

Data mining of optimum conditions to acquire bamboo micro-fiber with mechanical methods

K. Ogawa, T. Hirogaki, E. Aoyama & T. Katayama

Department of Mechanical Engineering, Doshisha University, Japan

Abstract

In the present report, we purposed a new mechanical method to obtain micro fibers from the bamboo. The data mining method was also used in order to elucidate the fiber quality from the factors obtained. As a result, it is clear that a shearing mechanical method is effective to obtain fine micro fibers from the bamboo.

Keywords: bamboo micro fiber, shearing mechanical method, data mining, natural board, hot press forming.

1 Introduction

Recently, problems have emerged because of the negative impact of the industrial products on the environment. Bamboo fibers have attracted attention because of a fast growth among various kinds of renewable natural fibers. There are previous reports dealing with the bamboo fiber reinforced plastics (BFRP) [1]-[4]. However, the used plastics, including BFRP, also impact the environment. We, therefore, have attracted attention to the natural boards made of micro bamboo fibers using the lignin as a matrix, formed by a hot press forming. An effective method to obtain the micro bamboo fibers is needed in order to a fine natural press forming boards. In the present report, first, we proposed a new mechanical method to obtain the micro fibers from a bamboo. However, the factors associated with the obtained fiber length remain not completely understood. It was necessary to apply the method to large amounts of data in order to identify quantitative tendencies. Second, we also applied a data mining method for obtained fiber quality data, with special attempts to elucidate the factors that influence the obtained fiber quality, using multiple regression and variable clustering analyses. That is, data mining methods to discover new



scientific knowledge from large amounts of stored data have been researched actively [5]. These new statistical methods are now widely used as large-scale data processing systems in order to mine valuable information.

2 Data mining

Data mining is a data processing technology which can be used to mine out valuable information from large-scale databases. In conventional statistical analyses, data is observed directly for data noise, which is then removed based on the judgment of an analyst. A feature of data mining is a method of removing noise by data cleansing. In conventional statistical analyses, the data must be sampled to match the purpose of the analysis. On the other hand, data mining can deal with data, which is stored without any plans. Moreover, it is possible to recycle the analysed result to attain new knowledge.

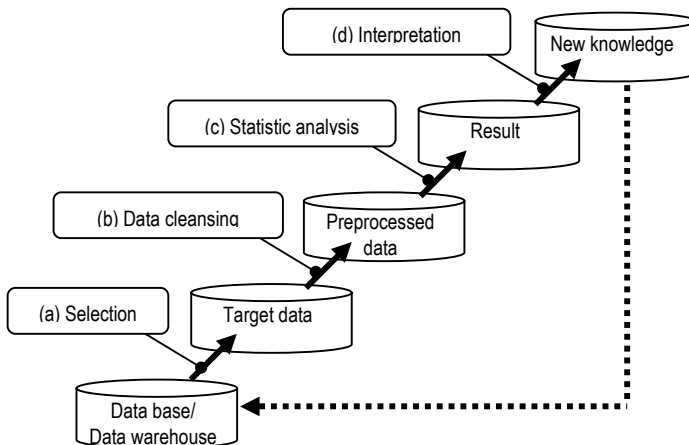


Figure 1: Data mining processes.

The flow of data mining is shown in Fig.1. It is possible that the data includes disunity of the unit or contains outlier values by such as from measurement or data-input mistakes. Such things are detrimental to quantitative evaluation of analysis results. Therefore, in order to remove these problems data, normalization of data and the outlier value test were carried out as data cleansing. A Smilnof Grabs statistical evaluation to detect outlier values was done as follows. First, T_i was calculated with eqn. (1) for X_{ij} that was the maximum (or minimum) value of measured data X_i . Second, T_i was compared to judgment point t at a significance level of 0.05 ($t=3.75$). If T_i was not lower than t , T_i was judged to be an outlier value. Next, this process was repeated in a similar way for a number of data $n-1$ until T_i was lower than t .

$$T_i = |X_{ij} - \bar{X}_i| \sqrt{U_i} \quad (1)$$

Here, $\overline{X_i}$ is the mean value of X_i , and U_i is the variance given by eqn. (2).

$$U_i = \frac{\sum_{i=1}^n (X_i - \overline{X_i})^2}{n-1} \quad (2)$$

Here, n is the number of data.

Multivariate analysis was applied for the statistical analysis method for preprocessed data. In multivariate analysis, a proper analysis method must be selected in proportion to the data type, as shown Table 1. Multiple regression analysis was therefore adopted because the predictor variables were numerical data such as the bamboo chips conditions, and the criterion variable was also numerical data such as weight of micro fibers obtained. However, the reliability of the factor analysis result becomes lower when a high correlation exists among predictor variables in multiple regression analysis. Thus, variable cluster analysis was applied as data cleansing in order to find the structure of predictor variables. The variable cluster analysis was performed as follows. Distance L_{gh} between variable g and variable h is defined with correlation coefficient r_{gh} between variable g and variable h as eqn. (3), and a cluster is paired with another cluster minimized L_{gh} . Here, L_{gh} becomes the minimum distance 0 when correlation coefficient r_{gh} is 1, and becomes the maximum distance 2 when correlation coefficient r_{gh} is -1 , because $-1 < r_{gh} < 1$. Ward's method was utilized for the calculation of distance after clusters were united. Ward's method is a method to unite pairs of clusters whose distance becomes minimum by minimizing D_{ko} , as shown in eqn. (4), so as to minimize the sum of squares of the centroid of clusters produced by amalgamation.

$$L_{gh} = 1 - r_{gh} \quad (3)$$

$$D_{ko}^2 = \frac{(m_p + m_o)D_{po}^2 + (m_q + m_o)D_{qo}^2 - m_o D_{pq}^2}{m_k + m_o} \quad (4)$$

Here, D_{ko} is distance between D_k and D_o . D_k is a cluster generated by uniting cluster D_p and cluster D_q . D_o show other clusters. m is the number of data.

Table 1: Analysis method and data type.

Analysis method	Used data type	
	Criterion variable	Predictor variable
Multiple regression analysis	Numerical	Numerical
Quantification theory type 1	Numerical	Category
Discriminant analysis	Category	Numerical
Quantification theory type 2	Category	Category

Backward elimination was performed for selection of predictor variables in the multiple regression analysis. The significant predictor variables can be



selected in backward elimination by judging with an F-value that shows the significance of predictor variables. Backward elimination was done as follows. First, the F-value was calculated for each predictor variable. Second, the predictor variable that has the minimum F-value of all was removed. Third, the F-value was calculated again for remained predictor variables and the predictor variable that has the minimum F-value of them was combed out. Next, this process was repeated in a similar way until that the F-value of all selected predictor variables became 2.0 or more, when all predictor variables were accepted generally as significance for the criterion variable.

3 **Acquirement of stored data for data mining**

3.1 **Machining condition for stored data**

Figure 2 shows the overview of machining equipment used to obtain the stored data. The equipment is called SHOKUSENKI made of Shinko Engineering Co., Ltd. Table 2 shows the properties of the equipment. The uniformed bamboo chips thrown into the hopper are carried to the end of the cylinder by the screw rotating motion. The rotating cutter in synchronizing with the screw chops the pressurized bamboo chips. Smaller bamboo chips are pushed out through many holes in poriferous slit to become bamboo fibers by shearing processes. The bamboo fibers are sorted out to three groups as follows using two kinds of sieves.

- Group A: Samples through the opening of 45μm square.
- Group B: Samples through the opening of 500μm square but of 45μm square.
- Group C: Samples not through the opening of 500μm square.

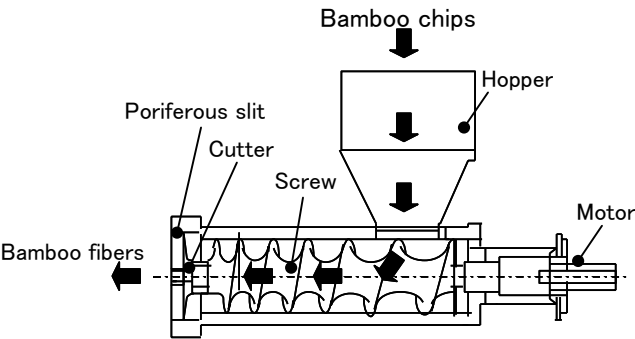


Figure 2: Overview of machining equipment.

Table 2: Properties of machining equipment.

Rotation speed of screw and cutter (rpm)	120
Number of cutting edge	2
Hole diameter of poriferous slit (mm)	7
Hole pitch of poriferous slit (mm)	8

Table 3: Experimental condition of bamboo chips.

Provided amount a experiment (g)	200
Cross section (mm square)	5
Length (mm)	25,50
Water content (wt%)	12,50,74
Providing speed (g/min)	12,20,40

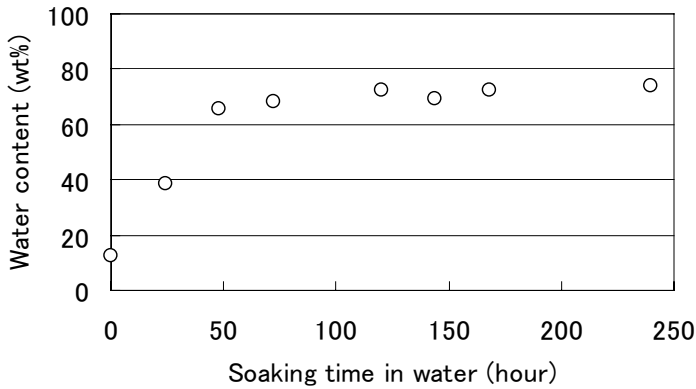


Figure 3: Relationship between soaking time in water and water content, in the case bamboo chips length 50mm.

The bamboo chips used are carved out from MOUSOUTIKU bamboo growing naturally in Japan. Table 3 shows the experimental conditions of the bamboo chips. The total amount of weight at one experiment thrown in is 200 g constant. The length l of the chip, the water content ε , and the providing speed v in hopper were changed. The water content is determined as eqn. (5).

$$\varepsilon = \frac{W - W_0}{W_0} \times 100 \quad (5)$$

Here, W is the weight of bamboo chips dried in room and W_0 , which is 200 g constant, is the weight of bamboo chips soaked in water. Fig. 3 shows the relation between the soaking time in water and the water content in case of using the 50mm length bamboo chips. In the present report, the soaking time was decided to control the water content from Fig. 3. So, the 25mm length bamboo chips were obtained by cutting in half the 50mm length bamboo chips after soaking.

3.2 Observation of bamboo fiber figuration

Figure 4 shows the overview of the bamboo chips provided and the bamboo fibers obtained. It can be seen that the bamboo chips are changed into small bamboo fibers. Fig. 5 shows SEM (Scanning Electron Micrograph) images of



bamboo fibers sorted out with sieves. It is clarified the μm order size fibers as sort of powder can be obtained in group A. Moreover, it can be seen that about $10\mu\text{m}$ diameter monofilaments also exist shown in Fig. 5 (a). In group B shown in Fig. 5 (b), the mm order size length fiber bundles can be seen. It is clear that the large size bulk fibers belong to group C.

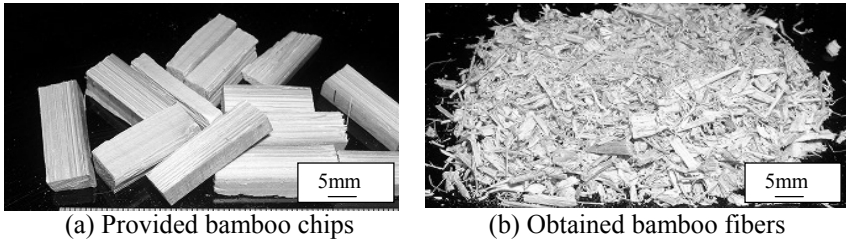


Figure 4: Provided bamboo chips and obtained bamboo fibers.

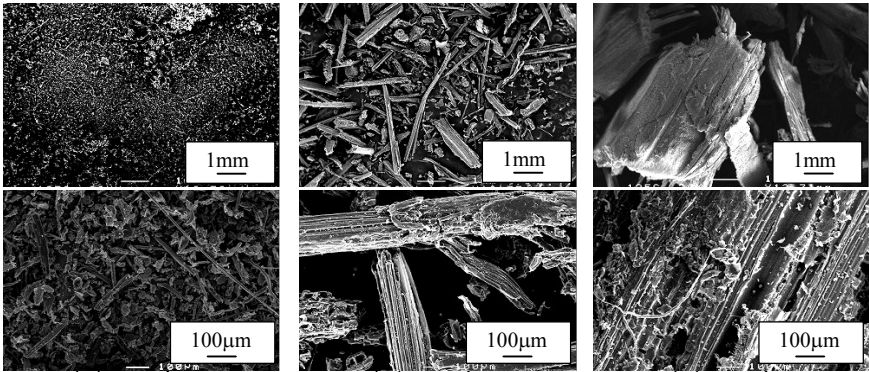


Figure 5: SEM micrographs of obtained bamboo fibers.

3.3 Influence of bamboo chip condition on bamboo fiber weight ratio

It is clear that various size fibers can be obtained from bamboo chips in section 3.2. Therefore, the investigation was carried out to evaluate the influences of the bamboo chip conditions on the bamboo fiber weight ratio. Here, the bamboo fiber weight ratio R means the ratio between the weight of each group sorted with sieves and the total weight of bamboo fibers obtained.

Figure 6 shows the relationship between water content and the weight ratio in case of the of bamboo chips length 50 mm and the providing speed 40 g/min. It is clear that the weight ratio of group C decrease with the water content increasing. Therefore, the weight ratio of group A and B grows with the water content increasing. Figure 7 shows the relationship between the providing speed and the weight ratio in case of the length of bamboo chips 50 mm and the water content 74 wt%. It seems that there is little influence of the providing speed on the weight ratio.

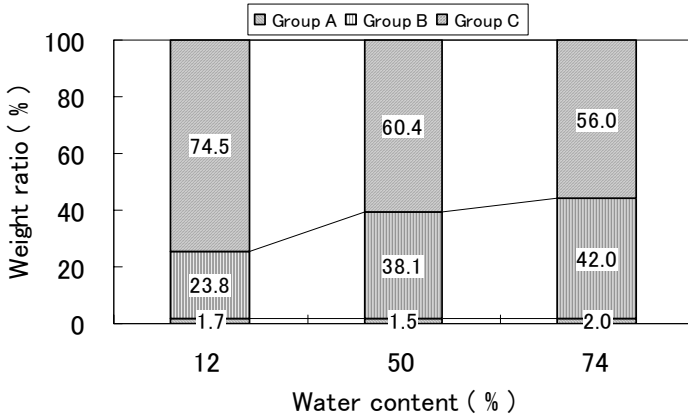


Figure 6: Relationship between water content and weight ratio.

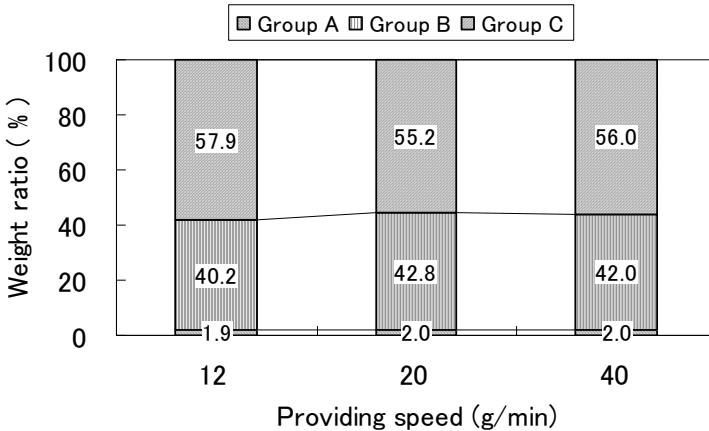


Figure 7: Relationship between providing speed and weight ratio.

4 Data mining of optimum condition

4.1 Target data for analysis

Smaller size fibers are needed to obtain the superior product so as FRP. Therefore, treated parameters for data mining are the group A weight ratio R_A as the criterion variable, the water content ε , the providing speed v (g/min), the length l (mm) of the bamboo chips and the index α as predictor variables. Here, the index α indicates the provided amount rate for total amount of bamboo chips in one experiment. The weight ratio was evaluated in three steps from the start to the end of providing. Therefore, the values index α are 66.7, 133.3 and 200.0 corresponding respectively to the first, the second and the third step.



4.2 Data cleansing

All target data were used because it was clarified that there is no outlier values from a result of a Smilnof Grabs statistical evaluation. Table 4 shows the single correlation coefficients among variables. The sign of the single correlation coefficient between the length l and the weight ratio R_A is minus. Therefore, l' was determined as the inverse number of l . That is, l' means the shortness of bamboo chips. As a result, the single correlation coefficient between the length l' and the weight ratio R_A became plus (0.0562).

Table 4: Single correlation coefficients among variables.

	R_A	l	ε	v	α
R_A	1.0000	-0.0562	0.2503	0.0950	0.6990
l	-0.0562	1.0000	-0.2163	-0.0937	0.0000
ε	0.2503	-0.2163	1.0000	0.0608	0.0000
v	0.0950	-0.0937	0.0608	1.0000	0.0000
α	0.6990	0.0000	0.0000	0.0000	1.0000

4.3 Data structure analysis by variable cluster analysis

Figure 8 shows the result of the variable cluster analysis from Table 4. In the tree diagram in Fig. 8, it means that variables combined at lower sides have higher correlation. It is found that there is the highest correlation between l' and ε .

If multiple regression analysis is performed using both l' and ε , no solution can be obtained or the reliability of analysis becomes lower because of multicollinearity. Therefore, either l' or ε must be removed from the predictor variables. A single correlation analysis for weight ratio R_A was then carried out. As a result, ε (single correlation coefficient 0.2503) was adopted, and l' (single correlation coefficient 0.056) was removed.

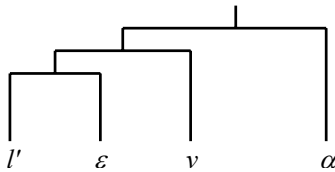


Figure 8: Cluster tree of predictor variables.

4.4 Factor analysis by multiple regression analysis

The multiple regression analysis is carried out with the water content ε , the providing speed v and the index α as predictor variables. As a result, the factors selected to be significant for R_A are ε and α . The partial regression coefficient of α is 0.6990 and that of ε is 0.25028. It is found that the sign of the partial regression coefficient of these factors is in good agreement each other. When the signs are different, it is necessary to re-evaluate the correlation among the



variables with variable cluster analysis, because the multi-colinearity exists between the variables. Therefore, it is obvious that the water content ε and the index α influence the weight ratio R_A . That is, increasing ε and α is effective to increase R_A . Moreover, α influences largely on R_A because the partial regression coefficient of α is about three times than that of ε .

4.5 Interpretation of α as a significant factor

The index α has not been examined as a factor of the weight ratio R_A heretofore. However, it is necessary to interpret the result to show the validity of statistically analysed data. The influence of the index α therefore investigated.

Figure 9 shows temperature distributions observed by a thermography, unless Fig. 9(a) indicated the view of equipment. It can be seen that the temperature increase according to the increasing the time. The index α increase as the time from the start going on. Therefore, it is found out that the increase of temperature is effective to increase weight ratio R_A . Though the water content of bamboo chips has been evaluated previously, a significant result of data mining is that a new factor hidden in the data was revealed.

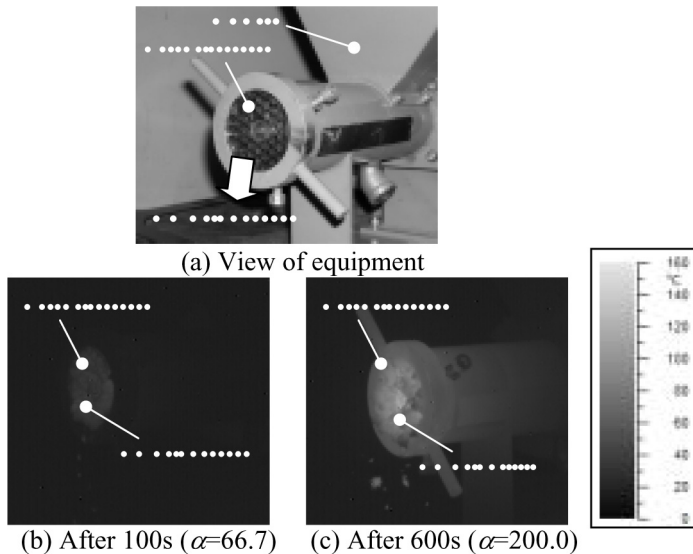


Figure 9: Temperature distribution observed by thermography.

5 Hot press forming using only micro bamboo fibers

It is clear that various size bamboo fibers can be obtain from the bamboo chips with a purposed mechanical method. So, hot press forming was carried out using obtained micro bamboo fibers utilizing lignin around the fibers as a matrix. The press temperature decided to 170 centigrade degrees considering the melting temperature of the bamboo lignin as a matrix. The forming pressure is 20 MPa

and the forming time is 10 minutes. Figure 10 shows the bamboo micro fiber board formed by hot press process. It is clear that the hot press forming using only bamboo micro fibers is possible under adequate temperature conditions.

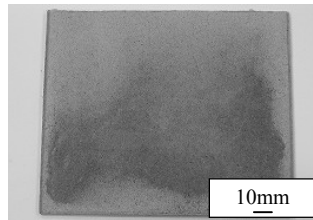


Figure 10: Bamboo fiber board using lignin as matrix.

6 Conclusion

Optimizing condition to acquire bamboo micro fibers using a mechanical method was researched using a data mining, which has been systematized as a statistical technique for discovering scientific information. The following conclusions were obtained.

- (1) It is possible to acquire fine micro bamboo fibers using mechanical shearing method.
- (2) Using the data mining method, it is clear that the water content of bamboo chips and the temperature have influence on the weight ratio of micro fibers.
- (3) Hot press forming using only bamboo micro fibers is possible under adequate temperature conditions.

References

- [1] H. Martikka and J. Katajisto, Study of Natural Fiber Reinforced Biodegradable Composite Materials for Designing Optimally Sustainable Products, *Proc. of the 23rd Riso International Symposium on Materials Science*, pp. 251-258, 2002.
- [2] Moe Moe Thwe and Kin Liao, Characterization of Bamboo-glass Fiber Reinforced Polymer Matrix Hybrid Composite, *Journal of Materials Science Letters*, 19, pp. 1873-1876, 2000.
- [3] A. Varada RaJulu, S. Allah Baksh, G. Ramachandra Reddy and K. Narasimha Chary, Chemical Resistance and Tensile Properties of Short Bamboo Fiber Reinforced Epoxy Composites, *Journal of Reinforced Plastics and Composites*, Vol. 17, No. 17, pp. 1507-1511, 1998.
- [4] Seema Jain, U. C. Jindal and Rakesh Kumar, Development and fracture mechanism of the bamboo/polyester resin composite, *Journal of Materials Science Letters*, 12, pp. 558-560, 1993.
- [5] Usama Fayyad, Gregory Piatetky-Shapio and Padhraic Smyth, From Data Mining to Knowledge Discovery in Databases, *AI Magazine*, pp.37-54, 1996.

