



**HAL**  
open science

## Applying wavelet-based functional approach in modelling tree taper

Subedi, Sharma

► **To cite this version:**

Subedi, Sharma. Applying wavelet-based functional approach in modelling tree taper. *Annals of Forest Science*, 2011, 68 (5), pp.1039-1048. 10.1007/s13595-011-0113-z . hal-00930682

**HAL Id: hal-00930682**

**<https://hal.science/hal-00930682>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Applying wavelet-based functional approach in modelling tree taper

Nirmal Subedi · Mahadev Sharma

Received: 15 September 2010 / Accepted: 7 February 2011 / Published online: 2 August 2011  
© INRA and Springer Science+Business Media B.V. 2011

## Abstract

• **Introduction** The wavelet-based functional approach was evaluated for modelling tree taper of jack pine (*Pinus banksiana* Lamb.) trees grown in the Canadian boreal forest region.

• **Objectives** Wavelet-based functional fixed and mixed-effects models were developed to predict tree taper, and these models were evaluated for their predictive accuracy using calibration and evaluation data sets.

• **Results and discussion** Diameters predicted using both fixed and mixed-effects taper models were unbiased for calibration data set as the 95% credible limits included 0 at all locations along the boles. The diameters predicted by these models for validation data set, however, were unbiased only at four out of 11 locations as the 95% credible limits of mean bias using fixed effects model did not include 0 at other locations.

• **Conclusion** The study concludes that the wavelet-based taper models are able to describe the taper of the trees used in fitting the model but are unable to capture the mean taper function of the trees not used in fitting the model.

**Keywords** *Pinus banksiana* · Jack pine · Stem profile model · Functional taper model · Bayesian taper model

## 1 Introduction

Forestry modellers are interested in estimating tree bole diameter or cross-sectional area of standing tree. Models that statistically predict the diameter or cross-sectional area at any height across the tree stem are called taper models, and these models are needed to accurately predict the amount of woody material contained in the tree. Accurate estimation of a tree volume is one of the key requirements for effective forest management planning. Taper models are also useful in determining stem quality and calculating merchantable volumes, log breakout and product recovery.

A number of pair-wise diameter and height measurements (stem profile data) of an appropriate number of sample trees of a particular species and physiographic growing region are required to fit a taper model. In most of the cases, these measurements are obtained by destructive sampling. The sequence of pair-wise diameter and height measurements on a tree are spatially correlated. However, the measurements among trees can be assumed independent as trees are usually sampled from a large number of plots representing the physiographic region of the species (Gregoire et al. 2000).

Many approaches have been developed for producing taper equations. Early taper equations were based on simple functions, e.g., Kozak et al. (1969), Ormerod (1973), Amidon (1984). Max and Burkhart (1976) introduced the segmented approach, in which a tree is divided into three distinct sections and a separate polynomial equation is used to describe each section. Liu (1980) divided the stem into a number of sections and used a spline smoothing approach to fit a model for each section. Later, Kozak (1988) and Newnham (1992) introduced the variable exponent model in which the exponent changes from the base to the tip of the tree to account for differing shapes along the stem

---

**Handling Editor:** Barry Gardiner

N. Subedi (✉) · M. Sharma  
Ministry of Natural Resources, Ontario Forest Research Institute,  
1235 Queen St East,  
Sault Ste. Marie, ON P6A 2E5, Canada  
e-mail: nirmal.subedi@ontario.ca

M. Sharma  
e-mail: mahadev.sharma@ontario.ca

profile. In addition, Sharma and Oderwald (2001) applied dimensional analysis and Bi (2000) used the Fourier transformation approach to estimate the exponent of taper equation.

Linear and nonlinear least squares regression techniques have been used to fit these models most of the time. However, stem profile data are longitudinal, consisting of multiple measurements from a single stem, which violates the fundamental least squares (Gauss Markov) assumptions of independence and equal distribution of errors with zero mean and constant variance (Gregoire et al. 1995; Dieguez-Aranda et al. 2006). Forest modellers have increasingly recognized the issues of autocorrelation and heteroscedasticity in stem profile data (Leites and Robinson 2004; Trincado and Burkhart 2006; Li and Weiskittel 2010) and are seeking more robust methods to model stem taper.

Mixed-effects models are considered very powerful and effective tools to counteract problems related to longitudinal data. These models include both fixed and random effects and can be used to estimate the covariance structure from longitudinal data (e.g., Schabenberger and Pierce 2002). Fixed effects relate to parameters averaged across a population, and random effects are associated with a specific sampling unit (Demidenko 2004). These models also distinguish between deterministic and stochastic model errors.

Most recent taper modelling studies rely on mixed-effects models. For example, taper models developed using a mixed model approach include those by Valentine and Gregoire (2001), Zhang et al. (2002), Garber and Maguire (2003), Leites and Robinson (2004), Trincado and Burkhart (2006), Koskela et al. (2006) and Sharma and Parton (2009). Mixed models provide the flexibility to estimate both population-averaged and subject-specific (tree) parameters and thus to more precisely estimate tree taper. However, Koskela et al. (2006) stated that the linear mixed model may not always be flexible enough to precisely describe a tree stem as it is strictly tied to the functional model forms. Nonlinear mixed-effects models, on the other hand, can be difficult to fit if the number of mixed-effects parameters is greater than 3.

Tree diameter inside bark (dib) or relative diameter inside bark (dib/dbh, where dbh is diameter outside bark at breast height (1.3 m)) that varies from the bottom to the top, can be regarded as a function and be treated as a unit of analysis (Ramsay and Silverman 1997; Besse et al. 2005; Ferraty and Vieu 2006). Tree taper measurements contain discrete pair-wise height and relative diameter measurements at certain locations along the length of the tree. Such functional taper data can be analysed using either nonparametric or wavelet-based approaches. The nonparametric approach is less restrictive, as it does not require equally spaced (either in time or space) measurements. However,

this approach has limitations when making predictions for a new subject for example a tree as there is neither a way to display nonparametric relationships nor is there structural information about it (Ferraty and Vieu 2006, pp 49–50). In contrast, the wavelet-based functional approach being a parametric approach, predictions about a new subject can be easily made. However, this approach requires equally spaced functional measurements. Thus, if the height and relative diameter are measured at equal intervals of relative height or time across many trees, these discrete measurements can be analysed using wavelet-based function approach.

The measurement locations or intervals can be regarded as a grid, and tree diameter measurements at various height locations for each tree can be regarded as a function. Figure 1 displays the relationship between jack pine trees relative height and its corresponding relative diameter. There are 11 measurements for each tree that are represented by circles. These circles are connected to obtain a curve. In addition, the functional approach provides specific model parameters for each equally spaced measurement grids (measurement locations). Thus, the functional data refers to the observations collected across curves or functions, where the curve is used as a basic unit of data analysis (Ramsay and Silverman 1997).

The functional mixed model is an emerging approach to model functional data with linear mixed effects. Recently, Morris and Carroll (2006) have developed a wavelet-based functional mixed model (WFMM) to analyse functional data. In this approach, the observed response functions (e.g. taper measurements) are first projected into wavelet space, and then the functional mixed model is fitted in the wavelet space. They are then re-projected back into data space once the model parameters are estimated across grid locations.

The WFMM has the advantage of both a functional approach and wavelet transformation. The major benefit of a functional approach to modelling is that it does not require specification of parametric form for curves. For example, a simple linear form may represent tree taper because separate model parameters are estimated to make prediction at each measurement (grid location). Similarly, the wavelet transformation provides compact support to efficient transformation of local features to wavelet space. Another dimension of WFMM is the parameter estimation in wavelet space. Wavelets provide orthogonal transformation and a unique opportunity to remove noise from the signal without losing information. The wavelet allows both a parsimonious and a flexible way to represent variance covariance matrix because of its whitening property (much smaller correlation among wavelet coefficients than original data). Hence, WFMM has attractive properties to model functional data.

The objectives of this study were to demonstrate the application of (1) a functional mixed model to tree taper modelling for jack pine (*Pinus banksiana* Lamb.) trees using the WFMM approach and (2) investigate its usefulness, advantages and limitations in taper modelling.

## 2 Wavelet-based functional mixed model

In this section, we briefly summarize the wavelet-based functional mixed model method introduced by Morris and Carroll (2006) and describe how to apply it to tree taper modelling. Functional mixed models are a recent development (Guo 2002; Antoniadis and Sapatinas 2007; Morris et al. 2003) and can be used to model both fixed and random functions in the same functional space. Functional mixed models can be thought of an extension of Laird and Ware’s (1982) random effects model for functional data, where the forms of the fixed and random effects functions are left completely unspecified (Morris and Carroll 2006). For example, let a sample of  $n$  curves  $Y_i(\mathbf{t})$ ,  $i=1, \dots, n$ , be observed on compact set  $T \in \mathfrak{R}^1$ , which is assumed without loss of generality to be  $[0, 1]$ . In the case of tree taper modelling, these can be assumed as the relative diameter measurements of a tree, which are equally spaced along the length of tree ( $\mathbf{t}$ ). Following Morris and Carroll (2006) and Morris et al. (2008), a functional mixed model for these functions can be represented as:

$$Y_i(\mathbf{t}) = \sum_{j=1}^p X_{ij}B_j(\mathbf{t}) + \sum_{k=1}^m Z_{ik}U_k(\mathbf{t}) + E_i(\mathbf{t}) \tag{1}$$

where  $X_{ij}$  and  $Z_{ik}$  are the covariate matrix and the design matrix, respectively, for functional (time or space ) fixed effects  $B_j(\mathbf{t})$  and random effects  $U_k(\mathbf{t})$ , respectively, and  $E_i(\mathbf{t})$  are residual error processes. We assume  $U_k(\mathbf{t})$  are independent and identically distributed (i.i.d.) Gaussian processes with mean zero and covariance surface  $Q(t_1, t_2)$ . Similarly,  $E_i(\mathbf{t})$  are i.i.d. Gaussian processes with mean zero and covariance surfaces  $S(t_1, t_2)$ .  $U_k(\mathbf{t})$  and  $E_i(\mathbf{t})$  are assumed independent.

Covariates  $\{X_{ij}, j = 1, 2, \dots, p\}$ , discrete or continuous, can be any factors that can explain the tree taper. Each functional coefficient  $B_j(\mathbf{t})$  describes the effect of the corresponding factor at location  $t$  of the taper. The covariate can include a column of 1’s to represent overall mean function, continuous or discrete variable of interest like relative height, function of dbh, height, stand density, or any interaction of these factors. Similarly, the functional random effects provide a flexible approach for modelling correlation among trees if trees are not independent. For example, sample plot or regional level random effect

functions can be specified when multiple trees are obtained from the same plot.

Fixed and random effects and the residual error process are all functional in functional mixed models. Morris et al. (2006) stated “heuristically [the WFMM] model fitting can be viewed as fitting separate mixed models at each time [or space point across the grid  $\mathbf{t}$ ], but with an extra layer added for regularization, that is, which borrows strength across observations within a function.”

A discrete version of model 1 is required if we have all trees relative diameters are measured on the same equally spaced grid  $\mathbf{t}$  of length  $T$ . Let  $\mathbf{Y}$  be the  $n \times T$  matrix containing the observed samples on grid with each row containing one observed sample on grid  $\mathbf{t}$ . A discrete form of this mixed model can be expressed as:

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{Z}\mathbf{U} + \mathbf{E} \tag{2}$$

Each row of the  $n \times T$  matrix  $\mathbf{Y}$  contains measurements from a tree on grid  $\mathbf{t}$ . The matrix  $\mathbf{X}$  is an  $n \times p$  design matrix of covariates, and  $\mathbf{B}$  is a  $p \times T$  matrix of *fixed effect functions* on grid  $\mathbf{t}$  with  $B_{ij}$  representing the effect of the covariate in column  $i$  of  $\mathbf{X}$  to response variable  $\mathbf{Y}$  at time or space  $t_j$ . The matrix  $\mathbf{U}$  is an  $m \times T$  matrix of random effect functions on the grid  $\mathbf{t}$ , and  $\mathbf{Z}$  is the corresponding  $n \times m$  design matrix.  $\mathbf{E}$  is an  $(n \times T)$  matrix containing the residual error process of the corresponding observed sample. We assume that the rows of  $\mathbf{U}$  are i.i.d. MVN (i.e., multivariate normal)  $(0, \mathbf{Q})$  and rows of  $\mathbf{E}$  are i.i.d. MVN $(0, \mathbf{S})$ , independent of  $\mathbf{U}$ , with  $\mathbf{Q}$  and  $\mathbf{S}$  being  $T \times T$  covariance matrices that are analogues of the covariance surfaces in Eq. 1 on the grid  $\mathbf{t}$  (Morris et al. 2008).

Note the definition of functional data (sampled on a fine grid) will not be too restrictive if the functions are observed across a fine grid (i.e. smaller steps relative to tree height). We can interpolate the observed data to obtain a common grid without changing it substantively (see Morris and Carroll (2006) for details on WFMM modelling assumptions and computation).

### 2.1 Model development

Following Morris and Carroll (2006), the nonparametric wavelet-based mixed model can be developed as follows:

1. Perform a *discrete wavelet transformation* (DWT) on the response vector for each observed curve to obtain corresponding empirical wavelet coefficients.
2. Fit a functional mixed model using Markov chain Monte Carlo (MCMC) methods to these empirical wavelet coefficients, yielding a posterior sample of model parameters for fixed effects and random effects functions, as well as the variance components.

3. Obtain the posterior distribution of the mean fixed and random effects on a fixed grid  $\mathbf{t}$  using *inverse discrete wavelet transformation* (IDWT) and use these posterior samples of model parameters to perform Bayesian estimation, inference and prediction.

Thus, the first step involves the right matrix multiplication of observed functional curves sampled over an equally spaced grid of size  $T$  by an orthogonal DWT matrix. Mathematically, it can be represented as  $\mathbf{D}=\mathbf{Y}\mathbf{W}'$ , where  $\mathbf{W}$  is a  $T\times T$  size orthogonal DWT matrix,  $\mathbf{Y}$  is a  $N\times T$  matrix containing each function (tree taper measurements) stacked as a row, and  $\mathbf{D}$  is  $N\times T$  matrix containing the empirical wavelet coefficients for all observed curves (note the superscript ' stands for *transpose*). Each row of  $\mathbf{D}$  contains empirical wavelet coefficients for the curve and the columns double indexed for scale  $j$  and location  $k$ , with  $j=1, \dots, J$  and  $k=1, \dots, K_j$ .

Wavelet coefficients are defined as the inner product of the observed function  $y(t)$  and the corresponding wavelet basis function. For example, let  $\mathbf{y}=(y_1, \dots, y_T)$  be a sample of a function at  $T$  equally spaced points with  $T$  assumed to be the power of 2. The vector  $\mathbf{y}$  can be transformed to wavelet space as  $\mathbf{d}=\mathbf{y}\mathbf{W}'$  where  $\mathbf{W}'=[V_j', W_1', W_2', \dots, W_j']$ . However, there exists a fast algorithm (DWT) for decomposing  $\mathbf{y}$  into a set of  $T$  wavelet coefficients (Mallat 1989; Percival and Walden 2000) but only in  $O(T)$  operations. The DWT decomposes the set of observations into sets of wavelet and scaling coefficients  $\mathbf{d}=(\mathbf{c}_j, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \dots, \mathbf{d}_j)$  where  $\mathbf{d}_j=\mathbf{y}\mathbf{W}_j'$  are wavelet coefficients at level or scale  $j$  and  $\mathbf{c}_j=\mathbf{y}\mathbf{V}_j'$  are the scaling coefficients (Percival and Walden 2000; Morris et al. 2006). For simplicity, both the wavelet and scaling coefficients are referred to as wavelet coefficients. For details on wavelets, see Daubechies (1992), Chui (1992) and Keinert (2004), and for details on DWT, see Percival and Walden (2000), Vidakovic (1999) and Ogden (1997).

## 2.2 Parameter estimation

WFMM estimates model parameters in a wavelet space using MCMC with vague priors for the variance and independent mixture priors for the elements of  $\mathbf{B}^*$  (\* sign used for parameter in wavelet space). The prior distribution of  $B_{ijk}^*$ , the wavelet coefficient at scale  $j$  and location  $k$  for fixed effect  $i$  are defined as

$$B_{ijk}^* = \gamma_{ijk}^* N(0, \tau_{ijk}) + (1 - \gamma_{ijk}^*) \delta_0, \quad (3)$$

$$\gamma_{ijk}^* = \text{Bernoulli}(\pi_{ij}),$$

where  $\delta_0$  is a point mass at 0 and  $\gamma_{ijk}^*$  is an indicator of

whether wavelet coefficient  $(j, k)$  is important for representing the signal for the fixed effect function  $i$ . Similarly,  $\pi_{ij}$  and  $\tau_{ijk}$  are prior probability and variance, respectively, for representing the fixed effect  $i$ . This prior is frequently used in Bayesian implementation of wavelet regression (Clyde et al. 1998; Abramovich et al. 1998) as it provides adoptive regularization of the fixed effect functions by nonlinearly shrinking them toward 0. In addition, regularization parameters  $\tau_{ijk}$  and  $\pi_{ij}$  determine the relative trade-off between variance and bias in the nonparametric estimation (Morris et al. 2006). These regularization parameters can either be specified or estimated using an *empirical Bayes* method (see Carlin and Louis 2000 for details).

A fully specified Bayesian model for functional data is defined after specifying distribution and priors for model parameters (fixed functional effects, random functional effects and variance–covariance vectors ( $\mathbf{Q}$  and  $\mathbf{S}$ )). The marginalized likelihood functions are used to obtain the posterior sample of model parameters. The MCMC scheme includes three major steps (Morris et al. 2006). First, fixed effect samples are drawn and the fixed effect  $i$  at level  $j$  and location  $k$  is estimated by integrating out the random effects. This will improve the mixing properties of the samples over a Gibbs sampler. Samples of between-curve covariance parameters  $\mathbf{Q}$  and  $\mathbf{S}$  were then obtained using a single random walk Metropolis–Hastings step. Finally, a series of Gibbs samples were taken to update the random effect wavelet coefficients. See Morris and Carroll (2006) for details about the parameter's conditional distributions.

Posterior samples of the model parameters in the wavelet space are projected back to the data space by applying IDWT. These samples are obtained for each fixed effect function  $\mathbf{B}_i(t)$  on the grid  $\mathbf{t}$  by applying IDWT to each posterior sample of the corresponding vector of wavelet coefficients  $\mathbf{B}_i^*=(B_{i11}^*, \dots, B_{iJK_j}^*)$  and random effects functions. Once the posterior samples of model parameters are estimated, Bayesian parameter estimation and inference can be made. However, we are interested in predicting tree taper across the tree height grid  $\mathbf{t}$  using a WFMM model in this study. Tree taper can be predicted using the posterior point estimate of model parameters.

## 3 Wavelet-based functional taper model

### 3.1 Model fitting

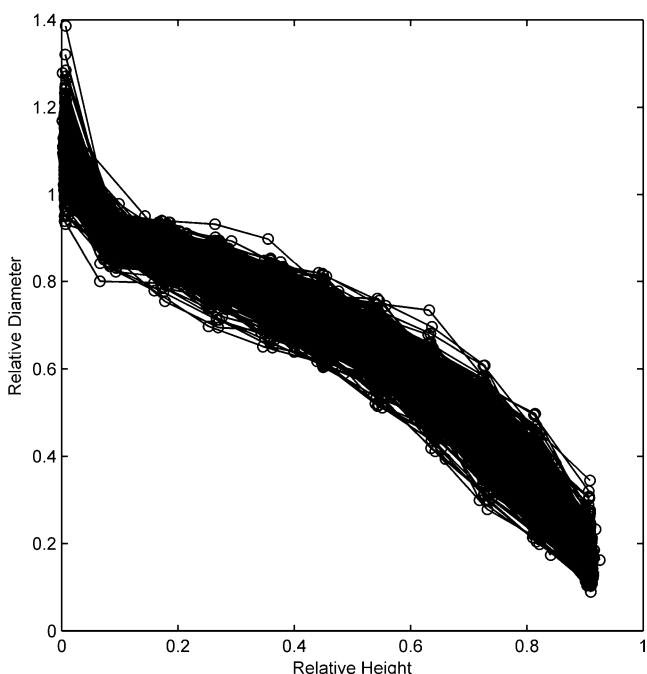
In a project to model stand-density effects on tree taper, Sharma and Parton (2009) destructively sampled 1,135 jack pine (*P. banksiana* Lamb.) trees across 25 even aged monospecific plantations covering the Canadian boreal forest region of Northern Ontario. The measurements made on those trees included tree diameter inside bark (dib) at



0.15, 0.5, 0.9 and 1.3 m (breast height) and at every 10th percentile of tree height above breast height for 80% of sampled trees and at every 5th percentile of tree height above breast height for the rest along with bark thickness at the breast height.

As mentioned earlier, WFMM requires equally spaced functional data. Therefore, we used the measurements at 0.15 and 1.3 m and every 10th percentile above breast height from each tree as these constituted approximately equally spaced measurements. This resulted in 11 measurements per tree or function. Following Percival and Walden (2000, p. 140), the 11 taper (dib/dbh) measurements for each tree were doubled to make 22 measurements (11 from bottom to top and another 11 from top to bottom) to minimize the boundary effects. Hence, the size of  $T$  was 22 for parameter estimation. However, only the first 11 grid point measurements were used for model evaluation.

Of 25 sites, 13 sites were randomly selected and trees from these sites were used for model calibration and the rest were used for model evaluation. For simplicity, trees with multiple missing measurements across the girds were not included in the analysis. This resulted in 6,149 pair-wise dib measurements from 559 trees for the calibration data set and 5,650 pairs from 516 trees for the evaluation data set. The mean dbh of tree was 17.86 cm, ranging from 6.1 to 34.3 cm, and the mean height was 16.13 m, ranging from 9.08 to 23.17 m for calibration data. Relative diameters of these trees plotted against their corresponding relative heights are displayed in Fig. 1. The mean dbh of the



**Fig. 1** An illustration of tree taper for jack pine trees data from Ontario, Canada ( $n=559$ )

evaluation data set was 16.74 cm ranging from 7.5 to 33.8 cm. Similarly, the height ranged from 7.93 to 22.3 m with the mean of 14.75 m for this data set.

We tried a variety of explanatory variables for both the fixed effects and random effects for wavelet-based linear functional taper models. Based on the smallest values of residual sum of squares and sum of absolute residuals, the following two models (Eqs. 4 and 5) were selected. However, we emphasize that developing a predictive taper model for jack pine was not the objective of this study. Rather, we attempted to explore the potential of wavelet application in forestry modelling. A row matrix of these models in the form of Eq. 2 can be represented as:

$$y(\mathbf{t}) = \left[ (h(\mathbf{t})/H) \quad \sqrt{dbh} \quad \sqrt{H} \right] \beta \tag{4}$$

$$y(\mathbf{t}) = \left[ (h(\mathbf{t})/H) \quad \sqrt{dbh} \quad \sqrt{H} \right] \beta + \left[ \sqrt{bt/dbh} \sqrt{bt/dbh} \right] \mathbf{U} \tag{5}$$

where  $y(\mathbf{t})$  is  $1 \times 22$  row matrix of tree relative diameter (dib/dbh),  $h(\mathbf{t})/H$  is  $1 \times 11$  row matrix with the relative height measurements across the tree height (at stump height, dbh and 1/10 of height above breast height) and dbh,  $H$  and  $bt$  are the tree diameter at breast height, total height and bark thickness, respectively.  $\beta$  and  $\mathbf{U}$  are fixed and random effects parameter matrix of size  $13 \times 22$  and  $2 \times 22$ , respectively. Because of replication of the tree relative diameter measurements, there were 22 columns in the parameter matrix. As the first 11 column of parameter matrix were similar to the remaining columns of parameter matrix, the first 11 columns of parameters were used to make inference about models fitting and evaluation (Table 1). As suggested by Morris and Carroll (2006), tree variables  $\sqrt{dbh}$ ,  $\sqrt{H}$  and  $\sqrt{bt/dbh}$  were standardized to have mean 0 and standard deviation 1.

We used Daubechies (1992) wavelet with two vanishing moments at  $J=2$  levels and periodic boundary conditions. The choice of wavelet and boundary condition were made by assessing the deviation of the diagonal eigenvalues of correlation matrix ( $\mathbf{w}^{-1/2} \mathbf{W}^2 \mathbf{w}^{-1/2}$ ,  $\mathbf{W}^2 = \mathbf{W}\mathbf{W}'$  and  $\mathbf{w} = \text{diag}(\mathbf{W}^2)$ ). The wavelet and boundary condition that provided minimum deviation in eigenvalues were preferred to maintain orthogonality, as far as possible, in wavelet transformation either from data to wavelet space or vice versa. We fitted the WFMM model using a computer programme, WFMM version 2.0.3.<sup>1</sup>

While specifying the MCMC simulation, we used 20,000 iterations as burn-in and ran the model for 300,000 iterations, keeping every 10th run. The Metropolis–Hastings acceptance probability for the variance com-

<sup>1</sup> Detailed information about the software is available at [http://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware.aspx?Software\\_Id=70](http://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware.aspx?Software_Id=70).

**Table 1** Posterior point estimates for wavelet-based linear functional fixed and mixed-effects model parameters

	$t1$	$t2$	$t3$	$t4$	$t5$	$t6$	$t7$	$t8$	$t9$	$t10$	$t11$
<b>Fixed effects model</b>											
$\beta_1$	-8.1767	-0.6487	0.3088	0.3665	0.4079	0.4789	-0.0984	-0.7540	-1.9242	-1.7409	-0.4306
$\beta_2$	0.6507	0.3712	-0.0394	-0.0578	-0.1264	-0.4401	-0.6447	-1.2369	-1.3200	-0.2891	0.1734
$\beta_3$	0.5194	0.5944	0.1147	0.4624	0.2063	0.2398	-0.1015	0.2849	-0.5086	0.2791	0.3881
$\beta_4$	0.4807	0.2705	0.5826	0.2133	0.2553	0.2768	0.4755	0.0608	0.4057	-0.0512	0.2725
$\beta_5$	0.2289	0.2970	0.2164	0.1596	0.4982	0.1009	0.1412	0.2504	0.1140	0.1885	-0.0437
$\beta_6$	-0.2353	0.5302	0.1209	0.3271	0.0164	0.3061	0.1145	0.2490	0.0980	0.0981	0.1277
$\beta_7$	0.4460	-0.0887	0.5162	-0.0920	0.1921	0.0643	0.2553	0.1811	-0.2930	0.2461	0.2437
$\beta_8$	1.0614	-0.2692	0.0663	0.1990	0.2842	0.0533	0.2852	0.0375	0.2368	0.0482	-0.1044
$\beta_9$	0.1478	0.1205	0.2785	0.1054	0.0850	0.2078	0.1032	0.2860	-0.2485	0.2719	-0.0108
$\beta_{10}$	0.0073	0.2972	-0.0481	0.3437	-0.1864	0.3412	-0.0144	0.1409	0.1688	-0.0854	0.3438
$\beta_{11}$	-0.1456	0.2875	0.0888	0.0513	0.3205	-0.0500	0.1137	-0.0465	0.5822	-0.0786	-0.3986
$\beta_{12}$	0.0154	-0.0029	-0.0086	-0.0133	-0.0153	-0.0154	-0.0179	-0.0213	-0.0257	-0.0265	-0.0217
$\beta_{13}$	0.0060	0.0205	0.0163	0.0169	0.0183	0.0186	0.0193	0.0202	0.0192	0.0410	0.0376
<b>Linear mixed-effects model</b>											
$\beta_1$	-7.9731	-0.4536	0.4698	0.6176	0.5871	0.6760	0.0566	-0.6395	-1.8999	-1.7287	-0.4599
$\beta_2$	0.7326	0.4771	-0.0690	0.0017	-0.1568	-0.4450	-0.6877	-1.3042	-1.2814	-0.3671	0.1497
$\beta_3$	0.3985	0.5895	0.1705	0.4658	0.1899	0.2609	-0.0661	0.2158	-0.3479	0.1919	0.3240
$\beta_4$	0.5137	0.2386	0.5552	0.1568	0.3134	0.2163	0.3739	0.1511	0.2973	0.0334	0.2510
$\beta_5$	0.1798	0.3414	0.2957	0.1863	0.3630	0.1903	0.1422	0.2886	0.0671	0.2021	-0.0213
$\beta_6$	-0.2027	0.4858	0.2079	0.2766	0.0224	0.3311	0.0603	0.3503	-0.1494	0.2225	0.1310
$\beta_7$	0.6441	-0.1926	0.5398	-0.1113	0.2828	0.0225	0.2986	0.1635	-0.3182	0.2072	0.3827
$\beta_8$	0.8008	-0.1615	0.1428	0.1258	0.3335	-0.0123	0.3754	0.0307	0.1094	0.1296	-0.0900
$\beta_9$	0.3216	0.0934	0.0555	0.1806	0.1891	0.0994	0.1615	0.1567	-0.0214	0.1106	0.0790
$\beta_{10}$	0.1125	0.1861	0.0281	0.2414	-0.0447	0.2706	-0.0223	0.1795	0.1148	-0.0050	0.2219
$\beta_{11}$	-0.3075	0.3997	0.0552	0.1691	0.0596	0.1357	0.0378	-0.0359	0.6916	-0.1238	-0.4430
$\beta_{12}$	0.0186	0.0004	-0.0058	-0.0107	-0.0131	-0.0133	-0.0165	-0.0203	-0.0254	-0.0269	-0.0224
$\beta_{13}$	0.0044	0.0190	0.0155	0.0152	0.0170	0.0170	0.0177	0.0195	0.0180	0.0409	0.0380
$U_1$	-0.0173	0.0001	0.0136	-0.0070	-0.0078	-0.0123	-0.0100	-0.0034	-0.0037	-0.0030	0.0071
$U_2$	-0.0173	0.0001	0.0136	-0.0070	-0.0078	-0.0123	-0.0100	-0.0034	-0.0037	-0.0030	0.0071

$t1, t2, \dots, t11$  represent tree height grids

ponent was between 0.2730 and 0.3428 for the fixed effect and 0.2747 and 0.4974 for the mixed-effects WFMM model. We used trace plots of the model parameters (not shown) to assess convergence. The trace plots revealed that the MCMC algorithm converged and mixed adequately.

To evaluate the performance of the WFMM model, we adopted the Sharma and Parton (2009) variable exponent taper model Eq. 6 as the benchmark. Hereafter, this model is referred to as the VET model. For simplicity, we fitted the VET model using the NLIN procedure of SAS 9.1 (SAS Institute Inc 2004).

$$\frac{d}{D} = \delta_0 \left( \frac{H-h}{H-h_D} \right) \left( \frac{h}{h_D} \right)^{\delta_1 + \delta_2 z + \delta_3 z^2} \quad (6)$$

where  $h_D$  is the breast height (1.3 m),  $z (= h/H)$  and  $\delta_i$ ,

$i=0,1, 2, 3$  are model parameters and other variables are defined earlier.

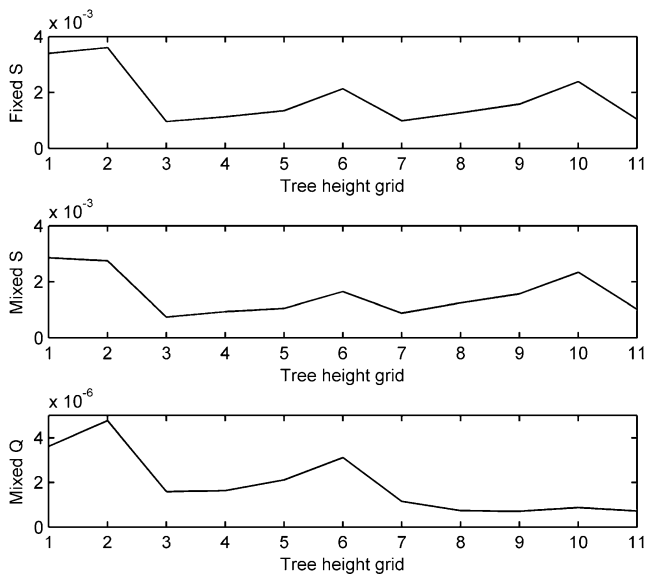
### 3.2 Model performance and evaluation

Model performance was evaluated based on predictions of dib across 11 equally spaced grid points (across tree height) for both the calibration and evaluation data set. We compared the predicted dib for the three models (Eqs. 4, 5 and 6) across the grid. To assess the residuals across the calibration and evaluation data sets, we first computed the posterior mean of WFMM model coefficients for each functional effect. Posterior point estimates of model coefficients were then used to predict dib. At each grid location, the bias in dib prediction is defined as the difference between observed and predicted stem dib. Since WFMM being a Bayesian model, we calculated the 95% credible limits of dib prediction bias

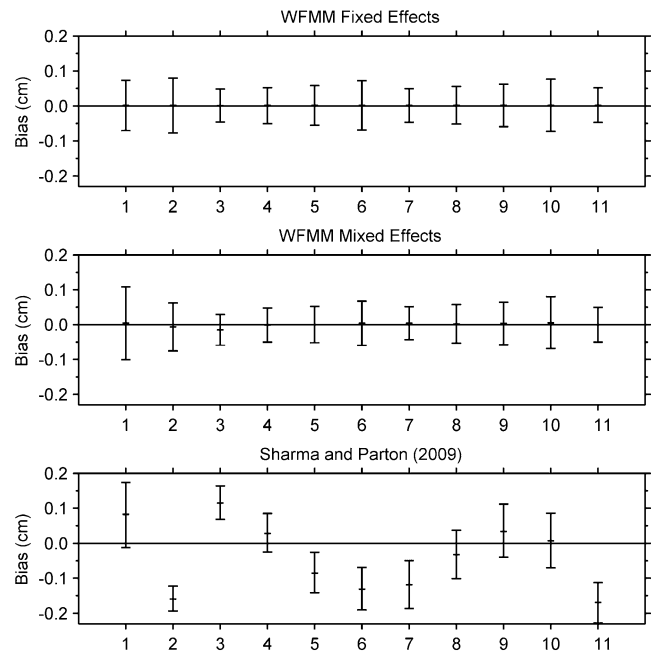
for both fixed and mixed-effects models. In the case of the VET model, we calculated studentized bootstrap 95% confidence interval of dib prediction bias using 10,000 bootstrap samples (DiCiccio and Efron 1996). Unlike a linear mixed model, no clear algorithm is available to predict the random coefficient for a wavelet-based linear mixed functional model for a new data set. Therefore, we were unable to calculate the predicted dibs for this model for the evaluation data set.

The point estimate of posterior mean of the fixed and mixed (fixed and random) effects WFMM model coefficients are given in Table 1. For every grid location, there were 13 model coefficients for fixed effects and two additional coefficients for mixed effects. The variance-covariance matrix of residuals ( $S$ ) for both fixed and mixed-effects WFMM and the random effects ( $Q$ ) for mixed-effects WFMM are given in Fig. 2. The variation of wavelet coefficients was larger at the first, second, sixth and 10th grid locations than at other locations for variance covariance  $S$  and at first, second and sixth grid locations for the random effect variance. The first grid location represents the tree section close to the ground, where larger variation was expected, and the other grid locations represented the regions of taper inflection. Thus, the diagonal variance covariance  $Q$  (random effects) and  $S$  (model residual) matrices figures clearly demonstrate that the WFMM model can easily accommodate the range of non-stationarity and heteroscedasticity in the data.

Dibs predicted by both the fixed and mixed-effects WFMM models were very accurate for the calibration data (Fig. 3). The mean bias 95% credible limits included the zero for all grid



**Fig. 2** Fixed effect covariance parameter ( $S$ ) and random effects covariance parameter ( $Q$ ) of fixed and mixed-effects wavelet-based taper model in relation to equally spaced tree stem height. Note: 1 to 11 along the  $x$ -axis represents the tree height grids



**Fig. 3** Dib prediction mean biases across tree stem grids and their 95% credible limits for the fixed and random effects wavelet-based functional models (WFMM) and mean biases across tree stem grids and their 95% studentized bootstrap interval for a variable exponent model for calibration data. Note: 1 to 11 along the  $x$ -axis represents the tree height grids

points implying that the predictions were unbiased. Hence, the WFMM model fitted the calibration data almost perfectly. In the case of VET model, the studentized bootstrap 95% confidence included 0 at 5 grid points.

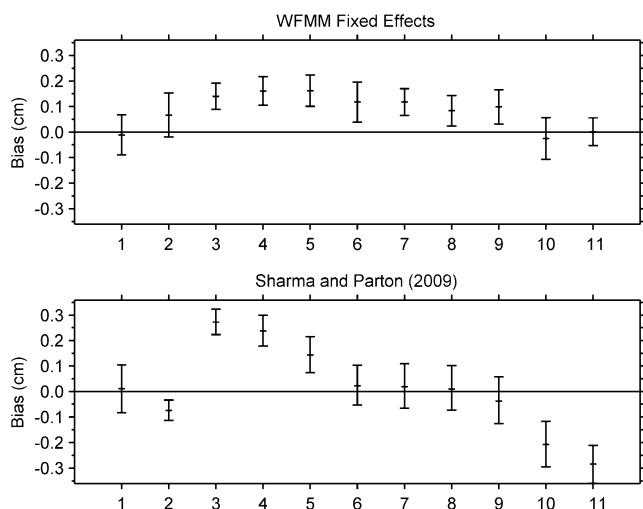
In case of the evaluation data, the WFMM fixed effects model dib prediction had a small bias for some of the height grid locations (Fig 4). Out of 11 grid points across tree stem, only four grid points prediction were unbiased. Similarly, VET model predictions for this datasets were unbiased at five grid point locations. However, the range of mean bias 95% credible limits varied from  $-0.1$  to  $0.22$  cm for the WFMM fixed effects model (Fig. 4). The dib predictions by WFMM model seems to be slightly less accurate for the evaluation data than for calibration data.

The WFMM fixed and mixed-effects models were further evaluated by generating stem profile curves for three randomly selected small-, medium- and large-sized trees (Fig. 5). Dibs predicted for these trees were very similar across tree height for both models. Predicted dibs were more accurate for small- and medium-sized trees than for large-sized tree.

### 3.3 Estimating diameter inside bark and tree volume

A functional model treats dib measurements across the grid as a unit of analysis, and it has a matrix of model parameters. It may seem that using this model is more complicated, but the





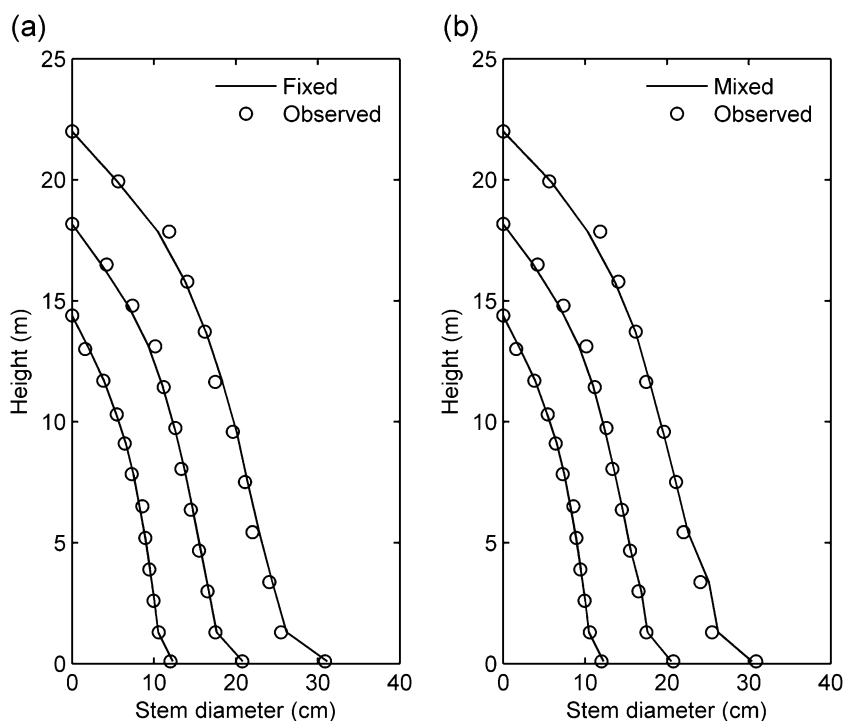
**Fig. 4** Dib prediction mean biases across tree stem grids and their 95% credible limits for the fixed effects wavelet-based functional model (WFMM) and mean biases across tree stem grids and their 95% studentized bootstrap intervals for variable exponent model for validation data. Note: 1 to 11 along the x-axis represents the tree height grids

use of this model is as simple as other traditional models. The fixed effect WFMM taper model can be used to estimate relative dib at any grid locations. For example, diameter inside bark can be estimated as

$$dib(\mathbf{t}) = dbh \cdot \hat{y}(\mathbf{t}), \tag{7}$$

where  $dib(\mathbf{t})$  is the predicted diameter inside bark at grid locations  $\mathbf{t}$ ,  $\hat{y}(\mathbf{t}) = [(h(\mathbf{t})/H) \sqrt{dbh} \sqrt{H}] \hat{\beta}$  and other terms

**Fig. 5** Tree profiles generated using **a** wavelet-based functional fixed effects model and **b** wavelet-based functional mixed-effects model for sample tree #253 (dbh=11.5 cm and ht=14.39 m), #115 (dbh=19.2 cm and ht=18.18 m) and #544 (dbh=28.3 cm and ht=22 m), respectively



as defined earlier. For illustration, let us take the example of the small diameter-sized tree as shown in Fig. 5. For this tree,  $[(h(\mathbf{t})/H) \sqrt{dbh} \sqrt{H}] = [0.0069, 0.0903, 0.1807, 0.2710, 0.3614, 0.4517, 0.5448, 0.6324, 0.7158, 0.8131, 0.9034, -1.5018, -0.7513]$ , where the first 11 numbers are tree relative height at equally spaced grid  $\mathbf{t}$  and the remaining two are standardized values (mean 0 and variance 1) of  $\sqrt{dbh}$  and  $\sqrt{H}$ . This row vector can be multiplied by model parameter  $\hat{\beta}$  (Table 1a) to get predicted values of tree taper across grid  $\mathbf{t}$ . Dib along the bole of the tree can then be obtained by the scalar multiplication of this taper value and dbh (Eq. 7). This results in the predicted dib (centimetres) across the grid  $\mathbf{t}$  as 12.25, 10.48, 10.04, 9.49, 8.91, 8.21, 7.47, 6.56, 5.39, 3.91 and 2.09, respectively.

The dib predictions across the height grid can be linearly interpolated to obtain diameter at any point along the bole. In order to calculate the dib at 5 m using Eqs. 6 and 7 for the trees shown in Fig. 5, for example, the dib across the tree height grids are predicted first. Two stem height grids close to 5 m are then identified, and their corresponding dibs and heights are noted. Dib at 5 m can now be calculated based on these dibs and heights using linear interpolation. In our examples, predicted dib for the small-, medium- and large-sized trees at 5 m height were 9.02, 15.40 and 23.12 cm and 9.04, 15.53 and 23.21 cm by using the fixed and mixed-effects WFMM, respectively. The volume of the tree can then be calculated using these predicted diameters and heights along the bole by applying Smalian’s formula (Avery and Burkhart 2002). Volumes (deci m<sup>3</sup>) of these small, medium and large trees are 67.824,

233.506 and 599.822, respectively, for the fixed and 68.140, 235.806 and 608.342, for the mixed-effects WFMM models, respectively.

#### 4 Discussion

This study demonstrated the application of wavelet and functional mixed models for tree taper modelling in forestry. When diameters are measured at equally spaced tree height, these measurements can be transformed into wavelet space and functional mixed model can be applied to estimate model parameters. Once model parameters are re-projected onto data space, these models can be used to predict tree diameters at any height and also volumes between any heights.

Diameters predicted by WFMM taper models were unbiased for the calibration data but were slightly biased for the evaluation data. However, the magnitude of bias was within the acceptable range (i.e. less than 0.16 cm). This implies that the wavelet approach could accurately mimic the data used in model fitting as required in image restoration but is unable to capture the mean function accurately to describe the taper of the trees not used in the parameter estimation.

The WFMM model works in the functional context, i.e., it fits  $T$  scalar linear mixed models simultaneously. Therefore, the fit could be unstable either due to collinearity in the design matrix or by having a small effective sample size for some of the variance components (Morris et al. 2006). We recognized this difficulty in this study. Unlike a linear mixed model, the current version of WFMM does not accept a single random effect due to assumptions related to variance covariance matrices  $Q$  and  $S$ . The assumptions on which these matrices are based sometimes limit the choice of random effects. In addition, if we choose multiple random effects, the fit could be poor as the model assigns larger variability to the random wavelet coefficients than to fixed effect wavelet coefficients. Therefore, the current version of WFMM requires judicious choice of random and fixed effects.

These limitations of the WFMM model forced us to choose bark thickness as the random effect. Since the bark thickness varies from one tree to another, the value of the dependent variable ( $y = \text{dib/dbh}$ ) for two trees may not be the same even if their dbhs are identical because of the differences in bark thickness. Therefore, the bark thickness was chosen as the random effect here. However, the model did not limit the random effects that are considered independent across rows. As the WFMM is at the early stages of development, this issue needs to be addressed by conducting more research and developing an efficient algorithm to support broader application. Another concern

with the WFMM model is the lack of a global functional test for formally assessing fixed and random effects. However, we found it to be very robust for the choice of fixed effects. When the Eq. 4 was fitted without  $H$  (tree height) in the design matrix  $X$ , this model also provided a better fit in terms of bias in predicting diameters.

The wavelet-based functional taper model presented provides a viable alternative to modelling taper across trees. Since the WFMM approach can be applied to any data that are equally spaced across time or space, this approach opens new possibilities for analysing longitudinal or taper data collected across time or space.

**Acknowledgements** This study was supported by Ontario Ministry of Natural Resources (OMNR) and the Forestry Futures Trust Enhanced Forest Productivity Science Program. The authors gratefully acknowledge the support of Jeffery Morris and Richard Herrick, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030 for their support to fit the WFMM mixed-effects model. The authors are grateful to Lisa Buse (OFRI) for editing the manuscript and John Parton (OMNR) for his support in collecting the data. They also thank the Associate Editor Barry Gardiner, whose insightful comments and questions have improved this article.

#### References

- Abramovich F, Sapatinas T, Silverman BW (1998) Wavelet thresholding via a Bayesian approach. *J R Stat Soc B* 60:725–749
- Amidon EL (1984) A general taper function form to predict bole volume for five mixed-conifer species in California. *For Sci* 30:166–171
- Antoniadis A, Sapatinas T (2007) Estimation and inference in functional mixed-effects models. *Comput Statist Data Anal* 51:4793–4813
- Avery TE, Burkhart HE (2002) *Forest measurements*. McGraw-Hill, New York
- Besse PC, Cordot H, Faivre R, Goulard M (2005) Statistical modeling of functional data. *Appl Stoch Model Bus* 21:165–173
- Bi H (2000) Trigonometric variable-form taper equations for Australian eucalyptus. *For Sci* 46:397–409
- Carlin BP, Louis TA (2000) *Bayes and empirical Bayes methods for data analysis*. Chapman and Hall, Boca Raton, 419p
- Chui CK (1992) *An introduction to wavelets*. Academic, Boston, 266 p
- Clyde M, Parmigiani G, Vidakovic B (1998) Multiple shrinkage and subset selection in wavelets. *Biometrika* 85:391–401
- Daubechies I (1992) *Ten lectures on wavelets*. SIAM, Philadelphia, 357 p
- Demidenko E (2004) *Mixed model theory and applications*. Wiley, Hoboken, 704 p
- DiCiccio TJ, Efron B (1996) Bootstrap confidence intervals. *Stat Sci* 11:189–228
- Dieguez-Aranda U, Castedo-Dorada F, Alvarez-Gonzalez JG, Rojo A (2006) Compatible taper function for Scots pine plantations in northwestern Spain. *Can J For Res* 36:1190–1205
- Ferraty F, Vieu P (2006) *Nonparametric functional data analysis: theory and practice*. Springer, New York, 258 p
- Garber SM, Maguire DA (2003) Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. *For Ecol Manage* 179: 507–522

- Gregoire TG, Schabenberger O, Barrett JP (1995) Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can J For Res* 25:137–156
- Gregoire TG, Schabenberger O, Kong F (2000) Prediction from an integrated regression equation: a forestry application. *Biometrics* 56:414–419
- Guo W (2002) Functional mixed effects models. *Biometrics* 58:121–128
- Keinert F (2004) Wavelets and multiwavelets. Chapman and Hall/CRC, Boca Raton, 269 p
- Koskela L, Nummi T, Wenzel S, Kivinen V-P (2006) On the analysis of cubic smoothing spline-based stem curve prediction for forest harvesters. *Can J For Res* 36:2909–2919
- Kozak A (1988) A variable-exponent taper equation. *Can J For Res* 18:1363–1368
- Kozak A, Munro DD, Smith JHG (1969) Taper functions and their application in forest inventory. *For Chron* 45:278–283
- Laird N, Ware JH (1982) Random-effect models for longitudinal data. *Biometrics* 38:963–974
- Leites LP, Robinson AP (2004) Improving taper equations of loblolly pine with crown dimensions in mixed-effects modeling framework. *For Sci* 50:204–212
- Li R, Weiskittel AR (2010) Comparison of model forms for estimating stem taper and volume in the primary conifer species in the Acadian Region of North America. *Ann For Sci* 67:16p
- Liu CJ (1980) Log volume estimation with spline approximation. *For Sci* 26:361–369
- Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intel* 11:674–693
- Max TA, Burkhart HE (1976) Segmented polynomial regression applied to taper equation. *For Sci* 22:283–289
- Morris JS, Carroll RJ (2006) Wavelet-based functional mixed models. *J R Stat Soc B* 68:179–199
- Morris JS, Vannucci M, Brown PJ, Carroll RJ (2003) Wavelet-based nonparametric modeling of hierarchical functions in colon carcinogenesis. *J Am Stat Assoc* 98:573–597
- Morris JS, Arroyo C, Coull BA, Ryan LM, Herrick R, Gortmaker SL (2006) Using wavelet-based functional mixed models to characterize population heterogeneity in accelerometer profiles: a case study. *J Am Stat Assoc* 101:1352–1364
- Morris JS, Brown PJ, Herrick RC, Baggerly KA, Coombes KR (2008) Bayesian analysis of mass spectrometry proteomic data using wavelet-based functional mixed model. *Biometrics* 64:479–489
- Newnham RM (1992) Variable-form taper functions for four Alberta tree species. *Can J For Res* 22:210–223
- Ogden RT (1997) Essential wavelets for statistical applications and data analysis. Birkhauser, Boston, 198 p
- Ormerod DW (1973) A simple bole model. *For Chron* 49:136–138
- Percival DB, Walden AT (2000) Wavelet methods for time series analysis. Cambridge University Press, New York, 594 p
- Ramsay JO, Silverman BW (1997) Functional data analysis. Springer, New York, 310 p
- SAS Institute Inc. (2004) SAS/STAT 9.1 user's guide. SAS Institute, Cary
- Schabenberger O, Pierce FJ (2002) Contemporary statistical models for the plant and soil sciences. CRC, Boca Raton, 790 p
- Sharma M, Oderwald RG (2001) Dimensionally compatible volume and taper equations. *Can J For Res* 31:797–803
- Sharma M, Parton J (2009) Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. *For Sci* 55:268–282
- Trincado G, Burkhart HE (2006) A generalized approach for modeling and localizing stem profile curves. *For Sci* 52:670–682
- Valentine HT, Gregoire TG (2001) A switching model of bole taper. *Can J For Res* 31:1400–1409
- Vidakovic B (1999) Statistical modeling by wavelets. Wiley, New York
- Zhang Y, Borders BE, Bailey RL (2002) Derivation, fitting, and implication of compatible stem taper-volume-weight system for intensively managed, fast growing loblolly pine. *For Sci* 48:595–607