axioms

MDPI

Article

Modeling Environmental Pollution Using Varying-Coefficients Quantile Regression Models under Log-Symmetric Distributions

Luis Sánchez [1], Germán Ibacache-Pulgar [2,3], Carolina Marchant [4,*] and Marco Riquelme [2]

1 Institute of Statistics, Universidad Austral de Chile, Valdivia 5110234, Chile; luis.sanchez@uach.cl
2 Institute of Statistics, Universidad de Valparaíso, Valparaíso 2360102, Chile; german.ibacache@uv.cl (G.I.-P.); marco.riquelme@uv.cl (M.R.)
3 Centro Interdisciplinario de Estudios Atmosféricos y Astroestadística, Universidad de Valparaíso, Valparaíso 2360102, Chile
4 Faculty of Basic Sciences, Universidad Católica del Maule, Talca 3480112, Chile
* Correspondence: cmarchant@ucm.cl

Abstract: Many phenomena can be described by random variables that follow asymmetrical distributions. In the context of regression, when the response variable Y follows such a distribution, it is preferable to estimate the response variable for predictor values using the conditional median. Quantile regression models can be employed for this purpose. However, traditional models do not incorporate a distributional assumption for the response variable. To introduce a distributional assumption while preserving model flexibility, we propose new varying-coefficients quantile regression models based on the family of log-symmetric distributions. We achieve this by reparametrizing the distribution of the response variable using quantiles. Parameter estimation is performed using a maximum likelihood penalized method, and a back-fitting algorithm is developed. Additionally, we propose diagnostic techniques to identify potentially influential local observations and leverage points. Finally, we apply and illustrate the methodology using real pollution data from Padre Las Casas city, one of the most polluted cities in Latin America and the Caribbean according to the World Air Quality Index Ranking.

Keywords: local influence techniques; log-symmetric distributions family; PM2.5 levels; quantile regression; semiparametric models

MSC: 62J20

Citation: Sánchez, L.; Ibacache-Pulgar, G.; Marchant, C.; Riquelme, M. Modeling Environmental Pollution Using Varying-Coefficients Quantile Regression Models under Log-Symmetric Distributions. *Axioms* **2023**, *12*, 976. https://doi.org/10.3390/axioms12100976

Academic Editor: Giovanni Nastasi

Received: 9 September 2023
Revised: 12 October 2023
Accepted: 13 October 2023
Published: 17 October 2023

1. Introduction

In the process of data modeling, it is common to utilize regression models that assume that the response variable follows a normal distribution, as this is well-established in theory. However, there are situations where using such models may not be appropriate, particularly when the response variable exhibits an asymmetric distribution and is restricted to the positive real line. Failing to account for this behavior can introduce bias in parameter estimates and the estimation of associated measures of variability; see [1]. To address the limitations associated with the assumption of normality, several authors have proposed alternative approaches that employ more flexible distributional assumptions. This allows for a better representation of the underlying data. Some examples of such approaches include the works of [2–7].

Vanegas and Paula [8] proposed a family of log-symmetric distributions, which are obtained by transforming symmetric distributed random variables whose probability density functions involve the exponential function. Some examples of log-symmetric distributions are the log-normal, log-power-exponential, log-Laplace, log-logistic, log-slash, log-hyperbolic, Birnbaum–Saunders (BS), and log-Student-t cases. This family of distributions includes special cases that exhibit bimodal behavior, as well as distributions with tails

that are either lighter or heavier than the log-normal distribution. Regression models based on log-symmetric distributions have been studied by Vanegas and Paula [1,9,10].

In many real-life phenomena, the focus of interest is often on modeling a specific quantile of the response variable rather than the mean, as commonly done in classical regression models. This is particularly relevant when the distribution of the response variable exhibits asymmetry, where the median becomes a more appropriate measure of central tendency for estimating the response. Another reason is that our interest can be to model the relation between another non-central position measure and the covariates. This happens, for example, when we want to analyze the relationship between the greater (or lower) values of the response variable and the covariates; see [11]. Therefore, quantile regression models are useful for modeling the relationship between a set of predictor variables and specific quantiles of a response variable. Unlike traditional regression models, quantile regression does not assume a specific distribution for the response variable [11,12]. However, if we introduce a distributional assumption, it is possible to formulate quantile regression models based on the reparameterization of the distribution using a quantile. This approach has been successfully applied by [5,6,13]. Quantile regression models based on reparametrized log-symmetric distributions by quantiles (QLS) have been recently developed by Saulo et al. [7], albeit from a purely parametric perspective.

Considering the inclusion of nonparametric functions in the modeling, it becomes possible to incorporate the nonlinear effects of covariates. Semiparametric models have been developed to address this, where linear structures are described by parametric components and nonlinear structures are described by nonparametric components. Therefore, these models offer better flexibility for modeling data than those using only a parametric approach. Semiparametric structures have been effectively utilized to represent nonlinear components, as demonstrated in previous studies such as [1,14–21]. Based on our literature review, it appears that no semiparametric quantile regression models based on log-symmetric distributions have been developed thus far.

For over 30 years, Chile has been grappling with a significant public health issue related to the contamination of respirable particulate matter, particularly during winter periods. In the context of Latin America and the Caribbean, Chile currently ranks second, following Peru, in terms of cities with the highest levels of fine particulate matter (PM2.5), as reported by the World Air Quality Index Ranking (https://bit.ly/3MXVP38; accessed on 20 August 2023). It is concerning to note that these levels often exceed both national and international regulations, highlighting the severity of the problem in terms of public health. Statistical models provide a valuable approach to understanding and describing air quality, enabling us to study the relative impact of atmospheric contaminants on human health and the urban environment. Periodic episodes of extreme air pollution concentrations can occur with certain atmospheric contaminants, varying with geographical and meteorological factors and dependent on changes in emission sources and types; see [22]. Considering this variability, air pollutant concentrations are treated as non-negative random variables. In general, the distribution of these variables is asymmetrical and exhibits positive skewness, aligning with the characteristics of log-symmetric distributions.

The primary aim of this article is to develop varying-coefficients quantile regression models based on the family of log-symmetric distributions. Our secondary objectives encompass the following: (i) to estimate the parameters of the model using the maximum penalized likelihood (MPL) technique and a back-fitting algorithm; (ii) to incorporate the nonparametric structure through natural cubic smoothing splines (iii) to calculate local influence techniques for model diagnostics by assess the normal curvatures under different perturbation scenarios; (iv) to implement the obtained outcomes computationally within the R programming environment; and (v) to apply these results to real data related to atmospheric pollutants in Padre Las Casas, a city in Chile recognized as one of the most contaminated cities in Latin America and the Caribbean, as per the World Air Quality Index Ranking (https://www.iqair.com/; accessed on 20 August 2023).

The remainder of this work is organized as follows. Section 2 presents the proposed model for varying-coefficients quantile regression based on log-symmetric distributions. In Section 3, we explain the parameter estimation procedure utilizing the MPL method and a back-fitting algorithm. Section 4 extends the local influence method to assess the potential impact of specific observations on the proposed model, including the derivation of the generalized leverage matrix. In Section 5, we apply the proposed model to analyze a real dataset, demonstrating its potential applications. Finally, in Section 6, we provide concluding remarks and suggestions for future research.

2. Log-Symmetric Varying-Coefficient Quantile Regression Models

In this section, we introduce the varying-coefficients quantile regression models based on the log-symmetric distribution family.

2.1. Formulation

Let $q \in (0,1)$ be a fixed number. We will denote by $Y \sim \text{QLS}(Q, \phi, g)$ to the log-symmetric distribution reparametrized by the q-quantile of Y (Q), where $\phi > 0$ is a power parameter and $g(\cdot)$ is the probability density function generator kernel; see [7]. Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables such that $Y_i \sim \text{QLS}(Q_i, \phi, g)$, for $i \in \{1, 2, \ldots, n\}$. We assume the semiparametric additive structure for Q_i given by

$$h(Q_i) = \boldsymbol{w}_i^\top \boldsymbol{\alpha} + x_{1i}\,\beta_1(t_{1i}) + \cdots + x_{si}\,\beta_s(t_{si}), \quad i \in \{1, 2, \ldots, n\}, \tag{1}$$

where $\boldsymbol{w}_i^\top = (1, w_{1i}, \ldots, w_{pi})$, $\boldsymbol{\alpha}$ is a $(p+1) \times 1$ unknown regression coefficients vector, with $p + 1 < n$, $\beta_1, \ldots, \beta_s$ are unspecified smooth real functions of the explanatory variable T_k that do not depend on $\boldsymbol{\alpha}$ or some other parameter. Also, x_{ji}, w_{ji} and t_{ji} are the values of covariates X_j, W_j and T_j for the ith observation, respectively. The function h has positive support and is at least twice differentiable, called the link function. The structure of the right side in Equation (1) defines the so-called partially varying-coefficients regression models; see [23]. Therefore, we have defined a partially varying-coefficient quantile regression model based on the family of log-symmetric distributions. Equation (1) can be written as

$$h(Q_i) = \boldsymbol{w}_i^\top \boldsymbol{\alpha} + \tilde{\boldsymbol{n}}_{1i}^\top \boldsymbol{\beta}_1 + \cdots + \tilde{\boldsymbol{n}}_{si}^\top \boldsymbol{\beta}_s, \quad i \in \{1, 2, \ldots, n\},$$

where $\tilde{\boldsymbol{n}}_{ki}^\top$ denote the ith row of the matrix $\tilde{\boldsymbol{N}}_k = \boldsymbol{X}^{(k)}\,\boldsymbol{N}_k$, $\boldsymbol{X}^{(k)} = \text{diag}\{x_{k1}, , \ldots, x_{kn}\}$, $\boldsymbol{N}_k$ is the incidence matrix $n \times r_k$ whose (i,l)-th element equals to the indicator function $I(t_{ki} = t_{kl}^0)$, $\boldsymbol{\beta}_k = (\zeta_{k1}, \ldots, \zeta_{kr_k})^\top$ is a $r_k \times 1$ vector called a vector of spline coefficients such that $\zeta_{kl} = \beta_k(t_{kl}^0)$, with t_{kl}^0 for $l \in \{1, \ldots, r_k\}$ representing the distinct and ordered values of the explanatory variable T_k usually called knots. For a similar formulation, see [16].

2.2. Penalized Log-Likelihood Function

The log-likelihood function for the proposed model in Equation (1) is given by

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log\left(\frac{\xi_{\text{nc}}}{y_i}\right) + \log(g(v_i^2)) - \frac{1}{2}\log(\phi),$$

where $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_s^\top, \phi)^\top$ and $v_i = (\log(y_i) - \log(Q_i) + \sqrt{\phi} z_q)/\sqrt{\phi}$. To address the identifiability issues of the regression coefficient $\boldsymbol{\alpha}$ and mitigate overfitting in the semiparametric modeling process, penalties are commonly incorporated into the smooth functions. The MPL method, initially introduced by Good and Gaskins [24] for estimating probability density curves, has been extended to the nonparametric regression context by researchers such as [25,26]. These extensions have provided effective solutions to handle the challenges of identifiability and overfitting in semiparametric models. This same approach is used to fit our model, optimizing the penalized log-likelihood function expressed as

$$\ell_{\text{P}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \ell(\boldsymbol{\theta}) - \sum_{k=1}^{s} \lambda_k^* J(\beta_k), \tag{2}$$

where $J(\beta_k)$ corresponds to a penalty function on the function β_k that regulates the lack of smoothness of the estimated curve. Assuming that the design points t_k^0 belong to the compact set $[a_k, b_k]$ and that the functions β_k's belongs to the Sobolev function space [27]

$$W_{[a_k,b_k]} = \left\{\beta_k : \beta_k \text{ and } \beta_k' \text{ are absolutely continuous on } [a_k, b_k], \text{ and } \int_{a_k}^{b_k} [\beta_k''(t_k)]^2\, dt_k < \infty\right\},$$

Then one way to measure the roughness of the function β_k over the interval $[a_k, b_k]$ is by their squared norm given by $J(\beta_k) = \|\beta_k\|^2 = \int_{a_k}^{b_k} [\beta_k''(t_k)]^2\, dt_k$. Green and Silverman [15] showed that $J(\beta_k) = \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k$, where $\boldsymbol{K}_k$ is a $r_k \times r_k$ non-negative definite matrix. Please note that both $\boldsymbol{\beta}_k$ and $\boldsymbol{K}_k$ are evaluated at the values belonging to the set of knots $\{t_{k1}^0, t_{k2}^0, \ldots, t_{kr_k}^0\}$, for $k \in \{1, 2, \ldots, s\}$, and therefore have finite dimensions. Taking $\lambda_k^* = \lambda_k/2$, we can obtain the maximum penalized likelihood estimator (MPLE) of $\boldsymbol{\theta}$, denoted by $\widehat{\boldsymbol{\theta}}$, maximizing

$$\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \ell(\boldsymbol{\theta}) - \sum_{k=1}^{s} \frac{\lambda_k}{2} \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k, \tag{3}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_s)^\top$ denotes an $s \times 1$ vector of smoothing parameters. Each $\lambda_k \geq 0$ measures the "rate of exchange" between goodness-of-fit and variability of the function β_k. In this scenario, the estimators of β_k's result in a cubic spline that is completely determined by the finite-dimensional set of knots $\{t_{k1}^0, t_{k2}^0, \ldots, t_{kr_k}^0\}$.

3. Parameter Estimation and Inference

In this section, we focus on estimating the parameters of the model described in Equation (1) and discuss aspects of statistical inference. We also provide a brief discussion on calculating the effective degrees of freedom and selecting smoothing parameters. To facilitate the parameter estimation process and associated inference, we have developed a routine in the R-project (https://www.r-project.org/; accessed on 15 May 2023).

3.1. Penalized Score Vector

First, we make the assumption that the function $\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})$ given in Equation (2) is regular, meaning that it has first and second partial derivatives with respect to the elements of the parameter vector $\boldsymbol{\theta}$. By performing partial derivative operations, we can express the score function for $\boldsymbol{\theta}$ in matrix form as follows:

$$\boldsymbol{U}_{\mathrm{p}}^\top(\boldsymbol{\theta}) = \frac{\partial \ell_{\mathrm{p}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\alpha}^\top}(\boldsymbol{\theta}) & \boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\beta}_1^\top}(\boldsymbol{\theta}) & \ldots & \boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\beta}_s^\top}(\boldsymbol{\theta}) & \boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\phi}^\top}(\boldsymbol{\theta}) \end{pmatrix}^\top,$$

where $\boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \boldsymbol{W}^\top \boldsymbol{D}_a \boldsymbol{z}$, $\boldsymbol{U}_{\mathrm{p}}^{\boldsymbol{\beta}_k}(\boldsymbol{\theta}) = \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_a \boldsymbol{z} - \lambda_k \boldsymbol{K}_k \boldsymbol{\beta}_k$, for $k \in \{1, \ldots, s\}$, and $U_{\mathrm{p}}^{\phi}(\boldsymbol{\theta}) = \mathrm{tr}(\boldsymbol{D}_b)$, with $\boldsymbol{D}_a = \mathrm{diag}\{a_1, \ldots, a_n\}$, $\boldsymbol{D}_b = \mathrm{diag}\{b_1, \ldots, b_n\}$, $\boldsymbol{z} = (z_1, \ldots, z_n)^\top$, $z_i = v_i r(v_i)/Q_i\sqrt{\phi}$, $b_i = r(v_i)\,\phi^{-3/2} v_i[\log(y_i) - \log(Q_i)]/2 - 1/2\phi$ and $a_i = 1/h'(Q_i)$, being $r(v_i) = -2g'(v_i^2)/g(v_i^2)$. Please note that g' represents the derivative of the function g.

3.2. Penalized Hessian Matrix

To obtain the penalized Hessian matrix, we need to compute the second derivate of $\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})$ with respect to each element of $\boldsymbol{\theta}$, i.e., $\partial^2 \ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})/\partial\theta_{j^*}\partial\theta_{l^*}$ for $j^*, l^* \in \{1, \ldots, p^*\}$ and $p^* = 2 + p + \sum_{k=1}^{s} r_k$. After performing some algebraic manipulations, we obtain the penalized Hessian matrix in the following form:

$$\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta}) = \frac{\partial^2 \ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{pmatrix} \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_1} & \cdots & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_s} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}} \\ \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_1^\top} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_1} & \cdots & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_s} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\phi}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_s^\top} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_s^\top} & \cdots & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\beta}_s^\top} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\phi}} \\ \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}^\top} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\phi}^\top} & \cdots & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\phi}^\top} & \ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\phi}\boldsymbol{\phi}} \end{pmatrix}, \tag{4}$$

with $\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \boldsymbol{W}^\top \boldsymbol{D}_c \boldsymbol{W}$, $\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_k} = \boldsymbol{W}^\top \boldsymbol{D}_c \tilde{\boldsymbol{N}}_k$, $\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}} = \boldsymbol{W}^\top \boldsymbol{D}_a \boldsymbol{m}$, $\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_k\boldsymbol{\phi}} = \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_a \boldsymbol{m}$, for $k \in \{1, \ldots, s\}$, $\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\phi}\boldsymbol{\phi}} = \mathrm{tr}(\boldsymbol{D}_d)$, and

$$\ddot{\boldsymbol{L}}_{\mathrm{p}}^{\boldsymbol{\beta}_k\boldsymbol{\beta}_{k'}^\top} = \begin{cases} \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_c \tilde{\boldsymbol{N}}_k - \lambda_k \boldsymbol{K}_k, & k = k' \\ \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_c \tilde{\boldsymbol{N}}_{k'}, & k \neq k' \end{cases}$$

where the matrices $\boldsymbol{D}_c = \mathrm{diag}\{c_1, \ldots, c_n\}$, $\boldsymbol{D}_a = \mathrm{diag}\{a_1, \ldots, a_n\}$ and vector $\boldsymbol{m} = (m_1, \ldots, m_n)^\top$, with c_i, a_i and m_i defined in Appendix A. The Hessian matrix presented in this section will be used in the construction of the normal curvature for the local influence method developed in Section 4.

3.3. Penalized Fisher Information Matrix

By taking the expectation of the matrix $-\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta})$ given in Equation (4), we derive the $p^* \times p^*$ penalized expected information matrix given by

$$\boldsymbol{J}_{\mathrm{p}}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_1} & \cdots & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_s} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}} \\ \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_1^\top} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_1} & \cdots & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_s} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\phi}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_s^\top} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\beta}_s^\top} & \cdots & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\beta}_s^\top} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\phi}} \\ \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}^\top} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_1\boldsymbol{\phi}^\top} & \cdots & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_s\boldsymbol{\phi}^\top} & \boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\phi}\boldsymbol{\phi}} \end{pmatrix}, \tag{5}$$

whose elements can be expressed as $\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \boldsymbol{W}^\top \boldsymbol{D}_v \boldsymbol{W}$, $\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\beta}_k} = \boldsymbol{W}^\top \boldsymbol{D}_v \tilde{\boldsymbol{N}}_k$, $\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\alpha}\boldsymbol{\phi}} = \boldsymbol{W}^\top \boldsymbol{D}_a \boldsymbol{s}$, $\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_k\boldsymbol{\phi}} = \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_a \boldsymbol{s}$, for $k \in \{1, \ldots, s\}$, $\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\phi}\boldsymbol{\phi}} = \mathrm{tr}(\boldsymbol{D}_u)$, and

$$\boldsymbol{J}_{\mathrm{p}}^{\boldsymbol{\beta}_k\boldsymbol{\beta}_{k'}^\top} = \begin{cases} \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_v \tilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k, & k = k' \\ \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_v \tilde{\boldsymbol{N}}_{k'}, & k \neq k', \end{cases}$$

where $\boldsymbol{D}_v = \mathrm{diag}\{e_1, \ldots, e_n\}$, $\boldsymbol{D}_u = \mathrm{diag}\{u_1, \ldots, u_n\}$ and $\boldsymbol{s} = (s_1, \ldots, s_n)^\top$, being $e_i = \mathbb{E}[-c_i]$, $s_i = \mathbb{E}[-m_i]$ and $u_i = \mathbb{E}[-d_i]$, with $\mathbb{E}[\cdot]$ denoting the expected value operator. This matrix will be utilized to approximate the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, as discussed in Section 3.5.

3.4. Iterative Process

The MPLE of $\boldsymbol{\theta}$ is obtained by maximizing the penalized log-likelihood function presented in Equation (3). Since the resulting estimation equation $\boldsymbol{U}_{\mathrm{p}}(\boldsymbol{\theta}) = \boldsymbol{0}$ is nonlinear, an iterative process is necessary to solve it. In this regard, we propose to employ the Fisher scoring algorithm, which updates $\boldsymbol{\theta}$ using the matrix equation

$$\boldsymbol{J}_{\mathrm{p}}(\boldsymbol{\theta}) \left[\boldsymbol{\theta}^{(m+1)} - \boldsymbol{\theta}^{(m)}\right] = \boldsymbol{U}_{\mathrm{p}}(\boldsymbol{\theta})^{(m)}, \qquad m = 0, 1, \ldots. \tag{6}$$

3.4.1. ϕ Unknown

After some algebraic operations, we obtain the following expressions for the iterative solutions for the case where ϕ unknown:

$$
\begin{aligned}
\boldsymbol{\alpha}^{(m+1)} &= (\boldsymbol{W}^\top \boldsymbol{D}_v^{(m)} \boldsymbol{W})^{-1} \boldsymbol{W}^\top \boldsymbol{D}_v^{(m)} \Big[\boldsymbol{\psi}_\alpha^{(m)} - \boldsymbol{D}_{v,a}^{(m)} \boldsymbol{s} \boldsymbol{\Phi}_\phi^{(m+1,m)} - \sum_{k=1}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\Phi}_{\beta_k}^{(m+1,m)} \Big], \\
\boldsymbol{\beta}_\ell^{(m+1)} &= (\tilde{\boldsymbol{N}}^\top \boldsymbol{D}_v^{(m)} \tilde{\boldsymbol{N}} + \lambda_k \boldsymbol{K})^{-1} \tilde{\boldsymbol{N}}^\top \boldsymbol{D}_v^{(m)} \Big[\boldsymbol{\psi}_{\beta_\ell}^{(m)} - \boldsymbol{D}_{v,a}^{(m)} \boldsymbol{s} \boldsymbol{\Phi}_\phi^{(m+1,m)} - \boldsymbol{W} \boldsymbol{\Phi}_\alpha^{(m+1,m)} \\
&\quad - \sum_{k=1, k\neq \ell}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\Phi}_{\beta_k}^{(m+1,m)} \Big], \qquad \ell \in \{1, \ldots, s\} \quad \text{and} \\
\phi^{(m+1)} &= \mathrm{tr}^{-1}(\boldsymbol{D}_u^{(m)} \Big[\mathrm{tr}(\boldsymbol{D}_b^{(m)}) + \mathrm{tr}(\boldsymbol{D}_u^{(m)}) \phi^{(m)} - \boldsymbol{s}^\top \boldsymbol{D}_a^{(m)} \boldsymbol{W} \boldsymbol{\Phi}_{\boldsymbol{\alpha}}^{(m+1,m)} \\
&\quad - \boldsymbol{s}^\top \boldsymbol{D}_a^{(m)} \sum_{k=1}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\Phi}_{\boldsymbol{\beta}_k}^{(m+1,m)} \Big] ,
\end{aligned}
$$

where $\boldsymbol{\psi}_\alpha^{(m)} = \boldsymbol{D}_{v,a}^{(m)} \boldsymbol{z}^{(m)} + \boldsymbol{W} \boldsymbol{\alpha}^{(m)}$ and $\boldsymbol{\psi}_{\boldsymbol{\beta}_\ell}^{(m)} = \boldsymbol{D}_{v,a}^{(m)} \boldsymbol{z}^{(m)} + \tilde{\boldsymbol{N}}_\ell \boldsymbol{\beta}_\ell^{(m)}$, with $\boldsymbol{D}_{v,a}^{(m)} = \boldsymbol{D}_v^{(m)^{-1}} \boldsymbol{D}_a^{(m)}$.

3.4.2. ϕ Known

When ϕ is known, it is possible to obtain simplified expressions for the iterative solutions of $\boldsymbol{\alpha}^{(m+1)}$ and $\boldsymbol{\beta}_\ell^{(m+1)}$. In this case, we have that

$$
\begin{aligned}
\boldsymbol{\alpha}^{(m+1)} &= (\boldsymbol{W}^\top \boldsymbol{D}_v^{(m)} \boldsymbol{W})^{-1} \boldsymbol{W}^\top \boldsymbol{D}_v^{(m)} \Big[\boldsymbol{r}_{v,a}^{(m)} - \sum_{k=1}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(m+1)} \Big], \quad \text{and} \\
\boldsymbol{\beta}_\ell^{(m+1)} &= (\tilde{\boldsymbol{N}}^\top \boldsymbol{D}_v^{(m)} \tilde{\boldsymbol{N}} + \lambda_k \boldsymbol{K})^{-1} \tilde{\boldsymbol{N}}^\top \boldsymbol{D}_v^{(m)} \Big[\boldsymbol{r}_{v,a}^{(m)} - \boldsymbol{W} \boldsymbol{\alpha}^{(m+1)} - \sum_{k=1, k\neq \ell}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(m+1)} \Big],
\end{aligned}
$$

for $\ell \in \{1, \ldots, s\}$, where $\boldsymbol{r}_{v,a}^{(m)} = \boldsymbol{D}_{(v,a)}^{(m)} \boldsymbol{z}^{(m)} + \boldsymbol{\eta}^{(m)}$, with $\boldsymbol{\eta}^{(m)} = \boldsymbol{W} \boldsymbol{\alpha}^{(m)} + \sum_{k=1}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(m)}$. It is possible to prove that these expressions correspond to the weighted back-fitting (Gauss-Seidel) iterations considering $\boldsymbol{r}_{v,a}^{(m)}$ as dependent modified variable and $\boldsymbol{D}_v$ as a matrix of weights that changes with each iteration of the process; see, for instance [28]. A general expression for these iterations is as follows:

$$
\boldsymbol{\beta}_\ell^{(m+1)} = \boldsymbol{S}_\ell^{(m)} \Big[\boldsymbol{r}_{v,a}^{(m)} - \sum_{k=0, k\neq \ell}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(m+1)} \Big], \quad \ell \in \{1, \ldots, s\}, \tag{7}
$$

where $\boldsymbol{r}_{v,a}^{(m)} = \boldsymbol{D}_{v,a}^{(m)} \boldsymbol{z}^{(m)} + \boldsymbol{\eta}^{(m)}$, with $\boldsymbol{\eta}^{(m)} = \sum_{k=0}^{s} \tilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(m)}$, $\tilde{\boldsymbol{N}}_0 = \boldsymbol{W}$, $\boldsymbol{\beta}_0 = \boldsymbol{\alpha}$, $\boldsymbol{S}_0^{(m)} = (\tilde{\boldsymbol{N}}_0^\top \boldsymbol{D}_v^{(m)} \tilde{\boldsymbol{N}}_0)^{-1} \tilde{\boldsymbol{N}}_0^\top \boldsymbol{D}_v^{(m)}$ and $\boldsymbol{S}_k^{(m)} = (\tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_v^{(m)} \tilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \tilde{\boldsymbol{N}}_k^\top \boldsymbol{D}_v^{(m)}$. A discussion about the consistency of the system of Equations (6) and the convergence of the back-fitting algorithm in (7) is given, for example, in [29].

3.5. Approximate Standard Errors

In this work, we propose to approximate the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ using the inverse of the penalized Fisher information matrix defined in Equation (5). In effect, an estimation of the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ is given by

$$
\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\theta}}) \approx \boldsymbol{J}_\mathrm{p}(\widehat{\boldsymbol{\theta}})^{-1} . \tag{8}
$$

Following [14], we can consider an approximate pointwise standard error band (SEB) for nonparametric functions β_k's to evaluate the accuracy of the estimators $\widehat{\beta}_k'$s for different locations within the range of interest. In our case, these approximate pointwise SEBs are provided by

$$
\mathrm{SEB}_{\mathrm{approx}} \left(\beta_k \left(t_l^0 \right) \right) = \widehat{\beta}_k \left(t_l^0 \right) \pm 2 \sqrt{\widehat{\mathrm{Var}} \left(\widehat{\beta}_k \left(t_l^0 \right) \right)},
$$

where $\mathrm{Var}\left(\widehat{\beta}_k(t_l)\right)$ is the l-th principal diagonal element of the matrix provided in Equation (8) for $l \in \{1, 2, \ldots, r_k\}$. Please note that t_l^0 corresponds to the knots associated with each variable with a nonparametric contribution to the model.

3.6. Effective Degrees of Freedom and λ_k's

The calculation of the degrees of freedom associated with the parametric and non-parametric contributions is based on the iterative process used in the parameters estimation of the proposed model. Assuming ϕ fixed, we have from the convergence of the iterative process that

$$\widehat{\boldsymbol{\beta}}_\ell = (\tilde{\boldsymbol{N}}^\top \widehat{\boldsymbol{D}}_v \tilde{\boldsymbol{N}} + \lambda_k \boldsymbol{K})^{-1} \tilde{\boldsymbol{N}}^\top \widehat{\boldsymbol{D}}_v \widehat{\boldsymbol{r}}^*_{v,a}, \quad \ell \in \{1, \ldots, s\},$$

where $\widehat{\boldsymbol{r}}^*_{v,a} = \widehat{\boldsymbol{r}}_{v,a} - \sum_{k=0, k\neq\ell}^{s} \tilde{\boldsymbol{N}}_k \widehat{\boldsymbol{\beta}}_k$, $\widehat{\boldsymbol{r}}_{v,a} = \widehat{\boldsymbol{D}}_{(a,v)} \widehat{\boldsymbol{z}} + \widehat{\boldsymbol{\eta}}$, $\widehat{\boldsymbol{\eta}} = \boldsymbol{W}\widehat{\boldsymbol{\alpha}} + \sum_{k=1}^{s} \tilde{\boldsymbol{N}}_k \widehat{\boldsymbol{\beta}}_k$ and $\widehat{\boldsymbol{z}} = (\widehat{z}_1, \ldots, \widehat{z}_n)^\top$, with z_i ($i \in \{1, 2, \ldots, n\}$) defined in Section 3.1. Note that $\boldsymbol{r}^*_{v,a}$ can be interpreted as a modified variable and $\boldsymbol{D}_v$ a weight matrix that is updated at each stage of the iterative process. From this, we define the effective degrees of freedom (edf) associated with the smooth functions as (see, for instance [14])

$$\mathrm{edf}(\lambda_k) = \mathrm{tr}\{\tilde{\boldsymbol{N}}(\tilde{\boldsymbol{N}}^\top \widehat{\boldsymbol{D}}_v \tilde{\boldsymbol{N}} + \lambda_k \boldsymbol{K})^{-1} \tilde{\boldsymbol{N}}^\top \widehat{\boldsymbol{D}}_v\}, \quad \ell \in \{1, \ldots, s\}.$$

Following Ibacache-Pulgar and Reyes [23], we choose the optimal smoothing parameter for each smooth function by specifying an appropriate $\mathrm{edf}(\lambda_k)$ value. Another way to select the λ_k's is to consider the Akaike Information Criterion (AIC). The idea is to minimize a function with respect to $\boldsymbol{\lambda}$ formulated as follows:

$$\mathrm{AIC}(\boldsymbol{\lambda}) = -2\ell_{\mathrm{p}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\lambda}) + 2(2 + p + \mathrm{edf}(\boldsymbol{\lambda})), \tag{9}$$

where $\ell_{\mathrm{p}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\lambda})$ denotes the penalized log-likelihood function evaluated at $\widehat{\boldsymbol{\theta}}$ for a fixed $\boldsymbol{\lambda}$ and $\mathrm{edf}(\boldsymbol{\lambda}) = \sum_{k=1}^{s} \mathrm{edf}(\lambda_k)$ denoting the number of effective parameters involved in the modeling of the smooth functions. A grid for different values of $\boldsymbol{\lambda}$ and its corresponding $\mathrm{AIC}(\boldsymbol{\lambda})$ are helpful to choose the suitable smoothing parameters. The criteria defined in Equation (9) can also be used to select the best model within the class of varying coefficients quantile regression models based on the log-symmetric family.

4. Diagnostic Analysis

In this section, we extend the local influence method for the model given in Equation (1) and derive the generalized leverage matrix, which allows us to assess the influence of each observed value of the response variable y_i on its corresponding predicted value $\hat{y}_i$.

4.1. Local Influence Analysis

Let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ be an $n \times 1$ vector of perturbations restricted to some open subset $\Omega \in \mathbb{R}^n$ and $\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega})$ be the logarithm of the perturbed penalized likelihood function. It is assumed that exists $\boldsymbol{\omega}_0 \in \Omega$, a vector of non-perturbation, such that $\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda} | \boldsymbol{\omega}_0) = \ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})$. To assess the influence of small perturbations on the MPL estimate $\widehat{\boldsymbol{\theta}}$, we can consider the displacement of the penalized likelihood, which is given by $\mathrm{LD}(\boldsymbol{\omega}) = 2\big(\ell_{\mathrm{p}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\lambda}) - \ell_{\mathrm{p}}(\widehat{\boldsymbol{\theta}}_\omega, \boldsymbol{\lambda})\big)$, where $\widehat{\boldsymbol{\theta}}_\omega$ is the MPL estimate under $\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega})$. The measure $\mathrm{LD}(\boldsymbol{\omega})$ is helpful for assessing the distance between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_\omega$. Cook [30] suggested studying the local behavior of $\mathrm{LD}(\boldsymbol{\omega})$ around $\boldsymbol{\omega}_0$. The procedure involves selecting a unit direction $\boldsymbol{d} \in \Omega$ with $|\boldsymbol{d}| = 1$ and plotting $\mathrm{LD}(\boldsymbol{\omega}_0 + a\boldsymbol{d})$ against $a \in \mathbb{R}$. This plot, known as a lifted line, can be characterized by considering the normal curvature $C_d(\boldsymbol{\theta})$ around $a = 0$. To determine the direction $\boldsymbol{d} = \boldsymbol{d}\mathrm{max}$ that corresponds to the largest curvature $Cd_{\max}(\boldsymbol{\theta})$, one can examine the index plot of $\boldsymbol{d}_{\max}$. This plot helps identify cases that, under small perturbations, may have a significant potential influence on $\mathrm{LD}(\boldsymbol{\omega})$. According to Cook [30], the normal curvature at the unit direction $\boldsymbol{d}$ can be expressed as

$$C_d(\boldsymbol{\theta}) = -2\big(\boldsymbol{d}^\top \boldsymbol{\Delta}_{\mathrm{p}}^\top \ddot{\boldsymbol{L}}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}} \boldsymbol{d}\big),$$

with $\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta}) = \partial^2 \ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$ and $\boldsymbol{\Delta}_{\mathrm{p}} = \partial^2 \ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}\,|\boldsymbol{\omega})/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. $\boldsymbol{\Delta}_{\mathrm{p}}$ is called a penalized perturbation matrix. Observe that $C_d(\boldsymbol{\theta})$ denotes the local influence on the estimate $\widehat{\boldsymbol{\theta}}$ after perturbing the model or data. Escobar and Meeker [31] proposed to study the normal curvature at the direction $\boldsymbol{d} = \boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is an $n \times 1$ vector with a one at the ith position and zeros at the remaining positions. Thus, the normal curvature, called the total local influence of the ith case, assumes the form $C_{e_i}(\boldsymbol{\theta}) = 2|c_{ii}|$, for $i \in \{1, \ldots, n\}$, where c_{ii} is the ith principal diagonal element of the matrix $\boldsymbol{C} = \boldsymbol{\Delta}_{\mathrm{p}}^\top \ddot{\boldsymbol{L}}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}}$.

Next, we present the perturbed penalized log-likelihood function for four perturbation schemes, namely case weight, response variable, power parameter, and explanatory variable perturbation. The matrix $\boldsymbol{\Delta}_{\mathrm{p}}$ for each case is presented in Appendix B.

1. The case-weight perturbation scheme considers the perturbed penalized log-likelihood function as
$$\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}\,|\,\boldsymbol{\omega}) = \sum_{i=1}^{n} \omega_i \ell_i(Q_i, \phi; y_i) - \sum_{k=1}^{s} \frac{\lambda_k}{2} \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k,$$
where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ is the vector of weights, with $0 \leq \omega_i \leq 1$ for $i \in \{1, \ldots, n\}$.
2. Regarding the response variable perturbation scheme, we consider an additive type of perturbation weighted by a scaling factor on the ith response variable, i.e., $y_i(\omega_i) = y_i + \omega_i s_{Y_i}$, where s_{Y_i} is a scale factor that can be the sample standard deviation of Y_i and $\omega_i \in \mathbb{R}$, for $i \in \{1, \ldots, n\}$. Then, the perturbed penalized log-likelihood function is written as
$$\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}\,|\,\boldsymbol{\omega}) \quad = \quad \sum_{i=1}^{n} \ell_i(Q_i, \phi; y_i(\omega_i)) - \sum_{k=1}^{s} \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k.$$
3. Initially, the model given in Equation (1) assumes that the power parameter is constant across observations. However, we can introduce a perturbation in the power parameter such that it is not constant between the observations, i.e., $Y_i \sim \mathrm{QLS}(Q_i, \phi_i, g)$, where $\phi_i = \phi/\omega_i$, with $\omega_i > 0$ for $i \in \{1, \ldots, n\}$. Under this perturbation scheme, the perturbed penalized log-likelihood function is constructed from the expression defined in Equation (3) with ϕ being replaced by ϕ_i.
4. The last perturbation scheme considered in this work consists of incorporating an additive type perturbation on one of the covariates $X_1, \ldots, X_s$, say X_l, given by $x_{li}(\omega_i) = x_{li} + \omega_i s_{x_l}$, where s_{x_l} is a scale factor that can be the sample standard deviation of X_l and $\omega_i \in \mathbb{R}$, for $i \in \{1, \ldots, n\}$. In this case, the perturbed penalized log-likelihood function can be expressed as
$$\ell_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}\,|\,\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(Q_i(\omega_i), \phi; y_i) - \sum_{k=1}^{s} \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k,$$
where $Q_i(\omega_i)$ is as given in Equation (1) replacing w_{li} for $w_{li}(\omega_i)$.

4.2. Generalized Leverage Matrix

The generalized leverage (GL) measures the influence of the observed value of the response variable y_i on its corresponding predicted value $\hat{y}_i$ based on the model given in Equation (1). Following the approach proposed by Wei et al. [32], the GL for $\hat{\boldsymbol{\theta}}$ can be computed using the lemma they provided. The expression for the GL is given by $\partial\widehat{\boldsymbol{y}}/\partial\boldsymbol{y}^\top = \boldsymbol{H}_{\boldsymbol{\theta}}(-\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta}))^{-1}\ddot{\ell}_{\boldsymbol{\theta}\boldsymbol{y}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, where $\boldsymbol{H}_{\boldsymbol{\theta}} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}^\top$, $\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta}) = \partial^2\ell_{\mathrm{p}}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$, $\ddot{\ell}_{\boldsymbol{\theta}\boldsymbol{y}} = \partial^2\ell_{\mathrm{p}}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{y}^\top$, $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$, with μ_i being the mean of the Y_i. Using the chain rule, we have
$$\frac{\partial\widehat{\boldsymbol{y}}}{\partial\boldsymbol{y}^\top} = \frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{Q}^\top}\frac{\partial\boldsymbol{Q}}{\partial\boldsymbol{\theta}^\top}(-\ddot{\boldsymbol{L}}_{\mathrm{p}}(\boldsymbol{\theta}))^{-1}\ddot{\ell}_{\boldsymbol{\theta}\boldsymbol{y}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.$$

Because of $\mu = \log(\lambda)$ and $Q = \lambda \exp(\sqrt{\phi}\, z_q)$, where z_q is the q-quantile of the distribution $S(0, 1, g)$ [7], we have $\mu = \log(Q) - \sqrt{\phi} z_q$. Therefore, $\partial \boldsymbol{\mu}/\partial \boldsymbol{Q}^\top = \text{diag}\{1/Q_1, \ldots, 1/Q_n\}$. Also, we can obtain the $n \times p^*$ matrix

$$\frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{\theta}^\top} = \left(\frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{\alpha}^\top}, \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{\beta}_1^\top}, \ldots, \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{\beta}_s^\top}, \frac{\partial \boldsymbol{Q}}{\partial \phi} \right) = \left(\boldsymbol{D}_a \boldsymbol{W} \; \boldsymbol{D}_a \tilde{\boldsymbol{N}}_1 \; \cdots \; \boldsymbol{D}_a \tilde{\boldsymbol{N}}_s \; \boldsymbol{0}_n \right),$$

where $\boldsymbol{0}_n$ is the $n \times 1$ null vector and $\ddot{\ell}_{\boldsymbol{\theta}\boldsymbol{y}} = \left(\boldsymbol{D}_\psi \boldsymbol{D}_a \boldsymbol{W} \; \boldsymbol{D}_\psi \boldsymbol{D}_a \tilde{\boldsymbol{N}}_1 \; \cdots \; \boldsymbol{D}_\psi \boldsymbol{D}_a \tilde{\boldsymbol{N}}_s \; \boldsymbol{\tau} \right)$ is a $n \times p^*$ matrix. Please note that the computation of the matrix $-\ddot{\boldsymbol{L}}_{\text{p}}(\boldsymbol{\theta})$ relies on the availability of the penalized Hessian matrix given in Equation (4). By utilizing this penalized Hessian matrix, we have all the necessary elements to calculate the GL matrix $\partial \widehat{\boldsymbol{y}}/\partial \boldsymbol{y}^\top$.

5. Real Data Analysis

In this section, we apply the model proposed in Section 2 to real pollution data from the Padre Las Casas Air Quality Monitoring Station (AQMS). The AQMS is situated in the commune of Padre Las Casas in the Araucanía region of southern Chile, approximately 695 km away from Santiago, the capital city of Chile. Padre Las Casas has gained notoriety for its elevated levels of pollution, particularly concerning PM2.5. It is recognized as one of the most heavily polluted cities in Latin America and the Caribbean, as indicated by the World Air Quality Index Ranking (https://bit.ly/3MXVP38; accessed on 20 August 2023). The average concentration of PM2.5 in Padre Las Casas exceeds the limits set by national and international regulations [22], highlighting the significance of analyzing this type of data and developing models that accurately capture its behavior.

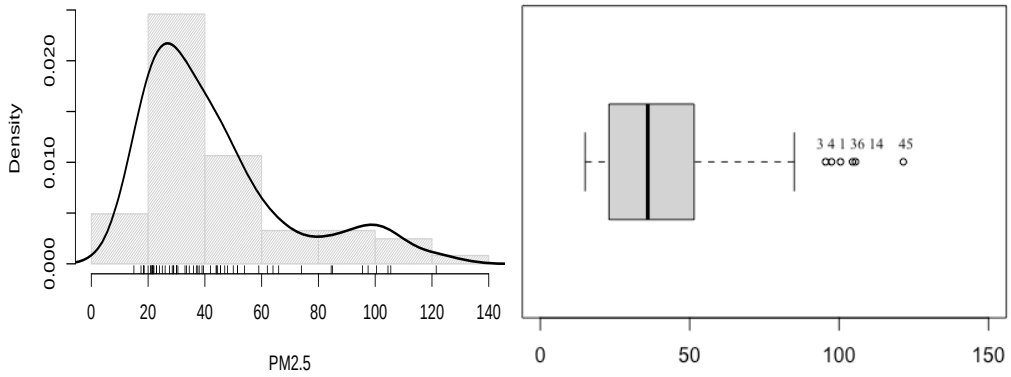
By studying the pollution data from the Padre Las Casas AQMS, we aim to gain insights into the underlying patterns and factors contributing to pollution levels. The proposed model will help us to describe and understand the behavior of pollution in this area, providing valuable information for monitoring and management purposes.

5.1. Exploratory Data Analysis

The dataset used in this analysis consists of hourly (h) average values for the months of June and July 2020, acquired from the Chilean Ministry of Environment (MMA) website (http://sinca.mma.gob.cl; accessed on 11 January 2022). The dataset includes measurements of various variables related to air pollution and meteorological conditions in Padre Las Casas. The considered random variables in this dataset are: (i) Median of PM2.5 concentrations: this variable represents the median concentration of fine particulate matter with a diameter less than 2.5 micrometers in micrograms per normal cubic meter ($\mu g/Nm^3$). PM2.5 is a commonly monitored pollutant and is known to have detrimental effects on human health; (ii) Median of PM10 concentrations: this variable represents the median concentration of particulate matter with a diameter smaller than 10 micrometers (PM10) in $\mu g/Nm^3$. PM10 includes both fine and coarse particles and is also considered a significant air pollutant; (iii) Ambient temperature (TEMP): this variable represents the temperature at the monitoring station in degrees Celsius. Temperature is an important meteorological parameter that can influence air quality and pollutant levels; (iv) Wind speed (WIND): this variable represents the speed of wind at the monitoring station in meters per second. Wind speed plays a crucial role in the dispersion and transport of pollutants in the atmosphere; (v) Relative air humidity (HR): this variable represents the percentage of moisture in the air at the AQMS. Humidity can affect atmospheric stability and the formation of certain pollutants. By analyzing these variables, we can gain insights into the relationship between air pollution levels and meteorological conditions in Padre Las Casas during the specified period.

In the exploratory data analysis (EDA) of the median PM2.5 concentrations recorded by the Padre Las Casas AQMS during June–July 2020, Figure 1a shows a histogram with density kernel estimation. This plot provides an overview of the distribution of the data, and permits us to visualize the shape of the empirical distribution. From the histogram, it appears that the distribution of the PM2.5 concentrations has a positive skewness, indicating

that most of the observations have lower values with a few extremely high values. Figure 1b presents a boxplot for the median PM2.5 concentrations. From the boxplot, we can see that there are some observations labeled as atypical data (#1, #3, #4, #14, #36, #45) that lie outside the whiskers. These observations deviate from the overall pattern of the data and may represent extreme or unusual values. This suggests that there may be some extreme pollution events or unusual conditions during the observed period. Based on the positive skewness of the empirical distribution and the presence of atypical data points, it is reasonable to consider using log-symmetrical distributions to model the PM2.5 concentrations. Log-symmetrical distributions can better capture the positive skewness and accommodate the potential presence of extreme values in the data.



(**a**) (**b**)

Figure 1. Histogram with density kernel estimation (solid black line) (**a**) and boxplot (**b**) for median PM2.5 concentrations recorded by Padre Las Casas AQMS during June–July 2020.

Table 1 provides descriptive statistics for the median PM2.5 concentrations recorded by the Padre Las Casas AQMS during June-July 2020. These statistics include measures of central tendency (mean, median), dispersion (range, standard deviation –SD–), as well as coefficients of skewness (CS) and kurtosis (CK). The descriptive statistics reveal that the median PM2.5 concentrations have a mean of 43.4 $\mu g/Nm^3$ and a median of 36.0 $\mu g/Nm^3$. The SD is relatively high, with a value of 26.0 $\mu g/Nm^3$, indicating substantial variability in the data. The CS is 1.3, indicating a positive skewness and confirming the observation from the histogram in Figure 1a. The positive skewness suggests that most of the observations have lower values, while a few extremely high values contribute to the right tail of the distribution. The CK is 0.8, which indicates a moderately peaked distribution compared to a normal distribution. Furthermore, as mentioned in the text, a significant quantity of levels that surpass the recommended Chilean thresholds for PM2.5, set at 50 $\mu g/Nm^3$. This suggests that the air pollution level in Padre Las Casas is dangerous from a toxicological perspective, posing potential health risks for the inhabitants of this commune in southern Chile. Overall, the descriptive statistics and Figure 1a,b provide evidence of the high pollution levels and the need for modeling approaches that can adequately capture the characteristics of the PM2.5 concentrations in this region.

Table 1. Descriptive statistics for median PM2.5 concentrations recorded by Padre Las Casas AQMS during June–July 2020.

Variable	*n*	Min	Max	Range	Mean	Median	SD	CS	CK
PM2.5	61	15	121.5	106.5	43.4	36.0	26.0	1.3	0.8

Figure 2 shows a correlation matrix for PM2.5, PM10, TEMP, WIND, and HR. From this figure, we detect: (i) a high positive association between PM2.5 and PM10 (Pearson coefficient of correlation equal to 0.99); (ii) medium negative association between PM2.5 and TEMP and WIND (Pearson coefficient of correlation equal to -0.70); (iii) low positive association between PM2.5 and HR (Pearson coefficient of correlation equal to 0.38). In

Figure 3, scatter plots depicting the explanatory variables, response variable, and potential interactions among the explanatory variables are presented. In Figure 3a, note that the relationship between PM2.5 and PM10 is linear, while in Figure 3b, the relationship between PM2.5 and WIND is not linear. Furthermore, Figure 3c,d imply that the explanatory variables TEMP and HR may be engaging with the WIND variable in a nonlinear manner.

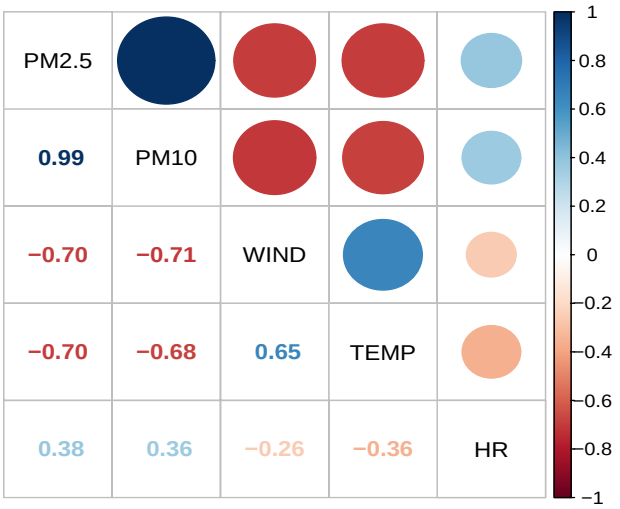


Figure 2. Correlation matrix displaying the respective Pearson correlation coefficient for the specified explanatory and response variables.

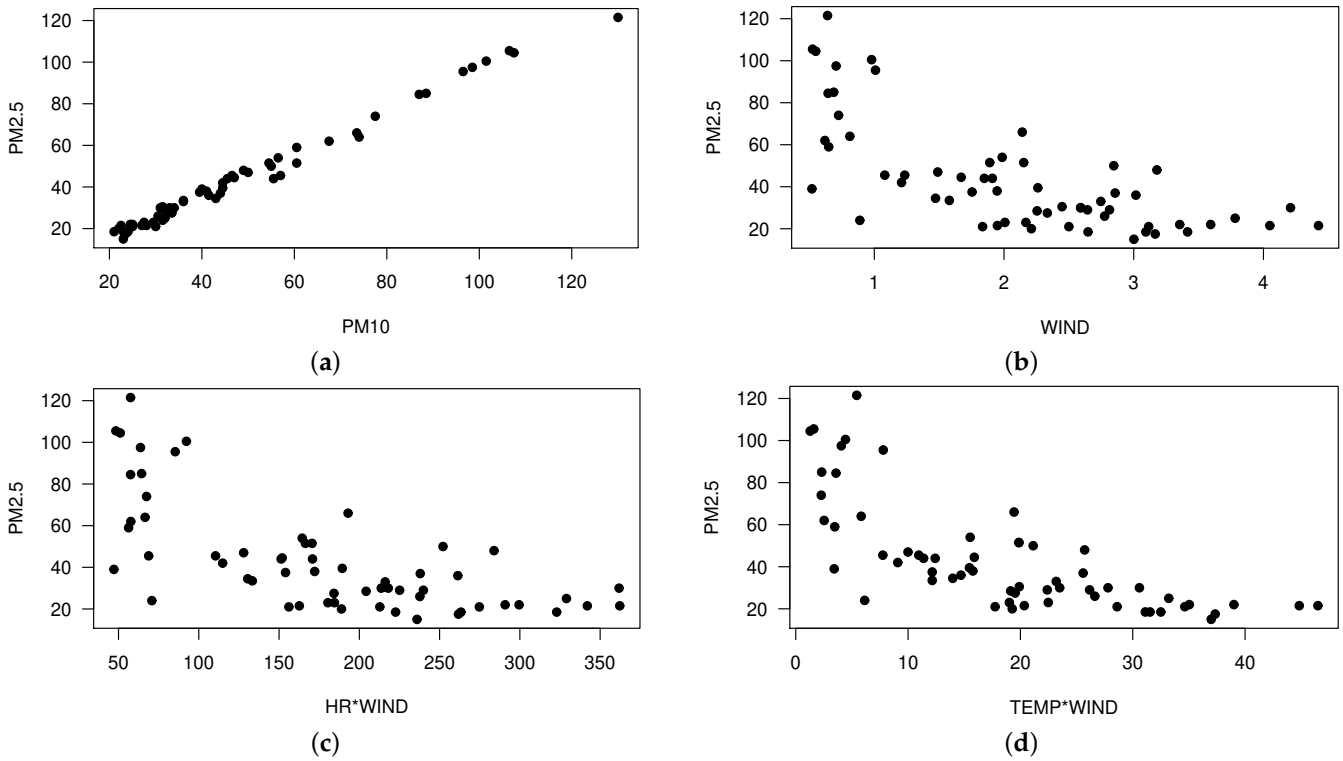


Figure 3. Scatter plots for median PM2.5 vs. PM10 concentrations (**a**); median PM2.5 vs. WIND (**b**); median PM2.5 vs. HR*WIND (**c**); and, median PM2.5 vs. TEMP*WIND (**d**) recorded by Padre Las Casas AQMS during June–July 2020.

5.2. Parameter Estimation

Based on the EDA and the observed relationships between the median PM2.5 concentration and the variables as PM10, WIND, TEMP, and HR, we suggest the following varying-coefficients quantile regression models to capture the trends:

$$\sqrt{Q_i} = \boldsymbol{w}_i^\top \boldsymbol{\alpha} + x_{1i}\beta_1(t_i) + x_{2i}\beta_2(t_i), \quad i \in \{1, 2, \ldots, 61\} \tag{10}$$

where $y_i \sim \text{QLS}(Q_i, \phi, g)$ with Student-t and normal PDF generator g, $\boldsymbol{\beta}$ represents the vector of regression coefficients, while $\boldsymbol{w}_i^\top = (1, w_{1i})^\top$ with w_{1i} denoting the values of the parametric covariate for the ith observation (PM10). The coefficients β_k(for $k \in \{1, 2\}$) correspond to unknown, smooth, and arbitrary functions of the explanatory variable t_i (WIND), which are linked to the explanatory variables x_{1i} (TEMP) and x_{2i} (HR) from the ith case. These varying-coefficients quantile regression models allow for a more flexible and comprehensive characterization of the relationships between the median PM2.5 concentration and the other variables, considering potential variations across quantiles.

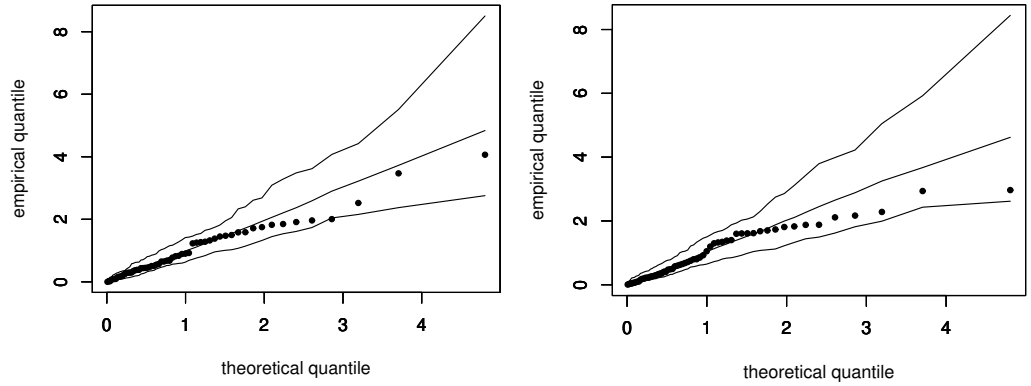
Table 2 presents the MPL estimates for the model parameters, their approximate standard errors (SEs), p-values obtained from a z-test, the AIC, selected smoothing parameters, and the degrees of freedom $\text{df}(\cdot)$ for the models defined by Equation (10). The best values of λ_1 and λ_2 were selected by considering a grid of values and choosing those that yielded a range of $\text{df}(\lambda_1)$ and $\text{df}(\lambda_2)$ within the range of $(4, 12)$, while minimizing the AIC value.

When comparing the results reported in Table 2, we observe that the estimates for α_0 and α_1 show similarity between both models, but the log-t model has smaller estimated standard errors (SEs) for these parameters compared to the log-normal model. Additionally, the estimated value of ϕ in the log-t model is smaller than that in the log-normal model. It is worth noting that based on the (AIC), the log-t model is preferred as it yields a lower AIC value.

Table 2. MPL estimates, SEs, p-values, AIC and selected smoothing parameters and $\text{df}(\cdot)$ of the indicated model.

Model	Parameter	Estimate	SE	p-Value	AIC
Log-normal	α_0	3.072	2.2×10^{-5}	<0.001	374.1
	α_1	0.068	1.1×10^{-3}	<0.001	
	ϕ	0.013	4.1×10^{-6}		
	λ_1	4034.3			
	λ_2	2.2×10^5			
	$\text{df}(\lambda_1)$	4.001			
	$\text{df}(\lambda_2)$	4.466			
Log-$t(\nu = 4)$	α_0	3.052	1.7×10^{-5}	<0.001	361.3
	α_1	0.070	8.3×10^{-4}	<0.001	
	ϕ	0.007	4.9×10^{-6}		
	λ_1	4034.3			
	λ_2	5.9×10^5			
	$\text{df}(\lambda_1)$	4.556			
	$\text{df}(\lambda_2)$	4.198			

To assess the distributional assumption made in the model, we examine the goodness-of-fit plots based on generalized Cox-Snell (GCS) residuals, as shown in Figure 4. Additionally, we provide the p-values associated with the Kolmogorov–Smirnov (KS) test, which are 0.73 for the log-normal model and 0.89 for the log-$t(\nu = 4)$ model. Based on the goodness-of-fit plots, the KS test, and the AIC, we can conclude that the log-$t(\nu = 4)$ model provides a better fit to the dataset. The log-t model captures the underlying distribution of the data more accurately compared to the log-normal model, as indicated by the higher p-value and better fit observed in the goodness-of-fit plots.



(**a**) Log-normal model (**b**) Log-t ($\nu = 4$) model

Figure 4. Goodness-of-fit plots with simulated envelope for GCS residual under the indicated model with the analyzed data set.

Figure 5 displays the plots of the partial residuals relative to the WIND covariate, with the superimposed estimated smooth functions β_1 (on the left) and β_2 (on the right). The behavior of the partial residuals (dots) in these plots appears reasonable, indicating that the fit of the log-$t(\nu = 4)$ varying-coefficients quantile regression model to the pollution dataset is adequate. The dots are closely aligned with the estimated curves, as expected, suggesting that the model captures the relationship between the WIND covariate and the partial residuals effectively. This agreement between the partial residuals and the estimated curves supports the appropriateness of the log-$t(\nu = 4)$ varying-coefficients quantile regression model for analyzing the pollution data.

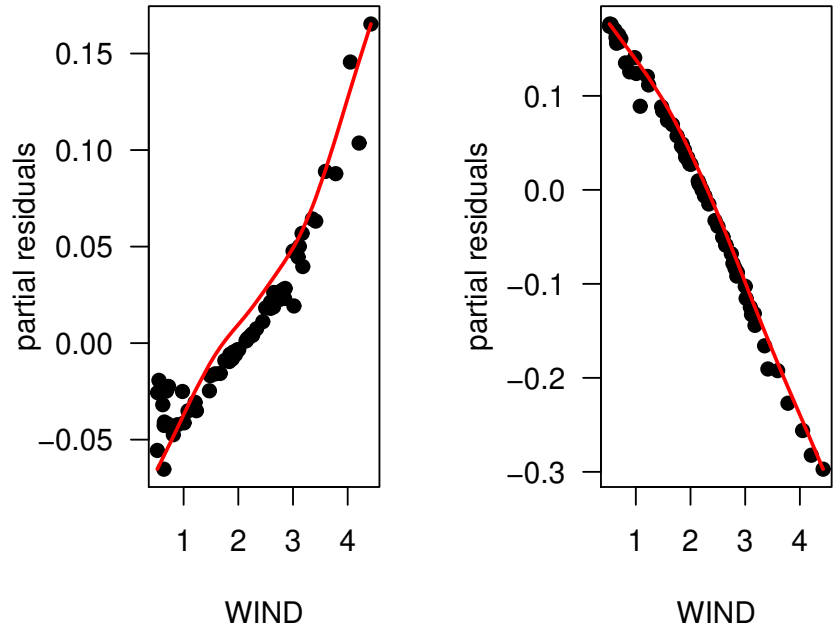


Figure 5. Plots of partial residuals in relation to the WIND covariate, with the estimated smooth functions β_1 (on the left) and β_2 (on the right) superimposed.

5.3. Diagnostic Analysis

In this section, we investigate the potential influence of individual observations using the local influence method for the selected varying-coefficients quantile regression model. We consider four perturbation schemes: case-weight perturbation, response variable perturbation, power parameter perturbation, and explanatory variable perturbation. Additionally, we examine the GL to assess the influence of each observed value on its own predicted

value. These analyses allow us to identify potentially influential cases and understand their impact on the selected model. Details on the local influence method and the perturbation schemes can be found in Section 4.2.

In Figure 6, we present index plots illustrating $C_i(\boldsymbol{\theta})$ as defined in Section 4.2 for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and ϕ under the case-weight perturbation (a,b,c,d), under response perturbation (e,f,g,h) and perturbation on the power parameter (i,j,k,l) schemes. Also, Figure 7 showcases the index plots of $C_i(\boldsymbol{\theta})$ when introducing perturbations in covariates X_1 (a, b, c, d) and X_2 (e, f, g, h). Despite different observations being detected as potentially influential, it is worth noting that there are four cases (#13, #18, #31, and #45) that consistently appear as potentially influential across multiple perturbation schemes. These cases exhibit characteristics that make them stand out and have a notable impact on the model results.

Figure 8 displays the GL plot, which assesses the influence of each observation on its own predicted value. From this plot, we observe that cases #45, #36, #14, #1, #3, #4 are potentially leverage points. These observations have response variable values that can exert a significant influence on their own predicted values. It is worth noting that these cases correspond to the outliers identified by the boxplot in Figure 1b. Their extreme values contribute to their influential nature within the model.

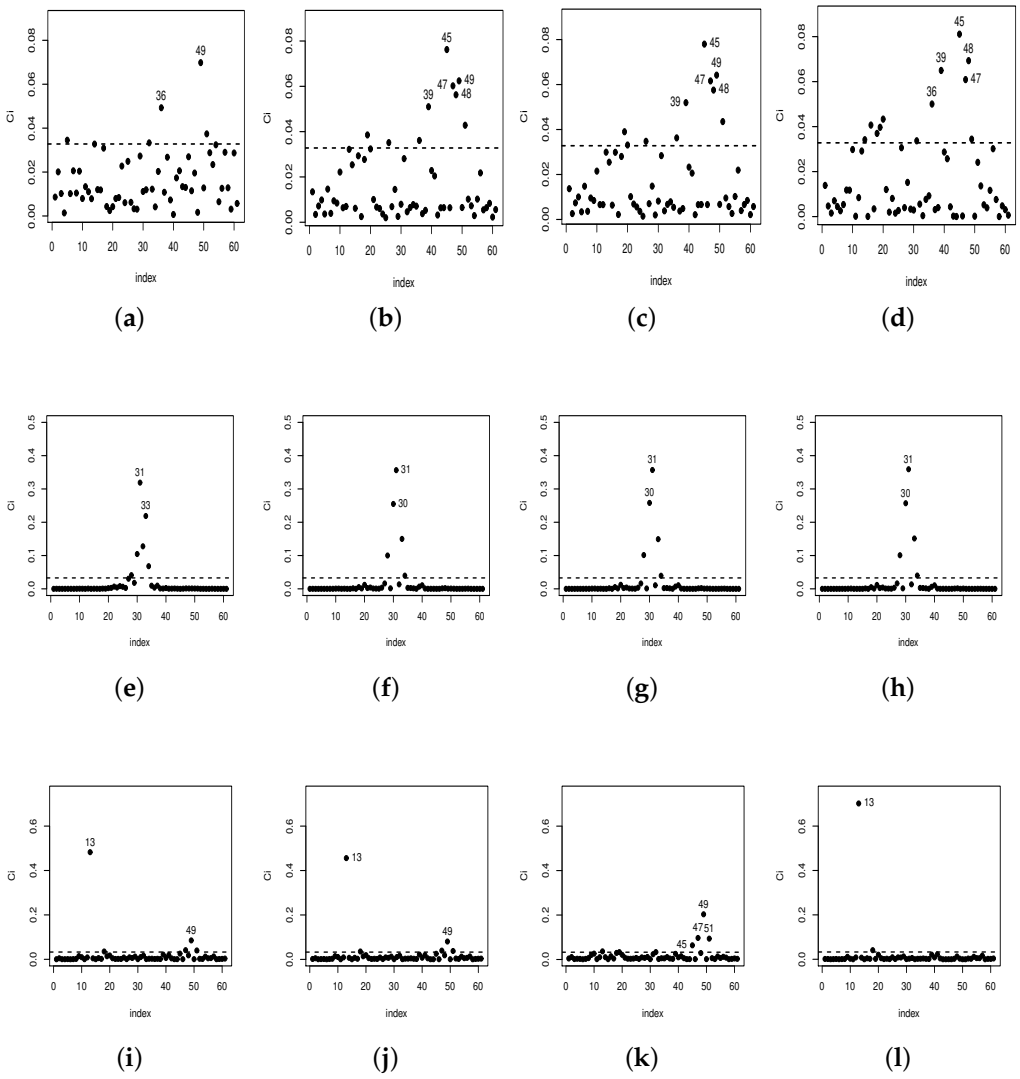


Figure 6. Case weight (**a**–**d**), response (**e**–**h**) and on the power parameter (**i**–**l**) perturbation for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and ϕ.

It is interesting to observe that the cases identified as potentially influential in the parametric component may not necessarily be detected in the nonparametric component, and vice versa. This indicates that different aspects of the data and model may be driving their influence in different ways. Additionally, the local influence analysis technique has successfully detected some atypical cases that were previously identified as outliers in Figure 1b. This reinforces the effectiveness of the local influence method in identifying observations that have a considerable impact on the model.

In the sense of evaluating the impact of these observations in the selected model, the subsets of cases {#13}, {#18}, {#31}, {#45}, {#13, #18}, {#13, #31}, {#13, #45}, {#18, #31}, {#18, #45}, {#31, #45}, and {#13, #18, #31}, {#13, #18, #45}, {#18, #31, #45} and {#13, #18, #31, #45} are removed and the model parameters are re-estimated. To determine the variation in the estimates of model parameters, we use the value of the relative changes (RCs) for each parameter. The RCs for each estimated parameter are calculated using the formula: $\mathrm{RC}_\theta = \left|(\hat{\theta}_j - \hat{\theta}_{j(i)})/\hat{\theta}_j\right| \times 100\%$, where $\hat{\theta}_j$ represents the MPLE of θ_j, and $\hat{\theta}_{j(i)}$ represents the MPLE of θ_j after removing the subset i of observations. Here, $j = 1, 2, 3$ with $\theta_1 = \alpha_0$, $\theta_2 = \alpha_1$, and $\theta_3 = \phi$.

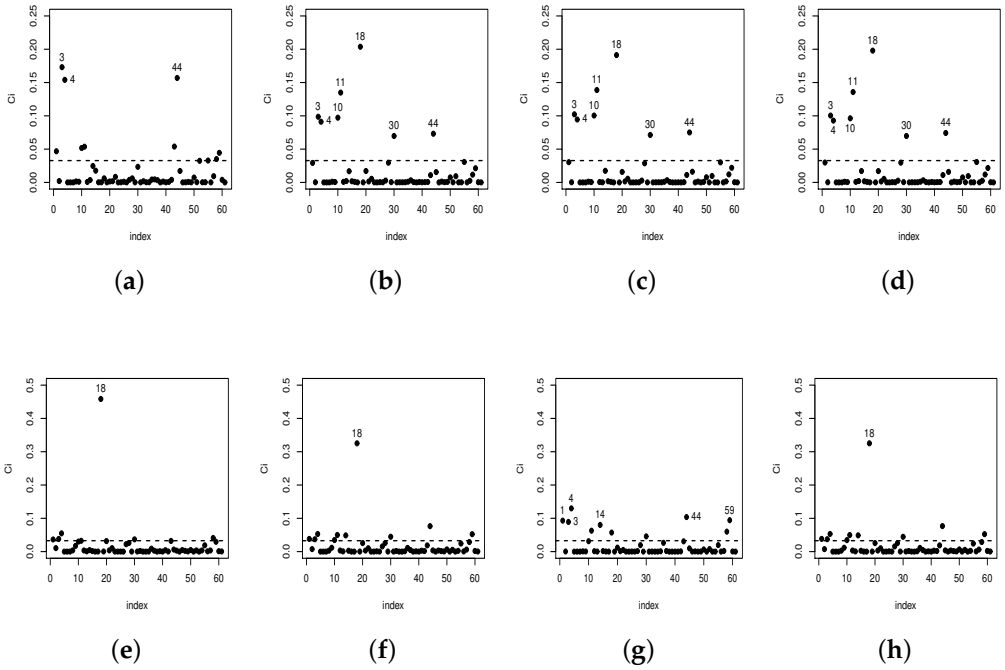


Figure 7. Perturbationin the covariate X_1 (**a**–**d**) and X_2 (**e**–**h**) scheme for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and ϕ.

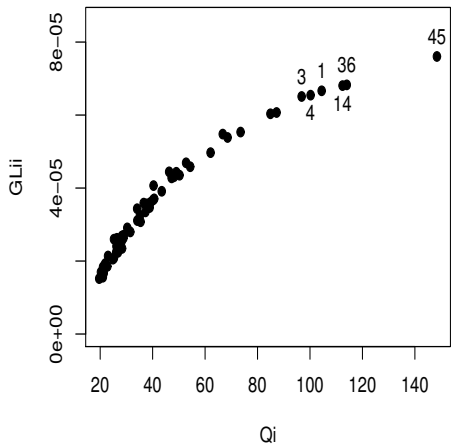


Figure 8. Generalized leverage.

Table 3 reports the values of RCs for the varying-coefficients quantile regression model after removing the indicated sets of cases. In this table, the individual elimination of cases #13 and #45 produces a RC in α_0 and α_1 of 5.1%, 4.7% and 5.3%, 5.6%, respectively, while the elimination of case #18 produces an RC in ϕ of 5.5%. In addition, note that set of cases {#13, #18} and {#13, #18, #31} produces the largest RCs in α_0, α_1 and ϕ.

During the analyzed period, it was observed that observation #45 had particularly high concentrations of PM2.5 and PM10 compared to other observations. On the other hand, observation purple #31 had a very low wind speed, close to the minimum recorded during the entire period. These observations exhibit extreme values in their respective covariates. When the sets of potentially influential cases {#13, #18, #31, #45} are excluded from the analysis, it is observed that their removal results in notable alterations solely in the estimation of ϕ, displaying a percentage change of 21.4%. This suggests that these observations have a notable influence on the estimation of the dispersion parameter ϕ in the model.

Table 3. RC (in %) on the MPL estimate of α_j and ϕ and respective p-values (in parenthesis) for varying-coefficients quantile regression model after removing the indicated sets of cases.

	Parameters			**Relative Changes**		
Removed Case	$\boldsymbol{\alpha_0}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\phi}$	$\mathbf{RC}_{\boldsymbol{\alpha_0}}$	$\mathbf{RC}_{\boldsymbol{\alpha_1}}$	$\mathbf{RC}_{\boldsymbol{\phi}}$
none	3.052 (<0.001)	0.069 (<0.001)	0.007			
{#13}	3.213 (<0.001)	0.066 (<0.001)	0.007	5.1%	4.7%	4.0%
{#18}	2.961 (<0.001)	0.072 (<0.001)	0.006	3.0%	3.4%	5.5%
{#31}	3.095 (<0.001)	0.069 (<0.001)	0.007	1.4%	1.1%	3.9%
{#45}	2.891 (<0.001)	0.073 (<0.001)	0.006	5.3%	5.6%	4.0%
{#13, #18}	3.415 (<0.001)	0.065 (<0.001)	0.006	11.9%	6.7%	18.0%
{#13, #31}	3.223 (<0.001)	0.066 (<0.001)	0.006	5.6%	4.7%	9.4%
{#13, #45}	3.093 (<0.001)	0.069 (<0.001)	0.006	1.4%	0.4%	13.0%
{#18, #31}	3.011 (<0.001)	0.071 (<0.001)	0.006	1.3%	2.2%	9.5%
{#18, #45}	2.901 (<0.001)	0.073 (<0.001)	0.006	4.9%	5.4%	10.9%
{#13, #18, #31}	3.488 (<0.001)	0.064 (<0.001)	0.005	14.3%	7.8%	20.5%
{#13, #18, #45}	3.005 (<0.001)	0.071 (<0.001)	0.006	1.5%	2.8%	17.4%
{#18, #31, #45}	2.960 (<0.001)	0.072 (<0.001)	0.006	3.0%	4.0%	14.5%
{#13, #18, #31, #45}	3.046 (<0.001)	0.071 (<0.001)	0.005	0.2%	1.9%	21.4%

Finally, in Table 3, while certain RCs exhibit considerable values, there are no substantial alterations in inference, as evidenced by the diminutive p-values (less than 0.001) associated with the parameter estimates. It is important to note that when observations detected as influential in the diagnostic plots are eliminated, it can lead to significant changes in the parameter estimates. This indicates that the well-known robustness properties of maximum likelihood estimates from Student-t models may not necessarily apply to other perturbation schemes. Therefore, it is crucial to conduct diagnostic examinations specific to each case to properly assess the influence of observations and ensure the reliability of the model estimates.

6. Discussion, Conclusions and Future Research

In this work, we propose new varying-coefficients semiparametric quantile regression models based on the family of log-symmetric distributions, following the approach of [5–7]. By reparametrizing the family of log-symmetric distributions using a quantile, we introduce new quantile models that offer greater flexibility in modeling data compared to the model proposed by Saulo et al. [7], as a nonparametric component has been added (Section 2). We develop parameter estimation based on the penalized likelihood function and propose a back-fitting iterative algorithm implemented in the R language (Section 3). Additionally, we discuss diagnostic techniques for detecting potentially influential local observations and identifying leverage points (Section 4). Please note that the local influence analysis reinforces the need for diagnostic evaluation. It has been observed that parameter estimators in this class of models tend to be sensitive to the presence of atypical or influential data points. To the best of our knowledge, techniques for detecting leverage points have not been developed for semiparametric quantile regression models.

We illustrate the methodology developed in this work using data associated with PM2.5 pollution in Padre Las Casas city for predicting the daily median of 1-h average values. We propose and fit two models (log-normal and log-$t(\nu = 4)$) and evaluate them using CGS residuals and their AIC values. The plots of CGS residuals and partial residuals show a good fit of the selected model (log-$t(\nu = 4)$) to the data. We also apply our diagnostic techniques to detect potentially influential cases and leverage points (Section 4.2); however, no inferential changes are observed in the parameter estimates.

Thus, among the accomplishments of this work, we can highlight: (i) The development of novel quantile regression models suitable for modeling data following asymmetric distributions, which can be added into the existing toolkit for quantile modeling; (ii) The expansion of our model beyond the one presented in [7], incorporating nonlinear structures arising from interaction effects. (iii) The derivation of analytical tools for identifying potentially influential observations and leverage points.

One limitation of our study is that the proposed models may not be suitable for describing other types of data, such as temporally or spatially correlated data, as well as censored data. In such cases, the utilization of multivariate distributions for the response variable, reparametrized by quantiles of marginal distributions, may be necessary. Another area for future investigation is conducting a simulation study to evaluate the distributional behavior of the residuals used in this study and exploring alternative types of residuals appropriate for this type of regression. This aspect has received limited attention in the existing literature. Furthermore, we aim to establish inferences about the model parameters through asymptotic analysis of specific estimators. Lastly, we intend to compare our model with others, including models proposed by [7,12]. These are additional areas that remain unexplored, and we plan to address these open questions in our future research.

Author Contributions: Conceptualization, L.S. and G.I.-P.; methodology, L.S., G.I.-P. and C.M.; software, L.S. and C.M.; validation, L.S. and C.M.; formal analysis, L.S. and C.M.; investigation, L.S., G.I.-P. and C.M.; resources, C.M. and M.R.; data curation, L.S. and C.M.; writing—original draft preparation, L.S., G.I.-P. and C.M.; writing—review and editing, L.S., G.I.-P., C.M. and M.R.; visualization, L.S. and C.M. All authors have read and agreed to the published version of the manuscript.

Funding: The research was partially funded by FONDECYT, project grant number 11190636 (C.M.) from the National Agency for Research and Development (ANID) of the Chilean government under the Ministry of Science, Technology, Knowledge and Innovation.

Data Availability Statement: Data and computational codes are available upon request from the authors.

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

Appendix A

Here, we present the quantities c_i, m_i, and d_i, involved in the definition of the Penalized Hessian matrix presented in Section 3.2. In fact, we have

$$\begin{aligned}
c_i &= \frac{1}{\sqrt{\phi}}\left[-\frac{r(v_i)}{Q_i^2\sqrt{\phi}} + \frac{v_i}{Q_i}\frac{\partial\, r(v_i)}{\partial Q_i} - \frac{v_i r(v_i)}{Q_i^2}\right]a_i^2 - z_i\, a_i \frac{h''(Q_i)}{(h'(Q_i))^2},\\
m_i &= \frac{1}{Q_i}\left[\frac{\phi^{-2}}{2}(\log(Q_i) - \log(y_i))\, r(v_i) + \frac{\partial r(v_i)}{\partial \phi} v_i\, \phi^{-1/2} - \frac{1}{2} v_i r(v_i)\, \phi^{-3/2}\right], \quad \text{and}\\
d_i &= \frac{1}{2}[\log(y_i) - \log(Q_i)]\phi^{-3/2}\left[v_i\, \frac{\partial\, r(v_i)}{\partial \phi} - \frac{3}{2} v_i\, r(v_i)\phi^{-1} - \frac{1}{2} r(v_i)(\log(y_i)\right.\\
&\quad \left. - \log(Q_i))\phi^{-3/2}\right] + \frac{1}{2\phi^2}.
\end{aligned}$$

In addition, the expression $\partial\, r(v_i)/\partial Q_i$ and $\partial\, r(v_i)/\partial \phi$ are, respectively,

$$\begin{aligned}
\frac{\partial\, r(v_i)}{\partial Q_i} &= 4\left[\frac{g''(v_i^2)g(v_i^2) - (g'(v_i^2))^2}{(g(v_i^2))^2}\right]\frac{v_i}{Q_i\sqrt{\phi}}, \quad \text{and}\\
\frac{\partial\, r(v_i)}{\partial \phi} &= 2\left[\frac{g''(v_i^2)g(v_i^2) - (g'(v_i^2))^2}{(g(v_i^2))^2}\right] v_i \phi^{-3/2}[\log(y_i) - \log(Q_i)].
\end{aligned}$$

Appendix B

Here we present the matrix $\mathbf{\Delta}_{\mathrm{p}}$ for four perturbation schemes, namely case weight, response variable, power parameter, and explanatory variable perturbation. In general, this matrix is defined as

$$\mathbf{\Delta}_{\mathrm{p}} = \begin{pmatrix} \mathbf{\Delta}_\alpha^\top & \mathbf{\Delta}_{\beta_1}^\top & \ldots & \mathbf{\Delta}_{\beta_s}^\top & \mathbf{\Delta}_\phi^\top \end{pmatrix}^\top.$$

Appendix B.1. Case-Weight Perturbation

Here, the elements of the matrix $\mathbf{\Delta}_{\mathrm{p}}$ are given by

$$\begin{aligned}
\mathbf{\Delta}_\alpha &= \boldsymbol{W}^\top \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\boldsymbol{z}},\\
\mathbf{\Delta}_{\beta_k} &= \tilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\boldsymbol{z}}, \quad \text{for} \quad k \in \{1, \ldots, s\},\\
\mathbf{\Delta}_\phi &= \widehat{\boldsymbol{b}},
\end{aligned}$$

with $\widehat{\boldsymbol{D}}_a$, $\widehat{\boldsymbol{D}}_{\boldsymbol{z}}$ and $\widehat{\boldsymbol{b}}$ correspond to $\boldsymbol{D}_{\boldsymbol{a}}$, $\boldsymbol{D}_{\boldsymbol{z}}$ and $\boldsymbol{b} = (b_1, \ldots, b_n)^\top$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0 = (1, \ldots, 1)^\top$, respectively.

Appendix B.2. Response Variable Perturbation

Under this perturbation schemes, the elements of the matrix $\mathbf{\Delta}_{\mathrm{p}}$ are given by $\mathbf{\Delta}_\alpha = \boldsymbol{W}^\top \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_\psi \widehat{\boldsymbol{D}}_\vartheta$, $\mathbf{\Delta}_{\beta_k} = \tilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_\psi \widehat{\boldsymbol{D}}_\vartheta$, for $k \in \{1, \ldots, s\}$, $\mathbf{\Delta}_\phi = \widehat{\boldsymbol{\tau}}^\top \boldsymbol{D}_\vartheta$, with $\widehat{\boldsymbol{D}}_\vartheta = \mathrm{diag}\{\widehat{\vartheta}_1, \ldots, \widehat{\vartheta}_n\}$, $\widehat{\boldsymbol{D}}_\psi = \mathrm{diag}\{\widehat{\psi}_1, \ldots, \widehat{\psi}_n\}$, and $\widehat{\boldsymbol{\tau}} = (\widehat{\tau}_1, \ldots, \widehat{\tau}_n)^\top$, with

$$\begin{aligned}
\widehat{\vartheta}_i &= s_{Y_i},\\
\widehat{\psi}_i &= \frac{1}{\widehat{\phi}\, \widehat{Q}_i\, y_i}\left[r(\widehat{v}_i) + \widehat{v}_i\, r'(\widehat{v}_i)\right],\\
\widehat{\tau}_i &= -\frac{\widehat{\phi}^{-3/2}}{2}\left[\left[v_i\, r'(\widehat{v}_i) + r(\widehat{v}_i)\right]\left[\frac{\log(y_i) - \log(\widehat{Q}_i)}{y_i\sqrt{\widehat{\phi}}}\right] + \frac{r(\widehat{v}_i)\, \widehat{v}_i}{y_i}\right], \quad i \in \{1, \ldots, n\},
\end{aligned}$$

and $r'(\widehat{v}_i) = dr(\widehat{v}_i)/d\widehat{v}_i$. In this case, $\widehat{v}_i$, $\widehat{Q}_i$ and $\widehat{\phi}$ are evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = (0, \ldots, 0)^\top$.

Appendix B.3. Power Parameter Perturbation

Considering the power parameter perturbation, the elements of the matrix $\mathbf{\Delta}_{\mathrm{p}}$ are given by $\mathbf{\Delta}_{\beta_k} = \tilde{\boldsymbol{N}}_k^{\top} \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\varpi}$, for $k \in \{1, \ldots, s\}$, $\mathbf{\Delta}_{\phi} = \widehat{\boldsymbol{\varphi}}^{\top}$, where $\widehat{\boldsymbol{D}}_{\varpi} = \mathrm{diag}\{\widehat{\varpi}_1, \ldots, \widehat{\varpi}_n\}$ and $\boldsymbol{\varphi} = (\widehat{\varphi}_1, \ldots, \widehat{\varphi}_n)^{\top}$, with $\varpi_i = -\widehat{\phi}\,\widehat{m}_i$ and $\widehat{\varphi}_i = -\widehat{\phi}\,\widehat{d}_i$, for $i \in \{1, \ldots, n\}$. Here, $\widehat{m}_i$ and $\widehat{d}_i$ correspond to m_i and d_i evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0 = (1, \ldots, 1)^{\top}$, respectively.

Appendix B.4. Explanatory Variable Perturbation

In this case, the elements of the matrix $\mathbf{\Delta}_p$ can be expressed as follows:

(*i*) for $l = k$,

$$\begin{aligned}
\mathbf{\Delta}_{\alpha} &= \boldsymbol{W}^{\top} \big(\widehat{\boldsymbol{D}}_{a'} \widehat{\boldsymbol{D}}_z + \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_c\big) \widehat{\boldsymbol{D}}_a \, s_{X_l} \, \widehat{\boldsymbol{D}}_{\tilde{\boldsymbol{N}}_l \boldsymbol{f}_l}, \\
\mathbf{\Delta}_{\beta_l} &= \tilde{\boldsymbol{N}}_l \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_z s_{X_l} + \tilde{\boldsymbol{N}}_l^{\top} \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\tilde{\boldsymbol{N}}_l \boldsymbol{f}_l} s_{X_l} \big(\widehat{\boldsymbol{D}}_{a'} \widehat{\boldsymbol{D}}_z + \widehat{\boldsymbol{D}}_c\big) - \lambda_l \boldsymbol{K}_l \widehat{\boldsymbol{\beta}}_l, \text{ for } k \in \{1, \ldots, s\}, \\
\mathbf{\Delta}_{\phi} &= \widehat{\boldsymbol{m}}^{\top} \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\tilde{\boldsymbol{N}}_l \boldsymbol{f}_l} s_{X_l};
\end{aligned}$$

(*ii*) for $l \neq k$,

$$\begin{aligned}
\mathbf{\Delta}_{\alpha} &= \boldsymbol{W}^{\top} \big(\widehat{\boldsymbol{D}}_{a'} \widehat{\boldsymbol{D}}_z + \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_c\big) \widehat{\boldsymbol{D}}_a \, s_{X_l} \, \widehat{\boldsymbol{D}}_{\tilde{N}_l \cdot \boldsymbol{f}_l}, \\
\mathbf{\Delta}_{\beta_l} &= \tilde{\boldsymbol{N}}_1^{\top} \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\tilde{N}_l \boldsymbol{\beta}_l} s_{X_l} \big(\widehat{\boldsymbol{D}}_{a'} \widehat{\boldsymbol{D}}_z + \widehat{\boldsymbol{D}}_c\big) - \lambda_1 \boldsymbol{K}_1 \widehat{\boldsymbol{\beta}}_1, \quad \text{for} \quad k \in \{1, \ldots, s\}, \\
\mathbf{\Delta}_{\phi} &= \widehat{\boldsymbol{m}}^{\top} \widehat{\boldsymbol{D}}_a \widehat{\boldsymbol{D}}_{\tilde{N}_l \boldsymbol{\beta}_l} s_{X_l}.
\end{aligned}$$

where $\boldsymbol{D}_{a'} = \mathrm{diag}\{a'_1, \ldots, a'_n\}$, with $a'_i = da_i/dQ_i$, and $\boldsymbol{D}_{\tilde{N}_l \boldsymbol{\beta}_l}$ is the diagonalization of the vector $\tilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l$. Here, $\boldsymbol{\omega}_0 = (0, \ldots, 0)^T$ corresponds to the vector of no perturbation.

References

1. Vanegas, L.; Paula, G. A semiparametric approach for joint modeling of median and skewness. *Test* **2015**, *24*, 110–135. [CrossRef]
2. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. A New Class of Skew-Normal Distributions. *Commun. Stat. Theory Methods* **2004**, *33*, 1465–1480. [CrossRef]
3. Paula, G.A.; Leiva, V.; Barros, M.; Liu, S. Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Appl. Stoch. Model. Bus. Ind.* **2012**, *28*, 16–34. [CrossRef]
4. Leiva, V.; Santos-Neto, M.; Cysneiros, F.J.A.; Barros, M. Birnbaum–Saunders statistical modelling: A new approach. *Stat. Model.* **2014**, *14*, 21–48. [CrossRef]
5. Sánchez, L.; Leiva, V.; Galea, M.; Saulo, H. Birnbaum-Saunders quantile regression and its diagnostics with application to economic data. *Appl. Stoch. Model. Bus. Ind.* **2021**, *37*, 53–73. [CrossRef]
6. Sánchez, L.; Leiva, V.; Marchant, C.; Saulo, H.; Sarabia, J.M. A new quantile regression model and its diagnostic analytics for a Weibull distributed response with applications. *Mathematics* **2021**, *9*, 2768. [CrossRef]
7. Saulo, H.; Dasilva, A.; Leiva, V.; Sanchez, L.; de la Fuente-Mella, H. Log-symmetric quantile regression models. *Stat. Neerl.* **2022**, *76*, 124–163. [CrossRef]
8. Vanegas, L.; Paula, G. Log-symmetric distributions: Statistical properties and parameter estimation. *Braz. J. Probab. Stat.* **2016**, *30*, 196–220. [CrossRef]
9. Vanegas, L.; Paula, G. An extension of log-symmetric regression models: R codes and applications. *J. Stat. Simul. Comput.* **2016**, *86*, 1709–1735. [CrossRef]
10. Ventura, M.; Saulo, H.; Leiva, V.; Monsueto, S.E. Log-symmetric regression models: Information criteria and application to movie business and industry data. *Appl. Stoch. Model. Bus. Ind.* **2019**, *35*, 963–977. [CrossRef]
11. Hao, L.; Naiman, D.Q. *Quantile Regression*; Sage Publications: London, UK, 2007.
12. Koenker, R.; Chernozhukov, V.; He, X.; Peng, L. *Handbook of Quantile Regression*; CRC Press: Boca Raton, FL, USA, 2018.
13. Noufaily, A.; Jones, M.C. Parametric quantile regression based on the generalized gamma distribution. *J. R. Stat. Soc. Ser.* **2013**, *62*, 723–740. [CrossRef]
14. Hastie, T.; Tibshirani, R. *Generalized Additive Models*; Chapman and Hall: New York, NY, USA, 1990.
15. Green, P.J.; Silverman, B.W. *Nonparametric Regression and Generalized Linear Models*; Chapman and Hall: Boca Raton, FL, USA, 1994.
16. Ibacache-Pulgar, G.; Paula, G.A.; Cysneiros, F.J.A. Semiparametric additive models under symmetric distributions. *Test* **2013**, *22*, 103–121. [CrossRef]
17. Ramires, T.; Ortega, E.; Hens, N.; Cordeiro, G.; Paula, G. A flexible semiparametric regression model for bimodal, asymmetric and censored data. *J. Appl. Stat.* **2018**, *45*, 1303–1324. [CrossRef]

18. Manghi, R.; Cysneiros, F.J.A.; Paula, G. Generalized additive partial linear models for analyzing correlated data. *Comput. Stat. Data Anal.* **2019**, *129*, 47–60. [CrossRef]
19. Oliveira, R.A.; Paula, G.A. Additive models with autoregressive symmetric errors based on penalized regression splines. *Comput. Stat.* **2021**, *36*, 2435–2466. [CrossRef]
20. Ferreira, C.; Montoril, M.; Paula, G. Partially linear models with p-order autoregressive skew-normal errors. *Braz. J. Probab. Stat.* **2022**, *36*, 792–806.
21. Cardozo, C.A.; Paula, G.A.; Vanegas, L.H. Generalized log-gamma additive partial linear models with P-spline smoothing. *Stat. Pap.* **2022**, *63*, 1953–1978. [CrossRef]
22. Cavieres, M.F.; Leiva, V.; Marchant, C.; Rojas, F. A methodology for data-driven decision making in the monitoring of particulate matter environmental contamination in Santiago of Chile. *Rev. Environ. Contam. Toxicol.* **2020**, *250*, 5–67.
23. Ibacache-Pulgar, G.; Reyes, S. Local influence for elliptical partially varying coefficient model. *Stat. Model.* **2018**, *18*, 149–174. [CrossRef]
24. Good, I.J.; Gaskins, R.A. Nonparametric roughness penalties for probability densities. *Biometrika* **1971**, *58*, 255–277. [CrossRef]
25. Silverman, B.W. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *J. R. Stat.* **1985**, *47*, 1–52. [CrossRef]
26. Green, P.J. Penalized likelihood for general semi-parametric regression models. *Int. Stat. Rev.* 1987, **55**, 245–259. [CrossRef]
27. Adams, R.A.; Fournier, J. Sobolev Spaces. In *Pure and Applied Mathematics*; Academic Press: Boston, MA, USA, 2003.
28. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. (Appl. Stat.)* **2005**, *54*, 507–554. [CrossRef]
29. Berhane, K.; Tibshirani, J. Generalized additive models for longitudinal data. *Can. J. Stat.* **1998**, *26*, 517–535. [CrossRef]
30. Cook, R.D. Assessment of local influence (with discussion). *J. R. Stat. Soc.* **1986**, *48*, 133–169.
31. Escobar, L.; Meeker, W. Assessing influence in regression analysis with censored data. *Biometrics* **1992**, **48**, 507–528. [CrossRef]
32. Wei, B.C.; Hu, Y.Q.; Fung, W.K. Generalized leverage and its applications. *Scand. J. Stat.* **1998**, *25*, 25–37. [CrossRef]