



Spline representation of irregular and malformed stem profiles of broadleaved tree species in White Carpathian Mountains

Karel Kuželka, Róbert Marušák

Česká zemědělská univerzita v Praze, Fakulta lesnická a dřevařská, Kamýcká 1176, 165 21 Praha 6 – Suchbátka, Česká republika, E-mail: kuzelka@fd.czu.cz, marusak@fd.czu.cz

Abstract: Kuželka K., Marušák R. 2012: Spline representation of irregular and malformed stem profiles of broadleaved tree species in White Carpathian Mountains. – *Beskydy*, 5 (2): 111–120

The paper deals with representation of individual stem profiles of beech and oak in White Carpathian Mountains, Czech Republic. Due to complicacy and irregularity of stem profiles of most trees common taper functions are not applicable for individual stem profile description. Selected spline functions were used to model the irregular stem profiles. To fit complicated stem profiles with a spline more input points are needed than for regular stems. The splines were computed using four fixed input points representing diameters in four conventional heights (base diameter, stump diameter, DBH and the height) and another two input points interactively added to describe the location of the most significant malformation. More input points were added in respect of the malformation position. To evaluate the accuracy of each spline a number of stems have been measured. For each stem diameters with interspaces of 0.2m have been measured from the base to the top. At each measurement position predicted diameters with the real value were compared. Five different statistical indicators were used to evaluate the accuracy of the spline models. The profiles were the best represented by Catmull-Rom spline with tangent vector length corrections. Previously frequently used natural cubic spline is not suitable for irregular stems. When interpolating complicated stem profiles it suffers from oscillation considerably more than for regular stems of coniferous species.

Key words: spline, stem profile, irregular stems, beech, oak

Introduction

The form and taper of tree stems has been studied for more than a century (e. g. Höjer 1903 in Kitikidou, Chatzilazarou 2008). By form is meant the shape of a stem or the dependence of diameter on height, while taper is defined as the rate of narrowing in diameter with respect to increase in height (Gray, 1956). Using taper functions volume of the stem can be precisely derived; taper function is also an important basis of estimation of sorting of the stem and of optimal utilization of the wood mass.

On assumption that stems of forest trees conform to the dimensions of predesignated geometric solids many simple taper models of

polynomial (eg. Matte 1949, Bruce 1968 et al., Goulding, Murray 1976), logarithmic (Demaeerschalk 1972), trigonometric (Thomas, Parresol 1991) or other (Biging 1984) forms have been developed for trees of a wide range of species and geographical areas. They were intended to provide reasonable estimates of upper stem diameters and volume of the stem with little input data.

Another approach doubts, that any pre-conceived general functional form can represent the taper properly (Grosenbaugh 1966, Brooks et al. 2008, Sharma, Parton 2009) because stem taper is an unstable factor and individual trees seem capable to assume infinite variety of shapes. Liu (1980) supposes that more rational approach

than fitting the taper to any preconceived functional form is to provide numerical technique that is capable of assuming various forms. Such approach is represented by segmented models (eg. Max, Burkhart 1976, Demaerschalk, Kozak 1977, Jiang et al. 2005), variable-exponent models (eg. Flewelling, Raynes 1993, Eerikäinen 2001, Lee et al. 2003, Li et al. 2012) or spline functions. In the last decade, most of the attention concerning stem profile modeling was paid to the variable-exponent models, which unlike splines have shape constraints, no risk of oscillation and due to their parametric form are easily applicable and generalizable. On the other hand, the splines are considerably more formable and can be applied to a single stem in contrast to the variable-exponent models, which need a number of stems for parameter values setoff.

The term spline is a general tag for a wide class of functions defined piecewise by polynomial segments. The cubic spline was many times used to describe the stem form (Goulding 1979, Lahtinen and Laasasenaho 1979, Liu 1980, Smaltschinski 1983, Figueiredo-Filho et al. 1996, Laasasenaho et al. 2005), to model bark thickness (Laasasenaho et al. 2005) or to predict the stem curve from lower stem measurement (Nummi and Möttönen 2004, Koskela 2006). From other splines Lahtinen (1988) used monotony preserving quadratic spline eliminating oscillations. All previous studies concerned only coniferous species. Utilizing splines for stem profile representation of broadleaved species may be more beneficial than for coniferous species due to higher complicity of stem profile of broadleaved species compared to coniferous species and due to high percentage of stems with irregular shape. Irregularities in broadleaved stem profiles are represented mainly by sudden diameter drop due to bifurcation of the main stem or large branching. Moreover such a large branching is usually preceded by local enlargement of the stem diameter, which violates the monotony of the stem profile and makes the profile even more complicated. Profiles of such irregular stems are difficult or even impossible to be described by common taper functions usually restricted to create a smooth monotone curve.

Material and Methods

In this study data from 43 individuals of European beech (*Fagus sylvatica* L.) and sessile oak (*Quercus petraea* [Matt.] Liebl.) were used. The trees were taken from several stands of different age located in protected landscape area Bílé

Karpaty (White Carpathians), Czech Republic. DBH of the trees varied between 168 mm and 635 mm; heights ranged from 14.6 m to 33.4 m. Diameters outside bark were measured on the felled trees with interspaces 0.1 m from the ground level to the breast height and with interspaces 0.2 m in the range from the breast height to the top of the stem. Distances from the tree base were measured using steel tape with precision 0.01 m, the diameters were recorded into electronic caliper with precision 1 mm.

The selection of splines used in this study is based on results of previous work (Kuželka 2011a, 2011b), where several splines were compared concerning their suitability to represent the stem profile of coniferous species. Three candidates were selected for further utilization: Catmull-Rom spline (Kochanek and Bartels 1984) with tangent vector corrections, natural cubic spline and B-spline of 2nd degree (Piegl and Tiller 1996, Linkeová 2007).

Catmull-Rom spline is an easily computable interpolation curve consisting of cubic segments. The way of connection of the segments ensures 1st degree continuity whereas the 2nd degree continuity is infringed in join points. This allows the curve to readily change the direction at the expense of slightly unnatural look. Natural cubic spline is defined by cubic segments with second degree continuity preserved throughout the curve and zero torsion in the endpoints. It has the minimal curvature among all twice continuously differentiable interpolating curves. These properties and a smooth natural run make the natural cubic spline the most exploited representative of splines. B-spline is the most important representative of approximation splines, which does not fit the input points, but it follows their geometry. Clamped B-spline, which exactly fits the first and the last point, was used in this study.

The basic set of input points for spline computation contained six input points. Four fixed input points represented diameters conventional heights (base diameter, stump diameter, DBH and the height). Another two input points interactively added to describe the location of the most significant malformation. A position of the seventh input point was optimized to maximize the accuracy of the curve. It is obvious, that the optimal position of the next input point depends on the location of the two inserted points. If they are located in the lower stem, it is necessary to place the point so that it controls the upper part of the curve. With shifting the inserted points upwards also the additional point must be shifted so that the input points are

evenly distributed along the stem. Above certain threshold height it is not useful to specify more the curve above the inserted points; the accuracy of the curve is improved by placing the point rather to the lower part of the stem.

To evaluate the accuracy of splines, residuals at each position of measured diameters were determined and five statistics were computed. Diameter bias (DB) shows, how much is the resulting curve shifted against the original data. It helps to determine if a curve systematically under- or overestimates the stem thickness. Mean absolute deviation (MAD) reflects the average distance between the predicted and the original diameters. Standard deviation of residues (SDR) helps to detect heterogeneity in residue values; it may detect oscillations. Mean squared residue (MSR) value discovers locally high deviation of the curve; high values of mean squared residue usually signify oscillations. Total volume difference (TVD) is the difference of total volume derived from the modeled spline curves and total volume derived from the measured profiles. The optimal positions of 7th input point were determined for each stem individually using multi-criteria optimizing method of aggregate objective function with ten criteria. The first five criteria were the medians of above mentioned statistics with weights set as 0.05, 0.05, 0.05, 0.1 and 0.25, respectively. The second five criteria were variances of the same magnitudes with identical weight distribution. The target function value is calculated as the weighted sum of the normed criteria values. The alternative with the lowest value of the target function is considered as the best solution. Using this method the optimal placement for each individual stem was obtained. Three positions with the lowest value of the target functions were considered in the next analyses. Dependency of the three best positions of seventh input point on the position of inserted points was searched using linear regression. With respect to the assumed type of dependence between the optimal positions of additional input point and the positions of the inserted points describing the stem irregularity, the resulting point set was split into two parts and linear regression was computed separately for each part (Fig. 1). The threshold value was determined in such way that the residual sum of squares is minimized.

Suitability of additional input point placement according to the linear regression was reviewed. Several options of the seventh input point were compared. Firstly, the seventh point was placed to the optimal position determined individually for each stem. Secondly, the point

was placed according to the linear regression. Thirdly, the point was placed in the center of the largest input point interspace. Fourthly (in tables labeled as Randomly I), the point was placed randomly near the center of the largest input point interspace (normal distribution with mean value equal to the center of the largest interspace and with such variance that the interspace covers the random point with probability 99.9%). Finally (in tables labeled as Randomly II), the point was placed randomly into one of the two largest interspaces (uniform distribution). As the control variant served spline computed from six points only.

Programs for computation of separate splines as well as for statistical evaluation were written in MATLAB R2012a. For statistical comparison analysis of variance and Tukey's HSD test were used.

Results

Expected relation between optimal position of the seventh point and positions of the inserted points was found only for Catmull-Rom spline. Therefore regression was fitted only for Catmull-Rom spline (Fig. 1). Coefficients of determination of both lines are very low (0.29 and 0.13 respectively), but the regression parameters are statistically proved to be significantly different from zero.

With natural cubic spline and B-spline the location of the optimal input points is in principle independent on the positions of the inserted input points and it does not allow building the regression. Therefore general optimal position of the seventh point for all stems was calculated. It was found that both for natural cubic spline and B-spline the curve is the most accurate if the seventh point is placed in 20% of the stem height. For evaluating different methods of placing the seventh point was the regression replaced by situating the additional point to the relative height 20% of the stem height.

For all splines, according to expectations the lowest error values are found in case that the point is placed to its optimal position. But rarely the errors significantly differ from errors of other variants of point placement.

As shows Tab. 1, the errors of Catmull-Rom splines with the additional point placed according to the linear regression, into the center of largest interspace or randomly near the center are approximately equal. Statistical tests indicate that there mostly are no significant differences between placing the seventh point to its optimal position, according to the regression and

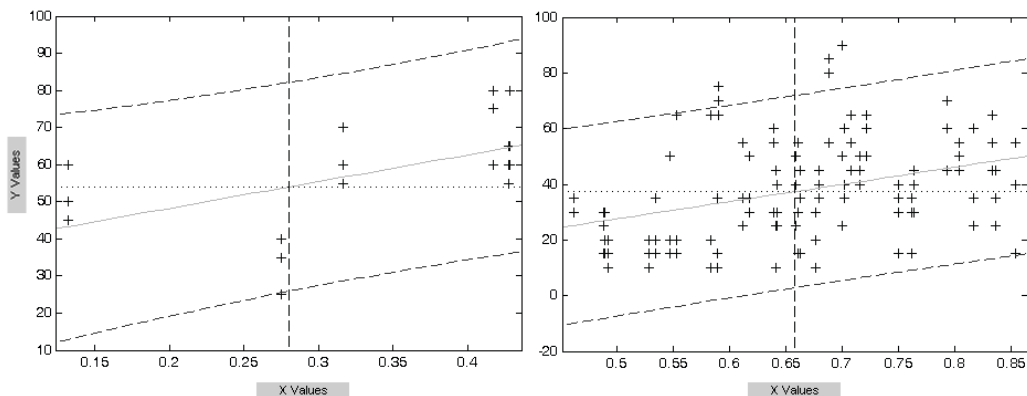


Fig. 1: Regression lines for Catmull-Rom spline. X values denote the relative height of the inserted points describing the stem irregularity. Y values are the optimal position of the seventh input point. Solid line is the regression line, dashed lines denote borders of 95 % confidence region.

Tab. 1: Comparison of variants of placement of the seventh point for Catmull-Rom spline. If values in a column are followed by the same letter, their spline types are not significantly different in the particular aspect.

Variant of 7 th point placement	DB (cm)	MAR (cm)	SDR (cm)	MSR (10 ⁻³ m ²)	TVD (%)
Optimal position	0.001a,b	0.394a	0.523a	0.033a	0.204a
Regression	-0.020a,b	0.465b	0.585a,b	0.045a,b	-0.051a
Center of largest interspace	0.039a	0.467b	0.576a,b	0.046a,b	1.076a
Randomly I (near center)	-0.045a,b	0.483b	0.600b	0.050a,b	-0.305a
Randomly II	-0.042a,b	0.527b,c	0.634b,c	0.063b,c	-0.357a
6-points spline	-0.089b	0.570c	0.665c	0.071c	-0.924a

to the center of the largest interspace. On the contrary, considerably worse are variants with only six points or with the seventh point placed randomly. The 6-point spline and splines with the seventh point placed according the regression and in the center of the largest interspace are shown in Fig. 2. Two characteristic stems with demonstrable irregularities were selected for the figure. The first stem is a representative of stems with a large branching in the middle of the stem height; the second stem illustrates an irregularity in the upper profile.

In natural cubic spline the seventh input point markedly reduces the high errors (Tab. 2) caused by uncontrollable oscillations (Fig. 3). The lowest error values are obtained by placing the seventh point to 20% of the stem height. But for all variants the oscillation is still very strong.

For B-spline the results of the comparison (Tab. 3.) show that placing the additional point to the relative height 20% significantly lowers the diameter bias and the total volume difference if compared with other variants of the additional

point placement. On the contrary, values of MAD, SDR and MSR are higher, in some cases significantly. While placing the point to the relative height 20% improves the shape of the curve in the lower stem, which is the most important for volume estimation, placing the point into the largest interspace between existing points results in more even distribution of input points and decrease of the deviations of the modeled curve from the real data.

Differences between individual spline types are also significant.

Discussion

Oscillations of natural cubic spline caused by its second degree continuity have been already many times reported (eg. Goulding 1979, Lahtinen 1988). Natural cubic spline may be suitable for describing smooth profiles of coniferous trees, but is not able to cope with complicated profiles of broadleaved trees. By this reason Goulding (1979) does not recommend

Tab. 2: Comparison of variants of placement of the seventh point for natural cubic spline. If values in a column are followed by the same letter, their spline types are not significantly different in the particular aspect.

Variant of 7 th point placement	DB (cm)	MAR (cm)	SDR (cm)	MSR (10 ⁻³ m ²)	TVD (%)
Optimal position	-0.765a	2.411a	2.873a	1.354a	5.431a
20% of stem height	-0.507a,b	2.572a	3.070a	1.573a	15.207a
Center of largest interspace	1.005b,c	2.952a	3.797a	2.453a	47.234b,c
Randomly I (near center)	0.938b,c	3.037a,b	3.822a,b	2.418a	49.829b
Randomly II	2.424c	4.176b	4.961b	5.671a	99.684c
6-points spline	4.911d	7.243c	7.782c	14.169b	209.835d

Tab. 3: Comparison of variants of placement of the seventh point for B-spline. If values in a column are followed by the same letter, their spline types are not significantly different in the particular aspect.

Variant of 7 th point placement	DB (cm)	MAR (cm)	SDR (cm)	MSR (10 ⁻³ m ²)	TVD (%)
Optimal position	0.098a,b	0.553a	0.760a,b	0.070a	0.973a,b
20% of stem height	0.033a	0.621b	0.811b	0.086a,b	-0.004a
Center of largest interspace	0.241c	0.610a,b	0.756a	0.079a,b	3.743c
Randomly I (near center)	0.204c	0.618b	0.774a,b	0.082a,b	2.848c
Randomly II	0.181a,b	0.654b	0.807b	0.089b,c	2.684b,c
6-points spline	0.163b,c	0.718c	0.857b	0.104c	2.506b,c

utilization of cubic splines for modeling profiles of malformed stems. Also Lahtinen (1988) recommends lowering the degree of continuity for more complicated profiles.

Also B-spline due to its approximation property is not suitable for complicated stem profiles, although profiles of coniferous trees were well represented using B-spline. It smoothes the driving polygon and therefore it is not able to fit the rapid diameter drop following branching. Good choice is the utilization of Catmull-Rom spline.

Optimal input point combinations proposed by Smaltschinski (1983) or Figueiredo-Filho (1996) are not applicable for malformed profiles. Determination of optimal position individually for each stem in practice is out of the question, because it requires measurement of many diameters from which one is selected as the best. Utilization of regression seems to be useless regarding identical results obtained by placing the additional input point around the center of the longest interspace. A comparison shows, that the average distance between the location determined by the regression and the center of the largest interspace is only 6% of the stem height (median value is only 2% of the stem height), which is a negligible distance. At the same time it

is proved, that input point position determined in this way is not significantly worse than the best position of individual stems.

It can be concluded, that for improving the accuracy the position of the seventh point is important. It should not be placed randomly, but equalizing the point distribution by placing the point in the center of the longest input point interspace can be a reasonable way.

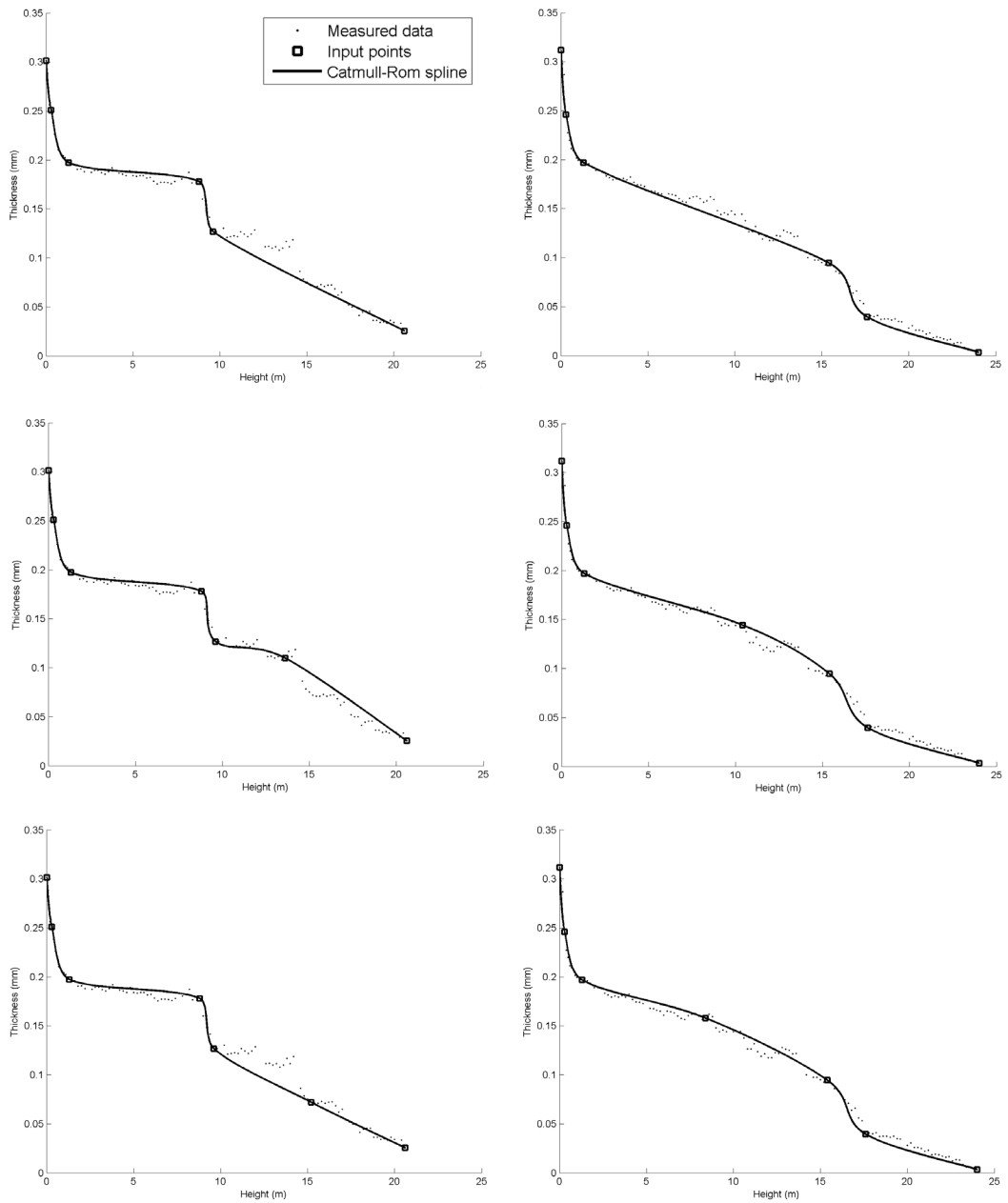


Fig. 2: Two stem profiles modeled by Catmull-Rom spline with 6 input points (top), with the seventh point placed according to the regression (middle) and with seventh point placed in the centre of the longest point interspace (bottom).

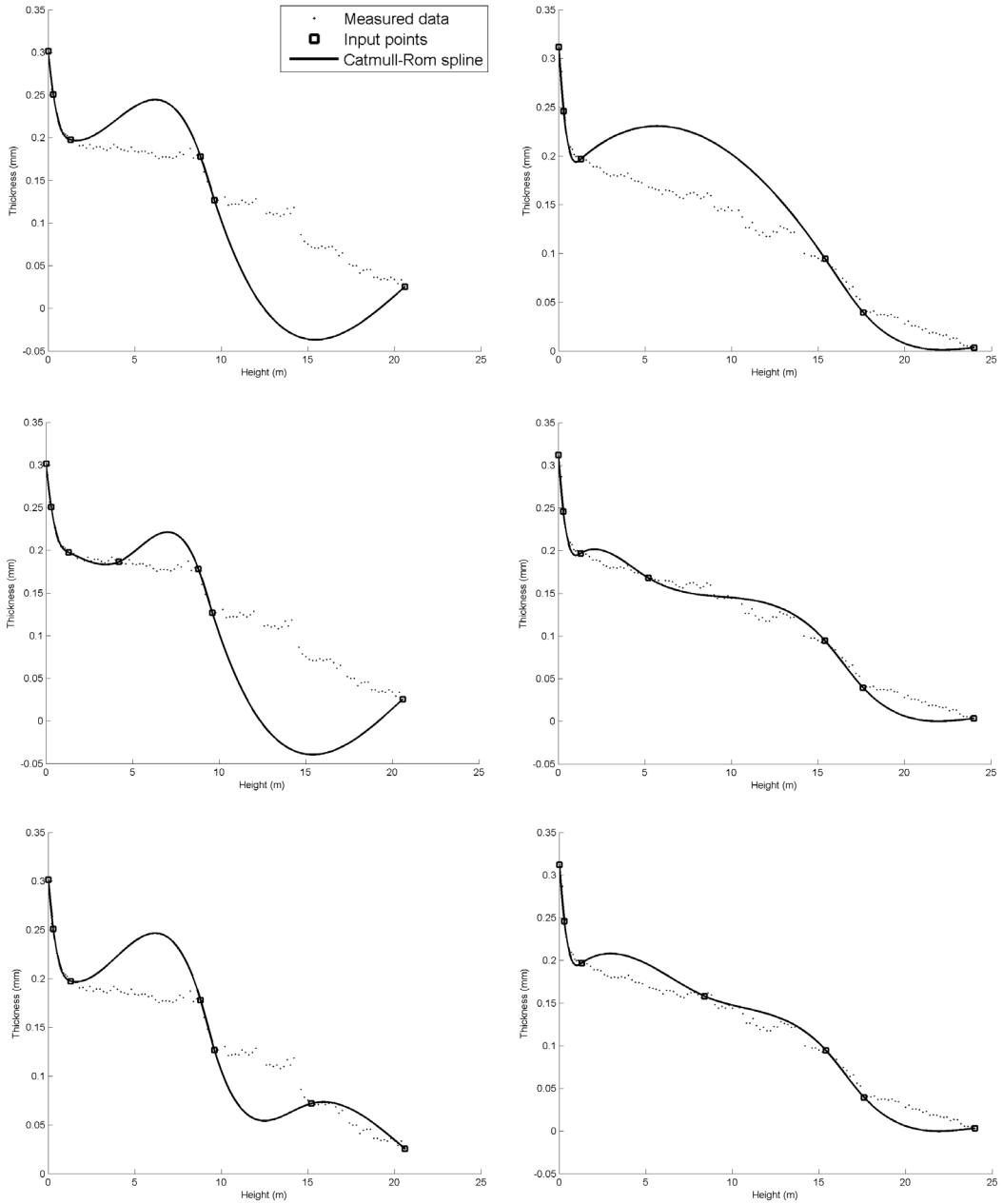


Fig. 3: Two stem profiles modeled by natural cubic spline with 6 input points (top), with the seventh point placed to 20 % of stem height (middle) and with seventh point placed in the centre of the longest point interspace (bottom).

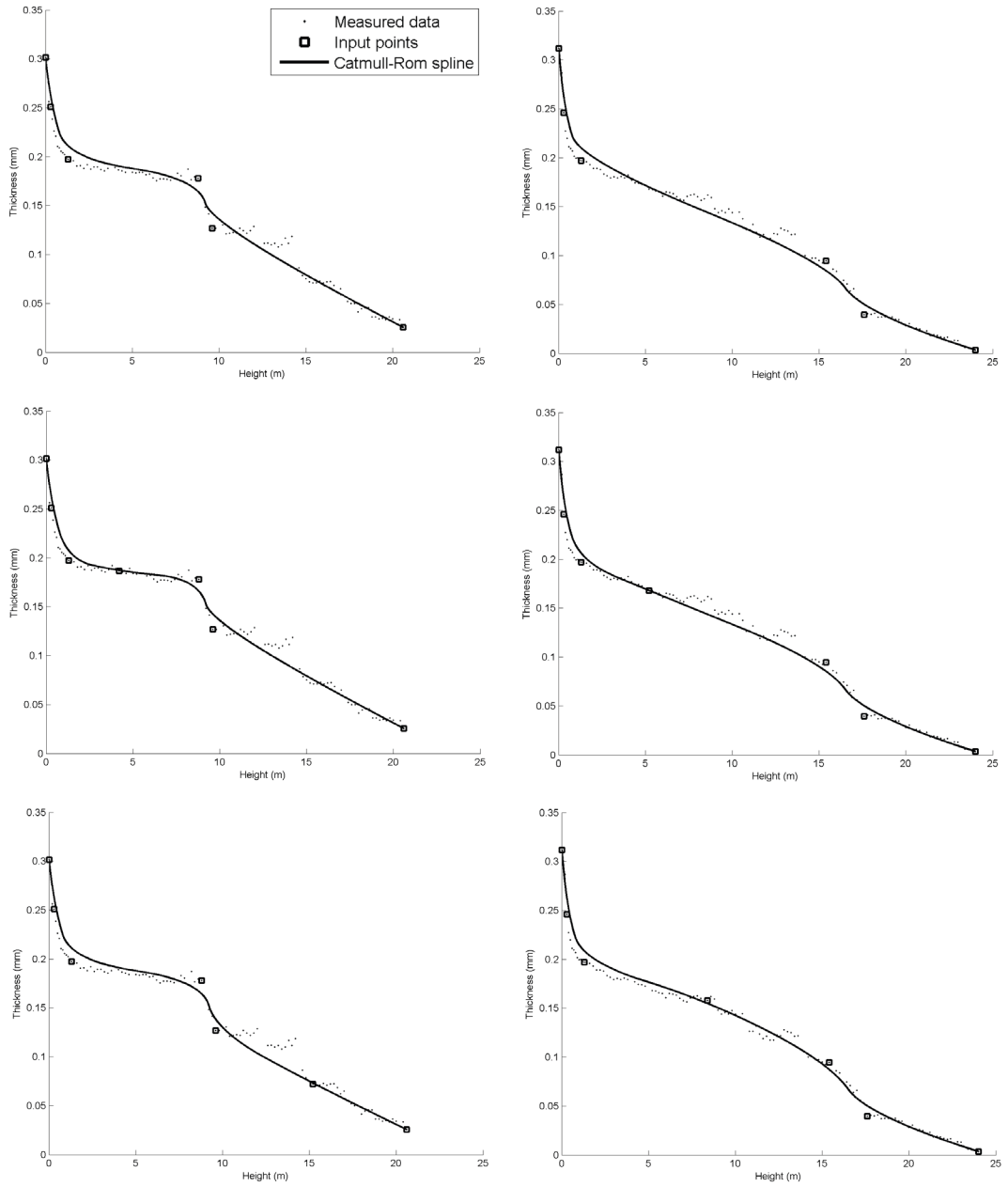


Fig. 4: Two stem profiles modeled by B-spline with 6 input points (top), with the seventh point placed to 20% of stem height (middle) and with seventh point placed in the centre of the longest point interspace (bottom).

Acknowledgment

This research was supported by the project of the Ministry of Agriculture of the Czech Republic No. QI102A079 „Research on biomass of broadleaved species“ and by the project of the University Internal Grant Agency IGA No. 4314013123117 „Development of a regression model of local stem profile using splines“.

References

- BIGING, G. S. 1984: Taper equations for second-growth mixed conifers of Northern California. *Forest Science*, 30: 1103–1117.
- BROOKS, J. R., JIANG, L., OZÇELİK, R. 2008: Compatible stem volume and taper equations for Brutian pine, Cedar of Lebanon, and Cilicica fir in Turkey. *Forest Ecology and Management*, 256: 147–151.
- BRUCE, D., CURTIS, R. O., VANCOEVERING, C. 1968: Development of a system of taper and volume tables for red alder. *Forest Science*, 14: 339–350.
- DEMAERSCHALK, J.P. 1972: Converting volume equations to compatible taper equations. *Forest Science*, 18: 241–245.
- DEMAERSCHALK, J. P., KOZAK, A. 1977: The whole-bole system: a conditioned dual-equation system for precise prediction of tree profiles. *Canadian Journal of Forest Research*, 7 (3): 488–497.
- EERIKÄINEN, K. 2001: Stem volume models with random coefficients for *Pinus kesiya* in Tanzania, Zambia, and Zimbabwe. *Canadian Journal of Forest Research*, 31 (5): 879–888.
- FIGUEIREDO-FILHO, A., BORDERS, B. E., HITCH, K.L. 1996: Number of diameters required to represent stem profiles using interpolated cubic splines. *Canadian Journal of Forest Research*, 26: 1113–1121.
- FLEWELLING, J. W., RAYNES, L. M. 1993: Variable-shape stem-profile predictions for western hemlock. Part I. Predictions from DBH and total height. *Canadian Journal of Forest Research*, 23: 520–536.
- GOULDING, C. J. 1979: Cubic spline curves and calculation of volume of sectionally measured trees. *New Zealand Journal of Forest Science*, 9: 89–99.
- GOULDING, C. J., MURRAY, J. C. 1976: Polynomial taper equations that are compatible with tree volume equations. *New Zealand Journal of Forest Science*, 5: 313–322.
- GRAY, H.R. 1956: The form and taper of forest-tree stems. *Imperial Forestry Institute Paper*, 32. Oxford, University Press, 88 pp.
- GROSENBAUGH, L. R. 1966: Tree form: Definition, interpolation, extrapolation. *The Forestry Chronicle*, 42 (4): 444–457.
- JIANG, L., BROOKS, J. R., WANG, J. 2005: Compatible taper and volume equations for yellow-poplar in West Virginia. *Forest Ecology and Management*, 213: 399–409.
- KITIKIDOU, K., CHATZILAZAROU, G. 2008: Estimating the sample size for fitting taper equations. *Journal of Forest Science*, 54 (4):176–182.
- KOCHANEK, D. H. U., BARTELS, R. H. 1984: Interpolating splines with local tension, continuity, and bias control. *SIGGRAPH Computers & Graphics* 18: 33–41.
- KOSKELA, L., NUMMI, T., WENZEL, S., KIVINEN V. P. 2006: On the analysis of cubic smoothing spline-based stem curve prediction for forest harvesters. *Canadian Journal of Forest Research*, 36: 2909–2919.
- KUŽELKA, K. 2011a: Stem curve modeling using spline functions. In: *Bulletin of Szent István University, Gödöllő 2011, Hungary, (special issue). April 28th–29th, 2011, ISSN 1586-4502*, p. 211.
- KUŽELKA, K. 2011b: Stem curve modeling using spline functions. In: *ICFFI News. International Centre of Forestry and Forest Industries*, 1 (13): 70.
- LAASASENAHO, J., MELKAS, T., ALDÉN, S. 2005: Modelling bark thickness of *Picea abies* with taper curves. *Forest Ecology and Management*, 206: 35–47.
- LAHTINEN, A., LAASASENAHO, J. 1979: On the construction of taper curves by using spline functions. *Communications Instituti Forestalis Fenniae*, 95 (8): 1–63.
- LAHTINEN, A. 1988: On the construction of monotony preserving taper curves. *Acta Forestalia Fennica*, 203: 1–34.
- LEE, W. K., SEO, J. H., SONM Y. M., LEE, K. H., VON GADOW, K. 2003: Modeling stem profiles for *Pinus densiflora* in Korea. *Forest Ecology and Management*, 172: 69–77.

- LI, R., WEISKITTEL, A., DICK, A. R., KERSHAW, J. A., AND SEYMOUR, R. S. 2012: Regional stem taper equations for eleven conifer species in the Acadian Region of North America: Development and assessment. *Northern Journal of Applied Forestry*, 29: 5–14.
- LINKEOVÁ, I. 2007: NURBS křivky. *NeUniformní Racionální B-Spline křivky* [NURBS curves: Non-Uniform Rational B-Spline curves]. Praha, Nakladatelství ČVUT, 208 s.
- LIU, C. J. 1980: Log volume estimation with spline approximation. *Forest Science*, 26: 361–369.
- MATTE, L. 1949: The taper of coniferous species with special reference to Loblolly pine. *The Forestry Chronicle*, 25: 21–31.
- MAX, T. A., BURKHART, H. E. 1976: Segmented polynomial regression applied to taper equations. *Forest Science*, 22: 283–289.
- NUMMI, T., MÖTTÖNEN, J. 2004: Prediction of stem measurements of Scots pine. *Journal of Applied Statistics*, 31: 105–114.
- PIEGL, L., TILLER, W. 1996: *The NURBS Book*. Springer-Verlag, Berlin, second edition, 646 pp.
- SHARMA, M., PARTON, J. 2009: Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. *Forest Science*, 55 (3): 268–282.
- SMALTSCHINSKI, T. 1983: Individuelle Baumstammform und cubische Spline Interpolation [Individual stem form and cubic spline interpolation]. *Allgemeine Forst und Jagdzeitung*, 155: 193–197.
- THOMAS, C. E. AND PARRÉSOL, B. R. 1991: Simple, flexible, trigonometric taper equations. *Canadian Journal of Forest Research*, 21: 1132–1137.