

Modelling Fuel Consumption of Cars by Method of Cauchit Quantile Regression

Besta Okey Onyegbuchulem¹, Chialuka Adline Onyegbuchulem², Obiageri Ogwo³,
Emmanuel Oliwe⁴

1,3 & 4. Department of Mathematics / Statistics, Dr. B.U. Nzeribe (State) Polytechnic Awo-
mamma, Imo State.

2. Department Mathematics, Alvan Ikoku University of Education Owerri, Imo State

Corresponding Author's Email Address: bokey@imopoly.net

ORCID ID: 0000-0002-3507-3109

doi: <https://doi.org/10.37745/ejsp.2013/vol14n14054>

Published October 03, 2026

Citation: Onyegbuchulem B.O., Onyegbuchulem C.A., Ogwo O., Oliwe E. (2026) Modelling Fuel Consumption of Cars by Method of Cauchit Quantile Regression, *European Journal of Statistics and Probability*, 14 (1) 40-54

Abstract: Fuel consumption is a major operational cost for vehicle owners and transport operators and is closely associated with energy efficiency and environmental sustainability. Conventional mean-based methods such as Ordinary Least Squares (OLS) regression may be inadequate for fuel consumption data characterized by skewness, heavy tails, and outliers. This study therefore applies Cauchit Quantile Regression (CQReg), a robust modelling technique that combines the distributional flexibility of quantile regression with the heavy-tailed properties of the Cauchy distribution, to model car fuel consumption. The study utilized data from 91 cars, with miles per gallon (MPG) as the response variable and car weight, length, wheelbase, width, engine size, and horsepower as explanatory variables. The CQReg model was fitted at the 0.05, 0.25, 0.50, 0.75, and 0.95 quantiles. Bootstrap resampling with 200 replications was employed to estimate standard errors, confidence intervals, and p-values. Model adequacy was assessed using pseudo- R^2 , Akaike Information Criterion (AIC), and residual diagnostics. The results revealed that car weight and length were the most important predictors of fuel consumption across the middle quantiles. At the 0.25 and 0.50 quantiles, both weight and length had statistically significant effects, while at the 0.75 quantile, weight and length also remained significant. However, none of the predictors was statistically significant at the 0.05 and 0.95 quantiles, suggesting greater variability at the extremes of fuel consumption. The CQReg models produced relatively low AIC values and small prediction errors, indicating satisfactory model performance. The study concludes that CQReg provides a robust and informative approach for modelling fuel consumption and is recommended for analyzing vehicle fuel consumption data containing outliers and heavy-tailed characteristics.

Keywords: Fuel consumption, Cauchit Quantile Regression, Cauchy distribution, Outliers, Heavy-tailed data, Bootstrap regression, Model selection, Fuel efficiency.

INTRODUCTION

Fuel consumption is one of the critical operational costs for vehicle owners, transport companies, and governments. Considering the rising cost of fuel prices, environmental challenges, and the demands for efficient energy generation, proper modelling and prediction of fuel consumption has become important. Average-based models like the Ordinary Least Squares (OLS) regression were mostly applied in this situation but because fuel consumption data are typically characterized by skewness, heavy tails, and outliers due to variations in driving conditions, vehicle maintenance, driver behavior, road terrain, and traffic congestion, (Onyegbuchulem et al.,2026).

To address these shortcomings, a robust statistical approach is needed. Quantile Regression (QReg) Model given as:

$$y = \beta_0^{(\tau)} + \beta_1^{(\tau)} x_i \quad (1)$$

provides a better alternative. This is due to the fact that not only that it can model cars' fuel consumption datasets characterized by skewness, heavy tails and outliers, it can equally model different quantiles of the conditional distribution of cars' fuel consumption dataset rather than just the mean. Furthermore, when the error distribution exhibits extreme heavy tails, the Cauchit distribution becomes relevant. The Cauchit Quantile Regression (CQRreg) method combines the robustness of quantile regression with the heavy-tail properties of the Cauchy distribution, making it suitable for modelling fuel consumption data that are prone to outliers.

Fuel consumption of cars depends on multiple factors. These can be broadly classified as:

1. Vehicle characteristics: engine size, weight, age, transmission type, fuel type, aerodynamics
2. Driver behavior: acceleration, braking, speed, idling time
3. Environmental/road factors: road grade, traffic density, weather, road surface
4. Operational factors: maintenance status, tire pressure, load

But for this study Weight of the car, length of the car, wheel base of the car, width of the car, engine size of the car and the car's horse power will be considered.

Early models such as the ordinary least square (OLS) model used deterministic equations and average-based regression which is sensitive to outliers and violation of one or more of its assumptions may result in a situation where the olsm estimator can no longer be said to be Best, linear, unbiased estimator (BLUE), (Onyegbuchulem et al.,2026a); (Onyegbuchulem et al.,2026b). In real-world fuel data, a few trips with aggressive driving or mechanical faults can produce extreme fuel consumption values that distort the mean model and this made the researchers to adopt this model. QReg Koenker and Bassett (1978), allows estimation of the relationship between predictors and any quantile $\tau \in (0,1)$ of the response variable. It is less sensitive to outliers and does not require assumptions about the error distribution.

As a predictive analysis, CQReg describes data and explains the relationship between one dependent variable and two or more independent variables. In CQReg analysis, the dependent variable is ordinal (statistically it is polytomous ordinal) and the independent variables are ordinal or continuous-level which means it can also be ratio or interval. CQReg model is applied for two major reasons which include causal analysis and forecasting an effect. Onyegbuchulem, (et al.,2026) stated that CQReg is used when outliers are present in the data. The idea behind the proposed CQReg is to equate the quantile function of cauchit distribution to the QReg model and solve simultaneously to obtain $F_x(x)$ from the new function which automatically becomes the quantile function of the proposed cauchit quantile regression model.

Some of the advantages of CQReg for fuel consumption modelling include:

- i. Robustness to outliers: Extreme fuel consumption values due to traffic jams, faulty sensors, or aggressive driving do not unduly influence the estimates.
- ii. Heavy-tail accommodation: Fuel data often have more extreme values than a normal distribution can capture. Cauchy errors fit better.
- iii. Full distributional insight: Instead of one mean model, CQReg can model the 5th, 25th, 50th, 75th and 95th quantiles, showing how factors affect low, average, and high fuel consumers differently.
- iv. No moment assumptions: Since Cauchy has no finite mean/variance, CQReg does not require these assumptions that OLS needs.

Despite the importance of fuel efficiency, most existing studies on vehicle fuel consumption in Nigeria and other developing countries still rely on mean regression models. These models often produce biased estimates because fuel consumption data are non-normal and contain outliers.

There is limited research applying CQReg to model fuel consumption of cars, especially using data from local contexts such as Imo State or Nigerian road conditions. This gap limits policymakers, transport planners, and vehicle owners from understanding how different factors affect not just average fuel use, but also low and high consumption levels.

this study is aimed to model fuel consumption of cars using the method of CQReg. Specifically, it will Objectively identify the key factors that influence fuel consumption of cars, fit a CQReg model at different quantiles of fuel consumption, determine which factors have significant effects on low, median, and high fuel consumption levels and lastly it will provide recommendations for improving fuel efficiency based on the model results. this study will helps in identifying driving and vehicle factors that lead to high fuel consumption, enabling cost savings; Provides data for policies on emissions, fuel standards, and public transport planning and Contributes to the literature on robust statistical modelling of transportation data in Nigeria.

This study will focus on modelling fuel consumption of petrol and diesel cars. Data will be collected from car owners, transport operators, or secondary sources within a defined area, e.g. Owerri metropolis, Imo State. The study will use variables such as Weight of the car, length of the car, wheel base of the car, width of the car, engine size of the car and the car's horse power. The analysis will be limited to the application of CQReg.

METHODOLOGY

Cauchy distribution

The Cauchy distribution with a common notation as $X \sim \text{Cauchy}(\theta, \lambda)$ where θ is the location parameter and $(\lambda > 0)$ is the scale parameter is coined after Augustin Cauchy, and it belongs to the family of stable distributions that is closed under the formation of sums of independent random variables, its expected value, the variance, skewness and kurtosis do not exist but its median is given as θ (Alzaatreh et. al; 2016). Cauchy distribution has been applied in various fields like mechanical and electrical theory, physical anthropology, measurement problems, risk and financial analysis. It was used by Stigler (1998) to derive an explicit expression for $P(Z_1 \leq 0, Z_2 \leq 0)$, where $(Z_1, Z_2)'$ follows the standard bivariate normal distribution, Johnson et al. (1994).also applied it to model the points of effect of a fixed straight line of particles released from a point source. It is called a Lorentzian distribution in physics, where it is defined as the energy of an unstable state distribution in quantum mechanics. Generally, the probability density function (PDF) of the Cauchy distribution is defined as:

$$(\pi\alpha)^{-1} \left[1 + \left\{ \frac{y-\theta}{\alpha} \right\}^2 \right]^{-1}, \quad -\infty < y < \infty \quad (2)$$

(Alshawarbeh et al.,2012)

The standard probability density function (pdf) of a Cauchy distribution is therefore:

$$\begin{aligned} f(y) &= \frac{\alpha/\pi}{\alpha^2 + y^2}, \text{ for } \theta = 0 \\ &= \pi^{-1} (1 + y^2)^{-1}, \text{ for } \alpha = 1, -\infty < y < \infty \end{aligned} \quad (3)$$

(Alshawarbeh et al.,2012)

Using the standard probability density function (PDF) of the Cauchy distribution, let

$$u = \frac{\alpha/\pi}{\alpha^2 + y^2}, \quad -\infty < y < \infty \quad (4)$$

the Cumulative density function of the Cauchy distribution is derived as:

$$\begin{aligned}
 F(u) &= \int_{-\infty}^y f(u) du & (5) \\
 &= \int_{-\infty}^y \frac{\frac{\alpha}{\pi}}{\alpha^2 + u^2} du \\
 &= \frac{\alpha}{\pi} \int_{-\infty}^y \frac{1}{\alpha^2 + u^2} du \\
 &= \frac{\alpha}{\pi} \left\{ \frac{1}{\alpha} \tan^{-1} \left(\frac{u}{\alpha} \right) \right\} \Big|_{-\infty}^y \\
 &= \frac{\alpha}{\pi} \times \frac{1}{\alpha} \left\{ \tan^{-1} \frac{y}{\alpha} - \tan^{-1} \left(\frac{-\infty}{\alpha} \right) \right\} \\
 &= \frac{1}{\pi} \left\{ \tan^{-1} \frac{y}{\alpha} - \tan^{-1} (-\infty) \right\} \\
 &= \frac{1}{\pi} \left\{ \frac{\pi}{2} + \tan^{-1} \left(\frac{y}{\alpha} \right) \right\}
 \end{aligned}$$

$$F(y) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{y}{\alpha} \right) \quad (6)$$

$$F(\infty) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(\infty) = 1 \quad (7)$$

Let $\alpha = 1$ in equation (6), the standard cumulative density function (cdf) of the cauchy distribution becomes

$$F(y) = \frac{1}{2} + \pi^{-1} \tan^{-1}(y), -\infty < y < \infty \quad (8)$$

(Alshawarbeh et al., 2012)

Alshawarbeh et al (2012), stated that if $X \sim N(0, 1)$ and $Y \sim N(0, 1)$ are independent random variable, then $\frac{X}{Y} \sim \text{Cauchy}(0, 1)$, random variables generated with this.

Cauchit quantile regression

CQReg is from the family of Cauchy distribution. The cdf inverse (F^{-1}) function is therefore derived theoretically as:

$$\frac{1}{2} + \pi^{-1} \tan^{-1} h[Q_y(\tau)] = Q_y(\tau) \quad (9)$$

$$\pi^{-1} \tan^{-1} h[Q_y(\tau)] = Q_y(\tau) - \frac{1}{2}$$

$$\tan^{-1} h[Q_y(\tau)] = \pi \left\{ Q_y(\tau) - \frac{1}{2} \right\}$$

$$h[Q_y(\tau)] = \tan \left[\pi \left\{ Q_y(\tau) - \frac{1}{2} \right\} \right] \quad (\text{Onyegbuchulem et al., 2026}) \quad (10)$$

Using the cdf inverse (F^{-1}) of the Cauchy function of equation (10), the CQReg model is therefore given as shown in equation (13), where $Q_y(\tau)$ is the cdf inverse (F^{-1}) of the distribution to be simulated and eventually the response variable. The next step is to equate, equation (10) to equation (1) and solve for $Q_y(\tau)$ simultaneously as shown in equation (11) to form a CQReg model as shown in equations (13).

$$h[Q_y(\tau)] = (\beta_0^{(\tau)} + \beta_1^{(\tau)} x_i) \quad (11)$$

$$\tan \left\{ \pi \left(Q_y(\tau) - \frac{1}{2} \right) \right\} = (\beta_0^{(\tau)} + \beta_1^{(\tau)} x_i)$$

$$\left\{ \pi \left(Q_y(\tau) - \frac{1}{2} \right) \right\} = \tan^{-1} (\beta_0^{(\tau)} + \beta_1^{(\tau)} x_i) \quad (12)$$

$$\begin{aligned} Q_y(\tau) - \frac{1}{2} &= \frac{1}{\pi} \left\{ \tan^{-1} (\beta_0^{(\tau)} + \beta_1^{(\tau)} x_i) \right\} \\ Q_y(\tau) &= \frac{1}{\pi} \left\{ \tan^{-1} (\beta_0^{(\tau)} + \beta_1^{(\tau)} x_i) \right\} + \frac{1}{2}, \quad \left. \begin{aligned} -\infty < Q_y(\tau) < \infty, \quad -0 < x_i < \infty, \quad 0 < \tau < 1 \end{aligned} \right\} \quad (13) \end{aligned}$$

$Q_y(\tau)$ = the response variable and the cdf inverse (F^{-1}) of the distribution to be estimated

x_i = the covariates to be simulated

β_0 = the intercept parameter

β_i = the unknown parameters

τ = specified quantiles of the model. This research examines the following quantiles: 0.05, 0.25, 0.5, 0.75, 0.95

Properties of Cauchit Quantile Regression

The properties of CQReg model include all the properties proposed by Koenker and Bassett, (1978). which are properties a to c and these include:

- a. Equivariance:** Let the quantile regression estimate for a given $\tau \in (0, 1)$ and observations (y, X) be denoted by $\hat{\beta}(\tau; y, X)$ then for any $p \times p$ nonsingular matrix A , $\gamma \in \mathbb{R}^p$ and $a > 0$ holds:
- i. $\hat{\beta}(\tau; ay, X) = a\hat{\beta}(\tau; y, X)$
 - ii. $\hat{\beta}(\tau; -ay, X) = -a\hat{\beta}(1-\tau; y, X)$
 - iii. $\hat{\beta}(\tau; y + X\gamma, X) = \hat{\beta}(\tau; y, X) + \gamma$
 - iv. $\hat{\beta}(\tau; y, XA) = A^{-1}\hat{\beta}(\tau; y, X)$
- b. Invariance to monotonic transformation:** Another basic property of quantile regression is monotone equivariance property. Let $f\{\hat{\beta}(\tau; y, X)\}$ be a nondecreasing function on $\mathbb{R}^p$ then it follows closely from the quantile function definition that for any random variable

$$h[Q_\tau(y/X)] = Q_\tau[h(y/X)]$$

This implies that if a monotone transformation h is applied (for an instance, the logarithmic function or the exponential function) to a random variable, the quantiles can be obtained by applying the same transformation to the quantile function (Hao L. & Naiman D.Q; 2007).

- c. Robbostness to Outliers:** This property means that when given a sample data $x_1, x_2, \dots, x_n$ with a sample median (md) The sample can be modified by changing one or more data values x_i that is above the median to any value(s) above the median, the same can also be done for data values below the median. This cannot affect the value of the median. (Hao L. & Naiman D.Q; 2007).

Quantile regression model fitness determinations:

For the purpose of this study, we are going to employ the following model fitness determination to ascertain the model fitness:

a. Coefficient of Determination

Huo & Naiman (2007) gave an analog of R^2 as:

$$R(p) = \frac{\tau \sum_{y_i \geq \hat{\beta}_0^{(\tau)} + \hat{\beta}_i^{(\tau)} x_i} |Q_y(\hat{\tau}) - \hat{\beta}_0^{(\tau)} - \hat{\beta}_i^{(\tau)} x_i| + (1-\tau) \sum_{y_i < \hat{\beta}_0^{(\tau)} + \hat{\beta}_i^{(\tau)} x_i} |Q_y(\hat{\tau}) - \hat{\beta}_0^{(\tau)} - \hat{\beta}_i^{(\tau)} x_i|}{\sum_{y_i \geq \hat{\beta}_0^{(p)}} \tau |Q_y(\hat{\tau}) - \hat{\beta}_0^{(\tau)}| + \sum_{y_i < \bar{y}} (1-\tau) |Q_y(\hat{\tau}) - \hat{\beta}_0^{(\tau)}|} \quad (14)$$

(Huo et al., 2013)

Where

$R(p)$ = the restricted coefficient of determination for quantile regression

$Q_y(\hat{\tau})$ = the response variable and the cdf inverse (F^{-1}) of the distribution to be estimated

x_i = the covariates to be simulated

$\hat{\beta}_0$ = the intercept parameter

$\hat{\beta}_i$ = the unknown parameters

τ = specified quantiles of the model. This research examines the following quantiles: 0.05, 0.25, 0.5, 0.75, 0.95

b. Model Selection Criteria

In order to select the appropriate model Akaike Information Criteria (AIC) was found to be adequate for the model selection criteria. This is given as:

$$AIC_c(\tau) = \left\{ 2n \ln \left(\frac{V^1(p)}{n} \right) + 2p(n/(n-p-1)) \right\} \quad (\text{Brain et al., 2005}) \quad (15)$$

where

p is the number of parameters,

n is the sample size,

2.7 Bootstrap for Standard Error

The bootstrap standard error is given as:

$$SE_{boot} = \sqrt{\frac{\sum_{m=1}^k (\hat{\beta}_{pi}^{*m} - \bar{\beta}^*)^2}{(k-1)}} \quad (\text{Wong., 2018}) \quad (16)$$

where

$\hat{\beta}_{pi}^{*m}$ is The Estimates of the regression coefficient for all the new samples

m is the no of bootstrapped sample

p is the p^{th} quantile

$\bar{\beta}^*$ denotes the mean of the bootstrapped regression coefficients and can be written as:

$$\bar{\beta}^* = \frac{\sum_{k=1}^k (\beta_{pi}^{*m})}{k} \quad (\text{Wong, 2018}) \quad (17)$$

c. Bootstrap for Confidence Interval

The confidence interval can be directly calculated from the non-parametric bootstrap resampling as:

$$CI = (\hat{\beta}_{pi} + z_{(1-\alpha/2)} SE_{boot}) \quad (\text{Wong, 2018}) \quad (18)$$

where

$z_{(1-\alpha/2)}$ is the $(1 - \alpha/2)$ quantile of a standard normal distribution.

SE_{boot} is the bootstrapped standard error.

d. Bootstrap for P – Value

To compute the T – Values for the estimated regression coefficient for all resampled datasets by using bootstrap is given as:

$$\hat{p}_{boot} = \frac{1}{k} \sum_{m=1}^k 1(T_{boot}^{*m} \geq T). \quad (\text{Wong, 2018}) \quad (19)$$

Where

$$T_{boot}^{*m} = \frac{\hat{\beta}_{pi}^{*m}}{SE_{boot}} = \frac{\hat{\beta}_{pi}^{*m}}{\sqrt{\frac{\sum_{m=1}^k (\hat{\beta}_{pi}^{*m} - \bar{\beta}^*)^2}{(k-1)}}$$

m denotes the m^{th} bootstrapped sample, $m = 1 - - - k$,

T_{boot}^{*m} is the m^{th} bootstrapped test statistic.

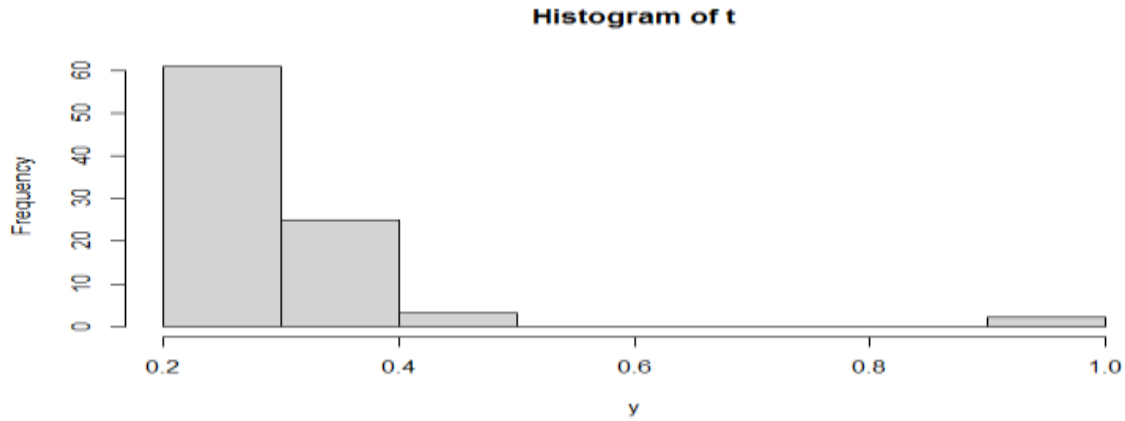
T is the test statistic of our original data

e. Software

Some of the packages used in this article are: Quantreg, Moments package and Rela package

DATA AND RESULTS

The experiment was done using real life data on fuel consumption for 91 cars measured on six explanatory variable with the design matrix given as $(n \times k)$, where $n = 91$, $k = 6$ and $x_i = (x_1, x_2, x_3, x_4, x_5, x_6)$ where x_1 is weight of the car, x_2 is length of the car, x_3 is width of the car, x_4 is Wheel base of the car, x_5 is Engine size of the car, and x_6 is the car's horse power, the response variable (y_i) is miles per gallon of fuel consumed by the car. The size of the sample is 91 but in order to estimate the standard error, the confidence interval and the p-values, the real data on fuel consumption were bootstrapped 200 times following Efron (1979). The data analysis was done using the model in equation (13):

**Fig 1:** Histogram of Cars' fuel consumption**Table 1: Estimated Parameters for the CQReg Model in Equation (13) Computed using Data on Cars' fuel consumption**

Quantiles	parameters	coefficient	Std error	T - value	Pr(> t)
0.05	intercept	0.01890	0.01875	1.00813	0.31629
	weight	0.00000	0.00000	0.09876	0.92157
	Length	0.00006	0.00004	1.63622	0.10554
	Wheel base	-0.00018	0.00015	-1.19987	0.23356
	width	-0.00040	0.00024	-1.66296	0.10005
	Engine Size	0.00125	0.00068	1.83661	0.06980
	Horse Power	-0.00002	0.00001	-1.39614	0.16635
0.25	intercept	-0.01699	0.00858	-1.98008	0.05097
	weight	-0.00001	0.00000	-4.16024	0.00008
	Length	0.00008	0.00002	3.21769	0.00184
	Wheel base	0.00002	0.00007	0.22930	0.81919
	width	0.00006	0.00009	0.65323	0.51539
	Engine Size	0.00024	0.00045	0.54877	0.58462
	Horse Power	0.00000	0.00001	-0.28229	0.77842
0.5	intercept	-0.01548	0.00636	-2.43319	0.01709
	weight	-0.00001	0.00000	-5.37093	0.00000
	Length	0.00006	0.00003	1.85380	0.04728
	Wheel base	0.00003	0.00007	0.38048	0.70455
	width	0.00009	0.00009	1.02585	0.30791
	Engine Size	0.00013	0.00042	0.31063	0.75685
	Horse Power	0.00000	0.00001	0.20961	0.83448
0.75	intercept	-0.00517	0.00772	-0.66995	0.50472
	weight	-0.00001	0.00000	-5.77556	0.00000
	Length	0.00001	0.00004	0.36953	0.01267

	Wheel base	0.00005	0.00009	0.55679	0.57915
	width	0.00006	0.00009	0.67536	0.50130
	Engine Size	0.00062	0.00058	1.07008	0.28765
	Horse Power	0.00000	0.00001	0.25664	0.79808
0.95	intercept	0.03252	0.11513	0.28248	0.77827
	weight	-0.00001	0.00001	-0.52314	0.60225
	Length	0.00014	0.00028	0.49358	0.62289
	Wheel base	-0.00037	0.00099	-0.36906	0.71301
	width	-0.00016	0.00103	-0.15139	0.88003
	Engine Size	0.00366	0.00424	0.86169	0.39131
	Horse Power	-0.00004	0.00008	-0.51268	0.60952

Table 2: Descriptive Analysis for the Residuals of the CQReg Model

quantiles	Skewness	kurtosis	mean	Median	RMSE	MSE	SD
0.05	4.58190	26.6010	0.00369	0.00250	0.00671	0.00005	0.00564
0.25	5.134375	31.4077	0.00167	0.00046	0.00568	0.00004	0.00546
0.5	5.141661	31.4474	0.00101	0.00000	0.00550	0.00003	0.00544
0.75	5.009144	30.3745	0.00000	-0.0007	0.00537	0.00003	0.00561
0.95	3.716878	21.1170	-0.0041	-0.0049	0.00694	0.00005	0.00607

Table 3: Model Comparison Results of the CQReg Model

Quantiles	0.05	0.25	0.5	0.75	0.95
Pseudo R	0.2944	0.3372	0.3215	0.2744	0.1524
AIC	-809.97	-835.97	-810.76	-738.45	-530.35

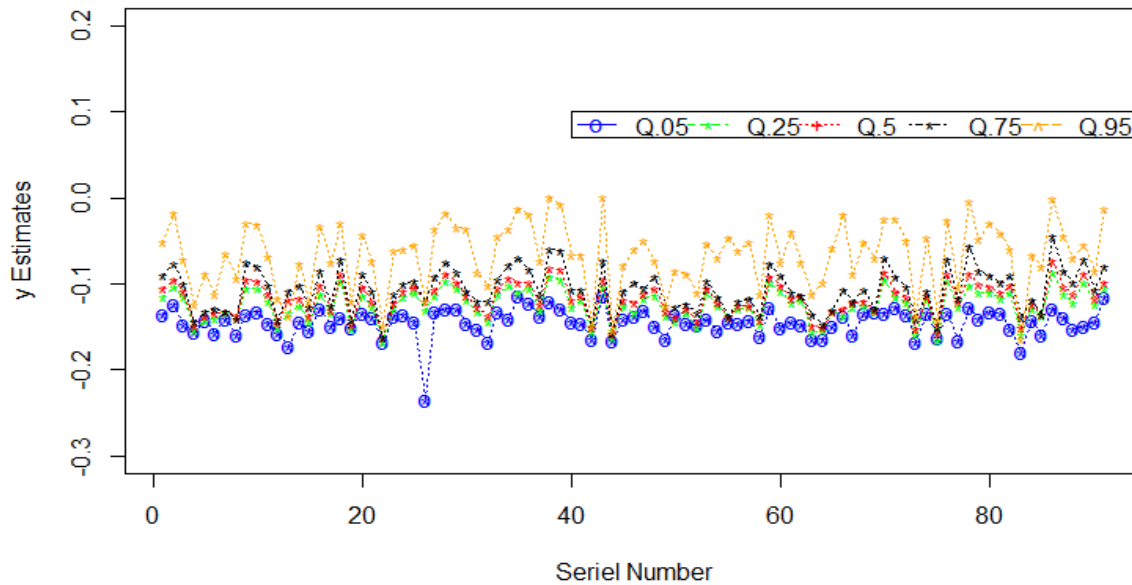
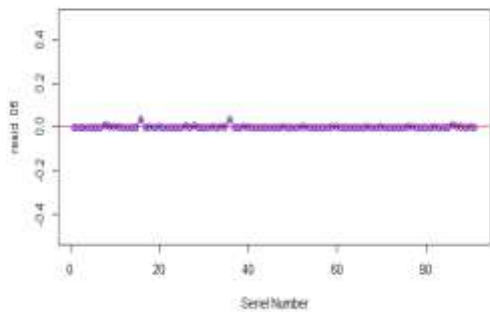
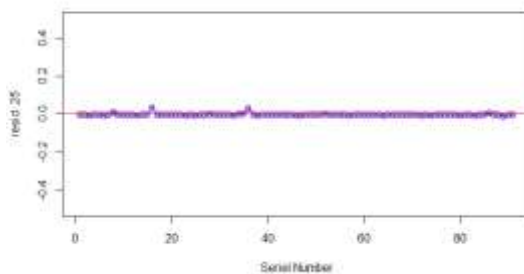


Fig.2: Plot for the y-Estimate of the CQReg Model



Residual for Quantile 0.05



Residual for Quantile 0.25

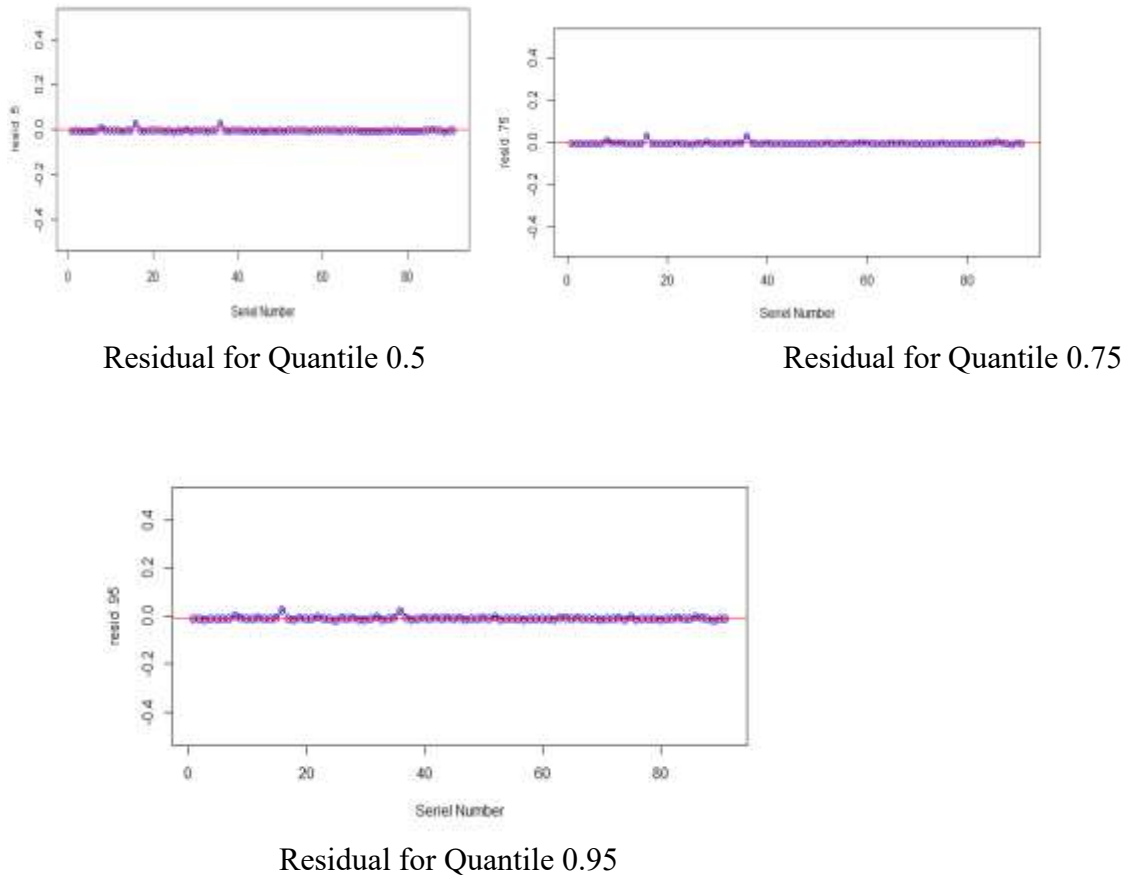


Fig.3: Plot for the Residual of Different Quantile for the CQReg Model

CONCLUSION

Data in Table 1 show that none of the parameters of the 5th quantile is significant, for the 25th and 50th quantiles both the intercept, weight and length of the car are all significant, for the 75th quantile both weight and length of the car are significant, and for the 95th quantile none of the parameters is significant.

This implies that for the 25th, 50th, and 75th quantiles of fuel consumption, weight and length of a car matter a lot. Width, wheelbase, engine size, and horsepower do not show a significant effect in that same middle range.

25th quantile are cars that already use relatively little fuel. heavier and longer cars still burn more fuel than lighter, shorter cars, the other factors don't add much explanatory power.

50th quantile is the typical car. Weight and length are the main drivers of fuel consumption. If a car gets longer or heavier, expect fuel use to go up. Engine size and HP aren't statistically driving differences at this level once weight and length are accounted for.

75th quantile are cars that already use relatively more fuel. The same pattern holds: weight and length keep pushing consumption up.

5th and 95th quantiles are not significant because they are the extremes, the most fuel-efficient and the least fuel-efficient cars. At those tails the data gets noisy. Fuel use there is often driven by outliers: unusual driving behavior, measurement error, specialty vehicles, or other factors not in the model. So, weight and length don't show a consistent, statistically significant relationship there, and neither do width, wheelbase, engine, or HP.

Result in Table 2 shows that the Skewness for all the study quantile are a little higher than zero while the kurtoses are very high above 3, the means are equal to the medians for all the study quantiles, the mean square errors (MSE), the root mean square errors (RMSE) and the standard deviation (SD) for all the study quantiles approximate to zero. The AIC for all the study quantiles in Table 3 show that the AIC is of low value. This implies that the residuals are normally distributed hence the CQReg Model is adequate for the study.

Following our conclusion, we therefore recommend the CQReg Model for study on fuel consumption of cars.

REFERENCES

- Alzaatreh, A, Lee, C, Famoye, F. A. andGhash, I (2016). The generalized Cauchy family of distributions with applications, *Statistical Distributions and Applications*, 3:12, DOI 10.1186/s40488-016-0050-3
- Alshawarbeh, E, Lee, C, Famoye, F; (2012), The beta-Cauchy distribution. *J. Probab. Stat. Sci.* 10(1), 41–57
- Brian, S. C., Noon, B.R.,Scherer, R.D., &John, J. K. (2017), Logistic quantile regression provides improved estimates for bounded. avian counts: A case study of California Spotted Owl fledgling production, *American Ornithological Society*. 134, 783–801, DOI: 10.1642/AUK-16-195.1
- Efron, B. (1979), Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1-26.
- Hao L. &Naiman,D.Q., (2007). Quantile Regression; 01-Hao.qxd. 3/13/2007.3.28
- Koenker, R. and Bassett, J., G. (1978). Regression quantiles. *Econometrica*, 46, 33-50.
- Nwakuya M. T., Nwabueze J. C., Onyegbuchulem, B. O., & Imoh, J. C. (2019), Response variable transformation for quantile regression model, *International Journal of Scientific & Engineering Research* (10) 8.
- Onyegbuchulem, B. O., Onyegbuchulem, C. A., Ogwo, O., & Oliwe, E. (2026a). Comparative Performance Stability of Two Cauchit Quantile Regression Models. *Asian Journal of Probability and Statistics*, 28(8), 118–130. <https://doi.org/10.9734/ajpas/2026/v28i8934>

Onyegbuchulem, B.O., Onyegbuchulem, C.A., Ogwo, O & Oliwe, E (2026b). Diagnostic tests for quantile regression with weibull response variables. Royal Statistical Society Nigeria Local Group 2026 Conference Proceeding

Stigler, S.M. (1989). Letter to the editor: normal orthant probabilities. *Am. Stat.* 43(4), 291

Wong, V (2018) Logistic quantile regression to evaluate bounded outcomes, Mathematical Statistics Stockholm University, <http://www.math.su.se>