# **FITTING DIAMETER DISTRIBUTIONS OF TROPICAL RAINFORESTS IN VIETNAM BY FIVE PROBABILITY FUNCTIONS**

## **Cao Thi Thu Hien1, Nguyen Hong Hai 1, Nguyen Phuc Truong2**

*1 Vietnam National University of Forestry*

*2 East Northern Sub Forest Inventory and Planning Institute*

#### **SUMMARY**

This study aims to fit diameter distribution of tropical rainforests in Vietnam and to select suitable continuous distributions. To investigate the diameter distribution of trees in tropical rainforests in Vietnam, data from 20 one – hectare plots were used. The following functions were tested: Beta, Weibull (two- and three-parameter), Lognormal, and Gamma. Parameters of distribution functions were estimated using the maximum likelihood estimation method. The best fits were selected by the Kolmogorov–Smirnov test. The results showed that the diameter distribution of the tropical rainforest was best described by the three – parameter Weibull distribution, followed by Log – normal distribution, while the Weibull 2P model fails in every case to adequately describe these frequency distributions. Estimated parameters  $\mu$ , β, and  $\alpha$  of Weibull 3P distribution ranged from 5.697 – 10.155; 1.029 – 1.606, and 6.921 – 15.234, respectively, while estimated parameters u and  $\delta$  of Log – normal distribution varied from  $2.804 - 2.891$  and  $6.921 - 15.234$ , respectively. The studied forests showed diameter distributions with decreasing number of trees for larger trees.

**Keywords: Continuous distribution, diameter structure, Kolmogorov–Smirnov test, maximum likelihood, tropical rainforest.**

### **1. INTRODUCTION**

Diameter distributions are crucial decision– making tools for forest management. They directly affect the choices concerning silvicultural and harvesting stages activities. For instance, timing and intensity of thinning and harvesting, as well as harvesting equipment are dependent on the diameter distributions (Robinson and Hamann, 2011). Furthermore, they are applied as inputs of growth models and sometimes are the subject of growth modeling themselves. Information on current diameter distribution of a forest stand allows prediction of its future structure which provides even better support for sustainable forest management (Vanclay, 1994; Podlaski, 2006).

Beta, Weibull, and Log–normal functions are classic models frequently applied for diameter distribution analysis in tropical forests (Bailey and Dell, 1973; Rennolls *et al*., 1985; Nanang, 1998; Palahí *et al*., 2007, Burkhart and Tomé, 2012). However, there are also other investigations highlight the potential of alternative models for predictions that support forest conservation and management (Wang and Rennolls, 2005; Podlaski and Zasada, 2008; Gorgoso-Varela and RojoAlboreca, 2014; Lima *et al*., 2014,; Podlaski and Roesch, 2014). As a result, there are many different probability density functions as the best in fitting diameter of trees. However, the current situation is that there is no clear resolution as to which model is the most suitable for tree distributional modelling (Wang and Rennolls, 2005). There is no theoretical reason why there should exist a best model for all situations. It might be that in one case a particular distribution will be found empirically to give the best fit, whilst in another case another model will be found to be empirically best. The only way in which it is meaningful to talk about the best distributional model is in terms of the most flexible of models in representational terms. Thus, the reason for this research is to compare five distributional models in terms of flexibility and ability on which best fit diameter of trees in tropical rainforests in Vietnam.

# **2. RESEARCH METHODOLOGY**

# **2.1. Data collection**

Stratified random sample method and simple random sampling was applied for the set up sample plots. The data were used from a project "Method of establishing a volume table for standing trees in natural forest, Vietnam"

(Vu Tien Hinh, 2012). Each plot has a square shape (100 m x 100 m) and is divided into twenty five 20 m x 20 m quadrats. From each quadrat, all the tree species having diameter at breast height (DBH) of  $\geq 6$  cm were recorded by their name and diameters, except the extremely rich forest, DBH of tree species having  $\geq 10$  cm were recorded by their name and diameters. General plot imformation is reported in Table 1.



**Table 1. Provinces, plots and forest states**

#### **2.2. Data analysis**

#### *2.2.1. Descriptive statistics*

Several general information on forest structure were computed for each sample plot, including: mean, standard deviation, variance, min, max, skewness and kurtosis (Machado *et al*. 2009).

#### *2.2.2. Fitting of diameter distribution model*

The data were analyzed on several probability density function (pdf), then ranked based on Kolmogorov - Smirnov test. From some previous researches (Carretero and Torres–Alvarez, 2013; Ige et al., 2013; Aigbe, 2014), the following functions were used: Beta, Weibull (two- and three-parameter), Lognormal, and Gamma.

- Beta distribution

This distribution is a continuous distribution and its pdf formula is as follows:

$$
F(x) = \frac{1}{B(p,q)} \frac{(x-a)^{p-1}(b-x)^{q-1}}{(b-a)^{p+q-1}}
$$
  
  $a \le x \le b, p > 0, q > 0$ 

- Two - parameter Weibull distribution

Weibull distribution was also presented as an cumulative frequency in which a is the starting point and b presents the curve's concavity degree, and c is the curve's factor form or shape index and its mathematical process is as follows:

$$
f(x)=1-e^{-\left(\frac{x}{\beta}\right)^{\alpha}}
$$

For  $x > 0$ ,  $\alpha > 0$  and  $\beta > 0$ . The parameters  $\alpha$ and  $\beta$  are referred to as scale and shape parameter, respectively.

- Three - parameter Weibull distribution

$$
f_X(x) = \frac{\alpha}{\beta} \left(\frac{x-\mu}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x-\mu}{\beta}\right)^{\alpha}}
$$
  
x > \mu, \alpha > 0, \beta > 0

Where  $\alpha$ ,  $\beta$  and  $\mu$  are shape, scale and location parameters, respectively.

- Lognormal distribution

This is continuous distribution and its natural logarithm has a normal distribution

$$
F(x) = \frac{1}{\sqrt{2\pi\delta x}} exp\left(-\frac{1}{2\delta^2} (Inx - \mu)^2\right)
$$
  

$$
\delta > 0, x > 0, 0 < \mu + \infty
$$

- Gamma distribution

This distribution is a continuous distribution and it has a good flexibility and its frequency curve in all modes has a lean towards right.

$$
F(x) = \frac{x^{\alpha-1}}{\beta T(\alpha)} \exp\left(\frac{-x}{\beta}\right)
$$
  
0 \le x \le +\infty \alpha, \beta > 0

### *2.2.3. Estimating parameters for diameter distribution models*

The parameters of the distribution are estimated by maximizing the likelihood of the sample.

The method of maximum likelihood is a commonly used procedure for the fitting distribution in forestry because it has very desirable properties. Estimation of the parameters by maximum likelihood has been found to produce consistently better goodnessof-fit statistics compared to the previous methods, but it also puts the greatest demands on the computational resources (Cao, McCarty, 2005).

#### *2.2.4. Test for the goodness of fit*

The goodness of fit in different class intervals was carried out using the Kolmogorov–Smirnov test according to the following expression:

$$
D_n = \frac{SUP_X|Fo_{(x)} - Fe_{(x)}|}{n}
$$

Where  $F<sub>o</sub>(x)$  is the accumulate observed frequency;

 $Fe<sub>(x)</sub>$  is the accumulate expected frequency; n is the number of observations;  $D_n$  is the

calculated D value.

Dn was compared with the value of the Kolmogorov–Smirnov table at a probability level of 95%.

This test was used to check the following hypotheses of the bilateral test:

Null hypothesis:  $H_0 =$  the observed diameters follow the proposed distributions

Alternative hypothesis:  $H_1$  = the observed diameters do not follow the proposed distributions.

Parameters for diameter distribution model were estimated using XLSTAT version 2015.5 software.

#### **3. RESULTS AND DISCUSSION**

## **3.1. Descriptive statistics of diameter at breast height**

The summary of the descriptive statistics of diameter for tropical rainforests in Vietnam are presented on Table 2.



**Table 2. Descriptive statistics of the data sets of diameter at breast height**

*N\* : Number of trees per hectar* 

The highest and lowest dbh values found were 25.9 cm and 13.5 cm, respectively (Table 2). The max DBH came from plot 20 in the extremely rich forest. The highest tree densities were counted in plot 13 of the rich forest, whereas the lowest was observed in the plot 1 of the poor forest. Positive kurtosis

values imply that diameter distributions are flatter than the normal curve. Skewness was shown positive distributions since the diameter distributions are located on the left side. This pattern means that the tails of diameter distribution extend to the right side or other words, means that considerable numbers of

trees concentrate in the lower diameter classes (Gadow, 1983).

**3.2. Diameter distribution model**

The summary of the goodness of fit of diameter distribution functions for tropical rainforest in Vietnam are presented on Table 3.





The Kolmogorov - Smirnov test indicates that the five distributions can provide good fits for the diameter data, because their p-value of calculated D-values were greater than significance level (Sig. level) 0.05, except plot 13 (in rich forest) and plot 20 (in extremely rich forest). This implies the null hypotheses were accepted for 18/20 distributions, meaning the data followed the specified distribution.

The distribution of diameter was well described by the Log–normal function for 8/20 plots, and by Beta, Weibull 2P, Gamma function for  $2/20$ ,  $0/20$ ,  $5/20$ , respectively. However, three – parameter Weibull distribution was more flexible in fitting the diameter data when tested with Kolmogorov - Smirnov because it has the highest p-value of calculated D-values for 18/20 plots.

A number of different distribution functions have been used to model diameter distributions, including Beta (Zohrer 1972; Li *et al*., 2002), two – parameter Weibull (Bailey and Dell, 1973; Rennolls *et al*., 1985), three – parameter Weibull, and Lognormal (Bliss and Reinker, 1964).

The Weibull distribution, introduced by Bailey and Dell (1973) as a model for diameter distributions, has been applied extensively in forestry due to (1) its ability to describe a wide range of unimodal distributions including reversed-J shaped, exponential, and normal frequency distributions, (2) the relative simplicity of parameter estimation, and (3) its closed cumulative density functional form (Bailey, Dell, 1973; Schreuder, Swank, 1974; Schreuder *et al*., 1979; Little, 1983; Rennolls *et al*., 1985; Mabvurira *et al*., 2002), and (4) its previous success in describing diameter frequency distributions within tropical rainforests (Le Sau, 1996; Dao Cong Khanh, 1996; Pham Quy Van, 2018; Cao Thi Thu Hien and Nguyen Hong Hai, 2018; Nguyen Quang Phuc, 2019).

Weibull distribution also was adjudged

more flexible in a research carried out in Gorazbon district of Kheyroudkenar forest by Namiranian (1990) and Mataji *et al*., (2000). Both scientists, using Kolmogorov-Smirnov

tests showed that Weibull distribution could determine diameter distribution of trees.

Table 4 shows the parameter values of the selected distribution functions.





The graphs of observed and estimated dbh class of the distribution functions show that there is no significant difference between the observed and predicted diameter frequencies (Figure 1). Some sample plots as according to the three – parameter Weibull distribution for observed frequency and estimated frequency was illustrated in Figure 1.

Figure 1 showed the distribution pattern of the dbh (m) of trees in tropical rainforests in Vietnam. The pattern shows that the majority of stems were concentrated in the first class or second class that is sufficient enough to replace trees in the upper dbh class in the future (i.e. when the big trees are harvested or when they die) (Aigbe and Omokhua, 2014) or another word, the type of decreasing distribution in tropical rainforests indicates that

the regeneration is continuously happening as a consequence of the species' ability to adapt to environments (Ferreira *et al*., 2015). The implication of this is that the forests are still undergoing regeneration and recruitment, which are vital indicators of forest health and vigour (Jimoh *et al*., 2011). This is consistent with other reports for two other tropical rainforests (Boubli *et al*., 2004; Bobo *et al*., 2006). In addition, the poor and the medium forests were lacking large stems (over 60-cm DBH). Trees with a DBH greater than 70 cm were only found in the rich and extremely rich forests. This is similar to findings of Pham Quy Van (2018), Cao Thi Thu Hien and Nguyen Hong Hai (2018), Nguyen Quang Phuc (2019), Nguyen Thuy Hong (2019).



**Figure 1. Diameter distributions and fitted curves of the models for several plots. Estimated frequency was from three – parameter Weibull distribution**

The results of this study are also consistent with the results of many previous researches for the distribution of diameter at breast height in natural forest in Vietnam, such as the study in evergreen broadleaf natural forest in Kon Ha Nung (Le Sau, 1996), in Huong Son (Dao Cong Khanh, 1996), for natural forest state IIIA in An Lao district, Binh Dinh province (Pham Quy Van, 2018), Cao Thi Thu Hien and Nguyen Hong Hai (2018), Nguyen Quang Phuc (2019) also concluded that Weibull 3P distribution showed good flexibility to describe

the diameter structure of the stand.

#### **4. CONCLUSION**

Tree diameter distributions play an important role in stand modelling. The area of rainforest in Vietnam showed diameter distributions with decreasing exponential curves and positive skewness. Using appropriate probability theories to predict trees distribution in tropical rainforest is important in estimation of productivity in different dbh class. In this study, probability distributions were applied to estimate the diameter

distribution, and statistical methods were used to provide diameter distribution models.

Three – parameter Weibull function was the most flexible in fitting the diameter data in tropical rainforest in Vietnam when tested with Kolmogorov – Smirnov test, followed by Log– normal function. The Weibull 2P function did not depicted adequate development for frequency estimation of diameter class.

#### **REFERENCES**

1. Aigbe H.I., Omokhua G.E., 2014. Modelling diameter distribution of the tropicao rainforest in Oban Forest Reserve. Journal of Environment and Ecology 5,  $10 - 143.$ 

2. Bailey R.L., Dell T.R., 1973. Quantifying diameter distributions with the Weibull function. Forest Science 19, 97–104.

3. Cao Q.V., McCarty S.M., 2005. Presented at the Thirteenth Biennial Southern Silvicultural Research Conference. Memphis, TN.

4. Carretero A.C. and Torres-Alvarez E., 2013. Modelling diameter distributions of Quercus suber L. stands in "Los Alcornocales" Natural Park (Cádiz– Málaga, Spain) by using the two parameter Weibull functions. Forest Syst 22: 15-24

5. Ferreira W.N., Lacerda C.F., Costa R.C., Medeiros F.S., 2015. Effect of water stress on seedling growth in two species with different abundances: the importance of Stress Resistance Syndrome in seasonally dry tropical forest. Acta Bot Bras 29, 375-382

6. Gorgoso-Varela J.J., Rojo – Alboreca A., 2014. Use of Gumbel and Weibull functions to model extreme values of diameter distributions in forest stands. Ann Forest Sci 71, 741-750.

7. Ige P.O., Akinyemi G.O., Abi E.E., 2013. Diameter Distribution Models for Tropical Natural Forest trees in Onigambari Forest Reserve. JNSR 3: 14-23.

8. Jimoh S.O., Adesoye P.O., Adeyemi A.A., Ikyaagba E. T., 2012. Forest Structure Analysis in the Oban Division of Cross River National Park, Nigeria. Journal of Agricultural Science and Technology B 2, 510-518.

9. Li, F., Zhang, L., Davis, C. J., 2002. Modeling the joint distribution of tree diameters and heights by bivariate generalized beta distribution. Forest Science, 48(1), 47–58.

10. Lima R.A.F., Batista J.L.F., Prado P.I., 2014. Modeling Tree Diameter Distributions in Natural Forests: An Evaluation of 10 Statistical Models. Forest Sci 61, 320- 327.

11. Mabvurira D., Maltamo M., Kangas A., 2002. Predicting and calibrating diameter distributions of Eucalyptus grandis (Hill) Maiden plantations in

Zimbabwe. New Forests 23, 207–223.

12. Machodo S.A., Augustynczik A.L.D., Nascimento R.G.M, Téo S.J., Miguel E.P., Figura M.A., Silva L.C.R., 2009. Diametric distribution functions in a fragment of mixed ombrophyllous Forest. Cienc Rural 39, 2428- 2434.

13. Mataji A., Hojati M., Namiranian M., 2000. The study of quantity distribution in diameter levels in natural forests by using probability distribution. Iran's natural resources magazine, 53(2), 165-171.

14. Namiranian M., 1990. The application of probability theories in determining trees distribution in different diameter levels. Iranian Journal natural resource, 44, 93-108.

15. Nanang D.M., 1998. Suitability of the normal, log–normal and Weibull distributions for fitting diameter distributions of Neem plantations in Northern Ghana. Forest Ecol Manag 103, 1-7.

16. Nguyen Quang Phuc, 2019. Some structural characteristics and tree species diversity of three natural forest states in Son La province*.* Master thesis, Vietnam National University of Forestry, Hanoi (Vietnamese)

17. Palahí M., Pukkala T., Blasco E., Trasobares A., 2007. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north–east of Spain). Eur J Forest Res 126, 563-571.

18. Pham Quy Van, 2018. Some structural characteristics and tree species diversity of natural forest status IIIA in An Lao district, Binh Dinh province, Journal of Forestry Science and Technology, No. 1, pages 69 - 78.

19. Podlaski R., 2006. Suitability of the selected statistical distributions for fitting diameter data in distinguished development stages and phases of near– natural mixed forests in the Swietokrzyski National Park (Poland). Forest Ecol Manag 236, 393-402.

20. Podlaski R., Zasada M., 2008. Comparison of selected statistical distributions for modelling the diameter distributions in near–natural Abies–Fagus forests in the Świętokrzyski National Park (Poland). Eur J Forest Res 127, 455-463.

21. Rennolls K., Geary D.N., Rollinson T.J.D., 1985. Characterizing diameter distributions by the use of the Weibull distribution. Forestry 58, 57–66.

22. Schreuder H.T., Swank W.T., 1974. Coniferous stands characterized with the Weibull distribution. Canadian Journal of Forest Research 4, 518–523.

23. Vanclay J.K., 1994. Modelling forest growth and yield: applications to mixed tropical forests. CAB International, Wallingford, UK, 330 p.

24. Zohrer, F., 1972. The beta distribution for best fit of stem diameter distribution. 3rd Conf. Advisory Group For. Stat. Proc. IUFRO, Institut National Recherche Agronomique, Paris.

# **MÔ PHỎNG PHÂN BỐ ĐƯỜNG KÍNH CỦA RỪNG MƯA NHIỆT ĐỚI Ở VIỆT NAM THEO NĂM HÀM XÁC SUẤT**

# **Cao Thị Thu Hiền1, Nguyễn Hồng Hải1, Nguyễn Phúc Trường2**

*1 Trường Đại học Lâm nghiệp 2 Phân viện Điều tra Quy hoạch Đông Bắc bộ*

## **TÓM TẮT**

Nghiên cứu này nhằm mô phỏng phân bố đường kính của rừng mưa nhiệt đới ở Việt Nam và lựa chọn phân bố phù hợp. Số liệu được thu thập từ 20 ô đo đếm, mỗi ô có diện tích 1 ha. Các hàm sau đây đã được thử nghiệm: Beta, Weibull (hai và ba tham số), Lognormal và Gamma. Các tham số của các hàm phân bố được ước tính bằng phương pháp tối đa hợp lý. Tiêu chuẩn Kolmogorov – Smirnov được sử dụng để lựa chọn hàm phân bố phù hợp nhất. Kết quả cho thấy phân bố đường kính của rừng mưa nhiệt đới được mô tả tốt nhất bằng phân bố ba tham số Weibull, tiếp theo là phân bố Long – normal, trong khi đó hàm Weibull 2P không thành công trong việc mô phỏng các phân bố này. Các tham số µ, β, và α của phân bố Weibull 3P lần lượt dao động trong khoảng 5,697 – 10,155; 1,029 – 1,606, and 6,921 – 15,234, trong khi đó các tham số  $\mu$  và  $\delta$  của phân bố Log – normal lần lượt dao động trong khoảng 2,804 – 2,891 và 6,921 – 15,234. Phân bố đường kính có xu hướng số lượng cây giảm dần khi cỡ đường kính tăng lên.

**Từ khóa: Cấu trúc đường kính, phân bố liên tục, rừng mưa nhiệt đới, tiêu chuẩn Kolmogorov–Smirnov, tối đa hợp lý.**

