Computers and Chemical Engineering 147 (2021) 107254

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Review

A review on robust M-estimators for regression analysis

D.Q.F. de Menezes[a,*], D.M. Prata[b], A.R. Secchi[a], J.C. Pinto[a]

[a] Programa de Engenharia Química, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, CEP 21941-972, Brasil
[b] Departamento de Engenharia Química e de Petróleo, Universidade Federal Fluminense, Niterói, RJ, CEP 24210-240, Brasil

A R T I C L E I N F O

Article history:
Received 11 August 2020
Revised 12 December 2020
Accepted 2 February 2021
Available online 5 February 2021

Keywords:
Regression analysis
Data reconciliation
Parameter estimation
Robust statistic
M-estimator

A B S T R A C T

Regression analysis constitutes an important tool for investigating the effect of explanatory variables on response variables. When outliers and bias errors are present, the weighted least squares estimator can perform poorly. For this reason, alternative robust techniques have been studied in several areas of science. However, often these different scientific communities are disconnected from each other, culminating in the scarcity of knowledge exchange among these areas. Thus, this paper presents a review on robust M-estimators in various knowledge areas. 50 (48 robust) M-estimators are illustrated, including the Weighted Least Squares estimator (non-robust), the Contaminated Normal estimator (quasi-robust), the Huber estimator (monotone), the Correntropy estimator (soft-redescending), the Smith estimator (hard-redescending), and the adaptive Barron and Generalized T-distribution. The mathematical functions that describe the estimators and their respective graphical forms are presented. The tuning constants of all these estimators, for 90%, 95%, 98%, and 99% relative efficiency levels in respect to the Normal distribution are also presented.

Contents

* Corresponding author.
E-mail address: dmenezes@coppe.ufrj.br (D.Q.F. de Menezes).

https://doi.org/10.1016/j.compchemeng.2021.107254

1. Introduction

Regression analysis is a statistical tool widely used in almost all areas of knowledge that seeks to fit a mathematical model to an experimental data set. This is particularly true for engineering, especially chemical process engineering. In these fields, the term "regression analysis" usually refers to problems of data reconciliation, parameter estimation, or both. Although there are several methods for obtaining the statistically coherent value of variables and estimating the parameters of the mathematical model based on available data, the weighted least squares (WLS) method is the one used most frequently and assumes that measurement errors follow the Normal distribution model (Prata et al., 2008a; Prata et al., 2009; Prata et al., 2009). The Normal distribution model, developed in 1733 by Abraham de Moivre, and later attributed to C. F. Gauss, is used most often because of its general acceptance, elegant statistical properties, and ease of calculation (Bellhouse et al., 2007; Zoubir et al., 2018).

Bard (1974), in his classic book on parameter estimation, points out the main reasons for using the Normal distribution model:

- Generally describes well the fluctuations of large number of experimental measurements;
- As the number of experimental disturbances increases, many distributions approach the Normal distribution (Central Limit Theorem);
- Having defined the mean and variance of a set of measurements and using the concepts of variational calculus, it becomes possible to show that the Normal distribution is the one that demands the input of minimum amount of additional information about a particular problem;
- Easy mathematical treatment, allowing the theoretical derivation of well-known sample statistics, including the t-Student, Chi-square (χ^2) and F-Fisher distributions.

Rey (1983) pointed out that the adoption of weighted least squares techniques was considered to be one of the best approaches for resolution of a regression problem. As a matter of fact, the "dogma" of normality has been widely accepted; observations that somehow denied this dogma were frequently regarded as wrong and discarded.

Unfortunately, the mathematical elegance that makes this classic method (Gaussian model) so popular depends on a series of assumptions that are often unrealistic or not always applicable, as actual sampled measurements can follow other underlying statistical distributions and can be contaminated by gross errors that do not necessarily follow the underlying distribution of measurements. The assumption of Normal distribution can be severely violated if one or more outliers are present in the measured data set, even though most of the data conforms to a Normal distribution, resulting in poor or deviated estimates, as presented by Tukey (1960). This author has shown that many estimators that perform optimally for data sampled from a Normal distribution perform poorly if minor changes are made to this distribution (contamination). Tukey (1960) anticipated that the classical regression procedure based on the Normal distribution, or WLS estimator, is excessively sensitive to small uncertainties about the idealized hypotheses. Long before, Legendre (1805), one of the pioneers in using the WLS regression method, explicitly predicted the necessity to reject spurious values (apud (Stigler, 1973)).

The need for "robustness" in data analysis appears in many of Tukey's works, notably in the 1962 article entitled "The Future of Data Analysis" (Tukey, 1962) specifically in the phrase:

"We need to tackle old problems in more realistic framework".

Thus, in the 1960s, a new field of mathematical statistics called **robust statistics** was born, whose foundations include the work of Tukey (1960), the article by Huber (1964) and the thesis by Hampel (1968) - the founding fathers of robust statistics.

According to Rey (1983) and Stigler (2010), the term "robustness" for use in the field of statistical mathematics was introduced by G.E.P. Box in 1953. Qualitatively, it can be understood as "insensitivity to small deviations from idealized hypotheses" for which the estimator is optimized (Huber and Ronchetti, 2009). The word "small" can have two different interpretations: the first refers to a small number of considerable divergences (spurious values) and the second to a considerable number of small divergences (non-normal distribution). Stigler (1973) commented that the term "robustness" refers to the sensitivity of procedures to deviations from idealized hypotheses, particularly the hypothesis of normality - provided it is employing well-defined procedures.

Robust statistics aims to provide methods that emulate conventional statistical methods but are not unduly affected by spurious values or other small deviations from the reference statistical distribution model. It is particularly important to note that, according to Box (1953):

"All models are wrong, but some are useful".

The three main classes of robust estimators are:

- L-estimators (linear combinations of order statistics of the observations);
- R-estimators (estimator based on waste ranking);
- M-estimators (generalizations of a Maximum Likelihood estimator).

Although there are other classes of robust estimators (D, S, τ, CM, GM, GS, MM, RM, LMS, and TS), these classes have been rarely applied to regression analyses in chemical engineering problems (Prata, 2009; Kodamana et al., 2018).

L-estimators are usually extremely simple and have robust statistics, such as the sample median and trimmed mean (α-trimmed mean). Astronomers from ancient Egypt and Rome usually discarded extreme data to average observations, indicating that they were the first to use the L-estimator α-trimmed mean (Vichare, 1993). R-estimators involve the classification of residues. The classification of a sample is a mapping of n real numbers, assuming the ordering of the data. Thus, the calculation of weights and interpretations are facilitated and generally resistant to outliers. The L- and R-estimators are therefore useful in robust statistics, such as descriptive statistics, statistical education, and when calculating the full statistical distribution of a particular problem becomes very costly. For these reasons, the L- and R-estimators play key roles in many nonparametric statistical approaches. However, currently, M-estimators are preferred, although they are computationally more expensive. M-estimators are simpler to handle because the shape of the estimator is fixed by a function. Indeed, definition of the robust properties of the L- and R-estimators are difficult to obtain *a priori*. However, this problem does not exist with M-estimators. If one decides to calculate a robust estimate with good model efficiency over a known distribution, then one must select a limited robust M-estimator with previously determined good efficiency over the known distribution. If one suspects that spurious values can be present with large deviations, then one must select a strongly robust M-estimator for correction or removal of these spurious measurements (Huber and Ronchetti, 2009; Jurečkova, 1984).

According to Albuquerque and Biegler (1996) and Prata et al. (2009), an advantage of M-estimators is that outlier detection and regression problems can be solved simultaneously, unlike sequential approaches based on regression residues. M-estimators are robust because of their intrinsic mathematical structure, which renders the estimator less sensitive to spurious deviations (Rey, 1983; Huber and Ronchetti, 2009). These estimators tend to value the majority of the data located around the mean and ignore the influence of spurious values (usually located far from the mean) simultaneously. Thus, an accurate regression can be performed even if nothing is known *a priori* about outliers or the structure of data errors. In addition, nothing prevents exploratory methods from being used to detect gross errors and obtain more information about the statistical properties of the data once residues are estimated. Although the main concepts of robust statistics have been formally developed recently, the first robust M-estimators were proposed during the early development of statistics as presented in the studies of Stephen Mack Stigler, who analyzed the history of statistics before the 1900's:

- Roger Joseph Boscovich (1757) used the absolute minimum value estimator, resulting from the hypothesis that measurement errors follow a Laplacian distribution (de Laplace, 1774) for linear regression of observed data (apud (Stigler, 1984));
- Daniel Bernoulli (1785), based on the unpublished paper of 1769, proposed the iterative mean reweighting algorithm, using a semi-circle function (apud (Stigler, 2010));
- Simon Newcomb (1886) used combinations of Normal distributions to solve regression problems (apud (Stigler, 1973));
- Smith (1888) developed the first robust M-estimator, the Smith estimator, that fits completely into the modern notation (apud (Stigler, 1980)).

However, it was G. Galileo in 1632 who first reported that measurement errors deserve systematic and scientific treatment (Zoubir et al., 2018). This shows that the discussion regarding the statistical treatment of real data has been in progress since the beginning of statistic. The literature shows that robust M-estimators constitute powerful tools for development of sampling strategies and solution of laboratory or industrial data regression problems (parameter estimation, data reconciliation, or both simultaneously), when idealized hypotheses are inaccurate, as they are capable of ignoring atypical values (spurious values), due to their mathematical structure. For these reasons, robust M-estimators have been applied in several fields, including:

- Electrical, electronic and telecommunications engineering (Merrill and Schweppe, 1971);
- Image processing (Charbonnier et al., 1997; Stewart, 1999; Arya et al., 2007);
- Econometrics and finance (Ronchetti and Trojani, 2001; Martin and Simin, 2003);
- Civil engineering, meteorology and geodesy (Krarup, 1980; Berberan, 1995; Wieser and Brunner, 2001);
- Chemical engineering (Özyurt and Pike, 2004; Prata et al., 2008a);
- Astronomy (Wu and Wu, 2005);
- Mechanical engineering (Pennacchi, 2008);
- Petrochemical engineering (Prata et al., 2010);
- Nuclear engineering (Valdetaro and Schirru, 2011);
- Industrial pharmaceutics (Liu et al., 2018; Su et al., 2019);
- Medical, biomedical and biotechnological applications (Tabatabai et al., 2014).

Based on the previous paragraphs, main purpose of the present work is surveying the applications of robust M-estimators for data regression problems (data reconciliation and parameter estimation) that are subject to gross errors. A collection of 48 robust models from numerous areas of science and covering a long period of time (from 1888 to 2019) are disclosed, including presentation of their main characteristics and performances when applied to different applications, mainly in the chemical engineering field. Besides, the underlying theory is presented, with emphasis on concepts associated with robust statistics, including discussions about the influence function (M-estimator derivative), discontinuous estimators, relationship between robustness and relative efficiency in respect to a reference distribution (generally Normal) and tuning of estimator efficiency. In particular, parameter tuning is discussed in depth (including the numerical stability), so that parameter values are presented for levels of 90%, 95%, 98% and 99% efficiency in respect to the Normal distribution for all analyzed M-estimators for the first time, including the computational codes used for tuning computations. This article is therefore expected to provide the scientific community with opportune directions when it comes to selecting appropriate M-estimators for a particular application.

2. Foundations of robust M-estimators

There are many classes of robust estimators. The most important class for the field of data rectification, as described in Albuquerque and Biegler (1996), Özyurt and Pike (2004) and Prata et al. (2010), is the class of M-estimators, which corresponds to a generalization of the Maximum Likelihood Estimator. As described by Prata (2009), no other class of robust estimators has been used so far in the field of data rectification.

An estimator can be developed as the result of the Maximum Likelihood formulation, after assuming the validity of a certain statistical distribution for measurement errors, as described below.

2.1. Maximum likelihood formulation

The Maximum Likelihood formulation was introduced by Fisher in a series of articles between 1912 and 1922 (Aldrich et al., 1997), although some of the basic ideas had been previously discussed by Daniel Bernoulli (Kendall, 1961) and Gauss (Hald et al., 1999), and is the procedure used most often to obtain estimators for data regression problems in different environments (experimental, laboratory or industrial).

Initially, it is necessary to assume that a population of measured data with random errors, ε_i (white Gaussian noise) related to the population of sampled measurements, follow the probability function f_i (where $z_i = \hat{z}_i + \varepsilon_i$, z_i are the measured variables, and $\hat{z}_i$ are the estimated variables or reconciled variables). If the random errors ε_i are uncorrelated, measurements are independent and the respective covariance matrix, which contains the variance values of the measured data, becomes diagonal. In this case, the joint probability functions f of observed fluctuations can be calculated as the product of the individual probability functions f_i. Thus, the maximum likelihood estimate can be defined as the one that maximizes the probability of observing experimental data f, according to Eq. (1).

$$\max_{\hat{\underline{z}}} f = \max_{\hat{\underline{z}}} \prod_{i=1}^{n} f_i \tag{1}$$

However, in practice, it is more convenient to work with the minimization of the likelihood function (Reilly and Carpani, 1963), according to Eq. (2).

$$\max_{\hat{\underline{z}}} \prod_{i=1}^{n} f_i = -\min_{\hat{\underline{z}}} \prod_{i=1}^{n} f_i \tag{2}$$

Eq. (2) can be rewritten with help of the logarithm operator as presented in Eq. (3).

$$\min_{\hat{\underline{z}}} -\ln\left[\prod_{i=1}^{n} f_i\right] = \min_{\hat{\underline{z}}} -\sum_{i=1}^{n} \ln\left[f_i\right] \tag{3}$$

Eq. (3) can be used to build different data regression estimators, including the most popular ones, as shown below.

Normal → WLS:

$$\max_{\hat{z}} \prod_{i=1}^{n} \left\{ \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\} \tag{4}$$

$$\min_{\hat{z}} - \sum_{i=1}^{n} \ln \left\{ \exp\left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\} \tag{5}$$

$$\min_{\hat{z}} \sum_{i=1}^{n} \frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \tag{6}$$

$$\min_{\hat{z}} \sum_{i=1}^{n} \frac{1}{2} \xi_i^2 \tag{7}$$

The standardized residue ξ_i corresponds to the difference between the measured value z_i and the estimated (reconciled) value $\hat{z}_i$ weighted by the standard deviation σ_i. Optimal $\hat{z}_i$ values are those that maximize the Likelihood function.

The Normal distribution results in the WLS estimator, presented in Eq. (7), also known as an L_2 estimator (Rey, 1983). The WLS estimator can usually be regarded as a suitable estimator when measurements are free of spurious values.

Contaminated Normal:

$$\min_{\hat{z}} \sum_{i=1}^{n} - \ln \left[\frac{(1-p)\exp\left(-\frac{\xi_i^2}{2}\right)}{\sigma_i \sqrt{2\pi}} + \frac{p \exp\left(-\frac{\xi_i^2}{2}\right)}{b\sigma_i \sqrt{2\pi}} \right] \tag{8}$$

Eq. (8) presents the estimator based on the Contaminated Normal distribution, also known as the bivariate distribution (Tjoa and Biegler, 1991), where p is the probability of finding spurious values (with $0 < p \leq 0.5$), and $b^2\sigma_i^2$ is the variance of contamination by spurious values (with $b > 1$).

Laplacian →LAV:

$$\min_{\hat{z}} \sum_{i=1}^{n} |\xi_i| \tag{9}$$

The Laplacian distribution, also known as the double exponential distribution, results in the Least Absolute Value (LAV) estimator, presented in Eq. (9), also known as the L_1 estimator (Rey, 1983).

Cauchy:

$$\min_{\hat{z}} \sum_{i=1}^{n} \ln\left(1 + \xi_i^2\right) \tag{10}$$

Eq. (10) presents the estimator based on the Cauchy distribution.

Fig. 1 illustrates the probability distribution functions - $f(\xi)$ - for the functions presented in this section: Normal, Contaminated Normal, Laplacian, and Cauchy, respectively.

2.2. Generalization of the maximum likelihood objective function

Huber (1964) proposed a generalization of the maximum likelihood objective function. In 1973 (Huber et al., 1973), Huber extended the idea of using M-estimators for solution of regression problems through minimization of a smooth, symmetrical and reasonably monotonic function of residuals -ρ- over optimal estimates, as presented in Eq. (11).

$$\min_{\hat{z}} \sum_{i=1}^{n} \rho\left(\frac{z_i - \mu_i}{\sigma_i}\right) = \min_{\hat{z}} \sum_{i=1}^{n} \rho(\xi_i) \tag{11}$$

It is important to emphasize that M-estimators should not be independent of the dispersion measure, σ_i (scale parameter) while they attempt to estimate the "mean" of the sample set of the variable i, μ_i (location parameter), with help of some sort of iterative procedure (Zoubir et al., 2018). In order to estimate the means simply and robustly, the median (**med(z)**) is frequently used to represent the average. In Eq. (11) it is assumed that the measurement errors are independent and not correlated, with diagonal covariance-variance matrix.

The evaluation of the scale can be performed with traditional statistics (Feital et al., 2014) or robustly, using the M-estimators iteratively or the robust L-estimators MADn (Normalized Median Absolute Deviation) (Hampel, 1974; Zoubir et al., 2018), Tn (Rousseeuw and Croux, 1992) or Sn and Qn (Rousseeuw and Croux, 1993), with help of direct or indirect methods. Indirect methods make use of the process model during the estimation of standard deviations, based on the available data set and assuming that deviations from the reference model provide appropriate estimates for measurement errors. However, finding a global optimum for the Location and Scale of non-Gaussian Maximum Likelihood functions still remains an open problem (Zoubir et al., 2018). Fig. 2 illustrates the M-Estimators - $\rho(\xi)$ - for the previously analyzed functions: Normal, Contaminated Normal, Laplacian, and Cauchy, respectively.

It must be pointed out that the development of most M-estimators is not based on a well-defined and previously known probability distribution. According to Rey (1983), the development of most estimators is based on a convenient mathematical structure, such as the Fair estimator (Fair, 1974).

Some common properties of ρ are:

- ρ is continuous, to render the numerical manipulation easier, although this is not needed for successful use and numerical implementation of the estimator;
- ρ is symmetrical ($\rho(\xi) = \rho(-\xi)$), although this does not necessarily represent the real behavior of measurement errors, which can be distributed asymmetrically around the mean;
- ρ must be strictly positive ($\rho(\xi) \geq 0$) and integrable in order to allow the statistical interpretation of the problem;
- ρ should be an increasing monotonous function of ξ_i ($\rho(\xi_i) \geq \rho(\xi_j)$, for $|\xi_i| > |\xi_j|$), indicating that the probability of measuring large errors decrease with the increase of the magnitude of the error, although this is not needed for successful use and numerical implementation of the estimator;
- $\rho(0) = 0$; (It is recommended that the tuning constants satisfy this constraint, Hoaglin et al. (1983, p. 366))
- As a consequence of the previous properties, ρ should be preferentially convex, although this is not needed for successful use and implementation of the estimator.

The convexity of ρ guarantee uniqueness of the obtained solution (global optimal) for problems described by linear models (Huber, 1981; Rey, 1983), or nonlinear problems that can be appropriately described by linear constraints after linearization.

When ρ is pseudo-convex or quasi-convex (generally called "non-convex"), Huber (1981) suggests the implementation of an iterative procedure that initiates the optimization procedure with another convex ρ function and then switches conveniently to the desired ρ function. This is the procedure used most often by statisticians, although the increase of the computational power and advances in mathematics, including the development of deterministic and non-deterministic global search methods, are exerting enormous influence on the evolution of numerical techniques proposed for solution of data regression problems. Particularly, in most cases data regression performed with robust estimators can only be regarded as robust when global optimization methods are used. This point has been formally discussed by Baselga (2007) using non-deterministic simulated annealing and genetic algorithm procedures as numerical tools for minimization of robust objective

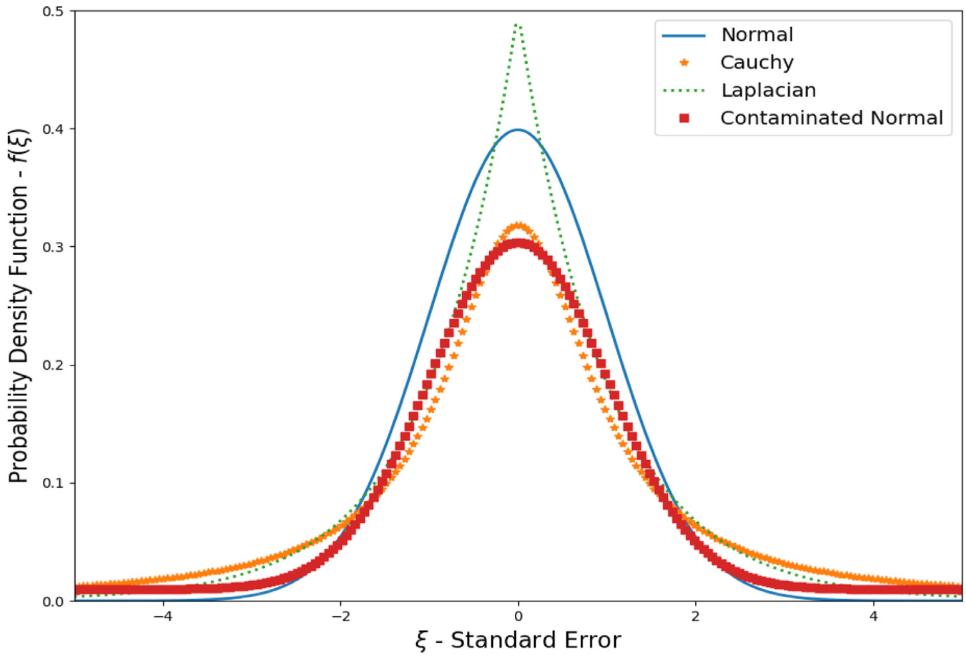


Fig. 1. Probability Distribution Functions $f(\xi)$ of some common estimators.

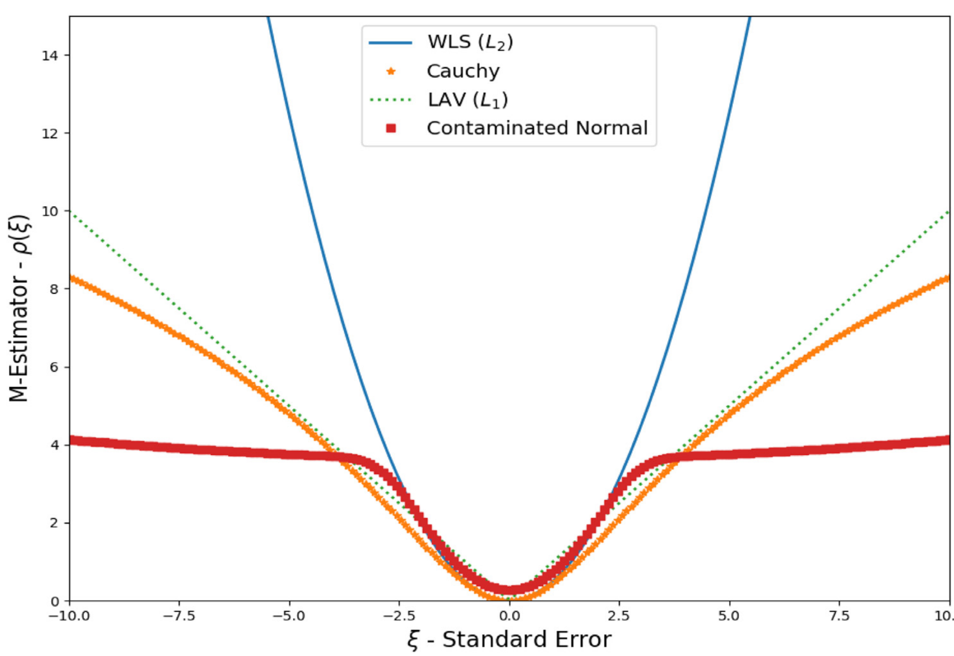


Fig. 2. Graphical representation of some estimators $\rho(\xi)$.

functions. For these reasons, as already said, the convexity of the M-estimator and the continuity of the derivatives are not mandatory, as pointed out and illustrated by Pennacchi (2008).

2.3. Influence function

The concept of influence function (or curve) was introduced by Hampel (1968) and constitutes an important measure of robustness, being generally classified as a measure of qualitative robustness (Hampel, 1971; 1974; Huber, 1981). For example, for an estimator to be considered robust, its influence function must be limited. In simple words, the influence function ψ corresponds to the weight (influence/impact) given to the effect of the magnitude of a spurious value (almost always measured in terms of multiples of the standardized residue ξ) on the obtained estimates. This function can be formally defined by Eq. (12), and is characterized by the first derivative of the objective function, ρ, in respect to the standardized residue, ξ.

$$\psi(\xi) = \frac{\partial \rho(\xi)}{\partial \xi} \tag{12}$$

The estimator WLS $[\frac{\xi^2}{2}]$ is not robust because its influence function is ξ, meaning that the influence of spurious values on the estimates is unlimited (lack of robustness) and increases proportionally with the increasing magnitude of the spurious value. On the other hand, the Welsch (Dennis and Welsch, 1978;

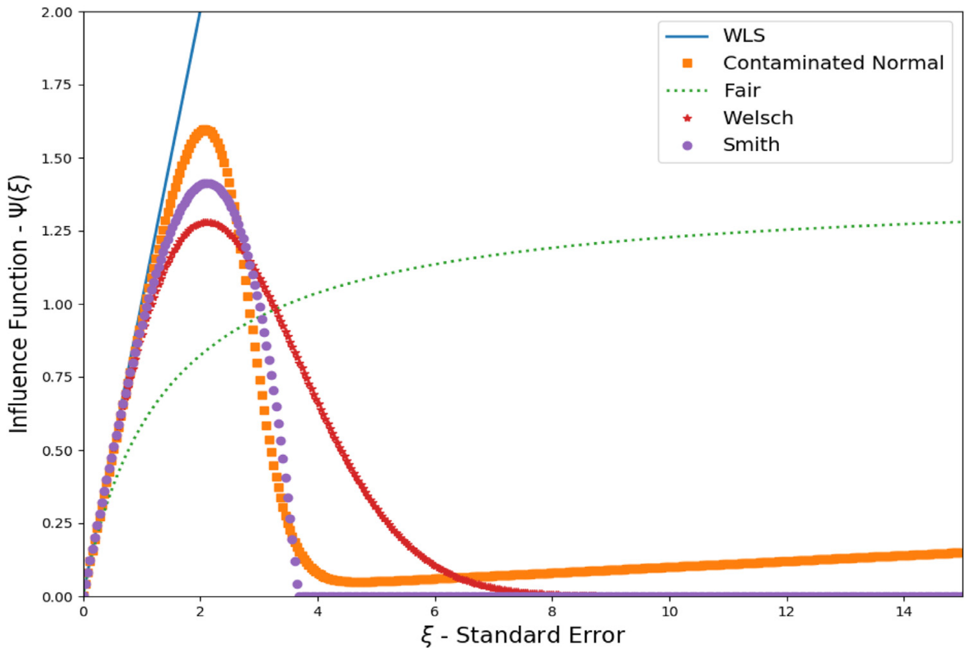


Fig. 3. Graphical representation of the influence functions of some analyzed estimators.

Rey, 1983) and Smith (Smith, 1888; Stigler, 1980) estimators present decreasing influence functions that can remove the negative effects of spurious values on estimates, even when the magnitudes of spurious values increase, explaining why these estimators are classified as **redescending**.

According to Holland and Welsch (1977), redescending estimators can be classified as **soft-redescending** (pseudo-convex) and **hard-redescending** (quasi-convex), depending whether the respective influence functions are approximately null and exactly null for spurious values of high magnitude. In the case of hard-redescending estimators, ψ is usually built with discontinuous segments and makes use of “if” clauses, having a finite point of rejection, such as the three-part Hampel estimator (Hampel, 1968).

The Fair estimator, on the other hand, can be influenced by spurious values, even if only to a limited extent, and is classified as a **monotone** estimator. However, the influence of spurious values in this case is much smaller than in the case of the WLS estimator. One must note that the ρ function of a monotonous estimator is convex. Rey (1983, p. 101) proposed the term “quasi-robust” to refer to estimators that are more robust than the traditional WLS estimator without being strictly robust.

Thus, based on the previous discussion, M-estimators can be classified in terms of the mathematical properties of the respective influence functions as:

- Non-robust (as the WLS estimator);
- quasi-Robust (as the Contaminated Normal estimator);
- Robust-Monotonous (as the Fair estimator);
- Robust-Soft-Redescending (as the Welsh estimator);
- Robust-Hard-Redescending (as the Smith estimator).

An illustrative graphical comparison of the mentioned influence functions is presented in Fig. 3.

It can be observed in Fig. 3 that the Contaminated Normal (quasi-robust) estimator has its influence function limited to approximately $\xi = 4$ (actually, $\xi = 4.742$); after that, the estimator becomes unlimited and behaves as a pseudo-robust estimator.

Based on the structure of the influence function, three other properties can be defined and discussed (Hampel et al., 1986):

- Gross-error sensitivity - γ - defined as the supreme (maximum) of the absolute value of the influence function. γ measures the maximum (negative) effect that a possible contamination on the measured value can exert on the estimator (Hoaglin et al., 1983, p. 358).
- Local-shift sensitivity - λ - defined as the derivative of the influence function at the analyzed point. λ measures the limit of influence on the estimator if a measured value deviates from its original value (leverage point) and evaluates the possible occurrence of numerical sensitivity and discontinuities in ψ.
- Rejection point - r_p - defined as the point where the influence function becomes exactly null. This is related to the complete nullification of the (negative) effect of the high magnitude spurious value on parameter and/or variable estimates. As explained previously, this can only be defined for hard-redescending estimators. In soft-redescending estimators, this value can be calculated as an approximation, if some sort of threshold value is defined for the influence function (Özyurt and Pike, 2004).

Fig. 4 presents an illustrative graphical explanation of these properties based on a hypothetical influence function.

Some common properties of ψ are (Hoaglin et al., 1983, p. 365):

- ψ is limited, which characterizes robustness (finite gross error sensitivity γ);
- ψ is continuous or piecewise continuous (finite local-shift sensitivity λ), to facilitate the implementation of numerical procedures, although this is not required for successful use and numerical implementation of robust estimators;
- ψ is an odd function ($\psi(-\xi) = -\psi(-\xi)$), which is related to the symmetrical nature of ro, although measurement errors are not necessarily symmetrical;
- ψ is approximately linear in the vicinities of the origin ($\psi(\xi) \approx k.\xi$, $k \neq 0$, for small ξ), which is related to the quadratic nature of $\rho(\xi)$ in the vicinities of the origin, although this property is not necessary for successful use and numerical implementation of robust estimators;

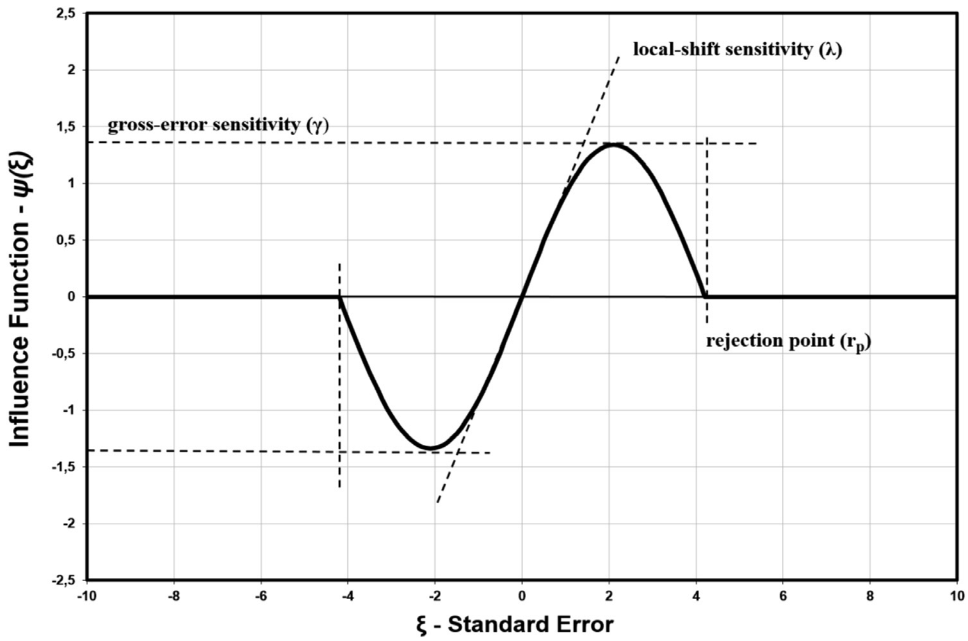


Fig. 4. Properties of an Influence Function (Hampel et al., 1986).

- the rejection point of ψ is finite (if resistance to large spurious values is desired).

2.4. First derivative of the influence function

Another very important function can be defined as the derivative of the influence function ψ in respect to the standardized residue ξ. This function corresponds to the second derivative of ρ in respect to the standardized residue ξ, as presented in Eq. (13).

$$\psi'(\xi) = \frac{\partial\psi(\xi)}{\partial\xi} = \frac{\partial^2\rho(\xi)}{\partial\xi^2} \tag{13}$$

Thus, ψ' corresponds to the Hessian matrix of $\rho(\xi)$. According to Hoaglin et al. (1983), ψ' has been rarely reported in the literature in the field of robust statistics, since the vast majority of analyzed regression problems have been solved with deterministic methods and making use of numerical approximations of the Hessian.

When ψ' is discontinuous, there is a lack of robustness concerning classical (deterministic) optimization procedures. To deal with possible discontinuities of ψ', it is possible to:

- Modify the discontinuous form of ψ' with smoothing functions using deterministic optimization procedures, as performed by Arora and Biegler (2001) for the Hampel's hard redescending estimator.
- Use non-deterministic optimization procedures.

2.5. Robustness and relative efficiency

The use of M-estimators in data regression problems usually involves the competitive relationship between efficiency and robustness (Albuquerque and Biegler, 1996), namely:

- (Relative) Efficiency: refers to the quality of the fit performed with an estimator when errors actually follow another distribution (reference), often assuming the Normal distribution as the appropriate reference.
- Robustness: refers to estimator performance over a variety of (nonnormal) error distributions.

This relationship can usually be manipulated through specification of tuning constants. Attention should be paid to the important statement by Albuquerque and Biegler (1996): "*The more robust an estimator is, the less efficient it is*".

It is particularly relevant to report the tunable parameter values when one intends to compare the performances of distinct robust estimators. As a matter of fact, comparisons must be performed for similar levels of relative efficiency and reference functions. If calculation of tunable parameter values results in lower relative efficiencies (that is, larger robustness), the analysis of removal of spurious values may become biased. This occurred recently with Xie et al. (2019) during tuning of the Xie estimator with relative efficiency of 95% in respect to the normal distribution. The value attributed to the tunable parameter was equal to 1.7134, resulting in relative efficiency of 91% instead of 95% (the correct value corresponds to 1.9597, as shown in Table 3). As a consequence, the Xie estimator presented the best overall performance, because it was more robust than the other analyzed estimators (tuned for relative efficiencies of 95%).

According to Prata (2009) and Prata et al. (2010), robust estimators are designed to perform very well when the distribution is contaminated and corrupted by spurious values, with a small loss of efficiency when the distribution is Normal and free of spurious values, which makes the WLS estimator the appropriate optimal estimator.

Recently, Valdetaro and Schirru (2011) proposed a method based on the Akaike criterion for tuning robust estimators based on the quality of the data set. This approach can be particularly useful in real-time applications of dynamic systems. As already said, the performance comparison of robust estimators should be based on the same relative efficiency, which sheds light on the importance of tuning methods (Rey, 1983; Özyurt and Pike, 2004).

Eq. (14) presents the mathematical definition of relative efficiency.

$$E_{ff}[\psi(\xi), f(\xi)] = \frac{V_f[\psi_f(\xi), f(\xi)]}{V[\psi(\xi), f(\xi)]} \tag{14}$$

where V_f is the asymptotic variance of the reference estimator and V is the asymptotic variance attributed to a particular M-estimator (Albuquerque and Biegler, 1996). On the other hand, M-estimators are consistent and asymptotically normally distributed, with asymptotic variance V defined in Eq. (15) (Huber, 1981, p. 103), (Hampel et al., 1986, p. 103), (Shevlyakov et al., 2008):

$$V[\psi(\xi), f(\xi)] = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[\int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi\right]^2} \quad (15)$$

where f is the true error distribution, ψ is the influence function and ψ' is the first derivative of the influence function.

Some desirable properties of ψ, f, ψ' and f' for calculation of relative efficiencies are (Shevlyakov et al., 2008), (Hampel et al., 1986, p. 126):

- f should be symmetrical and unimodal;
- f should be continuously differentiable and present a finite support $\{\xi : f(\xi) > 0\} = (-l, l)$ with $f(l) = f'(l) = 0$;
- For relative efficiency in respect to a reference distribution with density f, ψ should be proportional to $-[ln(f)]' = -\frac{f'}{f}$, i.e., $\psi \propto \psi_f$ (Hoaglin et al., 1983, p. 365);
- $\int_{-\infty}^{+\infty} \psi(\xi) f(\xi) d\xi = 0$ (i.e., ψ should be an odd function);
- $\int_{-\infty}^{+\infty} \psi^2(\xi) d\xi < \infty$ (i.e., ψ should be asymptotically or identically null when it tends to $\pm\infty$);
- $0 < \int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi = -\int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi < \infty$ (i.e., ψ' and f' should be limited).

As ψ' may be discontinuous, it may be convenient to rewrite Eq. (15) in the form of Eq. (16), making use of the previously defined desirable properties:

$$V[\psi(\xi), f(\xi)] = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[-\int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi\right]^2} = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[\int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi\right]^2} \quad (16)$$

2.6. Breakdown point

The Breakdown Point (BP) concept was introduced by Hampel (1968) and is another important measure of robustness, being used as a measure of quantitative robustness (Huber, 1981). The BP can be defined as the proportion of incorrect (measured) observations (spurious values) in a dataset that the robust regression technique can tolerate and still be used successfully. The higher the BP of an estimator, the more robust it is. Theoretically, the highest possible BP value is 0.5 (50%) (Rousseeuw and Leroy, 1987), because if more than half of the measurements are corrupted, it is not possible to distinguish between the true distribution and the distribution contaminated by spurious values. The BP of an M-estimator does not depend on its probability distribution (Rousseeuw and Leroy, 1987, p. 10), (Hoaglin et al., 1983, p. 370). Importantly, the WLS estimator has BP close to zero; therefore, a single spurious value causes significant deviations on the obtained estimates. This again reflects the extreme sensitivity of the WLS estimator to spurious values (Rousseeuw and Leroy, 1987). The BP of a very robust M-estimator is expected to be 0.5 (Huber, 1984), as these estimators can handle approximately 50% of spurious values in the data set. This has been asymptotically illustrated through simulation for the Biweight, Hampel, Andrews and Hyperbolic Tangent M-estimators (Zhang et al., 1998).

3. Application of robust M-estimators in data reconciliation

It is important to emphasize that the literature on Data Reconciliation (DR) and Multiple Gross Errors Detection (MGED) has usually recommended redescending M-estimators as the best options to reject the negative effects of spurious values (large deviations) on estimates of variables and parameters (Arora and Biegler, 2001; Özyurt and Pike, 2004), as these M-estimators showed better results in comparative studies. However, the use of robust M-estimators for DR and MGED is still being consolidated, as these estimators have been increasingly employed by researchers. Applications include analyses of very distinct problems, ranging from evaluation of linear models used to fit simulated data sets at steady state conditions to evaluation of highly nonlinear model constraints used to fit real dynamic industrial data in real time, for purposes of process monitoring and real time optimization. Table 1 summarizes some works that used robust M-estimators to solve DR and MGED problems.

4. The 48 robust M-estimators used in data regression problems

Many M-estimators have been proposed in the robust statistics literature and some of these have been used in data regression problems to minimize the negative influence of less frequent spurious values on estimates of variables, parameters or both together. Table 2 shows the analyzed robust estimators divided into 3 categories, with the respective reference numbers used in the present work.

It is important to emphasize that the present work displays the most complete set of robust estimators described in the literature and that no articles or books already published in the literature review more than 10 examples of robust M-estimators. Moreover, in the vast majority of the previously published articles, M-estimators are presented only in terms of the function ρ, or the function ψ, although frequently in terms of the **weight function** w, where $w(\xi) = \psi(\xi)/\xi$. Additionally, presentation of tunable parameters for different levels of relative efficiency is rare, as parameters are usually tuned to 95% efficiency in respect to the Normal distribution. Moreover, some estimators have seemingly never been tuned at all and available tuned parameter values are scattered in the literature. The presentation of illustrative plots for the ρ, ψ, and ψ' functions of the M-estimators has also been rare in the literature, making more difficult the numerical validation of M-estimators and respective tunable parameters.

Based on the previous remarks, the present review aims to facilitate the access to the M-estimator families of functions (ρ, ψ and ψ'), respective graphs and associated tuning parameters. Another objective of the present work is to motivate the use of some M-estimators that have been reportedly and successfully used in few areas of knowledge to solve data regression problems. Thus, we believe that the present review can facilitate the exchange of knowledge between the areas of statistics, geodesy, meteorology, engineering, economics, image processing, astronomy, among others. Additionally, this article aims to fill this gap by presenting and discussing the properties and use of 50 different M-Estimators (WLS, Contaminated Normal and 48 robust estimators).

Besides the estimators shown in Table 2, the 47-Barron and 48-GT (Generalized t-distribuition) estimators, which belong to the category of **General and Adaptive Robust Estimators**, are also considered (Barron, 2019; McDonald and Newey, 1988). The properties of these two adaptive estimators depend on certain tunable parameters so that the estimator obtains characteristics of monotonous, soft-redescending or hard-redescending estimators. These estimators are not subject to formal efficiency tuning, because the variation of the parameters can completely modify the characteristics of the estimator, instead of the simpler asymptotic adjustment to the variance of the reference distribution.

Tables C.5 and C.6 in Appendix C present the ρ functions (estimator) and ψ functions (Influence Function) as functions of the standardized residue ξ and their tuning constants, for all estimators mentioned previously. For better visualization of these functions, the following figures were elaborated:

Table 1
Examples of robust estimators applied in the DR literature.

Reference	System*	M-Estimators	Process(es)
Tjoa and Biegler (1991)	S-SS	Contaminated Normal	Heat exchanger
Johnston and Kramer (1995)	S-SS	Contaminated Normal and Lorentzian	Heat exchanger and flow network
Zhang et al. (1995)	R-SS	Contaminated Normal	Sulfuric acid plant
Albuquerque and Biegler (1996)	S-D	Contaminated Normal and Fair	Heat exchanger, tanks and hydrolysis
Chen et al. (1998)	R-SS	Fair and Lorentzian	Chemical reactor
Bourouis et al. (1998)	R-SS	Contaminated Normal	Multistage flash desalination plants
Mingfang et al. (2000)	S-SS	Kong	Adiabatic CSTR
Arora and Biegler (2001)	S-SS/D	Fair and Hampel	Steam metering and tanks
Özyurt and Pike (2004)	S/R-SS	Contaminated Normal, Cauchy, Fair, Hampel, Logistic e Lorentzian	Chemical reactor, steam metering, metallurgical grinding, heat-exchanger, sulfuric acid and alkylation
Ragot et al. (2005)	S-SS	Contaminated Normal	Component material balances
Wongrat et al. (2005)	S-SS	Hampel	Pai and Fisher (1988) problem
Lingke et al. (2006)	S-SS	Huber and Kong	Pai and Fisher (1988) problem
Faber et al. (2006, 2007)	R-SS	Kong	Coke-oven-gas purification
Schladt and Hu (2007)	R-SS	Contaminated Normal	Chemical reaction and distillation column
Alhaj-Dibo et al. (2008)	S-SS	Contaminated Normal	Mineral processing plant
Lid and Skogestad (2008a,b)	S/R-SS	Contaminated Normal	Catalytic reformer, pipe model and flash
Prata et al. (2008b)	S-D	Contaminated Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian e Welsch	Liebman et al. (1992) reactor (CSTR)
Prata et al. (2010)	R-D	Welsch	Industrial polypropylene reactor
Zhang et al. (2010)	S-SS	Zhang	Atmospheric tower, ethylene separation, air separation and steam metering
Valdetaro and Schirru (2011)	S/R-SS	Hampel	Thermal reactor power
Jin et al. (2012)	S-SS	Jin, Cauchy and Huber	Measurement network and Pai and Fisher (1988) problem
Chen et al. (2013)	S-SS	Zhang, Fair, Hampel and Correntropy	Atmospheric distillation tower and steam metering
Zhang et al. (2014)	R-SS	Correntropy	Air separation
Zhang and Chen (2015)	S-D	Correntropy	Polymerization of styrene
Llanos et al. (2015)	S-SS	Zhang, Biweight, Welsch and Correntropy	Steam metering and Pai and Fisher (1988) problem
Korpela et al. (2016)	S/R-SS	Welsch	Multi-fuel fired industrial boilers
Coimbra et al. (2017)	R-D	Welsch	Polymerizations of methyl methacrylate
da Cunha et al. (2017)	S-SS	Contaminated Normal, Bell, Huber, Lorentzian, Hampel, Welsch, Andrews, Smith, Jin, Biweight and Correntropy	Van de Vusse (1964) reactor
Wu et al. (2017)	S-SS	Wu	Crude oil distillation set
Llanos et al. (2017)	S-SS	Biweight, Huber, Hampel, Welsch and Correntropy	Steam metering and heat exchanger network
Valluru et al. (2018)	S/R-D	Fair and Hampel	Williams and Otto (1960) reactor and reactive distillation
do Valle et al. (2018)	S-SS/D	Contaminated Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian and Zhang (quasi-Weighted)	Collection of benchmark test problems
Liu et al. (2018)	R-D	Fair, Logistic, Lorentzian and Welsch	Pharmaceutical manufacturing
Su et al. (2019)	R-D	Welsch	Pharmaceutical manufacturing
Xie et al. (2019)	S/R-SS	Fair, Welsch, Cauchy e Xie	Measurement network and industrial evaporation
da Cunha et al. (2020)	S-SS	16 robust M-estimators	Chemical reactors problems

***S** - simulated data; **R** - real data; **SS** - steady state e **D** - dynamic.

Table 2
Analyzed Robust M-Estimators.

#- Monotone	#- Soft-redescending	#- Hard-redescending
1- LAV L_1	11- Alamgir	31- Asad
2- $L_1 - L_2$ (Charbonnier)	12- Bab-Hadiashar	32- Andrews
3- Fair	13- Bell	33- Biweight (Tukey-biweight)
4- Huber	14- Blake-Zisserman	34- Collins
5- Modified Huber	15- Cauchy (Lorentz)	35- Hampel
6- Kong	16- Correntropy	36- Hiperbolic Tangent
7- LnCosh	17- Danish	37- LQQ (Linear Quadratic Quadratic)
8- Logistic	18- Geman-McClure	38- Modified Asad-Qadir I
9- Müller	19- GGW (Generalized Gauss Weight)	39- Modified Asad-Qadir II
10- Zhang	20- Insha	40- Optimal
	21- Jin	41- Qadir
	22- Kumar	42- Smith
	23- Lorentzian	43- Talwar (Huber type-skipped mean)
	24- Merril-Schweppe (BDS)	44- Uk
	25- Ramsay	45- Yang I
	26- Sech	46- Yang II
	27- Welsch	
	28- Wu	
	29- Xie	
	30- Youssef	

Table 3
Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution.

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
0	Contaminated Normal	$b_{CN} = 10$; $p_{CN} = 0.235$[1]; $p_{CN} = 0.26449$[2]	Özyurt and Pike (2004)[1]; This work[2]	Jeffreys (1932)
1	LAV (L_1)	No tuning parameters. $E_{ff} = \pi/2 \approx 64\%$		de Laplace (1774)
2	$L_1 - L_2$	No tuning parameters. $E_{ff} = 95.75\%$		Charbonnier et al. (1997)
3	Fair	$c_F = 1.3998$	Rey (1983); Holland and Welsch (1977); Özyurt and Pike (2004)	Fair (1974)
4	Huber	$c_{Hu} = 1.345$	Rey (1983); Holland and Welsch (1977)	Huber (1964)
5	Modified Huber	$c_{MH} = 1.2107$	Rey (1983); Pennacchi (2008)	Rey (1983)
6	Kong	$c_k = 0.2$[3]; $c_k = 2.4541$[4]	Kong et al. (2000)[3]; This work[4].	Kong et al. (2000)
7	LnCosh	$c_{LC} = 0.83$	This work	Karal (2017)
8	Logistic	$c_L = 1.205$	Holland and Welsch (1977)	Verhulst (1838)
9	Müller	$c_{Mu} = 2.28302$	This work	Vichare (1993)
10	Zhang	$c_Z = 0.814$	This work	Zhang et al. (2010)
Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 2/5				
11	Alamgir	$c_{Al} = 2.37111$	This work	Alamgir et al. (2013)
12	Bab-Hadiashar	$c_{BH} = 1.81206$	Bab-Hadiashar et al. (2002)	Bab-Hadiashar et al. (2002)
13	Bell	$c_{Bl} = 2.1522$	This work	Bell (1980)
14	Blake-Zisserman	$c_{BZ} = 0.00075$	This work	Hartley and Zisserman (2003)
15	Cauchy	$c_C = 2.3849$	Rey (1983), Özyurt and Pike (2004)	Poisson (1824) (Apud (Stigler, 1974))
16	Correntropy	$c_{Co} = 2.1105$	This work	Liu et al. (2006)
17	Danish	$c_{Da} = 2.76705$; $D_{Da} = 3.91523$	This work	Berberan (1995)
18	Geman-McClure	$c_{GM} = 3.787376$	This work	Geman and McClure (1987)
19	GGW	$a_{GGW} = 1.38636$; $b_{GGW} = 1.50$; $c_{GGW} = 1.06282$	Koller and Stahel (2011)	Koller and Stahel (2011)
20	Insha	$c_I = 3.2296$	This work	Ullah et al. (2006)
21	Jin	$a_J = 0.65$; $c_J = 3.77312$	This work	Jin et al. (2012)
Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 3/5				
22	Kumar	$a_{Ku} = 1.3860$; $b_{Ku} = 2.7720$; $c_{Ku} = 4.80284$	This work	Kumar and Rao (2009)
23	Lorentzian	$c_{Lz} = 2.6$[5]; $c_{Lz} = 2.678$[6]	Özyurt and Pike (2004)[5]; This work[6]	Johnston and Kramer (1995)
24	Merril-Schweppe	$c_{MS} = 1.637$	This work	Merrill and Schweppe (1971)
25	Ramsay	$c_R = 0.3569$	This work	Ramsay (1977)
26	Sech	$c_{Sh} = 0.40497$	This work	Tabatabai et al. (2014)
27	Welsch	$c_W = 2.9846$	Rey (1983), Holland and Welsch (1977)	Dennis and Welsch (1978)
28	Wu	$a_{Wu} = 0.5$; $c_{Wu} = 2.521$	This work	Wu et al. (2017)
29	Xie	$c_X = 1.7134$[7]; $c_X = 1.9598$[8]	Xie et al. (2019)[7]; This work[8]	Xie et al. (2019)
30	Youssef	$c_Y = 2.6006$	This work	Youssef et al. (2013)
31	Asad	$c_{As} = 3.6175$	This work	Ali and Qadir (2005)
32	Andrews	$c_{An} = 1.338$	Rey (1983), Holland and Welsch (1977)	Andrews et al. (1972)
33	Biweight	$c_{Bi} = 4.6851$	Rey (1983), Pennacchi (2008)	Beaton and Tukey (1974)
Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 4/5				
34	Collins	$a_{Cl} = 1.6515$; $r_{Cl} = 4.0$; $q_{Cl} = 1.5908$; $D_{Cl} = 3.8948$	This work	Collins (1976)
35	Hampel	$a_H = 1.35$; $b_H = 2.70$; $c_H = 5.40$	Özyurt and Pike (2004)	Andrews et al. (1972)
36	Hyperbolic Tangent	$A = 0.7261$; $B = 0.8234$; $k = 4.5$; $r = 4.0$; $q = 1.5562$; $D = 3.9053$	This work	Hampel et al. (1981)
37	LQQ	$b = 1.4734$; $c = 0.9823$; $s = 1.5$; $a = 5.4025$; $D_1 = -2.5802$; $D_2 = 4.9047$	Koller and Stahel (2011)	Koller and Stahel (2011)
38	Modified Asad-Qadir I	$a_{AQI} = 3.3094$	This work	Ali and Qadir (2005)
Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 5/5				
39	Modified Asad-Qadir II	$a_{AQII} = 3.1665$	This work	Ali and Qadir (2005)

(continued on next page)

Table 3 (*continued*)

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
40	Optimal*	$\begin{cases} a_{Op} = 0.3225 \\ b_{Op} = 3.0263 \\ c_{Op} = 8.2006e^{-4} \\ d_{Op} = 5.0349e^{-3} \\ e_{Op} = 1.8418e^{-3} \end{cases}$	This work	Yohai and Zamar (1997)
41	Qadir	$a_Q = 4.6851$	This work	Ali et al. (2005)
42	Smith	$a_S = 3.6732$	This work	Smith (1888)
43	Talwar	$a_T = 2.7955$	Holland and Welsch (1977)	Hinich and Talwar (1975)
44	Uk	$a_{Uk} = 4.1323$	This work	Khalil (2012)
45	Yang I	$\begin{cases} c_{YaI} = 2.3708 \\ a_{YaI} = 3.5 \\ D_{YaI} = 3.7026 \end{cases}$	This work	Yang (1999)
46	Yang II	$\begin{cases} c_{YaII} = 1.3884 \\ a_{YaII} = 3.5 \end{cases}$	This work	(Yang, 1994)
47	Barron	Tuning not applied in this case. Shape parameters in Fig. 14		Barron (2019)
48	GT	Tuning not applied in this case. Shape parameters in Fig. 14		McDonald and Newey (1988)

E_{ff}: [1]=95.53%; [2]=95.00%; [3]=74.03%; [4]=95.00%; [5]=94.58%; [6]=95.00%; [7]=90.97%; [8]=95.00%; ∗ Polynomial approximation..

Table 4
Summary of properties and characteristics of analyzed robust M-estimators.

	Monotonous		Soft-redescending		Hard-redescending
	Without "if" clause	With "if" clause	Without "if" clause	With "if" clause	Always with "if" clause
Estimator (#)	#1, 2, 3, 6, 7, 8, 10	#4, 5, 9	#11, 13, 14, 15, 16, 18, 20, 23, 25, 26, 27, 29	#12, 17, 19, 21, 22, 24, 28, 30	#31 - 46
Function (ρ)	Convex		pseudo-convex (except #19 does not exist ρ form; it was built from ψ function)		quasi-convex
IF (ψ) - gross errors influence	moderately increases and asymptote to a constant value (except #1)		asymptote to zero		is exactly zero
Rejection Point (r_P) in ψ	No		Yes (can be approximated		Yes (exactly defined by the last parameter)
ψ discontinuous (Figs. 5–12)	#1	#9	Continuous	#17, 28, 30	#43, 46
ψ' discontinuous (Figs. 5–12)	#1	#4, 9	Continuous	#12, 17, 21, 22, 24, 28, 30	#32, 34, 35, 36, 42, 43, 45, 46
Tuning parameters	Just one, Except #1, 2	Just one	Just one	One, two (#17, 21, 28), three (#22)	One, two (#46), three (#35), four (#34), five(#40), six(#36, 37)
Difficulty of tuning	Easy	Easy, Moderate (#9)	Easy/Difficult (#19)	Easy, Moderate (#17, 21, 28, 30), Difficult (#22)	Easy, Moderate (#43, 35, 46), Difficult (#34), Hard (#40, 36, 37)

- Figs. 5–12 illustrate graphically the families of ρ, ψ and ψ' functions of all analyzed M-estimators;
- Fig. 13 illustrates the ρ and ψ functions of the Barron (General and Adaptive Loss Function) and GT (Generalized t-distribution) estimators.

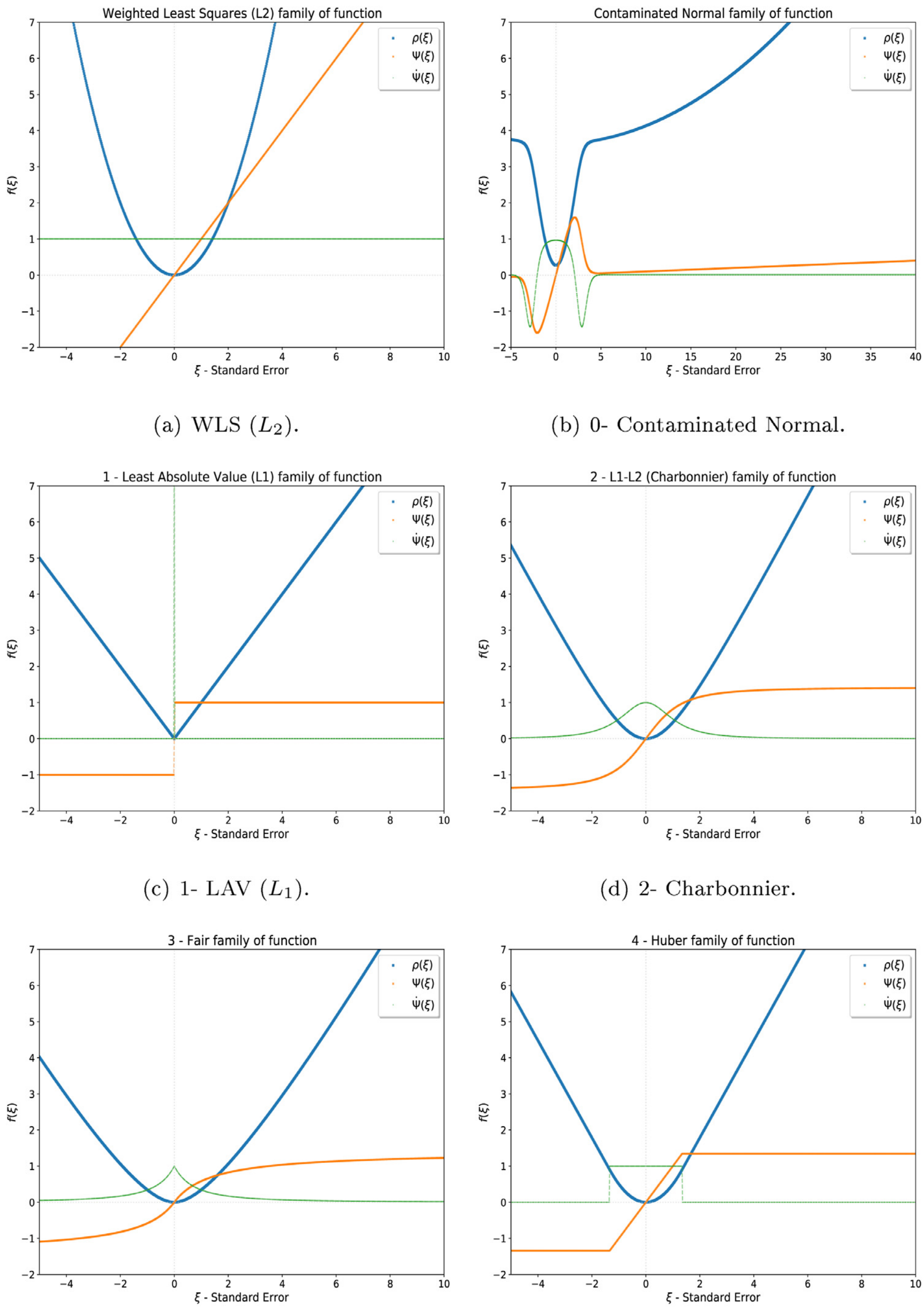
Properties of the analyzed monotonous and redescending robust M-estimators discussed previously and based on Figs 3 and 4 remain valid for all estimators represented in the present section. A detailed summary of the main characteristics of each analyzed M-estimator is provided in Table 4. Except for the WLS, Charbonnier, LAV, Barron and GT M-estimators, all figures were built for Relative Efficiency of 95% in respect to the Normal distribution, while the respective tuning constants are presented in Section 5.

5. The tuning parameters of the M-estimators

Table 3 presents the values of the tuning parameters for the analyzed M-estimators for a given relative efficiency in respect to the Normal distribution. The references for these parameters and the introductory reference for each analyzed M-estimator are also presented.

Looking at Figs. 13 and 14, we can understand why these estimators are classified as General and Adaptive. The Barron estimator can take the shape of 5 different estimators, ranging from the non-robust shape of the WLS to the soft-redescending Welsch, through the monotone Charbonnier ($L_1 - L_2$). The GT estimator can be adapted into 3 different estimators: LAV, WLS, and Cauchy.

As noted in Table 3, many of the M-estimators have two or more tuning parameters, so it is necessary to know the relationship among them for appropriate tuning. It is important to remember that the desired properties presented by the ρ, ψ and ψ' functions are fundamental for correct tuning. An estimator that is not well aligned with the desired properties may be tuned incorrectly, so that the desired relationship among some of the parameters must be known and kept constant for correct tuning. The desired properties of the functions that characterize the analyzed M-estimators, such as the Gross-Error Sensitivity (γ) and the Rejection Point (r_p) (Fig. 4), are fundamental for definition of the geometric properties of the considered function and calculation of its Relative Efficiency. Therefore, these properties may depend on one or more parameters, particularly in the case of hard-redescending M-estimators. In the present work, the Maple software version 15 was used for tuning the estimators. An illustrative example of the tuning procedure is detailed in Appendix A (analytic procedure for LAV estimator) and Appendix B (symbolic algebraic procedure), illustrating the tuning of the Fair, Welsch, and Hampel estimators. For M-estimators with two or more parameters, the tuning was performed according to:



(a) WLS (L_2).

(b) 0- Contaminated Normal.

(c) 1- LAV (L_1).

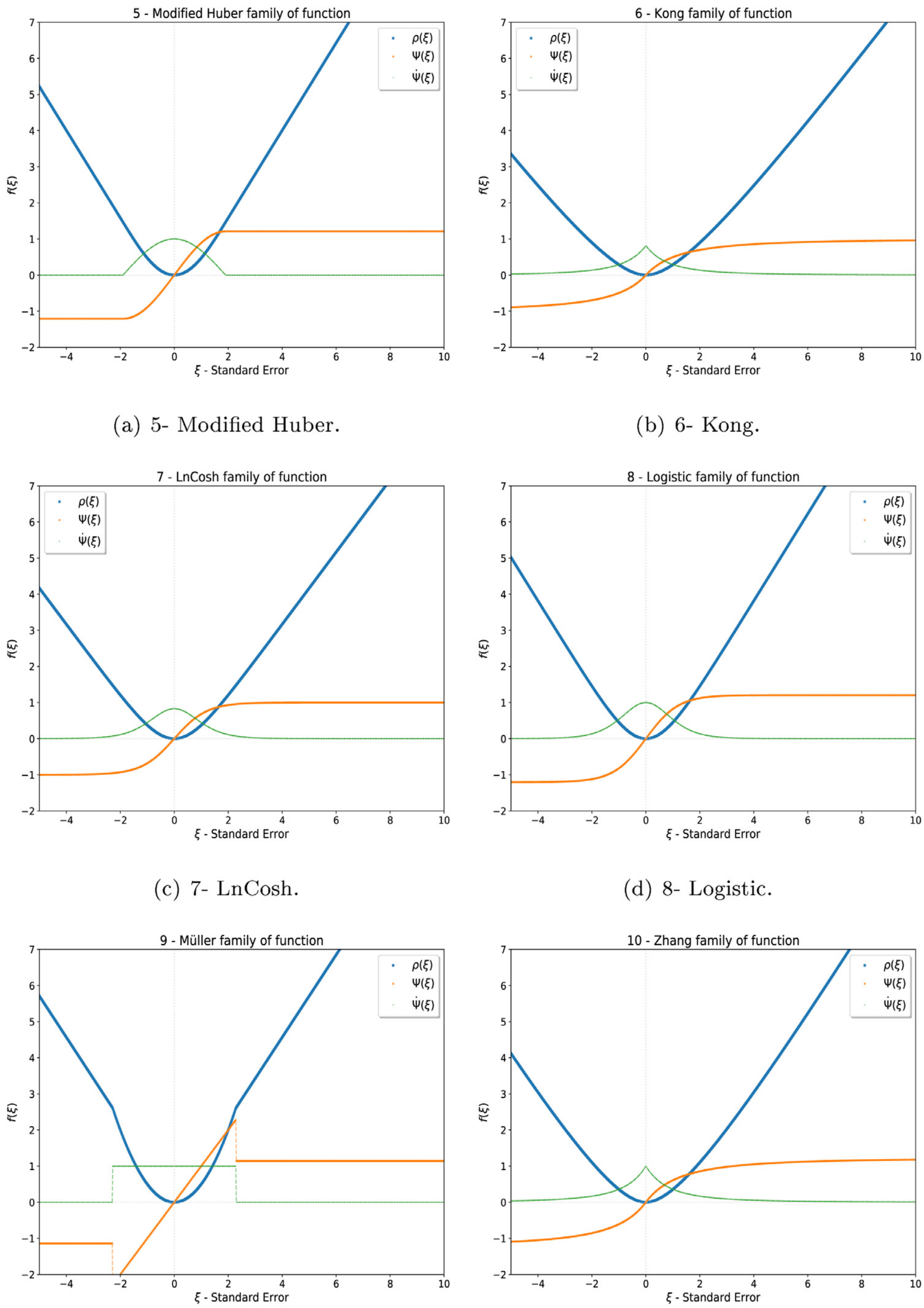
(d) 2- Charbonnier.

(e) 3- Fair.

(f) 4- Huber.

Fig. 5. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 1/8.

- Normal Contaminated: The constant b_{CN} represents the contamination of the variance by spurious values. The p_{CN} constant represents the probability of contaminated distribution. They have no fixed relationship with each other. Therefore, the estimator can be tuned by fixing one of the constants, in the case of this work the parameter $b_{CN} = 10$ was fixed.
- Danish: The D_{Da} constant is an auxiliary constant, and depends on the tuning constant c_{Da}. The estimators that comprise two or more functions must be adjusted in order to remove any discontinuity, according to the desired properties. For that, it is necessary to add auxiliary constants to correct these discontinuities. For this estimator, the auxiliary constant follows the following relationship: $D_{Da} = \rho_1(c_{Da}) - \rho_2(c_{Da})$.
- GGW: This estimator has 3 tuning parameters, but no definition has been found for the relationships among these parameters. Thus, 2 parameters were selected for tuning. The c_{GGW} parameter is the parameter that limits two parts of the characteristic estimator functions and was manipulated to keep the shape of the function approximately the same. The parameter $b_{GGW} = 1.5$ is fixed and the parameter a_{GGW} is used for tuning. It is important to remember that this estimator does not



(a) 5- Modified Huber. (b) 6- Kong.

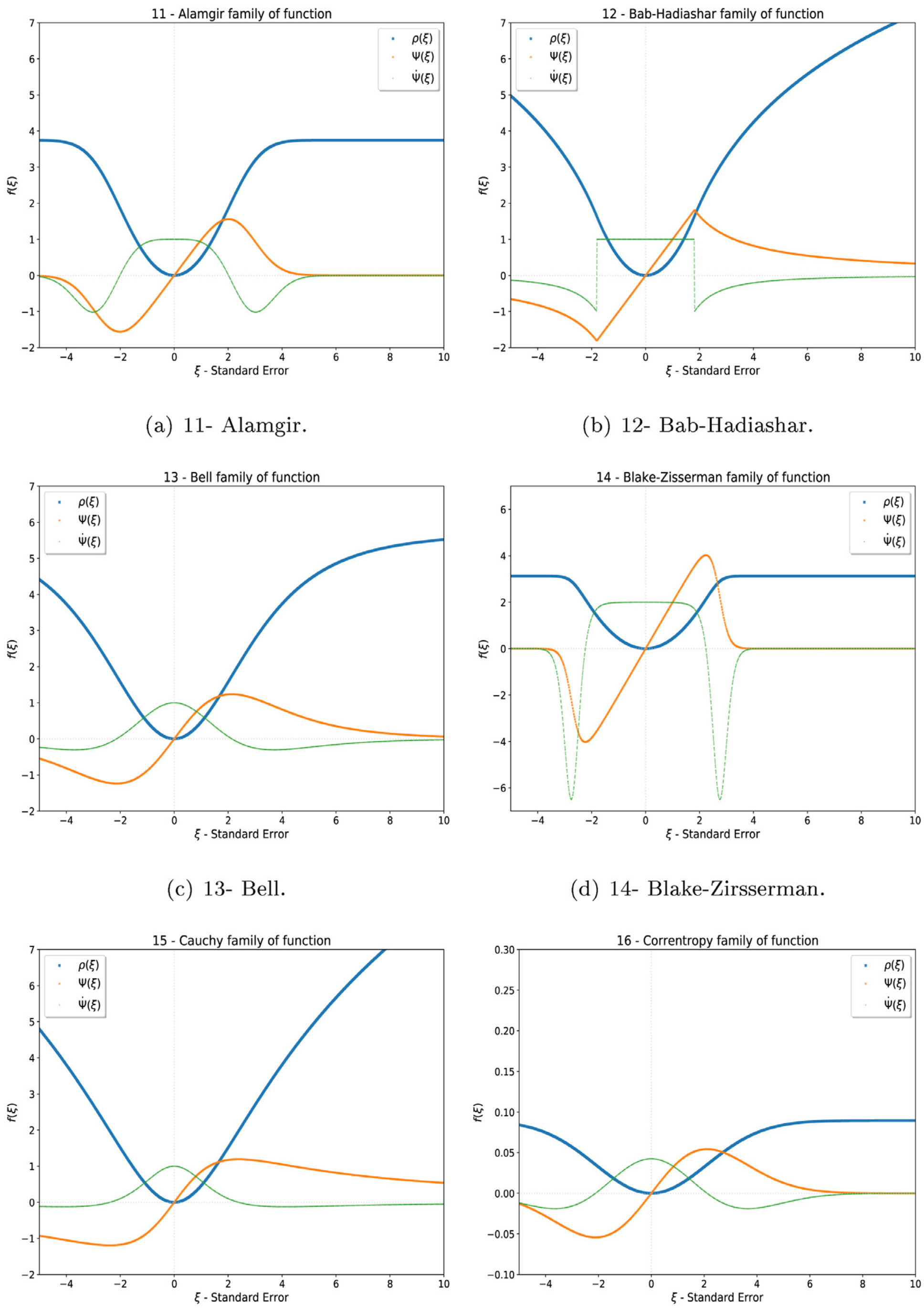
(c) 7- LnCosh. (d) 8- Logistic.

(e) 9- Müller. (f) 10- Zhang.

Fig. 6. Family of function (ρ, ψ, ψ') part 2/8.

present an analytical integral for the second part of the Influence Function (ψ_2). The graph of the ρ function of this estimator was built numerically.

- Jin: The parameter a_J was fixed at $a_J = 0.65$ as recommended by the authors of the original work. Therefore, the parameter c_J was used to tune the estimator.
- Kumar: This estimator presents 3 parameters and 3 parts of functions. One of them is the auxiliary parameter to adjust the desired properties, the parameter $D_{Ku} = \rho_2(b_{Ku}) - \rho_3(b_{Ku})$, which is a function of the parameters a_{Ku} and b_{Ku}. The a_{Ku} parameter is related to the b_{Ku} parameter in order to keep the geometrical properties of the function, defined in the present work in the form $a_{Ku} = b_{Ku}/2$.
- Wu: This estimator presents 2 tuning parameters, the parameters a_{Wu} and c_{Wu}. $0 < a_{Wu} < 1$ according to the reference work. Therefore, $a_{Wu} = 0.5$ was set and the parameter c_{Wu} was used for tuning.
- Collins: This estimator presents 4 parameters and 3 parts of functions. The D_{Cl} parameter is the auxiliary parameter, with $D_{Cl} = \rho_1(q_{Cl}) - \rho_2(q_{Cl})$. The q_{Cl} parameter represents the gross-error sensitivity, in addition to adjusting the continuity of the estimator. The calculation of this parameter constitutes



(a) 11- Alamgir. (b) 12- Bab-Hadiashar.

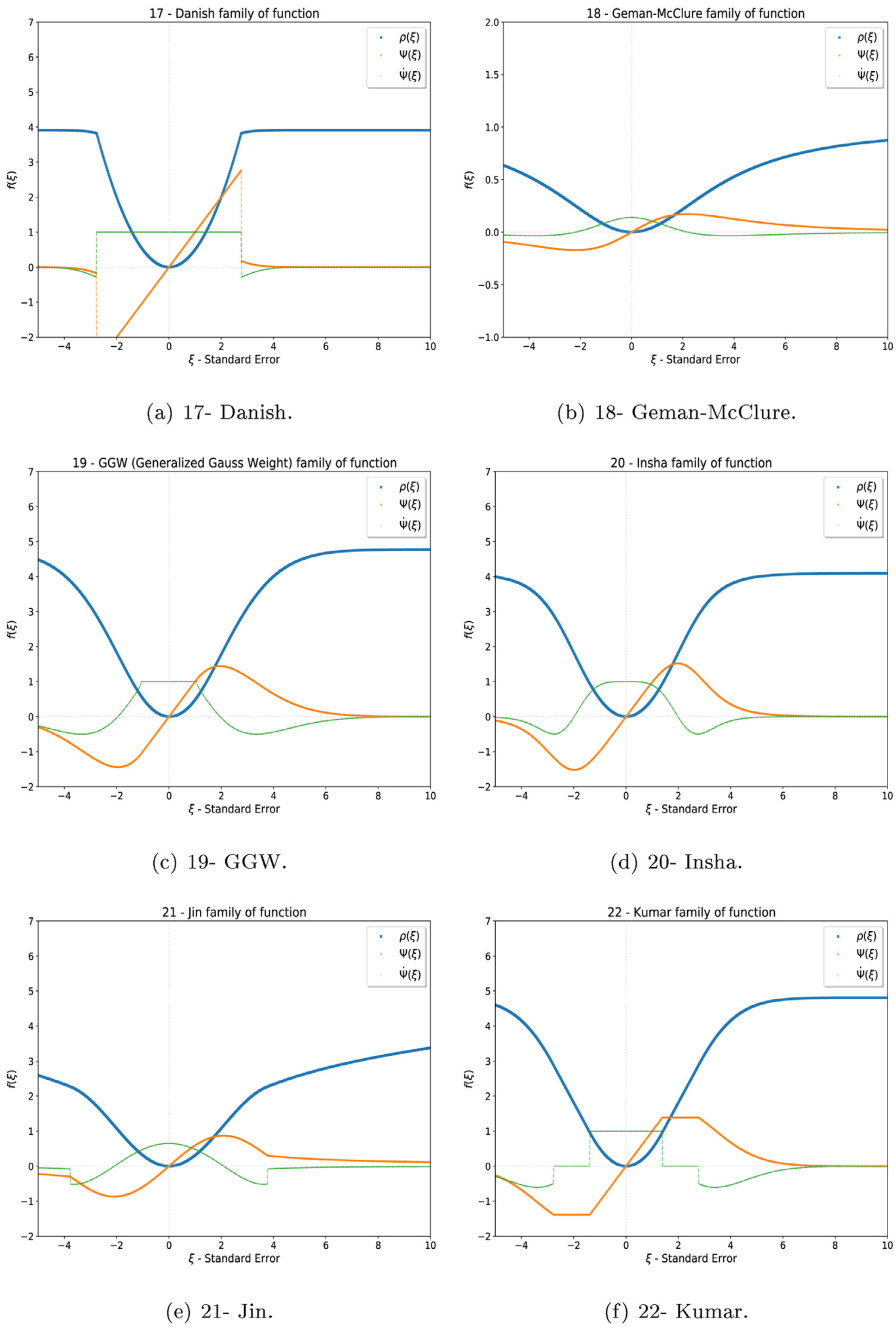
(c) 13- Bell. (d) 14- Blake-Zirsserman.

(e) 15- Cauchy. (f) 16- Correntropy.

Fig. 7. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 3/8.

an intermediate step within the tuning procedure, where q_{Cl} is the positive root of the equation ($\psi_2 - \psi_1 = 0$), ψ_i being the ith part of influence function. The r_{Cl} parameter (which represents the Rejection Point (r_p) of the estimator) must follow the geometric proportionality, varying according to the Relative Efficiency and fixed for tuning. Thus, for this work the tuning parameter was a_{Cl}.

- Hampel: This estimator presents 3 parameters and 4 parts of functions. The 3 parameters (a_H, b_H, c_H) have a geometric relationship widely used in the literature, namely: $c_H = 4a_H$ and $b_H = 2a_H$. The parameters must follow the relation $a_H < b_H < c_H$ to satisfy the geometric proportionality, therefore the relation most used in the literature and in this work is adequate. Thus, for this work the tuning parameter used was a_H.
- Hyperbolic Tangent: This estimator presents 6 parameters and 3 parts of functions. The D parameter is the auxiliary parameter, with $D = \rho_1(q) - \rho_2(q)$. The q parameter represents the gross-error sensitivity, in addition to adjusting the continuity of the estimator. The calculation of this parameter constitutes an intermediate step within the tuning procedure, where q is the positive root of the equation ($\psi_2 - \psi_1 = 0$), ψ_i being the ith part of influence function. The parameter r (which



(a) 17- Danish. (b) 18- Geman-McClure.

(c) 19- GGW. (d) 20- Insha.

(e) 21- Jin. (f) 22- Kumar.

Fig. 8. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 4/8.

represents the point of rejection of the estimator) must follow the geometric proportionality, varying according to the Relative Efficiency and fixed for a particular tuning. However, parameter r is related to parameter k, where $r \leq k$. Parameters A and B must obey the inequalities $0 < A < B < 1$. Thus, fixing k and B, and respecting the proposed relationships, only parameter A is left for tuning.

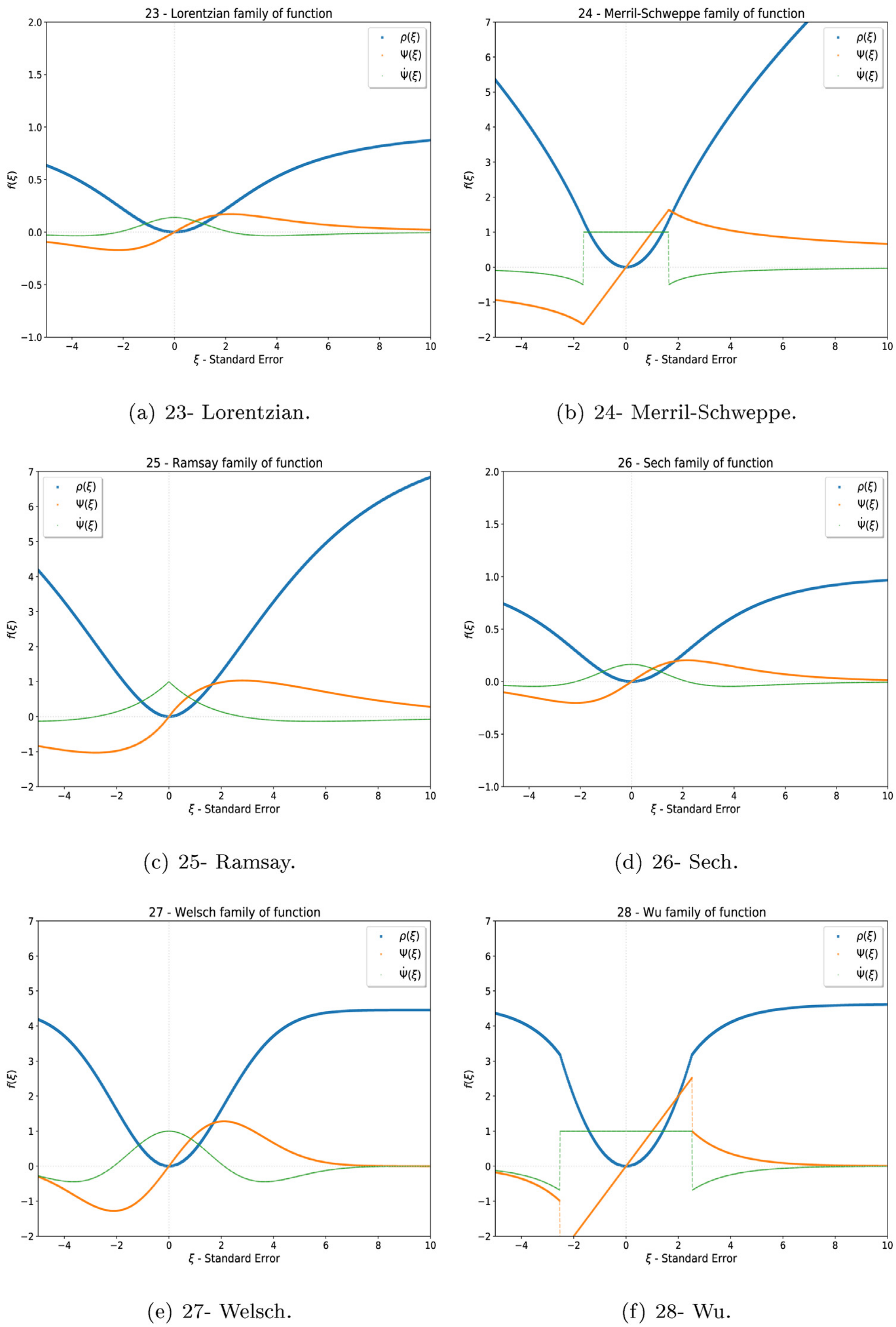
- LQQ: This estimator has 6 parameters and 4 parts of functions. The 3 auxiliary parameters are a, D_1 and D_2, being:

$$a = \frac{bs - 2b - 2c}{1 - s} \tag{17}$$

$$D_1 = \frac{(-b^3 + (-a - 3c)b^2 - 3c(a + c)b - 3ac^2 - c^3)s + (c + b)^3}{6a} \tag{18}$$

$$D_2 = -\frac{1}{3}a^2 s - \frac{1}{2}abs + ac + bc - \frac{1}{6}b^2 s + \frac{1}{2}c^2 + \frac{1}{3}a^2 + ab + \frac{1}{2}b^2 \tag{19}$$

The parameter s is fixed at $s = 1.5$, where the geometric relationship of the parameter is $s = 1 - \inf(\psi'(\xi))$.



(a) 23- Lorentzian.

(b) 24- Merril-Schweppe.

(c) 25- Ramsay.

(d) 26- Sech.

(e) 27- Welsch.

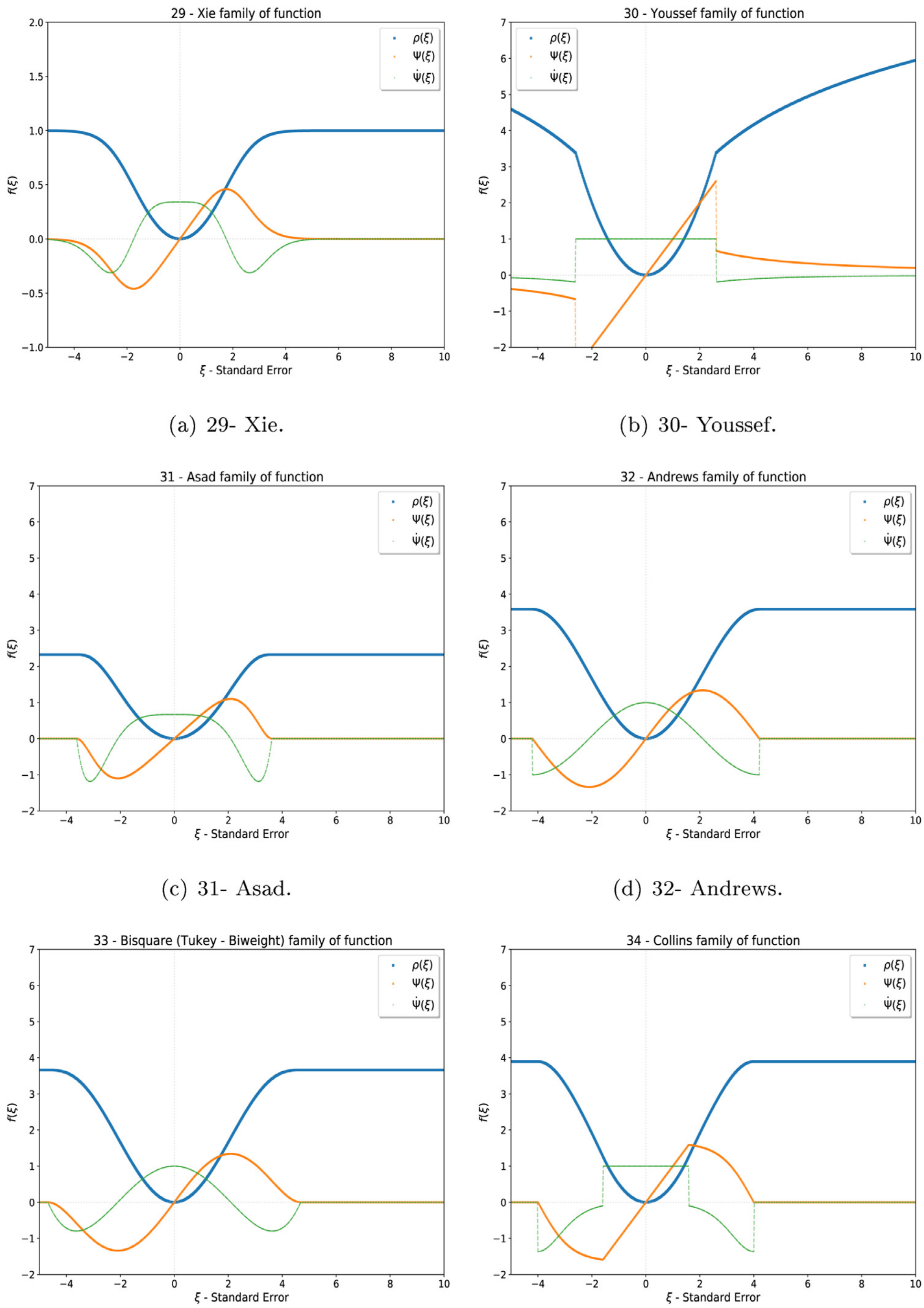
(f) 28- Wu.

Fig. 9. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 5/8.

Parameter c follows geometric proportionality, varying according to Relative Efficiency, as it is associated with sensitivity to gross error. Therefore, the tuning parameter is b.

- Optimal: This estimator has 5 parameters and 3 parts of function. The 5 parameters are the parameters of the polynomial interpolation of the original estimator, because the analytical function of the estimator was not found. The 5 parameters must follow the relationships that suit the desired properties of the characteristic functions of the M-estimator, such as continuity. This estimator can be tuned iteratively. Initially, c_{Op} and d_{Op} are estimated at $c_{Op} = 2e^{-4}$ and $d_{Op} = 2e^{-3}$. The a_{Op} and b_{Op} parameters depend on the other parameters through the following relationships:
  1. a_{Op} is the smallest positive root of the equation ($\psi_2 - \psi_1 = 0$), ψ_i being the ith part of influence function;
  2. b_{Op} is the positive root of the equation ($\psi_2 = 0$), ψ_i being the ith part of influence function.

  Thus, the parameters c_{Op} and d_{Op} must be readjusted, because they depend on a_{Op} and b_{Op}. The adjustments follow the relation:
  1. $c_{Op} = \frac{7a_{Op}^6 e_{Op} - 5a_{Op}^4 d_{Op}}{-3a_{Op}^2}$;



(a) 29- Xie. (b) 30- Youssef. (c) 31- Asad. (d) 32- Andrews. (e) 33- Biweight. (f) 34- Collins.

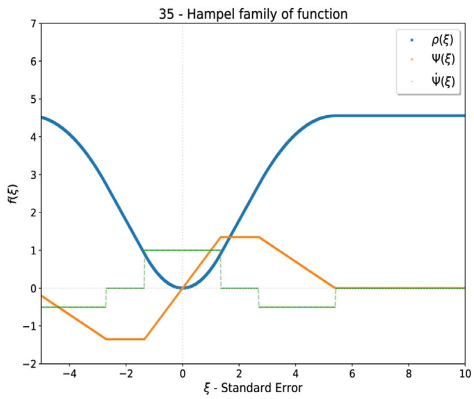
Fig. 10. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 6/8.

2. $d_{Op} = \frac{b_{Op}^7 e_{Op} + b_{Op}^3 c_{Op} - b_{Op}}{b_{Op}^5}$.

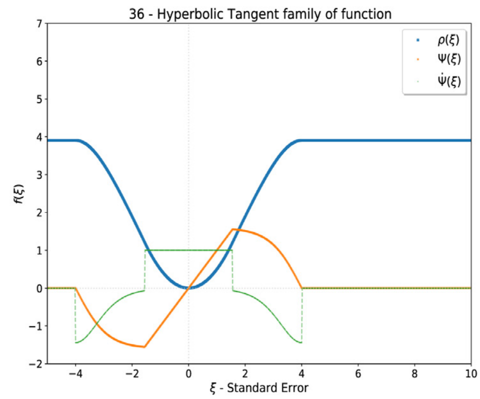
Thus, the only possible parameter for tuning is the e_{Op} parameter.

- Yang I: This estimator presents 3 parameters and 3 parts of functions. The D_{YaI} parameter is the auxiliary parameter, with $D_{YaI} = \rho_1(c_{YaI}) - \rho_2(c_{YaI})$. The parameter a_{YaI} (which represents the point of rejection of the estimator) must follow the geometric proportionality, varying according to the Relative Efficiency and fixed for a tuning. Thus, for this work the tuning parameter used was c_{YaI}.
- Yang II: This estimator presents 2 parameters and 3 parts of functions. The parameter a_{YaII} (which represents the point of rejection of the estimator) must follow the geometric proportionality, varying according to the Relative Efficiency and fixed for a tuning. Thus, for this work the tuning parameter used was c_{YaII}.

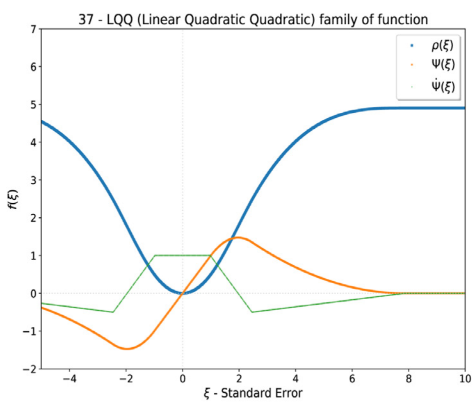
Additional information regarding the tuning constants for each estimator can be found in the respective original works, as shown in Table 3. Table D.7 in Appendix D shows the tunings for the Relative Efficiencies of 90, 98 and 99%, in respect to the Normal distribution. Most of these tunings are proposed here for the first time,



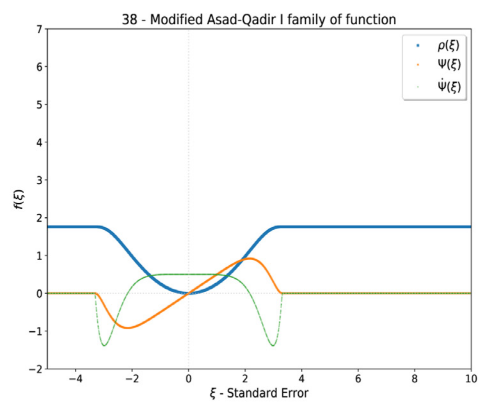
(a) 35- Hampel.



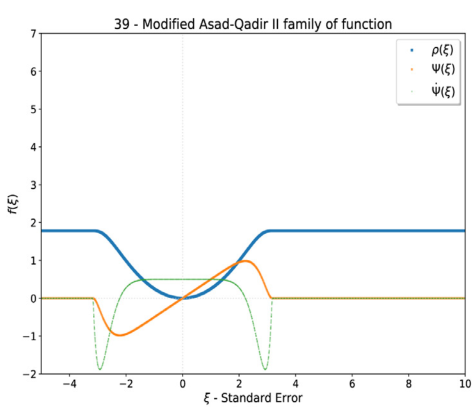
(b) 36- Hyperbolic Tangent.



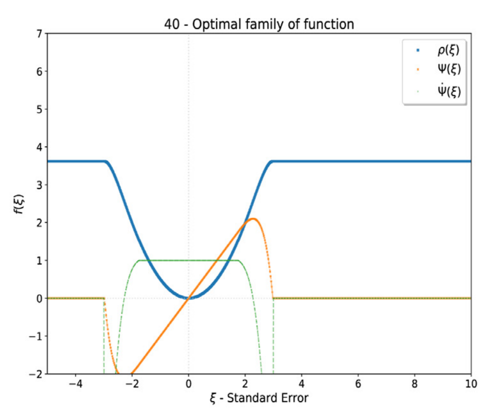
(c) 37- LQQ.



(d) 38- Modified Asad-Qadir I.



(e) 39- Modified Asad-Qadir II.

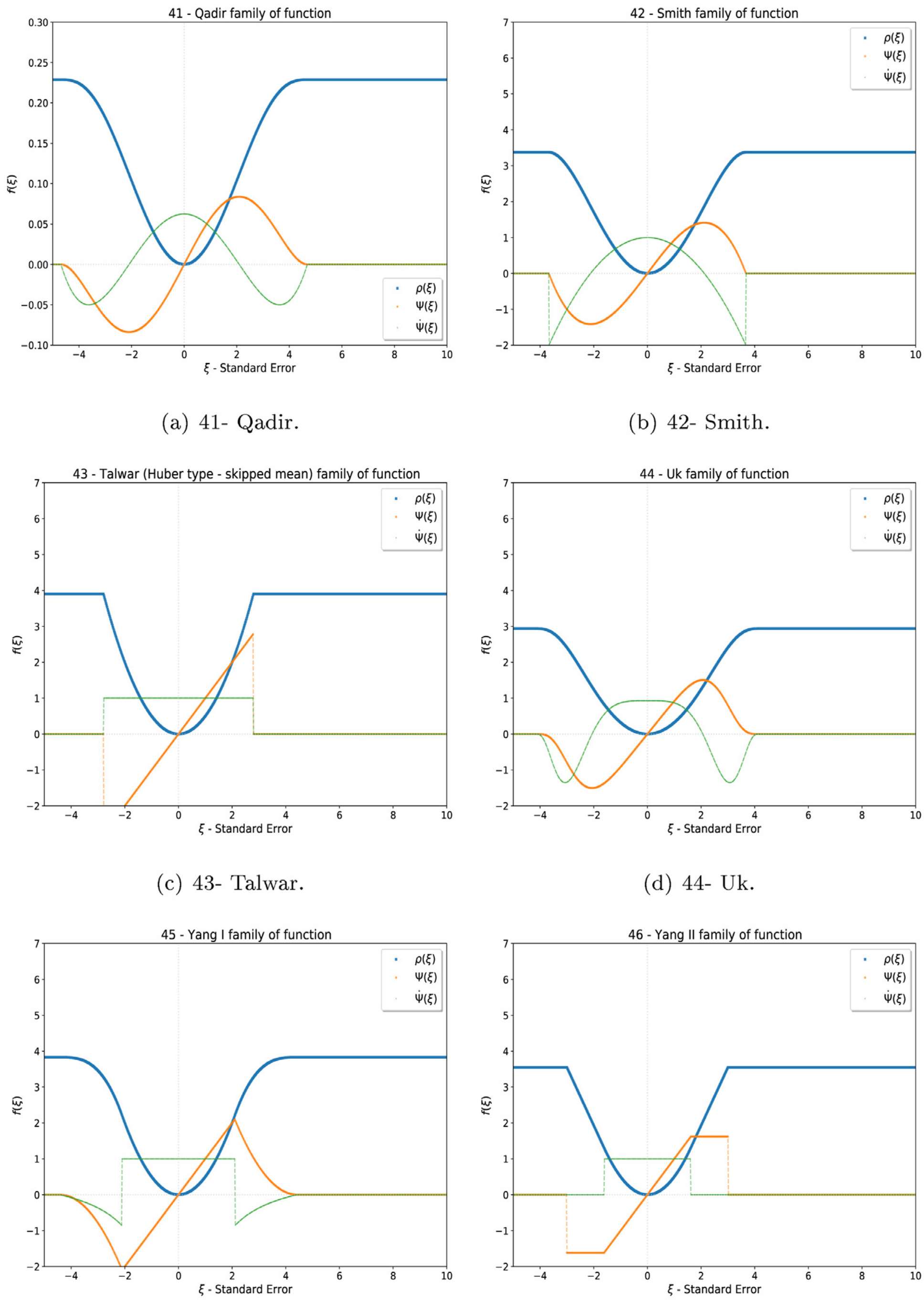


(f) 40- Optimal.

Fig. 11. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 7/8.

especially for the 98 and 99% efficiencies. All parameters presented in Table D.7 were tuned in the present work, even when they were available in the literature, to assure the consistency of the reported data.

It can be seen in Fig. 15 that the ρ functions (objective function) tend to become more convex when the relative efficiency increases for the estimators: (i) Contaminated Normal (Fig. 15a) - quasi-robust; (ii) Fair (Fig. 15b) - monotonous; (iii) Welsch (Fig. 15c) - soft-redescending; and (iv) Smith (Fig. 15d) - hard-redescending. This is because these estimators were tuned in respect to the Normal distribution, which is strictly convex. Notably, the Contaminated Normal (quasi-robust) estimator is more convex at the 99% efficiency level than at the 90% efficiency level (with $\xi = 3$), as its structure is based on the Normal distribution. It can also be observed that the robustness decreases when the relative efficiency increases, as indicated by the ψ functions (influence function). Again, this is because these estimators were tuned in respect to the Normal distribution (which is not robust). This is particularly evident in the case of the Smith estimator (hard-redescendig and non-convex), as the rejection point (r_p) increases at high efficiency levels; that is, it takes longer (in terms of the magnitude of the standard deviation) to remove the



(a) 41- Qadir. (b) 42- Smith.

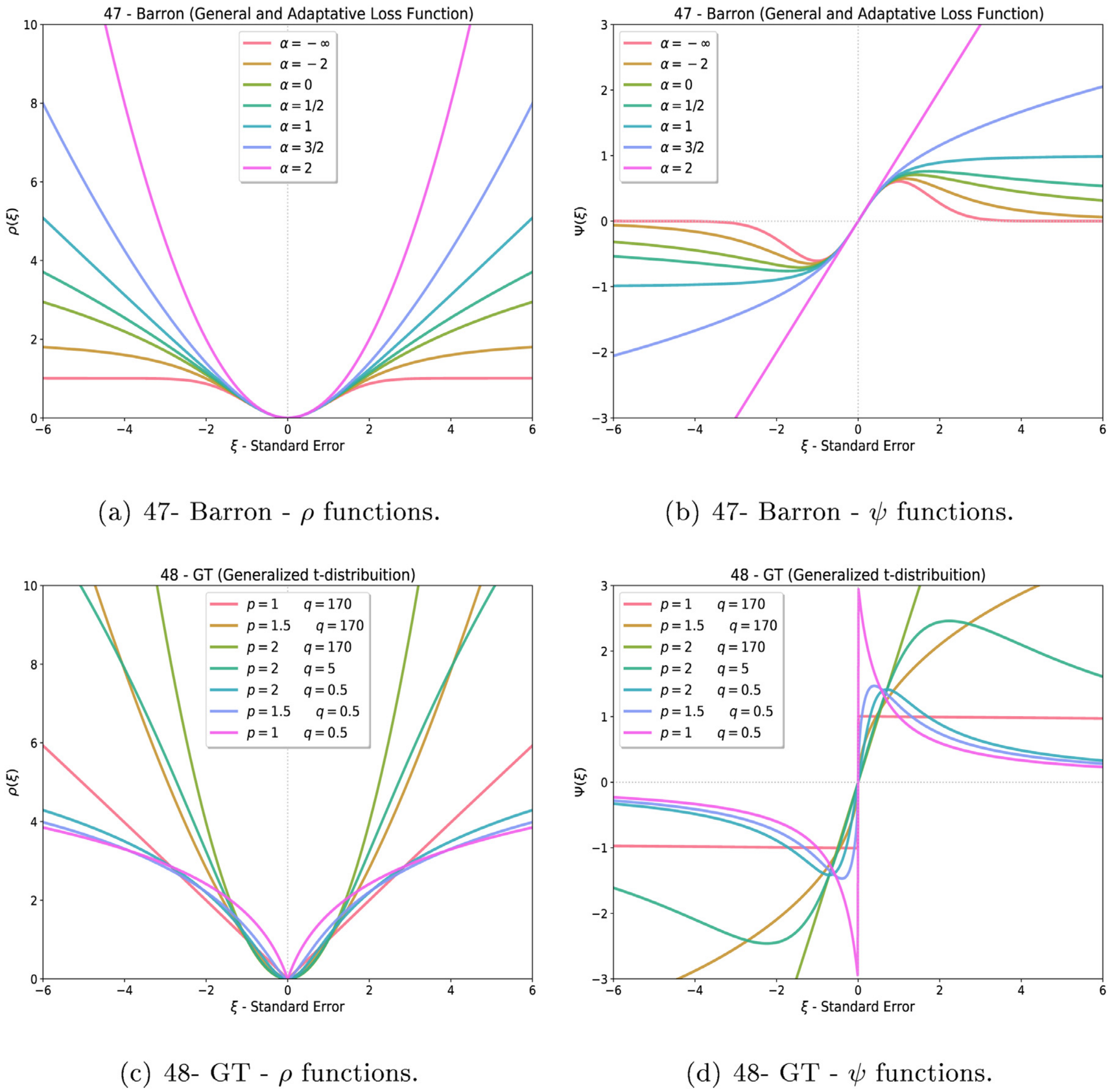
(c) 43- Talwar. (d) 44- Uk.

(e) 45- Yang I. (f) 46- Yang II.

Fig. 12. Families of analyzed functions for distinct estimators (ρ, ψ, ψ') part 8/8.

effect of a spurious value (less robust) in this case. Otherwise, the rejection point decreases at low efficiency levels; that is, the estimator descends steeply to nullify the effect of a spurious value, so that the estimator becomes more robust and less similar to the Normal distribution. For example, selecting a spurious value of magnitude $\xi = 4$, this estimator nullifies the negative effect on the estimates for the levels of 90–95% of relative efficiency, which does not occur for the 98-99% efficiency levels.

Therefore, as previously shown in Table 1 (which presents 37 M-estimator applications in the chemical engineering field) and based on the results reported in these works, it becomes possible to say that soft-redescending M-estimators described by exponential terms (without "if" clause) have received more attention from researchers so far. Besides, as reported by those authors, these estimators have presented the best overall performances in comparative studies (Welsch - Prata et al. (2008b); Correntropy - Chen et al. (2013); Correntropy/Biweight - Llanos et al. (2015, 2017); Welsch/Correntropy/Bell - da Cunha et al. (2017); Xie/Welsch - Xie et al. (2019)). Additionally, these estimators also performed well when the data sets were contaminated with random errors (at 95% level of efficiency in respect to the Normal distribution). As shown in Table 4 (which presents a summary of



(a) 47- Barron - ρ functions.

(b) 47- Barron - ψ functions.

(c) 48- GT - ρ functions.

(d) 48- GT - ψ functions.

Fig. 13. General and Adaptative Robust Estimators.

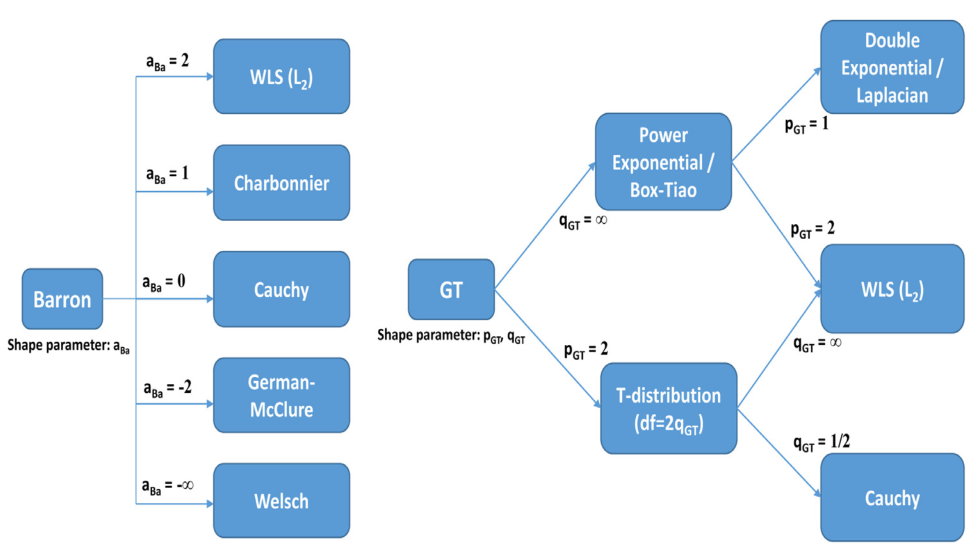
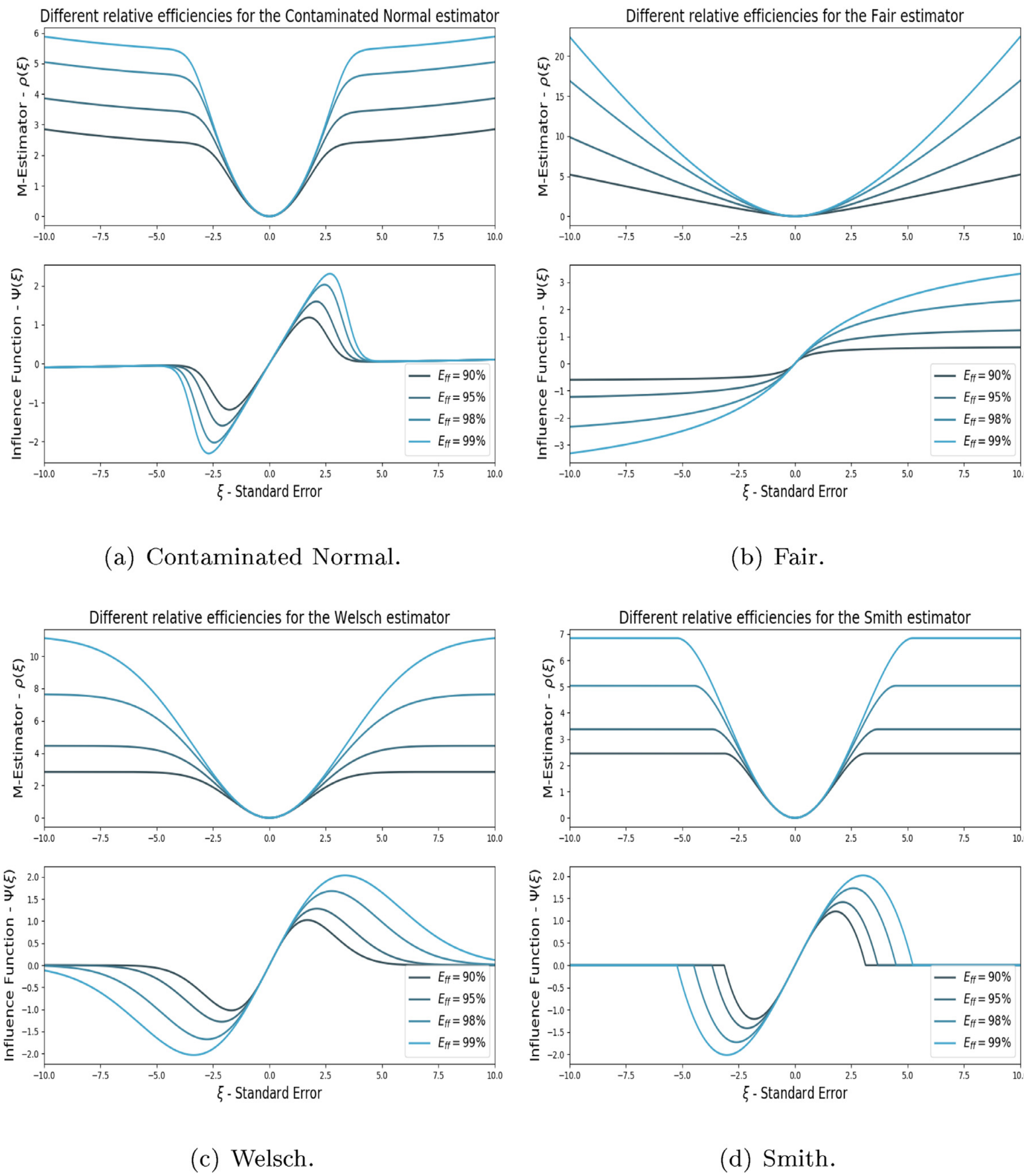


Fig. 14. Shape parameter of Barron and GT M-estimators.



(a) Contaminated Normal.

(b) Fair.

(c) Welsch.

(d) Smith.

Fig. 15. Characteristic functions of different estimators for distinct relative efficiency levels.

the properties and characteristics of the M-estimators presented in this work), the use of these estimators is also attractive because the respective IF's (ψ) are continuous and redescending, nullifying the negative effects of spurious values over the estimates. Besides, the derivatives of the IF's (ψ') are also continuous in these cases and do not require the use of smoothing functions when deterministic optimizations are carried out (Arora and Biegler, 2001; Wongrat et al., 2005). In addition, these estimators depend on a single tuning parameter, which can be useful for actual real time applications, when it is desired to adjust the tuning constant appropriately to data sets that may contain gross error frequencies that vary with the operation conditions (Prata et al., 2010; Valdetaro and Schirru, 2011). However, although many redescending robust M-estimators are available and have been used so far, characterization of relative advantages through detailed comparative studies involving benchmarking problems (particularly when nonlinear constraints and dynamic problems are considered) have not been developed yet. This indicates that special attention must be given to the redescending M-estimators presented in the present work and to the main features that may encourage the development of new ones in the near future (based on the desired properties that are revised and discussed here).

6. Conclusions

Although robust M-estimators have been applied in many fields of knowledge, including the fields of pure and applied statistics, and outperform the classical WLS estimator, robust regression methods have not been widely used yet. Several reasons may possibly help to understand this point:

- Robust estimation is computationally more intense than classical method (Hampel et al., 1986); however, due to the increasing computation capacity this possible reason is becoming less pertinent;

- For many robust estimators, especially those classified as re-descending, there is a need for use of global optimization methods, as some of these estimators are pseudo-convex or quasi-convex (both generally called “non-convex” in literature) (Baselga, 2007; Prata et al., 2009; Prata et al., 2010);
- The vast majority of popular statistical computational packages have not yet introduced robust methods (Stromberg et al., 2004);
- The opinion of many statisticians and researchers that classical statistical methods are robust enough (although they are not).

Most of the present review was devoted to presentation and use of M-estimators, because even though some applications of these estimators are focused specifically on robust statistics, electrical/electronic engineering, and signal processing, other areas of knowledge have also developed, studied and found use for M-estimators, particularly for solution of regression problems in the field of chemical engineering. For this reason, the present review presents for the first time a compilation of 50 (48 robust) estimators, with the respective characteristic functions (the ρ function, the ψ function and the ψ' function) and tunable parameters for efficiency levels of 90%, 95%, 98% and 99% in respect to the normal distribution. Importantly, many of these estimators had never been presented previously in graphical form, making more difficult the correct implementation and validation of the estimators. Besides, many of the analyzed estimators had been used few times (sometimes a single time), showing the importance of tuning and introducing them to the scientific communities. As a whole, this paper presents a collection of robust M-estimators that are spread across several areas of science that study data regression problems. Thus, the present work fills a gap and allows the knowledge exchange among these diverse scientific communities, providing a a large and well-documented number of case studies, results, and conclusions about the use of robust M-estimators are presented. Finally, it should be noted through Table 1, that the Contaminated Normal (quasi-robust), Welsch, Hampel, Fair, Lorentzian, Correntropy, and Cauchy M-estimators were the most used for regression analysis in chemical engineering problems, showing that the preference in this area represents a reduced set in the universe of M-estimators, such as the 49 M-estimators (Contaminated Normal + 48 robust estimators) presented in this work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Relative efficiency of LAV estimator

To illustrate the calculation of Relative Efficiency (Eff), the M-LAV estimator is selected because it does not contain any tuning constant. This estimator has an efficiency equal to $2/\pi$, approximately equal to 64% (Huber, 1981) in respect to the Normal distribution. Based on the following settings:

$$f_{Normal} = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2}\xi^2) \quad \text{(A.1)}$$

$$f'_{Normal} = -\xi\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2}\xi^2) \quad \text{(A.2)}$$

$$\psi_{Normal}(\xi) = \frac{d\rho(\xi)}{d\xi} = \frac{d}{d\xi}(\frac{1}{2}\xi^2) = \xi \quad \text{(A.3)}$$

$$\psi_{LAV}(\xi) = sgn(\xi) \quad \text{(A.4)}$$

and using Eqs. (14) and (16), it is possible to calculate the Efficiency of the LAV M-estimator in respect to to the Normal distribution, as presented in Eq. (A.5) through (A.8).

$$E_{ff}[\psi, f]_{LAV} = \frac{V_f[\psi_{Normal}, f_{Normal}]}{V[\psi_{LAV}, f_{Normal}]} \quad \text{(A.5)}$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\dfrac{\int_{-\infty}^{+\infty}\psi_{Normal}^2(\xi)f_{Normal}(\xi)d\xi}{\left[-\int_{-\infty}^{+\infty}\psi_{Normal}(\xi)f'_{Normal}(\xi)d\xi\right]^2}}{\dfrac{\int_{-\infty}^{+\infty}\psi_{LAV}^2(\xi)f_{Normal}(\xi)d\xi}{\left[-\int_{-\infty}^{+\infty}\psi_{LAV}(\xi)f'_{Normal}(\xi)d\xi\right]^2}} \quad \text{(A.6)}$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\dfrac{\frac{1}{\sigma\sqrt{2\pi}}\int_0^{+\infty}\xi^2\exp(-\frac{1}{2}\xi^2)d\xi}{(\frac{1}{\sigma\sqrt{2\pi}})^2\left[\int_0^{+\infty}\xi^2\exp(-\frac{1}{2}\xi^2)d\xi\right]^2}}{\dfrac{\frac{1}{\sigma\sqrt{2\pi}}sgn^2(\xi)\int_0^{+\infty}\exp(-\frac{1}{2}\xi^2)d\xi}{(\frac{1}{\sigma\sqrt{2\pi}})^2 sgn^2(\xi)\left[\int_0^{+\infty}\xi\exp(-\frac{1}{2}\xi^2)d\xi\right]^2}} \quad \text{(A.7)}$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\dfrac{\frac{\sqrt{2\pi}}{2}}{\left[\frac{\sqrt{2\pi}}{2}\right]^2}}{\dfrac{\frac{\sqrt{2\pi}}{2}}{[1]^2}} = \frac{\dfrac{1}{\frac{\sqrt{2\pi}}{2}}}{\dfrac{\sqrt{2\pi}}{2}} = \frac{1}{\frac{\sqrt{2\pi}}{2}}\frac{1}{\frac{\sqrt{2\pi}}{2}} = \frac{4}{2\pi} = \frac{2}{\pi} \quad \text{(A.8)}$$

It is important to observe that this is the basis for calculation of tuning constants of all estimators. The relative efficiency (90, 95, 98 or 99%) and the reference distribution (usually Normal distribution) are fixed. Numerical methods can be used to solve the resulting equation.

Appendix B. Example of tuning procedure in maple

Figs. B.16–B.18

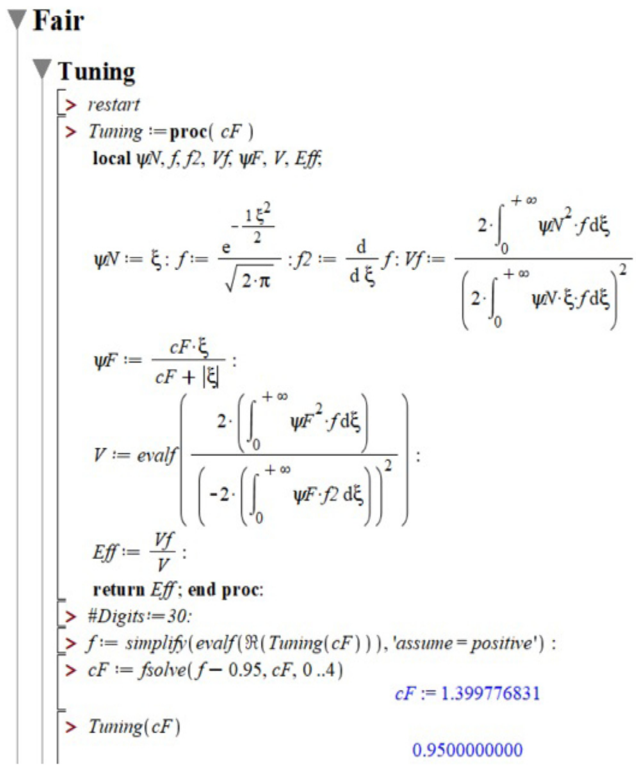


Fig. B.16. Tuning procedure for the Fair estimator.

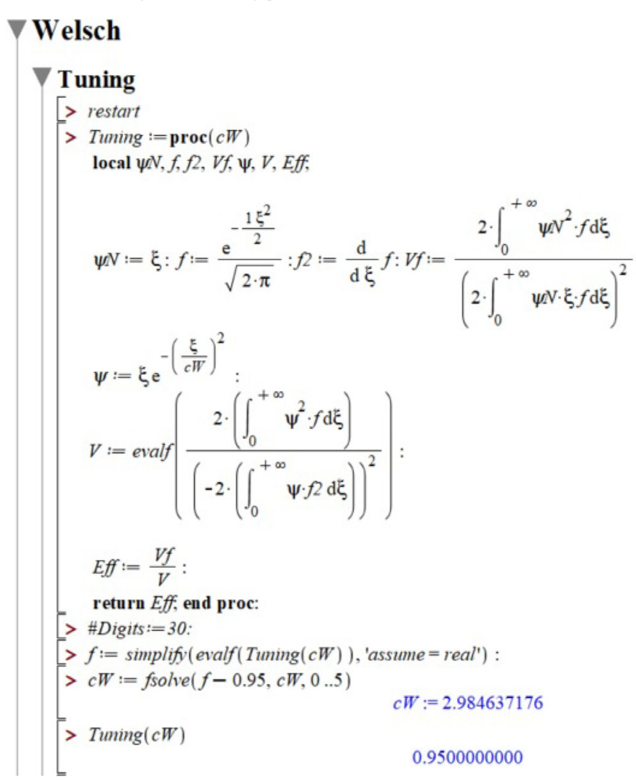


Fig. B.17. Tuning procedure for the Welsch estimator.

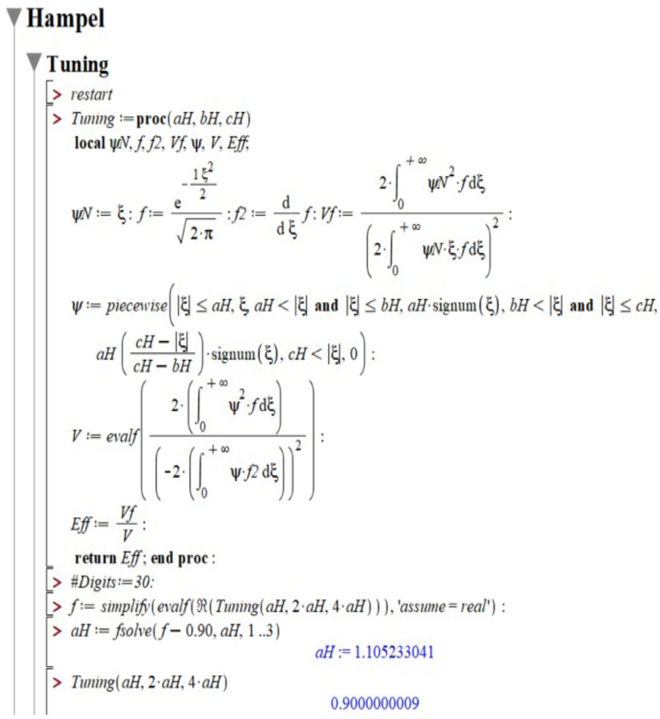


Fig. B.18. Tuning procedure for the Hampel estimator.

Appendix C. M-estimators functions - ρ and ψ

Table C.5
M-Estimators: $\rho(\xi)$.

#	$\boldsymbol{\rho(\xi)}$	Range
1	$\lvert\xi\rvert$	$\lvert\xi\rvert \le \infty$
2	$2\left(\sqrt{1+\frac{\xi^2}{2}}-1\right)$	$\lvert\xi\rvert \le \infty$
3	$c_F^2\left[\frac{\lvert\xi\rvert}{c_F}-\ln\left(1+\frac{\lvert\xi\rvert}{c_F}\right)\right]$	$\lvert\xi\rvert \le \infty$
4	$\begin{cases}\frac{\xi^2}{2}\\ c_{Hu}\lvert\xi\rvert-\frac{c_{Hu}^2}{2}\end{cases}$	$\lvert\xi\rvert \le c_{Hu}$ $\lvert\xi\rvert > c_{Hu}$
5	$\begin{cases}c_{MH}^2\left[1-\cos\left(\frac{\xi}{c_{MH}}\right)\right]\\ c_{MH}\lvert\xi\rvert+c_{MH}^2\left(1-\frac{\pi}{2}\right)\end{cases}$	$\lvert\xi\rvert \le \frac{\pi}{2}c_{MH}$ $\lvert\xi\rvert > \frac{\pi}{2}c_{MH}$
6	$\frac{\xi^2}{c_K+\lvert\xi\rvert}$	$\lvert\xi\rvert \le \infty$
7	$\frac{\ln(\cosh(c_{LC}\xi))}{c_{LC}}$	$\lvert\xi\rvert \le \infty$
8	$c_L^2\ln\left[\cosh\left(\frac{\xi}{c_L}\right)\right]$	$\lvert\xi\rvert \le \infty$
9	$\begin{cases}\frac{\xi^2}{2}\\ \frac{c_M\lvert\xi\rvert}{2}\end{cases}$	$\lvert\xi\rvert \le c_M$ $\lvert\xi\rvert > c_M$
10	$\frac{\xi^2}{2+c_Z\lvert\xi\rvert}$	$\lvert\xi\rvert \le \infty$
11	$\frac{4c_{Al}^2\left[1+3\exp\left(-\frac{\xi^2}{c_{Al}^2}\right)\right]}{3\left[1+\exp\left(-\frac{\xi^2}{c_{Al}^2}\right)\right]^3}-\frac{2c_{Al}^2}{3}$	$\lvert\xi\rvert \le \infty$
12	$\begin{cases}\frac{\xi^2}{2}\\ c_{BH}^2\left[\ln\left(\frac{\lvert\xi\rvert}{c_{BH}}\right)+\frac{1}{2}\right]\end{cases}$	$\lvert\xi\rvert \le c_{BH}$ $\lvert\xi\rvert > c_{BH}$
13	$\frac{5c_{Be}^2}{4}\left[1-\left(1+\frac{\xi^2}{5c_{Be}^2}\right)^{-2}\right]$	$\lvert\xi\rvert \le \infty$
14	$\log\left[\frac{1+c_{BZ}}{\exp(-\xi^2)+c_{BZ}}\right]$	$\lvert\xi\rvert \le \infty$
15	$\frac{c_C^2}{2}\ln\left(1+\frac{\xi^2}{c_C^2}\right)$	$\lvert\xi\rvert \le \infty$
M-Estimators:	$\rho(\xi)$ - part 2/5	
16	$\frac{1}{c_{Co}\sqrt{2\pi}}\left[1-\exp\left(-\frac{\xi^2}{2c_{Co}^2}\right)\right]$	$\lvert\xi\rvert \le \infty$

(continued on next page)

Table C.5 (*continued*)

#	$\boldsymbol{\rho(\xi)}$	Range
17	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{c_{Da}}{2}\exp\left(\frac{-\xi^2}{c_{Da}}\right)+D_{Da} \end{cases}$	$\lvert\xi\rvert \le c_{Da}$ $\lvert\xi\rvert > c_{Da}$
18	$\frac{\xi^2}{c_{GM}^2+\xi^2}$	$\lvert\xi\rvert \le \infty$
19	*There is no analytical integral of the influence function	
20	$\frac{c_I^2}{4}\left[\arctan\left(\frac{\xi^2}{c_I^2}\right)+\frac{c_I^2\xi^2}{c_I^4+\xi^4}\right]$	$\lvert\xi\rvert \le \infty$
21	$\begin{cases} \frac{c_J^2}{6}\left[1-\left(1-a_J\frac{\xi^2}{c_J^2}\right)^3\right] \\ \frac{c_J^2 a_J}{6}\left[3\left(1-a_J\right)^2\ln\left(\frac{\xi^2}{c_J^2}\right)+a_J^2-3a_J+3\right] \end{cases}$	$\lvert\xi\rvert \le c_J$ $\lvert\xi\rvert > c_J$
22	$\begin{cases} \frac{\xi^2}{2} \\ a_{Ku}\lvert\xi\rvert-\frac{a_{Ku}^2}{2} \\ -\frac{a_{Ku}\cdot b_{Ku}}{2}\exp\left(1-\frac{\xi^2}{b_{Ku}^2}\right)+D_{Ku} \end{cases}$	$\lvert\xi\rvert \le a_{Ku}$ $a_{Ku} < \lvert\xi\rvert \le b_{Ku}$ $\lvert\xi\rvert > b_{Ku}$
23	$1-\frac{1}{1+\left(\frac{\xi^2}{2c_{Lz}^2}\right)}$	$\lvert\xi\rvert \le \infty$
24	$\begin{cases} \frac{\xi^2}{2} \\ 2c_{MS}^{3/2}\sqrt{\lvert\xi\rvert}-1.5c_{MS}^2 \end{cases}$	$\lvert\xi\rvert \le c_{MS}$ $\lvert\xi\rvert > c_{MS}$
25	$\frac{1}{c_R^2}\left[1-\left(1+c_R\lvert\xi\rvert\right)\exp\left(-c_R\lvert\xi\rvert\right)\right]$	$\lvert\xi\rvert \le \infty$
26	$1-\mathrm{sech}(c_{Sh}\xi)$	$\lvert\xi\rvert \le \infty$
27	$\frac{c_W^2}{2}\left[1-\exp\left(-\frac{\xi^2}{c_W^2}\right)\right]$	$\lvert\xi\rvert \le \infty$
28	$\begin{cases} \frac{\xi^2}{2} \\ \frac{a_{Wu}^{\lvert\xi\rvert-c_{Wu}}}{\ln\left(a_{Wu}\right)}-\frac{1}{\ln\left(a_{Wu}\right)}+\frac{c_{Wu}^2}{2} \end{cases}$	$\lvert\xi\rvert \le c_{Wu}$ $\lvert\xi\rvert > c_{Wu}$
M-Estimators:	$\rho(\xi)$ - part 3/5	
29	$\frac{1-\exp\left(-\frac{\xi^2}{c_X^2}\right)}{1+\exp\left(-\frac{\xi^2}{c_X^2}\right)}$	$\lvert\xi\rvert \le \infty$
30	$\begin{cases} \frac{\xi^2}{2} \\ \ln\left(\xi^2+1\right)+\left[\frac{c_Y^2}{2}-\ln\left(c_Y^2+1\right)\right] \end{cases}$	$\lvert\xi\rvert \le c_Y$ $\lvert\xi\rvert > c_Y$
31	$\begin{cases} \frac{\xi^2}{45c_{As}^8}\left(3\xi^8-10c_{As}^4\xi^4+15c_{As}^8\right) \\ \frac{8c_{As}^2}{45} \end{cases}$	$\lvert\xi\rvert \le c_{As}$ $\lvert\xi\rvert > c_{As}$
32	$\begin{cases} c_{An}^2\left[1-\cos\left(\frac{\xi}{c_{An}}\right)\right] \\ 2c_{An}^2 \end{cases}$	$\lvert\xi\rvert \le \pi c_{An}$ $\lvert\xi\rvert > \pi c_{An}$
33	$\begin{cases} \frac{c_{Bi}^2}{6}\left[1-\left(1-\frac{\xi^2}{c_{Bi}^2}\right)^3\right] \\ \frac{c_{Bi}^2}{6} \end{cases}$	$\lvert\xi\rvert \le c_{An}$ $\lvert\xi\rvert > c_{An}$
34	$\begin{cases} \frac{\xi^2}{2} \\ -2\left\{\log\left[\cosh\left(\frac{a_{Cl}}{2}\left(\lvert\xi\rvert-r_{Cl}\right)\right)\right]\right\}+d_{Cl} \\ d_{Cl} \end{cases}$	$\lvert\xi\rvert \le q_{Cl}$ $q_{Cl} < \lvert\xi\rvert \le r_{Cl}$ $\lvert\xi\rvert > r_{Cl}$
35	$\begin{cases} \frac{\xi^2}{2} \\ a_H\lvert\xi\rvert-\frac{a_H^2}{2} \\ a_Hb_H-\frac{a_H^2}{2}+\frac{a_H\left(c_H-b_H\right)}{2}\left[1-\left(\frac{c_H-\lvert\xi\rvert}{c_H-b_H}\right)^2\right] \\ a_Hb_H-\frac{a_H^2}{2}+\frac{a_H\left(c_H-b_H\right)}{2} \end{cases}$	$\lvert\xi\rvert \le a_H$ $a_H < \lvert\xi\rvert \le b_H$ $b_H < \lvert\xi\rvert \le c_H$ $\lvert\xi\rvert > c_H$
36	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{2A}{B}\left\{\log\left[\cosh\left(\frac{(k-1)^{1/2}B}{2A^{1/2}}\left(\lvert\xi\rvert-r\right)\right)\right]\right\}+D \\ D \end{cases}$	$\lvert\xi\rvert \le q$ $q < \lvert\xi\rvert \le r$ $\lvert\xi\rvert > r$
M-Estimators:	$\rho(\xi)$ - part 4/5	
37	$\begin{cases} \frac{\xi^2}{2} \\ \frac{1}{6b}\left[3b\xi^2+s\left(c^3-3c^2\lvert\xi\rvert+3c\xi^2-\lvert\xi\rvert^3\right)\right] \\ \begin{cases} \frac{1}{6a}\left\{(s-1)\lvert\xi\rvert^3-3(s-1)(s_2)\xi^2+\right. \\ +\left[\left(3b^2+(3a+6c)b+6ac+3c^2\right)s+\right. \\ \left.\left.-3(c+b)^2\right]\lvert\xi\rvert\right\}+D_1 \end{cases} \text{where:} s_2=a+b+c \\ D_2 \end{cases}$	$\lvert\xi\rvert \le c$ $c < \lvert\xi\rvert \le c+b$ $c+b < \lvert\xi\rvert \le s_2$ $\lvert\xi\rvert > s_2$
38	$\begin{cases} \frac{\xi^2}{56a_{AQI}^{12}}\left(14a_{AQI}^{12}-7a_{AQI}^6\xi^6+2\xi^{12}\right) \\ \frac{9a_{AQI}^2}{56} \end{cases}$	$\lvert\xi\rvert \le a_{AQI}$ $\lvert\xi\rvert > a_{AQI}$
39	$\begin{cases} \frac{\xi^2}{180a_{AQII}^{16}}\left(45a_{AQII}^{16}-18a_{AQII}^8\xi^8+5\xi^{16}\right) \\ \frac{8a_{AQII}^2}{45} \end{cases}$	$\lvert\xi\rvert \le a_{AQII}$ $\lvert\xi\rvert > a_{AQII}$
40	$\begin{cases} \frac{\xi^2}{2} \\ \frac{\xi^2}{2}-\frac{c_{Op}\xi^4}{4}+\frac{d_{Op}\xi^6}{6}-\frac{e_{Op}\xi^8}{8} \\ \frac{b_{Op}^2}{2}-\frac{c_{Op}b_{Op}^4}{4}+\frac{d_{Op}b_{Op}^6}{6}-\frac{e_{Op}b_{Op}^8}{8} \end{cases}$	$\lvert\xi\rvert \le a_{Op}$ $a_{Op} < \lvert\xi\rvert \le b_{Op}$ $\lvert\xi\rvert > b_{Op}$
41	$\begin{cases} \frac{\xi^2}{96c_Q^4}\left(\xi^4-3c_Q^2\xi^2+3c_Q^4\right) \\ \frac{c_Q^2}{96} \end{cases}$	$\lvert\xi\rvert \le c_Q$ $\lvert\xi\rvert > c_Q$
42	$\begin{cases} \frac{c_S^2}{4}\left[1-\left(1-\frac{\xi^2}{c_S^2}\right)^2\right] \\ \frac{c_S^2}{4} \end{cases}$	$\lvert\xi\rvert \le c_S$ $\lvert\xi\rvert > c_S$

(*continued on next page*)

Table C.5 (*continued*)

#	$\rho(\xi)$	Range
43	$\begin{cases} \frac{\xi^2}{2} \\ \frac{c_T^2}{2} \end{cases}$	$\lvert\xi\rvert \le c_T$ $\lvert\xi\rvert > c_T$
M-Estimators: $\rho(\xi)$ - part 5/5		
44	$\begin{cases} \frac{3}{2}\sin\left(\frac{4}{9}\right)\left[\frac{\xi^{10}}{10c_{Uk}^8} - \frac{\xi^6}{3c_{Uk}^4} + \frac{\xi^2}{2}\right] \\ \frac{3}{2}\sin\left(\frac{4}{9}\right)c_{Uk}^2 \end{cases}$	$\lvert\xi\rvert \le c_{Uk}$ $\lvert\xi\rvert > c_{Uk}$
45	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{c_{YaI}(a_{YaI}-\lvert\xi\rvert)^3}{3(c_{YaI}-a_{YaI})^2} + D_{YaI} \\ D_{YaI} \end{cases}$	$\lvert\xi\rvert \le c_{YaI}$ $c_{YaI} < \lvert\xi\rvert \le a_{YaI}$ $\lvert\xi\rvert > a_{YaI}$
46	$\begin{cases} \frac{\xi^2}{2} \\ c_{YaII}\lvert\xi\rvert + \frac{c_{YaII}^2}{2} - c_{YaII}\lvert c_{YaII}\rvert \\ c_{YaII}\lvert a_{YaII}\rvert + \frac{c_{YaII}^2}{2} - c_{YaII}\lvert c_{YaII}\rvert \end{cases}$	$\lvert\xi\rvert \le c_{YaII}$ $c_{YaII} < \lvert\xi\rvert \le a_{YaII}$ $\lvert\xi\rvert > a_{YaII}$
47	$\frac{\lvert 2-a_{Ba}\rvert}{a_{Ba}}\left[\left(\frac{\left(\frac{\xi}{c_{Ba}}\right)^2}{\lvert 2-a_{Ba}\rvert} + 1\right)^{\frac{a_{Ba}}{2}} - 1\right]$	$\lvert\xi\rvert \le \infty$
48	$\ln\left(1 + \frac{\lvert\xi\rvert^{p_{GT}}}{q_{GT}}\right)\left(q_{GT} + \frac{1}{p_{GT}}\right)$	$\lvert\xi\rvert \le \infty$

Table C.6
Influence Function: $\psi(\xi)$.

#	$\psi(\xi)$	Range
1	$\text{sign}(\xi)$	$\lvert\xi\rvert \le \infty$
2	$\frac{\xi}{\sqrt{1+\frac{\xi^2}{2}}}$	$\lvert\xi\rvert \le \infty$
3	$\frac{c_F\xi}{c_F+\lvert\xi\rvert}$	$\lvert\xi\rvert \le \infty$
4	$\begin{cases} \xi \\ c_{Hu}\,\text{sign}(\xi) \end{cases}$	$\lvert\xi\rvert \le c_{Hu}$ $\lvert\xi\rvert > c_{Hu}$
5	$\begin{cases} c_{MH}\sin\left(\frac{\xi}{c_{MH}}\right) \\ c_{MH}\,\text{sign}(\xi) \end{cases}$	$\lvert\xi\rvert \le \frac{\pi}{2}c_{MH}$ $\lvert\xi\rvert > \frac{\pi}{2}c_{MH}$
6	$\frac{\xi(\lvert\xi\rvert+2c_K)}{(c_K+\lvert\xi\rvert)^2}$	$\lvert\xi\rvert \le \infty$
7	$\tanh(c_{LC}\xi)$	$\lvert\xi\rvert \le \infty$
8	$c_L\tanh\left(\frac{\xi}{c_L}\right)$	$\lvert\xi\rvert \le \infty$
9	$\begin{cases} \xi \\ \frac{c_M\,\text{sign}(\xi)}{2} \end{cases}$	$\lvert\xi\rvert \le c_M$ $\lvert\xi\rvert > c_M$
10	$\frac{\xi(4+c_Z\lvert\xi\rvert)}{(2+c_Z\lvert\xi\rvert)^2}$	$\lvert\xi\rvert \le \infty$
11	$\frac{16\xi\left[\exp\left(-\frac{\xi^2}{c_{Al}^2}\right)\right]^2}{\left[1+\exp\left(-\frac{\xi^2}{c_{Al}^2}\right)\right]^4}$	$\lvert\xi\rvert \le \infty$
12	$\begin{cases} \xi \\ \frac{c_{BH}^2}{\xi} \end{cases}$	$\lvert\xi\rvert \le c_{BH}$ $\lvert\xi\rvert > c_{BH}$
13	$\xi\left(1 + \frac{\xi^2}{5c_{Be}^2}\right)^{-3}$	$\lvert\xi\rvert \le \infty$
14	$\frac{2\xi\exp(-\xi^2)}{\left[\exp(-\xi^2)+c_{BZ}\right]\ln(10)}$	$\lvert\xi\rvert \le \infty$
15	$\frac{\xi}{1+\frac{\xi^2}{c_C^2}}$	$\lvert\xi\rvert \le \infty$
Influence Function: $\psi(\xi)$ - part 2/5		
16	$\frac{\xi}{c_{Co}^3\sqrt{2\pi}}\exp\left(-\frac{\xi^2}{2c_{Co}^2}\right)$	$\lvert\xi\rvert \le \infty$
17	$\begin{cases} \xi \\ \xi\exp\left(\frac{-\xi^2}{c_{Da}}\right) \end{cases}$	$\lvert\xi\rvert \le c_{Da}$ $\lvert\xi\rvert > c_{Da}$
18	$\frac{2\xi c_{GM}^2}{\left(c_{GM}^2+\xi^2\right)^2}$	$\lvert\xi\rvert \le \infty$
19	$\begin{cases} \xi \\ \xi\exp\left[-\frac{(\lvert\xi\rvert-c_{GGW})^{b_{GGW}}}{2a_{GGW}}\right] \end{cases}$	$\lvert\xi\rvert \le c_{GGW}$ $\lvert\xi\rvert > c_{GGW}$
20	$\xi\left(1 + \frac{\xi^4}{c_I^4}\right)^{-2}$	$\lvert\xi\rvert \le \infty$
21	$\begin{cases} a_J\xi\left(1 - a_J\frac{\xi^2}{c_J^2}\right)^2 \\ \frac{c_J^2\cdot a_J\left(1-a_J\right)^2}{\xi} \end{cases}$	$\lvert\xi\rvert \le c_J$ $\lvert\xi\rvert > c_J$
22	$\begin{cases} \xi \\ a_{Ku}\,\text{sign}(\xi) \\ \frac{a_{Ku}\xi}{b_{Ku}}\exp\left(1 - \frac{\xi^2}{b_{Ku}^2}\right) \end{cases}$	$\lvert\xi\rvert \le a_{Ku}$ $a_{Ku} < \lvert\xi\rvert \le b_{Ku}$ $\lvert\xi\rvert > b_{Ku}$
23	$\frac{\xi}{c_{Lz}^2\left(1+\frac{\xi^2}{2c_{Lz}^2}\right)^2}$	$\lvert\xi\rvert \le \infty$
24	$\begin{cases} \xi \\ \frac{c_{MS}^{3/2}\,\text{sign}(\xi)}{\sqrt{\lvert\xi\rvert}} \end{cases}$	$\lvert\xi\rvert \le c_{MS}$ $\lvert\xi\rvert > c_{MS}$
25	$\xi\exp(-c_R\lvert\xi\rvert)$	$\lvert\xi\rvert \le \infty$
26	$\frac{c_{Sh}\sinh(c_{Sh}\xi)}{\cosh(c_{Sh}\xi)^2}$	$\lvert\xi\rvert \le \infty$
27	$\xi\exp\left(-\frac{\xi^2}{c_W^2}\right)$	$\lvert\xi\rvert \le \infty$
28	$\begin{cases} \xi \\ \text{sign}(\xi)a_{Wu}^{(\lvert\xi\rvert-c_{Wu})} \end{cases}$	$\lvert\xi\rvert \le c_{Wu}$ $\lvert\xi\rvert > c_{Wu}$

(*continued on next page*)

Table C.6 (*continued*)

#	$\boldsymbol{\psi(\xi)}$	Range
Influence Function: $\psi(\xi)$ - part 3/5		
29	$\frac{4\xi \exp\left(-\frac{\xi^2}{c_X^2}\right)}{c_X^2\left[1+\exp\left(-\frac{\xi^2}{c_X^2}\right)\right]^2}$	$\lvert\xi\rvert \le \infty$
30	$\begin{cases} \xi \\ \frac{2\xi}{\xi^2+1} \end{cases}$	$\lvert\xi\rvert \le c_Y$ $\lvert\xi\rvert > c_Y$
31	$\begin{cases} \frac{2}{3}\xi\left(1-\frac{\xi^4}{c_{As}^4}\right)^2 \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_{As}$ $\lvert\xi\rvert > c_{As}$
32	$\begin{cases} c_{An}\sin\left(\frac{\xi}{c_{An}}\right) \\ 0 \end{cases}$	$\lvert\xi\rvert \le \pi c_{An}$ $\lvert\xi\rvert > \pi c_{An}$
33	$\begin{cases} \xi\left(1-\frac{\xi^2}{c_{Bi}^2}\right)^2 \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_{Bi}$ $\lvert\xi\rvert > c_{Bi}$
34	$\begin{cases} \xi \\ a_{Cl}\tanh\left[\frac{a_{Cl}}{2}(r_{Cl}-\lvert\xi\rvert)\right]\mathrm{sign}(\xi) \\ 0 \end{cases}$	$\lvert\xi\rvert \le q_{Cl}$ $q_{Cl} < \lvert\xi\rvert \le r_{Cl}$ $\lvert\xi\rvert > r_{Cl}$
35	$\begin{cases} \xi \\ a_H\,\mathrm{sign}(\xi) \\ \frac{a_H(c_H-\lvert\xi\rvert)}{c_H-b_H}\mathrm{sign}(\xi) \\ 0 \end{cases}$	$\lvert\xi\rvert \le a_H$ $a_H < \lvert\xi\rvert \le b_H$ $b_H < \lvert\xi\rvert \le c_H$ $\lvert\xi\rvert > c_H$
36	$\begin{cases} \xi \\ [A(k-1)]^{1/2}\tanh\left[\frac{(k-1)^{1/2}B}{2A^{1/2}}(r-\lvert\xi\rvert)\right]\mathrm{sign}(\xi) \\ 0 \end{cases}$	$\lvert\xi\rvert \le q$ $q < \lvert\xi\rvert \le r$ $\lvert\xi\rvert > r$
Influence Function: $\psi(\xi)$ - part 4/5		
37	$\begin{cases} \xi \\ \mathrm{sign}(\xi)\left[\lvert\xi\rvert-\frac{s}{2b}(\lvert\xi\rvert-c)^2\right] \\ \begin{cases} \mathrm{sign}(\xi)\left[c+b-\frac{bs}{2}+\right. \\ \left.+\frac{(s-1)}{a}\left(\frac{1}{2}(\lvert\xi\rvert-b-c)^2-a(\lvert\xi\rvert-b-c)\right)\right] \end{cases} \\ 0 \end{cases}$ where: $s_2=a+b+c$	$\lvert\xi\rvert \le c$ $c < \lvert\xi\rvert \le c+b$ $c+b < \lvert\xi\rvert \le s_2$ $\lvert\xi\rvert > s_2$
38	$\begin{cases} \frac{\xi}{2}\left(1-\frac{\xi^6}{a_{AQI}^6}\right)^2 \\ 0 \end{cases}$	$\lvert\xi\rvert \le a_{AQI}$ $\lvert\xi\rvert > a_{AQI}$
39	$\begin{cases} \frac{\xi}{2}\left(1-\frac{\xi^8}{a_{AQII}^8}\right)^2 \\ 0 \end{cases}$	$\lvert\xi\rvert \le a_{AQII}$ $\lvert\xi\rvert > a_{AQII}$
40	$\begin{cases} \xi \\ \xi-c_{Op}\xi^3+d_{Op}\xi^5-e_{Op}\xi^7 \\ 0 \end{cases}$	$\lvert\xi\rvert \le a_{Op}$ $a_{Op} < \lvert\xi\rvert \le b_{Op}$ $\lvert\xi\rvert > b_{Op}$
41	$\begin{cases} \frac{\xi}{16}\left(1-\frac{\xi^2}{c_Q^2}\right)^2 \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_Q$ $\lvert\xi\rvert > c_Q$
42	$\begin{cases} \xi\left(1-\frac{\xi^2}{c_S^2}\right) \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_S$ $\lvert\xi\rvert > c_S$
43	$\begin{cases} \xi \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_T$ $\lvert\xi\rvert > c_T$
Influence Function: $\psi(\xi)$ - part 5/5		
44	$\begin{cases} \frac{3}{2}\xi\left(1-\frac{\xi^4}{c_{Uk}^4}\right)^2\sin\left[\frac{2}{3}\left(1-\frac{\xi^4}{c_{Uk}^4}\right)^2\right] \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_{Uk}$ $\lvert\xi\rvert > c_{Uk}$
45	$\begin{cases} \frac{\xi^2}{2} \\ \frac{\xi\cdot c_{YaI}(a_{YaI}-\lvert\xi\rvert)^2}{\lvert\xi\rvert(a_{YaI}-c_{YaI})^2} \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_{YaI}$ $c_{YaI} < \lvert\xi\rvert \le a_{YaI}$ $\lvert\xi\rvert > a_{YaI}$
46	$\begin{cases} \xi \\ \frac{c_{YaII}\cdot\xi}{\lvert\xi\rvert} \\ 0 \end{cases}$	$\lvert\xi\rvert \le c_{YaII}$ $c_{YaII} < \lvert\xi\rvert \le a_{YaII}$ $\lvert\xi\rvert > a_{YaII}$
47	$\frac{\xi}{c_{Ba}^2}\left[\frac{\left(\frac{\xi}{c_{Ba}}\right)^2}{\lvert 2-a_{Ba}\rvert}+1\right]^{\frac{a_{Ba}}{2}-1}$	$\lvert\xi\rvert \le \infty$
48	$\frac{(p_{GT}\cdot q_{GT}+1)\lvert\xi\rvert^{p_{GT}-1}\mathrm{sign}(\xi)}{q_{GT}+\lvert\xi\rvert^{p_{GT}}}$	$\lvert\xi\rvert \le \infty$

Appendix D. Tuning parameters for E_{ff} = 90, 98, and 99%

Table D.7
Tuning parameters for E_{ff} = 90, 98 and 99%.

#	M-Estimators	Tuning Parameters	90%	98%	99%
0	Contaminated Normal	$\begin{cases} b_{CN} \\ p_{CN} \end{cases}$	10 0.514	10 0.09647	10 0.0441

(*continued on next page*)

Table D.7 (*continued*)

#	M-Estimators	Tuning Parameters	90%	98%	99%
3	Fair	c_F	0.6351	3.0421	4.958
4	Huber	c_{Hu}	0.9818	1.7459	2.0102
5	Modified Huber	c_{MH}	0.8590	1.6543	2.0062
6	Kong	c_k	1.221	4.998	7.915
7	LnCosh	c_{LC}	1.2778	0.5475	0.423
8	Logistic	c_L	0.7826	1.8265	2.364
9	Müller	c_{Mu}	1.9127	2.6861	2.951
10	Zhang	c_Z	1.638	0.4002	0.252
11	Alamgir	c_{Al}	2.031	2.819	3.165
12	Bab-Hadiashar	c_{BH}	1.5139	2.1586	2.3948
13	Bell	c_{Bl}	1.6703	2.8928	3.5533
14	Blake-Zisserman	c_{BZ}	$3.9e^{-3}$	$9.3e^{-5}$	$2.0e^{-5}$
15	Cauchy	c_C	1.7249	3.3962	4.2904
16	Correntropy	c_{Co}	1.6851	2.7651	3.35219
17	Danish	c_{Da}	2.4583	3.1180	3.3544
		D_{Da}	3.1268	4.93	5.6845
Tuning parameters for E_{ff} = 90, 98 and 99% - part 2/4					
18	Geman-McClure	c_{GM}	2.8937	5.1597	6.3809
19	GGW	a_{GGW}	1.0283	1.9523	2.3777
		b_{GGW}	1.5	1.5	1.5
		c_{GGW}	0.8709	1.3	1.5
20	Insha	c_I	2.7331	3.8834	4.3882
21	Jin	a_J	0.65	0.65	0.65
		c_J	3.11	4.7735	5.677
22	Kumar	a_{Ku}	1.1156	1.7521	2.0115
		b_{Ku}	2.2311	3.5042	4.0230
		D_{Ku}	3.1111	7.6745	10.1154
23	Lorentzian	c_{Lz}	2.0462	3.6485	4.5120
24	Merril-Schweppe	c_{MS}	1.3209	1.9999	2.2452
25	Ramsay	c_R	0.5266	0.2189	0.1527
26	Sech	c_{Sh}	0.5271	0.2985	0.2418
27	Welsch	c_W	2.3831	3.9104	4.7407
28	Wu	a_{Wu}	0.5	0.5	0.5
		c_{Wu}	2.1320	2.9261	3.1885
29	Xie	c_X	1.6706	2.3409	2.6359
30	Youssef	c_Y	2.1966	3.0050	3.2637
Tuning parameters for E_{ff} = 90, 98 and 99% - part 3/4					
31	Asad	c_{As}	3.1576	4.2103	4.6664
32	Andrews	c_{An}	1.1117	1.6930	2.0170
33	Biweight	c_{Bi}	3.8827	5.9207	7.0414
34	Collins	a_{Cl}	1.5112	1.9016	2.1827
		r_{Cl}	3.5	4.5	5.0
		q_{Cl}	1.3913	1.8759	2.1736
		D_{Cl}	2.8493	5.3769	7.1493
35	Hampel	a_H	1.1052	1.7515	2.0114
		b_H	2.2105	3.5030	4.0228
		c_H	4.4209	7.00609	8.0456
36	Hyperbolic Tangent	A	0.4174	0.8038	0.9173
		B	0.70	0.90	0.95
		k	4.0	5.0	5.5
		r	4.0	5.0	5.5
		q	1.1091	1.7874	2.0290
		D	3.0285	6.12261	7.7731
37	LQQ	b	1.2314	1.8262	1.9945
		c	0.8	1.2	1.4
		s	1.5	1.5	1.5
		a	4.4314	6.6262	7.5945
		D_1	−1.7555	−3.9058	−4.9878
		D_2	3.3206	7.4040	9.5731
Tuning parameters for E_{ff} = 90, 98 and 99% - part 4/4					
38	Modified Asad-Qadir I	a_{AQI}	2.9260	3.7751	4.1102
39	Modified Asad-Qadir II	a_{AQII}	2.8134	3.5855	3.8791
40	Optimal	a_{Op}	0.3279	0.3211	0.3206
		b_{Op}	2.7024	3.4216	3.7025
		c_{Op}	$8.1e^{-4}$	$4.9e^{-4}$	$3.3e{-}4$
		d_{Op}	$5.0e^{-3}$	$3.0e^{-3}$	$2.0e^{-3}$
		e_{Op}	$3.2e^{-3}$	$8.8e^{-3}$	$5.3e^{-3}$
41	Qadir	a_Q	3.8827	5.9207	7.0414
42	Smith	a_S	3.1316	4.4864	5.2317
43	Talwar	a_T	2.5003	3.1365	3.3682
44	Uk	a_{Uk}	3.5890	4.8374	5.3812
45	Yang I	c_{YaI}	2.1622	2.6646	2.8260
		a_{YaI}	3.0	4.0	4.5
		D_{YaI}	2.9414	4.7362	5.5699
46	Yang II	c_{YaII}	1.1018	1.7619	2.0140
		a_{YaII}	3.0	4.0	4.5

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.compchemeng.2021.107254.

CRediT authorship contribution statement

D.Q.F. de Menezes: Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **D.M. Prata:** Conceptualization, Methodology, Writing - review & editing. **A.R. Secchi:** Supervision, Writing - review & editing. **J.C. Pinto:** Supervision, Writing - review & editing.

References

Alamgir, A.A., Khan, S.A., Khan, D.M., Khalil, U., 2013. A new efficient redescending m-estimator: Alamgir redescending m-estimator. Res. J. Recent Sci. 2 (8), 79–91.
Albuquerque, J.S., Biegler, L.T., 1996. Data reconciliation and gross-error detection for dynamic systems. AIChE journal 42 (10), 2841–2856.
Aldrich, J., et al., 1997. Ra fisher and the making of maximum likelihood 1912–1922. Stat. Sci. 12 (3), 162–176.
Alhaj-Dibo, M., Maquin, D., Ragot, J., 2008. Data reconciliation: a robust approach using a contaminated distribution. Control Eng. Pract. 16 (2), 159–170.
Ali, A., Qadir, M.F., 2005. A modified m-estimator for the detection of outliers. Pak. J. Stat. Oper.Res. 1 (1).
Ali, A., Qadir, M.F., Salahuddin, A.Q., 2005. Regression outliers: new m-class ψfunctions based on Winsor's principle with improved asymptotic efficiency. In: Proceedings of the 8 th Islamic Countries conference on Statistical Sciences in. Citeseer.
Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., Tukey, J.W., 1972. Robust Estimates of Location: Survey and Advances. Princeton University Press.
Arora, N., Biegler, L.T., 2001. Redescending estimators for data reconciliation and parameter estimation. Comput. Chem. Eng. 25 (11), 1585–1599.
Arya, K., Gupta, P., Kalra, P.K., Mitra, P., 2007. Image registration using robust m-estimators. Pattern Recognit. Lett. 28 (15), 1957–1968.
Bab-Hadiashar, A., Suter, D., Hesami, R., 2002. Robust model fitting in pattern recognition. In: International Conference on Digital Image Computing Techniques and Applications 2002. Australian Pattern Recognition Society (APRS), pp. 358–363.
Bard, Y., 1974. Nonlinear Parameter Estimation. Academic press.
Barron, J.T., 2019. A general and adaptive robust loss function. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4331–4339.
Baselga, S., 2007. Global optimization solution of robust estimation. J. Surv. Eng. 133 (3), 123–128.
Beaton, A.E., Tukey, J.W., 1974. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics 16 (2), 147–185.
Bell, R.M., 1980. An adaptive choice of the scale parameter for m-estimators.
Bellhouse, D.R., Genest, C., et al., 2007. Maty's biography of abraham de moivre, translated, annotated and augmented. Stat. Sci. 22 (1), 109–136.
Berberan, A., 1995. Multiple outlier detection. a real case study. Surv. Rev. 33 (255), 41–49.
Bourouis, M., Pibouleau, L., Floquet, P., Domenech, S., Al-Gobaisi, D.M.K., 1998. Simulation and data validation in multistage flash desalination plants. Desalination 115 (1), 1–14.
Box, G.E., 1953. Non-normality and tests on variances. Biometrika 40 (3/4), 318–335.
Charbonnier, P., Blanc-Féraud, L., Aubert, G., Barlaud, M., 1997. Deterministic edge-preserving regularization in computed imaging. IEEE Trans. Image Process. 6 (2), 298–311.
Chen, J., Bandoni, A., Romagnoli, J.A., 1998. Outlier detection in process plant data. Comput. Chem. Eng. 22 (4–5), 641–646.
Chen, J., Peng, Y., Munoz, J.C., 2013. Correntropy estimator for data reconciliation. Chem. Eng. Sci. 104, 1019–1027.
Coimbra, J.C., Melo, P.A., Prata, D.M., Pinto, J.C., 2017. On-line dynamic data reconciliation in batch suspension polymerizations of methyl methacrylate. Processes 5 (3), 51.
Collins, J.R., 1976. Robust estimation of a location parameter in the presence of asymmetry. Ann. Stat. 68–85.
da Cunha, A.S., Peixoto, F.C., Prata, D.M., 2020. Robust data reconciliation in chemical reactors. Comput. Chem. Eng. 107170.
da Cunha, A.S., Santos, L.S., Peixoto, F.C., Prata, D.M., 2017. Robust data reconciliation in a chemical reactor through simulated annealing optimization. Latin Am. Appl. Res. 47, 131–136.
de Laplace, P., 1774. Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards. Mém. Acad. Roy. Sci. Paris 6, 353–371.
Dennis Jr., J.E., Welsch, R.E., 1978. Techniques for nonlinear least squares and robust regression. Commun. Stat.-Simul.Comput. 7 (4), 345–359.
do Valle, E.C., de Araújo Kalid, R., Secchi, A.R., Kiperstok, A., 2018. Collection of benchmark test problems for data reconciliation and gross error detection and identification. Comput. Chem. Eng..
Faber, R., Arellano-Garcia, H., Li, P., Wozny, G., 2007. An optimization framework for parameter estimation of large-scale systems. Chem. Eng. Process. Process Intensif. 46 (11), 1085–1095.
Faber, R., Li, B., Li, P., Wozny, G., 2006. Data reconciliation for real-time optimization of an industrial coke-oven-gas purification process. Simul. Modell. Pract. Theory 14 (8), 1121–1134.
Fair, R.C., 1974. On the robust estimation of econometric models. In: Annals of Economic and Social Measurement, Volume 3, number 4. NBER, pp. 667–677.
Feital, T., Prata, D.M., Pinto, J.C., 2014. Comparison of methods for estimation of the covariance matrix of measurement errors. Can. J. Chem. Eng. 92 (12), 2228–2245.
Geman, S., McClure, D.E., 1987. Statistical methods for tomographic image reconstruction. Bull. Int. Stat.Inst. 52 (4), 5–21.
Hald, A., et al., 1999. On the history of maximum likelihood in relation to inverse probability and least squares. Stat. Sci. 14 (2), 214–222.
Hampel, F.R., 1968. Contribution to the theory of robust estimation. University of California, Berkeley Ph.D. thesis.
Hampel, F.R., 1971. A general qualitative definition of robustness. Ann. Math. Stat. 1887–1896.
Hampel, F.R., 1974. The influence curve and its role in robust estimation. J. Am. Stat. Assoc. 69 (346), 383–393.
Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics: the Approach Based on Influence Functions. John Wiley & Sons.
Hampel, F.R., Rousseeuw, P.J., Ronchetti, E., 1981. The change-of-variance curve and optimal redescending m-estimators. J. Am. Stat. Assoc. 76 (375), 643–648.
Hartley, R., Zisserman, A., 2003. Multiple View Geometry in Computer Vision. Cambridge University Press.
Hinich, M.J., Talwar, P.P., 1975. A simple method for robust regression. J. Am. Stat. Assoc. 70 (349), 113–119.
Hoaglin, D.C., Mosteller, F., Tukey, J.W., 1983. Understanding Robust and Exploratory Data Analysis, Vol. 3. Wiley, New York.
Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. Commun. Stat.-Theory Methods 6 (9), 813–827.
Huber, P.J., 1964. Robust estimation of a location parameter. Ann. Math. Stat. 35 (1), 73–101.
Huber, P.J., 1981. Robust Statistics. John Wiley & Sons.
Huber, P.J., 1984. Finite sample breakdown of m-and p-estimators. Ann. Stat. 119–126.
Huber, P.J., Ronchetti, E.M., 2009. Robust Statistics, Vol. 10. A John Wiley & Sons, Inc., Publication.
Huber, P.J., et al., 1973. Robust regression: asymptotics, conjectures and Monte Carlo. Ann. Stat. 1 (5), 799–821.
Jeffreys, H., 1932. An alternative to the rejection of observations. Proc. R. Soc. London Ser.A 137 (831), 78–87.
Jin, S., Li, X., Huang, Z., Liu, M., 2012. A new target function for robust data reconciliation. Ind. Eng. Chem. Res. 51 (30), 10220–10224.
Johnston, L.P., Kramer, M.A., 1995. Maximum likelihood data rectification: steady state systems. AIChE J. 41 (11), 2415–2426.
Jurečkova, J., 1984. Rates of consistency of classical one-side tests. In: Robustness of Statistical Methods and Nonparametric Statistics. Springer, pp. 60–62.
Karal, O., 2017. Maximum likelihood optimal and robust support vector regression with lncosh loss function. Neural Netw. 94, 1–12.
Kendall, M.G., 1961. Daniel bernoulli on maximum likelihood. Biometrika 48 (1), 1–2.
Khalil, U., 2012. A new robust m-estimator with an application to non-stationary time series forecasting. Ph.D. thesis, Pakistan.
Kodamana, H., Huang, B., Ranjan, R., Zhao, Y., Tan, R., Sammaknejad, N., 2018. Approaches to robust process identification: a review and tutorial of probabilistic methods. J. Process Control 66, 68–83.
Koller, M., Stahel, W.A., 2011. Sharpening Wald-type inference in robust regression for small samples. Comput. Stat. Data Anal. 55 (8), 2504–2515.
Kong, M., Chen, B., Li, B., 2000. Simultaneous gross error detection and data reconciliation based on the robust estimation principle. J. Tsinghua Univ. 02.
Korpela, T., Suominen, O., Majanne, Y., Laukkanen, V., Lautala, P., 2016. Robust data reconciliation of combustion variables in multi-fuel fired industrial boilers. Control Eng. Pract. 55, 101–115.
Krarup, T., 1980. Gotterdammerung over least squares adjustment. In: Proc. 14th Congress of the International Society of Photogrammetry, Vol. 3, pp. 369–378.
Kumar, T.A., Rao, K.D., 2009. A new m-estimator based robust multiuser detection in flat-fading non-Gaussian channels. IEEE Trans. Commun. 57 (7), 1908–1913.
Lid, T., Skogestad, S., 2008. Data reconciliation and optimal operation of a catalytic naphtha reformer. J. Process Control 18 (3-4), 320–331.
Lid, T., Skogestad, S., 2008. Scaled steady state models for effective on-line applications. Comput. Chem. Eng. 32 (4–5), 990–999.
Liebman, M.J., Edgar, T.F., Lasdon, L.S., 1992. Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques. Comput. Chem. Eng. 16 (10–11), 963–986.
Lingke, Z., Hongye, S., Jian, C., 2006. A new method to solve robust data reconciliation in nonlinear process. Chin. J. Chem. Eng. 14 (3), 357–363.
Liu, J., Su, Q., Moreno, M., Laird, C., Nagy, Z., Reklaitis, G., 2018. Robust state estimation of feeding–blending systems in continuous pharmaceutical manufacturing. Chem. Eng. Res. Des. 134, 140–153.
Liu, W., Pokharel, P.P., Principe, J.C., 2006. Error entropy, correntropy and m-estimation. In: Machine Learning for Signal Processing, 2006. Proceedings of the 2006 16th IEEE Signal Processing Society Workshop on. IEEE, pp. 179–184.

Llanos, C.E., Sánchez, M.C., Maronna, R.A., 2017. Classification of systematic measurement errors within the framework of robust data reconciliation. Ind. Eng. Chem. Res. 56 (34), 9617–9628.
Llanos, C.E., Sanchéz, M.C., Maronna, R.A., 2015. Robust estimators for data reconciliation. Ind. Eng. Chem. Res. 54 (18), 5096–5105.
Martin, R.D., Simin, T.T., 2003. Outlier-resistant estimates of beta. Financ. Anal. J. 59 (5), 56–69.
McDonald, J.B., Newey, W.K., 1988. Partially adaptive estimation of regression models via the generalized t distribution. Econom. Theory 4 (3), 428–457.
Merrill, H.M., Schweppe, F.C., 1971. Bad data suppression in power system static state estimation. IEEE Trans. Power Apparatus Syst. (6) 2718–2725.
Mingfang, K., Bingzhen, C., Bo, L., 2000. An integral approach to dynamic data rectification. Comput. Chem. Eng. 24 (2–7), 749–753.
Özyurt, D.B., Pike, R.W., 2004. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Comput. Chem. Eng. 28 (3), 381–402.
Pai, C.C., Fisher, G.D., 1988. Application of broyden's method to reconciliation of nonlinearly constrained data. AIChE J. 34 (5), 873–876.
Pennacchi, P., 2008. Robust estimate of excitations in mechanical systems using m-estimators theoretical background and numerical applications. J. Sound Vibr. 310 (4–5), 923–946.
Prata, D.M., 2009. Reconciliação robusta de dados para monitoramento em tempo real. Ph.D. thesis, Rio de Janeiro, RJ.
Prata, D.M., Lima, E.L., Pinto, J.C., 2008. In-line monitoring of bulk polypropylene reactors based on data reconciliation procedures. Macromol. Symp. 271 (1), 26–37.
Prata, D.M., Lima, E.L., Pinto, J.C., 2009. Nonlinear dynamic data reconciliation in real time in actual processes. Comput. Aided Chem. Eng. 27, 47–54.
Prata, D.M., Pinto, J.C., Lima, E.L., 2008. Comparative analysis of robust estimators on nonlinear dynamic data reconciliation. Comput. Aided Chem. Eng. 25, 501–506.
Prata, D.M., Schwaab, M., Lima, E.L., Pinto, J.C., 2009. Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: application for an industrial polypropylene reactor. Chem. Eng. Sci. 64 (18), 3953–3967.
Prata, D.M., Schwaab, M., Lima, E.L., Pinto, J.C., 2010. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor. Chem. Eng. Sci. 65 (17), 4943–4954.
Ragot, J., Chadli, M., Maquin, D., 2005. Mass balance equilibration: a robust approach using contaminated distribution. AIChE J. 51 (5), 1569–1575.
Ramsay, J.O., 1977. A comparative study of several robust estimates of slope, intercept, and scale in linear regression. J. Am. Stat. Assoc. 72 (359), 608–615.
Reilly, P., Carpani, R., 1963. Application of statistical theory of adjustment to material balances. In: Proceedings of the 13th Canadian Chemical engineering Conference, pp. 21–23.
Rey, W., 1983. Introduction to Robust and Quasi-Robust Statistical Methods. Springer Science & Business Media.
Ronchetti, E., Trojani, F., 2001. Robust inference with GMM estimators. J. Econom. 101 (1), 37–69.
Rousseeuw, P., Leroy, A., 1987. Robust Regression and Outlier Detection.
Rousseeuw, P.J., Croux, C., 1992. Explicit scale estimators with high breakdown point. L1-Stat. Anal. Relat. Methods 1, 77–92.
Rousseeuw, P.J., Croux, C., 1993. Alternatives to the median absolute deviation. J. Am. Stat. Assoc. 88 (424), 1273–1283.
Schladt, M., Hu, B., 2007. Soft sensors based on nonlinear steady-state data reconciliation in the process industry. Chem. Eng. Process. Process Intensif. 46 (11), 1107–1115.
Shevlyakov, G., Morgenthaler, S., Shurygin, A., 2008. Redescending m-estimators. J. Stat. Plann. Inference 138 (10), 2906–2917.
Smith, R.H., 1888. True average of observations? Nature 37 (959), 464.
Stewart, C.V., 1999. Robust parameter estimation in computer vision. SIAM Rev. 41 (3), 513–537.
Stigler, S.M., 1973. Simon newcomb, percy daniell, and the history of robust estimation 1885–1920. J. Am. Stat. Assoc. 68 (344), 872–879.
Stigler, S.M., 1974. Studies in the history of probability and statistics. XXXIII Cauchy and the witch of Agnesi: an historical note on the cauchy distribution. Biometrika 375–380.
Stigler, S.M., 1980. Studies in the history of probability and statistics XXXVIII: R.H. Smith, a victorian interested in robustness. Biometrika 67 (1), 217–221.
Stigler, S.M., 1984. Studies in the history of probability and statistics XL Boscovich, Simpson and a 1760 manuscript note on fitting a linear relation. Biometrika 71 (3), 615–620.
Stigler, S.M., 2010. The changing history of robustness. Am. Stat. 64 (4), 277–281.
Stromberg, A., et al., 2004. Why write statistical software? The case of robust statistical methods. J. Stat. Softw. 10 (5), 1–8.
Su, Q., Bommireddy, Y., Shah, Y., Ganesh, S., Moreno, M., Liu, J., Gonzalez, M., Yazdanpanah, N., O'Connor, T., Reklaitis, G.V., et al., 2019. Data reconciliation in the quality-by-design (QbD) implementation of pharmaceutical continuous tablet manufacturing. Int. J. Pharmaceutics 563, 259–272.
Tabatabai, M., Kengwoung-Keumo, J., Eby, W., Bae, S., Manne, U., Fouad, M., Singh, K., 2014. A new robust method for nonlinear regression. J. Biom. Biostat. 5 (5), 211.
Tjoa, I.-B., Biegler, L.T., 1991. Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems. Comput. Chem. Eng. 15 (10), 679–690.
Tukey, J.W., 1960. A survey of sampling from contaminated distributions. Contrib. Probab. Stat. 2, 448–485.
Tukey, J.W., 1962. The future of data analysis. Ann. Math. Stat. 33 (1), 1–67.
Ullah, I., Qadir, M.F., Ali, A., 2006. Insha's redescending m-estimator for robust regression: a comparative study. Pak. J. Stat. Oper.Res. 2 (2), 135–144.
Valdetaro, E.D., Schirru, R., 2011. Simultaneous model selection, robust data reconciliation and outlier detection with swarm intelligence in a thermal reactor power calculation. Ann. Nucl. Energy 38 (9), 1820–1832.
Valluru, J., Patwardhan, S.C., Biegler, L.T., 2018. Development of robust extended Kalman filter and moving window estimator for simultaneous state and parameter/disturbance estimation. J. Process Control 69, 158–178.
Verhulst, P.F., 1838. Notice sur la loi que la population poursuit dans son accroissement. Correspondance Mathématique et Physique 10, 113–121.
Vichare, N.S., 1993. Robust Mahalanobis distances in power system state estimation. Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA Ph.D. thesis.
Van de Vusse, J., 1964. Plug-flow type reactor versus tank reactor. Chem. Eng. Sci. 19 (12), 994–996.
Wieser, A., Brunner, F.K., 2001. Robust estimation applied to correlated GPS phase observations. First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS, Zurich, Switzerland.
Williams, T.J., Otto, R.E., 1960. A generalized chemical processing model for the investigation of computer control. Trans. Am. Inst.Electr. Eng. Part I 79 (5), 458–473.
Wongrat, W., Srinophakun, T., Srinophakun, P., 2005. Modified genetic algorithm for nonlinear data reconciliation. Comput. Chem. Eng. 29 (5), 1059–1067.
Wu, H.-y., Wu, L.-d., 2005. A robust estimation method in orbit improvement. Chin. Astron. Astrophys. 29 (4), 430–437.
Wu, S., Xu, J., Liu, W., Wu, X., Gu, X., 2017. Data reconciliation based on an improved robust estimator and NT-MT for gross error detection. In: Advanced Computational Methods in Life System Modeling and Simulation. Springer, pp. 400–409.
Xie, S., Yang, C., Yuan, X., Wang, X., Xie, Y., 2019. A novel robust data reconciliation method for industrial processes. Control Eng. Pract. 83, 203–212.
Yang, Y., 1994. Robust estimation for dependent observations. Manuscripta Geod. 19 (1), 10–17.
Yang, Y., 1999. Robust estimation of geodetic datum transformation. J. Geodesy 73 (5), 268–274.
Yohai, V.J., Zamar, R.H., 1997. Optimal locally robust m-estimates of regression. J. Stat. Plann. Inference 64 (2), 309–323.
Youssef, A.H., El-Sheikh, A.A., Mohammed, E.T.H., 2013. New m-estimator objective function in simultaneous equations model (a comparative study). In: International Mathematical Forum, vol. 8. Citeseer, pp. 1007–1022.
Zhang, J., Li, G., et al., 1998. Breakdown properties of location *m*-estimators. Ann. Stat. 26 (3), 1170–1189.
Zhang, Z., Chen, J., 2015. Correntropy based data reconciliation and gross error detection and identification for nonlinear dynamic processes. Comput. Chem. Eng. 75, 120–134.
Zhang, Z., Chuang, Y.-Y., Chen, J., 2014. Methodology of data reconciliation and parameter estimation for process systems with multi-operating conditions. Chemom. Intell. Lab. Syst. 137, 110–119.
Zhang, Z., Pike, R.W., Hertwig, T.A., 1995. Source reduction from chemical plants using on-line optimization. Waste Manage. 15 (3), 183–191.
Zhang, Z., Shao, Z., Chen, X., Wang, K., Qian, J., 2010. Quasi-weighted least squares estimator for data reconciliation. Comput. Chem. Eng. 34 (2), 154–162.
Zoubir, A.M., Koivunen, V., Ollila, E., Muma, M., 2018. Robust Statistics for Signal Processing. Cambridge University Press.