REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 321–342
https://doi.org/10.57805/revstat.v22i3.507

Assessing Influence on Partially Varying-Coefficient Generalized Linear Model

Authors: Germán Ibacache-Pulgar ✉
– Institute of Statistics, Universidad de Valparaíso, Chile
– Center for Atmospheric Studies and Climate Change (CEACC), Universidad de Valparaíso, Chile
german.ibacache@uv.cl

Valeria Lira
– Institute of Statistics, Universidad de Valparaíso, Chile
valeria.s.lira@gmail.com

Cristian Villegas
– Exact Science Department, University of São Paulo, São Paulo, Brazil
clobos@usp.br

Received: August 2021 Revised: July 2022 Accepted: July 2022

Abstract:

- In this paper we discuss estimation and diagnostic procedures in partially varying-coefficient generalized linear models based in the penalized likelihood function. Specifically, we derive a weighted back-fitting algorithm to estimate the model parameters using smoothing spline. Moreover, we developed the local influence method to assess the sensitivity of maximum penalized likelihood estimators when small perturbations are introduced into the model or data. Finally, an example with real data of ozone concentration is given for illustration.

Keywords:

- *exponential family; maximum penalized likelihood estimators; likelihood displacement; semiparametric models; weighted back-fitting algorithm.*

AMS Subject Classification:

- 49A05, 78B26.

✉ Corresponding author

1. INTRODUCTION

Partially varying-coefficient generalized linear model (PVCGLM) is an extension of generalized linear model (GLM), and have received special attention in recent years. These models have the same characteristics as GLM (see, for instance, McCullagh and Nelder, 1989 [27]), in the sense of encompassing different families of distributions for the response variable, allowing for non-linear dependence between the mean of the response variable and the explanatory variables (linear predictor) through a link function, and allowing for non-constant variance in the data. In addition, PVCGLM have the flexibility to model explanatory variables effects that can contribute parametrically and explanatory variables effects in which the coefficients are allowed to vary as smooth functions of other variables (for example, time variable). The model is a very useful tool for exploring dynamic patterns in some scientific areas, such as environmental, epidemiology, medical science, ecology and so on; see Fan and Zhang (2008) [9], Finley (2011) [14], Ma *et al.* (2011) [26], Li *et al.* (2018) [24], and He *et al.* (2022) [18].

As was noted by some authors (see, for example, Ouwens *et al.*, 2001 [29]), GLM parameter estimators can be higly impacted by outlying observations. For this reason, diagnostic analysis is of fundamental importance in the statistical modelling of any data set. The main idea of the local influence technique, introduced by Cook (1986) [5], is to evaluate the sensitivity of parameter estimators when small perturbations are introduced in the assumptions of the model or in the data. Some of the works related to the technique of local influence applied to different regression models are the following. Thomas and Cook (1989) [33] extended the method of local influence proposed by Cook to generalized linear models, with the purpose to asses the effect of small perturbations in the data. Ouwens *et al.* (2001) [29] developed local influence to detect influential data structures under a generalized linear mixed model; specifically, they proposed a two-stage diagnostic procedure, the first to measure the influence of the subjects and the second to measure the influence of the observations. Zhu and Lee (2001) [35] extended the method of local influence for incomplete data based on the conditional expectation of the complete-data log-likelihood function, and applied the results to the generalized linear mixed model; see also Zhu and Lee (2003) [36]. Espinheira *et al.* (2008) [8] developed the local influence method for beta regressions model under different perturbation schemes. Rocha and Simas (2011) [31] extended the local influence method to a general formulation of the class of the beta regression models, whereas Ferrari *et al.* (2011) [12] derived the normal curvatures of local influence for beta regression models with varying dispersion. Ferreira and Paula (2016) [13] extended the local influence technique for different perturbation schemes considering a skew-normal partially linear model and Emami (2016) [6] applied local influence analysis to the Liu penalized least squares estimator.

In semiparametric context, Thomas (1991) [33] constructed local influence diagnostics to evaluate the sensitivity of the smoothing parameter estimate obtained by cross-validation criterion. Zhu *et al.* (2003) [36] and Ibacache-Pulgar and Paula (2011) [21] provide local influence measures to evaluate the sensitivity of the maximum penalized likelihood estimator in normal and Student-t partially linear models, respectively. Ibacache-Pulgar *et al.* (2012, 2013) [19, 20] derived the local influence curvature for elliptical semiparametric mixed and symmetric semiparametric additive models, respectively. Zhang *et al.* (2015) [34] and Ibacache-Pulgar and Reyes (2018) [22] developed local influence measures for normal and elliptical partially varying-coefficient models, respectively. Recently, Ibacache-Pulgar *et al.* (2021) [23] developed the local influence method to semiparametric additive beta regression models and Sanchez *et al.* (2021) [32] derived the normal curvature for a new quantile regression model.

The aim of this paper is to apply local influence to the PVCGLM. The paper is organized as follows. In Section 2, the PVCGLM is presented. A discussion on the process used to obtain the maximum likelihood (ML) estimator based on the penalized likelihood, the derivation of a back-fitting algorithm and some inferential result are given in Section 3. In Section 4 the main concepts of local influence are considered and normal curvatures for different perturbations schemes are derived. An illustration of the methodology is presented in Section 5. Finally, in Section 6, some concluding remarks are given.

2. STATISTICAL MODEL

In this section we present the PVCGLM and the penalized log-likelihood function used to carry out parameter estimation.

2.1. Formulation

Consider a data set that is composed of a response variable y_i, for $i \in \{1, ..., n\}$, that follows a distribution in the exponential family with density function

$$f_y(y_i; \theta_i, \phi) = \exp\left[\frac{y_i\theta_i - \psi(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right], \tag{2.1}$$

where θ_i is the canonical form of the location parameter and is a function of the mean μ_i, $a_i(\phi)$ is a known function of the unknown dispersion parameter ϕ (or a vector of unknown dispersion parameters), c is a function of the dispersion parameter and the responses, and ψ is a known function, such that the mean and variance of y_i are equals to $\mu_i = \mathrm{E}(y_i) = \partial\psi(\theta_i)/\partial\theta_i$ and $\mathrm{Var}(y_i) = a_i(\phi)\,\mathrm{V}_i$, with $\mathrm{V}_i = \mathrm{V}(\mu_i) = \partial^2\psi(\theta_i)/\partial\theta_i^2$, respectively. The PVCGLM is defined by Equation (2.2) and the following systematic component:

$$g(\mu_i) = \eta_i = \boldsymbol{w}_i^\top \boldsymbol{\alpha} + \sum_{k=1}^{s} \mathrm{x}_i^{(k)}\, \beta_k(\mathrm{t}_{k_i}), \tag{2.2}$$

where $\boldsymbol{w}_i$ is a $(p \times 1)$ vector of predictors variables, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)^\top$ is a vector of regression coefficients, $\beta_k(\cdot)$ for $k \in \{1, ..., s\}$ are unknown smooth arbitrary functions of t_k, associated with the predictor variable $\mathrm{x}_i^{(k)}$. Here, the superscript k refers to the relationship of the predictor variable x_i with the k-th nonparametric component. Note that Model (2.2) can be written in a matrix form as

$$\boldsymbol{\eta} = \boldsymbol{W}\boldsymbol{\alpha} + \sum_{k=1}^{s} \widetilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k\,, \tag{2.3}$$

where $\boldsymbol{W} = (\boldsymbol{w}_1^\top, ..., \boldsymbol{w}_n^\top)$, $\widetilde{\boldsymbol{N}}_k = \boldsymbol{X}^{(k)}\boldsymbol{N}_k$, $\boldsymbol{X}^{(k)} = \mathrm{diag}_{1\le i\le n}(\mathrm{x}_i^{(k)})$, $\boldsymbol{N}_k$ is an $(n\times r_k)$ incidence matrix with the (i,l)-th element equal to the indicator $I(\mathrm{t}_{k_i} = \mathrm{t}_{k_l}^0)$ with $\mathrm{t}_{k_l}^0$ denoting the distinct and ordered values of the explanatory variable t_k, and $\boldsymbol{\beta}_k = (\psi_{k_1}, ..., \psi_{k_r})^\top$ is a $(r_k \times 1)$ vector of parameters with $\psi_{k_l} = \beta_k(\mathrm{t}_{k_l}^0)$ for $l \in \{1, ..., r_k\}$.

2.2. Penalized log-likelihood function

Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_s^\top, \phi)^\top \in \boldsymbol{\Theta} \subseteq \mathcal{R}^{p^*}$, with $p^* = p + r + 1$ and $r = \sum_{k=1}^s r_k$, be the vector of unknown parameters associated to Model (2.1). Then, the log-likelihood function is given by

$$L(\boldsymbol{\theta}) = \sum_{i=1}^n L_i(\boldsymbol{\theta}), \tag{2.4}$$

where

$$L_i(\boldsymbol{\theta}) = \left[\frac{y_i\theta_i - \psi(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right]. \tag{2.5}$$

Since the β_k's belong to a space of infinite dimension and are considered parameters with respect to the expected value of y_i, it is necessary to define a restricted subspace for these functions so that the identifiability of the parameters holds. This choice typically depends on the domain of the function, on a priori knowledge of form of the function, on constraints to ensure identifiability, or simply on some specific application. In this paper, we will assume that the function β_k belongs to the Sobolev function space

$$\mathcal{W}_2^{(l)} = \{\beta_k : \beta_k, \beta_k^{(1)}, ..., \beta_k^{(l-1)} \text{abs. cont.}, \beta_k^{(l)} \in \mathcal{L}^2[a_k, b_k]\},$$

where $\beta_k^{(l)}(\mathrm{t}_k) = \mathrm{d}^l\beta(\mathrm{t}_k)/\mathrm{dt}_k^l$, with $\mathrm{t}_k^0 \in [a_k, b_k]$. To ensure the identifiability of the parameters and an adequate fit of the model, we incorporate a penalty term in the original log-likelihood function over each function β_k. In this way, we obtain a penalized version of the log-likelihood function of the form (see details in Green and Silverman, 1994 [15])

$$L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = L(\boldsymbol{\theta}) - \sum_{k=1}^s \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k\ , \tag{2.6}$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_s)^\top$ denotes a $(s\times 1)$ vector of smoothing parameters that controls the tradeoff between goodness of fit and the smoothness estimated functions, and $\boldsymbol{K}_k$ is a $(r_k \times r_k)$ nonnegative definite smoothing matrix associated with the k-th explanatory variable that only depends on the knots. In this case, the estimation of β_k leads to a smooth cubic spline with knots at the points $\mathrm{t}_{k_l}^0$ for $l \in \{1, ..., r_k\}$.

3. ESTIMATION AND INFERENCE

In this section we outlying the estimation of the parameters of the PVCGLM. Specifically, we propose an iterative process based on the Fisher score and back-fitting algorithms to estimate the regression coefficients and the nonparametric functions, and respective standard errors being obtained from the penalized Fisher information matrix. More details about estimation procedure can be found, for example, in Hastie and Tibshirani (1993) [16], Cai *et al.* (2000) [4], Fang and Huang (2005) [10] and Rigby and Stasinopoulos (2005) [30].

3.1. Weighted maximum penalized likelihood estimator

Assuming that the function (2.6) is regular with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_k$'s and ϕ, the penalized score function vector of $\boldsymbol{\theta}$ is given by

$$\mathbf{U}_{\mathrm{P}}(\boldsymbol{\theta}) = \frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\theta}}\,.$$

After some algebraic manipulations (see, for instance, Liu *et al.*, 2021 [25], for details of the calculation of derivatives of matrix or vectors), we obtain the following:

$$\begin{aligned}
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\alpha}} &= \boldsymbol{W}^{\top}\boldsymbol{T}(\boldsymbol{y}-\boldsymbol{\mu})\,,\\
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\beta}_k} &= \widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{T}(\boldsymbol{y}-\boldsymbol{\mu}) - \lambda_k\boldsymbol{K}_k\boldsymbol{\beta}_k \qquad k\in\{1,...,s\}\,,\\
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \phi} &= \sum_{i=1}^{n} -(\mathrm{a}_i(\phi))^{-2}\{y_i\theta_i-\psi(\theta_i)\} + \sum_{i=1}^{n} c'(y_i,\phi),
\end{aligned}$$

where $\boldsymbol{W}$ is a $(n\times p)$ matrix whose i-th row is $\boldsymbol{w}_i^{\top}$, $\boldsymbol{T}=\mathrm{diag}_{1\le i\le n}\big((\mathrm{a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)V_i^{-1}\big)$ with $V_i=V(\mu_i)=\partial^2\psi(\theta_i)/\partial\theta_i^2$ the variance function, $\mathrm{a}_i(\phi)$ is a function of ϕ, $\boldsymbol{y}=(y_1,...,y_n)^{\top}$, $\boldsymbol{\mu}=(\mu_1,...,\mu_n)^{\top}$ and $c'(y_i,\phi)=\partial c(y_i,\phi)/\partial\phi$. To estimate $\boldsymbol{\theta}$, we have to solve $\mathbf{U}_{\mathrm{P}}(\boldsymbol{\theta})=\mathbf{0}$. However, the estimating equations are nonlinear and require an iterative method. For example, maximum penalized likelihood (MPL) estimator for $\boldsymbol{\theta}$ can be performed by using the Fisher scoring algorithm. Let $\boldsymbol{\beta}_0=\boldsymbol{\alpha}$, $\widetilde{\boldsymbol{N}}_0=\boldsymbol{W}$, and $\boldsymbol{\lambda}$ fixed. Then, the Fisher scoring algorithm is given by

$$\begin{pmatrix}
\boldsymbol{I} & \boldsymbol{S}_0^{(u)}\widetilde{\boldsymbol{N}}_1 & \dots & \boldsymbol{S}_0^{(u)}\widetilde{\boldsymbol{N}}_s\\
\boldsymbol{S}_1^{(u)}\widetilde{\boldsymbol{N}}_0 & \boldsymbol{I} & \dots & \boldsymbol{S}_1^{(u)}\widetilde{\boldsymbol{N}}_s\\
\vdots & \vdots & \ddots & \vdots\\
\boldsymbol{S}_s^{(u)}\widetilde{\boldsymbol{N}}_0 & \boldsymbol{S}_s^{(u)}\widetilde{\boldsymbol{N}}_1 & \dots & \boldsymbol{I}
\end{pmatrix}
\begin{pmatrix}
\boldsymbol{\beta}_0^{(u+1)}\\ \boldsymbol{\beta}_1^{(u+1)}\\ \vdots\\ \boldsymbol{\beta}_s^{(u+1)}
\end{pmatrix}
=
\begin{pmatrix}
\boldsymbol{S}_0^{(u)}\boldsymbol{z}^{(u)}\\ \boldsymbol{S}_1^{(u)}\boldsymbol{z}^{(u)}\\ \vdots\\ \boldsymbol{S}_s^{(u)}\boldsymbol{z}^{(u)}
\end{pmatrix}, \tag{3.1}$$

where $\boldsymbol{z}^{(u)}=(\boldsymbol{y}-\boldsymbol{\mu}^{(u)})+\left(\sum_{k=0}^{s}\widetilde{\boldsymbol{N}}_k\boldsymbol{\beta_k}^{(u)}\right)$ and

$$\boldsymbol{S}_k^{(u)} = \begin{cases}
(\widetilde{\boldsymbol{N}}_0^{\top}\boldsymbol{M}^{(u)}\widetilde{\boldsymbol{N}}_0)^{-1}\widetilde{\boldsymbol{N}}_0^{\top}\boldsymbol{M}^{(u)} & k=0\\
(\widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{M}^{(u)}\widetilde{\boldsymbol{N}}_k+\lambda_k\boldsymbol{K}_k)^{-1}\widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{M}^{(u)} & k\in\{1,...,s\}\,,
\end{cases}$$

where $\boldsymbol{M} = \operatorname{diag}_{1\le i\le n}\big((\mathrm{a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)^2 V_i^{-1}\big)$. Consequently, the weighted back-fitting (Gauss-Seidel) iterations that are used to solve the equations system (3.1) take the form

$$\boldsymbol{\beta}_k^{(u+1)} = \boldsymbol{S}_k^{(u)}\left(\boldsymbol{z}^{(u)} - \sum_{l=0,l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\right), \tag{3.2}$$

for $u \in \{0,1,...\}$. On the other hand, the MPL estimator of the dispersion parameter, $\widehat{\phi}$, can be obtained by solving the following iterative process:

$$\phi^{(u+1)} = \phi^{(u)} - \mathrm{E}\left\{\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\phi^2}\right\}^{-1} \frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\phi}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(u)}},$$

for $u \in \{0,1,...\}$.

Algorithm 1 – Joint iterative process for estimating the parameters of the PVCGLM.

(i) Initialize:

- **(a)** Provide values for $\boldsymbol{\beta}_0^{(0)}, \boldsymbol{\beta}_1^{(0)}, ..., \boldsymbol{\beta}_s^{(0)}$.
- **(b)** Get starting value for ϕ by using the fitted values from (a).
- **(c)** From the current value $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\beta}_0^{(0)\top}, \boldsymbol{\beta}_1^{(0)\top}, ..., \boldsymbol{\beta}_s^{(0)\top}, \phi^{(0)})^\top$ obtaining the weight matrix $\boldsymbol{M}^{(0)}$. Then, obtain

$$\begin{aligned}
\boldsymbol{z}^{(0)} &= (\boldsymbol{y} - \boldsymbol{\mu}^{(0)}) + \left(\sum_{k=0}^{s} \widetilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(0)}\right),\\
\boldsymbol{S}_0^{(0)} &= (\widetilde{\boldsymbol{N}}_0^\top \boldsymbol{M}^{(0)} \widetilde{\boldsymbol{N}}_0)^{-1} \boldsymbol{N}_0^\top \boldsymbol{M}^{(0)},\\
\boldsymbol{S}_k^{(0)} &= (\widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M}^{(0)} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M}^{(0)}, \qquad k \in \{1,...,s\}.
\end{aligned}$$

(ii) Step 1: Iterate repeatedly by cycling between the following equations:

$$\begin{aligned}
\boldsymbol{\beta}_0^{(u+1)} &= \boldsymbol{S}_0^{(u)}\left(\boldsymbol{z}^{(u)} - \sum_{l=1}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\right),\\
\boldsymbol{\beta}_1^{(u+1)} &= \boldsymbol{S}_1^{(u)}\left(\boldsymbol{z}^{(u)} - \widetilde{\boldsymbol{N}}_0 \boldsymbol{\beta}_0^{(u+1)} - \sum_{l=2}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\right),\\
&\vdots\\
\boldsymbol{\beta}_s^{(u+1)} &= \boldsymbol{S}_s^{(u)}\left(\boldsymbol{z}^{(u)} - \sum_{l=0}^{s-1} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u+1)}\right),
\end{aligned}$$

for $u \in \{0,1,...\}$. Repeat (ii) replacing $\boldsymbol{\beta}_j^{(u)}$ by $\boldsymbol{\beta}_j^{(u+1)}$ until convergence criterion $\Delta_u(\boldsymbol{\beta}_j^{(u+1)}, \boldsymbol{\beta}_j^{(u)}) = \sum_{j=0}^{s} \parallel \boldsymbol{\beta}_j^{(u+1)} - \boldsymbol{\beta}_j^{(u)} \parallel / \sum_{j=0}^{s} \parallel \boldsymbol{\beta}_j^{(u)} \parallel$ is below some small threshold (Hastie and Tibshirani, 1990 [17]).

(iii) Step 2: For current values $\boldsymbol{\beta}_j^{(u+1)}$ for $j \in \{0,1...,s\}$, obtaining $\phi^{(u+1)}$ by using

$$\phi^{(u+1)} = \phi^{(u)} - \mathrm{E}\left\{\frac{\partial^2 L_{\mathrm{p}}(\phi,\boldsymbol{\lambda})}{\partial\phi\partial\phi}\right\}^{-1} \frac{\partial L_{\mathrm{p}}(\phi,\boldsymbol{\lambda})}{\partial\phi}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(u)}}.$$

(iv) Iterating between steps (ii) and (iii) by replacing $\boldsymbol{\beta}_j^{(0)}$ for $j \in \{0,1...,s\}$ and $\phi^{(0)}$ by $\boldsymbol{\beta}_j^{(u+1)}$ and $\phi^{(u+1)}$, respectively, until convergence.

Note that the system of equations (3.1) is consistent and the back-fitting algorithm (3.2) converges to a solution for any starting values if the weights matrix involved is symmetric and positive definite. Additionally, the solution is unique when there is not concurvity in the data, that is, nonlinear dependencies among the predictor variables. However, in the presence of concurvity, the starting functions will determine the final solution, while in presence exact concurvity is highly unlikely, except in the case of symmetric smoothers with eigenvalues in [0,1]; see, for instance, Berhane and Tibshirani (1998) [1]. The summary, the solution of the estimating equation system (3.1) to obtain the MPL estimates of $\boldsymbol{\theta}$ may be attained by iterating between a weighted back-fitting algorithm with weight matrix $\boldsymbol{M}$ and a Fisher score algorithm to obtain ML estimation of ϕ, which is equivalent to the iterative process in Algorithm 1.

3.2. Standard error of MPL estimator

Similarly to the classical theory of generalized linear models, the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ can be approximated through the inverse of Fisher information matrix obtained from penalized log-likelihood function, $L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})$. Assuming that the penalized log-likelihood function (2.6) is twice differentiable with respect to $\boldsymbol{\theta}$, we have that the penalized Fisher information matrix is given by

$$\boldsymbol{\mathcal{I}}_{\rm p} \;=\; -{\rm E}\left(\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\right).$$

This matrix assumes the following diagonal structure in blocks:

$$\boldsymbol{\mathcal{I}}_{\rm p}(\boldsymbol{\theta}) \;=\; \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\rm p}^{\alpha\beta_k}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{I}}_{\rm p}^{\phi\phi}(\boldsymbol{\theta}) \end{pmatrix},$$

where

$$\boldsymbol{\mathcal{I}}_{\rm p}^{\alpha\beta_k}(\boldsymbol{\theta}) \;=\; \begin{pmatrix} \boldsymbol{W}^\top\boldsymbol{M}\boldsymbol{W} & \boldsymbol{W}^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1 & \ldots & \boldsymbol{W}^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s \\ \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\boldsymbol{W} & \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1+\lambda_1\boldsymbol{K}_1 & \ldots & \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\boldsymbol{W} & \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1 & \ldots & \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s+\lambda_s\boldsymbol{K}_s \end{pmatrix}$$

and

$$\boldsymbol{\mathcal{I}}_{\rm p}^{\phi\phi}(\boldsymbol{\theta}) \;=\; \sum_{i=1}^{n} -2({\rm a}_i(\phi))^{-3}(\mu_i\theta_i-\psi(\theta_i)) - \sum_{i=1}^{n}{\rm E}(c''(y_i,\phi)),$$

with $c''(y_i,\phi)=\partial^2 c(y_i,\phi)/\partial\phi^2$ for $i\in\{1,...,n\}$. Therefore, the approximate variance--covariance matrix of $\widehat{\boldsymbol{\theta}}$ and an approximate pointwise standard error band (SEB) for $\beta_k(\cdot)$, that allows us to assess the accuracy of $\widehat{\beta_k}(\cdot)$ at different locations within the range of interest, are given by

$$\widehat{\rm Cov}(\widehat{\boldsymbol{\theta}}) \;\approx\; \boldsymbol{\mathcal{I}}_{\rm p}^{-1}\big|_{\widehat{\boldsymbol{\theta}}},$$

$${\rm SEB}_{\rm approx}(\beta_k({\rm t}_l^0)) \;=\; \widehat{\beta}_k({\rm t}_l^0) \pm 2\sqrt{\widehat{\rm Var}(\widehat{\beta}_k({\rm t}_l^0))} \qquad l\in\{1,...,r_k\},$$

where ${\rm Var}(\widehat{\beta}_k({\rm t}_l))$, for $k\in\{1,...,s\}$, is the l-th principal diagonal element of the corresponding block-diagonal matrix of $\boldsymbol{\mathcal{I}}_{\rm p}^{-1}$.

3.3. Effective degrees of freedom and smoothing parameters

In the iterative process defined in the Equation (3.2), considering ϕ as known, we can write the expression of the estimator of $\boldsymbol{\beta}_k$ at step u as

$$\boldsymbol{\beta}_k^{(u+1)} = (\widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M} \boldsymbol{z}^{*(u)} \qquad k \in \{1, ..., s\}, \tag{3.3}$$

where $\boldsymbol{z}^{*(u)} = \boldsymbol{z}^{(u)} - \sum_{l=0, l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}$. From the convergence of the iterative process given in the Equation (3.3), we obtain

$$\widehat{\boldsymbol{\beta}}_k = (\widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widehat{\boldsymbol{z}^*} \qquad k \in \{1, ..., s\},$$

where $\widehat{\boldsymbol{z}^*} = \Big[(\boldsymbol{y} - \widehat{\boldsymbol{\mu}}) + \big(\sum_{k=0}^{s} \widetilde{\boldsymbol{N}}_k \widehat{\boldsymbol{\beta}}_k\big)\Big] - \sum_{l=0, l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \widehat{\boldsymbol{\beta}}_l$. In this paper we define the effective degrees of freedom (df) associated with the smooth functions as (see, for instance, Hastie and Tibshirani, 1990 [17])

$$\text{edf}(\lambda_k) = \text{tr}\big\{\widetilde{\boldsymbol{N}}_k (\widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}}\big\}.$$

Following Ibacache-Pulgar and Reyes (2018) [22], we choose the optimal smoothing parameter for each smooth functions by specifying an appropriate $\text{edf}(\lambda_k)$ value.

4. LOCAL INFLUENCE

In this section we obtain the normal curvature for PVCGLM. Specifically, the Hessian and perturbations matrices for different perturbations schemes.

4.1. The method

To assess the influence of minor perturbations on the MPL estimator of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}}$, we can consider the likelihood displacement $\text{LD}(\boldsymbol{\omega}) = 2\Big[L_\text{P}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\lambda}) - L_\text{P}(\widehat{\boldsymbol{\theta}}_\omega, \boldsymbol{\lambda})\Big] \geq 0$, where $\widehat{\boldsymbol{\theta}}_\omega$ is the MPL estimador under the perturbed penalized log-likelihood function, denoted by $L_\text{P}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega})$, and $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^\top$ be an n-dimensional vector of perturbations restricted to some open subset $\boldsymbol{\Omega} \in \mathcal{R}^n$. It is assumed that there exists $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$, a vector of no perturbation, such that $L_\text{P}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega}_0) = L_\text{P}(\boldsymbol{\theta}, \boldsymbol{\lambda})$. Cook (1986) [5] suggests to study the local behavior of $\text{LD}(\boldsymbol{\omega})$ around $\boldsymbol{\omega}_0$ selecting a unit direction $\boldsymbol{\ell} \in \Omega$ ($\|\boldsymbol{\ell}\| = 1$), and then to consider the plot of $\text{LD}(\boldsymbol{\omega}_0 + a\,\boldsymbol{\ell})$ (called lifted line) against a, where $a \in \mathcal{R}$. Each lifted line can be characterized by considering the normal curvature $C_\ell(\boldsymbol{\theta})$ around $a = 0$. The suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_\text{max}$ corresponding to the largest curvature $C_{\ell_\text{max}}(\boldsymbol{\theta})$. The index plot of $\boldsymbol{\ell}_\text{max}$ may reveal those observations that under small perturbations exercise notable influence on $\text{LD}(\boldsymbol{\omega})$. According to Cook (1986) [5], the normal curvature at the unit direction $\boldsymbol{\ell}$ is given by $C_\ell(\boldsymbol{\theta}) = -2\big[\boldsymbol{\ell}^\top \boldsymbol{\Delta}_\text{p}^\top \boldsymbol{L}_\text{p}^{-1} \boldsymbol{\Delta}_\text{p} \boldsymbol{\ell}\big]$, which represents the local influence on $\widehat{\boldsymbol{\theta}}$ after perturbing the model or data, where $\boldsymbol{L}_\text{p}$ is the Hessian matrix evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\Delta}_\text{p} = \partial^2 L_\text{P}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega}) / \partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top$ is the perturbation matrix evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Escobar and Meeker (1992) [7] proposed to study the normal curvature at the direction $\boldsymbol{\ell} = \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i$ is an n-dimensional vector with 1 at the i-th position and zeros at the remaining positions. In this case, the normal curvature, called total local influence of the i-th individual, takes the form $C_{\boldsymbol{\epsilon}_i}(\boldsymbol{\theta}) = 2|\,\mathrm{c}_{ii}\,|$ for $i \in \{1, ..., n\}$, where c_{ii} is the i-th principal diagonal element of the matrix $\boldsymbol{C} = \boldsymbol{\Delta}_\mathrm{p}^\top \boldsymbol{L}_\mathrm{p}^{-1} \boldsymbol{\Delta}_\mathrm{p}$. In order to have a invariant curvature under uniform change of scale, Poon and Poon (1999) [28] proposed the conformal normal curvature defined as

$$B_\ell(\boldsymbol{\theta}) = \frac{C_\ell(\boldsymbol{\theta})}{2\sqrt{\mathrm{tr}(\boldsymbol{\Delta}_\mathrm{p}^\top \boldsymbol{L}_\mathrm{p}^{-1} \boldsymbol{\Delta}_\mathrm{p})^2}} = -\frac{\boldsymbol{\ell}^\top \boldsymbol{\Delta}_\mathrm{p}^\top \boldsymbol{L}_\mathrm{p}^{-1} \boldsymbol{\Delta}_\mathrm{p} \boldsymbol{\ell}}{\sqrt{\mathrm{tr}(\boldsymbol{\Delta}_\mathrm{p}^\top \boldsymbol{L}_\mathrm{p}^{-1} \boldsymbol{\Delta}_\mathrm{p})^2}} \,.$$

This curvature is characterized to allow for any unit direction $\boldsymbol{\ell}$ that $0 \leq B_\ell(\theta) \leq 1$. A suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_{\max}$ corresponding to the largest curvature $B_{\ell_{\max}}(\boldsymbol{\theta})$ or, alternatively, to evaluate the normal curvature at the direction $\boldsymbol{\ell} = \boldsymbol{\epsilon}_i$ and analyse the index plot of $B_{\boldsymbol{\epsilon}_i}(\boldsymbol{\theta})$.

4.2. Derivation of normal curvature

The perturbation schemes that are considered in the analysis of local influence depend on the structure of the proposed model (see, for instance, Billor and Loynes, 1993 [2]), and can be classified into two broad groups: perturbation to the model (in order to study modifications in the assumptions) or in the data. For example, we might be interested in perturbing the response or the explanatory variables. The reasons for considering such perturbation schemes are, for example, the existence of outliers or measures with measurement errors, respectively. However, the perturbation scheme to be considered should be formulated in a way that responds to questions previously established by the researcher. We will present in what follows expressions of the $\boldsymbol{L}_\mathrm{p}$ and $\boldsymbol{\Delta}_\mathrm{p}$ matrices for some perturbations schemes.

Hessian matrix

Let $\boldsymbol{L}_\mathrm{p}$ $(p^*\times p^*)$ be the Hessian matrix with (j^*,ℓ^*)-th element given by $\partial^2 L_\mathrm{p}(\boldsymbol{\theta},\boldsymbol{\lambda})/\partial\theta_{j^*}\theta_{\ell^*}$ for $j^*, \ell^* \in \{1, ..., p^*\}$, where $p^* = p + r + 1$, with $r = \sum_{k=1}^s r_k$. After some algebraic manipulations we find

$$\begin{aligned}
\frac{\partial^2 L_\mathrm{p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}^\top} &= -\boldsymbol{W}^\top \widetilde{\boldsymbol{M}} \boldsymbol{W}\,, \\
\frac{\partial^2 L_\mathrm{p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\boldsymbol{\beta}_k\partial\boldsymbol{\beta}_k^\top} &= \begin{cases} -\widetilde{\boldsymbol{N}}_k^\top \widetilde{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k - \lambda_k \boldsymbol{K}_k & k = k' \\ -\widetilde{\boldsymbol{N}}_k^\top \widetilde{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_{k'} & k \neq k', \end{cases} \\
\frac{\partial^2 L_\mathrm{p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\phi^2} &= \sum_{i=1}^n 2(\mathrm{a}_i(\phi))^{-3}(y_i\theta_i - \psi(\theta_i)) + \sum_{i=1}^n c''(y_i,\phi))\,,
\end{aligned}$$

$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\beta}_k^\top} = -\boldsymbol{W}^\top\widetilde{\boldsymbol{M}}\widetilde{\boldsymbol{N}}_k\,,$$
$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\alpha_j\partial\phi} = -\sum_{i=1}^{n}({\rm a}_i(\phi))^{-2}\left\{(y_i-\mu_i)V_i^{-1}\frac{\partial\mu_i}{\partial\eta_i}\,{\rm w}_i\right\},$$

$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\psi_{k_l}\partial\phi} = -\sum_{i=1}^{n}({\rm a}_i(\phi))^{-2}\left\{(y_i-\mu_i)V_i^{-1}\frac{\partial\mu_i}{\partial\eta_i}\,{\rm n}_{k_{il}}\right\},$$

where $c''(y_i,\phi)=\partial^2 c(y_i,\phi)/\partial\phi^2$, $\widetilde{\boldsymbol{M}}={\rm diag}_{1\le i\le n}\big(({\rm a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)^2V_i^{-1}\rho_i\big)$, $\rho_i=\kappa(\mu_i)/\{g'(\mu_i)^2V_i\}$, with $\kappa(\mu_i)=1+(y_i-\mu_i)\{V_i'/V_i+g''(\mu_i)/g'(\mu_i)\}$ and $g'(\mu_i)={\rm d}\eta_i/{\rm d}\mu_i$, and ${\rm n}_{k_{li}}$ denotes the (i,l)-th element of the matrix $\boldsymbol{N}_k$.

Cases-weight perturbation

Let us consider the attributed weights for the observations in the penalized log-likelihood function as

$$L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega})-\sum_{k=1}^{s}\frac{\lambda_k}{2}\,\boldsymbol{\beta}_k^\top\boldsymbol{K}_k\boldsymbol{\beta}_k\,,$$

where $L(\boldsymbol{\theta}|\boldsymbol{\omega})=\sum_{i=1}^n\omega_i\,L_i(\boldsymbol{\theta})$, $\boldsymbol{\omega}=(\omega_1,...,\omega_n)^\top$ is the vector of weights, with $0\le\omega_i\le 1$, and $\boldsymbol{\omega}_0=(1,...,1)^\top$ the vector of no perturbation. Differentiating $L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$, we obtain after some algebraic manipulation

$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \boldsymbol{W}^\top\widehat{\boldsymbol{D}}_\tau\,,$$
$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\beta}_k\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \widetilde{\boldsymbol{N}}_k^\top\widehat{\boldsymbol{D}}_\tau \qquad k\in\{1,...,s\}\,,$$
$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\phi\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \widehat{\boldsymbol{u}}^\top\,,$$

where $\boldsymbol{D}_\tau={\rm diag}_{1\le i\le n}(\tau_i)$ and $\boldsymbol{u}=(u_1,...,u_n)^\top$, with $\tau_i=({\rm a}_i(\phi))^{-1}(y_i-\partial\psi(h(\eta_i))/\partial h(\eta_i))\cdot$ $\cdot\,\partial h(\eta_i)/\partial\eta_i$, $h(\eta_i)=\psi'^{-1}(\eta_i)$, $\psi'^{-1}(\cdot)$ denoting the inverse function of $\psi'(\cdot)$, $u_i=-({\rm a}_i(\phi))^{-2}\cdot$ $\cdot\,(y_ih(\eta_i)-\psi(h(\eta_i))+{\rm c}'(y_i,\phi)\boldsymbol{e}_{in}^\top$, and $\boldsymbol{e}_{in}$ a vector with 1 at the i-th position and zero elsewhere.

Response variable perturbation

In general, the response variable can be perturbed in two ways:

$$y_{i\omega}=\begin{cases} y_i+\ \omega_i & \text{additive perturbation} \qquad i\in\{1,...,n\}\\ y_i\times\ \omega_i & \text{multiplicative perturbation}\,.\end{cases}$$

In this paper we consider $y_{i\omega}=y_i+\omega_i$, where $\boldsymbol{\omega}=(\omega_1,...,\omega_n)^\top$ is the vector of perturbations and $\boldsymbol{\omega}_0=(0,...,0)^\top$ the vector of no perturbation. The perturbed penalized log-likelihood function is constructed from expression (2.6) with y_i replaced by $y_{i\omega}$, that is,

$$L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}) - \sum_{k=1}^{s} \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k \ ,$$

where $L(\cdot)$ is given by Equation (2.4) with $y_{i\omega}$ in the place of y_i. Differentiating $L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and ω_i we obtain, after some algebraic manipulation, that

$$\begin{aligned}
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{W}^\top \widehat{\boldsymbol{D}}_c \ , \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_c \qquad k \in \{1, ..., s\} \ , \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \phi \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \widehat{\boldsymbol{d}}^\top \ ,
\end{aligned}$$

where $\boldsymbol{D}_c = \mathrm{diag}_{1\le i\le n}(c_i)$ and $\boldsymbol{d} = (d_1, ..., d_n)^\top$, with $c_i = \partial h(\eta_i)/\partial \eta_i$ and $d_i = -(\mathrm{a}_i(\phi))^{-2} \cdot \cdot (h(\eta_i)\boldsymbol{e}_{in}^\top + c'(y_{i\omega}, \phi)/\partial \omega_i)$, with $\boldsymbol{e}_{in}$ denoting a vector with 1 at the i-th position and zero elsewhere..

Explanatory variable perturbation

The explanatory variable can be perturbed in two ways:

$$\mathrm{w}_{i\omega} = \begin{cases} \mathrm{w}_{i\omega} + \omega_i & \text{additive perturbation} \qquad i \in \{1, ..., n\} \\ \mathrm{w}_{i\omega} \times \omega_i & \text{multiplicative perturbation} \ . \end{cases}$$

Here the d-th explanatory variable, assumed continuous, is perturbed by considering the additive perturbation scheme, namely $\mathrm{w}_{id\omega} = \mathrm{w}_{id} + \omega_i$, where $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^\top$ is the vector of perturbations such as $\omega_i \in \mathcal{R}$. The vector of no perturbation is given by $\boldsymbol{\omega}_0 = (0, ..., 0)^\top$. The perturbed penalized log-likelihood function is given by

$$L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}) - \sum_{k=1}^{s} \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k \ ,$$

where $L(\cdot)$ is given by Equation (2.4) with $\mu_{i\omega} = g^{-1}(\eta_{i\omega})$ in the place of μ_i, for $\eta_{i\omega} = \boldsymbol{w}_{i\omega}^\top \boldsymbol{\alpha} + \sum_{k=1}^{s} \mathrm{x}_i^{(k)}\, \beta_k(\mathrm{t}_{k_i})$, with w_{id} replaced by $\mathrm{w}_{id\omega}$. Differentiating $L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and ω_i we obtain

$$\begin{aligned}
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{e}_p \widehat{\boldsymbol{\tau}}^\top - \alpha_d \boldsymbol{W}^\top \widehat{\boldsymbol{D}}_b \ , \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{e}_p \widehat{\boldsymbol{\tau}}^\top - \alpha_d \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_b \qquad k \in \{1, ..., s\} \ , \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial \phi \partial \boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\ \boldsymbol{\omega}=\boldsymbol{\omega}_0} &= -\sum_{i=1}^{n} (\mathrm{a}_i(\phi))^{-2} \left\{ y_i \frac{\partial(\eta_{i\omega})}{\partial \omega_i} - \frac{\partial \psi(h(\eta_{i\omega}))}{\partial \omega_i} \right\} \boldsymbol{e}_{in}^\top,
\end{aligned}$$

where $\boldsymbol{\tau} = (\tau_1, ..., \tau_n)^\top$, $\boldsymbol{D}_b = \mathrm{diag}_{1\le i\le n}(b_i)$ and $\boldsymbol{e}_p$ is a vector with 1 at the p-th position and zero elsewhere, $\tau_i = (\mathrm{a}_i(\phi))^{-1}(y_i - \partial\psi(h(\eta_{i\omega}))/\partial h(\eta_{i\omega}))\partial h(\eta_{i\omega})/\partial \eta_{i\omega}$ and $b_i = (\mathrm{a}_i(\phi))^{-1} \cdot \cdot (y_i - \partial\psi(h(\eta_{i\omega}))/\partial h(\eta_{i\omega}))\partial^2 h(\eta_{i\omega})/\partial \eta_{i\omega}^2 - (\partial^2 \psi(h(\eta_{i\omega}))/\partial^2 h(\eta_{i\omega}))(\partial h(\eta_{i\omega})/\partial \eta_{i\omega})^2$.

5. APPLICATION

In this section, we illustrate the applicability of the PVCGLM and the local influence method through an application based on a set of real data. For our analysis, we consider the Poisson distribution.

5.1. Data set and problem statement

To motivate the use of the PVCGLM and the local influence method developed in this work, we consider a set of real data from a study conducted in the city of Los Angeles during 1976 (see, for instance, Breiman and Friedman, 1985 [3] and Faraway, 2006 [11]) with the purpose of describing the relationship between the outcome variable O3 (concentration of ozone per hour in Upland, CA, measured in parts per million (ppm) and a set of nine explanatory variables, for a sample of 330 days. The description of such variables is as follows. VH (pressure height 500 millibar, measured at the base of the air force of Vandenberg, in meters), WIND (wind speed, in miles per hour), HUM (humidity in percentage), TEMP (sandburg Air Base temperature, in Celsius), IBH (inversion base height, in foot), DPG (dagget pressure gradient, in mmHg), IBT (inversion base temperature, in Fahrenheit), VIS (visibility, in miles), DAY (calendar day).

5.2. Model fit

In our application we will consider only four explanatory variables, specifically, the variables VIS, TEMP, IBT and DAY. Figure 1 contains the dispersion graphs between the outcome variable and each one of the explanatory variables.

We see in Figure 1a that the relationship between O3 and the explanatory variable VIS is approximately linear, whereas the relationship between O3 and DAY appear to be nonlinear (Figure 1b). Note that there is a significant increase in the level of O3 from January to July with a decrease until December. This suggests that the incorporation of a quadratic or nonparametric term in the model can account for the behavior of O3 over time. On other hand, Figures 1c and 1d suggest that the explanatory variables TEMP and IBT might be interacting with the variable DAY in nonlinear fashion. Figures 2a and 2b shows the graph of autocorrelation and partial autocorrelation, respectively. Following the same analysis of Faraway (2006) [11], in this work we will no consider the possible temporal correlation for O3.

Initially, we will fit a GLM assuming that the response variable O3 follows a Poisson distribution with mean μ_i and logarithmic link function considering different structures of the linear predictor for the explanatory variables VIS, TEMP, IBT and DAY (see Table 1).

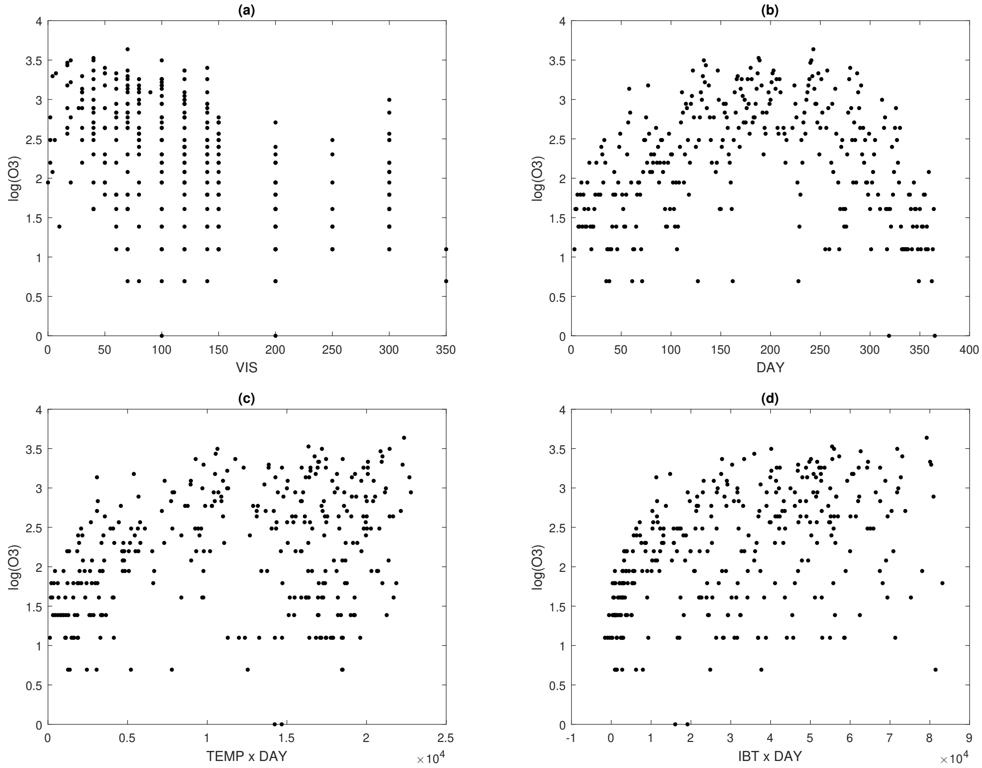

Figure 1: Scatter plots: log(O3) versus VIS (a), log(O3) versus DAY (b), log(O3) versus TEMP × DAY (c) and log(O3) versus IBT × DAY (d).

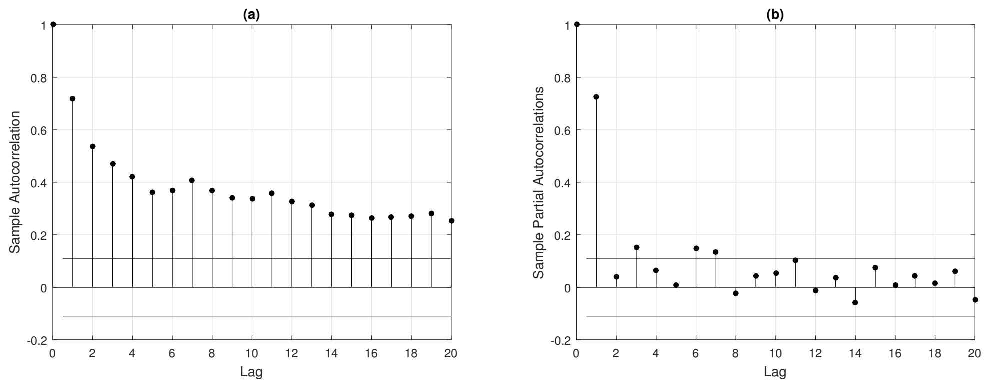

Figure 2: Autocorrelation (a) and partial autocorrelation (b) for Ozone data.

Table 1: Different structures of the linear predictor for the explanatory variables VIS, TEMP, IBT and DAY assuming that the response variable O3 $\sim$ Poisson(μ_i).

Model	Systematic component $g(\mu_i) = \log(\mu_i)$
I	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i$
II	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + \alpha_4 \mathrm{DAY}_i$
III	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + f(\mathrm{DAY}_i)$
IV	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + \alpha_4 \mathrm{DAY}_i + \alpha_5 \mathrm{TEMP}_i \times \mathrm{DAY}_i + \alpha_6 \mathrm{IBT}_i \times \mathrm{DAY}_i$
V	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \mathrm{TEMP}_i \beta_1(\mathrm{DAY}_i) + \mathrm{IBT}_i \beta_2(\mathrm{DAY}_i)$

For Model I, only the individual effect of the VIS, TEMP and IBT explanatory variables were considered. In Model II, the individual effects of these three covariates plus the effect of the DAY variable were incorporated in a linear manner, whereas in the Model III the individual effect of the DAY explanatory variable is included nonlinearly by using a smooth function. Model IV considers the individual contributions of VIS, TEMP, IBT and DAY explanatory variables, plus the interaction effects of the TEMP and IBT explanatory variables with the DAY variable. Finally, Model V corresponds to a PVCGLM where the explanatory variables TEMP and IBT interact with the variable DAY in nonlinear fashion. Table 2 contains the ML and MPL estimates associated with the parametric component for the five fitted models; the respective standard errors appear in parentheses.

Table 2: AIC, R^2, ML and MPL estimates (standard error) for all five fitted models to the Ozone data.

Parameters	Model				
	I	II	III	IV	V
α_0	0.65 (0.11)	0.79 (0.11)	1.09 (0.16)	0.88 (0.21)	1.18 (0.17)
α_1	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)
α_2	0.02 (0.00)	0.03 (0.00)	0.01 (0.00)	0.02 (0.00)	—
α_3	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	—
α_4	—	−0.001 (0.00)	—	−0.001 (0.00)	—
α_5	—	—	—	−0.00 (0.00)	—
α_6	—	—	—	0.00 (0.00)	—
AIC	1890.71	1861.27	1752.56	1863.66	1735.76
R^2	0.682	0.691	0.752	0.690	0.754

It should be noted that the p-values (omitted here) associated with the parameters of each fitted model are smaller than 0.05, thus indicating the contributions of the individual and interaction effects are statistically significant. Note also that the parameter estimates (associated with the parametric component) obtained from the different fitted models are quite similar and accurate. The last two rows of the Table 2 shows the Akaike Information Criterion (AIC) and R^2 values, respectively. It is clear that the PVCGLM, for which the $\text{AIC}(\lambda_1, \lambda_2) = 1735.76$, presents the best fit to the Ozone data, followed by Model III with an $\text{AIC} = 1752.56$, which is confirmed by the QQ-plots presented in Figure 3; see, specifically, Figures 3(c) and 3(e). Note also that the R^2 associated with our model is higher than Models I, II and IV, and slightly higher that Model III.

For the PVCGLM the estimates of the smoothing parameters λ_1 and λ_2 as well as the corresponding df's were obtained by the procedure proposed by Ibacache-Pulgar *et al.* (2013) [20], and are presented in Table 3. Figures 4(a) and 4(b) show the estimated smooth functions under PVCGLM and the corresponding approximate SEB (dashed curves).

Table 3: Fit summary for smoothing components under PVCGLM fitted to Ozone data.

	Smooth function	
	β_1(DAY)	β_2(DAY)
df(λ_k)	6.894	7.228
λ_k	89050.050	5886.339

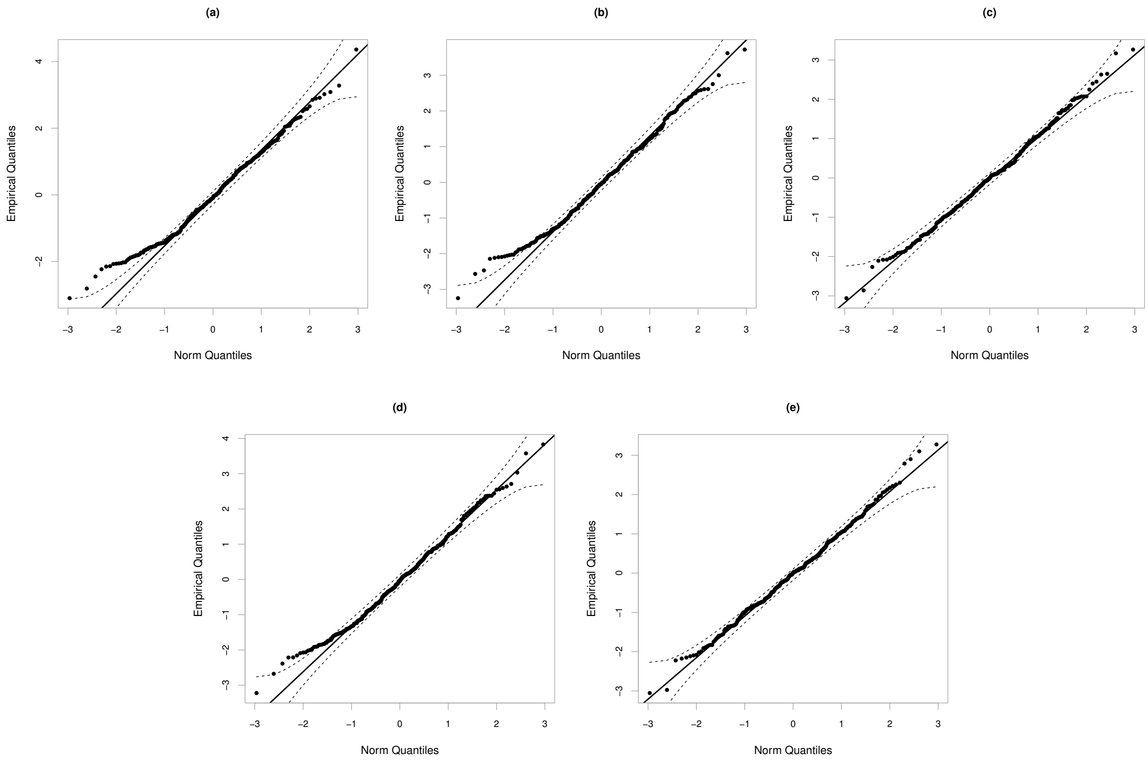

Figure 3: Normal probability plots to Ozone data: Model I (a), Model II (b), Model III (c), Model IV (d) and Model V (e).

Note that the plots confirm the nonlinear trends of the interaction effects between (TEMP, DAY) and (IBT, DAY).

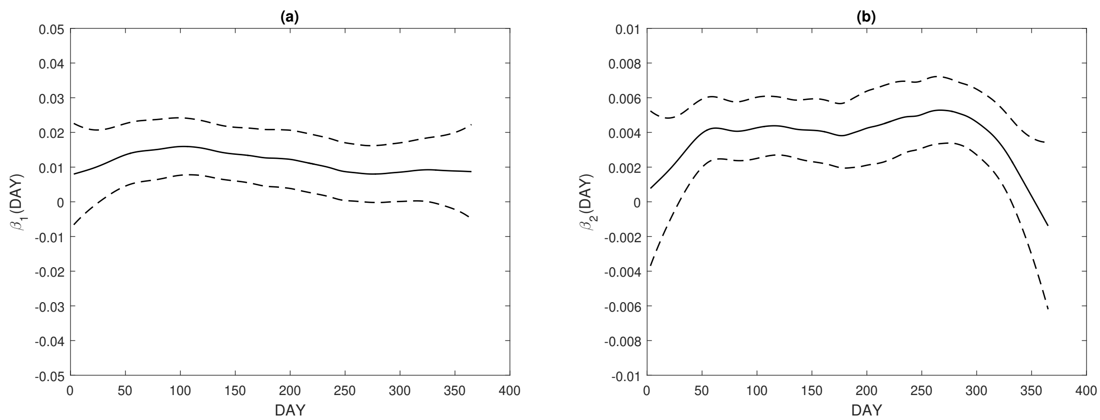

Figure 4: Plots of the estimated smooth functions β_1 (a) and β_2 (b) and their approximate pointwise SEB denoted by the dashed lines, Ozone data.

5.3. Local influence analysis

As mentioned earlier, the measure $\mathrm{LD}(\boldsymbol{\omega})$ is useful for assessing the distance between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_\omega$. In order to identify influential potentially observations on MPL estimators under the fitted PVCGLM model to Ozone data, we present some index plots of $B_i = B_{\boldsymbol{e}_i}(\boldsymbol{\gamma})$, for $\boldsymbol{\gamma} = \boldsymbol{\alpha}, \boldsymbol{\beta}_k$ and $k \in \{1, 2\}$.

Case-weight perturbation

Figure 5 shows the index plot B_i for the case-weight perturbation scheme under the fitted model. Note at Figure 5, that observations #125, #219, #167 and #258 are more influential for the MPL estimator $\widehat{\boldsymbol{\alpha}}$, whereas observations #219, #221 and #222 are influential for the MPL estimator $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$, respectively. When we introduce an additive perturbation to the response variable, the results are analogous to those observed under the case-weight perturbation scheme, and therefore the graphs are omitted.

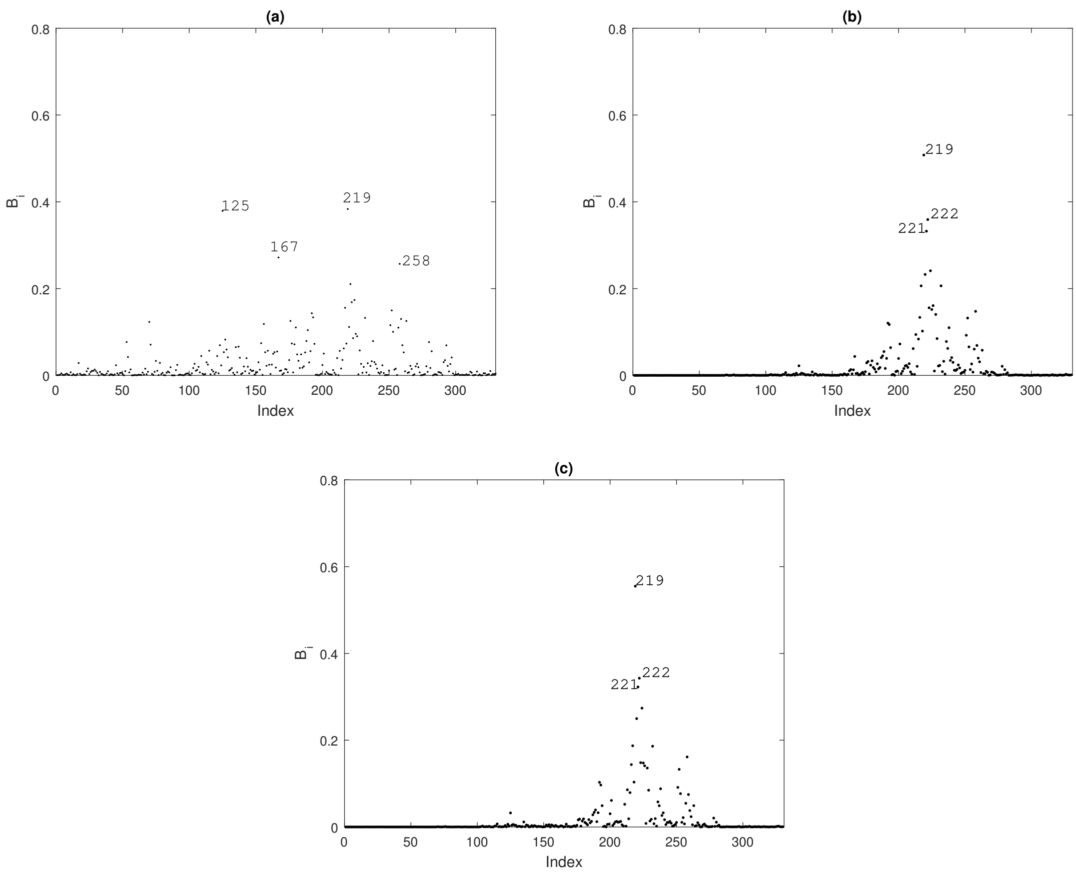

Figure 5: Index plots of B_i for assessing local influence on $\widehat{\boldsymbol{\alpha}}$ (a), $\widehat{\boldsymbol{\beta}}_1$ (b) and $\widehat{\boldsymbol{\beta}}_2$ (c) considering case-weight perturbation, Ozone data.

Explanatory variable additive perturbation

By perturbing the explanatory variable in an additive way, it becomes clear that s observations #125, #219 and #167 are more influential for the MPL estimator $\widehat{\boldsymbol{\alpha}}$, whereas observations #219, #221 and #222 are influential for the MPL estimator $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$, respectively; see Figure 6.

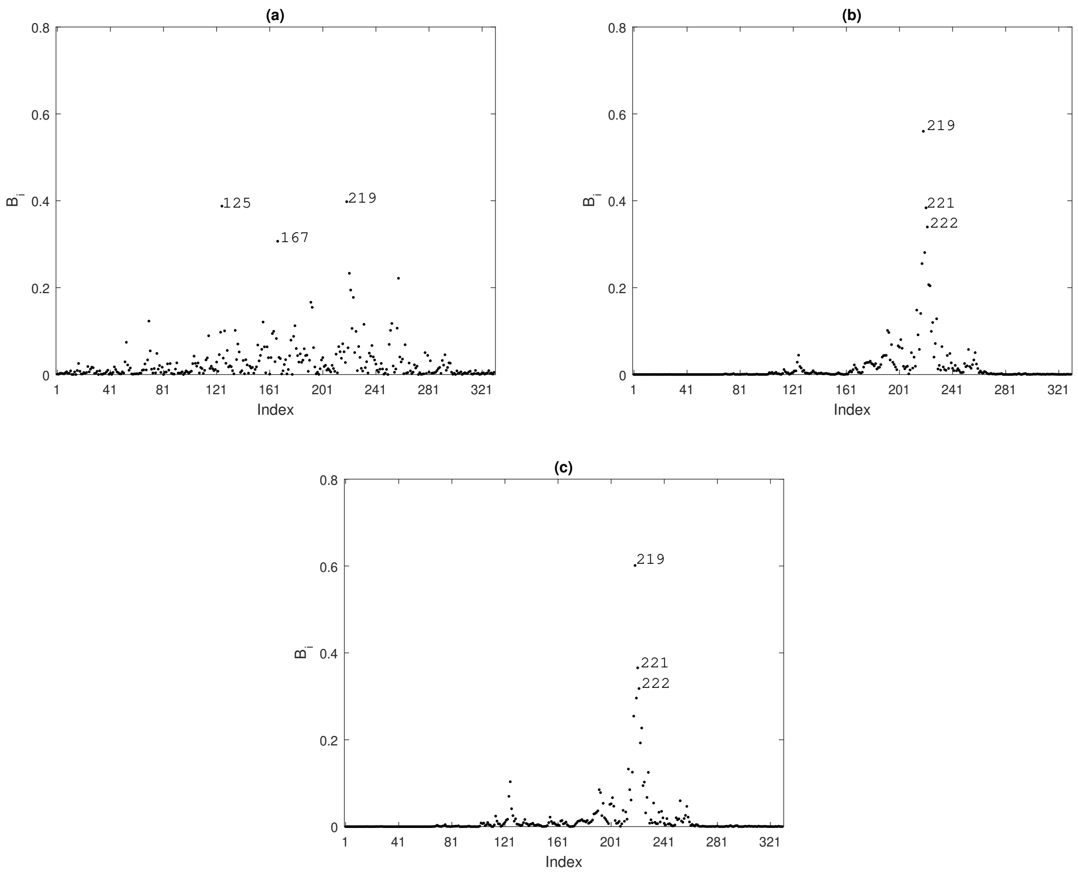

Figure 6: Index plots of B_i for assessing local influence on $\widehat{\boldsymbol{\alpha}}$ (a), $\widehat{\boldsymbol{\beta}}_1$ (b) and $\widehat{\boldsymbol{\beta}}_2$ (c) considering explanatory variable perturbation, Ozone data.

Based on the local influence analysis, we conclude that the MPL estimators of the regression coefficient and of the smooth functions are sensitive to perturbations introduced into the data or to the model. In addition, this analysis revealed that the observations that were detected as influential for the parametric component are not necessarily influential for the nonparametric component, and vice versa. For instance, under the case-weight perturbation scheme, observations #125, #219, #167 and #258 were detected as influential for the parametric component. However, of these three observations, only observation #219 is indicated as influential for the nonparametric component, in addition to observations #221 and #222. In general, similar results were obtained when the explanatory variable is additively perturbed.

5.4. Confirmatory analysis

In order to investigate the impact on the model inference when the observations detected as potentially influential in the diagnostic analysis are removed, we present the relative changes (RCs) in the MPL estimate of α_j for $j \in \{1,2\}$ after removing from the data set the influential potentially observations (%). The RC is defined as $\mathrm{RC}_\xi = |(\widehat{\xi} - \widehat{\xi}_{(\mathrm{I})})/\widehat{\xi}| \times 100\%$, where $\widehat{\xi}_{(\mathrm{I})}$ denotes the MPL estimate of ξ, with $\xi = \alpha_j$, after the corresponding observation(s) are removed according to the set I. Table 4 presents the RCs in the regression coefficient estimates after removing the observations indicated as potentially influential for the parametric component of the model.

Table 4: Relative changes (in %) in the MPL estimates of α_j under the PVCGLM.

Dropped observation	Parameters		Relative changes	
	α_0	α_1	RC_{α_0}	RC_{α_1}
125	1.17365	−0.001616	0.977	1.635
167	1.16798	−0.001592	1.455	0.125
219	1.18324	−0.001626	0.167	2.264
258	1.21007	−0.001622	2.096	2.013
125–167	1.17727	−0.001623	0.672	2.075
125–219	1.18273	−0.001628	0.211	2.389
125–258	1.52686	−0.001638	28.823	3.019
167–219	1.17701	−0.001603	0.694	0.817
167–258	1.53185	−0.001614	29.245	1.509
219–258	1.17689	−0.001609	0.703	1.195
125–167–219	1.15265	−0.001637	2.748	2.955
167–219–258	1.17625	−0.001593	0.758	0.189
125–167–219–258	1.51397	−0.001654	27.737	4.025

On the other hand, Table 5 shows the RCs observed in the estimation of the regression coefficient once the observations detected as potentially influential for the nonparametric component of the model are excluded.

Table 5: Relative changes (in %) in the MPL estimates of α_j under the PVCGLM considering the observations detected as influential on the nonparametric component.

Dropped observation	Parameters		Relative changes	
	α_0	α_1	RC_{α_0}	RC_{α_1}
none	1.18	−0.001		
219	1.183	−0.002	0.167	2.264
221	1.161	−0.002	2.041	1.132
222	1.552	−0.002	30.949	2.075
219–221	1.142	−0.002	3.641	1.635
219–222	1.521	−0.002	28.330	4.779
221–222	1.151	−0.002	2.865	1.258
219–221–222	1.186	−0.002	0.092	2.955

Considering these results, we conclude that, although some RCs are large, inferential changes are not detected. It is interesting to notice from Tables 4 and 5 the coherence with the local influence diagnostic shown previously. For instance, removal of the observations sets $\mathrm{I} = \{167, 258\}$ and $\mathrm{I} = \{125, 258\}$, which contain observations detected as influential potentially for the parametric component, leads to significant changes in the MPL estimates, mainly in $\hat{\alpha}_0$, of the order of 29.245% and 28.823%, respectively; see Table 4.

Note also that the individual removal of observation #258 produces a RC of order of 2.096%. On the other hand, the removal of the observations set $\mathrm{I} = \{219, 222\}$, whose observations were detected as influential potentially for nonparametric component, leads to significant changes in the MPL estimate of α_0, 28.330%. It is also observed that the removal of observation #222 produces a RC of 30.949%. This indicates the need of a diagnostic examination. The changes produced in the estimates of the smooth functions are presented in Figure 7.

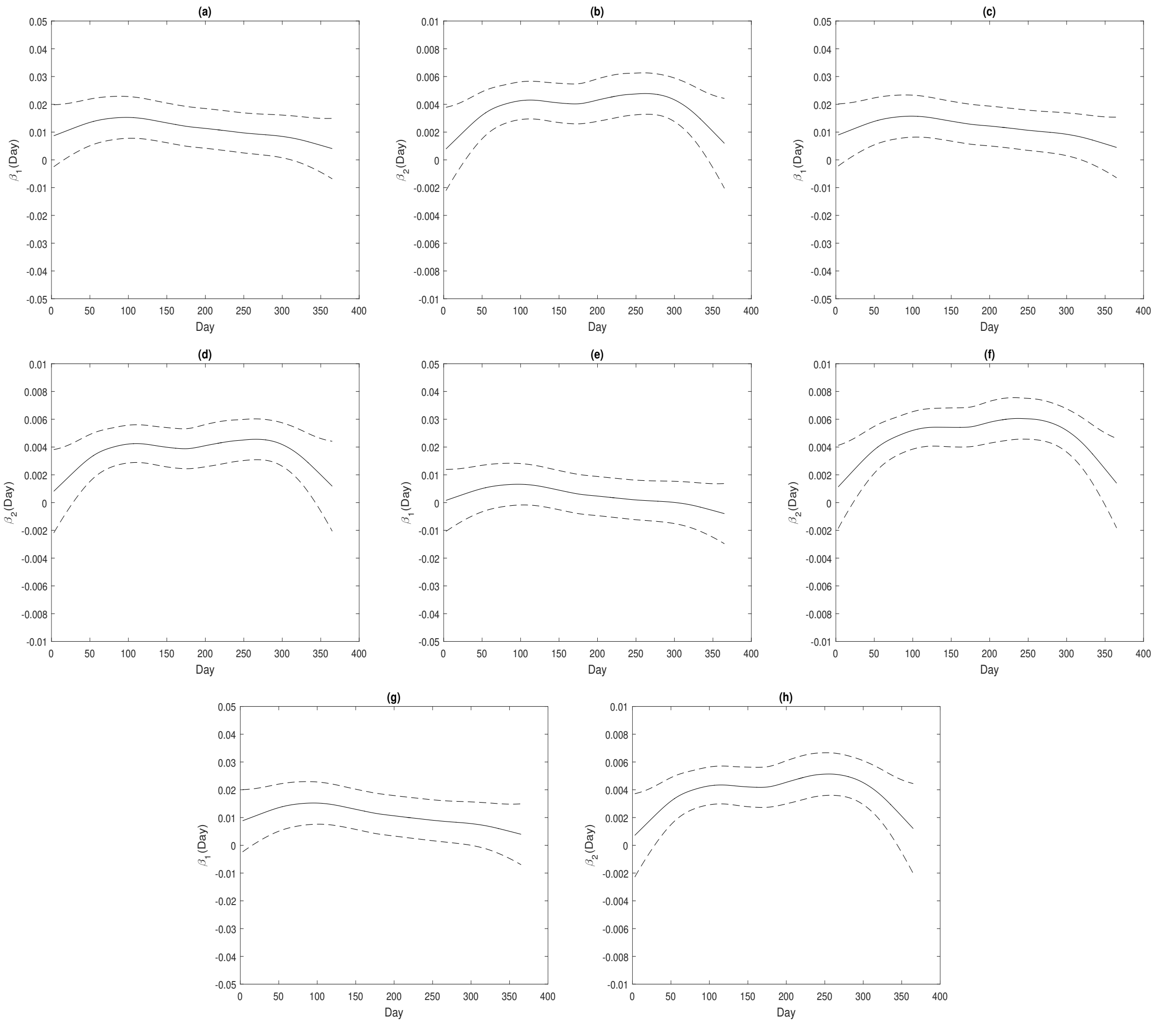

Figure 7: Plots of estimated smooth functions, $\widehat{\beta}_1$ and $\widehat{\beta}_2$, for the Ozone data and their approximate pointwise SEB denoted by the dashed lines: excluding observations #219 and #221 (a)–(b), excluding observations #219 and #222 (c)–(d), excluding observations #221 and #222 (e)–(f), excluding observations #219, #221 and #222 (g)–(h).

5.5. Computational aspects and summary of our methodology

The fitted models, quantile-quantile plots (qqplot) with simulated envelopes, and local influence were done in Matlab version R2015a and are available via email for people interested in replicating our analyses. Additionally, it is important to note that there are at least two libraries in the free R software that can be used to fit our models, for example, the mgcv library (`https://cran.r-project.org/web/packages/mgcv/index.html`) and gamlss (`https://cran.r-project.org/web/packages/gamlss/index.html`). However, there is no R library that performs local influence for the models studied in our work. Next, we summarize all the stages of our methodology through an algorithm (Algorithm 2).

Algorithm 2 – Some guidelines for applying the analysis of local influence on the PVCGLM.

1. Make a scatterplot and analyze the trend of the variables. Depending on the trend of your data you should use linear, quadratic, or polynomial function (parametric function). Alternatively, you could use a non-linear parametric form or nonparametric function (cubic spline for instance).

2. Decide if your response variable is a discrete random variable (Bernoulli, Binomial, Poisson, etc) or continuous (Normal, Gamma, etc) belonging to the exponential family. After that, decide which is the best option for your link function (log, square root, inverse, logit, probit, etc) and try different parametric, nonparametric, or semiparametric for the systematic component of a generalized linear model.

3. Choose the best model based on some criterion such as R-square or AIC.

4. Apply the local influence method and if you have some outlying observation study the relative changes deleting some observations. If you do not have outlying observations, make some conclusions about your data set.

6. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

In this paper we study some aspects of the partially varying-coefficient generalized linear models. Specifically, we derive a weighted back-fitting iterative process to estimate the regression coefficients, the smooth functions and the dispersion parameter associated with our model. The variance-covariance matrix of the maximum penalized likelihood estimators was approximated by the inverse of penalized Fisher information matrix, and the effective degrees of the freedom of the nonparametric components were calculated from the estimates obtained in convergence of the iterative process. Furthermore, we extended the local influence method and obtained closed expressions for the Hessian matrix and the perturbation matrix under different perturbation schemes. We performed a statistical data analysis with a real data set on ozone concentration and some meteorological variables. The study showed the advantage of incorporating a semiparametric additive term when there are predictors whose interactions contribute nonlinearly to the model, and the utility of the local influence method to detect influential observations on the maximum penalized likelihood estimators. One of the main limitations of our model is the absence of a structure that allows modeling the correlation in those data sets that have a time component, this being one of the main lines of research to be developed. In addition, we believe that the exploration of new perturbation schemes is necessary, mainly in the interaction components and the smoothing parameter.

Finally, we recommend the use of partially varying-coefficient generalized linear models and the local influence method when the response variable belongs to the exponential family and the interactions between the explanatory variables can be modelled through smooth functions, and our interest is to evaluate the sensitivity of the maximum penalized likelihood estimator.

ACKNOWLEDGMENTS

The authors thanks Editors and Referees for their suggestions which allowed us to improve the presentation of this work. G. Ibacache-Pulgar acknowledges funding support by grant FONDECYT 11130704, Chile.

REFERENCES

[1] Berhane, K. and Tibshirani, R. (1998). Generalized additive models for longitudinal data, *Canadian Journal of Statistics*, **26**, 517–535.

[2] Billor, N. and Loynes, R. (1993). Local influence: a new approach, *Communications in Statistics – Theory and Methods*, **22**, 1595–1611.

[3] Breiman, L. and Friedman, J. (1985). Estimating optimal transformations for multiple regression and correlation, *Journal of the American Statistical Association*, **80**, 580–598.

[4] Cai, Z.; Fan, J. and Li, R. (2000). Efficient estimation and inferences for varying-coefficient models, *Journal of the American Statistical Association*, **95**, 888–902.

[5] Cook, R. (1986). Assessment of local influence (with discussion), *Journal of the Royal Statistical Society B*, **48**, 133–169.

[6] Emami, H. (2017). Local influence for Liu estimators in semiparametric linear models, *Statistical Papers*, **19**, 529–544.

[7] Escobar, L. and Meeker, W. (1992). Assessing local influence in regression analysis with censored data, *Biometrics*, **48**, 507–528.

[8] Espinheria, P.; Ferrari, S. and Cribari-Neto, F. (2008). On beta regression residuals, *Journal of Applied Statistics*, **35**, 407–419.

[9] Fan, J. and Zhang, W. (2008). Statistical methods with varying coefficient models, *Statistics and Interface*, **1**, 179–195.

[10] Fan, J. and Huang, T. (2005). Profile likelihood inferences on semiparametric varying-coefficient partially linear models, *Bernoulli*, **11**, 1031–1057.

[11] Faraway, J. (2006). *Extending the Linear Model with R Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Chapman and Hall.

[12] Ferrari, S.; Espinheira, P. and Cribari-Neto, F. (2011). Diagnostics tools in beta regression with varying dispersion, *Statistica Neerlandica*, **65**, 337–351.

[13] Ferreira, C. and Paula, G. (2016). Estimation and diagnostic for skew-normal partially linear models, *Journal of Applied Statistics*, **44**, 3033–3053.

[14] Finley, A. (2011). Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence, *Methods in Ecology and Evolution*, **2**, 142–154.

[15] Green, P. and Silverman, B. (1994). *Nonparametric Regression and Generalized Linear Models*, Chapman and Hall, Boca Raton.

[16] Hastie, T. and Tibshirani, R. (1993). Varying-coefficient models, *Journal of the Royal Statistical Society B*, **55**, 757–796.

[17] Hastie, T. and Tibshirani, R. (1990). Generalized additive models, *Statistical Science*, **3**, 297–318.

[18] He, K.; Wang, Y.; Su, W. and Yang, H. (2022). A varying-coefficient regression approach to modelling the effects of wind speed on the dispersion of pollutants, *Environmental and Ecological Statistics*. http://doi.org/10.1007/s10651-022-00535-6

[19] Ibacache-Pulgar, G.; Paula, G. and Galea, M. (2012). Influence diagnostics for elliptical semiparametric mixed models, *Statistical Modelling*, **12**, 165–193.

[20] Ibacache, G.; Paula, G. and Cysneiros, F. (2013). Semiparametric additive models under symmetric distributions, *TEST*, **22**, 103–121.

[21] Ibacache, G. and Paula, G. (2011). Local Influence for student-t partially linear models, *Computational Statistics and Data Analysis*, **55**, 1462–1478.

[22] Ibacache-Pulgar, G. and Reyes, S. (2018). Local influence for elliptical partially varying-coefficient model, *Statistical Modelling*, **18**, 149–174.

[23] Ibacache-Pulgar, G.; Figueroa-Zuñiga, J. and Marchant, C. (2021). Semiparametric additive beta regression models: inference and local influence diagnostics, *REVSTAT – Statistical Journal*, **19**, 255–274.

[24] Li, J.; Xia, X.; Wong, W.E. and Nott, D. (2018). Varying-coefficient semiparametric model averaging predicction, *Biometrics*, **14**, 1417–1426.

[25] Liu, S.; Leiva, V.; Zhuang, D.; Ma, T. and Figueroa-Zuñiga, J.I. (2021). Matrix differential calculus with applications in the multivariate linear model and its diagnostics, *Journal of Multivariate Analysis*, pages in press available at https://doi.org/10.1016/j.jmva.2021.104849

[26] Ma, S.; Yang, L.; Romero, R. and Cui, Y. (2011). Varying coefficient model for gene-environment interacction: a non-linear look, *Bioinformatics*, **15**, 2119–2126.

[27] McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, 2nd ed., London: Chapman and Hall.

[28] Poon, W. and Poon, Y. (1999). Conformal normal curvature and assessment of local influence, *Journal of the Royal Statistical Society B*, **61**, 51–61.

[29] Ouwens, M.; Tan, F. and Berger, M. (2001). Local influence to detect influential data structures for generalized linear mixed models, *Biometrics*, **57**, 1166–1172.

[30] Rigby, R. and Stasinopoulos, D. (2005). Generalized additive models for location, scale and shape, *Appl. Statist.*, **54**, 507–554.

[31] Rocha, A. and Simas, A. (2011). Influence diagnostics in a general class of beta regression models, *TEST*, **20**, 95–119.

[32] Sanchez, L.; Leiva, V.; Helto, S.; Marchant, C. and Sarabia, J.M. (2021). A new quantile regression model and its diagnostics analytics for a weibull distributed response with applications, *Mathematics*, **9**, 2768. http://doi.org/10.3390/math9212768

[33] Thomas, W. (1991). Influence diagnostics for the cross-validated smoothing parameter in spline smoothing, *Journal of the American Statistical Association*, **9**, 693–698.

[34] Zhang, J.; Zhang, X.; Ma, H. and Zhiya, C. (2015). Local influence analysis of varying coefficient linear model, *Journal of Interdisciplinary Mathematics*, **3**, 293–306.

[35] Zhu, H. and Lee, S. (2001). Local influence for incomplete-data models, *Journal of the Royal Statistical Society B*, **63**, 111–126.

[36] Zhu, H. and Lee, S. (2003). Local influence for generalized linear mixed models, *Canadian Journal of Statistics*, **31**, 293–309.