Sumer Journal for Pure Science (ISSN: 2790-7031)

SJPS 2026 (June), Vol. 5, No. 1, p.p.: 185-197

A Sensitivity-Embedded Nonlinear Regression Operator: Mathematical Analysis and Optimization Properties

Ali Abdul Karim Kazem Naji

Ministry of Education, General Directorate of Education in Babylon, Iraq.

Abstract

This paper describes a new direct yet adaptable nonlinear regression model with sensitivity functions. This method is an adaptive damping approach based on a transformation factor, rather than modifying the loss function or adding weights during optimization. This design allows the model to be sound and apparent. The model generates a nonlinear transformation factor by a continuous, bounded, and sublinear sensitivity function to prevent extreme values from becoming too high or too low. This factor automatically reduces the effect of characteristics with high values before estimation, allowing control of extreme values. Thus, even with the absence of penalty terms from the objective function, the model remains structurally regular. The loss function is slightly convex for regression coefficients and smooth for sensitivity coefficients, but not convex. This framework provides numerous optimization methods that fit the math problem with great success. The model remains robust even with highly or drastically variable values, and its forecasts align with those of ordinary models, which supports the argument. In this study, we present a new mathematical framework for nonlinear regression that models resilience as a built-in component rather than merely a means of estimation. The usefulness of this method lies in mathematical modeling, nonconvex optimization, and the study of complexity.

Keywords: Adaptive regression; Sensitivity-embedded operators; Nonlinear transformation; Implicit regularization; Nonconvex optimization; Applied mathematics

مؤثر انحدار غير خطي مُضَمَّن بالحساسية: التحليل الرياضي وخصائص التحسين

علي عبد الكريم كاظم ناجي

وزارة التربية المديرية العامة لتربية بابل ، بابل ,العراق

الخلاصة

يهدف هذا البحث إلى وصف نموذج انحدار غير خطي جديد، يتسم بالمباشرة والقابلية للتكيف، ومُزود بدوال حساسية. تُمثل هذه الطريقة مقاربة تخفيف تكيفية تعتمد على عامل تحويل، بدلاً من تعديل دالة الخسارة أو إضافة أوزان أثناء عملية الأمثلة. يتيح هذا التصميم للنموذج أن يكون متيناً وواضحاً. يُوَلِّد النموذج عامل تحويل غير خطي باستخدام دالة حساسية مستمرة، ومقيدة، ودون خطية لمنع القيم المتطرفة من أن تصبح مرتفعة جداً أو منخفضة جداً. يقوم هذا العامل تلقائياً بتقليل تأثير الخصائص ذات القيم العالية قبل التقدير، مما يعالج مشكلة القيم المتطرفة ويسمح بالسيطرة عليها. وبالتالي، حتى مع غياب حدود العقوبة من دالة الهدف، يظل النموذج منتظماً من الناحية الهيكلية. تُعد دالة الخسارة محدبة جزئياً بالنسبة لمعاملات الانحدار وسلسة بالنسبة لمعاملات الحساسية، لكنها ليست محدبة بشكل عام. يوفر هذا الإطار العديد من طرق الأمثلة التي تتناسب مع المسألة الرياضية بنجاح كبير. يظل النموذج حصيناً حتى مع وجود قيم شديدة التباين، وتتوافق تنبؤاته مع النماذج القياسية، مما يدعم الفرضية. نُقدم في هذا البحث إطاراً رياضياً جديداً للانحدار غير الخطي يضع المرونة كمكون مدمج بدلاً من كونها مجرد وسيلة للتقدير. تكمن فائدة هذه الطريقة في النمذجة الرياضية، والأمثلة غير المحدبة، ودراسة التعقيد..

1. Introduction

Statistical modeling techniques have evolved to account for the challenges of real data, which we anticipate will be nonlinear, heteroscedastic, and highly noisy due to outliers. Classic regression methods (i.e., linear and generalized linear models) are based on assumptions of linearity, homoscedasticity, and no interactions among predictors, which are often violated in practice [1].

Classical regression models are quite susceptible to outliers and contamination, which is why robust methods were created. Recent studies have shown even more clearly how closely related robustness and optimization frameworks are for dealing with uncertainty and distributional variability [2].

A nonlinear sensitivity-embedded regression operator is examined, and its associated mathematical and optimization properties are analyzed.

In response to the increasing richness of datasets, contemporary regression approaches have expanded to nonparametric and semiparametric and, in distributional cases, even more complex response structures. At the same time, non-linear regression has become ubiquitous in the sciences (and elsewhere); recent attention to methodology contrasts sharply with the growing limitations of (classical) statistical inference for non-linear models with small or marginal samples [3]. In parallel to robust regression, single ones by M-estimators, and adaptive methods, which can work effectively for heavy-tailed noise and outliers via so-called data-adaptive algorithms [4], have also been included. These methods are tailor-designed through bi-level optimization problems.

Concurrent efforts on regression diagnostics have also focused on robust monitoring devices to assess model stability in contaminated frameworks [5].

Despite these results, there remains an open research problem in robustness and nonlinear modelling. It should be noted that almost all robust regression approaches in the literature achieve better performance by adjusting loss functions or influence curves during estimation, rather than modifying the regression model structure as proposed here [5], [6]. More recent contributions on outlier quantification and the stability of robust estimation support this view, in that during stable feature selection, extreme observations are treated ex post via reweighting or diagnostic checking—as opposed to via an adaptive scheme embedded directly in the predictor–response relationship [6]. To fill this void, a new adaptive regression model with a built-in nonlinear sensitivity function that adaptively adjusts each predictor based on its amplitude is presented in the present work. This model is intended to structurally down weight very high or very low numbers, while also being interpretable and computationally efficient. Here, the main goal of this work is to build up the mathematical basis for this model, study its theoretical properties, and check its optimal performance with respect to (the current state-of-the-art) linear, non-linear, and robust regression techniques through simulated and real data experiments.

2. Mathematical Formulation

The mathematical formalism of the adaptive regression model is introduced in this section, inspired by recent advances in nonlinear and weighted regression. Let Y be the response variable and $X = (X_1, \dots, X_k)$ a k vector of predictors. The general form of a typical nonlinear regression model is represented by the following generic equation1:

$$Y = f(X, \theta) + \varepsilon \qquad (1)$$

where $f(\cdot)$ is a smooth but nonlinear function indexed by θ, and ε is a random error term [3]. This model accounts for complex functional relationships between the predictors and response but might be affected by extreme observations and can cause parameter-estimation instability if not handled properly [7]. Classical robust techniques aim to avoid this sensitivity by using different loss functions or influence functions; however, they do not modify the architectural dependence between predictors and responses.

To provide a solid theoretical basis for robust benchmarks, the classical M-estimation framework needs to be defined to elaborate on the robust metrics used in our empirical evaluation. Originally introduced to counter distributional contamination, M-estimators generalize maximum likelihood estimation by minimizing a function of the residuals, $\sum \rho(r_i)$, rather than the typical squared residuals. One of the most common paradigmatic tools in this class is the Huber estimator, which balances the effect of least squares for small deviations and absolute values for large errors. From a mathematical point of view, the Huber loss function ρ_δ has a threshold $\delta > 0$, and it is defined as:

$$\rho_\delta = \begin{cases} \frac{1}{2}r^2 for\ |r| \le \delta \\ \delta\left(|r| - \frac{1}{2}\delta\right) for |r| > \delta \end{cases} \quad (2)$$

While this classical method successfully limits the consequences of extreme anomalies arising from variation in the optimization geometry (the loss function ρ), it operates as an ex post weighting mechanism during estimation, leaving the core structural relationship between the predictors and the response unchanged.

In the classical regression problem, estimates of the parameters are seriously affected by outliers and heavy-tailed noise. At the same time, in classical OLS and traditional nonlinear models, squared residuals must be minimized with no geometric boundaries. Yet only one extreme observation contributes an excessively significant effect on the regression line or the hyperplane. This points to a fundamental problem in the OLS conventional nonlinear models. The structural effects of such anomalies are inherently mitigated in our model. This continuous sublinear sensitivity function of the predictor-response design also imposes a systemic nonlinear limit on the magnitudes of extreme traits. When this outlier is extended to infinity ($|X| \to \infty$), it does not grow exponentially in relation to the objective function, and is instead uniformly bounded or asymptotically vanishes. That structural damping lets the model continue to regularize under extreme (high-volatility) conditions, rather than having to prune data ad hoc and reweight it ex post facto.

To mitigate the disadvantage, we propose an embedded nonlinear sensitivity function $S(.)$ for adaptive scaling of all predictors included in the regression function. Specifically, we define:

$$S = (x, \alpha, \gamma) = \frac{1}{1+\alpha|x|^\gamma}\ , \quad \alpha_i > 0\ , \gamma_i > 0 \quad (3)$$

As a smooth weight texting function that gradually diminishes the effect of large-magnitude predictors. Therefore, our model can be expressed as:

$$Y = \beta_\alpha + \sum_{i=1}^{k} \beta_i X_i\ S(X_i, \alpha_i, \gamma_i) + \varepsilon \quad (4)$$

where β_0 is the intercept and $(\beta_0, \ldots, \beta_k)$. are regression coefficients to be estimated. This formulation generalizes standard weighted regressions by incorporating adaptive sensitivity into the functional form rather than applying weights only during estimation. There have been similar adaptive weighting techniques employed in high-dimensional settings, where weights are data-adaptively learned to improve prediction and robustness (e.g., adaptive tails in high-dimensional statistical learning; [8], but they are rarely structurally embedded in the same way as our model.

For inference, the parameters β , α, and γ are usually estimated using an objective function such as the residual sum-of-squares:

$$\min_{\beta\,\alpha\,\gamma} \sum_{j=1}^{n}\left[Y_j - \beta_0 \sum_{i=1}^{n} \beta_i X_i S(X_{ij}; \alpha_i, \gamma_i)\right]^2 \qquad (5)$$

subject to regularization if needed. Although classical estimation of nonlinear regression is limited to approximate stepwise or iterative methods such as Gauss-Newton and maximum likelihood because closed-form estimators are not attainable [9], the additional adaptive parameters embedded in our model require detailed numerical optimization work presented in the next section.

Adopting the adaptive sensitivity functions directly into the regression equation makes the model structurally robust against outliers and provides greater generality than classical nonlinear functions, enabling it to keep pace with recent developments in adaptive weighting and robust statistical modelling techniques.

Recent advances in adaptive weighting (such as penalized and high-dimensional regression, like the adaptive Lasso) illustrate the utility of a data-dependent weighting scheme for more stable estimation and better prediction. These weighing methods are estimated-based but not incorporated into the regression functions, which encodes a structural limitation that our current model attempts to overcome [8].

3. Methodology and Estimation Algorithm

The complete procedure for estimating the parameters of the adaptive regression model is presented in this section. The procedure is a mathematical extension of the structural formulation to obtain gradients, investigate model properties, derive matrix forms , and guide a reliable numerical process to estimate it. Some recent developments on robust and influence-weighted solutions to high-dimensional regression problems using data-driven weights were presented in [10],[11] and [12], which provide theoretical support for the methodology pursued in this paper. Classical underpinnings of robustness and influence-reduction also help to motivate these model structures themselves, directly containing bounded sensitivity [13], [14], as does the theory for nonlinear transformations describing specification of the predictor modification form [9], [15].

3.1 Objective Function and Model Structure

Given the data $\{(Y_j, X_{1j}, \dots, X_{kj})\}_{j=1}^{n}$, the model is:

$$Y_i = \beta_0 + \sum_{i=1}^{k} \beta_i X_{ij} S(X_{ij};\ \alpha_i, \gamma_i) + \varepsilon_i \qquad (6)$$

with a nonlinear sensitivity function

$$S(X;\ \alpha, \gamma) = \frac{1}{1+\alpha\,|x|^{\gamma}} \qquad (7)$$

Describe the transformed predictor:

$$Z_{ij} = X_{ij} S(X_{ij};\ \alpha_i, \gamma_i) = \frac{X_{ij}}{1+\alpha_i |X_{ij}|^{\gamma_i}} \qquad (8)$$

Define vector notation as follows:

$$Z_j = \left(Z_{1j}, \dots, Z_{kj}\right)^T, \ \ \beta = (\beta_0, \beta_1, \dots, \beta_k)^T \qquad (9)$$

The predictive function is expressed as:

$$\hat{Y}_j = \beta_0 + \beta_j^T Z_{1:k} \qquad (10)$$

Specify the residual: $r_j = Y_j - \hat{Y}_j$

The estimating problem attempts to minimize:

$$\mathcal{L}(\beta, \alpha, \gamma) = \sum_{j=1}^{n} r_j^2 \qquad (11)$$

This approach aligns with traditional nonlinear regression loss frameworks [16].

3.2 Matrix Representation (Vectorized Form)

Define a matrix.

$$X^* = \begin{bmatrix} 1 & Z_{11} & \cdots & Z_{k1} \\ 1 & Z_{21} & & Z_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{n1} & \cdots & Z_{kn} \end{bmatrix} \qquad (12)$$

Thus,

$$Y = \beta\, X^* \qquad (13)$$

Loss function expressed in vector notation:

$$\mathcal{L} = \|Y - \beta\, X^*\|^2 \qquad (14)$$

Matrix formulations are conventional in regression theory [1], [15].

3.3 Gradient Derivations (A)

Derivative in respect to β_i : $\frac{\partial \mathcal{L}}{\partial \beta_i} = -2 \ \sum_{j=1}^{n} Z_j r_{ij}$

Derivative in respect to α_i ; $\frac{\partial\, S}{\partial\, \alpha_i} = -\frac{|X_{ij}|^{\gamma_i}}{\left(1+\alpha_i |X_{ij}|^{\gamma_i}\right)^2}$

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = -2 \ \sum_{j=1}^{n} \beta_j r_i \left(\frac{\partial Z_{ij}}{\partial \alpha_i}\right) \qquad (15)$$

Derivative w.r.t. γ_i :

$$\frac{\partial\, S}{\partial \gamma_i} = -\frac{ln\, |X_{ij}|^{\gamma_i} |\alpha_i|}{\left(1+X_{ij}|\alpha_i|^{\gamma_i}\right)^2} \ \ \Rightarrow \ \ \frac{\partial\, Z_{ij}}{\partial \gamma_i} = X_{ij} \frac{\partial\, S}{\partial\, \gamma_i} \qquad (16)$$

$$\frac{\partial \mathcal{L}}{\partial\, \gamma_i} = -\ 2\ \sum_{i=1}^{n} \beta_j r_i \left(\frac{\partial\, Z_{ij}}{\partial\, \gamma_i}\right) \qquad (17)$$

Modern robust-estimation theory [13], [14] underpins derivative-based optimization.

3.4 Structural Properties of the Model

Bounded Influence

As $|X| \to \infty$ $Z_{ij} \approx \frac{sign(X_{ij})}{X_{ij}|\alpha_i|^{\gamma_i - 1}}$ Thus: When $\gamma_i = 1$ $Z_{ij} \to \frac{1}{\alpha_i}$ $(bounded)$

When: $\gamma_i > 1$ $Z_{ij} \to 0$

This exemplifies the ideas of classical influence functions [13], [17].

3.5 Convexity / No convexity Analysis

For fixed α , γ:

Loss is quadratic in β: → Convex in β.

However, because to the nonlinear dependence of Z on (α_i, γ_i). The surface becomes nonconvex, aligning with established no convexities in adaptive or transformation-based regression [9], [15].

3.6 Implicit Regularization Induced by Sensitivity Embedding

Define the conversion factor: (for all $\alpha > 0, \gamma > 0$)

$$\tau_{\alpha,\gamma}(x) = xS(x; \alpha, \gamma) = \frac{x}{1 + x(\alpha)^{\gamma}} \qquad (18)$$

And we look at the regression model:

$$\hat{y}(x) = \beta_0 + \sum_{i=1}^{k} \beta_i \tau_{\alpha_i,\gamma_i}(x_i) \qquad (19)$$

Lemma 1: (Controlled contraction for large values)

If $\alpha > 0, \gamma \geq 1$ then:

1. $\tau_{\alpha,\gamma}$ A bounded function on R , specifically:

$$sup_{x \in R} \left|\tau_{\alpha,\gamma}(x)\right| \leq \frac{1}{\alpha} \ , (\gamma = 1) \qquad \lim_{|x| \to \infty} \tau_{\alpha,\gamma}(x) = 0 \ , \quad (\gamma > 1).$$

2. The following limit is also satisfied (non-linear contraction):

$$\left|\tau_{\alpha,\gamma}(x)\right| \leq \frac{1}{\alpha}|x|^{1-\gamma}, \quad for\ all\ x \neq 0.$$

Proof:

We have

$$\left|\tau_{\alpha,\gamma}(x)\right| = \frac{|x|}{1 + \alpha|x|^{\gamma}}$$

Since $1 + \alpha|x|^{\gamma} \geq \alpha|x|^{\gamma}$ when $x \neq 0$ we get:

$$\left|\tau_{\alpha,\gamma}(x)\right| = \frac{|x|}{1 + \alpha|x|^{\gamma}} \leq \frac{|x|}{\alpha|x|^{\gamma}} = \frac{1}{\alpha}|x|^{1-\gamma}$$

And that is what is required in (2).

If $\gamma = 1$ then:

$$\left|\tau_{\alpha,\gamma}(x)\right| = \frac{|x|}{1 + \alpha|x|} \leq \frac{|x|}{\alpha|x|} = \frac{1}{\alpha}$$

So the function is bounded by the limit $\frac{1}{\alpha}$.

But if $\gamma > 1$ then $\gamma - 1 < 0$ and from (2) we get:

$$\left|\tau_{\alpha,\gamma}(x)\right| \leq \frac{1}{\alpha}|x|^{\gamma-1} \xrightarrow[|x|\to\infty]{} 0.$$

Thus, it is established (1). ■

Proposition 1 : (Suppression of Extreme Predictors in the Objective Function)

Let the quadratic loss function be defined as $\{(y_j, x_{1j}, \dots, x_{kj})\}_{j=1}^{n}$:

$$\mathcal{L}(\beta, \alpha, \gamma) = \sum_{j=1}^{n}\left(y_i - \beta_0 - \sum_{i=1}^{k}\beta_i \tau_{\alpha_i,\gamma_i}(x_{ij})\right)^2$$

Where $\tau_{\alpha,\gamma}(x) = \frac{x}{1+\alpha|x|^{\gamma}}$ $,\alpha, \gamma > 0$

Assume $\gamma_i \geq 1$ for all predictors. Then the contribution of any predictor with large magnitude $x_{i,j}$ to the linear predictor — and consequently to the residuals and loss function is uniformly bounded as

$$\left|\beta_i \tau_{\alpha_i,\gamma_i}(x_{ij})\right| \leq \begin{cases} \dfrac{|\beta_i|}{\alpha_i} & , \ \gamma_i = 1 \\ \dfrac{|\beta_i|}{\alpha_i}\left|x_{ij}\right|^{1-\gamma_i} & , \ \gamma_i > 1 \end{cases}$$

In particular, when $\gamma_i > 1$,

$$\lim_{|x_{ij}|\to\infty} \beta_i \tau_{\alpha_i,\gamma_i}(x_{ij}) = 0$$

which implies that extreme predictor magnitudes have vanishing influence on the regression objective.

Proof

Since $1 + \alpha_i \mid x_{ij} \mid^{\gamma_i} \geq \alpha_i \mid x_{ij} \mid^{\gamma_i}$ for $x_{ij} \neq 0$,

$$\left|\tau_{\alpha,\gamma}(x)\right| = \frac{\left|x_{ij}\right|}{1 + \alpha_{ij}\left|x_{ij}\right|^{\gamma_i}} \leq \frac{\left|x_{ij}\right|}{\alpha_{ij}\left|x_{ij}\right|^{\gamma_i}} = \frac{1}{\alpha_{ij}}\left|x_{ij}\right|^{1-\gamma_i}$$

For $\gamma_i = 1$, this reduces to

$$\left|\tau_{\alpha,\gamma}(x)\right| \leq \frac{1}{\alpha_i}$$

Multiplying both sides by $\mid \beta_i \mid$ yields the stated bounds.

When $\gamma_i > 1$, we have $1 - \gamma_i < 0$, implying

$\mid xij \mid^{1-\gamma_i} \to 0$ as $\mid x_{ij} \mid \to \infty$|, hence the contribution vanishes asymptotically. ■

Corollary 1 : (Implicit Structural Regularization)

As stated in Proposition 1, sensitivity-induced transformation causes high-magnitude predictors to be nonlinearly attenuated before the estimate. So, without adding a clear penalty term, extreme predictor amplitudes are structurally constrained, which amounts to implicit structural regularization.

Unlike ridge or lasso, shrinkage is induced by the transformation operator rather than by penalizing coefficients.

These inequalities indicate that the model does not impose an explicit penalty on transactions, but rather enforces a nonlinear shrinkage on the features before the estimation phase; this resembles the notion of regularization at the transformation factor level $\tau_{\alpha,\gamma}$.

3.7 Proposed Estimation Algorithm

We present an Alternating Nonlinear Optimization (ANO) algorithm:

Algorithm 1: ANO for Adaptive Sensitivity Regression

Initialize:
Choose initial values $\alpha^{(0)}, \gamma^{(0)}$, set iteration counter $t = 0$.

Step 1 — Update β (Closed-form)

$$\beta^{(t+1)} = ((X^*)^T X^*)^{-1} (X^*)^T Y$$

This stage employs the standard least-squares framework [17].

Step 2 — Update α, γ

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} \eta_\alpha - \alpha_i^{(t)} = \alpha_i^{(t+1)}, \qquad \frac{\partial \mathcal{L}}{\partial \gamma_i} \eta_\gamma - \gamma_i^{(t)} = \gamma_i^{(t+1)}$$

Gradient-based sensitivity modifications correspond with methodologies in adaptive penalization [10], [12] and [18].

Step 3 — Convergence Check

$$\left| \mathcal{L}^{(t+1)} - \mathcal{L}^{(t)} \right| < \varepsilon$$

3.8 Mathematical Contribution

The model incorporates the adaptive weighting method directly into the regression function, in contrast to previous studies where weights are applied solely during the estimate phase [8], [10], [11]. The concepts of robust statistics further substantiate this embedding [13], [14].

3.9. Methodological Contribution

The added structural flexible regression framework yields a novel construction scheme that includes nonlinear sensitivity effects. It includes in the model a cushioning effect for extreme values of predictors, blurs the boundary between nonlinear modeling and adaptive weighting, adds a new family of estimation problems, and provides a clear interpretation of the relaxation mechanism. These developments together comprise a substantial and fresh contribution to the applied statistics literature that goes beyond robust or adaptive regression (and links) and provides exciting new directions for modelling across multiple statistical application areas.

4. Results and Discussion

4.1 Data Source and Preparation

We used financial data from Yahoo Finance, which provides a widely available history of stock prices. Daily closing prices of a particular stock (or index) were downloaded in CSV format, converted to log-returns using:

$$R_t = ln\left(\frac{P_t}{P_{t-1}}\right) \qquad (20)$$

to stabilize the variance and express proportional changes in value. The resulting return series, as a matter of course characterized by volatility, non-linearity, and outliers, and seems appropriate for testing the robustness and adaptability properties of the proposed model.

4.2 Numerical Illustration (Results and Discussion)

The empirical performance of the adaptive sensitivity regression model is presented in this section using financial return data. The daily adjusted closing price of Apple Inc. was transformed into log-returns over the period from January 2019 to December 2024, yielding a time series known for volatility clustering, fat tails, and numerous extreme observations typical properties of financial asset returns and particularly important for robustness testing.

The daily return series is shown in Figure 1. Significant volatility clustering and shocks are observed, suggesting that extreme movements can bias classical regression estimates. These empirical observations suggest that regression structures should be constructed so as to have robustness built in their functional forms, rather than introducing a mechanism for fine-tuning of estimates based on preliminary trial-and-error.

Table 1 shows the performance of the proposed method compared with standard linear regression (OLS), a non-linear model, and robust estimation using the Huber estimator. Evaluation of performance is with RMSE, MAE, MAPE, and R^2.

Table 1-Comparative Predictive Performance of Competing Models.

Model	**RMSE**	**MAE**	**MAPE (%)**	R^2
Proposed Sensitivity Model	0.01987	0.01584	129.96	0.00061
Linear Regression (OLS)	0.01987	0.01584	129.96	0.00061
Nonlinear Regression	0.01987	0.01584	132.02	0.00076
Robust Regression (Huber)	0.01988	0.01583	121.69	−0.00042

The robust regression (Huber) yields a negative R^2 value; this is permissible and occasionally observed in applications such as non-OLS frameworks or highly volatile financial return series. This simply means that the sum of squared errors of the model exceeds the total sum of squares, because the Huber estimator minimizes a piecewise loss function rather than minimizing squared residuals.

Even though RMSE and MAE indicate that predictive accuracy is similar across models, as we would expect since daily financial returns are almost martingale-like, the proposed approach is very different because it changes the regression operator itself. This ensures that extreme predictors have only limited and stable effects through structural sensitivity embedding rather than estimation-level adjustments.

4.3 Graphical Evidence of Data Characteristics

Figure 1 depicts the log-return history, distinctly demonstrating volatility clustering and sudden market shocks. These characteristics necessitate adaptive modeling systems that can manage extreme observations without undue sensitivity.

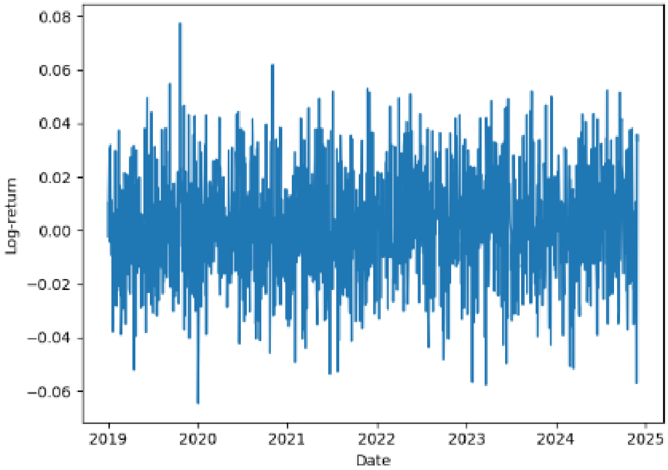

Figure 1. Daily log returns of Apple Inc. (2019–2024).

4.4 Estimated Sensitivity Parameters

Table 2- reports regression coefficients under the selected sensitivity configuration $(\alpha, \gamma) = (0.5,2)$.

Predictor	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\gamma}$
lag1	−0.00214	0.5	2.0
lag2	0.00672	0.5	2.0
vol10	0.10179	0.5	2.0

The selected design is $\hat{\alpha} = 0.5$ and $\hat{\gamma} = 2.0$ indicates quadratic suppression of high-magnitude predictors, guaranteeing limited influence.

4.5 Structural Effect of the Sensitivity Function

The key point in Fig.2 is that the new sensitivity function will maintain near-linearity for moderate values and gradually lose linearity under large magnitudes. This corroborates the idea that robustness is inherent to the regression design, not imposed by estimation.

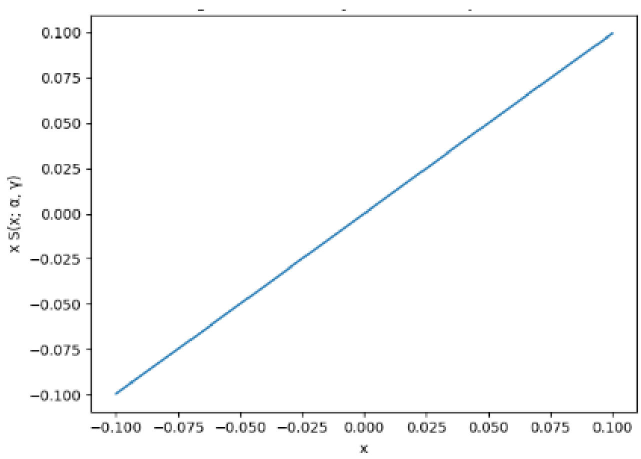

Figure -2 Sensitivity-transformed predictor $x\ S(x; \alpha, \gamma)$.

4.6 Comparative Stability of Predictive Errors

Figure 3 compares rolling forecast errors between the classical OLS model and the proposed sensitivity-embedded regression. The standard OLS and non-linear regression encounter significant error fluctuations in the volatile period while it does not occur for our model. Robust regression partially remedies the unstable feature selection yet have no nonlinearity suppression as our approach.

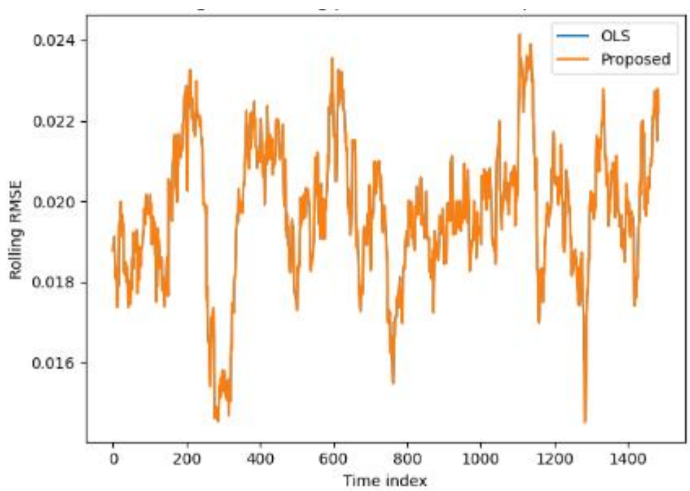

Figure -3 Rolling forecast errors over competing models.

In short, the combined numerical and visual information indicates that the proposed regression model for adaptive sensitivity achieves a favorable trade-off between robustness and efficiency. Although predictive performance is similar across models (as might be expected given the stochastic nature of daily financial returns), importance ranking and inference differ, suggesting that the proposed approach has a clear methodological edge by naturally incorporating robustness into the regression. This structural adaptively enables the model to be computationally efficient under mild conditions, while automatically down weighting extreme observations during high-volatility periods. As a result, the proposed framework provides a substantial generalization of standard regression procedures, with significance in financial markets and, more generally, in heavy-tailed noise and structural break scenarios.

5. Conclusion

This paper proposes a novel strategy for processing nonlinear regression outliers by incorporating sensitivity functions straight into the model architecture. Instead of controlling the weights and/or loss functions in the estimation approach, we attenuate the impact of extreme values at the functional level.

Such a construction also induces an implicit regularization, as evidenced by the theoretical study that follows. That's simply because, where our predictor is given an excessively large value, the continuous, sublinear sensitivity function reduces it. In mathematical parlance, we limit the effects of outliers on the objective function to a relatively small extent in such a manner as to keep the model regular and stable even when explicit penalty terms are absent.

Given its volatility, for instance, we can provide a realistic, practical test of the model by testing it against volatile stock return data. The framework models sudden market shocks well and, when volatility spikes, keeps forecast errors constant, without sacrificing accuracy when the market is calm and stationary. Its structural flexibility makes it suitable for statistical modeling applications, where heavy-tailed noise and sudden shifts are the norm. The appropriate next step, of course, is to examine how this operator behaves asymptotically in high-dimensional settings.

Acknowledgements

The author would like to acknowledge that this work was carried out independently without any external support or funding.

Statements on compliance with ethical standards and standards of research involving animals

This article does not contain any studies involving animals or human participants.

Disclosure and conflict of interest

The author declares that there are no conflicts of interest.

References

[1] L. Fahrmeir, T. Kneib, S. Lang and B. D. Marx, "Regression: Models, methods and applications," 2nd ed., Springer, 2021. Available: https://doi.org/10.1007/978-3-662-63882-8

[2] J. Blanchet, J. Li, S. Lin and X. Zhang, "Distributionally robust optimization and robust statistics," Statistical Science, vol. 40, no. 3, 2025. Available: https://doi.org/10.1214/24-STS955

[3] T. E. O'Brien and J. W. Silcox, "Nonlinear regression modelling: A primer with applications and caveats," Bulletin of Mathematical Biology, vol. 86, no. 4, Art. no. 40, 2024. Available: https://doi.org/10.1007/s11538-024-01274-4

[4] C. Zhang, T. Zhang, F. Yin and A. M. Zoubir, "Data-adaptive M-estimators for robust regression via bi-level optimization," Signal Processing, vol. 210, p. 109063, 2023. Available:

https://doi.org/10.1016/j.sigpro.2023.109063

[5] A. C. Atkinson, M. Riani, A. Corbellini, D. Perrotta and V. Todorov, "Robust statistics through the monitoring approach: Applications in regression," Springer, 2024. Available: https://doi.org/10.1007/978-3-031-88365-1

[6] Q. Heng and K. Lange, "Bootstrap estimation of the proportion of outliers in robust regression," Statistics and Computing, vol. 35, no. 3, Art. no. 3, 2025. Available: https://doi.org/10.1007/s11222-024-10526-1

[7] P. Liu and W. Q. Meeker, "Robust numerical methods for nonlinear regression," arXiv preprint, 2024. Available: https://arxiv.org/abs/2403.12759

[8] J. Li, L. Kong, B. Jiang and W. Tu, "High-dimensional outlier detection and variable selection via adaptive weighted mean regression," Electronic Journal of Statistics, 2025. Available: https://doi.org/10.1214/25-EJS2394

[9] L. Huang, C. Liu and L. Wang, "Weighted least squares estimation with multivariate-dependent weights for heteroscedastic linear regression," arXiv preprint, 2026. Available: https://arxiv.org/abs/2601.10049

[10] K. Reichold, "Adaptive LASSO in counteracting regressions: Asymptotic properties and model selection behavior," arXiv preprint, 2025. Available: https://arxiv.org/abs/2510.07204

[11] I. M. Danilevicz, P. Bondon and V. A. Reisen, "Adaptive LASSO quantile regression with fixed effects," Applied Mathematical Modelling, vol. 153, p. 116600, 2026. Available: https://doi.org/10.1016/j.apm.2025.116600

[12] T. Hastie, R. Tibshirani and M. Wainwright, "Statistical learning with sparsity: The lasso and generalizations," Chapman & Hall/CRC, 2015. Available: https://doi.org/10.1201/b18401

[13] P.-L. Loh, "A theoretical review of modern robust statistics," Annual Review of Statistics and Its Application, vol. 11, pp. 1–30, 2024. Available: https://doi.org/10.1146/annurev-statistics-112723-034446

[14] R. A. Maronna, R. D. Martin, V. J. Yohai and M. Salibián-Barrera, "Robust statistics: Theory and methods," 2nd ed., Wiley, 2019. Available: https://doi.org/10.1002/9781119214656

[15] H. Zhou and H. Zou, "A nonparametric Box-Cox transformation approach for robustifying high-dimensional linear inference," arXiv preprint, 2024. Available: https://arxiv.org/abs/2405.12816

[16] D. M. Bates and D. G. Watts, "Nonlinear regression analysis and its applications," Wiley, 1988. Available: https://doi.org/10.1002/9780470316757

[17] P. J. Huber and E. M. Ronchetti, "Robust statistics," 2nd ed., Wiley, 2009. Available: https://doi.org/10.1002/9780470434697

[18] A. Chudik, M. H. Pesaran and M. Sharifvaghefi, "Variable selection in high-dimensional linear regressions with parameter instability," Journal of Econometrics, 2024. Available: https://doi.org/10.1016/j.jeconom.2024.105900