https://doi.org/10.31871/WJIR.12.2.4

World Journal of Innovative Research (WJIR)
ISSN: 2454-8236, Volume-12, Issue-2, February 2022 Pages 01-09

An Alternative to Generalized Linear Models with Application

Mangga A.M, Nwaosu S. C., Ahmed A

***Abstract*— In model building, forecasting is the ultimate goal since it assists in policy implementation and further research work. This paper aimed at establishing a robust quantile regression analysis. Appropriate model and parameters of the model with their statistical test for effective prediction was established. It has been established that quantile regression relative to ordinary least squares produce regression estimates that are more robust against outliers.From 5th to 99th percentile, results showed that, at 25th percentile the number of tillers and plant height are more significant with p-value of 0.044 and 0.001 at 0.05 level of significance compare to others. The generalized linear models considered showed insignificance with p-value 1.000 and 0.760 at 0.05 level of significance.However, quantile regression tells us what happened as we move from the smallest to the highest quantile in estimating the goodness of fit test for the model for proper forecasting also, the model has been establishedat 25th percentiles.(Yield = -2419.596 + 1.352(Tillers) + 45.322(Height)) with the best yield which is significant and the best amongst others.**

***Index Terms*— Quantile, Model, Percentile, Tillers and Height.**

I. Introduction

Quantile Regression (QR) modelshave provided a valuable tool in economics, finance, and statistics as a way of capturingheterogeneous effects of covariates on the outcome of interest, exposing a wide variety offorms of conditional heterogeneity under weak distributional assumptions. Importantly, Quantile Regression also provides a framework for robust inference. Applying quantile regression to count data presents logical and practical complicationswhich are usually solved by artificially smoothing the discrete responsevariable through jittering.Other recent approaches include that of Congdon (2017), in which the asymmetric Laplace distribution is combined with a Poisson model in a Bayesian framework, and the model-based quantile regression of Padellini and Rue (2018), in which quantiles are mapped to the parameters of a generalized linear model identified by a continuous version of a valid count distribution. Tzavidis et al. (2015) proposed a semiparametric M-quantile approach for counts that extends theideas of Cantoni and Ronchetti (2001) and Breckling and Chambers (2001). These methods avoid jittering, but depend on a limited choice of predefined parametric models. [1]

Frumento and Bottai (2016, 2017) suggested using a fully parametric approach and reformulated model. Quantile regression is a type of regression analysis used in statistics and econometrics. Whereas the Method of Least Squares (MLS) results in estimates of the conditional mean of the response variable given certain values of the predictor variables$[E(y/x)]$, quantile regression aims at estimating either the conditional median or other quantiles of the response variable. Essentially, quantile regression is the extension of Linear Regression and we use it when the conditions of linear regressions are not applicable.

Quantile regression methods provide an alternative approach for robust inference. Rather than re- lying exclusively on a single measure of conditional central tendency, the quantile regression approach allows the investigator to explore a range of conditional quantile functions thereby exposing a variety of forms of conditional heterogeneity. [3].

Constant coefficient linear time series models have played an enormously successful role in statistics, and gradually various forms of random coefficient time series models, have also emerged as viable competitors in particular fields of application. One variant of the latter class of models, although perhaps not immediately recognizable as such, is the linear quantile regression model. This model has received considerable attention in the theoretical literature, and can be easily estimated with the quantile regression methods proposed in Koenker and Bassett (1978). Curiously, however, all of the theoretical work dealing with this model (that we are aware of) focuses exclusively on the i.i.d. innovation case that restricts the autoregressive coefficients to be independent of the specified quantiles. In this project work we seek to relax this restriction and consider linear quantile autoregression models who's autoregressive (slope) parameters may vary with quantiles $\tau \in [0, 1]$. We hope that these models might expand the modeling options for time series that display asymmetric dynamics or local persistency.

Considerable recent research effort has been devoted to modifications of traditional constant coefficient dynamic models to incorporate a variety of heterogeneous innovation effects. An important motivation for such modifications is the introduction of asymmetries into model dynamics. It is widely acknowledged that many important economic

Mangga A.M, Department of Mathematical Sciences, Abubakar Tafawa Balewa University, P. M. B. 0248, Bauchi, Nigeria.

Nwaosu S. C., Department of Mathematics/Statistics,University of Agriculture, Makurdi, Nigeria.

Ahmed A, 2Department of Mathematical Sciences, Abubakar Tafawa Balewa University,P. M. B. 0248, Bauchi, Nigeria.

variables may display asymmetric adjustment paths Enders and Granger (1998)).
We believe that quantile regression methods can provide an alternative way to study asymmetric dynamics and local persistency in time series. We propose a quantile autoregression (QAR) model in which autoregressive coefficients may take distinct values over different quantiles of the innovation process. We show that some forms of the model can exhibit unit-root-like tendencies or even temporarily explosive behavior, but occasional episodes of mean reversion are sufficient to ensure stationarity. The models lead to interesting new hypotheses and inference apparatus for time series, and data saddled with outliers [4].

II. MATERIALS AND METHODS

Quantile Regression Model

For better understanding, we start quantile regression from the basic idea of linear regression.
Apart from the mean, the lower and upper quantile are also important. A regression model does not capture the pattern of the situation. To better understand the quantile regression, we take a leave from linear regression. The quantile regression model estimates the potential differential effect of a covariate on various quantiles in the conditional distribution. For the linear regression model, we have

$Y_i = \beta_0 + \beta_1 x_1 + \mathcal{E}_i$ … (1)

From (3.1) above, $\mathcal{E}_{ii}$s identically, independently and normally distribute with mean zero and unknown variance σ^2. Owing to the fact that the error is normally distributed with mean zero, the function $\beta_0 + \beta_1 x$ is fitted to the data Corresponding to conditional mean of y given by $E[y|x]$ and is always interpreted as mean in the population of y values corresponding to the fixed value of x. Given p to denote proportion, that is (0< p <1). Then the corresponding quantile regression for the equation (1) above is

$Y_i = \beta_0^{(p)} + \beta_1^{(p)} x_i + \mathcal{E}_i^{(p)}$ … (2)

This indicates that the proportion of the population having scores below the quantile at p. The p^th conditional quantile $\mathcal{E}_i$ is defined as

$Q^{(p)}(y_i|x_i) = \beta_0^{(p)} + \beta_1^{(p)} x_i$ … (3)

Equation (3) above represent the conditional p^th quantile which can be determine by the quantile specific parameters $\beta_0^{(p)}$ and $\beta_1^{(p)}$ which are specific value of the covariate x_i.
Again, Let Y be a random variable with a distribution function F_Y and π be a real number between 0 and 1. The π^{th} quantile of F_Y denoted as $q_y(\pi)$ is the solution to $F_Y(q) = \pi$ which is given as

$q_y(\pi) = F_Y^{-1}(\pi) = \inf\{y: F_Y(y) \geq \pi\}$

where o <π < 1 is the quantile level.
The π^{th} quantile of F_Y can be obtained by minimizing the following function with respect to q

$$\pi \int_{y-q} |y-q|\, dF_y(y) + (1-\pi) \int_{y<q} |y-q|\, dF_y(y)$$

$$\pi \int_{y-q} (y-q)dF_y(y) - (1-\pi) \int_{y<q} (y-q)dF_y(y)$$

Applying the first order condition for minimization problem which is by taking its partial derivatives with respect to q and equate the result to zero

$$\{-\pi \int_{y>q} (y-q)dFy(y) - [-(1-\pi)] \int_{y<q} (y-q)dFy(y)\} = 0 \quad \dots (5)$$

$$-\pi \int_{y>q} (y-q)dFy(y) + (1-\pi) \int_{y<q} (y-q)dFy(y) = 0 \quad \dots (6)$$

Substituting in the limit we have (7) and opening the bracket will yield (8)

$-\pi[1-F_y(q)] + (1-\pi)F_y(q) = 0$ … (7)

$$-\pi + \pi Fy(q) + Fy(q) - \pi Fy(q) = 0$$

$-\pi + F_y(q) = 0$

$\pi = F_y(q)$ … (8)

Equation (8) solution is the πth quantile of F_y

Conditional Quantile

Using Chung-Ming, (2007) model, suppose we have Y as a response and X is the p dimensional predictor. $F_y(Y|X = x) = P(Y \leq y|X = x)$ denote the conditional cumulative density function of Y given $X = x$ then, the πth

https://doi.org/10.31871/WJIR.12.2.4 World Journal of Innovative Research (WJIR)
ISSN: 2454-8236, Volume-12, Issue-2, February 2022 Pages 01-09

conditional quantile of Y can be denoted as $Q\pi(Y|X=x) = inf\{y: Fy(y|x) \geq \pi\}$ furthermore, if the random variable y depend on x that is event y happening conditioning on another random variable x is $Fy|x(y)$, its πth quantile can be given as

$$Q_{y|x}(\pi) = F_{y|x}^{-1}(\pi) * Q_{y|x} \quad \ldots (9)$$

Equation (9) is a function of X, solving it by minimization and at the same time applying the first other condition yield

$$\min_q [\pi \int_{y>q} |y-q| \, dF_{y|x}(y) + (1-\pi) \int_{y<q} |y-q| \, dF_{y|x}(y)] \quad \ldots (10)$$

$Q_{Y|x}(0.5)$ is the conditional median which represents the center is the point of symmetry of $Q_{Y|x}$. If π is close to zero $Q_{Y|x}(\pi)$ is called the left tail of $F_{Y|x}$. Also if $Q_{Y|x}(\pi)$ is a linear function $X'\beta$, (when q is substitute with X'β) unknown up to the parameter vector β then equation (10) becomes

$$\min_q [\pi \int_{x>X'\beta} |x - X'\beta| \, dF_{x|y}(x) + (1-\pi) \int_{x<X'\beta} |x - X'\beta| \, dF_{X|y}(x)] \quad \ldots (11)$$

Using the same principle as (8) above, we have

$Q_{Y|x}(\pi) = X'\beta_{\pi}$ the solution is denoted as β_{π} which is the π^{th} conditional quartile.

Parameters Estimation

Before estimating the parameters, let first of all look at ordinary least square. In OLS, we minimize the sum of squares of the error (the error term) and thereafter find the optional vaue of β_0 and β_1 that is

$$OLS = \min \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 x_i)]^2 \ldots (12)$$

For the Quantile regression, we replace the square with absolute therefore its minimizes the absolute least deviation (LAD)

$$LAD = \min \sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)| \ldots (13)$$

The error term $e_i = y_i - (\beta_0 + \beta_{1x_i})$ and hence the Quantile regression denoted by τ which is an extension least absolute deviation (LAD) can be expressed as

$$\tau = \sum_{i=1}^{n} \{Q|e_i| + (1-Q)|e_i|\} \quad \ldots (14)$$

replacing the value of $e_i = y_i - (\beta_0 + \beta_1 x_i)$ in equation (14) and we have

$$\tau = \sum_{i=1}^{n} \{Q|y_i - (\beta_0 + \beta_1 x_i)| + (1-Q)|y_i - (\beta_0 + \beta_1 x_i)|\}$$

$$\tau = \sum_{i=1}^{n} Q|yi - (\beta 0 + \beta 1xi)| + \sum_{i=1}^{n} (1-Q)|yi - (\beta o + \beta_0 xi)| \quad \ldots (15)$$

it convenience to write (15) as

$$\tau = \min_{\beta_0,\beta_1} \sum_{i:y_i \geq \beta_0+\beta_1 x_i} Q\,|\,y_i - (\beta_0+\beta_1 x_i)\,| + \min_{\beta_0,\beta_1} \sum_{i:y_i \leq \beta_0+\beta_1 x_i} (1-Q)\,|\,y_i-(\beta_0+\beta_1 x_i)\,| \quad \ldots (16)$$

For OLS, we find the β_0 and β_1 by differentiation, maximum likelihood estimation but in quantile regression, we make use of linear programming (simplex algorithm, integer point algorithm or simplex algorithm).

We formulate quantile regression problem in a way analogous to the formulation of least square (conditional mean) regression. Let Y be a random variable with some distribution function and a sample y_i,i=1,2,…,n. The median of the set of sample $y_i$can be defined as the solution of the minimization problem.

$$\min_{\beta} \sum_{i}^{n} |\,y_i - \beta\,|, \beta \varepsilon R \quad \ldots (17)$$

Estimating the value of β in (17) above, that is θ^{th} sample of y_i then we have

$$\min_{\beta} \{\sum_{y \geq \beta}^{n} Q\,|\,y_i-\beta\,| + \sum_{y<\beta}^{n} (1-Q)\,|\,y_i-\beta\,|\}, \beta \varepsilon R \quad \ldots (18)$$

Where Q is penalty imposed on under prediction and 1-Q is penalty imposed on over prediction. From linear regression model, let consider a set of random variables y_i,i€[1,n],n€N that are paired with a set of X={x_i}, i=1,2,…,n and y_i is a realization of Y and hence we have

$$\min_{\beta_0\beta_1} \sum_{i \in \{i:y_{i \geq} \beta_0+\beta_1 x_i} Q\,|\,y_i - (\beta_0+\beta_1 x_i) + \min_{\beta_0,\beta_1} \sum_{i \in \{i:y_i<\beta_0+\beta_1 x_i\}} (1-Q)\,|\,y_i-(\beta_0+\beta_1 x_i)\,| \quad \ldots (19)$$

Since β can be β_0,β_1 and therefore β€R becomes β_0,β_1€R. Converting the (19) to linear programming problem, we introduce a non-negative variable s_i and r_i for which the following equation holds

$$y_i - (\beta_0+\beta_1 x_i) + s_i = 0, i \in \{i: y_i \geq \beta_0+\beta_1 x_i\} \quad \ldots (20)$$

$$s_i = 0, i \notin \{i: y_i \geq \beta_o+\beta_1 x_i\}$$

$$(\beta_0+\beta_1 x_i) - y_i + r_i = 0, i \in \{i: y_i < \beta_0+\beta_1 x_i\} \quad \ldots (21)$$

$$r_i = 0, i \in \{i: y_i < \beta_0+\beta_1 x_i\}$$

Since s_i and r_i>0 on complementary sets. We can re-write (20) and (21)

$$y_i - (\beta_0+\beta_1 x_i) + s_i - r_i = 0 \quad \ldots (22)$$

$$s_i \geq 0, r_i \geq 0, i \in [1,n] \quad \ldots (23)$$

Then (19) can be express as

$$\min_{s_i,r_i,\beta_0,\beta_1} \{\sum_{i \in \{i:y_i \geq \beta_0+\beta_1 x_i\}} Qs_i + \sum_{i \in \{i:y_i<\beta_0+\beta_1 x_i\}} (1-Q)r_i\} \quad \ldots (24)$$

Since s_i≥0, r_i≥0 then minimization function of (24) above becomes

https://doi.org/10.31871/WJIR.12.2.4
World Journal of Innovative Research (WJIR)
ISSN: 2454-8236, Volume-12, Issue-2, February 2022 Pages 01-09

$$\min_{s_i, r_i, \beta_0, \beta_1} \{\sum_{i=1}^{n} Qs_i + \sum_{i=1}^{n} (1-Q) r_i\} \quad \ldots (25)$$

Equation (22), (23) and (25) are the linear programming formulation of the quantile regression of (19) above. In summary, the estimation of parameters in quantile regression is done via linear programming.

III. RESULTS AND DISCUSSION

Table Summary of Quantile Regression

Yield		Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
0.05.	HEIGHT	26.82915	10.95797	2.45	0.018	4.771909	48.8864
	TILLERS	-.3909365	.9412576	-0.42	0.680	-2.28559	1.503717
	_CONS	-1281.001	674.3538	-1.90	0.064	-2638.405	76.40236
0.10	HEIGHT	28.42516	11.71113	2.43	0.019	4.851876	51.99845
	TILLERS	.3078091	.977165	0.32	0.754	-1.659122	2.27474
	_CONS	-1434.758	715.5499	-2.01	0.051	-2875.085	5.569423
0.15	HEIGHT	26.15	12.83133	2.04	0.047	.3218785	51.97811
	TILLERS	1.449999	.8440589	1.72	0.093	-.249003	3.149002
	_CONS	-1393.05	748.4139	-1.86	0.069	-2899.529	113.4293
0.20	HEIGHT	38.93838	13.3661	2.91	0.006	12.03381	65.84295
	TILLERS	1.416289	.7238217	1.96	0.056	-.0406884	2.873266
	_CONS	-2080.622	756.3746	-2.75	0.008	-3603.125	-558.1194
0.25.	HEIGHT	45.32166	12.67594	3.58	0.001	19.80633	70.83699
	TILLERS	1.351982	.6517332	2.07	0.044	.0401111	2.663853
	_CONS	-2419.596	706.9043	-3.42	0.001	-3842.521	-996.6714
0.30.	HEIGHT	54.33097	11.80222	4.60	0.000	30.57434	78.0876
	TILLERS	.8941423	.8731224	1.02	0.311	-.8633619	2.651647
	_CONS	-2864.586	658.998	-4.35	0.000	-4191.081	-1538.092
0.35	HEIGHT	47.49151	11.00345	4.32	0.000	25.34271	69.6403
	TILLERS	.4441779	.9960077	0.45	0.658	-1.560682	2.449037
	_CONS	-2364.226	636.2711	-3.72	0.001	-3644.973	-1083.479
0.40	HEIGHT	46.71641	11.45156	4.08	0.000	23.66561	69.7672
	TILLERS	.9175528	1.161073	0.79	0.433	-1.419565	3.254671
	_CONS	-2320.985	669.2884	-3.47	0.001	-3668.192	-973.7769
0.45	HEIGHT	52.90792	11.87167	4.46	0.000	29.01148	76.80436
	TILLERS	.791709	1.418251	0.56	0.579	-2.063081	3.646499
	_CONS	-2629.956	719.3343	-3.66	0.001	-4077.901	-1182.011
0.50	HEIGHT	53.96568	11.90502	4.53	0.000	30.00212	77.92924
	TILLERS	.7645777	1.750539	0.44	0.664	-2.759075	4.288231
	_CONS	-2681.051	745.561	-3.60	0.001	-4181.788	-1180.315
0.55	HEIGHT	50.95909	13.0534	3.90	0.000	24.68397	77.23422
	TILLERS	.2454546	2.039178	0.12	0.905	-3.859198	4.350107
	_CONS	-2356.946	824.8908	-2.86	0.006	-4017.365	-696.5264
0.60	HEIGHT	45.91223	14.40138	3.19	0.003	16.92376	74.90071
	TILLERS	-.6259455	2.630528	-0.24	0.813	-5.920924	4.669033
	_CONS	-1812.901	950.4585	-1.91	0.063	-3726.075	100.2722
0.65	HEIGHT	48.35139	14.48095	3.34	0.002	19.20276	77.50003
	TILLERS	-.7718782	2.710668	-0.28	0.777	-6.228169	4.684413
	_CONS	-1905.715	965.1009	-1.97	0.054	-3848.362	36.93283
0.70	HEIGHT	49.12381	16.04288	3.06	0.004	16.83117	81.41646
	TILLERS	-1.22	3.406536	-0.36	0.722	-8.077001	5.637002
	_CONS	-1814.533	1051.535	-1.73	0.091	-3931.164	302.0966

0.75	HEIGHT |	44.68595	18.02229	2.48	0.017	8.408963	80.96294
	TILLERS |	-1.643614	4.110853	-0.40	0.691	-9.918332	6.631104
	_CONS |	-1438.929	1215.085	-1.18	0.242	-3884.769	1006.911
0.80	HEIGHT |	46.99695	18.95242	2.48	0.017	8.847708	85.14618
	TILLERS |	2.018369	4.573416	0.44	0.661	-7.18744	11.22418
	_CONS |	-1909.834	1339.197	-1.43	0.161	-4605.498	785.8297
0.85	HEIGHT |	48.55745	19.45285	2.50	0.016	9.400898	87.714
	TILLERS |	6.913523	5.021539	1.38	0.175	-3.19431	17.02136
	_CONS |	-2421.605	1358.815	-1.78	0.081	-5156.759	313.5478
0.90	HEIGHT |	31.27261	18.56168	1.68	0.099	-6.090104	68.63533
	TILLERS |	5.18504	5.012381	1.03	0.306	-4.90436	15.27444
	_CONS |	-952.3943	1300.671	-0.73	0.468	-3570.51	1665.721
0.95	HEIGHT |	37.65052	19.71045	1.91	0.062	-2.024554	77.32559
	TILLERS |	1.540521	4.698095	0.33	0.744	-7.916253	10.9973
	_CONS |	-766.5238	1170.506	-0.65	0.516	-3122.63	1589.582
0.99	HEIGHT |	41.86546	20.02067	2.09	0.042	1.565931	82.16499
	TILLERS |	-.8680157	3.855182	-0.23	0.823	-8.628094	6.892062
	_CONS |	-643.6884	1081.384	-0.60	0.555	-2820.401	1533.024

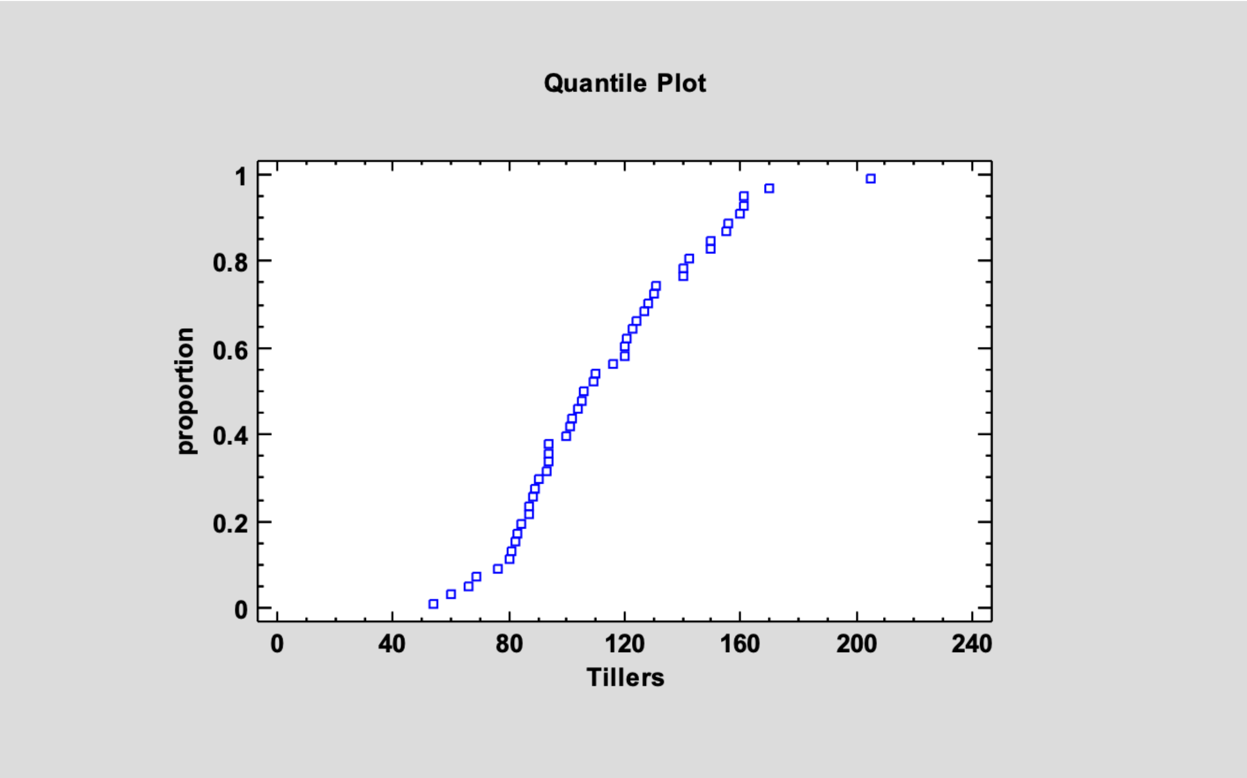

Fig. 4.1: Quantile Plot showing the distribution of the number of Tillers

The above procedure creates a plot that shows the proportion of data below each observed value of the number of Tillers.

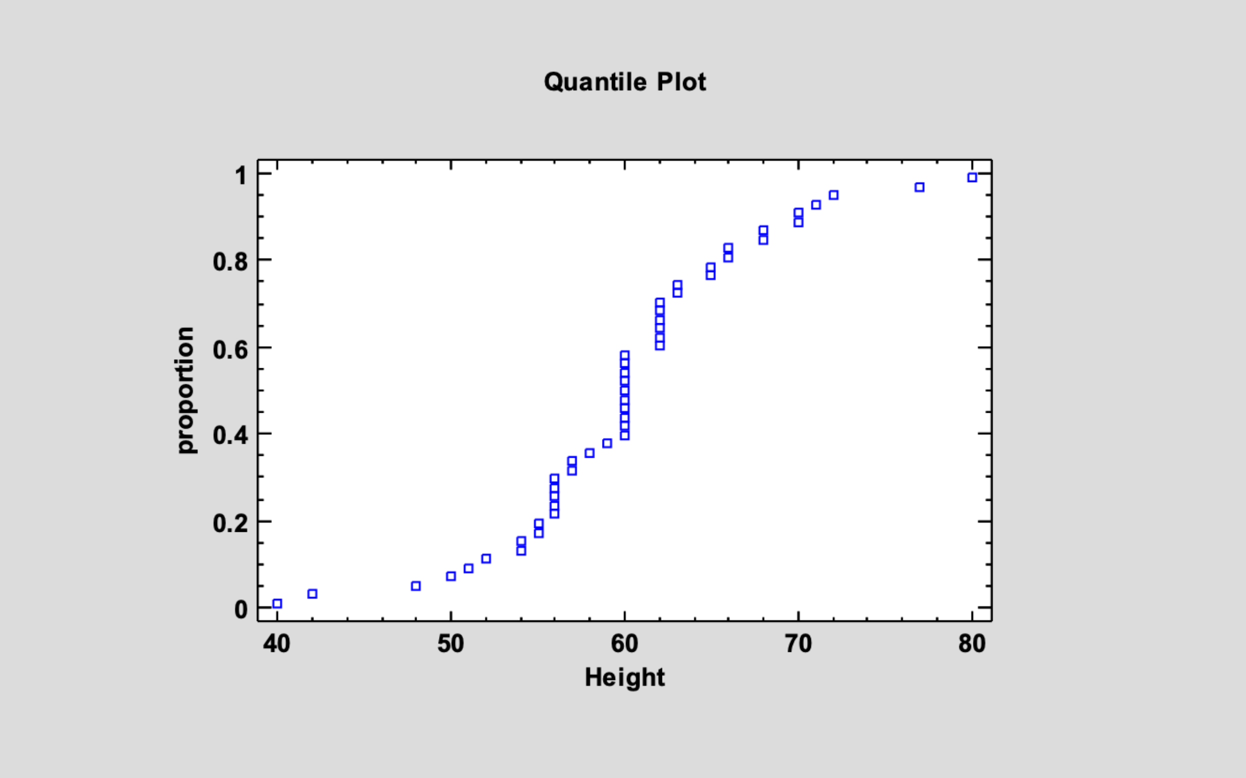

Fig. 4.2: Quantile Plot showing the distribution of Height (cm)

The above procedure creates a plot that shows the proportion of data below each observed value of the number of Tillers.

Discussion of Results

The quantiles refer to the general case of dividing a dataset or population into quarters. In this work we made use of twenty (20) quantiles, namely: 5th quantile, 10th quantile, 15th quantile, 20th quantile, 25th quantile, 30th quantile, 35th quantile, 40th quantile, 45th quantile, 50th quantile (Median), 55th quantile, 60th quantile, 65th quantile, 70th quantile, 75th quantile, 80th quantile, 85th quantile, 90th quantile, 95th quantile and 99th quantile using STATA Version 15 Statistical software. The research modelled the Grain Yield (kg/ha) as a function of the number of Tillers and the height (cm) of the plant.

From the summary table above, the result of 5th quantile showed that the number of tillers is insignificant with p-value of 0.680 while plant height is significant with p-value of 0.018 at 0.05 level of significance. The result of 10th quantile showed that the number of tillers is insignificant with p-value of 0.754 while plant height is significant with p-value of 0.019 at 0.05 level of significance. The result of 15th quantile showed that the number of tillers is insignificant with p-value of 0.093 while plant height is significant with p-value of 0.047 at 0.05 level of significance.

The result of 20th quantile showed that the number of tillers and plant height are insignificant with p-value of 0.056 and 0.006 respectively at 0.05 level of significance.The result of 25th quantile showed that the number of tillers and plant height are significant with p-value of 0.044 and 0.001 respectively at 0.05 level of significance.The result of 30th quantile showed that the number of tillers is insignificant with p-value of 0.311 while plant height is significant with p-value of 0.000 at 0.05 level of significance.The result of 35th quantile showed that the number of tillers is insignificant with p-value of 0.658 while plant height is significant with p-value of 0.000 at 0.05 level of significance.

The result of 40th quantile showed that the number of tillers is insignificant with p-value of 0.433 while plant height is significant with p-value of 0.000 at 0.05 level of significance. The result of 45th quantile showed that the number of tillers is insignificant with p-value of 0.579 while plant height is significant with p-value of 0.000 at 0.05 level of significance.The median Regression result showed that the number of tillers is insignificant with p-value of 0.664 while plant height is significant with p-value of 0.000 at 0.05 level of significance.The result of 55th quantile showed that the number of tillers is insignificant with p-value of 0.905 while plant height is significant with p-value of 0.000 at 0.05 level of significance. The result of 60th quantile showed that the number of tillers is insignificant with p-value of 0.813 while plant height is significant with p-value of 0.003 at 0.05 level of significance.The result 65th quantile showed that the number of tillers is insignificant with p-value of 0.777 while plant height is significant with p-value of 0.002 at 0.05 level of significance.

The result of the 70th quantile showed that the number of tillers is insignificant with p-value of 0.722 while plant height is significant with p-value of 0.004 at 0.05 level of significance.The result of the 75^{th} quantile showed that the number of tillers is insignificant with p-value of 0.691 while plant height is significant with p-value of 0.017 at 0.05 level of significance.The result of 80^{th} quantile showed that the number of tillers is insignificant with p-value of 0.661 while plant height is significant with p-value of 0.017 at 0.05 level of significance.The result 85^{th} quantile showed that the number of tillers is insignificant with p-value of 0.175 while plant height is significant with p-value of 0.016 at 0.05 level of significance.The result of the 90^{th} quantile showed that the number of tillers is insignificant with p-value of 0.306 while plant height is significant with p-value of 0.099 at 0.05 level of significance.The result of the 95^{th} quantile showed that the number of tillers and plant height are insignificant with p-value of 0.744 and 0.062 respectively at 0.05 level of significance.The result of the 99^{th} quantile showed that the number of tillers and plant height are insignificant with p-value of 0.823 and 0.042 respectively at 0.05 level of significance.

Analysis of Poisson Regression Model

Generalized linear models No. of obs = 49
Optimization : ML Residual df = 46

Variance function: V(u) = u [Poisson]
Link function : g(u) = ln(u) [Log]

AIC = 8.325254
Log likelihood = -200.9687141 BIC = -179.0237

Yield	Observed Coef.	Bstrap * Std. Err.	z	P>|z|	[95% Conf.	Interval]
Height	3.86e-15	2.03e-10	0.00	1.000	-3.97e-10	3.97e-10
Tillers	7.12e-16	3.55e-11	0.00	1.000	-6.96e-11	6.96e-11
_cons	-2.74e-13	1.30e-08	-0.00	1.000	-2.54e-08	2.54e-08
ln(Yield)	1 (exposure)					

Analysis of Gaussian Regression Model

Generalized linear models No. of obs= 49
Optimization : ML Residual df = 46
Variance function: V(u) = 1 [Gaussian]
Link function : g(u) = u [Identity]

AIC = 15.14927
Log likelihood = -368.1570605 BIC = 9633368

Yield	Observed Coef.	Bstrap * Std. Err.	z	P>|z|	[95% Conf.	Interval]
Height	42.69765	10.21391	4.18	0.000	22.67876	62.71655
Tillers	.4219653	1.380386	0.31	0.760	-2.283541	3.127472
_cons	-1820.388	647.7526	-2.81	0.005	-3089.96	-550.8161
ln(Yield)	1 (exposure)					

In generalized linear models

Poisson regression model, the result showed that the number of tiller and plant height are insignificant with p-value of 1.000 and 1.000 respectively at 0.05 level of insignificance.

For Gaussian regression model, the result showed that the number of tiller is insignificant and plant height is significant with p-value of 0.000 and 0.760 respectively at 0.05 level of significance.

IV. CONCLUSION

In this research work, we have demonstrated the potential use of quantile regression in dealing with rainfed barley observation nursery. Using real dataset, the result showed in table 4.5 above, that the number of tillers and plant height are significant with p-value of 0.044 and 0.001 respectively at 0.05 level of significance. Meanwhile the generalized linear model, the poisson regression model and Gaussian regression model is not significant. Quantile regression is a robust regression to outliers compared to generalized linear regression models.Traditional means regression models like Generalized Linear Model (GLM) are not able to capture the entire distribution of the data.

However, from quantile 5 to99, the quantile regression tells us what happened as we move from the smallest to the highest quantile. In estimating the goodness of fit test for the model for proper forecasting. The model has been established in equation 25th Quantile with the best yield which is significant and the best amongst others. Quantiles regression will be able to provide not just a single value for estimation but for numerous quantiles.

REFERENCES

[1] André F. B. M., JosmarMazucheli, Marcelo Bourguignon (2020). A parametric quantile regression approach for modelling zero-or-one inflated double bounded data. *Biometrical Journal.*

[2] Aznarte, J. L. (2017). Probabilistic forecasting for extreme no 2 pollution episodes. Environ. Pollut. 229, 321–328

[3] Baione F. & Biancalana D. (2019). An individual risk model for premium calculation based on quantile: a comparison between generalized linear models and quantile regression. *North American Actuarial Journal* **23**, 573–590. [Taylor & Francis Online], [Web of Science ®], [Google Scholar]

[4] Baione F., Biancalana D. & De Angelis P. (2019). A quantile regression approach for the analysis of the diversification in non-life premium risk. *Soft Computing* **24**, 8523–8534. Doi: 10.1007/s00500-019-04291-x. [Crossref], [Web of Science ®], [Google Scholar]

[5] Bai, L., Wang, J., Ma, X. & Haiyan, L. (2018). Air pollution forecasts: An overview. Int. J. Environ. Res. Public Health 15(4).

[6] Baur, D., Saisana, M. & Schulze, N. (2004). Modelling the efects of meteorological variables on ozone concentration: A quantile regression approach. Atmos. Environ. 38(28), 4689–4699.

[7] Bergmeir, C., Hyndman, R.J. & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. Comput. Stat. Data Anal. 120, 70–83.

[8] Biancalana D. (2017). Unapproccio quantile regression per la tariffazionedanni, basatosu un modello a due parti. Ph.D. Thesis. https://iris.uniroma1.it/retrieve/handle/11573/1209293/944777/Tesi%20dottorato%20Biancalana.pdf. [Google Scholar]

[9] Brian,S.C.(2003).Quantile RegressionModelsof AnimalHabitat relationships.Ph.D. ThesisColoradoState UniversityFortCollins, ColoradoSpring.194pp

[10] Dong A., Chan J. & Peters G. (2015). Risk margin quantile function via parametric and non-parametric Bayesian approaches. *ASTIN Bulletin* **45**(3), 503–550. doi: 10.1017/asb.2015.8. [Crossref], [Web of Science ®], [Google Scholar]

[11] Fabio Baione&DavideBiancalana (2020). An application of parametric quantile regression to extend the two-stage quantile regression for ratemaking. Scandinavian actuarial journal (2020) vol 2021 (6).

[12] Frees E. W. (2010). *Regression modeling with actuarial and financial applications*. New York: Cambridge University Press. [Google Scholar]

[13] Frees E. W., Jin X. & Lin X. (2013). Actuarial applications of multivariate two-part regression models. *Annals of Actuarial Science* **7**, 258–287. [Crossref], [Google Scholar]

[14] Frumento P. (2017). QRCM: quantile regression coefficients modeling. R package version 2.1. https://cran.r-project.org/package=qrcm. [Google Scholar]

[15] Frumento P. & Bottai M. (2016). Parametric modeling of quantile regression coefficient functions. *Biom* **72**, 74–84. doi:10.1111/biom.12410. [Crossref], [Google Scholar]

[16] Frumento P. & Bottai M. (2017). Parametric modeling of quantile regression coefficient functions with censored and truncated data. *Biometrics* **73**(4), 1179–1188. doi: 10.1111/biom.12675 [Crossref], [PubMed], [Web of Science ®], [Google Scholar]

[17] Fuzi M. F. M., Jemain A. A. & Noriszura I. (2016). Bayesian quantile regression model for claim count data. *Insurance, Mathematics & Economics* **66**, 124–137. [Crossref], [Web of Science ®], [Google Scholar]

[18] Heras A., Moreno I. & Vilar-Zanón J. L. (2018). An application of two-stage quantile regression to insurance ratemaking. *Scandinavian Actuarial Journal* **9**, 753–769. [Taylor & Francis Online], [Google Scholar]

[19] Hong, T. *et al.* (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. Int. J. Forecast. 32(3), 896–913.

[20] Hothorn, T., Kneib, T. &Bühlmann, P. (2014). Conditional transformation models. J. R. Stat. Soc. B 76(1), 3–27.

[21] Jang, Y., Kim, J. H., Lee, H., Lee, K. &Ahn, S. A. (2018). A quantile regression approach to explain the relationship of fatigue and cortisol, cytokine among Koreans with Hepatitis b. Sci. Rep. 8(1), 16434.

[22] Ke, G. et al. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Adv. Neural Inf. Process. Syst. 30, 3146–3154.

[23] Koenker R. (2005). *Quantile regression*. Econometric Society Monograph Series, Vol. **38**. Cambridge: Cambridge University Press. [Crossref], [Google Scholar]

[24] Koenker R. (2015). Quantreg: quantile regression. R package version 5.19. http://CRAN.R-project.org/package=quantreg. [Google Scholar]

[25] Lebotsa, M. E. et al. (2018). Short-term electricity demand forecasting using partially linear additive quantile regression with an application to the unit commitment problem. Appl. Energy 222, 104–118.

[26] Mangalova, E. &Shesterneva, O. (2016). K-nearest neighbors for gefcom2014 probabilistic wind power forecasting. Int. J. Forecast. 32(3), 1067–1073.

[27] Martínez-Silva, I., Roca-Pardiñas, J. &Ordóñez, C. (2016). Forecasting SO2 pollution incidents by means of quantile curves based on additive models. Environmetrics 27(3), 147–157.