# Modelling height-diameter relationships of *Pinus radiata* plantations in Canterbury, New Zealand

**Weizhong Zhao[1,3], Euan G. Mason[1] and Jennifer Brown[2]**

**Abstract**

This paper describes the modelling of height-diameter relationships for *Pinus radiata* D. Don at stand and individual tree levels in Canterbury, New Zealand. Sixteen functional forms were evaluated for stands that varied considerably in stand age, site index, altitude and the number of trees sampled. The Petterson equation with exponent -5 and the two-parameter Richards' equation both gave the smallest average mean square error. The inclusion of stand age, site index and altitude into the Petterson equation reduced mean square error by 72% in the regional height-diameter model.

**Keywords: height-diameter equations; stand and regional level; *Pinus radiata*.**

## Introduction

Height-diameter relationships play an important role in forest mensuration systems. Stand mean height, mean top height (height, estimated with a height vs diameter curve, of the 100 trees with the largest diameters in a one-hectare area), and mean height for each diameter class, are usually calculated directly from equations of tree height (h) on diameter at breast-height outside of bark (d). Volume per hectare is an accumulation of individual tree volumes often derived from a two-dimensional function using measured d and estimated h.

The best attributes for height-diameter equations are for the equations to pass through the x-axis at $h$=1.40 m and to attain an asymptote for large values of diameter and always have a positive slope ( Curtis, 1967; Garman *et al.*, 1995). The Petterson equation (Schmidt, 1967)

$$h = 1.40 + (\alpha + \beta/d)^{(-2.5)}$$

is an example of a height-diameter equation with these properties. Although mathematically the parameters could adopt negative values, this is very unlikely given a set of real data. The estimated height, $h$ is 1.40 m when $d$ approaches zero; the asymptote is $1.40 + \alpha^{-2.5}$ when $d$ approaches + infinity; and the derivative of the curve is $2.5\beta/[d^2(\alpha + \beta/d)^{-3.5}]$, which is always greater than zero because $d$, $\alpha$ and $\beta$ would be all positive.

Linearity is another desirable property of height-diameter equations especially in permanent sample plot (PSP) database computational systems because coefficients can be solved explicitly and uniquely with simple algorithms. Early height-diameter equations were forms of linear, and linearized, equations (Henriksen, 1950; Myers, 1966; Curtis, 1967), and many of these are still in use today. The Petterson equation can be easily transformed into linear form and has been widely used in New Zealand (McEwen, 1978, Goulding, 1995, Woollons 2003). As the power of computers develops a number of non-linear equations have been tried because they are often more flexible than linear equations (Larsen and Hann, 1987; Wang and Hann, 1988; Arabatzis and Burkhart, 1992; Huang *et al.*, 1992; Dolph *et al.*, 1995).

There are many published height-diameter equations and sixteen are described in Table 1. All of the equations in Table 1 have positive slopes and all but equations 11 and 12 pass through the x-axis at h=1.4. All equations except 9 and 10 have an asymptote at large diameters. Although none of the equations is linear, equations 1 to 6 can be transformed to a linear form. Equations 13 to 16 are three-parameter forms, and all others contain only two parameters.

Height-diameter equations can be used to describe the growth relationship at a plot (Curtis, 1967; Garcia, 1974), stand, or a regional level where it is used to predict individual tree heights (Larsen & Hann 1987; Wang & Hann 1988, Huang *et al.*, 1992). A stand is a basic forestry unit of continuous area within which site conditions and management regimes are relatively uniform and it is common practice in some forestry companies that height-diameter equations be estimated at the stand level. Therefore, robust equations which fit well under a wide range of stands are desirable.

When a regional height-diameter model is used usually height measurements are not required and heights are predicted by identifying and incorporating relevant factors. For example, the relationship between height and diameter in a region may vary with tree species, stand age, site characteristics, genetics, stocking, and silvicultural treatment (Trorey, 1932; Buford, 1986; Larsen and Hann, 1987; Wang and Hann, 1988; Dolph, 1989; Knowe, 1994).

The objectives of the study were to estimate height-diameter equations that were robust at the stand-level, and to predict tree height from auxiliary variables for estimating a suitable regional model for radiata pine grown in Canterbury, New Zealand.

## Methods

### *Data*

The data used in the study were PSP measurements within Selwyn Plantation Board Limited's (SPBL) estate. The estate includes approximately 10,000 ha of stocked plantations 90% of which is comprised of even-aged radiata pine stands (SPBL, 1998). Plantations cover plains, hills and coastal sands between the Waimakariri and Rakaia Rivers, Canterbury. Altitudes of plantations range from 2 to 600 metres above sea level. Site indices (mean top height

[1] *School of Forestry, University of Canterbury Private Bag 4800, Christchurch, New Zealand*

[2] *Biomathematics Research Centre, University of Canterbury Private Bag 4800, Christchurch, New Zealand*

[3] *Weizhong Zhao now works for Intercontinental Information Science, Eli Lilly Australia.*

*Table 1: Height-diameter equations and example references*

| No. | Equation | Example Reference |
|---|---|---|
| 1 | $h=bh+\alpha\, d/(\beta+d)$ | Bates & Watts 1980; Huang ***et al*. 1992** |
| 2 | $h=bh+d^2/(\alpha+\beta d)^2$ | Huang *et al*. 1992 |
| 3 | $h=bh+(\alpha+\beta/d)^{-2.5}$ | Schmidt 1967 |
| 4 | $h=bh+(\alpha+\beta/d)^{-5}$ | Modified with the above |
| 5 | $h=bh+(\alpha+\beta/d)^{-8}$ | Modified with the above |
| 6 | $h=bh+\alpha(1-\exp(\beta d))$ | Meyer 1940; Huang *et al*. 1992 |
| 7 | $h=bh+\alpha\,(1+1/d)^{(-\beta)}$ | Curtis 1967; Huang *et al*. 1992 |
| 8 | $h=bh+\exp(\alpha+\beta/d)$ | Schumacher 1939; Curtis 1967; Buford 1986; Arabatzis & Burkhart 1992; Huang *et al*. 1992 |
| 9 | $h=bh+\alpha(\ln(1+d))\beta$ | Modified from an equation in Curtis 1967 |
| 10 | $h=bh+\alpha d\beta$ | Arabatzis & Burkhart 1992 |
| 11 | $h=\alpha+\beta\exp(-0.08d)$ | Garcia 1974 |
| 12 | $h=\alpha+\beta\,(d+10)^{-1}$ | Garcia 1974 |
| 13 | $h=bh+\alpha(1-\exp(\beta d))\gamma$ | Richards 1959; Huang *et al*. 1992; Garman 1995 |
| 14 | $h=bh+\alpha(1-\exp(\beta d\gamma))$ | Yang *et al*. 1978; Huang *et al*. 1992 |
| 15 | $h=bh+\alpha/(1+\beta^{-1}\, d^{-}\gamma)$ | Huang *et al*. 1992 |
| 16 | $h=bh+\exp(\alpha+\beta\, d^{-}\gamma)$ | Larsen & Hann 1987; Flewelling & De Jong 1994; Lappi 1996 |

Note: h = height of trees in m, bh = breast height (1.40 m in NZ); d = diameter at breast height in cm; $\alpha$, $\beta$, $\gamma$ = parameters in the equations.

*Table 2: Fit of height-diameter equations*

| No | Equation form | Stand level | | |
|---|---|---|---|---|
| | | Average MSE | Std.dev | Rank |
| 1 | $h=bh+\alpha\, d/(\beta+d)$ | 1.3750 | 1.3012 | 10 |
| 2 | $h=bh+d^2/(\alpha+\beta\, d)^2$ | 1.3695 | 1.2941 | 8 |
| 3 | $h=bh+(\alpha+\beta/d)^{-2.5}$ | 1.3689 | 1.2935 | 5 |
| 4 | $h=bh+(\alpha+\beta/d)^{-5}$ | 1.3684 | 1.2933 | 2 |
| 5 | $h=bh+(\alpha+\beta/d)^{-8}$ | 1.3685 | 1.2936 | 3 |
| 6 | $h=bh+\alpha(1-\exp(\beta d))$ | 1.3670 | 1.2930 | 1 |
| 7 | $h=bh+\alpha(1+1/d)^{(-\beta)}$ | 1.3686 | 1.2938 | 4 |
| 8 | $h=bh+\exp(\alpha+\beta/d)$ | 1.3689 | 1.2945 | 5 |
| 9 | $h=bh+\alpha(\ln(1+d))\beta$ | 1.3870 | 1.3174 | 11 |
| 10 | $h=bh+\alpha d\beta$ | 1.4024 | 1.3404 | 12 |
| 11 | $h=\alpha+\beta\exp(-0.08\, d)$ | 1.3694 | 1.3099 | 7 |
| 12 | $h=\alpha+\beta\,(d+10)^{-1}$ | 1.3704 | 1.2944 | 9 |
| 13 | $h=bh+\alpha(1-\exp(\beta\, d))\gamma$ | - | - | - |
| 14 | $h=bh+\alpha(1-\exp(\beta d\gamma))$ | - | - | - |
| 15 | $h=bh+\alpha/(1+\beta^{-1}\, d^{-}\gamma)$ | - | - | - |
| 16 | $h=bh+\exp(\alpha+\beta\, d^{-}\gamma)$ | - | - | - |

*- At the stand level, too few data were available in several stands to allow convergence in three-parameter equations 13 to 16. No compatible result is available for comparison.*

*Table 3: Regression result of parameter estimates vs. explanatory variables*

| | | Parameter $\alpha$ ($R^2$**=0.83**) | | | | Parameter $\beta$ ($R^2$=0.10) | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | DF | Coefficient | Std. Error | T | Prob>\|T\| | Coefficient | Std. Error | T | Prob>\|T\| |
| Intercept | 1 | 0.80979 | 0.04958 | 16.34 | 0.0001 | 0.26516 | 1.09710 | 0.242 | 0.8093 |
| Age | 1 | 0.58063 | 0.11016 | 5.271 | 0.0001 | -0.90022 | 2.43785 | -0.369 | 0.7124 |
| Altitude | 1 | 0.00001 | 0.00001 | 1.312 | 0.1915 | 0.00086 | 0.00025 | 3.406 | 0.0008 |
| Site Index | 1 | -0.13844 | 0.01257 | -11.01 | 0.0001 | 0.34307 | 0.27826 | 1.233 | 0.2194 |
| Stocking(N) | 1 | -0.00004 | 0.00004 | -1.101 | 0.2726 | -0.00028 | 0.00082 | -0.345 | 0.7302 |
| Age x N | 1 | 0.00017 | 0.00015 | 1.118 | 0.2652 | 0.00053 | 0.00338 | 0.157 | 0.8756 |
| Error | 161 | | | | | | | | |

at age 20 years) range from 15 to 28 m. When stands were measured stockings were mostly between 400 and 1000 stems per hectare (*stems/ha*) but a few stands had stockings as high as 2250 stems/ha.

An inventory system of permanent sample plots was established by the company during the 1960's and re-measurements have been carried out regularly over the last 30 years. Stand ages included in the PSP database used were between eight and 29 years.

For this study we used data from 168 stands. Each stand had been measured approximately three times, at 3 yearly intervals. In total 168 stands were measured three times, resulting in 529 separate surveys. Stands were measured by multiple plots within each stand. The number of plots per stand ranged from 1 to 41. Plots were mostly 0.04 ha in size, but some were 0.02 ha. Diameters were measured on all trees within the plots but heights of a sample of eight trees per plot were measured at any given re-measurement. In total there were data from 24790 trees. Trees selected for height measurements were chosen to cover the range of the tree diameters in the plot. Other variables measured were age, site index stems/ha, altitude. Data from plots within stands were aggregated to provide data at a stand-level

### *Data analysis*

The sixteen height-diameter equations (Table 1) were fitted to data for each stand and measurement time to select the best model. Data were fitted using using non-linear procedures of the SAS software (SAS Institute Inc., 1990). Mean square error (MSE) was used as a measure of fit. Plots of residuals were examined for bias and residuals were tested for normality using proc UNIVARIATE is SAS software. A lower mean square error (MSE) suggests a better fit in terms of minimum variance. Although the individual MSE for each equation will vary from stand to stand we considered a good equation should fit well on average, i.e., be robust over a range of stands and a range of ages. We used the average of the mean square errors as a measure of overall performance of an equation. The average MSE for all equations were compared and ranked in the study to select the best equation.

To predict tree heights from auxiliary variables for use in the regional model we selected the best equation identified in step 1 described above and regressed the estimated coefficients from fits for individual measurement times against stand age ($T$), site index ($SI$), stocking (stems/ha) and altitude ($Alt$) using least squares regression. For this analysis we used data from only one survey of each stand. The use of one data-set per stand avoids potential problems in underestimating the regression variance from taking repeated samples from sample units (Neter and Wasserman, 1974; West *et al.*, 1984; West, 1995). The data-set for each stand was randomly chosen from the repeat surveys of the stand to avoid bias.

A regional height-diameter model was estimated by using the auxiliary variables as predictors of the parameters in the best equation. Data for this model building was the individual tree measurements.

### Results

The average MSE were similar for all the equations where data were fitted. The average MSE ranged from 1.3670 for equation 6 to 1.4024 for equation 10 (Table 2). For four of the equations, equation 13 to 16, convergence was not possible because there were too few data for some stands. Equations 13 to 16 were three parameter equations.

The equation with the lowest average MSE was equation 6. The equation with the second smallest average MSE was equation 4. However, equation 4, an equation that is a member of the Petterson equation family can be transformed to a linear form. Given the small difference between the MSE of equation 6 and 4, and that equation 4 has the desirable property of being able to be linearised we chose equation 4 as the best model for stand-level data. The linear form of equation 4 is:

$Y = \alpha d + \beta$, where $Y = \dfrac{d}{(h-1.40)^{0.2}}$

The linear regression equations using auxiliary stand variables to predict the parameter estimates for the linear form of equation 4 were:

$\alpha = a_0 + a_1 T^{-0.5} + a_2 Alt + a_3 \ln(SI) + a_4\, spha + a_5\, T{\cdot}spha$

and

$\beta = b_0 + b_1 T^{-0.5} + b_2 Alt + b_3 \ln(SI) + b_4\, spha + b_5\, T{\cdot}spha$. (Table 3)

The auxiliary stand variables, stand age and site index, were significantly related to the $\alpha$ parameter estimates. The parameter $\alpha$ increases with increasing stand age and decreasing site index. Overall the auxiliary variables explained 83% of the variation in the parameter $\alpha$ ($R^2$ = 0.83) The ability to predict the parameter $\beta$ was poor ($R^2$ = 0.10) and only altitude was significantly related with $\beta$.

The auxiliary site variables; stand age, site index and altitude were incorporated into equation 4 by using them to predict the parameters $\alpha$ and $\beta$. The non-linear of equation 4 was used because non-linear procedures can often lead to a smaller MSE and avoid back-transforming bias in prediction. The model with the auxiliary variables was:

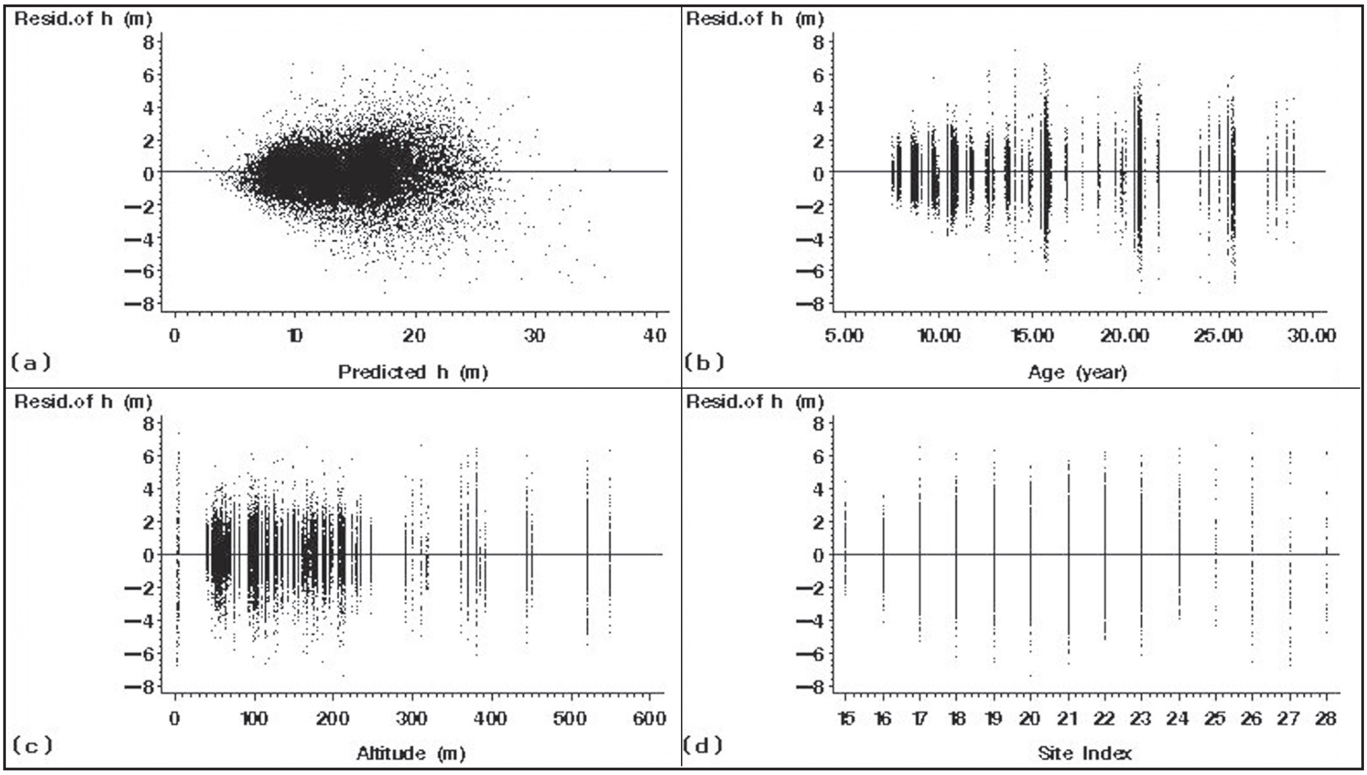
$h = 1.40 + [0.695955 + 0.666983\, T^{-0.5} - 0.106771\, \ln(SI) + (0.954201 + 0.000741 Alt)/d]^5$

When tree diameter alone was used to estimate tree height (i.e., the non-modified form of equation 4) the MSE was 6.1394. A reduction of 50.9% of MSE was gained when age was introduced. MSE reduced a further 39.5% when site index was added. Altitude was less important but statistically significant, further reducing MSE by 6.4%. The final MSE with the three auxiliary variables was 1.7052. There was no apparent bias in the residuals (Figure 1). Ninety-eight percent of residuals were within ±3.5 m and 90% were within ±2.0 m.

### Discussion

Of the sixteen height-diameter equations evaluated there was little difference in average MSE at a stand level. The two-parameter Richards' equation (equation 6) and the Petterson equation with exponent -5 (equation 4) gave the

*Figure 1: Residual pattern of the regional height equation (21) with full data*
*(a) Residual vs. predicted value;* *(b) Residual vs. age;*
*(c) Residual vs. altitude;* *(d) Residual vs. site index*




smallest average MSE for stands varying considerably in age, site index, altitude and the number of trees sampled. The latter equation form, in particular, was chosen as the best because it has the desired properties of passing through the x-axis at when height=1.4 m, approaching an exponent, having a positive slope, being transformable to a linear form and leading to a small fitted mean square error. Three parameter equations offered limited benefits for many stands at huge computational expense and failed to converge when fitting for a few stands. Although the exponent of 5 worked best, the fit was only slightly better than with a more commonly used exponent of 2.5, a finding that agrees with that of Woollons (2003).

Repeated measures of 168 stands were used for comparing equations, but not for estimating effects of site and stand variables on coefficients. This may have led to apparently better fits to equations, but it did not, however, affect the relative comparisons of fits between equation forms.

The use of auxiliary stand variables, stand age, site index altitude reduced the MSE of the chosen height-diameter equation. With this approach no samples of height measurements are needed when using the model. The inclusion of the three auxiliary variables resulted in a reduction of MSE of 72%. It was expected that 90% of predicted tree heights should be within ±2.0 m.

The final model (equation 17) reflects the relationships of radiata pine within SPBL's estate. It is likely that the equation is local, and future studies should be aimed at using hybrid modelling techniques to provide a more generally applicable height versus diameter equation.

We recomend the Petterson family of equations for height-diameter modelling. The advantages of the equations are that they provide relatively unbiased fits to height vs diameter relationships in stands subject to competition and they can be fitted it linear form, thus facilitating mechanical height vs diameter estimation in large databases. Until now their use has been restricted and their robustness for modelling has been infrequently acknowledged.

## Acknowledgements

We are indebted to Dr C. J. Goulding and Mr A. Gordon in NZ Forest Research Institute for their indications of references and detailed descriptions of Petterson equation application in New Zealand. Acknowledgements are also expressed to Mr W. P. Studholme and Mr H. W. Stevenson in SPBL for the use of their database. Dr Helge Dzierzon and Mr Horacio Bown, the referees of this paper, and Dr A. G. D. Whyte provided valuable comments.

## References

Arabatzis, A.A., Burkhart, H.E., 1992, An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations, *Forest Science* 38: 192-198

Bates, D.M., and Watts, D.G., 1980, Relative curvature measures of nonlinearity, *Journal of Royal Statistical Society* 42:1-16

Buford, M. A., 1986, Height-diameter relationships at age

15 in Loblolly Pine seed sources, *Forest Science* 32: 812-818

Curtis, R. O., 1967, Height-diameter and height-diameter-age equations for second-growth Douglas-fir, *Forest Science* 13: 365-375

Dolph, L.D., Mori, S. R. and Oliver, W. W., 1995, Height-diameter relationships for conifer species on the Blacks Mountain experimental forest, US Pacific Southwest Research Station, Research note PSW-RN-418

Dolph, L.D., 1989, Height-diameter equations for young-growth Red-fir in California and Southern Oregon, US Pacific Southwest Research Station, Research note PSW-408

Flewelling, J.W., and De Jong, R., 1994, Considerations in simultaneous curve fitting for repeated height-diameter measurements, *Canadian Journal of Forest Research* 18(2): 195-201

Garcia, O., 1974, Height-diameter equations for *Pinus radiata*, Instituto Forestal, Chile, Nota Tecnica No.19

Garman, S. L., Acker, S.A., Ohmann, J. L., and Spies T. A., 1995, Asymptotic height-diameter equations for twenty-four species in Western Oregon, Forest Research Laboratory, Oregon State University, Research Contribution 10

Goulding, C.J., 1995, Measurement of trees, In D. Hammond (editor): *NZIF Forestry Handbook*, New Zealand Institute of Forestry (Inc.), 1995, Christchurch, page 106

Henriksen, H.A., 1950, Height-diameter curve with logarithmic diameter: brief report on a more reliable method of height determination from height curves, introduced by the State Forest Research Branch, Dansk Skovforen Tidsskr, 35: 193-202

Huang, S., Titus, S. J., & Wiens, D. P., 1992, Comparison of nonlinear height-diameter functions for major Alberta tree species, *Canadian Journal of Forest Research* 22: 1297-1304

Knowe, S.A., 1994, Effect of competition control treatments on height-age and height-diameter relationships in young Douglas-fir plantations, *Forest Ecology & Management* 67:101-111

Lappi, J., 1997, A longitudinal analysis of height/diameter curves, *Forest Science* 43(4): 555-570

Larsen, D. R. & Hann, D. W., 1987, Height-diameter equations for seventeen tree species in southwest Oregon, Oregon State University, Forestry Research Laboratory, Research paper 49

McEwen, A. D., 1978, N. Z. forest service computer system for permanent sample plots, In FRI Symposium - 'Mensuration for Management Planning of Exotic Forest Plantations', Symposium No. 20: 235-252

Myers, C.A., 1966, Height-diameter curves for tree species subject to stagnation, Forest Service, USDA, Rocky Mountain Forest and Range Experiment Station, Research note RM-69

Myers, H.A., 1940, A mathematical expression for height curve, *Journal of Forestry* 38: 415-420

Neter, J., and Wasserman, W., 1974, 'Applied Linear Statistical Models', RICHARD D. IRWIN, INC., Georgetown, Ontario, USA, 1974, page 352-370

Richards, F.J., 1959, A flexible growth function for empirical use, *Journal of Experimental Botany* 10: 290-300

SAS Institute Inc., 1990, SAS/STAT user's guide, Cary, NC, version 6(2): 1135-1194

Schmidt, A. von, 1967, Der rechnerische ausgleich von bestandeshohenkurven, Forstwissenschaftliches Zentrat Blatt 86: 370-382

Schumacher, F.X., 1939, A new growth curve and its application to timber yield studies, *Journal of Forestry* 37:819-820

SPBL, 1998, Annual Report, Selwyn Plantation Board Limited, Christchurch, New Zealand, 40 pages

Trorey, L.G., 1932, A mathematical method for the construction of diameter height curves based on site, *The Forestry Chronicle* 8:121-132

Wang, C.H. & Hann, D.W., 1988, Height-diameter equations for sixteen tree species in the central western Willamette valley of Oregon, Oregon State University, Forestry Research Laboratory, Research paper 51

West, P.W., Ratkosky, R.A., and A.W. Davis, 1984, Problems of hypothesis testing of regressions with multiple measurements from individual sampling units, *Forest Ecology and Management* 7: 207-224

West, P.W., 1995, Application of regression analysis to inventory data with measurements on successive occasions, *Forest Ecology and Management* 71: 227-234

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(3): 15–18

Yang, R. C., Kozak, A., and Smith, J.H.G., 1978, The potential of Weibull-type functions as a flexible growth curve, *Canadian Journal of Forest Research* 8: 424-431