Stochastic differential equation approach of height-diameter equations of individual trees

Petras Rupšys[1,2]

[1]*Institute of Forest Management and Wood Sciences, Aleksandras Stulginskis University, Kaunas, LT-53361, Lithuania*
[2]*Centre of Mathematics, Physics and Information Technologies, Aleksandras Stulginskis University, LT-53361, Lithuania*
petras.rupsys@asu.lt

Abstract: *In this paper we use a stochastic differential equation to describe the dynamic evolution of the height of an individual tree. The first model is defined by Gompertz shape stochastic differential equation. The second model is defined by Gompertz stochastic differential equation with a threshold parameter. This model can be considered as an extension of the three parameter stochastic Gompertz model with the addition of a fourth parameter. The parameters are estimated by considering discrete sampling of the diameter and height and by using maximum likelihood procedure. Two developed models were employed to compare predicted values with observed values of a height. Performance statistics for developed height-diameter equations included statistical indexes, Shapiro-Wilk test and normal probability plot. We used the data of tropical Atlantic moist forest trees in southeastern Brazil (Scaranello et al., 2012) to validate our modelling technique. Results indicated that our model is able to capture the behaviour of tree height quite accurately. All results were implemented in a symbolic algebra system MAPLE.*

Keywords: diameter, height, mean, stochastic differential equation, threshold parameter, transition density function.

Introduction

In most applications foresters are interested in predicting tree height of a particular tree if only diameter is known. Tree height-diameter relationship establishes quantitative relations between two key characteristic dimensions of trees and is an important component for describing vertical stand structure and for estimating stand volume and site quality. This relationship varies between tree species and stands. Numerous height-diameter mathematical equations of sigmoidal and concave-shaped have been developed using only diameter outside bark at breast height as the predictor variable (Arabatzis and Burkhart, 1992; Rupšys and Petrauskas, 2010c; Petrauskas et al., 2011; Rupšys, 2012; Scaranello et al., 2012). It is well known that many parameters of a tree, including stem density, age, are essential for the dynamic behaviour of the size of a tree height (Hummel, 2000; Ouzennou et al., 2008; Prieditis et al., 2012). Stochasticity may as well play an interesting role in the dynamic behaviour of tree parameters (Rupšys et al., 2007, 2011; Rupšys and Petrauskas, 2010a,b, 2012).

Diameter and height dynamics is affected by many processes and varies among stands. The base assumption of traditionally used regression models is that the observed variations from the regression curve are constant at different values of a diameter would be realistic if the variations were due to measurement errors. Instead, it is unrealistic, as the variations are due to random changes on growth rates induced by random environmental perturbations. Stochastic differential equations models do not have such weakness (Rupšys and Petrauskas, 2012). The modelling of the height-diameter process leads to an equation for the stochastic variable - height, such as a stochastic differential equation, or for an equation which predicts how the probability density function for the height changes in diameter. Stochastic processes are ubiquitous in the physical, biological and social sciences (Allen, 2007).

Stochastic height-diameter dynamics models allow us to reduce the unexplained variability of a height. In recent decades, few models have been put forward to explain stochastic behaviour of diameter and height (Rupšys and Petrauskas, 2010a, b, 2012; Rupšys et al., 2007, 2011). These dynamics are basically the classical deterministic logistic growth dynamics, extended by a level-dependent diffusion term. In reality, external factors such as climate, terrain, the presence of other tree species, and indeed any factor which has an uncertain influence on tree height, will also affect the intrinsic growth rate. This can be modelled by adding an external random term to the intrinsic growth rate, α, which represents this environmental stochasticity. Although many refinements and extensions are possible, the basic dynamics model for height process $H(d),\ d \geq 0$ can be described by the Itô's (1942) univariate stochastic differential equation

$$dH(d) = \mu(H(d),\theta)dd + \sigma(H(d),\theta)dW(d) \tag{1}$$

where $W(d),\ d \geq 0$ is a standard Brownian motion. Intuitively, we interpret the term $dW(\cdot)$ as ecological and environmental noise. Parametric approach assume that the drift $\mu(H(d),\theta)$ and diffusion $\sigma(H(d),\theta)$ are known functions except for an unknown parameter vector θ. Examples include (Rupšys and Petrauskas, 2010a,b, 2012; Rupšys et al., 2007, 2011). Parametric stochastic differential equations often provide a convenient way to describe the dynamics of tree data, and a great deal of effort has been expended searching for efficient ways to estimate model parameters. Maximum likelihood is typically the estimator of choice (Rupšys et al., 2007; Rupšys and Petrauskas, 2010a, b).

Following the recent trends in stochastic differential equations, we develop a stochastic height-diameter model using the Gompertz shape tree growth. Multivariate models can deal, for instance, with multiple explanatory factors (diameter, stem, basal area) in asset tree height.

This article is an attempt to design a model for predicting the height of a tree from its diameter-varying univariate distribution. For modelling the stochastic process of a height, we use the family of the Gompertz shape stochastic differential equations that are reducible to an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930).

The aim of this study is to put forward the advantages of using stochastic differential equations in the analysis of height-diameter curves and to show how an adequate model can be made. In this paper attention is restricted to homogeneous stochastic differential equation in the Gompertz type (Rupšys et al., 2007; Rupšys and Petrauskas, 2009, 2010a), whose solution produces the regression term of the fixed effects model. We also discuss how transition density function can be used to construct maximum likelihood estimators. We present an application of stochastic differential equation approach dealing with the study of the height-diameter dynamics of tropical trees. A MAPLE program was implemented to carry out the calculations required for this study.

Materials and methods

We model the dynamic of a height as a stochastic process over diameter. In this study, we select to use the deterministic ordinary differential equation of the Gompertz type (Gompertz, 1825) as the basis of our new developed stochastic height-diameter model. The change in tree height, h(d), is described using the ordinary differential equation

$$\frac{dh(d)}{dd} = \alpha h(d) - \beta h(d)\ln\left(h(d)\right), \tag{2}$$

where α is the intrinsic growth rate of the height and β is the growth deceleration factor. The parameters α and β characterize the evolution of a height of different tree species and stands. The formula describing deterministic Gomperz shape height-diameter trajectory admits the form of a sigmoidal function

$$h(d) = \exp\left(\frac{\alpha}{\beta} - \left(\frac{\alpha}{\beta} - \ln(h_0)\right)\exp(-\beta d)\right),\ d \in \left[0; D_0\right], \tag{3}$$

where $h_0 = h(0) = 1.37$. From the solution one can easily see the non-trivial equilibrium point $h(\propto) = \exp\left(\frac{\alpha}{\beta}\right)$ representing the largest height size that a tree can tolerate (i.e. carrying capacity). There also exists an inflection point $h^* = \exp\left(\frac{\alpha}{\beta} - 1\right)$ corresponding to the maximum growth rate of a height, which reflects the self regulation effect by an intrinsic growth control mechanism.

There are alternative ways of introducing stochasticity in the behaviour of a height. In this work, we approximate the randomness in the operation of tree heights as a standard Brownian motion. Therefore, we convert the complete deterministic model defined by Eq.(2) for the tree height into a stochastic model assuming that the intrinsic growth rate varies in diameter according to

$$\alpha(d) = \alpha + \sigma\varepsilon(d), \tag{4}$$

where α is the constant mean value of $\alpha(d)$, σ is the diffusion coefficient, and $\varepsilon(d)$ is a Gaussian white noise process. We describe the height, H(d), using stochastic differential equation of the form

$$dH(d) = \left[\alpha H(d) - \beta H(d)\ln(H(d))\right]dd + \sigma H(d)dW(d),\ P\left(H(0) = 1.37\right) = 1,\ d \in \left[0; D_0\right]. \tag{5}$$

By Ito’s lemma Eq.(5) implies that the exponent transformation ψ ≡ ln(h) follows the Ornstein-Uhlenbeck process. This transformation changes the state-space R^+ into R and allows us to obtain the non-conditional probability density function for the considered height process, resulting in

$$f\left(h, d\right) = \frac{1}{h\sqrt{2\pi v(d)}}\exp\left(-\frac{1}{2v(d)}\left(\ln h - \mu(d)\right)^2\right), \tag{6}$$

which corresponds with a lognormal distribution, where

$$\mu(d) = \ln(1.37) * e^{-\beta d} + \left(\alpha - \frac{\sigma^2}{2}\right)\left(\frac{1-e^{-\beta d}}{\beta}\right), \tag{7}$$

$$v(d) = \frac{1-e^{-2\beta d}}{2\beta}\sigma^2. \tag{8}$$

The non-conditional mean trend and variance functions of the stochastic height process is given by the following expressions, respectively, by

$$h(d) = \exp\left(\ln(1.37)e^{-\beta d} + \frac{1-e^{-\beta d}}{\beta}\left(\alpha - \frac{\sigma^2}{2}\right) + \frac{\sigma^2}{4\beta}\left(1-e^{-2\beta d}\right)\right), \tag{9}$$

$$va(d) = \exp\left(2\left(\ln(1.37)e^{-\beta d} + \frac{1-e^{-\beta d}}{\beta}\left(\alpha - \frac{\sigma^2}{2}\right)\right) + \frac{\sigma^2}{2\beta}\left(1-e^{-2\beta d}\right)\right)\left(\exp\left(\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta d}\right)\right) - 1\right). \tag{10}$$

Next we propose a new stochastic Gompertz shape height-diameter model with a threshold parameter. This model can be considered as an extension of the three parameter stochastic Gompertz process with the addition of a fourth parameter (Gutierrez et al., 2006; Rupšys et al., 2011). We describe the height, H(d), by stochastic differential equation of the form

$$dH(d) = \left[\alpha(H(d)-\gamma) - \beta(H(d)-\gamma)\ln(H(d)-\gamma)\right]dd + \sigma(H(d)-\gamma)dW(d),\ P(H(0)=1.37)=1,\ d \in [0; D_0]. \tag{11}$$

The non-conditional probability density function for the considered height process (Eq. (11)) is defined in the following form

$$f^t(h,d) = \frac{1}{(h-\gamma)\sqrt{2\pi v^t(d)}}\exp\left(-\frac{1}{2v^t(d)}\left(\ln(h-\gamma) - \mu^t(d)\right)^2\right) \tag{12}$$

which corresponds with a lognormal distribution, where

$$\mu^t(d) = \ln(1.37-\gamma)e^{-\beta d} + \frac{1-e^{-\beta d}}{\beta}\left(\alpha - \frac{\sigma^2}{2}\right), \tag{13}$$

$$v^t(d) = \frac{1-e^{-2\beta d}}{2\beta}\sigma^2. \tag{14}$$

The non-conditional mean trend and variance functions of the height process is given by the following expressions, respectively, by

$$h^t(d) = \gamma + \exp\left(\ln(1.37-\gamma)e^{-\beta d} + \frac{1-e^{-\beta d}}{\beta}\left(\alpha - \frac{\sigma^2}{2}\right) + \left(\frac{\sigma^2}{4\beta}\left(1-e^{-2\beta d}\right)\right)\right), \tag{15}$$

$$va^t(d) = \exp\left(2\left(\ln(1.37-\gamma)e^{-\beta d} + \frac{1-e^{-\beta d}}{\beta}\left(\alpha - \frac{\sigma^2}{2}\right)\right) + \frac{\sigma^2}{2\beta}\left(1-e^{-2\beta d}\right)\right)\left(\exp\left(\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta d}\right)\right) - 1\right). \tag{16}$$

Computation of the parameter estimators

The drift and diffusion parameters α, β and σ are estimated by means of the maximum likelihood procedure using discrete sampling and non-conditional probability density functions (Eqs.(6), (12)), as we assume that all observations (measured trees) are independent. Let us consider a discrete sample of the process ($h_1, h_2, \ldots, h_n$) at the diameters ($d_1, d_2, \ldots, d_n$). Under the initial condition $P(H(0)=1.37)=1$, the associate likelihood function can be obtained by the following expressions, respectively

$$L(\alpha,\beta,\sigma) = \prod_{i=1}^{n} f(h_i,d_i) = \prod_{i=1}^{n}\frac{1}{h_i\sqrt{2\pi v(d_i)}}\exp\left(-\frac{1}{2v(d_i)}\left(\ln h_i - \mu(d_i)\right)^2\right), \tag{17}$$

$$L(\alpha,\beta,\sigma,\gamma) = \prod_{i=1}^{n} f^t(h_i,d_i) = \prod_{i=1}^{n}\frac{1}{(h_i-\gamma)\sqrt{2\pi v^t(d_i)}}\exp\left(-\frac{1}{2v^t(d_i)}\left(\ln(h_i-\gamma) - \mu^t(d_i)\right)^2\right). \tag{18}$$

Data

We focus on the modelling of a tropical Atlantic forest tree data set. Here, we analyse tropical forest tree height-diameter database of 280 individual tree height and diameter measurements across plots along an altitudinal gradient published by Scaranello et al. (2012). Our aim is to improve understanding of tropical tree variability and reduce uncertainty of tree height estimates at the altitudinal scale. Summary statistics for diameter outside bark at breast height (D) and total height (H) of all trees used for parameters estimate are presented in Table 1.

Table 1

Summary statistics of the data set (diameter at breast height (D) and total height (H))

Altitude	Count	Variable	Min	Max	Mean	St. Dev.
Sea level	61	D (cm)	4.8	76.9	20.4	15.3
		H (m)	3.0	19.0	10.2	4.2
100 m	73	D (cm)	6.0	75.1	30.5	20.0
		H (m)	4.0	22.0	11.7	4.7
400 m	77	D (cm)	4.9	79.0	30.6	20.7
		H (m)	4.0	25.0	11.3	4.9
1,000 m	79	D (cm)	4.9	100.4	28.6	24.3
		H (m)	3.5	30.0	13.4	6.6
All levels	280	D (cm)	4.8	100.4	27.8	20.9
		H (m)	3.0	30.0	11.7	5.3

Results and discussion

Using the estimation data set presented in Table 1, the parameters of height-diameter models defined by stochastic differential equations (5), (11), were estimated by the maximum likelihood procedure (Eqs. (17), (18)). Estimation results are presented in Table 2.

In general, both stochastic height-diameter models produced relatively high root square mean errors (3.49 m and 3.38 m, respectively) and explained a relatively low proportion of the total variation in observed values of the tree height, accounting for only 57.7% and 60.3%. Nevertheless, these results may not be surprising since the height-diameter relationships found in the data were highly variable and scattered (see Table 1 and Fig. 1).

As we can see in Table 2, the relationship between stem height and diameter are altered by environmental conditions. Stand-specific characteristics such as soil type, nutrient status, elevation cause parameters to differ across stands. Thus, specific stands may have what are generally termed “random parameters” in mixed effects model terminology. Equations (5) and (11) can be altered by adding stand-specific random effects to the population fixed effects parameters to produce stand-specific parameters in the following form

Table 2

Estimated parameters of both models applied to the height-diameter analysis data set

Models	Altitude	Parameters			
		α	β	σ	γ
Eq. (5)	Sea level	0.3872	0.1439	0.1503	-
	100 m	0.3405	0.1236	0.1348	-
	400 m	0.3491	0.1289	0.1456	-
	1,000 m	0.3876	0.1299	0.1386	-
	All levels	0.3698	0.1326	0.1521	-
Eq. (11)	Sea level	0.2908	0.0939	0.0684	-7.0171
	100 m	0.2118	0.0616	0.0368	-14.2301
	400 m	0.2098	0.0628	0.0451	-11.9547
	1,000 m	0.2318	0.0671	0.0496	-10.4416
	All levels	0.2232	0.0644	0.0440	-14.3375

$$\alpha(d) = \alpha + u_{1i}, \tag{19}$$

$$\beta(d) = \beta + u_{2i}, \tag{20}$$

$$\gamma(d) = \gamma + u_{3i}, \tag{21}$$

where u_{1i}, u_{2i}, u_{3i} (i=1, 2, 3, 4) - stand-specific random effects, assumed to be independent and normally distributed with 0 mean and constant variance ($u_{ki}\sim N(0,\sigma^2_k)$, k=1,2,3). Additionally, a covariance, σ_{kl}, can be assumed to exist between u_{ki} and u_{li}.

The influence of the altitude within Atlantic forests on the height-diameter mean and standard deviation curves is illustrated in Fig. 1 and Fig. 2.

For the evaluation of goodness-of-fit of our developed stochastic Gompertz shape height-diameter models (Eqs. (5), (11)) we used the Shapiro-Wilk statistic and normal probability plot. The p-values of the Shapiro-Wilk statistic were 0.005 and 0.195, respectively. The normal probability plots of the pseudo-residuals using the estimates of parameters presented in Table 1 showed that both new developed height-diameter models fit not too bad.

The coefficient of variation is usually used as a measure of precision for the dispersion of data sets and is also often used to compare numerical distributions measured on different scales. The coefficient of variation of a tree height measures the variability of the tree height relative to its mean and it relates the mean and standard deviation by expressing the standard deviation as a percentage of a mean. To further discuss the results of the paper, we provide the coefficient of variation that might to be of considerable interest to compare the dispersion of tree height running at the diameter d, which is defined by

$$CV(d)=\frac{\sqrt{va(d)}}{h(d)}\cdot 100 \qquad (22)$$

Fig.3 shows plot of the coefficient of the variation against diameter. In both cases the coefficient of variation of tree height monotonically evolve to stationary coefficient of variation.

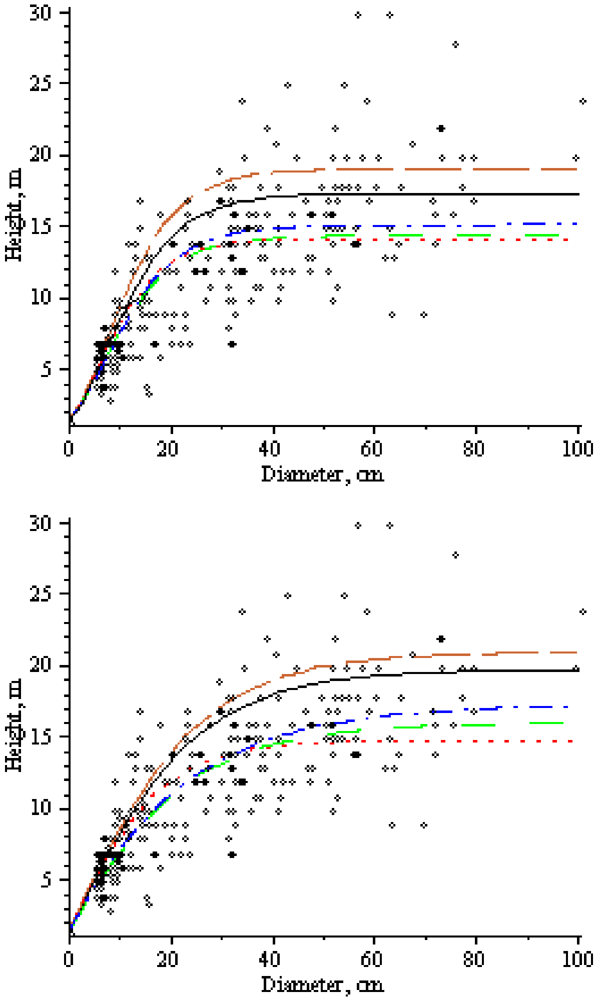

Fig. 1. **Plot of the mean dynamics of a tree height with the parameterization data sets:** in the left – Eq. (5), in the right – Eq. (11), using parameterization data set of sea level altitude – dot (red), using parameterization data set of 100 m altitude – dash dot (blue), using parameterization data set of 400 m altitude – space dash (green), using parameterization data set of 1000 m altitude – long dash (gold), using parameterization data set of all altitudes – solid line (black).

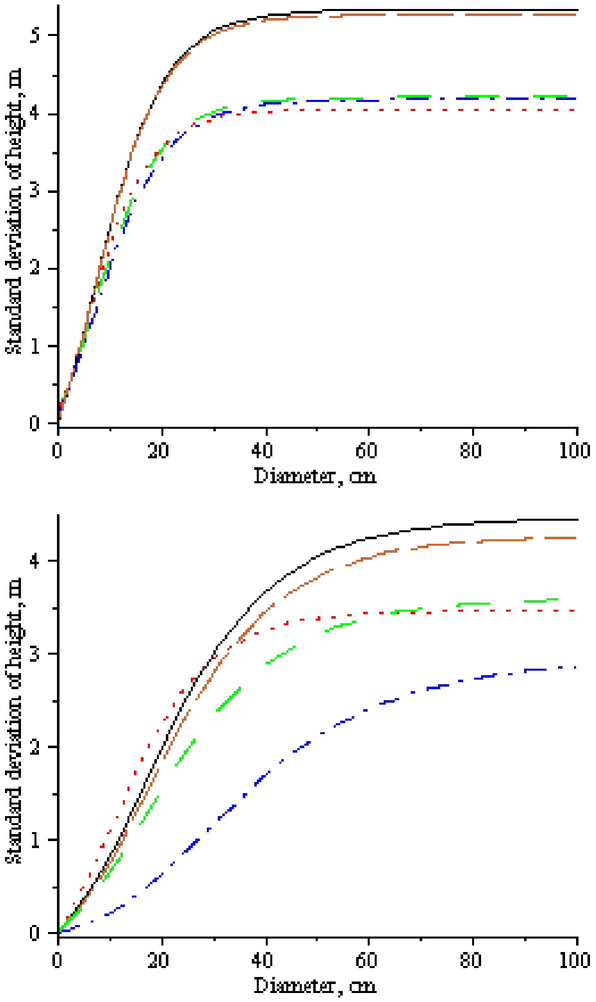

Fig. 2. **Plot of the standard deviation dynamics of a tree height with the parameterization data sets:** in the left – Eq. (5), in the right – Eq. (11), using parameterization data set of sea level altitude – dot (red), using parameterization data set of 100 m altitude – dash dot (blue), using parameterization data set of 400 m altitude – space dash (green), using parameterization data set of 1000 m altitude – long dash (gold), using parameterization data set of all altitudes – solid line (black).

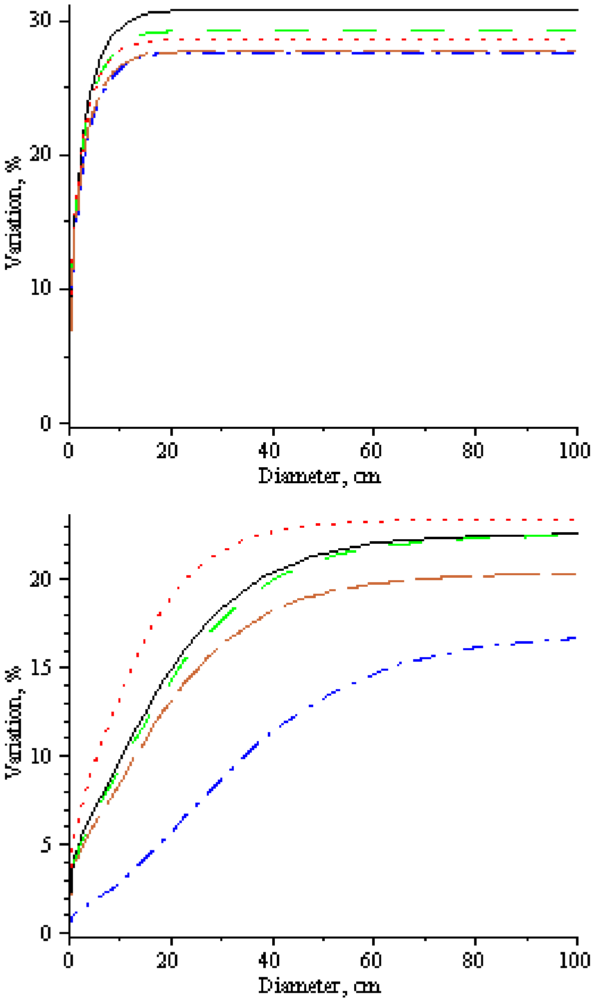

Fig. 3. **Plot of the variation dynamics of a tree height with the parameterization data sets:** in the left – Eq. (5), in the right – Eq. (11), using parameterization data set of sea level altitude – dot (red), using parameterization data set of 100 m altitude – dash dot (blue), using parameterization data set of 400 m altitude – space dash (green), using parameterization data set of 1000 m altitude – long dash (gold), using parameterization data set of all altitudes – solid line (black).

Conclusion

The new height-diameter models were developed using stochastic Gompertz shape differential equations. Comparison of the predicted height values calculated using stochastic differential equations (5), (11) with the observed values revealed a comparable predictive power of the stochastic height model (11).

The developed stochastic models may be recommended both for their ease of fitting procedures and the biological interpretations of the relevant parameters.

The stochastic differential equations approach allows us to incorporate new tree variables, mixed-effect parameters and new forms of stochastic dynamics. The accuracy of the height-diameter-dependent non-conditional density functions (Eqs. (6), (12)) depends on the amount of information available from the stand. Our methodology extends some way to inclusion of the basal-area or/and density of a stand as an exogenous factor or as an independent variable.

The variance functions developed here can be applied generate weights in every linear and nonlinear least squares regression height model by the weighted least squares form.

Finally, stochastic differential equation methodology may be of interest in diverse of areas of research that are far beyond the modelling of a tree height.

References

Allen**,** E., 2007. *Modeling with Itô Stochastic Differential Equations*, Springer, 230p.

Arabatzis, A.A. and Burkhart, H.E., 1992. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. *Forest Science*, 38, pp.192-198.

Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil. Trans. Roy. Soc.*, 115, pp.513-585.

Gutierrez, R., Gutierrez-Sanchez, R., Nafidi, A. and Ramos, E., 2006. A new stochastic Gompertz diffusion process with threshold parameter: Computational aspects and applications. *Applied Mathematics and Computation*, 183, pp.738–747.

Hummel, S., 2000. Height, diameter and crown dimensions of *Cordia alliodora* associated with tree density. *Forest Ecology and Management*, 127, pp.31–40.

Itô, K., 1942. On stochastic processes. *Japan Journal of Mathematics*, 18, pp.261-301.

Ouzennou, H., Pothier, D. and Raulier, F., 2008. Adjustment of the age–height relationship for uneven-aged black spruce stands. *Canadian Journal of Forest Research*, 38, pp.2003–2012.

Petrauskas, E., Rupšys, P. and Memgaudas, R., 2011. Q-exponential variable form of a stem taper and volume models for Scots pine (*Pinus Sylvestris*) in Lithuania. *Baltic Forestry*, 17, pp. 118-127.

Prieditis, G., Šmits, I., Archipova I., Daģis, S. and Dubrovskis, D., 2012. Allometric models for predicting tree diameter at breast height. In: Ramos R.A.R., Straupe I., Panagopoulos T. (eds) Recent Researches in Environment, Energy Systems and Sustainability. *8th WSEAS International Conference on Energy, Environment, Ecosystems and Sustainable Development*, WSEAS LLC, pp.105-110.

Rupšys, P., 2012. On the Use of q-exponential functions for developing stem profile and volume models. *International Journal of Biological Engineering*, 2(5), pp.48-55.

Rupšys, P., Bartkevičius, E. and Petrauskas, E., 2011. A univariate stochastic Gompertz model for tree diameter modeling. *Trends in Applied Sciences Research*, 6, pp.134-153.

Rupšys, P. and Petrauskas, E., 2012. Analysis of height curves by stochastic differential equations. *International Journal of Biomathematics*, 5(5), 1250045. doi No: 10.1142/S1793524511001878.

Rupšys, P. and Petrauskas, E., 2010a. The bivariate Gompertz diffusion model for tree diameter and height distribution. *Forest Science*, 56, pp.271-280.

Rupšys, P. and Petrauskas, E., 2010b. Quantifying tree diameter distributions with one-dimensional diffusion processes. *Journal of Biological Systems*, 18, pp.205-221.

Rupšys, P. and Petrauskas, E., 2010c. Development of q-exponential models for tree height, volume and stem profile. *International Journal of the Physical Sciences*, 5(15), pp.2369-2378.

Rupšys, P. and Petrauskas, E., 2009. Forest harvesting problem in the light of the information measures. *Trends in Applied Sciences Research*, 4, pp.25-36.

Rupšys, P., Petrauskas, E., Mažeika, J. and Deltuvas, R., 2007. The Gompertz type stochastic growth law and a tree diameter distribution. *Baltic Forestry*, 13, pp.197-206.

Scaranello, M.A., Alves, L.F., Vieira, S.A., Camargo, P.B., Joly, C.A. and Martinelli, L.A., 2010. Height-diameter relationships of tropical Atlantic moist forest trees in southeastern Brazil. *Scientia Agricola*, 69, pp.26-37.

Uhlenbeck, G.E. and Ornstein, L.S., 1930. On the theory of Brownian Motion. *Physical Review*, 36, pp.823–841.