A new tree height model for cut-to-length *Pinus radiata* stems and its prediction error distribution

Dr. Huiquan Bi

ForMetrics Pty. Ltd. &
Forest Science, NSWDPI

huiquan.bi@formetrics.com.au

NSW GOVERNMENT | Primary Industries

A new tree height model for cut-to-length (CTL) *Pinus radiata* stems and its prediction error distribution

1. Introduction
2. Data and CTL simulations
3. Model development and validation
4. Characterizing prediction errors through the Burr Type XII distribution
5. Conclusions

Introduction

1. CTL harvesters as a major source of big data for forest management

   Modern cut-to-length (CTL) harvesters have been widely utilised to improve log harvesting productivity in plantations. Equipped with a GPS receiver and a computerized harvester head, they constantly capture, accrue and provide a daily flow of spatially explicit and timestamped data on log production and assortment as well as detailed diameter and log length measurements of harvested stems over large operational areas.

2. Harvesters can capture the total log length but not the total height of individual trees.

   The lack of total tree height data represents a stumbling block in the full integration of harvester data with conventional inventory data, remote sensing imagery and LiDAR data.

3. Transformation of big data into valuable data for forest management

   Without the full data integration, maximum value extraction from harvester data cannot be attained, preventing the transformation of big harvester data into valuable data for forest management.

Introduction

4. Estimating tree height of CTL stems - a first step in harvester data analytics

   For the most effective use of harvester data, a necessary first step is to estimate the total height of each harvested tree that was bucked, measured and recorded by the harvester head.

5. Previous methods of tree height estimation

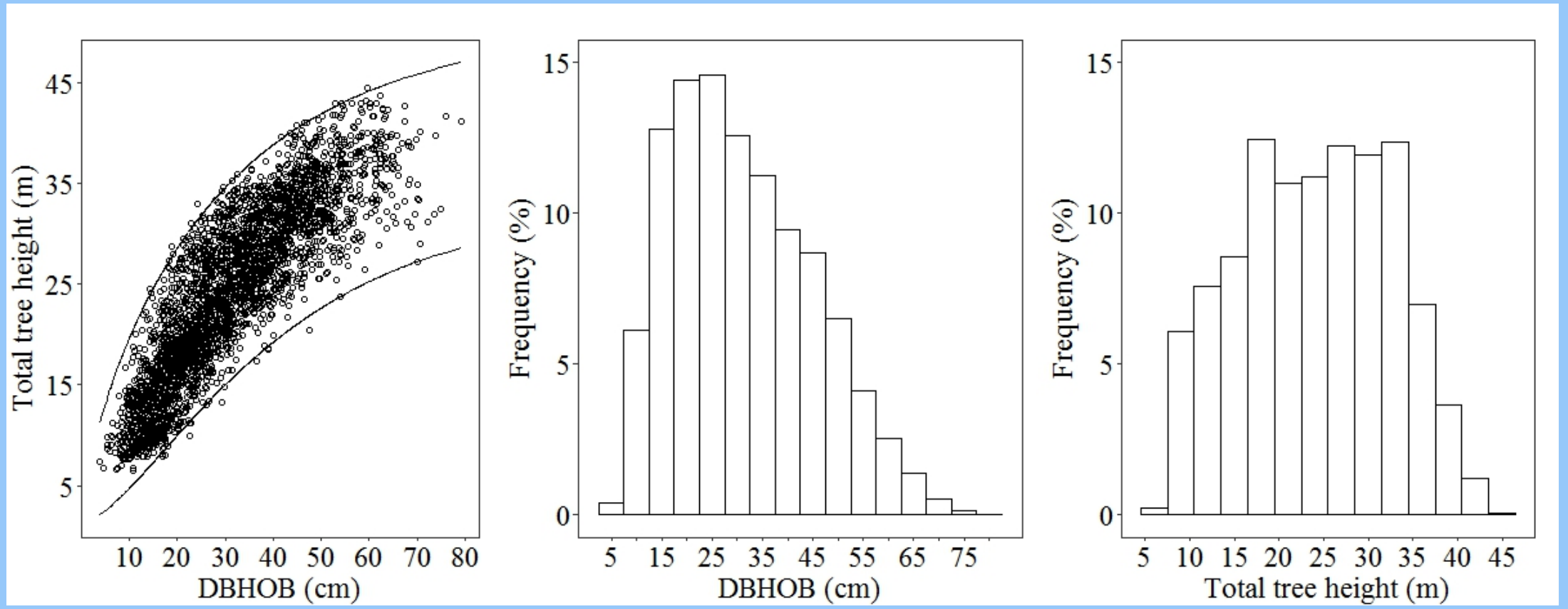
   (a) Varjo's (1995) log linear model predicts the length of unprocessed treetop from DBH, total log length and SED of the top log. It appeared in a short two-and-half page section of an internal research paper of the Finnish Forest Research Institute. The essential elements of model development such as model formulation, variable selection, parameter estimation, model testing and statistical validation were not documented.

   (b) Iterative search algorithm through a taper equation.

6. A new tree height model is needed for cut-to-length (CTL) *Pinus radiata* stems

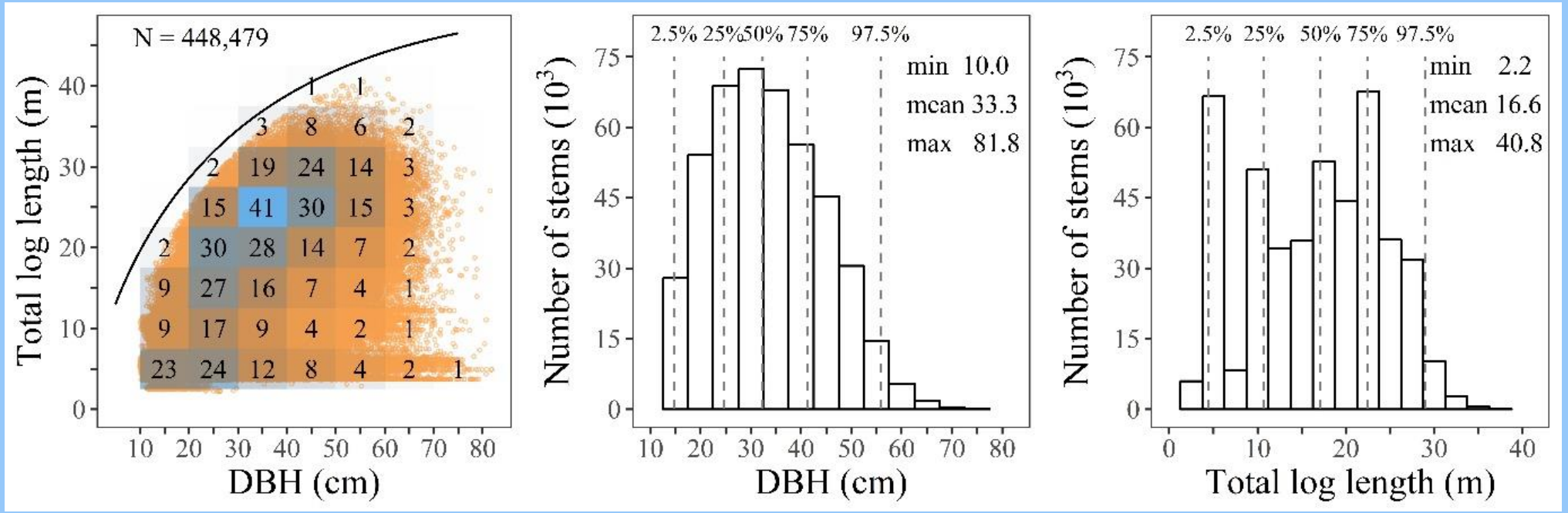
Data

1. Taper Data



Data

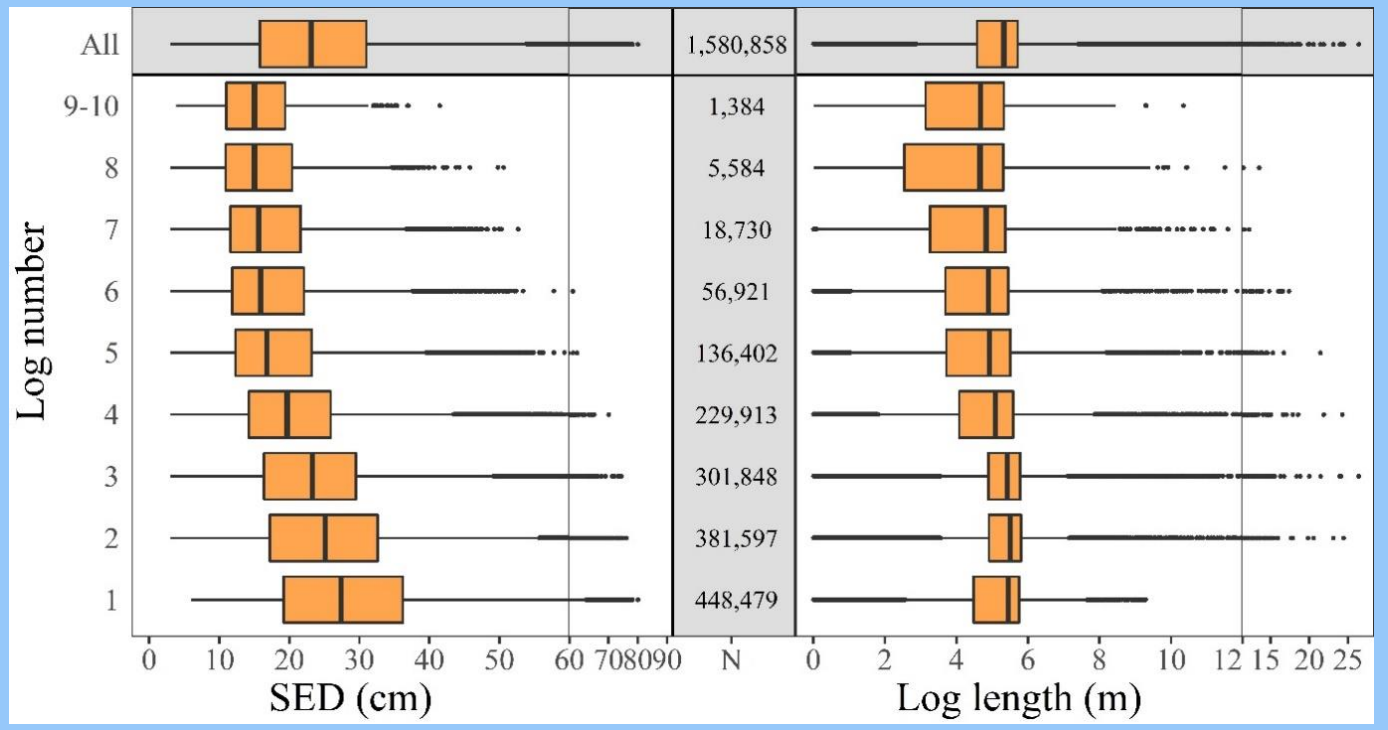
2. Harvester Data



Total log length (L) in relation to DBH under the curve of the maximum attainable merchantable height for the 0.448 million stems drawn on the left as clustered heatmaps. The numbers in the grid cells indicate the number of stems in thousands. The corresponding frequency distributions of DBH (center) and L (right) were shown together with characteristic percentiles and descriptive statistics.

Data

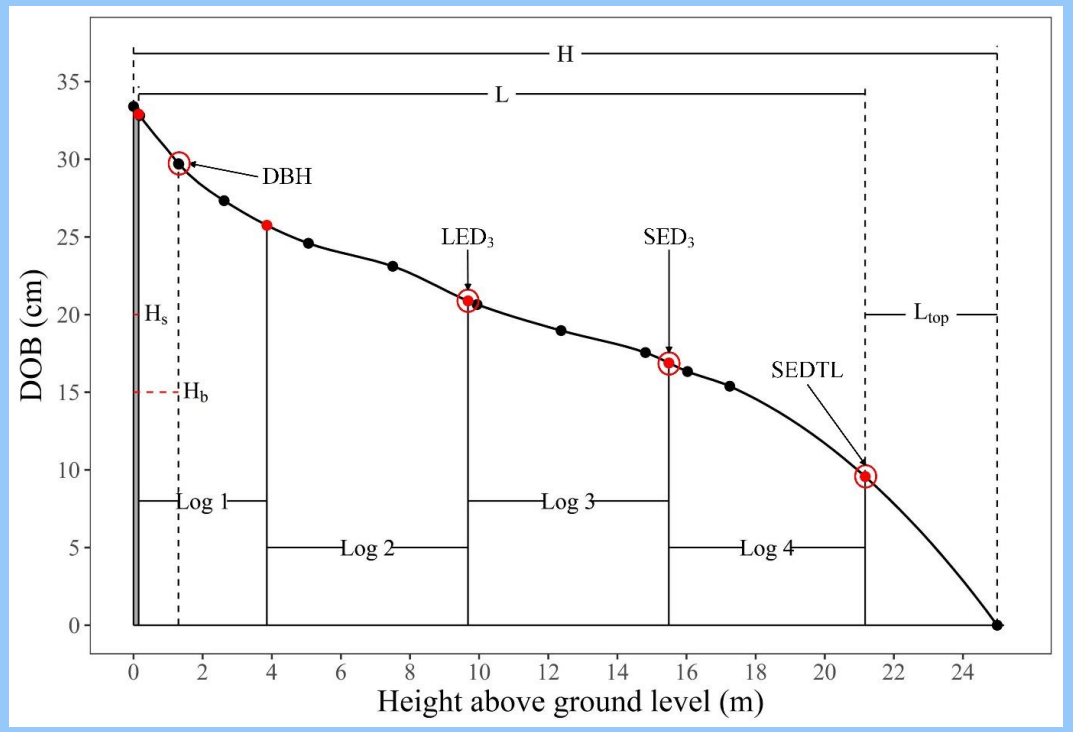
2. Harvester Data

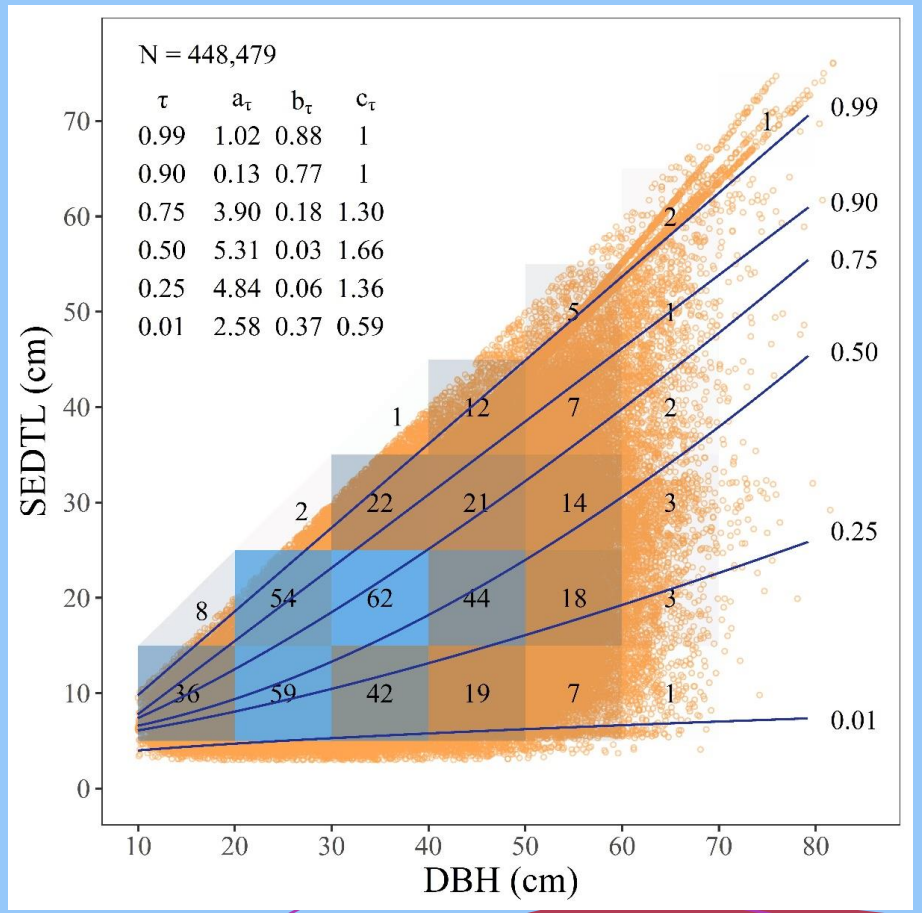


Boxplots of SED (left) and log length (right) across the sequential log numbers for the 1.58 million logs (including waste sections) cut-to-length from the 0.448 million stems contained in the screened harvester data set. The numbers in the middle vertical stripe indicate the number of logs across the sequence. The boxplots in the top horizontal stripe are for all the logs combined.

Log cutting simulations

generated about 186000 CTL stems





$$SI = \sqrt{(DBH_H - DBH_T)^2 + (L_H - L_T)^2 + (SEDTL_H - SEDTL_T)^2}$$

Model development and validation

1. Model derivation
   (a) Considered the relative stem profile of the part of the stem above breast height.
   (b) Followed the approach of Bi (2000) in the construction of the trigonometric variable-form taper model.
   (c) Model in both linear and nonlinear form:

Linear: $$lnh = \left(a_1 + a_2 L + a_3\sqrt{T} + a_4\sqrt{d^3} + a_5\sqrt{DBH^3}\right) ln(1-d)$$

Nonlinear: $$h = (1-d)^{a_1 + a_2 L + a_3\sqrt{T} + a_4\sqrt{d^3} + a_5\sqrt{DBH^3}}$$

$$H = \frac{L + H_s - H_b}{(1-d)^{a_1 + a_2 L + a_3\sqrt{T} + a_4\sqrt{d^3} + a_5\sqrt{DBH^3}}} + H_b$$

Notation:

DBH	diameter at breast height overbark in cm;
H	total tree height from ground level to tip of the tree in m;
H_s	the average stump height of 0.15 m;
H_b	the defined breast height of 1.3m above ground level;
L	total log length, i.e. the sum of lengths of logs and waste sections of a stem in m;
$SEDTL$	SED of the top log in cm, the smallest SED of a cut stem;
d	$= (SEDTL/DBH)$, relative diameter that takes any value between 0 and 1;
h	$= (L + H_s - H_b)/(H - H_b)$, relative height above breast height;
T	$= (DBH - SEDTL)/L$, average taper over total log length.

Model development and validation

2. Parameter estimation

- Non-independent errors

  Because the CTL simulations processed each taper tree 6 times using 6 values of $SEDTL_\tau$, there was an inherent correlation among the residuals from the same tree.

- Heteroskedasticity

  The 6 values of L for each taper tree represented different proportions of its total tree height H. As L decreased and $SEDTL_\tau$ increased, the magnitude of residual variation became increasingly larger, presenting a clear case of heteroskedasticity.

- Estimation

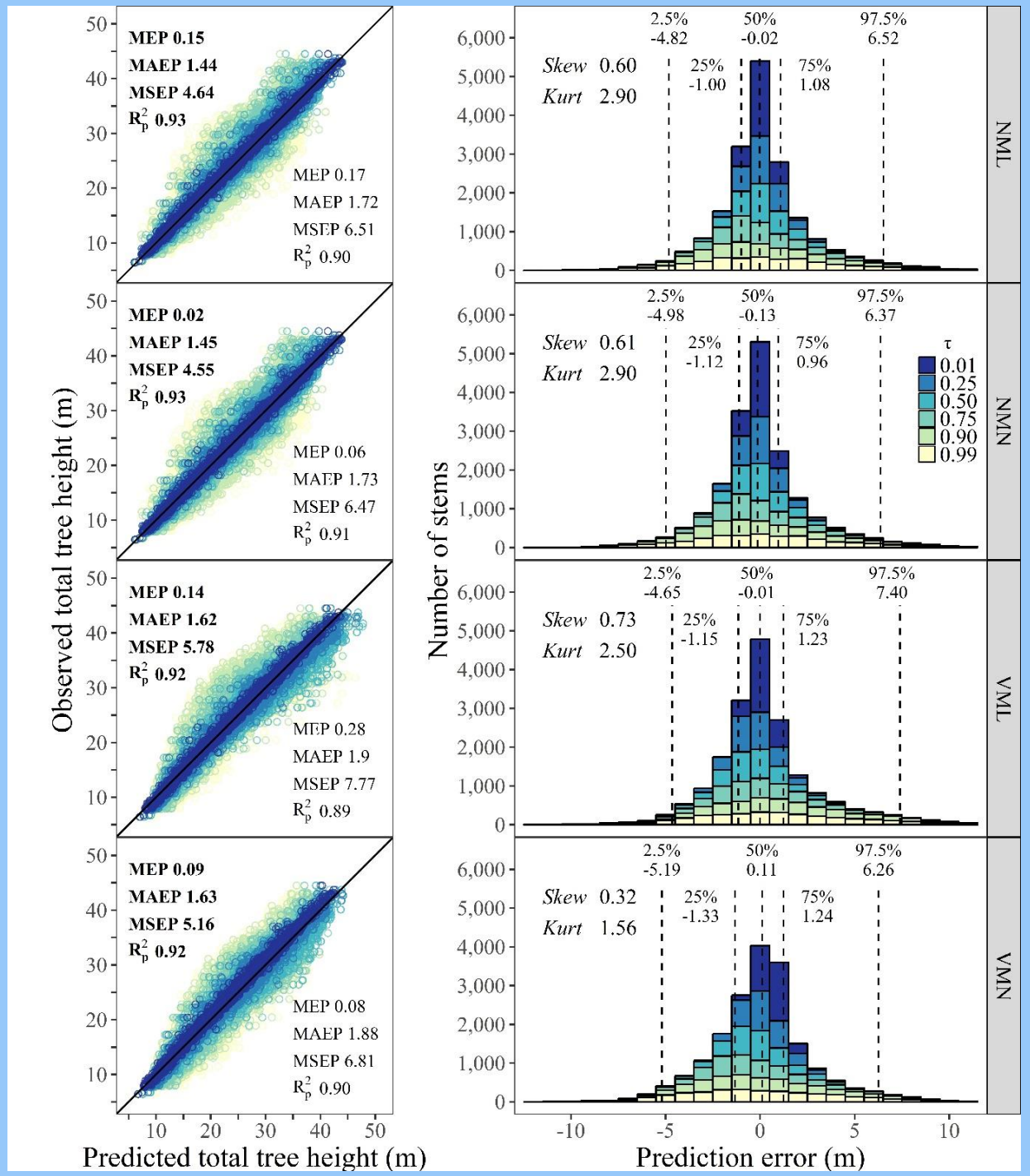
  System of 6 equations estimated by Generalized Method of Moments (GMM).

$$
\begin{aligned}
H_1 &= f(X,\theta) + \varepsilon_1 \\
H_2 &= f(X,\theta) + \varepsilon_2 \\
&\vdots \\
H_6 &= f(X,\theta) + \varepsilon_6
\end{aligned}
$$

Model development and validation

3. Evaluating and comparing prediction accuracy

- Leave-one-tree-out — This cross-validation approach was adopted to obtain prediction errors from and for trees that were independent of the model building process.
- Benchmarking statistics — Five benchmarking or validation statistics commonly used in forest growth and yield modelling.
- Models compared — Our new model and Varjo's (1995) equation
- Model Forms — Linear and nonlinear



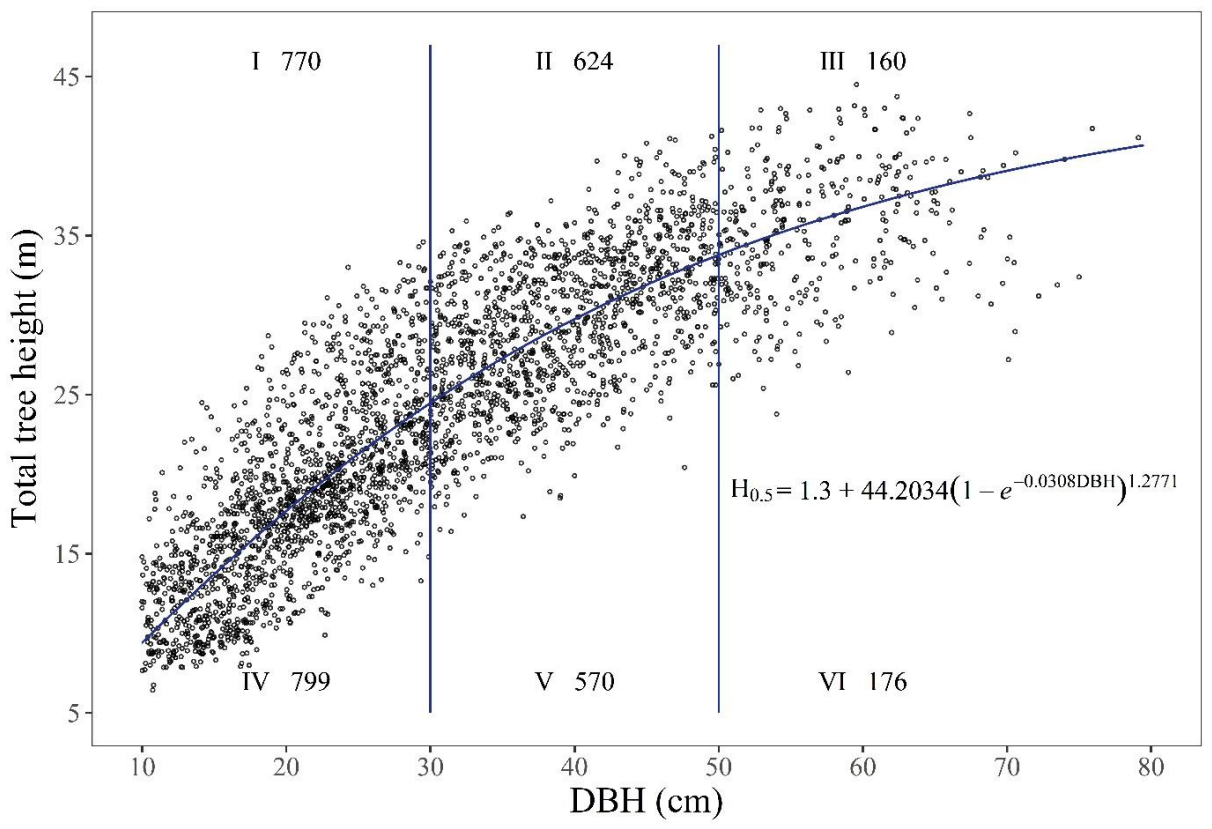
Model development and validation

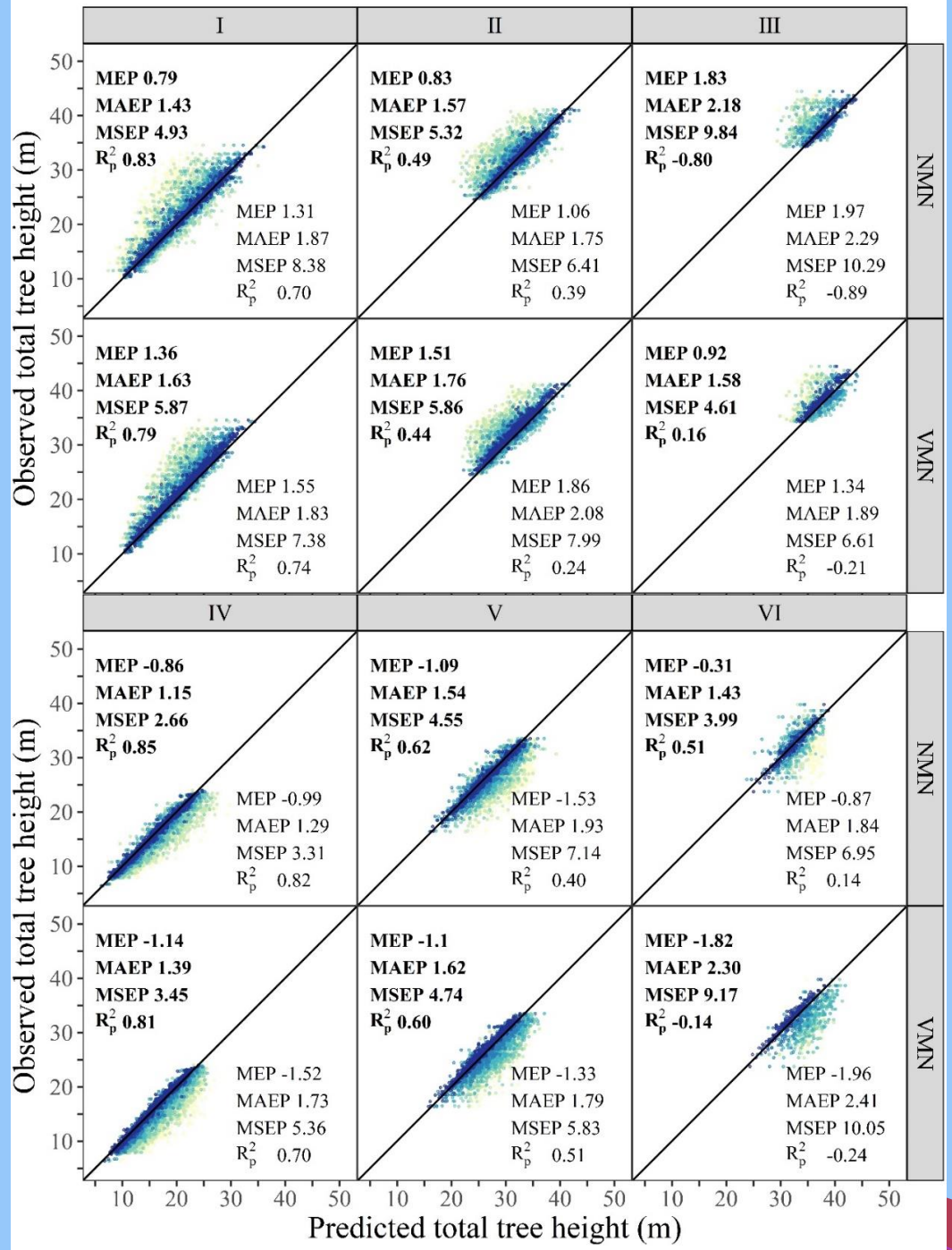
3. Evaluating and comparing prediction accuracy

Our new model in its nonlinear form was the least biased and most precise in predicting the total tree height of cut-to-length Pinus radiata stems.

Model development and validation

3. Evaluating and comparing prediction accuracy





relative diameter d	N	Median Skewness Kurtosis	MEP (m)	MAEP (m)	MSEP	R_p^2
$d \leq 0.20$	1817	-0.15	-0.10	0.48	0.39	0.99
		0.88	0.81	0.85	1.06	0.97
		2.78	(-0.12)	(0.57)	(0.37)	
$0.20 < d \leq 0.30$	954	0.02	0.05	0.46	0.40	0.98
		0.97	0.51	0.64	0.72	0.97
		6.44	(0.10)	(0.71)	(0.55)	
$0.30 < d \leq 0.40$	2599	-0.03	0.04	0.88	1.42	0.98
		0.70	0.30	0.99	1.75	0.97
		2.29	(0.14)	(0.89)	(0.81)	
$0.40 < d \leq 0.50$	3081	-0.12	0.00	1.08	2.14	0.96
		0.60	-0.03	1.27	2.84	0.95
		1.91	(-0.09)	(0.85)	(0.75)	
$0.50 < d \leq 0.60$	741	-0.07	0.22	1.10	2.68	0.98
		1.33	-0.66	1.48	4.00	0.96
		3.55	(-0.33)	(0.74)	(0.67)	
$0.60 < d \leq 0.70$	3125	-0.43	-0.03	1.89	6.02	0.91
		0.76	-0.37	2.13	7.22	0.90
		0.93	(0.08)	(0.89)	(0.83)	
$0.70 < d \leq 0.80$	3197	-0.41	0.12	2.77	12.29	0.83
		0.51	0.14	2.79	12.30	0.83
		0.11	(0.85)	(0.99)	(1.00)	
$0.80 < d \leq 0.90$	2194	-0.22	-0.02	3.13	15.80	0.79
		0.26	0.17	3.17	15.18	0.80
		0.26	(-0.14)	(0.99)	(1.04)	
$0.90 < d \leq 0.95$	873	0.40	0.69	3.06	15.86	0.61
		0.18	-0.37	3.07	14.20	0.65
		0.14	(-1.85)	(0.99)	(1.12)	
$0.95 < d < 0.98$	13	3.14	3.43	3.79	26.42	0.04
		0.83	2.40	2.89	14.05	0.49
		0.35	(1.43)	(1.31)	(1.88)	

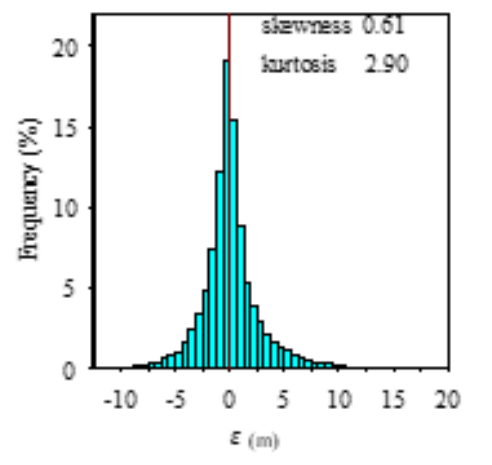
Model development and validation

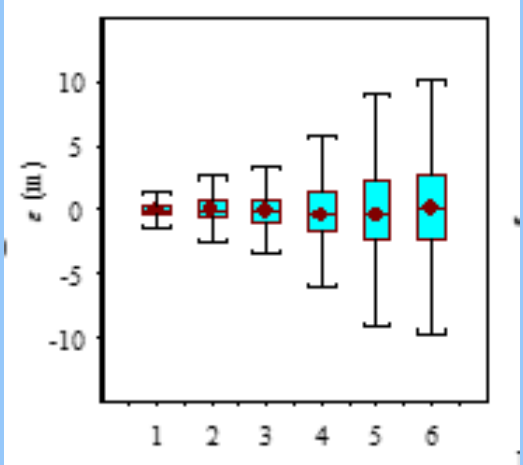
3. Evaluating and comparing prediction accuracy

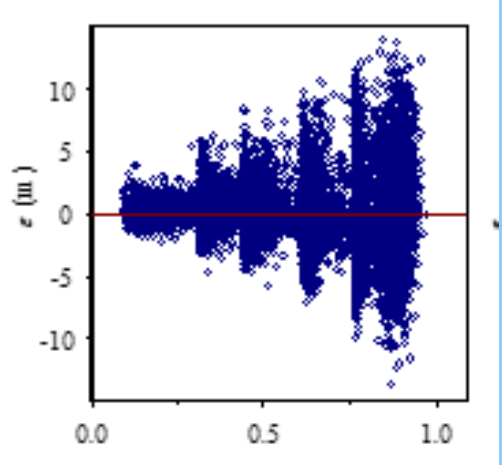
Our new model in its nonlinear form also performed better than the nonlinear form of Varjo's (1995) model across relative SEDTL (top end) diameter classes where d≤0.9.

Characterizing prediction errors through a PDF

- The distribution of prediction errors was highly leptokurtic with an excess kurtosis of 2.90. When the benchmarking statistics were examined locally across nine intervals of d, the MSEP was found to increase from 0.39 when $d \leq 0.20$ to 15.86 when $0.91 < d \leq 0.95$, an increase of almost 40 folds, while the bias of prediction, as indicated by MEP, was negligible across the intervals.







- These results indicated that the prediction error (ε) was highly heteroskedastic and its variance was conditional upon d and possibly other predictor variables in the model. Such conditional heteroskedasticity meant that the conditional variance of ε was a function of the predictor variables. Therefore, a skedastic or weighting function must be first derived to delineate the pattern of heteroskedasticity and adequately quantify the conditional variation of ε. Such a function would be used for weighting ε so that the weighted prediction errors (ε_w) could become much less heteroskedastic and more homoscedastic before being characterized through a probability density function (PDF).

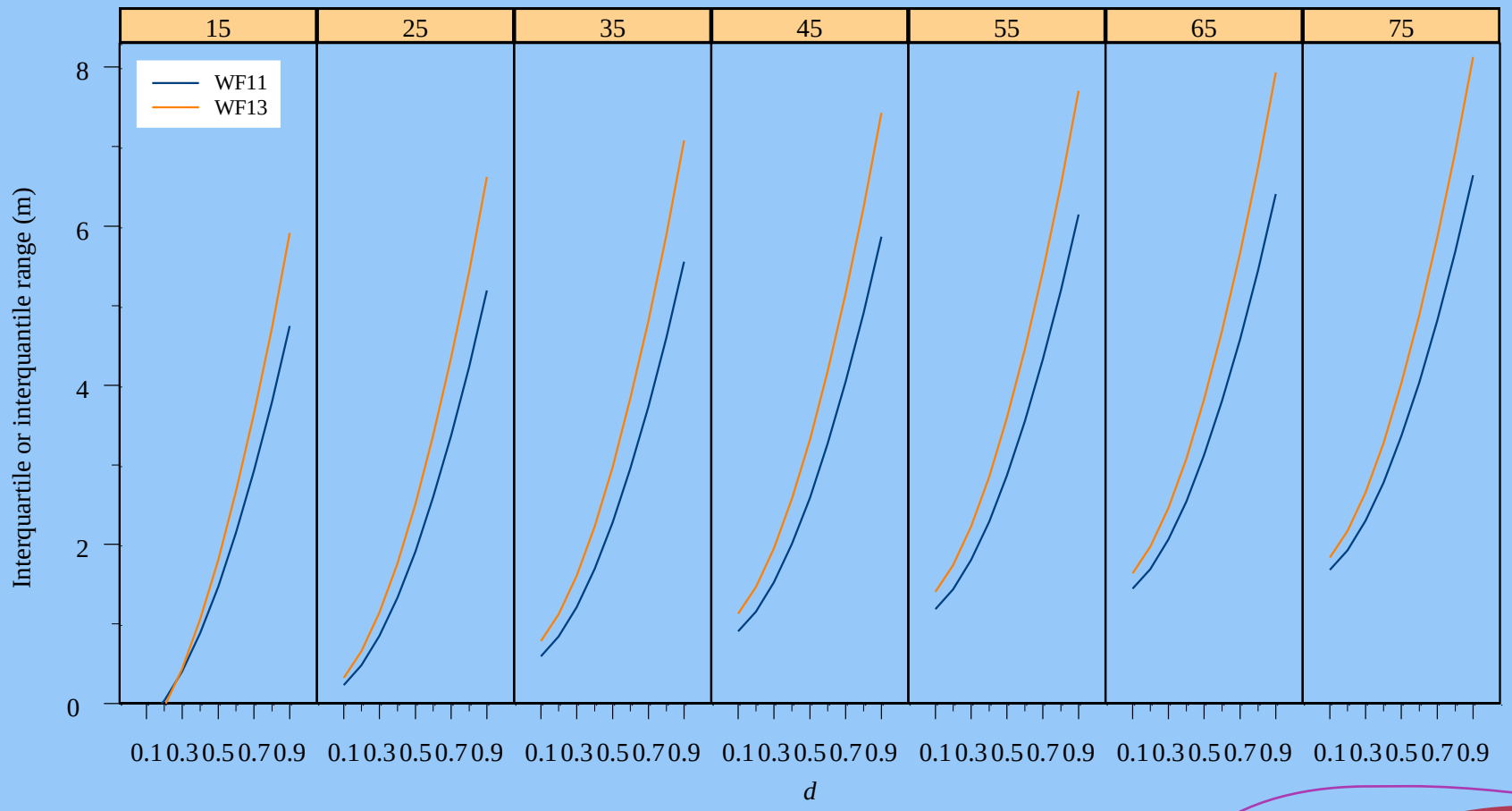
Characterizing prediction errors through a PDF

- Skedastic and weighting functions

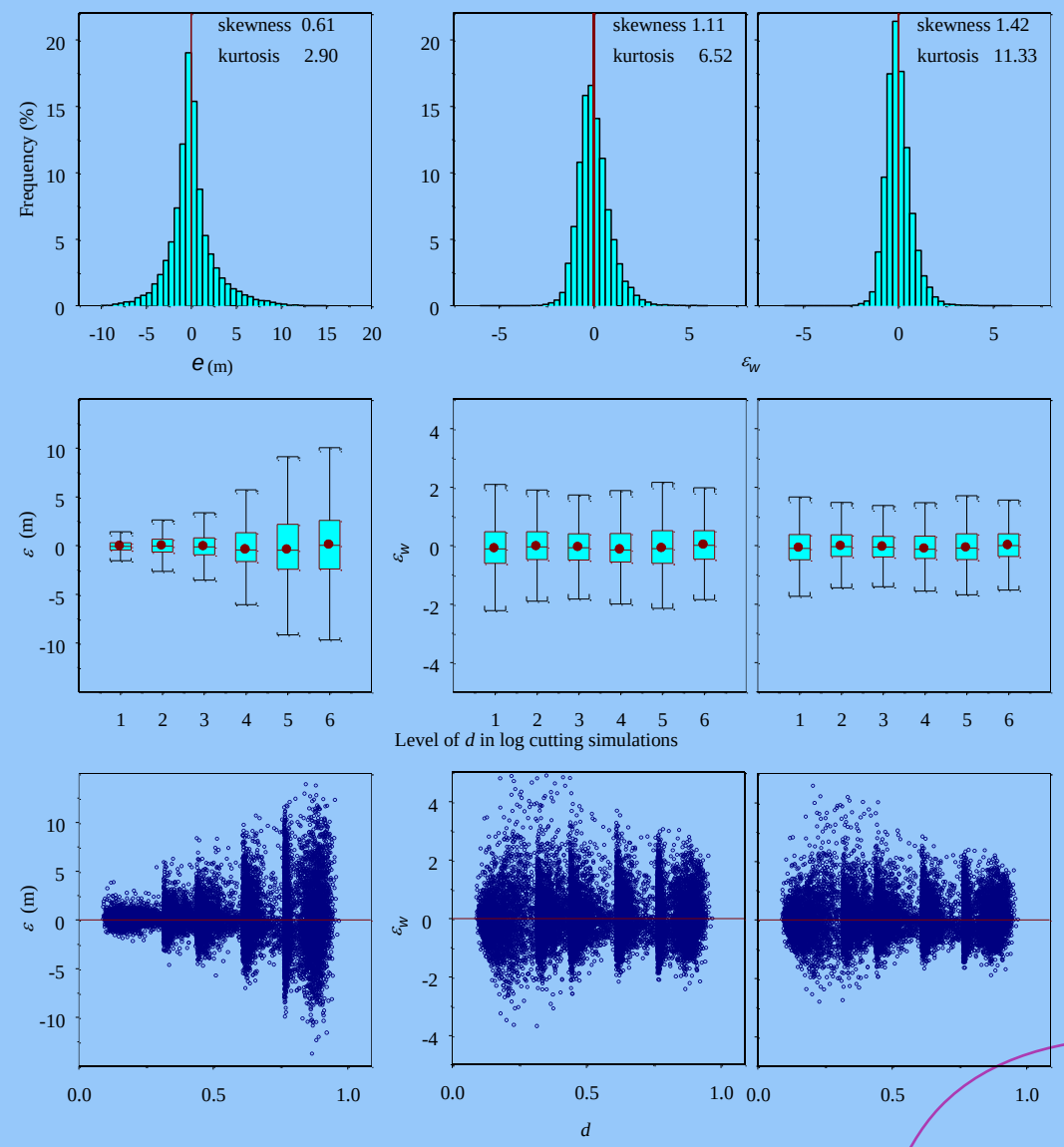
No.	$w(d, DBH)$	q	Eqn	p<0.05	Rank
1	$\sqrt{e^{\beta_0} d^{\beta_1} DBH^{\beta_2}}$		(2)	7.3/2.3	6/6
2	$e^{\beta_0} d^{\beta_1} DBH^{\beta_2}$		(3)	6.3/2.2	4/5
3	$\beta_{0,1-q} d^{\beta_{1,1-q}} DBH^{\beta_{2,1-q}} - \beta_{0,q} d^{\beta_{1,q}} DBH^{\beta_{2,q}}$	0.10	(5)	24.7/6.2	14/14
4	$\beta_{0,1-q} d^{\beta_{1,1-q}} DBH^{\beta_{2,1-q}} - \beta_{0,q} d^{\beta_{1,q}} DBH^{\beta_{2,q}}$	0.20	(5)	20.5/5.4	13/13
5	$\beta_{0,1-q} d^{\beta_{1,1-q}} DBH^{\beta_{2,1-q}} - \beta_{0,q} d^{\beta_{1,q}} DBH^{\beta_{2,q}}$	0.25	(5)	16.6/4.1	12/12
6	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} DBH - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} DBH$	0.10	(7)	7.3/2.4	6/7
7	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} DBH - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} DBH$	0.20	(7)	9.3/2.6	11/11
8	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} DBH - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} DBH$	0.25	(7)	7.3/2.5	6/9
9	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} \sqrt{DBH} - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} \sqrt{DBH}$	0.10	(7)	9.0/2.5	10/9
10	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} \sqrt{DBH} - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} \sqrt{DBH}$	0.20	(7)	6.1/2.1	3/3*
11	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} \sqrt{DBH} - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} \sqrt{DBH}$	0.25	(7)	3.2/1.9	1/1
12	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} ln(DBH) - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} ln(DBH)$	0.10	(7)	8.8/2.4	9/7
13	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} ln(DBH) - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} ln(DBH)$	0.20	(7)	5.3/2.0	2/2
14	$\beta_{0,1-q} + \beta_{1,1-q} d^{\beta_{2,1-q}} + \beta_{3,1-q} ln(DBH) - \beta_{0,q} - \beta_{1,q} d^{\beta_{2,q}} - \beta_{3,q} ln(DBH)$	0.25	(7)	7.2/2.1	5/3*

Characterizing prediction errors through a PDF

- Shape of the two selected weighting functions



Characterizing prediction errors through a PDF



$$\varepsilon_w = \frac{\varepsilon}{w(d, DBH)}$$

$$\varepsilon_{w+} = \varepsilon_w + C_I$$

Characterizing prediction errors through the Burr Type XII distribution

The three-parameter Burr Type XII (BXII) distribution:

$$f(x) = \frac{ck}{\alpha}\left(\frac{x}{\alpha}\right)^{c-1}\left(1+\left(\frac{x}{\alpha}\right)^{c}\right)^{-(k+1)}$$

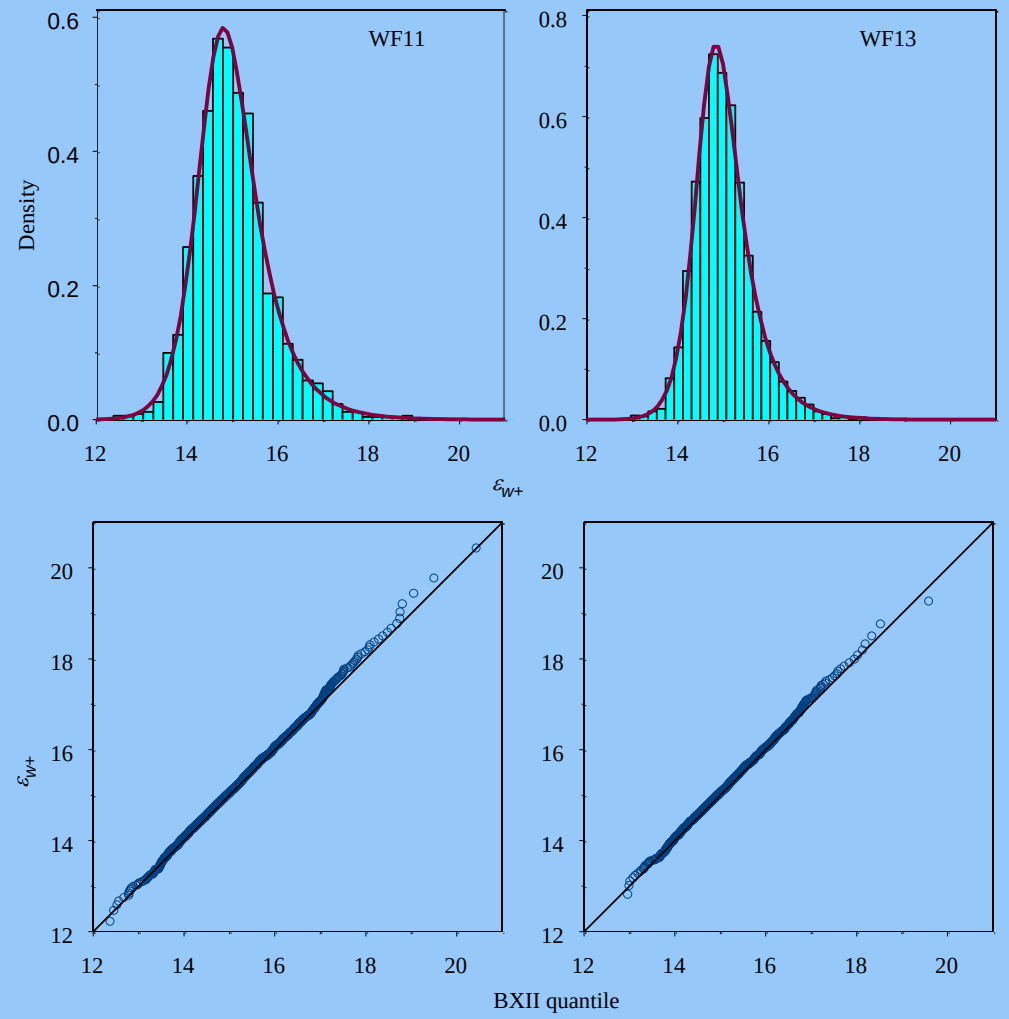
where $x > 0$ is a non-negative continuous random variable, $\alpha > 0$ is the scale parameter, $c > 0$ and $k > 0$ are the shape parameters of the distribution. The mean and variance of the three-parameter BXII distribution are given by:

$$E(x) = \frac{\alpha}{c}\beta\left(\frac{1}{c}, k-\frac{1}{c}\right), \quad \alpha > 1, c > 1$$

$$Var(x) = \frac{2\alpha^2}{c}\beta\left(\frac{2}{c}, k-\frac{2}{c}\right) - \frac{\alpha^2}{c^2}\left(\beta\left(\frac{1}{c}, k-\frac{1}{c}\right)\right)^2, \quad \alpha > 2, c > 2$$

Because the BXII distribution is a positive continuous distribution, the weighted prediction errors were shifted to the positive domain by adding a positive integer: $\varepsilon_{w+}=\varepsilon_w + C_I$

Characterizing prediction errors through the Burr Type XII distribution



Top: Density distribution histograms of a sample of 3099 weighted and shifted prediction errors (ε_{w+}) derived through the two best performing weighting functions (WF11 and WF13) overlayed with their fitted BXII PDF curves.

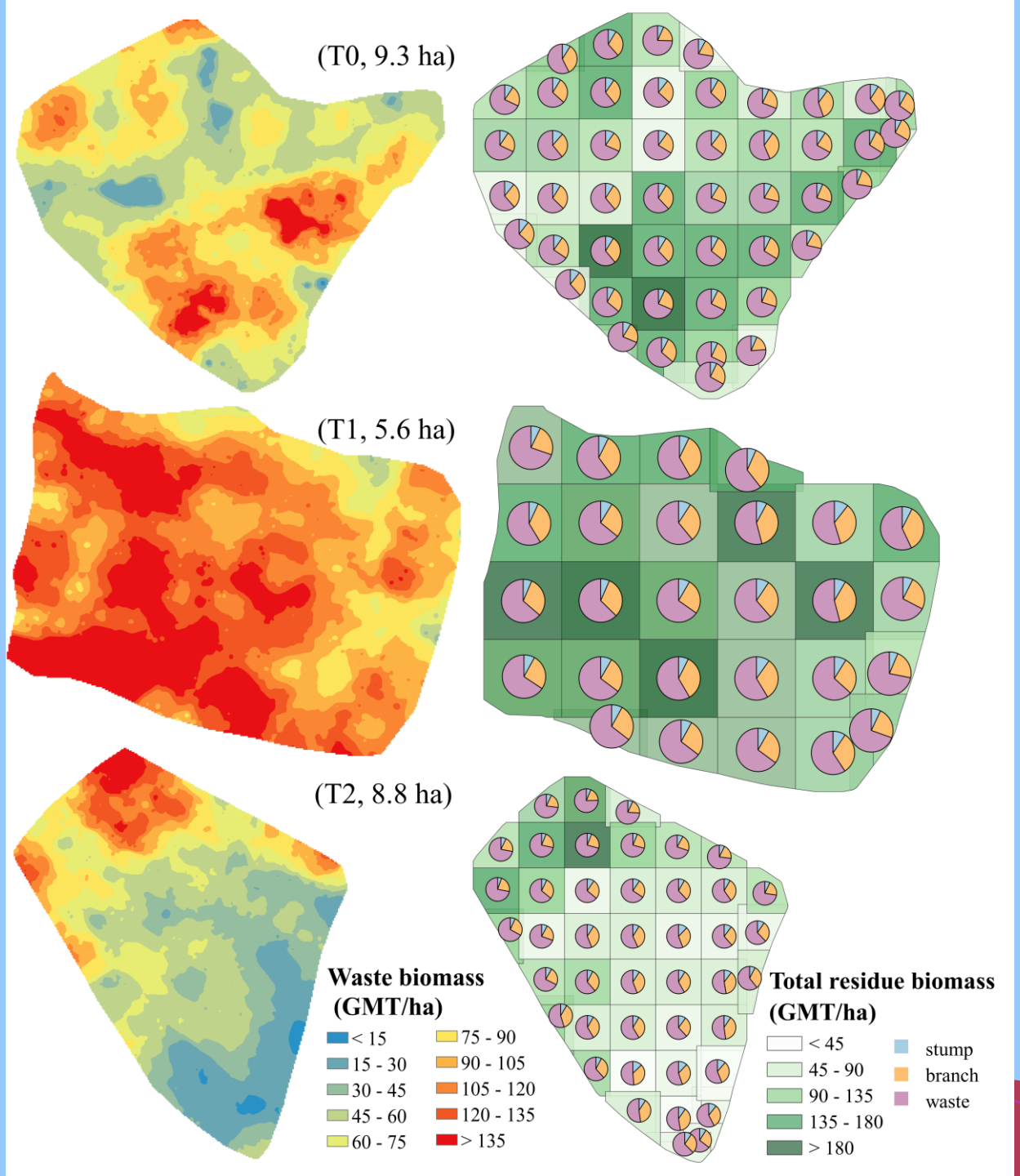
Bottom: Corresponding Q-Q plots of the empirical quantiles of ε_{w+} and the quantiles of the fitted BXII distributions.

Conclusions

1. The tree height model will facilitate and widen the utilization of harvester data far beyond the current limited use of monitoring log yield and assortment only.
2. It will enable the full integration of harvester data with conventional inventory data, remote sensing imagery and LiDAR data for the development of harvester-based inventory systems, for the prediction of attributes of individual trees, stands and forests, and for the estimation of product recovery and residue biomass in radiata pine plantations.
3. Accurately estimated total tree height will make harvester data a potential source of taper data to supplement the conventional destructive taper sampling in the field.
4. The model will also facilitate (1) the screening and exploratory analysis of harvester data, (2) the calibration and estimation of bark thickness, (3) the mapping of site index, (4) the development site-specific height-diameter curves, and (5) the post-thinning assessment of diameter and height distributions of retained versus removed stems.
5. Our characterized prediction error distribution would help to unlock the full potential of the tree height model in many research and practical applications, particularly when uncertainty assessments, statistical inferences and error propagations are needed in harvester data analytics.
6. The tree height model will help the transformation of big harvester data into valuable data for forest management.

Example of model applications

Mapping harvest residue biomass in unthinned, T1 and T2 radiata plantations



Publications

Lu, K., Bi, H., Watt, D., Strandgard, M., Li, Y. (2018) Reconstructing the size of individual trees using log data from cut-to-length harvesters in *Pinus radiata* plantations: a case study in NSW, Australia. Journal of Forestry Research 29:13-33.

Shan, C., Bi, H., Watt, D., Li, Y, Strandgard, M., Ghaffariyan M.R. (2021) A new model for predicting the total tree height of stems cut-to-length by harvesters in Pinus radiata plantations Journal of Forestry Research 32:21-41

Li W., Bi H., Watt D., Li Y., Ghafariyan, M.R. (2021) Estimation and spatial mapping of residue biomass following CTL harvesting in Pinus radiata plantations: an application of harvester data analytics. Forests 13(3), p.428.

Cao X., Bi H., Watt D., Li Y. (2023) Characterizing prediction errors of a new tree height model for cut-to-length Pinus radiata stems through the Burr Type XII Distribution Journal of Forestry Research. https://doi.org/10.1007/s11676-023-01632-3