computer sciences and mathematics forum

MDPI

Proceeding Paper

Introducing Robust Least-Squares Methods in the Detrended Fluctuation Analysis †

Eric Grivel *,‡ and **Pierrick Legrand** ‡

IMS UMR CNRS 5218, Bordeaux INP, University of Bordeaux, 351 Cours de la Liberation, 33400 Talence, France; pierrick.legrand@ensc.fr

* Correspondence: eric.grivel@ims-bordeaux.fr

† Presented at the 12th International Conference on Time Series and Forecasting, Canaria, Spain, 15–17 July 2026.

‡ These authors contributed equally to this work.

Abstract

Detrended fluctuation analysis (DFA) is a widely used method for estimating the Hurst exponent of time series exhibiting nonstationary trends. In its standard formulation, DFA relies on the least-squares (LS) criterion at several stages of the procedure, including trend removal and Hurst exponent estimation. In practice, the estimation is often sensitive to the practitioner's choice of local trend sizes, which can significantly affect the estimated Hurst exponent. Moreover, the LS criterion may be poorly suited in the presence of outliers or limited sample sizes. To address these limitations, we propose robust extensions of the DFA. They are based either on a least trimmed squares method or on an iterative weighted LS approach. In the latter case, the weights are defined using Cook distance. This allows all scales to be considered without requiring any a priori selection of local trend sizes. This strategy also improves the stability of the regression step without introducing too many tuning parameters. Illustrations are provided to demonstrate the relevance of the proposed approaches, especially in the presence of outliers.

Keywords: detrended fluctuation analysis; Hurst exponent; robust estimation; Cook distance; regression

1. Introduction

In time-series analysis, scaling properties such as self-similarity are often used to characterize the structure of stochastic signals across time scales. More particularly, a time-continuous stochastic process $y(t)$ is said to be self-similar with parameter H, called the Hurst exponent, if it satisfies the following scaling relation:

$$y(Nt) \stackrel{d}{=} N^H y(t) \tag{1}$$

where N is an integer and d denotes equality in distribution.

Moreover, the parameter H characterizes how the statistical properties of the process change with scale. This definition leads to a relationship for the variance or the standard deviation. Indeed, if $F^2(N)$ denotes the variance of the scaled process $y(Nt)$, it follows that $F^2(N) = N^{2H}F^2(1)$ or equivalently:

$$\log(F(N)) = H\log(N) + \log(F(1)) \tag{2}$$

Beyond self-similarity, the Hurst exponent also provides information about the regularity and memory of a process [1]. Brownian noise, pink noise, and white noise have Hurst

check for updates

Academic Editors: Olga Valenzuela, Fernando Rojas, Luis Javier Herrera, Hector Pomares and Ignacio Rojas

Published: 19 October 2026

Copyright: © 2026 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.

exponents equal to 0.5, 0 and -0.5, respectively. In addition, short-memory processes are characterized by a Hurst exponent equal to 0.5 [2].

Estimating the Hurst exponent can help distinguish between short-memory processes, where correlations decay rapidly (in other words, the correlation function is absolutely summable; this is true in the case of autoregressive and moving average (ARMA) processes), and long-memory processes characterized by persistent dependence across time scales (in other words, the correlation function is no longer absolutely summable; this is the case for autoregressive fractionally integrated moving average (ARFIMA) processes, fractional Brownian motion processes, fractional Gaussian noises, etc.). Therefore, it can support prior model selection. In a machine learning framework, H can be used as an input feature to enrich data representation with information related to temporal regularity and persistence. Consequently, H can be useful in various applications, including biomedical engineering, aeronautics, fault detection, and econometrics.

Over the last three decades, the detrended fluctuation analysis (DFA) [3] has become a standard method for estimating the Hurst exponent of signals exhibiting nonstationary trends. The method relies on the least-squares (LS) criterion at two stages:

1. The estimation of the piecewise linear trend of the integrated and centered signal. This detrending step is used to estimate $F(N)$, also called the square root of the mean power of the detrended integrated signal.
2. The estimation of the Hurst exponent H based on (2) for different values of N.

Due to its simplicity, the method has been applied on a broad range of signals in fields such as physiology and psychology, mathematics, and signal processing. In recent years, numerous contributions have extended and refined the DFA. They can be classified into several families:

- The first one includes variants of DFA whose purpose is to address alternative ways to estimate the trend of the integrated signal. Thus, higher-order DFA replaces the piecewise linear trend with higher-degree polynomials. To reduce discontinuities between local trends, which may induce spurious effects [4], adaptive corrections have been proposed, leading to the adaptive fractal analysis (AFA) [5]. More recently, continuity constraints and regularization have been explicitly incorporated into the estimation criterion. This includes continuous DFA (CDFA) and regularized DFA (RDFA), the latter being related to Tikhonov regularization or Hodrick–Prescott filtering [6–8]. If the above approaches rely on an *a priori* model of the trend, alternative methods avoid explicit trend modeling and, instead, rely on low-pass filtering. This is the case of the detrending moving average (DMA) methods with finite impulse response filters [9], including backward, central, and forward variants [10]. A third family, including higher-order DMA [11] and LOWESS-based detrending [12], bridges the gap between model-based and filter-based methods.
- DFA has been extended to multi-fractal time series [13], fast implementations [14], and the joint analysis of two signals through detrended cross-correlation analysis [15].
- Several studies have focused on the theoretical properties of DFA, including the behavior of the fluctuation function and the frequency response of the detrending step [7,16–29].
- Practical limitations of DFA have been discussed. In particular, to improve robustness, averaging strategies based on sliding-window decompositions have been proposed [29].

Moreover, the estimation of H may be sensitive to the selection of the values of N made by the practitioner. In [30], Kantelhardt reported that small values of N could lead to underestimated fluctuation functions in the data sets under study. We first attempted

to address this issue by extending the model described in (2) and by exploring several strategies based on an additive function designed to capture the drop-off observed for small values of N. Depending on the model complexity, parameter estimation was based on standard LS methods or expectation–maximization (EM)-type algorithms. However, these approaches yielded limited improvement in practice.

More fundamentally, estimating a local trend from a small number of points using standard LS methods is statistically unreliable. This motivated us to focus our attention on robust LS approaches, which provide more stable estimates and reduce the influence of outliers [31].

Thus, a great deal of interest has been paid to robust regression methods in the literature for the last 50 years. The following overview, which is not exhaustive, presents some of the main families of robust regression methods. In 1977, Holland et al. proposed an iteratively weighted least squares (IWLS) procedure in which weights are set inversely proportional to the squared residuals and are updated until convergence. In [32], weighting functions such as Huber, "logistic" , Fair and Talwar were compared.

L-estimators, where "L" refers to linear combination of order statistics of the measurements, were based on linear combinations of ordered residuals, often obtained after reducing the influence of extreme residuals through trimming or winsorization.

Then, M-estimators [33], whose name comes from "generalized maximum likelihood", were based on the minimization of the following quantity:

$$\sum_{i=1}^{M} \rho\left(\frac{r_i(\theta)}{\sigma}\right) \tag{3}$$

where θ is the vector of p parameters to be estimated, r_i is the residual related to the ith observation, M denotes the number of observations, ρ is a robust loss function such as the Huber or Tukey function, and σ is a scale parameter. Differentiating (3) yields the first-order optimality conditions, which can be reformulated as a weighted LS problem with weights depending on the current residuals and updated iteratively via an IWLS scheme. The breakdown point, which measures the maximal fraction of contaminated observations that an estimator can handle before being driven to arbitrarily large values, is however low.

Introduced by Rousseeuw [34], the least median of squares (LMS) estimator is defined as the parameter vector that minimizes the median of the squared residuals. This amounts to considering the hth order statistic of the ordered squared residuals, $r^2_{(h)}$, where $h = \lfloor (M + p + 1)/2 \rfloor$, with $\lfloor . \rfloor$ the floor function. Note that this estimator has a breakdown point equal to $(\lfloor (M - p)/2 \rfloor + 1)/M$. However, the LMS is computationally demanding in practice. Approximate implementations have therefore been proposed, notably through the PROGRESS algorithm. Moreover, the LMS converges at a rate of $M^{-1/3}$, meaning that the estimation error decreases relatively slowly with sample size, whereas the LS estimator achieves a rate equal to $M^{-1/2}$. A closely related alternative is the least trimmed squares (LTS) estimator, which minimizes the sum of the h smallest squared residuals. When $h = \lfloor (M + p + 1)/2 \rfloor$, LTS attains the same maximal breakdown point as LMS. The above robust methods can be combined with a weighted LS (WLS) approach. In that case, given initial robust estimates, binary weights can be defined using a predefined threshold. In [35], Gervini et al. proposed to adjust this threshold, leading to a robust and efficient WLS estimators (REWLSE) that simultaneously attain the maximum breakdown point and full asymptotic efficiency under normal errors. In [36], the tilted LS estimator does not use a robust estimator as an initial step but is based on an empirically tilted log-likelihood and the Kullback–Leibler divergence between the weight distribution and the uniform distribution. There are many other works related to LTS approaches: In [37], the authors develop two algorithms to compute the LTS estimate in multiple linear regressions. Influ-

ence functions, which quantify the effect of an infinitesimal outlier on an estimator, and asymptotic variances are studied in [38]. Moreover, LTS approaches were extended to nonlinear models [39], to LS estimation under constraint [40] and to the multivariate case in [41,42].

Concerning S-estimators, where "S" stands for "scale", they aim at minimizing the residual scale σ_s by considering the following equality:

$$\frac{1}{M}\sum_{i=1}^{M}\rho\left(\frac{r_i(\theta)}{\sigma_s}\right) = b \tag{4}$$

where b is a fixed calibration constant. S-estimators may also achieve breakdown points close to 50%. Under Gaussian assumptions, they exhibit lower efficiency (related to the Fisher information and the variance of the estimator) than maximum likelihood or standard LS estimators in the absence of contamination.

Starting from a high-breakdown estimator such as the LTS or an S-estimator, a single M-estimation correction can be applied to reduce the variance of the estimates while maintaining robustness. This led to one-step M-estimators. Finally, MM-estimators combine the high-breakdown properties of S-estimators with the improved statistical efficiency of M-estimation procedures.

Note that several extensions have been proposed to adapt robust regression methods to large-scale settings (see for instance [43,44]). The main purpose of these extensions is to reduce computational complexity. Robust and regularized LS approaches can also be combined within a unified framework [45].

In this work, we adopt two strategies to make DFA more robust: the first one is based on the LTS while the second is based on an iterative weighted least-squares (IWLS) approach. Several strategies for defining the weights exist. Based on preliminary investigations, we decided to focus our attention on the method based on Cook distance [46]. Note that the robust methods are used to estimate the local trends of the integrated centered signal or/and the Hurst exponent H. While DFA methods and robust estimation techniques are individually well known and have been extensively studied, the originality of the present work stems from their joint use in the context of Hurst exponent estimation. In addition, robust estimation techniques can be easily incorporated into existing DFA-based frameworks. Therefore, the proposed combination has the advantage of being straightforward to implement in practice.

The remainder of this paper is organized as follows. Section 2 recalls the main steps of the standard DFA. Then, in Section 3, we present the new variants. Simulation results are then given in Section 4 with data without or with outliers. After the conclusions, Appendix A recalls the Cook distance.

2. Standard DFA and Its Higher-Order Variants

In the following, the subscript $_d$ indicates that the quantity is associated with the DFA method based on polynomial local trends of degree $d \geq 1$.

The five steps of the methods applied on M consecutive samples $\{y(m)\}_{m=1,\ldots,M}$ are the following:

Step 1: =The profile y_{int} is computed by centering and integrating the signal y:

$$y_{int}(m) = \sum_{i=1}^{m}(y(i) - \mu_y) \text{ with } \mu_y = \frac{1}{M}\sum_{m=1}^{M} y(m) \tag{5}$$

Step 2: For a given value of N selected in the interval $[3; \lfloor \frac{M}{2} \rfloor]$ and introducing $L = \lfloor \frac{M}{N} \rfloor$, the first LN samples of the profile trend are estimated by decomposing the profile

into L non-overlapping sub-parts of length N, denoted as $\{y_{int,l}(n)\}_{l=1,\dots,L}$ with $n \in [\![1;N]\!]$. With a vector form, the dth-degree polynomial trend vector of the lth segment can be expressed $\forall l \in [\![1;L]\!]$ as follows:

$$\underline{\mathrm{T}}_{d,l} = \mathrm{A}_{d,l}\theta_{d,l} = \begin{bmatrix} 1 & \cdots & ((l-1)N+1)^d \\ \vdots & & \vdots \\ 1 & \cdots & (lN)^d \end{bmatrix} \begin{bmatrix} a_{0,l} \\ \vdots \\ a_{d,l} \end{bmatrix} \tag{6}$$

Step 3: Given $\underline{\Theta}_d = \left[\underline{\theta}_{d,1}^T \dots \underline{\theta}_{d,L}^T\right]^T$ and the block diagonal matrix $A_d = diag\left[\mathrm{A}_{d,1} \quad \dots \quad \mathrm{A}_{d,L}\right]$, the parameters of the local trends stored in $\hat{\Theta}_d$ are deduced by minimizing the LS criterion $\left|\left|\underline{Y}_{int,LN} - A_d\underline{\Theta}_d\right|\right|^2$:

$$\hat{\underline{\Theta}}_d = (A_d^T A_d)^{-1} A_d^T \underline{Y}_{int,LN} \tag{7}$$

Using (7), the global trend vector T_d satisfies:

$$\underline{T}_d = A_d\hat{\Theta}_d = A_d(A_d^T A_d)^{-1} A_d^T Y_{int,LN} = P_d\underline{Y}_{int,LN} \tag{8}$$

with $\underline{Y}_{int,LN}$ the column vector storing the first LN values of $y_{int}(m)$, $P_d = diag\left(\left[\mathrm{P}_{d,1}, \cdots, \mathrm{P}_{d,L}\right]\right)$ and $\{\mathrm{P}_{d,l}\}_{l=1,\dots,L}$ the orthogonal projectors onto the space spanned by the column vectors of $\{\mathrm{A}_{d,l}\}_{l=1,\dots,L}$, defined by:

$$\mathrm{P}_{d,l} = \mathrm{A}_{d,l}\left({\mathrm{A}_{d,l}}^T\mathrm{A}_{d,l}\right)^{-1}{\mathrm{A}_{d,l}}^T \tag{9}$$

Step 4: Given (8), the residual vector $\underline{R}_d = \underline{Y}_{int,LN} - \underline{T}_d$ is computed. Then, the square of the fluctuation function $F_d^2(y,N)$ is obtained as the power of the residual:

$$F_d^2(y,N) = \frac{1}{LN}||\underline{R}_d||^2 \tag{10}$$

Step 5: The previous steps are repeated for different values of N selected in the interval $[3; \lfloor \frac{M}{2} \rfloor]$. Due to (2), each pair $(\log(F_d(y,N)), \log(N))$ that is obtained should satisfy:

$$\log(F_d(y,N)) = b_d + \alpha\log(N) + \epsilon_d(N) \tag{11}$$

with b_d the vertical intercept and $\epsilon_d(N)$ the error variable. Therefore α is estimated in the LS sense using the different values of N. It is however not equal to the Hurst exponent H, but to $= H + 1$ due to the integration step.

3. Our Proposals

3.1. First Suggestion: Using a Robust LS Method

At Steps 3 and 5 of the DFA described above, the problem to be addressed in its generic form is the following: given a set of I points defined by their coordinates $(x_i, y_i)_{i=1,\dots,I}$, let us express $\underline{Y}$ the column vector storing the set of values $\{y_i\}_{i=1,\dots,I}$ as follows:

$$\underline{Y} = A\underline{\theta} + \underline{\epsilon} \tag{12}$$

where $\underline{\theta}$ is the model-parameter vector and the matrix A stores the vectors $\{\underline{x}_i\}_{i=1,\dots,M}$ defined from $\{x_i\}_{i=1,\dots,M}$. For instance, $\underline{x}_i^T = \left[x_i \quad 1\right]^T$ where $x_i = \log(N_i)$ with $N_i = 3, \dots, \lfloor \frac{M}{2} \rfloor$ when addressing Step 5. The LS estimation of $\underline{\theta}$ consists of minimizing $J = ||\underline{Y} - A\underline{\theta}||_2^2$. If $A^T A$ is invertible, $\underline{\hat{\theta}}_{LS} = (A^T A)^{-1} A^T \underline{Y}$. Instead of using a standard LS criterion, we propose to consider two robust LS methods:

1. The LTS proceeds as follows: After initializing the model-parameter vector estimated with the LS estimate ($\hat{\underline{\theta}}^{(0)} = \underline{\theta}_{LS}$), the residuals for the data are computed: $r_i^{(k)} = y_i - \underline{x}_i^T \hat{\underline{\theta}}^{(k)}$. A fixed proportion of the data (defined by the practitioner), referred to as the "inliers" and corresponding to the smallest squared residuals, are considered to estimate the model-parameter vector using the LS criterion. This process is then iterated if necessary.
2. The IWLS method is based on iteratively updating weights until convergence of the model parameters. Thus, after the initialization and residual computation (as in the TLS method), the weights $w_i^{(k)}$ are computed using the Cook distance, as recalled in Appendix A (note that Huber loss function or Tukey loss function could have been considered, but they are not developed in this article due to space constraints and the desire to limit the number of hyperparameters). The estimate of the model-parameter vector is updated as follows:

$$\hat{\underline{\theta}}^{(k+1)} = (A^T W^{(k)} A)^{-1} A^T W^{(k)} \underline{Y} \tag{13}$$

with $W^{(k)}$ the diagonal matrix defined from the weights $w_i^{(k)}$. This process is repeated until the normalized difference between two consecutive estimations is smaller than a threshold.

It should be noted that two cases will be considered: the robust methods will be applied only at Step 5 or used at Steps 3 and 5 of the standard DFA. When applied at Step 5, they have the advantage of avoiding any *a priori* choice of the values of N.

3.2. Second Suggestion: Unbiased Estimate of $F_d^2(y, N)$

At Step 4, the definition (10) of $F_d^2(y, N)$ can be explained as follows: the values of $\underline{R}_d$ are decimated by a factor N. The biased empirical estimates of the variances of the N resulting decimated sequences are then computed and averaged. We propose a slightly different expression of $F_d^2(y, N)$ by considering the average of the unbiased empirical variance estimates. These quantities will be referred to as the biased (B) and unbiased (UB) estimates of $F_d^2(y, N)$.

The difference between the logarithm of the UB estimate and that of the B estimate of $F_d^2(y, N)$ is equal to $\log(\frac{L}{L-1})$. As a consequence, the difference between the corresponding estimate of the Hurst exponent can be written as $\begin{bmatrix} 1 & 0 \end{bmatrix} (A^T A)^{-1} A^T \delta \underline{F}$ with $\delta \underline{F} = \begin{bmatrix} \lfloor \frac{M}{3} \rfloor & \dots & 2 \\ \lfloor \frac{M}{3} \rfloor - 1 & & \end{bmatrix}^T$. The entries of $\delta \underline{F}$ increase. In addition, the first column of the matrix A consists of increasing values ranging from $\log(3)$ to $\log(\lfloor \frac{M}{2} \rfloor)$ whereas the second column is a column of ones. Therefore, the estimate of H based on the UB estimate of $F_d^2(y, N)$ is necessarily larger than the B estimate.

4. Simulation Results

In addition to the standard DFA, the following methods are studied: (i) the approach in which the local trends are estimated using the LS criterion, but H is estimated with a robust LS method (LTS or the Cook-distance based IWLS method), (ii) the approach in which both the local trends and H are estimated with a robust LS method.

Several simulations were conducted on different sets of data. A subset of the obtained results for data without or with outliers is presented to provide illustrative examples and objective performance measures. These elements are intended to give the reader a clear overview of the behavior of the methods.

4.1. Application on Data with No Outlier

Let us consider a Brownian noise with a variance equal to 4 around a decaying exponential trend. In Table 1, the results are given for different number M of samples available varying from 125 to 2000. For LTS, the proportion is set at 70%. Some illustrations are also given in Figures 1 and 2.

Using the unbiased estimates of $F_d^2(y, N)$ does not necessarily improve the accuracy of the estimation of H when using the LS criterion at Step 3 of the DFA. However, when it is combined with the IWLS approach at Steps 3 and 5, the result is often the most accurate.

Table 1. Mean $\pm$ standard deviation of the estimation of $1 + H$ for the various methods for different numbers M of samples available, based on 1000 trials. True result: $1 + H = 1.5$. Bold values correspond to the best result.

M	**LS/B/LS**	**LS/B/IWLS**	**IWLS/B/IWLS**	**LS/B/LTS**	**LTS/B/LTS**
125	1.43 ± 0.16	1.43 ± 0.18	1.29 ± 0.19	$\mathbf{1.46 \pm 0.19}$	1.55 ± 0.24
250	1.44 ± 0.14	1.44 ± 0.15	1.35 ± 0.16	1.46 ± 0.18	1.54 ± 0.24
500	1.43 ± 0.14	1.43 ± 0.14	1.38 ± 0.15	1.45 ± 0.17	1.47 ± 0.24
1000	1.44 ± 0.12	1.44 ± 0.12	1.41 ± 0.13	1.46 ± 0.16	1.45 ± 0.22
2000	1.45 ± 0.12	1.45 ± 0.13	1.43 ± 0.13	$\mathbf{1.47 \pm 0.16}$	1.45 ± 0.21
M	**LS/UB/LS**	**LS/UB/IWLS**	**IWLS/UB/IWLS**	**LS/UB/LTS**	**LTS/UB/LTS**
125	1.58 ± 0.16	1.59 ± 0.18	$\mathbf{1.46 \pm 0.19}$	1.60 ± 0.19	1.70 ± 0.25
250	1.57 ± 0.14	1.57 ± 0.15	$\mathbf{1.50 \pm 0.16}$	1.59 ± 0.18	1.67 ± 0.23
500	1.56 ± 0.14	1.56 ± 0.14	$\mathbf{1.51 \pm 0.15}$	1.58 ± 0.18	1.61 ± 0.22
1000	1.55 ± 0.12	1.56 ± 0.12	$\mathbf{1.53 \pm 0.13}$	1.58 ± 0.16	1.57 ± 0.19
2000	1.56 ± 0.12	1.56 ± 0.13	1.55 ± 0.13	1.59 ± 0.17	1.57 ± 0.20

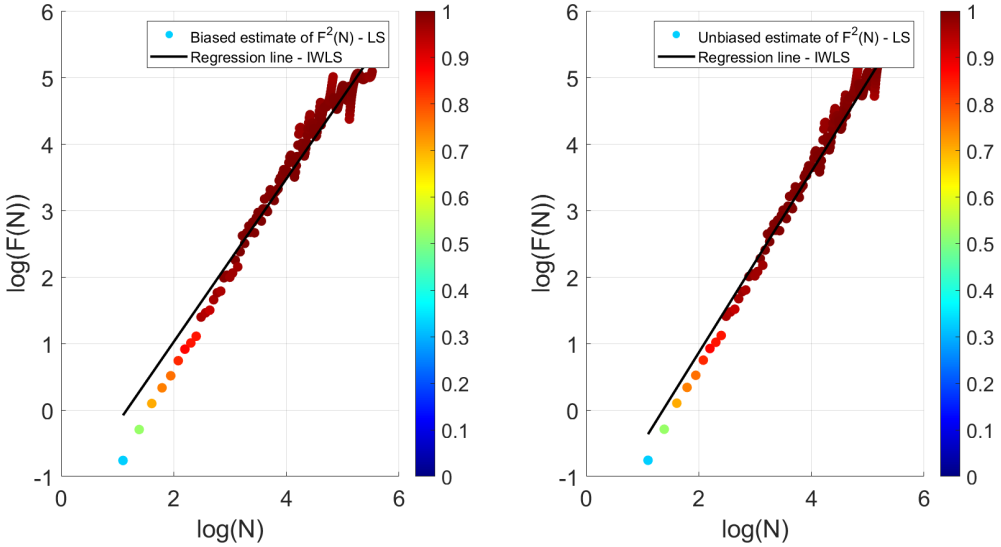


Figure 1. LS/B/IWLS and LS/UB/IWLS methods using the biased and unbiased estimations of $F_d^2(y, N)$. IWLS approach used only for Step 5. The colormap represents the weights for each point. 500 samples.

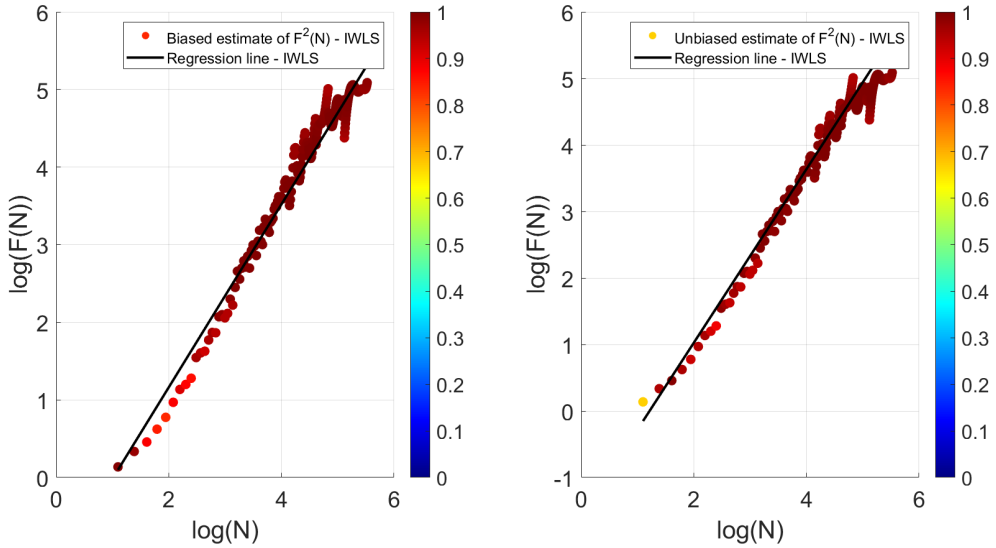


Figure 2. IWLS/B/IWLS and IWLS/UB/IWLS methods using the biased and unbiased estimations of $F_d^2(y, N)$. IWLS approach used for Steps 3 and 5. The colormap represents the weights for each point. 500 samples.

4.2. Application on Data Contaminated by Outliers

Let us now consider a signal that includes a slowly decreasing exponential trend and a Weierstrass function. Pointwise outliers are added, with a probability of occurrence of 4%. Their magnitudes are Gaussian distributed with zero mean and a standard deviation equal to 4 times that of the original signal. Note that the Weierstrass function is a real-valued function continuous everywhere but differentiable nowhere. It was generated using the Matlab R2021a toolbox FracLab [47] (it uses a stochastic version of $W(x) = \sum_{n=1}^{+\infty} \lambda^{(-nH)} \sin(2\pi\lambda^n x)$ with $\lambda \geq 2$, by introducing two random parameters, c_n and a_n, uniformly distributed in $[0, 1]$, to obtain $\hat{W}(x) = \sum_{n=1}^{+\infty} c_n \lambda^{(-nH)} \sin(2\pi(\lambda^n x + a_n)))$ with H equal to 0.3 (for illustration purposes)—see Figure 3. Some results are given in Figures 4–6 for $M = 1000$ and in Table 2. Results confirm the relevance of robust methods.

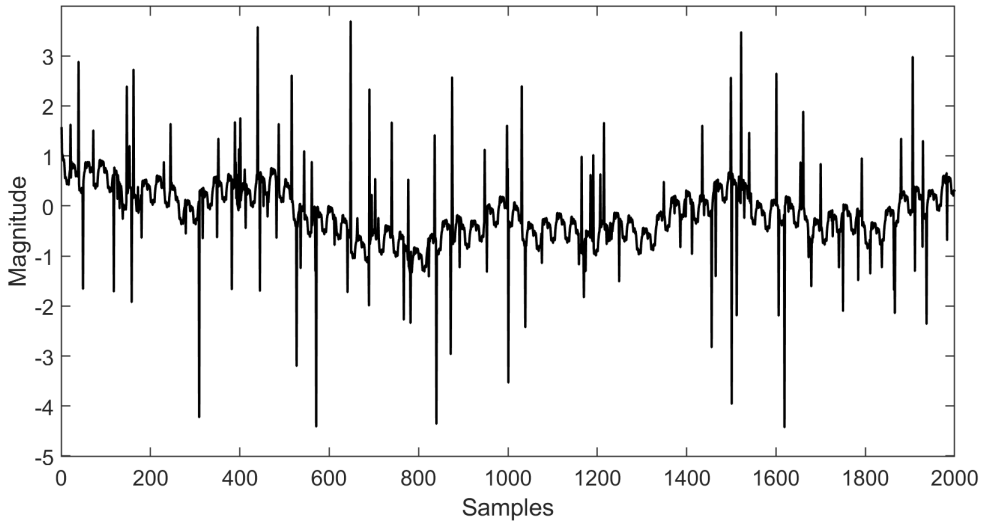


Figure 3. Realization of a signal under study contaminated by outliers.

Table 2. Mean ± standard deviation of the estimation of $1 + H$ for the 10 methods for different numbers M of samples (with outliers) available, based on 1000 trials. True result: $1 + H = 1.3$. Bold values correspond to the best result.

M	**LS/B/LS**	**LS/B/IWLS**	**IWLS/B/IWLS**	**LS/B/LTS**	**LTS/B/LTS**
1000	1.09 ± 0.03	1.11 ± 0.03	1.10 ± 0.03	1.05 ± 0.08	1.21 ± 0.16
2000	1.14 ± 0.01	1.15 ± 0.02	1.14 ± 0.02	1.26 ± 0.11	1.39 ± 0.11
M	**LS/UB/LS**	**LS/UB/IWLS**	**IWLS/UB/IWLS**	**LS/UB/LTS**	**LTS/UB/LTS**
1000	1.21 ± 0.03	1.23 ± 0.03	1.22 ± 0.03	1.24 ± 0.07	$\mathbf{1.35 \pm 0.14}$
2000	1.25 ± 0.01	$\mathbf{1.27 \pm 0.02}$	$\mathbf{1.27 \pm 0.02}$	1.24 ± 0.07	1.45 ± 0.18

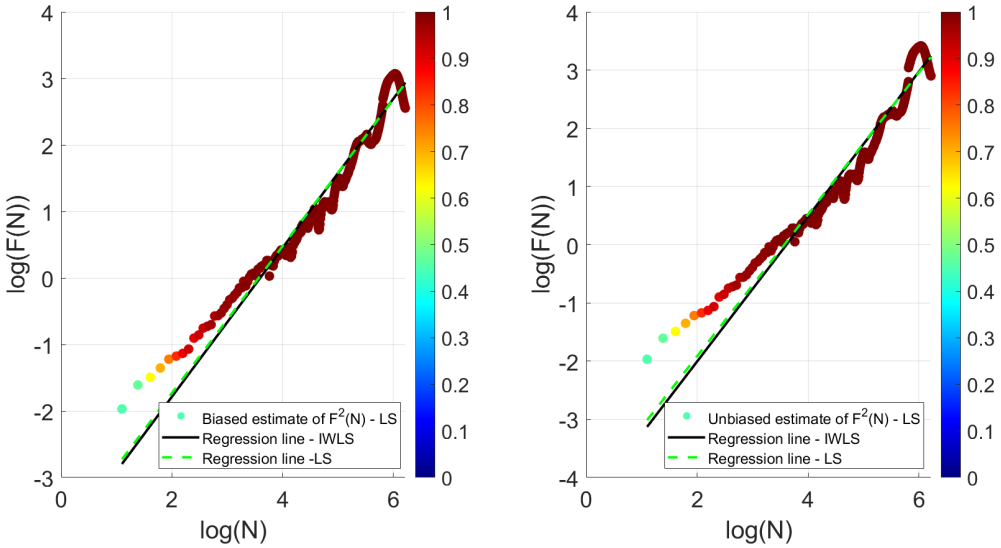


Figure 4. LS/B/IWLS and LS/UB/IWLS methods using the biased and unbiased estimations of $F_d^2(y, N)$. IWLS approach only used for Step 5 on 1000 samples with outliers. The colormap represents the weights for each point.

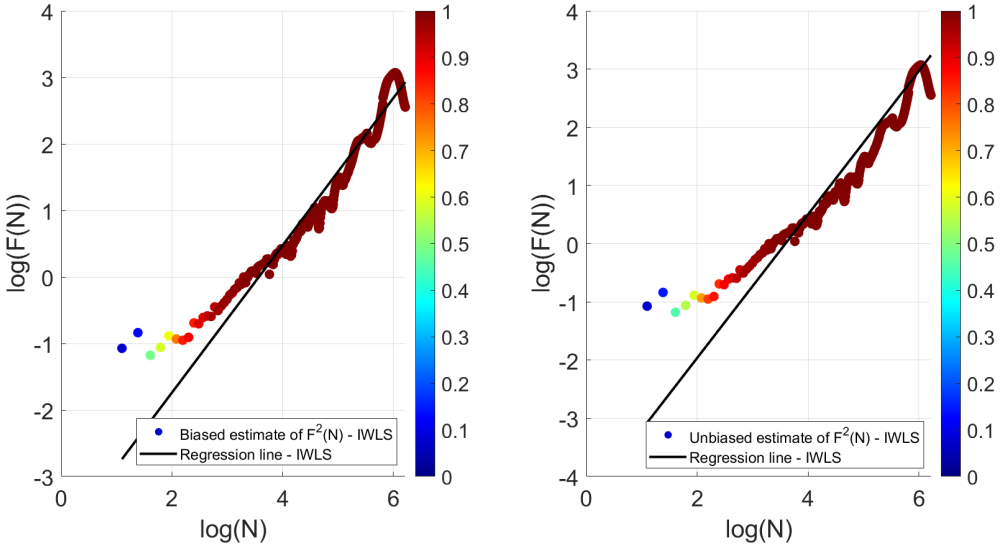


Figure 5. IWLS/B/IWLS and IWLS/UB/IWLS methods using the biased and unbiased estimations of $F_d^2(y, N)$. IWLS approach used for both Steps 3 and 5 on 1000 samples with outliers. The colormap represents the weights associated with each point.

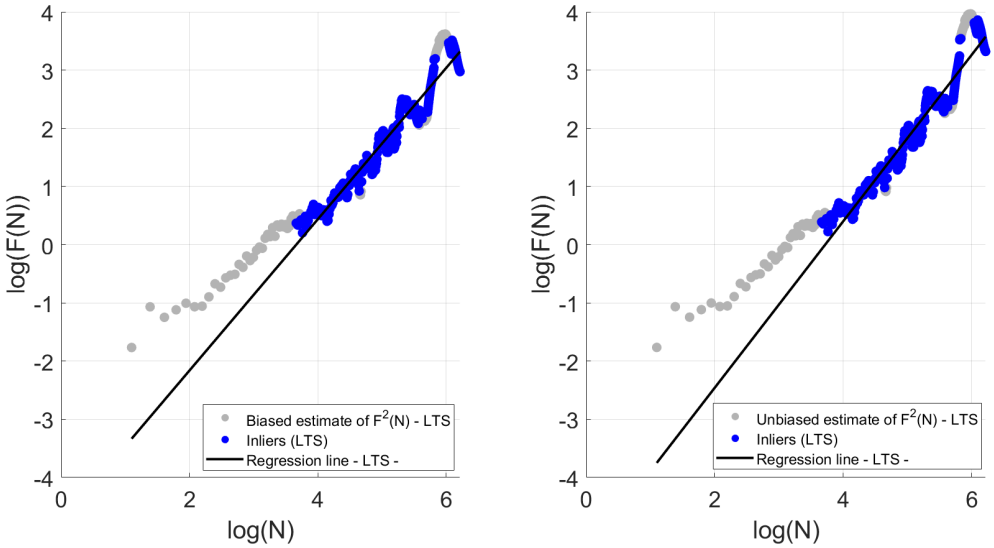


Figure 6. LTS/B/LTS and LTS/UB/LTS methods using the biased and unbiased estimations of $F_d^2(y, N)$. LTS approach used for both Steps 3 and 5 on 1000 samples with outliers. The colormap represents the weights for each point.

5. Conclusions and Perspectives

The proposed approaches consider all values of N starting from $N = 3$, without any *a priori* selection. This can be done by using robust LS methods. It also makes the method less dependent on practitioner tuning. Based on the various simulations we performed, using the IWLS only to estimate the Hurst exponent from the estimations of the fluctuation functions seems relevant. It is true that the computation cost is increased when replacing a standard LS estimation by a robust one: When dealing with IWLS method, the closed-form LS solution is replaced by an iterative sequence of weighted least-squares problems solved until convergence. Note that this approach could also be applied to variants such as CDFA. Although the methods considered in this work are well established and have been widely used, the literature on robust approaches remains active, encompassing alternative families such as S-estimators and MM-estimators and more recent developments such as [48]. In this context, future work will aim to evaluate these variants and their integration within DFA-based frameworks.

Author Contributions: Conceptualization, E.G.; methodology, E.G.; software, E.G.; validation, E.G. and P.L.; formal analysis, E.G.; investigation, E.G. and P.L.; writing—original draft preparation, E.G. and P.L.; writing—review and editing, E.G. and P.L.; visualization, E.G.; supervision, E.G.; project administration, E.G. and P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset generated and analyzed during the current study is available from the corresponding author on a reasonable request.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AFA	Adaptive Fractal analysis
DMA	Detrended Moving Average
DFA	Detrended Fluctuation Analysis
EM	Expectation Maximization
LOWESS	Locally Weighted Scatterplot Smoothing
LS	Least Squares
RDFA	Regularized Detrended Fluctuation Analysis

Appendix A

Given the LS estimate, let us introduce the residuals for each point by:

$$r_i = y_i - \underline{x}_i^T \hat{\underline{\theta}}_{LS} \tag{A1}$$

Cook distance is defined by [46]:

$$D_i = \frac{(\hat{\underline{\theta}}_{LS} - \hat{\underline{\theta}}_{(i)})^T A^T A (\hat{\underline{\theta}} - \hat{\underline{\theta}}_{(i)})}{p\hat{\sigma}^2} \tag{A2}$$

with $\hat{\underline{\theta}}_{(i)}$ the LS estimate of the model parameters without considering the *ith* point and p the parameter-vector size.

Let us find another expression for D_i. To this end, let us first recall that if the matrix B is invertible and u, v are vectors, one gets:

$$(B + uv^T)^{-1} = B^{-1} - \frac{B^{-1}uv^TB^{-1}}{1 + v^TB^{-1}u} \tag{A3}$$

Let us now build $A_{(i)}$ and $\underline{Y}_{(i)}$ by removing the *ith* row of A and $\underline{Y}$. In that case, one has:

$$\hat{\theta}_{(i)} = (A_{(i)}^T A_{(i)})^{-1} A_{(i)}^T \underline{Y}_{(i)} \tag{A4}$$

Let us rewrite (A4) having in mind that $A_{(i)}^T A_{(i)} = A^T A - \underline{x}_i \underline{x}_i^T$ and $A_{(i)}^T \underline{Y}_{(i)} = A^T \underline{Y} - x_i y_i$. Applying (A3) on $(A^T A - x_i x_i^T)^{-1}$, one gets:

$$(A^T A - x_i x_i^T)^{-1} = (A^T A)^{-1} + \frac{(A^T A)^{-1} x_i x_i^T (A^T A)^{-1}}{1 - h_i} \tag{A5}$$

with $h_i = x_i^T (A^T A)^{-1} x_i$. Combining (A4) and (A5) leads to:

$$\hat{\theta}_{(i)} = (A^T A - \underline{x}_i \underline{x}_i^T)^{-1} (A^T \underline{Y} - \underline{x}_i y_i) \tag{A6}$$

$$= \left[(A^T A)^{-1} + \frac{(A^T A)^{-1} \underline{x}_i \underline{x}_i^T (A^T A)^{-1}}{1 - h_i} \right] (A^T \underline{Y} - \underline{x}_i y_i) \tag{A7}$$

Using (A1) and after mathematical simplifications, one gets:

$$\hat{\theta}_{(i)} = \hat{\underline{\theta}}_{LS} - \frac{(A^T A)^{-1} \underline{x}_i r_i}{1 - h_i} \tag{A8}$$

Removing the *ith* point depends on its residual r_i and h_i. Combining (A2) and (A8) gives another expression of D_i:

$$D_i = \frac{r_i^2 h_i}{p\hat{\sigma}^2 (1 - h_i)^2} \tag{A9}$$

D_i is large if the point has a large residual.

D_i can then be used to define a weight $w_i = \frac{1}{1+D_i}$ or $w_i = exp(-\alpha D_i)$ with α that can be set to a predefined value. Alternatively, one can examine the histogram of D_i and choose α such that a certain proportion of the data receives a weight below 0.5 for example.

References

1. Mandelbrot, B.B.; Van Ness, J.W. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Rev.* **1968**, *10*, 422–437. [CrossRef]
2. Lillo, F.; Farmer, J.D. The long memory of the efficient market. *Stud. Nonlinear Dyn. Econom.* **2004**, *8*, 35. [CrossRef]
3. Peng, C.K.; Buldyrev, S.V.; Havlin, S.; Simons, M.; Stanley, H.E.; Goldberger, A.L. Mosaic organization of DNA nucleotides. *Phys. Rev. E* **1994**, *49*, 1685–1689. [CrossRef] [PubMed]
4. Kiyono, K.; Tsujimoto, Y. Time and frequency domain characteristics of detrending-operation-based scaling analysis: Exact DFA and DMA frequency responses. *Phys. Rev. E* **2016**, *94*, 012111. [CrossRef] [PubMed]
5. Riley, M.A.; Bonnette, S.; Kuznetsov, N.; Wallot, S.; Gao, J. A Tutorial Introduction to Adaptive Fractal Analysis. *Front. Physiol.* **2012**, *3*, 371. [CrossRef] [PubMed]
6. Hodrick, R.J.; Prescott, E.C. Postwar U.S. Business Cycles: An Empirical Investigation. *J. Money Credit Bank.* **1997**, *29*, 1–16. [CrossRef]
7. Berthelot, B.; Grivel, E.; Legrand, P.; André, J.M.; Mazoyer, P. Regularized DFA to study the gaze position of an airline pilot. In Proceedings of the 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands, 18–21 January 2021.
8. Berthelot, B.; Grivel, E.; Legrand, P.; André, J.M.; Mazoyer, P. Alternative Ways to Compare the Detendred Fluctuation Analysis and its Variants. Application to Visual Tunneling Detection. *Digit. Signal Process.* **2020**, *108*, 102865. [CrossRef]
9. Alessio, E.; Carbone, A.; Castelli, G.; Frappietro, V. Second-order moving average and scaling of stochastic time series. *Eur. Phys. J. B* **2002**, *27*, 197–200. [CrossRef]
10. Xu, L.; Ivanov, P.C.; Hu, K.; Chen, Z.; Carbone, A.; Stanley, H.E. Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Phys. Rev. E* **2005**, *71*, 051101. [CrossRef] [PubMed]
11. Arianos, S.; Carbone, A.; Türk, C. Self-similarity of higher-order moving averages. *Phys. Rev. E* **2011**, *84*, 046113. [CrossRef] [PubMed]
12. Berthelot, B.; Grivel, E.; Legrand, P. New Variants of DFA based on LOESS and LOWESS methods: Generalization of the detrended moving average. In Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, Ontario, Canada, 6–11 June 2021.
13. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Bunde, A.; Havlin, S.; Stanley, H.E. Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series. *Phys. A Stat. Mech. Its Appl.* **2002**, *316*, 87–114. [CrossRef]
14. Tsujimoto, Y.; Miki, Y.; Watanabe, E.; Hayano, J.; Yamamoto, Y.; Nomura, T.; Kiyono, K. Fast algorithm of long-range cross-correlation analysis using Savitzky-Golay detrending filter and its application to biosignal analysis. In Proceedings of the 2017 International Conference on Noise and Fluctuations (ICNF), Vilnius, Lithuania, 20–23 June 2017.
15. Podobnik, B.; Stanley, H.E. Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. *Phys. Rev. Lett.* **2008**, *100*, 084102. [CrossRef] [PubMed]
16. Taqqu, M.S.; Teverovsky, V.; Willinger, W. Estimators for long range dependence: An empirical study. *Fractals* **1995**, *3*, 785–788. [CrossRef]
17. Taqqu, M.S.; Teverovsky, V. On Estimating the Intensity of Long Range Dependence in Finite and Infinite Variance Time Series. In *A Practical Guide to Heavy Tails: Statistical Techniques and Applications;* 1996; pp. 177–217.
18. Maraun, D.; Rust, H.W.; Timmer, J. Tempting long-memory on the interpretation of DFA results. *Nonlinear Processes Geophys.* **2004**, *11*, 495–503. [CrossRef]
19. Crato, N.; Linhares, R.R.; Lopes, S.R.C. Statistical properties of detrended fluctuation analysis. *J. Stat. Comput. Simul.* **2010**, *80*, 625–641. [CrossRef]
20. Kiyono, K. Establishing a direct connection between detrended fluctuation analysis and Fourier analysis. *Phys. Rev. E* **2015**, *92*, 042925. [CrossRef] [PubMed]
21. Höll, M.; Kantz, H. The relationship between the detrendend fluctuation analysis and the autocorrelation function of a signal. *Eur. Phys. J. B* **2015**, *88*, 327. [CrossRef]
22. Höll, M.; Kantz, H. The fluctuation function of the detrended fluctuation analysis — investigation on the AR(1) process. *Eur. Phys. J. B* **2015**, *88*, 126. [CrossRef]
23. Höll, M.; Kantz, H.; Zhou, Y. Detrended fluctuation analysis and the difference between external drifts and intrinsic diffusion-like non-stationarity. *Phys. Rev. E* **2016**, *94*, 042201. [CrossRef] [PubMed]

24. Kiyono, K. Theory and applications of detrending-operation-based fractal-scaling analysis. In *Proceedings of the International Conference on Noise and Fluctuations (ICNF)*; IEEE: New York, NJ, USA, 2017; pp. 1–4.
25. Løvsletten, O. Consistency of detrended fluctuation analysis. *Phys. Rev. E* **2017**, *96*, 012141. [CrossRef] [PubMed]
26. Höll, M.; Kiyono, K.; Kantz, H. Theoretical foundation of detrending methods for fluctuation analysis such as detrended fluctuation analysis and detrending moving average. *Phys. Rev. E* **2019**, *99*, 033305. [CrossRef] [PubMed]
27. Berthelot, B.; Grivel, E.; Legrand, P.; Donias, M.; André, J.M.; Mazoyer, P.; Ferreira, T. 2D Fourier Transform Based Analysis Comparing the DFA with the DMA. In Proceedings of the 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019.
28. Sikora, G.; Höll, M.; Gajda, J.; Kantz, H.; Chechkin, A.; Wylomanska, A. Probabilistic properties of detrended fluctuation analysis for Gaussian processes. *Phys. Rev. E* **2020**, *101*, 032114. [CrossRef] [PubMed]
29. Grivel, E.; Berthelot, B.; Legrand, P.; Giremus, A. DFA-based abacuses providing the Hurst exponent estimate for short-memory processes. *Digit. Signal Process.* **2021**, *116*, 103102. [CrossRef]
30. Kantelhardt, J.W.; Koscielny-Bunde, E.; Rego, H.H.A.; Havlin, S.; Bunde, A. Detecting long-range correlations with detrended fluctuation analysis. *Phys. A Stat. Mech. Its Appl.* **2001**, *295*, 441–454. [CrossRef]
31. Atkinson, A. Residuals and Influence in Regression. *Biometrics* **1983**.
32. O'Leary, D. Robust regression computation using iteratively reweighted least squares. *SIAM J. Matrix Anal. Appl.* **1990**, *11*, 466–480. [CrossRef]
33. Hampel, F.R.; Ronchetti, E.M.; Rousseeuw, P.; Stahel, W.A. *Robust Statistics. The Approach Based on Influence Functions*; John Wiley: Hoboken, NJ, USA, 1986.
34. Rousseeuw, P.; Leroy, A.M. *Robust Regression and Outlier Detection*; Wiley: Hoboken, NJ, USA, 1987.
35. Gervini, D.; Yohai, V.J. A Class of robust and fully efficient regression estimators. *Anntials Stat.* **2002**, *30*, 583–616. [CrossRef]
36. Mua, B.; Bai, E.W.; Zheng, W.X. Tilted least squares robust estimators. *Automatica* **2026**, *184*, 112699. [CrossRef] [PubMed]
37. Agulló, J. New algorithms for computing the least trimmed squares regression estimator. *Comput. Stat. Data Anal.* **2001**, *36*, 4425–4439. [CrossRef]
38. Jann, B.; Verardi, V.; Vermandele, C. Influence functions for LTS and reweighted LTS estimators based on moment conditions. *Stat. Probab. Lett.* **2026**, *234*, 110693. [CrossRef]
39. Prochazka, B. Regression quantiles and trimmed least squares estimator in the nonlinear regression model. *Comput. Stat. Data Anal.* **1988**, *6*, 385–391. [CrossRef]
40. Li, L.M. An algorithm for computing exact least-trimmed squares estimate of simple linear regression with constraints. *Comput. Stat. Data Anal.* **2005**, *48*, 717–734. [CrossRef]
41. Agullo, J.; Croux, C.; Van Aelst, S. The multivariate least-trimmed squares estimator. *J. Multivar. Anal.* **2008**, *99*, 311–338. [CrossRef]
42. Jung, K. Multivariate least-trimmed squares regression estimator. *Comput. Stat. Data Anal.* **2005**, *48*, 307–316. [CrossRef]
43. Rousseeuw, P.J.; Van Driessen, K. Computing LTS Regression for large datasets. *Data Min. Knowl. Discov.* **2006**, *12*, 29–45. [CrossRef]
44. Xie, Z.; Feng, X.; Chen, X. Partial least trimmed squares regression. *Chemom. Intell. Lab. Syst.* **2022**, *221*, 104486. [CrossRef]
45. Bai, L.; Zhang, J.; Qiao, L. A nonconvex sparse recovery method for DOA estimation based on the trimmed lasso. *Digit. Signal Process.* **2024**, *153*, 104628. [CrossRef]
46. Cook, R. Detection of Influential Observation in Linear Regression. *Technometrics* **1977**, *19*, 15–18. [CrossRef] [PubMed]
47. Legrand, P. Denoising Interpolation by Analysis of the hölderian Regularity: Application to the Modelisation of Road/Tyre Friction. PhD Thesis, Central School of Nantes, Nantes, France, 2004.
48. Nguyen, T.; Welsch, R. Outlier detection and least trimmed squares approximation using semi-definite programming. *Comput. Stat. Data Anal.* **2010**, *54*, 3212–3226. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.