The Use of Stochastic Differential Equations to Describe Stem Taper and Volume

EDMUNDAS PETRAUSKAS[2,*], EDMUNDAS BARTKEVIČIUS[1], PETRAS RUPŠYS[2,3] AND ROMAS MEMGAUDAS[2]

Aleksandras Stulginskis University

[1]Institute of Forest Biology and Silviculture, [2]Institute of Forest Management and Wood Science, [3]Centre of Mathematics, Physics and Information Technology,

Aleksandras Stulginskis University, Studenų g. 11, Akademija, Kaunas district. LT – 53361 LITHUANIA, Phone +370 37 752276; Email: petras.rupsys@asu.lt

Petrauskas, E., Bartkevičius, E., Rupšys, P. and Memgaudas, R. 2013. The Use of Stochastic Differential Equations to Describe Stem Taper and Volume. *Baltic Forestry* 19(1): 143–151.

Abstract

An approach combining the information generated from different stochastic differential equations was developed to improve the predictive qualities of stem taper and volume. The stochastic differential equations and the stem taper and volume models were fitted to data from Scots pine and Norway spruce trees that were collected from across the entire Lithuanian territory. New models deduced from the Gompertz and Ornstein-Uhlenbeck shape stochastic differential equations were tested against the classical Kozak's stem taper model, q-exponential segmented stem taper model, classical Schumacher-Hall's volume model, and q-exponential volume model based on allometric and geometric concepts. Comparison of the predicted stem taper and stem volume values with those obtained using regression fixed-effects models demonstrated the predictive power of the stochastic differential equations models.

Key words: stochastic differential equation, stem taper model, volume model

Introduction

Predictive forestry is a specific application of the field of mathematical modelling to describing the behaviour of an individual tree and stand under a given set of environmental conditions. Height and volume predictive models have been widely accepted as informative tools that provide quick and cost-effective assessments of tree growth for product development, risk analysis, and recreation purposes. In this work, we study individual tree growth models that take into account the effects of random environmental perturbations on growth. Traditionally, the relationship between volume, height and diameter has been modelled based on simple linear and nonlinear regressions (Tabacchi et al. 2011, Serinaldi et al. 2012). The base assumption of these regression models is that the observed variations from the regression curves that are independent of covariate values would be realistic if these variations were due to measurement errors. However, this assumption is unrealistic, as these variations are due to random changes in growth rates induced by random environmental perturbations. Stem volume and stem phytomass data always exhibit heteroscedasticity (Parresol 1999, Tabacchi et al. 2011), the error variances are not constant across all observations. With these considerations in mind, we present a methodology for taper analysis as an alternative to other statistical techniques. In this paper, the developed novel stochastic differential equation model is not affected by the limitations described above.

Taper equations are widely used in forestry to estimate the diameter at any given height along a tree bole and, therefore, to calculate the total or merchantable stem volume (Alegria and Tome 2011, Özēelik et al. 2011). The classical taper model has been extensively studied to predict various diameter dynamics behaviours. One crucial element in taper models is the functional response that describes the relative diameter of tree stem consumed per relative height for given quantities of diameter at breast height D and total tree height H. The most commonly studied taper relations range from simple taper functions to more complex forms (Kozak et al. 1969, Demaerschalk 1972, Max and Burkhart 1976, Kozak 2004, Rupšys and Petrauskas 2010c, Petrauskas et al. 2011). Taper curve data consist of repeated measurements of a continuous diameter growth process relative to the height of individual trees. These longitudinal data have two characteristics that complicate their statistical analysis:

a) within-individual tree correlation that includes the data measured on a single tree and b) independent but extremely high variability between the experimental taper curves obtained for different trees. Mixed-effects models incorporate the variability between individual trees using the expression of the model's parameters in terms of both fixed and random effects. Random effects are conceptually random variables. They are modelled as such by describing their distributions. A large number of mixed-effects taper models have been published (see, Yang et al. 2009, Trincado and Burkhart 2006, Westfall and Scott 2010, Özēelik et al. 2011).

The increasing popularity of mixed-effects models can be attributed to their ability to model total variation by splitting the variation within- and between-individual tree components. We propose to model these variations and nonlinearities using stochastic differential equations that are deduced from the standard deterministic growth function by adding random variations to the growth dynamics (Suzuki 1971, Tanaka 1986, Rupšys et al. 2007, 2011, Rupšys and Petrauskas 2010a-b, 2012). We thus consider stochastic differential equation models with drift and diffusion terms that can depend linearly or nonlinearly on state variables. The basis of the work is a segmented model of the tree taper, which uses different models for different parts of the stem to overcome local bias. In this paper, an effort has been made to develop two stochastic differential equation segmented stem taper models. We assume that the lower part of a tree trunk can be modelled by a nonlinear dynamical system with multiplicative noise (called the Gompertz stochastic differential equation), in which the random perturbations of the growth rate are proportional to the relative diameter. Additionally we assume, that the middle and upper parts can be modelled by a linear dynamical system with additive noise (called the Ornstein-Uhlenbeck stochastic differential equation), in which the perturbations in the relative diameter do not depend upon the relative height.

The aim of this study is to communicate the advantages of using stochastic differential equations in the analysis of taper models and to demonstrate how an adequate model can be developed. In this paper, attention is restricted to homogeneous stochastic differential equations of the Gompertz, and Ornstein-Uhlenbeck types (Gutiérrez et al. 2008, Rupšys and Petrauskas 2010a, Picchini et al. 2011), the solutions of which produce the regression terms of the fixed effects model.

Materials and methods

Data

We focus on the modelling of Scots pine (*Pinus Sylvestris*) and Norway spruce (*Picea Abies*) tree data sets. Scots pine and Norway spruce tree stands dominate Lithuanian forests, grow on *Arenosols* and *Podzols* forest sites and cover 725,500 ha, and 427,000 ha, respectively. Stem measurements for 598 Scots pine trees and 1021 Norway spruce trees were used for stem volume and taper model' analyses. All data were collected during 1979-2008 across the entire Lithuanian territory, except for Kuršių Nerija National Park (latitude, 53°54' - 56°27' N; longitude, 20°56' - 26°51' E; altitude, 10 - 293 m). Mean temperatures vary from -16.4 C° in winter to +22 C° in summer. Precipitation is distributed throughout the year, although it occurs predominantly in summer, the average precipitation is, approximately, 680 mm per year. Temporary circle test plots were placed in each of 42 Lithuanian state forests in randomly selected clear-cutting areas. The diameter over the bark and the diameter under the bark of each stem in a plot were measured at regular intervals, starting from the diameter of the root collar and at heights 1 metre, 1.3 metres, 3 metres, 5 metres, etc. All section measurements include 7,783 data points (Scots pine trees) and 12,999 (Norway spruce trees). The diameter was measured to an accuracy of 1 mm. A random sample of 300 Scots pine trees (out of the total sample of 598 trees) was selected for model estimation, and the remaining data set of 298 Scots pine trees was utilised for model validation. The Norway spruce sample data set of 1,021 trees was also randomly divided: 497 trees were used for model estimation and 524 were used for model validation. Summary statistics for the diameter over the bark at breast height (D), total height (H), volume (V) and age (A) of all the trees used for model estimation and validation are presented in Table 1.

Table 1. Summary statistics of the data sets for the Scots pine and Norway spruce trees across Lithuania

Data	Number of trees	Min	Max	Mean	St. Dev.	Number of trees	Min	Max	Mean	St. Dev.
	Estimation					Validation				
Scots pine										
D(cm)	300	6.3	53.8	24.2642	9.8518	298	3.8	58.5	24.9694	10.0412
H (m)	300	5.6	34.5	20.5669	5.4710	298	4.2	35.3	21.5621	5.8824
V(m^3)	300	0.0109	3.2139	0.5832	0.5745	298	0.0031	3.1957	0.6449	0.5642
A (yr)	300	23	161	77.1733	25.7327	298	19	140	71.6	29.6
Norway spruce										
D cm)	497	7.8	52.4	22.0887	8.5145	524	7.0	49.8	21.7330	8.2975
H (m)	497	7.6	33.1	20.5564	5.4908	524	7.7	33.9	20.5863	5.1744
V(m^3)	497	0.0183	2.9775	0.5285	0.4853	524	0.0230	2.9945	0.5113	0.4652
A (yr)	497	34	120	67.2937	19.2336	524	30	150	66.7233	20.4064

In our analysis, we interpret the stem volume calculated by equation (1) as the observed volume

$$V_i = \frac{\pi}{40000}\left(\sum_{k=1}^{n_i-2}\frac{(d_{ik}^2 + d_{ik+1}^2 + d_{ik}\cdot d_{ik+1})\cdot L_{ik}}{3} + \frac{d_{in_i-1}^2 \cdot L_{in_i-1}}{3}\right), \quad (1)$$

where d_{ik} is the diameter over the bark (cm) for section k of tree i and L_{ik} is the section length (m).

Stochastic Differential Equation Framework

The stochastic differential equations stem taper model, which accounts for the variations from the deterministic predictions that occur in a given longitudinal data series of individual tree growth was used to fit the longitudinal data series of the growth of each individual tree. The model used in this work incorporated environmental stochasticity, which accounts for variability in the diameter growth rate that arises from external factors (such as soil structure, water quality and quantity, and levels of various soil nutrients) that equally affect all the trees in the stands. In the diffusion stem taper model, the changes in the diameter over the bark between two consecutive heights are represented by a scalar diffusion process *Y(x)* that is indexed by relative height *x* and given by the Itō (1942) stochastic differential equation. Consider a one-dimensional continuous process *Y(x)* evolving in *M* different experimental units (trees) randomly chosen from a theoretical population (tree species). We suppose that the dynamics of the relative diameter $Y^i = d/D^i$ subject to the relative height $x^i = h/H^i$ is expressed by a stochastic differential equation, where is the diameter over the bark at any given height *h*, *D* is the diameter at breast height over the bark, *H* is the total tree height from ground to tip. The first utilised diffusion process of relative diameter dynamics is defined using the following Gompertz form (Gutiérrez et al. 2008, Rupšys and Petrauskas 2010a)

$$dY^i(x^i) = [\alpha_G Y^i(x^i) - \beta_G Y^i(x^i)\ln(Y^i(x^i))]dx^i + \sigma_G Y^i(x^i)dW_G^i(x^i), \quad P(Y^i(x_0^i) = y_0^i) = 1, \quad i = 1,...,M, \quad (2)$$

where $Y^i(x^i)$ is the value of the process at the relative height $x^i \geq x_0^i$ and α_G, β_G and σ_G are fixed effects parameters (identical for the entire population of trees). The $W_G^i(x^i)$, i=1,…,M – are mutually independent standard Brownian motions. The second model of relative diameter dynamics is defined using the following three parameter Ornstein-Uhlenbeck form (Picchini et al. 2011)

$$dY^i(x^i) = \left(\alpha_O - \frac{Y^i(x^i)}{\beta_O}\right)dx^i + \sigma_O dW_O^i(x^i), \quad P(Y^i(x_0^i) = y_0^i) = 1, \quad i = 1,...,M, \quad (3)$$

where α_O, β_O and σ_O are fixed effects parameters (identical for the entire population of trees) and $W_O^i(x^i)$, i=1,…,M – are mutually independent standard Brownian motions.

Assume that tree i is measured at n_i+1 discrete relative height points $(x_0^i, x_1^i..., x_{n_i}^i)$ and relative diameter points $(y_0^i, y_1^i,..., y_{n_i}^i)$, $y^i(x_j^i) = y_j^i$, i=1,…,M. The transition probability density functions of the random variable $Y^i(x_j^i)|Y(x_{j-1}^i) = y_{j-1}^i$ of two different relative diameter stochastic processes defined by Eq. (2)-(3) can be deduced in the following form because the stochastic Gompertz process is lognormal (Gutiérrez et al. 2008)

$$p_G(y_j^i, x_j^i | y_{j-1}^i, x_{j-1}^i, \alpha_G, \beta_G, \sigma_G) = \frac{1}{y_j^i\sqrt{2\pi v_G(x_j^i, x_{j-1}^i)}}, \quad (4)$$
$$\times \exp\left(-\frac{1}{2v_G(x_j^i, x_{j-1}^i)}\left(\ln y_j^i - \mu_G(x_j^i, x_{j-1}^i, y_{j-1}^i)\right)^2\right),$$

where

$$\mu_G(x_j^i, x_{j-1}^i, y_{j-1}^i) = \ln y_{j-1}^i e^{-\beta_G(x_j^i - x_{j-1}^i)} + \frac{1 - e^{-\beta_G(x_j^i - x_{j-1}^i)}}{\beta_G}\left(\alpha_G - \frac{\sigma_G^2}{2}\right), \quad (5)$$

$$v_G(x_j^i, x_{j-1}^i) = \frac{1 - e^{-2\beta_G(x_j^i - x_{j-1}^i)}}{2\beta_G}\sigma_G^2, \quad (6)$$

and because the stochastic Ornstein-Uhlenbeck process is normal (Picchini et al. 2011)

$$p_O(y_j^i, x_j^i | y_{j-1}^i, x_{j-1}^i, \alpha_O, \beta_O, \sigma_O) = \frac{1}{\sqrt{2\pi v_O(x_j^i, x_{j-1}^i)}} \times \exp\left(-\frac{(y_j^i - \mu_O(x_j^i, x_{j-1}^i, y_{j-1}^i))^2}{2v_O(x_j^i, x_{j-1}^i)}\right), \quad (7)$$

where

$$\mu_O(x_j^i, x_{j-1}^i, y_{j-1}^i) = y_{j-1}^i \exp\left(-\frac{x_j^i - x_{j-1}^i}{\beta_O}\right) + \alpha_O\beta_O\left(1 - \exp\left(-\frac{x_j^i - x_{j-1}^i}{\beta_O}\right)\right), \quad (8)$$

$$v_O(x_j^i, x_{j-1}^i) = \frac{\sigma_O^2\beta_O}{2}\left(1 - e^{-\frac{2(x_j^i - x_{j-1}^i)}{\beta_O}}\right). \quad (9)$$

The conditional mean and variance functions $m(x^i|\cdot)$ and $w(x^i|\cdot)$ (x^i is the relative height of *ith* tree) of the stochastic processes (2)-(3) are

$$m_G(x^i | y_0^i, \alpha_G, \beta_G, \sigma_G) = y_0^i e^{-\beta_G x^i} \exp\left(\frac{1 - e^{-\beta_G x^i}}{\beta_G}\left(\alpha_G - \frac{\sigma_G^2}{2}\right) + \frac{\sigma_G^2}{4\beta_G}\left(1 - e^{-2\beta_G x^i}\right)\right) \quad (10)$$

$$w_G(x^i | y_0^i, \alpha_G, \beta_G, \sigma_G) = \exp\left(2\left(\ln y_0^i e^{-\beta_G x^i} + \frac{1 - e^{-\beta_G x^i}}{\beta_G}\left(\alpha_G - \frac{\sigma_G^2}{2}\right)\right) + \frac{\sigma_G^2}{2\beta_G}\left(1 - e^{-2\beta_G x^i}\right)\right) \times \left(\exp\left(\frac{\sigma_G^2}{2\beta_G}\left(1 - e^{-2\beta_G x^i}\right)\right) - 1\right) \quad (11)$$

for the stochastic Gompertz process (Gutiérrez et al. 2008), and for the stochastic Ornstein-Uhlenbeck process, the conditional mean and variance functions $m(x^i|\cdot)$, $w(x^i|\cdot)$ are (Picchini et al. 2011)

$$m_O(x^i | y_0^i, \alpha_O, \beta_O) = y_0^i \exp\left(-\frac{x^i}{\beta_O}\right) + \alpha_O\beta_O\left(1 - \exp\left(-\frac{x^i}{\beta_O}\right)\right) \quad (12)$$

$$w_O(x^i|\beta_O,\sigma_O)=\frac{\sigma_O^2\beta_O}{2}\left(1-e^{-\frac{2x^i}{\beta_O}}\right) \tag{13}$$

In this paper, a segmented stochastic taper process was used that consists of three stochastic differential equations defined by Eq. (2)-(3). This process conforms to the paradigm of stem taper curve that marks three different stem sections along the bole (two points of inflection); the lower section corresponds to a neiloid shape, the middle section corresponds to a parabolic shape, and the upper section corresponds to a conic shape. Max and Burkhart (1976) proposed a segmented polynomial model that uses two joining points to link the three different stem sections. Following from this, the non-continuous at the joining point 0.75 stem taper stochastic differential equations Model 1 is defined in the lower section by the Gompertz form Eq. (2), the middle section by the Ornstein-Uhlenbeck form Eq. (3), and the upper section by the Ornstein-Uhlenbeck form Eq. (3)

$$dY^i(x^i)=\begin{cases}[\alpha_G Y^i(x^i)-\beta_G Y^i(x^i)\ln(Y^i(x^i))]dx^i+\sigma_G Y^i(x^i)dW_G^i(x), \\ P(Y^i(0)=y_0^i)=1, x^i\le 0.15, \\ \left(\alpha_{O1}-\frac{Y^i(x^i)}{\beta_{O1}}\right)dx^i+\sigma_{O1}dW_{O1}^i(x^i), \\ P(Y^i(0.15)=m_G(0.15|y_0^i,\alpha_G,\beta_G,\sigma_G)=1, 0.15<x^i\le 0.75, \\ \left(\alpha_{O2}-\frac{Y^i(x^i)}{\beta_{O2}}\right)dx^i+\sigma_{O2}dW_{O2}^i(x^i), x^i>0.75, P(Y^i(1)=0)=1,\end{cases} \tag{14}$$

where α_G, β_G, σ_G, α_{O1}, β_{O1}, σ_{O1}, α_{O2}, β_{O2}, σ_{O2} are fixed effects parameters (identical for the entire population of trees) and $W_G^i(x^i)$, $W_{O1}^i(x^i)$, $W_{O2}^i(x^i)$, i=1,...,M – are mutually independent standard Brownian motions. The joining points were selected at 0.15 and 0.75 for both tree species, as the fit statistics produced the best values for these points. These values of the joining points are very close to the values utilised by Max and Burkhart (1976).

Using Eq. (14) and assuming that the stem butt was free ($P(Y^i(0)=\gamma)=1$, (γ is an additional fixed effect parameter identical for the entire population of trees) we define stem taper Model 2.

Therefore, we need to estimate α_G, β_G, σ_G, α_{O1}, β_{O1}, σ_{O1}, α_{O2}, β_{O2}, σ_{O2}, γ using all the data in $\underline{y}$, $\underline{x}$, simultaneously, where $\underline{y}=(\underline{y}^1,\underline{y}^2,...,\underline{y}^M)$, $\underline{x}=(\underline{x}^1,\underline{x}^2,...,\underline{x}^M)$, $\underline{y}^i=(y_0^i,y_1^i,...,y_{n_i}^i)$, $\underline{x}^i=(x_0^i,x_1^i,...,x_{n_i}^i)$. The Model 1 proposed in this paper uses the one tree-specific prior relative diameter y_0^i (this known initial condition requires that the diameter over the bark is measured at a stem height of 0 m).

In the latter, we define an approximation of the trajectories of the diameter' and its variance' for Models 1-2 using the following form

$$d_1(h,D,H,d_0)=\begin{cases}D\cdot m_G(\frac{h}{H}|\frac{d_0}{D},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G), & \frac{h}{H}\le 0.15, \\ D\cdot m_O(\frac{h}{H}-0.15|m_G(0.15|\frac{d_0}{D},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G),\hat{\alpha}_{O1},\hat{\beta}_{O1}), & 0.15<\frac{h}{H}\le 0.75, \\ D\cdot m_O(1-\frac{h}{H}|0,\hat{\alpha}_{O2},\hat{\beta}_{O2}), & \frac{h}{H}>0.75,\end{cases} \tag{15}$$

$$w_1(h,D,H,d_0)=\begin{cases}D^2\cdot w_G(\frac{h}{H}|\frac{d_0}{D},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G), & \frac{h}{H}\le 0.15, \\ D^2\cdot(w_G(0.15|\frac{d_0}{D},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G)+w_O(\frac{h}{H}-0.15|\hat{\beta}_{O1},\hat{\sigma}_{O1})), & 0.15<\frac{h}{H}\le 0.75, \\ D^2\cdot w_O(1-\frac{h}{H}|\hat{\beta}_{O2},\hat{\sigma}_{O2}), & \frac{h}{H}\ge 0.75,\end{cases} \tag{16}$$

$$d_2(h,D,H)=\begin{cases}D\cdot m_G(\frac{h}{H}|\hat{\gamma},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G), & \frac{h}{H}\le 0.15, \\ D\cdot m_O(\frac{h}{H}-0.15|m_G(0.15|\hat{\gamma},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G),\hat{\alpha}_{O1},\hat{\beta}_{O1}), & 0.15<\frac{h}{H}\le 0.75, \\ D\cdot m_O(1-\frac{h}{H}|0,\hat{\alpha}_{O2},\hat{\beta}_{O2}), & \frac{h}{H}>0.75,\end{cases} \tag{17}$$

$$w_2(h,D,H)=\begin{cases}D^2\cdot w_G(\frac{h}{H}|\hat{\gamma},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G), & \frac{h}{H}\le 0.15, \\ D^2\cdot(w_G(0.15|\hat{\gamma},\hat{\alpha}_G,\hat{\beta}_G,\hat{\sigma}_G)+w_O(\frac{h}{H}-0.15|\hat{\beta}_{O1},\hat{\sigma}_{O1})), & 0.15<\frac{h}{H}\le 0.75, \\ D^2\cdot w_O(1-\frac{h}{H}|\hat{\beta}_{O2},\hat{\sigma}_{O2}), & \frac{h}{H}\ge 0.75,\end{cases} \tag{18}$$

where the parameter estimators $\hat{\alpha}_G$, $\hat{\beta}_G$, $\hat{\sigma}_G$, $\hat{\alpha}_{O1}$, $\hat{\beta}_{O1}$, $\hat{\sigma}_{O1}$, $\hat{\alpha}_{O2}$, $\hat{\beta}_{O2}$, $\hat{\sigma}_{O2}$, $\hat{\gamma}$ are obtained using a maximum likelihood procedure.

Maximum Likelihood Estimators and Data Fitting

In this paper, we apply the theory of a one-stage maximum likelihood estimator for the stochastic differential equations stem taper Models 1-2. To perform the maximum likelihood estimators for all the trees (i=1,...,M), an extra point $x_l^i=0.15$, $y_l^i=m_G(0.15|y_0^i,\alpha_G,\beta_G,\sigma_G)$, for k=1, or $y_l^i=m_G(0.15|\gamma,\alpha_G,\beta_G,\sigma_G)$, for k=2 is utilised. As both models have closed form transition probability density functions (4), (7), the log-likelihood function for the stem taper Models 1-2 is given as

$$L_k(\theta^k)=\sum_{i=1}^{M}\Big(\sum_{j=1}^{j=l(x_l^i=0.15)}\ln(p_G(y_j^i,x_j^i|y_{j-1}^i,x_{j-1}^i,\alpha_G,\beta_G,\sigma_G)) \tag{19}$$
$$+\sum_{\substack{j=l+1\\ x_j^i<0.75}}\ln(p_O(y_j^i,x_j^i|y_{j-1}^i,x_{j-1}^i,\alpha_{O1},\beta_{O1},\sigma_{O1}))+\sum_{x_j^i\ge 0.75}\ln(p_O(y_j^i,1-x_j^i|y_{j+1}^i,x_{j+1}^i,\alpha_{O2},\beta_{O2},\sigma_{O2}))$$

where: $\theta^1 = \alpha_G, \beta_G, \sigma_G, \alpha_{O1}, \beta_{O1}, \sigma_{O1}, \alpha_{O2}, \beta_{O2}, \sigma_{O2}$,
$\theta^2 = \alpha_G, \beta_G, \sigma_G, \alpha_{O1}, \beta_{O1}, \sigma_{O1}, \alpha_{O2}, \beta_{O2}, \sigma_{O2}, \gamma$.

An essential feature of a stem taper model is its ability to reproduce not only the diameters over the bark at any given height but also the merchantable volume. Thus, to assess the performance of our developed stem taper Models 1-2, we present in this section two alternative regression stem taper models and two alternative regression stem volume models. We shall also fit the parameters of the alternative models to estimation data set, with the aim of comparing our developed model with both the regression stem taper and volume models. The alternative stem taper and volume models were fitted to the estimation data using a least-squares technique. All calculations were implemented in the symbolic computational language MAPLE.

Two alternative models (described below) were used to predict the stem volume: the three-parameter Schumacher-Hall (1933) model (Eq. (20)) and the six-parameter q-exponential model (Eq. (21)) developed by Rupšys and Petrauskas (2010c)

$$V = \beta_1 D^{\beta_2} H^{\beta_3}, \tag{21}$$

$$V = \beta_1 H^{\beta_2}\left[\beta_3 - \frac{\beta_4}{\beta_5}(1-\exp((1-\beta_6)\beta_5 D))\right]_+^{\frac{1}{1-\beta_6}}, \tag{22}$$

where $[a]_+ = \begin{cases} a, & if\ a \ge 0, \\ 0, & if\ a < 0 \end{cases}$, β_1-β_6 are parameters estimated from the data.

Two alternative models (described below) were used to predict the stem taper: the variable-exponent single continuous function nine-parameter model (22) developed by Kozak (2004) and the q-exponential segmented eight-parameter model (23) (Rupšys and Petrauskas 2010c)

$$d = \beta_1 D^{\beta_2} H^{\beta_3} X^{\beta_4 z^4 + \beta_5 \exp(-D/H) + \beta_6 X^{0.1} + \beta_7 D^{-1} + \beta_8 H^Q + \beta_9 X}, \tag{22}$$

where $X = \frac{1-(h/H)^{1/3}}{1-(p)^{1/3}}$, $z = \frac{h}{H}$, $Q = 1 - z^{1/3}$, $p = 1.3/H$, β_1-β_9 are parameters estimated from the data;

$$d = \beta_1 D^{\beta_2} \begin{cases} \beta_3 (z-1) + \beta_4 (z^2 - 1), & if\ z \ge \alpha_1 \\ \left[\beta_5 - \frac{\beta_6}{\beta_7}(1-\exp((1-\beta_8)\beta_7 z))\right]_+^{\frac{1}{1-\beta_8}}, \end{cases} \tag{23}$$

where β_1-β_8 are parameters estimated from the data.

The performance statistics of the stem taper equations for the diameter and the volume included four statistical indices: mean absolute prediction bias (MAB), precision (P), the least squares-based Akaike' (1974) information criterion (AIC), and a coefficient of determination (R^2)

$$MAB = \sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|, \tag{24}$$

$$P = \left(\frac{1}{n-1} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 + \left(\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right) \right)^2 \right)^{1/2}, \tag{25}$$

$$AIC = n \ln(MSE) + 2p, \quad MSE = \frac{1}{n-p} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2, \tag{26}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{27}$$

where n is the total number of observations used to fit the stem volume and taper models, is the number of model parameters, and y_i, $\hat{y}_i$, and $\bar{y}$ are the measured, estimated and average values of the dependent variable (stem volume, diameter over the bark), respectively. The AIC can generally be used for the identification of an optimum model in a class of competing models (Akaike 1974). The first term on the right hand side of the AIC (Eq. (26)) is a measure of the lack-of-fit of the chosen model, while the second term measures the increased unreliability of the chosen model due to the increased number of model parameters.

To assess the standard errors of the maximum likelihood estimators for stochastic differential equations stem taper Models 1-2, a study of the Fisher (1922) information matrix was performed. The asymptotic variance of the maximum likelihood estimator is given by the inverse of the Fisher' information matrix, which is the lowest possible achievable variance among the competing estimators. By defining $p_k(\theta^k) \equiv \ln(L_k(\theta^k))$, where k=1,2, $L_k(\theta^k)$ is defined by Eq. (19), the vector

$p_k(\theta^k)' \equiv \frac{\partial p_k(\theta^k)}{\partial \theta^k}$, and the matrix $p_k(\theta^k)'' \equiv \left[\frac{\partial^2 p_k(\theta^k)}{\partial \theta_i^k \partial \theta_j^k} \right]^T$,

we get that $n^{1/2}\left(\hat{\theta}_n^k - \theta^k \right) \xrightarrow{d} N(0, [i(\theta^k)]^{-1})$, where the Fisher' information matrix is

$$i(\theta^k) = E(p'(\theta^k) p'(\theta^k)^T) = -Ep''(\theta^k)). \tag{28}$$

The standard errors of the maximum likelihood estimators are defined by the diagonal elements of the matrix $[i(\theta^k)]^{-1}$, k=1,2.

Results

Using the estimation data set, the parameters of stochastic differential equations stem taper Models

1-2 were estimated by the maximum likelihood procedure and the parameters of the regression stem volume and taper models (20)-(23) were estimated by the least squares estimation technique. For the q-exponential segmented taper model defined by equation (23), a joint point was calculated at 0.52 for Scots pine trees and 0.46 for Norway spruce trees, as the fit statistics produced the optimum values at these points. Estimation results are presented in Table 2. All parameters of the Models 1-2 are highly significant (p<0.001).

The dashed line in Figure 2 that was generated by the Nadaraya-Watson (1964) kernel regression indicates the bias between the observed and predicted volumes. The stem volume predictions calculated by taper Model 1 also exhibited some biases when the predicted volume was more than 2.0 m^3 (for both tree species) but these biases are smaller than those exhibited by the other volume models. Graphical diagnostics of the residuals for the stem volume predictions indicated that the residuals calculated using the stochastic differential equations stem taper Model 1 had more homogeneous variance than the other models.

Table 2. Estimated parameters (standard errors in parentheses) of all models applied to the stem analysis data sets*

Eq.	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
					Scots pine					
M. 1	-2.2877 (0.0504)	20.0825 (0.7822)	0.4059 (0.0096)	-0.3914 (0.0909)	2.1953 (0.6134)	0.1815 (0.0033)	2.5249 (0.0464)	.3529 (0.0173)	0.1919 (0.0047)	-
M.2	-2.2753 (0.0795)	20.5382 (1.0411)	0.4081 (0.0100)	-0.3755 (0.9920)	2.0418 (0.5375)	0.1837 (0.0033)	2.5249 (0.0464)	0.3529 (0.0173)	0.1919 (0.0047)	1.2031 (0.0155
(20)	$5.8*10^5$ ($4.2*10^6$)	1.8804 (0.0276)	0.9723 (0.0450)	-	-	-	-	-	-	-
(21)	0.0044 ($3.5*10^4$)	0.9762 (0.0328)	-1.2917 (0.9863)	0.2342 (0.0234)	0.0238 (0.0068)	0.1254 (0.1856)	-	-	-	-
(22)	0.9279 (0.0204)	0.9403 (0.0057)	0.0846 (0.1006)	0.4081 (0.0103)	0.0291 (0.0357)	0.3887 (0.0089)	-0.7610 (0.2283)	-0.0024 ($9.4*10^4$)	0.0536 (0.0150)	-
(23)	1.4205 (0.0154)	0.9290 (0.0031)	1.3403 (0.0744)	-1.6591 (0.0453)	0.3485 (0.0271)	-4.1716 (0.1518)	-0.7398 (0.0263)	13.1119 (0.8011)	-	-
					Norway spruce					
M. 1	-1.3307 (0.0224)	24.9216 (0.6785)	0.2743 (0.0051)	-0.7862 (0.0125)	25.0479 (6.8668)	0.1541 (0.0020)	2.4675 (0.0358)	0.3682 (0.0146)	0.1994 (0.0035)	-
M. 2	-1.4196 (0.0556)	28.9738 (1.3727)	0.2855 (0.0066)	-0.7963 (0.0111)	37.8298 (10.992)	0.1550 (0.0020)	2.4675 (0.0358)	0.3682 (0.0146)	0.1914 (0.0035)	1.2445 (0.0176)
(20)	$3.0*10^5$ ($2.9*10^6$)	1.7267 (0.0245)	1.3678 (0.0432)	-	-	-	-	-	-	-
(21)	0.0022 (0.0004)	1.3797 (0.0299)	0.5622 (0.381)	0.1401 (0.0107)	-0.0475 (0.0841)	0.7016 (0.2374)	-	-	-	-
(22)	0.9206 (0.0184)	0.9343 (0.0060)	0.0983 (0.099)	0.4560 (0.0093)	-0.4205 (0.0333)	0.4500 (0.0091)	1.5364 (0.1913)	0.0223 (0.0009)	-0.2021 (0.0141)	-
(23)	1.1591 (0.0119)	0.9709 (0.0031)	0.5586 (0.0518)	-1.2562 (0.0325)	0.0006 (0.0006)	-1.2883 (0.0908)	-0.6867 (0.0182)	30.7587 (3.9916)	-	-

* $\beta_1 = \alpha_G$, $\beta_2 = \beta_G$, $\beta_3 = \sigma_G$, $\beta_4 = \alpha_{O1}$, $\beta_5 = \beta_{O1}$, $\beta_6 = \sigma_{O1}$, $\beta_7 = \alpha_{O2}$, $\beta_8 = \beta_{O2}$, $\beta_9 = \sigma_{O2}$, $\beta_{10} = \gamma$

To test the reliability of all the tested stem taper models, the observed and predicted volume values for the sampled trees were calculated by Eq. (1). Table 3 lists the fit statistics for the taper and volume models.

Another way to evaluate and compare the stem taper and volume models is to examine the graphics of the residuals at different predicted diameters and volumes and to plot the Nadaraya-Watson (1964) kernel regression. The residuals are the differences between the measured and predicted diameters over the bark. Positive residuals indicate underestimation, and negative residuals indicate overestimation. Residual plots of all the fitted taper models are presented in Figure 1 for Scots pine and Norway spruce trees. The residuals suggested that Model 1 behaves similarly but produced better fit statistics than the other tested volume models. The residuals of Model 1 with fixed tree butt and top are clustered at 0 for the butt and top sections of the stems. Distributions of the residuals are similar for both species.

Taper profiles for three randomly selected Scots pine trees (diameters over the bark at breast heights of 6.3 cm, 17.0 cm, 40.7 cm, and total tree heights of 6.8 m, 21.1 m, 30.3 m) and for Norway spruce trees (diameters over the bark at breast heights of 9.9 cm, 28.0 cm, 37.0 cm, and total tree heights of 14.0 m, 20.1 m, 26.0 m) were constructed using stochastic differential equations stem taper Model 1 and are plotted in Figure 3. Figure 3 includes the stem taper curves and the standard deviation curves. It is clear that all of the taper profiles followed the stem data very closely. Graphical examination of these taper profiles leads to the conclusion that stem taper Model 1 with fixed stem bottom describes taper profile quite well.

Discussion and conclusions

For volume calculation, Lithuanian foresters use the stem form-factor model created by Kuliešis (1972),

Table 3. Fit statistics for all the tested stem taper and volume models*

Equation	Estimation					Validation				
	MAB	P	AIC	$\bar{R}^2$	Count.	MAB	P	AIC	$\bar{R}^2$	Count
Scots pine										
Taper models										
(22)	0.9703	1.3889	34036	0.9841	3821	0.9162	1.3022	34932	0.9866	3962
(23)	0.9371	1.3807	33996	0.9843	3821	0.8833	1.2976	34895	0.9867	3962
M. 1	**0.9036**	**1.3372**	**33754**	**0.9853**	3821	**0.8578**	**1.2880**	**34845**	**0.9869**	3962
M. 2	1.0612	1.5307	34778	0.9807	3821	0.9896	1.4418	35729	0.9836	3962
Volume models										
(20)	**0.0400**	0.0674	98	0.9862	300	**0.0393**	0.0605	27	0.9886	298
(21)	0.0403	0.0693	99	0.9862	300	0.0411	0.0628	55	0.9876	298
(22)	0.0401	0.0675	109	0.9859	300	0.0395	0.0611	42	0.9882	298
(23)	**0.0400**	0.0681	106	0.9860	300	0.0394	0.0611	41	0.9883	298
M. 1	**0.0400**	**0.0639**	**71**	**0.9876**	300	0.0400	**0.0597**	**18**	**0.9892**	298
M. 2	0.0443	0.0732	153	0.9837	300	0.0454	0.0664	90	0.9862	298
Norway spruce										
Taper models										
(22)	0.9464	1.4377	60083	0.9815	6336	1.0062	1.5722	64692	0.9770	6663
(23)	0.9575	1.4927	60549	0.9801	6336	1.0416	1.6614	65391	0.9745	6663
M. 1	**0.9099**	**1.3360**	**59153**	**0.9840**	6336	**0.9501**	**1.4135**	**63624**	**0.9815**	6663
M. 2	1.0758	1.6924	62143	0.9744	6336	1.1428	1.8821	66804	0.9685	6663
Volume models										
(20)	0.0394	0.0625	**334**	**0.9834**	497	0.0446	0.0729	539	0.9755	524
(21)	0.0393	**0.0623**	338	**0.9834**	497	0.0444	0.0733	553	0.9750	524
(22)	**0.0381**	0.0628	348	0.9831	497	**0.0428**	**0.0711**	**520**	**0.9766**	524
(23)	0.0404	0.0673	411	0.9808	497	0.0468	0.0789	627	0.9713	524
M. 1	0.0423	0.0638	360	0.9827	497	0.0485	0.0779	596	0.9730	524
M. 2	0.0498	0.0706	460	0.9789	497	0.0506	0.0830	668	0.9691	524

* The best values of fit statistics for all the taper and volume models are in bold

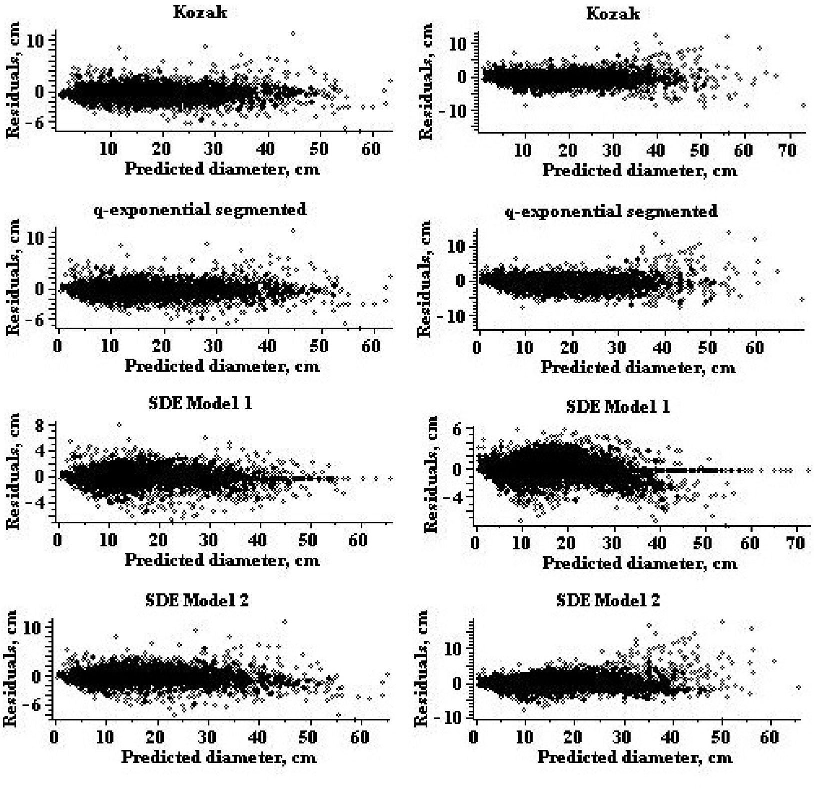

Figure 1. Residuals for the taper models: left – Scots pine trees; right – Norway spruce trees

which was the first attempt to identify the effects of the different factors influencing the stem taper function using multi-factorial ANOVA. It was highlighted that trees of one species have no permanent stem taper curve and that stem form depends on the tree growing conditions, the total height of the tree, DBH, and the length of the crown. As was shown by Lejeune et al. (2009), the low reliability of the stem taper models in the upper bole section can be explained by the lack of diameter measurements in the upper bole sections.

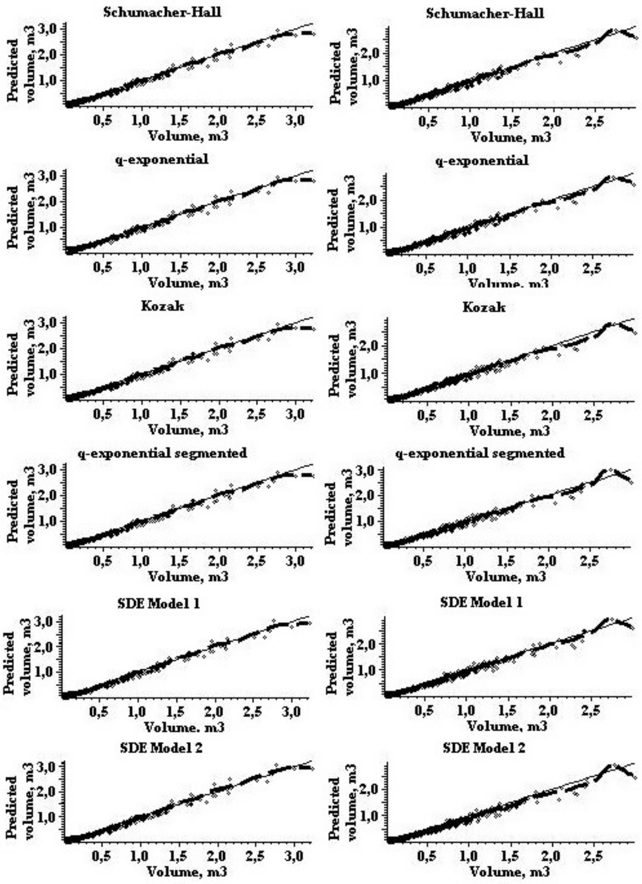

Figure 2. Nonparametric kernel regression curves for the volume models: left – Scots pine trees; right – Norway spruce trees

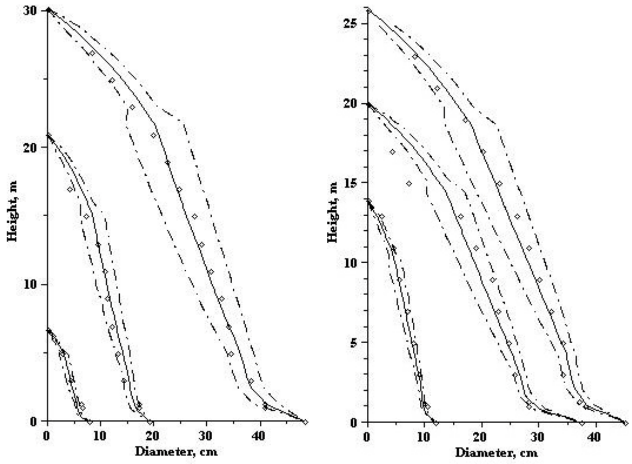

Figure 3. Stem tapers for three randomly selected trees generated using the stochastic differential equations Model 1: left – Scots pine trees; right – Norway spruce trees.

The Schumacher-Hall (Eq.(20)) and q-exponential (Eq. (21)) volume models had very similar fit statistics for both the estimation and validation data sets and for both species. Both of the base taper regression models (22)-(23) result in very similar fit statistics for the both estimation and validation data sets and for both species. The q-exponential segmented model (23) performed the best for Scots pine trees, while the Kozak model (23) performed slightly better for Norway spruce trees.

The three stem taper models, Eq. (22), Eq. (23), and Model 1, had very similar fit statistics. The best values of the fit statistics were produced by stem taper Model 1 with fixed tree butt for the both estimation and validation data sets and for both species. The volume predictions by stem taper Model 1 also produced the best fit statistics for Scots pine trees.

The new taper models were developed using stochastic differential equations. Comparison of the predicted stem taper and stem volume values calculated using stochastic differential equations Models 1-2 with the values obtained using the existing regression models revealed a comparable predictive power of stem taper Model 1 with fixed stem bottom.

The developed stem taper Model 1 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 variance functions developed here can be applied generate weights in every linear and nonlinear least squares regression stem taper model 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 tree taper.

References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716-723.

Alegria, C. and Tome, M. 2011. A set of models for individual tree merchantable volume prediction for Pinus pinaster Aiton in central inland of Portugal. *European Journal of Forest Research* 130:871–879. doi: 10.1007/s10342-011-0479-3.

Demaerschalk, J. 1972. Converting volume equations to compatible taper equations. *Forest Science* 18:241–245.

Fisher, R.A. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A-Mathematical, Physical and Engineering Sciences* 222: 309-368.

Gutiérrez, R., Gutiérrez-Sánchez, R. and Nafidi, A. 2008. Trend analysis using nonhomogeneous stochastic diffusion processes. Emission of CO2; Kyoto protocol in Spain. *Stochastic Environmental Research Risk Assessment* 22: 57–66.

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

Kozak, A., Munro, D.D. and Smith, J.G. 1969. Taper functions and their application in forest inventory. *Forest Chronicle* 45: 278-283.

Kozak, A. 2004. My last words on taper equations. *Forest Chronicle* 80: 507–515.

Kuliešis, A., Kenstavičius, J., Grigaliūnas, J. and Garbinčius, A. 1983. Lietuvos TSR pagrindinių medžių rūšių stiebų formrodžių (f) priklausomybės nuo skersmens (D) ir aukščio (H) matematinių modelių parametrai [Stem form (f) dependence on diameter (D) and height (H) parameters of mathematical models of main tree species stems of Lithuanian SSR]. Miško taksuotojo žinynas. Vilnius p. 34-35.

Lejeune, G., Ung, C.-H., Fortin, M., Guo, X. J., Lambert, M.-C. and Ruel, J.-C. 2009. A simple stem taper model with mixed effects for boreal black spruce. *European Journal of Forest Research* 128: 505–513.

Max, T.A. and Burkhart, H.E. 1976. Segmented polynomial regression applied to taper equations. *Forest Science*, 22: 283–289.

Nadaraya, E.A. 1964. On estimating regression. *Theory of Probability and their Applications* 9: 141–142.

Özçelik, R., Brooks, J.R. and Jiang, L. 2011. Modeling stem profile of Lebanon cedar, Brutian pine, and Cilicica fir in Southern Turkey using nonlinear mixed-effects models. *European Journal of Forest Research* 130:613–621.

Parresol, B.R. 1999. Assessing tree and stand biomass: a review with examples and, critical comparisons. *Forest Science* 45(4): 573-593.

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(1): 118-127.

Picchini, U., Ditlevsen, S. and De Gaetano, A. 2011. Practical estimation of high dimensional Stochastic differential mixed-effects models. *Computational Statistics & Data Analysis* 55(3): 1426-1444.

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

Rupšys, P., Petrauskas, E., Mažeika, J. and Deltuvas, R. 2007. The Gompertz type stochastic growth law and a tree diameter distribution. *Baltic Forestry* 13(2): 197-206.

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

Rupšys, P. and Petrauskas, E. 2010b. Quantyfying tree diameter distributions with one-dimensional diffusion processes. *Journal of Biological Systems* 18: 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: 2369-2378.

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.

Schumacher, F.X. and Hall F.D.S. 1933. Logarithmic expression of timber tree volume. *Journal of Agricultural Research* 47: 719-734.

Serinaldi, F., Grimaldi, S., Abdolhosseini, M., Corona, P. and Cimini, D. 2012. Testing copula regression against benchmark models for point and interval estimation of tree wood volume in beech stands. *European Journal of Forest Research* 131: 1313-1326.

Suzuki, T. 1971. Forest transition as a stochastic process. *Mit Forstl Bundesversuchsanstalt Wien* 91: 69-86.

Tabacchi, G., Di Cosmo, L. and Gasparini, P. 2011. Aboveground tree volume and phytomass prediction equations for forest species in Italy. *European Journal of Forest Research* 130: 911-934.

Tanaka, K. 1986. A stochastic model of diameter growth in an even-aged pure forest stand. *Journal of the Japanese Forest Society* 68: 226-236.

Trincado, J. and Burkhart, H.E. 2006. A generalized approach for modeling and localizing profile curves. *Forest Science* 52: 670-682.

Westfall, J.A. and Scott, C.T. 2010. Taper models for commercial tree species in the Northeastern United States. *Forest Science* 56: 515-528.

Yang, Y., Huang, S. and Meng, S.X. 2009. Development of a tree-specific stem profile model for White spruce: a nonlinear mixed model approach with a generalized covariance structure. *Forestry* 82: 541-555**.**

Received 19 April 2012
Accepted 07 June 2013

ИСПОЛЬЗОВАНИЕ СТОХАСТИЧЕСКИХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ДЛЯ ОПИСАНИЯ ОБРАЗУЮЩЕЙ СТВОЛА И ОБЪЕМА

Э. Пятраускас, Э. Барткевичиус, П. Рупшис, Р. Мемгаудас

Резюме

Подход, объединяющий информацию, генерированную различными стохастическими дифференциальными уравнениями, разработан для повышения точности прогнозирования образующей ствола и объема. Стохастические модели дифференциальных уравнений образующей ствола и объема были разработаны на данных сосны и ели обыкновенных, собранных по всей территории Литвы. Новые модели выведены из стохастических дифференциальных уравнений Гомпертца и Орнштейна-Уленбека сравнены с классической моделью образующей ствола Козака и q-экспоненциальной сегментированной образующей ствола а также с моделью объема Шумахер'а-Хал'а и q-экспоненциальной моделью объема, основанной на аллометрической и геометрической концепциях. Сравнение разработанной образующей ствола и объема с моделями, основанными на регрессионной анализе, показали лучную мощность прогнозирования стохастических дифференциальных уравнений.

Ключевые слова: стохастические дифференциальные уравнения, модели образующей ствола, модели объема.