ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

NETWORKING PROJECT PLANNING AND SCHEDULING OF BUILDING CONSTRUCTION USING ROBUST REGRESSION ANALYSIS BASED ON ITERATIVE REWEIGHTED LEAST SQUARES

Udensi, Chiege .N, Lasisi K.E, Abdulrasheed. B, Mathias Ojochide Enemali

Department of Mathematical Sciences, Faculty of Science, Abubakar Tafawa Balewa University, Bauchi.

DOI: https://doi.org/10.5281/zenodo.13379601

Published Date: 27-August-2024

Abstract: **Effective project planning and scheduling are crucial for the successful execution of building construction projects. Traditional scheduling methods often encounter challenges such as uncertainty in task durations and resource constraints. This research explores the application of robust regression analysis, specifically iterative reweighted least squares (IRLS), as a method to enhance the accuracy and reliability of project planning and scheduling networks. Robust regression techniques are employed to mitigate the impact of outliers and deviations from normal distributions, common in construction project data. The study focuses on developing a framework that integrates robust regression into the scheduling process, aiming to improve the estimation of task durations, resource allocation, and overall project timelines. Through the case study, the method use is critical path method (CPM), project evaluation review technique (PERT) which identified the critical paths and probability of finishing the project on time. The research extends to regression analysis using ordinary least squares to determine the relationship between the independent variables (Time optimistic, Time most likely and Time pessimistic) to the dependent variable and identify outliers in the data.MM Estimation was use to minimize the outliers. The efficacy of the proposed approach is demonstrated, highlighting its potential to enhance the efficiency and resilience of project management practices in building construction.**

Keywords: **Regression Analysis, Iterative reweighted Least Squares, Planning and Scheduling.**

1. INTRODUCTION

The construction industry is characterized by its inherent complexity, with projects often plagued by uncertainties, delays, and cost overruns. Effective project planning and scheduling are crucial for mitigating these risks and ensuring project success. Traditional project management methodologies, while valuable, often rely on deterministic approaches that may not adequately account for the inherent variability and uncertainties in construction processes.

In recent years, there has been a growing interest in employing advanced statistical techniques to enhance project management practices. Robust regression analysis, a statistical method designed to handle outliers and influential data points, offers a promising avenue for improving the accuracy and reliability of project planning and scheduling. By incorporating iterative reweighted least squares, a robust regression algorithm, it is possible to develop models that are more resilient to data anomalies and provide more accurate predictions of activity durations.

Regression analysis, a form of predictive modelling, is a statistical process that investigates the relationship between a dependent variable and one or more independent variables. In the context of construction project management, these variables could range from resource allocation and task dependencies to different construction phases. By examining these relationships, regression analysis can provide valuable insights into how various factors impact project timelines, thereby enabling project managers to make informed decisions.

This research aims to develop a robust framework for networking project planning and scheduling in the context of building construction. By leveraging the power of robust regression analysis, the study seeks to create a model that can accurately predict activity durations, considering the various factors influencing project performance. This model will be instrumental in improving project planning, resource allocation, and risk management, ultimately leading to enhanced project outcomes.

Problem statement:

Effective scheduling and planning are critical components of successful building project management. Project planning and scheduling plays a central role in predicting both the time and cost aspects of a project. M. Mostafa (2023): The application of PERT and CPM method in building construction industry to organise all the project's schedule systematic studied how to use PERT and CPM methods in the construction industry to reduce project completion times and obtain a proper project model through PERT and CPM based networks but could not consider robust regression analysis based on iterative reweighted least squares. Traditional methods often fall short in addressing complexities and uncertainties inherent in construction projects, leading to delays, cost overruns, and resource misallocations. To enhance the reliability and accuracy of project schedules, this study investigates the application of robust regression analysis based on iterative reweighted least squares in the context of building project management as a framework that integrates robust regression into the scheduling process, aiming to improve the estimation of task durations, resource allocation, and overall project timelines.

Aim of the Research

The research investigated the critical paths, project completion time discrepancies, and robust regression analysis based on iteratve reweighted least squares to identify outliers in the network data, important predictors and maximize time spent in the construction process of storey building lecture classes.

Objective the Research

1. Identify and analyse the critical path and associated activities involved in the construction process of storey building lecture halls.

2. Evaluate the discrepancy between the estimated project completion time and the industry-standard benchmark, focusing on the probability of completing the project within the specified timeframe.

3 Apply robust regression analysis based on iteratve reweighted least squares to assess the relationship between the duration of activities (To, Tm, Tp) and the overall project completion time in the construction process of storey building lecture halls, identifying any significant predictors and minimizing the outliers of project duration.

4. Examine the effectiveness of the software computation of each activity duration using robust regression analysis based on iterative reweighted least squares with the manually computed network analysis on the project completion time.

2. LITERATURE REVIEWS

By Robbins & Coulter (2012) disclose that management includes coordinating and supervising one's work so that its activities can run effectively and efficiently. This lead to the implementation of four management functions namely planning, coordinating, leading, and controlling. The project management in the arts and sciences for planning, organizing, monitoring and controlling all aspects of the project to achieve project goals and targets safely in accordance with the agreed schedule, budget, and performance criteria (Radujković & Sjekavica, 2017). By Al-Hajj & Zraunig (2018) described project management practices try to be able to complete projects efficiently to minimize costs and be able to achieve external goals related to customer needs. The main task of project management is to plan and control the project schedule (Zareei, 2018). Berjis et al. (2020) in his paper project planning is an instrument used to develop work plans taking into account of various aspects of each activity and to predict project status during the project life cycle. However,

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

Project planning is divided into several stages, namely activity planning, an order of activities, resource allocation, scheduling, and project floatation. The success of project management will affect the success of the project, because the success of a project depends on the success of the project management and the success of the final product produced. Gomes & Romão,(2016). A project is said to be successful or unsuccessful depending on whether or not the project meets the standards of time, cost, and quality on the project (Gomes & Romão, 2016).

Atin & Lubis (2019) stated that scheduling the planning of project activities outlined in a work schedule which describes the sequence in the process of work on an activity accompanied by the time of comment and end of work in an activity. This scheduling is useful as a guideline for each work unit of activity against time constraints in starting and ending a task.

OUTLIERS IN MULTIPLE LINEAR REGRESSION

The multiple linear regression model in terms of the observations can be written as matrices notation by $Y = X\beta + \varepsilon$, where y is an $n\times 1$ vector of observed response values, X is the $n\times p$ matrix of the predictor variables, β is the $p\times 1$, and ε is the $n\times 1$ vector of random error terms. The aim of regression analysis is to find the estimates of unknown parameters. The OLS is used to find the best estimate of β's with the least squares criterion which minimizes the sum of squared distances of all of the points from the actual observation to the regression surface . It often happens in practice that an assumed normal distribution model (e.g., a location model or a linear regression model with normal errors) holds approximately in that it describes the majority of observations, but some observations follow a different pattern or no pattern at all. In the case when the randomness in the model is assigned to observational errors which was the first instance of the use of the least-squares method, the reality is that while the behavior of many sets of data appeared rather normal, this held only approximately, with the main discrepancy being that a small proportion of observations were quite atypical by virtue of being far from the bulk of the data. Behavior of this type is common across the entire spectrum of data analysis and statistical modeling applications. Such atypical data or even a single outlier can have a large distorting influence on a classical statistical method that is optimal under the assumption of normality or linearity. The primary purpose of robust regression techniques is to fit a model that describes the information in the majority of the data. This general definition implies that these techniques should perform well on both with outliers and on without outliers.

Ordinary least square (OLS) is considered as a best technique in model selection only under some assumptions are met (Zuur et al., 2009). The problem in the dataset occurs in case of outlier are present in the dataset. The linear regression estimates got effected in presence of outliers. The efficiency of OLS get reduced in such kind of dataset. So for such kind of problems, robust methods are available for handling the issue as (Gad & Qura, 2016) reviewed in their study the different types of robust methods for handling the outliers. Many kind of robust estimators are available as maximum likelihood type estimators (M estimators), modified M estimators (MM) and estimators of scale (S) estimators (Susanti et al.,2014). But mostly researchers preferred M estimators (Sinova & Van Aelst, 2018). The main purpose of the robust regression is to provide efficient estimates even in case of outliers Draper and Smith (1998). In robust M estimators, the weighted functions are reduced at the tails in comparison of the least squae estimators in which weight one is given to all observations (Stuart, 2011). Robust type of estimators are used by (Dupuis & Victoria, 2013) for developing the variance inflation factor (VIF) regression for dealing with outliers. later on (Amini & Roozbeh, 2016) introduced robust ridge regression with the help of some robust estimtors for the problems of outliers in the dataset. One of the work was of Lukman et al.(2017) , a comparison was made from them for M, MM, LTS, LAD, OLS consisted of six economic variables from 1947 to 1962. Later on, the shrinkage robust estimators are developed by (Norouzirad et al., 2017) for combined problem of multicollinearity and outliers. In previous research, there are many types of robust estimators were developed but the most common type is the M estimators due to its advantages and properties (Sinova & Van Aelst, 2018).

Robust regression is a method used when the distribution of the residual is not normally distributed and there are some outliers which affect the model (Susanti et al., 2014). It detects the outliers and provides better results (Chen, 2002). A common method of robust regression is the M estimate, introduced by Huber (1973), which is as efficient as Ordinary Least Square (OLS), and is considered the simplest approach. The Least Trimmed Squares (LTS) estimation was introduced by Rousseeuw (1984), and is a high breakdown value method. So, too, is the S estimation, another high breakdown value method with a higher statistical efficiency than LTS estimation (Rousseeuw & Yohai, 1984). The S

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

estimation is used to minimize the dispersion of residuals. The MM estimation ,a special type of M estimation introduced by Yohai (1987), combines high breakdown value estimation and efficient estimation.

MM- ESTIMATOR

Yohai (1983) proposed the class of MM- estimator in the straight relapse setting. MM-estimator has turned out to be progressively prominent and is maybe now the most ordinarily utilized vigorous relapse method. They consolidate a high breakdown point (half) with great effectiveness. The "MM" in the name alludes to the way that in excess of one M-estimation methodology is utilized to ascertain the last gauges. Typically, one begins with a profoundly strong relapse estimator, traditionally an S-estimator. At that point one can utilize the scale dependent on this starter fit alongside a superior tuned ρ capacity to acquire a more productive M estimator of the relapse parameter. A MM-estimator of ρ is then arrangement of an M-type condition.

3. METHODOLOGY

The collected data undergone comprehensive analysis using regression analysis techniques to explore the relationships between variables and develop predictive models for project planning and scheduling. The following steps outline the data analysis process:

Networking Analysis: Network analysis techniques such as Critical Path Method (CPM) and Program Evaluation Review Technique (PERT) will be utilized to visualize project dependencies and optimize scheduling strategies. The Critical Path Method (CPM) is a mathematically grounded algorithm for scheduling project activities, crucial for effective project management. Any project with interdependent activities can apply this scheduling method

Critical Path Method assist in the management of projects in two different ways: The forward and backward pass.

The forward pass calculation is obtained with the formula below:

$$ES_j = Maxi\,(ES_i + D_{ij}) \qquad \forall\ ij \qquad [1]$$

After the forward pass computation, the earliest completion and the latest completion of activities are obtained from the backward pass computation with the formula stated below:

$$LC_i = Min\lfloor LC_j - D_{ij} \rfloor\ \forall_{ij}\ \text{activities} \qquad [3]$$

THE PERT APPROACH :

This is called programmable Evaluation and review techniques. This deals with problem of uncertain activities. The statistical analysis to apply in order to estimate or determine time of each activity concerning the project used 3-time estimates. Namely are: to: The optimistic time (T_o) ;The pessimistic time (T_p) The most likely time (T_m) .

Optimistic estimate: the time the activity would take if things did go well.

Pessimistic estimate: the time the activity would take if things did not go well.

Most likely estimate: the consensus best estimate of the activity’s duration.

The duration of an activity is calculated using the following formula:

$$T_e = \frac{(T_o + 4T_p + T_p)}{6} \qquad [4]$$

Where,

t_e is the expected time,

The standard deviation, which is a good measure of the variability of each activity is calculated by the rather simplified formula:

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

$$S = \frac{(T_p - T_o)}{6} \qquad [5]$$

ROBUST REGRESSION

Robust regression is a regression method that is used when the distribution of residual is not normal or there are some outliers that affect the model. This method was used to analyze the network data that were affected by outliers such that the resulting models stouted against outliers. The researcher set regression models and test the common assumption that the regression assumptions violated, the transformation seemed unlikely to eliminate or weaken the influence of outliers which eventually became biased predictions. Under these circumstances, Robust regression was used to detect outliers and provide results that are resistant to the outliers.

MM ESTIMATION

we look for first partial derivative $\hat{\beta}_M$ to β so that

$$\sum_{j=1}^{n} x_{ij}\psi\left(y_i - \sum_{j=0}^{k} x_{ij}\beta \right) = 0, \text{ j} = 0,1,...,\text{k}$$

Where $\psi = \rho^1$, x_{ij} is the i-th observation on the j-th independent variable and x_i =1

We take c = 4.685 for Tukey's bisquare weighted function. So equation (10) becomes

$$\sum_{j=1}^{n} x_{ij} w_i\left(y_i - \sum_{j=0}^{k} x_{ij}\beta \right) = 0, \text{ j} = 0,1,...,\text{k}$$

3.5.4. MM-estimation Algorithm:

1. Initial estimates of the coefficients $\hat{\beta}^{(1)}$and corresponding residuals $e_i^{(1)}$are taken from a highly resistant regression (i.e., a regression with a breakdown point of 50%). Therefore, S-estimation will be use with a bisquare weight function.

2. The residuals $e_i^{(1)}$ from the initial estimation at Stage 1 are used to compute an M-estimation of the scale of the residuals, $\hat{\sigma}_e$.

3. The initial estimates of the residuals $e_i^{(1)}$ from Stage (1) and of the residual scale $\hat{\sigma}_e$ from Stage (2) are used in the first iteration of weighted least squares to determine the M-estimates of the regression coefficient

where the w_i is bisquare weights function.

4. New weights are calculated w_i , using the residuals from the initial WLS (Step 3).

5. Keeping constant the measure of the scale of the residuals from Step 2, Steps 3 and 4 are continually reiterated until convergence

4. DATA PRESENTATION, ANALYSIS AND DICUSSIONS

4.1 INTRODUCTION

In this chapter, data analysis and discussion of the results are presented.

In the realm of project management, efficiency and precision are paramount (Haughey, 2020). As the complexities of modern construction projects continue to evolve, the need for robust planning and scheduling methodologies becomes increasingly evident (Kerzner, 2018). Within this context, the application of the Critical Path Method (CPM) and Program Evaluation Review Techniques (PERT) emerges as a cornerstone for effective project planning and scheduling.

To determine the critical path, the project network needs to be calculated using the forward and backward pass procedures.

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

Table 1: for the forward pass on CPM

Coefficient	Calculations of coefficient	Results
ES_1	**0**	**0**
ES_2	$(ES_1 + D_{1,2})$	4
ES_3	(ES_3+ $D_{3,4}$)	9
ES_4	(ES_3+ $D_{3,4}$)	13
ES 5 ⋮ ES_{55}	{ ES_3+ $D4_{,5}$ } Max{ ES_8+ $D_{8,10}$, ES_9 + $D_{9,10......D54,55}$}	16 336

The table 1: it shows that the process is increases at the same time number are in increase order**.**

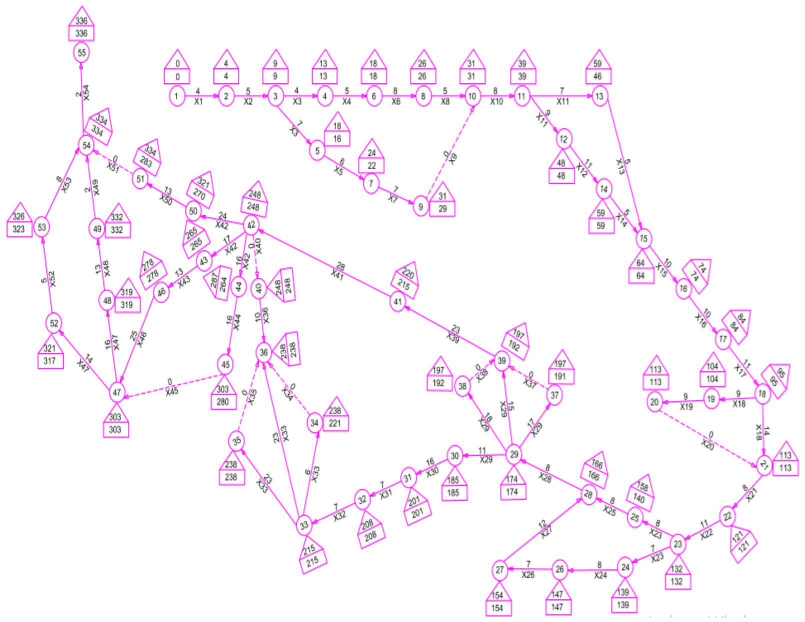
Table 2: for the Backward pass on CPM

Coefficient	Calculations of coefficient	Results
LF_{55}	336	336
LF_{54}	(LF_{55}- $D_{55,54}$)	334
LF_{53}	(LF_{54}- $D_{54,53}$)	326
LF_{52}	(LF_{53}- $D_{53,52}$)	329
⋮ LF_1	0	0

The table 1: it shows that the process is decrease at the same time number are decrease in order**.**

FIGURE 1:

A C.P.M NETWORK DIAGRAM OF A STOREY BUILDING LECTURE CLASSES



Let us calculated the Critical Path

1→2→3→4→6→8→10→11→12→14→15→16→17→18→19→20→21→22→23→24→

26→27→28→29→30→31→32→33→35→36→40→42→43→46→47→48→49→54→55

4+5+4+5+8+5+8+9+11+5+10+10+11+9+9+0+8+11+7+8+7+12+8+11+16+7+7+23+0+10+0+17+13+25+16+13+2+2 =336

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

4.2: Descriptive Statistics for Three Estimates of a Storey Building Lecture Classes.

The table below shows the descriptive statistics for three estimates namely time optimistic, time most likely and time pessimistic with their mean, standard deviation and variance.

Table 3: below shows the descriptive statistics for three estimates namely time optimistic, time most likely and time pessimistic with their mean, standard deviation and variance.

S/N	Activity	Predecessor	Duration in days	To	Tm	Tp	Mean	Sd	Variance
1	Site Preparation	-	4	3	4	6	4.1666667	0.5	0.25
2	Site Clearing	X1	5	4	5	6	5	0.333333	0.111111
3	Levelling of the site	X2	4	3	4	5	4	0.333333	0.111111
4	Line out	X3	7	5	7	9	7	0.666667	0.444444
5	Excavation	X3	5	4	5	7	5.1666667	0.5	0.25
6	Preparation of steel cage	X4	6	5	6	9	6.3333333	0.666667	0.444444
7	PCC	X5	8	6	8	10	8	0.666667	0.444444
8	Column line out	X6	7	5	7	9	7	0.666667	0.444444
9	Footing	X7	5	4	5	7	5.1666667	0.5	0.25
10	Column	X8, X9	8	6	8	10	8	0.666667	0.444444
11	Ground beam	X10	9	7	9	12	9.1666667	0.833333	0.694444
12	Curing	X11	7	5	7	8	6.8333333	0.5	0.25
13	Brick work up to pinch beam	X11	11	9	11	12	10.833333	0.5	0.25
14	Murum filing	X12	5	4	5	6	5	0.333333	0.111111
15	Compaction	X13, X14	5	3	5	6	4.8333333	0.5	0.25
16	Pinch beam	X15	10	7	10	11	9.6666667	0.666667	0.444444
17	Installation of septic tank and plumbing pipe	X16	10	8	10	12	10	0.666667	0.444444
18	Bed concrete	X17	11	10	11	13	11.166667	0.5	0.25
19	Column line out and column starter	X18	9	7	9	10	8.8333333	0.5	0.25
20	Column casting up to 7 feet	X19	9	7	9	10	8.8333333	0.5	0.25
21	Brick work up to 7 Feet	X18, X19	14	12	14	15	13.833333	0.5	0.25
22	Seal Casting	X20	8	6	8	11	8.1666667	0.833333	0.694444
23	Lintel	X21	11	8	11	13	10.833333	0.833333	0.694444
24	Loft and Lintel	X23	7	5	7	9	7	0.666667	0.444444
25	Porch slab	X23	8	6	8	9	7.8333333	0.5	0.25
26	Brick level up to slab level	X24, X25	7	6	7	9	7.1666667	0.5	0.25
27	Slab bean casting	X26	12	10	12	14	12	0.666667	0.444444
28	Centring and shuttering	X25, X27	8	6	8	11	8.1666667	0.833333	0.694444
29	Slab casting	X28	8	7	8	9	8	0.333333	0.111111
30	Head room column	X29	11	9	11	13	11	0.666667	0.444444
31	Head room brick work	X29, X30	16	14	16	18	16	0.666667	0.444444
32	Head room Lintel	X31	7	6	7	8	7	0.333333	0.111111
33	Head room slab	X32	7	6	7	9	7.1666667	0.5	0.25
34	Wall Putting	X33	6	5	6	7	6	0.333333	0.111111
35	External Plastering	X33, X36	23	20	23	26	23	1	1
36	Internal plastering	X33, X35	23	20	23	26	23	1	1
37	Plumberng	X40	17	15	17	20	17.166667	0.833333	0.694444
38	Electrical Work	X29	18	15	18	21	18	1	1
39	Steel Railing	X29,X37,X38	15	14	15	17	15.166667	0.5	0.25
40	Water Proofing	X36	10	8	10	11	9.8333333	0.5	0.25
41	Flooring	X39	23	22	23	25	23.166667	0.5	0.25
42	Roofing	X41	28	24	28	32	28	1.333333	1.777778
43	POP	X42	17	14	17	19	16.833333	0.833333	0.694444
44	Siling window fixing	X42	16	14	16	17	15.833333	0.5	0.25
45	Door fixing	X44	16	14	16	17	15.833333	0.5	0.25
46	Fittings	X43	13	10	13	16	13	1	1
47	Paintings	X45,X46	25	22	25	27	24.833333	0.833333	0.694444

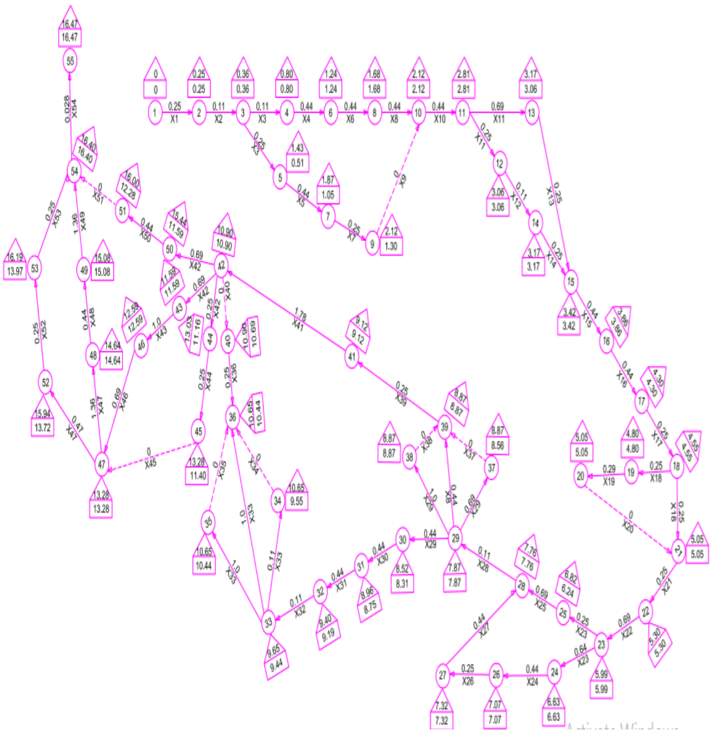
International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

48	Borehole	X47	16	14	16	18	16	0.666667	0.444444
49	Drainages	X48	13	10	13	17	13.166667	1.166667	1.361111
50	Fence	X42	24	21	24	26	23.833333	0.833333	0.694444
51	Land scaping	X50	13	11	13	15	13	0.666667	0.444444
52	Equipment and furnishing	X47	14	12	14	16	14	0.666667	0.444444
53	Testing	X52	5	3	5	6	4.8333333	0.5	0.25
54	Clean up	X53	8	7	8	10	8.1666667	0.5	0.25
55	Hand over	X54	2	1	2	2	1.8333333	0.166667	0.027778

From table 3 above, the mean, standard deviation and variance were increasing and reducing concurrently due to the effect of the three **variable estimates** compared to the CPM that has one variable estimate

FIGURE 2:

A PERT NETWORK DIAGRAM OF A STOREY BUILDING LECTURE CLASSES



The network shows arcs,nodes ,forward pass, backward pass of the variances and the critical paths.

The Critical Path **using PERT techniques were identified as**

1→2→3→4→6→8→10→11→12→14→15→16→17→18→19→20→21→22→23→24→26→27→28→29→38→39→41→42→43→46→47→48→49→54→55

And the variance of the critical path is

0.25+0.11+0.11+0.44+0.44+0.44+0.44+0.25+0.11+0.25+0.44+0.44+0.25+0.25+0.29+0.25+0.69+0.64+0.44+0.25+0.44+0.11+1.00+0+0+1.78+0.69+1.0+1.69+1.36+0.44+1.36+0.028

=16.47

The Z score was used to determine the probability of completing the project within 336 days:

$$z = \frac{x - \mu}{\delta}$$

z = (336 - 204.01) / 4.08

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

z = 32.35 / 4.08

z ≈ 7.94

δ = √16.47 =4.08

Assuming that the probability that the project completion time i.e ≤ 336

$$C = \frac{336-204.01}{4.08} = 32.35$$

P (e ≤ 336) = p (z ≤ 32.35) = 0.9554 from normal table

Therefore, based on the mean and variance of the critical path activities, the probability that the project completion time is less than or equal to 336 days is approximately 0.9994 or 96% chance of completing the project within 336 days indicating a high likelihood of completing the project within that time frame

4.3 MODIFY PERT USING USING ROBUST REGRESSION ANALYSIS BASED ON ITERATVE REWEIGHTED LEAST SQUARES ANALYSIS OF A STOREY BUILDING LECTURE CLASSES

4.3.1. MODIFIED PERT R ANALYSIS OF A STOREY BUILDING LECTURE CLASSES

Applying Linear Regression Analysis, we assessed the relationship between the duration of activities (To, Tm, Tp) and the overall project completion time in the construction process of a storey building using R Software, identifying any significant predictors of project duration.

The regression model

$$y = \alpha + \beta_1 t_o + \beta_2 t_m + \beta_3 t_p + \ell_i \qquad [18]$$

Table 4. Descriptive statistics of the three dependent variables To,Tm,Tp

Variables	Mean	Median	Variance	S.D
Duration in Days	11.16	9.00	36.80606	6.0668
To	9.255	7.000	31.41549	5.604952
Tm	11.16	9.00	36.80606	6.0668
Tp	13.04	11.00	42.18384	6.494909

Table 5: Table of OLS Estimation Result

Variables	**Coefficient**	**S.E**	**t-value**	**p-value**
Constant	9.590e-16	1.123e-166.778e	8.537e+00	2.13e-11 ***
TO	5.175e-16	6.778e-17	7.635e+00	5.42e-10 ***
TM	1.000e+00	9.211e-17	1.086e+16	<2e-16***
TP	-2.705e-17	5.173e-17	-5.230e-01	0.603
RSE	**2.766e-16**			

Source: Authors' computation aided by R package v 4.1.1.

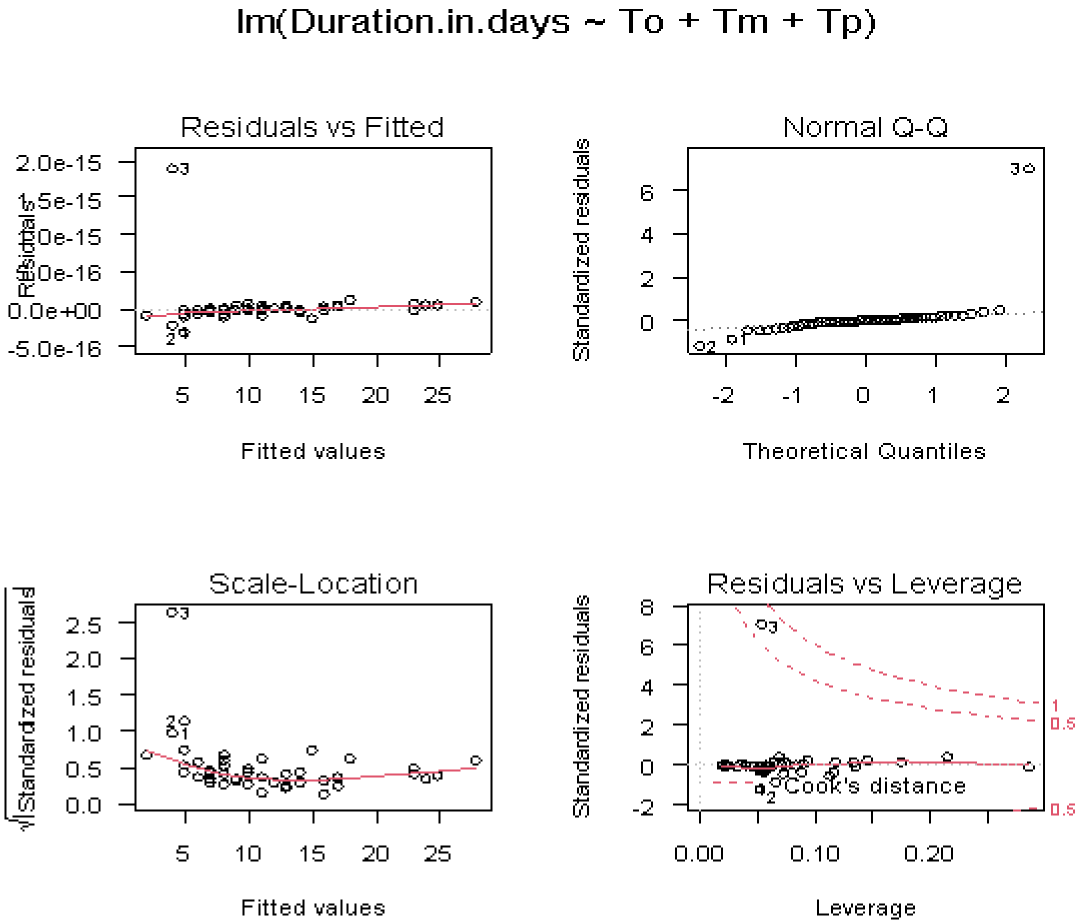
From Table 4.5 above, the result shows that To and Tm are statistically significant because the p values are all less than 0.05 while the Tp is not statistically significant since (p>0.05). Tm emerges as a strong predictor of project duration. This suggests that variations in the Tm duration of activities have a direct and proportional effect on the project's total duration. Therefore, minimizing delays or optimizing the duration of activities classified as Tm can lead to improvements in project timelines. This implies that variations in the optimistic and pessimistic durations of activities do not significantly influence the overall project completion time, holding other variables constant. While these durations contribute to estimating the expected duration of tasks, they may not directly impact the project's final completion time as much as the Tm duration.

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

4.4. Outlier Detection and Checking of some Classical Assumptions

Outliers and violations of distributional assumptions such as normality, linearity, homoscedasticity are common in many areas of research. These issues might introduce substantial bias in the analysis and potentially lead to grossly incorrect inferences. The ordinary least squares method is quite sensitive for outlying observations, both for dependent and independent variables and can have an adverse effect on the estimate. In higher dimensional data, these outlying observations can remain unnoticed. So, after running a regression analysis, you should check if the model works well for the data. One of the ways to check if the model works well is by plotting the residuals, it shows how poorly a model represents a data and could also reveal unexplained patterns in the data by the fitted model. Therefore, we are going to plot the residuals of the regression estimate from table 4.5 above to check for outliers, leverage points, influential observation and classical assumptions.

Figure 3. DIAGONISTIC TEST OF THE REGRESSION RESIDUALS



In the plotted graph above, the robust regression analysis based on iteratve reweighted least squareswas plotted and the result showed that there is linear relationship between the variables (To Tm Tp) and time of finishing the project. The variation around the estimated regression line is constant suggesting that the assumption of equal error variance is reasonable.

Normal Q-Q: was tested to check the normality of residuals. And the data set showed that the majority of the residuals followed the straight dashed line. We tested for scale location ,the data showed that the residuals spread randomly along the horizontal line with equal spread points to satisfy the assumption of homoscedasticity of residuals but in this case it showed that they are heteroscedasticity. we also tested for the residual vs. Leverage to identify outlier or influential value in our dataset and the data showed that there is an outlier. We plotted a diagonisict graph to determine the type of outlier in the cook distance.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 1, pp: (44-58), Month: April 2024 – September 2024, Available at: www.paperpublications.org

4.5. DIAGONISTIC PLOT TO DETERMINE OUTLIERS IN THE COOK DISTANCE

4.5.1. OUTLIER

Outliers can be identified by examining the standardized residual, which is residual divided by its estimated standard error. Observations whose standardized residuals are greater than 3 in absolute value are possible outliers.

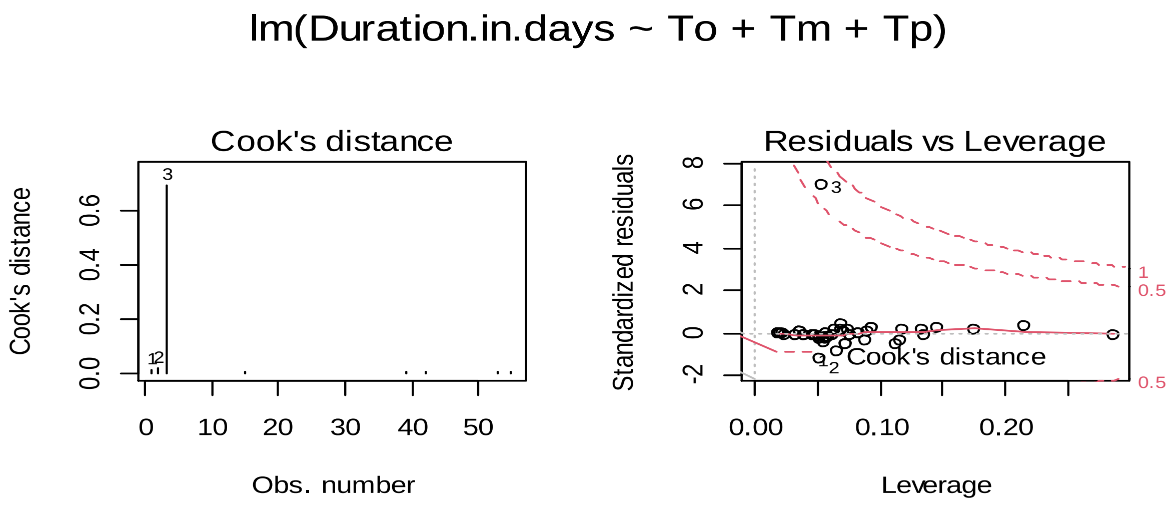


Figure 4: Cook’s distance and Residuals vs Leverage plots

Source: Authors’ computation aided by R package v 4.1.1.

4.5.2. HIGH LEVEARAGE POINTS

A data point has high leverage, if it has extreme predictor x values. This can be detected by exarmining the leverage statistic. A value of this statistic above $2(p+1)/n$ indicates an observation with high leverage where, p is the number of predictors and n is is the number of observations. So in our own case we have $2(3+1)/55 = 0.1$, now looking at the plot above we can see that observation 3 is grater than the leverage statistic 0.1. So we conclude that there is high leverage point on the explanatory variables. Therefore, a Robust regression is required for a better result.

Table 6: MM-Estimation Result

Variables	Coefficient	S.E	t-value	p-value
Constant	9.590e-16	1.123e-166.778e	8.537e+00	2.13e-11 ***
TO	4.175e-16	5.778e-17	7.635e+00	5.42e-10 ***
TM	1.000e+00	6.211e-17	1.086e+16	<2e-16***
TP	-2.705e-18	3.173e-17	-5.230e-01	0.0435*
RSE	**1.326e-17**			

Source: Authors’ computation aided by R package v 4.1.1.

From Table 6 above, the result shows that To, Tm and Tp are statistically significant because the p values are all less than 0.05 meaning (p<0.05). Tm emerges as a strong predictor of project duration. This suggests that variations in the Tm duration of activities have a direct and proportional effect on the project's total duration. Therefore, minimizing delays or optimizing the duration of activities classified as Tm can lead to improvements in project timelines. This implies that variations in the optimistic and pessimistic durations of activities do not significantly influence the overall project completion time, holding other variables constant. While these durations contribute to estimating the expected duration of tasks, they may not directly impact the project's final completion time as much as the Tm duration. Likewise, the application of MM-estimator improves the result significantly and also tackles the effect of outlier in the dat \

5. RECOMMENDATION AND CONCLUSION

Recommendation

Project managers can leverage this information to conduct risk assessments and contingency planning. By identifying activities with high variability, they can proactively manage risks and develop contingency plans to mitigate potential delays. This reviews the importance of considering both statistically significant and insignificant variables in project management, as they collectively contribute to a more comprehensive and robust project schedule analysis.

In the final analysis, the robust regression analysis based on iteratve reweighted least squaresunderscores the statistical significance of medium activity durations Tm in predicting the overall project completion time. The coefficient associated with Tm in the regression equation provides a quantifiable measure of its impact. For instance, a positive coefficient indicates that an increase in Tm would result in a proportional increase in the project duration, holding all other variables constant.

The optimization of Tm durations is thus a key factor in enhancing project efficiency. By minimizing the Tm durations while maintaining the quality of work, project managers can effectively reduce the overall project timeline, leading to fewer delays.

However, it’s equally important to manage project risks related to optimistic (To) and pessimistic (Tp) durations. While these variables may not have a statistically significant impact on the project duration, their role in estimating the range of possible activity durations is crucial. This range, represented by the spread between To and Tp, provides a measure of uncertainty or risk associated with each activity.

Project managers can use this information to conduct risk assessments and contingency planning. By identifying activities with a wide range of possible durations (high To - Tp spread), they can proactively manage risks and develop strategies to mitigate potential delays.

In conclusion, through the statistical lens of robust regression analysis based on iteratve reweighted least squares, we see that the optimization of Tm durations and the effective management of project risks related to To and Tp are pivotal in enhancing project efficiency, reducing delays, minimizing outliers and improving overall project outcomes.

Conclusion

Integrating robust regression analysis based on iterative reweighted least squares and networking practices into construction project management enables stakeholders to enhance decision-making, optimize resource allocation, mitigate risks and minimize outliers ultimately contributing to the successful and timely delivery of construction projects. By leveraging these methodologies, project managers can navigate the complexities of construction projects more effectively, ensuring adherence to timelines, budgets, and quality standards. Embracing innovation in project management methodologies is essential for addressing the evolving challenges of the construction industry and driving sustainable growth and development.

It is now clear that, through the statistical lens of robust regression analysis based on iteratve reweighted least squares, we see that the optimization of Tm durations and the effective management of project risks related to To and Tp are pivotal in enhancing project efficiency, reducing delays, minimize outliers and improving overall project outcomes.

Limitations

Limitations of the regression analysis, including data constraints, model assumptions, and generalizability, will be acknowledged.

REFERENCES

[1] Abdel-Hamid, T. K., & Sherif, Y. S. (2015). *Project Management for Engineering and Construction* (3rd ed.). Springer.

[2] Al-Hajj, A., & Zraunig, M. M. (2018). The Impact of Project Management Implementation on the Successful Completion of Projects in Construction. International Journal of Innovation, Management and Technology, 9(1), 21–27. https://doi.org/10.18178/ijimt.2018.9.1.781

[3] Almetwally, E. M. and Almongy, H. M.,Comparison Between M-estimation, S estimation, and MM Estimation Methods of Robust Estimation with Application and Simulation, International Journal of Mathematical Archive-9 (11), 2018, 55-63

[4] Alpu, O. and Samkar, H. (2010). Liu Estimator based on an M Estimator. Turkiye Klinikleri Biostat, 2(2):49-53

[5] Altoryman, 2014; González et al., 2014; Hussein and Klakegg, 2014; (Hussein and Klakegg, 2014 Analysis, 104, 88-100 and Shrinkage. Electronic Journal of Statistics, 3, 384-415.

[6] Atin, S., & Lubis, R. (2019). Implementation of Critical Path Method in Project Planning and Scheduling. IOP Conference Series: Materials Science and Engineering, 662(2), 1–6. https://doi.org/10.1088/1757899X/662/2/022031

[7] Badruzzaman, F. H., Fajar, M. Y., Rohaeni, O., Gunawan, G., & Harahap, E. (2020). CPM and PERT Technique Efficiency Model For Child Veil Production. International Journal of Scientific and Technology Research, 9(4), 1470–1476.

[8] Bai,X.,Yao,W.,Boyer,J.E.(2012).Robustfittingofmixtureregressionmodels.ComputationalStatistics andDataAnalysis56:2347–2359.

[9] Berjis, N., Shirouyehzad, H., & Jouzdani, J. (2020). A New Approach To Determine The Weights Of Project Activities Using Data Envelopment Analysis: A Case Of Mobarakeh Steel Company. International Journal of Managing Projects in Business, 13(6), 1187–1217. https://doi.org/10.1108/IJMPB-12-2019-0304

[10] Calp MH, Akcayol MA (2017). Optimization of project scheduling activities in dynamic CPM and PERT networks using genetic algorithms. Süleyman Demirel University Journal of Natural and Applied Sciences.;22(2):615-627.

[11] Candraningtyas, Sherli., dkk. 2013. Regresi Robust MM-Estimator Untuk Penanganan Pencilan Pada Regresi Linear Berganda. Jurnal Gaussian Vol.2 No.4 Tahun 2013, Hal. 395-404. UNDIP.

[12] Chen, C. 2002. Robust Regression and Outlier Detection with the ROBUSTREG Procedure. Paper 265-27. North Carolina: SAS Institute.

[13] Chun Yua and Weixin Yaob a School o Statistics ,Jiangxi University of Finance and Economics, Nanchang, China; b Department of Statistics, University of California, Riverside ,California, USA CommunicationsinStatistics-SimulationandComputation41:1282–1300.

[14] Cristóbal, J. R. S. (2017). The S-Curve Envelope as a Tool for Monitoring and Control of Projects. Procedia Computer Science, 121, 756–761. https://doi.org/10.1016/j.procs.2017.11.097

[15] Department of Mathematics, College of Science, Mustansiriyah University, Baghdad, Iraq mmaahhaa938@gmail. com; hudamath@uomustansiriyah.edu.iq

[16] Fleming, Q. W., & Koppelman, J. M. (2016). *Earned Value Project Management* (4th ed.). Project Management Institute.

[17] Göksu A, Catovic S. (2012). Implementation of critical path method and project evaluation and review technique. 3rd International Symposium on Sustainable Development, Sarajevo. 2012;205-212.

[18] Gomes, J., & Romão, M. (2016). Improving Project Success: A Case Study Using Benefits and Project Management. Procedia Computer Science, 100, 489–497. https://doi.org/10.1016/j.procs.2016.09.187

[19] Heizer, J.; Render, B. & Munson (2017) Production & Operations Management: strategic and tactical decisions. 4th Edition. USA, New Jersey: Prentice Hall.

[20] Heritier, S., Cantoni, E., Copt, S. and Victoria-Feser, M.-P. (2009). Robust Methods in Biostatistics, Wiley Series in Probability and Statistics, John Wiley & Sons, New York.

[21] Hussein, Y. A. and Abdalla, A. A. (2012). Generalized Two stages Ridge Regression Estimator for Multicollinearity and Autocorrelated errors. Canadian Journal on Science and Engineering Mathematics, 3(3), 79-85.

[22] International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS) Vol. 2, Issue 3, pp: (45-54), Month: July 2015 - September 2015, Available at: www.paperpublications.org Page | 45 Paper Publications

[23] J ay L. Devore, Kenneth N. Berk, Matthew A. Carlton (2021). Overview and Descriptive Statistics. Modern Mathematical Statistics with Applications,ISBN: 978-3-030-55155-1

[24] Jack L, Okeke O.C, Okechukwu SI, Akinola AO. (2016) Project management: A system approach to planning, implementation, monitoring and evaluation. International Journal of Advanced Academic Research: Social & Management Sciences. 2(11):65-79.

[25] Jahufer, A. (2013). Detecting Global Influential Observations in Liu Regression Model. Open Journal of Statistics, 3, 5-11

[26] James R. Knaub (2012). Properties of Weighted Least Squares Regression for Cut off Sampling in Establishment Surveys. Energy Information Administration.

[27] Kerzner, H. (2017). *Project Management: A Systems Approach to Planning, Scheduling, and Controlling* (12th ed.). Wiley.

[28] Koller, M. and Stahel, W. A. (2011). Sharpening wald-type inference in robust regression for small samples, Computational Statistics & Data Analysis 55(8): 2504–2515

[29] Konior, Jaroslaw, & Szóstak, M. (2020). Methodology of Planning the Course of the Cumulative Cost Curve in Construction Projects. Sustainability (Switzerland), 12(6), 1–22. https://doi.org/10.3390/su12062347

[30] Konior, Jarosław, & Szóstak, M. (2020). The S-Curve as a Tool for Planning and Controlling of Construction Process-Case Study. Applied Sciences (Switzerland), 10(2071), 1–14. https://doi.org/10.3390/app10062071

[31] Lermen, F. H., Morais, M. de F., Matos, C., Röder, R., & Röder, C. (2016). Optimization of Times and Costs of Project of Horizontal Laminator Production Using PERT/CPM Technical. Independent Journal of Management & Production, 7(3), 833–853. https://doi.org/10.14807/ijmp.v7i3.423

[32] Lucko, G., & Kljajic, M. (2017). Utilization of PERT and CPM methods in project management

[33] Lukman, A. F, Arowolo, O. and Ayinde, K. (2014). Some Robust Ridge Regession for Handling Multicollinearity and Outliers. International Journal of Sciences: Basic and Applied Research (IJSBAR). 16(2), 192-202.

[34] Lukman, A. F. (2015). Review and Classifications of the Ridge Parameter Estimation Techniques. Unpublished M.phil./ PhD thesis..

[35] M.A. and KALAI R. (Eds.), *Robustness in OR-DA*, Annales du LAMSADE no. 7,

[36] M.A. and KALAI R. (Eds.), *Robustness in OR-DA*, Annales du LAMSADE no. 7,

[37] M.A. and KALAI R. (Eds.), *Robustness in OR-DA*, Annales du LAMSADE no. 7,

[38] Maha, I. l; Huda A. R. (2021). Robust Regression Methods / a Comparison Study

[39] Maronna, R., Martin, D. and Yohai, V. (2006). Robust Statistics: Theory and Methods, Wiley Series in Probability and Statistics, John Wiley & Sons.

[40] Maronna,R.A.,Martin,R.D.,Yohai,V.J.(2006).RobustStatistics.NewYork:JohnWiley.

[41] Nandal, N., & Nandal, N. (2019). BSCQUAL: A Measuring Instrument of Service Quality for

[42] Naranjo, J. D., Hettmansperger, T. P. (1994). Bounded influence rank regression. Journal of the Royal Statistical Society, Series B56:209–220.

[43] O'Brien, J. J., & Plotnick, F. L. (1968). Critical path planning and scheduling: Mathematical basis. *Operations Research*, 16(1), 101-122.

[44] Olive, D. J., Hawkins, D. M. (2011). Practical High Breakdown Regression. Available at: http://lagrange. math.siu.edu/Olive/pphbreg.pdf.

[45] Park, Y., Kim, D., Kim, S. (2012). Robust regression using data partitioning and M-estimation.

[46] Pastiarsa, M. (2015). Manajemen Proyek Konstruksi Bangunan Industri : Perspektif Pemilik Proyek (Cet. 1). Teknosain

[47] Pinto, J. K., & Slevin, D. P. (1987). Critical factors in successful project implementation. *IEEE Transactions on Engineering Management*, (4), 22-27.

[48] Radujković, M., & Sjekavica, M. (2017). Project Management Success Factors. Procedia Engineering, 196, 607–615. https://doi.org/10.1016/j.proeng.2017.08.048

[49] Ritzman, L. J.; Krajewski, L. P. (2012) Operations Management: strategy and analysis. 5th Edition. USA: Addison-Wesley Publishing Company.

[50] Robbins, S. P., & Coulter, M. (2012). Management (Eleventh E). Prentice Hall.

[51] Robust linear regression : A review and comparison

[52] Robust Regression Methods / a Comparison Study

[53] Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M. and Maechler, M. (2011). robustbase: Basic Robust Statistics. R package version 0.7-0

[54] Ruckstuhl, A. F., Henne, S., Reimann, S., Steinbacher, . M., Vollmer, M. K., O'Doherty, S., Buchmann, B. and Hueglin, C. (2012). Robust extraction of baseline signal of atmospheric trace species using local regression, Atmospheric Measurement Techniques 5(11): 2613–2624. URL: http://www.atmos-meas-tech.net/5/2613/2012/

[55] Schuyler, J. R., & Muckstadt, J. A. (1976). The critical path method: A solution to the management of complex research and development projects. *Operations Research*, 24(6), 1017-1037.

[56] Shafiq, M., Amir, W. M., & Zafakali, N. S. (2017). JMASM 46: Algorithm for Comparison of Robust Regression Methods In Multiple Linear Regression By Weighting Least Square Regression (SAS). Journal of Modern Applied Statistical Methods, 16(2), 490-505. doi: 10.22237/jmasm/1509496020

[57] She, Y. (2009). Thresholding-Based Iterative Selection Procedures for Model Selection

[58] She, Y. and Owen, A. B. (2011). Outlier Detection Using Nonconvex Penalized regression. Journal of American Statistical Association, 106, 626-639

[59] Sugiyono. (2013). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. CV Alfabeta.

[60] Sulaimon Mutiu O. (2015). Application of Weighted Least Squares Regression in Forecasting the B-Schools. International Journal of Psychosocial Rehabilitation, Vol. 23, Issue 04, 1574-1589 Universite Paris-Dauphine, 2007.

[61] Wang, Q. and Yao, W. (2012). An adaptive estimation of MAVE. Journal of Multivariate

[62] Yao,W.,Li,L.(2014).A New Regression Model: Modal linear regression. Scandinavian Journal of Statistics 41:656–671.

[63] Yao,W.,Lindsay,B.G.,Li,R.(2012).Localmodalregression.JournalofNonparametricStatistics24:647– 663.

[64] Yao,W.,Zhao,Z.(2013).Kerneldensitybasedlinearregressionestimates.CommunicationsinStatisticsTheoryandMethods 42:4499–4512.

[65] Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. The AnnalsofStatistics 15:642–656.

[66] You,J.(1999).AMonteCarlocomparisonofseveralhighbreakdownandefficientestimators.ComputationalStatisticsandDataAnalysis30:205–219.

[67] Yu,C.,Chen,K.,Yao,W.(2015).Outlierdetectionandrobustmixturemodelingusingnonconvexpenalizedlikelihood.JournalofStatisticalPlanningandInference164:27–38.

[68] Yuan, A., De Gooijer, J. G. (2007). Semiparametric regression with kernel error model. Scandinavian JournalofStatistics 34:841–869.

[69] Zareei, S. (2018). Project Scheduling for Constructing Biogas Plant using Critical Path Method. Renewable and Sustainable Energy Reviews, 81, 756–759. https://doi.org/10.1016/j.rser.2017.08.025