Dhaka Univ. J. Sci. 74(1): 138-159, 2026 (January)
Prof. M. Muhibur Rahman Memorial Issue
ISSN 1022 - 2502, E-ISSN 2408- 8528
DOI: https://doi.org/10.3329/dujs.v74i1.83640

Assessment of Influential Points Detection in Gamma-Pareto Regression Residuals Using Diagnostic Measure, Difference of Fits

Nasir Saleem[1]*, Atif Akbar[1], Muhammad Aman Ullah[1], and Khuban Meral[2]

Department of Statistics, Bahauddin Zakariya University (BZU), Multan, Pakistan, 60800
Department of English, Bahauddin Zakariya University (BZU), Multan, Pakistan, 60800

(Received : 9 August 2025; Accepted : 1 January 2026)

Abstract

The principal diagnosis process in order to achieve reliable output of the regression is the influential analysis. Same applies in case of the generalized linear model. In the present paper, the practical comparison of the functioning of different residuals of the Gamma-Pareto regression model is made in order to define the points of influence. The residuals of Gamma-Pareto regression model are further divided into two i.e., standardized and adjusted residuals. The said two residuals have both undergone difference of fits and consequently comparison of these residuals in the finding of influential points were carried out by simulation and Ardennes related data set. Simulation result shows that in case dispersion parameter is very low, likelihood residuals will perform better compared to the others and all of the adjustment type of residuals will perform identically but not superior to the standardized. Although, in case of large values of dispersion parameter, all the standardized residuals behave similarly, and they are superior to likelihood residuals in revealing the presence of influential points.

Keywords: Deviance; DFFITS; GLM; G-PRM; Identity link function; Influential Points; Likelihood; Pearson; Standardized and Adjusted Residuals

I. Introduction

Very serious effects can be caused through single observation on regression analysis. It can provide inaccurate approximations of coefficients and presents false statistical description of covariance matrix. Examples of these are the logistic and probit regression responses. Another variation of the GLM was mentioned by McCullagh and Nelder[33]. It can be used, when the response variable is positively skewed and mean proportional to variance with constant coefficient of variation. This model is commonly applied in the reliability or survival analysis (Lawless,[31].

Practically, non-influential observation is not true in the GLM (see Hardin and Hilbe,[30]. Cook[17] proposed diagnostics of the linear models (LMs). These influence diagnostics were explained by Belsley et al.15 in variety of dimensions. The problem of influence diagnostics in the GLM is still the central topic of consideration and discussion (see Preisser and Qaqish,[36]; Pregibon[35]; Vanegas et al.,[39]; Williams,[40]; Zhu and Lee, [32]. To check the influence of an outcome, pearson residuals are commonly used on the influence diagnostics. In addition, this author (Williams[40]) demonstrated the effect of diagnostics based on deviance residuals.

It has been stated that residuals play a significant role in both the diagnosis modeling and the determination of the goodness of fit. The LM only tests the model diagnostics with raw residuals, whereas the GLM has an excessive number of structures of residuals such as Pearson, deviance and likelihood. There are three most typical types of residuals that one can use in influence diagnostics of the GLM: the Pearson ones, the deviance ones and the likelihood ones. These residuals do not follow the normality criteria that is they take up different probability distribution. Therefore, residuals cannot be zero as this would be used to fulfill this assumption of normality. The two existing significant theories of adjusted residuals are the adjusted deviance residuals as provided by Pierce and Schafer[34] and adjusted Pearson residuals as suggested by Cordeiro[20] based on Cox and Snell[21]. These theories are directed at achieving the normality. Nevertheless, these findings are also obtained when analyzing the adjusted Pearson residuals in the exponential nonlinear model family (see Simas and Cordeiro, [37] as well as in the beta regression (Anholeto et al., [10]. The regression diagnostics plays a significant role in the model building of linear regression. Regression diagnostics is one of the strategies to detect the unusual observations (Amin et al.[5-7]. Other literature that is available regarding the detection of influential observations in generalized linear regression model includes that of Venezuela et al. (2007), Amin et al.[5], Amin et al.[9], and Amin et al.[8], respectively, on influential observation in the gamma regression model (GRM).

Gamma-Pareto distribution (G-PD) is a novel distribution. The G-PD was created by Alzaatreh et al. [2]. G-PD Hanum et al.[27, 28] defines an element of the exponential family distribution. Therefore, the regression modeling of Gamma-Pareto regression model (G- PRM) could be developed under GLM. Hanum et al.[27, 28] develop analytically GLM G-PD. GLM Gamma is founded on the Gamma distribution (GD), and is utilized regularly. The right skew data are commonly modelled when GLM Gamma is employed as the analysis tool. Alzaatreh et al.[2] stated mathematical associations between G-PD and GD. The idea of using G-PD to model and predict extreme monthly rainfall was proposed by Hanum et al. [26], and hence this is rational as the G-PD is a

[*]Author for correspondence. e-mail: nasirsaleem160@gmail.com

development of the GD. The G-PD is a statistical distribution which unites the features of the Gamma and the Pareto distributions. It's often used to model heavy-tailed data, such as financial returns or insurance claims. It is one such GLM which uses the G-PD as the distribution of the response variable of the regression. This enables the model to take account of heavy-tailedity of the data.

The non-normal response regression models are mostly expressed in the form of GLM. Hanum et al [27, 28] examined correlation between explanatory variable and a simulated G-PD where the allocated response variable took a Gamma distribution based GLM. Hanum et al.[29] wrote about modeling of the Gamma-Pareto distributed data to estimate monthly rainfall via use of the TRMM data to fit using GLM Gamma. According to Zheng et al.[42], it is significant to chart the Shifted Gamma-Generalized Pareto Distribution model in order to regulate the safety continuum and estimate the crashes. The new Log-Gamma-Pareto Distribution is proposed by Ashour et al.[11]. Alzaatreh and Ghosh[1] presented an introduction of a new Gamma-Pareto (IV) distribution and its implementation. Andrade, de Figueiredo, and de Andrade[23] explained Gamma generalized Pareto distribution and the application on the survival analysis. Alzaghal[3] applied Exponentiated Gamma-Pareto distribution to the susceptibility to cancer of the bladder. Weighted Gamma-Pareto distribution was discussed by Dar et al.[22] and also state its use. The Pearson residuals have been used extensively in influence diagnostics, when checking the influential observations. Cordeiro[20] suggests the adjusted Pearson residuals, which though they are the ones of Cox and Snell[21]. The aims of these theories are to attain normality. Simas and Cordeiro [37] carried out a study based on the exponential family of non-linear models and concluded that analysis of the adjusted Pearson residuals (APR) will yield a similar result. Some of the approaches in the literature to diagnose significant observation or observations to the LM are namely Cook and Weisberg [19], Atkinson[13], Cook [18], and Chatterjee and Hadi [16]. Lee[32] on the other an approach of evaluating partial influence in the GLM was provided. One of the methods under which Thomas and Cook[38] suggested to analyze impact on the GLM regression coefficients was the impact of a variable in a regression analysis. Amin et al.[43-44] talked about the control charts on Beta regressions residuals using different link functions: application to the data of thermal power plants. The Gamma response model was first invented by the Aslam et al. (2024), and they dealt with the usage of the Shewhart ridge profiling.

We learn influence diagnostics in terms of the Pearson residuals (PR), deviance residuals (DR) and likelihood residuals (LR) of their standardized and adjusted form respectively in the G-PRM. The given study is applicable because standardized and adjusted residuals of Difference of fit (DFFITS) influence diagnostics are needed. This paper is organized as follows: section wise discuss about the G-PRM and its estimation method and presents the G-PRM residuals and calculation of standardized and adjusted residuals. The third Section looks at the influence diagnostics of the G-PRM standardised and adjusted residuals. In Section next results are summarized of a simulation study and real measures of influence of the G-PRM using regularized residuals as well as round off residuals. Finally, it is end of the research work provided in last Section.

II. Material and Methods

Notation and Estimation Techniques for the Gamma-Pareto Regression Model

According to Alzaatreh et al.[2] probability density function (pdf) of G-PD is provided by:

$$f(y;\alpha,\beta,\gamma)=\frac{\gamma^{-1}}{\beta^{\alpha}\Gamma(\alpha)}\left(log\left(\frac{y}{\gamma}\right)\right)^{\alpha-1}\left(\frac{y}{\gamma}\right)^{-\left(\frac{1}{\beta}+1\right)} \quad (1)$$

with α, β,γ> 0 and y > γ.

The mean and variance of G-P distribution are, $E\big(a(y)\big)=\alpha\beta$, $V(a(y))=\alpha\beta^2$ respectively and $a(y)$ is a function of $f(y)$, where y is a G-P random variable. According to Hanum et al.[27,28], With parameters, Eq. (1) can be modified $\alpha=\frac{1}{\phi}$ and $\beta=\mu\phi$. The Gamma Pareto density for y under these conditions is given by

$$f(y;\mu,\phi)=\frac{\gamma^{-1}}{(\mu\phi)^{\frac{1}{\phi}}\Gamma\left(\frac{1}{\phi}\right)}\left(log\left(\frac{y}{\gamma}\right)\right)^{\frac{1}{\phi}-1}\left(\frac{y}{\gamma}\right)^{-\left(\frac{1}{\mu\phi}+1\right)} \quad (2)$$

with $y\geq 0,\ \mu>0\ and\ \phi>0$. Since the mean and variance of y are $E(y)=\mu$ and $V(y)=\phi V(\mu)=\phi\mu^2$.

For the ith observation, let $x_{i1},x_{i2},\dots,x_{ip}$ represent the p non-stochastic regressors. Following that, the G-PRM for the response variable y mean is provided.

According to Hanum et al. [27,28], Link function g in GLM is $g(\mu_i)=X_i^T\beta=\eta_i$ where $\mu_i=E\big(a(y_i)\big)$, which are presented in Table 1.

Table 1. Different link function of Gamma Pareto Distribution

Link function	Form of link function	Reference
Inverse link function	$\mu=\frac{1}{X'\beta}$	Hanum et al. (2016)
Identity link function	$\mu=X'\beta$	
Log link function	$\mu=\log(X'\beta)$ $\mu=e^{X'\beta}$	

Estimation of Parameters for the Glm Gamma-Pareto Regression Model

Finding the likelihood function's derivative with respect to β_j is the first step in estimating the parameter β_j using maximum likelihood and τi is the function of $f(y)$. By Eq. (2)

$$\frac{\partial l}{\partial \beta_j} = U_j = \sum_{i=1}^{N}\left[\frac{\partial l_i}{\partial \beta_j}\right] = \sum_{i=1}^{N}\left[\frac{\partial l_i}{\partial \tau_i}\frac{\partial \tau_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right] \quad (3)$$

Now

$$\frac{\partial l_i}{\partial \tau_i} = a(y)b'(\tau) + c'(\tau) = \beta^{-2}\left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)$$

$$\frac{\partial \tau_i}{\partial \mu_i} = \frac{1}{\frac{\partial \mu_i}{\partial \tau_i}} = \frac{1}{\frac{\partial \alpha\beta}{\partial \beta}} = \frac{1}{\alpha}$$

$$\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i}x_{ij}$$

Where $\frac{\partial \mu_i}{\partial \eta_i}$ based on the GLM's link function. So, the score for β_j in GLM Gamma-Pareto is

$$\frac{\partial l}{\partial \beta_j} = U_j = \sum_{i=1}^{N}\alpha^{-1}\beta^{-2}\left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)\frac{\partial \mu_i}{\partial \eta_i}x_{ij} \quad (4)$$

Lastly, the jth score is presented.

$$U_j = \sum_{i=1}^{N}\left[var\left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)\right]^{-1}\left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)\frac{\partial \mu_i}{\partial \eta_i}x_{ij}$$

The variance U_j is

$$var(U_j) = \zeta_{jk} = \sum_{i=1}^{N}\frac{x_{ij}x_{ik}}{\left[var\left(log\left(\frac{y_i}{\gamma}\right)\right)\right]}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 = X^TWX$$

Where,

$$W = \frac{1}{\left[var\left(log\left(\frac{y_i}{\gamma}\right)\right)\right]}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$$

Since the estimators of $\beta_j$is not in close form.

Iterative weighted least squares (IWLS) were proposed by (Dobson et al.,[24] as a method for estimating β_j.

It's the IWLS.

$$X^TWX\, b^{(m)} = X^TWz$$

$$b^{(m)} = (X^TWX)^{-1}(X^TWz) \quad (5)$$

And now, Using W and $var(U_j)$ for G-P and obtained the iteration for β_j as, i is a number of observations $i = 1,2,3,\dots,n$. and j are a number of parameters $j = 1,2,3,\dots,p \text{ and } k \neq j$

$$X^TWX\, b^{(m)} = \sum_{k=1}^{p}\sum_{i=1}^{N}\frac{x_{ij}x_{ik}}{\left[var\left(log\left(\frac{y_i}{\gamma}\right)\right)\right]}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 b_k^{(m-1)} + \frac{\left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)x_{ij}}{\left[var\left(log\left(\frac{y_i}{\gamma}\right)\right)\right]}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)$$

$$z_i = \sum_{i=1}^{N}x_{ij}b_k^{(m-1)} + \left(log\left(\frac{y_i}{\gamma}\right) - \mu_i\right)\frac{\partial \mu_i}{\partial \eta_i}$$

In the literature there are nowadays numerous possibilities of types of residuals in GLM (Hardin and Hilbe, [30]). And we only used Pearson, deviance and likelihood residuals are adjusted form and standardized respectively.

The Pearson residuals in the G-PRM is given by

$$R_{i(pr)} = \frac{y_i - \hat{\mu}}{\hat{\mu}} \quad (6)$$

$g(\mu_i) = \eta_i = \frac{1}{X^T\beta}$ and fitted model, $\hat{\mu} = \eta_i = \frac{1}{X^T\beta}$

The standardized Pearson residuals is present by using Eq. (6)

$$R_{i(spr)} = \frac{R_{i(pr)}}{\sqrt{\phi(1-h_{ii})}} \quad (7)$$

Since h_{ii} is the ith diagonal element of the hat matrix $H = W^{\frac{1}{2}}X(X^TWX)^{-1}X^TW^{\frac{1}{2}}$

The adjusted Pearson residuals is defined by using Eq. (6)

$$R_{i(apr)} = \frac{R_{i(pr)} - \hat{E}(R_{i(pr)})}{\sqrt{\hat{V}(R_{i(pr)})}} \quad (8)$$

$\hat{E}(R_{i(pr)})$ and $\hat{V}(R_{i(pr)})$ are the expected value and variance of $R_{i(pr)}$ respectively.

The adjusted Pearson residuals approximately follow a normal distribution (Cordeiro,[20]).

Same, the deviance residuals for the G-PRM are given by

$$R_{i(dr)} = sign(y_i - \hat{\mu})\sqrt{|d_i|} \quad (9)$$

where $d_i = -2\left\{\ln\left(\frac{y_i}{\hat{\mu}}\right) - \left(\frac{y_i - \hat{\mu}}{\hat{\mu}}\right)\right\}$ and sign$(y_i - \hat{\mu})$ is signum function, which is defined as

$$sign(y_i - \hat{\mu}) = \begin{cases} + \; if \; y_i > \hat{\mu} \\ 0 \; if \; y_i = \hat{\mu} \\ - \; if \; y_i < \hat{\mu} \end{cases}$$

Eq. (9) is used to present the standardized deviance residuals.

$$R_{i(sdr)} = \frac{R_{i(dr)}}{\sqrt{\phi(1-h_{ii})}} \quad (10)$$

Adjusted residuals were first introduced by (Cox and Snell,[21]). According to (Cordeiro,[20]) and (Pierce and Schafer,[34]), the adjusted deviance residuals for both methods. The adjusted deviance residuals are defined by using Eq. (9)

$$R_{i(adr)} = \frac{R_{i(dr)} - \hat{E}(R_{i(dr)})}{\sqrt{\hat{v}(R_{i(dr)})}} \quad (11)$$

$\hat{E}(R_{i(dr)})$ and $\hat{V}(R_{i(dr)})$ are the expected value and variance of $R_{i(dr)}$ respectively.

The adjusted deviance residuals are normal distributed by (Pierce and Schafer's,[34].

Similarly, the likelihood residuals (McCullagh and Nelder,[33]) and the G-PRM are given by

$$R_{i(lr)} = sign(y_i - \hat{\mu})\sqrt{h_{ii}(R_{i(spr)})^2 + (1-h_{ii})(R_{i(sdr)})^2} \quad (12)$$

Eq. (12) is used to define the standardized likelihood residuals.

$$R_{i(slr)} = \frac{R_{i(lr)}}{\sqrt{\phi(1-h_{ii})}} \quad (13)$$

Eq. (12) is used to define the adjusted likelihood residuals.

$$R_{i(alr)} = \frac{R_{i(lr)} - \hat{E}(R_{i(lr)})}{\sqrt{\hat{V}(R_{i(lr)})}} \quad (14)$$

$\hat{E}(R_{i(lr)})$ and $\hat{V}(R_{I(lr)})$ are the expected value and variance of $R_{i(lr)}$ respectively.

II. Influence Diagnostics in Gamma-Pareto Regression Model

The LM has great importance in determination of model estimates and inferences in an agreement with what was observed by (Atkinson,[12]. These bad values may be influential and they may make a difference or they may be outliers. An outlier is produced by an outlier in response variable but an influential observation is one that is produced by an outlier in explanatory variable. A few of them are commented on G-PRM influence diagnostics since even the GLM using Pearson, Deviance and Likelihood (uncentered and adjusted) residuals is yet to be considered. The reason behind this is that, checking of the GLM influence diagnostics with the various GLM residuals has attracted small interest. The residuals topic of GLM was initially studied (Pregibon,[35]. The GLM influence assessment tools are calculated off the various GLM residuals.

The DFFITS is a diagnostic measure that attracted a great deal of attention to the literature referred to as influence: it is defined as part of the scaled difference between the fitted value of the entire data set and the fitted value of the data set after deleting the ith observation.

$$DFFITS_i = \frac{y_i - \hat{y}_i}{\sqrt{\hat{\phi}_i h_{ii}}} \quad (15)$$

Eq. (15) can also be written as

$$DFFITS_i = \frac{\widehat{W}_{ii}^{\frac{1}{2}} x_i^T (y_i - \hat{y}_i)}{\sqrt{\hat{\phi}_i h_{ii}}} \quad (16)$$

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (17)$$

The DFFITS for standardized Pearson residuals used Eq. (7)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (18)$$

$$t_{i(spr)} = R_{i(spr)}\sqrt{\frac{n-p-1}{n-p-(R_{i(spr)})^2}} \quad (18.1)$$

The DFFITS for adjusted Pearson residuals used Eq. (8)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (19)$$

$$t_{i(apr)} = R_{i(apr)}\sqrt{\frac{n-p-1}{n-p-(R_{apr})^2}} \quad (19.1)$$

The DFFITS for standardized deviance residuals used Eq. (10)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}}$$

$$t_{i(sdr)} = R_{i(sdr)}\sqrt{\frac{n-p-1}{n-p-(R_{i(sdr)})^2}}$$

The DFFITS for adjusted deviance residuals used Eq. (11)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}}$$

$$t_{i(adr)} = R_{i(adr)}\sqrt{\frac{n-p-1}{n-p-(R_{adr})^2}}$$

The DFFITS for standardized likelihood residuals used Eq. (13)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}}$$

$$t_{i(slr)} = R_{i(slr)}\sqrt{\frac{n-p-1}{n-p-(R_{I(slr)})^2}}$$

The DFFITS for adjusted likelihood residuals used Eq. (14)

$$DFFITS_i = |t_i|\sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (23)$$

$$t_{i(alr)} = R_{i(alr)}\sqrt{\frac{n-p-1}{n-p-(R_{i(alr)})^2}} \quad (23.1)$$

where $hii = diag(H)$ is the ith hat matrix H diagonal element for the G-PRM McCullagh and Nelder [33], $H = \widehat{W}^{\frac{1}{2}} X\left(X^T \widehat{W} X\right)^{-1} X^T \widehat{W}^{\frac{1}{2}}$. These diagonal elements can be used in diagnostics of influences, and they are called also the leverages. The leverages make an indication in order to call prescription of other diagnostic measures. In case of small-sized data, an observation is said to be influential when the value of DFFITS exceeds one (Chatterjee and Hadi, [16]). In the case of large data sets, an observation is considered influential when the ith value of DFFITS exceeds $2\sqrt{\frac{p+1}{n}}$ (Belsley et al., [15]). DFFITS fortifies the effect of the ith influential observation on the values of the fitted and estimated values. Similarly, we can replace any other form of standardized and smoothed residuals of G-PRM with the intent of locating influential observation. In calculating DFFITS we adopt the same cut-off point of using standardized and adjusted G-PRM residuals so as to compare the results with standardized and adjusted residuals that are used in customary usage.

III. Numerical Results

The Simulation Study

This section should show through a simulation that the G-PRM standardized and adjusted residuals are effective in terms of influence diagnostics. The independent variables include five powerful points. The following Monte Carlo scheme is considered to compare the performance of the G-PRM residuals with DFFITS. We employed the generation of response variable that is a Gamma Preto of algorithm of Hanum et al. [27, 28]. Data generation is done as follows: $y_i \sim G-P(\alpha,\beta,\gamma)$, where $\hat{\mu} = E(y_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3}$, $i = 1, 2, \ldots, n$ is mean function and ϕ is dispersion parameter $\phi =$ 0.04, 0.11, 0.17, 0.33, 0.67, 2, 5, 10 which is thought to have arbitrary values. For the true parameters, we choose the following arbitrary values as $\beta_0 = 0.05, \beta_1 = 0.0025,\ \beta_2 = 0.005\ and\ \beta_3 = 0.0001$ (Amin et al. [5-6]) and γ is minimum value of response variable. In this case, the design matrix X has no influential points of sample sizes n = 25, 50, 100 and 200 generated as $X_i \sim N(-1,1)$, $i = 1,2,\ldots,n;\ and\ j = 1,2,3,$ and then we make 5th, 10th, 15th, 20th, 25th, points in the X as $X_{ij} = \alpha_0 + X_{ij}$, $i = 5, 10, 15, 20\ and\ 25, and\ j = 1,2,3,$ where $\alpha_0 = \overline{X}_j$ +100. In the estimation of G-PRM the link function employed is identity link.

Table 1. Difference of fits (DFFITS) influence detection (%) using Gamma-Pareto regression model (G-PRM), residuals and diagnostic measures,

		Difference of fits with G-PRM Standardized residuals			Difference of fits with G-PRM Adjusted residuals		
Influence Points		$DFFITS_{(i)R_{spr}}$	$DFFITS_{(i)R_{sdr}}$	$DFFITS_{(i)R_{slr}}$	$DFFITS_{(i)R_{apr}}$	$DFFITS_{(i)R_{adr}}$	$DFFITS_{(i)R_{alr}}$
n			ϕ =0.04				
25	**5**	92.8	92.8	99.6	57.3	65.4	91.6
	10	84.7	84.8	97.9	80.6	80.5	85.1
	15	73.5	74.3	95.6	69.6	67.5	74.5
	20	60.5	62.6	88.1	57	55.1	62.7
	25	47.2	54.6	81.9	44.5	47.4	54.1
n			ϕ =0.04				
50	**5**	92.7	92.7	99.3	70.1	73.1	93.4
	10	88.7	88.8	97.7	85.2	85.4	88.7
	15	83.1	82.4	97.1	80.3	79.4	83.3
	20	78.3	78.3	93.8	76.2	74.7	78.5
	25	71.5	72.8	91.9	69.2	67.9	71.8
n			ϕ =0.04				
100	**5**	94.7	94.8	99.4	78.8	83.1	94.1
	10	90.8	90.9	98.8	87.9	88.6	90.5
	15	88.5	88.6	97.5	85.9	85.7	88.8
	20	83.7	83.6	96.7	81.4	80.4	83.7
	25	85	85.3	96.5	83.2	82.5	84.6
n			ϕ =0.04				
200	**5**	96.1	96.1	99.4	86.3	86.4	96.1
	10	93.4	93.3	99	91.1	90.9	93.1
	15	92	92.1	97.7	90.7	89.9	91.8
	20	90.8	91.1	98.5	89.8	89.8	90.9
	25	87.9	88	97.5	87.1	86.4	88.3
n			ϕ =0.11				
25	**5**	93.5	93.5	99.7	56.4	67.4	92.5
	10	85.6	85.3	97.7	80.2	79.8	85.8
	15	75.7	76.6	96.6	72	70.8	76.2
	20	59	61.6	88.2	54.8	54.1	60.9
	25	46.4	55.2	83.9	43.8	48.5	53.1
n			ϕ =0.11				
50	**5**	93.2	93.2	99.2	69	71.8	93.1
	10	90.1	90.1	98.2	85.8	85.5	90.5
	15	81	80.7	96.4	77.7	77.1	81.4
	20	78.5	79.8	94.7	76.5	75	78.9
	25	73.6	75.3	91.7	71.5	70	74.3

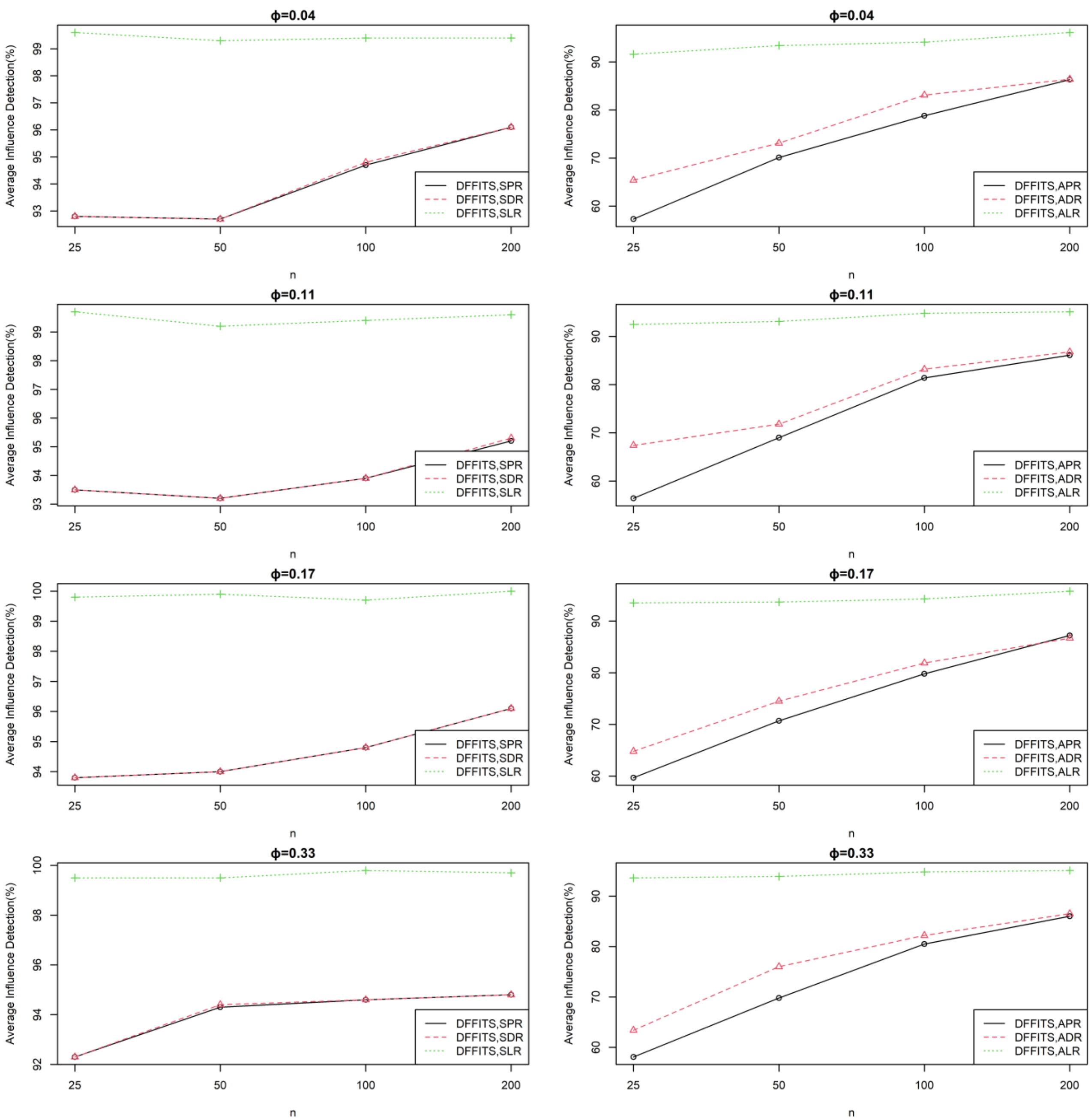
n			ϕ =0.11				
100	**5**	93.9	93.9	99.4	81.4	83.2	94.8
	10	90.6	90.4	98.3	86.8	87	90.3
	15	87.9	87.7	97.6	85.7	85.6	88.3
	20	83.9	84.3	96.4	82	81.7	84.1
	25	82.1	82.5	96.2	80.3	79.1	82.2
n			ϕ =0.11				
200	**5**	95.2	95.3	99.6	86.1	86.8	95.1
	10	92.4	92.4	98.5	90.5	90.4	92.6
	15	91.2	91.2	98	89.3	89.4	91.4
	20	91	91	97.2	89.9	89.8	90.7
	25	88	88	97.3	86.3	85.9	87.8
n			ϕ =0.17				
25	**5**	93.8	93.8	99.8	59.7	64.8	93.5
	10	82.5	82.4	97.6	75.8	76.5	82.3
	15	76.3	77.7	94.9	71.9	70.1	76.7
	20	63.1	64.6	91.3	59.4	58.3	65
	25	48.9	55.6	84.5	45.8	47.2	55.4
n			ϕ =0.17				
50	**5**	94	94	99.9	70.7	74.5	93.7
	10	87.2	87.3	98.1	81.7	82.4	87.2
	15	80.4	80.9	95.5	77.8	76.8	80.3
	20	76.7	77.4	92.6	73.5	72.1	77.1
	25	74.8	75.3	92.8	72.9	71.2	75.3
n			ϕ =0.17				
100	**5**	94.8	94.8	99.7	79.8	81.9	94.3
	10	90.9	90.9	98.4	88.3	88.2	91.1
	15	88.1	88.3	97.4	86.1	85.4	88.1
	20	83.2	83.7	95.8	81.3	80.4	82.9
	25	82.5	82.9	95.2	80.1	79.6	81.7
n			ϕ =0.17				
200	**5**	96.1	96.1	100	87.2	86.7	95.8
	10	93.5	93.4	98.7	91.7	91.6	93.2
	15	92	92	98.6	89.6	89.7	91.4
	20	90.8	90.9	97.7	89.2	88.7	90.4
	25	89.9	90.1	97.6	88.8	88.6	89.7
n			ϕ =0.33				
25	**5**	92.3	92.3	99.5	58.1	63.4	93.6
	10	84.3	84.9	97.8	79.8	80	84.8
	15	73	72.9	93.1	69.1	67.8	73.5
	20	60.4	63.9	89.8	57.1	56.9	63.2
	25	46.5	54.7	83.4	44.2	46.8	53.4
n			ϕ =0.33				
50	**5**	94.3	94.4	99.5	69.8	76	93.9
	10	89.8	89.9	98.4	85.8	86.3	89.6
	15	82.4	82.7	97.2	80.1	79.4	82.5
	20	78.6	79.4	94.5	76.6	75	78.8
	25	70.2	71.7	92.7	67.6	66.1	71.1
n			ϕ =0.33				
100	**5**	94.6	94.6	99.8	80.5	82.2	94.8
	10	91.3	91.3	98.2	88.8	88.5	91.6
	15	88.2	88	97.4	85.4	84.7	88.1
	20	85.9	85.8	97.2	83.8	83	86.1
	25	82.2	83	96.5	80.4	79.7	82.8
n			ϕ =0.33				
200	**5**	94.8	94.8	99.7	86	86.5	95.1
	10	93	92.8	98.9	90.7	91.6	93.3
	15	90.5	90.3	98.5	88.8	88.7	90.6

	20	88.6	88.7	97.2	87	86.5	88.5
	25	89.9	89.9	98	88.3	88.3	89.5
n			ϕ =0.67				
25	**5**	92.7	92.7	99.7	55.6	63.6	92
	10	83.9	83.4	96.8	79	79.2	84.6
	15	73.3	74.1	93.9	69.6	68.6	73.6
	20	62	65.4	89	56.5	58.3	65.2
	25	48.7	54	82.8	46	46.6	53.5
n			ϕ =0.67				
50	**5**	93.3	93.3	99.3	69.9	73.9	92.9
	10	86.7	86.8	98.5	82.9	82.9	87
	15	82.9	82.7	96.5	79.3	78.7	82.9
	20	78.9	79.4	96.2	75.5	74.4	79
	25	71.2	73.5	93.1	70	68.5	71.5
n			ϕ =0.67				
100	**5**	94.6	94.6	99.5	81.4	81.2	94.5
	10	92	91.9	97.4	89.3	88.9	92.5
	15	88.1	88.2	97.6	85.9	85.7	88.3
	20	85	84.9	96.9	82.8	82.1	85
	25	81.7	82.3	95.8	80.3	79.5	82.1
n			ϕ =0.67				
200	**5**	95.8	95.8	99.4	86.9	87.8	95.2
	10	93.3	93.4	98.9	91.8	91.9	93.4
	15	91.2	91.3	98.5	89.2	89.4	91
	20	90.2	90.4	98.5	88.9	88.8	90.6
	25	87.7	87.9	97.6	86.6	86.1	87.3
n			ϕ =2				
25	**5**	91.4	91.4	99.5	57.5	64.1	91
	10	82.5	82.2	97	78.9	79	82.8
	15	71.9	72.2	93.9	68.2	67.2	72.2
	20	60.7	63.3	90.9	57.8	57.2	62.9
	25	48.3	57.6	86.3	45.4	50.2	56.9
n			ϕ =2				
50	**5**	92.6	92.6	99.6	70.6	74.5	92.5
	10	88	88.1	98.2	84.2	83.9	88.2
	15	84.6	84.5	97.7	81.2	80.8	84.7
	20	77.4	77.7	93.8	75.1	74.1	77.5
	25	73.7	74.6	92.7	70.5	69.7	74.2
n			ϕ =2				
100	**5**	94.1	94.1	99.5	79.7	80.8	95
	10	89.9	90	97.7	86.7	86.8	90
	15	87.4	86.9	97.8	84.9	83.9	87.4
	20	84.2	84.5	97	82.4	81.4	84
	25	82	82.9	96.5	80.9	80	82.4
n			ϕ =2				
200	**5**	95.3	95.3	99.7	85.8	87.1	94.7
	10	93	93	99	90.4	90.1	93.1
	15	90.8	90.8	97.8	89.3	89.4	90.9
	20	89.6	89.7	98.2	88.4	88.1	89.8
	25	88.3	88.2	97.1	87.4	87.4	88.8
n			ϕ =5				
25	**5**	90.6	90.6	99.2	57.5	62	90.9
	10	85.2	85.2	97.3	79.6	79.8	85.3
	15	73.4	74.8	94	69.9	68.7	73.7
	20	61.3	64.9	89.6	58.1	57.6	64.1
	25	47.2	54.3	82.3	44	46.1	52.1
n			ϕ =5				
50	**5**	94.6	94.6	99.5	69.8	74	94.1

	10	88.2	88	98.5	83.4	83.5	88.8
	15	81.8	82.6	95.6	78.4	77.6	82
	20	77.2	78.5	94.9	74.7	74.3	77.4
	25	72.6	74	92.2	70.1	69.1	73.3
n			ϕ =5				
100	**5**	95.5	95.5	99.7	81.7	81.9	95.6
	10	90.5	90.4	98.9	88	87.7	90.8
	15	87.9	88.3	97.7	86	85.3	87.7
	20	85.4	86	96.7	84	83.1	85.5
	25	83.3	82.9	95.8	81.2	80.3	83.6
n			ϕ =5				
200	**5**	96	95.9	99.8	87.1	88.1	95.7
	10	93.7	93.7	98.6	91.1	91.3	93.5
	15	89.8	90	97.7	87.9	87.3	89.4
	20	89.4	89.3	97.3	87.6	87.6	89.3
	25	89.4	89.6	97.6	88.4	88.2	89.6
n			ϕ =10				
25	**5**	93.4	93.4	99.7	58.5	64.8	93.2
	10	82.9	83.5	97.2	79.4	78.9	83.8
	15	73.2	73.2	94	68.5	68	73.9
	20	61.8	64.5	88.9	58.3	58.1	64.8
	25	46.5	53.8	84.1	43.3	47.2	52.6
n			ϕ =10				
50	**5**	94	94.1	99.5	70.2	72.5	93
	10	88	87.8	97.9	84.8	84.6	87.9
	15	81.7	82	95.8	79.5	78.7	82.3
	20	79.3	79.5	95.6	76.1	75	79.6
	25	72.9	74.9	92.7	70.3	68.9	74
n			ϕ =10				
100	**5**	94.4	94.3	99.9	79	82.1	94.7
	10	90.2	90.1	98.2	86.8	87.2	90.4
	15	87.2	87.2	96.6	85.2	84.4	87.4
	20	88.1	88.2	97.6	85.9	84.9	87.7
	25	81.7	81.7	95.8	80	79.5	81.8
n			ϕ =10				
200	**5**	96.6	96.6	100	86.9	87	96.6
	10	93.2	93.1	98.6	90.7	90.6	93
	15	91.7	91.7	97.8	89.4	89.9	91.7
	20	89.9	89.7	98.2	88.4	87.8	90.5
	25	88.2	88	97.4	87	86.6	88.1

These G-PRM residuals are performed through the G-PD generated samples to determine the generated influential points. In order to determine whether to hold the influential point detection (in percentages) of the DFFITS in standardized and adjusted format with each G-PRM residuals in standardized and adjusted form, this simulation is repeated 10000 times. These simulations are conducted with the help of the R-software. Table 1, indicates the percentages of both residents identified by the two types of residents using the DFFITS, method. All of the G-PRM standardized residuals except the likelihood residuals detect the same influential points using the above-mentioned formulation. If ϕ is less than 0.04 then likelihood-based influence diagnostics perform better than other G-PRM standardized residuals. While DFFITS using likelihood, residuals perform as well as the G-PRM standardized residuals for $\phi \geq 0.04$ due to a higher percentage of influence detection than the other cases. Conversely, every adjusted form of G-PRM residuals performs the same in terms of identifying influential points, but it still performs worse than the standardized residuals for $\phi < 0.04$. While not outperforming the standardized G-PRM residuals, all adjusted versions of the residuals perform similarly for $\phi \geq 0.04$. Additionally, it is discovered that influence detection rate using likelihood residuals is low for $\phi \geq 0.04$ and improves with adjusted residuals as the dispersion parameter value rises.

Figures 1-5 displays the simulation results (for n = 25, 50, 100, and 200 and influential observations 5, 10, 15, 20 and 25) graphically. Because these residuals produce different results, we plot DFFITS for a graphical comparison using the standardized Pearson, deviance, and likelihood residuals, as well as the adjusted Pearson residuals For comparison's sake, we take into account one of the other forms of all the standardized residuals and adjusted residuals since they produce comparable outcomes. Moreover, note that the influence analysis of the G-PRM residuals other than the influence analysis of the G-PRM residuals that are different does not affect the increase in the sample size. Another fact which is clear is that G-PRM dispersion parameter is always less than one. Therefore, the importance of likelihood residuals in the influence diagnostics in G-PRM is more evident as compared to the other G-PRM residuals.



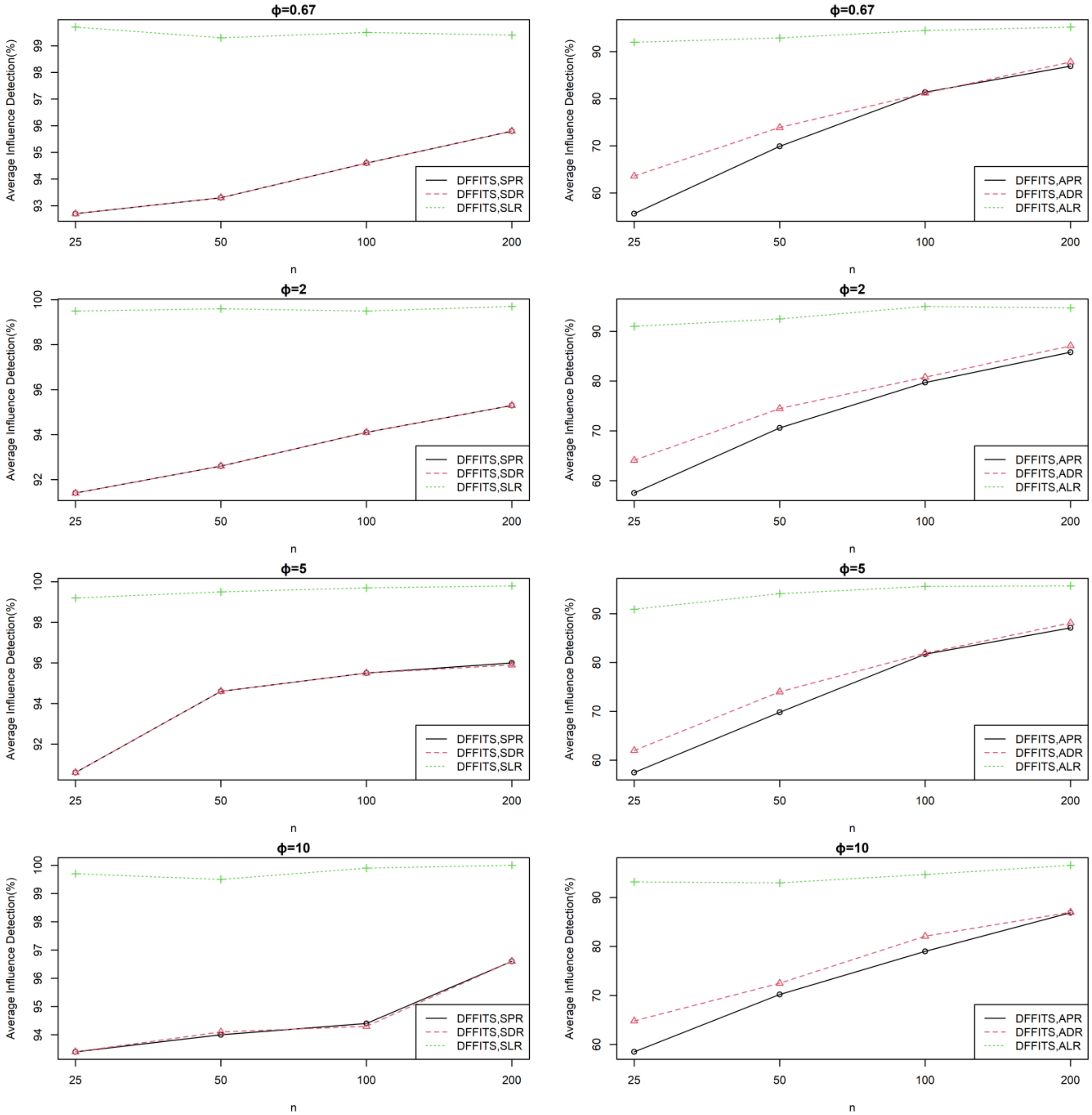
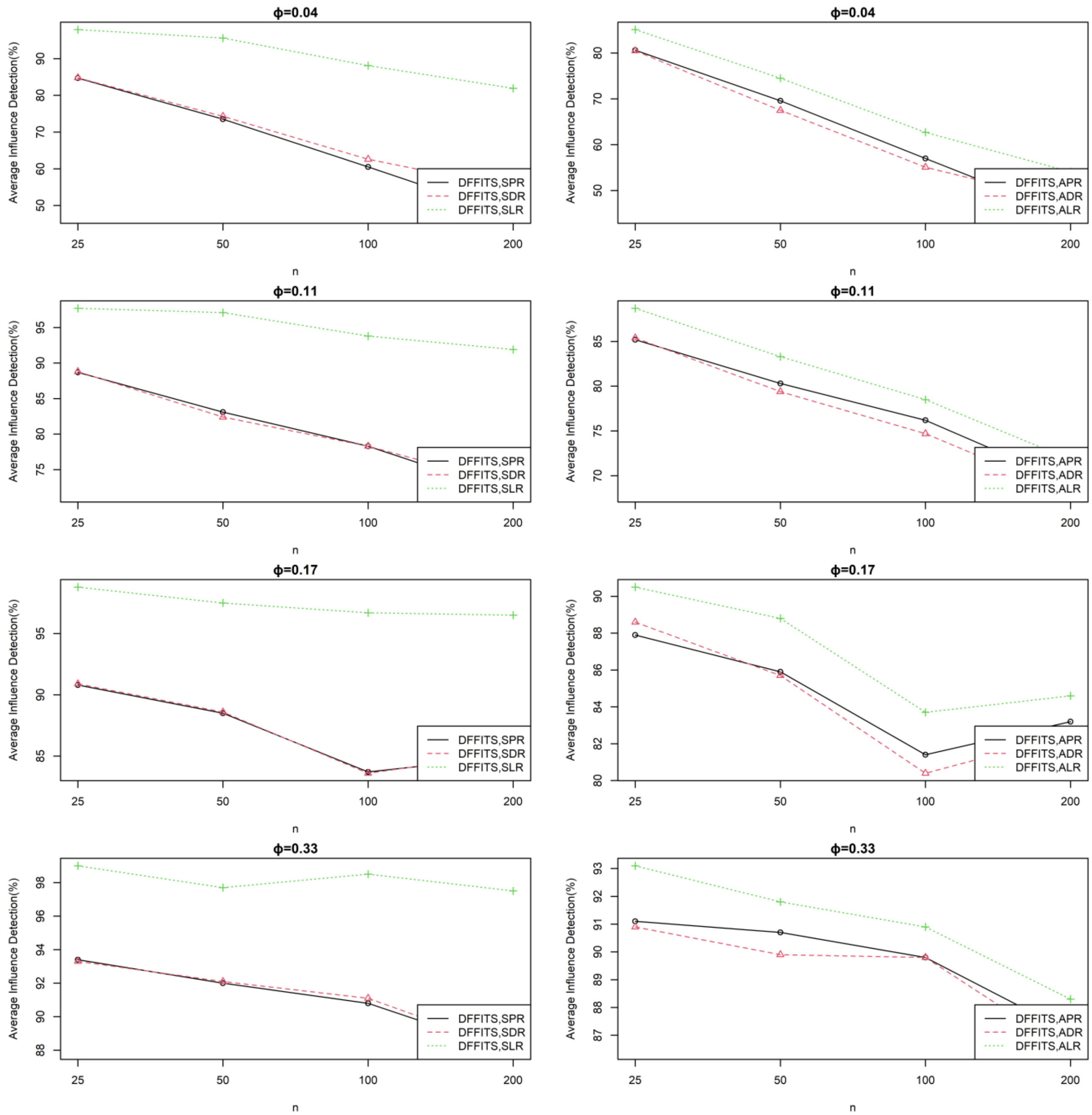


Fig. 1. Performance of the Gamma-Pareto regression model residuals using DFFITS, when influential observation is 5.


ϕ=0.04
ϕ=0.11
ϕ=0.17
ϕ=0.33
Average Influence Detection(%)
n
DFFITS,SPR
DFFITS,SDR
DFFITS,SLR
DFFITS,APR
DFFITS,ADR
DFFITS,ALR

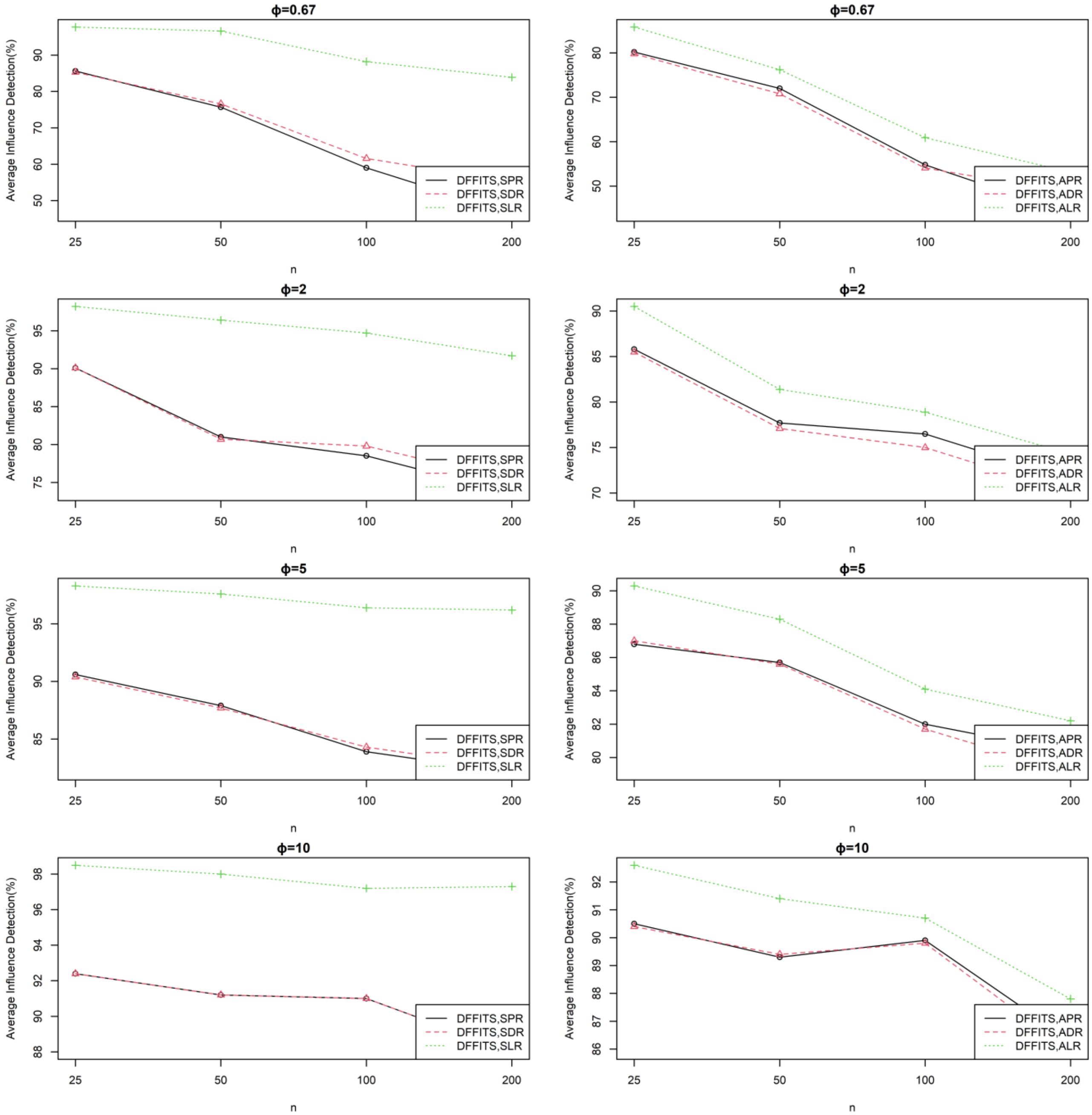
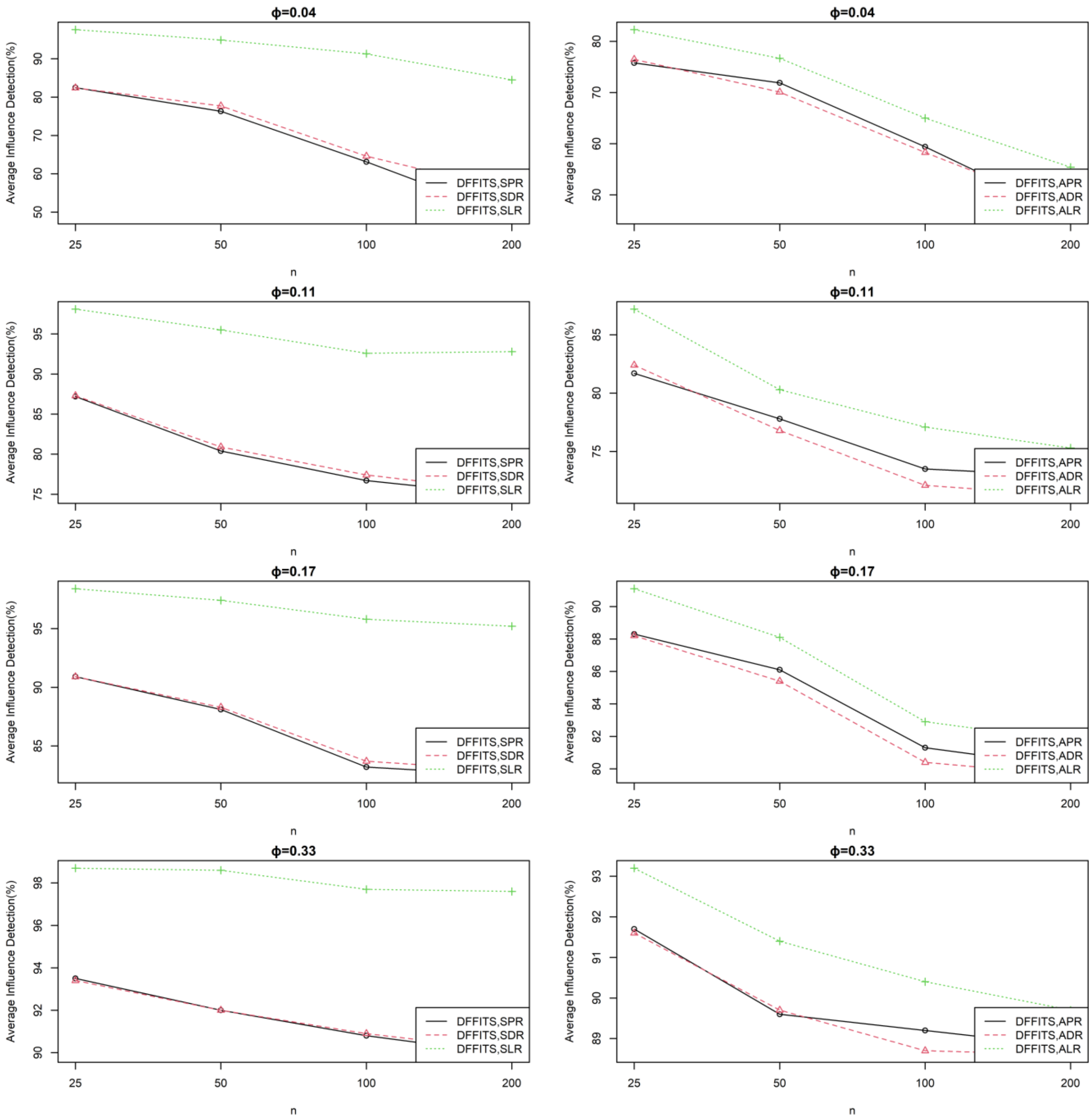


Figure 2. Performance of the Gamma-Pareto regression model residuals using DFFITS, when influential observation is 10.


ϕ=0.04
ϕ=0.11
ϕ=0.17
ϕ=0.33
Average Influence Detection(%)
n
DFFITS,SPR
DFFITS,SDR
DFFITS,SLR
DFFITS,APR
DFFITS,ADR
DFFITS,ALR

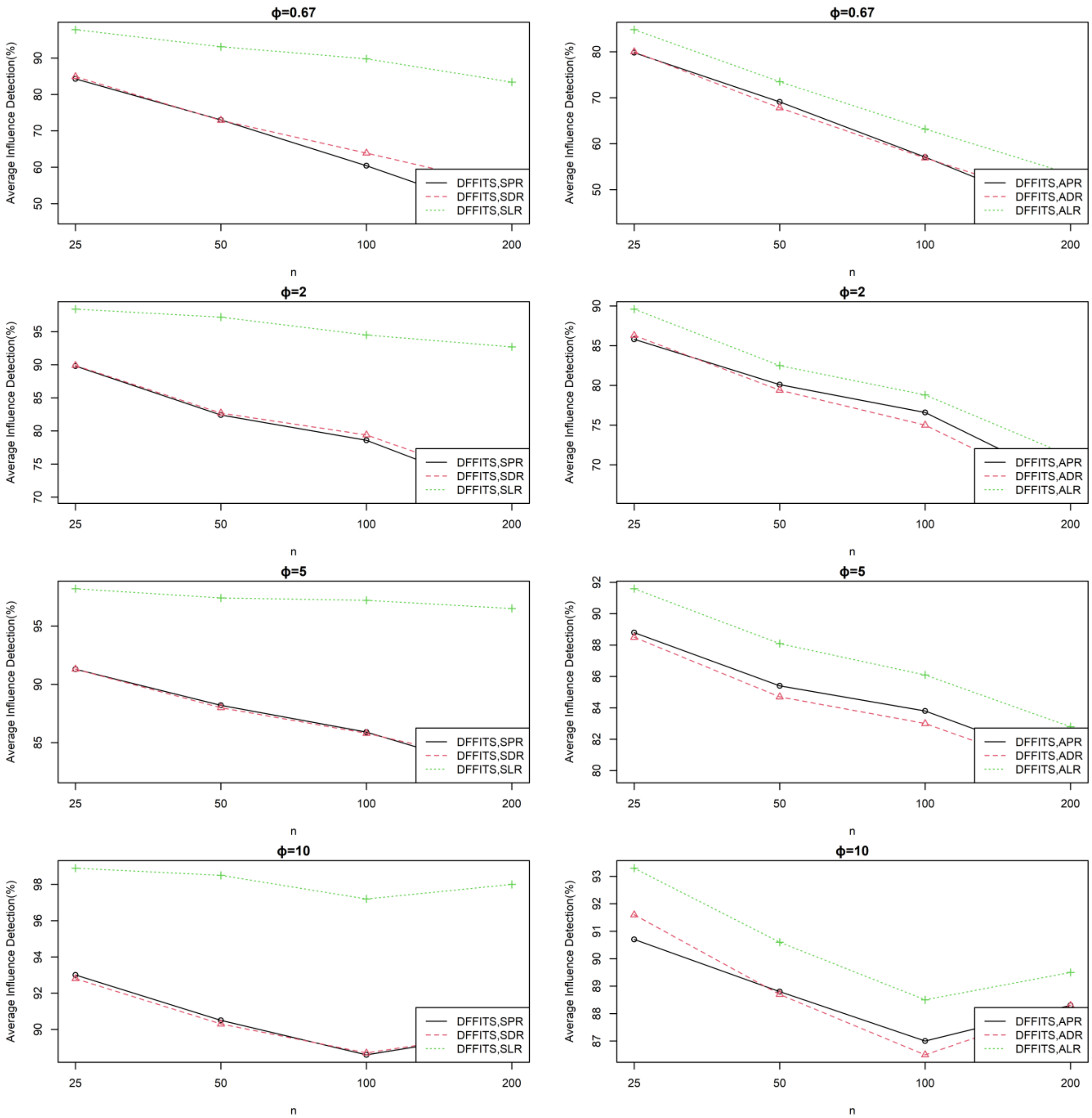
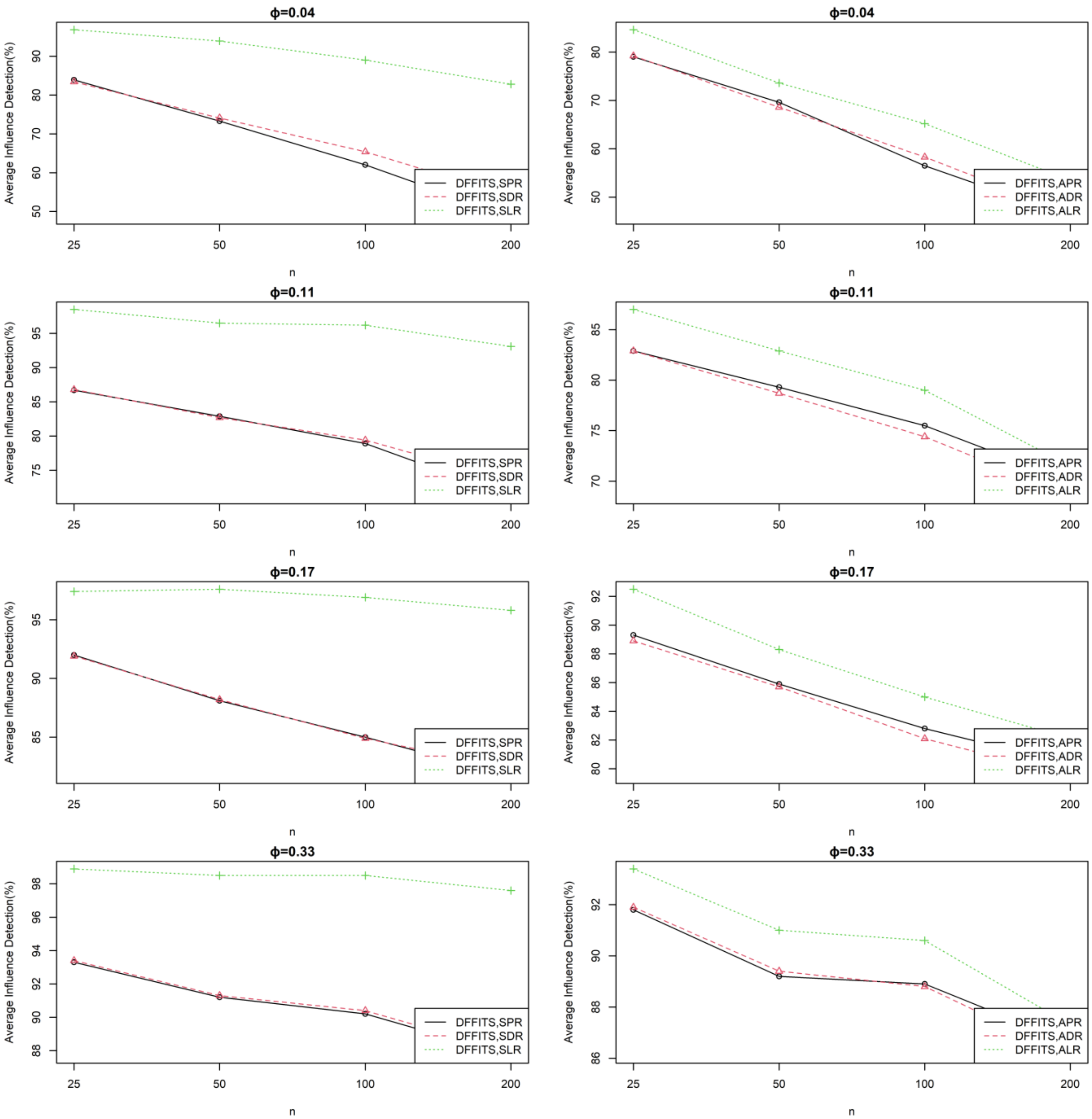


Fig. 3. Performance of the Gamma-Pareto regression model residuals using DFFITS, when influential observation is 15.


ϕ=0.04
ϕ=0.04
ϕ=0.11
ϕ=0.11
ϕ=0.17
ϕ=0.17
ϕ=0.33
ϕ=0.33
Average Influence Detection(%)
n
DFFITS,SPR
DFFITS,SDR
DFFITS,SLR
DFFITS,APR
DFFITS,ADR
DFFITS,ALR

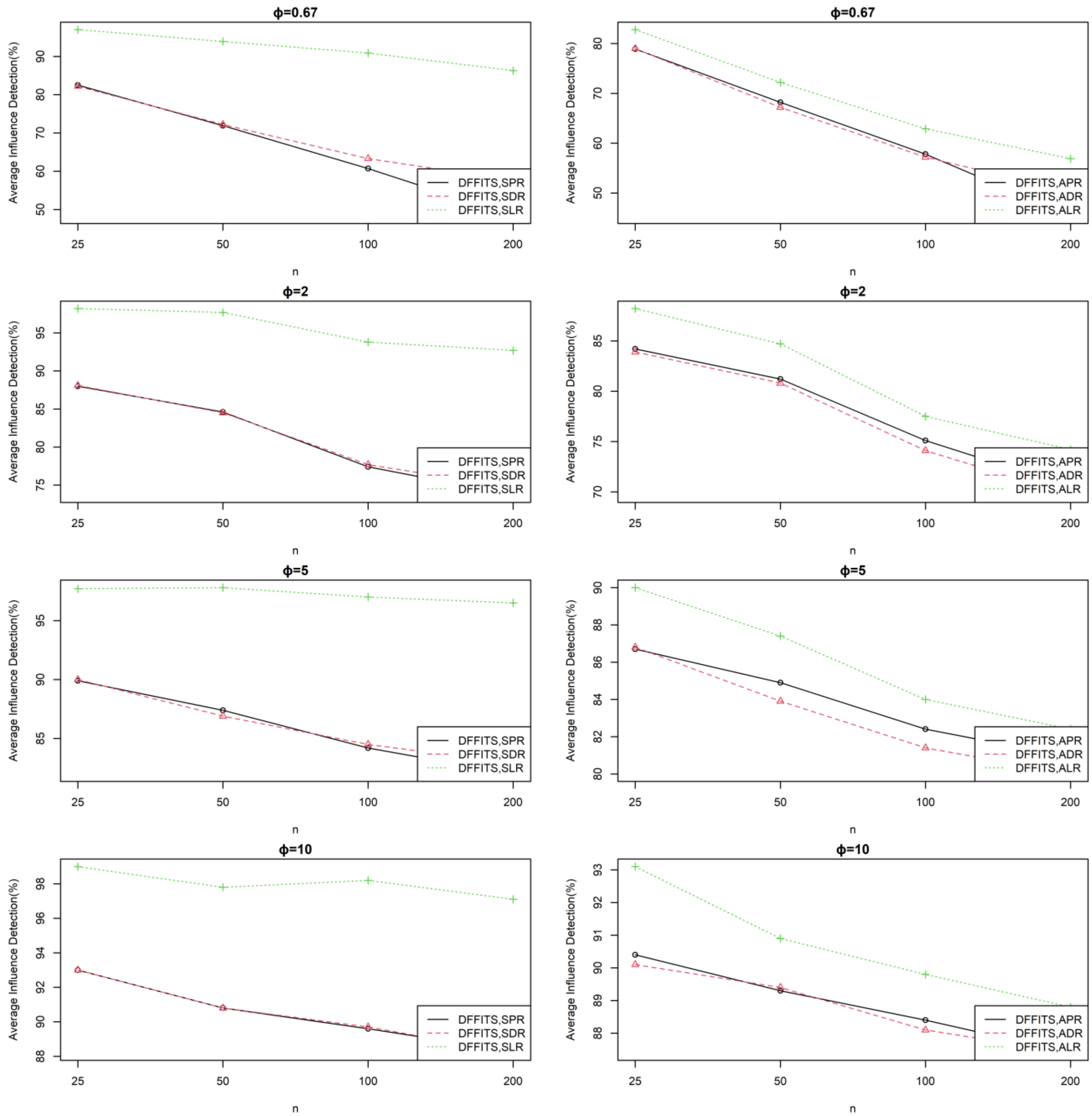
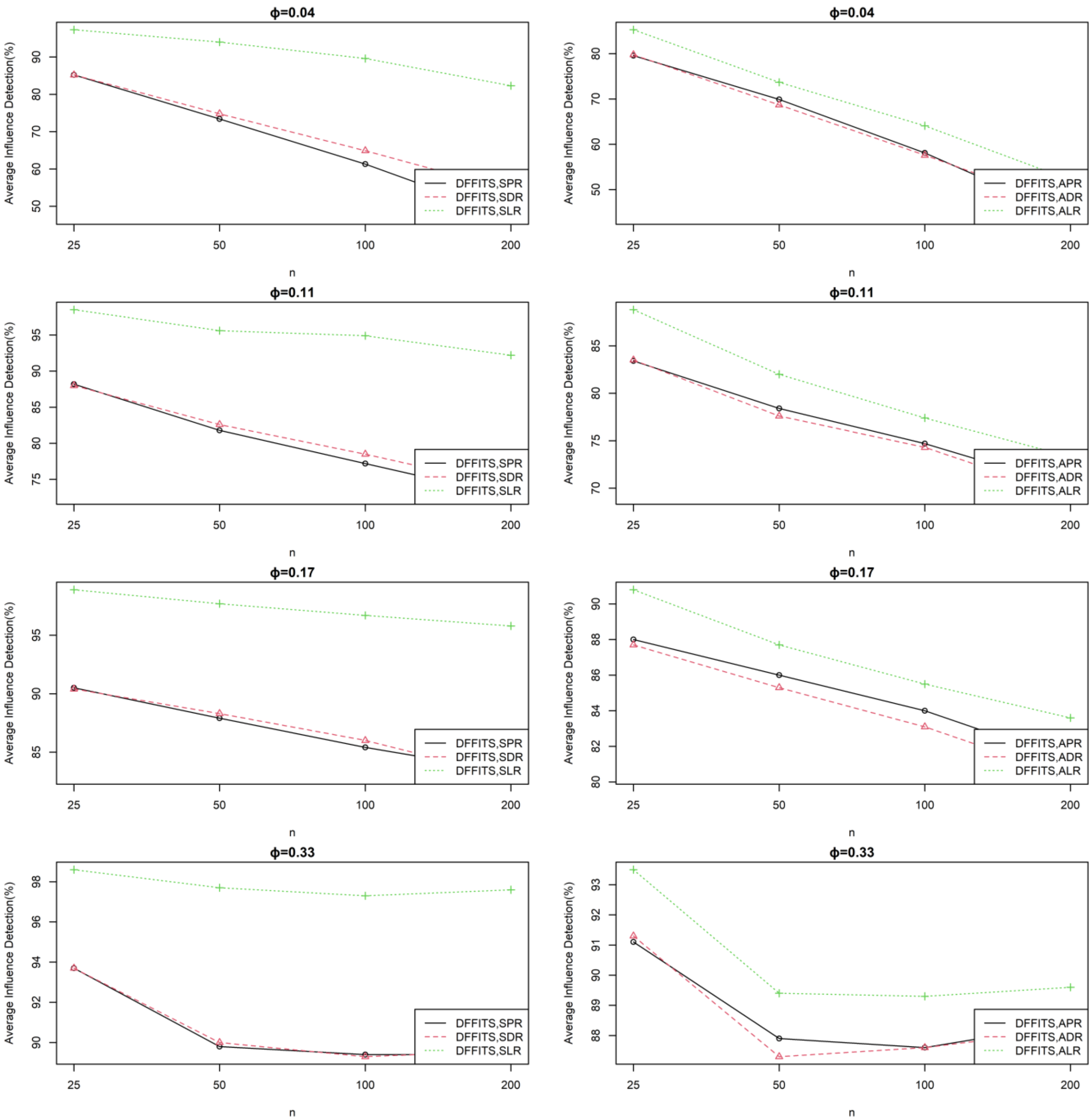


Fig. 4. Performance of the Gamma-Pareto regression model residuals using DFFITS, when influential observation is 20.


ϕ=0.04
Average Influence Detection(%)
DFFITS,SPR
DFFITS,SDR
DFFITS,SLR
n
ϕ=0.04
Average Influence Detection(%)
DFFITS,APR
DFFITS,ADR
DFFITS,ALR
n
ϕ=0.11
ϕ=0.11
ϕ=0.17
ϕ=0.17
ϕ=0.33
ϕ=0.33

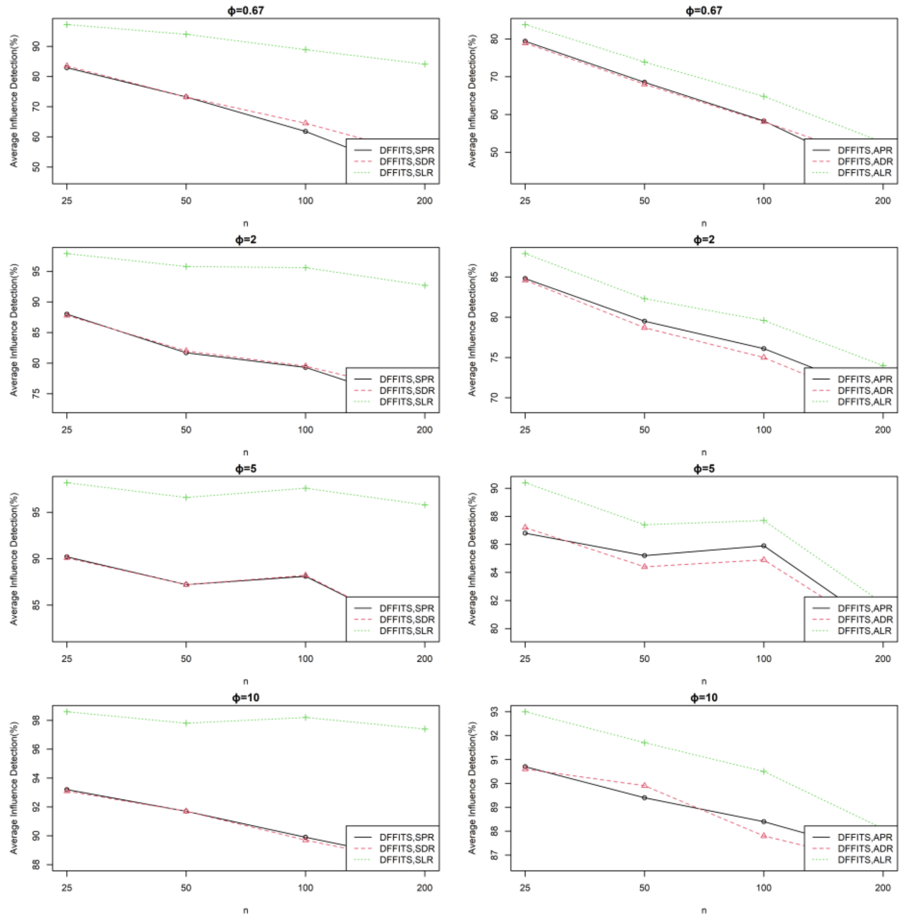


Fig. 5. Performance of the Gamma-Pareto regression model residuals using DFFITS, when influential observation is 25.

IV. Application: Ardennes Data

This time we applied a real-life application to gauge the success of the proposed approach. To do so, we have selected the ARDENNES data set and it was retrieved in Barnard et al. [14] and Amin et al. [4]. This data set's primary purpose was to identify the initial etch biopsy, i.e., starting point of an extracted layer of incisor enamel (Y) according to two explanatory variables for the data collected from 55 children. The etched depth (X1) was the first explanatory variable of those explanatory variables and was an expression in millimeters of the amount of calcium removed as a result of the etch biopsy. The second explanatory variable was the age of the children (X2) that was changed in the decimal system instead of the years and months one. However, the abnormal normal distribution does not aptly suit such a data set due to the positive skewness trend of the dependent variable. The probability distribution of the dependent variable is tested in the first step to determine the required regression model based on the analysis of Anderson Darling Cramer Mises, and Pearson chi square test (see Zhang, [41]; Evan et al., [25]; etc.). According to Table 2 we realize that data set ARDENNES suit G-PD perfectly. Thus we analyze influential points using the G-PRM and evaluate the performance of the G-PRM residuals. Based on the distribution fitting tests, we noted that G-PR model is closer to fitting this data set, the results can be seen in Table 2.

Table 2. Goodness of Fit Distribution Tests for ARDENNES Data.

Goodness of fit test (GFT)		Probability Distribution						
		Gamma	Pareto	Gamma-Pareto	Weibull	Weibull-Pareto	Normal	Normal-Pareto
Anderson-Darling (AD)	Statistic	0.3553	0.8110	0.5832	0.5883	0.4897	1.0945	3.0023
	P-value	0.4699	0.0790	**0.6145**	0.1279	0.3451	0.0066	0.6340
Cramer-von Mises (CVM)	Statistic	0.0539	0.9823	0.5941	0.0725	0.1231	0.1405	0.7231
	P-value	0.4617	0.0012	**0.7980**	0.2542	0.0023	0.0311	0.0003
Pearson chi-square (PCS)	Statistic	5.5454	11.430	15.967	7.7273	4.889	10.636	2.7350
	P-value	0.5937	0.5612	**0.8070**	0.3572	0.4432	0.1553	0.0654

Gamma-Pareto Distribution (G-PD).

Table 3. Detection of influential points (IP) with G-PRM Residuals

IP detection methods	G-PRM residuals		
	Pearson	Deviance	Likelihood
Index plots standardized residuals	5, 6, 48, 52	5, 6, 29, 48	1, 4, 5, 6, 7, 9, 10, 13, 16, 23, 28, 29, 30, 48, 52
Index plots adjusted residuals	5, 6, 48, 52	5, 6, 48	1, 4, 5, 6, 7,9, 10, 23, 28, 29, 30, 37, 39, 41, 43, 48, 52

Influential points (IP), Gamma-Pareto regression model (G-PRM)

To determine the best regression model, we first test the dependent variable's probability distribution using the Anderson–Darling (AD), Cramer–von Mises (CVM), and Pearson chi-square tests (for more information, see Zhang [41]). Table 2 concludes that ARDENNES data are quite appropriate when fitted to the G-PD. Hence, the performance of G-PRM residuals is compared and the G-PRM is used in detecting influential points. In order to determine the critical points according to DFFITS we split the G-PRM residuals into standardized and adjusted ones. Then we compare residuals of each and every data set. These above significant points detection methods have been forwarded in table 3 based on the standardized and adjusted G-PRM residuals. Table 3 shows that the 1^{th}, 4^{th}, 5^{th}, 6^{th}, 7^{th}, 9^{th}, 10^{th},13^{th}, 16^{th}, 23^{th}, 28^{th}, 29^{th}, 30^{th}, 37^{th}, 39^{th}, 41^{th}, 43^{th}, 48^{th}, 52^{th} points are influential for the given data. Our applied methods use the standardized and adjusted residuals to identify these points. Index plots are used to display these results. In Figures 6 (a-f) and 7, an index plot for DFFITS is displayed. Table 4 shows that the 5^{th} and 48^{th} points are the most influential, having a significant impact on the estimates of β_0 and β_1, respectively. Removing these points results in a greater improvement in the coefficient of determination (R^2 Efron). The 5^{th} and 48^{th} points were found to be influential in Amin et al. [5] GLM analysis of the same data. Moreover, in Table 5, the model coefficient summary is presented for full data with influential observation and after deleting, IP without influential observations. The aim is to detect influential observation using standard Statistical procedures leads to the best estimation and then forecasting of extracted layer of incisor enamel provided that etched depth and children age. The more the data is free from anomalies, valid the forecasting.

Table 4. Absolute percentage relative change in the G-PRM estimates after deleting IP and R^2_{Efron}

Influential points	Percentage relative change in the G-PRM estimates			R^2_{Efron}(0.7549)
	β_0	β_1	β_2	
1	3.54644	9.57323	6.171059	0.3470
4	24.27527	3.283381	19.2096	0.4547
5	59.84056	8.539992	47.96717	0.2568
6	4.677016	21.42682	20.67624	0.4894
7	13.85418	8.243382	5.460256	0.4757
9	5.775091	9.057438	6.615505	0.3459
10	24.6791	1.72639	23.51832	0.1843
13	13.23746	1.447899	10.24986	0.7479
16	21.20687	3.666324	18.35904	0.7477
23	15.55794	9.037685	23.87241	0.9458
28	16.60801	0.459533	17.90606	0.4959
29	40.84103	11.38833	26.68242	0.9385
30	10.54146	4.537828	16.29676	0.4757
37	0.703639	4.50067	6.240335	0.7478
39	9.448029	4.726265	4.393694	0.4799
41	10.62651	6.595186	3.75295	0.8948
43	5.024344	6.693485	1.906002	0.7947
48	72.48761	0.343848	66.85726	0.8928
52	22.83891	8.268594	28.64051	0.8948

(0.7549) represent the R^2_{Efron} of the full data
Influential points (IP), Gamma-Pareto regression model (G-PRM)

Table 5. The G-PRM summary with and without influential observations

Variables	Full data				After deleting, IP			
	Estimate	SE	Z	P-value	Estimate	SE	Z	P-value
Constant	0.000734	0.000798	0.920294	0.011669	0.000902	0.00079	1.141903	0.018827
X_1	0.0003	0.00012	2.501856	0.015539	0.000325	0.000117	2.789297	0.007409
X_2	-0.5902	0.5827	-1.09113	0.080248	-0.00012	0.58150	-1.42429	0.000452

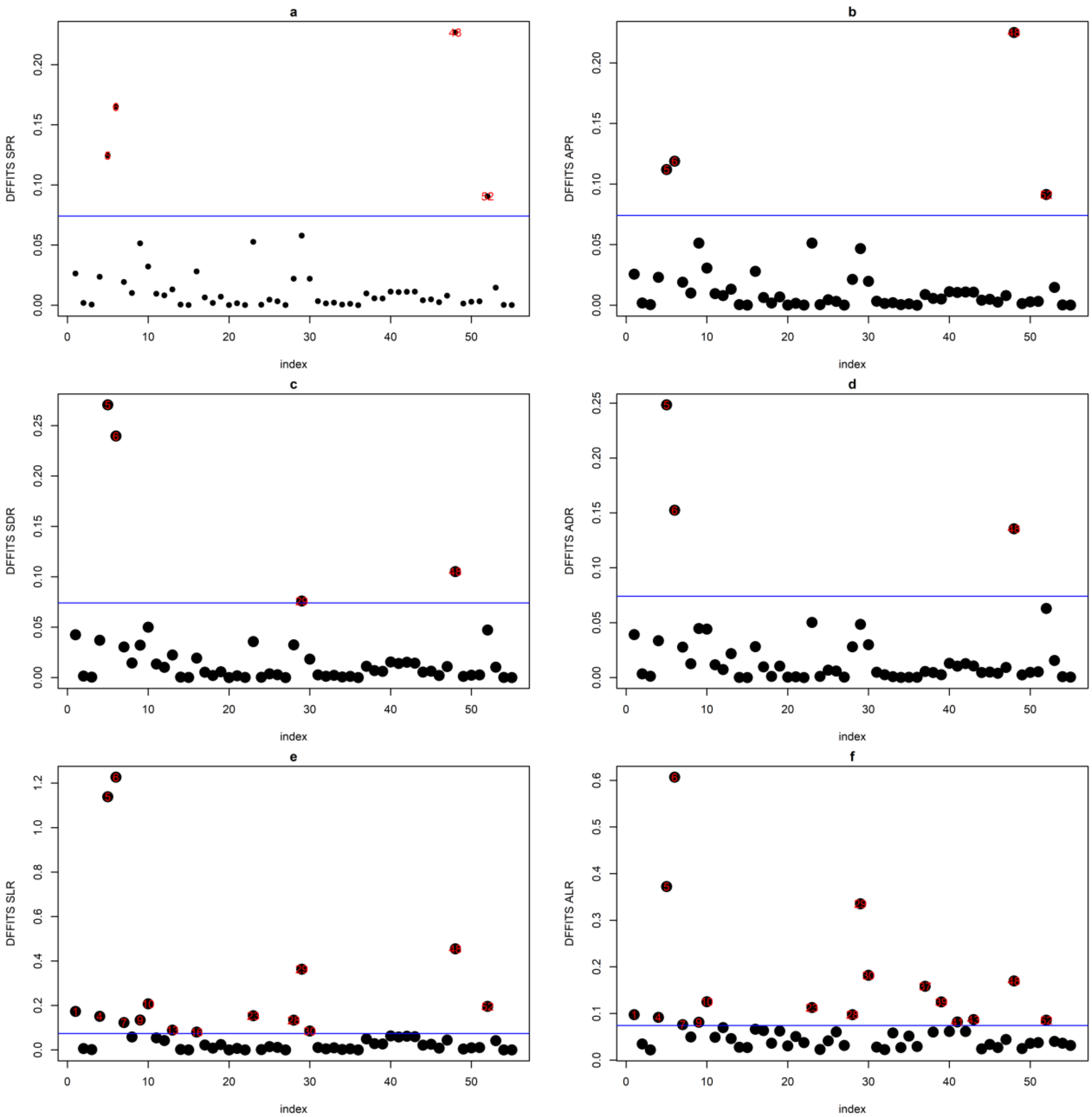


Fig. 6. Performance of the Gamma-Pareto regression model residuals using DFFITS ARDENNES Real data

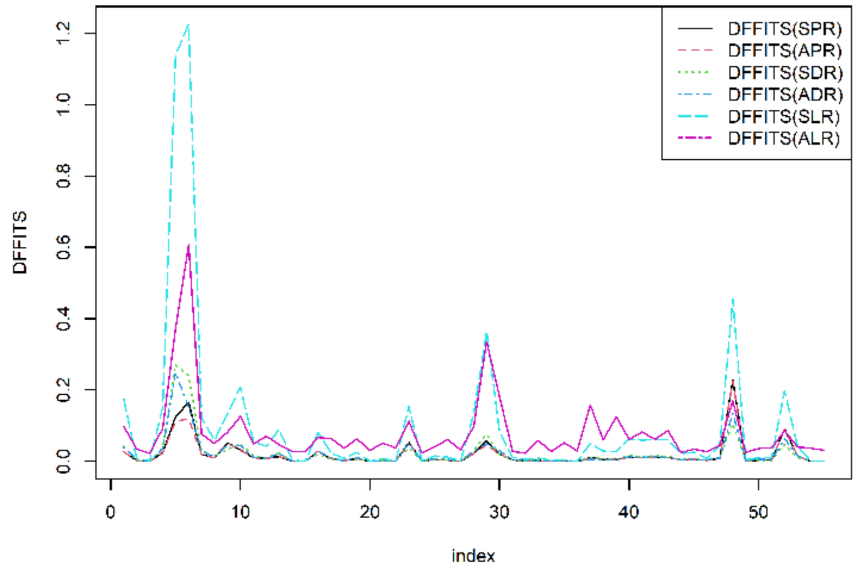


Fig. 7. Index plot of DFFITS

V. Conclusion

The influential points affect the estimates, predicted values and inferences in regression modeling. It is essential to test the response variable's distribution before modeling. The G-PRM, a particular kind of GLM, is employed if the response variable's probability distribution is the G-PD. In this paper we determine the influential points with the different types of graphical technique and with the different G-PRM residuals of DFFITS techniques. One would then compare using a real data set and a simulation study. We find that, the likelihood residuals are useful in detecting influential points compared to the other G-PRM standardized residuals when the dispersion is lower. When we are assessing their influence, the adjusted residuals of G-PRM are performing equally, and they do not outperform the likelihood residuals. The same can be said as far as all the standardized G-PRM residuals excluding the likelihood residuals carry out comparatively with the influence diagnostics in larger dispersion. In large samples with a dispersion close to one all standardized G-PRM residuals are approximately equal in terms of influence diagnostics and superior in influence diagnostics to all adjusted G-PRM residuals. Nonetheless, all the adjusted G-PRM residuals act the same when used in influence diagnostics, and they outperform (namely, they tend to be much smaller in magnitude) when compared to standardized G-PRM residuals when the value of dispersion parameter is large. As per the findings, it is also suggested that likelihood residuals could be used as opposed to other forms of residuals in the G-PRM in order to study the influence diagnostics of the G-PRM. Research recommendations in future, there are some dimensions that have not been touched yet. This paper discusses the results of influence diagnostics in the Gamma-Pareto regression model using identity link function. They may then be extrapolated to the influence diagnostics of the various GLM residuals and varied link functions of the G-PRM.

Funding. No funding resource is available for this study.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

1. Alzaatreh, A., & I. Ghosh, (2016). A study of the Gamma-Pareto (IV) distribution and its applications. Communications in Statistics-Theory and Methods, 45(3), 636-654.
2. Alzaatreh, A., F., Famoye, & C. Lee, (2012). Gamma-Pareto distribution and its applications. Journal of Modern Applied Statistical Methods, 11(1), 7.
3. Alzaghal, A. (2020). The exponentiated gamma-Pareto distribution with application to bladder cancer susceptibility. J. Stat. Appl. Probab. Lett, 7, 39-56.
4. Amin, M., S., Afzal, M. N., Akram, A. H., Muse, A. H., Tolba, & T. A. Abushal, (2022). Outlier detection in gamma regression using Pearson residuals: Simulation and an application. AIMS MATHEMATICS, 7(8), 15331-15347.
5. Amin, M., M., Amanullah, & M. Aslam, (2016). Empirical evaluation of the inverse Gaussian regression residuals for the assessment of influential points. Journal of Chemometrics, 30(7),394-404.
6. Amin, M., M., Amanullah, & G. M. Cordeiro, (2017). Influence diagnostics in the gamma regression model with adjusted deviance residuals. Communications in Statistics-Simulation and Computation, 46(9), 6959-6973.
7. Amin, M., M., Amanullah, M., Aslam, & M. Qasim, (2019). Influence diagnostics in gamma ridge regression model. *Journal of Statistical Computation and Simulation*, *89*(3), 536-556.

8. Amin, M., M., Faisal, & M. N. Akram, (2021). Influence diagnostics in the inverse Gaussian ridge regression model: Applications in chemometrics. *Journal of Chemometrics*, *35*(6), e3342.
9. Amin, M., M., Qasim, M., Amanullah, & S. Afzal, (2020). Performance of some ridge estimators for the gamma regression model. *Statistical papers*, *61*, 997-1026.
10. Anholeto, T., M. C., Sandoval, & D. A. Botter, (2014). Adjusted Pearson residuals in beta regression models. *Journal of Statistical Computation and Simulation*, *84*(5), 999-1014.
11. Ashour, S. K., D. R., Said, & M. A. A Fahim, (2014). The Log-Gamma-Pareto Distribution.
12. Atkinson, A. C. (1981). Two graphical displays for outlying and influential observations in regression. Biometrika, 68(1), 13-20.
13. Atkinson, A. C. (1985). Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis.
14. Barnard, T. E., K. S., Booksh, R. G., Brereton, D. H., Coomans, S. N., Deming, Y., Hayashi, ... & Z. Wang, (1995). Chemometrics in environmental chemistry-statistical methods (Vol. 2). Springer.
15. Belsley,D. A., E., Kuh, R. Welsch, (1980). Regression Diagnostics: Identifying InfluentialData and Sources of Collinearity. New York:Wiley.
16. Chatterjee S, Hadi AS (1988). Sensitivity Analysis in Linear Regression. Wiley: New York.
17. Cook, R. D. (1977). Detection of influential observation in linear regression. Technometrics, 19(1), 15-18.
18. Cook, R. D. (1986). Assessment of local influence. Journal of the Royal Statistical Society Series B: Statistical Methodology, 48(2), 133-155.
19. Cook, R. D., & S. Weisberg, (1982). Residuals and influence in regression.
20. Cordeiro, G. M. (2004). On Pearson's residuals in generalized linear models. Statistics & probability letters, 66(3), 213-219.
21. Cox, D. R., & E. J. Snell, (1968). A general definition of residuals. Journal of the Royal Statistical Society: Series B (Methodological), 30(2), 248-265.
22. Dar, A. A., A., Ahmed, & J. A. Reshi, (2020). weighted gamma-pareto distribution and its application. Pakistan Journal of Statistics, 36(4).
23. De Andrade, T. A., L. M. Z., Fernandez, F. G., Silva, & G. M. Cordeiro, (2017). The gamma generalized pareto distribution with applications in survival analysis.
24. Dobson, A. J., & A. G. Barnett, (2002). An introduction to generalized linear models. CRC press.
25. Evans, D. L., J. H., Drew, & L. M. Leemis, (2008). The distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling test statistics for exponential populations with estimated parameters. *Computational probability applications*, 165-190.
26. Hanum, H. (2015). Modeling Extreme Rainfall with Gamma. Applied Mathematical Sciences, 9(121), 6029-6039.
27. Hanum, H., A. H., Wigena, A., Djuraidah, & I. W. Mangku, (2016). Developing generalized linear model of Gamma-Pareto Distribution. Far east Mathematical Sciences Journal. Link finction
28. Hanum, H., A. H., Wigena, A., Djuraidah, & I. W. Mangku, (2016). Modeling Gamma-Pareto distributed data using GLM Gamma. Glob. J. Pure Appl. Math, 12(4), 3569-3575.
29. Hanum, H., A. H., Wigena, A., Djuraidah, & I. W. Mangku, (2017). The application of modeling gamma-Pareto distributed data using GLM gamma in estimation of monthly rainfall with TRMM data. Sriwijaya Journal of Environment, 2(2), 40-45.
30. Hardin, J. W., J. M. Hilbe, (2012). Generalized Linear Models and Extensions. 3rd ed. USA: Stata Press Publication.
31. Lawless, J. F. (2003). *Statistical Models andMethods for Life Time Data*. NewYork: Wiley.
32. Lee AH (1986). Assessing partial influence in generalized linear models. Biometrics; 44: 71–77.
33. McCullagh, P., J.A. Nelder, (1989). Generalized Linear Models. London: Chapman & Hall.
34. Pierce, D. A., & D. W. Schafer, (1986). Residuals in generalized linear models. Journal of the American Statistical Association, 81(396), 977-986.
35. Pregibon, D. (1981). Logistic regression diagnostics. The annals of statistics, 9(4), 705-724.
36. Preisser, J. S., & B. F. Qaqish, (1996). Deletion diagnostics for generalised estimating Equations. Biometrika, 83(3), 551-562.
37. Simas, A. B., & G. M. Cordeiro, (2009). Adjusted Pearson residuals in exponential family nonlinear models. Journal of Statistical Computation and Simulation, 79(4), 411-425.
38. Thomas, W., & R. D. Cook, (1989). Assessing influence on regression coefficients in generalized linear models. Biometrika, 76(4), 741-749.
39. Vanegas, L. H., L. M., Rondón, & G. M. Cordeiro, (2013). Diagnostic tools in generalized Weibull linear regression models. *Journal of Statistical Computation and Simulation*, *83*(12), 2315-2338.
40. Williams, D. (1987). Generalized linear model diagnostics using the deviance and single case deletions. Journal of the Royal Statistical Society: Series C (Applied Statistics), 36(2), 181-191.
41. Zhang, J. (2001). Powerful goodness-of-fit and multi-sample tests. Toronto: York University.
42. Zheng, L., K., Ismail, & X. Meng, (2014). Shifted Gamma-Generalized Pareto Distribution model to map the safety continuum and estimate crashes. Safety Science, 64, 155-162.
43. Amin, M., A., Noor, & T. Mahmood, (2024). Beta regression residuals-based control charts with different link functions: an application to the thermal power plants data. *International Journal of Data Science and Analytics*, 1-13.
44. Aslam, M. Z., M., Amin, T., Mahmood, & M. Nauman Akram, (2024). Shewhart ridge profiling for the Gamma response model. *Journal of Statistical Computation and Simulation*, *94*(8), 1715-1734.