



Canadian Journal of Forest Research
Revue canadienne de recherche forestière

**A new method for capturing stem taper variation for trees
of diverse morphological types**

Journal:	<i>Canadian Journal of Forest Research</i>
Manuscript ID	cjfr-2016-0018.R1
Manuscript Type:	Article
Date Submitted by the Author:	18-Mar-2016
Complete List of Authors:	MacFarlane, David; Michigan State University Weiskittel, Aaron; University of Maine
Keyword:	taper, form, branching architecture, mixed-effects modeling, morphology



1 A new method for capturing stem taper variation for trees of diverse morphological types.

2 David W. MacFarlane^{1*} and Aaron R. Weiskittel²

3 ¹Department of Forestry, Michigan State University, East Lansing Michigan, 48840

4 ²University of Maine, School of Forest Resources, Orono, ME 04469

5 *corresponding author; email: macfar24@msu.edu

6 **Abstract**

7 Understanding variation in tree stem form is fundamental to both ecological and economic
8 assessments of forest ecosystem structure and function. Stem taper models (STMs) are widely
9 used to describe tree form, but it can be challenging to apply them to trees with stems that
10 diverge from an idealized norm, often leading to exclusion of many trees from stem taper
11 studies. Here, new 'whole-tree' form type classes are advanced, as simple and useful
12 groupings for capturing stem form variation of trees of diverse morphological types, and tested
13 with a large tree data set without exclusion criteria. New form type classes explained much
14 more of main stem form variation than knowledge of tree species, while 'merchantable form
15 types' explained most stand-to-stand level variation. Broad-leaved species were much more
16 likely to have complex stem forms than needle-leaved, but species 'evergreenness' was a very
17 weak predictor of stem form variation, when tree and stand –level form variation was accounted
18 for. A new, generalized framework for stem taper modeling is demonstrated, using both species
19 and merchantable form types to capture tree-level random effects. New form types and the STM
20 approach are relatively easy to apply and should be relatively simple to integrate into any
21 conventional forest inventory system. Overall, the study demonstrates the importance of
22 including and accounting for the diversity of observed stem forms in developing STMs.

23 **Keywords:** taper, form, branching architecture, mixed-effects modeling, morphology

24 1. Introduction

25 Stem taper models (STMs) describe the changing shape of the 'main' stem of a tree,
26 from ground to tree top. This enables them to be used to explore and predict variation in stem
27 morphology for trees of diverse species, because stem taper reflects factors such as ecological
28 conditions and forest management history (Assmann 1970, Muhairwe et al. 1994, Niklas 1995,
29 Valentine and Gregoire, 2001, Bravo-Oviedo et al. 2014). STMs have long been applied to
30 forest inventory data, because they allow for estimation of merchantable stem volume to a
31 variable top diameter as well as merchandizing the stem into various products (Kozak et al.
32 1969). More recently, they have been applied to simultaneously estimate tree volume and
33 biomass (Jordan et al. 2006, Zakrzewski and Duchesne 2012, Ver Planck and MacFarlane
34 2015), which can allow for extension of timber volume inventories to ecological studies and
35 provide quantitative data for understanding synergies and tradeoffs between different types of
36 forest ecosystem services, such as carbon storage versus timber production. For example, the
37 US Forest Inventory and Analysis program currently links national timber stocks inventory data
38 to the national forest carbon inventory, by using a component ratio method (Domke et al. 2012)
39 that predicts both total tree and tree component biomass (e.g., branches, leaves), as ratios of a
40 tree's merchantable main stem volume.

41 Despite the long and successful history of development and application of STMs, it can
42 be challenging to apply them to trees with complex stem forms, namely trees that fork to some
43 significant degree or have major limbs that otherwise distort the idealized, continuous main stem
44 form that is typically assumed in STMs. For example, MacFarlane (2010) and Ver Planck and
45 MacFarlane (2014) showed that taper may drop precipitously at branch nodes, especially if the
46 branches are large relative to the main stem. Tree modelers have taken two general
47 approaches to address this challenge: (1) exclude trees which do not conform well to the
48 assumed model, or (2) incorporate different tree branching architectures into STMs.

49 The prevailing approach is to fit STMs to a sample population of trees that have been
50 censored to some degree by excluding trees which diverge from some idealized stem form.
51 Many research papers report excluding 'irregular' trees from data sets used to fit STMs. For
52 example, Burkhart (1977) reported including only single-stemmed trees, while others excluded
53 trees with forks (Forslund 1982, Barrio Anta et al. 2007) or trees with multiple stems or crooked
54 boles (Özçelik 2008). McTague and Bailey (1987) excluded trees with forks, ramicorn branches
55 (large, high-angled branches) or 'foxtails' (sections of the stem which grow without producing
56 any lateral branches) and Newnham (1992) reported excluding 2293 of 7367 (31% of) trees,
57 which had 'defects', when fitting taper models to four Alberta tree species, without specifying
58 what those defects were. Sometimes, entire types of stands have been excluded from STM
59 studies. For example, Burkhart and Walton (1985) chose plantations that had not been thinned,
60 burned, or pruned, and were free of severe insect or disease damage, to assess the influence of
61 crown variation on the taper of loblolly pine (*Pinus taeda* L.). Such censoring could be a
62 significant problem whenever STMs fit to censored data are later applied to populations of trees
63 covering a larger range of tree forms or stand conditions, which is generally the case.

64 Some researchers have sought approaches which recognize that a model of the main
65 stem of a tree should be referenced somehow to the whole tree it is part of, whose other parts
66 (namely branches) influence its form (MacFarlane 2010, Ver Planck and MacFarlane 2014).
67 Muhairwe (1994) demonstrated that simply including a measure of crown size as a predictor
68 variable could help explain some of the variation in tree shape, in even-aged, fully-stocked,
69 lodgepole pine (*Pinus contorta*) stands, where we would expect fairly uniform crown geometry.
70 One solution offered for teak (*Tectona grandis*) explicitly parameterized the heights of forks and
71 the relative change in diameter at forking points into the STM (Adu-Bredu et al. 2008). While
72 effective, this latter model requires a number of additional measurements on every tree, which
73 would be time consuming and prone to measurement error, and the model allows for only two

74 forks in a tree. Two other approaches model continuous shifting allocation between a main stem
75 and branches, above and below a relative crown height, but require some prior estimate of
76 volume (Ver Planck and MacFarlane 2014) or the centroid of volume (MacFarlane 2010) for
77 each tree. Zakrzewski (2011) modeled the cumulative distribution of volume from a taper
78 model, including wood in both the main stem and branches, but his approach did not allow for
79 separation of main stem volume from branch volume; the two components were treated as a
80 composite sum.

81 Our literature review revealed that, despite a long history of applying STMs in forestry
82 and recent efforts to link changes in tree form to changes in tree biomass and forest carbon
83 storage (e.g., Chave et al. 2014, MacFarlane 2015), little research has been conducted to
84 document the potential problem of excluding trees with diverse forms, when fitting STMs.
85 Further, no generalized, low-cost, simply-to-apply solution has been advanced to model stem
86 taper over the very broad range of tree forms that can be observed. A simple, and possibly
87 effective approach would be to identify 'whole-tree' form types (meaning ones that reference
88 both the main stem and branches), which trees could be relatively easily assigned to and would
89 capture major differences in main stem form variation, with specific reference to the relative
90 'branchiness' of the tree. To address these issues, we: (1) develop whole-tree form type
91 classes that provide useful groupings for capturing stem form variation; (2) assess the relative
92 abundance of trees of different form type classes, within and between species, from a large tree
93 data set with no specific exclusion criteria; (3) explore the hierarchical structure of variation in
94 tree form at the tree, stand and species levels over a large spatial domain, and (4) discuss
95 implications of the former on development and application of STMs to trees of diverse stem form
96 and branching architecture.

97

98 2. Materials and Methods

99 2.1. Study area and tree data

100 The study area included public forest lands in the state of Michigan, USA. These forests
101 cover about 1.7 million ha of Michigan's 14.7 million ha land area and encompass a diverse
102 array of glacial landforms and forest community types, within temperate broad-leaf and mixed
103 needle-leaf-broad-leaf biomes (Dickmann, 2004). Stands were selected at 31 test locations to
104 cover a diverse set of species, tree sizes and stand conditions and involved all Michigan
105 Department of Natural Resources (MDNR) management regions across the state. Trees were
106 selected from 12 of the locations using horizontal point sampling on a 64 m square grid. At the
107 other 19 locations, trees were selected systematically, sequentially selecting sample trees of
108 desired species within size classes, until the desired size range of target species were sampled
109 from the stand. In all cases, trees were selected without exclusion of trees with complex or
110 irregular stem forms. In total, 1828 trees of 21 species were destructively sampled covering a
111 wide range of sizes, from a minimum size of 11.7 cm to a maximum of 101.6 cm diameter at
112 breast height (DBH; 1.3 m) (Table 1).

113 Trees were felled and dissected by professional loggers and measured. Stem diameter
114 at breast height (DBH, cm) was measured before felling and total height (H, m) was measured
115 after felling (species averages are shown in Table 1). The main stem and branches of each tree
116 were cut into approximately 2.5 m sections to a 10 cm top diameter outside bark (DOB),
117 although shorter and longer section lengths were occasionally included. Beginning with the
118 stump, the main stem diameter outside bark was measured at the ends of cut sections along
119 with the length of each section all the way to the top of the tree. The main stem was followed to
120 the top of the tree, selecting the largest and most vertical stem at a fork to be the main stem
121 (aka the "dominant" stem, *sensu* VerPlanck and MacFarlane 2014). Any branch with a basal

122 diameter of at least 10 cm with at least 2.5 m long, sound section was measured to a 10 cm top
123 DOB, in the same manner as the main stem.

124 *2.2. Whole-tree form types defined*

125 Botanists have worked for decades to define generalized architectural models for plant
126 form, based on the relative arrangement of plant parts, which reflect both endogenous (genetic)
127 and exogenous (environmental) processes that define plant form at any point in time
128 (Barthélémy and Caraglio 2007). Here, we sought to define analogous architectural form types
129 for trees, to capture major, visually-apparent differences in the allocation of wood to the main
130 stem versus branches. The form classes were also designed to be consistent with common
131 forest inventory practices, to give them practical as well as scientific value for tree form
132 evaluation.

133 *2.2.1. Large Branch Type (LBTs)*

134 MacFarlane (2010) theorized that only relatively large branches should have a significant
135 impact on main stem form, so the simplest classification is binary, determining whether a tree
136 has relatively large branches, or not. This Large Branch Type (LBT) takes on a value of zero if
137 the tree has no large branches and a value of one if it does. The definition of a 'large' branch
138 may vary from region to region, but can be made consistent within a forest inventory system by
139 linking it to merchantability standards applied within that system. Here, the minimum size of a
140 'large' branch is one with a minimum basal diameter = 10 cm and containing at least one sound,
141 pulp-sized log to a minimum top DOB = 10 cm. This standard meets the average minimum
142 specifications for a merchantable section of a tree in Michigan forestry practices (MDNR 2013)
143 and the USA in general (e.g., Clark et al. 1991, Jenkins et al. 2003, Woodall et al 2011). For
144 this study, trees were assigned an LBT = 1 if they contain a branch with at least 2.5 m long

145 branch section, with a minimum basal diameter of 10 cm and a minimum top DOB also = 10; or
146 LBT = 0, otherwise. All study trees were assigned an LBT (see Table 2).

147 2.2.2. Merchantable Branch Type (MBTs)

148 The LBT can be made more descriptive and useful for volume inventory by
149 differentiating between large and very large branches, which can correspond in size to branches
150 that could be used for the two major types of wood products: pulp and saw products,
151 respectively. To be classed as having branch saw wood present, a tree needed to have at least
152 one branch section at least 2.5 m long, with a minimum basal diameter = 20 cm and at least one
153 sound log to a minimum top DOB also = 20 cm; this is also consistent with regionally and
154 nationally -used minimum merchantability standards cited above. Trees are assigned to MBTs
155 using a two-digit code system, with the first digit representing the absence / presence of pulp
156 log-sized branches and the second indicting the absence / presence of saw log-sized branch
157 sections. So, there are four possible MBTs for a tree: 0-0 = no merchantable branch wood; 1-0
158 = pulp wood in branches, but no saw wood in branches; 0-1 = no pulp wood in branches, but
159 saw wood in branches; and 1-1 both pulp and saw wood in branches. All study trees were
160 assigned an MBT (see Table 2).

161 2.2.3. Merchantable Form Types (MFTs)

162 LBTs and MBTs allow for identification of trees with or without large, merchantable
163 branches, or of different branch types, but have no code representing the main stem. Since
164 trunk taper is expected to change as a tree increases in size (Niklas 1995), form types can be
165 further differentiated by categorizing the main stem into two merchantable types, just as the
166 branches were under the MBT system. So, a four-digit Merchantable Form Types (MFT) code
167 was developed to create a variety of tree form types relevant to forest inventory methods,
168 depending on whether or not the tree contains merchantable wood in some or all of the

169 following four categories: (1) main stem pulp, (2) main stem saw, (3) branch pulp, or (4) branch
170 saw. A value of “1” is recorded when that part-product is present and “0” when it is not. So, a
171 tree with a code of “1-0-1-0”, would have pulp wood volume in both the main stem and
172 branches, but would not have any saw wood volume in either tree component. All study trees
173 were assigned an MFT (see Table 3).

174 2.3. Analytical Methods

175 Our data set consisted of measurements of the change in main stem cross-sectional
176 area, as a function of height above ground from the stump to the top of the tree, computed from
177 the stem diameter measurements taken along the stem. We used a multi-level mixed-effects
178 modeling framework for analysis (Venables and Ripley 2002), recognizing the hierarchical
179 structure of our data, where correlated within-tree variation in stem cross-sectional area
180 (computed from the stem diameter measurements of the stem) represented the finest level of
181 variation. When specifying the hierarchy, we sought to investigate tree-level form variation from
182 the most general to the most specific effects on tree form, which was, from top to bottom:
183 evergreenness > spp > form type > stand > tree. Since genetics has a top-down effect
184 constraining tree form (Barthélémy and Caraglio 2007, Dardick et al. 2013), ‘evergreenness’
185 (e.g., Ducey 2012) was used as a the coarsest group, which species were nested in, reflecting
186 the expected difference between trees with a decurrent (broadleaf) versus excurrent (needle-
187 leaf) stem form due to expected differences in apical dominance between conifers and
188 angiosperms (Wilson 2000). Species-level effects were further refined by form type groupings
189 (LBT, MBT, MFT), which represented within-species variation in form. Finally, the random effect
190 of local environment on tree form (Xiong et. al. 2010) was captured by the specific stand
191 conditions the trees were drawn from.

192 In our first analytical step, we computed main stem volume for each tree from the taper
193 measurements using tree-specific mixed-effects B-spline regression; a semi-parametric method
194 described by Kublin et al. (2013), that was fit using the *TapeR* package (Kublin and
195 Breidenbach 2013) in R v3.2.3 (R Core Team 2015). Then, we estimated the relative
196 contribution of each grouping variable in the hierarchy to stem form variation, using main stem
197 volume as an integrating variable to capture stem form differences. This was accomplished with
198 the *varcomp* function within the *APE* package (Paradis et al. 2004) in R, which computes the
199 variance components from a fitted mixed-effects model object.

200 Our next step was to use non-linear mixed effects modeling (*nlme*, Venables and Ripley
201 2002) to fit a parametric taper model to the whole data set (all trees) and then adjust model
202 coefficients based on random effects, which were allowed to vary based on how each individual
203 tree diverged from the trend over all the diverse tree types represented in our data. There were
204 several reasons for this approach. First, there was insufficient data to fit a separate model for
205 every combination of species and form type, despite the very large data set we had (see Tables
206 2 and 3). A combined model with dummy variables for each species and form type would have
207 created an enormous number of parameters and coefficients to estimate; even a reduced
208 version with only evergreenness and LBT failed to converge on a solution. Simply allowing the
209 intercept to vary by species or form type resulted in grossly oversimplified effects of form type or
210 species.

211 We chose the highly flexible segmented polynomial model of Max and Burkhart (1976) to
212 explore how well the form types captured major differences in tree form, within and between
213 species, because this model is well-tested, it has been previously been shown to work with
214 predicted random effects (Cao and Wang 2011), and allowed for evaluating the influence of tree
215 form on various stem segments.

216 The segmented polynomial model we used was in the form:

$$217 \quad Y = \beta_1(X - 1) + \beta_2(X^2 - 1) + \beta_3(\alpha_1 - X)^2 I_1 + \beta_4(\alpha_2 - X)^2 I_2 \quad \text{eq. 1}$$

218 where Y is the relative squared diameter outside bark (d^2 , cm^2) at a height above ground h (m) =
 219 $(d_h/\text{DBH})^2$, which hereafter is referred to as relative cross-sectional area outside bark (*relcob*); X
 220 is the relative height above ground (*relh*) = h/H ; and α_i and β_i are coefficients to be estimated,
 221 with α_i representing the joining points of the segments of the equation. I_1 and I_2 are indicator
 222 variables denoting the relative position of the joints with respect to the i^{th} cross-sectional position
 223 in the upper and bottom portions of the tree, respectively.

$$224 \quad I_1 = 1, \text{ if } \alpha_1 \geq X, \text{ else } I_1 = 0.$$

$$225 \quad I_2 = 1, \text{ if } \alpha_2 \geq X, \text{ else } I_2 = 0.$$

226 Max and Burkhardt (1976) allow for α_i to be specified or estimated as free coefficients.

227 We tested a model where we set the value of α_2 at the relative height of DBH (\overline{rbh}), because
 228 this simplified model fitting procedures, and because α_2 is typically very close to \overline{rbh} point when
 229 estimated independently (α_1 by contrast is highly variable, typically occurring somewhere
 230 between 50 to 90% of total tree height). Additionally, since the general form of the taper model
 231 (eq. 1) is relative to DBH there is a natural inflection point at DBH, where $Y = 1$; points below
 232 that are generally $Y > 1$ and above that $Y < 1$.

233 The DBH-segmented Max and Burkhardt (1976) model is:

$$234 \quad Y = \beta_1(X - 1) + \beta_2(X^2 - 1) + \beta_3(\alpha_1 - X)^2 I_1 + \beta_4(\overline{rbh} - X)^2 I_2 \quad \text{eq. 2}$$

$$235 \quad I_1 = 1, \text{ if } \alpha_1 \geq X, \text{ else } I_1 = 0$$

$$236 \quad I_2 = 1, \text{ if } \overline{rbh} \geq X, \text{ else } I_2 = 0$$

237 We fit both models (eq. 1 and eq. 2) to the data and used a likelihood ratio test to determine if
238 leaving α_2 as a free coefficient was superior to assigning it to \overline{rbh} .

239 To fit the models above, we used the *NLME* package (Bates et al. 2015) in R, specifying
240 tree-level random effects nested within each stand. In addition, hierarchical data tend to have
241 within-subject correlation and are also likely to have within subject residual heteroscedasticity.
242 We dealt with both by inclusion of a continuous first-order autoregressive (CorCAR1) correlation
243 structure and a variance power (varPower) weighting structure as a both function of relative
244 height to address within-subject correlation and heteroscedasticity, respectively. However, both
245 were found to not significantly improve model fit (tested using a likelihood ratio comparison) and
246 consequently, not included in the final model.

247 We explored different combinations of model coefficients α_i and β_i as random effects to
248 determine which was contributing the most to unexplained variation. We looked for the best
249 combination of coefficients to assign random effects to, which allowed models to converge and
250 which improve the model fit as indicated by a likelihood ratio test. In the final stage of modeling,
251 we predicted the random effects for each tree within each stand from the best model, as linear
252 functions of species and form types.

253

254 **3. Results**

255 *3.1. Diversity in tree form types*

256 A broad range of tree form types were found within and among species over a range of
257 tree sizes (Tables 2 and 3). As expected most needle-leaf trees had few branches large
258 enough to meet minimum merchantability standards (only 3% of trees examined had an LBT =
259 1, Table 2). *Pinus strobus*, however, was the exception among conifers, with about 14%

260 containing large branches (Table 3). By contrast, almost half of all broad-leaved trees we
261 observed had at least one large, merchantable branch and about 17% of those were large
262 enough to produce saw logs (Table 2).

263 Over all trees, the five most common MFTs were: 1-0-0-0: which corresponds roughly to
264 smaller ("pole"-sized) trees (smallest DBH = 12 cm) without relatively large branches; 1-0-1-0:
265 smaller trees (DBH \geq 15 cm) with a major fork; 1-1-0-0: larger (saw-timber-sized trees without
266 large branches, smallest with DBH = 21 cm); 1-1-1-0: larger trees with large branches (DBH \geq
267 21 cm); and 1-1-1-1: large trees with saw-log sized branches (smallest individual had DBH = 31
268 cm). Allometrically, a tree has to be large enough to grow a relatively large branch, and the
269 bigger the ratio of the largest branch to the main stem, the more dramatic the fork (MacFarlane
270 2010). For example, the smallest tree with a pulp-sized branch (category 1-0-1-0) had a branch
271 fork with a basal diameter almost as large as the main stem.

272 3.2. *The relative contribution of form types to explaining tree form variation*

273 Analysis of variance components of main stem volume showed that knowing the species
274 or species group ('evergreenness') explained very little of the difference in main stem form and
275 volume, when compared to tree-to-tree and within-stand variation, suggesting a very high level
276 of intra-specific variation in stem form, across a broad range of species and forest communities
277 (Fig. 1a). Knowing that a tree, within a species, has a relatively large branch (LBT), large
278 enough to be merchantable, added considerably to explaining tree to tree variation (Fig. 1b).
279 Furthermore, including information related to the saw timber potential of the branch (MBT),
280 increased the degree of variation explained (Fig. 1b). MFTs were the most informative; knowing
281 the MFT of a tree explained about half of the tree-to-tree variation in main stem volume, but it
282 also helped considerably to clarify intra-specific variation, such that they combine to explain
283 about two-thirds of within-tree form variability (Fig. 1d). Further, with trees assigned to a MFT

284 much of the variation between stands was accounted for (note the relatively shallow slope from
285 'MFT' to 'stand' in Fig. 1d, relative to that for species, LBT and MBT, Fig.1 a-c, respectively).

286 3.3. Divergence of trees of different form type and species from a general all-species STM

287 The DBH-segmented Max and Burkhart (1976) model (eq. 2) was superior by all metrics
288 to the standard model (eq. 1, Table 4), so it was used for all subsequent model fittings. Only
289 having to estimate one joining point simplified estimation of fixed and random effects on other
290 model coefficients. The best model that converged was one with random effects on coefficients
291 β_1 , β_2 , and β_4 , as indicated by AIC, BIC, and a likelihood ratio test (Table 5). Both MFT and
292 species were significantly ($p < 0.0001$) and linearly correlated with all three coefficients and
293 model fitting generated using linear adjustment factors for each coefficient depending on the
294 species and MFT (Table 6).

295 The final generalized model selected for all trees of all species and form types was:

$$296 Y = (\beta_1 + \mu_{1(ijk)})(X - 1) + (\beta_2 + \mu_{2(ijk)})(X^2 - 1) + \beta_3(\alpha_1 - X)^2 I_1 + (\beta_4 + \mu_{4(ijk)})(\overline{rbh} - X)^2 I_2 \quad \text{eq. 3}$$

$$297 I_1 = 1, \text{ if } \alpha_1 \geq X, \text{ else } I_1 = 0$$

$$298 I_2 = 1, \text{ if } \overline{rbh} \geq X, \text{ else } I_2 = 0$$

299 where $\mu_{1(ijk)}$, $\mu_{2(ijk)}$, and $\mu_{4(ijk)}$ are the predicted random effects for a tree in stand k of MFT j and
300 species i . Coefficients and fit statistics for eq. 3 are shown in Table 6.

301 After fitting, eq. 3 was used to explore model behavior. To apply the model, stand-level
302 random effects were set to zero and the predicted random effects terms for both MFT and
303 species were added to the random effects intercepts which were then added to the fixed effects
304 terms in the models (Table 6).

305 For example, the predicted random effect term μ_1 was added to the fixed effect
306 coefficient β_1 (eq. 3). The intercept for the predicted random effect = 0.0710 (Table 6). If the tree
307 was an *Acer saccharum* tree with a MFT of 1-0-1-0 we added - 0.1606 (for species, Table 6)
308 plus - 0.1714 (for MFT, Table 6), meaning $\mu_1 = 0.0710 + [- 0.1606] + [- 0.1714] = - 0.3320$, which
309 is added to the fixed-effect coefficient β_1 (= -3.9845, Table 6) + [-0.3320] = - 4.3165.

310 Looking at form variation within a species, against the background of variation in all
311 species (Fig. 2), it can be seen that e.g., *Q. rubra* trees with relatively large branches had
312 considerably more taper than *Q. rubra* trees without them, such that saw-sized trees with saw-
313 sized branches (MFT = 1-1-1-1) were more similar in stem form to pole-sized trees with pole-
314 sized branches (1-0-1-0) than the latter were to pole-sized trees without large branches (1-0-0-
315 0). Comparing two species, e.g., *A. saccharum* and *T. americana*, which tend to co-occur on a
316 variety of mesic upland sites in the region, we can see a case where species differences were
317 important, but only in magnifying differences caused by branching (Fig. 3). According to our
318 model (eq. 3) and the underlying data (see Fig. 3), the main stem of *T. Americana* tends to taper
319 less than that of sugar maple, on average, even when forks or other major branching effects are
320 accounted for. However, the “regular” form stems (1-0-0-0) of both species are more similar to
321 each other than “forked” individuals (1-0-1-0) of the same species, further demonstrating that
322 intraspecific stem form variation tended to be much greater than interspecific variation (as
323 shown in Fig. 1). The model (eq. 3) clearly shows that the net effect of branching is to divert
324 volume from the main stem into branches, increasing stem taper and reducing accumulated
325 volume in the main stem.

326

327

328

329 4. Discussion

330 4.1. *The value of form type classes which consider both stem and branches*

331 Foresters have long recognized that trees have different stem forms and have sought
332 ways to capture this, as a way to improve stem volume estimation, but also to fundamentally
333 understand variability in tree form. For example, Assmann (1970) chronicled almost a century of
334 theories which attempt to describe tree-to-tree stem form variation, and noted (on p. 64) “the
335 problem of form and form factor”, which is simply the fact that tree-to-tree stem form varies
336 within wide limits and, for each tree, it responds dynamically to changing environments (e.g.,
337 thinning, see Assmann 1970, p. 61). Much of the early research on form resulted in the advent
338 of ‘form classes’, which look at the ratio of some upper stem diameter, typically to top of the first
339 log (e.g., Girard Form Class, Avery and Burkhart 2015) to DBH. These early form classes
340 generally differ from modern STMs, because they focus mainly on modeling taper lower down in
341 the tree’s stem, where the most economically valuable parts are. Since STMs seek to profile
342 stem form from the base to the top of the tree using a continuous mathematical function
343 (Zakrzewski 1999), they should require more complex form type classifications, unless they are
344 applied only to trees with regular form, where it might be reasonable to assume that form
345 differences captured lower extend all the way to the top of the tree.

346 Here, we experimented with multiple new form-type classes, which recognize that the
347 main stem is embedded in, and an inextricable part of, a complex tree branching network
348 (Barthélémy and Caraglio 2007). Our results suggest that ‘regular’ stem form might even reflect
349 a special case for many species. Again quoting Assmann’s (1970) seminal text: “As compared
350 with the many diverse and, in some cases, bizarre outlines of broad-leaved trees, conifers
351 present regular forms.” Our study generally supports Assmann’s (1970) assertion, with a large
352 amount of data collected over a range of species and forest conditions, although *Pinus strobus*

353 appears to diverge considerably from the general pattern for conifers. This species is a mid-
354 tolerant conifer that often regenerates naturally under an intolerant hardwood canopy (Bebber et
355 al 2004) and tends to have excessive forking due to a white pine weevil (*Pissodes strobe*),
356 which kills terminal shoots and releases branches from apical dominance (Stiel 1979). Thus,
357 while observed differences between needle-leaved and broad-leaved, could be expected due to
358 coarse-level taxonomic differences in apical dominance (Wilson 2000), needle-leaved species
359 may not have regular form under natural forest conditions. For example, Xiong et. al. (2010)
360 showed that genetics is important in determining forking in *Pinus taeda*, but also that local
361 environmental / silvicultural factors tended to explain most of the forking variation. The other two
362 pine species we examined (*P. resinosa* and *P. banksiana*) were drawn mainly from only a few
363 plantations, where spacing and density were kept more uniform during growth, so our data set
364 may actually underestimate how branchy pines are. Certainly, the fact that so many published
365 STM studies of needle-leaved species report excluding forking or branchy trees (e.g., McTauge
366 and Bailey 1987), suggests that such trees are not uncommon in the general population.

367 Our study demonstrates that the new form type classes: LBT, MBT, and MFT, provide a
368 highly useful system for capturing variation in stem form, within and between species, and MFTs
369 appear to explain most of the variation caused by different stand / site conditions. MFTs offer
370 the most flexible system for classifying trees, because they explicitly consider the size of both
371 main stem and branches. Additionally, the MFT coding allows some unusual tree forms to be
372 specifically identified. For example, our data set also included some very large individuals of
373 *Pinus strobus* (up to 100 cm DBH), which had a unusual form type, MFT = 0-1-0-0, which
374 characterizes a tree with only saw-timber-sized sections of the main stem, which rises up to a
375 top that breaks into many small branches (Table 3). A STM would typically predict that a large
376 tree with saw timber lower in the stem, would also have pulp wood higher up in the main stem
377 above it.

378 MFTs might also allow for the possibility of capturing trees with defects, with
379 consideration of the size of the tree. For, example a tree with a code of “0-0-0-0” could either be
380 too small to meet merchantability requirements (< 10 cm DBH) or a large-enough tree (typically
381 ≥ 12 cm DBH by US standards), but with significant defects which limited the merchantability of
382 the section. So, aside from helping to better predict stem volume these new form types can be
383 used to better characterize form diversity in growing stock, which could be related to growth and
384 mortality trends.

385 A potential problem with any classification system is misclassification, but the form type
386 classes described here are easy to implement. In fact, the Michigan DNR already implements a
387 more complex system in their forest volume inventory, in that foresters are already trained to
388 count the number of saw and pulp logs in both the main stem and branches of trees (MDNR
389 2013). MFTs are a simplification, in that those log counts are reduced to a binary, presence or
390 absence assessment for each category of wood products. Consequently, this classification
391 system could be implemented right away in Michigan, USA and a similar system could likely be
392 developed for any forest inventory system that encounters trees of diverse form. Even simpler
393 are the MBT or LBT classes. They provide less information for differentiating trees, but may
394 have lower misclassification error, but could be used in lieu of MFTs, if misclassification error
395 were to outweigh gains of using a more complex form type.

396 One limitation of the form types presented here is that there is only one MFT that can
397 accommodate trees smaller than 10 cm DBH: 0-0-0-0. Previous studies have suggested that
398 small / young trees have different forms than mature trees (Niklas 1991, Ter-Mikaelian et al.
399 2004), but it should be acknowledged that small-tree form can be influenced by branching too.
400 Kerr and Boswell (2001), e.g., surveyed regenerating forests across the United Kingdom and
401 found that 69% of young *Fraxinus excelsior* trees had at least one fork and 29% had more than
402 one. The level of forking varied widely from site to site and was hypothesized to be related to

403 terminal bud damage from frost and bud moths. Kerr and Boswell (2001) also noted that such
404 forks tend to persist, leading to a reduction in the saw timber volume produced in the lower stem
405 of these trees later in life. So, clearly small trees can have complex forms that warrant more
406 than a simple taper modeling approach and the form types of larger trees often reflect the
407 evolution of form types set early in the life of the tree. In combination with other results, this
408 highlights the potential benefits of a more generalized classification approach, where, e.g.,
409 branch size is expressed relative to the size of the main stem, instead of in absolute terms.

410 *4.2. Hierarchical structure of variation in tree form at the tree, stand and species levels*

411 Our data and analyses indicate that tree-to-tree variation in form is enormous over the
412 wide range of species and stand conditions we examined. It was surprising how little
413 ‘evergreenness’ and species mattered as grouping variables, in light of this variation. The
414 practical implication is that if a tree is randomly drawn from the landscape and measured for
415 stem form and volume, knowing whether it is a needle leaved vs. broad-leaved tree, or even its
416 species, does not tell us much more about its form. By contrast, knowing the particular stand it
417 was drawn from tells us quite a lot, because the ecological conditions and historic stand
418 dynamics have an apparently large effect on how that tree grew. This supports the idea that
419 species-specific volume models are not superior to ‘composite’ or mixed species ones, because
420 so-called species effects are often confounded with effects of the stand conditions from which
421 the species data were drawn; “taper is greatly influenced by the past history of the stand”
422 (Gevorkiantz and Olsen 1955). Here, we used mixed effects modeling to disarticulate species
423 and stand effects and used our new form types to capture growth-history effects at the individual
424 tree level, which explained most of the stand-level influence on tree form.

425 Since MFTs did not capture all the stand-to-stand variation in stem form, our hierarchical
426 analysis indicated other residual uncertainty from stand level factors. Stand density should

427 affect taper to some degree and also act to reduce lateral branching when density is high
428 (Nielsen and Gerrand 1999). Garber and Maguire (2003) showed that including height-diameter
429 ratios as a model predictor could to some extent account for stand to stand variation in stem
430 taper, because it tends to be correlated with stand density and it is a good proxy for crown ratio,
431 at least in stands with simple crown architecture. However, they also concluded that spacing
432 had little appreciable influence on the taper of some species (e.g., *Abies grandis*) and
433 underscored the importance of developing site and species-specific equations. By contrast, our
434 goal was to reduce the need for stand / site specific equations. Muhairwe et al. (1994) explored
435 the explanatory power of site productivity class and age into STMs developed for several
436 species, but concluded that the variables contributed only marginally to improving the model.
437 Stand-to-stand / site-to-site variation in stem form may be confounded with other factors, such
438 as local genetic effects on taper (Sherill et al. 2004, Gomat et al. 2011), but we would expect
439 these to be evident in the form type and species combinations represented in the stand.
440 Ultimately, it simply may not be possible to explain all the stand-to-stand variation in tree form.

441 *4.3. Implications for development and application of STMs to trees of diverse form and*
442 *branching architecture.*

443 The results of this study suggest that STM approaches would benefit more from fitting
444 them by form type, then by species, but that it would be most useful to fit MFT-specific models
445 for every species. In terms of this study, that would mean filling a data matrix of trees covering
446 21 species by 16 MFTs. Over even larger spatial domains, such as a national forest inventory,
447 fitting a set of form-by-species models would require an enormous data set, where trees with a
448 variety of non-regular forms were sought out. Such data sets are currently rare and will likely
449 remain so given the high cost of sampling for stem taper/volume.

450 Instead of trying to fit a model for every combination of species and form type, we chose
451 to use a universal hierarchical modeling approach, which could draw from the strength of all our
452 data. This method also respects and accounts for within-tree and within-stand variation in taper
453 measurements, by treating them as deviation from the general population (i.e. random effects).
454 After accounting for this hierarchical variation, our universal model predicts how each tree
455 diverges from the all-tree trend, based on its species and MFT. This approach bears similarities
456 to the STM approach of Cao and Wang (2011), except that an actual stem measurement was
457 needed at the mid-point of the stem to localize the model for each tree under their approach.
458 Indeed, calibrating taper models in a mixed-effects modeling framework is a promising new
459 approach (Sabatia and Burkhart 2015), but often relies on upper stem diameter measurements
460 which can be difficult to measure accurately. Our approach, in a sense, estimates the random
461 divergence of members of a species from the all-species trend, but also independently captures
462 the average form deviation of a tree of any species, depending on a simple assignment of trees
463 to categories related to the relative size of the main stem and branches. This novel approach
464 appears to be quite effective for exploring intra- and inter-specific variation in stem form, but
465 likely needs more testing before it is implemented operationally for forest volume estimation.
466 Certainly, there appears to be an opportunity to further explore stand-level predictors of form,
467 with the opportunity to account for tree-level differences, captured by form type classes.

468 Another major implication of this study is that the exclusion of trees without 'regular' form
469 represents a significant omission from the STM literature, and also likely translates into a
470 significant bias in many published STMs and most historical data sets relating to stem taper.
471 Our results show that STMs based on censored data will overestimate main stem volume for the
472 general population of trees, with the magnitude of the bias dependent upon what proportion of
473 trees in the general population are 'irregular' in form. If STMs are used to compute stem
474 volumes that are later extrapolated to whole-tree mass estimates, e.g., biomass expansion

475 factors (Lehtonen et al. 2004, Jalkanen et al. 2005, Skovsgaard and Nord-Larsen 2012, Domke
476 et al. 2012), censoring bias could significantly skew forest biomass and carbon stock inventories
477 estimated from stem volume. Overestimation of main stem volume would presumably then
478 cause over-estimation in whole-tree biomass (Lehtonen et al. 2004). Biomass expansion
479 factors should be particularly sensitive to the relative proportion of volume in the main stem
480 versus the branches (MacFarlane 2011).

481 Given the high degree of plasticity in tree morphology, despite species-specific
482 constraints on branching architecture (Dardick et al. 2013, MacFarlane 2015), and given
483 significant influences of silvicultural practices (e.g., thinning, Assmann 1970) and environment
484 (e.g., frost, bud moths, Kerr and Boswell 2001) on tree form, it is not unreasonable to suppose
485 that the idealized trees that modelers seek for STM development may not actually represent the
486 typical tree encountered during forest inventory. This means that for STMs fit with exclusions to
487 be applied without bias, trees would need to be categorized as belonging to, or not belonging to,
488 the exclusion group. We recommend that future studies should eschew exclusion criteria and
489 expand taper data sets to include so-called irregular trees. This would allow model developers
490 to either fit a general STM to trees of all forms, or take an approach, such as the one described
491 here, to categorize form differences between different types of trees. Minimally, one could fit the
492 taper model to a population of “included” vs. “excluded” trees and report the difference.

493

494 **Acknowledgements**

495 The authors would like to thank the Michigan Department of Natural Resources (MDNR)
496 for contributing significant resources to amassing the data used in this study. In particular Doug
497 Heym, who has been an advocate for innovation in forest inventory in Michigan. Part of DWM's
498 time was supported with funds from Michigan AgBioResearch, through the USDA National

499 Institute of Food and Agriculture, with additional support from the MDNR. A portion of ARW's
500 time was supported by the Maine Agricultural and Forest Experimental Station. Both DWM and
501 ARW received support from the USDA Forest Service, Northern Research Station, Forest
502 Inventory and Analysis Program, NRS-05 (FIA).

503 **References**

504 Adu-Bredu, S., Bi. A.F.T., Bouillet, J-P., Me, M.K., Kyei, S.Y., and Saint-Andre, L. 2008. An
505 explicit stem profile model for forked and un-forked teak (*Tectona grandis*) trees in West Africa.
506 *Forest Ecology and Management*, 255: 2189–2203.

507 Assmann, E. 1970. The Principles of Forest Yield Study. Pergamon Press, Oxford, U.K. 506. p.

508 Barrio Anta, M., Diéguez-Aranda, U., Castedo-Dorado, F., González, J.G.A., and von Gadow, K.
509 2007. Merchantable volume system for pedunculate oak in northwestern Spain. *Annals of*
510 *Forest Science*, Volume 64 (5): 511-520.

511 Barthelemy, D. and Caraglio, Y. 2007. Plant architecture: a dynamic, multilevel and
512 comprehensive approach to plant form, structure and ontogeny. *Ann. Bot.* 99, 375–407.

513 Bebber, D.P., Thomas, S.C., Cole, W.G., Balsillie, D. 2004. Diameter increment in mature
514 eastern white pine *Pinus strobus* L. following partial harvest of old-growth stands in Ontario,
515 Canada. *Trees*, 18:29–34.

516 Bravo-Oviedo A, del Río M, Calama R, Valentine HT. 2014. New approaches to modelling
517 cross-sectional area to height allometry in four Mediterranean pine species. *Forestry*, 87:399–
518 406.

519 Burkhart, H. E. 1977. Cubic foot volume of loblolly pine to any merchantable top limit. *Southern*
520 *Journal of Applied Forestry*, 1: 7-9.

- 521 Burkhart , H.E.and Walton, S.B. 1985. Incorporating crown ratio into taper equations for loblolly
522 pine trees. *Forest Science*, 31, 478 – 484.
- 523 Cao, Q.V., and J. Wang. 2011. Calibrating fixed- and mixed-effects taper equations. *Forest*
524 *Ecology and Management*, 262: 671–673.
- 525 Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C,
526 Duque, A., Eid, T., Fearside, P.M., Goodman, R.C., Henry, M., Martínez-Yrizar, A., Mugasha,
527 W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, A., Ortiz-
528 Malavassi, E., Péliissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G. 2014.
529 Improved allometric models to estimate the aboveground biomass of tropical trees. *Global*
530 *Change Biology*, 20: 3177-3190.
- 531 Clark, Alexander, III; Souter, Ray A.; Schlaegel, Bryce E. 1991. Stem Profile for Southern
532 Equations for Southern Tree Species. Res. Pap. SE-282. Asheville, NC: U.S. Department of
533 Agriculture, Forest Service, Southeastern Forest Experiment Station. 117 p.
- 534 Dardick, C., Callahan, A., Horn, R., Ruiz, K.B., Zhebentyayeva, T., Hollender, C., Whitaker, M.,
535 Abbott, A., and Scorza, R. 2013. PpeTAC1 promotes the horizontal growth of branches in peach
536 trees and is a member of a functionally conserved gene family found in diverse plants species.
537 *The Plant Journal*, 75: 618–630.
- 538 Domke, G.M., Woodall, C.W., Smith, J.E., Westfall, J.A., McRoberts, R.E., 2012. Consequences
539 of alternative tree-level biomass estimation procedures on U.S. forest carbon stock estimates.
540 *Forest Ecology and Management*, 270, 108-116.
- 541 Ducey, M.J. 2012. Evergreenness and wood density predict height–diameter scaling in trees of
542 the northeastern United States. *Forest Ecology and Management*, 279, 21-26.

- 543 Forslund, R.R. 1982. A geometrical volume model based on the location of the centre of
544 gravity of the bole. *Canadian Journal of Forest Research*, 12: 215-221.
- 545 Neilsen, W.A., and Gerrand, A.M. 1999. Growth and branching habit of *Eucalyptus nitens* at
546 different spacing and the effect on final crop selection. *Forest Ecology and Management*, 123 :
547 217-229.
- 548 Gevorkiantz, S.R., and Olsen, L.P. 1955. Composite volume tables for timber and their
549 application in the Lake States. Technical Bulletin No. 1104, Jan., 1955, USDA, Washington,
550 D.C.
- 551 Gomat, H.Y., Deleporte P., Moukini R., Mialounguila G., Ognouabi N., Saya A.R., Vigneron P.,
552 Saint André L.. 2011. What factors influence the stem taper of *Eucalyptus*: Growth,
553 environmental conditions, or genetics? *Annals of Forest Science*, 68: p. 109-120.
- 554 Jalkanen, A., Mäkipääb, R., Ståhl, Göran, Lehtonen, 2005. A. and Petersson, H. Estimation of
555 the biomass stock of trees in Sweden: comparison of biomass equations and age-dependent
556 biomass expansion factors. *Annals or Forest Science*, 62: 845–851.
- 557 Jordan, L., Souter, R., Parresol, B., and Daniels, R.F. 2006. Application of the algebraic
558 difference approach for developing self-referencing specific gravity and biomass equations.
559 *Forest Science*, 52: 81–92.
- 560 Kerr, G.and Boswell , R.C. 2001 The influence of spring frosts, ash bud moth (*Prays fraxinella*)
561 and site on forking of young ash (*Fraxinus excelsior*) in southern Britain. *Forestry*. 74, 29 – 40 .
- 562 Kozak, A, Munro, D.D. and Smith, J. H. G. 1969. Taper functionsa and their application in forest
563 inventory. *Forestry Chronicle*, 45: 1-6.

- 564 Kublin E, Breidenbach (2013) TapeR—Flexible tree taper curves based on semiparametric
565 mixed models. R-package version 0.3.0. Available on CRAN: [http://cran.r-](http://cran.r-project.org/web/packages/TapeR/)
566 [project.org/web/packages/TapeR/](http://cran.r-project.org/web/packages/TapeR/).
- 567 Lehtonen, A., Cienciala, E., Tatarinov, F., and Mäkipää, R. 2007. Uncertainty estimation of
568 biomass expansion factors for Norway spruce in the Czech Republic. *Annals of Forest Science*,
569 64: 133-140.
- 570 MacFarlane, D.W. 2010. Predicting branch to bole volume scaling relationships from varying
571 centroids of tree bole volume. *Canadian Journal of Forest Research*, 40: 2278–2289.
- 572 MacFarlane, D.W. 2015. A generalized tree component biomass model derived from principles
573 of variable allometry. *Forest Ecology and Management*, 354: 43-55.
- 574 MacFarlane, D.W. 2011. Allometric scaling of branch volume in hardwood trees in Michigan,
575 USA: implications for improvements in above-ground forest carbon biomass inventories. *Forest*
576 *Science*, 57: 451– 459.
- 577 MacFarlane, D.W., Kuyah, S., Mulia, R., Dietz, J., Muthuri, C., and Van Noordwijk, M. 2014.
578 Evaluating a non-destructive method for calibrating tree biomass equations derived from tree
579 branching architecture. *Trees*, 28: 807–817.
- 580 Max, T.A. and Burkhart, H.E. 1976. Segmented polynomial regression applied to taper
581 equations. *Forest Science*, 22, 283–289.
- 582 McTague, J.P., and Bailey, R.L. 1987. Simultaneous total and merchantable volume equations
583 and a compatible taper function for loblolly pine. *Canadian Journal of Forest Research*, 17: 87-
584 92.

- 585 MDNR. 2013. Product Standards and Cruising Manual. IC4057, Michigan Department of Natural
586 Resources, Forest Resources Division. 07/03/2013. 44 p.
- 587 Muhairwe, C.K. 1994. Tree form and taper variation over time for interior lodgepole pine.
588 *Canadian Journal of Forest Research*, 24:1904-1913.
- 589 Muhairwe, C.K., LeMay, V.M., and Kozak, A. 1994. Effects of adding tree, stand, and site
590 variables to Kozak's variable-exponent taper equation. *Canadian Journal of Forest Research*,
591 24: 252–259.
- 592 Newnham, R.M. 1992. Variable form taper functions for four Alberta tree species. *Canadian*
593 *Journal of Forest Research*, 22: 210-223.
- 594 Niklas, K.J. 1995. Size-dependent allometry of tree height, diameter and trunk-taper. *Annals of*
595 *Botany*, 75: 217–227.
- 596 Ozcelik, R. 2008. Comparison of formulae for estimating tree bole volumes of *Pinus sylvestris*.
597 *Scandinavian Journal of Forest Research*, 23(5): 412-418.
- 598 Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2015). nlme: Linear and nonlinear
599 mixed effects models. R package version 3.1-122, <http://CRAN.R-project.org/package=nlme>.
- 600 Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
601 language. *Bioinformatics*, 20: 289-290.
- 602 R Development Core Team. 2015. R: a language and environment for statistical computing. R
603 Foundation for Statistical Computing, Vienna, Austria.
- 604 Sabatia, C.O., and Burkhart, H. E. 2015. On the use of upper stem diameters to localize a
605 segmented taper equation to new trees. *Forest Science*, 61: 411-423.

- 606 Skovsgaard, J.P. and Nord-Larsen, T. 2012. Biomass, basic density and biomass expansion
607 factor functions for European beech (*Fagus sylvatica* L.) in Denmark. *European Journal of*
608 *Forest Research*, 131:1035–1053.
- 609 Stiehl W.M. 1979. Releasing unweevilled white pine to ensure first log quality of final crop.
610 *Forestry Chronicle*, 55:142–143.
- 611 Sherill, J.R. Mullin, T.R., Bullock, B.P., McKeand, S.E. and Purnell, R.C. 2004. Forest Genetics
612 and Tree Breeding in the Age of Genomics: Progress and Future, IUFRO Joint Conference of
613 Division 2, Conference Proceedings. 1-5 November 2004, edited by B. Li and S.E. McKeand,
614 Volume: Pp. 457
- 615 Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S (Fourth Edition). New
616 York: Springer-Verlag.
- 617 Valentine, H.T., and Gregoire, T.G. 2001. A switching model of bole taper. *Canadian Journal of*
618 *Forest Research*, 31: 1400-1409.
- 619 Ver Planck N.R., MacFarlane, D.W. 2015. A vertically integrated whole-tree biomass Model.
620 *Trees*, 29: 449–460.
- 621 Ver Planck N.R., MacFarlane, D.W. 2014. Modelling vertical allocation of tree stem and branch
622 volume for hardwoods. *Forestry*; 00, 1–16 doi:10.1093/forestry/cpu007.
- 623 Wilson, B.F. 2000. Apical control of branch growth and angle in woody plants. *American Journal*
624 *of Botany* 87(5): 601–607.
- 625 Xiong, J.S., Isik, F., McKeand, S.E., and Whetten. R.W. 2010. Genetic variation of stem forking
626 in loblolly pine. *Forest Science*, 56(5):429–436.

627 Zakrzewski, W.T. 2011 Estimating wood volume of the stem and branches of sugar maple (*Acer*
628 *saccharum* marsh.) using a stem profile model with implicit height. *Forest Science*, 57, 117–133.

629 Zakrzewski, W.T., and Duchesne, I. 2012. Stem biomass model for jack pine (*Pinus banksiana*
630 Lamb.) in Ontario. *Forest Ecology and Management*, 279:112–120, 2012.

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

Draft

Table 1. Study tree attributes, listed by species.

species	# stands	# trees	mean DBH (cm) [sd, min, max]	mean Height (m) [sd, min, max]
<i>Abies balsamea</i>	1	37	20.7 [5.7, 11.9, 39.4]	16.3 [3.5, 11.6, 25.6]
<i>Picea glauca</i>	1	2	35.8 [14.3, 18.0, 46.7]	20.5 [6.2, 12.9, 25.2]
<i>Pinus banksiana</i>	2	152	20.6 [5.7, 11.7, 41.4]	13.2 [2.3, 6.9, 20.1]
<i>Pinus resinosa</i>	2	198	35.2 [4.4, 17.5, 48.3]	22.2 [1.9, 12.7, 26.5]
<i>Pinus strobus</i>	5	56	51.9 [25.7, 14.7, 101.6]	23.8 [7.0, 12.0, 34.0]
All Needle-leaved	7	445	30.3 [14.8, 11.7, 101.6]	19.2 [5.5, 6.9, 34.0]
<i>Acer rubrum</i>	14	219	33.0 [12.2, 11.9, 69.6]	21.6 [3.7, 8.9, 32.1]
<i>Acer saccharinum</i>	1	2	37.0 [6.6, 30.2, 43.2]	24.9 [1.5, 23.3, 26.2]
<i>Acer saccharum</i>	14	397	35.3 [11.8, 11.9, 73.7]	23.7 [3.2, 13.1, 31.7]
<i>Amelanchier arborea</i>	1	1	34.3 [—, 34.3, 34.3]	23.4 [—, 23.4, 23.4]
<i>Betula alleghaniensis</i>	4	46	33.5 [12.2, 16.5, 63.8]	20.7 [2.9, 12.0, 27.5]
<i>Betula papyrifera</i>	6	70	27.3 [9.1, 12.4, 57.4]	21.3 [2.8, 13.2, 26.2]
<i>Fagus grandifolia</i>	7	96	32.7 [12.0, 11.7, 57.2]	20.8 [3.9, 12.3, 30.4]
<i>Fraxinus americana</i>	9	80	38.4 [14.7, 15.0, 77.7]	26.5 [3.9, 16.6, 34.5]
<i>Ostrya virginiana</i>	1	2	24.4 [1.7, 22.9, 26.2]	19.2 [0.3, 18.9, 19.6]
<i>Populus balsamifera</i>	1	2	30.6 [3.6, 27.2, 34.3]	24.0 [1.3, 22.8, 25.3]
<i>Populus grandidentata</i> and <i>P. tremuloides</i> ¹	5	113	29.0 [7.6, 14.7, 54.6]	23.5 [3.0, 15, 31.7]
<i>Prunus serotina</i>	9	36	30.7 [7.2, 15.2, 43.9]	22.1 [3.3, 11.2, 26.7]
<i>Quercus alba</i>	4	60	35.7 [9.2, 11.9, 58.7]	21.8 [3.2, 7.8, 25.7]
<i>Quercus ellipsoidalis</i>	3	9	34.2 [6.6, 23.4, 44.7]	16.7 [3.1, 12.6, 22.9]
<i>Quercus rubra</i>	5	112	33.5 [9.2, 13.2, 56.9]	24.7 [3.6, 14.8, 32.4]
<i>Tilia americana</i>	9	138	37.2 [9.8, 15.2, 78.5]	24.3 [3.2, 13.2, 30.2]
All broad-leaved	29	1383	32.9 [11.4, 11.7, 78.5]	23.2 [3.7, 7.8, 34.5]
All trees	31	1828	32.3 [7.1, 11.7, 101.6]	21.6 [4.7, 6.9, 34.5]

¹these two *Populus* species were not consistently differentiated in the field so they were combined.

Table 2. Study trees classified into Large Branch Types (LBT) and Merchantable Branch-Types (MBT), listed by species.

species	# trees	Large Branch Type (LBT)		% in LBT		Merchantable Branch Type (MBT)			
		0	1	0	1	0-0	0-1	1-0	1-1
<i>Abies balsamea</i>	37	37	—	100%	0%	37	—	—	—
<i>Picea glauca</i>	2	2	—	100%	0%	2	—	—	—
<i>Pinus banksiana</i>	152	149	3	98%	2%	149	—	3	—
<i>Pinus resinosa</i>	198	194	4	98%	2%	194	—	3	1
<i>Pinus strobus</i>	56	48	8	86%	14%	48	—	2	6
All conifers	445	430	15	97%	3%	430	—	8	7
<i>Acer rubrum</i>	219	127	92	58%	42%	127	1	69	22
<i>Acer saccharinum</i>	2	1	1	50%	50%	1	—	1	—
<i>Acer saccharum</i>	397	161	236	41%	59%	161	—	210	26
<i>Amelanchier arborea</i>	1	—	1	0%	100%	—	—	1	—
<i>Betula alleghaniensis</i>	46	18	28	39%	61%	18	1	20	7
<i>Betula papyrifera</i>	70	47	23	67%	33%	47	1	19	3
<i>Fagus grandifolia</i>	96	63	33	66%	34%	63	1	28	4
<i>Fraxinus americana</i>	80	36	44	45%	55%	36	—	27	17
<i>Ostrya virginiana</i>	2	2	—	100%	0%	2	—	—	—
<i>Populus balsamifera</i>	2	2	—	100%	0%	2	—	—	—
<i>Populus grandidentata</i> and <i>P. tremuloides</i> ¹	113	102	11	90%	10%	102	—	11	—
<i>Prunus serotina</i>	36	18	18	50%	50%	18	—	17	1
<i>Quercus alba</i>	60	12	48	20%	80%	12	—	38	10
<i>Quercus ellipsoidalis</i>	9	7	2	78%	22%	7	—	1	1
<i>Quercus rubra</i>	112	40	72	36%	64%	40	—	52	20
<i>Tilia americana</i>	138	80	58	58%	42%	80	1	49	8
All hardwoods	1383	716	667	52%	48%	716	5	543	119
All trees	1828	1146	682	63%	37%	1146	5	551	126

¹these two *Populus* species were not consistently differentiated in the field so they were combined.

Table 3. Study trees classified into Merchantable Form Types (MFT), listed by species.

species	# trees	Merchantable Form Type (MFT)									
		0-0-0-0	0-1-0-0	0-1-1-0	0-1-1-1	1-0-0-0	1-0-1-0	1-1-0-0	1-1-0-1	1-1-1-0	1-1-1-1
<i>Abies balsamea</i>	37	—	—	—	—	36	—	1	—	—	—
<i>Picea glauca</i>	2	—	—	—	—	1	—	1	—	—	—
<i>Pinus banksiana</i>	152	1	—	—	—	126	—	22	—	3	—
<i>Pinus resinosa</i>	198	—	—	—	—	75	1	119	—	2	1
<i>Pinus strobus</i>	56	—	16	—	1	16	—	16	—	2	5
All needle-leaved	445	1	16	—	1	254	1	159	—	7	6
<i>Acer rubrum</i>	219	—	—	1	1	67	16	60	1	52	21
<i>Acer saccharinum</i>	2	—	—	—	—	—	—	1	—	1	—
<i>Acer saccharum</i>	397	1	1	1	2	94	26	65	—	183	24
<i>Amelanchier arborea</i>	1	—	—	—	—	—	—	—	—	1	—
<i>Betula alleghaniensis</i>	46	—	—	—	4	9	—	9	1	20	3
<i>Betula papyrifera</i>	70	—	—	—	—	25	1	22	1	18	3
<i>Fagus grandifolia</i>	96	—	8	1	—	33	2	22	1	25	4
<i>Fraxinus americana</i>	80	—	—	—	—	11	2	25	—	25	17
<i>Ostrya virginiana</i>	2	—	—	—	—	2	—	—	—	—	—
<i>Populus grandidentata</i> and <i>P. tremuloides</i> ¹	113	—	5	1	—	50	1	47	—	9	—
<i>Populus balsamifera</i>	2	—	—	—	—	1	—	1	—	—	—
<i>Prunus serotina</i>	36	—	—	—	—	9	3	9	—	14	1
<i>Quercus alba</i>	60	—	—	1	—	6	1	6	—	36	10
<i>Quercus ellipsoidalis</i>	9	—	—	—	—	2	—	5	—	1	1
<i>Quercus rubra</i>	112	—	—	—	1	17	—	23	—	52	19
<i>Tilia americana</i>	138	—	2	2	1	21	2	57	1	45	7
All broad-leaved	1383	1	16	7	9	347	54	352	5	482	110
All trees	1828	2	32	7	10	601	55	511	5	489	116
% trees		0.1%	1.8%	0.4%	0.5%	32.9%	3.0%	28.0%	0.3%	26.8%	6.3%

¹these two *Populus* species were not consistently differentiated in the field so they were combined.

Table 4. Model fit statistics for the standard Max and Burkhart (1976) taper model with two free joining points (eq. 1) and a DBH-segmented version that fixes one joining point at the mean relative breast height for the population (eq. 2).

Model	df	AIC	BIC	log.lik.	like.ratio	p-value
eq. 1	9	-16203	-16133.54	8110.775		
eq. 2	8	-26587	-26525.39	13301.807	10382.07	< 0.0001

Draft

Table 5. Fit statistics using the Max and Burkhart (1976) taper model with various random effects on eq. 2 coefficients. The best model is presented in bold.

Coef.	AIC	BIC	LogLik
β_1	-16980	-16918	8498
β_2	-16647	-16585	8331
β_3	<i>below step halving factor</i>		
β_4	-26587	-26525	13301
α_1	<i>below step halving factor</i>		
$\beta_1 + \beta_4$	-29392	-29299	14708
$\beta_2 + \beta_4$	-30395	-30302	15209
$\beta_1 + \beta_2 + \beta_4$	-33034	-32894	16535

Draft

Table 6. Model coefficients with standard error in parentheses and fit statistics for eq. 3 (ME = mean error, MAE = mean absolute error, RMSE = root mean square error). *Abies balsamea* with Merchantable Form Type (MFT) 0-0-0-0 was used as the reference level (Intercept) for comparing random effects of the other species and MFTs to.

	ME	MAE	RSME
eq. 3	0.0005	0.0726	0.1419
Fixed effects	β_1	β_2	β_3
	-3.9846 (0.2055)	2.0105 (0.1242)	-1.865 (0.1093)
	β_4	α_1	α_2
	138.4753 (7.7222)	0.7001 (0.0124)	0.0673 (0.0001)
Random effects			
	μ_1	μ_2	μ_4
Intercept	0.0710 (0.1927)	-0.0785 (0.1903)	15.0346 (90.9481)
MFT	μ_1	μ_2	μ_4
0-0-0-0	—	—	—
0-1-0-0	0.2305 (0.1915)	-0.2163 (0.1896)	6.7301 (91.1553)
0-1-1-0	-0.0294 (0.2079)	0.0192 (0.2059)	40.2638 (99.1691)
0-1-1-1	-0.1827 (0.2026)	0.1888 (0.2005)	-80.5889 (96.3921)
1-0-0-0	0.2306 (0.1833)	-0.2228 (0.1815)	45.3814 (87.4401)
1-0-1-0	-0.1714 (0.1877)	0.1777 (0.1858)	-13.1429 (89.4219)
1-1-0-0	0.1748 (0.1837)	-0.1652 (0.1819)	11.8544 (87.6034)
1-1-0-1	0.3186 (0.2189)	-0.3000 (0.2166)	1.1661 (104.0487)
1-1-1-0	0.0183 (0.1837)	-0.0147 (0.1819)	-10.3019 (87.6118)
1-1-1-1	-0.2473 (0.1847)	0.2461 (0.1830)	-14.9249 (88.1492)

species	μ_1	μ_2	μ_4
<i>Abies balsamea</i>	—	—	—
<i>Picea glauca</i>	-0.1740 (0.1858)	0.1781 (0.1840)	-3.7965 (88.8311)
<i>Pinus banksiana</i>	-0.1447 (0.0807)	0.1364 (0.0788)	-46.5734 (34.6997)
<i>Pinus resinosa</i>	-0.0192 (0.1077)	0.0192 (0.1032)	-59.9879 (41.3147)
<i>Pinus strobus</i>	-0.3074 (0.0825)	0.312 (0.0806)	-61.4255 (35.8098)
<i>Acer rubrum</i>	-0.3478 (0.0585)	0.3472 (0.0576)	-52.9227 (26.8991)
<i>Acer saccharinum</i>	-0.5638 (0.1950)	0.5741 (0.1930)	-185.2351 (92.4421)
<i>Acer saccharum</i>	-0.2317 (0.0564)	0.2271 (0.0557)	-9.9501 (26.1222)
<i>Amelanchier arborea</i>	-0.4268 (0.2638)	0.4318 (0.2613)	-154.4622 (125.8714)
<i>Betula alleghaniensis</i>	-0.3238 (0.0788)	0.3097 (0.0775)	75.5961 (35.7510)
<i>Betula papyrifera</i>	-0.0152 (0.0655)	0.0189 (0.0646)	7.9239 (30.1749)
<i>Fagus grandifolia</i>	-0.1753 (0.0681)	0.1694 (0.0670)	17.6663 (30.9635)
<i>Fraxinus americana</i>	-0.0883 (0.0655)	0.0946 (0.0645)	-24.8026 (30.0410)
<i>Ostrya virginiana</i>	-0.2719 (0.1916)	0.2549 (0.1897)	211.1978 (91.2025)
<i>Populus balsamifera</i>	-0.0299 (0.1858)	0.0458 (0.1840)	-116.2716 (88.8311)
<i>Populus grandidentata</i> and <i>P. tremuloides</i> ¹	0.1431 (0.0545)	-0.1294 (0.0538)	-67.1723 (25.5528)
<i>Prunus serotina</i>	-0.0344 (0.0764)	0.0392 (0.0752)	-58.874 (34.9918)
<i>Quercus alba</i>	-0.3668 (0.0708)	0.3648 (0.0697)	6.1543 (32.5058)
<i>Quercus ellipsoidalis</i>	-0.1890 (0.1145)	0.1716 (0.1126)	-16.5648 (52.2819)
<i>Quercus rubra</i>	-0.1243 (0.0650)	0.1272 (0.0640)	16.3635 (29.9453)
<i>Tilia americana</i>	-0.0098 (0.0560)	0.0230 (0.0553)	-38.9422 (26.2258)

Figure captions

Figure 1. Relative components of variation in observed main stem outside bark volume from stem taper data from different grouping variables. Variance components are relative to tree to tree variation which describes 100% of the variation (variance component = 1) in the data used to fit the model. (a) shows a case when form types are omitted and (b-d) show the contribution of adding LBT (Large Branch Types), MBT (Merchantable Branch Types) and MFT (Merchantable Form Types) as grouping variables.

Figure 2. A scatter plot of the relative cross-sectional area outside bark (*relcob*) at different relative height locations (*relh*), from 1828 trees of 21 species (light gray circles), with data from *Quercus rubra* (*Quru*) trees superimposed (dark gray circles). The predicted taper curves from DBH-segmented polynomial stem taper model with variable random effects are shown for *Quru* for four different MFTs (Merchantable Form Types), representing a range of forms. The solid line is the fixed effects model for the entire population of trees, from which *Quru* trees of differing form diverge (eq. 3).

Figure 3. A scatter plot of the relative cross-sectional area outside bark (*relcob*) at different relative height locations (*relh*), with data from *Acer saccharum* (*Acsa*) (light gray circles) *Tilia americana* (*Tiam*) trees superimposed (asterisks). The predicted taper curves from DBH-segmented polynomial stem taper model with variable random effects are shown for *Acsa* and *Tiam* for two different MFTs (Merchantable Form Types): 1-0-0-0 and 1-0-1-0. The solid line is the fixed effects model for the entire population of trees, from which each diverges based on predicted random effects (eq. 3).

Fig. 1

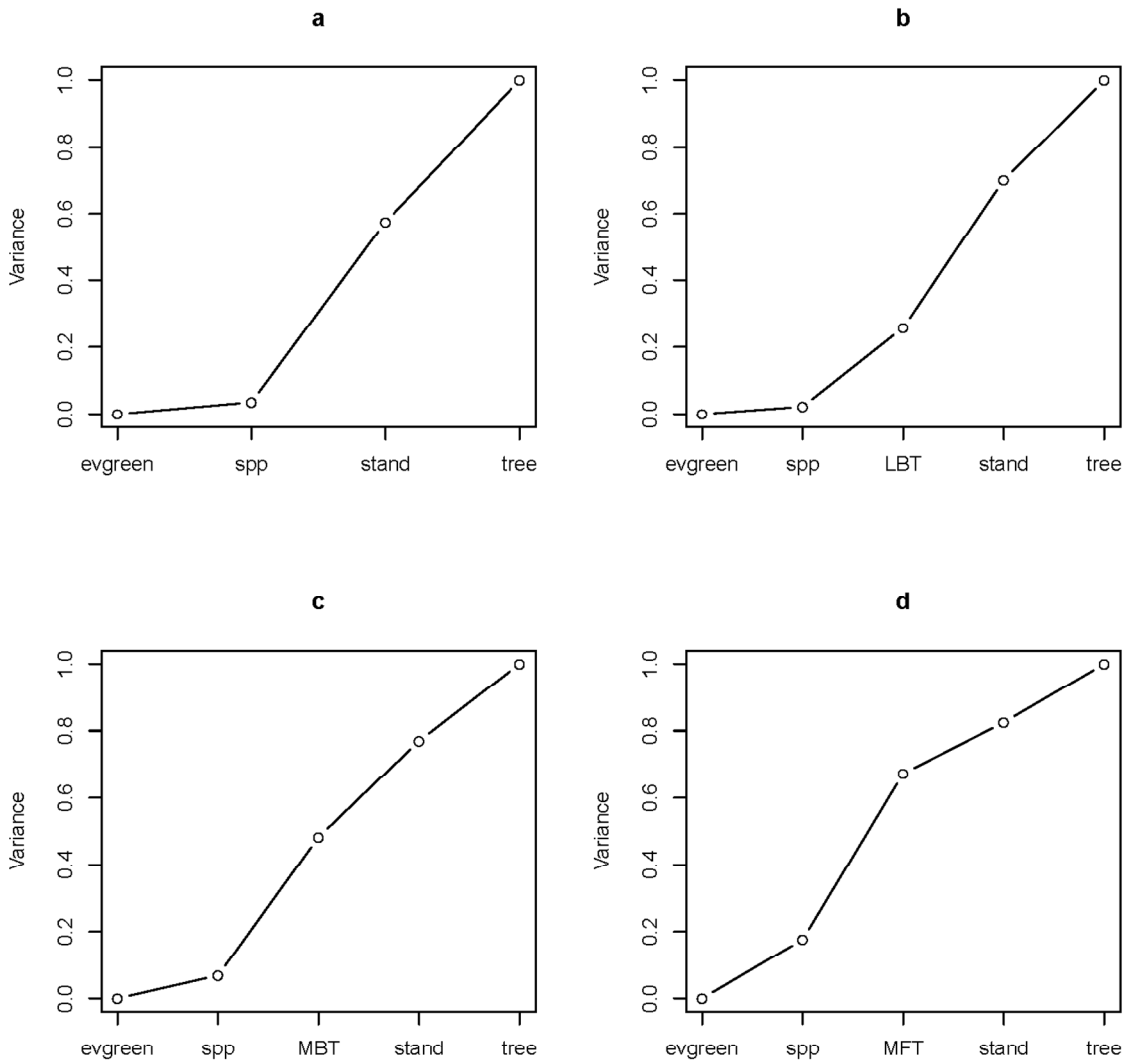


Fig. 2

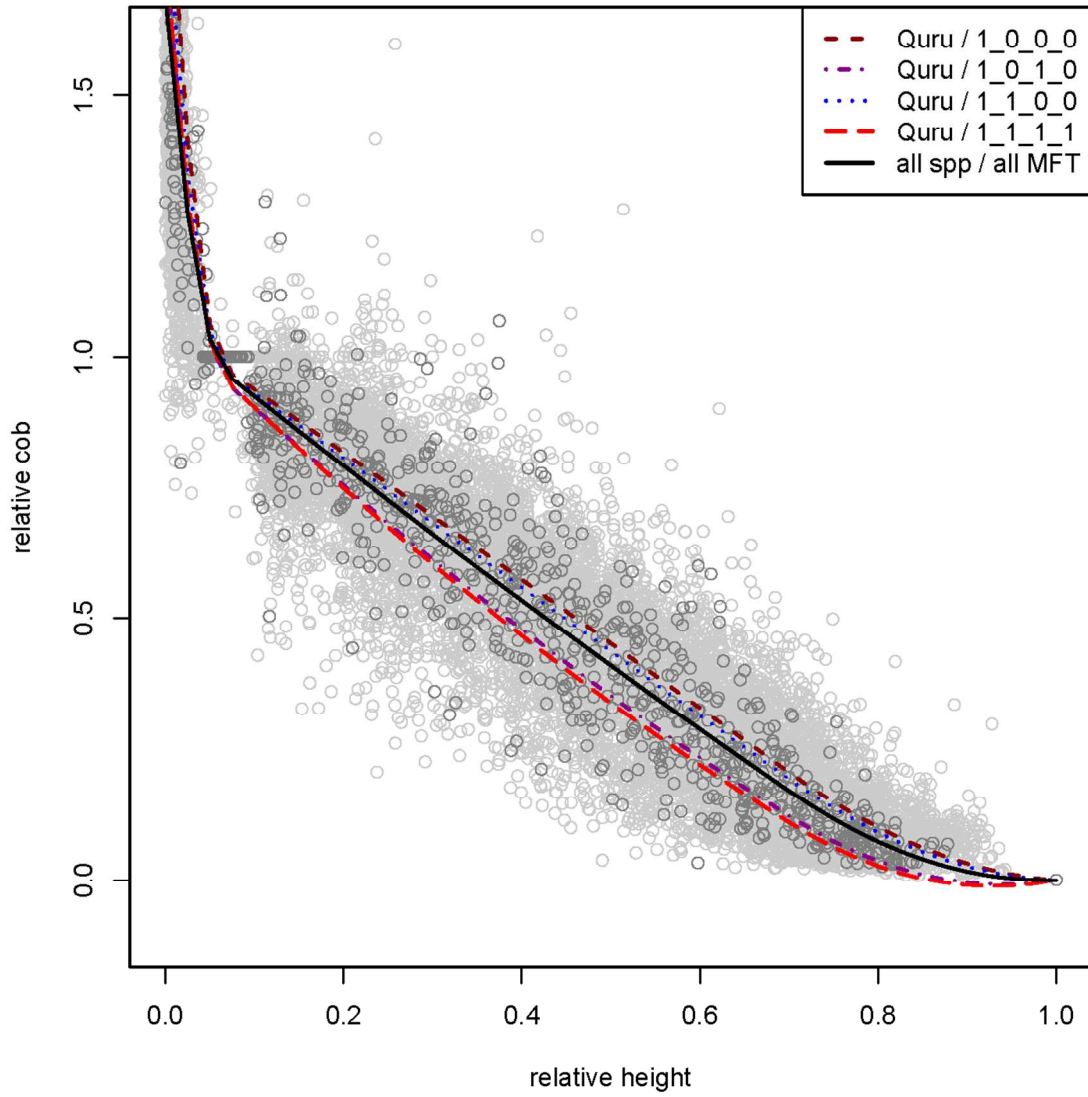


Fig. 3

