Minna Pulkkinen

On Non-Circularity of Tree Stem Cross-Sections: Effect of Diameter Selection on Cross-Section Area Estimation, Bitterlich Sampling and Stem Volume Estimation in Scots Pine
Abstract


In the common methods of forest mensuration, including stem volume models and Bitterlich sampling, stem cross-sections are assumed to be circular. In nature this assumption is never exactly fulfilled. Errors due to non-circularity have been presumed to be small and unimportant but studied little: theoretical and empirical studies exist on cross-section area estimation, but errors in stem volume estimation have not been investigated at all, and errors in Bitterlich sampling are theoretically known only for stand basal area estimation. In the theoretical part of this study, we developed methods for quantifying the systematic and sampling errors that 22 common ways of selecting diameter within non-circular cross-sections induce (i) in area estimates by the circle area formula, (ii) in stand total estimates by Bitterlich sampling, and (iii) in stem volume estimates by a volume equation, by a cubic-spline-interpolated stem curve, and by a generalised volume estimator. In the empirical part, based on the digital images of 709 discs taken at 6–10 heights in 81 Scots pine stems from different parts of Finland, we investigated the variation in cross-section shape, and demonstrated the magnitude of the errors presented in the theoretical part. We found that non-circularity causes systematic overestimation of area and volume, and inflicts potentially systematic error on stand total estimates by Bitterlich sampling. In our data these effects were small, but the finding is not generalisable due the skewed size distribution and poor geographical representativeness of the data. We recommend using diameter derived from girth for both tree and stand level estimation, as it involves no sampling error and produces clearly the most stable systematic errors.

Keywords forest mensuration, cross-section, non-circularity, basal area, stem volume, Bitterlich sampling, Scots pine
E-mail minna.pulkkinen@iki.fi
Received 21 September 2012 Accepted 12 December 2012
Available at http://www.metla.fi/silvafennica/full/sf46/sf465747.pdf

Cover photo by Metla/Erkki Oksanen
Contents

Acknowledgements ...................................................................................................751
List of Symbols .........................................................................................................753

1 Introduction ........................................................................................................ 757

Theory

2 Geometrical and Statistical Concepts ..............................................................761

3 Estimation of Cross-Section Area .....................................................................769
  3.1 Foundations: Source of Randomness and Measures of Estimator Performance ..........................................................769
  3.2 Effect of Non-Convexity ..............................................................................770
  3.3 Estimators Based on Diameters and Circle Area Formula .........................770
    3.3.1 Girth Diameter: Mean Diameter Derived from Convex Perimeter ....771
    3.3.2 Random Diameters with Uniform Direction Distribution ................772
    3.3.3 Minimum and Maximum Diameters ..................................................780
  3.4 Estimators Based on Radii ...........................................................................781

4 Bitterlich Sampling ....................................................................................783
  4.1 Estimation of Stand Totals............................................................................784
    4.1.1 Principle under Assumption of Circular Cross-Sections ....................785
    4.1.2 Effect of Non-Circularity of Cross-Sections ......................................787
    4.1.3 Special Case: Relative Basal Area .....................................................795
  4.2 Bitterlich Diameters: Diameters Measured Parallel or Perpendicular to Plot Radius Direction .................................................799
    4.2.1 Direction Distributions of Bitterlich Diameters .................................800
    4.2.2 Estimation of Cross-Section Area by Bitterlich Diameters and Circle Area Formula ........................................................803

5 Estimation of Stem Volume ...............................................................................808
  5.1 Practical Volume Estimator: Laasasenaho Volume Equation .......................808
  5.2 General Volume Estimator: Definite Integral of Cross-Section Area Estimation Function .............................................................811
  5.3 Volume Estimator Based on Non-Parametric Stem Curve: Definite Integral of Interpolating Cubic Splines .........................................................814

Practice

6 Material ............................................................................................................... 819

7 Methods ............................................................................................................... 826
  7.1 Cross-Section Shape .....................................................................................826
    7.1.1 Shape of Convex Closure .................................................................827
    7.1.2 Non-Convexity .................................................................................829
    7.1.3 True Shape .......................................................................................830
  7.2 Estimation of Cross-Section Area ...............................................................832
  7.3 Estimation of Stem Volume ..........................................................................836
    7.3.1 True Volume .....................................................................................838
    7.3.2 Laasasenaho Volume Equation .........................................................839
    7.3.3 Cubic-Spline-Interpolated Stem Curve ..............................................842
    7.3.4 General Volume Estimator ...............................................................844
  7.4 Estimation of Stand Totals in Bitterlich Sampling .......................................846
8 Results with Discussion ........................................................................................................ 849
  8.1 Caveats on Data ........................................................................................................ 849
  8.2 Cross-Section Shape .................................................................................................. 851
    8.2.1 Shape of Convex Closure ................................................................................ 851
    8.2.2 Non-Convexity ................................................................................................... 860
    8.2.3 True Shape ......................................................................................................... 863
  8.3 Estimation of Cross-Section Area ............................................................................... 868
    8.3.1 Reference Estimator Based on Girth Diameter ............................................. 868
    8.3.2 Comparison of Area Estimators ...................................................................... 872
  8.4 Estimation of Stem Volume ........................................................................................ 888
    8.4.1 Laassanaho Volume Equation ......................................................................... 892
    8.4.2 Cubic-Spline-Interpolated Stem Curve .......................................................... 901
    8.4.3 General Volume Estimator ............................................................................... 912
  8.5 Estimation of Stand Totals in Bitterlich Sampling .................................................... 918

Synthesis

9 Summary and Conclusions ............................................................................................. 926

References .......................................................................................................................... 935

Appendices .......................................................................................................................... 940
  Appendix A Support Function of Closed Convex Set or Curve in Plane .............. 940
  Appendix B Horvitz-Thompson Theorem ................................................................. 944
  Appendix C Expressing Boundary Co-ordinates, Radii and Total Area of Inclusion Region of Tree in Bitterlich Sampling in Terms of Support Function of Breast Height Cross-Section ................................................................. 947
  Appendix D Expressing Probability Distribution of Bitterlich Diameter Direction in Terms of Support Function of Breast Height Cross-Section of Tree ......................................................................................... 951
  Appendix E Within-Tree Expectations of Volume Predictions by Laassanaho Volume Equation with Diameter Selection Methods 1–5, 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90 ................................................................. 956
  Appendix F Within-Tree Variances of Volume Predictions by Laassanaho Volume Equation with Diameter Selection Methods 1–5, 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90 ................................................................. 961
  Appendix G Computation of Diameters, Breadths and Radii from Vector Image of Cross-Section ........................................................................................................... 971
  Appendix H Computation of Contour Co-ordinates, Sector Areas and Total Area of Inclusion Region in Bitterlich Sampling from Vector Image of Cross-Section ........................................................................... 976
  Appendix I Computation of Scale in Vector Image of Cross-Section ................. 980
  Appendix K Examples of Association between Some Shape Indices and Actual Shapes of Cross-Sections ........................................................................................................... 984
Acknowledgements

Many people gave invaluable help in this study. The topic was suggested to me by Juha Lappi, who first became a supervisor and then a friend. I wish to thank him warmly and sincerely for his pertinent advice and lucid comments, particularly in the earlier phase of the work, for his alert action in the final publication phase of the work, and for the many profound discussions that we have had throughout the years.

The work began as a MSc project at the then Faculty of Forestry of the University of Joensuu (nowadays the School of Forest Sciences of the University of Eastern Finland). The disc data were generously provided by the VAPU project of the Finnish Forest Research Institute led by Kari T. Korhonen at the Joensuu Research Station. The discs were photographed by Matti Hälinen, using the rack system constructed by the staff in the Imaging Centre of the University; these staff also printed the photographs. With the vectorisation of the raster images of the scanned photographs, Janne Soimasuo and Ari Turkia lent a crucially helping hand. I am indebted to all these persons for their indispensable contribution.

All along the way, Bertil Matérn inspired the work with his excellent early (1956) paper. He was also kind enough to elucidate, in an exchange of letters, the issues in the paper that I was not able to fathom out myself. Further, he generously placed at my disposal some older literature that is hard to get hold of. I wish I could have completed the work before his demise, so that he could have appraised the extension of his original ideas, particularly those about the effect of non-circularity on Bitterlich sampling.

The ideas on Bitterlich sampling started to germinate during my working in the spatial statistics group led by Antti Penttininen at the Department of Statistics of the University of Jyväskylä, and during my half-year visit to the spatial statistics group at the Department of Mathematical Sciences of Aalborg University. I am grateful to these two statistics communities for sowing the seeds of the ideas in my mind, as well as for setting good examples of rigorous statistical thinking and presentation.

I carried out almost all the computation work after the MSc thesis, as well as a considerable part of writing, at the Vantaa Research Centre of the Finnish Forest Research Institute. The group of forest yield and growth study, and later also the group of forest technology, became my hearth and home there. My thanks are due to the researchers and other staff for their approving and encouraging attitude, and to the Institute for very good working facilities. Particularly, I am indebted to Carl-Gustaf Snellman for his assistance in computational problems, and to Jari Hynynen for arranging additional funding after the grant awarded for the work ran out. I also gratefully acknowledge Sointu Virkkala for drawing Figure 22. At that time, Timothy Gregoire was kind to send me the copies of some highly relevant works, especially those by G. R. Grosenbaugh that would have been difficult to obtain otherwise, for which I express my gratitude.

Outside the research institutions, Harri Hypén played an important role. I am especially obliged to him for one significant brainstorm, during a car drive to Eastern Finland, in which we sketched the algorithm for computing the inclusion regions of trees in Bitterlich sampling from the vector images of the cross-section contours. I also gratefully acknowledge him for drawing the original version of Figure 25.

I resumed writing, after a considerable break, while working at the Departments of Forest Ecology and Forest Resource Management (nowadays merged into one, the Department of Forest Sciences) of the University of Helsinki. I wish to give my heartfelt thanks to Annikki Mäkelä and Pauline Stenberg, my superiors, for their encouragement and support, which they gave also in the concrete form of arranging funding. My thanks are also due to the colleagues in the forest ecology modelling group, from whom I learnt much of doing research work, particularly the necessity of always asking what one’s results really
mean and imply, which then benefitted the writing of the thesis. Pertti Hari deserves special thanks for initiating and persisting with regular mentoring discussions, as a result of which the focus of the work was changed from empirical to theoretical findings. Finally, I wish to give recognition to all the people at the Department for making it such a lively and versatile community, to which it was a pleasure to belong, also outside the working hours.

In the finalisation phase, Christophe Godin’s help and support was indispensable. I am much obliged to him for the discussions that helped shape the introductory and summary chapters, as well as for the pivotal technical solutions for converting the manuscript into the format required by the publisher.

Juha Heikkinen and Göran Ståhl, the official pre-examiners, gave a substantial amount of their time for reviewing the thesis. I wish to thank them sincerely for their work. Matti Maltamo, the professor of forest mensuration formally responsible for the defence process, deserves thanks for his upholding attitude and prompt aid in all the technicalities related to the process.

Numerous institutions funded the study: the COMAS Graduate School of the University of Jyväskylä, the Academy of Finland, the Faculty of Forestry of the University of Joensuu / the School of Forest Sciences of the University of Eastern Finland, and the Foundation for Research of Natural Resources in Finland (grants number 1552/99, 1578/00, 1603/01) are hereby gratefully acknowledged.

Throughout the years, my near and dear, close and far, have tried to make me see, understand and remember the beauty of life, in its myriad aspects. I wish to give them my deepest thanks for this. As Mephistopheles puts it in his well-known words in J. W. von Goethe’s play Faust: “Grau, theurer Freund, ist alle Theorie, und grün des Lebens goldner Baum [My friend, all theory is grey, and green the golden tree of life].”
# List of Symbols

### Cross-Section

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>area of a cross-section</td>
</tr>
<tr>
<td>$A_C$</td>
<td>area of the convex closure of a cross-section</td>
</tr>
<tr>
<td>$\hat{A}_j$</td>
<td>estimator of cross-section area based on the circle area formula and</td>
</tr>
<tr>
<td></td>
<td>diameter selection method $j$</td>
</tr>
<tr>
<td>$\hat{A}_{Gj}$</td>
<td>generalisation of estimator $\hat{A}_j$ for $n$ similarly selected</td>
</tr>
<tr>
<td></td>
<td>diameters</td>
</tr>
<tr>
<td>$(A_C-\bar{A})/A_C$</td>
<td>relative convex deficit of a cross-section</td>
</tr>
<tr>
<td>$(\bar{A}_0-A_C)/A_C$</td>
<td>relative isoperimetric deficit of a cross-section</td>
</tr>
<tr>
<td>$B(\theta)$</td>
<td>breadth of a cross-section in direction $\theta$, with respect to the centre</td>
</tr>
<tr>
<td></td>
<td>of gravity; $B: [0, \pi) \to (0, \infty)$</td>
</tr>
<tr>
<td>$B_C(\theta)$</td>
<td>breadth of the convex closure of a cross-section in direction $\theta$,</td>
</tr>
<tr>
<td></td>
<td>with respect to the centre of gravity; $B_C: [0, \pi) \to (0, \infty)$</td>
</tr>
<tr>
<td>$b_c/a_c$</td>
<td>girth-area ellipse ratio of a cross-section; axis ratio of the ellipse that</td>
</tr>
<tr>
<td></td>
<td>has the same perimeter and area as the convex closure of a cross-section</td>
</tr>
<tr>
<td>$C$</td>
<td>girth of a cross-section; perimeter of the convex closure of a cross-section</td>
</tr>
<tr>
<td>$C_p$</td>
<td>perimeter of a cross-section</td>
</tr>
<tr>
<td>$CV_D$</td>
<td>coefficient of variation of diameter in a cross-section; $CV_D=\sigma_D/\mu_D$</td>
</tr>
<tr>
<td>$D(\theta)$</td>
<td>diameter of a cross-section in direction $\theta$; $D: [0, \pi) \to (0, \infty)$</td>
</tr>
<tr>
<td>$D_A$</td>
<td>true area diameter of a cross-section; diameter that produces the area of</td>
</tr>
<tr>
<td></td>
<td>a cross-section when substituted in the circle area formula; $D_A=2[A/\pi]^{1/2}$</td>
</tr>
<tr>
<td>$D_{Ac}$</td>
<td>convex area diameter of a cross-section; diameter that produces the area of</td>
</tr>
<tr>
<td></td>
<td>the convex closure of a cross-section when substituted in the circle area</td>
</tr>
<tr>
<td></td>
<td>formula; $D_{Ac}=2[A_C/\pi]^{1/2}$</td>
</tr>
<tr>
<td>$D_{max}$</td>
<td>maximum diameter of a cross-section</td>
</tr>
<tr>
<td>$D_{min}$</td>
<td>minimum diameter of a cross-section</td>
</tr>
<tr>
<td>$F_\xi(\xi; \alpha)$</td>
<td>cumulative distribution function of diameter direction $\xi$ parallel to</td>
</tr>
<tr>
<td></td>
<td>plot radius in Bitterlich sampling with viewing angle $\alpha$</td>
</tr>
<tr>
<td>$f_\xi(\xi; \alpha)$</td>
<td>probability density function of diameter direction $\xi$ parallel to plot</td>
</tr>
<tr>
<td></td>
<td>radius in Bitterlich sampling with viewing angle $\alpha$</td>
</tr>
<tr>
<td>$p(\theta)$</td>
<td>support function of the convex closure of a cross-section; $p: [0, 2\pi) \to</td>
</tr>
<tr>
<td></td>
<td>(0, \infty)$</td>
</tr>
<tr>
<td>$R(\theta)$</td>
<td>radius of a cross-section in direction $\theta$ from the centre of gravity;</td>
</tr>
<tr>
<td></td>
<td>$R: [0, 2\pi) \to (0, \infty)$</td>
</tr>
<tr>
<td>$R_C(\theta)$</td>
<td>radius of the convex closure of a cross-section in direction $\theta$ from</td>
</tr>
<tr>
<td></td>
<td>the centre of gravity; $R_C: [0, 2\pi) \to (0, \infty)$</td>
</tr>
<tr>
<td>$R_{max}$</td>
<td>maximum radius of a cross-section</td>
</tr>
<tr>
<td>$R_q$</td>
<td>quadratic mean of the observed radii in a cross-section</td>
</tr>
<tr>
<td>$R$</td>
<td>landmark configuration of a cross-section; collection of the observed radii</td>
</tr>
<tr>
<td></td>
<td>in a cross-section</td>
</tr>
<tr>
<td>$R^*$</td>
<td>centred pre-shape of a landmark configuration of a cross-section; $R^<em>=(R^</em>(j·1°))_{j=0, 1, \ldots, 359}$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>general symbol of direction; diameter direction with uniform distribution</td>
</tr>
<tr>
<td></td>
<td>in $[0, \pi)$</td>
</tr>
<tr>
<td>$\theta_{D_{max}}$</td>
<td>direction of $D_{max}$</td>
</tr>
<tr>
<td>$\theta_{D_{min}}$</td>
<td>direction of $D_{min}$</td>
</tr>
<tr>
<td>$\theta_{R_{max}}$</td>
<td>direction of $R_{max}$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>plot radius direction in Bitterlich sampling</td>
</tr>
<tr>
<td>$\xi$</td>
<td>diameter direction parallel to plot radius in Bitterlich sampling; $\xi \sim</td>
</tr>
<tr>
<td></td>
<td>$F_\xi(\xi; \alpha)$ with viewing angle $\alpha$</td>
</tr>
</tbody>
</table>
μ_D\(D\) girth diameter of a cross-section; within-cross-section expectation of diameter over the uniform direction distribution in \([0, \pi)\); \(\mu_D = C/\pi\)

μ_D(ξ) mean Bitterlich diameter, parallel to plot radius, in a cross-section; within-cross-section expectation of diameter over the distribution of the diameter direction ξ parallel to plot radius in Bitterlich sampling

μ_D(ξ+π/2) mean Bitterlich diameter, perpendicular to plot radius, in a cross-section; within-cross-section expectation of diameter over the distribution of the diameter direction ξ+π/2 perpendicular to plot radius in Bitterlich sampling

σ_D^2 diameter variance in a cross-section; within-cross-section variance of diameter over the uniform direction distribution in \([0, \pi)\)

σ_D(ξ)^2 variance of Bitterlich diameters, parallel to plot radius, in a cross-section; within-cross-section variance of diameter over the distribution of the diameter direction ξ parallel to plot radius in Bitterlich sampling

σ_D(ξ+π/2)^2 variance of Bitterlich diameters, perpendicular to plot radius, in a cross-section; within-cross-section variance of diameter over the distribution of the diameter direction ξ+π/2 perpendicular to plot radius in Bitterlich sampling

γ_D(φ) within-cross-section covariance of the diameters intersecting at angle φ over the uniform direction distribution in \([0, \pi]\); diameter autocovariance function in a cross-section; \(\gamma_D(0, \pi/2) = 0\)

ρ_D(φ) within-cross-section correlation of the diameters intersecting at angle φ over the uniform direction distribution in \([0, \pi]\); diameter autocorrelation function in a cross-section; \(\rho_D(0, \pi/2) = 1\)

ρ_D(π/2) within-cross-section correlation of perpendicular diameters over the uniform direction distribution in \([0, \pi)\)

ρ_D(ξ)(π/2) within-cross-section correlation of the diameters parallel and perpendicular to plot radius in Bitterlich sampling

**Stem**

A(h) area of the cross-section of a stem at height h; cross-section area function; \(A : [0, H] \rightarrow [0, \infty)\)

\(\hat{A}_j(h)\) estimator of the area of the cross-section of a stem at height h based on the circle area formula and diameter selection method j; area estimation function; \(\hat{A}_j : [0, H] \rightarrow [0, \infty)\)

D(θ, h) diameter in direction θ at height h, \(\theta \in [0, \pi)\)

D_j(h) diameter selected with method j at height h

D_A(h) true area diameter \(D_A\) at height h; true stem curve; \(D_A : [0, H] \rightarrow [0, \infty)\)

D_{Ac}(h) convex area diameter \(D_{Ac}\) at height h

\(\bar{D}_j\) vector of diameters selected with method j at the observation heights H in a stem

\(D_A\) vector of true area diameters at the observation heights H in a stem

\(D_{Ac}\) vector of convex area diameters at the observation heights H in a stem

H length of a stem; height of a tree determined from the ground level

H vector of the observation heights in a stem

V true volume of a stem

\(\hat{V}\) estimated true volume of a stem; \(\hat{V} = \hat{V}_S\)

\(V_C\) convex volume of a stem (approximately the volume of the convex hull of a stem)

\(\hat{V}_C\) estimated convex volume of a stem; \(\hat{V}_C = \hat{V}_{CS}\)
\( \hat{V}_L \) best estimate of the volume of a stem by the re-estimated Laasasenaho three-variable volume equation involving true area diameters

\( \hat{V}_{CL} \) best estimate of the convex volume of a stem by the re-estimated Laasasenaho three-variable volume equation involving convex area diameters

\( \hat{V}_{Lj} \) estimator of stem volume based on the re-estimated Laasasenaho three-variable volume equation, involving true area diameters, and on diameters selected with method j

\( \hat{V}_{CLj} \) estimator of convex stem volume based on the re-estimated Laasasenaho three-variable volume equation, involving convex area diameters, and on diameters selected with method j

\( \hat{V}_S \) best estimate of the volume of a stem by cubic-spline-interpolated stem curve obtained from true area diameters

\( \hat{V}_{CS} \) best estimate of the convex volume of a stem by cubic-spline-interpolated stem curve obtained from convex area diameters

\( \hat{V}_{Sj} \) estimator of stem volume by cubic-spline-interpolated stem curve obtained from diameters selected with method j

\( \hat{V}_{Gj} \) estimator of stem volume by general stem volume estimator involving diameters selected with method j

\( \mu_{\hat{A}_j}(h) \) mean function of area estimation process \( \{\hat{A}_j(h), h \in [0, H]\} \) in a stem; \( \mu_{\hat{A}_j}:[0, H] \rightarrow [0, \infty), \mu_{\hat{A}_j}(h)=E[\hat{A}_j(h)] \)

\( \mu_{\Delta \hat{A}_j}(h) \) mean function of area estimation error process \( \{\hat{A}_j(h)-A(h), h \in [0, H]\} \) in a stem; \( \mu_{\Delta \hat{A}_j}:[0, H] \rightarrow \mathbb{R}, \mu_{\Delta \hat{A}_j}(h)=E[\hat{A}_j(h)-A(h)] \)

\( \sigma_{\hat{A}_j}^2(h) \) variance function of area estimation process \( \{\hat{A}_j(h), h \in [0, H]\} \) in a stem; \( \sigma_{\hat{A}_j}^2:[0, H] \rightarrow [0, \infty), \sigma_{\hat{A}_j}^2(h)=\text{Var}[\hat{A}_j(h)] \)

\( \gamma_{\hat{A}_j}(h, k) \) covariance function of area estimation process \( \{\hat{A}_j(h), h \in [0, H]\} \) in a stem; \( \gamma_{\hat{A}_j}:[0, H] \times [0, H] \rightarrow \mathbb{R}, \gamma_{\hat{A}_j}(h, k)=\text{Cov}[\hat{A}_j(h), \hat{A}_j(k)] \)

\( \rho_{\hat{A}_j}(h, k) \) correlation function of area estimation process \( \{\hat{A}_j(h), h \in [0, H]\} \) in a stem; \( \rho_{\hat{A}_j}:[0, H] \times [0, H] \rightarrow [-1, 1], \rho_{\hat{A}_j}(h, k)=\gamma_{\hat{A}_j}(h, k)/[\sigma_{\hat{A}_j}(h)\sigma_{\hat{A}_j}(k)] \)

**Stand**

\( A_{Ci} \) area of the convex closure of the breast height cross-section of tree i

\( \hat{A}_{ji} \) estimator of the area of the breast height cross-section of tree i based on the circle area formula and diameter selected with method j

\( D_i \) breast height diameter of tree i with circular breast height cross-section

\( G \) relative basal area of a tree stand; \( G=\sum_{i \in I} A_{Ci}/|L| \) in stand I in region L

\( \hat{G}_{HT} \) Horvitz-Thompson estimator of G, based on a sample taken from the tree stand

I tree stand in L; set of the trees growing in L and reaching above breast height

L region of interest in Bitterlich sampling

|L| area of L

\( M_i(\alpha) \) inclusion region of tree i in Bitterlich sampling with viewing angle \( \alpha \)

\( M_i(\alpha)^b \) sector of \( M_i(\alpha) \) edged by the rays emanating in directions a and b from the centre of gravity of the breast height cross-section of tree i

|\( M_i(\alpha) \)| area of \( M_i(\alpha) \)

\( r_i(\alpha) \) radius of the inclusion region of tree i with circular breast height cross-section in Bitterlich sampling with viewing angle \( \alpha \)

\( s_q(\alpha) \) sample of trees selected with Bitterlich sampling with viewing angle \( \alpha \) at a uniformly randomly located point Q
Y total amount of a characteristic of interest in a tree stand; $Y=\sum_{i\in I}Y_i$ in stand $I$ in region $L$.

$\hat{Y}_{HT}$ Horvitz-Thompson estimator of $Y$, based on a sample taken from the tree stand.

$\alpha$ viewing angle in Bitterlich sampling, relascope angle.

$\kappa_i(\alpha)$ basal area factor of tree $i$ in Bitterlich sampling with viewing angle $\alpha$; $\kappa_i: (0, \pi) \rightarrow (0, 1)$, $\kappa_i(\alpha)=A_C/|M_i(\alpha)|$.

$\pi_i(\alpha)$ inclusion probability of tree $i$ growing in $L$ in Bitterlich sampling with viewing angle $\alpha$; $\pi_i(\alpha)=|M_i(\alpha)|/|L|$.

**General**

$\Pr\{A\}$ probability of event $A$.

$S_3(a; A, B)$ interpolating cubic spline based on the vector $B$ of values observed at the vector $A$ of one-dimensional locations.

$\delta_i(A)$ random indicator variable; $\delta_i(A)=1$, if $i \in A$, and $\delta_i(A)=0$, if $i \notin A$.

**Diameter Selection Methods within Cross-Section**

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Girth diameter; mean diameter $\mu_D=C/\pi$ of over the uniform direction distribution in $[0, \pi)$</td>
</tr>
<tr>
<td>1</td>
<td>“Random” diameter; diameter taken in a uniformly distributed direction in $[0, \pi)$</td>
</tr>
<tr>
<td>2</td>
<td>Arithmetic mean of the diameter in 1 and the diameter perpendicular to it</td>
</tr>
<tr>
<td>3</td>
<td>Geometric mean of the diameters in 2</td>
</tr>
<tr>
<td>4</td>
<td>Arithmetic mean of two “random diameters”; arithmetic mean of two diameters taken independently in uniformly distributed directions in $[0, \pi)$</td>
</tr>
<tr>
<td>5</td>
<td>Geometric mean of the diameters in 4</td>
</tr>
<tr>
<td>6</td>
<td>Arithmetic mean of the minimum diameter and the maximum diameter</td>
</tr>
<tr>
<td>7</td>
<td>Geometric mean of the diameters in 6</td>
</tr>
<tr>
<td>8</td>
<td>Arithmetic mean of the minimum diameter and the diameter perpendicular to it</td>
</tr>
<tr>
<td>9</td>
<td>Geometric mean of the diameters in 8</td>
</tr>
<tr>
<td>10</td>
<td>Arithmetic mean of the maximum diameter and the diameter perpendicular to it</td>
</tr>
<tr>
<td>11</td>
<td>Geometric mean of the diameters in 10</td>
</tr>
<tr>
<td>1\xi</td>
<td>Diameter taken parallel to plot radius in Bitterlich sampling</td>
</tr>
<tr>
<td>2\xi</td>
<td>Arithmetic mean of the diameters in taken parallel and perpendicular to plot radius in Bitterlich sampling</td>
</tr>
<tr>
<td>3\xi</td>
<td>Geometric mean of the diameters in 2\xi</td>
</tr>
<tr>
<td>4\xi</td>
<td>Arithmetic mean of the diameter taken parallel to plot radius in Bitterlich sampling and a diameter taken independently in a uniformly distributed direction in $[0, \pi)$</td>
</tr>
<tr>
<td>5\xi</td>
<td>Geometric mean of the diameters in 4\xi</td>
</tr>
<tr>
<td>1\xi90</td>
<td>Diameter taken perpendicular to plot radius in Bitterlich sampling</td>
</tr>
<tr>
<td>4\xi90</td>
<td>Arithmetic mean of the diameter taken perpendicular to plot radius in Bitterlich sampling and a diameter taken independently in a uniformly distributed direction in $[0, \pi)$</td>
</tr>
<tr>
<td>5\xi90</td>
<td>Geometric mean of the diameters in 4\xi90</td>
</tr>
</tbody>
</table>

$\min$ minimum diameter

$\max$ maximum diameter
1 Introduction

In the common methods of forest mensuration, the cross-sections of tree stems are assumed to be circular: The area of a stem cross-section is usually estimated with the circle area formula. The volume of a stem is typically estimated as the solid of revolution of a stem curve, or with a volume equation constructed with a combination of circle-base geometric solids (cones, paraboloids, neiloids) as the starting point. Trees to be measured in forest inventories are often selected with Bitterlich sampling (relascope sampling), where the inclusion probabilities, required in the estimator of the stand total of any characteristic of interest, are estimated by assuming that the breast height cross-sections of the selected trees are circular. The basal area of a stand is estimated either as the sum of the treewise estimates of cross-section area given by the circle area formula, or, as is most often the case in practical forestry, by multiplying the number of trees selected in Bitterlich sampling with the so called basal area factor; the latter is a special case of the stand total estimation in Bitterlich sampling and thus assumes circularity on the breast height cross-sections of the trees.

In nature, however, the cross-sections of tree stems are hardly ever exactly circular. Non-circularity occurs due to defects and deformations inflicted by pathogens or mechanical damages, but also without them, in entirely healthy and undamaged trees. Asymmetric growing space affecting the access of a tree to light has been suggested to be one potential cause of non-circularity: with uneven spatial distribution of light, the crown would develop asymmetrically, and the stem should then compensate the resulting mass imbalance by increased wood formation in the direction of the torsional moment. Observations have been reported both for (e.g. Isomäki 1986, Robertson 1991) and against (e.g. Bucht 1981, Bouillet and Houllier 1994) this idea. Wind is another factor proposed to induce non-circularity (Banks 1973, Grace 1977). The mechanism would essentially be similar to the one suggested above: winds blowing continuously from the same direction would cause a torsional moment, which the stem should then counteract by forming more wood in the direction of the wind. Observations supporting this idea have been reported, for example, by Müller (1958a) and Robertson (1986, 1990, 1991). Through the same mechanism, growing in a steep slope or in a leaning position could also result in non-circular stems (Pawsey 1966, Loetsch et al. 1973). Regular heavy snow loads, in turn, could encourage trees to develop straight stems with circular cross-sections to resist the bending forces of the load (Professor emeritus Pertti Hari, personal communication). In general, eccentric radial growth has often been found to be associated with reaction wood formation (e.g. Burdon 1975, Harris 1977, Robertson 1991). The factors listed above could explain the age-related variation in non-circularity: old trees tend to have more irregular cross-sections than young ones, simply because they have been exposed to inflicting growing conditions for a longer time. However, even if growing in similar conditions, some species appear to be more non-circular than others: according to Kärkkäinen (2003), broad-leaved species in Finland are in general more non-circular than coniferous ones; Loetsch et al. (1973), in turn, mention Tectona grandis, Carpinus betulus and Robinia pseudacacia as the species with very irregular cross-sections.

Following from the circularity assumption, the key characteristic of a stem cross-section is its diameter. Yet in a non-circular cross-section there is no single diameter value but diameter varies with direction. In practice, diameter is typically measured with a caliper, which gives the distance between two parallel tangents (that are perpendicular to the direction in which the diameter is being measured), or with a girth tape, by dividing the perimeter measurement by π, which gives the mean of the calipered diameter over all directions (as we will see in Chapter 2). It is important to differentiate between diameter variation due to non-circular shape and diameter variation caused by measurement errors (arising from faulty
handling of instruments, adverse measurement conditions, psychological factors etc.); in literature this distinction has not always been made, but genuine measurement errors have erroneously been ascribed to non-circularity (Matérn 1990).

Variation in diameter within a cross-section of a tree causes variation in the output of such a cross-section area or volume estimator where the diameter is used as an input variable: the estimate of area or volume depends on the selection of diameter. This selection involves the choice of the measurement direction, the number of diameters to be measured and the type of mean (geometric, arithmetic, quadratic) to be applied to the measured diameters. Volume models often involve diameters at more than one height in the stem; the more heights are involved the more complex variation non-circularity potentially induces in the volume estimates. The within-tree variation in estimator output can be characterised by bias and variance. Bias is a measure of the systematic error due to non-circularity: it is the deviation of the mean of all possible estimates, obtained with all possible outcomes of the diameter selection, from the true value, or in other words, the mean error expected over all possible outcomes of the diameter selection. Variance of all possible estimates around their mean, in turn, quantifies the sampling error related to non-circularity; from it we can derive the likely range of the error in the estimate associated with one outcome of the diameter selection.

In Bitterlich sampling, variation in diameter within the breast height cross-section causes variation in the estimate of the inclusion probability of a tree, since the estimator of the probability, based on the circularity assumption, involves the breast height diameter as its input variable. In addition (as we will see in Chapter 4; not shown before), non-circularity inflicts another separable error component in the inclusion probability estimate, dependent only of the cross-section shape and independent of diameter selection. The combination of these errors then decides how the tree contributes, due to its non-circularity, to the bias of the stand total estimator of any characteristic of interest.

In his path-breaking study, Matérn (1956) derived, without postulating anything about the shape of a cross-section, the bias and the approximative variance for the area estimators based on the circle area formula and some common diameter selection methods (diameter derived from perimeter, diameter calipered in a random direction, mean of this diameter and its perpendicular, or mean of two diameters calipered in random directions). On the basis of the bias formulae, he could show that these area estimators systematically overestimate true cross-section area. Furthermore, he developed theory on the effect of non-circularity on basal area estimation with Bitterlich sampling and was able to establish that the overestimating bias induced by non-circularity is practically (with the commonly used small viewing angles) the same as we would get by calipering every stem in one randomly chosen direction. In a later work (1990), he then applied this theory in data that consisted of contour drawings made on one hundred discs sawn on about forty Scots pine and Norway spruce stems. Besides Matérn’s work, no other theoretical developments concerning the effect of within-cross-section diameter variation on area estimation are to be found in literature. Likewise, the effect of non-circularity on stem volume estimation or on estimation of other stand totals than basal area in Bitterlich sampling appear hitherto theoretically unexplored.

Many empirical studies addressing non-circularity have focused on the shape of cross-sections, investigating with simple shape indices to what extent cross-sections deviate from a circle and how this deviation is related to position of the cross-section in the stem, tree species, silvicultural treatments, growing conditions etc. (e.g. Renvall 1923, Solbraa 1939, Williamson 1975, Kellogg and Barber 1981, Okstad 1983, Mäkinen 1998). Another large set of empirical studies have concerned cross-section area estimation with the circle area formula and different diameter selection methods, reporting differences in area estimates between the diameter selection methods and, more recently with the emergence of less
laborious area measurement methods, also the errors with respect to true area (e.g. Kennel 1959, Chacko 1961, Kärkkäinen 1975a, Biging and Wensel 1988, Gregoire et al. 1990). In many of these studies, in particular earlier ones, data consist of field measurements with a caliper or a girth tape and are thus very likely to contain measurement errors. A clearly distinct branch of studies, making use of ample data of cross-section radii provided by scanners employed in sawmills, have tested different models of cross-section shape with the aim of adapting sawing patterns to maximise the sawing yield (e.g. Skatter and Høibø 1998, Saint-André and Leban 2000). Practically no empirical studies seem to exist concerning the effect of within-cross-section diameter variation on volume estimation nor concerning the effect of non-circularity on Bitterlich sampling.

Nonetheless, the effects of non-circularity are worth investigating, even though other sources of error (sampling errors, measurement errors, model misspecification errors, estimation errors in model parameters, model residual errors) probably induce much more uncertainty in the results of forest inventories: on the basis of Matérn’s theoretical results, it is realistic to anticipate that non-circularity might inflict systematic errors also in volume estimates and stand total estimates by Bitterlich sampling, and that these, although presumably small in magnitude, can then cumulate into considerable errors in large area inventories. Further, in research purposes where one often strives for eliminating confounding factors, taking non-circularity effect into account may clarify and consolidate the results of analyses; particularly, when estimating growth with the difference between diameters, cross-section areas or stem volumes at two time points, the potentially asymmetric growth in non-circular cross-sections need be heeded.

This study concerns the effects of non-circularity on (i) cross-section area estimation with the circle area formula, (ii) stand total estimation in Bitterlich sampling with the circularity assumption, and (iii) stem volume estimation by a standard three-variable volume equation, by a non-parametric stem curve often applied in research purposes, and by a theoretical general volume estimator based on a cross-section area estimation function. The estimators considered are commonly used for standing trees or felled sample trees and involve dimensions of stems that can be measured from outside with the usual measurement equipment (calipered diameters, perimeters given by a tape, height obtained with a hypsometer or a tape). The diameter selection methods included in the examination are such that they exist in literature and are used, or could in principle be used, in practice. The first aim of the study was to develop further the existing theory: to derive the missing statistical properties of the area estimators under within-cross-section diameter variation, to devise methods for estimating similar statistical properties for the volume estimators, and to unravel theoretically how non-circularity affects Bitterlich sampling, without postulating anything about cross-section shape. The second aim of the study was to investigate with reasonable data how cross-section shape varies in Scots pine and of what magnitude the above-mentioned theoretically established effects of non-circularity can be in practice.

In the theoretical part of the work, we (i) derived the within-cross-section bias, approximative variance and true variance for the area estimators based on the circle area formula and five common diameter selection methods (involving diameters calipered in randomly chosen directions and their perpendiculurs; methods already addressed by Matérn), as well as their generalisations up to n diameters, (ii) described how non-circularity of breast height cross-sections influences inclusion probabilities of trees and hence stand total estimation in Bitterlich sampling, (iii) derived the non-uniform direction distributions for the diameters measured parallel or perpendicular to plot radius in Bitterlich sampling, and presented the within-cross-section bias, approximative variance and true variance for the area estimators based on the circle area formula and eight diameter selection methods involving these diameters, and (iv) presented methods for estimating the within-tree bias and variance of
the three types of volume estimators (volume equation, non-parametric stem curve, theoretical general volume estimator based on a cross-section area estimation function) with the diameter selection methods considered in (i) and (iii).

The empirical part of the work relied on the data of 709 discs taken at 6–10 heights in 81 healthy Scots pine stems in different parts of Finland; to eliminate measurement errors, the characteristics of the cross-sections were computed from digital images. With these data, we (i) investigated the variation in cross-section shape in different parts of the stems by means of some scalar and functional indices, (ii) estimated the within-cross-section biases and variances of the area estimators based on the circle area formula and 22 diameter selection methods (including those considered in the theoretical part), (iii) estimated the within-tree biases and variances of the three types of volume estimators with the same 22 diameter selection methods, and (iv) estimated, still with the same 22 diameter selection methods, the tree-specific errors caused by non-circularity in the inclusion probabilities in Bitterlich sampling, imparting the contribution of each tree in the bias of a stand total estimate. Owing to some defects in the data (non-probabilistic sampling of trees with uneven spatial distribution over Finland, skewed size distribution with small trees highly over-represented, debarking of discs before imaging), the empirical results cannot be generalised to any defined population (such as the Scots pine trees in Finland) or do not compare with the usual characteristics (that include bark) but should rather be taken as illustrations.

As a final remark, non-circularity was here studied as a static geometric phenomenon: the biological processes forming cross-section shape and the factors influencing these processes (site conditions, competition, management history of the stand etc.) were beyond the scope of this study.
2 Geometrical and Statistical Concepts

In this chapter, we introduce the notions of convex closure and support function, the latter of which we then use to define the key concept of diameter. Then we describe diameter selection within a cross-section of a tree stem as a random experiment involving sampling from diameter direction distribution and present some characteristics to summarise the diameter distribution within a cross-section. Further, we illustrate with artificial examples how these characteristics relate to the shape of the convex closure of a cross-section. Finally, we define radius and breadth, two other dimensions of a cross-section. Much of the notation employed later in the thesis is set up in this chapter.

Let us consider a cross-section perpendicular to the longitudinal axis of a tree stem as a closed and bounded set in $\mathbb{R}^2$. This set is convex if for every pair of points in it the line segment connecting the points is also contained in it. The convex closure of the set is the intersection of all the closed convex sets in $\mathbb{R}^2$ that contain the set (Fig. 1); from the definition it immediately follows that the cross-section is convex if and only if it equals its convex closure (Santaló 1976, Kelly and Weiss 1979). The boundary of the convex closure forms a closed convex curve (Santaló 1976), the length of which is here referred to as the convex perimeter of the cross-section and denoted by $C$. The area of the convex closure is here termed the convex area of the cross-section and denoted by $A_C$, whereas for the area of the cross-section the expression true area and denotation $A$ is used. Clearly, every non-convex cross-section is smaller in area and larger in perimeter than its convex closure, that is, always $A \leq A_C$ and true perimeter $\geq C$.

The concept of convex closure is important in forest mensuration, because the commonly used measuring instruments — girth tape, caliper and relascope — do not detect the possible non-convexity of a cross-section but only provide information on its convex closure: A girth tape stretched around a stem contours the convex closure of a cross-section and gives the convex perimeter as its reading. Caliper and relascope measurements, in turn, are based on observing the tangents of the convex closure of a cross-section (Matérn 1956, Matérn 1990, Loetsch et al. 1973, Bitterlich 1984).

Support function, first applied by Matérn (1956) in a context similar to this, is an invaluable tool for relating a convex closure with its tangents by a straightforward mathematical expression. In order to define the function, we first need to set a rectangular planar coordinate system: the origin is chosen to be an interior point $O$ of the convex closure, and a reference direction (the direction of the positive x-axis) is fixed; as usual, we measure the angles anticlockwise with respect to the reference direction. Now the value $p(\theta)$ of the support function $p: [0, 2\pi) \to (0, \infty)$ of the convex closure is defined as the length of the normal drawn in direction $\theta$ from the origin $O$ to the tangent of the convex closure (Fig. 2) (Matérn 1956, 1990, Santaló 1976; see Kelly and Weiss 1979, Rockafellar 1970, Stoyan and Stoyan 1994, and Webster 1994 for a more general definition). Different selections of $O$ naturally result in different functions (which cannot be transformed to the same form by a simple
phase transition, as is the case with different x-axis direction selections); in other words, the support function of a set is always defined with reference to some inner point of the set.

A necessary and sufficient condition for any twice differentiable nonnegative function \( p(\cdot) \) to be a support function of a convex closure is that \( p(\theta) + p''(\theta) > 0 \) for all \( \theta \) in \([0, 2\pi)\) (Matérn 1956, Santaló 1976). There is a one-to-one correspondence between the shape of a convex closure and its support function, that is, the boundary curve of the closure determines \( p(\cdot) \) uniquely and vice versa (Rockafellar 1970, Webster 1994). As suggested above, the family of all the tangents of a convex closure is easy to express in terms of the support function; from the tangents, in turn, a parametric representation for the rectangular Cartesian co-ordinates of the boundary is straightforward to derive; then, knowing the boundary co-ordinates, we can express the perimeter and the area of the convex closure in terms of the support function as

\[
C = \int_0^{2\pi} p(\theta) d\theta ,
\]

and

\[
A_c = \frac{1}{2} \int_0^{2\pi} p(\theta)[p(\theta) + p''(\theta)] d\theta = \frac{1}{2} \int_0^{2\pi} [p(\theta)^2 - p'(\theta)^2] d\theta
\]

(Matérn 1956, Santaló 1976, Stoyan and Stoyan 1994). For a more detailed discussion and derivation of these results, see Appendix A.

The crux of the usefulness of the support function in our context is that it lends itself so naturally to the diameter definition common in forestry: the diameter of a cross-section is the continuous function \( D:[0, \pi) \rightarrow (0, \infty) \),

\[
D(\theta) = p(\theta) + p(\theta + \pi) .
\]

The diameter in direction \( \theta \) is thus the distance between the two parallel tangents of the convex closure of the cross-section drawn in direction \( \theta + \pi/2 \), or, equivalently, the length of the orthogonal projection of the convex closure of the cross-section in direction \( \theta \) (Fig. 3). Accordingly, a diameter defined in this way corresponds to a calipered diameter in practical forest mensuration on one hand (Matérn 1956, 1990), and to the general concept of the width of a closed set in Euclidian n-spaces on the other hand (Kelly and Weiss 1979, Stoyan and Stoyan 1994). Note that there is a clear distinction between this definition and
the usual topological concept of the diameter of a set, which is defined as the supremum of the Euclidian distances between all the pairs of the points in the set (see e.g. Kelly and Weiss 1979, Santaló 1976). Although the support function is always defined with respect to some interior point O of the cross-section, the diameter function is invariant of the selection of this point.

Measuring a diameter of a non-circular cross-section means sampling from an infinite “diameter population” within the cross-section. This can be carried out by selecting a diameter direction θ within the interval \([0, \pi)\). If θ is sampled randomly, implying that θ is a random variable with some probability distribution, also the diameter \(D(\theta)\) becomes a random variable with some probability distribution (as \(D(\theta)\) is merely a continuous transformation of \(\theta\) determined by the form of the cross-section). The diameter moments and central moments, which characterise the diameter distribution within the cross-section, can then be expressed by means of the diameter function and the direction distribution: for \(k \in \mathbb{Q}_+\),

\[
E[D(\theta)^k] = \frac{1}{\pi} \int_0^{\pi} D(\theta)^k f_\theta(\theta) \, d\theta
\]  

and

\[
E \left\{ \left( D(\theta) - E[D(\theta)] \right)^k \right\} = \frac{1}{\pi} \int_0^{\pi} \left( D(\theta) - E[D(\theta)] \right)^k f_\theta(\theta) \, d\theta ,
\]

where \(f_\theta(\theta)\) is the probability density function of \(\theta\). Note that although we in the following consider diameter moments in a cross-section primarily over the uniform direction distribution in \([0, \pi)\), this distribution is, although perhaps the most natural choice, by no means the only option. For example, if the breast height cross-section of a tree deviates from the circular shape, measuring the breast height diameter parallel or perpendicular to the plot radius direction in relascope sampling corresponds to selecting the diameter direction from a particular non-uniform direction distribution (see Chapter 4, Section 4.2).

As the probability density function of the uniform direction distribution is \(1/\pi\) for \(\theta \in [0, \pi)\) and zero elsewhere, the mean diameter in a cross-section over this distribution becomes

\[
\mu_D = E[D(\theta)] = \frac{1}{\pi} \int_0^{\pi} D(\theta) \, d\theta
\]

(cf. Matérn 1956, 1990, Stoyan and Stoyan 1994). Between the mean diameter and the convex perimeter \(C\) there exists the following quite practical relation (obvious from Eqs. 1, 3, and 6):

\[
C = \frac{2\pi}{\theta} \int_0^{\pi} p(\theta) \, d\theta = \int_0^{\pi} \left[ p(\theta) + p(\theta + \pi) \right] \, d\theta = \int_0^{\pi} D(\theta) \, d\theta = \pi \mu_D
\]
The diameter variance in a cross-section over the uniform direction distribution is given by

$$\sigma_D^2 = \text{Var}[D(\theta)] = E\left\{ \left( D(\theta) - E[D(\theta)] \right)^2 \right\} = \frac{1}{\pi} \int_0^\pi (D(\theta) - \mu_D)^2 d\theta \quad (8)$$

Variance is a special case of covariance, as $\sigma_D^2 = \gamma_D(0)$. Further, the correlation between diameters intersecting at angle $\phi$, that is, the diameter autocovariance function $\gamma_D: [0, \pi/2] \rightarrow [0, \infty)$ at point $\phi$ taken over the uniform direction distribution, is defined as

$$\gamma_D(\phi) = \text{Cov}[D(\theta), D(\theta + \phi)] = E\left\{ (D(\theta) - E[D(\theta)]) \left( D(\theta + \phi) - E[D(\theta + \phi)] \right) \right\}$$

$$= \frac{1}{\pi} \int_0^\pi (D(\theta) - \mu_D)(D(\theta + \phi) - \mu_D)d\theta \quad (9)$$

Variance is a special case of covariance, as $\sigma_D^2 = \gamma_D(0)$. Further, the correlation between diameters intersecting at angle $\phi$, that is, the diameter autocorrelation function $\rho_D: [0, \pi/2] \rightarrow [-1, 1]$ at point $\phi$, $\phi \in [0, \pi/2]$, is expressed as

$$\rho_D(\phi) = \frac{\text{Cov}[D(\theta), D(\theta + \phi)]}{\sqrt{\text{Var}[D(\theta)] \text{Var}[D(\theta + \phi)]}} = \frac{1}{\sigma_D^2 \pi} \int_0^\pi (D(\theta) - \mu_D)(D(\theta + \phi) - \mu_D)d\theta \quad (10)$$

It is essential to notice that diameter information — however complete and precise concerning the diameter function $D(\cdot)$ — is in general insufficient for making inference about the exact shape and area of a non-circular cross-section. Besides non-convexity, which is ignored in diameter information by definition, one may also encounter problems with the family of constant-width convex sets, also referred to as orbiforms (a name given by Leonhard Euler; Tiercy 1920, Matérn 1956), in which the diameter is constant in all directions (i.e., $\sigma_D^2 = 0$). In addition to the circle, the family comprises, for example, the Reuleux polygons (Fig. 4): given a regular (i.e., equiangular and equilateral) polygon with an odd number of sides, the corresponding Reuleux polygon is formed by the circular arcs that are subtended by the sides of the linear polygon and whose centres are the opposite vertices of the linear polygon (Santaló 1976). While the circle is the largest in area of the orbiforms of equal diameter (Hadwiger 1957), the Reuleux triangle is the smallest. By Cauchy’s theorem (Eq. 7), the orbiforms of equal diameter are also isoperimetric, that is, their perimeters are equal (this result is sometimes referred to as Barbier’s theorem after the French mathematician Joseph Emile Barbier).

The numerical examples provided by Matérn (1956) (Fig. 5, Table 1) illustrate how the above-mentioned diameter characteristics (computed over the uniform direction distribution) relate to the shape of the convex closure of a cross-section (Table 1, Figs. 6, 7 and 8). The mean diameter $\mu_D$ is merely a size parameter, whereas the diameter variance $\sigma_D^2$ reflects both size, shape and smoothness of the convex closure (Stoyan and Stoyan 1994); the effect
of size can naturally be partialled out from $\sigma_D^2$ by using the diameter coefficient of variation $C_{VD}=\sigma_D/\mu_D$. Even if the diameter autocorrelation function $\rho_D(\cdot)$ does not uniquely relate to the diameter function $D(\cdot)$ — different non-orbiform convex shapes with the same $\rho_D(\cdot)$ can be found (e.g. shapes B and C in Fig. 8; Stoyan and Stoyan 1994) — it evidently characterises some aspects of the shape of a cross-section. Simple inference can be made from the value of $\rho_D(\cdot)$ at the end point $\pi/2$ of the domain, that is, from the autocorrelation of perpendicular diameters (Matérn 1956, 1990, personal communication in November 28, 1995): For an ellipse, $\rho_D(\pi/2)$ is very near $-1$ — in fact, $\rho_D(\pi/2)$ is a function of the axis ratio in the way that it tends to $-1$, although very slowly (cf. Matérn 1990, p. 16), as the axis ratio tends to 1. For a circle and other orbiforms, $\sigma_D^2=0$, and $\rho_D(\pi/2)$ thus becomes indefinite. Finally, for a square, $\rho_D(\pi/2)$ tends to $+1$. The fact that the angle between the minimum and maximum diameters is $\pi/2$ for the ellipse and $\pi/4$ for the square makes these results intuitively plausible. Note that $\rho_D(\pi/2)$ bears also some practical meaning, since if two diameters are measured in a tree, they are usually calipered at right angles to each other.

As mentioned before, the support function and the diameter function derived from it give no information on non-convexity, as they are defined for the convex closure of a cross-section. A polar co-ordinate representation of the boundary of a cross-section provides means for distinguishing between a non-convex cross-section and its convex closure and examining non-convexity. We set the rectangular planar co-ordinate system as before, by choosing an interior point O of the cross-section as the origin and by fixing the positive x-axis. For uniqueness, we need to assume that any ray emanating from O intersects the boundary only once; regarding tree cross-sections, this assumption of star-shapedness seems feasible, that is, the origin inside the cross-section can practically always be selected so that the condition is fulfilled. Now the polar co-ordinate representation of the boundary

---

**Fig. 4.** Four examples of orbiforms — the Reuleux triangle and the Reuleux pentagon in the upper row. The percentage indicates the proportion of the area to the area of the isoperimetric circle (Matérn 1956, Santaló 1976).

**Fig. 5.** Six examples of convex shapes provided by Matérn (1956) to illustrate the relation between geometrical shape and some diameter characteristics. Shape A is an ellipse with axis ratio 0.8. See Table 1 and Figs. 6, 7 and 8 for different characteristics of the shapes.
of the cross-section is a continuous function \( R : [0, 2\pi) \rightarrow (0, \infty) \), where \( R(\theta) \) is the uniquely determined radius of the cross-section in direction \( \theta \), that is, the distance between \( O \) and the point where the ray emanating from \( O \) at angle \( \theta \) intersects the boundary (Fig. 3). By means of \( R(\cdot) \), the perimeter of the cross-section can be expressed as

\[
C_p = \int_0^{2\pi} \sqrt{R(\theta)^2 + R'(\theta)^2} \, d\theta
\]  

(11)

and the area of the cross-section as

\[
A = \frac{1}{2} \int_0^{2\pi} R(\theta)^2 \, d\theta
\]  

(12)


Table 1. Support function \( p(\theta) \), diameter coefficient of variation \( \text{CV}_D \), ratio between minimum and maximum diameters \( D_{\text{min}}/D_{\text{max}} \), and correlation coefficient \( \rho_D(\pi/2) \) of the diameters intersecting at right angles for the shapes in Fig. 5 according to Matérn (1956).

<table>
<thead>
<tr>
<th>Shape</th>
<th>( p(\theta) )</th>
<th>( \text{CV}_D(%) )</th>
<th>( D_{\text{min}}/D_{\text{max}} )</th>
<th>( \rho_D(\pi/2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>((100\cos^2\theta + 64\sin^2\theta)^{1/2})</td>
<td>7.82</td>
<td>0.800</td>
<td>-0.9985</td>
</tr>
<tr>
<td>B</td>
<td>(9 + \cos(2\theta))</td>
<td>7.86</td>
<td>0.800</td>
<td>-1.0000</td>
</tr>
<tr>
<td>C</td>
<td>(16 + \cos(2\theta) + \cos(3\theta))</td>
<td>4.42</td>
<td>0.882</td>
<td>-1.0000</td>
</tr>
<tr>
<td>D</td>
<td>(32 + 2\cos(2\theta) + \cos(3\theta) + \cos(4\theta))</td>
<td>4.94</td>
<td>0.871</td>
<td>-0.6000</td>
</tr>
<tr>
<td>E</td>
<td>(35 + 2\cos(2\theta) + 2\sin(4\theta))</td>
<td>5.71</td>
<td>0.817</td>
<td>0.0000</td>
</tr>
<tr>
<td>F</td>
<td>(16 + \cos(4\theta))</td>
<td>4.42</td>
<td>0.882</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Fig. 6. Diameter functions \( D(\theta), \theta \in [0, \pi) \), scaled with the mean diameter \( \mu_D \), for the six example shapes in Fig. 5. The positive x-axis, with respect to which the direction \( \theta \) is determined, runs horizontally through the centre of gravity of the shape; \( \theta \) increases anticlockwise.
The sum $B(\theta) = R(\theta) + R(\theta + \pi)$ is here termed the *breadth* of the cross-section in direction $\theta$, $\theta \in [0, \pi)$; it is the length of the chord passing through the point O in direction $\theta$ (Fig. 3). Note that the term breadth has sometimes (e.g. Santaló 1976) been used as a synonym for the usual meaning of the width of a set, which, as said before, corresponds to the concept of diameter in this study (in Santaló 1976, however, the width of a convex set is defined as the least of the breadths, i.e., the concept of minimum diameter in this study).

Obviously, the radius and breadth functions depend on the selection of the reference point O. Two examples of a non-arbitrary choice for O are the centre of gravity and the Steiner point (the centre of the mass, distributed over the surface of a convex body, with density...
equal to the Gaussian curvature; see Hazewinkel 1992). Given \( O \), the radius function, as opposed to the diameter function, determines the shape of the cross-section uniquely; further, if \( O \) is the centre of gravity of the cross-section, the radius function being constant implies that the cross-section be a circle (Matérn 1956, Santaló 1976, Stoyan and Stoyan 1994).
3 Estimation of Cross-Section Area

In this chapter, we consider estimating the area of a tree stem cross-section with the circle area formula and somehow selected diameters; each area estimator is thus characterised by the diameter selection method, comprising both the way in which the diameters are selected and the way in which the selected diameters are combined (typically averaged) into the circle formula input. We study how variation in diameter within a non-circular cross-section is reflected to area estimates produced by different estimators; ultimately, we want to find out what can be said about the performance of different estimators without assuming anything about the shape of a cross-section. In the end, we briefly discuss area estimation based on radii.

3.1 Foundations: Source of Randomness and Measures of Estimator Performance

Within a cross-section, we regard the randomness in an area estimator as arising from the procedure of selecting diameters from a fixed, albeit infinite, “diameter population” (cf. the discussion in Chapter 2). Measuring a diameter in a direction sampled from the uniform distribution over $[0, \pi)$ means carrying out simple random sampling in the diameter population, whereas taking a diameter in a direction sampled from some non-uniform distribution means performing random sampling with unequal selection probabilities. Further, measuring an additional diameter perpendicular, or at any fixed angle, to a random diameter is systematic sampling with a random starting point. Yet, naturally, the diameter selection need not involve any randomness at all: taking fixed diameters, such as the maximum or the minimum diameter, or the mean diameter (the girth measurement divided by $\pi$), is just selective sampling that entails no randomness when carried out in a fixed diameter population. The area estimators based on a sample of random diameters are of course also random variables, whereas non-randomly selected diameters result in non-random area estimators when viewed from within a cross-section.

The sampling distribution of an area estimator within a cross-section is determined with respect to the diameter sampling design, that is, over all possible samples of diameters. If we select diameters by sampling their directions from a known direction distribution, we obtain the area estimator distribution via this direction distribution and the diameter function. The area estimator can be thought to be composed of a systematic part and a random part:

$$\hat{A}(\theta) = \mathbb{E}_{\theta} [\hat{A}(\theta)] + \epsilon(\theta),$$

where $\mathbb{E}_{\theta} [\hat{A}(\theta)]$ is the within-cross-section expectation of the area estimator taken over the diameter direction distribution and $\epsilon(\theta)$ is a random error term with zero expectation (and with the distribution determined by the diameter direction distribution and the diameter function). The argument $\theta$ here refers to the source of randomness in general; it can be thought of, for example, as a random vector containing the directions of the diameters included in the estimator.

The area estimation error now consists of the bias of the estimator and of the random error term:

$$\hat{A}(\theta) - A = \{\mathbb{E}_{\theta} [\hat{A}(\theta)] - A\} + \epsilon(\theta).$$

The bias, measuring how far the expected value of the estimator is from the true area, represents the systematic error associated with the estimator; note that this is the “mean error”
to be expected over repeated diameter samplings within a cross-section, not an error being realised in each individual sampling. The random error term, in turn, stands for the sampling error resulting from the fact that a diameter sample does not perfectly represent the diameter population but that there is variation between diameter samples that causes variation in estimator values. The usual measure for the magnitude of the sampling error \( \varepsilon(\theta) \), also termed the precision of the estimator, is its variance or standard deviation. The mean squared error 

\[
\mathbb{E}_\theta \{ [\hat{A}(\theta) - A]^2 \} = \{ \mathbb{E}_\theta [\hat{A}(\theta)] - A \}^2 + \mathbb{E}_\theta [\varepsilon(\theta)^2]
\]

(15)

combining the squared bias with the precision is a commonly used measure for the accuracy of an estimator (e.g. Lindgren 1976).

Trivially, non-random area estimators — those based on girth or fixed diameters, for example — are unaffected by sampling errors: for them, \( \mathbb{E}_\theta [\hat{A}(\theta)] = \hat{A}(\theta) \) and \( \text{Var}_\theta [\hat{A}(\theta)] = 0 \), and the mean squared error is reduced to the squared bias, that is, to the squared estimation error \( [\hat{A}(\theta) - A]^2 \).

An alternative to the design-based thinking discussed above could be a model-based approach where cross-sections were regarded as realisations of random figures or as stochastic deformations of a template curve and where stochastic models for the invariant parameters of the random contour functions were then established and estimated in empirical analysis (see e.g. Stoyan and Stoyan 1994 and Hobolth and Jensen 1999). In this thesis, however, we will examine the different area estimators expressly in a design-based way at the within-cross-section level. What was discussed above in terms of diameter-based estimators naturally apply to radius-based estimators as well.

### 3.2 Effect of Non-Convexity

As already mentioned in Chapter 2, the difference \( A_C - A \) between the convex area and the true area of a cross-section is always nonnegative. Aptly, Matérn (1956) termed this difference the convex deficit of a cross-section.

From the practical point of view, non-convexity is rather an insidious source of error, since it cannot be observed by a girth tape or a caliper commonly used for measuring standing trees. Thus nothing besides non-negativity can be inferred about convex deficit in an ordinary area estimation situation. It then becomes a valid question whether we had better use the convex area, instead of the true area, as the reference when computing the within-cross-section bias of an area estimator.

### 3.3 Estimators Based on Diameters and Circle Area Formula

As suggested many times above, an intuitively appealing and the most commonly used way to estimate cross-section area is to apply the circle area formula

\[
\hat{A} = \frac{\pi}{4} D^2 ,
\]

(16)

where \( D \) is the diameter of the circle that the cross-section is assumed to equate with. As already discussed, \( D \) can be chosen in a number of ways within a cross-section. Firstly,
Pulkkinen On Non-Circularity of Tree Stem Cross-Sections: Effect of Diameter Selection …

D may be a single measurable diameter: either random, that is, a diameter the direction of which is a random variable usually chosen to be uniformly distributed over \([0, \pi]\); or systematically sampled, such as the diameter perpendicular to a random diameter; or fixed, such as the minimum diameter or the maximum diameter or the mean diameter derived from the convex perimeter. Secondly, D may be the arithmetic, geometric, or quadratic mean of two or more randomly or systematically sampled or fixed diameters. Using the geometric mean of two diameters in the circle area formula implies the assumption of ellipticity, as this estimator yields the area of an ellipse with the axis lengths equal to the diameters used in the geometric mean. Employing the quadratic mean of two or more diameters, in turn, corresponds to estimating the area as the arithmetic mean of the areas of the circles that have the diameters used in the quadratic mean. Thirdly, D may as well be some other expression constructed from random and non-random diameters, such as the “diameter” involved in the area estimator where the geometric or the quadratic mean of two or more area estimators of the types mentioned above is written in the form of a circle area formula (Matérn 1956, 1990, Loetsch et al. 1973, Kärkkäinen 1974, 1975a, 1984).

In this study, we confine ourselves to estimators where D is a single measurable diameter or a mean of two or more measurable diameters. About the mutual relations between the three types of means, it is useful to remember the following two results: First, for positive variables \(X_i\), \(i=1, \ldots, n\),

\[
\left( \prod_{i=1}^{n} X_i \right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^{n} X_i \leq \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \right)^{\frac{1}{2}},
\]

that is, the geometric mean of is never greater than the arithmetic mean, which in turn is never greater than the quadratic mean (see Hardy et al. 1988 for the proofs); the equality between the means holds if all the variables \(X_i\) are equal. This implies that the area estimate given by the circle area formula with the geometric mean of unequal diameters be always less than the estimate obtained with the arithmetic mean of the same diameters, which in turn be always less than the estimate obtained with the quadratic mean of the diameters. Second, in the case of two positive variables \(X_1\) and \(X_2\), there exists the following relation between the squares of the quadratic, arithmetic and geometric means:

\[
\frac{X_1^2 + X_2^2}{2} - 2\left( \frac{X_1 + X_2}{2} \right)^2 + X_1X_2 = 0. \tag{18}
\]

This implies that the area estimate based on one of the three diameter means be straightforwardly obtained from the estimates based on the other two (Matérn 1956).

3.3.1 Girth Diameter: Mean Diameter Derived from Convex Perimeter

The non-random area estimator

\[
\hat{A}_0 = \frac{\pi}{4} \mu_0^2 = \frac{C^2}{4\pi}, \tag{19}
\]

where the mean diameter \(\mu_0=C/\pi\) (over the uniform direction distribution) derived from the convex perimeter \(C\) of the cross-section, termed here the girth diameter, is substituted
in the circle area formula, yields the area of the isoperimetric circle, that is, the area of the circle that has the perimeter equal to the convex perimeter of the cross-section.

For a non-circular cross-section, this estimator overestimates the convex area irrespective of the shape of the cross-section, because the circle has the largest area among the isoperimetric figures in plane (Hadwiger 1957). Matérn (1956) termed the nonnegative difference $\hat{A}_0 - A_C$ the isoperimetric deficit of a cross-section. Regardless of the shape of a cross-section, this deficit can be shown to have the following lower bound depending on diameter variance $\sigma_D^2$ within the cross-section:

$$\hat{A}_0 - A_C \geq \frac{3\pi}{4} \sigma_D^2 \quad (20)$$

(see Matérn 1956 for the proof).

Interestingly, $\hat{A}_0$ serves as a baseline for many estimators based on random diameters: in cross-sections with nonnegative correlation between perpendicular diameters, the overestimation error in $\hat{A}_0$ is a lower bound for the within-cross-section bias in those estimators, as we will see in the next section.

3.3.2 Random Diameters with Uniform Direction Distribution

Estimators Involving One or Two Diameters

In the way paved by Matérn (1956), we next consider the area estimators that are of the same form

$$\hat{A} = \frac{\pi}{4} D(\cdot)^2 \quad (21),$$

but where $D(\cdot)$ is now

1. random diameter $D(\theta), \theta \sim \text{Uniform}(0, \pi)$ ($\hat{A}_1$)
2. arithmetic mean of a random diameter $D(\theta)$ and the diameter $D(\theta + \pi/2)$ perpendicular to it ($\hat{A}_2$)
3. geometric mean of $D(\theta)$ and $D(\theta + \pi/2)$ ($\hat{A}_3$)
4. arithmetic mean of two independent random diameters $D(\theta_1)$ and $D(\theta_2)$, $\theta_1$, $\theta_2 \sim \text{Uniform}(0, \pi)$ i.i.d. ($\hat{A}_4$)
5. geometric mean of $D(\theta_1)$ and $D(\theta_2)$ ($\hat{A}_5$).

First we focus on the systematic errors, that is, on the within-cross-section biases of these estimators. With the notation introduced before — $\mu_D$ denoting the diameter mean, $\sigma_D^2$ denoting the diameter variance, and $\rho_D(\pi/2)$ denoting the correlation between perpendicular diameters within a cross-section — the expectations of the estimators over the uniform diameter direction distribution become as follows:

$$E(\hat{A}_i) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{4} \sigma_D^2$$

$$= \hat{A}_0 + \frac{\pi}{4} \sigma_D^2 \quad (22),$$
Pulkkinen On Non-Circularity of Tree Stem Cross-Sections: Effect of Diameter Selection ...

\[ E(\hat{A}_2) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{8} \sigma_D^2 \left[ 1 + \rho_D \left( \frac{\pi}{2} \right) \right] \]

\[ = \hat{A}_0 + \frac{\pi}{8} \sigma_D^2 \left[ 1 + \rho_D \left( \frac{\pi}{2} \right) \right], \quad (23) \]

\[ E(\hat{A}_3) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{4} \sigma_D^2 \rho_D \left( \frac{\pi}{2} \right) \]

\[ = \hat{A}_0 + \frac{\pi}{4} \sigma_D^2 \rho_D \left( \frac{\pi}{2} \right), \quad (24) \]

\[ E(\hat{A}_4) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{8} \sigma_D^2 \]

\[ = \hat{A}_0 + \frac{\pi}{8} \sigma_D^2 , \quad (25) \]

and

\[ E(\hat{A}_5) = \frac{\pi}{4} \mu_D^2 \]

\[ = \hat{A}_0 \]

(cf. Matérn 1956, 1990). (The expectations are obtained by writing the area estimators as sums of squared diameters and diameter products, by taking the expectations separately on each term in the sum, and by applying to these the usual rules that relate means, variances and correlations to each other: \( E[D(\theta)^2] = E[D(\theta+\pi/2)^2] = \mu_D^2 + \sigma_D^2 \), \( E[D(\theta)D(\theta+\pi/2)] = \mu_D^2 + \sigma_D^2 \rho_D(\pi/2) \), \( E[D(\theta_1)D(\theta_2)] = E[D(\theta_1)]E[D(\theta_2)] = \mu_D^2 \). Clearly, the estimator \( \hat{A}_0 \) based on the convex perimeter of a cross-section, which in the previous section was found to overestimate the convex area of a cross-section, now makes a convenient reference. Interestingly, the estimator \( \hat{A}_5 \) based on the geometric mean of two independent random diameters comprises a bias equal to that of the estimator \( \hat{A}_0 \). Moreover, the bias of the estimator \( \hat{A}_2 \) is the arithmetic mean of the biases of the estimators \( \hat{A}_1 \) and \( \hat{A}_3 \) (Matérn 1956). In Table 2, the expectations of the estimators are given for the six example shapes of Fig. 5.

<table>
<thead>
<tr>
<th>Shape</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{A}_0/A ) (‰)</td>
<td>1019</td>
<td>1019</td>
<td>1022</td>
<td>1017</td>
<td>1030</td>
<td>1030</td>
</tr>
<tr>
<td>( E(\hat{A}_1)/A ) (‰)</td>
<td>1025</td>
<td>1025</td>
<td>1024</td>
<td>1020</td>
<td>1034</td>
<td>1032</td>
</tr>
<tr>
<td>( E(\hat{A}_2)/A ) (‰)</td>
<td>1019</td>
<td>1019</td>
<td>1022</td>
<td>1018</td>
<td>1032</td>
<td>1032</td>
</tr>
<tr>
<td>( E(\hat{A}_3)/A ) (‰)</td>
<td>1013</td>
<td>1013</td>
<td>1020</td>
<td>1016</td>
<td>1030</td>
<td>1032</td>
</tr>
<tr>
<td>( E(\hat{A}_4)/A ) (‰)</td>
<td>1022</td>
<td>1022</td>
<td>1023</td>
<td>1019</td>
<td>1032</td>
<td>1031</td>
</tr>
</tbody>
</table>
The mutual ranking of the estimators in terms of bias obviously depends on the values of \( \sigma_D^2 \) and \( p_D(\pi/2) \) within a cross-section. With the condition \( \sigma_D^2 > 0 \) — that is, orbiforms excluded — and by recalling that \( -1 \leq p_D(\pi/2) \leq 1 \), we can compose the following \( p_D(\pi/2) \)-dependent comparisons between the estimator expectations:

\[
\begin{align*}
&\text{\( p_D(\pi/2) = -1 \):} & E(\hat{A}_3) < E(\hat{A}_0) = E(\hat{A}_4) < E(\hat{A}_2) < E(\hat{A}_1) \\
&\text{\( -1 < p_D(\pi/2) < 0 \):} & E(\hat{A}_3) < E(\hat{A}_0) = E(\hat{A}_4) < E(\hat{A}_2) < E(\hat{A}_1) \\
&\text{\( p_D(\pi/2) = 0 \):} & \hat{A}_0 = E(\hat{A}_3) = E(\hat{A}_4) = E(\hat{A}_2) < E(\hat{A}_1) \\
&\text{\( 0 < p_D(\pi/2) < \frac{1}{2} \):} & \hat{A}_0 = E(\hat{A}_3) < E(\hat{A}_4) < E(\hat{A}_2) < E(\hat{A}_1) \\
&\text{\( \frac{1}{2} \leq p_D(\pi/2) < 1 \):} & \hat{A}_0 = E(\hat{A}_3) < E(\hat{A}_4) < E(\hat{A}_2) < E(\hat{A}_1) \\
&\text{\( p_D(\pi/2) = 1 \):} & \hat{A}_0 = E(\hat{A}_3) < E(\hat{A}_4) < E(\hat{A}_2) = E(\hat{A}_1)
\end{align*}
\]

We find here that with the nonnegative values of \( p_D(\pi/2) \), all the estimators have expectations greater or equal to that of \( \hat{A}_0 \); in consequence, all the estimators yield a positive bias, that is, they systematically overestimate the area of a cross-section. With the negative values of \( p_D(\pi/2) \) this does not hold for the estimator \( \hat{A}_3 \); however, \( E(\hat{A}_3) \) attains its smallest value when \( p_D(\pi/2) = -1 \), and by combining this minimum with the lower bound previously obtained for \( \hat{A}_0 \) (Eq. 20) we get

\[
E(\hat{A}_1) \geq \hat{A}_1 - \frac{\pi}{4} \sigma_D^2 \geq \left( A_c + \frac{3\pi}{4} \sigma_D^2 \right) - \frac{\pi}{4} \sigma_D^2 = A_c + \frac{\pi}{2} \sigma_D^2,
\]

which shows that also \( \hat{A}_3 \) always systematically overestimates the convex area of a cross-section (Matérn 1956). We hence conclude that regardless of the shape of a non-circular cross-section all these estimators systematically overestimate the convex area of the cross-section. In the case of orbiforms this conclusion also stands, since a circle is the largest in area among the orbiforms (cf. Fig. 4). (Note that this systematic overestimation motivates the omission of the estimators based on the quadratic mean of diameters, which yield the largest estimates by definition, from our considerations.)

The comparison above also shows that regardless of the value of \( p_D(\pi/2) \) and thus irrespective of the shape of a cross-section, the estimator \( \hat{A}_1 \) based on only one random diameter yields the largest bias. The best strategy in terms of bias minimisation, however, depends on the shape of a cross-section: with negatively correlated perpendicular diameters \( -1 < p_D(\pi/2) < 0 \), the use of the geometric mean is preferred and, given this, the second diameter is recommended to be taken at right angles to the first one instead of picking it randomly; with mildly positively correlated perpendicular diameters \( 0 < p_D(\pi/2) < 1/2 \), the preference of the geometric mean still pertains, but now the second diameter is more advisable to be taken independently at random direction; and finally, with the highly positively correlated perpendicular diameters \( 1/2 < p_D(\pi/2) < 1 \), the priority is given to measuring both the diameters independently at random directions, and, conditional on this, to employing the geometric mean. The equalities in the change points are also interesting: if \( p_D(\pi/2) = -1 \) (ellipse-like cross-sections, cf. shapes A–C in Fig. 5 and Table 2), employing the arithmetic mean of two perpendicular diameters is equivalent, in terms of bias, of using the geometric mean of two random diameters; if \( p_D(\pi/2) = 0 \) (cf. shape E in Fig. 5 and Table 2), the way of choosing diameters does not matter, but the estimators based on the same type of mean yield an equal bias; if \( p_D(\pi/2) = 1/2 \), using the geometric mean of two perpendicular diameters equals, in terms of bias, using the arithmetic mean of two random diameters; and, if \( p_D(\pi/2) = 1 \) (square-like cross-sections, cf. shape F in Fig. 5 and Table 2), the estimators based on two perpendicular diameters yield as large a bias as the estimator based on only one random diameter.
The optimal selection of an estimator may, however, contribute to bias reduction only marginally, since, quoting Matérn (1956), "the average difference between the 'worst' and the 'best' estimate cannot be greater than the average difference between the 'best' estimate and the area of the convex closure": if \( 0 \leq \rho D(\pi/2) \leq 1 \), then \( \hat{A}_1 \) is the worst and \( \hat{A}_0 \) is the best estimator, and

\[
E(\hat{A}_1 - \hat{A}_0) = E(\hat{A}_1) - \hat{A}_0 = \left( \hat{A}_0 + \frac{\pi}{4} \sigma_D^2 \right) - \hat{A}_0 = \frac{\pi}{4} \sigma_D^2 \\
< \frac{3\pi}{4} \sigma_D^2 = \left( \hat{A}_c + \frac{3\pi}{4} \sigma_D^2 \right) - \hat{A}_c \\
\leq \hat{A}_0 - \hat{A}_c = E(\hat{A}_0 - \hat{A}_c) \quad \text{if} \quad \rho D(\pi/2) \leq 1
\]

if, in turn, \(-1 \leq \rho D(\pi/2) < 0\), then \( \hat{A}_1 \) is the worst and \( \hat{A}_3 \) is the best estimator, and

\[
E(\hat{A}_1 - \hat{A}_3) = E(\hat{A}_1) - E(\hat{A}_3) = \left( \hat{A}_0 + \frac{\pi}{4} \sigma_D^2 \right) - \left[ \hat{A}_0 + \frac{\pi}{4} \sigma_D^2 \rho_D \left( \frac{\pi}{2} \right) \right] \\
= \frac{\pi}{4} \sigma_D^2 \left[ 1 - \rho_D \left( \frac{\pi}{2} \right) \right] \\
\leq \frac{\pi}{2} \sigma_D^2 = \left( \hat{A}_c + \frac{\pi}{2} \sigma_D^2 \right) - \hat{A}_c \\
\leq E(\hat{A}_1) - \hat{A}_c = E(\hat{A}_3 - \hat{A}_c)
\]

(in both these elaborations, Eq. 20 giving the lower bound for \( \hat{A}_0 \) is employed, and in the latter derivation also Eq. 27 giving the lower bound for \( \hat{A}_3 \) is used).

From the systematic errors we proceed into the sampling errors quantified by the within-cross-section variances of the estimators over the uniform direction distribution. Except for the estimator \( \hat{A}_5 \) based on the geometric mean of two independent diameters, the variances cannot be expressed in terms of the simple parameters \( \mu_D, \sigma_D^2 \), and \( \rho_D(\pi/2) \), but involve higher moments and product moments of diameters, which make them somewhat difficult to compare with each other:

\[
\text{Var}(\hat{A}_1) = \frac{\pi^2}{16} \left\{ E[D(\theta)^4] - \left( \mu_D^2 + \sigma_D^2 \right)^2 \right\}, \\
\text{Var}(\hat{A}_2) = \frac{\pi^2}{16} \left\{ \frac{1}{8} E[D(\theta)^4] + \frac{1}{2} E[D(\theta)^3] \left[ D(\theta + \frac{\pi}{2}) \right] + \frac{3}{8} E[D(\theta)^2] \left[ D(\theta + \frac{\pi}{2}) \right]^2 \right\} \\
- \frac{1}{4} \left\{ 2\mu_D^2 + \sigma_D^2 \left[ 1 + \rho_D \left( \frac{\pi}{2} \right) \right] \right\}^2
\]
\[
\text{Var}(\hat{A}_1) = \frac{\pi^2}{16} \left\{ E \left[ \left( D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)\right)^2 \right] - \left[ \mu_D^2 + \sigma_D^2 \rho_D \left( \frac{\pi}{2} \right) \right]^2 \right\}, \quad (32)
\]

\[
\text{Var}(\hat{A}_4) = \frac{\pi^2}{16} \left\{ \frac{1}{8} E[D(\theta)^4] + \frac{1}{2} \mu_D E[D(\theta)^2] + \frac{3}{8} \left( \mu_D^2 + \sigma_D^2 \right)^2 \right\} - \frac{1}{4} \left( 2\mu_D^2 + \sigma_D^2 \right)^2, \quad (33)
\]

and

\[
\text{Var}(\hat{A}_5) = \frac{\pi^2}{16} \left\{ \left( \mu_D^2 + \sigma_D^2 \right)^2 - \mu_D^4 \right\}. \quad (34)
\]

(The variances are obtained by applying the definition of variance Var(X) = E(X^2) - E(X)^2 straightforwardly to the squared mean of diameters in the estimator, by writing the expressions inside the expectation operators as sums of diameter powers and products of diameter powers, by taking the expectations separately on each term in the sums, and by applying to these the usual rules that relate means, variances and correlations to each other: E[D(\theta)^2] = E[D(\theta + \pi/2)^2] = \mu_D^2 + \sigma_D^2, E[D(\theta)D(\theta + \pi/2)] = \mu_D^2 + \sigma_D^2 \rho_D(\pi/2), E[D(\theta)^aD(\theta + \pi/2)^b] = E[D(\theta)^a]E[D(\theta + \pi/2)^b], E[D(\theta_1)^aD(\theta_2)^b] = E[D(\theta_1)^a]E[D(\theta_2)^b].) We can see that if \( \rho_D(\pi/2) = 0 \), the estimators \( \hat{A}_2 \) and \( \hat{A}_4 \), on one hand, and the estimators \( \hat{A}_3 \) and \( \hat{A}_5 \), on the other hand, coincide in terms of variance quite as we saw them to do in terms of bias; in other words, if perpendicular diameters in a cross-section are uncorrelated, the way of selecting the diameters does not influence the sampling error but the type of mean used in the area estimator does. In Table 3, the standard deviations (square roots of variances) of the estimators are given for the six example shapes of Fig. 5.

By the procedure known as the delta method (see e.g. Casella and Berger 1990) — based on the linearisation of a non-linear function by Taylor series expansion — the following variance approximations expressible in terms of only \( \mu_D, \sigma_D^2 \), and \( \rho_D(\pi/2) \) are found:

\[
\text{Var}(\hat{A}_1) = \frac{\pi^2}{4} \mu_D^2 \sigma_D^2
\]

\[
= \pi \hat{A}_0 \sigma_D^2, \quad (35)
\]

\[
\text{Var}(\hat{A}_4) = \text{Var}(\hat{A}_5) = \frac{\pi^2}{8} \mu_D^2 \sigma_D^2
\]

\[
= \frac{\pi}{2} \hat{A}_0 \sigma_D^2, \quad (36)
\]

\[
\text{Var}(\hat{A}_4) = \text{Var}(\hat{A}_5) = \frac{\pi^2}{8} \mu_D^2 \sigma_D^2
\]

\[
= \frac{\pi}{2} \hat{A}_0 \sigma_D^2. \quad (37)
\]
Pulkkinen On Non-Circularity of Tree Stem Cross-Sections: Effect of Diameter Selection …

(For the estimators $\hat{A}_1$–$\hat{A}_4$, Matérn (1956, 1990) presented, without reference to the delta method or any derivation, approximate standard deviations, the squares of which are seen to correspond to these approximate variances.) The nonnegative difference between the actual and approximate variance of the estimator $\hat{A}_5$

\[
\text{Var}(\hat{A}_5) - \text{Var}(\hat{A}_5) = \frac{\pi^2}{16} \left[ \left( \mu_D^2 + \sigma_D^2 \right)^2 - \mu_D^4 \right] - \frac{\pi^2}{8} \mu_D^2 \sigma_D^2 = \frac{\pi^2}{16} \sigma_D^4 \quad (38)
\]

suggests that these approximate variances might underestimate the true ones; in the example shapes of Fig. 5, however, the approximate variances practically equal the true variances (Table 3).

Unlike the biases and the true variances, the approximate variances are invariant of the type of the mean employed in the estimator — only the number and the way of selecting the diameters influence the approximate sampling error of the estimator; consequently, in terms of approximate variance, it would appear natural to favour the estimators based on the geometric mean, as they yield smaller bias than those based on the arithmetic mean.

By the possible values of $\rho_D(\pi/2)$, the variance approximations can be ranked as follows:

<table>
<thead>
<tr>
<th>$\rho_D(\pi/2)$</th>
<th>$\text{Var}(\hat{A}_2)$</th>
<th>$\text{Var}(\hat{A}_3)$</th>
<th>$\text{Var}(\hat{A}_4)$</th>
<th>$\text{Var}(\hat{A}_5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1 &lt; \rho_D(\pi/2) &lt; 0$</td>
<td>$\text{Var}(\hat{A}_2)$ = $\text{Var}(\hat{A}_3)$ &lt; $\text{Var}(\hat{A}_4)$ = $\text{Var}(\hat{A}_5)$ &lt; $\text{Var}(\hat{A}_1)$</td>
<td>$\text{Var}(\hat{A}_2)$ = $\text{Var}(\hat{A}_3)$ = $\text{Var}(\hat{A}_4)$ = $\text{Var}(\hat{A}_5)$ &lt; $\text{Var}(\hat{A}_1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_D(\pi/2) = 0$</td>
<td>$\text{Var}(\hat{A}_2)$ = $\text{Var}(\hat{A}_3)$ = $\text{Var}(\hat{A}_4)$ = $\text{Var}(\hat{A}_5)$ &lt; $\text{Var}(\hat{A}_1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0 &lt; \rho_D(\pi/2) &lt; 1$</td>
<td>$\text{Var}(\hat{A}_4)$ = $\text{Var}(\hat{A}_5)$ &lt; $\text{Var}(\hat{A}_2)$ = $\text{Var}(\hat{A}_3)$ &lt; $\text{Var}(\hat{A}_1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_D(\pi/2) = 1$</td>
<td>$\text{Var}(\hat{A}_4)$ = $\text{Var}(\hat{A}_5)$ &lt; $\text{Var}(\hat{A}_2)$ = $\text{Var}(\hat{A}_3)$ = $\text{Var}(\hat{A}_1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This comparison suggests that when measuring two diameters in a cross-section, one should take the second diameter at right angles to the first one, if $-1 < \rho_D(\pi/2) < 0$, and at random direction, if $0 < \rho_D(\pi/2) < 1$, in order to minimise the approximate sampling error. Noteworthily, this policy agrees with minimising the bias, as was shown by the previous comparison of the within-cross-section expectations of the estimators.

The equalities in the change points of the above comparison are also somewhat interesting: If perpendicular diameters in a cross-section are uncorrelated ($\rho_D(\pi/2) = 0$), also the effect of the diameter selection method vanishes: all the two-diameter estimators $\hat{A}_2$, $\hat{A}_3$, $\hat{A}_4$, and $\hat{A}_5$ have the same approximate sampling error, half of that of the estimator $\hat{A}_1$. If, in turn, perpendicular diameters are fully positively correlated ($\rho_D(\pi/2) = 1$), taking the
second diameter at right angles to the first one benefits none compared to measuring only one
diameter but the approximate variance is halved only by measuring the two diameters inde-
pendently. Finally, if perpendicular diameters are fully negatively correlated ($\rho_D(\pi/2)=-1$),
the approximate sampling error for the estimators $\hat{A}_2$ and $\hat{A}_3$ disappears (cf. shapes B and
C in Table 3); hence, for ellipse-like cross-sections, measuring two perpendicular diameters
should always lead to an approximately sampling-error-free area estimate.

Generalisation to $n$ Diameters

The area estimators based on a mean of two diameters can naturally be generalised to involve
countably many diameters. For the generalised estimators involving the arithmetic mean,
the within-cross-section expectations are expressible in terms of the simple parameters $\mu_D$,
$\sigma_D^2$, and $\rho_D(\pi/2)$: the expectation of the estimator based on the arithmetic mean of
the $n$ independent random diameters (a generalisation of $\hat{A}_4$ and thus denoted by $\hat{A}_{G4}$) becomes

$$E(\hat{A}_{G4}) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{4n} \sigma_D^2$$

$$= \hat{A}_0 + \frac{\pi}{4n} \sigma_D^2,$$

and the expectation of the estimator based on the arithmetic mean of $n/2$, $n=2k$, $k\in\mathbb{N}$,
independent random diameters and their $n/2$ perpendiculars (a generalisation of $\hat{A}_2$ and
thus denoted by $\hat{A}_{G2}$) is found to be

$$E(\hat{A}_{G2}) = \frac{\pi}{4} \mu_D^2 + \frac{\pi}{4n} \sigma_D^2 \left[ 1 + \rho_D \left( \frac{\pi}{2} \right) \right]$$

$$= \hat{A}_0 + \frac{\pi}{4n} \sigma_D^2 \left[ 1 + \rho_D \left( \frac{\pi}{2} \right) \right],$$

(cf. the commonly known theorem about the mean of quadratic forms, to be found in e.g.
Kendall and Stuart 1979). If perpendicular diameters are uncorrelated, these general esti-
mators clearly coincide in terms of bias. Further, as $n$ tends to infinity, the expectations of
both the general estimators tend to $\hat{A}_0$; the biases of the estimators can hence be reduced
by measuring more diameters, but the biases can never fall below that of $\hat{A}_0$.

The expectations of the generalised estimators based on the geometric mean of more than
two diameters cannot be expressed in terms of $\mu_D$, $\sigma_D^2$, and $\rho_D(\pi/2)$, but involve “fractional
diameter moments”, that is, expectations of fractional powers of diameters over the uniform
direction distribution, which make them difficult to compare with those presented above:
the expectation of the estimator based on the geometric mean of $n$ independent random
diameters (a generalisation of $\hat{A}_5$ and thus denoted by $\hat{A}_{G5}$) becomes

$$E(\hat{A}_{G5}) = \frac{\pi}{4} \left\{ E \left[ D(\theta)^2 \right]^\frac{2}{n} \right\}^n$$

$$= \hat{A}_0 \left\{ E \left[ D(\theta)^2 \right]^\frac{2}{n} \right\}^n,$$

and the expectation of the estimator based on the geometric mean of $n/2$ independent
random diameters and their $n/2$ perpendiculars (a generalisation of $\hat{A}_3$ and thus denoted
by $\hat{A}_{G3}$) is found to be
\[
E(\hat{A}_{G3}) = \frac{\pi}{4} \left\{ \mathbb{E} \left[ D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)^n \right] \right\}^\frac{1}{n}.
\]

(42)

For the generalised area estimators based on the arithmetic mean, the variances become somewhat intricate:

\[
\text{Var}(\hat{A}_{G4}) = \frac{\pi^2}{16n^2} \left\{ n \mathbb{E}[D(\theta)^4] + 4n(n-1)\mu_D \mathbb{E}[D(\theta)^3] + n(n-3)\left(\mu_D^2 + \sigma_D^2\right)^2 \right. \\
+ \left. \left[ 6n(n-1)(n-2) - 2n^2(n-1) \right] \mu_D^2 \left(\mu_D^2 + \sigma_D^2\right) \right. \\
+ \left. \left[ n(n-1)(n-2)(n-3) - n^2(n-1)^2 \right] \mu_D^4 \right\},
\]

(43)

and

\[
\text{Var}(\hat{A}_{G2}) = \frac{\pi^2}{16n^2} \left\{ n \mathbb{E}[D(\theta)^4] + 4n \mathbb{E}[D(\theta)^3 D\left(\theta + \frac{\pi}{2}\right)] + 3n \mathbb{E}[D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)^2] \right. \\
+ \left. 4n(n-2)\mu_D \left(\mathbb{E}[D(\theta)^3] + 3\mathbb{E}[D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)] \right) \right. \\
+ \left. 3n(n-2) \left\{ 2\mu_D^2 + \sigma_D^2 \left[ 1 + \rho_D \left(\frac{\pi}{2}\right) \right] \right\}^2 \right. \\
+ \left. 6n(n-2)(n-4)\mu_D^2 \left\{ 2\mu_D^2 + \sigma_D^2 \left[ 1 + \rho_D \left(\frac{\pi}{2}\right) \right] \right\} \right. \\
+ \left. n(n-2)(n-4)(n-6)\mu_D^4 \right. \\
- \left. \left\{ n \left\{ 2\mu_D^2 + \sigma_D^2 \left[ 1 + \rho_D \left(\frac{\pi}{2}\right) \right] \right\} + n(n-2)\mu_D^2 \right\} \right\},
\]

(44)

For the generalisations based on the geometric mean, the variance expressions are far simpler on the face but again relatively uninformative if one wants to compare them to those above:

\[
\text{Var}(\hat{A}_{G3}) = \frac{\pi^2}{16} \left\{ \mathbb{E} \left[ D(\theta)^4 \right]^n - \left\{ \mathbb{E} \left[ D(\theta)^2 \right]^n \right\} \right\},
\]

(45)

and

\[
\text{Var}(\hat{A}_{G3}) = \frac{\pi^2}{16} \left\{ \mathbb{E} \left[ D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)^2 \right]^n - \left\{ \mathbb{E} \left[ D(\theta)^2 D\left(\theta + \frac{\pi}{2}\right)^2 \right]^n \right\} \right\}.
\]

(46)
Contrary to these actual variances, the delta method approximations are both uncomplicated and interpretative:

\[
\text{Vr}(\hat{A}_{G4}) = \text{Vr}(\hat{A}_{G5}) = \frac{\pi^2}{4n} \mu_D^2 \sigma_D^2
\]

\[
= \frac{\pi}{n} \hat{A}_0 \sigma_D^2,
\]

and

\[
\text{Vr}(\hat{A}_{G2}) = \text{Vr}(\hat{A}_{G1}) = \frac{\pi^2}{4n} \mu_D^2 \sigma_D^2 \left[1 + \rho_D \left(\frac{\pi}{2}\right)\right]
\]

\[
= \frac{\pi}{n} \hat{A}_0 \sigma_D^2 \left[1 + \rho_D \left(\frac{\pi}{2}\right)\right].
\]

Trivially, by measuring more diameters one reduces the true and the approximate variances of the generalised estimators, up to zero as \(n\) tends to infinity. Otherwise the generalised estimators share the properties previously demonstrated for the two-diameter estimators: If perpendicular diameters are uncorrelated in a cross-section, the generalised estimators coincide pairwise in terms of true sampling errors (\(\text{Var}(\hat{A}_{G3}) = \text{Var}(\hat{A}_{G5}) = \text{Var}(\hat{A}_{G2}) = \text{Var}(\hat{A}_{G4})\)) and all together in terms of approximate sampling errors (\(\text{Vr}(\hat{A}_{G3}) = \text{Vr}(\hat{A}_{G5}) = \text{Vr}(\hat{A}_{G2}) = \text{Vr}(\hat{A}_{G4})\)). Further, in the case of an elliptic cross-section, for which \(\rho_D(\pi/2)\) is near \(-1\), the approximate variances of \(\hat{A}_{G2}\) and \(\hat{A}_{G3}\) become close to zero independent of the value of \(n\), and thus increasing the number of diameter measurements within the cross-section should not make much difference in terms of the approximate sampling error.

### 3.3.3 Minimum and Maximum Diameters

From the estimators involving random diameters with uniformly distributed directions, we proceed to discussing six fixed estimators based on the minimum and maximum diameters of a cross-section. The estimators are still of the form

\[
\hat{A} = \frac{\pi}{4} D^2,
\]

where \(D\) is now, following Matérn (1956),

6. arithmetic mean of the minimum diameter \(D_{\min}\) and the maximum diameter \(D_{\max}\) in the cross-section \(\hat{A}_6\)
7. geometric mean of \(D_{\min}\) and \(D_{\max}\) \(\hat{A}_7\)
8. arithmetic mean of \(D_{\min}\) and its perpendicular \(\hat{A}_8\)
9. geometric mean of \(D_{\min}\) and its perpendicular \(\hat{A}_9\)
10. arithmetic mean of \(D_{\max}\) and its perpendicular \(\hat{A}_{10}\)
11. geometric mean of \(D_{\max}\) and its perpendicular \(\hat{A}_{11}\).

Note that the estimators \(\hat{A}_8--\hat{A}_{11}\) may not always be well-defined — consider, for example, the situation where \(D_{\min}\) is attained in two different directions but the diameters perpendicular to these diameters are unequal (Matérn 1956). If the estimators are well-defined,
they are also fixed, as the randomness induced by a selection between the alternative diameters is avoided. In Table 4, the values of the estimators are given for the six example shapes of Fig. 5.

Since the geometric mean is never greater than the arithmetic mean for nonnegative measurements and since $D^2$ is a monotonously increasing function for $D>0$, the above estimators can be ranked pairwise as $\hat{\Theta}_6 \geq \hat{\Theta}_7$, $\hat{\Theta}_8 \geq \hat{\Theta}_9$, and $\hat{\Theta}_{10} \geq \hat{\Theta}_{11}$. In fact, the larger is the difference between the diameters to be averaged, the more pronounced becomes the reducing effect of using the geometric mean instead of the arithmetic one: if we denote by $\beta$ the ratio of the smaller diameter to the larger one, the ratio between the estimator involving the arithmetic mean and the estimator involving the geometric mean becomes $(\beta+1)^2/4\beta$ (cf. Matérn 1990), which is a monotonously decreasing function of $\beta$ in $(0, 1]$. Furthermore, since the diameter perpendicular to $D_{\text{max}}$ cannot be smaller than $D_{\text{min}}$ in a cross-section and, similarly, since the diameter perpendicular to $D_{\text{min}}$ cannot be larger than $D_{\text{max}}$, we can straightforwardly infer that $\hat{\Theta}_{10} \geq \hat{\Theta}_6 \geq \hat{\Theta}_8$, and $\hat{\Theta}_{11} \geq \hat{\Theta}_7 \geq \hat{\Theta}_9$. The equalities hold for the cross-sections that have $D_{\text{max}}$ and $D_{\text{min}}$ at right angles to each other. For ellipses, the estimators $\hat{\Theta}_7$, $\hat{\Theta}_9$ and $\hat{\Theta}_{11}$ naturally give the true area, whereas, contrary to what Matérn (1956) claims, the estimators $\hat{\Theta}_6$, $\hat{\Theta}_8$ and $\hat{\Theta}_{10}$ yield the arithmetic mean of the true area and the estimate involving the quadratic mean of $D_{\text{max}}$ and $D_{\text{min}}$.

However, no general inference independent of cross-section shape can be made on how these area estimators relate to the estimators based on girth or one or more random diameters with uniformly distributed directions; unlike those estimators, the estimators involving extreme diameters may also underestimate the convex area of a cross-section (cf. estimators $\hat{\Theta}_8$ and $\hat{\Theta}_9$ with shapes D–F in Table 4).

3.4 Estimators Based on Radii

Lastly, we digress a little from the circle area formula and diameters and briefly discuss area estimation based on radial information. In the following, we assume a cross-section to be star-shaped so that all the radii from a suitably chosen interior point $O$ of the cross-section can be uniquely determined (cf. Chapter 2).

The true area of a cross-section, expressed as a function of the polar co-ordinate representation $R(\cdot)$ of the contour of a cross-section (see Chapter 2), can alternatively be viewed as the product of $\pi$ and the expectation of the squared radius over the uniform direction distribution in $[0, 2\pi)$:

$$A = \frac{1}{2} \int_0^{2\pi} R(\theta)^2 d\theta = \pi \frac{1}{2\pi} \int_0^{2\pi} R(\theta)^2 d\theta = \pi E[R(\theta)^2]. \quad (50)$$

Table 4. Area estimates produced by the estimators $\hat{\Theta}_6$–$\hat{\Theta}_{11}$ for the shapes in Fig. 5, expressed in permille of true area (cf. Matérn 1956).

<table>
<thead>
<tr>
<th>Shape</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\Theta}_6/A$ (%)</td>
<td>1013</td>
<td>1019</td>
<td>1022</td>
<td>1066</td>
<td>1030</td>
<td>1030</td>
</tr>
<tr>
<td>$\hat{\Theta}_7/A$ (%)</td>
<td>1000</td>
<td>1006</td>
<td>1018</td>
<td>1061</td>
<td>1020</td>
<td>1026</td>
</tr>
<tr>
<td>$\hat{\Theta}_8/A$ (%)</td>
<td>1013</td>
<td>1019</td>
<td>1022</td>
<td>986</td>
<td>921</td>
<td>905</td>
</tr>
<tr>
<td>$\hat{\Theta}_9/A$ (%)</td>
<td>1000</td>
<td>1006</td>
<td>1018</td>
<td>985</td>
<td>919</td>
<td>905</td>
</tr>
<tr>
<td>$\hat{\Theta}_{10}/A$ (%)</td>
<td>1013</td>
<td>1019</td>
<td>1022</td>
<td>1082</td>
<td>1145</td>
<td>1163</td>
</tr>
<tr>
<td>$\hat{\Theta}_{11}/A$ (%)</td>
<td>1000</td>
<td>1006</td>
<td>1018</td>
<td>1078</td>
<td>1143</td>
<td>1163</td>
</tr>
</tbody>
</table>
These expressions suggest an area estimator of the type

\[ \hat{A} = \frac{1}{2} \sum_{i=1}^{n} R(\theta_i)^2 \frac{2\pi}{n} = \frac{\pi}{n} \sum_{i=1}^{n} R(\theta_i)^2, \]  

(51)

where \( R(\theta_i) \) is the radius of a cross-section measured in direction \( \theta_i \), \( i = 1, 2, ..., n \). In the first expression, the definite integral of \( R(\theta)^2 \) is obviously approximated by the sum of \( n \) rectangles with the width \( 2\pi/n \) and the height \( R(\theta_i)^2 \). This corresponds to estimating the area of the cross-section with the sum of \( n \) disjoint, equally wide circular sectors (the area of a circular sector with angle \( \Delta \theta_i = 2\pi/n \) and side length \( R(\theta_i) \) is \( \Delta \theta_i R(\theta_i)^2/2 = \pi R(\theta_i)^2/n \), and summing up \( n \) such sectors then results in the second expression). This expression is straightforwardly seen to be \( \pi \) times the squared quadratic mean of \( n \) radii \( R(\theta_i) \), a natural estimator for \( \pi \) times the expectation of the squared radius over the uniform direction distribution.

The application of this estimator does not require that the radii be measured equidistantly. Just as the diameters in the estimators based on the circle area formula can be chosen in many ways, also for the radii various selection strategies are naturally available (see e.g. Gregoire and Valentine 1995). The remarkable property of the estimator of Eq. 51 is that it is unbiased if the marginal distribution of each radius direction \( \theta_i \) is uniform over \([0, 2\pi)\) (Matérn 1956, Gregoire and Valentine 1995). This condition naturally holds if the directions are independently sampled from the uniform distribution over \([0, 2\pi)\), in which case it is also easy to derive analytically the exact variance for the estimator and also to find an unbiased estimator for that variance (Gregoire and Valentine 1995). However, the variance is directly related to the variation in the radii used and thus considerably affected by the choice of the location of \( O \); for decreasing the variance of the estimator, centralising the location of \( O \) as well as applying systematic sampling strategies for radii have been suggested (Gregoire and Valentine 1995).

In sum, the bias in area estimation can be avoided, if the cross-sections can be observed “from inside” instead of “from outside”, that is, if radii instead of diameters can be measured. For standing trees this is not feasible with the standard measurement equipment, but to estimating the area of end sections of logs — or the area of the crown projection of a tree — for instance, this approach certainly pertains (Matérn 1956).
4 Bitterlich Sampling

Bitterlich sampling — also referred to as relascope sampling, angle-count sampling, horizontal point sampling, variable radius plot sampling, and plotless sampling — was introduced in forestry by the Austrian forester Walter Bitterlich (1948a, 1948b) as a simple counting technique for estimating relative basal area of a tree stand. Nowadays it is also extensively used for sampling trees in forest inventories. The idea in the method is to count, or pick into the sample, those trees in a stand that subtend an angle greater or equal to a fixed angle α when viewed horizontally at breast height from a randomly located point in the stand (Fig. 9). An estimate for the relative basal area at breast height is then obtained by simply multiplying the tally of trees by an α-dependent constant. The procedure may be repeated at several randomly located viewing points in the stand.

There are several alternative ways to formalise Bitterlich sampling. In a model-based approach, the sampling procedure is considered under a model assumed for the structure of the tree stand: in such a model, tree locations are typically regarded as a realisation of a (stationary) spatial point process and tree diameters are treated either as independent random variables with assumed (identical) probability distributions or as the marks of the point process with a mark distribution function involving spatial dependence between the trees; the viewing points need not then be randomly located, but the statistical properties of the estimators of stand totals (typically relative basal area or stem volume) based on angle-count samples are derived given the assumed model (e.g. Holgate 1967, Sukwong et al. 1971, Matérn 1972, Oderwald 1981, and Penttinen 1988). In this study, however, we adopt the usual design-based approach (refer to common textbooks such as e.g. de Vries 1986, Schreuder et al. 1993, and Gregoire and Valentine 2008), where the spatial pattern of tree locations in the region of interest is regarded as fixed and where the randomness in estimators is considered to arise only from the choice of the viewing point locations.

Within this design-based context, there are then at least three different unequal probability sample designs for finite population and one formulation for infinite population to choose among (Eriksson 1995); these differ from each other by the definition of sampling unit and population — whether the sampling unit is considered to be the area from which a tree is seen by relascope, the actual tree, or the viewing point in which relascope is used, and,

![Fig. 9. Selection of trees in region L by Bitterlich sampling with viewing angle α: the trees (white cross-sections) that subtend an angle greater or equal to α, when viewed at breast height from a randomly selected point Q in L, are included in the sample. The inclusion region M_i(α) of the tree i with diameter D_i is outlined with the dashed line.](image-url)
accordingly, whether the population to be sampled then becomes finite or infinite. In this study, we adhere to the classical Grosenbaugh design.

In the Grosenbaugh design, the population is finite and consists of the trees reaching above the breast height in the region of interest. A sample is then composed of the trees selected at one randomly located viewing point, and sampling at k independently randomly located viewing points in the region results in k independent samples. The probability of being included in a sample, termed the inclusion probability — which in this case equals the probability of being selected at each view, referred to as the selection probability — varies according to tree. More precisely — and first realised by the American forester Lewis R. Grosenbaugh (1958), hence the name of the design (Eriksson 1995) — Bitterlich sampling is a method of PPS sampling, that is, sampling with probability proportional to size, meaning that each sampling unit is included in a sample with probability proportional to a covariate, in this case the cross-section area at breast height, which is positively correlated with the variable of interest, in this case typically stem volume.

The estimation of population totals in PPS sampling relies upon the Horvitz-Thompson theorem, which is a beautifully general theory for design-based inference from probability samples (Overton and Stehman 1995). The Horvitz-Thompson estimation is based on the inclusion probabilities of the sampling units, by the inverses of which the observed values of the characteristics of interest are weighted. The theorem is explained in more detail in Appendix B.

Our interest focuses on how non-circularity in the breast height cross-sections of trees affects stand or tree characteristic estimation based on Bitterlich sampling. In order to avoid the complications irrelevant from this point of view, we make the following three assumptions, which are independent of our chosen sampling strategy consisting of the Grosenbaugh design and the Horvitz-Thompson estimator: First, we assume the ground level to be horizontal. Second, we assume the viewer’s eye to be at breast height (1.3 m) from the ground level and the viewer to sight trees horizontally. And third, we assume the region of interest to be such that the edge effects on the inclusion probabilities of trees need not be considered (for example, we assume the borders of the region to be set in the manner that each tree taken to belong to the stand is located farther than its inclusion radius (Eq. 52) from the borders; for some methods of dealing with the edge effect in Bitterlich sampling, see Grosenbaugh 1958, Haga and Maezawa 1959, Barrett 1964, and Schmid 1969).

4.1 Estimation of Stand Totals

We start by explaining how stand totals, that is, sums of tree characteristics within a stand, are estimated from a Bitterlich sample of trees, when the tree cross-sections at breast height are circular. Then we investigate the errors that emerge when the estimators based on the circularity assumption are applied to trees with non-circular of cross-sections. Finally, we examine the estimation of relative basal area in both these aspects.

In addition to the general simplifying presumptions stated above, we assume the characteristics of the trees in the sample to be measured or estimated without error. For the notation, let $L$ denote the region of interest in the ground plane, $|L|$ the area of this region, and $I$ the set of the trees growing in $L$ and reaching above breast height. Further, let $Q$ denote a uniformly randomly located viewing point in $L$ and $\alpha \in (0, \pi)$ the viewing angle used in a relascope.
4.1.1 Principle under Assumption of Circular Cross-Sections

If the cross-sections of the trees are circular at breast height, the maximum distance determined from the tree pith, at which the tree \( i \in I \) with breast height diameter \( D_i \) is taken into the Bitterlich sample becomes

\[
\rho_i(\alpha) = \frac{D_i}{2} \cdot \frac{1}{\sin(\alpha/2)}
\]

by a straightforward trigonometric consideration (Fig. 9). In other words, the inclusion region \( M_i(\alpha) \) within which the tree \( i \in I \) is observed in an angle greater or equal to \( \alpha \) is a circle with radius \( r_i(\alpha) \) and area

\[
|M_i(\alpha)| = \pi r_i(\alpha)^2 = \frac{\pi}{4} D_i^2 \cdot \frac{1}{\sin^2(\alpha/2)}
\]

(Fig. 9). The ratio \( \kappa_i(\alpha) \) of the cross-section area to the inclusion area for tree \( i \in I \) is now seen to depend only on the viewing angle \( \alpha \) and thus have the same value for all the trees in \( I \):

\[
\kappa_i(\alpha) = \frac{\pi}{4} D_i^2 \cdot \frac{1}{|M_i(\alpha)|} = \frac{\pi}{4} D_i^2 \cdot \frac{1}{\sin^2(\alpha/2)}
\]

the ratio is referred to as the basal area factor. (Note that the cross-section area and inclusion area are here assumed to be expressed in the same units, so that the units cancel out in \( \sin^2(\alpha/2) \); to express the basal area factor in \( \text{m}^2/\text{ha} \) or \( \text{ft}^2/\text{ac} \), \( \sin^2(\alpha/2) \) has to be multiplied with 10 000 or 43 560, respectively.)

The inclusion probability \( \pi_i(\alpha) \) of a tree \( i \in I \) is the probability that the viewing point \( Q \) is located in the inclusion region \( M_i(\alpha) \); on account of the uniform distribution of \( Q \) in \( L \), the probability is given by the simple area ratio

\[
\pi_i(\alpha) = \frac{|M_i(\alpha)|}{|L|} = \frac{\pi}{4} D_i^2 \cdot \frac{1}{\sin^2(\alpha/2)|L|}
\]

(Note that the inclusion region is here defined to contain the ground-level cross-section of the stem, although the viewing point \( Q \) cannot be located within the cross-section.) Accordingly, the probability of a tree to be included in a sample is proportional to its cross-section area (squared diameter) at breast height and inversely proportional to the area of the region of interest; further, the smaller the viewing angle (and, consequently, the smaller the basal area factor) the larger the probability, as \( \sin^2(\alpha/2) \) is monotonously increasing for \( \alpha \in (0, \pi) \).

Let now \( s_Q(\alpha) \subseteq I \) denote the sample of trees obtained with the relascope angle \( \alpha \) at the viewing point \( Q \in L \). From each tree \( i \) in \( s_Q(\alpha) \), we measure the breast height diameter \( D_i \) and the value \( Y_i \) of the characteristic of interest. As is apparent from Eq. 55, we need the diameter measurements for the inclusion probabilities of the sample trees, which then enable us to apply the Horvitz-Thompson theorem for an unbiased estimation of the total amount of the characteristic of interest in the region \( L \). The sum \( Y = \Sigma_{i \in I} Y_i \) is estimated by
\[
Y_{HT} = \sum_{i \in s_Q(\alpha)} \frac{Y_i}{\pi_i(\alpha)} = |L| \sin^2(\alpha/2) \sum_{i \in s_Q(\alpha)} \frac{Y_i}{4D_i^2},
\]

(56)

that is, as a weighted sum of the sample tree measurements with the inverses of the cross-section areas (squared diameters) as weights and with a quantity depending on the viewing angle (or the basal area factor) and the area of the region of interest as the scaling coefficient.

The randomness in \(Y_{HT}\) stems from the placement of the viewing point \(Q\) in the region \(L\), as this determines the sample \(s_Q(\alpha)\). By the Horvitz-Thompson theorem (Appendix B), \(Y_{HT}\) is unbiased with respect to the sampling design, that is, \(E(\hat{Y}_{HT}) = Y\), where the expectation is taken over the probability distribution of all possible samples \(s_Q(\alpha)\) resulting from all possible locations of \(Q\).

The variance of \(Y_{HT}\) depends on the spatial pattern of the trees in the stand, as it involves the joint inclusion probabilities for all the pairs of the trees in the population (Appendix B). The joint inclusion probability \(\pi_{ij}(\alpha)\) for trees \(i, j \in I\) is the probability that both the trees are included in a sample, and in Bitterlich sampling this probability is proportional to the overlapping area of the tree inclusion regions. For the unbiased variance estimator given in Appendix B (Eq. B9) and based on a sample of trees taken at one viewing point, these probabilities need not only be known but also positive for all the pairs of trees in the sample. In a Bitterlich sample, the positivity requirement is invariably fulfilled, as the inclusion regions of the trees in the same sample necessarily overlap. The requirement of knowing the spatial pattern of the trees in the sample can be circumvented by repeating Bitterlich sampling at \(k\) randomly and independently selected viewing points \(Q_j \in L, j = 1, ..., k\): The estimators \(\hat{Y}_{HT}[s_{Q_j}(\alpha)]\) obtained from the \(k\) samples are independent and identically distributed, and hence have the same variance \(\text{Var}(\hat{Y}_{HT}[s_{Q_j}(\alpha)]) = \text{Var}(\bar{Y}_{HT})\) for all \(Q_j\). An unbiased estimator for this variance is given by the sample variance of the \(k\) estimators:

\[
\text{Var}(\hat{Y}_{HT}) = \frac{1}{k} \sum_{j=1}^{k} \left( \hat{Y}_{HT}[s_{Q_j}(\alpha)] - \bar{Y}_{HT} \right)^2,
\]

(57)

where \(\bar{Y}_{HT}\) is the mean of the \(k\) independent estimators \(\hat{Y}_{HT}[s_{Q_j}(\alpha)]\). The population total \(Y\) is estimated with \(\bar{Y}_{HT}\), the variance of which is \(\text{Var}(\bar{Y}_{HT}) = \text{Var}(\hat{Y}_{HT})/k\). An unbiased estimator of the variance of \(\bar{Y}_{HT}\) then becomes

\[
\text{Var}(\bar{Y}_{HT}) = \frac{1}{k} \text{Var}(\hat{Y}_{HT}) = \frac{1}{k(k-1)} \sum_{j=1}^{k} \left( \hat{Y}_{HT}[s_{Q_j}(\alpha)] - \bar{Y}_{HT} \right)^2
\]

(58)


A model-based approach provides an alternative way of estimating the variance of a stand total estimator — in fact, the model-based studies referred to in the beginning of this chapter expressly address this problem. Under a certain class of models for the random mechanism generating forest stands, analytical expressions for the variance can be found; under other models, the variance can be estimated via simulations.
4.1.2 Effect of Non-Circularity of Cross-Sections

Since Bitterlich sampling operates with the tangents of the convex closures of breast height cross-sections, non-convexity in the cross-sections is unobservable and thus does not influence the selection of trees. Therefore, when examining how non-circularity in the breast height cross-sections of trees affects stand total estimation by Bitterlich sampling, we confine ourselves to considering the convex closures of the cross-sections of arbitrary non-circular shape.

If the convex closure of the breast height cross-section of a tree \( i \in I \) is not a circle, the maximum distance at which the tree is counted with a certain viewing angle \( \alpha \) varies by the direction in which the tree is being viewed. Ergo, the inclusion region becomes non-circular (Fig. 10; cf. Grosenbaugh 1958, Bitterlich 1984). The size and shape of the inclusion region depend not only on the size and shape of the breast height cross-section but also on the viewing angle \( \alpha \) used (Fig. 11). Using the support function \( p_i(\cdot) \) of the breast height cross-section in tree \( i \), now defined in a co-ordinate system with the origin set in the centre of gravity of the convex closure of the cross-section (or in the pith of the cross-section), we attain the following parametric representation for the length of the radius of the inclusion region:

\[
\begin{align*}
    r_i(\theta; \alpha) &= \frac{1}{\sin \alpha} \sqrt{p_i(\theta)^2 + p_i(\theta + \pi - \alpha)^2 + 2p_i(\theta)p_i(\theta + \pi - \alpha) \cos \alpha}, \\
    \text{where } \alpha &\in (0, \pi) \text{ and } \theta \in [0, 2\pi) (\text{see Appendix C for the derivation}).
\end{align*}
\]

(Note that in this parametrisation, \( \theta \) does not indicate the direction of the radius, but the direction is a function of \( \theta \) and \( \alpha \) (Appendix C); consequently, the area of the inclusion region is not obtained straightforwardly by integrating \( r_i(\theta; \alpha)^2 \) from 0 to \( 2\pi \) (cf. Eq. 12 in Chapter 2). We see that as \( \alpha \) tends to zero, the numerator of \( r_i(\theta; \alpha) \) tends to \( p_i(\theta) + p_i(\theta + \pi) = D_i(\theta) \) (and the denominator tends to 0); accordingly, for the trees with an orbiform breast height cross-section (constant diameter function \( D(\cdot) \) at breast height), the limiting shape of the inclusion region is a circle (e.g. Fig. 11 B). As \( \alpha \) tends to \( \pi \), the viewer draws closer and closer to the stem, and the radii \( r_i(\theta; \alpha) \) are finally found to tend to the radii of the convex closure of the breast height cross-section; in other words, the inclusion region tends to the convex closure of the breast height cross-section (cf. Matérn 1956, p. 24).
The area of the inclusion region of the tree \( i \in I \) can be expressed as

\[
|M_i(\alpha)| = \frac{1}{\sin^2 \alpha} \left[ \int_0^{2\pi} \left( p_i(\theta)^2 + p_i(\theta)p_i(\theta+\pi-\alpha) \cos \alpha + p_i(\theta)p_i'(\theta+\pi-\alpha) \sin \alpha \right) d\theta \right]^{\pi \theta} \]

(cf. Matérn 1956; see Appendix C for the derivation). If \( p_i(\cdot) \) is multiplied with a constant, so is \( p_i'(\cdot) \), and \( |M_i(\alpha)| \) thus becomes multiplied with the square of the constant. A similar reasoning pertains to the convex area of the breast height cross-section

\[
A_{C_i} = \frac{1}{2} \int_0^{2\pi} \left( p_i(\theta)^2 - p_i'(\theta)^2 \right) d\theta
\]

(Eq. 2 in Chapter 2). In the basal area factor \( \kappa_i(\alpha) \) for tree \( i \), then, defined as

\[
\kappa_i(\alpha) = \frac{A_{C_i}}{|M_i(\alpha)|},
\]

the multiplying constant cancels out. Now if \( p_i(\cdot) \) is defined with reference to the centre of gravity of the convex closure of the breast height cross-section, the size of the cross-section influences \( p_i(\cdot) \) only by a multiplying constant. Consequently, the basal area factor of a tree is a function of the viewing angle \( \alpha \) that depends on the shape but not on the size of the breast height cross-section.

For example, if the breast height cross-section of the tree \( i \) is a circle with radius \( R_i \), for which \( p_i(\theta)=R_i \) for all \( \theta \), the convex area of the cross-section is naturally \( A_{C_i} = \pi R_i^2 \), and the area of the inclusion region is given by
\[ |M_1(\alpha)| = \frac{1}{\sin^2(\alpha)} \int_0^{2\pi} (R_i^2 + R_i^2 \cos \alpha) \sin \alpha \mid d\theta = \frac{2\pi R_i^2 (1 + \cos \alpha)}{\sin^2 \alpha}, \]  

whereby the basal area factor becomes

\[ \kappa_i(\alpha) = \frac{A_{Ci}}{|M_1(\alpha)|} = \frac{\sin^2 \alpha}{2(1 + \cos \alpha)} = \sin^2(\alpha/2), \]  

– a result already established, in a different way, in the previous section (cf. Eq. 54; \( R_i = D_i/2, \ \cos \alpha = 2 \cos^2(\alpha/2) - 1, \ \sin \alpha = 2 \sin(\alpha/2) \cos(\alpha/2) \)). Patently, the ratio does not depend on the radius \( R_i \) of the cross-section but is \( \sin^2(\alpha/2) \) for every tree with circular cross-section at breast height. For any other convex shape irrespective of size, the basal area factor is then another function of \( \alpha \), although not necessarily very much different from \( \sin^2(\alpha/2) \) (cf. Fig. 12).

As in the circular case (cf. Eq. 55), the inclusion probability of the non-circular tree \( i \in I \) is proportional to the convex area of the cross-section:

\[ \pi_1(\alpha) = \frac{|M_1(\alpha)|}{|L|} = \frac{A_{Ci}}{\kappa_i(\alpha)|L|}, \]  

where the proportionality coefficient \( 1/\kappa_i(\alpha) \) just varies according to the shape of the cross-section. The Horvitz-Thompson estimator of the sum \( Y=\Sigma_{i \in I} Y_i \) then becomes

\[ \hat{Y}_{HT} = \sum_{i \in \omega_j(\alpha)} \frac{Y_i}{\pi_1(\alpha)} = |L| \sum_{i \in \omega_j(\alpha)} \kappa_i(\alpha) \frac{Y_i}{A_{Ci}}, \]  

that is, a weighted sum of the sample tree measurements with \( \kappa_i(\alpha)/A_{Ci} \) as weights and with the area of the region of interest as the scaling coefficient (cf. Eq. 56).

---

**Fig. 12.** Basal area factor \( \kappa(\alpha) = A_{Ci}/|M(\alpha)| \), where \( A_{Ci} \) is the convex area of the breast height cross-section and \( |M(\alpha)| \) is the inclusion area of the tree, as a function of viewing angle \( \alpha \) for the six example shapes in Fig. 10, with reference to the basal area factor \( \sin^2(\alpha/2) \) of a circle.
Yet in practice, both the shapes and the convex areas of the cross-sections usually remain unknown, and the inclusion probabilities are estimated by applying the formula of the circular case (Eq. 55) also to non-circular cross-sections. The resulting estimator (cf. Eq. 56)

\[
\hat{Y} = \sum_{i \in s_Q(\alpha)} \frac{Y_i}{\hat{\pi}_i(\alpha)}
\]

\[
= \sum_{i \in s_Q(\alpha)} \frac{\pi_i}{4D_i^2} \cdot \frac{Y_i}{\sin^2(\alpha/2)|L|}
\]

\[
= |L| \sin^2(\alpha/2) \sum_{i \in s_Q(\alpha)} \frac{Y_i}{\pi_i(\alpha) / 4D_i^2}
\]

where \(D_i\) now denotes the diameter selected in some fashion at breast height in a non-circular tree \(i\), is not necessarily unbiased with respect to the sampling design.

To unveil the possible bias caused by the estimated inclusion probabilities, we bring in the random variable

\[
\delta[s_Q(\alpha)] = \begin{cases} 
1, & i \in s_Q(\alpha) \\
0, & i \notin s_Q(\alpha)
\end{cases}
\]

indicating whether a tree \(i \in I\) is included in the sample \(s_Q(\alpha)\); obviously, \(E\{\delta[s_Q(\alpha)]\} = \pi_i(\alpha)\), the expectation being taken over the probability distribution of all possible samples (see Appendix B). Now

\[
E(\hat{Y}) = E \left[ \sum_{i \in s_Q(\alpha)} \frac{Y_i}{\hat{\pi}_i(\alpha)} \right]
\]

\[
= E \left[ \sum_{i \in I} \frac{\delta[s_Q(\alpha)] Y_i}{\hat{\pi}_i(\alpha)} \right]
\]

\[
= \sum_{i \in I} E \left[ \delta[s_Q(\alpha)] \right] \frac{Y_i}{\hat{\pi}_i(\alpha)}
\]

\[
= \sum_{i \in I} Y_i \frac{\pi_i(\alpha)}{\hat{\pi}_i(\alpha)}
\]

\[
= \sum_{i \in I} Y_i \frac{A_{ci} \cdot 1}{\pi_i(\alpha)|L|} \frac{1}{\sin^2(\alpha/2)|L|}
\]

\[
= \sum_{i \in I} Y_i \frac{A_{ci}}{\pi_i(\alpha)} \frac{1}{D_i^2} \frac{\sin^2(\alpha/2)}{\kappa_i(\alpha)}
\]

Here we see that the possible bias in the estimator ensues from two tree-specific flaws in the estimated inclusion probabilities: First, the diameter \(D_i\) selected to be measured in the
non-circular breast height cross-section of tree $i$ may not yield the true convex area $A_{Ci}$ when substituted in the circle area formula. Second, the deviation of the true basal area factor $\kappa_i(\alpha) = A_{Ci}/M_i(\alpha)$ from that of the circle, that is, from $\sin^2(\alpha/2)$, produces an error the magnitude of which depends on the relascope angle $\alpha$ and the shape of the breast height cross-section. (Note that if the selection of diameters $D_i$ involves randomness, $E(\hat{Y})$ becomes a random variable the expectation of which over the (multidimensional) diameter direction distribution should be used as the measure for the goodness of the estimator.)

The two-fold errors resulting from the estimated inclusion probabilities tend to be systematic and of opposite signs. With the diameter derived from the girth, the circle area formula was previously (Section 3.3.1) found to overestimate the convex area of a cross-section of any shape; this is the case also when the maximum diameter is used. This overestimation in all the trees then causes $\hat{Y}$ to underestimate $Y$ on average. Also with random diameters, the circle area formula was found to overestimate the convex area on average, that is, in terms of the expectation over the uniform direction distribution within cross-section (Section 3.3.2). However, $A_{Ci}/E_0[\pi D_i(\theta)^2/4] \leq 1$ does not imply that $E_0\{A_{Ci}/[\pi D_i(\theta)^2/4]\} \leq 1$. Yet for the example shapes in Fig. 10, also the latter inequality holds with the diameter selection methods 1–5 (Table 5), but the order of $E_0\{A_{Ci}/[\pi D_i(\theta)^2/4]\}$ does not follow the order of $A_{Ci}/E_0[\pi D_i(\theta)^2/4]$.

The basal area factor $\kappa_i(\alpha)$, in turn, can be shown to be not larger than $\sin^2(\alpha/2)$ irrespective of the shape of the breast height cross-section of tree $i$: Matérn (1956) proved that for any (convex) shape of breast height cross-section

$$\left|\sin^2(\alpha / 2) \left|M_i(\alpha) - \hat{A}_{0i}\right|\right| \leq \hat{A}_{0i} - A_{Ci},$$

(70)

where $\hat{A}_{0i}$ is the area estimator based on the convex perimeter at breast height in the tree $i$; we can rephrase this (setting $|M_i(\alpha)|=A_{Ci}/\kappa_i(\alpha)$ according to Eq. 62) to get

$$1 \leq \frac{\sin^2(\alpha / 2)}{\kappa_i(\alpha)} \leq 1 + \frac{2(\hat{A}_{0i} - A_{Ci})}{A_{Ci}},$$

(71)

The basal area factor being smaller than $\sin^2(\alpha/2)$ in all the trees then causes $\hat{Y}$ to overestimate $Y$ on average.

Table 5. Within-cross-section expectation of $A_{Ci}/[\pi D_i(\theta)^2/4]$ for the example shapes in Fig. 10, when the diameter $D(\theta)$ is measured with methods 1–5 (used in area estimators $\hat{A}_1$–$\hat{A}_5$, see Section 3.3.2). For a reference, the ratio obtained with the girth diameter ($A_{Ci}/\hat{A}_0$, diameter selection method 0) is shown. The values are given as 100-fold.

<table>
<thead>
<tr>
<th>Diameter selection method</th>
<th>Shape</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>98.16</td>
<td>96.00</td>
<td>97.85</td>
<td>98.29</td>
<td>97.06</td>
<td>97.07</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>100.00</td>
<td>96.00</td>
<td>98.43</td>
<td>98.98</td>
<td>98.02</td>
<td>97.64</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>98.16</td>
<td>96.00</td>
<td>97.85</td>
<td>98.43</td>
<td>97.54</td>
<td>97.64</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>98.77</td>
<td>96.00</td>
<td>98.04</td>
<td>98.62</td>
<td>97.70</td>
<td>97.64</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>99.07</td>
<td>96.00</td>
<td>98.14</td>
<td>98.64</td>
<td>97.54</td>
<td>97.36</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>98.77</td>
<td>96.00</td>
<td>98.04</td>
<td>98.62</td>
<td>97.70</td>
<td>97.64</td>
<td></td>
</tr>
</tbody>
</table>
In Fig. 13, the ratio $\sin^2(\alpha/2)/\kappa(\alpha)$ is drawn as a function of the viewing angle $\alpha$ for the six example shapes of Fig. 10 (cf. Matérn 1956, p. 25).

In Fig. 13, the ratio $\sin^2(\alpha/2)/\kappa(\alpha)$ is drawn as a function of $\alpha$ for the six example shapes of Fig. 10. With the viewing angles used in practice ($\alpha$ around 1–2°, not more than 5°), each tree with such a non-circular breast height cross-section and with the area $A_{Ci}$ correctly estimated with $\pi D_i^2/4$ appears to contribute to the stand total estimate with an amount that is 2.5–4 % too large. For ellipses, the overestimation increases substantially along with eccentricity (Fig. 14; cf. the results by Grosenbaugh (1958) showing that for an ellipse with axis ratio 0.5 the overestimation amounts to over 25 % with $\alpha = 1.74^\circ$, 3.81°, 45° or 90°).

As for more exotic cross-section shapes, we can deduce from the results given by Bitterlich (1984) that for a half circle the overestimation becomes 38 % with $\alpha = 2.29^\circ$.

Grosenbaugh (1958) examined the combined error $\sin^2(\alpha/2)/\kappa(\alpha) \cdot A_{Ci}/(\pi D_i^2/4)$ in the ellipses with axis ratio 0.9 or 0.5 by employing the girth diameter or the quadratic, arithmetic or geometric mean of $D_{\text{min}}$ and $D_{\text{max}}$ as $D_i$ and by applying the values 90°, 45°, 3.81° and 1.74° for $\alpha$. He found that in all those cases, the combined error was larger or equal to 1,
implying an overestimating bias in \( \hat{Y} \). Not unexpectedly, the quadratic mean yielded the smallest combined error (in fact, an error of negligible magnitude when the two smallest values of the viewing angle were used): the quadratic mean in \( \pi D_i^2/4 \) results in the largest overestimation of \( A_{\text{Ci}} \), which then most effectively counterbalances the overestimation of \( \kappa_i(\alpha) \) by \( \sin^2(\alpha/2) \) (which, in turn, does not depend on diameter selection). In Table 6, the combined errors produced by the diameter selection methods 0–11 and viewing angles 1° and 30° are given for the example shapes of Fig. 10. Manifestly, Grosenbaugh’s findings on elliptical cross-section cannot be generalised, but with different shapes and diameter selection methods the combined error may very well result in underestimation. In a majority of the shapes, however, the most common diameter selection methods 0–5 result in overestimation smaller than 1%. Within the range applied, the viewing angle appears to influence the error only negligibly.

Table 6. Contribution of a sample tree to bias in stand total estimation from Bitterlich sample \( (\sin^2(\alpha/2) / \kappa(\alpha) \cdot A_{\text{Ci}} / (\pi D_i^2/4), \text{see Eq. 69}) \), if the breast height cross-section of the tree is one of the shapes in Fig. 10 and if its diameter is measured with methods 0–11 used in area estimators \( \hat{A}_0 - \hat{A}_{11} \) (see Section 3.3). Results with viewing angles \( \alpha=1^\circ \) and \( \alpha=30^\circ \) (in parentheses) are given; with \( \alpha=2^\circ \) and \( \alpha=5^\circ \), the results were similar to those obtained with \( \alpha=1^\circ \). For the diameter selection methods 1–5, the expectation of \( A_{\text{Ci}} / (\pi D_i^2/4) \) over the uniform direction distribution (Table 5) was used in the computation. The values are given as 100-fold.

<table>
<thead>
<tr>
<th>Diameter selection method</th>
<th>Shape</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.61</td>
<td>100.00</td>
<td>100.20</td>
<td>100.24</td>
<td>100.33</td>
<td>100.20</td>
</tr>
<tr>
<td></td>
<td>(100.75)</td>
<td>(99.87)</td>
<td>(100.19)</td>
<td>(100.29)</td>
<td>(100.40)</td>
<td>(100.24)</td>
</tr>
<tr>
<td>1</td>
<td>102.50</td>
<td>100.00</td>
<td>100.79</td>
<td>100.95</td>
<td>101.32</td>
<td>100.79</td>
</tr>
<tr>
<td></td>
<td>(102.64)</td>
<td>(99.87)</td>
<td>(100.79)</td>
<td>(100.99)</td>
<td>(101.40)</td>
<td>(100.83)</td>
</tr>
<tr>
<td>2</td>
<td>100.61</td>
<td>100.00</td>
<td>100.20</td>
<td>100.39</td>
<td>100.82</td>
<td>100.79</td>
</tr>
<tr>
<td></td>
<td>(100.76)</td>
<td>(99.87)</td>
<td>(100.19)</td>
<td>(100.43)</td>
<td>(100.90)</td>
<td>(100.83)</td>
</tr>
<tr>
<td>3</td>
<td>101.24</td>
<td>100.00</td>
<td>100.39</td>
<td>100.58</td>
<td>100.99</td>
<td>100.79</td>
</tr>
<tr>
<td></td>
<td>(101.38)</td>
<td>(99.87)</td>
<td>(100.38)</td>
<td>(100.62)</td>
<td>(101.06)</td>
<td>(100.83)</td>
</tr>
<tr>
<td>4</td>
<td>101.55</td>
<td>100.00</td>
<td>100.49</td>
<td>100.60</td>
<td>100.82</td>
<td>100.49</td>
</tr>
<tr>
<td></td>
<td>(101.69)</td>
<td>(99.87)</td>
<td>(100.48)</td>
<td>(100.65)</td>
<td>(100.90)</td>
<td>(100.54)</td>
</tr>
<tr>
<td>5</td>
<td>101.24</td>
<td>100.00</td>
<td>100.39</td>
<td>100.58</td>
<td>100.99</td>
<td>100.79</td>
</tr>
<tr>
<td></td>
<td>(101.38)</td>
<td>(99.87)</td>
<td>(100.38)</td>
<td>(100.62)</td>
<td>(101.06)</td>
<td>(100.83)</td>
</tr>
<tr>
<td>6</td>
<td>101.23</td>
<td>100.00</td>
<td>100.20</td>
<td>95.71</td>
<td>100.33</td>
<td>100.20</td>
</tr>
<tr>
<td></td>
<td>(101.38)</td>
<td>(99.87)</td>
<td>(100.19)</td>
<td>(95.75)</td>
<td>(100.40)</td>
<td>(100.24)</td>
</tr>
<tr>
<td>7</td>
<td>102.50</td>
<td>100.00</td>
<td>100.59</td>
<td>96.16</td>
<td>101.35</td>
<td>100.20</td>
</tr>
<tr>
<td></td>
<td>(102.64)</td>
<td>(99.87)</td>
<td>(100.58)</td>
<td>(96.20)</td>
<td>(101.43)</td>
<td>(100.63)</td>
</tr>
<tr>
<td>8</td>
<td>101.23</td>
<td>100.00</td>
<td>100.20</td>
<td>103.45</td>
<td>112.19</td>
<td>114.00</td>
</tr>
<tr>
<td></td>
<td>(101.38)</td>
<td>(99.87)</td>
<td>(100.19)</td>
<td>(103.50)</td>
<td>(112.27)</td>
<td>(114.05)</td>
</tr>
<tr>
<td>9</td>
<td>102.50</td>
<td>100.00</td>
<td>100.59</td>
<td>103.56</td>
<td>112.46</td>
<td>114.00</td>
</tr>
<tr>
<td></td>
<td>(102.64)</td>
<td>(99.87)</td>
<td>(100.58)</td>
<td>(103.60)</td>
<td>(112.54)</td>
<td>(114.05)</td>
</tr>
<tr>
<td>10</td>
<td>101.23</td>
<td>100.00</td>
<td>100.20</td>
<td>94.26</td>
<td>90.25</td>
<td>88.75</td>
</tr>
<tr>
<td></td>
<td>(101.38)</td>
<td>(99.87)</td>
<td>(100.19)</td>
<td>(94.30)</td>
<td>(90.32)</td>
<td>(88.79)</td>
</tr>
<tr>
<td>11</td>
<td>102.50</td>
<td>100.00</td>
<td>100.59</td>
<td>94.61</td>
<td>90.42</td>
<td>88.75</td>
</tr>
<tr>
<td></td>
<td>(102.64)</td>
<td>(99.87)</td>
<td>(100.58)</td>
<td>(94.65)</td>
<td>(90.49)</td>
<td>(88.79)</td>
</tr>
</tbody>
</table>
If we choose to measure the girth diameter at breast height from each tree included in Bitterlich sample, we obtain the following bounds for the expectation of $\hat{Y}$ (the stand total estimator involving inclusion probabilities that were estimated as if the cross-sections were circular, Eq. 67):

$$\sum_{i \in I} Y_i \frac{A_{C_i}}{A_{\hat{0}_i}} \leq E(\hat{Y}) \leq \sum_{i \in I} Y_i \left(2 - \frac{A_{C_i}}{A_{\hat{0}_i}}\right)$$  \hspace{1cm} (72)

(combine Eqs. 69 and 71 and set $\pi D_i^2/4 = \pi \mu D_i^2/4 = \pi (C_i/\pi)^2/4 = \hat{A}_{0i}$). The bounds of the bias are now ascribed to the deviations of the cross-section area estimates $\hat{A}_{0i}$ from the convex areas $A_{C_i}$ at breast height in the stems of the stand. The lower bound evidently falls below the true value $Y = \sum_{i \in I} Y_i$, and hence $\hat{Y}$ with girth diameters may even systematically underestimate $Y$.

If, in turn, we choose to measure from each tree in the sample one random diameter with uniformly distributed direction, we attain a point approximation for the expectation of $\hat{Y}$: Matérn (1956) proved that for any (convex) shape of breast height cross-section

$$\left| \sin^2(\alpha/2) |M_i(\alpha)| - E_\theta(\hat{A}_{1i}) \right| \leq \left[ E_\theta(\hat{A}_{1i}) - A_{C_i} \right] \tan^2(\alpha/2) \cdot \frac{1}{\kappa_i(\alpha)} \leq \left[ E_\theta(\hat{A}_{1i}) - A_{C_i} \right] \tan^2(\alpha/2)$$  \hspace{1cm} (73)

where $E_\theta(\hat{A}_{1i})$ is the expectation, over the uniform direction distribution, of the area estimator $\hat{A}_1$ involving one random diameter at breast height in tree $i$; by rephrasing this, the basal area factor $\kappa_i(\alpha)$ is found to satisfy

$$\frac{E_\theta(\hat{A}_{1i})}{A_{C_i}} - \tan^2(\alpha/2) \left[ \frac{E_\theta(\hat{A}_{1i})}{A_{C_i}} - 1 \right] \leq \sin^2(\alpha/2) \frac{1}{\kappa_i(\alpha)} \leq \frac{E_\theta(\hat{A}_{1i})}{A_{C_i}} + \tan^2(\alpha/2) \left[ \frac{E_\theta(\hat{A}_{1i})}{A_{C_i}} - 1 \right].$$  \hspace{1cm} (74)

With the usually employed relascope angles $\alpha$ of the magnitude $1–2^\circ$, $\tan^2(\alpha/2)$ assumes the values $0.000076–0.00030$, and hence $\sin^2(\alpha/2)/\kappa_i(\alpha)$ becomes very close to $E_\theta(\hat{A}_{1i})/A_{C_i}$. Consequently,

$$E(\hat{Y}) = \sum_{i \in I} Y_i \frac{\sin^2(\alpha/2)}{\kappa_i(\alpha)} \frac{A_{C_i}}{A_{\hat{1}_i}} = \sum_{i \in I} Y_i \frac{E_\theta(\hat{A}_{1i})}{A_{C_i}} \frac{A_{C_i}}{A_{\hat{1}_i}} = \sum_{i \in I} Y_i \frac{E_\theta(\hat{A}_{1i})}{A_{\hat{1}_i}}$$  \hspace{1cm} (75)

that is, the bias in the stand total estimator $\hat{Y}$ is attributable to the deviation of the area estimates based on one random diameter from their within-cross-section expectations in the stems of the stand.
When the viewing point is located close to the boundary of the inclusion region of a tree with non-circular (breast height) cross-section, it is not a straightforward task to check whether the tree should be included in the Bitterlich sample, as the critical distance varies according to the viewing direction (cf. Eq. 59). Contrary to what Grosenbaugh (1958) reasoned, the convention of measuring diameter $D$ perpendicular to the viewing direction and computing the critical distance with the circle formula (Eq. 52) as

$$r(\alpha) = \frac{D}{2} \cdot \frac{1}{\sin(\alpha / 2)}$$

(76)

appears to give a good approximation with the viewing angles used in practice ($\alpha$ around 1–2°, not more than 5°) (Fig. 15). For ellipses, Grosenbaugh (1958) recommended computing the critical distance as the radius of a circle with the same area as the inclusion area (for ellipses this can be easily determined as a function of $D_{\min}$, $D_{\max}$ and $\alpha$); however, if applied only when the viewing point appears to be near the inclusion region boundary (and not routinely to all sample trees), this practice is likely to result in an inclusion region that is the union of the true inclusion region and the approximating circle (Fig. 16) and is thus not something to advocate.

4.1.3 Special Case: Relative Basal Area

We complete our examination of stand total estimation by Bitterlich sampling by applying the theory presented above to the estimation of relative basal area. As non-convexity of cross-sections cannot be observed with relascope, we define the relative basal area of a stand
If the cross-sections of the trees are circular at breast height — let $D_i$ again stand for the breast height diameter of a tree $i \in I$ — the relative basal area can be expressed as

$$G = \frac{1}{|L|} \sum_{i \in I} \frac{\pi}{4} D_i^2,$$

and its Horvitz-Thompson estimator becomes (cf. Eq. 56)

$$\hat{G}_{HT} = \frac{1}{|L|} \sum_{i \in \alpha(\omega)} \frac{\pi}{4} D_i^2 \frac{\pi}{\pi_1(\alpha)}$$

$$= \frac{1}{|L|} |L| \sin^2(\alpha/2) \sum_{i \in \alpha(\omega)} \frac{\pi}{4} D_i^2$$

$$= n \sin^2(\alpha/2),$$

(79)
where \( n \) is the number of trees in the sample \( s_Q(\alpha) \) obtained at point \( Q \in L \) with viewing angle \( \alpha \). Thus, an unbiased (with respect to the sampling design) estimate for the relative basal area is indeed attained by just counting the number of trees subtending the relascope angle — the fact perceived by Walter Bitterlich, at first only empirically (Bitterlich 1984)!

The result follows from the fact that the inclusion area, and thereby also the inclusion probability, of each sampled tree \( i \) is proportional to \( \pi D_i^2/4 \) with \( 1/\sin^2(\alpha/2) \) as the proportionality coefficient. (Note that \( G_{HT} \) here is unitless, i.e., the cross-section areas and the area of the region \( L \) are assumed to be expressed in the same units; as mentioned before, the basal area factor \( \sin^2(\alpha/2) \) is usually multiplied by 10000 or 43 560 so as to express \( G_{HT} \) in m\(^2\)/ha or ft\(^2\)/ac, respectively.)

If counting of trees is repeated at \( k \) randomly and independently selected viewing points \( Q_j \in L, j=1, \ldots, k \), \( G \) is unbiasedly estimated with the mean \( \bar{G}_{HT} \) of the \( k \) independent estimators \( \hat{G}_{HTj} \); an unbiased estimator for the variance of \( \bar{G}_{HT} \) is given by

\[
\text{Vår}(\bar{G}_{HT}) = \frac{1}{k(k-1)} \sum_{j=1}^{k} (\hat{G}_{HTj} - \bar{G}_{HT})^2
\]

\[
= \frac{\sin^2(\alpha/2)}{k(k-1)} \sum_{j=1}^{k} (n_j - \bar{n})^2,
\]

where \( \bar{n} \) is the arithmetic mean of the numbers of trees \( n_j \) counted at the \( k \) separate viewing points.

If the convex closures of the breast height cross-sections of the trees are non-circular, the basal area estimator \( \hat{G} = n \sin^2(\alpha/2) \) is biased (when considered over the probability distribution of all possible tree samples):

\[
\text{E}(\hat{G}) = \sin^2(\alpha/2) \text{E}(n)
\]

\[
= \sin^2(\alpha/2) \text{E} \left\{ \sum_{i=1}^{k} \delta_i [s_Q(\alpha)] \right\}
\]

\[
= \sin^2(\alpha/2) \sum_{i=1}^{k} \pi_i(\alpha)
\]

\[
= \sin^2(\alpha/2) \sum_{i=1}^{k} \frac{|M_i(\alpha)|}{|L|}
\]

\[
= \frac{1}{|L|} \sum_{i=1}^{k} \sin^2(\alpha/2) |M_i(\alpha)|
\]

\[
= \frac{1}{|L|} \sum_{i=1}^{k} \sin^2(\alpha/2) \kappa_i(\alpha)
\]

\[
= \frac{1}{|L|} \sum_{i=1}^{k} \hat{A}_i(\alpha).
\]

Clearly, the bias of \( \hat{G} \) owing to the non-circular shapes of tree cross-sections manifests itself in the deviation of \( \hat{A}_i(\alpha) \) from \( A_{Ci} \) (Matérn 1956), that is, in the deviation of \( \sin^2(\alpha/2)/\kappa_i(\alpha) \) from 1. As \( \sin^2(\alpha/2)/\kappa_i(\alpha) \) was found to be larger or equal to 1 (Eq. 71) for all convex shapes, \( \hat{G} \) tends to overestimate \( G \) systematically.

The inequalities proved by Matérn (1956) and introduced in the previous section (Eqs. 70 and 73) give bounds for the overestimating bias in \( \hat{G} \) (originally, the inequalities were derived...
expressly for the case of estimating relative basal area by relascope). The first inequality
\[ |\tilde{A}(\alpha) - \tilde{A}_0| \leq \tilde{A}_0 - A_{ci} \Leftrightarrow A_{ci} \leq \tilde{A}(\alpha) \leq A_{ci} + 2(\tilde{A}_0 - A_{ci}) \] (82)
implies that, irrespective of the shape of a cross-section, \( \tilde{A}(\alpha) \) be always greater or equal to the convex area of the cross-section, and that the error of \( \tilde{A}(\alpha) \) never exceed double the isoperimetric deficit. Accordingly, no matter what shapes breast height cross-sections assume,
\[ G \leq E(\hat{G}) \leq G + \frac{2}{|L|} \sum_{i} (\tilde{A}_0 - A_{ci}) \] (83)
The second inequality
\[ |\tilde{A}(\alpha) - E_{\theta}(\tilde{A}_1)| \leq \left[ E_{\theta}(\tilde{A}_1) - A_{ci} \right] \tan^2(\alpha / 2) \] (84)
shows that in practice — \( \alpha \) being of the magnitude 1–2°, and \( \tan^2(\alpha/2) \) thus 0.000076–0.00030 — \( \tilde{A}(\alpha) \) becomes very close to the within-cross-section expectation \( E_{\theta}(\tilde{A}_1) \) of the estimator based one random diameter. Accordingly,
\[ E(\hat{G}) = \frac{1}{|L|} \sum_{i} E_{\theta}(\tilde{A}_1) \] (85)
which implies that, regardless of the shapes of breast height cross-sections, the overestimating bias of the basal area estimator by relascope be approximately equal to the bias obtained by caliper all the trees in \( I \) in a random direction and applying the circle area formula (Matérn 1956).
Relative basal area may, of course, be estimated using diameter measurements, if they are available on the trees included in the Bitterlich sample. Then it is possible to apply different area estimators, say \( \hat{A}_X \) and \( \hat{A}_Y \), in the inclusion probability estimation versus the cross-section area estimation: the estimator with estimated inclusion probabilities becomes (cf. Eq. 67)
\[ \hat{G} = \frac{1}{|L|} \sum_{i \in \theta_{\alpha}(x)} \hat{A}_Y / \hat{A}_X \] (86)
and its expectation over the probability distribution of all possible tree samples is (cf. Eq. 69)
For elliptical cross-sections, Grosenbaugh (1958) recommended using the geometric mean of $D_{\min}$ and $D_{\max}$ in the circle area formula for $A_Y$ and the quadratic mean ditto for $A_X$ to minimise the bias. This is a most natural choice: the geometric mean in the circle area formula yields the true area $A_C$ of an ellipse, whereas the quadratic mean gives the largest overestimation and thus minimises the combined error $\sin^2(\alpha/2)/\kappa(\alpha) \cdot A_C/A_X$.

### 4.2 Bitterlich Diameters: Diameters Measured Parallel or Perpendicular to Plot Radius Direction

When a tree to be investigated is selected by Bitterlich sampling, it is a common practice to measure its breast height diameter parallel or perpendicular to the plot radius, that is, parallel or perpendicular to the line segment from the viewing point to the assumed centre of gravity, or pith, of the breast height cross-section (Fig. 17). This diameter is then usually regarded as “random”, implying a correspondence to the diameter with the uniform direction distribution, which is the way in which “random diameter” is usually understood in colloquial language. Although practically sensible and seemingly sound, this practice involves a likely pitfall if the breast height cross-section of the tree is non-circular: Plot radius direction in Bitterlich sampling is a random variable the value of which depends on the outcome of the random experiment of placing the viewing point in the inclusion region. If the inclusion region of a stem deviates from a circle, the probability of a viewing point to be located in a certain direction when viewed from the tree pith varies according to the direction. In Fig. 17, for example, it is more probable to place the viewing point in the way that the plot radius direction becomes $t_Q$ than in the way that it becomes $t_R$ — simply because the line segment from the pith to the inclusion region boundary is longer through the point Q than through the point R. Accordingly, the distribution of the plot radius direction — and, in consequence, the distribution of the diameter measurement direction — is not uniform over $[0, 2\pi)$ — or over $[0, \pi)$, respectively — when viewed from the pith of the tree. In other words, contrary to the usual implicit assumption or belief, the diameter taken parallel or perpendicular to relascope plot radius direction does not correspond to a random diameter with the uniform direction distribution, if the cross-section of a tree is non-circular. In Fig. 17, for example, measuring diameters near $D(t_Q)$ is more probable than measuring diameters near $D(t_R)$ if the measurement practice is to take the diameter parallel to the plot radius, and less probable if the practice is to take the diameter perpendicular to the plot radius.

In the following, we first derive the direction distributions for these Bitterlich diameters, that is, for the diameters measured parallel or perpendicular to plot radius in Bitterlich sampling. Then we derive the expectations, variances and approximate variances of the area estimators similar to those dealt with in Section 3.3.2 but with Bitterlich diameters involved.
4.2.1 Direction Distributions of Bitterlich Diameters

The tree-specific probability distribution of the plot radius direction $\tau \in [0, 2\pi)$ is obtained from the inclusion region $M(\alpha)$ of the tree via geometrical probability. The density mass of $\tau$ between directions $t_1$ and $t_2$, $t_1 < t_2$, is the probability that the viewing point is located in the sector of $M(\alpha)$ edged by the rays emanating from the tree pith in directions $t_1$ and $t_2$ (Fig. 18); since the viewing point is uniformly randomly located in the region of interest, the probability equals the area of the sector divided by the total area of $M(\alpha)$, that is,

$$\Pr \{ t_1 \leq \tau \leq t_2 ; \alpha \} = \frac{|M(\alpha)_{t_1}^{t_2}|}{|M(\alpha)|}. \quad (88)$$

The plot radius directions $\tau$ and $\tau + \pi$ result in the same diameter of the cross-section, and hence the probability that the direction $\xi \in [0, \pi)$ of the diameter taken parallel to plot radius is between $t_1$ and $t_2$ is the sum of the probabilities of $\tau$ being between $t_1$ and $t_2$ or between $t_1 + \pi$ and $t_2 + \pi$ (Fig. 19), that is,

$$\Pr \{ t_1 \leq \xi \leq t_2 ; \alpha \} = \Pr \{ t_1 \leq \tau \leq t_2 ; \alpha \} + \Pr \{ t_1 + \pi \leq \tau \leq t_2 + \pi ; \alpha \}$$

$$= \frac{|M(\alpha)_{t_1}^{t_2}|}{|M(\alpha)|} + \frac{|M(\alpha)_{t_1+\pi}^{t_2+\pi}|}{|M(\alpha)|}. \quad (89)$$

Setting $t_1 = 0$ and $t_2 = \pi$ and restricting $\tau \in [0, \pi)$, we obtain the cumulative distribution function of $\xi$:
\begin{equation}
F_\xi(t; \alpha) = \Pr\{\xi \leq t; \alpha\}
= \Pr\{0 \leq \xi \leq t; \alpha\}
= \Pr\{0 \leq \tau \leq t; \alpha\} + \Pr\{\pi \leq \tau \leq \pi + t; \alpha\}
= \frac{\left| M(\alpha)^{t_1}_\tau \right|}{\left| M(\alpha) \right|} + \frac{\left| M(\alpha)^{t_2}_\pi \right|}{\left| M(\alpha) \right|}.
\end{equation}

Although the inclusion region boundary co-ordinates (Eq. C9 in Appendix C) or radii (Eq. 59) can straightforwardly be expressed as functionals of the support function \( p(\cdot) \), the sector area of the inclusion region cannot, or at least this is rather troublesome; the difficulty lies in the determination of the integration limits (see Appendix D). Therefore, when computing the Bitterlich direction distribution, we resort to numerical integration.

The probability density function \( f_\xi(t; \alpha) \) may be approximated with
\begin{equation}
f_\xi(t; \alpha) \approx \frac{F_\xi(t + \Delta t; \alpha) - F_\xi(t; \alpha)}{\Delta t},
\end{equation}
where \( \Delta t \) is small (0.5°, for example). For the example shapes of Fig. 10, the approximated density functions with viewing angles 1°, 20° and 30° are shown in Fig. 20.

The direction distribution of the diameter taken \textit{perpendicular} to plot radius in Bitterlich sampling is obtained from \( F_\xi(t; \alpha) \) by a location shift of \(-\pi/2\):
\begin{equation}
F_{\xi + \pi/2}(t; \alpha) = \Pr\{\xi + \pi/2 \leq t; \alpha\} = \Pr\{\xi \leq t - \pi/2; \alpha\} = F_\xi(t - \pi/2; \alpha).
\end{equation}
In the example shapes of Fig. 10, taking the Bitterlich diameter perpendicular to plot radius results in larger (or equal; shape F) values on average than taking it parallel to plot radius (Table 7), even though the differences are rather small (about 2.5 % in the ellipse A, less than 1 % in the other shapes). Compared to the diameters with the uniform direction distribution, the Bitterlich diameters parallel to plot radius are on average smaller in a majority
Fig. 21. (Approximated) cumulative distribution functions of the Bitterlich diameters taken parallel \( F_{D(\xi)}(d) \) (black) or perpendicular \( F_{D(\xi+\pi/2)}(d) \) (grey) to plot radius for the six example shapes in Fig. 10. The distributions were determined with viewing angles \( \alpha=1^\circ \) (continuous line) and \( \alpha=30^\circ \) (dashed line). For reference, the cumulative distribution function of the diameter with uniformly distributed direction within \([0, \pi)\) is also shown (thin black line).

of the shapes, whereas the Bitterlich diameters perpendicular to plot radius are on average larger in all the shapes. Interestingly, in all the shapes but one, both types of the Bitterlich diameters seem to have a smaller variance than the diameters with the uniform direction distribution. The correlations between the Bitterlich diameters are the same (up to the third decimal) as the correlations between perpendicular diameters over the uniform direction distribution (Table 1; note that the shape B in this table is not the orbiform in Fig. 10 but the ellipse-like shape in Fig. 5).

From the direction distribution, we get the diameter distribution in a general manner: \( F_{D(\xi)}(d) = \Pr\{D(\xi) \leq d; \alpha\} \), where \( \Pr\{D(\xi) \leq d; \alpha\} \) is the probability mass of directions in which diameter is not larger than \( d \). For the example shapes of Fig. 10, the Bitterlich diameter distributions are shown in Fig. 21.

4.2.2 Estimation of Cross-Section Area by Bitterlich Diameters and Circle Area Formula

As mentioned above, Bitterlich diameters (i.e., the diameters measured parallel or perpendicular to plot radius in Bitterlich sampling) are often used as if they were random diameters with the uniform direction distribution. As the direction distributions of the Bitterlich diameters are actually not uniform if the breast height cross-sections of trees are non-circular, we consider the cross-section area estimators similar to \( \hat{A}_1 \text{--} \hat{A}_5 \) discussed in Section 3.3.2 but now involving Bitterlich diameters. The estimators are of the form

\[
\hat{A} = \frac{\pi}{4} D(\cdot)^2 ,
\]  

(93)
where $D(\cdot)$ is

1. $\xi$ diameter $D(\xi)$ taken parallel to plot radius direction, $\xi \sim F_\xi(\xi; \alpha)$ ($\hat{A}_{1\xi}$)
2. $2\xi$ arithmetic mean of diameters $D(\xi)$ and $D(\xi + \pi/2)$ taken parallel and perpendicular to plot radius direction ($\hat{A}_{2\xi}$)
3. $3\xi$ geometric mean of $D(\xi)$ and $D(\xi + \pi/2)$ ($\hat{A}_{3\xi}$)
4. $4\xi$ arithmetic mean of $D(\xi)$ and an independent random diameter $D(\theta)$, $\xi \sim F_\theta(\xi; \alpha)$, $\theta \sim \text{Uniform}(0, \pi)$ ($\hat{A}_{4\xi}$)
5. $5\xi$ geometric mean of $D(\xi)$ and $D(\theta)$ ($\hat{A}_{5\xi}$)

Hardly surprisingly, the within-cross-section expectations, variances and variance approximations of these estimators closely resemble those of the random estimators $\hat{A}_{1\xi} - \hat{A}_{5\xi}$ (Eqs. 22–26 and 30–37 in Section 3.3.2). If we denote the expectations of $D(\xi)$ and $D(\xi + \pi/2)$ taken over the distribution $F_\xi(\xi; \alpha)$ by $\mu_{D(\xi)}$ and $\mu_{D(\xi + \pi/2)}$, the variances by $\sigma_{D(\xi)}^2$ and $\sigma_{D(\xi + \pi/2)}^2$, and the correlation by $\rho_{D(\xi), D(\xi + \pi/2)}$ (note that for brevity, the viewing angle $\alpha$ affecting the distribution is omitted in the notation) — and recall that $\mu_D$ and $\sigma_D^2$ stand for the expectation and the variance of $D(\theta)$ over the uniform distribution of $\theta$ — the expectations of the estimators become as follows:

\[ E(\hat{A}_{1\xi}) = \frac{\pi}{4} \mu_{D(\xi)}^2 + \frac{\pi}{4} \sigma_{D(\xi)}^2, \quad \text{(94)} \]

\[ E(\hat{A}_{2\xi}) = \frac{\pi}{4} \left( \mu_{D(\xi)} + \mu_{D(\xi + \pi/2)} \right)^2 - \frac{\pi}{16} \left( \sigma_{D(\xi)}^2 + \sigma_{D(\xi + \pi/2)}^2 + 2\sigma_{D(\xi)} \sigma_{D(\xi + \pi/2)} \rho_{D(\xi)} \left( \frac{\pi}{2} \right) \right), \quad \text{(95)} \]

\[ E(\hat{A}_{3\xi}) = \frac{\pi}{4} \mu_{D(\xi)} \mu_{D(\xi + \pi/2)} + \frac{\pi}{4} \sigma_{D(\xi + \pi/2)} \rho_{D(\xi)} \left( \frac{\pi}{2} \right), \quad \text{(96)} \]

\[ E(\hat{A}_{4\xi}) = \frac{\pi}{4} \left( \mu_{D(\xi)} + \mu_{D} \right)^2 + \frac{\pi}{16} \left( \sigma_{D(\xi)}^2 + \sigma_{D}^2 \right), \quad \text{(97)} \]

and

\[ E(\hat{A}_{5\xi}) = \frac{\pi}{4} \mu_{D(\xi)} \mu_{D}. \quad \text{(98)} \]

Unlike with the expectations of the random estimators $\hat{A}_{1\xi} - \hat{A}_{5\xi}$, it is not easy to see whether these expectations overestimate or underestimate the convex area of a cross-section (because the relationship between $\pi \mu_{D(\xi)}^2/4$, or $\pi \mu_{D(\xi + \pi/2)}^2/4$, and $A_C$ is not straightforward to derive with no assumptions about cross-section shape; cf. Eq. 20 in Section 3.3.2). The exact variances are given by

\[ \text{Var}(\hat{A}_{1\xi}) = \frac{\pi^2}{16} \left[ E[D(\xi)^4] - \left( \mu_{D(\xi)}^2 + \sigma_{D(\xi)}^2 \right)^2 \right], \quad \text{(99)} \]
Finally, the variance approximations by the delta method become
\[
\begin{align*}
\text{Var}(\hat{A}_{22}) &= \frac{\pi^2}{16} \left\{ \frac{1}{16} E[D(\xi)^4] + \frac{1}{16} E\left[D\left(\xi + \frac{\pi}{2}\right)^4\right] + \frac{1}{4} E\left[D(\xi)^2 D\left(\xi + \frac{\pi}{2}\right)^2\right] \\
&\quad + \frac{1}{4} E\left[D(\xi) D\left(\xi + \frac{\pi}{2}\right)^3\right] + \frac{3}{8} E\left[D(\xi)^3 D\left(\xi + \frac{\pi}{2}\right)^2\right] \\
&\quad - \frac{1}{16} \left[\left(\mu_{D(\xi)} + \mu_{D(\xi+\pi/2)}\right)^2 + \sigma^2_{D(\xi)} + \sigma^2_{D(\xi+\pi/2)}\right] \\
&\quad + 2\sigma_{D(\xi)} \sigma_{D(\xi+\pi/2)} \rho_{D(\xi)} \left(\frac{\pi}{2}\right)^2 \right\} ,
\end{align*}
\]
(100)

\[
\begin{align*}
\text{Var}(\hat{A}_{3\xi}) &= \frac{\pi^2}{16} E\left[D(\xi)^3 D\left(\xi + \frac{\pi}{2}\right)^2\right] \\
&\quad - \left[\mu_{D(\xi)} \mu_{D(\xi+\pi/2)} + \sigma_{D(\xi)} \sigma_{D(\xi+\pi/2)} \rho_{D(\xi)} \left(\frac{\pi}{2}\right)^2\right] ,
\end{align*}
\]
(101)

\[
\begin{align*}
\text{Var}(\hat{A}_{4\xi}) &= \frac{\pi^2}{16} \left\{ \frac{1}{16} E[D(\xi)^4] + \frac{1}{16} E[D(\theta)^4] + \frac{1}{4} \mu_D E[D(\xi)^3] \\
&\quad + \frac{1}{4} \mu_{D(\xi)} E[D(\theta)^3] \right\} \left[ \left(\mu_{D(\xi)} + \mu_{D(\xi+\pi/2)}\right)^2 + \sigma^2_{D(\xi)} + \sigma^2_{D(\xi+\pi/2)}\right] \\
&\quad - \frac{1}{16} \left[\left(\mu_{D(\xi)} + \mu_D\right)^2 + \sigma^2_{D(\xi)} + \sigma^2_D\right] ,
\end{align*}
\]
(102)

and
\[
\begin{align*}
\text{Var}(\hat{A}_{5\xi}) &= \frac{\pi^2}{16} \left[\mu^2_{D(\xi)} + \sigma^2_{D(\xi)}\right] \left(\mu^2_D + \sigma^2_D\right) - \mu^2_{D(\xi)} \mu^2_D ,
\end{align*}
\]
(103)

Finally, the variance approximations by the delta method become
\[
\begin{align*}
\text{Vär}(\hat{A}_{1\xi}) &= \frac{\pi^2}{4} \mu^2_{D(\xi)} \sigma^2_{D(\xi)} ,
\end{align*}
\]
(104)

\[
\begin{align*}
\text{Vär}(\hat{A}_{22}) &= \frac{\pi^2}{16} \left[\mu_{D(\xi)} \mu_{D(\xi+\pi/2)} + \mu_{D(\xi+\pi/2)} \mu_D\right] \left[\sigma^2_{D(\xi)} + \sigma^2_{D(\xi+\pi/2)} + 2\sigma_{D(\xi)} \sigma_{D(\xi+\pi/2)} \rho_{D(\xi)} \left(\frac{\pi}{2}\right)\right] ,
\end{align*}
\]
(105)

\[
\begin{align*}
\text{Vär}(\hat{A}_{3\xi}) &= \frac{\pi^2}{16} \left[\mu^2_{D(\xi)} \sigma^2_{D(\xi+\pi/2)} + \mu^2_{D(\xi+\pi/2)} \sigma^2_D\right] \\
&\quad + 2\mu_{D(\xi)} \mu_{D(\xi+\pi/2)} \sigma_{D(\xi)} \sigma_{D(\xi+\pi/2)} \rho_{D(\xi)} \left(\frac{\pi}{2}\right) ,
\end{align*}
\]
(106)
\[ \text{Vár}(\hat{A}_4) = \frac{\pi^2}{16} \left( \frac{\mu_D(\hat{\xi}) + \mu_D}{2} \right)^2 \left( \sigma^2_{D(\hat{\xi})} + \sigma^2_D \right), \] (107)

and

\[ \text{Vár}(\hat{A}_5) = \frac{\pi^2}{16} \left( \mu^2_D \sigma^2_{D(\hat{\xi})} + \mu^2_D \sigma^2_D \right). \] (108)

Note that these variance approximations depend not only on the way of selecting the diameters — as with the estimators \( \hat{A}_1 - \hat{A}_5 \) — but also on the type of mean employed in the estimator. For the estimator \( \hat{A}_{55} \), the approximate variance is easily seen to underestimate the true variance, exactly as was the case with the estimator \( \hat{A}_5 \) involving two diameters with the uniform direction distribution (cf. Eq. 38 in Section 3.3.2).

In practice, it is often more convenient to measure the (first) diameter perpendicular to plot radius than parallel to it. This practice results in three more area estimators of the form above to be considered: the modifications \( \hat{A}_{1\xi90}, \hat{A}_{4\xi90} \) and \( \hat{A}_{5\xi90} \) of \( \hat{A}_1, \hat{A}_4 \) and \( \hat{A}_5 \) involving \( D(\xi+\pi/2) \) instead of \( D(\xi) \). The within-cross-section expectations and variances of these estimators are obtained from those of \( \hat{A}_1, \hat{A}_4 \) and \( \hat{A}_5 \) by substituting the moments of \( D(\xi+\pi/2) \) for the moments of \( D(\xi) \).

In Tables 8 and 9, the expectations, variances and approximate variances of the estimators \( \hat{A}_1 - \hat{A}_{55}, \hat{A}_{1\xi90}, \hat{A}_{4\xi90} \) and \( \hat{A}_{5\xi90} \) are given for the example shapes in Fig. 10; the distributions of \( \xi \) (and \( \xi+\pi/2 \)) were determined with viewing angles 1° and 30°. None of the

<table>
<thead>
<tr>
<th>Shape</th>
<th>A</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{E}(\hat{A}_{1\xi90})/A ) (%)</td>
<td>1000</td>
<td>1016</td>
<td>1014</td>
<td>1034</td>
<td>1040</td>
</tr>
<tr>
<td>(1003)</td>
<td>(1017)</td>
<td>(1014)</td>
<td>(1031)</td>
<td>(1036)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{2\xi90})/A ) (%)</td>
<td>1019</td>
<td>1022</td>
<td>1019</td>
<td>1037</td>
<td>1038</td>
</tr>
<tr>
<td>(1019)</td>
<td>(1022)</td>
<td>(1018)</td>
<td>(1034)</td>
<td>(1034)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{3\xi90})/A ) (%)</td>
<td>1013</td>
<td>1020</td>
<td>1018</td>
<td>1037</td>
<td>1040</td>
</tr>
<tr>
<td>(1013)</td>
<td>(1020)</td>
<td>(1017)</td>
<td>(1034)</td>
<td>(1036)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{4\xi90})/A ) (%)</td>
<td>1010</td>
<td>1019</td>
<td>1016</td>
<td>1032</td>
<td>1035</td>
</tr>
<tr>
<td>(1011)</td>
<td>(1020)</td>
<td>(1016)</td>
<td>(1031)</td>
<td>(1033)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{5\xi90})/A ) (%)</td>
<td>1006</td>
<td>1018</td>
<td>1014</td>
<td>1030</td>
<td>1034</td>
</tr>
<tr>
<td>(1008)</td>
<td>(1019)</td>
<td>(1014)</td>
<td>(1029)</td>
<td>(1032)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{1\xi90})/A ) (%)</td>
<td>1050</td>
<td>1032</td>
<td>1030</td>
<td>1047</td>
<td>1040</td>
</tr>
<tr>
<td>(1046)</td>
<td>(1031)</td>
<td>(1028)</td>
<td>(1043)</td>
<td>(1036)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{4\xi90})/A ) (%)</td>
<td>1034</td>
<td>1027</td>
<td>1024</td>
<td>1039</td>
<td>1035</td>
</tr>
<tr>
<td>(1033)</td>
<td>(1026)</td>
<td>(1023)</td>
<td>(1037)</td>
<td>(1033)</td>
<td></td>
</tr>
<tr>
<td>( \text{E}(\hat{A}_{5\xi90})/A ) (%)</td>
<td>1031</td>
<td>1026</td>
<td>1022</td>
<td>1037</td>
<td>1034</td>
</tr>
<tr>
<td>(1029)</td>
<td>(1025)</td>
<td>(1021)</td>
<td>(1035)</td>
<td>(1032)</td>
<td></td>
</tr>
</tbody>
</table>
estimators systematically underestimates the true area. As could be expected on the basis of diameter means (Table 7), the estimators involving the Bitterlich diameter parallel to plot radius ($\hat{A}_{1\xi}$, $\hat{A}_{4\xi}$ and $\hat{A}_{5\xi}$) yield smaller overestimating biases than those involving the Bitterlich diameter perpendicular to plot radius ($\hat{A}_{1\xi90}$, $\hat{A}_{4\xi90}$ and $\hat{A}_{5\xi90}$). In all the shapes but F, the former estimators perform better, in terms of bias, than the estimators that were earlier found to be the best ($\hat{A}_0$ and $\hat{A}_3$; Table 8 vs. Table 2). As to the estimator variance, the type of the Bitterlich diameter involved in the estimator appears to have practically no effect (Table 9). The approximated variances virtually equal the true variances, which in turn hardly deviate from the variances obtained with the diameters with the uniform direction distribution (Table 9 vs. Table 3).

Table 9. Square roots of the variances (Eqs. 99–103; Sd) and approximate variances (Eqs. 104–108; $\tilde{S}_d$) of the area estimators $\hat{A}_{1\xi}$–$\hat{A}_{5\xi}$, $\hat{A}_{1\xi90}$, $\hat{A}_{4\xi90}$ and $\hat{A}_{5\xi90}$ for the example shapes in Fig. 10 (the orbiform B with constant diameter in every direction is omitted, because in it the variances are identically zero), expressed in permille of true area. The distributions of the directions $\xi$ and $\xi+\pi/2$ of the Bitterlich diameters were determined with the viewing angle $\alpha=1^\circ$; with $\alpha=30^\circ$ the results were the same (up to a difference of 1 permille in a handful of cases).

<table>
<thead>
<tr>
<th>Shape</th>
<th>A</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sd($\hat{A}_{1\xi}$)/A (%)</td>
<td>157</td>
<td>90</td>
<td>101</td>
<td>118</td>
<td>91</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{1\xi}$)/A (%)</td>
<td>156</td>
<td>90</td>
<td>99</td>
<td>118</td>
<td>91</td>
</tr>
<tr>
<td>Sd($\hat{A}_{2\xi}$)/A (%)</td>
<td>4</td>
<td>0</td>
<td>45</td>
<td>83</td>
<td>91</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{2\xi}$)/A (%)</td>
<td>4</td>
<td>0</td>
<td>45</td>
<td>83</td>
<td>91</td>
</tr>
<tr>
<td>Sd($\hat{A}_{3\xi}$)/A (%)</td>
<td>9</td>
<td>1</td>
<td>44</td>
<td>83</td>
<td>91</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{3\xi}$)/A (%)</td>
<td>5</td>
<td>0</td>
<td>45</td>
<td>83</td>
<td>91</td>
</tr>
<tr>
<td>Sd($\hat{A}_{4\xi}$)/A (%)</td>
<td>112</td>
<td>64</td>
<td>71</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{4\xi}$)/A (%)</td>
<td>112</td>
<td>64</td>
<td>70</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>Sd($\hat{A}_{5\xi}$)/A (%)</td>
<td>112</td>
<td>64</td>
<td>71</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{5\xi}$)/A (%)</td>
<td>112</td>
<td>64</td>
<td>70</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>Sd($\hat{A}_{1\xi90}$)/A (%)</td>
<td>157</td>
<td>90</td>
<td>106</td>
<td>118</td>
<td>91</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{1\xi90}$)/A (%)</td>
<td>159</td>
<td>90</td>
<td>104</td>
<td>118</td>
<td>91</td>
</tr>
<tr>
<td>Sd($\hat{A}_{4\xi90}$)/A (%)</td>
<td>112</td>
<td>64</td>
<td>73</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{4\xi90}$)/A (%)</td>
<td>113</td>
<td>64</td>
<td>72</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>Sd($\hat{A}_{5\xi90}$)/A (%)</td>
<td>113</td>
<td>64</td>
<td>72</td>
<td>83</td>
<td>64</td>
</tr>
<tr>
<td>$\tilde{S}<em>d$($\hat{A}</em>{5\xi90}$)/A (%)</td>
<td>113</td>
<td>64</td>
<td>72</td>
<td>83</td>
<td>64</td>
</tr>
</tbody>
</table>
5 Estimation of Stem Volume

Analogously to the previous discussion on cross-section area estimation in Chapter 3, we now want to explore how diameter variation in non-circular cross-sections is reflected, via different diameter selection methods, to stem volume estimates or predictions. We consider three volume estimation methods — a volume equation, a theoretical general estimator (a definite integral of a cross-section area estimation function), and a definite integral of a non-parametric stem curve (a special case of the general estimator) — in which the input consists of tree height and diameters taken at two or more known (non-random) heights. Our interest is in the application of the volume estimation methods: we want to quantify how much of the estimation bias and error variance is attributable to uncertainty in the explanatory variables (diameters) and therefore seek to distinguish the error component due to diameter variation from the components related to other sources of uncertainty, such as model specification, parameter estimation and residual variation. The theoretical discussion in this chapter is a preliminary to the empirical part of the study, where the three types of volume estimators are investigated with real stems.

5.1 Practical Volume Estimator: Laasasenaho Volume Equation

The volume equations constructed by Laasasenaho (1982) are commonly used in Finland for predicting stem volume of standing trees from basic field measurements. We consider here the three-variable equation, with which the volumes of the sample trees are predicted for example in the Finnish national forest inventory (Tomppo et al. 1997). The model is of the form

\[ V = \beta_1 D(\cdot, 1.3)^2 + \beta_2 D(\cdot, 1.3)^2 H + \beta_3 D(\cdot, 1.3) H + \beta_4 D(\cdot, 1.3)^2 H^2 + \beta_5 D(\cdot, 1.3)^2 + D(\cdot, 6)D(\cdot, 6) + \beta_6 D(\cdot, 6)^2 (H - 6) + \varepsilon_L, \]

where \( V \) denotes the stem volume, \( D(\cdot, 1.3) \) is the breast height diameter, \( D(\cdot, 6) \) is the upper diameter taken at the height of 6 m, \( H \) stands for the stem height, and \( \varepsilon_L \) is a random error with certain properties discussed in more detail below. As indicated by the denotations \( D(\cdot, 1.3) \) and \( D(\cdot, 6) \), the diameters in the model have not been explicitly specified (in which direction they should be measured, whether they should be computed as a mean of two or more diameters etc.). We fix them here as those yielding the true cross-section areas when substituted in the circle area formula, denoted by \( D_A(1.3) \) and \( D_A(6) \):

\[ V = \beta_1 D_A(1.3)^2 + \beta_2 D_A(1.3)^2 H + \beta_3 D_A(1.3) H + \beta_4 D_A(1.3)^2 H^2 + \beta_5 D_A(1.3)^2 + D_A(6)D_A(6) + \beta_6 D_A(6)^2 (H - 6) + \varepsilon_L. \]

This specification, albeit not very feasible in practice, is well motivated by the geometrical background of the model (see Laasasenaho 1982, p. 35–39) and also serves our purpose of extracting the effect of non-circularity on prediction error.

The model was constructed on the basis of data on 5053 trees sampled from all over Finland (Laasasenaho 1982), and hence it may be considered to express the dependence of stem volume on diameters and height in the Finnish tree population, with different parameter values \( \beta_1, ..., \beta_6 \) for different species. This population is finite but so large that sampling from it more or less corresponds to sampling from a theoretical infinite population.
Randomness arises from choosing a tree in the population, and both the response variable and the explanatory variables may be regarded as random variables, the joint distribution of which is generated by repeated independent samplings from the population. As usual with random explanatory variables, the distributional properties of the error term $e_L$ are defined conditional on the explanatory variables: given any possible combination of values of $D_A(1.3)$, $D_A(6)$ and $H$ in the population, $e_L$ is assumed to have zero expectation and variance proportional to $D_A(1.3)^4H^2$ (Laasasenaho 1982, p. 36 and p. 43–44); further, $e_L$’s of separate trees are assumed to be independent. From the zero conditional expectation of $e_L$, it follows that the expectation of $e_L$ is zero also marginally, that is, over the joint distribution of $D_A(1.3)$, $D_A(6)$ and $H$ in the population.

We assume the equation to express the relation correctly, that is, the functional form as well as the values of the coefficients $\beta_1, \ldots, \beta_6$ to be true in the population of our interest, and wish to use it to predict the volume of a (randomly selected) tree in this population, when the two diameters and the height of the tree are known. The best predictor is the model without the error term, that is, the conditional expectation of the volume:

$$\hat{V}_L = \frac{V}{L} = E[L \mid D_A(1.3), D_A(6), H]$$

$$= \beta_1 D_A(1.3)^2 + \beta_2 D_A(1.3)^3 H + \beta_3 D_A(1.3)^4 H + \beta_4 D_A(1.3)^5 H^2 + \beta_5 D_A(1.3)^6 + \beta_6 D_A(6)^2 (H - 6).$$

(111)

As there is no estimation error in the parameter values, the prediction error $V - \hat{V}_L$ naturally equals the random error $e_L$ and follows the assumptions made on $e_L$ (i.e., $E[V - \hat{V}_L \mid D_A(1.3), D_A(6), H] = 0$ and $Var[V - \hat{V}_L \mid D_A(1.3), D_A(6), H] \propto D_A(1.3)^4H^2$).

If we then view volume prediction within the tree of interest, the setting becomes very similar to that of cross-section area estimation in Chapter 3: volume is now a fixed property of the tree that we estimate by a fixed estimator $\hat{V}_L$, and the estimation error $V - \hat{V}_L$ becomes a non-random scalar (a realisation of $e_L$). We adopt this within-tree view for a while to deal with volume predictors involving random diameters.

Consider estimating the stem volume of the tree by using some diameters other than $D_A(1.3)$ and $D_A(6)$, measured without error at the heights of 1.3 m and 6 m. Denote these diameters by $D(\theta, 1.3)$ and $D(\theta, 6)$, $\theta$ referring here to the diameter direction selection method in general. As the height $H$ is known, the estimator can now be written as a function of the diameters only:

$$\hat{V}_L(\theta) = c_1 D(\theta, 1.3)^2 + c_2 D(\theta, 1.3)^3 + c_3 D(\theta, 1.3)D(\theta, 6) + c_4 D(\theta, 6)^2,$$

(112)

where $c_1 = \beta_1 + \beta_2 H + \beta_4 H^2 + \beta_5$, $c_2 = \beta_3 H$, $c_3 = \beta_5$, and $c_4 = \beta_5 + \beta_6 (H - 6)$. If the selection of $D(\theta, 1.3)$ and $D(\theta, 6)$ involves randomness, the estimator becomes a random variable, the within-tree distribution of which is determined by the diameter selection method and the shape of the cross-sections at the heights of 1.3 m and 6 m. As with random area estimators (Section 3.3.2), the random volume estimator can be thought to consist of a systematic part and a stochastic part — the within-tree expectation over the diameter direction distribution and a random sampling error with zero expectation over the same distribution:

$$\hat{V}_L(\theta) = \frac{V}{L} = E[L \mid \hat{V}_L(\theta)] + \nu_L(\theta).$$

(113)

As the estimator is a linear combination of the second and third powers of diameters and their cross-product, its within-tree expectation and variance are obtained by means of the
corresponding diameter moments and product moments taken over the diameter direction distribution:

\[
E_{\theta}[\hat{V}_L(\theta)] = c_1E_{\theta}[D(\theta,1.3)^2] + c_2E_{\theta}[D(\theta,1.3)^3] \\
+ c_3E_{\theta}[D(\theta,1.3)D(\theta,6)] + c_4E_{\theta}[D(\theta,6)^2], 
\]

(114)

and

\[
\text{Var}_{\theta}[v_L(\theta)] = \text{Var}_{\theta}[\hat{V}_L(\theta)] \\
= E_{\theta}[\hat{V}_L(\theta)^2] - \left\{ E_{\theta}[\hat{V}_L(\theta)] \right\}^2 \\
= c_2^2E_{\theta}[D(\theta,1.3)^6] + 2c_1c_2E_{\theta}[D(\theta,1.3)^5] + c_2^2E_{\theta}[D(\theta,1.3)^4] \\
+ 2c_2E_{\theta}[D(\theta,1.3)^5D(\theta,6)] + 2c_2E_{\theta}[D(\theta,1.3)^4D(\theta,6)^2] \\
+ 2c_2E_{\theta}[D(\theta,1.3)^3D(\theta,6)^3] + (c_3^2 + 2c_1c_3)E_{\theta}[D(\theta,1.3)^2D(\theta,6)^2] \\
+ 2c_3E_{\theta}[D(\theta,1.3)^3D(\theta,6)^2] + c_3^2E_{\theta}[D(\theta,6)^4] \\
- \left\{ E_{\theta}[\hat{V}_L(\theta)] \right\}^2. 
\]

(115)

Each diameter selection method can in principle be applied either \textit{dependently} or \textit{independently} at the two observation heights within the stem: in the former, the diameter direction is selected at breast height, and the upper diameter is then measured in the same direction; in the latter, the diameter directions at the two heights are selected independently. In Bitterlich sampling, however, the idea of selecting diameters parallel or perpendicular to plot radius independently at the two heights is unfeasible (the plot radius direction is determined only once, from one viewing point at the breast height level); therefore, only dependent selection can be considered with the methods involving Bitterlich diameters. In the above expressions of within-tree expectation and variance, independent and dependent selection differ from each other only in terms of the product moments \(E[D(\theta,1.3)^kD(\theta,6)^p]\), \(k, p \in \mathbb{Q}^+\), which in the independent case reduce to the products of moments \(E[D(\theta,1.3)^k]E[D(\theta,6)^p]\). In Appendices E and F, more elaborate method-specific versions of the above general expressions of within-tree expectation and variance are given for the diameter selection methods involving randomness and considered previously in area estimation (methods 1–5 involving the uniform direction distribution, see Section 3.3.2; methods 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90 involving Bitterlich diameter direction distribution, see Section 4.2.2).

No matter if we are viewing volume determination at the within-tree level (volume estimation) or at the population level (volume prediction), the volume error \(V - \hat{V}_L(\theta)\) may be thought to consist of two components — the error term inherent in the model (the difference between the true volume \(V\) and the best predictor/estimator \(\hat{V}_L\)) and the error due to diameter selection (the difference between the best predictor/estimator \(\hat{V}_L\) and the predictor/estimator \(\hat{V}_L(\theta)\) obtained with the particular diameter selection method \(\theta\)); further, the error due to diameter selection can be divided into two components — the within-tree bias \(\hat{V}_L - E_{\theta}[\hat{V}_L(\theta)]\) taken with respect to the best predictor/estimator and a random sampling error \(v_L(\theta)\):
\[V - \hat{V}_L(\theta) = (V - \hat{V}_L) + \left[\hat{V}_L - \hat{V}_L(\theta)\right]\]
\[= \epsilon_L + \left\{\hat{V}_L - \left[E_\theta[\hat{V}_L(\theta)] + \nu_L(\theta)\right]\right\}\]
\[= \epsilon_L + \left\{\hat{V}_L - E_\theta[\hat{V}_L(\theta)]\right\} - \nu_L(\theta) .\]

At the within-tree level (conditional on the selected tree), only the within-tree diameter sampling error \(\nu_L(\theta)\) is a random variable (with zero expectation, variance \(\text{Var}_\theta[\nu_L(\theta)] = \text{Var}[\hat{V}_L(\theta)]\)), and the distribution determined by the diameter sampling method and the shapes of the cross-sections at 1.3 m and 6 m), whereas the model error \(\epsilon_L\) and the within-tree systematic error due to diameter selection \(\hat{V}_L - E_\theta[\hat{V}_L(\theta)]\) are non-random scalars and may in principle be compared to each other. At the population level, however, all the prediction error components are random variables. Yet we have only made assumptions on the model error \(\epsilon_L\). In any event, the prediction bias \(E[V - \hat{V}_L(\theta)]\) at the population level consists of only the between-trees expectation \(E\{\hat{V}_L - E_\theta[\hat{V}_L(\theta)]\}\) of the within-tree bias due to diameter selection, as \(E(\epsilon_L) = 0\) by assumption and \(E[\nu_L(\theta)] = E_{\text{tree}}[E_\theta[\nu_L(\theta) | \text{tree}] = 0\). The prediction error variance \(\text{Var}[V - \hat{V}_L(\theta)]\), in turn, comprises the variances of all the error components \((\epsilon_L, \hat{V}_L - E_\theta[\hat{V}_L(\theta)]\) and \(\nu_L(\theta)\) taken over their marginal distributions as well as double the pairwise covariances computed over their joint distribution.

Above we assumed the model (Eq. 110) to be true. If, however, the true values for the parameters \(\beta_1, ..., \beta_6\) were unknown and we had to use estimates obtained from any sub-population, the prediction error \(V - \hat{V}_L\) would also contain a component due to the parameter estimation error (see e.g. Fox 1984; cf. Gregoire and Williams 1992): even if the parameters were unbiasedly estimated, the prediction bias at the population level would be zero only when averaged over (infinitely many) repeated parameter estimations; furthermore, the prediction error variance would increase by a non-zero term depending on the variances and covariances of the parameter estimates and the values of the predicting variables.

5.2 General Volume Estimator: Definite Integral of Cross-Section Area Estimation Function

The true volume of a stem is attained as a definite integral of the cross-section area function — a continuous and bounded function \(A: [0, H] \rightarrow [0, \infty)\) expressing how the cross-section area perpendicular to the vertical stem axis changes along the position in the axis:

\[V = \int_0^H A(h) \, dh .\]

Substituting for the area function an estimate \(\hat{A}(h; \theta)\) (where \(\theta\) just generally refers to the features of the function distinguishing it from other area estimation functions) results in a volume estimator of the form

\[\hat{V}_L(\theta) = \int_0^H \hat{A}(h; \theta) \, dh .\]

Estimators of this kind, where no presumptions (besides boundedness) are necessarily made on \(\hat{A}(h; \theta)\), are here referred to as general volume estimators. (We can not require continuity of \(\hat{A}(h; \theta)\), if we want to allow it to be constructed from random diameters selected independently of each other at all the heights \(h \in [0, H]\); see the discussion below.)
Unlike with the volume equation earlier, we now choose to disregard the tree population level — although many population level models for the area estimation function (cross-section area taper models) are to be found in the literature (for a statistical consideration of prediction by such models, without uncertainty in the explanatory variables, refer e.g. to Gregoire et al. 2000). Instead, we adopt the within-tree view and simply regard the stem volume as a fixed property of our tree of interest, which we then want to estimate by means of an estimated cross-section area function (typically derived from either fixed or random diameters taken at predetermined heights) and the tree height. Within-tree randomness in the volume estimator then arises from randomness in the estimated area function (typically stemming from the use of random diameters in the construction of the function).

If \( \hat{A}(h; \theta) \) is a random function (\( \theta \) now referring to the source of randomness, typically the selection method of random diameter directions), it can equivalently be regarded as a continuous parameter stochastic process \( \{\hat{A}(h; \theta), h \in [0, H]\} \) within the tree (cf. Rao 1979, p. 2). We are interested in the definite integrals of this process (to get the volume estimator) and of its transformation, the area estimation error process \( \{\hat{A}(h; \theta) - A(h), h \in [0, H]\} \) (to get the volume estimation error). The definite integrals are well-defined (as limits, in the sense of convergence in mean square, of the sequences of approximating sums (Riemann sums) over \([0, H]\) or any of its sub-intervals), if the processes satisfy the following sufficient (but not necessary!) conditions (Parzen 1962, p. 78–79): First, the processes should have finite second moments. In our case, this is necessarily true, as both \( \{\hat{A}(h; \theta)\} \) and \( \{A(h)\} \) are bounded. Second, the processes should have continuous mean and covariance functions \( \mu_{\hat{A}}(h; \theta) = \mathbb{E}_\theta[\hat{A}(h; \theta)] \), \( \mu_{\hat{A}A}(h; \theta) = \mathbb{E}_\theta[\hat{A}(h; \theta) - A(h)] \) and \( \gamma_{\hat{A}}(h, k; \theta) = \text{Cov}_\theta[\hat{A}(h; \theta), \hat{A}(k; \theta)] \) = \text{Cov}_\theta[\hat{A}(h; \theta) - A(h), \hat{A}(k; \theta) - A(k)] = \gamma_{\hat{A}A}(h, k; \theta) \). Although \( \hat{A}(h; \theta) \) based on diameters taken in independent random directions at each height is not necessarily continuous, its mean and covariance functions are, as they consist of diameter means, variances and covariances that change with height within a stem smoothly, with no discontinuity. Hence, we can assume that our processes fulfil also the second condition.

For processes satisfying the above conditions, the expectation and variance of their definite integrals are given by the definite integrals of the mean and covariance functions (Parzen 1962, p. 79). Accordingly, for the area estimation process and the area estimation error process

\[
E_\theta[\hat{V}_G(\theta)] = E_\theta \left[ \int_0^H \hat{A}(h; \theta) \, dh \right]
= \int_0^H \mu_{\hat{A}}(h; \theta) \, dh,
\]

\[
E_\theta[\hat{V}_G(\theta) - V] = E_\theta \left[ \int_0^H \hat{A}(h; \theta) \, dh - \int_0^H A(h) \, dh \right]
= E_\theta \left[ \int_0^H [\hat{A}(h; \theta) - A(h)] \, dh \right]
= \int_0^H \mu_{\hat{A}A}(h; \theta) \, dh,
\]

and
In other words, the within-tree expectation and variance of a general volume estimator are attained by integrating the mean and covariance functions of the underlying area estimation process; further, the within-tree bias is obtained by integrating the mean function of the area estimation error process.

A theoretical upper bound for the within-tree variance of a general volume estimator is attained by assuming that the elements of the area estimation process are fully positively correlated, that is, by assuming that $\rho_{\hat{A}(h), \hat{A}(k)} = \text{Corr}_{\theta}[\hat{A}(h; \theta), \hat{A}(k; \theta)] = 1$ for all $h$ and $k$ in $[0, H]$. In this case, the covariance function of the area estimation process becomes

$\gamma_{\hat{A}}(h,k; \theta) = \frac{\sigma_{\hat{A}}^2(h; \theta)}{\sigma_{\hat{A}}^2(k; \theta)}$,  

(122)

where $\sigma_{\hat{A}}^2:[0, H] \rightarrow [0, \infty)$, $\sigma_{\hat{A}}^2(h; \theta) = \text{Var}_{\theta}[\hat{A}(h; \theta)]$ is the variance function of the area estimation process, and this yields

$\text{Var}_{\theta}[\hat{V}_G(\theta) \mid \rho_{\hat{A}}(h, k) = 1] = \left[ \int_0^H \sqrt{\gamma_{\hat{A}}(h; \theta)} \, dh \right]^2$  

(123)

as the volume estimator variance.

An area estimation process with uncorrelated elements makes another interesting special case. Envisage a process where the correlation between elements $\hat{A}(h; \theta)$ and $\hat{A}(k; \theta)$ decreases with an increasing distance between $h$ and $k$, finally to vanish with distances larger than some threshold value. It can be shown that as the threshold distance is diminished toward zero, that is, as the correlations between elements nearer and nearer each other are made vanish, the variance of the corresponding volume estimator (the variance of the integral of the process) tends to zero. The “limiting process” with mutually uncorrelated elements, that is, with $\text{Corr}_{\theta}[\hat{A}(h; \theta), \hat{A}(k; \theta)] = 0$ for all $h \neq k$ in $[0, H]$, is necessarily discontinuous and possesses a discontinuous covariance function with non-zero values $(\sigma_{\hat{A}}^2(h; \theta))$ only within the diagonal line $k=h$. In our tree stem context, this kind of process would arise, for example, if $\hat{A}(h; \theta)$ was based on diameters taken in independent random directions at all the heights $h \in [0, H]$. In practice, however, this process has no natural construction: we may well generate individual process elements (area estimators) that are uncorrelated with each other — by independent selection of random diameter directions, for example — but only sparsely, at a finite number of observation heights, not at the infinite number of all possible heights. To obtain a volume estimate, we need to know the values of the area estimation process at all the heights $h \in [0, H]$, and these we attain by assuming continuity — and, thus, some degree of correlation — between our finite number of observed values. Ergo, from
a practical point of view, an area estimation process with uncorrelated elements is just an abstraction that cannot indeed exist in our tree stem context.

A natural practice would be to assume a model for the area estimation function $\hat{A}(h; \theta)$ in a tree and derive from it (empirical) approximations of $\mu_{\hat{A}}(h; \theta)$, $\gamma_{\hat{A}}(h, k; \theta)$, and $\sigma_{\hat{A}}^2(h; \theta)$ (cf. Section 5.3). In the empirical part of this study, however, we will construct crude approximations of $\mu_{\hat{A}}(h; \theta)$, $\gamma_{\hat{A}}(h, k; \theta)$, and $\sigma_{\hat{A}}^2(h; \theta)$ in a data-driven manner, without an explicit model for $\hat{A}(h; \theta)$: We will consider area estimation processes where the area estimator at each height is one of the area estimators discussed in Sections 3.3 and 4.2.2. For each process, we will approximate $\mu_{\hat{A}}(h; \theta)$, $\gamma_{\hat{A}}(h, k; \theta)$, and $\sigma_{\hat{A}}^2(h; \theta)$ by just interpolating between discrete observations of the process mean, covariance and variance in the stem — between the elements of a vector containing the within-cross-section biases ($E_\theta[\hat{A}(h; \theta) - A(h)]_{h \in H}$), a matrix containing the between-cross-sections covariances ($\text{Cov}_\theta[\hat{A}(h; \theta), \hat{A}(k; \theta)]_{h, k \in H}$), and a vector containing the within-cross-section variances ($\text{Var}_\theta[\hat{A}(h; \theta)]_{h \in H}$) of the area estimator observed at fixed heights $H = \{h(1), h(2), \ldots, h(m)\}$ in the stem. Evidently, a unique area estimation function $\hat{A}(h; \theta)$ that would exactly correspond to the obtained $\mu_{\hat{A}}(h; \theta)$, $\gamma_{\hat{A}}(h, k; \theta)$, and $\sigma_{\hat{A}}^2(h; \theta)$ may not even exist, whereas there are likely to be many whose mean, covariance and variance functions are reasonably adequately approximated by the obtained functions.

5.3 Volume Estimator Based on Non-Parametric Stem Curve: Definite Integral of Interpolating Cubic Splines

Irrespective of the variation in cross-section shape in the stem, the area estimation function in the integral expression of stem volume (Eq. 117) can equivalently be written as $A(h) = \pi D_A(h)^2/4$, where $D_A: [0, H] \rightarrow [0, \infty)$ is a continuous and bounded function expressing the true area diameter (i.e., the diameter that yields the true cross-section area when substituted in the circle area formula) at height $h$. The true volume of a stem is thus given by

$$V = \frac{\pi}{4} \int_0^H D_A(h)^2 dh,$$

and $D_A(\cdot)$ is referred to as the true stem curve.

A plethora of methods have been suggested to predict or estimate $D_A(\cdot)$ (see Sterba 1980, and, for instance, Laasasenaho 1982 and Lappi 1986). In this study, we consider cubic spline interpolation between diameters measured at several non-random heights in a stem: the stem curve is composed of piecewise defined 4th order (3rd degree) polynomials of height, one polynomial for each interval between two adjacent measurement heights; each polynomial is required to pass through the endpoints of its domain (the observed (height, diameter)-points), and the smoothness of the whole interpolant is ensured by requiring its first and second derivatives to be continuous over the whole domain [0, H], that is, also in every point of junction of the polynomials. The number of diameter measurements permitting, this approach has quite commonly been used in Finland to estimate stem volume for research purposes (see e.g. Laasasenaho 1982, Lappi 1986, Ojansuu 1993, and Mäkinen et al. 2002).

We are now to examine how diameter selection within cross-sections at predetermined heights affects volume estimation by cubic-spline-interpolated stem curves. As with the family of general volume estimators in Section 5.2, the members of which these estimators patently are, we adopt the within-tree view: we regard the stem volume as a fixed property
of a tree and randomness in volume estimation as arising from diameter selection within the cross-sections at the non-random observation heights.

We omit discussing the potential difficulties in the practical use of interpolating cubic splines — how to determine the initial values optimally and how to avoid non-monotonicity or oscillation, for example, which aspirations manifest themselves in the empirical rules about the required number and distribution (along the vertical axis) of the measurement heights (see Lahtinen and Laasasenaho 1979, and Lahtinen 1988) — but simply assume that we have at our disposal an adequate set of error-free diameter measurements from an adequate set of fixed heights.

Let us denote by $S_3(h; D, H)$ the interpolating cubic spline based on a vector of diameters $D=(D(h(1)), D(h(2)), ..., D(h(m)))$ taken (by some selection method) at predetermined heights $H=(h(1), h(2), ..., h(m))$. Presumably the best stem volume estimate is given by the stem curve obtained from the observed true area diameters $D_A=(D_A(h(1)), D_A(h(2)), ..., D_A(h(m)))$:

$$\hat{V}_s = \frac{\pi}{4} \int_0^H S_3(h; D_A, H)^2 dh .$$

(125)

The estimation error in $\hat{V}_s$ now results from the interpolation procedure, that is, from the deviation of the estimated stem curve $S_3(\cdot ; D_A, H)$ from the true stem curve $D_A(\cdot)$ in other points than those included in $D_A$.

In practice, not knowing the cross-section areas at the observation heights, we have to do with other diameters than the true area ones; let us denote a vector of such diameters by $D(\theta)=(D(\theta(1)), D(\theta(2)), ..., D(\theta(m)))$, $\theta$ referring to the diameter selection method.

If the selection method involves randomness, $D(\theta)$ is a random vector and the interpolated stem curve $S_3[h; D(\theta), H]$ becomes a random function. Naturally, the volume estimator

$$\hat{V}_s(\theta) = \frac{\pi}{4} \int_0^H S_3[h; D(\theta), H]^2 dh$$

(126)

is then also a random variable. Following the theory presented in Section 5.2 for the family of general volume estimators, we could now attempt to estimate the within-tree expectation and variance of this estimator by integrating the mean and covariance functions $\mu_A(h; \theta)$ and $\gamma_A(h, k; \theta)$ of the area estimation process $\{A(h; \theta), h \in [0, H]\} = \{\pi S_3[h; D(\theta), H]^2/4, h \in [0, H]\}$ in a stem. However, the functions appear to be difficult to derive analytically, as the coefficients determining the values of $S_3[h; D(\theta), H]$ between any two elements of $H$ (the observation heights) are nonlinear functions of the elements of $D(\theta)$. Thus, instead of pursuing the mean and covariance functions of $\{\pi S_3[h; D(\theta), H]^2/4, h \in [0, H]\}$, we consider estimating the within-tree expectation and variance of $\hat{V}_s(\theta)$ by means of a number of repeated estimations (realisations of the volume estimator).

In principle, to obtain a realisation $\hat{V}_s(\theta_i)$ of the volume estimator, we draw a vector of diameter directions $\theta_i=(\theta_i(h(1)), \theta_i(h(2)), ..., \theta_i(h(m)))$ from the joint distribution of the diameter directions at the fixed observation heights (the distribution is specific to the diameter selection method), measure the corresponding diameters $D(\theta_i)=(D(\theta_i(h(1)), h(1)), D(\theta_i(h(2)), h(2)), ..., D(\theta_i(h(m)), h(m)))$ at different heights, interpolate between the diameters, and integrate the square of the resulting stem curve realisation $S_3[h; D(\theta_i), H]$.

If the diameter selection method involves such dependence between the diameters that the measurement direction chosen at one height determines the measurement directions at the other heights, that is, if all the elements of the direction vector $\theta_i$ are transformations of the direction $\theta_i(h_k)$ associated to one observation height $h_k \in H$, the multidimensional direction distribution reduces into a one-dimensional one. This is the case, for example,
when the diameters at different heights are measured in the same direction decided upon at breast height (cf. the dependent selection in Section 5.1). The expectation and variance of the volume estimator are then obtained by numerical integration over the one-dimensional direction distribution: we select values for \( \theta_i(h_k) \), \( i=1, \ldots, n \), equidistantly within \([0, \pi)\), estimate from the direction distribution a weight \( w_i \) associated to each \( \theta_i(h_k) \) as the probability of selecting a direction not farther than half the equi-interval from \( \theta_i(h_k) \), derive from each \( \theta_i(h_k) \) the direction vector \( \theta_i \) and determine the corresponding diameter vector \( D(\theta_i) \), compute the volume estimator realisation \( \hat{V}_S(\theta_i) \) associated to each \( D(\theta_i) \), and then estimate the within-tree expectation and variance of the volume estimator as the weighted mean and variance of the \( n \) realisations:

\[
E_w[\hat{V}_S(\theta)] = \frac{1}{n} \sum_{i=1}^{n} w_i \hat{V}_S(\theta_i) = \frac{1}{n} \sum_{i=1}^{n} w_i \left( \frac{\pi}{4} \int_{0}^{H} S_3[h; D(\theta_i), H]^2 dh \right),
\]

and

\[
\text{Var}_w[\hat{V}_S(\theta)] = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{V}_S(\theta_i) - E_w[\hat{V}_S(\theta)] \right)^2.
\]

Naturally, if the one-dimensional direction distribution is uniform, all the weights \( w_i \) are equal and the expectation and variance can be estimated with simple (non-weighted) mean and variance.

If the diameter selection method involves no simplifying dependence between the diameters at the different heights, we resort to Monte Carlo integration, where the direction vectors \( \theta_i \), \( i=1, \ldots, n \), are sampled independently from the multidimensional direction distribution. In this case, the within-tree expectation and variance of the volume estimator are then estimated simply with the mean and variance of the \( n \) realisations (Robert and Casella 2000):

\[
E_w[\hat{V}_S(\theta)] = \frac{1}{n} \sum_{i=1}^{n} \hat{V}_S(\theta_i),
\]

and

\[
\text{Var}_w[\hat{V}_S(\theta)] = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{V}_S(\theta_i) - E_w[\hat{V}_S(\theta)] \right)^2.
\]

Interestingly, there is a straightforward connection between these estimates (Eqs. 127–130) and the estimates obtained by integrating the mean and covariance functions \( \mu_A(h; \theta) \) and \( \gamma_A(h, k; \theta) \) of the area estimation process \( \{\hat{A}(h; \theta), \ h \in [0, H]\} = \{\pi S_3[h; D(\theta), H]^2/4, \ h \in [0, H]\} \) (Eqs. 119 and 121), namely, the above estimates imply certain approximations for \( \mu_A(h; \theta) \) and \( \gamma_A(h, k; \theta) \): Estimating the expectation of \( \hat{V}_S(\theta) \) as the (weighted) mean of the volume estimates computed from a sample of diameter vectors (Eqs. 127 and 129) corresponds to approximating \( \mu_A(h; \theta) \) at each \( h \) by the (weighted) mean of a sample of area estimation function realisations (stem curve realisations) \( \hat{A}(h; \theta_i) = \pi S_3[h; D(\theta_i), H]^2/4, \ i=1, \ldots, n \).
Let \( \hat{E}_e[\hat{V}_S(\theta)] \) be the estimate of the volume of the stem section at a given angle \( \theta \) based on the diameter selection method.

\[
\hat{E}_e[\hat{V}_S(\theta)] = \frac{\sum_{i=1}^n w_i \hat{V}_S(\theta_i)}{\sum_{i=1}^n w_i} \\
= \frac{\sum_{i=1}^n w_i \left[ \int_0^H \hat{A}(h; \theta_i) dh \right]}{\sum_{i=1}^n w_i} \\
= \frac{\int_0^H \left[ \sum_{i=1}^n w_i \hat{A}(h; \theta_i) \right] dh}{\sum_{i=1}^n w_i} \\
= \int_0^H \hat{\mu}_A(h) dh \\
= \int_0^H \mu_A(h) dh \\
= E_e[\hat{V}_G(\theta)].
\]

Similarly, estimating the variance of \( \hat{V}_S(\theta) \) as the (weighted) variance of the volume estimates computed from a sample of diameter vectors (Eqs. 128 and 130) corresponds to approximating \( \gamma_A(h, k; \theta) \) by the (weighted) covariance computed at each \( (h, k) \) from a sample of area estimation function realisations (stem curve realisations):

\[
\text{Var}_e[\hat{V}_S(\theta)] = \frac{\sum_{i=1}^n w_i \left[ \hat{V}_S(\theta_i) - \hat{E}_e[\hat{V}_S(\theta)] \right]^2}{\sum_{i=1}^n w_i} \\
= \frac{\sum_{i=1}^n w_i \left[ \int_0^H \hat{A}(h; \theta_i) dh - \int_0^H \hat{\mu}_A(h) dh \right]^2}{\sum_{i=1}^n w_i} \\
= \frac{\sum_{i=1}^n w_i \left[ \int_0^H \left[ \hat{A}(h; \theta_i) - \hat{\mu}_A(h) \right] \int_0^H \left[ \hat{A}(k; \theta_i) - \hat{\mu}_A(k) \right] dh dk \right]}{\sum_{i=1}^n w_i} \\
= \int_0^H \int_0^H \gamma_A(h, k) dh dk \\
= \int_0^H \gamma_A(h, k) dh dk \\
= \text{Var}_e[\hat{V}_G(\theta)].
\]

In other words, at each point \( (h, k) \), the definite integrals in \( \mu_A(h; \theta) = E_e[\hat{A}(h; \theta)] \) and \( \gamma_A(h, k; \theta) = \text{Cov}_e[\hat{A}(h; \theta), \hat{A}(k; \theta)] \) taken over the multidimensional direction distribution (the joint distribution of the diameter directions at the fixed observation heights) are approximated with sums, sums of the squares or sums of the products of \( n \) area estimates \( \hat{A}(h; \theta_i) = \pi S_j[h; D(\theta_i), H]^{3/4} \) based on \( n \) direction vector realisations \( \theta_i \) drawn from the distribution.
In Section 5.2 we discovered that as the elements of a stochastic area estimation process become uncorrelated, the within-tree variance of the corresponding generalised volume estimator becomes zero; moreover, we noted that such a discontinuous process is an abstraction that cannot really be generated in a tree stem context (by e.g. independent diameter selection at a finite number of observation heights). Evidently, cubic spline interpolation between uncorrelated (independently selected) random diameters at the observation heights does not result in a volume estimator with zero variance, often quite the contrary: the area estimators at the observation heights are of course mutually independent, but the interpolated curve imposes continuity and induces short-distance dependence (non-zero covariance between heights close to each other) for the area estimators at the heights in between (via spline coefficients involving common diameters), thus making the integral of the resulting covariance function deviate from zero.

As with the volume equation in Section 5.1, the volume estimation error $V - \hat{V}_S(\theta)$ can be divided into the two distinct additive components — the error $V - \hat{V}_S = \varepsilon_S$ inherent in the best estimator $\hat{V}_S$ and the error $\hat{V}_S - \bar{V}_S(\theta)$ attributable to diameter selection by method $\theta$ — the latter of which further consists of two components — the within-tree bias $\bar{V}_S - E_0[\bar{V}_S(\theta)]$ with respect to the best estimator and a random sampling error $\nu_S(\theta)$:

$$V - \hat{V}_S(\theta) = (V - \hat{V}_S) + [\bar{V}_S - \hat{V}_S(\theta)]$$

$$= \varepsilon_S + \left\{ \bar{V}_S - \left\{ E_0[\bar{V}_S(\theta)] + \nu_S(\theta) \right\} \right\}$$

(133)

At the within-tree level only $\nu_S(\theta)$ is a random variable (with zero expectation and variance $\text{Var}_0[\nu_S(\theta)] = \text{Var}_0[\bar{V}_S(\theta)]$), whereas at the population level all the estimation error components are random variables. The interpretation of the error $\varepsilon_S$ contained in the best non-parametric estimator $\bar{V}_S$, however, does not straightforwardly parallel that of the model error term $\varepsilon_L$ of the parametric volume equation: In our “model” $V = \bar{V}_S + \varepsilon_S$, we have incorporated no distributional assumptions on $\varepsilon_S$ that would then have governed the estimation of $\bar{V}_S$. In particular, the population expectation of $\varepsilon_S$ may well deviate from zero, whereupon the estimation bias $E[V - \bar{V}_S(\theta)]$ at the population level may involve not only the between-trees expectation $E\{\bar{V}_S - E_0[\bar{V}_S(\theta)]\}$ of the within-tree bias due to diameter selection but also the population expectation $E(\varepsilon_S)$.
6 Material

The material of the study consists of a small subset of the sample trees felled and measured in 1991 for the nationwide tree quality investigation (VAPU) of the Finnish Forest Research Institute. In the investigation, the forest stand population of Finland was divided into the strata of Scots pine, Norway spruce and silver/pubescent birch dominated stands, in each of which the following four-stage sampling was then performed: First, a subset of 20 clusters were systematically chosen among the clusters of plots placed systematically all over Finland in the 7th National Forest Inventory (for the sampling design of the inventory, see Kuusela and Salminen 1980 and Kuusela et al. 1986). Second, two plots among the total of 28 in each cluster were systematically selected for tree sampling. Third, 20 tally trees were picked in each selected plot (first the four trees growing closest to the plot centre, and then the 16 dominant or codominant trees next closest to the centre); the trees had to fulfill certain selection criteria (living trees of species Scots pine, Norway spruce or silver/pubescent birch with no visible biotic or abiotic defects and with breast height diameter not less than 7 cm). Fourth, six sample trees to be felled in each plot were randomly chosen among the 20 tally trees (three trees among all the tally trees and three trees among the tally trees of dominant species and dominant crown class). Among the sample trees taken from the stratum of the Scots pine dominated stands, a judgement sample was then drawn for this study, resulting in a total of 81 trees from 16 plots in 11 clusters (Fig. 22, Table 10).

The geographical distribution of the selected trees was somewhat uneven in both the north–south and east–west directions, in which some noteworthy variation in stem form is known to occur in Finland (e.g. Lappi 1986): in Southern Finland (Fig. 22) 26 stems were measured on six plots, while in Northern Finland 55 stems were taken on ten plots; further, east of the centre meridian 27° of the Finnish national uniform co-ordinate system, only 14 stems were measured on three plots, whereas west of it 67 stems were taken on 13 plots.

All the 16 plots were located in naturally regenerated one-storey stands; while ten of these were pure Scots pine stands, four contained some Norwegian spruce mixture and two silver/pubescent birch mixture. A majority of the stands grew on mineral soil (Table 10); of
the seven stands located in peatland, five were drained — in two of these the drainage had been completed so recently that the drainage effect could not yet be discovered in ground vegetation or in trees, whereas in the other three the drainage effect was clearly visible although the ground vegetation was still characterised by the original peatland type. The sites were mainly of the two fertility classes — medium or quite poor — that are the most typical of Scots pine stands in Finland; the three deviating stands grew on rich sites (Table 10). Except for two advanced stands, where one or two thinnings had been carried out, the stands were young stands or advanced seedling stands, where no silvicultural treatments had been applied or cuttings performed since the regeneration (Table 10).

Naturally, given the stage of development of the stands, pulpwood-size trees dominated in the size distribution of the trees (Fig. 23): a total of 72 trees fell into that category (breast height diameter $\leq 17$ cm, height $\leq 13$ m), whereas only eight trees could be considered sawlog-size (breast height diameter $\geq 22$ cm, height $\geq 20$ m). Further, one tree (from the plot 304 in the northernmost cluster) appeared to be a slight anomaly with its breast height diameter of 20 cm and height of 13 m; however, this tree was included only in the examination of cross-section shape and cross-section area estimation (and not in the examination of stem volume estimation or Bitterlich sampling), as not more than two discs were measured from it.

Of the large number of characteristics (on stem dimensions, stem quality, growth, crown structure and biomass; Valtakunnallisen puututkimuksen (VAPU) ja kasvuunvaihtelututkimuksen maastotyöohjeet 1991) measured in the field on each felled sample tree, only the following were employed in this study: plot radius direction with respect to N–S direction

### Table 10. Some site and stand characteristics of the plots in which the 81 investigated trees were felled

<table>
<thead>
<tr>
<th>Plot</th>
<th>Cluster</th>
<th>Ground class$^1$</th>
<th>Site fertility$^2$</th>
<th>Stand development stage$^3$</th>
<th>G (m$^2$/ha)</th>
<th>$H_{\text{dom}}$ (m)</th>
<th>$D_{\text{gM}}$ (cm)</th>
<th>Number of felled trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>301</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>6.0</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>303</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>7.0</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>304</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>9.5</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>305</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5.5</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>306</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4.3</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>307</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>14</td>
<td>12.0</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>308</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4.8</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>310</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>9.0</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>311</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>13</td>
<td>9.5</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>312</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>8.5</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>313</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>9.0</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>316</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>6.6</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>317</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>12</td>
<td>9.0</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>318</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>24</td>
<td>24.0</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>321</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>22</td>
<td>22.0</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>322</td>
<td>11</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>9.5</td>
<td>12</td>
<td>5</td>
</tr>
</tbody>
</table>

$^1$ 1 mineral soil
2 spruce/hardwoods peatland
3 pine peatland
4 open bog

$^2$ 1 very rich
2 rich
3 medium
4 quite poor
5 poor
6 very poor

$^3$ 1 open regeneration site
2 young seedling stand
3 advanced seedling stand
4 young (thinning) stand
5 advanced (thinning) stand
6 mature stand
7 shelterwood stand
8 seedtree stand
at breast height (i.e., compass bearing of the tree taken from the plot centre to the assumed pith of the tree at breast height), tree height determined with respect to the ground level, and stump height ditto.

In each of the 81 stems, discs of thickness of 3 cm were sawn at two fixed — 1.3 m and 6 m — and eight relative — 1%, 2.5%, 7.5%, 15%, 30%, 50%, 70% and 85% — heights in the way that the lower surface of the disc was located at the cutting height determined from the ground level; were there branches at the proposed height, the cutting height was shifted to the nearest location where a disc free from branches could be obtained, and this new height was recorded. On each disc, a mark indicating the plot radius direction was painted. A complete set of ten discs could not, even in theory, be obtained in every stem: 16 trees were shorter than 6 m, and no disc at the height of 6 m hence existed in them; in two trees, a relative height (15% or 30%) also coincided with a fixed one (1.3 m). Besides, a number

---

**Table 11.** Distribution of the 81 stems of this study according to the number of discs obtained in the stem.

<table>
<thead>
<tr>
<th>Number of discs obtained in the stem</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stems</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>25</td>
<td>30</td>
<td>21</td>
</tr>
</tbody>
</table>

**Table 12.** Number of discs obtained at each observation height in the 81 stems of this study (cf. the diagonal of the Table 13).

<table>
<thead>
<tr>
<th>Height</th>
<th>1%</th>
<th>2.5%</th>
<th>7.5%</th>
<th>15%</th>
<th>30%</th>
<th>50%</th>
<th>70%</th>
<th>85%</th>
<th>1.3 m</th>
<th>6 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of discs</td>
<td>28</td>
<td>75</td>
<td>77</td>
<td>79</td>
<td>80</td>
<td>79</td>
<td>81</td>
<td>81</td>
<td>80</td>
<td>51</td>
</tr>
</tbody>
</table>

---

**Fig. 23.** Frequency distributions of breast height diameter (taken perpendicular to plot radius direction; A) and tree height (B) in the data of the 81 trees investigated in this study.
of discs, especially those taken at 1% height, were broken when sawing or transporting to the laboratory or had to be discarded due to inadequate identification information. In all, a total of 709 cross-sections representing 711 observation heights were included in the study; see Tables 11, 12 and 13 for the summaries of the distribution of the discs between the stems and Fig. 24 for the disc size distributions at the relative observation heights.

The discs were photographed, and the photographs were turned into vectorised digital images, from which the characteristics of the cross-sections were then computed (cf. e.g.
Thies and Harvey 1979, Biging and Wensel 1984, Drake et al. 1988, Jonsson 1992, Saint-André 1998, and Saint-André and Leban 2000). This approach enabled us to exploit shape information to as large an extent as we ever desired and helped us eliminate measurement errors, which could badly confound the possibly subtle influences of non-circularity. Before photographing, the discs were debarked up to the cambium layer in order to eliminate the problems that bark irregularities could cause in the computational extraction of the characteristics from the images. Thus, all the characteristics computed from the images are under-bark.

The discs were photographed in a laboratory with fixed lighting by using an ordinary 25-mm film camera with a 50-mm objective and panchromatic black-and-white film of the sensitivity of 100 ASA; in addition to the flashlight attached to the camera, three extra flashlights with “slave switches” (i.e., flashlights that react, with a marginal delay, to the actual camera flashlight) were employed to increase the contrast between the disc and the background (Fig. 25). A ruler showing the scale and a slip of paper with a mark indicating the plot radius direction were placed on the same plane as the lower surface of the disc (Fig. 26 A). Two fixed distances between the camera lens and the plane under the disc were applied — 60 cm for the discs with the diameter up to 15 cm, and 100 cm for the discs larger than that. In order to take into account the possible variation in disc thickness, which would affect the distance between the camera lens and the upper surface of the disc (Fig. 25) later needed in scale determination, thickness was measured (in mm) in four points at the edge of each disc (at regular rotation angle intervals of 90° starting from the plot radius direction).

The film negatives were developed and printed to photographs (Fig. 26 A) with some extra contrasting between the disc edge and the light background (i.e., with some artificial increase in tone value differences). The photographs were transformed into grey-scale digital raster images with an optical scanner of the resolution of 300 dpi; it was expressly owing to the low resolution of the apparatus available that the scanning was performed on the prints instead of the negatives. In each raster image, the length of a 10-cm piece of the ruler beside the disc (Fig. 26 A) was measured (with pixel edge length as the unit) for scale computation, after which the image was manually cropped to contain only the disc and the mark indicating the plot radius direction (Fig. 26 A). In order to facilitate the computational feature extraction, the grey-scale pixels were then classified into two categories — the “information” (black) and the “background” (white) (Fig. 26 B): in each image, the threshold for the intensity values of the information class pixels was set by visual assess-
ment with the aim of locating the edge of the cross-section correctly and forming connected
sets of pixels; due to the high and homogeneous contrasts in the photographs, very little
adjustment in the threshold value was needed between the images. Unfortunately, the pith
locations were not recorded in the procedure.

From each classified raster image, the boundary pixels of the cross-section and of the
mark indicating the plot radius direction were extracted, and the image consisting of these
pixels was vectorised (i.e., transformed into a set of co-ordinates of discrete contour points)
with smoothing: at each change in direction, the two midpoints of the outer edge of the
corner pixel were taken, and the line segment connecting the points was used to outline
the corner in the resultant vector image (Fig. 27, Fig. 26 C). Further, the vector image of
the convex closure of the cross-section was computed as the convex hull of the discrete
contour points of the cross-section.

From the vector image of a cross-section, the following characteristics were determined:
centres of gravity of the cross-section (as a substitute for the pith) and the mark indicating
the plot radius direction; area of the cross-section; 360 radii from the centre of gravity of
the cross-section at regular rotation angle intervals of 1° starting from the N–S direction;
and 180 breadths ditto. (To check the effect of the smoothing in vectorisation, the centre
of gravity and the area of the cross-section were also computed from the non-vectorised
classified raster image: the differences were negligible in all the cross-sections.) From the
vector image of the convex closure, in turn, the following properties were computed: area
of the convex closure; convex perimeter (girth); 360 radii from the centre of gravity of the cross-section at regular rotation angle intervals of 1° starting from the N–S direction; 180 breadths ditto; 180 diameters ditto; and, with viewing angles 1.146°, 1.621°, 2.292° and 3.624° (corresponding in circular cross-sections to the basal area factors 1, 2, 4 and 10 m²/ha), the contour of the inclusion region in a discretised form consisting of 3600 points; the areas of the 360 sectors of the inclusion region, of angular width 1° and with the first sector midline in the N–S direction; and the total area of the inclusion region.

The boundary pixel identification and extraction, the vectorisation and the computation of centres of gravity, convex perimeter, and true and convex area were performed with the GRASS 4.1 and Mathematica 2.2 software (GRASS User’s Reference Manual 1993, Guide to Standard Mathematica Packages 1993). For the computation of radii, breadths, diameters, and inclusion area, Fortran code was written, the basic ideas of which are presented in Appendices G and H.

In the computation of the characteristics from the images, the pixel edge length was used as a natural measurement unit. While this unit was fully appropriate for the examination of the relative quantities related to the shape of cross-sections and the errors of various area estimation methods, the true scale was needed for the investigation of volume estimation errors. The scale was computed for each cross-section by means of the basic lens formula in optics, the four thickness measurements taken on each disc, and the length measurement of a 10-cm piece of the ruler in the unclassified raster image. See Appendix I for a detailed account of the scale computation.
7 Methods

With the 709 cross-sections from the 81 Scots pine stems in the data, we empirically investigated (i) the variation in cross-section shape, (ii) the effect of the within-cross-section variation in diameter on cross-section area estimates given by the estimators discussed in Sections 3.3 and 4.2.2, (iii) the influence of the within-cross-section variation in diameter on stem volume estimates given by Laassanenaho volume equation, cubic-spline-interpolated stem curve and generalised volume estimator considered in Chapter 5, and (iv) the bias inflicted by non-circular cross-section shape on stand total estimators in Bitterlich sampling as discussed in Section 4.1.2.

As the stems in the data were few in number and did not make up any actual probability sample (the sampling design was not probabilistic in the last stage of subsampling among the trees felled for the nationwide tree quality investigation; see Chapter 6), the data did not really permit confirmatory analysis (hypothesis testing, model building) concerning any meaningful Scots pine population. Instead, we contented ourselves with explorative analysis, considering the empirical distributions of relevant within-cross-section and within-tree characteristics in our data. To enable meaningful, size-independent comparisons between cross-sections or stems, the characteristics were usually turned relative (expressed as percentages). The hierarchical structure of the data resulting from the multi-stage sampling in the data collection (cross-sections being interdependent within a stem, stems interdependent within a plot, plots interdependent within a cluster) was largely disregarded in these considerations: the distributions of the within-cross-section characteristics were examined in the subsets formed by height classes as well as in the set of all the cross-sections, whereas the distributions of the within-tree characteristics were studied in the set of all the stems.

As already pointed out in Chapter 6, the discs were debarked before photographing, and, hence, all our empirical examinations pertain to cross-sections and stems without bark.

7.1 Cross-Section Shape

Variation in cross-section shape was examined in three aspects: (i) variation in shape of convex closure, (ii) amount and directional location of non-convexity, and (iii) variation in true shape. The interest in the shape of convex closure was motivated by our interest in area and volume estimation based on measurements made from the outside of a tree: with caliper, tape and height meter, no non-convexity can be registered but only the convex closure of a tree cross-section is observed.

The scalar and functional characteristics (shape indices) computed for each cross-section to analyse shape are summarised in Tables 14 and 15, respectively.

Table 14. Scalar shape indices computed for each cross-section.

<table>
<thead>
<tr>
<th>Index</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{min}/D_{max}$</td>
<td>Ratio between minimum and maximum diameters</td>
</tr>
<tr>
<td>$CVD$</td>
<td>Diameter coefficient of variation</td>
</tr>
<tr>
<td>$\rho_D(\pi/2)$</td>
<td>Correlation between perpendicular diameters</td>
</tr>
<tr>
<td>$b_e/b_a$</td>
<td>Girth-area ellipse ratio</td>
</tr>
<tr>
<td>$</td>
<td>\theta_{D_{min}}-\theta_{D_{max}}</td>
</tr>
<tr>
<td>$(A_C-A)/A_C$</td>
<td>Relative convex deficit</td>
</tr>
<tr>
<td>$(\hat{A}_0-A_C)/A_C$</td>
<td>Relative isoperimetric deficit; relative bias of estimator $\hat{A}_0$ based on girth diameter $C/\pi=\mu_D$ with respect to convex cross-section area</td>
</tr>
</tbody>
</table>
Mainly for comparison with previous empirical studies (especially that by Matérn in 1990), we investigated the shapes of the convex closures of the cross-sections by means of the following three simple diameter-based indices (which are invariant of scale, translations and rotations, i.e., independent of the size as well as the position and the orientation of the cross-section in the chosen rectangular co-ordinate system): the ratio \( D_{\text{min}}/D_{\text{max}} \) between the minimum and the maximum diameters, with which the deviation of cross-section shape from a circle has mostly been examined in previous empirical studies; the coefficient of variation of diameter \( CVD = \sigma_D/\mu_D \) (the ratio of diameter standard deviation to diameter mean), which better than the extreme diameters describe the magnitude of diameter variation in a cross-section and help detect the possible orbiforms in the data; and the correlation \( \rho_D(\pi/2) \) between perpendicular diameters (the diameter autocorrelation at angle \( \pi/2 \)), which gives information on possible ellipticity or square-shapedness of cross-sections (see Chapter 2).

These indices were estimated from the 180 diameters computed systematically (at regular rotation angle intervals of 1° and starting from the N–S direction) in each cross-section (see Chapter 6). \( D_{\text{min}} \) and \( D_{\text{max}} \) were determined as the minimum and the maximum of these diameters. In \( CVD \), the exact value for the mean diameter \( \mu_D \) was attained by dividing the convex perimeter \( C \) by \( \pi \), whereas the standard deviation \( \sigma_D \) was estimated as the square root of the variance of the 180 diameters:

\[
\hat{\sigma}_D = \frac{1}{180} \sum_{j=0}^{179} [D(j \cdot 1°) - \mu_D]^2
\]

(for unbiasedness, the denominator \( n=180 \) instead of the usual \( n–1=179 \) was used in this sample variance, as the population mean was not estimated by a sample mean but known exactly). Further, the correlation between perpendicular diameters was estimated as

\[
\hat{\rho}_D\left(\frac{\pi}{2}\right) = \frac{1}{\hat{\sigma}_D^2} \frac{1}{90} \sum_{j=0}^{89} \{D(j \cdot 1°) - \mu_D\} \{D((j+90)\cdot 1°) - \mu_D\}
\]

An ellipse has often been assumed to approximate cross-section shape better than a circle, with the ratio \( D_{\text{min}}/D_{\text{max}} \) as a simple estimate for the ellipse axis ratio. We tried also another axis ratio estimate, that of an ellipse with the perimeter and area equal to the convex
perimeter $C$ and convex area $A_C$ of the cross-section: since $C=\pi[3(a_e+b_e)/2-(a_e b_e)^{1/2}]$ and $A_C=\pi a_e b_e$ for an ellipse with $a_e$ and $b_e$ as the lengths of the long and the short semi-axes, the axis ratio estimate, here referred to as the *girth-area ellipse ratio*, became

$$\frac{b_e}{a_e} = \frac{A_C}{\pi a_e^2}, \quad (136)$$

where

$$a_e = \frac{1}{3} \left( \frac{C}{\pi} + \sqrt{\frac{A_C}{\pi}} \right) + \frac{1}{9} \left( \frac{C}{\pi} + \sqrt{\frac{A_C}{\pi}} \right)^2 - \frac{A_C}{\pi} \quad (137)$$

(Stoyan and Stoyan 1994); it is not difficult to see that for a circle $b_e/a_e=1$. (Note that there are of course many other alternatives for estimating the axis ratio of the approximating ellipse; these include e.g. the ratio of $D_{\text{max}}$ and the diameter perpendicular to it, $A_C/(\pi D_{\text{max}}^2/4)$ derived from $D_{\text{max}}$ and $A_C$ via the ellipse area formula, the ratio of the semi-axis lengths estimated by a least squares fitting of the ellipse contour equation to the discretely observed contour co-ordinates, and the ratio of the semi-axis lengths estimated by a least squares fitting of the ellipse radius autocovariance function to the sample autocovariances computed from the discretely observed radii; see e.g. Stoyan and Stoyan 1994, Skatter and Høibø 1998, Saint-André and Leban 2000).

Estimating the ellipse ratio with $D_{\text{min}}/D_{\text{max}}$ or $b_e/a_e$ disregards the requirement that the principal axes be at right angles to each other. To find out whether the extreme diameters would be perpendicular to each other in our cross-sections, we computed the absolute angle $|\theta_{D_{\text{min}}} - \theta_{D_{\text{max}}}| (\in [0, 90\degree])$ between $D_{\text{min}}$ and $D_{\text{max}}$. In addition, we examined the change in the orientation of the extreme diameters along with height in a stem by comparing the directions of $D_{\text{min}}$ and $D_{\text{max}}$ at each height to those at the lowest observed height in the stem.

The association between the size and shape of convex closure was examined by ordinary product moment correlations and accompanying scatterplots between the mean diameter $\mu_D$ and the shape indices. Also the interrelations of the indices were examined by pairwise correlations and scatterplots.

Since an ellipse has been much used as a model for cross-section shape, we finally compared the values of the shape indices in the cross-sections with their values in ellipses. This with keeping in mind that the indices impart by no means unique or unambiguous information on shape, but visually quite dissimilar shapes may assume very similar index values (for instance, think of a rhombus and an ellipse with the same $D_{\text{min}}/D_{\text{max}}$, or consider the shapes C and F with the same $CV_D$ in Fig. 5 in Chapter 2). Hence, the information given by the indices can only evidence against the ellipticity hypothesis but not for it (for an extensive discussion on some indices and their philosophy, see e.g. Exner 1987 and Stoyan and Stoyan 1994).

### Directional Variation in Diameter

Diameter variation with respect to direction within a cross-section was studied by means of the relative deviation $\{D(\cdot) - D_{\text{Ac}}\}/D_{\text{Ac}}$ of the discretely observed diameter function $D(j \cdot 1\degree)$, $j=0, \ldots, 179$, from the convex area diameter $D_{\text{Ac}}=2(A_C/\pi)^{1/2}$ yielding the convex area when substituted in the circle area formula. The choice of $D_{\text{Ac}}$, instead of, say, the mean diameter $\mu_D$, as the reference was motivated by the idea that error in area estimation might be equated
with error in diameter selection: besides diameter variation, the relative differences would then provide information on the preferable choice of diameter direction in area estimation.

The diameter direction was expressed with respect to the N–S direction or with respect to the direction of \( D_{\text{max}} \) and set to increase anticlockwise. With the N–S direction as the reference direction, the idea was to study whether large-scale exogenous factors (prevailing winds, amount of solar radiation etc.) could affect trees in a somewhat uniform way regardless of the very local growing conditions and thus beget any pattern in the diameter variation. With the direction of \( D_{\text{max}} \) (which was considered more stable than the other natural option, the direction of \( D_{\text{min}} \)) as the reference direction, in turn, the idea was to examine if the diameter variation would show an internal pattern reflecting a degree of similarity in reactions of trees to their varying local conditions.

The empirical distribution of the discretely observed functionals \( \frac{[D(\cdot) - D_{\text{Ac}}]}{D_{\text{Ac}}} \) in each height class was summed up simply with the mean, median and variance functions (Ramsay and Silverman 1997): these were estimated with the pointwise averages, medians and variances across the cross-sections in the height class (i.e., with the averages, medians and variances of the cross-sectionwise \( \frac{[D(j \cdot 1^\circ) - D_{\text{Ac}}]}{D_{\text{Ac}}} \) at every observed direction \( j \cdot 1^\circ \), \( j=0, 1, \ldots, 179 \)).

### 7.1.2 Non-Convexity

#### Convex Deficit

As a measure of the amount of non-convexity, the relative convex deficit \( \frac{[A_{\text{C}} - A]}{A_{\text{C}}} \), where \( A \) is the true area and \( A_{\text{C}} \) is the convex area of the cross-section, was computed in each cross-section. Sometimes this characteristic has also been regarded as a shape index (for instance, it is referred to as “convexity ratio” in Stoyan and Stoyan 1994 and as “roundness” in Glasbey and Horgan 1995). Its relation to the scalar shape indices discussed above on one hand, and to the size of a cross-section (the mean diameter \( \mu_D \)) on the other hand, were examined by pairwise correlations and accompanying scatterplots.

#### Directional Variation in Breadth

In each cross-section, the occurrence of non-convexity with respect to direction was investigated by means of the relative difference \( \frac{[B_{\text{C}}(\cdot) - B(\cdot)]}{B} \) between the discretely observed breadth function of the convex closure \( B_{\text{C}}(j \cdot 1^\circ), j=0, \ldots, 179 \), and that of the cross-section \( B(j \cdot 1^\circ), j=0, \ldots, 179 \), with the mean breadth \( B \) of the cross-section as the denominator. The breadth functions were employed instead of radius functions in order to “encapsulate” the directional information, as in a standard measurement situation we would be interested in the non-convexity occurring in “both ends” of the diameter measurement. As in the examination of diameter variation, the direction was determined both with respect to the N–S direction and in respect of the direction of \( D_{\text{max}} \). Also similarly to the diameter variation case, the empirical distribution of the relative difference functions of individual cross-sections in each height class was summed up by the mean, median and variance functions computed pointwise over the cross-sections.
7.1.3 True Shape

The true shapes of cross-sections were studied by means of the discretely observed radius functions $R(j \cdot 1^\circ), j=0, ..., 359$; in other words, the (discretely, but more densely) observed contour of each cross-section was reduced to 360 equiangular points in polar coordinates. In shape analysis literature, such points are referred to as (pseudo)landmarks and the set of them depicting a figure as a landmark configuration (Stoyan and Stoyan 1994, Dryden and Mardia 1998). We denote the landmark configuration vector of cross-section $i$ by $R_i=(R_i(j \cdot 1^\circ))_{j=0, 1, ..., 359}$.

Prior to estimating the average shape and examining the shape variability in our set of cross-sections, we rendered the landmark configurations with different locations, orientations and sizes commensurate by turning each configuration $R_i$ into its centred pre-shape $R_i^*=(R_i^*(j \cdot 1^\circ))_{j=0, 1, ..., 359}$ (Dryden and Mardia 1998) with the Euclidian similarity transformations (translations, rotations and scaling): First, the configuration was centred, that is, the origin of the Cartesian co-ordinate system was set into the centre of gravity of the configuration; actually, this involved no action, as (not surprisingly considering the large number of landmarks in the configuration) the centre of gravity of the configuration was found to virtually coincide with that of the cross-section (computed from all the contour points of the vector image; see Chapter 6) relative to which the radius function had been defined. Second, the configuration was rotated in the way that either the radius to the north of the origin or the maximum radius $R_{\text{max}}$ coincided with the positive y-axis; the advantage of fixing the reference axis along with $R_{\text{max}}$ was that similar or congruent shapes would be considered equal after the similarity transformations. Third, the configuration was scaled to constant size (centroid size) by dividing the landmarks (radii) by their quadratic average in the cross-section:

$$R_i^*(j \cdot 1^\circ) = \frac{R_i(j \cdot 1^\circ)}{\overline{R}_{qi}},$$

where

$$\overline{R}_{qi} = \sqrt[360]{\frac{1}{360}} \sum_{j=0}^{359} R_i(j \cdot 1^\circ)^2;$$

in other words, the configuration of the cross-section $i$ expressed in polar co-ordinates was dilated by the factor $1/\overline{R}_{qi}$; as a result, in each cross-section, the square root of the sum of squared Euclidean distances from the dilated landmarks to the centre became constant $360^{1/2}$ (this is how the centroid size is defined; Dryden and Mardia 1998).

The deviation of each centred pre-shape configuration from the unit circle was investigated with the difference

$$R_i^*(j \cdot 1^\circ) - 1 = \frac{R_i(j \cdot 1^\circ) - \overline{R}_{qi}}{\overline{R}_{qi}},$$

where

$$j=0, ..., 359.$$ The empirical distribution of these pre-shape deviations (discretely observed relative radius error functions with respect to the quadratic mean of the radii) in each height class was summarised with the mean, median and variance functions computed pointwise over the cross-sections in the height class.
Mean Shape

In each height class, we estimated the mean configuration $\mu_{R^*} = (\mu_{R^*(j\cdot1^\circ)})_{j=0,1,\ldots,359}$ with the pointwise mean of the pre-shape configurations of the n cross-sections in the class:

$$\hat{\mu}_{R^*}(j\cdot1^\circ) = \frac{1}{n} \sum_{i=1}^{n} R^*_i(j\cdot1^\circ) \quad (j=0, \ldots, 359),$$

(141)

The deviation of this estimated mean configuration from the unit circle

$$\hat{\mu}_{R^*}(j\cdot1^\circ) - 1 = \frac{1}{n} \sum_{i=1}^{n} (R^*_i(j\cdot1^\circ) - \hat{R}_q i) \quad (j=0, \ldots, 359),$$

(142)

By estimating the mean shape in this way, we assumed that the pre-shape configurations $R^*$ in each height class follow the model

$$R^* = \mu_{R^*} + \varepsilon,$$

(143)

where $\varepsilon = (\varepsilon(j\cdot1^\circ))_{j=0,1,\ldots,359}$ is the vector of independent random errors with zero mean (cf. Dryden and Mardia 1998, p. 88). The model implies that each pre-shape configuration be generated from the common mean by just random disturbance (measurement error) and not with any translations, rotations or rescaling.

(As opposed to full generalised Procrustes analysis, where several landmark configurations are matched to their mean shape with translation, rotation and scaling, and where this unknown mean shape is estimated as the configuration that minimises the sum of the so-called full Procrustes distances between it and the observed configurations (Dryden and Mardia 1998, p. 87–92), we performed here a restricted version, where we reduced translation into mere superimposition of centres of gravity and rotation into alignment of N–S or $R_{max}$ directions.)

Shape Variability

In each height class, we explored shape variability around the estimated mean shape with a principal component analysis of the 360 residual variables $R^*_i(j\cdot1^\circ) - \hat{\mu}_{R^*}(j\cdot1^\circ)$, $j=0,1,\ldots,359$ (cf. Stoyan and Stoyan 1994, Dryden and Mardia 1998; for the principal component analysis in general, refer e.g. to Dillon and Goldstein 1984, Jolliffe 1986, or Jackson 1991). The analysis involved estimating the covariance matrix of the variables and then extracting its positive eigenvalues and the corresponding eigenvectors. The use of the covariance matrix, as opposed to the correlation matrix, was justified by the variation in the variables being of the same order of magnitude. The elements of the covariance matrix were estimated with the sample covariances of the variables over the n cross-sections at the particular height; the estimation was based on considerably different numbers of observations, as the number of cross-sections in the height classes varied from 28 to 81 (see Table 12 in Chapter 6). The n–1 positive eigenvalues of the estimated covariance matrix gave the variances $\text{Var}(\text{PC}(k))$ of the n–1 principal components $\text{PC}(k)$, $k=1,\ldots,n–1$, and the corresponding eigenvectors gave the coefficients $a_{kj}$ of the 360 variables in the n–1 principal components, $j=0,1,\ldots,359, k=1,\ldots,n–1$.

Each estimated principal component $\text{PC}(k) = \sum_j [R^*_i(j\cdot1^\circ) - \hat{\mu}_{R^*}(j\cdot1^\circ)]a_{kj}$, $k=1,\ldots,n–1$, was visualised in plane by drawing the following contours:
\[
\hat{\mu}_n(j \cdot 1^\circ) + c \cdot a_{(k)j} \sqrt{\text{Var}(PC(k))}
\]

(144)

for \(c = \pm 2, \pm 4, \pm 6\) and \(j = 0, 1, \ldots, 359\). (Usually, the values for the coefficient \(c\) are selected in the way that they would cover the full potential range of the standardised principal component scores \(c(k)i = PC(k)i/\sqrt{\text{Var}(PC(k))}\), \(k = 1, \ldots, n-1, i = 1, \ldots, n\), in the data: if the variables \(R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ)\), \(j = 0, 1, \ldots, 359\), follow a multivariate normal distribution, the standardised principal component scores approximately follow the normal distribution with mean 0 and variance 1, for which 99.7% of the probability mass is within \([-3, 3]\); accordingly, \(c = \pm 1, \pm 2, \pm 3\) would be an adequate range. In our data, however, the shape variability was found to be so small that the range of \(c\) had to be magnified in order to easily visualise the effect of each principal component. See Dryden and Mardia 1998, p. 48–49 and p. 97, for a more detailed explanation of the rationale behind this common way of visualisation.) The idea was to illustrate the magnitude of the variability around the mean shape in each principal component estimated from our set of cross-sections. Since the correlation of the \(j\)th variable with the \(k\)th principal component is known to be

\[
\text{Corr}\left[R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ), PC(k)\right] = a_{(k)j} \sqrt{\text{Var}(PC(k))}/\sqrt{\text{Var}^2[R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ)]},
\]

(145)

the visualisation contours can equivalently be written as

\[
\hat{\mu}_n(j \cdot 1^\circ) + c \cdot \sqrt{\text{Var}(R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ)) \cdot \text{Corr}(R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ), PC(k))}.
\]

Hence, another interpretation of the visualisation is that the contour portrays the correlations of the radius residual variables to the principal component, each correlation being weighted with \(c\) times the standard deviation of the pertinent radius residual variable.

In order to detect severely deviating cross-sections (outliers) and their effect on the shape analysis at each height, the standardised principal component scores \(c(k)i = PC(k)i/\sqrt{\text{Var}(PC(k))}\) of the \(n\) cross-sections \((i = 1, \ldots, n)\) for each principal component \((k = 1, \ldots, n-1)\) were examined; they were plotted against the within-cross-section means of the 360 pre-shape residuals \(\Sigma[R^*(j \cdot 1^\circ) - \hat{\mu}_R*(j \cdot 1^\circ)]/360\) (good indicators of within-cross-section radial variation and, thus, of non-circularity) as well as against the sizes of the cross-sections as indicated by the expected diameters \(\mu_D\).

7.2 Estimation of Cross-Section Area

The effect of within-cross-section diameter variation on area estimation was examined with the area estimators for which the theoretical properties had previously been established (see Sections 3.3 and 4.2.2). The estimators were of the form

\[
\hat{A} = \frac{\pi}{4} D(\cdot)^2,
\]

(147)

where \(D(\cdot)\) was

0. girth diameter \(C/\pi = \mu_D\) (mean diameter of over the uniform direction distribution) \((\hat{A}_0)\)
1. a random diameter with the uniform direction distribution within \([0, \pi]\) \((\hat{A}_1)\)

832
The comparison between the estimators actually meant a comparison between different ways of selecting diameter under the circularity assumption (estimators $\hat{A}_0$, $\hat{A}_1$, $\hat{A}_2$, $\hat{A}_4$, $\hat{A}_{1\xi}$, $\hat{A}_{2\xi}$, $\hat{A}_{4\xi}$, $\hat{A}_{1\xi90}$, $\hat{A}_{4\xi90}$, $\hat{A}_6$, $\hat{A}_8$, $\hat{A}_{10}$, $\hat{A}_{\text{min}}$, $\hat{A}_{\text{max}}$) and different ways of determining principal axes under the ellipticity assumption (estimators $\hat{A}_3$, $\hat{A}_5$, $\hat{A}_{3\xi}$, $\hat{A}_{5\xi}$, $\hat{A}_{5\xi90}$, $\hat{A}_7$, $\hat{A}_9$, $\hat{A}_{11}$, $\hat{A}_{\text{min}}$, and $\hat{A}_{\text{max}}$), were included in the consideration mainly to obtain the lower and upper bounds for the estimates.

The estimators were compared by means of their estimated relative within-cross-section bias, standard deviation (the square root of the variance) and root mean squared error (RMSE, the square root of the sum of the variance and the squared bias; cf. Eq. 15 in Section 3.1); these characteristics, computed in each cross-section, are summarised in Table 16. Whether the errors should be expressed relative to the true area or to the convex area is not a straightforward question: as many times pointed out, non-convexity cannot be discerned by caliper or tape and is thus a matter of indirect observation of cross-section shape rather than a question of diameter measurement error or diameter sampling error; hence, it may appear unjustifiable to include into the area estimation error the bias induced by non-convexity that is inherent in the diameter population and thus present in the population total however precisely all the members of the population were ever measured. We circumvented this problem by proportioning the estimation error both to the true area and to the convex area.

For the fixed estimators ($\hat{A}_0$, $\hat{A}_6$–$\hat{A}_{11}$, $\hat{A}_{\text{min}}$, and $\hat{A}_{\text{max}}$), the relative within-cross-section biases were directly given by the relative estimation errors $|\hat{A} - A|/A$ and $|\hat{A} - A_C|/A_C$. The within-cross-section variances were naturally zero, and the relative root mean squared error thus equalled the absolute value of the relative bias. $D_{\text{min}}$ and $D_{\text{max}}$ were determined as the minimum and the maximum of the 180 diameters computed systematically (at regular rota-
tion angle intervals of 1° and starting from the N–S direction) in each cross-section; also the diameters perpendicular to \( D_{\text{min}} \) and \( D_{\text{max}} \) were then taken among the 180 systematic diameters.

For the random estimators (\( \hat{A}_1, \ldots, \hat{A}_5, \hat{A}_1\theta, \hat{A}_4\theta, \hat{A}_5\theta \)), the relative biases \( \frac{[E(\hat{A}) - A]}{A} \) and \( \frac{[E(\hat{A}) - AC]}{AC} \), the relative standard deviations \( \frac{\text{Var}(\hat{A})^{1/2}}{A} \) and \( \frac{\text{Var}(\hat{A})^{1/2}}{AC} \), and relative approximate standard deviations \( \hat{\text{Var}}(\hat{A})^{1/2} \) were estimated by inserting the estimated diameter moments and product moments in the analytical expressions of the within-cross-section expectations \( E(\hat{A}) \), variances \( \text{Var}(\hat{A}) \) and approximate variances \( \hat{\text{Var}}(\hat{A}) \) (Eqs. 22–26, 30–34 and 35–37 in Section 3.3.2; Eqs. 94–98, 99–103 and 104–108 in Section 4.2.2). As for the approximate variances, the interest was in their deviation from the true variances: in practice, we would prefer to use the approximate expressions for their simplicity (they only involve diameter means \( \mu_D \), \( \mu_{D(\xi)} \) and \( \mu_{D(\xi+\pi/2)} \), variances \( \sigma_D^2 \), \( \sigma_{D(\xi)}^2 \) and \( \sigma_{D(\xi+\pi/2)}^2 \), and correlations \( \rho_D(\pi/2) \) and \( \rho_{D(\xi)(\pi/2)} \) between perpendicular diameters), if only they do not deviate too much from the true variances.

For the random estimators involving only diameters with the uniform direction distribution (\( \hat{A}_1, \ldots, \hat{A}_5 \)), the mean diameter \( \mu_D \) was obtained from the convex perimeter as \( C/\pi \), whereas the other moments and product moments were estimated as simple means from the 180 systematic diameters in each cross-section: \( \sigma_D^2 \) and \( \rho_D(\pi/2) \) were obtained with Eqs. 134 and 135, and the other (product) moments generally from

\[
\hat{E}\left[D(\theta)^{k}D(\theta+\pi/2)^{p}\right] = \frac{1}{180}\sum_{j=0}^{179}D(j\cdot1^{\circ})^{k}D((j+90)\cdot1^{\circ})^{p}.
\]

(148)

k=0, ..., 4, p=0, ..., 4. (Note that for \( j+90=180 \), \( D((j+90)\cdot1^{\circ})=D((j-90)\cdot1^{\circ}) \).

For the random estimators involving Bitterlich diameters (diameters parallel or perpendicular to plot radius in Bitterlich sampling; \( \hat{A}_{1\xi}, \hat{A}_{4\xi}, \hat{A}_{1\theta}, \hat{A}_{4\theta}, \hat{A}_{3\theta}, \hat{A}_{5\theta} \)), the non-uniform direction distributions were determined for each cross-section using (i) the inclusion region of the individual cross-section or (ii) the inclusion region of the breast height cross-section.

### Table 16. Characteristics estimated for each cross-section to examine the performance of different area estimators (combinations of the circle area formula and a diameter selection method).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( [E(\hat{A}_j) - A]/A, j=0-11, 1\xi-5\xi, 1\xi90, 4\xi90, 5\xi90 )</td>
<td>Relative within-cross-section bias of estimator ( \hat{A}_j ) (circle area formula and diameter selection method ( j )) with respect to true or convex cross-section area</td>
</tr>
<tr>
<td>( \text{Var}(\hat{A}_j)^{1/2}/A, j=1-5, 1\xi-5\xi, 1\xi90, 4\xi90, 5\xi90 )</td>
<td>Within-cross-section standard deviation of estimator ( \hat{A}_j ) relative to true or convex cross-section area</td>
</tr>
<tr>
<td>( \hat{\text{Var}}(\hat{A}_j)^{1/2}/A, j=1-5, 1\xi-5\xi, 1\xi90, 4\xi90, 5\xi90 )</td>
<td>Delta method approximation of within-cross-section standard deviation of estimator ( \hat{A}_j ) relative to true or convex cross-section area</td>
</tr>
<tr>
<td>( \text{RMSE}(\hat{A}_j/A), j=0-11, 1\xi-5\xi, 1\xi90, 4\xi90, 5\xi90 )</td>
<td>Relative within-cross-section root mean squared error of estimator ( \hat{A}_j ) with respect to true or convex cross-section area; ( \text{RMSE}(\hat{A}_j/A) = \left{ [E(\hat{A}_j) - A]^2/A^2 + \text{Var}(\hat{A}_j)/A^2 \right}^{1/2} )</td>
</tr>
</tbody>
</table>

For the random estimators (\( \hat{A}_1, \ldots, \hat{A}_5, \hat{A}_1\xi, \hat{A}_4\xi, \hat{A}_5\xi, \hat{A}_1\theta, \hat{A}_4\theta, \hat{A}_5\theta \)), the non-uniform direction distributions were determined for each cross-section using (i) the inclusion region of the individual cross-section or (ii) the inclusion region of the breast height cross-section.
of the tree in which the cross-section was located. The first approach is naturally unfeasible in practice (plot radius direction in Bitterlich sampling cannot be determined separately for different observation heights), and it was considered mainly as a “thinking experiment” — to examine area estimation errors in the case that all the 709 cross-sections of the data had been taken at breast height. The inclusion regions were determined with the viewing angles 1.146°, 1.621°, 2.292° and 3.624° (corresponding, with circular cross-sections, to the basal area factors 1, 2, 4, and 10 m²/ha, respectively). The direction distributions were estimated in a discretised form, with a point probability (probability mass) associated to each of the 180 systematic diameter directions starting from the N–S direction: for each direction j·1°, j=0, ..., 179, the probability mass w(j·1°) was estimated as the summed area of the two sectors, of angular width 1°, around the directions j·1° and (j+180)·1° divided by the total inclusion area (see Appendix H, Fig. H2). The expectations of the Bitterlich diameters were then estimated as weighted means

$$\hat{\mu}_{D(\xi)} = \sum_{j=0}^{179} w(j\cdot1^\circ) D(j\cdot1^\circ) ,$$

$$\hat{\mu}_{D(\xi+\pi/2)} = \sum_{j=0}^{179} w(j\cdot1^\circ) D[(j+90)\cdot1^\circ] ,$$

(149)

the variances as weighted variances

$$\hat{\sigma}_{D(\xi)}^2 = \sum_{j=0}^{179} w(j\cdot1^\circ) \left[ D(j\cdot1^\circ) - \hat{\mu}_{D(\xi)} \right]^2 ,$$

$$\hat{\sigma}_{D(\xi+\pi/2)}^2 = \sum_{j=0}^{179} w(j\cdot1^\circ) \left[ D[(j+90)\cdot1^\circ] - \hat{\mu}_{D(\xi+\pi/2)} \right]^2 ,$$

(150)

the correlation as

$$\hat{\rho}_{D(\xi)} \left( \frac{\pi}{2} \right) = \frac{1}{\hat{\sigma}_{D(\xi)} \hat{\sigma}_{D(\xi+\pi/2)}} \cdot \sum_{j=0}^{89} w(j\cdot1^\circ) \left[ D(j\cdot1^\circ) - \hat{\mu}_{D(\xi)} \right] \left[ D[(j+90)\cdot1^\circ] - \hat{\mu}_{D(\xi+\pi/2)} \right] ,$$

(151)

and the other moments generally as

$$\hat{E} \left[ D(\xi)^k D\left( \xi + \frac{\pi}{2} \right)^p \right] = \sum_{j=0}^{179} w(j\cdot1^\circ) D(j\cdot1^\circ)^k D[(j+90)\cdot1^\circ]^p ,$$

(152)

k=0, ..., 4, p=0, ..., 4. (Note that the probability masses w(j·1°), j=0, ..., 179, sum to one, and that for j+90≥180, D[(j+90)·1°]=D[(j–90)·1°].)

The distributions of the relative within-cross-section biases, standard deviations and RMSEs were examined in the whole set of the cross-sections and in the subsets determined by height classes. Whether the performance of an area estimator was related to cross-section size was examined by computing the correlations of the cross-section areas with the biases, standard deviations and RMSEs of each estimator. Since the relative bias of
the reference estimator $\hat{A}_0$ with respect to the convex area (i.e., the relative isoperimetric deficit $[\hat{A}_0 - A_C]/A_C$) may also be regarded as a shape index (as e.g. in Stoyan and Stoyan 1994, where it is referred to as “area-perimeter ratio”), its correlations with the scalar shape indices of Section 7.1 (Table 14) and with the cross-section size indicated by the mean diameter $\mu_D$ were examined.

7.3 Estimation of Stem Volume

The effect of diameter selection on volume estimation was investigated by applying the 22 diameter selection methods listed in Section 7.2 to the three volume estimation methods (Laasasenaho volume equation, cubic-spline-interpolated stem curve and generalised volume estimator) discussed in Chapter 5. The diameter selection methods were applied both *dependently* and *independently* at the separate observation heights within a stem: in the former, the diameter direction was selected at breast height, and the diameters at the other heights were then measured in the same direction; in the latter, the diameter directions were selected at each height independently of the other heights. However, with Bitterlich diameters (diameter selection methods 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90) only dependent selection was applied, as independent selection could not be considered realistic: in Bitterlich sampling, it is not feasible to think of viewing a tree from different viewing points at different heights. The direction distributions of the Bitterlich diameters in each tree were then naturally determined from the inclusion region of the breast height cross-section (i.e., the distribution obtained at breast height was applied also to the cross-sections at the other heights in the tree). The inclusion regions of the trees were determined with the viewing angles 1.146°, 1.621°, 2.292° and 3.624° (corresponding, with circular cross-sections, to the basal area factors of 1, 2, 4, and 10 m²/ha, respectively).

For each resulting volume estimator $\tilde{V}_{Xj}$ (a combination of the volume estimation method $X$ and the diameter selection method $j$), the within-tree expectation $E(\tilde{V}_{Xj})$ and variance $\text{Var}(\tilde{V}_{Xj})$ were estimated. From these, the relative within-tree mean squared error (MSE), consisting of the squared relative within-tree bias and the relative within-tree variance, was computed with respect to a volume-estimation-method-specific reference volume $\bar{V}_X$:

$$E\left(\frac{(\tilde{V}_{Xj} - \bar{V}_X)^2}{\bar{V}_X^2}\right) = \frac{E(\tilde{V}_{Xj}) - \bar{V}_X}{\bar{V}_X}^2 \text{Var}(\tilde{V}_{Xj}) \frac{\bar{V}_X^2}{\bar{V}_X^2}.$$  \hspace{1cm} (153)

(Note that for the Laasasenaho volume equation, the prediction errors were now defined as the opposite numbers of those in the theoretical considerations in Section 5.1; note also that for the general volume estimator, the within-tree bias was estimated “directly” without computing the within-tree expectation of the estimator.) The idea of varying the reference volume according to the volume estimation method was to separate out, in the best possible way, the diameter selection effect from the other sources of error in the volume estimator. The within-tree characteristics computed for each volume estimator in each tree are summarised in Table 17; the variances and the MSEs were considered in the square root scale (as standard deviations and root mean squared errors, RMSEs), to make them easy to interpret.

The investigation was carried out on the 79 stems within which 7 or more cross-sections had been observed; this requirement was due to the method used for estimating the true stem volume (see Section 7.3.1). However, the Laasasenaho volume equation could be examined only with the 50 stems in which the discs at both the heights of 1.3 m and 6 m had been observed.
Table 17. Characteristics estimated for each stem to examine the performance of different volume estimators (combinations of a volume estimation method and a diameter selection method). In independent selection (IS), diameter measurement directions at different observation heights within a stem were selected independently of each other; in dependent selection (DS), diameters at all observation heights were measured in the direction selected at breast height.

<table>
<thead>
<tr>
<th>Volume estimation method</th>
<th>Diameter selection methods</th>
<th>Characteristic</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laasasenaho (1982)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>volume equation re-</td>
<td>j=0, 1–11 (IS, DS),</td>
<td>[E(\hat{V}_Lj)−\hat{V}_L]/\hat{V}_L,</td>
<td>Relative within-tree bias of prediction (by the equation and diameter selection method j) with respect to the best prediction by the equation</td>
</tr>
<tr>
<td>estimated</td>
<td>1ξ–5ξ, 1290, 4ξ90, 5ξ90</td>
<td>E(\hat{V}<em>{CLj}−\hat{V}</em>{CL})/\hat{V}_{CL}</td>
<td></td>
</tr>
<tr>
<td>(i) with true area</td>
<td>(IS, DS), 1ξ–5ξ, 1290,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diameters and estimated</td>
<td>4ξ90, 5ξ90 (DS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>true volume and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) with convex area</td>
<td>j=1–5 (IS, DS), 1ξ–5ξ,</td>
<td>Var(\hat{V}_Lj)^{1/2}/\hat{V}_L,</td>
<td>Within-tree standard deviation of prediction relative to the best prediction by the equation</td>
</tr>
<tr>
<td>diameters and estimated</td>
<td>4ξ90, 5ξ90 (DS)</td>
<td>Var(\hat{V}<em>{CLj})^{1/2}/\hat{V}</em>{CL}</td>
<td></td>
</tr>
<tr>
<td>convex volume</td>
<td>j=0, 1–11 (IS, DS), 1ξ–5ξ</td>
<td>RMSE(\hat{V}_Lj)/\hat{V}_L,</td>
<td>Relative within-tree root mean squared error of prediction with respect to the best prediction by the equation; RMSE(\hat{V}_Lj)/\hat{V}_L= {[E(\hat{V}_Lj)−\hat{V}_L]^2+Var(\hat{V}_Lj)</td>
</tr>
<tr>
<td></td>
<td>4ξ90, 5ξ90 (DS)</td>
<td>RMSE(\hat{V}<em>{CLj})/\hat{V}</em>{CL}</td>
<td></td>
</tr>
<tr>
<td>Cubic-spline-</td>
<td>j=0, 1–11 (IS, DS), 1ξ–5ξ</td>
<td>[E(\hat{V}_{Sj})−\hat{V})/\hat{V},</td>
<td>Relative error of the best prediction by the equation with respect to estimated true or convex volume (relative residual); error inherent in the Laasasenaho equation</td>
</tr>
<tr>
<td>interpolated stem</td>
<td>4ξ90, 5ξ90 (DS)</td>
<td>[E(\hat{V}<em>{SJj})−\hat{V}</em>{CL})/\hat{V}_{CL}</td>
<td></td>
</tr>
<tr>
<td>curve</td>
<td>j=1–5 (IS, DS), 1ξ–5ξ,</td>
<td>Var(\hat{V}_{Sj})^{1/2} /\hat{V},</td>
<td>Relative within-tree bias of the estimator (the stem curve with diameter selection method j) with respect to estimated true or convex volume</td>
</tr>
<tr>
<td></td>
<td>4ξ90, 5ξ90 (DS)</td>
<td>Var(\hat{V}<em>{SJj})^{1/2} /\hat{V}</em>{CL}</td>
<td></td>
</tr>
<tr>
<td>General volume</td>
<td>j=0, 1–5 (DS), 6–11 (IS,</td>
<td>E(\hat{V}_{Gj})−\hat{V})/\hat{V},</td>
<td>Relative within-tree root mean squared error of the estimator with respect to estimated true or convex volume; RMSE(\hat{V}<em>{Gj})/\hat{V}= {[E(\hat{V}</em>{Gj})−\hat{V})^2+Var(\hat{V}_{Gj})</td>
</tr>
<tr>
<td>estimator</td>
<td>1ξ–5ξ, 1290, 4ξ90, 5ξ90</td>
<td>E(\hat{V}<em>{Gj})−\hat{V}</em>{CL})/\hat{V}_{CL}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4ξ90, 5ξ90 (DS)</td>
<td>Var(\hat{V}_{Gj})^{1/2} /\hat{V},</td>
<td>Within-tree standard deviation of the estimator relative to estimated true or convex volume</td>
</tr>
<tr>
<td></td>
<td>j=1–5, 1ξ–5ξ, 1290, 4ξ90,</td>
<td>Var(\hat{V}<em>{Gj})^{1/2} /\hat{V}</em>{CL}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5ξ90 (DS)</td>
<td>RMSE(\hat{V}_{Gj}),\hat{V}<em>j/\hat{V}</em>{CL},</td>
<td></td>
</tr>
<tr>
<td></td>
<td>j=0, 1–5 (DS), 6–11 (IS,</td>
<td>RMSE(\hat{V}_{Gj}),\hat{V}<em>j/\hat{V}</em>{CL},</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1ξ–5ξ, 1290, 4ξ90, 5ξ90</td>
<td>RMSE(\hat{V}_{Gj}),\hat{V}<em>j/\hat{V}</em>{CL},</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4ξ90, 5ξ90 (DS)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The distributions of the relative within-tree biases, standard deviations and RMSEs produced by the different volume estimators were examined in the set of all the stems. Further, the associations of these characteristics with tree size (as expressed by the estimated true stem volume) as well as with growing site location (in terms of the cluster co-ordinates in N-S and E-W directions) were examined by computing their correlations and by plotting them with respect to each other.

7.3.1 True Volume

The true stem volume $V$ between the stump height $h_s$ and the total height $H$ (both measured in the field with respect to the ground level, see Chapter 6) was estimated for each tree with the method that was regarded as the most accurate among those available: from the true area diameters $D_A=(D_A(h))_{h=H}=(2[A(h)/\pi]^{1/2})_{h=H}$ observed at heights $H=(h_1, \ldots, h_{n_0}, H)$, a stem curve $S_3(h; D_A, H)$ was estimated with interpolating cubic splines, and the volume was then obtained as the solid of revolution of the curve as

$$\tilde{V} = \int_{h_s}^{H} S_3(h; D_A, H)^2 \, dh.$$  \hspace{1cm} (154)

(Instead of the symbol $\tilde{V}_S$ used in Eq. 125 in Section 5.3, the volume estimate is here denoted as $\tilde{V}$, because it will serve as the reference volume not only for the cubic-spline-interpolated stem curve method but also for the generalised volume estimator and be used to assess the model error component independent of diameter selection in the Laasasenaho volume equation). The computations were performed with the Fortran 77 subroutines written by Mr Carl-Gustaf Snellman in the Finnish Forest Research Institute and based on the work of Lahtinen and Laasasenaho (1979).

By their empirical experience, Lahtinen and Laasasenaho (1979, p. 54) recommended that the stem curve estimation be based on not fewer than 7 observation points along the stem. In the 79 trees included in the volume investigation, the number of observation heights varied between 7 and 10, and thus, with the tree height $H$ and the assumed top diameter $D_A(H)=0.4 \text{ cm}$ included, the interpolation was based on 8–11 observation points. The initial values (required for a unique solution of the system of the spline equations) were determined via estimated second derivatives of the stem curve at the lowest and highest observation heights (cf. Lahtinen and Laasasenaho 1979, p. 19, 31); the derivatives were estimated by approximating the stem curve at both ends of the stem with interpolating parabolas passing through the three lowermost and uppermost observation points.

As the squared cubic splines consist of piecewise polynomials of the observation heights, the squared stem curve could be integrated analytically. In addition to the total volume, the volume of the stem segment between the lowermost and the uppermost observation heights $h(1)$ and $h(m)$ was estimated; this partial volume, also denoted by $\tilde{V}$, was required as the reference volume for the general volume estimators. In the same way, the total and partial convex volumes, both denoted by $\tilde{V}_C$, were estimated for each stem from the convex area diameters $(D_{Ac}(h))_{h=H}=(2[A_c(h)/\pi]^{1/2})_{h=H}$.

In 11 stems of the total of 79, there were one or two intervals between adjacent observation heights where the true area was larger at the upper height than at the lower height; in 10 of these 11 stems this was the case also with the convex area. Such non-monotonicity might result in an aberrant oscillation in a cubic spline function. However, the taper curves obtained for the said 11 stems were regarded as acceptable — the three worst ones are shown...
in Fig. 28 — and thus none of the known improvements for the algorithm (Lahtinen 1988 and 1993) were considered necessary to implement.

### 7.3.2 Laasasenaho Volume Equation

Stem volume estimation for standing trees was imitated by applying the 22 different diameter selection methods (see Section 7.2) to the Laasasenaho (1982) three-variable volume equation where the originally unspecified diameters at the heights of 1.3 m and 6 m were fixed as the true area diameters (cf. the discussion in Section 5.1):

\[
V = \beta_1 D_A (1.3)^2 + \beta_2 D_A (1.3)^2 H + \beta_3 D_A (1.3)^3 H + \beta_4 D_A (1.3)^2 H^2 \\
+ \beta_5 [D_A (1.3)^2 + D_A (1.3) D_A (6) + D_A (6)^2] + \beta_6 D_A (6)^2 (H - 6) + \varepsilon_L .
\]

(155)

In Laasasenaho’s original specification, the stem volume was defined to extend from the top of the tree either to the highest root collar affecting cutting, or, if the collar did not exist or was located below the height of 10 cm, to the height of 10 cm (Laasasenaho 1982). This is how the cutting point (stump height \( h_s \)) of each tree was determined also in this study, and thus the estimated true volume in this study (see Section 7.3.1) agreed well with the definition of the stem volume in the original work.
Instead of employing Laasasenaho’s parameter estimates based on the data of 2326 Scots pine sample trees from all over Finland, we re-estimated the parameters from the 50 stems of our data in which the cross-section had been observed at both the heights of 1.3 m and 6 m; this was to make the error inherent in the model (the random error term \( \varepsilon_L \), also referred to as the model error) as small as possible. The estimation was performed both without and with the assumption of cross-section convexity: in the former, the estimated true volume \( \hat{V} \) was employed as the response variable and the true area diameters \( D_A(1.3) \) and \( D_A(6) \) (as well as the height \( H \)) were used as the explanatory variables; in the latter, the response variable was the estimated true convex volume \( \hat{V}_C \) and the explanatory variables comprised the convex area diameters \( D_{Ac}(1.3) \) and \( D_{Ac}(6) \). As the error terms were assumed heteroscedastic (variance proportional to \( D_A(1.3)^4 H^2 \), or to \( D_{Ac}(1.3)^4 H^2 \); refer to Section 5.1), a weighted least squares estimation was performed with \( 1/D_A(1.3)^4 H^2 \), or \( 1/D_{Ac}(1.3)^4 H^2 \), as the weight of each observation (stem); this corresponds to the ordinary least squares estimation of a form factor model carried out in the original work (Laasasenaho 1982, p. 43–44). As in the original study, the within-plot and within-cluster interdependencies of the stems, which at least in principle breached the model assumptions, were disregarded in the estimation. Judging from the residual plots, the models fitted well in the data; also, the relative root mean squared errors (the standard deviations of the relative residuals adjusted with the number of parameters) were in concordance with that in the original data (3.59% and 3.60% vs. 3.53%; Laasasenaho 1982, p. 43). The original and the re-estimated parameter values are given in Table 18. In the following, we will deal with the volume estimation without the convexity assumption; the particularities related to the estimation with the convexity assumption will be considered in the end of the section.

We assumed the re-estimated model to be correct and thus disregarded the effects of using estimated parameters instead of the true ones (cf. the discussion in Section 5.1). As the reference volume (to which the estimators by different diameter selection methods were to be compared) we used the best estimate

\[
\hat{V}_L = \hat{\beta}_1 D_A(1.3)^2 + \hat{\beta}_2 D_A(1.3)^2 H + \hat{\beta}_3 D_A(1.3)^3 H + \hat{\beta}_4 D_A(1.3)^2 H^2
+ \hat{\beta}_5 \left[ D_A(1.3)^2 + D_A(1.3)D_A(6) + D_A(6)^2 \right] + \hat{\beta}_6 D_A(6)^2 (H - 6)
\]

(156)

this was because we wanted to distinguish the error component caused by diameter selection from the error component inherent in the model (the model error that is present even if the
diameters were measured “correctly”, i.e., even if the true area diameters at the heights of 1.3 m and 6 m in the stem were known). The model error in each stem was then estimated as the relative error of this best estimate $\tilde{V}_L$ with respect to the estimated true volume $\hat{V}$ (i.e., as the relative residual $(\hat{V} - \hat{V})/\hat{V}$ of the estimated model). The purpose of the diameter specification as true area diameters (to comply with the estimation of true volume) and the parameter re-estimation was expressly to make this error as small as possible, that is, to set the reference volume as close to our estimated true volume as possible.

Applying the different diameter selection methods to the volume equation resulted in the volume estimators (here written as a function of the diameters only)

$$\hat{V}_{Lj} = \hat{c}_1 D_{j}(1.3)^2 + \hat{c}_2 D_{j}(1.3)^3 + \hat{c}_3 D_{j}(1.3)D_{j}(6) + \hat{c}_4 D_{j}(6)^2$$

(157)

where $\hat{c}_1 = \hat{\beta}_1 + \hat{\beta}_2 H + \hat{\beta}_4 H^2 + \hat{\beta}_5$, $\hat{c}_2 = \hat{\beta}_3 H$, $\hat{c}_3 = \hat{\beta}_5$, and $\hat{c}_4 = \hat{\beta}_5 + \hat{\beta}_6 (H - 6)$, and where $D_{j}(1.3)$ and $D_{j}(6)$ are the diameters selected with method $j$ at the heights of 1.3 m and 6 m.

With the diameter selection methods involving only fixed diameters (methods 0, 6–11, min and max; independent and dependent selection), there was naturally no within-tree variation in the estimator values, whereby the within-tree expectation of the estimator equalled the constant estimate itself (i.e., $E(\hat{V}_{Lj}) = \hat{V}_{Lj}$ and $\text{Var}(\hat{V}_{Lj}) = 0$).

With the diameter selection methods involving random diameters (methods 1–5, independent and dependent selection; methods 1ξ–5ξ 1ξ90, 4ξ90 and 5ξ90, dependent selection), the within-tree expectation and variance of the volume estimator were estimated by means of estimated diameter moments and product moments:

$$\hat{E}[\hat{V}_{Lj}] = \hat{c}_1 \hat{E}[D_{j}(1.3)^2] + \hat{c}_2 \hat{E}[D_{j}(1.3)^3] + \hat{c}_3 \hat{E}[D_{j}(1.3)D_{j}(6)] + \hat{c}_4 \hat{E}[D_{j}(6)^2]$$

and

$$\text{Var}[\hat{V}_{Lj}] = \hat{c}_2^2 \hat{E}[D_{j}(1.3)^6] + 2\hat{c}_1 \hat{c}_2 \hat{E}[D_{j}(1.3)^4 D_{j}(6)] + \hat{c}_3^2 \hat{E}[D_{j}(1.3)^4 D_{j}(6)] + 2\hat{c}_2 \hat{c}_3 \hat{E}[D_{j}(1.3)^3 D_{j}(6)^2] + 2\hat{c}_1 \hat{c}_3 \hat{E}[D_{j}(1.3)^3 D_{j}(6)] + (\hat{c}_3^2 + 2\hat{c}_1 \hat{c}_4) \hat{E}[D_{j}(1.3)^2 D_{j}(6)^2] + 2\hat{c}_2 \hat{c}_4 \hat{E}[D_{j}(1.3)D_{j}(6)^4] + \hat{c}_4^2 \hat{E}[D_{j}(6)^4] - \{\hat{E}[\hat{V}_{Lj}] \}^2$$

(159)

refer to Appendices E and F for the more detailed method-specific expressions. The diameter moments and product moments were estimated as explained in Section 7.2: With the methods 1–5 involving the uniform direction distribution, the moment estimates were obtained as simple means from the 180 systematic diameters in each cross-section. With methods 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90 involving the non-uniform direction distributions of the Bitterlich diameters, the moments were estimated as weighted means from the 180 systematic diameters, the weights being determined as the relative sector areas of the inclusion region of the breast height cross-section.

The assumption of cross-section convexity influenced the reference volume, the estimated model error, and the within-tree expectations and variances of the estimators: The convex reference volume $\tilde{V}_{CL}$ was obtained from the volume equation estimated with true convex volume $V_C$ and the convex area diameters, and the relative model error $(\tilde{V}_{CL} - \tilde{V}_C)/\tilde{V}_{CL}$
was naturally determined as the relative residual of this model. Since the diameters of a non-convex cross-section and its convex closure coincide, the within-tree expectation and variance of the estimator $\hat{V}_{CL,j}$ produced by each diameter selection method $j$ were obtained by using the same diameters and diameter moment estimates as in the non-convex case, but with the parameter estimates $\hat{c_1}, \hat{c_2}, \hat{c_3}$ and $\hat{c_4}$ of the volume equation estimated with the true convex volume and the convex area diameters.

### 7.3.3 Cubic-Spline-Interpolated Stem Curve

Stem volume estimation for a felled sample tree was mimicked by applying each of the 22 diameter selection methods (see Section 7.2) at the 7-10 observation heights in a stem and constructing the cubic-spline-interpolated stem curves from the resulting diameter vectors (with the assumed diameter of 0.4 cm at the top of the tree added in them). Let us denote with $D_j$ the diameter vector obtained with the diameter selection method $j$, with $H$ the vector containing the observation heights (plus the tree height) and with $S_3[h; D_j, H]$ the interpolated stem curve. The volume estimate by each diameter selection method was then obtained as the solid of revolution of the stem curve from the recorded stump height $h_s$ to the height $H$ of the tree:

$$\hat{V}_{S_j} = \frac{\pi}{4} \int_{h_s}^{H} S_3[h; D_j, H]^2 \, dh.$$  \quad (160)

Technically the volume estimation was carried out in the same manner as the estimation of the true volume (see Section 7.3.1; however, potential oscillation in the stem curves due to diameter increase within short distances upward was not controlled here). All the 79 stems with cross-sections observed at 7 or more heights were included in the investigation.

With the diameter selection methods involving fixed diameters (methods 0, 6–11, min and max; independent and dependent selection), there was no within-tree variation in the volume estimates (i.e., $\text{Var}(\hat{V}_{S_j})=0$), and the within-tree expectation was given by the constant volume estimate itself (i.e., $E(\hat{V}_{S_j})=\hat{V}_{S_j}$).

With the diameter selection methods involving random diameters (methods 1–5, independent and dependent selection; methods $1_\xi$–$5_\xi$, $190$, $4_\xi90$ and $5_\xi90$, dependent selection), the diameter vector selection was repeated several times in each stem in order to capture the within-tree variation in volume estimates. The “population” from which the diameter vectors were sampled varied according to the diameter selection method: it consisted of all the diameter vectors that could be formed by the diameter selection method in question from the systematically measured diameters, 180 at each observation height, in the stem.

The **dependent selection** of diameters at the different heights in a stem restricted the diameter vector population so that all the vectors therein could be included in the sample: with the methods where only one diameter direction within a cross-section was actually selected and where the other direction, if taken, was fully determined by the first one (methods 1–3, $1_\xi$–$3_\xi$ and $1_\xi90$), there were 180 possible diameter vectors; with methods where two diameter directions within a cross-section were selected independently of each other (methods 4, 5, $4_\xi$, $5_\xi$, $4_\xi90$ and $5_\xi90$), the number of possible diameter vectors was $180^2$, as for each first diameter direction selection there were 180 alternatives as the second choice. Using these 180 or $180^2$ diameter vectors meant selecting the one diameter direction or the pair of diameter directions equidistantly (by every 1°) within $[0°, 180°)$ or $[0°, 180°)\times[0°, 180°)$. To each diameter vector, the estimated selection probability of the one diameter direction or the pair of diameter directions was associated as a weight.
The probability was determined by means of the diameter direction distribution(s): with one direction selection, the weight was the probability of selecting a direction within 0.5° from the chosen direction (the probability mass within 0.5° from the common direction of the diameters in the diameter vector); with two independent direction selections within a cross-section, the weight was the product of such probabilities related to the two chosen directions. The within-tree expectation and variance of each volume estimator (diameter selection method) were then estimated as the weighted mean and variance of the volume estimates obtained with the diameter vectors of the method (cf. Eqs. 127 and 128 and the discussion on numerical integration in Section 5.3). (Thus, with interrelated diameter selection methods (1, 1ξ and 1ξ90; 2, 2ξ and 2ξ90; 3, 3ξ and 3ξ90 etc.), the volume estimates were computed using the same diameter vectors, but the within-tree expectation and variance were estimated with different weights.) With only the uniform direction distribution involved (methods 1–5), the weights associated to each diameter vector became equal, and the weighted mean and variance of the volume estimates were simplified to the simple arithmetic mean and sample variance. With the non-uniform direction distributions of the Bitterlich diameters involved in the first direction selection (methods 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90), the weights became unequal and were obtained from the relative sector areas of the inclusion region of the breast height cross-section of the tree (cf. Section 7.2).

In the independent selection of diameters at the different heights in a stem (applied only to the diameter selection methods 1–5 involving the uniform direction distribution), the diameter vector population grew so large that all the vectors therein could not be used: with methods 1–3 the number of possible diameter vectors amounted to 180m and with methods 4–5 to 1802m, m being the number of observation heights, total height excluded. Instead, we drew upon Monte Carlo integration where 180 2 diameter direction vectors were to be sampled independently from the multidimensional direction distribution of each diameter selection method. With the uniform marginal distributions of the independent directions, the multidimensional direction distributions became also uniform, implying that all the diameter vectors of generated by the method had the same probability to be chosen. Therefore a simple random sample of 1802 vectors was drawn from the diameter vector population of each diameter selection method, with replacement for practical ease; in practice, the sampling was carried out by selecting randomly at each observation height the required one or two diameters among the 180 alternatives and by repeating this procedure 1802 times for each stem. The within-tree expectation and variance of the volume estimator were then estimated as the mean and sample variance of the volume estimates computed from the selected diameter vectors (cf. Eqs. 129 and 130 and the discussion on Monte Carlo integration in Section 5.3).

As the natural reference volume, we employed the estimated true volume \( \hat{V} \) computed with the similar stem curve approach from the true area diameters (see Section 7.3.1). Unlike with the volume equation, the error contained in the reference volume could not now be quantified, as the true volumes of the stems were unknown and no more precise estimates were available. Consequently, the within-tree biases and variances produced by the different diameter selection methods could only be compared with each other; no notion about their magnitude with respect to the “error intrinsic in the volume estimation method” could be gained.

Since the diameters are the same for a non-convex cross-section and its convex closure, assuming cross-section convexity influenced only the reference volume: with the convexity assumption, we employed the convex volume estimate \( \hat{V}_C \) computed from the convex area diameters (see Section 7.3.1) as the reference.
Finally, we considered estimating the volume of the stem segment from the lowermost observation height \( h(1) \) to the uppermost one \( h(m) \) by means of general volume estimators that involved area estimation functions \( \hat{A}_j(h) = \pi D_j(h)^2/4 \) arising from the application of each of the 22 diameter selection methods (see Section 7.2; \( j \) refers to the diameter selection method) and the circle area formula at all the heights \( h \in [0, H] \) in a stem:

\[
\hat{V}_{Gj} = \int_{h(1)}^{h(m)} \hat{A}_j(h) \, dh = \frac{\pi}{4} \int_{h(1)}^{h(m)} D_j(h)^2 \, dh .
\]

(161)

We omitted to define explicitly the values of the area estimation functions between the observation heights (and to estimate the stem volumes actually) but focused on “direct” estimation of the bias, covariance and variance functions \( \mu_{\hat{A}_j}(h) \), \( \gamma_{\hat{A}_j}(h, k) \) and \( \sigma_{\hat{A}_j}^2(h) \) of the area estimation process \( \{\hat{A}_j(h), h \in [0, H]\} = \{\pi D_j(h)^2/4, h \in [0, H]\} \) of each diameter selection method in each tree. As explained in Section 5.2, the within-tree bias and variance of each general volume estimator (diameter selection method) in each tree could then be estimated by integrating the estimated treewise bias and covariance functions:

\[
\hat{E}[\hat{V}_{Gj} - V] = \int_{h(1)}^{h(m)} \hat{\mu}_{\hat{A}_j}(h) \, dh = \hat{E}[\hat{A}_j(h) - \Lambda(h)] \, dh ,
\]

(162)

and

\[
\text{Var}(\hat{V}_{Gj}) = \int_{h(1)}^{h(m)} \int_{h(1)}^{h(m)} \hat{\gamma}_{\hat{A}_j}(h, k) \, dh \, dk = \int_{h(1)}^{h(m)} \int_{h(1)}^{h(m)} \text{Cov}[\hat{A}_j(h), \hat{A}_j(k)] \, dh \, dk .
\]

(163)

Further, an estimate for the theoretical upper bound of the within-tree variance (the variance conditional on fully correlated area estimation process) was given by the definite integral of the estimated variance function:

\[
\text{Var}[\hat{V}_{Gj} | \rho_{\hat{A}_j}(h, k) = 1] = \phi_{\hat{A}_j}^2(h) \int_{h(1)}^{h(m)} \sqrt{\hat{\sigma}_{\hat{A}_j}^2(h)} \, dh = \left\{ \int_{h(1)}^{h(m)} \sqrt{\text{Var}[\hat{A}_j(h)]} \, dh \right\}^2 .
\]

(164)

As in the cubic-spline-interpolated stem curves before, all the 79 stems with 7 or more observed cross-sections were included in the examination. For each diameter selection method \( j \) in each stem, the bias, variance and covariance functions were estimated from the
discrete observations on the area estimator \( \hat{A}_j(h) = \pi D_j(h)^2/4 \) along the stem, that is, from the vectors of the estimated within-cross-section biases \( (\hat{E}[\hat{A}_j(h)] - A(h))_{h \in H} \) and variances \( (\text{Var}[\hat{A}_j(h)])_{h \in H} \) at the 7–10 observation heights \( H = (h_1, h_2, ..., h_{m}) \) and from the matrix of the estimated between-cross-sections covariances \( (\text{Cov}[\hat{A}_j(h), \hat{A}_j(k)])_{h, k \in H} \) at the 42–90 two-height combinations \( H \times H \) in the stem.

With the diameter selection methods involving only fixed diameters (methods 0, 6–11, min, max; independent and dependent selection), the variances and covariances of the area estimators at separate heights were naturally zero and the biases were given by the differences between the area estimates and the true area at each height.

With the diameter selection methods involving random diameters, only dependent selection was considered: With Bitterlich diameters (methods 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90), independent selection was not feasible, as noted before. With diameters having the uniform direction distribution (methods 1–5), independent selection was to result, by definition, in the same within-cross-section biases and variances of the area estimators as dependent selection, as well as in zero covariances between the area estimators at separate heights, thus providing no information for covariance function estimation (refer to the discussion on an area estimation process with uncorrelated elements in Section 5.2). In all the cross-sections, the biases and the variances of the area estimators had already been estimated when considering cross-section area estimators (Section 7.2). Similarly, the covariances of the area estimators at different heights could be expressed in terms of diameter moments and product moments, which were then estimated from the 180 systematic diameters and their products in each cross-section (cf. Section 7.2): With the methods involving the uniform direction distribution, the moment estimates were obtained as simple means. With the methods involving the non-uniform direction distributions of the Bitterlich diameters, the moment estimates were obtained as weighted means, the weights being determined as the relative sector areas of the inclusion region of the breast height cross-section.

From the discrete estimated values, the bias, variance and covariance functions of the area estimation process in each tree were estimated by linear interpolation. Linear interpolation was employed for simplicity and for its assuredly reasonable behaviour: as there were 56 functions to be estimated in each of the 79 stems (3 functions for each of the methods 1–5, 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90 with dependent selection; 1 function for the method 0; and 1 function for each of the methods 6–11, min and max with both dependent and independent selection), the use of smoother higher order interpolating splines requiring control and adjustment for possible non-feasible oscillations was regarded as too laborious an approach.

The linear functions could be integrated analytically. The integration limits were set as the lowermost and the uppermost observation heights \( h(1) \) and \( h(m) \), because the functions could not be justifiably extrapolated to cover the whole stem from the stump height to the very top. Thus, unlike with the other volume estimation methods, the volume bias and variance were computed only for that part of the stem within which cross-sections had been observed. As the reference volume, the estimated true volume \( \hat{V} \) of the stem segment between the lowermost and the uppermost observation heights (see Section 7.3.1) was used.

The assumption of cross-section convexity did not influence the actual area estimation processes \( \{\hat{A}_j(h), h \in H\} \) (based on diameters and the circle area formula, the area estimators \( \hat{A}_j(h) \) were the same whether the cross-sections were assumed convex or not) but affected the area estimation error processes \( \{\hat{A}_j(h) - A_C, h \in H\} \), where the errors were now taken with respect to the convex area \( A_C \). Consequently, with the convexity assumption, only the bias functions of the 30 area estimation processes (22 diameter selection methods with dependent selection, 8 methods with independent selection) needed to be re-estimated from the biases taken with respect to the convex area in each stem; the integrals of these functions
were then thought to yield the estimated biases $\hat{V}_{Gj} - V_C$ of the general volume estimators with respect to true (partial) convex volume $V_C$. The reference volume, to which the estimated within-tree volume biases and variances were proportioned, was naturally also changed for the estimated partial convex volume $\tilde{V}_C$ (see Section 7.3.1).

7.4 Estimation of Stand Totals in Bitterlich Sampling

In Bitterlich sampling, applying circularity assumption to non-circular cross-sections was found to inflict potential bias in stand total estimators, due to two tree-specific faults in the estimated inclusion probabilities (see Eq. 69 in Section 4.1.2): (i) deviation of the true basal area factor of the tree $\kappa(\alpha) = A_C/|M(\alpha)|$ (the ratio of the convex area of the breast height cross-section to the inclusion area) from that of a circle $\sin^2(\alpha/2)$, and (ii) estimation of the convex area of the breast height cross-section with estimator $\hat{A}_j = \pi D_j^2/4$ based on diameter $D_j$ measured in the cross-section by some method $j$. With the viewing angle $\alpha$ values of 1.146°, 1.621°, 2.292° and 3.624° (corresponding, with circular cross-sections, to the basal area factor values of 1, 2, 4, and 10 m²/ha, respectively), these errors were investigated in the 80 breast height cross-sections of the data. As a “thinking experiment”, the examination was then extended to all the 709 cross-sections, using their own inclusion regions as if they were all breast height cross-sections.

The deviation of the true basal area factor from that of a circle, expressed as the ratio $\sin^2(\alpha/2)/\kappa(\alpha)$, was first computed separately in each cross-section. In the estimation of the relative basal area of a stand by Bitterlich sampling, $\sin^2(\alpha/2)/\kappa(\alpha) - 1$ straightforwardly indicates the contribution of the tree to the bias of the basal area estimator, since

$$E(\hat{G}) - G = \frac{1}{|L|} \sum_{i \in L} A_C \left[ \frac{\sin^2(\alpha/2)}{\kappa(\alpha)} - 1 \right]$$

(cf. Eqs. 77 and 81 in Section 4.1.3). To check Matérn’s approximate theoretical result that, with viewing angle $\alpha$ of the magnitude 1°, this bias is the same as we would get by calipering every stem in the stand in a randomly chosen direction, that is,

$$E(\hat{G}) - G = \frac{1}{|L|} \sum_{i \in L} \left[ E_\theta(\hat{A}_1) - A_C \right]$$

(cf. Eq. 85 in Section 4.1.3), $\sin^2(\alpha/2)/\kappa(\alpha) - 1$ in each cross-section was compared to the relative within-cross-section bias $[E_\theta(\hat{A}_1) - A_C]/A_C$ of the area estimator $\hat{A}_1$ based on one random diameter, computed already earlier (see Section 7.2).

The combined effect $\sin^2(\alpha/2)/\kappa(\alpha) A_C/(\pi D_j^2/4)$ was computed with the 22 diameter selection methods (see Section 7.2) in each cross-section. With the methods involving randomness (1–5, 1ξ–5ξ, 1ξ90, 4ξ90 and 5ξ90), the expectation of the effect

$$E \left[ \frac{\sin^2(\alpha/2)}{\kappa(\alpha)} \cdot \frac{A_C}{\pi D_j^2/4} \right] = \frac{4}{\pi} \sin^2(\alpha/2) |M(\alpha)| E \left[ \frac{1}{D_j^2} \right]$$

(167) over the diameter direction distribution was considered. The fractional moments $E[1/D_j^2]$ were estimated in a similar way as moments and product moments in Section 7.2 (cf. Eqs. 148, 149 and 152). With the methods involving the uniform direction distribution (1–5),
the moments were estimated as the simple means of the 180 or 180^2 systematic diameters raised to the power of –2 in each cross-section:

\[ \hat{E}[D(\theta)^{-2}] = \frac{1}{180} \sum_{j=0}^{179} D(j \cdot 1°)^{-2} , \]

\[ \hat{E} \left\{ \left[ \frac{D(\theta) + D(\theta + \pi / 2)}{2} \right]^2 \right\} = \frac{1}{180} \sum_{j=0}^{179} \left\{ \frac{D(j \cdot 1°) + D[(j+90) \cdot 1°]}{2} \right\}^{-2} , \]

\[ \hat{E} \left\{ \left[ \sqrt{D(\theta)D(\theta + \pi / 2)} \right]^2 \right\} = \frac{1}{180} \sum_{j=0}^{179} \left[ D(j \cdot 1°)^{-1}D[(j+90) \cdot 1°]^{-1} \right] , \]

\[ \hat{E} \left\{ \left[ \frac{D(\theta_j) + D(\theta_{j+90})}{2} \right]^2 \right\} = \frac{1}{180^2} \sum_{j=0}^{179} \sum_{k=0}^{179} \left[ D(j \cdot 1°)^{-1}D[(j+90) \cdot 1°]^{-1} \right]^{-2} , \]

\[ \hat{E}[D(\xi)^{-2}] = \sum_{j=0}^{179} w(j \cdot 1°)D(j \cdot 1°)^{-2} , \]

\[ \hat{E} \left\{ \left[ \frac{D(\xi) + D(\xi + \pi / 2)}{2} \right]^2 \right\} = \sum_{j=0}^{179} w(j \cdot 1°) \left\{ \frac{D(j \cdot 1°) + D[(j+90) \cdot 1°]}{2} \right\}^{-2} , \]

\[ \hat{E} \left\{ \left[ \sqrt{D(\xi)D(\xi + \pi / 2)} \right]^2 \right\} = \sum_{j=0}^{179} w(j \cdot 1°)D(j \cdot 1°)^{-1}D[(j+90) \cdot 1°]^{-1} , \]

\[ \hat{E}[D(\xi + \pi / 2)^{-2}] = \sum_{j=0}^{179} w(j \cdot 1°)D[(j+90) \cdot 1°]^{-2} . \]

(Note that for \( j+90 \geq 180 \), \( D[(j+90) \cdot 1°] = D[(j–90) \cdot 1°] \).) With the methods involving the non-uniform direction distributions of the Bitterlich diameters (1ξ–3ξ, 1ξ90), the moments were estimated as the weighted means of the diameter powers, the weights being determined as the relative sector areas of the inclusion region:

\[ \hat{E}[D(\theta)^{-2}] = \sum_{j=0}^{179} \sum_{k=0}^{179} w(j \cdot 1°)D[(j+90) \cdot 1°]^{-1} \]

\[ \hat{E} \left\{ \left[ \frac{D(\theta_j) + D(\theta_{j+90})}{2} \right]^2 \right\} = \sum_{j=0}^{179} \sum_{k=0}^{179} \left[ D(j \cdot 1°)^{-1}D[(j+90) \cdot 1°]^{-1} \right]^{-2} , \]

\[ \hat{E} \left\{ \left[ \sqrt{D(\theta_j)D(\theta_{j+90})} \right]^2 \right\} = \sum_{j=0}^{179} \sum_{k=0}^{179} \left[ D(j \cdot 1°)^{-1}D[(j+90) \cdot 1°]^{-1} \right]^{-2} . \]

(Note that the weights \( w(j \cdot 1°), j=0, ..., 179 \), sum to one.) And with the methods involving both types of diameters (4ξ–5ξ, 4ξ90–5ξ90), the moments were estimated as the simple means of the weighted means (which in the case of the geometric mean of independent diameters turned into the product of the weighted and the simple mean):
The distributions of \(\sin^2(\alpha/2)/\kappa(\alpha)\) and \(E[\sin^2(\alpha/2)/\kappa(\alpha)\cdot A_C/(\pi D_j^2/4)]\) were viewed in the set of the breast height cross-sections and in the set of all the cross-sections. The relation between \(\sin^2(\alpha/2)/\kappa(\alpha)\) and cross-section shape, as expressed by the scalar shape indices (see Table 14 in Section 7.1), was studied with pairwise correlations and accompanying scatterplots.