Don't blame the tree volume equations

Richard Woollons

When predicted and actual clearfall volume per hectare estimates differ appreciably in a region, it is common for the applicable tree volume equation to be blamed for the discrepancies. Here it is argued that provided the volume function is well constructed and stands are adequately sampled, then total volume per hectare can be confidently estimated to within five per cent – too small an error to account for the outcome of many reconciliation studies.

The problem

An important study for many forest organisations is a comparison of predicted and realised volume yields. Estimates are made of net predicted volume per hectare as estimated from inventories or growth models, and the corresponding realised volume assessed by weighbridge tonnes converted to a volume estimate. There is a very reasonable question in relation to management – to what extent do the two figures agree? The trouble begins, at least for forest researchers responsible for the volume equations construction, when two estimates differ substantially.

It is the author's experience that immediate suspicion is almost always cast at the first link in the chain – the tree volume equation. Ironically, the problem should simplify when the two estimates are far apart. Differences of say 15 to 20 per cent can only be caused by major and usually obvious factors. Nevertheless, the equation is frequently considered suspect *per se*. Alternatively doubts are expressed as to whether the function can perform adequately over wide areas or in slightly unusual situations, for example, in high altitude stands or perhaps for pruned trees.

In this article one volume equation is discussed. The model is shown to possess a very high precision and is accurate over several regions. Its usage with pre-harvest inventories is examined. Two sources of error are recognised: (1) a prediction error associated with the volume equation *per se*; and (2) a sampling error associated with laying out inventory plots, then estimating total stem volume per hectare. These errors are summarised and discussed.

The volume equation

The data from which the volume equation is constructed was first acquired by the then Carter Holt Harvey Forests in 1995. Sectional log measurements were taken from 684 trees in forests extending from South Auckland (Athrenree, Maramarua and Tairua forests), Riverhead and Woodhill forests, North Auckland (Maharangi and Mangawhai forests) up to Northland (Whangarei and Dargaville forests). Riverhead and Woodhill are retained as regions in that they represent singular forests with historical nutrition problems and are reputed to have unusual stem forms. In 2004, 152 additional trees were measured in adjoining forests that up till now had not been sampled.

The sectional measurement methodology adopted followed that of Whyte (1971). Measurements were secured at 0.2, 0.7 and 1.4 metres, thereafter in 40 to 150 centimetre taper steps. Sectional diameters, over-bark and inside-bark were taken, the latter after stripping the bark. Sectional volumes were estimated through the conoid formula:

$$v = (\pi / 12)\, l\, (d_1^2 + d_2^2 + d_1 d_2) \qquad (1)$$

where in (1)

v = total stem volume, inside bark (m^3)

l = log length (m)

d_1, d_2 = large-end, short-end diameter (m)

Note that when short log lengths are utilised, it becomes almost irrelevant what sectional volume equation is assumed. Whyte (1971) Table 1 summarises the data.

Table 1. Summary of volume equation data

Volume(m^3) / Region	trees	mean	min.	max.
Northland	142	0.78	0.02	3.72
North Auckland	230	0.93	0.01	4.70
South Auckland	229	1.16	0.01	4.98
Riverhead	114	1.18	0.02	4.11
Woodhill	139	1.25	0.01	4.17

A set of candidate volume equations is summarised by Clutter et al. (1992). The models come down to variations of two basic forms:

$$v = \beta_0 + \beta_1 d^2 h + \beta_2 d^2 + \beta_3 h \qquad (2)$$

and

$$v = \beta_0 + \beta_1 d^{\beta 2} h^{\beta 3} \qquad (3)$$

where in (2) and (3)

v = total stem volume, inside-bark (m^3)

d = diameter breast height (cm)

h = total stem height (m)

β_i = model coefficients.

Several variants of equations (2) and (3) were tried. The diameter squared term in (2) was found to be totally non-significant and discarded. Ultimately, three models were considered:

$$v = \beta_0 + \beta_1 d^2 h \qquad (4)$$

$$v = \beta_0 + \beta_1 d^2 h + \beta_2 h \qquad (5)$$

and model (3) above.

The respective residual sum-of-squares for the three models are:

(3) 8.34 (4) 8.62 and (5) 8.35.

Inclusion of the height variable in (5) relative to (4) improves the precision by three per cent, but models (3) and (5) are essentially equivalent. Because the latter model is simpler to examine by virtue of being a linear form, model (5) was chosen. Inspection of the residual values by region shows that equation (5) is essentially unbiased. Figure 1 gives a histogram of the residual values broken down by region.

Equation (4) assumes that the tree form factor (tree volume/tree basal area x stem height) is a constant. For older trees this is plausible, but for young trees the form factor increases rapidly and (4) cannot therefore give an ideal fit. Figure 2 shows a plot of tree form factor and tree height, and the rise in form factor for heights less than around 10 metres is appreciable. This accounts for the need for the additional height term in equation (5).

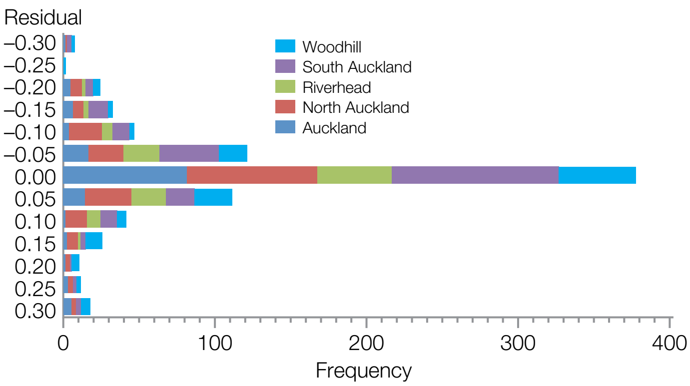

Figure 1. Residuals from equation (5) by region

Volume function prediction error

By most criteria volume function (5) is an excellent model, but nevertheless it is still subject to a prediction error. This is because the three coefficients β_0, β_1 and β_2 are estimated from the sample data by least-squares techniques so any prediction will have an error. It cannot be assumed, for example, that the relationship between inside-bark volume and the function of diameter and height is entirely free of bias. An exact formula for the predicted errors is available (Draper & Smith, 1998). These increase for data with diameter and height values progressively further away from the mean diameter/height centroid ($\bar{d}$, $\bar{h}$).

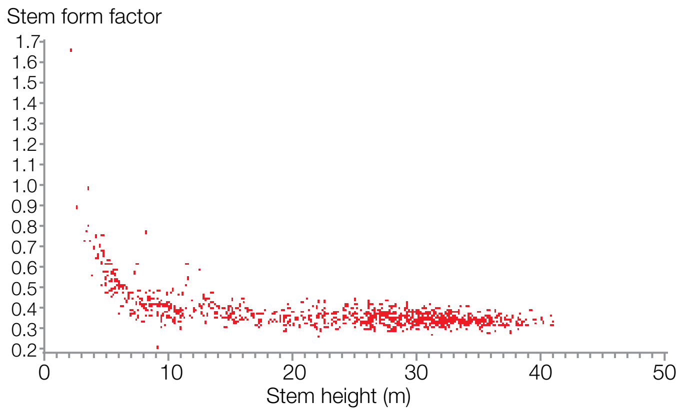

Figure 2. Tree form factor and stem height

For this data the average volume is 1.04 cubic metres, but in this study the equation is utilised to estimate clearfall volumes so it is appropriate to estimate the prediction error in relation to these (larger) volumes. For volumes of 2.00 cubic metres the average predicted error (sp) is ±0.005 cubic metres, equivalent to 0.25 per cent. Thus, the prediction error is trivial.

Inventory sampling error

To obtain estimates of predicted volumes per hectare in clearfall stands it is necessary to sample the stands through pre-harvest inventories. Usually plots are established via a grid, all trees are measured for breast height diameter, and a sub-sample of trees is measured within each plot to produce a height-diameter (breast height) regression. This is so all stems can be assigned a height and thus tree volumes can be estimated through a volume equation.

Given that sufficient trees, say 15 to 20, are sampled for height and they are chosen across the diameter range, then although the precision of any regression is not especially high (Woollons, 2003), it is a reasonable assumption that on average the tree heights are measured without any appreciable bias. Conversely, since only a very small proportion of any stand is actually surveyed, the prediction of the stand volume per hectare must be associated with a significant sampling error. If the layout of the plots is systematic, as it commonly is, then strictly no error formula exists (Shiver and Borders, 1996), but in practice the formula appropriate for simple random sampling is often substituted:

$$s_{\bar{y}} = \frac{s}{\sqrt{n}}\sqrt{\left(1 - \frac{n}{N}\right)} \qquad (6)$$

where in (6)

$s_{\bar{y}}$ = standard error of the mean

n = sample size

N = population size

s = standard deviation of the sample data.

Precision associated with (6) is a function of the population size, the variability of the response size, and critically the sample size. Stand areas can vary from one to two hectares out to several hundred hectares so with plot sizes usually 0.04 to 0.08 hectares, the contribution of the fraction (1– n/N) is usually negligible. However, the standard deviation can vary substantially.

To illustrate the use of (6), data was obtained from 11 Northland region *Pinus radiata* stands. An average of 28 (eight to 42) inventory plots (0.04 to 0.08 hectares) were established in stands ranged from six to 56 hectares. The data is summarised in Table 2.

Table 2. Summary of inventories

Inv	Area (ha)	Ave. V/ha	Std. err.	C.V. %	No. plots
1	34	775	52	7	29
2	27	791	37	5	27
3	45	686	40	6	37
4	36	718	30	4	35
5	19	685	27	4	25
6	15	794	32	4	16
7	28	622	26	4	32
8	29	633	25	4	30
9	23	741	40	5	25
10	6	661	62	9	7
11	36	514	17	3	43

Substituting the relevant values for each stand in (6) resulted in an average standard error of ±35 cubic metres per hectare. The average total stem volume per hectare is 690 cubic metres per hectare, equivalent to a sampling error of 5.1 per cent.

Discussion and conclusion

The analyses clearly demonstrate that the error associated with applying a suitable volume function to tree diameter/height data is very small – 0.25 per cent. The function utilised here operates satisfactorily over widespread regions, including Riverhead and Woodhill, two forests that anecdotally have unusual tree form. The dataset was well stratified by age and silviculture, including pruning, so a premise that the volume function could prove inaccurate in unusual conditions is unfounded. The sampling error is considerably higher, around five per cent, but considering the inherent variability in older radiata pine stands and the low sampling fraction usually utilised this is not an excessive figure.

These analyses substantiate that total volume per hectare can be estimated in radiata pine stands with an error of no more than around five per cent. When considering estimates of realised volume per hectare, a complete analysis would include the utilised breakage function and the adopted realisation factor, but the major contributor by far is the volume equation. For example, the break-point of the tree is usually around two-thirds of total height, so the volume to break height accounts for around 90 per cent of total volume.

When differences between realised and predicted volumes are at 20 per cent, management need to consider other factors aside from the volume function. Two candidates would be major errors when calculating net stocked areas or not recognising that the clear-fell yields are being transported to more than one outlet.

Acknowledgement

The data on which the volume function is based was supervised and collected by Mr Lindsay McKendrick, then senior technical officer of Carter Holt Harvey Forests, Northland. Subsequent editing of the data showed the entire set was essentially free of any substantial field errors.

References

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H. and Bailey, R.C. 1992. *Timber Management: A Quantitative Approach*. Krieger Publishing Company, Malabar, Florida, US.

Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis* (3rd Edn). New York: John Wiley and Sons.

Shiver, B.D. and Borders, B.E. 1996. *Sampling Techniques for Forest Resource Inventory*. New York: John Wiley and Sons.

Whyte, A.G.D. 1971. Sectional Measurement of Trees: A Rationalised Method. *New Zealand Journal of Forestry Science*, 1: 74–80.

Woollons, R.C. 2003. Examination of Mean Top Height Definitions and Height Estimation Equations for *Pinus Radiata* in New Zealand. *New Zealand Journal of Forestry*, 48: 15–18.

Richard Woollons is a semi-retired forest statistician with the School of Forestry, University of Canterbury. Before this he worked at Tokoroa, initially with NZ Forest Products and then Carter Holt Harvey Forests. Email: richard.woollons@canterbury.ac.nz.