Jurnal Manajemen Hutan Tropika Vol. XI No. 2 : 15-27 (2005) **Artikel (*Article*)**

PREDICTING SPATIAL DISTRIBUTION OF STAND VOLUME USING GEOSTATISTICS

(Pendugaan Sebaran Spasial Volume Tegakan Menggunakan Metode Geostatistika)

TATANG TIRYANA[1]

ABSTRACT

Pendugaan volume tegakan hutan alam biasanya dilakukan dengan metode penarikan contoh (sampling) non-spasial. Dalam penelitian ini, dikembangkan metode pendugaan spasial dengan pendekatan geostatistika yang diterapkan untuk menduga dan memetakan sebaran volume tegakan di areal HPH Labanan, Kalimantan Timur. Pendugaan dilakukan dengan menggunakan metode kriging berdasarkan model spherical variogram dan data dari 1090 plot contoh. Keakuratan hasil pendugaan diuji melalui validasi dengan menggunakan 272 plot contoh yang berbeda. Hasil analisis menunjukkan bahwa volume tegakan memiliki korelasi spasial dengan keragaman yang cukup besar, namun tidak dijumpai adanya kecenderungan arah sebarannya. Selain itu, dapat dibuktikan bahwa pendugaan dengan metode kriging dapat memberikan nilai dugaan volume tegakan dengan akurasi yang cukup tinggi. Hal ini menunjukkan bahwa pendekatan geostatistika dapat digunakan sebagai metode alternatif dalam pendugaan dan pemetaan sebaran volume tegakan.

Keywords: stand volume, sampling technique, geostatistics, spatial dependence variogram, kriging

INTRODUCTION

In any forest management activities, data and information concerning stand parameters, such as stand diameter, stand height, basal area, stand density, and stand volume, are absolutely required. Commonly, these data are obtained by conducting forest inventory in a particular area or in the whole forest area. In general, forest inventory is conducted to obtain information on the quantity, quality and condition of the forest resources that are being managed (Husch et al., 2002), particularly in order to estimate timber stocks.

Among the other stand parameters, stand volume is the most commonly used parameter to quantify the timber stocks, particularly in a production forest. Usually, a stand volume estimation is required to determine an appropriate annual allowable cut (AAC) in a

[1] Lecturer and researcher at Department of Forest Management, Faculty of Forestry, Bogor Agricultural University (IPB), Kampus IPB Darmaga Bogor, Phone/Facs. 62-251-621244, Email: tangtir@yahoo.com

forest management unit (FMU). Moreover, by knowing the stand volume, the forest managers will also be able to estimate total revenue that can be generated from their concession.

Commonly, estimating stand volume for a large forest area is conducted by applying a sampling technique. In Indonesia, the standard sampling technique for conducting forest inventory in a natural production forest is systematic transect sampling with random start (Tiryana, 2003). This technique is considered to be an effective way in estimating stand volume because it takes into account variations of the stand volume which may vary from site to site as the contour lines shift, and also because of its practicability in the fields (Shiver and Borders, 1996; Vries, 1986). The average and total stand volume are then estimated based on stand volume of each sampling unit (i.e. transect with a given area) using non-spatial estimators such as ratio or regression estimator (Vries, 1986).

Despite its applicability for estimating stand volume in a large forest area, the sampling technique as described above has a main drawback since it can not provide a means to estimate stand volume at unsampled forest areas; whereas the stand volume distribution on each forest site is absolutely required to manage the forest properly. In addition, either the ratio or regression estimator as commonly used in the sampling technique is a biased estimator particularly when the sample size is small (Cochran, 1977). These facts indicate the need for developing another technique to spatially estimate stand volume distribution. In this context, geostatistical methods (such as kriging) provide a novel approach to predict the stand volume at unsampled locations based on spatial dependences among the observed samples (Nielsen and Wendroth, 2003; Rossiter, 2004; Saborowski and Jansen, 2002). In a forest stand, the spatial dependence may occur since site productivity has continuity in space, hence the stand volume would have spatial continuity as well (Nanos et al., 2004). Accordingly, the geostatistical methods would be appropriate to be used in predicting spatial distribution of the stand volume.

This paper describes the use of geostatistics as an alternative method to predict spatial distribution of stand volume in the natural production forest of Labanan concession, East Kalimantan, Indonesia. More specifically, the objective of this study was threefold:

1) to model spatial dependence and trend of the stand volume distribution.
2) to analyze spatial factors affecting variation of the stand volume.
3) to map stand volume distribution using the kriging methods.

METHODS

Data

This study used forest inventory data from Labanan concession, East Kalimantan, Indonesia. The data were collected by the company (i.e. PT. Inhutani I) in 1997 in order to provide a forest resources database for managing the natural production forests in the concession.

Forest inventory in the Labanan concession was conducted by applying systematic transect sampling with a spacing of about 5 km with azimuth 315^{o}. The sample plots were located along transects with a spacing 100 m (see *Figure 1*). In each plot, tree measurements (i.e. dbh–diameter at breast height–and tree species) were conducted using three nested subplots, i.e. 0.125 ha for dbh >50 cm, 0.04 ha for dbh 20-49 cm, and 0.0125 for dbh 10-19 cm. Coordinates of centre plot were also recorded using GPS. The total size of sample was 1538 plots which covered approximately 0.24% of the total area of 81224.79 ha (Gunawan, 2002). From the tree measurements, stand volume of each plot was determined using a volume table.

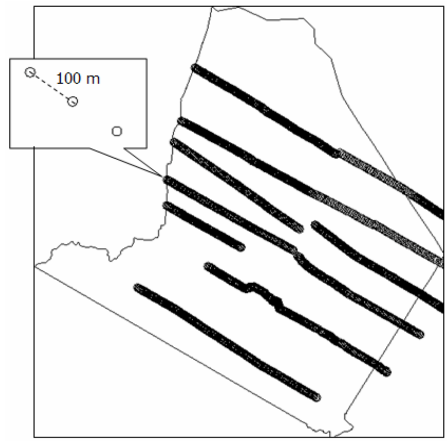

Figure 1. Layout of the sample plots in the study area

Methods

In this study, the geostatistical approaches were used to model spatial distribution of the stand volume as well as to predict stand volume at unsampled locations. In details, the methods used in this study can be described as follows:

Analyzing the data set

The data set consists of 1362 sample plots located inside the study area. This large amount of data was divided into two independent datasets, i.e. 1090 plots for modeling and 272 plots for validation purposes. The descriptive statistics was used to summarize both datasets. In addition, analysis of variance (anova) was also used to analyze some spatial factors affecting the variation of stand volume, namely slope, elevation, and geographical coordinate (x,y) which were derived from DEM (Digital Elevation Model) of the study area. The significant factors, then, were used as auxiliary variables for predicting stand volume using the kriging method.

Modelling spatial dependence and trend of the stand volume distribution

In a geostatistical analysis, spatial dependence of the stand volume can be modelled using variogram. Basically, a variogram is a statistical tool to measure spatial correlation between the two samples as a function of their separation distance, which is calculated from sample data using the following formula (Cressie, 1991; Nanos et al., 2004; Rossiter, 2004; Saborowski and Jansen, 2002):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x_i) - z(x_i + h) \right]^2 \tag{1}$$

where: $\hat{\gamma}(h)$ is semivariance of the stand volume for distance h, $N(h)$ is the number of point pairs within distance h, whereas $z(x_i)$ and $z(x_i+h)$ are the stand volume at locations x_i and x_i+h, respectively.

The best fitted variogram for the stand volume data was selected from the most commonly used variogram models as follows (Nangendo et al., 2002; Nielsen and Wendroth, 2003; Rossiter, 2004; Webster and Oliver, 2001):

- The spherical model: $$\gamma(h) = \begin{cases} c_0 + c_1 \left(\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3 \right) & \text{, for: } h < a \\ c_0 + c_1 & \text{, for: } h \geq a \end{cases} \tag{2}$$
- The exponential model: $$\gamma(h) = c_0 + c_1 \left(1 - e^{-\frac{h}{a}} \right) \tag{3}$$
- The Gaussian model: $$\gamma(h) = c_0 + c_1 \left(1 - e^{-\left(\frac{h}{a}\right)^2} \right) \tag{4}$$

where: c_0 is nugget effect, which represents unexplained variability at distance close to zero; c_1 is sill, which represents variability when the observations become independence; a is range, which represents distance in which the spatial dependence would no longer exist. *Figure 2* illustrates such variogram features. The modeling of variograms and other geostatistical computations were performed using the Gstat package of the R software (Pebesma, 2004; see also: www.r-project.org).

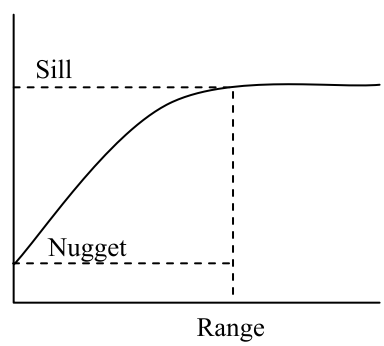

Figure 2. A theoretical variogram

The spatial dependence of the stand volume may occur in omni-direction (called isotropy) or tend to a certain direction only (called anisotropy). The anisotropic phenomenon was analyzed by plotting the directional variograms at different angles. In addition, the variogram surface tool of the ILWIS software was also used to confirm whether or not the anisotropy existed in the stand volume distribution.

Predicting the stand volume distribution

One of the objectives of this study was to obtain a map, showing the stand volume distribution of the study area. It can be obtained by using a geostatistical method known as kriging. There were two kriging methods used in this study, namely ordinary kriging and universal kriging.

The ordinary kriging was used to predict stand volume at unsampled locations by using the selected variogram model obtained from the previous step and the stand volume data in the neighborhood of an estimated location (Nielsen and Wendroth, 2003; Wackernagel, 1998). Meanwhile, the universal kriging was used to improve prediction of the stand volume by using an auxilary variable as complement to the other variables used in the ordinary kriging. The analysis of variance in the previous step would reveal which factors, i.e. slope, elevation, or coordinate, that could be used as an appropriate auxiliary variable.

Validating the predictions

To assess validity of the predictions produced by the krigings, a validation was carried out by comparing the predictions with 272 independent sample plots. The following statistics were used to assess reliability of the predictons (Rossiter, 2004; Webster and Oliver, 2001):

- Bias or mean error (ME), which should be zero (0):

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(z(x_i) - \hat{z}(x_i)\right) \quad (5)$$

- Root mean squared error (RMSE), which should be low:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(z(x_i) - \hat{z}(x_i)\right)^2} \quad (6)$$

- Mean squared deviation ratio (MSDR), which should be one (1):

$$MSDR = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(z(x_i) - \hat{z}(x_i)\right)^2}{\hat{\sigma}^2(x_i)} \quad (7)$$

where: z_i is observed stand volume, $\hat{z}_i$ is predicted stand volume, n is number of sample plots, and $\hat{\sigma}^2(x_i)$ is kriging variances.

RESULTS AND DISCUSSION

Characteristic of the stand volume and auxiliary variables

Table 1 shows descriptive statistics of the stand volume and the auxiliary variables (i.e. slope and elevation) of the 1090 sample plots.

Table 1. Descriptive statistics of the stand volume and topographical factors

Statistic	Variable		
	Volume (m^3/ha)	Slope (%)	Elevation (m)
Minimum	1.73	0.00	25.00
Mean	**159.65**	**14.23**	**132.01**
Maximum	620.77	176.60	250.00
Standard deviation	90.31	18.23	53.28

The stand volume had large variation, i.e. ranged from 1.73 m^3/ha to 620.80 m^3/ha, which means that the timber stocks were not evenly distributed in the study area. Based on the Shapiro-Wilk test, i.e. W=0.9505 and p-value <2.2e-16, at 5% of confidence level there was an evidence that the stand volume data were not normally distributed as seen in *Figure 3*. To facilitate the geostatistical analyses, which require normality of the data, the stand volume data were transformed into the squared root scale.

For the slope and elevation, it can be seen that the sample plots were mostly located in flat areas. Indeed, the Labanan forest is a low land tropical forest.

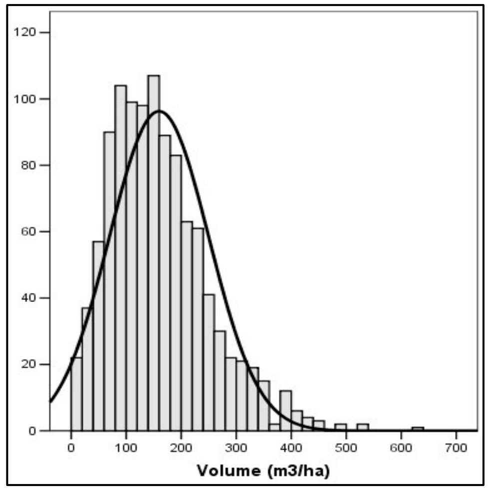

Figure 3. Distribution of the stand volume

Factors affecting the stand volume distribution

Table 2 shows analysis of variance for the auxiliary variables, i.e. slope, elevation, and the geographical coordinate (x,y).

Table 2. Analysis of variance of the auxiliary variables

Factor	**Analysis of variance**		
	Adjusted R^2 (%)	**F-test**	**p-value**
Slope	0.39	5.242	0.0222*
Elevation	0.07	0.736	0.3912
Geographical coordinate (x, y)	**6.13**	**36.56**	**4.297e-16****

*significant at 5% of confidence level, **highly significant at 1% of confidence level

The results showed that, at 5% confidence level, the slope had significant effect to the variation of stand volume. However, due to only 0.39% of the total variance which can be explained by the slope, it could not be used as an auxiliary variable for the universal kriging. Likewise, the elevation had no significant effect to explain variation of the stand volume. The main reason is that the sample plots were taken from the almost uniform conditions with little variations in their slope and elevation, since the Labanan forest is located in a low land area. Obviously, only the geographical coordinate had significant effect in which it explained about 6.13% of the total variation of the stand volume. Although, its adjusted-R^2 value was not so high, it could be used as an auxiliary variable to improve prediction of the stand volume using the universal kriging.

Spatial dependence of the stand volume

As mentioned before, the variogram can be used to model spatial dependence of the observed attributes. Table 3 presents the parameters of the three variogram models that were used to explain spatial dependence of the stand volume.

Table 3. Parameters of the variogram models

Parameter	Variogram model		
	Spherical	Exponential	Gaussian
Nugget	9.65	9.50	9.92
(total) Sill	12.80	13.21	12.71
Range (m)	**6870**	3578	2886
Sum squared error (SSErr)	0.0021	0.0019	0.0017

All the variogram models have low SSErr values, but they are not significantly different. Although the Gaussian model has lowest SSErr, it has the shortest range compared to the others. Therefore, it was considered that the spherical model is better than the others because it has longer range, hence it would explain better the spatial dependence of the stand volume. The spherical model was also used by Nanos et al. (2004) to develop a model for predicting spatially the stand height–diameter relationship. Visually, this variogram model is depicted in *Figure 4*.

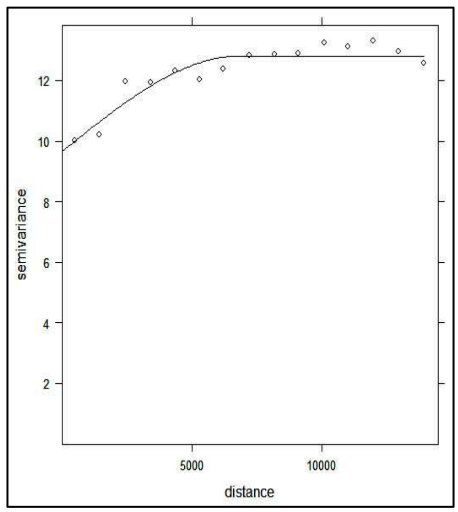

Figure 4. The spherical variogram of the stand volume

Based on the spherical variogram, it can be concluded that the longest distance in which spatial dependence of stand volume would still exist is approximately 6870 m with the total variability at this distance onwards is 12.80 $(m^3/ha)^2$ (in squared-root scale). While, at distance close to zero there is still high variability (nugget effect), i.e. 9.65 $(m^3/ha)^2$ (in squared-root scale).

This large nugget effect indicates that in each sample plot, the variability of stand volume is relatively high. This is due to the fact that in each plot there were mixed trees with three different diameter classes (i.e. trees with diameter 10-19 cm, 20-49, and ≥50 cm), whereas the stand volume was calculated by aggregating volume of these trees. Another reason is that the size of sampling unit (i.e. inventory plot) was too small to capture spatial variability in the heterogeneous tropical forest. Indeed, in the standard forest inventory for the natural forests, the common sampling support is continuous transects instead of clustered plots (Tiryana, 2003; Vries, 1986).

Anisotropic issue

Figure 5 shows the directional variograms at the nine different angles. It seems that the anisotropy would occur at direction 112.5^0 because it has longest distance (approximately 16973 m with nugget 9.05 and total sill 12.07) as compared to the others. However, it does not provide sufficient evidence since we only judged it visually. Further analysis using the variogram surface (the result is not shown here) revealed that obviously there was no pattern showing anisotropy, such as an ellipse-like pattern in a certain direction.

Therefore, we concluded that there was no anisotropy found in the stand volume distribution. It means that there were no spatial variability changes with direction of the stand volume.

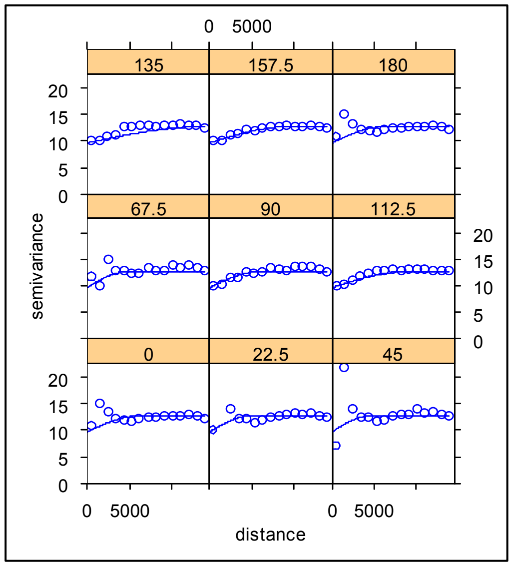

Figure 5. The directional variograms of the stand volume

This is possibly due to the sample plots which were not evenly distributed in the field, so that they could not be able to capture spatial variability of stand volume in all possible directions. Indeed, the sample plots were located along transects at spacing 100 m between plots and about 5 km between transects.

Prediction of the stand volume

The stand volume distribution of the whole study area was predicted using the ordinary kriging and the universal kriging. The results were prediction maps as shown in Figure 6 and Figure 7.

Obviously, there are some differences in both prediction maps, particularly in the areas away from the sample plots. From Table 4, it can be seen that the ordinary kriging produced predictions higher than the universal kriging, although the differences seem not so significant. In addition, both krigings produced the prediction errors as summarized in Table 5. It can be seen that the errors produced by the universal kriging are lower than those of the ordinary kriging, particularly for the areas around the sample plots.

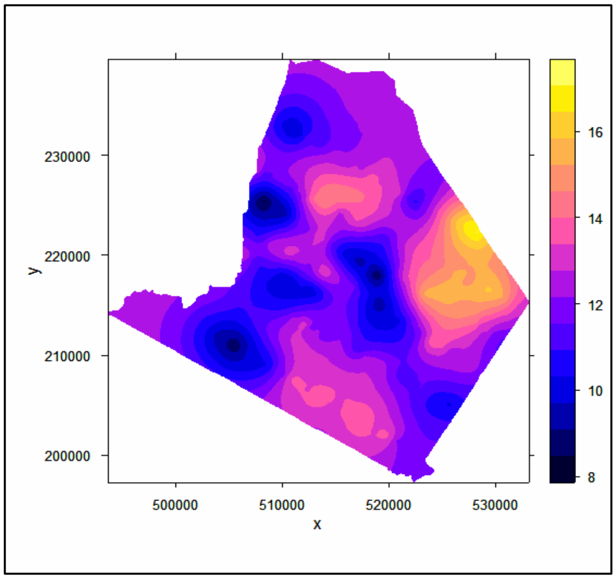

Figure 6. Prediction map of the stand volume obtained from the ordinary kriging (not to scale). The values are squared-root of stand volume in m^3/ha.

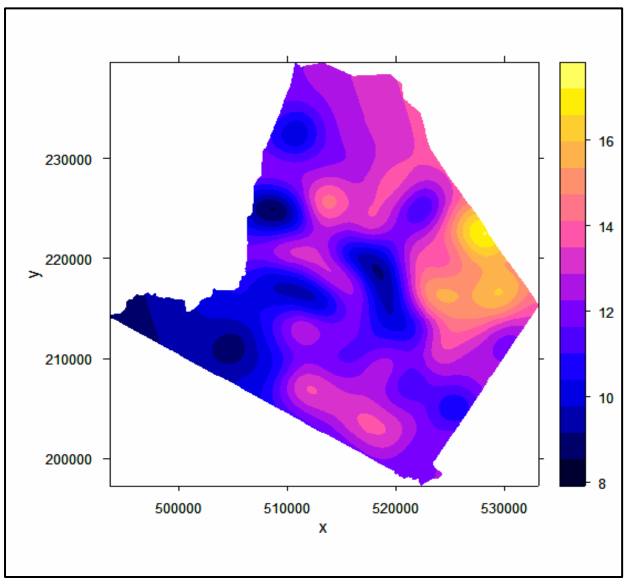

Figure 7. Prediction map of the stand volume obtained from the universal kriging (not to scale). The values are squared-root of stand volume in m^3/ha.

Table 4. Descriptive statistics of the predictions obtained from the ordinary kriging and the universal kriging

Kriging method	Statistics					
	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Ordinary kriging	8.46	11.33	12.18	12.21	12.94	17.07
Universal kriging	8.53	11.03	12.13	12.08	13.10	17.20
Difference *)	-1.57	-0.26	0.01	0.12	0.35	3.65

*) the differences were calculated from all prediction values of both krigings

Table 5. Descriptive statistics of the prediction errors produced by the ordinary kriging and the universal kriging

Kriging method	Statistics					
	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Ordinary kriging	10.21	10.65	11.06	11.29	11.75	12.95
Universal kriging	10.26	10.55	11.02	11.32	11.91	13.66
Difference *)	-0.71	-0.12	0.004	-0.024	0.08	0.22

*) the differences were calculated from all prediction error values of both krigings

The results revealed that predictions of the stand volume using the ordinary kriging and the universal kriging seem not so much different. However, the universal kriging performed better than the ordinary kriging. It was proven that by incorporating the auxiliary variable (i.e. geographical coordinate) into model, the prediction could be improved. In this case, the coordinate gives the advantage in adjusting local variability of the stand volume so that the prediction errors around sample plots would become lower. However, for the unsampled area or for the areas away from the sample plots, the universal kriging cannot well-predict the stand volume due to the absences of local trend, so that the variance of prediction is higher for those areas.

Validation of the predictions

Table 6 shows the results of validation using 272 independent sample plots. Obviously, there were no so much differences regarding accuracy of the predictions obtained from the ordinary kriging and the universal kriging. However, it seems that the universal kriging performed better than the ordinary kriging, particularly in terms of the mean error (ME) and the mean squared deviation ratio (MSDR). The validity measures confirmed that both krigings produced low bias (mean error), low RMSE, and low (i.e. close to one) MSDR.