SELECTION OF VEGETATIVE AND GENERATIVE CHARACTERS OF ARABICA COFFEE BY USING SEQUENTIAL PATH ANALYSIS AND STRUCTURAL EQUATION MODELS

Seleksi Karakter Vegetatif dan Generatif Kopi Arabika melalui Penggunaan Analisis Lintasan Bertahap dan Model Persamaan Struktural

EDI WARDIANA dan DIBYO PRANOWO

Balai Penelitian Tanaman Industri dan Penyegar
Jalan Raya Pakuwon km 2 Parungkuda, Sukabumi 43357
email : edwardiana@yahoo.com

(Diterima: 17-1-2013; Direvisi: 2-6-2014; Disetujui: 5-5-2014)

ABSTRACT

Interrelations among vegetative, generative and yield characters in coffee plantation is important in breeding and selection programs. Interrelationships among these characters are the causal model and it can be analyzed by Sequential Path Analysis (SPA). This research was carried out at Pakuwon Experimental Station located at 450 m above sea level, in Latosol soil type with B type of climate, from December 2010 to April 2012. The objectives of this research is to analyze the direct or indirect influence of several vegetative and generative characters on yield characters of Arabica coffee through the application of SPA and Structural Equation Models (SEM). The observation method with systematic sampling on 40 Arabica coffee plants of Kartika 1 and Kartika 2 varieties were used in this study. The results showed that the number of cherries of Arabica coffee var. Kartika 1 and 2 population in Pakuwon Experimental Station were affected directly by generative and indirectly by vegetative characters. Plant height and stem girth can be used as positive selection criteria for high yielding at vegetative phase, while the character of width canopy can be used as negative selection criteria. In generative phase, the character of quantity of productive branches and the amount of berries and flower cluster can be used as positive selection criteria.

Keywords: Coffea arabica, selection, vegetative, generative, yield, sequential path analysis, structural equation models

INTRODUCTION

Interrelationship of vegetative and generative characters, as well as yield in plants is important in breeding and selection programs. By through this information, the selection of high production of annual crops can be carried out in earlier phase by low cost. To investigate the interrelationship among plant characters, a correlation method is often used as a general approach. However, the correlation is only limited to find out the relationship between two variables and it could not explain the causal relationship among variables. The causal relationship between two variables can be analyzed by using regression models. There are limitations of this analysis due to inability to explain the phenomenon of direct and indirect effect between two variables, because all independent and dependent variables were analyzed simultaneously.

Sewal Wright, found a theory or method of multivariate analysis, known as Path Analysis (PA). This analysis is a combination of regression and correlation. It analyze the phenomenon of direct and indirect effects of independent variables to dependent variables. In general regression model, a direct effect is standardized coefficient or better known as "beta coefficient" or "beta weight" (GASPERSZ, 1992).

In PA with conventional approaches, when the vegetative and generative characters as independent variables analyzed simultaneously on yield character as
dependent variable, the researchers sometimes lose information about the direct relationship between vegetative and yield characters, therefore, the purpose to accelerate the selection cycle was not achieved. In this analysis, all predictor variables were analyzed simultaneously on the first-order, furthermore direct and indirect effect on the response variable were examined (MOKHTASSI et al., 2006). Therefore, the Sequential Path Analysis (SPA) is needed to find out the causal relationship between vegetative, generative and yield characters. The formulation model are sequentially based on the plant development cycle, starting from vegetative characters followed by generative and yield characters. Through this model, between dependent and independent variables was analyzed to form an endogenous variables, that is commonly referred to as mediation or moderation (GARSON, 2008; VAN ACKER and WITLOX, 2010).

Theoretically, based on plant development cycle, it can be known that vegetative characters precedes generative characters, as well as vegetative and generative characters for yield characters. Therefore, in the SPA model the vegetative characters can be used as independent variable for the generative and yield characters, as well as generative characters for yield characters. The relationships between these characters is a causal models and is generally described as “single-headed arrow” and the notation of "pi", is not the correlation models that generally described by “double-headed arrow” and the notation of "ri" (GARSON, 2008; ANTONAKIS et al., 2010).

The result of SPA can be confirmed by the other multivariate analysis, namely Structural Equation Modeling (SEM). SEM is the confirmatory analysis models that can analyze the interrelationships among complex variables, and the construction of their specification models based on the theories and research results that had been well-developed (PUI-WA LEI and QIONG WU, 2007). The PA and SPA models have also been used previously by MARANDU et al. (2004) in coffee; FIROUZABADI et al. (2011) in sugar beet; MALEKI et al. (2011) in tobacco; and MOHAMMADI et al. (2012) in safflower. The use of SPA and SEM is very useful for the selection and breeding programs of coffee plant. The objectives of this study were to analyze the vegetative and generative characters affecting the yield of Arabica coffee by using the SPA and SEM.

MATERIALS AND METHODS

The research was carried out using observation method at Pakuwon Experimental Station, 450 m above sea level (asl), Latosol type of soil and B type of climate (SCHMIDT and FERGUSSON), beginning from December 2010 to April 2012. The plant materials used in this study were two varieties of Arabica coffee (KARTIKA 1 and KARTIKA 2; 40 plants for each variety) which were planted in December 2010 at spacing of 2 x 3 m, and using kluwak plants (Pangium edule Reinw. Ex Blume) as shade trees with plant spacing of 6 x 6 m. The data collected from 20 samples of each variety (50% of the total population) determined by systematic sampling.

The variables observed consisted of 10 characters (6 of vegetative character, 2 of generative characters, 1 of yield components character, and 1 of yield character) as follows:

Vegetative characters:
1. Plant height (PH); measured from the ground to the top of shoots (cm),
2. Stem girth (SG); measured at 5 cm distance from the ground surface (cm),
3. Canopy width (CW); the average value of canophy width measured at the outer canopy of north-south and west-east (cm),
4. Canopy shape (CS); the ratio of PH on CW,
5. Number of unproductive branch per tree (NUPB); counted all unproductive of primary branches per tree,
6. Number of internodes per branch (NI); the average number of internodes per branch from three samples of branches started from the bottom.

Generative characters:
7. Number of productive branch per tree (NPB); counted all the productive primary branches per tree,
8. Number of fruit and flower clusters per branch (NFC); the average number of flowers and fruits per cluster from three samples of branches started from the bottom.

Yield component characters:
9. Number of cherries per branch (NCB); the average number of cherries per branch from three samples of branches from the bottom.

Yield characters:
10. Number of cherries per tree (NCT); an estimate value by multiplication of NCB and NPB, and then multiplied by 80%.

The data were analyzed by the SPA method and their formulation models are sequentially based on the plant development cycle as follows : (1) yield character as dependent variable and all of vegetative, generative and yield component characters as independent variables, (2) yield component character as dependent variable and all of...
vegetative and generative characters as independent variables, and (3) generative characters as dependent variable and all of vegetative characters as independent variables. In each phase of SPA all predictor variables are organized for the first-, second-, and third-order analysis respectively as has been done by FIROUZABADI et al. (2011); MALEKI et al. (2011); and MOHAMMADI et al. (2012). Whereas, in a conventional approach of PA, all predictor variables are organized as first-order and then their direct and indirect effect on the dependent variable were analyzed (MOKHTASSI et al., 2011).

To avoid the multicollinearity effects, all independent variable in the regression models were selected by stepwise method (GASPERSZ, 1992; ZHONG et al., 2012). Subsequent analysis, the model that had been formed by SPA then confirmed by SEM. Finally, to investigate the advantage of SPA then it was compared to the conventional approach of PA. The data analysis was performed by statistical software of SPSS and AMOS version 17 (WUENSCH, 2006).

### RESULT AND DISCUSSION

#### First-order Analysis

Regression in first-order analysis (Appendix - step1) shows that among of all vegetative, generative and yield components of Arabica coffee as independent variables, there were only two characters selected namely NCB and NPB. Both characters have positive direct effect on NCT, and the magnitude of these effect more lower than the correlation value (Table 1; Figure 1). Correlation is different from the causal, because correlation can not describe the causal relationship between the response variable with the predictor variables (MOHAMMADI et al., 2012), and can not explain the direct and indirect associations (OPGEN-RHEIN and STRIMMER, 2007). The correlation only indicates the relationship between the two variables, whereas the causation can define and measure the magnitude of the causal-effect of one variable to another. KARADAG (2012), argued that the correlation coefficient does not always define the cause-effect relationship. This can be accurately done only with path analysis.

#### Second-order Analysis

Regression in second-order analysis between yield component characters (NCB) as dependent variables and vegetative and generative characters (NFC and NPB) simultaneously as independent variables, there was only one character selected namely NFC (Appendix - step 2). The NFC has positive direct effect on the NCB and equal to their correlation values (Table 2; Figure 2). The other research in coffee plants had been carried out by PRIYONO and SUMIRAT (2012), showed that the number of flower cluster per tree has positive correlation with the number of cherries per cluster and total weight of cherries and beans.
Table 2. Direct and indirect effect of NFC on NCB

<table>
<thead>
<tr>
<th>Code of Characters</th>
<th>Direct effect (p)</th>
<th>Indirect effect</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFC</td>
<td>0.67 **</td>
<td>0.00</td>
<td>0.67 **</td>
</tr>
</tbody>
</table>

Notes: ** significant at 1% level

Keterangan: ** nyata pada taraf 1%

---

Third-order Analysis

Regression in third-order analysis between generative characters (NFC) as dependent variables and all vegetative characters (NUPB, PH, SG and CW) and the other of generative characters as antecedent (NPB), there were only two characters selected, NPB and CW (Appendix – step 3a). NPB has positive direct effect and CW has negative direct effect on NFC (Table 3; figure 3), while NUPB, PH and SG have no direct effect on NFC. These results were consistent with the other research result in coffee plants which showed a positive correlation between number of productive branch per tree and number of flower cluster per tree (PRIYONO and SUMIRAT, 2012). While, the results of the other studies showed that there were no relationship between girth and number of fruit and flowers cluster (ANIM-KWAPONG and ADOMAKO, 2010; ANIM-KWAPONG et al., 2011).

The other paths in third-order analysis were stepwise regression of NPB as dependent variable and NUPB, PH, SG, and CW as independent variables, showed that there were only three characters selected namely NUPB, SG, and PH (Appendix – step 3b). NUPB has negative direct effect, while SG and PH have positive direct effect on NPB (Table 4; Figure 4). Stem is one of carbohydrates source that support the generative growth and yield of coffee plant. On the other hand, the higher the plants the greater the probability to form productive branch. If the plant height is assumed to be constant, then the number of unproductive

Table 3. Direct and indirect effect of NPB and CW on NFC

<table>
<thead>
<tr>
<th>Code of Characters</th>
<th>Direct effect (p)</th>
<th>Indirect effect</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPB</td>
<td>0.58 **</td>
<td>-0.06</td>
<td>0.52 **</td>
</tr>
<tr>
<td>CW</td>
<td>-0.33 *</td>
<td>0.12</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

Notes: * and ** significant at 5 and 1% level respectively

Keterangan: * dan ** masing-masing nyata pada taraf 5 dan 1%
branches would be contrary to the number of productive branches. MARANDU et al. (2004), argued that the stem diameter and plant height are the important characters in the selection of coffee plants, because they have positive correlation with yield characters. Similarly, the opinion of MONTAGNON et al. (2001), which states that the selection for vigoritas of coffee plants can be done on the girth and plant height. The other research result also showed a positive relationship between the characteristics of stem diameter as well as the length and diameter of the primary branch of Robusta coffee (NIKHILA et al., 2008).

Table 4. Direct and indirect effect of NUPB, SG, and PH on NPB

<table>
<thead>
<tr>
<th>Code of Characters</th>
<th>Direct effect (p)</th>
<th>Indirect effect</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUPB</td>
<td>-0.54 **</td>
<td>-0.06</td>
<td>-0.60 **</td>
</tr>
<tr>
<td>SG</td>
<td>0.46 **</td>
<td>-0.09</td>
<td>0.37 *</td>
</tr>
<tr>
<td>PH</td>
<td>0.22 *</td>
<td>0.30</td>
<td>0.52 **</td>
</tr>
</tbody>
</table>

Notes: * and ** significant at 5 and 1% level respectively

Keterangan: * dan ** masing-masing nyata pada taraf 5 dan 1%
Analysis of Confirmation

Confirmatory analysis was performed by SEM and their formulation models based on the merging of the third-order of SPA that discussed previously. These formulation models are presented in Figure 5. The estimate value of causal relationship between the two characters by SEM are presented in Table 5.

The estimate values between the two analysis models (SPA and SEM) were similar. This indicated that the model of interrelation between characters developed based on SPA can be confirmed appropriately by SEM. Both models were valid and fit with the existing model in their populations. The result explained through the parameters of $\chi^2$, GFI, AGFI, NFI, RFI, CFI, and RMSEA that are presented in Table 5. A model is considered suitable or fit if the value of Chi-square ($\chi^2$) is insignificant ($p > 0.05$); value of GFI, AGFI, CFI, NFI, and TLI between 0.90-1.00; and value of RMSEA < 0.08 (HOE, 2008; HOOPER et al., 2008; BARRETT, 2007).

![Path diagram of the confirmatory model based on SEM](image)

**Figure 5. Path diagram of the confirmatory model based on SEM**

**Gambar 5. Diagram lintas model konfirmasi berdasarkan SEM**

**Table 5. Estimate value of direct effect and test result of SEM models**

<table>
<thead>
<tr>
<th>Causal relationship</th>
<th>Estimate value</th>
<th>Parameters for testing model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hubungan kausal</td>
<td>Nilai estimasi</td>
<td>Parameter untuk pengujian model</td>
</tr>
<tr>
<td>NUPB → NPB</td>
<td>-0.54 **</td>
<td>$\chi^2 = 0.95$; GFI = 0.96;</td>
</tr>
<tr>
<td>SG      → NPB</td>
<td>0.46 **</td>
<td>AGFI = 0.90; NFI = 0.97;</td>
</tr>
<tr>
<td>PH      → NPB</td>
<td>0.22 *</td>
<td>RFI = 0.95; CFI = 1.0;</td>
</tr>
<tr>
<td>CW      → NFC</td>
<td>-0.33 *</td>
<td>RMSEA = 0.00</td>
</tr>
<tr>
<td>NPB → NFC</td>
<td>0.58 **</td>
<td></td>
</tr>
<tr>
<td>NPB → NCT</td>
<td>0.39 **</td>
<td></td>
</tr>
<tr>
<td>NFC → NCB</td>
<td>0.67 **</td>
<td></td>
</tr>
<tr>
<td>NCB → NCT</td>
<td>0.79 **</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * and ** significant at 5 and 1% level respectively
Keterangan: * dan ** masing-masing nyata pada taraf 5 dan 1%

Based on the analysis of SPA and SEM it can be confirmed that several important vegetative and generative characters in Arabica coffee can be used as a selection criteria for high yielding materials. The vegetative characters (NUPB, PH, SG, and CW) indirectly effect on the yield component (NCB) and yield (NCT) through generative character (NPB and NFC), while the generative characters have direct effect. Therefore, it can be argued that in vegetative phase the positive selection can be done on the PH and SG characters. While the negative selection
can be done only on the CW character, because NUPB is identical to the NPB with a negative index. In generative phase, the positive selection can be done on the NFC and NPB characters.

A limitation of PA in conventional approach is it only able to measure the direct effect of one variable to other variables statistically, but it can not explain the causal relationship among those variables systematically and sequentially. In Figure 6, PH directly effect on NCT but theoretically it is an indirect effect because PH is a vegetative characters while NCT is a yield characters. There are three generative characters (NPB, NFC and NCB) as endogenous variables between PH and NCT (Figure 5) and commonly known as mediating or moderating variables. Mediation and moderation are two theories for refining and understanding a causal relationship. Mediating variable is a third variable that links a cause and an effect, whereas moderating variable modifies a causal effect. Statistically, the mediating variable were correlated whereas the moderating variable uncorrelated with the independent variable (WU and ZUMBO, 2008).

In conventional approach of PA for all variables indicated that PH, CS, NPB, and NCB significantly direct effect, while SG, CW, NUPB, NI, and NFC not significantly effect on NCT (Figure 6). Whereas, according to the SPA and SEM there are only two characters (NPB and NCB) wich directly effect on NCT (Table 5; Figure 5). On the other hand, according to the PA with the conventional approach it’s known that SG, CW, NUPB, NI, and NFC did not affect on NCT so that the purpose to shorten the plant selection cycle cannot be achieved. Whereas according to the SPA and SEM, NUPB, SG and PH has indirect effect on NCT through NPB, and CW has indirect effect on NCT through NFC and NCB. Therefore, by SPA and SEM the interrelationships among vegetative, generative and yield characters of Arabica coffee can be investigated systematically and sequentially.

**CONCLUSIONS**

The number of cherries of Arabica coffee var. Kartika 1 and 2 population in Pakuwon Experimental Station, were directly influenced by generative character and indirectly affected by vegetative characters. Plant height and stem...
girth can be used as positive selection criteria for high yielding at vegetative phase, while negative selection was used on the canopy width. In generative phase, the number of productive branches, number of cherries and flower cluster can be used as positive selection criteria.

ACKNOWLEDGEMENTS

We are grateful to Mr. Dede Suhendar as Field Assistant at Pakuwon Experimental Station of the Indonesian Industrial and Beverage Crops Research Institute (IIBCRI) in supporting of research implementation and collecting data in the field.

REFERENCES


EDU WARDIANA et al. : Selection of vegetative and generative characters of Arabica

WUENSCH, K.L. 2006. Conducting a path analysis with SPSS/AMOS. Dept. of Psychology, East Carolina University Greenville, NC 27858-4353. 15p. [accessed on 7-3-2008].

### APPENDIX

Stepwise regression for the first-, second-, and third-order analysis

Regression berturut untuk analisis pada orde pertama, kedua, dan ketiga

<table>
<thead>
<tr>
<th>Step of analysis</th>
<th>Regression models</th>
<th>Unstandardized coefficient Koefisien tidak terbakukan</th>
<th>Standardized coefficient Koefisien terbakukan</th>
<th>t-value Nilai-t</th>
<th>Significance Signifikansi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B Standard error Galat baku</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Constant</td>
<td>1363.51 82.57</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NCB</td>
<td>14.60 0.43</td>
<td>0.79</td>
<td>33.85</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>NPB</td>
<td>95.79 5.77</td>
<td>0.39</td>
<td>16.62</td>
<td>**</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>-33.25 23.73</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NFC</td>
<td>21.50 3.85</td>
<td>0.67</td>
<td>5.58</td>
<td>**</td>
</tr>
<tr>
<td>3a</td>
<td>Constant</td>
<td>4.21 1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NPB</td>
<td>0.24 0.06</td>
<td>0.58</td>
<td>4.37</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>CW</td>
<td>-0.02 0.01</td>
<td>-0.33</td>
<td>-2.39</td>
<td>*</td>
</tr>
<tr>
<td>3b</td>
<td>Constant</td>
<td>0.35 6.50</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NUPB</td>
<td>-0.73 0.14</td>
<td>-0.54</td>
<td>-5.35</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>SG</td>
<td>0.15 0.03</td>
<td>0.46</td>
<td>4.57</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>PH</td>
<td>4.36 2.00</td>
<td>0.22</td>
<td>2.18</td>
<td>*</td>
</tr>
</tbody>
</table>

Notes: * and ** significant at 5 and 1% level respectively

Keterangan: * dan ** masing-masing nyata pada taraf 5 dan 1%