International Journal of Engineering and Technology
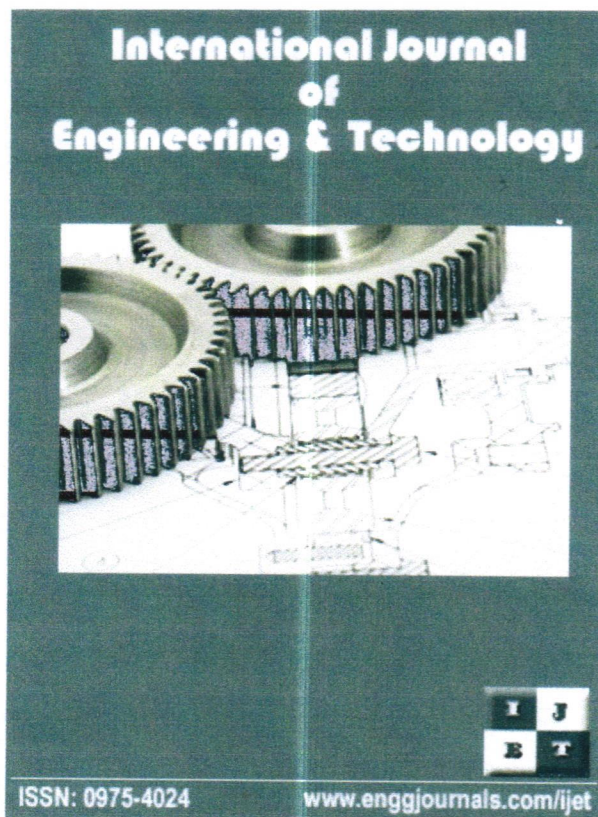
Home

IJET Topics

Call for Papers 2020

Author Guidelines

Special Issue

Current Issue

Articles in Press

Archives

Editorial Board

Reviewer List

Publication Ethics and Malpractice statement

Authors Publication Ethics

Policy of screening for plagiarism

Open Access Statement

Terms and Conditions

Contact Us

Aims and Scope

International Journal of Engineering and Technology (IJET) is a scholarly open access, peer-reviewed, interdisciplinary, quarterly and fully refereed journal focusing on theories, methods and applications in Engineering and Technology. IJET covers all areas of Engineering and Technology, publishing refereed original research articles and technical notes. IJET reviews papers approximately two to four months of submission date and publishes accepted articles on the forthcoming issue upon receiving the final versions.

All submitted articles should report original, previously unpublished research results, experimental or theoretical, and will be peer-reviewed. Articles submitted to the journal should meet these criteria and must not be under consideration for publication elsewhere. Manuscripts should follow the style of the journal and are subject to both review and editing.

All the papers in the journal are also available freely with online full-text content and permanent worldwide web link. The abstracts will be indexed and available at major academic databases.

IJET welcomes author submission of papers concerning any branch of the Engineering and Technology and their applications in business, industry and other subjects. The subjects covered by the journal includes but not limited to:

Aeronautical/ Aerospace Engineering, Agricultural Engineering, Automation Engineering, Automobile Engineering, Ceramic Engineering, Chemical Engineering, Civil Engineering, Collaborative Engineering, Communication Engineering, Complexity in Applied Science and Engineering, Computational Science and Engineering, Computer Aided Engineering and Technology, Computer Applications in Technology, Computer Engineering, Continuing Engineering Education and Life-Long Learning, Control Theory, Data Mining and Bioinformatics, Design Engineering, Electrical Engineering, Electronic and Electrical Engineering, Embedded Systems, Engineering Management and Economics, Environmental Engineering, Forensic Engineering, Industrial Engineering, Information Systems and Management, Information Technology, Instrumentation Engineering, Intelligent Engineering Informatics, Knowledge Engineering and Data Mining, Manufacturing Engineering, Marine Engineering, Materials Engineering, Mechanical Engineering, Mechatronics, Metallurgical Engineering, Microengineering, Mining Engineering, Nuclear Engineering, Petroleum Engineering, Process Systems Engineering, Production Engineering, Remote Monitoring, Sensor Network, Soft-computing and Engineering Education, Software Engineering, Strategic Engineering Asset Management, Structural Engineering, Textile Engineering, Transportation.

Frequency: 6 issues per year

e-ISSN: 0975-4024 (online Version)

p-ISSN: 2319-8613 (Print Version)

Subject Category: Engineering and Technology

Published by: Engg Journals Publications

International Journal of Engineering & Technology

ISSN: 0975-4024   www.enggjournals.com/ijet

I J E T

Home   Journal Rankings   Country Rankings   VIZ Tools   Help   About Us

International Journal of Engineering and Technology

Country   India - ⅢⅢ SIR Ranking of India

Subject Area and   Engineering
Category   Engineering (miscellaneous)

Publisher   Engg Journals Publications

Publication type   Journals

ISSN   23198613, 09754024

Coverage   2009-2016 (cancelled)

Scope   International Journal of Engineering and Technology (IJET) is a scholarly open access,
                       peer-reviewed, interdisciplinary, quarterly and fully refereed journal focusing on theories, methods
                       and applications in Engineering and Technology. IJET covers all areas of Engineering and
                       Technology, publishing refereed original research articles and technical notes. IJET reviews
                       papers approximately two to four months of submission date and publishes accepted articles on
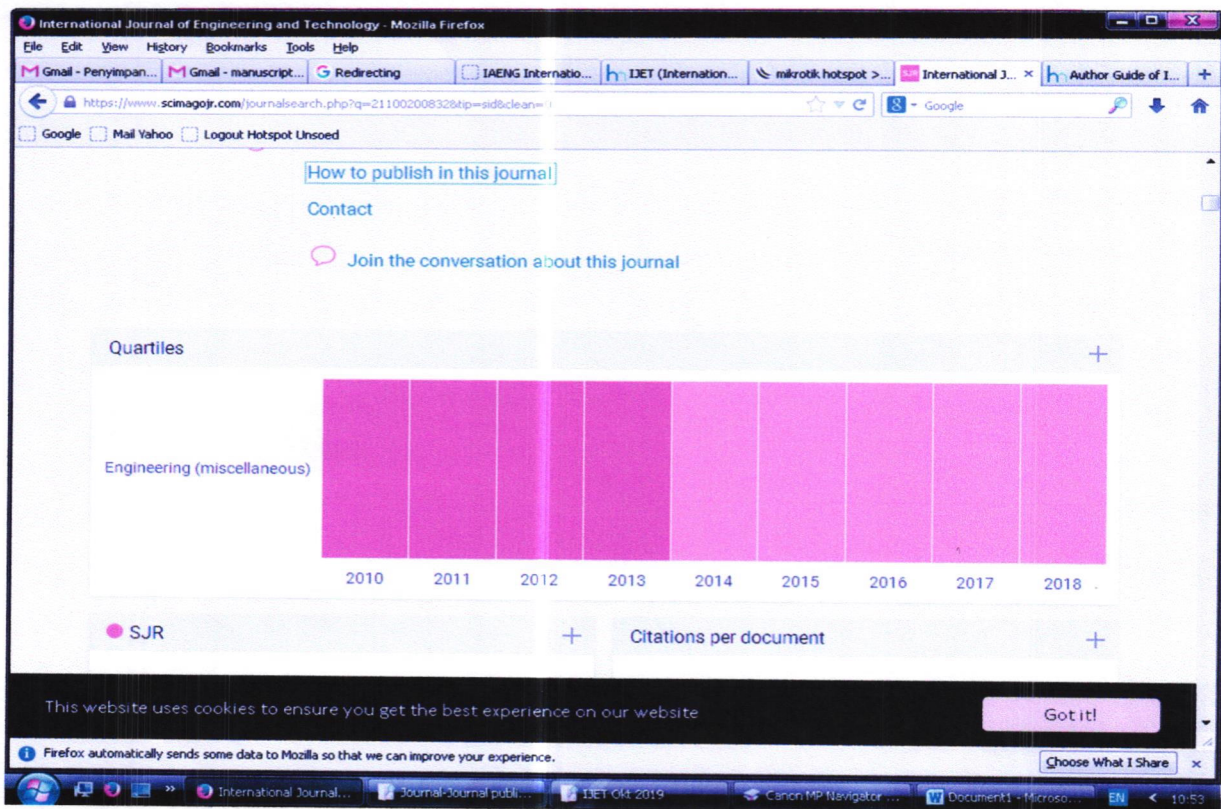
20

H Index

IJET Editorial Board

Lucknow-226002, UP INDIA

VELAYUTHAM .P
Associate Professor, Adhiparasakthi Engineering College, Melmaruvathur

M. DANWE Raïdandi
Head, The College of Technology, University of DOUALA, Cameroon.

VIVEK S. DESHPANDE
Assistant Professor, MIT College of Engineering, Pune, India

Umbarkar Anantkumar Janardan
Assistant Professor, Walchand College of Engineering, SANGLI.

Jifeng Wang
University of Illinois at Urbana-Champaign

ANUJ K. GUPTA
Associate Prof., RIMT - Institute of Engineering & Technology, Mandi
Gobindgarh

SUKUMAR SENTHILKUMAR
Post Doctoral Fellow, Universiti Sains Malaysia, Pulau Pinang-11800
[Penang], Malaysia.

Maniya Kalpesh Dudabhai
Lecturer/ Asst. Professor, C.K.Pithawalla college of Engg.& Tech.,Surat

Harish Kumar
Scientist in Force & Hardness Standard Group of National Physical
Laboratory, New Delhi

P.VELAYUTHAM
Assistant Professor, Adhiparasakthi Engineering College, Melmaruvathur

Prasenjit Chatterjee
Assistant Professor, MCKV Institute of Engineering, a Co-ed P.G. Degree
Engineering College, Liluah, Howrah, West Bengal

Mariya Petrova Aleksandrova
Assistant Professor and Department Researcher, Technical University of
Sofia, Faculty of Electronic Engineering and Technologies, Sofia, Kliment
Ohridski blv, 8

Süleyman KORKUT
Associate Professor, Duzce University, Duzce-Turkey.

Sina Kazemian
Senior Geotechnical Engineer, SCG Consultant Company, Selangor,
Malaysia

U C srivastava
Asst.Professor, Amity University, Noida, U.P-201301-India

Ch.V.Subbarao
Assistant Professor, University of Petroleum and Energy studies-
Rajahmundry.

Antonella Petrillo
Contract Researcher in Industrial Plant with specific skills in Multicriteria
Decision Analysis and Decision Support System, University of Cassino

Dr. Debasish Basak
Senior Principal Scientist, Central Institute of Mining and Fuel Research,
Barwa Road, Dhanbad – 826015.

FAHIMUDDIN.SHAIK

Assistant Professor, Annamacharya Institute of Technology & Sciences, Rajampet, A.P.

DR. DEWAN MUHAMMAD NURUZZAMAN
Professor, Dhaka University of Engineering and Technology (DUET), Gazipur, Gazipur-1700, Bangladesh.

Dr V S GIRIDHAR AKULA
Professor and Principal At Vemu Inst of Tech, P Kotha Kota Chittor dist

Prasenjit Chatterjee
Assistant Professor, MCKV Institute of Engineering, a Co-ed P.G. Degree Engineering College, Liluah, Howrah, West Bengal

Amin Paykani
Lecturer in Goldis English Institute, Tabriz.

Dr. Sameer Mahadeorao Wagh
Assistant Professor, Laxminarayan Institute of Technology, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur

Dhahri Amel
Research professor, Gafsa University-Faculty of Sciences

Dr.Waleed Khalil Ahmed
Instructor, United Arab Emirates University

DR. RAJNEESH KUMAR GUJRAL
Assoc. Professor, M.M. Engineering College, M.M university Mullana, Ambala-133 203, Haryana, India"

Ing. Norbert Herencsar
Contact person of the LLP Erasmus Bilateral Agreement, Czech Republic and Karadeniz Technical University, Trabzon, Turkey.

Dr.R.SATHISHKUMAR
Professor and Head, , Kodaikanal Institute of Technology, Dindigul Dist, India

Dr.V.Vaithiyanathan
Associate Dean – Research, School of Computing, SASTRA University, Thanjavur 613 401

SEDDIK BRI
Professor – Researcher, Materials and Instrumentations (MIN), Electrical Engineering Department, High School of Technology, Meknes - Morocco

Dr. Saurabh Pal
Sr. Lecturer, VBS Purvanchal Univ. Jaunpur

R.MANIKANDAN
Senior Assistant Professor, SASTRA University, Thanjavur, India

Professor.K.S.Zakiuddin
Dean Academics (PG), Professor & Head, Dept of Mech Engg,Nagpur.

K.V.L.N.ACHARYULU
Associate Professor, Bapatla Engineering College,Bapatla.

Dr. R. S. SHAJI
Professor at Noorul Islam University, Kumaracoil.

Dr. Roli Pradhan
Assistant Professor, Maulana Azad National Institute of Technology

Sunny Narayan
Assistant professor & Departmental coordinator, Indus University, India.

Dr. A. SENTHIL KUMAR
Professor, Velammal Engineering College, Chennai

Dr. C. LAKSHMI DEVASENA
Assistant Professor, IBS, Hyderabad

Terms and Conditions of International Journal of Engineering and Technology (IJET)

Terms and Conditions

The International Journal of Engineering and Technology (IJET) is web site ("Site") is registered with the National Science Library (NSL), ISSN(online) 0975-4024 ISSN(Print) 2319-8613 ("IJET," "we," "us" or "our").

By accessing or using the Site, you agree to be bound by the terms and conditions below ("Terms and Conditions"). If you do not agree with these Terms and Conditions, please do not use this Site.

IJET reserves the right to change, modify, add or remove portions of these Terms and Conditions in its sole discretion at any time and without prior notice. Please check this page periodically for any modifications. Your continued use of this Site following the posting of any changes will mean that you have accepted the changes.

Copyrights and Limitations on Use
All content in this Site, including site layout, design, images, programs, text and other information (collectively, the "Content") is the property of IJET and its affiliated companies or licensors and is protected by copyright and other intellectual property laws.

You may not copy, display, distribute, modify, publish, reproduce, store, transmit, create derivative works from, or sell or license all or any part of the Content, products or services obtained from this Site in any medium to anyone, except as otherwise expressly permitted under applicable law or as described in these Terms and Conditions or relevant license or subscriber agreement.

You may print or download Content from the Site for your own personal, non-commercial use, provided that you keep intact all copyright and other proprietary notices. You may not engage in systematic retrieval of Content from the Site to create or compile, directly or indirectly, a collection, compilation, database or directory without prior written permission from IJET.

You agree that you will not use any robots, spiders, crawlers or other automated downloading programs or devices to: (i) continuously and automatically search or index any Content, unless authorized by us; (ii) harvest personal information from the Site for purposes of sending unsolicited or unauthorized material; or (iii) cause disruption to the working of the Site. If this Site contains robot exclusion files or robot exclusion headers, you agree to honor them. Any questions about whether a particular use is authorized and any requests for permission to publish, reproduce, distribute, display or make derivative works from any Content should be directed to IJET.

You may not use the services on the Site to publish or distribute any information (including software or other content) which is illegal, which violates or infringes upon the rights of any

other person, which is abusive, hateful, profane, pornographic, threatening or vulgar, which contains errors, viruses or other harmful components, or which is otherwise actionable at law. IJET may at any time exercise editorial control over the content of any information or material that is submitted or distributed through its facilities and/or services.

Publication Fees Refund Policy
We do not provide any refund under any circumstance for any manuscript published online in our journal.

Indexing in Databases
IJET provide all the data at free of charge to major indexing databases like Google Scholar, DOAJ, EBSCO, **SCOPUS, Thomson Reuter etc**. We can't guarantee that the data provided by us will get indexed in their databases. It's upto the individual database provider's decision to include or reject the data provided by us.

Trademarks
All trademarks appearing on this Site are the property of their respective owners.

Links to Other Sites
The Site may contain hyperlinks to other sites or resources that are provided solely for your convenience. IJET is not responsible for the availability of external sites or resources linked to the Site, and does not endorse and is not responsible or liable for any content, advertising, products or other materials on or available from such sites or resources. Transactions that occur between you and any third party are strictly between you and the third party and are not the responsibility of IJET. Because IJET is not responsible for the availability or accuracy of these outside resources or their contents, you should review the terms and conditions and privacy policies of these linked sites, as their policies may differ from ours.

Disclaimer of Warranties and Liability
Neither IJET, its affiliates, nor any third-party content providers or licensors makes any warranty whatsoever, including without limitation, that the operation of the Site will be uninterrupted or error-free; that defects will be corrected; that this Site, including the server that makes it available, is free of viruses or other harmful components; as to the results that may be obtained from use of the Content or other materials on the Site; or as to the accuracy, completeness, reliability, availability, suitability, quality, non-infringement or operation of any Content, product or service provided on or accessible from the Site.

THIS SITE AND ALL CONTENT, PRODUCTS AND SERVICES INCLUDED IN OR ACCESSIBLE FROM THIS SITE ARE PROVIDED "AS IS" AND WITHOUT WARRANTIES OR REPRESENTATIONS OF ANY KIND (EXPRESS, IMPLIED AND STATUTORY, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF TITLE AND NONINFRINGEMENT AND THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE), ALL OF WHICH IJET DISCLAIMS TO THE FULLEST EXTENT PERMITTED BY LAW. YOUR USE OF THE SITE IS AT YOUR SOLE RISK.

NEITHER IJET NOR ANY OF ITS AFFILIATES OR LICENSORS SHALL BE LIABLE TO YOU OR ANYONE ELSE FOR ANY LOSS OR INJURY, CAUSED IN WHOLE OR PART BY ITS NEGLIGENCE OR CONTINGENCIES BEYOND ITS CONTROL IN PROCURING, COMPILING, INTERPRETING, REPORTING OR DELIVERING INFORMATION THROUGH THE SITE. IN NO EVENT WILL IJET, ITS AFFILIATES OR LICENSORS BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DECISION MADE OR ACTION TAKEN BY YOU IN RELIANCE ON SUCH INFORMATION. IJET AND ITS AFFILIATES AND LICENSORS SHALL NOT BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DAMAGES (INCLUDING, WITHOUT LIMITATION, CONSEQUENTIAL, SPECIAL, INCIDENTAL, INDIRECT, OR SIMILAR DAMAGES) EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

IJET neither endorses nor takes responsibility for any products, goods or services offered by outside vendors through our services or advertised on our system.

Systems Reliability

IJET aims to keep the Site available twenty-four (24) hours a day, seven (7) days a week and to maintain saved information. However, due to technical failures, acts of God or routine maintenance, availability may be limited and/or information may be lost. IJET shall not be liable for lost information or non-availability of the services.

Indemnification

You hereby agree to indemnify, defend and hold IJET, its directors, officers, shareholders, parents, subsidiaries, affiliates, agents and licensors harmless from and against any and all liability, losses, damages and costs, including, without limitation, reasonable attorneys' fees, arising from your use of the Site or Content.

Governing Law and Venue

These terms and conditions shall be governed by and construed in accordance with the laws of India, without regard to its conflicts of law principles. You hereby submit to and agree that the sole jurisdiction and venue for any actions that may arise under or in relation to the subject matter hereof shall be the courts located in Chennai, India.

A BEST MODEL ON MULTIPLE LINEAR REGRESSION

B. Pratikno [#1], I.P. Sulaeman [#2], D. Sopanti [#3] and Supriyono [#4]

 # First-Third
Department of Mathematics
Faculty of Mathematics and Natural Science
Jenderal Soedirman University, Purwokerto, Indonesia.
*Second Company
Fourth
Open University, UPBJJ Purwokerto, Indonesia
[1] bpratikto@gmail.com
[2] indahpns7@gmail.com
[3] diahsopanti.ds@gmail.com
[4] supriyono@ecampus.ut.ac.id

Abstract—We study the selection variables (predictors) on multiple linear regression model (MLRM). The best subset (C_p Mallow) and the stepwise (forward selection and backward elimination) methods are used to identify a best model on the MLRM. A simulation study is conducted using secondary data on Y *(percentage of poverty line) and four predictors X_1 (jobless), X_2 (population growth rate), X_3 (life expectancy), and X_4 (length of study)*. The result showed that the best model of the forward selection and backward elimination are similar, X_1 and X_3, are significant, with the model $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3$, but the forward selection (with 2 steps) is more efficient than backward elimination (with 3 steps). Unfortunately, the best subset method has a different selection variables, here X_2 is also significant, so the model is $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$. Finally, we conclude that the best model is $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3$. This is due to the both variables X_1 and X_3 are available on the three tests, namely (1) correlation testing (partial test), (2) stepwise method, and (3) best subset method.

Keyword - Best subset, Cp, MLRM, and stepwise.

I. INTRODUCTION

Generally, inferences about population parameter can be drawn from sample, and increasing a number of sample will affect significantly to the quality of inferences (Soejoeti, 2010). More detail about this, Bancrof [21] already studied to improve the inferences population using non-sample prior information (NSPI) from trusted sources. Therefore, we need sampling to get the eligible and representative sample (n) from population (N), with small error on level of significance (α), $\alpha = 0.01, 0.05$ and or 0.10. Following, Bhatacharya and Johnson [12], Walpole and Myers [19] and Bluman [1], the sample size (n) is then formulated by

$$n \geq \left(\frac{Z_{\alpha/2} \sigma}{d} \right)^2,$$ where d is determined by researcher, $Z_{\alpha/2}$ is from standard normal distribution, and σ^2

is unknown variance. Furthermore, from the pair of data sample $(x_i, y_i), i = 1, 2, \cdots, n$, we then estimate the linier regression model of the multiple regression model (MLRM) as $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p$, where x_i is a predictor and y_i is a response. Note that the general model of the MLRM is given as $y_i = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + e_i$. To ensure that the model is significant, we then test the significant model of the $\hat{y}_i$ to be recommended to users. Moreover, the model ($\hat{y}_i$) is generally tested using three criteria, namely (1) the coefficient determination (R^2), (2) analysis of variance (Anova) with F test, and or (3) partial (Draper and Smith, [21]). In this case, some assumptions of the regression model are also tested, such as (1) normality test and random error of $e_i \sim N(0, \sigma^2)$, (2) autocorrelation of the error, (3) heteroscedasticity and (4) multicollinearity problem among predictors.

In addition, the paper is focused on the best subset and stepwise methods in getting the best model on the MLRM. Here, we noted many authors have been studied about the selection variables on regression model, such as Draper and Smith [21], Gujarati [7], Gujarati [8], Ghozali [13], Kutner et al.[15], Montgomery [5], Montgomery [6], Bhattacharyya and Johnson [12], Greene [22], Baltagi [2], Baltagi [3], Mendenhall [23],

Mendenhall and Sincich [24], Bluman [1], Huggins and Staudte [20], Iriawan and Astuti [16], Graybil and Iyer [10], Walpole and Myers [18], Walpole et al. [19] and Pratikno [4]. Here, Pratikno [4] have been studied testing of the hypothesis tests on regression models using power and size of the hypothesis testing.

In this paper, the introduction is presented in Section 1. The regression model, best subset and stepwise methods are given in Section 2. A simulation is then obtained in Section 3. Section 4 described the conclusion of the research. document is a template.

II. REGRESSION MODEL, BEST SUBSET AND STEPWISE METHODS

2.1. Multiple Regression Model

For an n pair of observations on p independent variables $(X_1,...,X_p)$ and one dependent variable (Y), (X_{ij}, Y_i), for $i=1,2,...,n$ and $j = 1,2,...,p$, the multiple regression model (in matrix) is given by

$$Y = X\beta + e \tag{1}$$

Here, $\beta = (\beta_0, \beta_1, ..., \beta_p)'$ is a $(p+1)$ dimensional column vector of unknown regression parameters, $Y = (y_1,...,y_n)'$ is $(n \times 1)$ vector of response variables, X is a $n \times (p+1)$ matrix of know fixed values of the independent variables an e is the error term which is assumed to be identically and independently distributed as $N_n(0, \sigma^2 I_n)$. Here, I_n is the identity matrix of order n and σ^2 is the common variance of the error variables. Following Montgomery [5] and Graybil and Iyer [10], the estimate coefficient regression model in matrix is then given as

$$\hat{\beta} = (X^t X)^{-1} X^t Y \tag{2}$$

Furthermore, we test the hypothesis testing of the, $H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0$. Here, Anova *(F test)* is used, and it is presented in Table 1.

Table 1. Anova of the MLRM with p independent variables

Source of Variance	Degree of freedom (df)	Sum Square (SS)	Mean Square (MS)	F*
Regression	p	SSR	MSR	$F^* = \dfrac{MSR}{MSE} > F_{\alpha;p;n-(p+1)}$ W
Error	$n-(p+1)$	SSE	$MSE = s^2$	e then reject H_0
Total	$n-1$	SST		

Note that *SSE* is sum square error, *SST* is sum square total and *MSE* is mean square error. Following Montgomery [5], we must test the assumptions of the MLRM, such as (1) the multicollinearity is tested using variance inflation factor *(VIF)*, that is *VIF* > 10, or *F* test is significant but *t* is not. We guarantee that there is no autocorelation of the error term and there is no heteroscedasticity. Moreover, Gujarati [7] showed that if the plot of error versus X_i going to be large (see Figure 1.), then X_i and Y_i must be divided by X_i in order to get the eligible data.

Figure 1. Plot between *error* e_i and weighted X_i
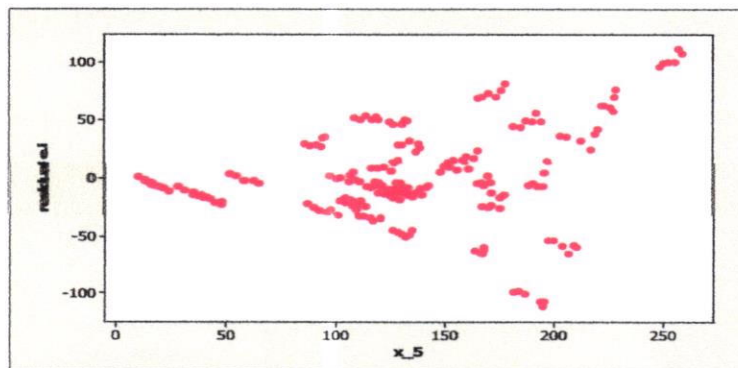
2.2. Stepwise Regression Method

The stepwise regression method generally consists of forward selection and backward elimination (Soejoeti, [25]). Here, we want to choose a small subset from the larger set (large set of candidate predictor variables) so that the resulting regression model is good predictive ability. In this method, we enter and remove predictors

until there is no justifiable reason to enter or remove more. First step, we fit each of the one-predictor models, that is, regress y on x_1, regress y on x_2,..., regress y on x_p. The first predictor is the predictor that has the smallest t-test (or high correlation (r) or significant in F test). Similarly, 2nd step, we suppose x_1 was the "best" one predictor, then we fit each of the two-predictor models with x_1 in the model, that is, regress y on (x_1, x_2), regress y on (x_1, x_3),..., and y on (x_1, x_p). Again, the second predictor is the predictor that has the smallest t-test (or high correlation (r) or significant in F test). But, we must consider to remove one of them if the model on the 2nd step, $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$, is not significant. For example, if the $\beta_1 = 0$ has become not significant, remove x_1. The procedure is stopped when adding an additional predictor has no significant t-test more. Following, [20], the first step of the backward elimination procedure is to allow us to fit the full model of the MLRM. We then eliminate one by one using correlation (r) criteria (and or t or F test) in testing hypothesis of the coefficient regression parameters. Here, we need many steps in getting the eligible coefficient regression parameters model of the MLRM.

2.3. Best Subset Method

The best subset method is used to find the eligible predictors (X) in the MLRM model, with $n > p$. Here, we select the subset of predictors that do the best at meeting some objective criterion. Following Draper and Smith [21], we follow several steps to get the eligible predictor in the model, that are:

(1) consider and choose the highest R^2 : $R^2 = 1 - \dfrac{SSE}{SST}$,

(2) determine the maximum R^2_{adj} : $R^2_{adj} = 1 - \dfrac{(n-1)SSE}{(n-p)SST}$,

(3) choose the small C_p and C_p is close to p, where $C_p = \dfrac{SSE}{MSE} - n + 2p$, and

(4) finally, we must choose the small s, where s is square root of MSE .

III. A SIMULATION STUDY

Following Subsections 2.1 to 2.3, we then simulate the model of response Y (percentage of poverty line) and four predictors X_1 (jobless), X_2 (population growth rate), X_3 (life expectancy), and X_4 (length of study). The full model of the data simulation and their Pearson correlation (Table 2.) are then given as, respectively,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4 = 108 + 0.5 x_1 - 0.8 x_2 - 1.6 x_3 + 0.9 x_4, \text{ and}$$

Table 2. Pearson correlation

	X_1	X_2	X_3	X_4
Y	0.8	0.5	-0.8	-0.5.

From Table 2, it is clear that X_1 and X_3 have high correlation, so they are significant and eligible to the model. To make sure that model is really good, we then analysis it using both methods, namely forward and backward *selection methods as follow.*

The output of the procedure of the stepwise forward selection and backward elimination are presented in Table 3. and Table 4. (Suleman, [14]).

Table 3. Stepwise Regression (2 steps): Y versus X_1, X_2, X_3, X_4 with $\alpha = 0.05$

Step	1	2
Constant	123.7	89.7
X_3	-1.7	-1.2
T-Value	-3.8	-5.2
p-Value	0.005	0.001
X_1		0.37
T-Value		5.2
p-Value		0.001
s	0.63	0.31
R-sq	64.5	92.6
R-sq (adj)	60.0	90.4
Mallow-C_p	25.3	2.5

Table 4. Stepwise Regression (3 steps): Y versus X_1, X_2, X_3, X_4 with $\alpha = 0.05$

Step	1	2	3
Constant	108.0	89.7	89.9
X_1	0.46	0.48	0.39
T-Value	4.0	4.3	5.2
p-Value	0.011	0.005	0.001
X_2	-0.8	-0.60	
T-Value	-1.2	-1.1	
p-Value	0.27	0.31	
X_3	-1.6	-1.2	-1.2
T-Value	-2.6	-5.3	-5.2
p-Value	0.049	0.002	0.001
X_4	0.9		
T-Value	0.7		
p-Value	0.55		
s	0.32	0.30	0.31
R-sq	94.3	93.8	92.6
R-sq (adj)	89.8	90.8	90.4
Mallow-C_p	5.0	3.4	2.5

Table3. and Table 4. showed that the eligible model is similar, namely $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3 = 90 + 0.4x_1 - 1.2x_3$. Here, we got two steps on the forward selection model but not in backward elimination model (3 steps). We then conclude that forward selection method is more efficient than backward elimination method.

Moreover, we then analysis the data using the best subsets method using Minitab software, and the result is given in the Table 5. (Suleman, [14]).

B. Pratikno et al. / International Journal of Engineering and Technology (IJET)

Table 5. Best Subsets Procedure

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	X_1	X_2	X_3	X_4
1	64.5	60.0	25.3	0.633			x	
1	63.8	59.2	25.9	0.639	x			
2	92.6	90.4	2.5	0.310	x		x	
2	86.7	82.8	7.7	0.415	x			x
3	93.8	90.8	3.4	0.304	x	x	x	
3	92.6	88.9	4.5	0.334	x		x	x
4	94.3	89.8	5.0	0.320	x	x	x	x

Here, we also presented the output of the best subset method in the case of poverty line (Y) and X_2 (population growth rate), X_3 (life expectancy), and X_4 (length of study) on different data (see Table 5., Sopanti, [9]).

Table 6. Output Best Subset Model

Vars	R-Sq	R-Sq (adj)	Mallows Cp	S	X_2	X_3	X_4
1	97,9	97,7	4,4	0,2882	x		
1	63,7	60,0	209,6	1,2072		x	
2	98,6	98,3	2,1	0,2460	x		x
2	98,5	98,2	3,0	0,2594	x	x	
3	98,7	98,2	4,0	0,2588	x	x	x

From Table 6., we see that X_2 and X_3 are the best significant predictors to the model of the MLRM. This is due to the C_p close to $p=3$ (even it is not the smallest), and the $R^2 = 98.5$ (2nd highest), maximum $R_{adj}^2 = 98.2$, $C_p = 3.0$ close to p ($p = 3$) and small $s = 0.25943$. In this case, we note that the variables name are: X_2 is population growth rate, X_3 is life expectancy, and X_4 is length of study. We then conclude that the best model is $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$.

IV. CONCLUSION

The paper studied the selection variables (predictors) in multiple regression model (MLRM) using best subset and stepwise regression methods. The result showed that the best model of the stepwise (forward and backward) method are similar (X_1 and X_3, are significant, and the model is $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3$), but the forward selection (with 2 steps) is more efficient than backward elimination (with 3 steps). Unfortunately, the best subset method has a different selection variables, where X_2 is also significant, so the model is $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$. Here, we note that the output of the $R^2 = 93.8$ (2nd highest), maximum $R_{adj}^2 = 90.8$, $C_p = 3.4$ close to p ($p=4$) and $s=0.30445$, are the suitable indicators of the criteria. We therefore (finally) conclude that the best model is, $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3$. This is due to the X_1 and X_3 variables are available on the three tests, namely (1) correlation testing (partial test), (2) stepwise method, and (3) best subset method.

REFERENCES

[1]   A. G. Bluman, 1997. Elementary Statistics: A Step by Step Approach. 3rd edition, Mc Graw-Hill.

[2]   B. H. Baltagi. 2005. Econometric Analysis of Panel Data 3th edition. England: John Willey and Sons, Ltd.

[3]   B. H. Baltagi. 2008. Econometrics 4th edition. Verlag Berlin Heidelberg: Springer.
[4]   B. Pratikno. 2012, Tests Of Hypothesis For Linear Regression Models With Non Sample Prior Information, Dissertation, University of Shouthern Queensland, Australia.
[5]   D. C. Montgomery. 1996. Regression Analysis. John and Wiley Son, USA
[6]   D. C. Montgomery, E. A. Peck and G. G. Vining. 2012. Introduction to Linear Regression Analysis, 5th Edition. Canada: John Wiley and Sons, Inc.

[7]   D.N. Gujarati. 2004. Basic Econometrics 4th edition. New York: McGraw-Hill
[8]   D.N. Gujarati. 2016. Dasar-Dasar Ekonometrika terjemahan. Jakarta: Erlangga.
[9]   D. Sopanti. 2019. Analisis Faktor Yang Berpengaruh Terhadap Persentase Kemiskinan Di Provinsi Jawa Barat Dengan Metode Regresi Linier Berganda, Laporan Kerja Praktek (Unpblished)
[10] F.A. Graybil, and H.K. Iyer. 1992. Regression Analysis. John & Wiley Son, USA.
[11] G.A.F. Seber and J. L. Alan (2003). Linear Regression Analysis, Second Edition. New Jersey: John Wiley dan Sons, Inc.
[12] G.K. Bhattacharyya and R.A. Johnson. 1977. Statistical Concepts and Methods, John and Wiley Son, USA.
[13] I. Ghozali. (2016). Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23, Ed. 8 Cetakan ke VIII. Semarang : Badan Penerbit Universitas Diponegoro.
[14] I.P. Suleman. 2019. Regresi Linier Berganda Dan Aplikasinya Pada Pemodelan Data Kemiskinan Di Kabupaten Tasikmalaya, Laporan Kerja Praktek (Unpblished)
[15] M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li. 2005. Aplied Linear Statistical Models, 5th Edition. New York: McGraw-Hill Companies, Inc.
[16] N. Iriawan and S.P. Astuti. 2006. Mengolah Data Statistik dengan Mudah Menggunakan Minitab 14. Yogyakarta: ANDI.
[17] N. R. Draper and H. Smith. 1998. Applied Regression Analysis. Jhon Wiley & Sons, USA
[18] R.E. Walpole and R.H. Myers. 1997. Probability and Statistics for Engineers and Scientist, Mc. Milan, USA.
[19] R.E. Walpole, R.H. Myers, S.L. Myers and K. Ye. 2012. Probability and Statistics for Engineers and Scientist. Pearson, USA.
[20] R. Huggins and R. Staudte. 2002. Biostatistics. La Trobe University, Melbourne.
[21] T. A. Bancrof. 1944. On biases in estimation due to the use of the preliminary tests of significance, Annals of Mathematical Statistics, 15(1944), 190-204.
[22] W.H. Greene. 2004. Econometric Analysis. Prentice Hall: New Jersey
[23] W. Mendenhall. 2012. Regression Analysis (7th ed). Amerika Serikat : University of Florida.
[24] W. Medenhall and T. Sincich. 2004. A Second Course in Statistic: Regression Analisis, Edisi ketujuh. New York: Prentice Hall.
[25] Z. Soejoeti. 2010. Metode Statisik I, Universitas Terbuka, Jakarta.

as Draper and Smith [21], Gujarati [7], Gujarati [8], Ghozali [13], Kutner et al.[17], Montgomery [5], Montgomery [6], Bhattacharyya and Johnson [12], Greene [22], Baltagi [2], Baltagi [3], Mendenhall [23],

Mendenhall and Sincich [24], Bluman [1], Huggins and Staudte [20], Iriawan and Astuti [16], Graybil and Iyer [10], Walpole and Myers [18], Walpole et al [19] and Pratikno [4]. Here, Pratikno [4] have been studied testing of the hypothesis tests on regression models using power and size of the hypothesis testing.

In this paper, the introduction is presented in Section 1. The regression model, best subset and stepwise methods are given in Section 2. A simulation is then obtained in Section 3. Section 4 described the conclusion of the research. document is a template.

II. REGRESSION MODEL, BEST SUBSET AND STEPWISE METHODS

2.1. Multiple Regression Model

For an n pair of observations on p independent variables $(X_1,...,X_p)$ and one dependent variable (Y), $(X_i, Y_i$, for $i=1,2,...,n$ and $j=1,2,...,p$, the multiple regression model (in matrix) is given by

$$Y = X\beta + e \tag{1}$$

Here, $\beta = (\beta_0, \beta_1,...,\beta_p)'$ is a $(p+1)$ dimensional column vector of unknown regression parameters, $Y = (y_1,...,y_n)'$ is a $(n \times 1)$ vector of response variables, X is a $n \times (p+1)$ matrix of know fixed values of the independent variables an e is the error term which is assumed to be identically and independently distributed as $N_n(0, \sigma^2 I_n)$. Here, I_n is the identity matrix of order n and σ^2 is the common variance of the error variables. Following Montgomery [5] and Graybil and Iyer [10], the estimate coefficient regression model in matrix is then given as

$$\hat{\beta} = (X'X)^{-1} X'Y \tag{2}$$

Furthermore, we test the hypothesis testing of the, $H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0$. Here, Anova ($F$ test) is used, and it is presented in Table 1.

Table 1. Anova of the MLRM with p independent variables

Source of Variance	Degree of freedom (df)	Sum Square (SS)	Mean Square (MS)	F^*
Regression	p	SSR	MSR	$F^* = \dfrac{MSR}{MSE} > F_{\alpha(p,n-1(p+1))}$
Error	$n-(p+1)$	SSE	$MSE = s^2$	
Total	$n-1$	SST		e then reject H_0

Note that SSE is sum square error, SST is sum square total and MSE is mean square error. Following Montgomery [5], we must test the assumptions of the MLRM, such as (1) the multicollinearity is tested using variance inflation factor (VIF), that is $VIF > 10$, or F test is significant but t is not. We guarantee that there is no autocorrelation of the error term and there is no heteroscedasticity. Moreover, Gujarati [7] showed that if the plot of error versus X_i going to be large (see Figure 1.), then X_i and Y_i must be divided by X_i in order to get the eligible data.

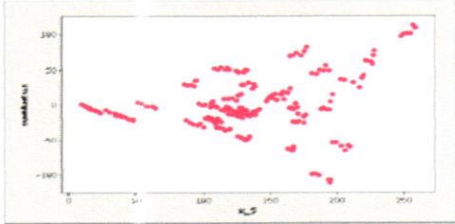Figure 1. Plot between error e_i and weighted X_i

2.2. Stepwise Regression Method

The stepwise regression method generally consists of forward selection and backward elimination (Soejoeti, [25]). Here, we want to choose a small subset from the larger set (large set of candidate predictor variables) so that the resulting regression model is good predictive ability. In this method, we enter and remove predictors

until there is no justifiable reason to enter or remove more. First step, we fit each of the one-predictor models, that is, regress y on x_1, regress y on x_2,..., regress y on x_p. The first predictor is the predictor that has the smallest t-test (or high correlation (r) or significant in F test). Similarly, 2nd step, we suppose x_1 as the "best" one predictor, then we fit each of the two-predictor models with x_1 in the model, that is, regress y on (x_1, x_2), regress y on (x_1, x_3), ... and y on (x_1, x_p). Again, the second predictor is the predictor that has the smallest t-test (or high correlation (r) or significant in F test). But, we must consider to remove one of them if the model on the 2nd step, $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$, is not significant. For example, if the $\beta_1 = 0$ has become not significant, remove x_1. The procedure is stopped when adding an additional predictor has no significant t-test more. Following, [20], the first step of the backward elimination procedure is to allow us to fit the full model of the MLRM. We then eliminate one by one using correlation (r) criteria (and or t or F test) in testing hypothesis of the coefficient regression parameters. Here, we need many steps in getting the eligible coefficient regression parameters model of the MLRM.

2.3. Best Subset Method

The best subset method is used to find the eligible predictors (X_i) in the MLRM model, with $n > p$. Here, we select the subset of predictors than do the best at meeting some objective criterion. Following Draper and Smith [21], we follow several steps to get the eligible predictor in the model, that are:

(1) consider and choose the highest R^2: $R^2 = 1 - \dfrac{SSE}{SST}$.

(2) determine the maximum R_{adj}: $R_{adj}^2 = 1 - \dfrac{(n-1)SSE}{(n-p)SST}$.

(3) choose the small C_p and C_p is close to p, where $C_p = \dfrac{SSE}{MSE} - n + 2p$, and

(4) finally, we must choose the small s, where s is square root of MSE.

III. A SIMULATION STUDY

Following Subsections 2.1 to 2.3, we then simulate the model of response Y (percentage of poverty line) and four predictors X_1 (jobless), X_2 (population growth rate), X_3 (life expectancy), and X_4 (length of study). The full model of the data simulation and their Pearson correlation (Table 2.) are then given as, respectively.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4 = 108 - 0.5 x_1 - 0.8 x_2 - 1.6 x_3 + 0.9 x_4, \text{ and}$$

Table 2. Pearson correlation

	X_1	X_2	X_3	X_4
Y	0.8	0.5	-0.8	-0.5

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3.$$ Here, we note that the output of the $R^2 = 93.8$ (2nd highest), maximum $R_{adj}^2 = 90.8$, $C_p = 3.4$ close to p ($p=4$) and $s=0.30445$, are the suitable indicators of the criteria. We therefore (finally) conclude that the best model is, $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3$. This is due to the X_2 and X_4 variables are available on the three tests, namely, (1) correlation testing (partial tests), (2) stepwise method, and (3) best subset method.

REFERENCES

[1] A. G. Bluman. 1997. Elementary Statistics: A Step by Step Approach. 3rd edition. Mc Graw Hill.

[2] B. H. Baltagi. 2005. Econometric Analysis of Panel Data. 3rd edition. England: John Willey and Sons, Ltd.

[3] B. H. Baltagi. 2008. Econometrics. 4th edition. Verlag Berlin Heidelberg: Springer.

[4] B. Pratikno. 2012. Tests Of Hypothesis For Linear Regression Models With Non Sample Prior Information. Dissertation. University of Southern Queensland, Australia.

[5] D. C. Montgomery. 1996. Regression Analysis. John and Wiley Son. USA.

[6] D. C. Montgomery, E. A. Peck and G. G. Vining. 2012. Introduction to Linear Regression Analysis. 5th Edition.Canada: John Wiley and Sons. Inc.

[7] D.N. Gujarati. 2004. Basic Econometrics. 4th edition. New York: McGraw Hill.

[8] D.N. Gujarati. 2016. Dasar-Dasar Ekonometrika terjemahan. Jakarta: Erlangga.

[9] D. Supatmi. 2019. Analisis Faktor Yang Berpengaruh Terhadap Persentase Kemiskinan Di Provinsi Jawa Barat Dengan Metode Regresi Linier Berganda. Laporan kerja Praktek (Unpublished).

[10] F.A. Graybil, and H.K. Iyer. 1992. Regression Analysis. John & Wiley Son. USA.

[11] G.A.J. Seber and J. S. Alan (2003. Linear Regression Analysis. Second Edition. New Jersey: John Wiley dan Sons, Inc.

[12] G.K. Bhattacharyya and R.A. Johnson. 1977. Statistical Concepts and Methods. John and Wiley Son. USA.

[13] I. Ghozali. (2016) Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23. Ed. 8 Cetakan ke VIII. Semarang : Badan Panerbit Universitas Diponegoro.

[14] I.P. Suleman. 2019. Regresi Linier Berganda Dan Aplikasinya Pada Pemodelan Data Kemiskinan Di Kabupaten Tasikmalaya. Laporan Kerja Praktek (Unpublished).