Available online at www.sciencedirect.com

jmr&t Journal of Materials Research and Technology
www.jmrt.com.br

Original article

Material selection of natural fibre using a stepwise regression model with error analysis

Muhammad Noryani [a,c,e], *Salit Mohd Sapuan* [a,b,*], *Mohammad Taha Mastura* [d], *Mohd Yusoff Moh Zuhri* [a], *Edi Syams Zainudin* [a]

[a] *Advanced Engineering Materials and Composites Research Centre, Department of Mechanical and Manufacturing Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[b] *Laboratory of Biocomposite Technology, Institute of Tropical Forestry and Forest Products (INTROP), Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[c] *Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
[d] *Faculty of Mechanical and Manufacturing Engineering Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
[e] *Centre of Advanced Research on Energy, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*

ARTICLE INFO

Article history:
Received 30 October 2018
Accepted 18 February 2019
Available online 25 May 2019

Keywords:
Material selection
Natural fibre
Stepwise regression
Error analysis

ABSTRACT

The nature of natural fibre such as it is lightweight, recyclable, biodegradable and gives a high performance in relation to its mechanical properties makes this material an excellent alternative to currently used materials in the manufacture of automotive components. The significant mechanical properties are identified using the best statistical model suggested by stepwise regression in this study. The estimation and error analysis of the response variables are discussed to select the best natural fibre for automotive component applications. The results using statistical measurement indicate that tensile strength is the most significant mechanical property for all the selected natural fibres. The final ranking that considered high performance score and minimum error analysis for a hand-brake lever application found that coir, kenaf and cotton are the top three candidates with average scores of 4, 4.5 and 5, respectively. The statistical model presented in this study can be used in multiple applications. In fact, this approach is helpful to the design engineer when huge amounts data are involved.

1. Introduction

Industry wide, mass reduction strategies have become increasingly important in recent years. In the United States, the automotive industry is required to meet the Corporate Average Fuel Economy (CAFE) standard for each car sale [1]. In 2011, 13 large automakers such as Ford, GM, BMW, Honda, Hyundai, Mazda, Nissan, Toyota and Volvo agreed with the regulations for vehicles from model year 2017 until 2025 [2]. This guideline was also implemented in the Kingdom of Saudi Arabia, effective by early 2016 [3]. In recent years, researchers have come up with an innovative idea to use natural fibre

* *Corresponding author.*
E-mail: sapuan@upm.edu.my (S.M. Sapuan).
https://doi.org/10.1016/j.jmrt.2019.02.019

as a substitute material in the automotive industry to comply with the global standard. The benefits of this material, such as it is lightweight, recyclable, biodegradable, economical and has good mechanical properties, mean that it is an excellent alternative to the currently used materials to manufacture automotive components in the industry [4–8]. A number of studies have shown that this natural fibre has its own strengths that can fulfil the product design requirements, especially in the automotive industry [9–12]. Not just in the automotive industry, but in other industries the demand for natural fibre composites is also increasing year by year, e.g., in construction, aerospace, textiles, medicine, packaging, and electrical and electronics applications [6,13,14]. To replace steel-based materials with natural fibre in automotive components can overcome or reduce problems such as heavy mass, high cost of materials and manufacturing, high fuel consumption and corrosion issues over a certain period that are always a challenge for the design engineer to consider as a whole in the manufacturing process [15–17].

Material selection is one of the important processes in automotive assembly. The design engineer should select an appropriate suitable material that fulfils the criteria of the design and function of the automotive component. Numerous selection strategies can be used to transfer the inputs to the final output in different multi-criteria decision-making (MCDM) tools, as shown in Fig. 1. It should be noted that there is no specific procedure to select the best natural fibre in automotive applications. The traditional method used manual checking on the product design specification that match to the materials performance. A lot of literature discusses the strengths and limitations of the conventional material selection methods such as analytical hierarchy process (AHP), analytical network process (ANP), technique of ranking preferences by similarity of the ideal solutions (TOPSIS), quality function deployment (QFD), preference ranking organisational method for enrichment evaluation (PROMETHEE) and Vlse Kriterijumska Optimizacija Kompromisno Resenje (VIKOR) [18–26]. Just focusing on the weaknesses, the weighting process in AHP is very subjective, based on the judgement of the user. Furthermore, TOPSIS ignores the relationship and interdependency between the attributes throughout the process selection, which can produce inconsistent decision-making. Very simple tools such as PROMETHEE make the final decision not rational and big data is needed for some tools such as VIKOR, which is one of the limitations in MCDM tools. Quality Function Development is one of MCDM tools that can translate customer requirement into design requirement, but the subjective preference also occur in correlation matrix in house of quality matrix diagram [26,27]. Many studies have introduced single and integrated modelling in decision-making [28]; a single model focuses on one criterion, otherwise multiple criteria are considered in the study. With regard to the latter type of model, the effect of each criterion on the model is analysed and the final decision is dependent on the criteria or input considered in the study [29].

The huge number of materials that exist in the world means that it takes up the design engineers' valuable time and increases costs to select the best material. A novel statistical approach to select a natural fibre composite was introduced by Noryani et al. [20]. By using the statistical package for social science (SPSS) programme, the significant mechanical properties are identified, and the statistical model is constructed for the appropriate data set. This approach can reduce the time and cost throughout the process [31,32]. In the previous study, three types of composites are measured using statistical analysis such as correlation coefficient, coefficient of determination, significant testing, multicollinearity testing and analysis of variance. The present study is motivated by the need to take into consideration the natural fibre itself before it becomes the final composite. In addition, the precision of estimating the performance score for natural fibre is considered compared to the earlier study.

In this study, an improved strategy for the material selection process from the previous work is discussed [20] where another measurement such as error analysis is considered with the estimation values. The performance of alternative natural fibres is analysed by using statistical measurement. The significant mechanical properties are identified to estimate the performance score (*PS*) for each alternative using the best statistical model suggested by stepwise regression. To select the best natural fibre, the product design specification of a hand-brake lever is used as a case study. Finally, the error analysis for the estimation can increase the trust in the final decision material. In fact, the potential and suitability of this approach can be applied to different automotive components in the industry.

2. Methodology

The data collection of mechanical properties for natural fibre in the previous study is an initial stage in this methodology. The screening process is started by using stepwise regression to identify the best statistical model of the natural fibre by excluding the irrelevant models. In this stage, the coefficient of the regressor is expressed by using least squares estimation with significant parameters to the model constructed. The product design specification is used to estimate the *PS* of the natural fibre. The screening process is performed after calculating the error analysis for all candidates. The errors analysis in this study are mean absolute error (*MAE*), mean squared error (*MSE*) and root mean squared error (*RMSE*).

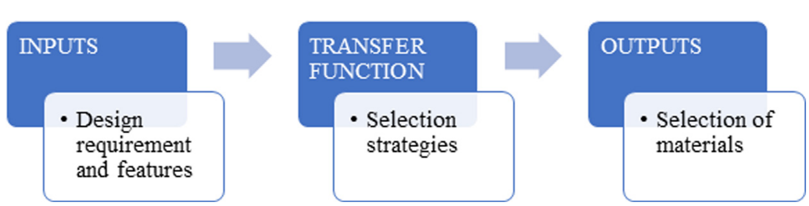

Fig. 1 – The process of material selection [30].

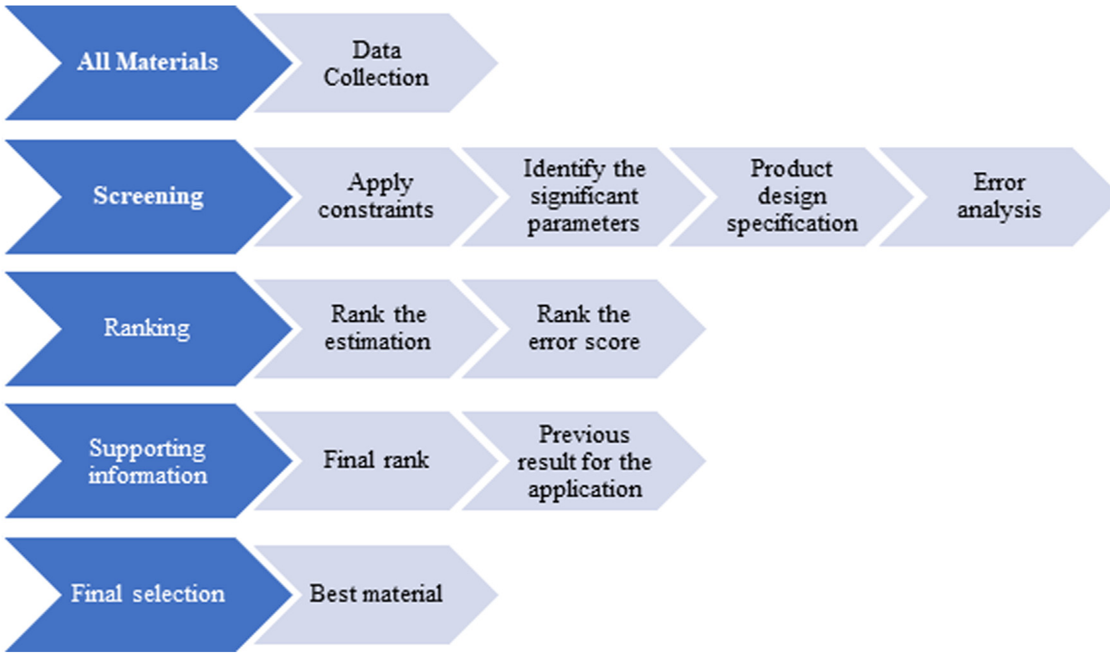

Fig. 2 – The methodology of the material selection process.

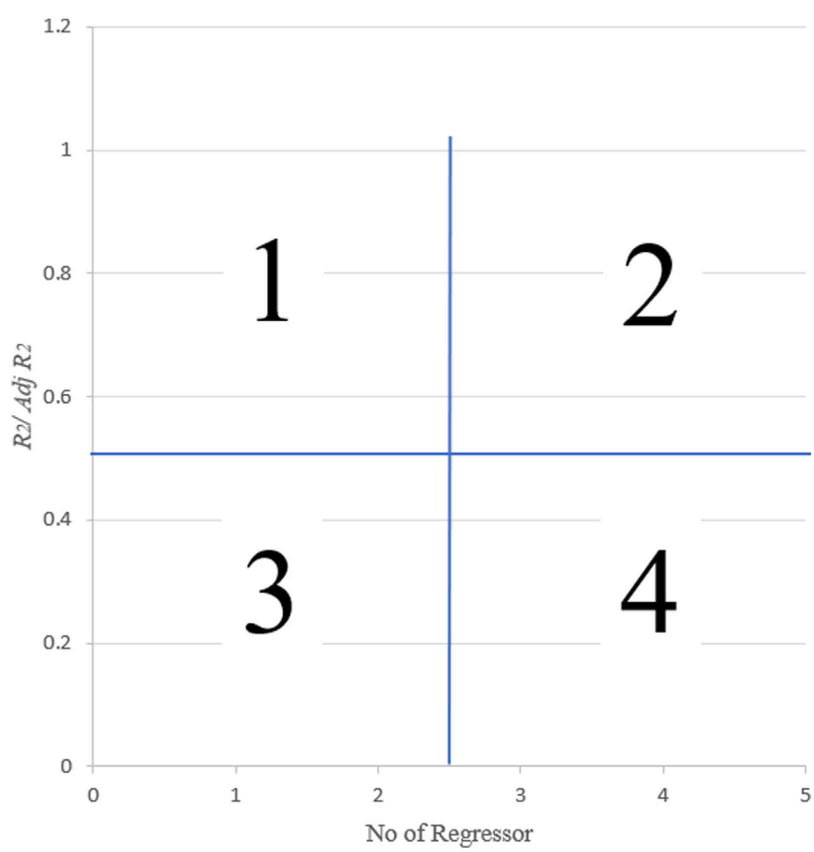

Fig. 3 – Quarter selection position of best natural fibre statistical model.

The second stage is ranking the result from the first stage. Identifying the maximum score of the estimation and the minimum score of the error analysis is the objective in this stage. The third stage is supporting information such as final ranking from the second stage and previous result for the application. Figs. 2 and 3 illustrates the three stages of the selection process for the material starting from all materials to the final selection of the specific material in this study.

In this methodology, each stage is continuously checked to ensure the decision-makers reach the best final decision [33].

2.1. Data collection

Mechanical properties of natural fibre from the previous study are used as secondary data in this study. Here, the implementation of the average value can reduce the dispersion of the data. Twelve types of natural fibre are selected. Table 1 shows the density (x_1), tensile strength (x_2), Young's modulus (x_3) and elongation at break (x_4) used in this study. These mechanical properties are the main criteria preferred in material selection for automotive components as mentioned by Noryani et al. [32]. Performance score (PS) is used as the response variable as calculated by using Eq. (1) [20]. Some of the cost of the material is not available (n/a). Therefore, the data is excluded in the analysis.

$$PS = \sum\nolimits_{i=1}^{n} x_i \quad i = 1, 2, ..., n \tag{1}$$

2.2. Stepwise regression

Forward selection and backward elimination are processes in stepwise regression. Forward selection starts with the assumption that there are no regressors in the model except the intercept. The process is followed by inserting the regressors into the model one at a time to find the optimal subset in the model. The largest simple correlation to the response variable (y) is considered into the equation; the second regressor considered into the equation also has a high partial correlation towards y after adjusting the effect of the first regressor entered into the model. The F-statistics in Eq. (2) that illustrate x_2 have a high partial correlation when x_1 is already in the model [40]:

$$F = \frac{SS_R\,(x_2|x_1)}{MS_{RES}\,(x_1, x_2)} \tag{2}$$

where SS_R is the sum of the square of the regression and MS_{RES} is the mean square of the residual.

If this F value exceeds F_{IN}, then the regressor is added to the model. In general, the regressor having a high partial correlation with y which considers the effect of another regressor already in the model is entered into the model. The process stops when the F-statistics do not exceed F_{IN} or the last regressor is added to the model. An opposite direction is a practice in backward elimination; the process starts with all regressors being included in the model. F_{OUT} is used to exclude the regressor that has the smallest partial correlation into the model. The summary of the process of forward selection and backward elimination is shown in Table 2.

2.3. Least squares estimation

Here, a standard approach in regression analysis is used. The general form implemented for the stepwise regression is shown in Eq. (3). It can be simple or multiple linear regression; the number of regressors will identify the type of regression.

$$y_i = \beta_0 + \beta_1 x_i + \ldots + \varepsilon_i \quad i = 1, \ldots, n. \tag{3}$$

where y_i is the response variable (PS) x_i are the regressor variables, $\beta_1, \beta_2, \ldots, \beta_n$ are partial regression coefficient, ε_i is an error term and the subscript i indexes a particular observation.

The variation of the PS is explained by the regressors by calculating R^2 and $Adj\ R^2$ using Eqs. (4) and (5).

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{Res}}{SS_T} \tag{4}$$

$$AdjR^2 = 1 - \frac{SS_{Res}/df_e}{SS_T/df_T} \tag{5}$$

where SS_R is the sum of the square of the regression; SS_{Res} is the sum of the square of the residual; SS_T is the sum of the square of the total; df is the degree of freedom ($n-1$ or $n-p-1$).

2.4. Measuring model performance using error analysis

Error analysis of the response variable (PS) is analysed by using three types of error, which are mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE), to identify the minimum error of the estimation of PS for each alternative natural fibre. The minimum error is required to guide the decision-maker in the process of selection with precision, less bias and result in a highly accurate solution.

2.4.1. Mean absolute error (MAE)

The simplest error used for assessing the fitness of the model. The advantages of this error are its simplicity, and it is easy to understand and to calculate. This type of error is most preferred compared to the others. MAE can be calculated using Eq. (6).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} e_i \tag{6}$$

2.4.2. Mean squared error (MSE)

Another common error used to measure the accuracy of the estimation. Large errors may occur in this type of error because of the square function. MSE can be calculated using Eq. (7).

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n} \tag{7}$$

2.4.3. Root mean squared error (RMSE)

Another favourite error used in inferential statistics. Mostly, using this type of error produces small errors compared to MSE. RMSE can be calculated by using Eq. (8).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}} \tag{8}$$

for Eq. (6) until Eq. (8) which $e_i = \left|y_i - \hat{y}_i\right|$, where y_i is the actual observed value, $\hat{y}_i$ is the estimated value and n is the number of sample error in the model.

Table 1 – Properties of natural fibre [5,11,12,34–39].

Natural fibre	Mechanical properties				
	Cost (MYR/kg)	Density (g/cm^3)	Tensile strength (MPa)	Young's modulus (GPa)	Elongation at break (%)
Banana	n/a	1.4	721.5	29.5	7
		1.35	500	12	5.25
		1.35	355	33.8	5.3
Bagasse	n/a	1.25	256	22.05	1.1
		1.25	256	22.05	1.1
		1.2	290	17	1.1
		1.2	155	23.4	1.01
Cotton	6.45	1.55	543.5	9.05	6.5
	9.68	1.55	543.5	9.25	6.5
	12.9	1.55	543.5	9.05	7.5
	6.45	1.5	442	9.05	6.5
	9.68	1.55	543.5	9.05	6.5
	12.9	1.51	400	12	7.5
	6.45	1.55	442	16.75	6.83
Ramie	6.04	1.28	669	86	4.6
	6.24	1.5	469	86	2.6
	6.14	1.5	669	57	2.6
	6.14	1.28	445	76.25	2.5
	6.14	1.28	700	76.25	2
	6.14	1.5	700	24.5	2.5
	6.14	1.5	560	44	2.8
	6.14	1.5	500	94.7	2.8
	6.14	1.42	669	68.09	2.8
Coir	0.78	1.2	212	5	27.5
	1.17	1.2	175.5	5	33.2
	1.56	1.31	162.5	5	33.2
	0.78	1.31	162.5	4.4	30
	1.17	1.2	175	4.4	20
	1.56	1.25	220	5	30
	0.78	1.2	175	6	27.5
	1.17	1.17	153	5	28.8
	1.56	1.23	179.4	6.5	28.8
Pineapple	n/a	1.74	1020	71	14.5
		1.2	513.5	1.44	2
		1.5	898.5	82	14.5
		1.2	513.5	1.44	10.3
Flax	9.68	1.5	922.5	25.75	2.4
	10.05	1.5	1087.5	18	2.25
	9.87	1.45	1171.5	53.5	2.25
	9.87	1.5	672	65.3	2.2
	9.87	1.5	690	27	2.95
	9.87	1.45	1172.5	27.6	2.25
	9.87	1.5	690	65.3	2.95
	9.87	1.4	1150	27.6	1.4
	9.87	1.5	690	70	2.95
	9.87	1.45	922.5	27.6	2.6
	9.87	1.48	916.85	63.8	2.42
Hemp	3.24	1.45	690	65	2.8
	6.45	1.5	830	46	3
	4.67	1.45	585	64	2.25
	4.79	1.48	644.5	56.75	1.6
	4.79	1.45	585	35	1.6
	4.79	1.48	690	56.75	1.6
	4.79	1.48	725	70	1.6
	4.79	1.48	690	70	2.06
	4.79	1.49	805.5	70	2.06
	4.79	1.47	693.9	70	2.06
Jute	2.26	1.38	586.5	21.5	1.9
	2.89	1.4	277	32.5	1.4
	1.17	1.4	596.5	43	1.65
	2.11	1.4	560	55	1.65
	2.11	1.46	596.5	26.5	1.65
	2.11	1.3	583	43	1.65

– Table 1 (Continued)

Natural fibre	Mechanical properties				
	Cost (MYR/kg)	Density (g/cm^3)	Tensile strength (MPa)	Young's modulus (GPa)	Elongation at break (%)
	2.11	1.4	560	26.5	1.8
	2.11	1.3	583	20	1.65
	2.11	1.46	600	26.5	1.48
	2.11	1.3	583	36.5	1.65
	2.11	1.4	550	33.1	1.65
Sisal	2.03	1.43	495	17.4	5
	3.89	1.4	681	18.7	4.5
	2.96	1.42	531.5	23.5	2.25
	2.96	1.5	573	15.7	4.5
	2.96	1.42	531.5	23.5	2.25
	2.96	1.5	573	15.7	2.5
	2.96	1.33	650	38	2.25
	2.96	1.5	573	15.7	4.5
	2.96	1.48	575.5	15.7	3.47
Kenaf	1.15	1.5	737.5	40.5	2.55
	1.22	1.19	361	57	2.1
	1.95	1.4	576.5	33.75	1.6
	1.44	1.45	930	53	2.1
	1.44	1.4	576.5	33.75	1.6
	1.44	1.3	930	53	4.25
	1.44	1.3	612.5	53	1.6
	1.44	1.2	930	53	2.25
	1.44	1.34	585	33.5	2.26
Bamboo	2.28	0.75	216.5	29	1.3
	3.60	0.85	470	21.5	3.1
	5.10	0.85	470	21.5	3.1
	3.66	0.85	185	14	2.5

Table 2 – The condition of forward selection and backward elimination in stepwise regression.

Condition	Forward selection	Backward elimination
Start	No regressor in the model except intercept	All regressors in the model
Add and remove	Example: → y, intercept (β_0) → y, β_0, x_1 → y, β_0, x_1, x_2 → y, β_0, x_1, x_2, x_3 → y, β_0, x_1, x_2, x_3, x_4 . . .. . . Add the regressor when: F-statistics > F_{IN} Example: $\frac{SS_R(x_2\|x_1)}{MS_{RES}(x_1,x_2)} > F_{IN}$ or P-value<α(0.05)	Example: → y, β_0, x_1, x_2, x_3, x_4 → y, β_0, x_1, x_2, x_3 → y, β_0, x_1, x_2 → y, β_0, x_1 → y, β_0 Remove the regressor when: F-statistics < F_{OUT} Example: $\frac{SS_R(x_2\|x_1)}{MS_{RES}(x_1,x_2)} < F_{OUT}$ or P-value > α(0.05)
Stop	No regressor to be added	Until all P-value ≤ α

F-statistics: F value calculated from Eq. (2), F_{IN}: critical point added, F_{OUT}: critical point remove, P-value: significant value, α: standard error.

2.4.4. *Ranking the error score and final rank*

Each error score is ranked (R) from small to large for alternative natural fibres. The rank of average score (R_{AS}) and the final rank are calculated using Eqs. (9) and (10), respectively. By using *AS* score, the final error rank for alternative natural fibres is finalised. Rank number 1 is the best alternative natural fibre with a small number of errors.

$$R_{AS} = \frac{R_{MAE} + R_{MSE} + R_{RMSE}}{3} \quad (9)$$

$$\text{Final rank} = \frac{R_{AS} + R_{PS}}{2} \quad (10)$$

3. Results and discussion

In this part, the result suggested by stepwise regression is discussed. The significant statistical model with the coefficient is shown in the least squares method. The best and consistent natural fibre statistical models are demonstrated in quarter R^2 and *Adj* R^2 figures. R^2 describes the variation of response

Table 3 – The best model of natural fibre suggested by stepwise regression.

Natural fibre	Model	Regressors in model	SS_{RES}	MS_{RES}	R^2	$Adj\ R^2$	P-value
Banana	1	x_2	285.423	285.423	0.996	0.992	0.041
	2	$x_1\ x_2$	0	0	1	1	0.000
Bagasse	1	x_2	10.974	5.487	0.999	0.998	0.001
	2	$x_2\ x_4$	0	0	1	1	0
Bamboo	1	x_2	72.782	36.391	0.999	0.999	0.000
Coir	1	x_2	140.744	20.106	0.965	0.961	0.000
	2	$x_2\ x_4$	5.088	0.848	0.999	0.998	0.000
	3	$x_2\ x_3 x_4$	0.873	0.175	1	1	0.000
Cotton	1	x_2	79.264	15.853	0.997	0.996	0.000
	2	$x_2\ x_4$	26.315	6.579	0.999	0.998	0.000
Flax	1	x_2	2987.07	331.89	0.992	0.991	0.000
	2	$x_2\ x_3$	1.216	0.152	1	1	0.000
	3	$x_2\ x_3 x_4$	0.076	0.011	1	1	0.000
Hemp	1	x_2	1135.779	141.972	0.982	0.98	0.000
	2	$x_2\ x_3$	3.824	0.546	1	1	0.000
Jute	1	x_2	1093.873	121.541	0.987	0.986	0.000
	2	$x_2\ x_3$	0.766	0.096	1	1	0.000
Kenaf	1	x_2	734.109	104.873	0.998	0.998	0.000
	2	$x_2\ x_3$	3.334	0.556	1	1	0.000
	3	$x_2\ x_3 x_4$	0.276	0.055	1	1	0.000
Pineapple	1	x_2	579.027	289.514	0.998	0.997	0.001
Ramie	1	x_2	3345.848	477.978	0.955	0.949	0.000
	2	$x_2\ x_3$	3.261	0.544	1	1	0.000
	3	$x_2\ x_3 x_4$	0.1	0.02	1	1	0.000
Sisal	1	x_2	294.484	42.069	0.99	0.989	0.000
	2	$x_2\ x_3$	4.8	0.8	1	1	0.000
	3	$x_2\ x_3 x_4$	0.317	0.063	1	1	0.000

variable, (PS) is explained by the regressor and $Adj\ R^2$ is the measurement for multiple (more than one) regressors. Estimation of performance score of the alternative natural fibres is plotted and the final natural fibre is selected by the comparison of both the estimation of PS and the error analysis.

3.1. *Best statistical model suggested by stepwise regression*

The previous study discussed the ability of stepwise regression to produce a significant statistical model without carrying out all possible simple and multiple linear regression [20]. The significant parameter is identified by constructing this statistical model [41]. In this study, there are 15 possible regression models for each natural fibre, but only the significant models are discussed. The main advantage of this method is that valuable time can be saved. By using stepwise regression, only the significant model is proposed by adding and excluding all the regressors based on the conditions discussed in Table 2. Table 3 shows the significant statistical model for each natural fibre that will be chosen as our best model for further decision-making for manufacturing the hand-brake parking lever. The minimum number of models proposed by the stepwise regression was for bamboo and pineapple. Coir, flax, kenaf, ramie and sisal were the natural fibres that had a maximum number of models suggested by the stepwise regression.

3.1.1. *Least squares estimates for best model*

The sum of squares method is used to find the coefficient of the regressor to the model. An appropriate model can be written for each model by using the coefficient appropriately, as shown in Table 4. Tensile strength (x_2) is the most significant regressor in the model, as shown in Table 4, followed by Young's modulus (x_3) and elongation at brake (x_4). Throughout the process, density (x_1) is the only significant parameter for banana.

3.1.2. *Selection of best statistical model using determination of coefficient score (R^2 and $Adj\ R^2$)*

The variation of the dependent variable in this study, which is PS, is measured by R^2 for a single regressor and $Adj\ R^2$ for multiple regressors in the model. The score of R^2 and $Adj\ R^2$ for all possible statistical models for some natural fibres is shown from Figs. 4–9. The area of each figure is divided into quarters, as shown in Fig. 3, to select the best model for estimation purposes. The selection is based on the highest value of $R^2/Adj\ R^2$ with minimum number of regressors. This condition can benefit the industry as less cost and time is required to estimate the performance of natural fibre in the industry with a reliable statistical model. The model is selected based on the position in the figure. The first quarter in Fig. 3 is the best condition, with a high score of $R^2/Adj\ R^2$ with fewer regressors, while the fourth quarter is the worst model, with the lowest score of $R^2/Adj\ R^2$ with more regressors.

Most of the best statistical models for natural fibre show x_2 (tensile strength) is the significant consistent mechanical property in the model with the highest R^2 value. The optimal parameter is identified; other studies used Taguchi design of experiments to optimise the parameters for experimental work [42]. For example, 99.7% of the variation of PS for cotton is explained by tensile strength. Tensile strength also explained the variation of the PS for coir, flax, hemp, jute and kenaf by 96.5%, 99.2%, 98.2%, 98.7% and 99.8% respectively. The consistent finding that concludes x_2

Table 4 – The coefficient of the regressor into the model.

Natural fibre	Regressors in model	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
Banana	x_2	31.874	–	1.001	–	–
	$x_1\ x_2$	−1328.355	1053.555	0.849	–	–
Bagasse	x_2	31.861	–	0.965	–	–
	$x_2\ x_4$	−107.578	–	0.85	–	154.889
Bamboo	x_2	24.34	–	1.012	–	–
Coir	x_2	38.491	–	0.988	–	–
	$x_2\ x_4$	5.214	–	1.004	–	1.057
	$x_2\ x_3 x_4$	1.046	–	1.001	1.072	1.027
Cotton	x_2	42.369	–	0.971	–	–
	$x_2\ x_4$	−11.998	–	0.989	–	6.712
Flax	x_2	105.879	–	0.946	–	–
	$x_2\ x_3$	15.322	–	0.998	0.999	–
	$x_2\ x_3 x_4$	11.188	–	1	1	1.051
Hemp	x_2	40.711	–	1.04	–	–
	$x_2\ x_3$	4.660	–	1.008	0.964	–
Jute	x_2	41.201	–	0.995	–	–
	$x_2\ x_3$	6.802	–	0.998	0.986	–
Kenaf	x_2	39.047	–	1.017	–	–
	$x_2\ x_3$	3.552	–	1.002	0.998	–
	$x_2\ x_3 x_4$	3.335	–	1.001	0.985	0.906
Pineapple	x_2	–832.581	–	1.178	–	–
Ramie	x_2	135.727		0.904		
	$x_2\ x_3$	8.853		1.001	1.01	
	$x_2\ x_3 x_4$	8.203		0.999	0.998	0.945
Sisal	x_2	0.676		1.048		
	$x_2\ x_3$	3.941		1.011	0.871	
	$x_2\ x_3 x_4$	1.404		1.008	0.959	0.822

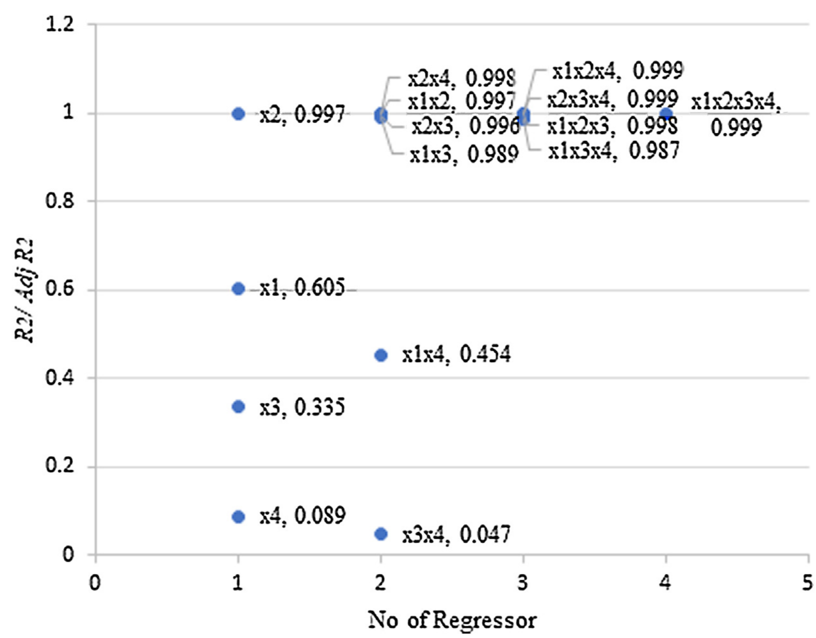

Fig. 4 – The score of R^2 and *Adj* R^2 for cotton natural fibre.

(tensile strength) is the most significant parameter in this study can help the estimation process compared to the previous study that found multiple significant mechanical properties for the composite [20,43]. Another study working on the variation of tensile strength using R^2 by Weibull statistics also found a high variation of sisal fibre of different gauge lengths [44]. Tensile properties of the materials are an important regressor to study, as shown in Table 4 where x_2 is the most significant regressor in the statistical models.

Previous studies on tensile strength of natural fibre performance show the significance of this mechanical property [45,46]. The performance of the tensile strength of natural fibre is affected by the process parameter, chemical treatment,

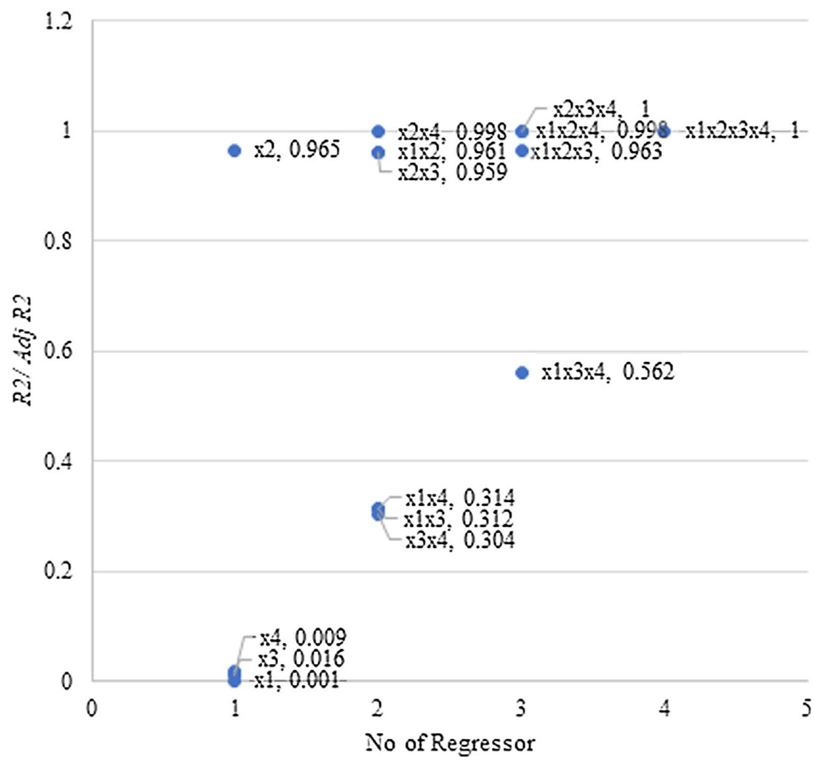

Fig. 5 – The score of R^2 and $Adj\ R^2$ for coir natural fibre.

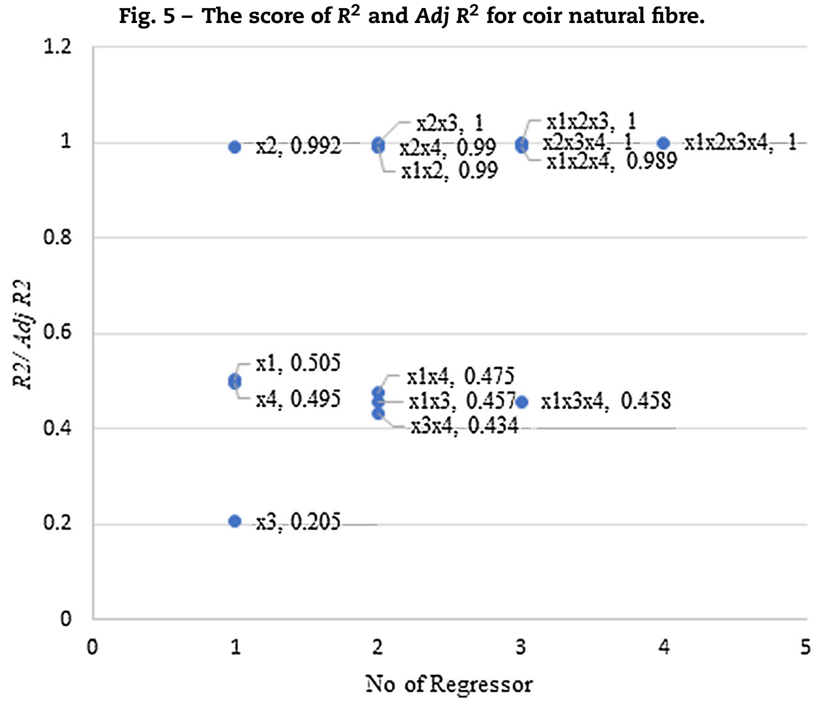

Fig. 6 – The score of R^2 and $Adj\ R^2$ for flax natural fibre.

chemical composition, fibre loading and fibre orientation [47–50].

3.1.3. *Natural fibre selection using estimation of performance score (PS)*

A statistical framework by Noryani et al. [51] is referred to. The process to select the most suitable material is essential after evaluation the performance of the different natural fibres. The estimation of the *PS* based on the product design specification for a hand-brake lever mentioned by Patel and Sarawade [52] is used in this study. Estimation and prediction by using a model can cut the cost and time in the manufacturing industry [53,54]. To manufacture the hand-brake lever, 460 MPa is required for tensile strength. This value is used in the statistical model proposed in Table 4. The final score of *PS* of

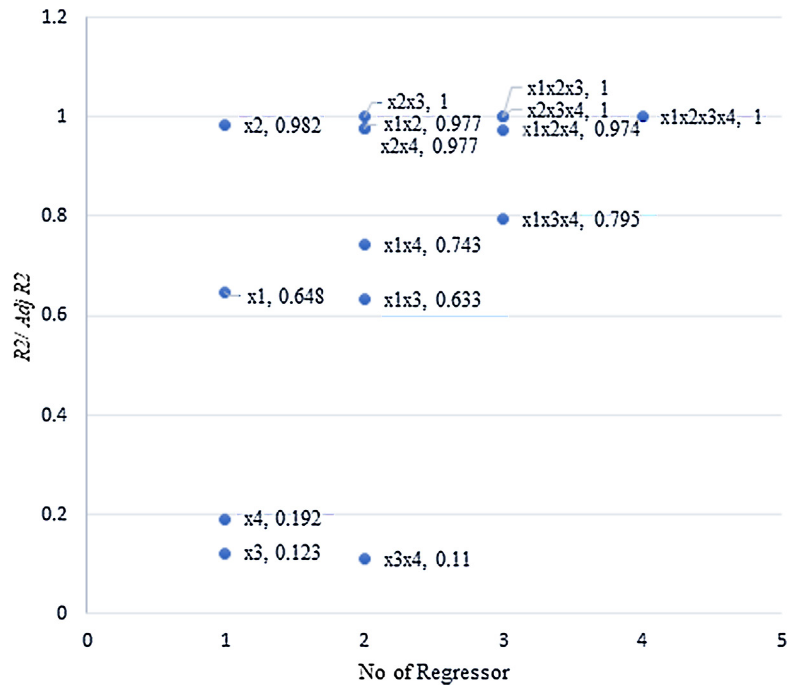

Fig. 7 – The score of R^2 and *Adj* R^2 for hemp natural fibre.

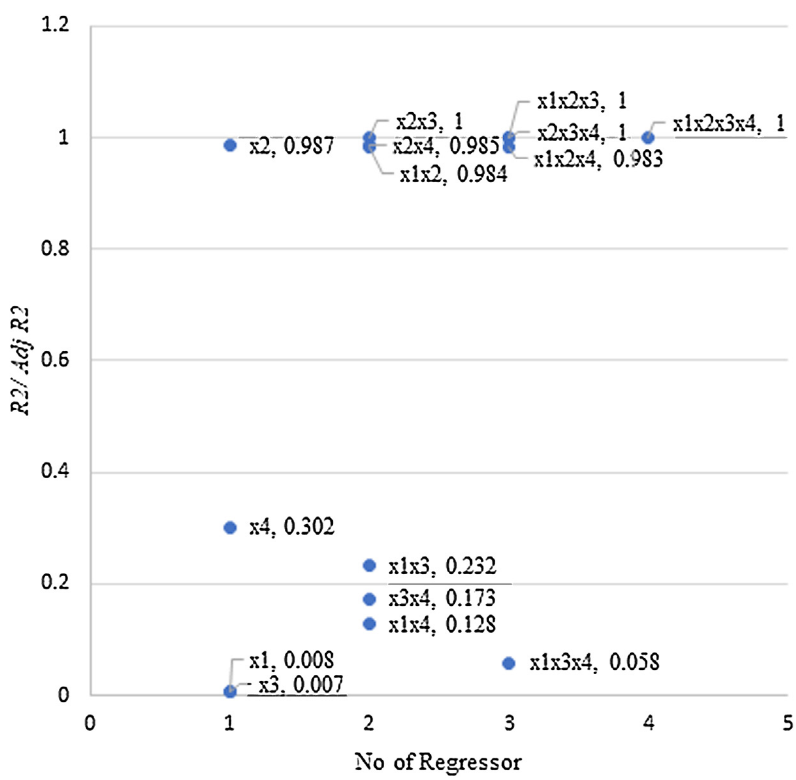

Fig. 8 – The score of R^2 and *Adj* R^2 for jute natural fibre.

the alternative natural fibre is shown in Fig. 10. This quantitative measurement can increase the accuracy in the decision-making and overcome the bias and subjective preference [55]. The best natural fibre for the hand-brake lever in this study was ramie, which scored the maximum value of PS, which is 551.57, compared to other fibres. Flax, hemp, kenaf and jute are the other top fibres for manufacturing a hand-brake lever, scoring 541.04, 519.11, 506.87 and 498.90 respectively.

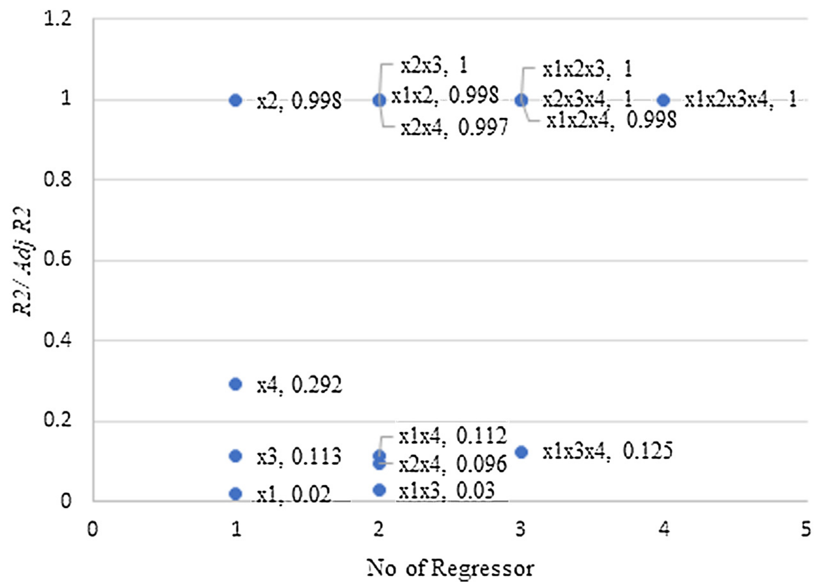

Fig. 9 – The score of R^2 and $AdjR^2$ for kenaf natural fibre.

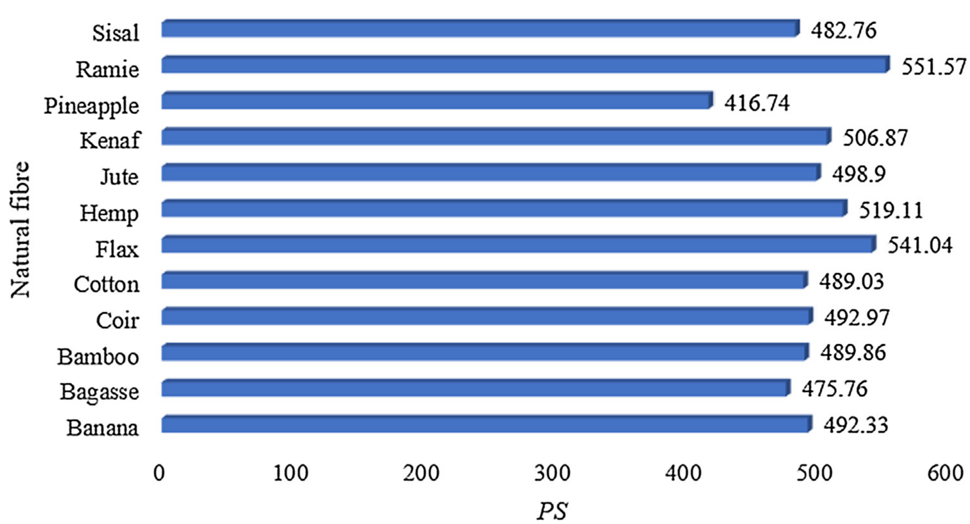

Fig. 10 – The performance score (PS) for the natural fibres.

3.1.4. *Error analysis*

Minimum error in estimation is used to help the decision-maker to select the final material in the automotive industry. The accuracy of the measurement is determined by calculating the error [56]. Analysis of error for the statistical model can produce trustworthy estimation results for the application. Table 5 shows the score of *MAE*, *MSE* and *RMSE* for the natural fibres.

3.1.5. *Comparison for final selection of natural fibre*

An improved material selection process with consideration of the error analysis of the estimation can improve decision-making. An ascending order is used to rank the highest value of *PS* while, for the error analysis, the minimum error is prioritised. Table 6 shows the final ranking to select the best natural fibre for manufacturing the hand-brake lever application. Three types of error show a consistent measurement of the error from the estimation and observed value. Most of the ranking is constant, such as cotton is ranked as the lowest error for all types of error; the values of the error are 2.75, 11.34 and 3.37, as shown in Table 5. The same goes for other natural fibres such as coir, bamboo, sisal, kenaf, flax, ramie, bagasse and pineapple.

As shown in Table 6, the best natural fibre suggested for manufacturing the hand-brake lever was the coir. The second-best candidate is kenaf fibre, followed by cotton fibre. Flax, hemp, jute, bamboo and ramie are placed in the same rank as the fourth most suitable candidate material for this application. In contrast, using the Analytic Hierarchy Process, kenaf was found to be the best natural fibre to be hybridised with glass fibre reinforced polymer composites for a similar automotive component application [11]. Based on error analysis, cotton, coir and bamboo fibres are the top three that score minimum error on the estimation process of *PS*. Pineapple and bagasse score the highest error in the estimation of *PS*. A high

Table 5 – Error analysis using *MAE*, *MSE* and *RMSE*.

Natural fibre	y_i	$\hat{y}_i$	e_i	e_i^2	*MAE*	*MSE*	*RMSE*	*AS*
Banana	759.4	754.09	5.30	28.14	9.10	95.15	9.75	38.0
	518.6	532.37	−13.77	189.72				
	395.45	387.30	8.22	67.58				
Bagasse	280.4	249.21	31.20	973.13	27.7	803.94	28.35	286.6
	280.4	249.21	31.20	973.13				
	309.3	278.07	31.23	975.25				
	180.61	163.46	17.15	294.26				
Bamboo	249.83	243.44	6.39	40.86	3.36	18.21	4.27	8.61
	499.05	499.98	−0.93	0.86				
	500.55	499.98	0.57	0.32				
	206.01	211.56	−5.55	30.80				
Coir	246.48	247.95	−1.47	2.15	2.78	15.64	3.95	7.46
	216.07	211.89	4.19	17.51				
	203.57	199.04	4.53	20.51				
	198.99	199.04	−0.05	0.00				
	201.77	211.39	−9.62	92.56				
	257.81	255.85	1.96	3.84				
	210.48	211.39	−0.91	0.83				
	189.14	189.66	−0.51	0.27				
	217.49	215.74	1.75	3.07				
Cotton	567.05	570.11	−3.06	9.35	2.75	11.34	3.37	5.81
	570.48	570.11	0.37	0.14				
	574.5	570.11	4.39	19.29				
	465.5	471.55	−6.05	36.61				
	570.28	570.11	0.17	0.03				
	433.91	430.77	3.14	9.87				
	473.58	471.55	2.03	4.12				
Flax	961.83	978.56	−16.73	280.03	14.59	271.62	16.48	100.9
	1119.3	1134.65	−15.35	235.75				
	1238.57	1214.12	24.45	597.90				
	750.87	741.59	9.28	86.10				
	731.32	758.62	−27.30	745.24				
	1213.67	1215.06	−1.39	1.94				
	769.62	758.62	11.00	121.02				
	1190.27	1193.78	−3.51	12.31				
	774.32	758.62	15.70	246.52				
	964.02	978.56	−14.54	211.53				
	994.42	973.22	21.20	449.48				
Hemp	762.49	758.31	4.18	17.46	8.97	113.62	10.66	44.4
	886.95	903.91	−16.96	287.68				
	657.37	649.11	8.26	68.21				
	709.12	710.99	−1.87	3.50				
	627.84	649.11	−21.27	452.46				
	754.62	758.31	−3.69	13.62				
	802.87	794.71	8.16	66.57				
	768.33	758.31	10.02	100.38				
	883.84	878.43	5.41	29.26				
	772.22	762.37	9.85	97.08				
Jute	613.54	624.77	−11.23	126.08	8.17	99.48	9.97	39.2
	315.19	316.82	−1.63	2.64				
	643.72	634.72	9.00	81.03				
	620.16	598.40	21.76	473.45				
	628.22	634.72	−6.50	42.23				
	631.06	621.29	9.77	95.53				
	591.81	598.40	−6.59	43.44				
	608.06	621.29	−13.23	174.93				
	631.55	638.20	−6.65	44.24				
	624.56	621.29	3.27	10.72				
	588.26	588.45	−0.19	0.04				
Kenaf	783.2	789.08	−5.88	34.63	8.07	81.60	9.03	32.9
	422.51	406.18	16.33	266.54				
	615.2	625.35	−10.15	102.97				
	987.99	984.86	3.13	9.82				
	614.69	625.35	−10.66	113.58				

– Table 5 (*Continued*)

Natural fibre	y_i	$\hat{y}_i$	e_i	e_i^2	*MAE*	*MSE*	*RMSE*	*AS*
	989.99	984.86	5.13	26.35				
	669.84	661.96	7.88	62.10				
	987.89	984.86	3.03	9.20				
	623.54	633.99	−10.45	109.24				
Pineapple	1107.24	368.98	368.98	545,029.30	752.21	565,966	752.3	18,915
	518.14	−227.68	−227.68	556,244.49		.52	1	7
	996.5	225.85	225.85	593,898.34				
	526.44	−227.68	−227.68	568,693.96				
Ramie	766.92	740.50	26.42	697.86	16.23	371.77	19.28	135.7
	565.34	559.70	5.64	31.78				
	736.24	740.50	−4.26	18.17				
	531.17	538.01	−6.84	46.74				
	785.67	768.53	17.14	293.88				
	734.64	768.53	−33.89	1148.33				
	614.44	641.97	−27.53	757.74				
	605.14	587.73	17.41	303.21				
	747.45	740.50	6.95	48.26				
Sisal	520.86	519.44	1.42	2.03	4.90	32.72	5.72	14.5
	709.49	714.36	−4.87	23.76				
	561.63	557.69	3.94	15.54				
	597.66	601.18	−3.52	12.39				
	561.63	557.69	3.94	15.54				
	595.66	601.18	−5.52	30.47				
	694.54	681.88	12.66	160.38				
	597.66	601.18	−3.52	12.39				
	599.11	603.80	−4.69	22.00				

Table 6 – The ranking for final decision-making.

Natural fibre	Method of selection					
	MAE	*MSE*	*RMSE*	R_{AS}	R_{PS}	Final rank
Banana	8	6	6	6	7	9
Bagasse	11	11	11	11	11	11
Bamboo	3	3	3	3	8	4
Coir	2	2	2	2	6	1
Cotton	1	1	1	1	9	3
Flax	9	9	9	9	2	4
Hemp	7	8	8	8	3	4
Jute	6	7	7	6	5	4
Kenaf	5	5	5	5	4	2
Pineapple	12	12	12	12	12	12
Ramie	10	10	10	10	1	4
Sisal	4	4	4	4	10	10

number of errors may occur due to insufficient data on these types of materials.

4. Conclusion

Selection of the best natural fibre for the automotive component of a hand-brake lever was performed using a new statistical approach based on the best statistical model suggested by stepwise regression. The consideration of both high score of estimation on *PS* using best statistical model and minimum error analysis using *MAE*, *MSE* and *RMSE* through the material selection process in this study can increase precision in the decision-making process. The human error and bias in preference that may occur can also be reduced during the selection process. Twelve types of natural fibre were evaluated, considering four mechanical properties: density, tensile strength, Young's modulus and elongation at break. It was found that the tensile strength is the major parameter that influences the performance score of the alternative natural fibres in the statistical models constructed. The product design specification to manufacture a hand-brake lever was used to estimate the performance score for each natural fibre. Coir, kenaf and cotton fibre were highlighted as the top three candidate materials with maximum *PS* score and minimum error for the hand-brake lever application. Consistent findings were shown in three different types of error in the error analysis. Error analysis was used to increase the confidence level and trust in the final result through the stepwise regression. The major advantage of this method is the time and cost saving where the important parameter is considered with a strengthened statistical model. In addition, stepwise regression also gives more than one significant statistical model that can be used in different applications and allows systematic and detailed analysis with appropriate product design specification information. The most suitable natural fibre is finalised quickly using this statistical optimisation, especially when larger mechanical properties are involved.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to thank Universiti Putra Malaysia for the opportunity to conduct this study as well as Universiti Teknikal Malaysia Melaka and Ministry of Education of Malaysia for providing the scholarship award and grant scheme Hi-COE (6369107) to the principal author in this project.

REFERENCES

[1] Carley S, Duncan D, Esposito D, Graham JD, Siddiki S, Zirogiannis N. Rethinking Auto Fuel Economy Policy-Technical and Policy Suggesrions for the 2016–17 Midterm Reviews; 2016.
[2] Nishino K. Development of Fuel Economy Regulations and Impact on Automakers; 2017.
[3] Al-alawi BM, Bradley TH. Analysis of corporate average fuel economy regulation compliance scenarios inclusive of plug in hybrid vehicles. Appl Energy 2014;113:1323–37.
[4] Maciel NOR, Ferreira JB, Vieira JS, Ribeiro CGD, Lopes FPD, Margem FM, et al. Comparative tensile strength analysis between epoxy composites reinforced with curaua fiber and glass fiber. J Mater Res Technol 2018:8–12.
[5] George M, Chae M, Bressler DC. Composite materials with bast fibres: structural, technical, and environmental properties. Prog Mater Sci 2016;83:1–23.
[6] Balakrishnan P, John MJ, Pothen L, Sreekala MS, Thomas S. Natural fibre and polymer matrix composites and their applications in aerospace engineering. In: Rana S, Fangueiro R, editors. Advanced composite materials for aerospace engineering. Elsevier Ltd.; 2016. p. 365–83.
[7] Ahmad F, Choi HS, Park MK. A review: natural fiber composites selection in view of mechanical, light weight, and economic properties. Macromol Mater Eng 2014;300(1):10–24.
[8] Sapuan SM. Composite materials. In: Sapuan SM, editor. Composite materials. 1st ed. Butterworth-Heinemann: Elsevier Inc.; 2017. p. 57–93.
[9] Hambali A, Sapuan SM, Ismail N, Nukman Y. Material selection of polymeric composite automotive bumper beam using analytical hierarchy process. J Cent South Uni Technol 2010;17:244–56.
[10] Al-Oqla FM, Sapuan SM. Natural fiber reinforced polymer composites in industrial applications: feasibility of date palm fibers for sustainable automotive industry. J Clean Prod 2014;66:347–54.
[11] Mansor MR, Sapuan SM, Zainudin ES, Nuraini AA, Hambali A. Hybrid natural and glass fibers reinforced polymer composites material selection using Analytical Hierarchy Process for automotive brake lever design. Mater Des 2013;51:484–92.
[12] Mastura MT, Sapuan SM, Mansor MR, Nuraini A. Environmentally conscious hybrid bio-composite material selection for automotive anti-roll bar. Int J Adv Manuf Technol 2017;89(5–8):2203–19.
[13] Pandey JK, Ahn SH, Lee CS, Mohanty AK, Misra M. Recent advances in the application of natural fiber based composites. Macromol Mater Eng 2010;295(11):975–89.
[14] Gurunathan T, Mohanty S, Nayak SK. A review of the recent developments in biocomposites based on natural fibres and their application perspectives. Compos Part A Appl Sci Manuf 2015;77:1–25.
[15] Michalos G, Makris S, Papakostas N, Mourtzis D, Chryssolouris G. Automotive assembly technologies review: challenges and outlook for a flexible and adaptive approach. CIRP J Manuf Sci Technol 2010;2(2):81–91.
[16] Kleiner M, Geiger M, Klaus A. Manufacturing of lightweight components by metal forming. CIRP Ann – Manuf Technol 2003;52(2):521–42.
[17] Chen JX, Chouw N. Performance of natural fibre reinforced polymer-concrete bridge piers in earthquakes. In: Proceeding of the second international conference on performance-based and life-cycle structural engineering (PLSE 2015). 2015. p. 1329–36.
[18] Amirjan M, Khorsand H, Siadati MH, Eslami Farsani R. Artificial neural network prediction of Cu–Al_2O_3 composite properties prepared by powder metallurgy method. J Mater Res Technol 2013;2(4):351–5.
[19] Dursun M, Karsak EE. A QFD-based fuzzy MCDM approach for supplier selection. Appl Math Model 2013;37(8):5864–75.
[20] Noryani M, Sapuan SM, Mastura MT, Zuhri MYM, Zainudin ES. A statistical framework for selecting natural fibre reinforced polymer composites based on regression model. Fibers Polym 2018;19(5):1039–49.
[21] Jahan A, Edwards KL, Bahraminasab M. Multi-criteria decision analysis for supporting the selection of engineering materials in product design. second edn Butterworth/Heinemann/Oxford: Elsevier; 2016.
[22] Aruldoss M, Lakshmi TM, Venkatesan VP. A survey on multi criteria decision making methods and its applications. Am J Inf Syst 2013;1(1):31–43.
[23] Velasquez M, Hester PT. An analysis of multi-criteria decision making methods. An analysis of multi-criteria decision making methods. Int J Oper Res 2013;10:56–66.
[24] Vinodh S, Jeya Girubha R. PROMETHEE based sustainable concept selection. Appl Math Model 2012;36(11): 5301–8.
[25] Liu H-C, Mao L-X, Zhang Z-Y, Li P. Induced aggregation operators in the VIKOR method and its application in material selection. Appl Math Model 2013;37(9): 6325–38.
[26] Mastura MT, Sapuan SM, Mansor MR, Nuraini AA. Materials selection of thermoplastic matrices for 'green' natural fibre composites for automotive anti-roll bar with particular emphasis on the environment. Int J Precis Eng Manuf Technol 2017;5(1):111–9.
[27] Mayyas A, Shen Q, Mayyas A, Abdelhamid M, Shan D, Qattawi A, et al. Using quality function deployment and analytical hierarchy process for material selection of body-in-white. Mater Des 2011;32(5):2771–82.
[28] Wahab OA, Bentahar J, Otrok H, Mourad A. A survey on trust and reputation models for Web services: single, composite, and communities. Decis Support Syst 2015;74:121–34.
[29] Barfod MB, Salling KB, Leleur S. Composite decision support by combining cost–benefit and multi-criteria decision analysis. Decis Support Syst 2011;51:167–75.
[30] Ashby MF, Bréchet YJM, Cebon D, Salvo L. Selection strategies for materials and processes. Mater Des 2004;25(1):51–67.

[31] Khmaissia F, Frigui H, Sunkara M, Jasinski J, Garcia AM, Pace T, et al. Accelerating band gap prediction for solar materials using feature selection and regression techniques. Comput Mater Sci 2018;147:304–15.
[32] Noryani M, Sapuan SM, Mastura MT, Zuhri MYM, Zainudin ES. Material selection criteria for natural fibre composite in automotive component: a review. In: IOP Conf Ser Mater Sci Eng. 2018. p. 368.
[33] Hasić F, De Smedt J, Vanthienen J. Augmenting processes with decision intelligence: principles for integrated modelling. Decis Support Syst 2018;107:1–12.
[34] Mustafa A, Abdollah MF, Shuhimi FF, Ismail N, Amiruddin H, Umehara N. Selection and verification of kenaf fibres as an alternative friction material using weighted decision matrix method. Mater Des 2015;67:577–82.
[35] Alkbir MFM, Sapuan SM, Nuraini AA, Ishak MR. Fiber properties and crashworthiness parameters of natural fiber-reinforced composite structure: a literature review. Compos Struct 2016;148:59–73.
[36] Pickering KL, Efendy MGA, Le TM. A review of recent developments in natural fibre composites and their mechanical performance. Compos Part A Appl Sci Manuf 2015;83:98–112.
[37] Majeed K, Jawaid M, Hassan A, Abu Bakar A, Abdul Khalil HPS, Salema AA, et al. Potential materials for food packaging from nanoclay/natural fibres filled hybrid composites. Mater Des 2013;46:391–410.
[38] Petrone G, Meruane V. Mechanical properties updating of a non-uniform natural fibre composite panel by means of a parallel genetic algorithm. Compos Part A Appl Sci Manuf 2017;94:226–33.
[39] Misnon MI, Islam MM, Epaarachchi JA, Lau K. Potentiality of utilising natural textile materials for engineering composites applications. Mater Des 2014;59:359–68.
[40] Satari SZ, Chuan ZL, Misni F, Moslim NH, Muhammad N, Zakaria R. Linear regression & correlation. In: Pahang PUM, editor. Applied statistics module. 2015. p. 268–311.
[41] Rezayat M, Gharechomaghlu M, Mirzadeh H, Parsa MH. A comprehensive approach for quantitative characterization and modeling of composite microstructures. Appl Math Model 2016;40(19–20):1–6.
[42] Muhammad I, Sana Ullah SM, Han DS, Ko TJ. Selection of optimum process parameters of biomachining for maximum metal removal rate. Int J Precis Eng Manuf – Green Technol 2015;2(4):307–13.
[43] Noryani M, Sapuan SM, Mastura MT, Zuhri MYM, Zainudin ES. Stepwise regression for kenaf reinforced polypropylene composite. Proc Mech Eng Res Day 2018;2018:48–9.
[44] Abir MR, Kashif SM, Razzak A. Tensile and statistical analysis of sisal fibers for natural fiber composite manufacture. Adv Mater Res 2015;1115:349–52.
[45] Glória GO, Teles MCA, Lopes FPD, Vieira CMF, Margem FM, Gomes MA, et al. Tensile strength of polyester composites reinforced with PALF. J Mater Res Technol 2017;6(4):401–5.
[46] Facca AG, Kortschot MT, Yan N. Predicting the tensile strength of natural fibre reinforced thermoplastics. Compos Sci Technol 2007;67:2454–66.
[47] Chen G, Ozden UA, Bezold A, Broeckmann C. A statistics based numerical investigation on the prediction of elasto-plastic behavior of WC–Co hard metal. Comput Mater Sci 2013;80:96–103.
[48] Nasir AAA, Azmi AI, Khalil ANM. Measurement and optimisation of residual tensile strength and delamination damage of drilled flax fibre reinforced composites. Measurement 2015;75:298–307.
[49] Le N, Van Den Oever M, Budtova T. Statistical analysis of fibre size and shape distribution after compounding in composites reinforced by natural fibres. Compos Part A 2011;42:1542–50.
[50] Ismail AE, Aziz MACA. Tensile strength of woven yarn kenaf fiber reinforced polyester composites. J Mech Eng Sci 2015;9:1695–704.
[51] Noryani M, Sapuan SM, Mastura MT. Multi-criteria decision-making tools for material selection of natural fibre composites: a review. J Mech Eng Sci 2018;12(1): 3330–53.
[52] Patel MV, Sarawade SS. Design and weight optimization of parking brake lever. Proc Int Conf Emerging Trends Eng Manage Res 2017;3(8):367–74.
[53] Gasparini VM, De Castro LFA, Quintas ACB, Moreira VEDS, Viana AO, Andrade DHB. Thermo-chemical model for blast furnace process control with the prediction of carbon consumption. J Mater Res Technol 2017;6(3): 220–5.
[54] Jang DY, Jung J, Seok J. Modeling and parameter optimization for cutting energy reduction in MQL milling process. Int J Precis Eng Manuf – Green Technol 2016;3(1):5–12.
[55] Al-Oqla FM, Sapuan SM, Ishak MR, Nuraini AA. A decision-making model for selecting the most appropriate natural fiber – polypropylene-based composites for automotive applications. J Compos Mater 2015;50(4): 543–56.
[56] Rabizadeh E, Saboor Bagherzadeh A, Rabczuk T. Adaptive thermo-mechanical finite element formulation based on goal-oriented error estimation. Comput Mater Sci 2015;102:27–44.