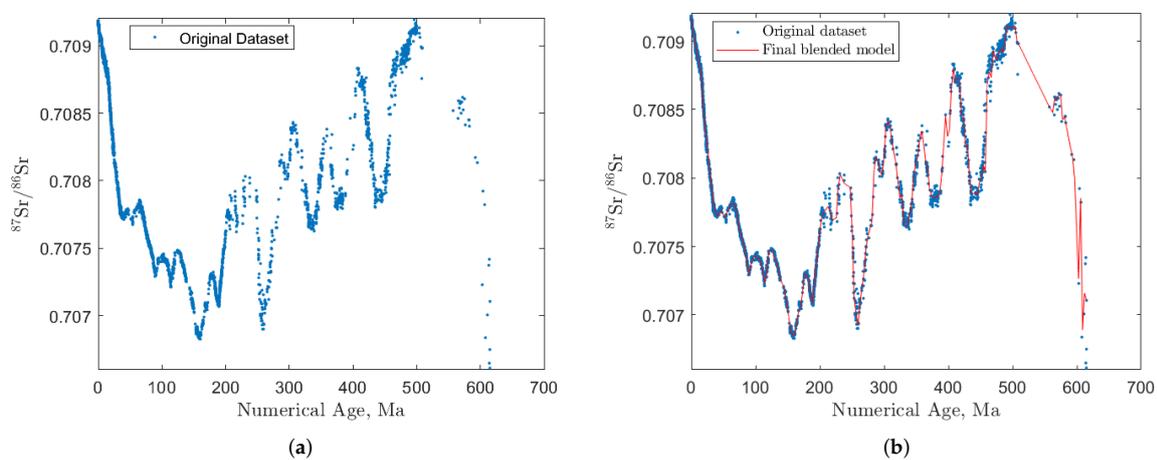


Figure 13. Results of non-isotonic modeling obtained on V3 (whole range).

In addition, the missing values in the V3 dataset have been filled-in as a result of interpolation by the regression model. Figure 14 focuses the view on the missing records, and Table 2 contains their numerical values.

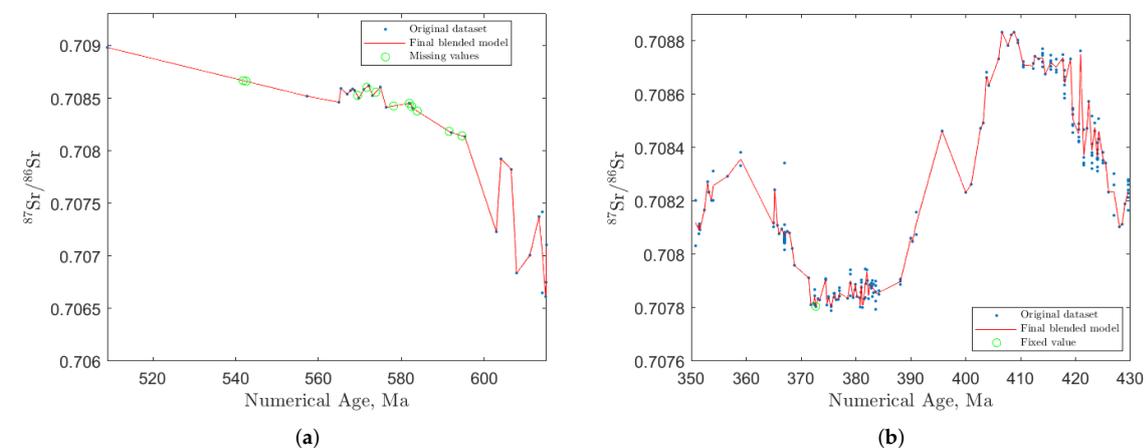


Figure 14. Representation of the missing values on V3: (a) Records that are missing a y attribute, (b) Record with a 0 in the y attribute.

Table 2. Pairs of values that fill-in gaps in V3. The first 11 columns refer to missing values, the last column refers to the pair (x, y) that had the original y attribute equal to 0.

Numerical Age (Ma)	541.8	542.5	569.4	571.6	573.8	578.1
Strontium Isotopes Ratio	0.7087	0.7087	0.7085	0.7086	0.7086	0.7084
Numerical Age (Ma)	581.9	582.5	583.8	591.5	594.6	372.7
Strontium Isotopes Ratio	0.7085	0.7084	0.7084	0.7082	0.7081	0.7078

6. Conclusions

The present paper dealt with nonlinear, non-monotonic data regression by isotonic statistical bivariate regression. Since isotonic regression may only cope with monotonic relationships, the present paper introduced an efficient algorithm to perform a reversible data-transformation to convert non-monotonic data to monotonic. Upon performing statistical regression by an isotonic regression technique previously devised by one of the authors, the obtained monotonic data model is brought back to its original domain by applying a reversed transformation.

The devised algorithm was applied to different non-monotonic data-sets, either synthetic and natural, to test its capabilities and to investigate on their sensitivity to different choices of its free parameters.

In addition, the proposed novel non-isotonic regression method was applied to a collection of data about strontium-isotope-based marine stratigraphy and the obtained results were compared to those obtained by a LOWESS method. The results of this comparison revealed that the devised non-isotonic statistical bivariate regression method compares favorably with the LOWESS method as it infers a model which appears to be more adherent to the data and less bound by smoothness/continuity constraints, yet being in excellent agreement with the LOWESS fit according to a Diebold-Mariano statistical significance test. By applying the inferred model as an interpolation tool, the proposed method was also shown to be able to fill-in gaps in the original data sets.

Author Contributions: Conceptualization, S.F.; Data curation, S.F.; Formal analysis, D.P.; Investigation, D.P. and M.M.; Methodology, S.F.; Software, D.P., M.M. and S.F.; Supervision, S.F.; Writing—original draft, D.P. and M.M.; Writing—review & editing, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors wish to thank John M. McArthur for very kindly sharing the marine isotope stratigraphy data used in the paper [17] and for sharing comments on an earlier version of the present paper. The authors also wish to gratefully thank the anonymous reviewers whose comments and suggestions contributed significantly to improve the scientific content of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Friedrich, L.; Curti, M.; Gysen, B.; Lomonova, E. High-Order Methods Applied to Nonlinear Magnetostatic Problems. *Math. Comput. Appl.* **2019**, *24*, 19. [[CrossRef](#)]
2. Hedar, A.R.; Deabes, W.; Almarashi, M.; Amin, H. Evolutionary Algorithms Enhanced with Quadratic Coding and Sensing Search for Global Optimization. *Math. Comput. Appl.* **2020**, *25*, 7. [[CrossRef](#)]
3. Xie, S.; Lawniczak, A.; Gan, C. Modeling and Analysis of Autonomous Agents' Decisions in Learning to Cross a Cellular Automaton-Based Highway. *Computation* **2019**, *7*, 53. [[CrossRef](#)]
4. Fiori, S. Fast Statistical Regression in Presence of a Dominant Independent Variable. *Neural Comput. Appl.* **2013**, *22*, 1367–1378. [[CrossRef](#)]

5. Fiori, S. A Comprehensive Comparison of Algorithms for the Statistical Modelling of Non-monotone Relationships via Isotonic Regression of Transformed Data. *Int. J. Data Anal. Tech. Strateg.* **2019**, *11*, 29–57. [[CrossRef](#)]
6. Fiori, S.; Gong, T.; Lee, H. Bivariate Nonisotonic Statistical Regression by a Lookup Table Neural System. *Cogn. Comput.* **2015**, *7*, 715–730. [[CrossRef](#)]
7. Akbar, I.; Igasaki, T. Drowsiness Estimation Using Electroencephalogram and Recurrent Support Vector Regression. *Information* **2019**, *10*, 217. [[CrossRef](#)]
8. Bethea, R.; Duran, B.; Boullion, T. *Statistical Methods for Engineers and Scientists*; Marcel Dekker: New York, NY, USA, 1985.
9. Helyer, B.; Courtney, M. An Improved Power Law for Nonlinear Least-Squares Fitting? *Data* **2017**, *2*, 31. [[CrossRef](#)]
10. Huang, Z.; Huang, G.; Chen, Z.; Wu, C.; Ma, X.; Wang, H. Multi-Regional Online Car-Hailing Order Quantity Forecasting Based on the Convolutional Neural Network. *Information* **2019**, *10*, 193. [[CrossRef](#)]
11. Kushiro, N.; Fukuda, A.; Kawatsu, M.; Mega, T. Predict Electric Power Demand with Extended Goal Graph and Heterogeneous Mixture Modeling. *Information* **2019**, *10*, 134. [[CrossRef](#)]
12. Xu, L.; Li, C.; Xie, X.; Zhang, G. Long-Short-Term Memory Network Based Hybrid Model for Short-Term Electrical Load Forecasting. *Information* **2018**, *9*, 165. [[CrossRef](#)]
13. Pan, J.J.; Mahmoudi, M.; Baleanu, D.; Maleki, M. On Comparing and Classifying Several Independent Linear and Non-Linear Regression Models with Symmetric Errors. *Symmetry* **2019**, *11*, 820. [[CrossRef](#)]
14. González Ayestarán, R. Support Vector Regression for the Modeling and Synthesis of Near-Field Focused Antenna Arrays. *Electronics* **2019**, *8*, 1352. [[CrossRef](#)]
15. Zhang, S.; Zhou, T.; Sun, L.; Liu, C. Kernel Ridge Regression Model Based on Beta-Noise and Its Application in Short-Term Wind Speed Forecasting. *Symmetry* **2019**, *11*, 282. [[CrossRef](#)]
16. Giles, S.; Fiori, S. Glomerular Filtration Rate Estimation by a Novel Numerical Binning-Less Isotonic Statistical Bivariate Numerical Modeling Method. *Information* **2019**, *10*, 100. [[CrossRef](#)]
17. McArthur, J.; Howarth, R.; Bailey, T. Strontium Isotope Stratigraphy: LOWESS Version 3: Best Fit to the Marine Sr-Isotope Curve for 0–509 Ma and Accompanying Look-up Table for Deriving Numerical Age. *J. Geol.* **2001**, *109*, 155–170. [[CrossRef](#)]
18. Björck, S.; Denegård, B.; Sandgren, P. The Marine Stratigraphy of the Hanö Bay, SE Sweden, Based on Different Sediment Stratigraphic Methods. *Geologiska Föreningen i Stockholm Förhandlingar* **1990**, *112*, 265–280. [[CrossRef](#)]
19. Boespflug, X.; Long, B.; Occhietti, S. CAT-scan in Marine Stratigraphy: A Quantitative Approach. *Mar. Geol.* **1995**, *122*, 281–301. [[CrossRef](#)]
20. Brett, C. Sequence Stratigraphy, Biostratigraphy, and Taphonomy in Shallow Marine Environments. *PALAIOS* **1995**, *10*, 597–616. [[CrossRef](#)]
21. Seibold, E. Stratigraphy Quo Vadis: Marine Stratigraphy from Continents and Oceans. Available online: <http://archives.datapages.com/data/specpubs/history2/data/a119/a119/0001/0000/0001.htm> (accessed on 9 March 2020).
22. Ramsay, J.O. Estimating Smooth Monotone Functions. *J. R. Stat. Soc. B Stat. Methodol.* **1998**, *60*, 365–375. [[CrossRef](#)]
23. Guedj, B.; Srinivasa Desikan, B. Kernel-Based Ensemble Learning in Python. *Information* **2020**, *11*, 63. [[CrossRef](#)]
24. Kim, J.M.; Wang, N.; Liu, Y.; Park, K. Residual Control Chart for Binary Response with Multicollinearity Covariates by Neural Network Model. *Symmetry* **2020**, *12*, 381. [[CrossRef](#)]
25. Li, B.; Zhang, Y.; Yuan, L.; Xi, X. Study on the Low Velocity Stability of a Prostate Seed Implantation Robot's Rotatory Joint. *Electronics* **2020**, *9*, 284. [[CrossRef](#)]
26. Eckerle, K. Circular Interference Transmittance Study. Available online: <http://www.itl.nist.gov/div898/strd/nls/data/eckerle4.shtml> (accessed on 9 March 2020).
27. Bataille, C.; Bowen, G. Mapping $^{87}\text{Sr}/^{86}\text{Sr}$ Variations in Bedrock and Water for Large Scale Provenance Studies. *Chem. Geol.* **2012**, *304–305*, 39–52. [[CrossRef](#)]

28. Howarth, R.; McArthur, J. Statistics for Strontium Isotope Stratigraphy: A Robust LOWESS Fit to the Marine Sr-Isotope Curve for 0 to 206 Ma, with Look-up Table for Derivation of Numeric Age. *J. Geol.* **1997**, *105*, 441–456. [[CrossRef](#)]
29. Diebold, F.X. Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. Available online: <https://www.nber.org/papers/w18391.pdf> (accessed on 9 March 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).