

RESEARCH ARTICLE

Open Access



Multi-sensor modelling of a forest productivity index for radiata pine plantations

Michael S. Watt^{1*}, Jonathan P. Dash², Pete Watt³ and Santosh Bhandari³

Abstract

Background: An understanding of how plantation productivity varies spatially is important for forest planning, management and projection of future plantation yields and returns. The 300 Index is a volume productivity index developed for *Pinus radiata* D.Don that has been widely used within New Zealand to assess site productivity. Although the 300 Index is routinely characterised at the stand level, little research has investigated if remotely sensed data sources can be used in combination with environmental layers to precisely predict this metric at fine spatial resolution.

Methods: This study uses an extensive dataset obtained from *P. radiata* plantations in the central North Island, New Zealand. Using this dataset, the objective of this research was to compare the precision of parametric and non-parametric models of the 300 Index that included explanatory variables extracted from aerially acquired light detection and ranging (LiDAR), satellite imagery (RapidEye) at 5-m resolution or environmental layers and combinations of these three data sources. Models were constructed both with and without stand age as an explanatory variable as managers may not always have access to stand age. A total of 28 models (14 data sources \times two model methods) were constructed using data from 433 plots. Precision and bias of these models was determined using an independent dataset of 60 plots.

Results: Of the non-parametric methods tested (*k*-most similar neighbour (*k*-MSN), *k*-nearest neighbour (*k*-NN)), *k*-NN using an optimised value of *k* most precisely predicted the 300 Index for 11 of the 14 constructed models. The use of *k*-NN was found to be more precise than parametric models when age was not available but of overall similar precision to parametric models when stand age was available as a predictor. For models including stand age, the inclusion of LiDAR resulted in the most precise model (mean $R^2 = 0.789$; root mean square error (RMSE) = $2.48 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$) while for models without stand age, metrics extracted from both satellite imagery and environmental layers produced the most precise model of the 300 Index ($R^2 = 0.65$; RMSE = $3.21 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$).

Conclusions: Results clearly show that models constructed from LiDAR provide the most precise means of estimating the 300 Index. However, in many situations, LiDAR is too expensive to acquire or stand age, which is used as a reference for linking LiDAR to 300 Index, is not available as an independent variable. Under these circumstances, results show that precise models can be constructed from variables derived from the combination of satellite imagery and environmental surfaces.

Keywords: 300 Index, ALS, LiDAR, Climate, Geospatial layers, Plantation forestry, Productivity layers, RapidEye, Satellite imagery

* Correspondence: michael.watt@scionresearch.com

¹Scion, PO Box 29237, Fendalton, Christchurch, New Zealand

Full list of author information is available at the end of the article

Background

Understanding and quantifying how environmental factors influence tree growth and site quality is of considerable importance to forest managers. Knowledge of site quality serves as an important tool for operational planning and is a major contributing factor for the formulation of management regimes, the prediction of harvest yields and the determination of the economic value of a forest. Site quality can be estimated directly from environmental factors, but it is generally more accurate to infer site quality from some characteristic of the current tree crop on a given site.

Site Index is the most common measure of forest site productivity used worldwide. This variable, which expresses the height of dominant and/or co-dominant trees at a given age (Skovsgaard and Vanclay 2008), is a useful measurement as it is relatively unaffected by management intervention. Despite the prevalent use of Site Index by forest managers and researchers, a site quality metric based on volume, rather than height, may more accurately reflect site quality as volume directly links to stand productivity and integrates both height and diameter growth. This contention is reinforced by previous analysis of stand density trials covering a wide range in Site Index. These analyses showed that, at a common age and stand density, Site Index is only weakly related to basal area, demonstrating that Site Index provides only a partial measure of site productivity for *Pinus radiata* D. Don (*P. radiata*) (Kimberley et al. 2005).

The 300 Index is a volume productivity index for *P. radiata* that provides a more accurate means to ascertain site productivity than Site Index. The 300 Index is calculated by adjusting field plot measurements for age, stand density and silvicultural history to give the mean annual increment (expressed as $\text{m}^3 \text{ha}^{-1} \text{year}^{-1}$) at age 30 for a reference regime where all crop trees are pruned to 6 m in a timely fashion, thinned to the final stocking density (300 stems ha^{-1}) at the completion of the final pruning. Research within New Zealand *P. radiata* plantations has shown that accurate and unbiased values for the 300 Index can be obtained using measurements taken from stands differing in age or stand density from those of the 300 Index standard regime (30 years and 300 stems ha^{-1} , Kimberley et al. 2005). The mean 300 Index for New Zealand is $27.4 \text{ m}^3 \text{ha}^{-1} \text{year}^{-1}$.

Stand-level estimates of the 300 Index are commonly made by averaging plot values, and these stand estimates are used to guide forest management. The resulting estimate aggregations may not be adequate to detect local growing variations within stands resulting from interactions between climate, soil, topography and genetic factors (Véga and St-Onge 2009; Saremi et al. 2014). Consequently, predictive modelling methods should be developed to produce spatial

layers that provide a finer spatial resolution for estimates of the 300 Index.

National layers of the 300 Index with moderate accuracy have been developed from site descriptors that can easily be viewed and analysed in a geographic information system (GIS). This approach has considerable merit as forest managers generally have access to GIS and can use these layers to understand the future yields of their resource. The last two decades have seen rapid increases in the number of spatial layers covering a diverse range of climatic (Leathwick and Stephens 1998; Leathwick et al. 2003; Tait et al. 2006) and edaphic variables (Watt and Palmer 2012). These environmental layers have been successfully used to develop spatial layers describing the 300 Index for *P. radiata* at a resolution of 100 m^2 (Palmer et al. 2009a).

Three-dimensional point cloud data acquired by airborne laser scanning (ALS, also referred to as light detection and ranging (LiDAR)) provides a valuable means of characterising forest canopy structure. Since the early applications of LiDAR to forestry management (Nilsson 1996), the technology has been widely used to spatially quantify variation in stand structure at a range of resolutions (e.g. Holmgren et al. 2003; Hyypä et al. 2001; Lim and Treitz 2004; Popescu et al. 2004; Watt et al. 2013a). Several studies have shown that data acquired through ALS surveys can also be used to predict Site Index (Rombouts et al. 2010; Watt et al. 2015; Packalén et al. 2011). This is to be expected as ALS data provides highly accurate information on canopy height. However, no research has investigated the utility of ALS for predicting the 300 Index. As the 300 Index is a composite of both stand height and diameter increment, the utility of ALS data is less certain.

Data from satellite imagery typically has a coarser resolution and contains less structural detail of the forest canopy than ALS. However, the use of satellite-derived metrics provides a cost-effective means of predicting vegetation attributes in many cases (Shamsoddini et al. 2013). Numerous studies have shown that multispectral satellite data can be used to map important forest attributes, such as tree height and volume when the data include various combinations of spectral information, vegetation ratios and texture metrics. Results are improved when imagery of a greater resolution is applied (McRoberts and Tomppo 2007; Donoghue and Watt 2006; Shamsoddini et al. 2013; Ozdemir 2008; Kayitakire et al. 2006; Watt et al. 2013b). Previous research has shown that moderate-resolution satellite imagery (5 m) can be useful for monitoring Site Index (Watt et al. 2015) but we are unaware of any research that has investigated the utility of such data for predicting the 300 Index.

Managers of planted forests are likely to have access to a wide variety of spatial information. Although stand age

is an important determinant of the 300 Index, this variable is often not easily available at a national scale. Consequently, there is considerable interest in determining the utility of a variety of different sources of information, in the development of productivity layers such as the 300 Index. A review of the literature indicates that no previous research has compared the precision of models of the 300 Index created from satellite imagery or ALS data, or various combinations of these data sources, supplemented by auxiliary environmental data. It is considered that a comprehensive comparison using these three different sources of information should be undertaken using common parametric and non-parametric modelling approaches. Models that include or exclude age as a determinant of the 300 Index should also be developed to test whether the 300 Index can be predicted in the absence of stand age.

This study used data obtained from a *P. radiata* plantation located in the central North Island of New Zealand. The objective of the research was to compare the accuracy of parametric and non-parametric models of the 300 Index models that included explanatory variables extracted from combinations of ALS data and satellite imagery supplemented with auxiliary environmental data. It was hypothesised that the supplementary environmental data may contain useful information on factors that influence site productivity that could not be detected using remote sensing. This study was designed to identify the most precise model for forest managers to predict the 300 Index based on data from remote sensing and auxiliary data.

Methods

Study site

Data was acquired from Kaingaroa forest which is located in the central North Island of New Zealand (Fig. 1). Kaingaroa is New Zealand's largest contiguous planted forest covering around 180,000 ha. The majority of the forest occupies the pumice plateau within the central North Island and has generally flat topography. The northern part of the forest is characterised by rolling hills and areas of steeper terrain. The terrain gradually slopes upwards towards the forest's southern extent leading to a notable gradient in productivity. The forest soils are classified as Orthic Pumice belonging to the Kaingaroa series (Hewitt 1993) with those in the north of the forest derived from Tarawera ash. The dataset was restricted to stands of *P. radiata*, which cover 92 % of the total forested area.

Field measurement

Grid sampling was used to locate 493 field plots throughout the study forest. Plots were located at the intersections of a grid that had a randomised start point and orientation. The sampling unit was a 0.06-ha

bounded, circular field plot. A survey-grade global positioning system (GPS) was used to fix the centre of each plot. This fix was differentially corrected using local base stations following acquisition. Diameter at breast height (dbh) was measured for all trees within each field plot. Tree height was measured for a subset of plot trees that were free from excessive lean or malformation and were selected from across the dbh range. The silvicultural history and stand age for each plot was extracted from the stand records maintained by the forest manager.

Derivation of the 300 Index

The 300 Index can be estimated from plot measurements of basal area, mean top height and stand density at a known age when the stand silvicultural history has been recorded (Watt et al. 2010). The productivity index is estimated using numerous models including a stand-level basal area growth model, a height/age function, a mortality function, a stand-volume function and a thinning function. These models have been embedded within a software package (West et al. 2013) that was used to estimate the 300 Index value for each field plot. These values served as the response variable for the modelling work in this study.

Candidate predictor variables

The methods used to derive candidate predictor variables for this study were the same as those described in Watt et al. (2015).

ALS data

An ALS survey was completed in early 2014 using an Optech Pegasus scanner to collect a discrete, small footprint dataset. The data were collected with a pulse rate frequency of 100 kHz, a maximum scan angle of 12° off nadir and a swath overlap of 25 %. This resulted in a dataset with a footprint size of 0.25 m and an average pulse density of 11.46 points m⁻². Returns were classified according to the ASPRS standard LiDAR point classes including 2—ground, 3—low vegetation, 4—medium vegetation, 5—high vegetation, 6—building and 7—low point (noise). Classification was automated using the TerraScan module of the TerraSolid software product (TerraSolid Ltd, Helsinki, Finland). Subsequent manual inspection and reclassification, where required, was used to improve the classification accuracy. The automated classifications were only adjusted when, following inspection by a technician, they were deemed to have been clearly classified erroneously.

Within the field plot boundaries, LiDAR metrics were extracted including height percentiles (P5ht, P10ht, P20ht, ..., P95ht, m); the mean (H_{mean} , m) and maximum height (H_{max} , m); several metrics describing the LiDAR height distribution through the canopy

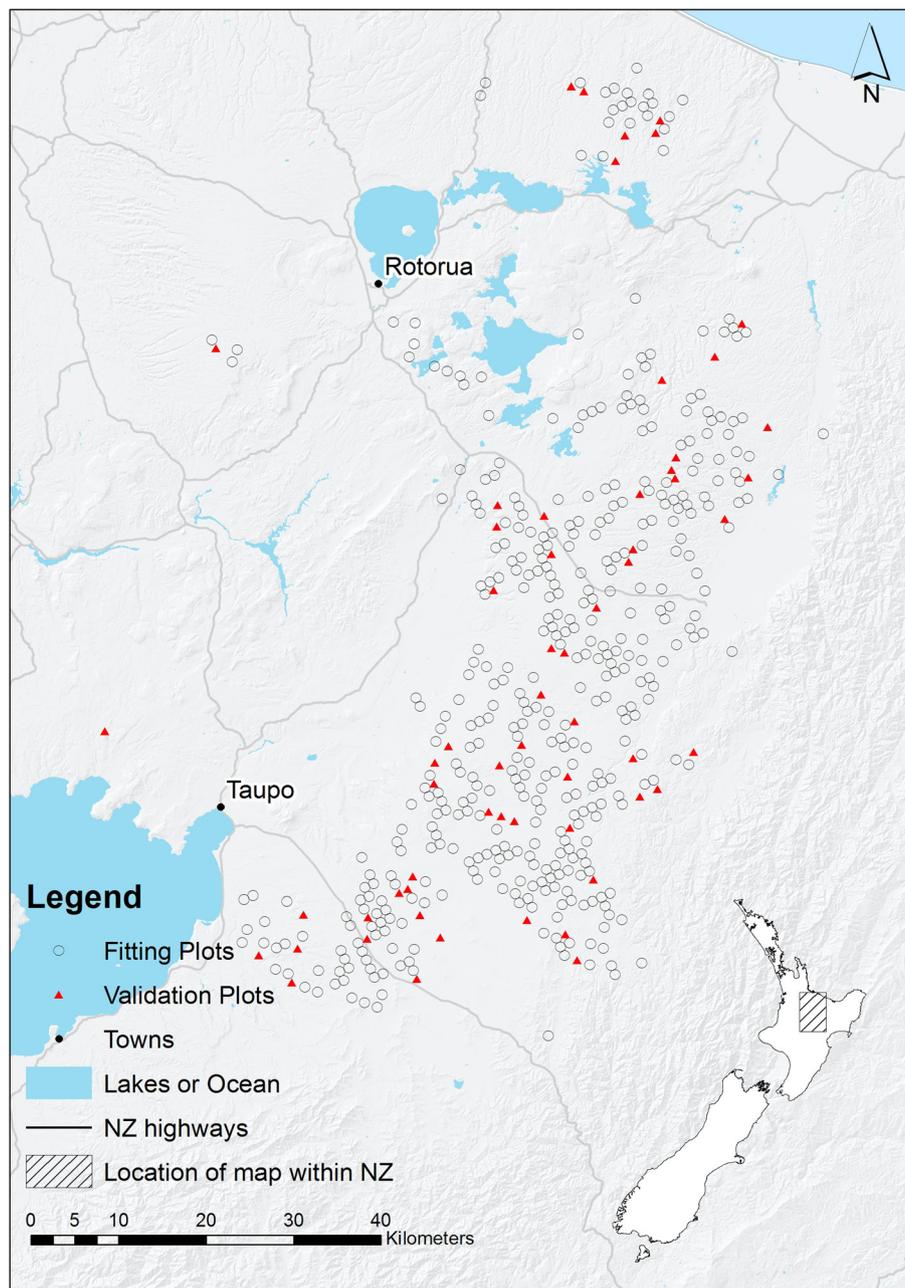


Fig. 1 Map showing the location of the plots used for fitting and validation of the tested models

(skewness, coefficient of variation, standard deviation (SD) and kurtosis); and measures of canopy density such as the percentage of returns reaching within 0.5 m of the ground (P_{ZERO} , %) and the percentage returns above 0.5 m (P_{COVER} , %).

RapidEye data

The RapidEye satellite system is a constellation of five satellites carrying identical sensors, all of which were launched at the end of 2008 (RapidEye AG 2011). Each

sensor collects electromagnetic radiation data in five wavelength bands that include band 1—blue (440–510 nm), band 2—green (520–590 nm), band 3—red (630–685 nm), band 4—red edge (690–730 nm) and band 5—near infrared (NIR) (760–850 nm) at a spatial resolution of 6.5 m, which is resampled to 5 m during pre-processing prior to delivery.

A level 3A ortho product provided by RapidEye was used in this study. The level 3A ortho product underwent a range of pre-processing stages including the application

of radiometric, sensor and geometric corrections. It was also aligned to a cartographic map projection (the Universal Transverse Mercator (UTM)) with the default geometric correction based on Ground Control Points derived from DigitalGlobe 2-m satellite imagery and the New Zealand 25-m digital elevation model (DEM). The intention of the ortho-correction process was to remove impacts of topographic distortions in the imagery. The process ensured that the satellite image conformed to a map projection and included corrections for terrain displacement.

RapidEye images that covered the study area were acquired from January 16 and 28, 2014. The images were delivered as 16-bit digital numbers, and reflectance values were calculated from each band of the RapidEye imagery using the ENVI 4.7 image processing software.

Once converted, the various vegetation indices (Table 1) were calculated from the reflectance and applied to the extent of the images. Texture metrics were calculated using four different window sizes— 3×3 , 5×5 , 15×15 and 25×25 pixels, a consistent displacement of 1 pixel and direction of 135° . The following summary provides an overview of the indices and metrics calculated.

The following vegetation indices were computed from the RapidEye data: Normalised Difference Vegetation Index (NDVI); Enhanced Vegetation Index (EVI); Simple Ratio (SR); Green Ratio (GR); Red Edge Ratio (RE); Vegetation Index (VI); and Brightness. Equations for these ratios are given in Table 1.

Textural attributes, developed by Haralick et al. (1973), are commonly utilised in remote sensing studies. We used the most relevant grey-level co-occurrence matrix (GLCM) textural attributes for remote sensing applications as described in the literature (Baraldi and Parmiggiani 1995; Solberg 1999; Lu 2005; Tuominen and Pekkarinen 2005; Kayitakire et al. 2006). These attributes included the mean (ME); variance (VAR); standard deviation (SD); contrast (CON); angular second moment (ASM); entropy (ENT); homogeneity (HOM); energy (EN); correlation (COR); and dissimilarity (DISS). All textural metrics were calculated using the *gldm* package in the R statistical environment (Zvoleff 2014).

Table 1 Vegetation indices used within the modelling process

Index	Equation
NDVI	$(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$
EVI	$2.5(\text{NIR} - \text{Red})/(\text{NIR} + 6 \text{ Red} - 7.5 \text{ Blue} + 1)$
RE	$\text{Green} \times \text{Red Edge}/\text{Red}$
SR	NIR/Red
GR	NIR/Green
VI	Green/Red
Brightness	$(\text{Blue} + \text{Green} + \text{Red} + \text{Red Edge} + \text{NIR})/5$

Mean values for spectral (individual bands and indices) and textural metrics from the pixels that were spatially coincident with the plots were extracted for the analyses.

Auxiliary environmental data

Auxiliary environmental data were extracted from GIS layers that included terrain attributes (Palmer 2008); environmental layers (Leathwick et al. 2003); and monthly and annual climate data (Mitchell 1991; Leathwick et al. 2002). Environmental variables that served as candidate predictor variables for inclusion in predictive models included mean annual and monthly air temperature, relative humidity, solar radiation, vapour pressure deficit and rainfall. These surfaces had a spatial resolution of 100 m^2 or less. A spatial soil water balance model developed for *P. radiata* (Palmer et al. 2009c) was used to determine mean annual and seasonal root-zone water storage (W) for all plot locations. Fractional available root-zone water storage was determined from these data and the maximum available root-zone water storage, W_{max} , as W/W_{max} . Soil fertility was represented by soil C:N ratio, which provides a useful index of nitrogen mineralisation (Watt and Palmer 2012).

Analyses

Overview

Some of the analytical methodology used here was similar to that in a previous study comparing parametric and non-parametric methods for predicting Site Index for *P. radiata* using combinations of data derived from environmental surfaces, satellite imagery and airborne laser scanning (Watt et al. 2015). From the full dataset of 493 plots, 60 plots were randomly selected and used for model validation. The remaining 433 plots were used for the model-fitting process. Separate models were created from the following categories of data that all included age as a predictor: (i) RapidEye band spectral values; (ii) RapidEye vegetation indices; (iii) RapidEye textural metrics; (iv) all RapidEye metrics (a combination of categories (i) to (iii)); (v) data from environmental layers; (vi) all RapidEye metrics and data from environmental layers; (vii) Site Index predicted from LiDAR (as Site Index is likely to be related to 300 Index); and (viii) Site Index predicted from LiDAR, along with all other LiDAR metrics in combination with environmental layers and all RapidEye metrics.

Models without stand age were also created using predictors from categories (i) to (vi), but not categories (vii) and (viii) as it is not practical to predict Site Index using LiDAR information without age. These 14 different models were created using both parametric and non-parametric methods.

Predictions of Site Index using LiDAR for models (vii) and (viii) were made for each field plot. This was achieved

by first predicting plot mean top height using the 99th LiDAR height percentile, P_{99} , from the following equation:

$$\text{Mean top height} = 2.138 + 0.9927 P_{99} \quad (1)$$

An appropriate transformation of a height age function was then used to convert this estimate of plot mean top height, along with plot age from the stand record system, into an estimate of Site Index.

Model precision was assessed for parametric and non-parametric models from the validation dataset using the squared Pearson correlation coefficient (R^2) and root mean square error (RMSE). Model bias was assessed through the mean error, with error for each observation in the validation dataset defined as the measured value minus the predicted value. Bias was also assessed through examination of plots between measured and predicted values.

Parametric modelling

All parametric analyses were undertaken using the general linear model procedure within the SAS software package (SAS-Institute-Inc. 2008). Multiple regression models were used to develop the 14 models of the 300 Index, described above. For each model, variables were introduced sequentially starting with the variable that exhibited the strongest correlation, until further additions were not significant, or did not substantially improve the model precision. Variable selection was undertaken manually, one variable at a time, and plots of residuals were examined prior to variable addition to ensure that the variable was included in the model using the least biased functional form. The number of independent variables in each model ranged from two to five for the models without age and three to five for the models with age. The most precise model had four independent variables including age.

Model bias was examined through plotting predicted against measured values of the 300 Index and residual values of the 300 Index (measured 300 Index – predicted 300 Index) against predicted values and all independent variables in the model.

The most important variables were identified through comparison of F -values across all of the 14 models. Where a variable was included in more than one model, these F -values were averaged for that variable.

Non-parametric modelling

Non-parametric k -nearest neighbour (k -NN) and k -most similar neighbour (k -MSN) modelling techniques were used to impute the 300 Index for the validation plots based on the measurements in the reference dataset. Nearest neighbour approaches are widely used statistical techniques that have gained considerable popularity for

use in forest assessment (e.g. McRoberts 2012) since their first application to forestry (Tomppo and Katila 1991). Under nearest neighbour estimation, variables of interest (Y) are imputed for target elements, commonly pixels covering an area of interest, where Y has not been measured. Imputation of Y is based on auxiliary variables (X) that are known for all elements in the population (N) and are correlated with Y . A subset of the elements in N have paired observations of both X and Y and are referred to as the reference dataset. Imputation for a target element is estimated as a function of k Y values in the reference dataset that have X values that are closest, using a measure of statistical proximity, to the X values of the target element (Magnussen and Tomppo 2014).

All k -NN and k -MSN modelling was completed in the `yalImpute` package (Crookston and Finley 2008) of the R statistical software environment (R Development Core Team 2014). The k -NN models were developed using the random forest distance metric. Under the random forest approach, observations are considered similar if they tend to converge in the same terminal node in a suitably constructed collection of classification and regression trees (Breiman 2001; Liaw and Wiener 2012). The metric used to define statistical distance is calculated as one minus the proportion of trees where a target observation is in the same terminal node as a reference observation (Crookston and Finley 2008). An alternative approach to assigning donor observations is through using the k -MSN distance metric, which is derived using canonical correlation analysis to produce a weighting matrix for the selection of donors from the fitting dataset (Moeur and Stage 1995). The value of k was varied for each imputation, and the value that minimised the RMSE within the validation dataset was selected for prediction. The imputed value (Y) for a given target was calculated using the distance-weighted average of the k value nearest to that from reference observations. Models developed using a k value of 1 were also produced for comparison in this analysis.

A measure of variable importance was estimated by taking a scaled composite of the importance scores outputted by random forests. Importance was estimated by permuting the values for each predictor variable and assessing the impact of the permutation on the model error. Importance scores were estimated for all predictors that were included in the final non-parametric models of the 300 Index. To account for the random element, 200 iterations were completed and the importance scores for each variable were averaged.

Results

Variation in data

The mean and range were relatively similar between the fitting and validation datasets for key explanatory

variables used in the models. Stand age averaged 16.6 years, ranging from 2.8 to 37.8 years (Table 2). Of the five RapidEye spectral bands, the green band exhibited the greatest variation, ranging 40-fold within the fitting dataset. Variation in the ratios was greatest for the GR with values ranging 31-fold.

Environmental variation was relatively high across the study site. Mean annual air temperature averaged 11.2 °C ranging from 9.27 to 14.1 °C, while total annual rainfall averaged 1420 mm ranging from 1119 to 2216 mm. The majority of the terrain over which measurements were made was relatively flat with slope averaging 5.24°. The slope ranged from 0° to 31.9° (Table 2) indicating that some measurements were taken on much steeper sites.

LiDAR metrics ranged widely, with P_{95} averaging 21.1 m and ranging 16-fold from 0.79 to 43.4 m (Table 2). Mean

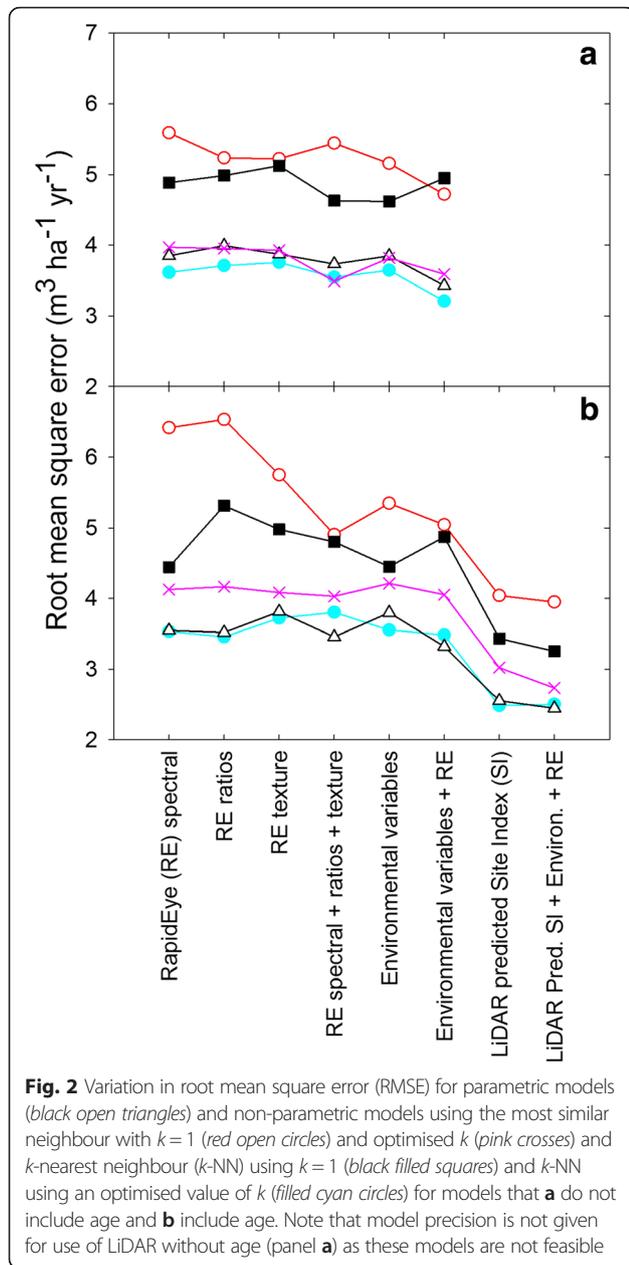
values for P_{05} and P_{50} were, respectively, 0.35 m (range 0.015–6.81 m) and 14.0 m (range 0.16–33.8 m).

Model comparison

For all 14 non-parametric models, the use of k -NN with an optimised value of k produced a RMSE that was considerably lower (Fig. 2) than for models where $k=1$ (mean gains in RMSE = 1.21 m³ ha⁻¹ year⁻¹, range = 0.75–2.01 m³ ha⁻¹ year⁻¹), or the use of the k -MSN with $k=1$ (mean gains = 1.83 m³ ha⁻¹ year⁻¹, range = 1.10–3.08 m³ ha⁻¹ year⁻¹) or an optimised value of k was used (mean gains = 0.37 m³ ha⁻¹ year⁻¹, range –0.056–0.711 m³ ha⁻¹ year⁻¹) (Fig. 2). The mean value of optimised k for models using random forest was 18 (range 8–48) while for k -MSN models, the mean value of k was 35 (range 9–58). Given the substantial increases in precision obtained using k -NN with an optimised value for k ,

Table 2 Variation in stand age, RapidEye reflectance and ratios, key environmental variables and key LiDAR metrics for the fitting ($n = 433$) and validation datasets ($n = 60$)

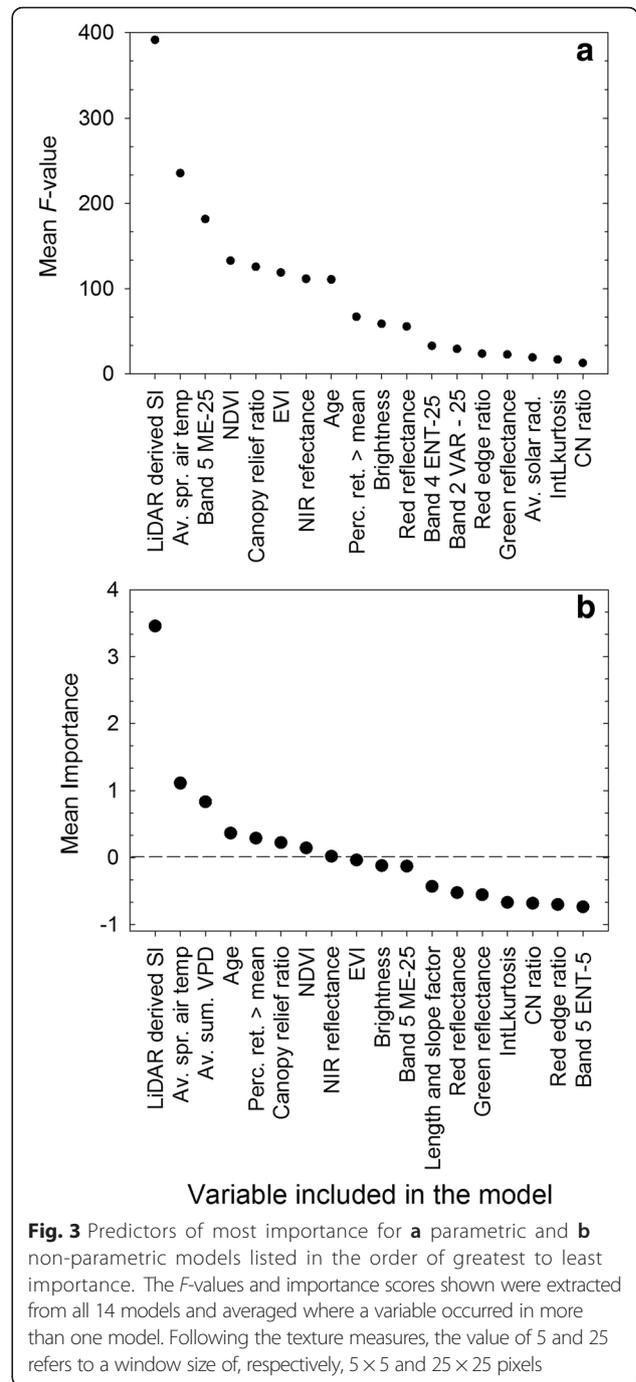
Variable	Fitting dataset		Validation dataset	
	Mean	Range	Mean	Range
Stand age (years)	16.6	2.84–37.8	17.8	3.89–32.8
RapidEye spectral				
Blue (band 1)	0.0736	0.058–0.122	0.0719	0.059–0.114
Green (band 2)	0.0031	6×10^{-4} –0.024	0.0029	8×10^{-4} –0.013
Red (band 3)	0.0373	0.024–0.123	0.0349	0.026–0.111
Red edge (band 4)	0.0787	0.061–0.149	0.0768	0.058–0.136
Near infrared (band 5)	0.235	0.156–0.366	0.235	0.180–0.329
RapidEye ratios				
EVI	0.546	0.226–0.817	0.551	0.276–0.753
NDVI	0.726	0.284–0.846	0.737	0.340–0.842
RE	0.0065	0.002–0.045	0.0065	0.002–0.029
SR	6.81	1.78–12.0	7.17	2.03–11.7
GR	98.9	10.9–336	103	15.8–279
VI	0.0814	0.020–0.324	0.0826	0.025–0.314
Brightness	0.0855	0.066–0.134	0.0843	0.071–0.117
Environmental variables				
Av. air temp. (°C)	11.2	9.27–14.1	11.2	9.86–13.7
Av. spr. air temp. (°C)	10.5	8.47–13.0	10.6	9.01–12.6
Av. rad. (MJ m ² day ⁻¹)	19.2	12.1–21.1	19.2	13.8–21.3
Av. spr. rad. (MJ m ² day ⁻¹)	23.9	15.1–25.8	23.9	18.4–26.0
Ann. rainfall (mm)	1420	1119–2216	1444	1158–2144
Slope (degrees)	5.24	0–31.9	6.50	0–20.6
LiDAR metrics				
P_{05} (m)	0.35	0.015–6.81	0.41	0.027–3.95
P_{50} (m)	14.0	0.16–33.8	15.6	0.258–30.7
P_{80} (m)	18.2	0.37–39.8	20.0	1.60–37.6
P_{95} (m)	21.1	0.79–43.4	22.9	2.54–41.4



further results were compared using this non-parametric method to parametric models for prediction of the 300 Index. There is a small computational cost associated with larger values of k , but this was easily offset by the increased model precision.

Of the three sources of information used, models generated from environmental data were the least precise in general, with RMSE values averaging $3.71 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$ (range $3.55\text{--}3.85 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$). The most important environmental predictor was mean air temperature during spring (Fig. 3).

Models created from satellite imagery generally had a slightly lower RMSE than those created from



environmental data (Fig. 2). Of the satellite imagery-derived models, the most precise were generated using individual band values (mean RMSE $3.64 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$) followed by vegetation indices (mean RMSE $3.67 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$). Textural metrics produced the least precise models (mean RMSE $3.79 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$) (Fig. 2). The use of all available metrics from satellite imagery resulted in models with equivalent precision

to those created using individual band values (mean RMSE = 3.64 m). For parametric models, the most important bands were NIR, red reflectance and green reflectance (Fig. 3). Addition of age to these models markedly improved the RMSE of parametric models (mean reductions in RMSE of 0.28 m³ ha⁻¹ year⁻¹) with lesser effect on non-parametric models (mean reductions in RMSE of 0.03 m³ ha⁻¹ year⁻¹).

Further modest gains occurred when metrics derived from satellite imagery were added to environmental variables (Fig. 2). These models had mean RMSE of 3.62 m³ ha⁻¹ year⁻¹. The addition of age as a variable markedly improved the RMSE of parametric (mean reductions in RMSE of 0.64 m³ ha⁻¹ year⁻¹) and to a lesser extent non-parametric models (mean reductions in RMSE of 0.23 m³ ha⁻¹ year⁻¹).

Of the data sources investigated, models developed with LiDAR data that included age had the highest precision (Fig. 2). The mean RMSE for the two model types was 2.52 m³ ha⁻¹ year⁻¹ (range 2.48–2.55 m³ ha⁻¹ year⁻¹). Both *F*-values and model importance scores show that Site Index derived from LiDAR to be the most useful predictor of the 300 Index of the LiDAR metrics (Fig. 3). The addition of environmental variables and metrics derived from satellite imagery to these models provided modest precision gains, with RMSE declining by on average 0.04 m³ ha⁻¹ year⁻¹. The model with the highest precision was a parametric model using stand age that included variables derived from all three data sources and had an *R*² of 0.790 and RMSE of 2.45 m³ ha⁻¹ year⁻¹.

For all six models that did not include age, non-parametric models were more precise than parametric models (Fig. 2). Expressed as a percentage of the parametric RMSE, the RMSE for non-parametric models was on average 94.6 % (range 92.9–97.0 %). For models that did include age, non-parametric models were more precise for five of the eight models, but overall, the mean RMSE between the two model types was very similar (Fig. 2).

Non-parametric models were more biased than parametric models for all but two of the constructed models (Fig. 4). For models without age, mean error averaged 0.530 and 0.384 m³ ha⁻¹ year⁻¹, respectively, for non-parametric and parametric models while for models with age, mean error averaged 0.613 and 0.451 m³ ha⁻¹ year⁻¹, respectively, for non-parametric and parametric models.

Inclusion of stand age as an explanatory variable was far more effective when included within parametric models than non-parametric models. Expressed as a percentage of the RMSE for models without age, the RMSE for models with age averaged 94.5 % (range 88.2–98.8 %) for parametric models and averaged 101.0 % (range 93.2–107.0 %) for non-parametric models. Results show

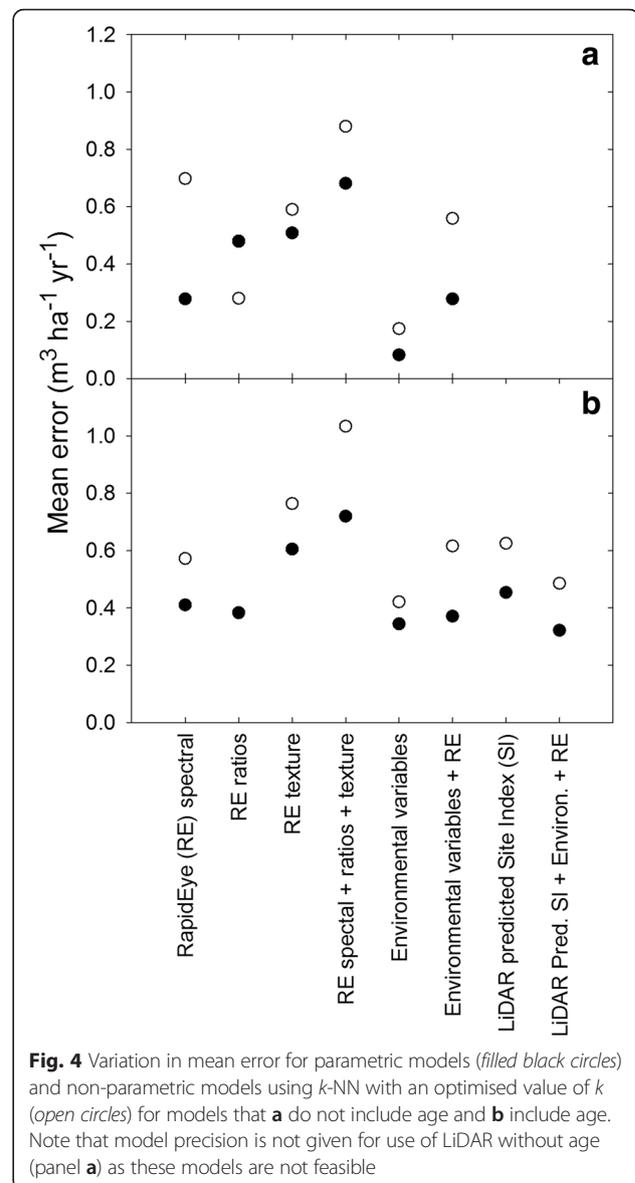


Fig. 4 Variation in mean error for parametric models (filled black circles) and non-parametric models using *k*-NN with an optimised value of *k* (open circles) for models that **a** do not include age and **b** include age. Note that model precision is not given for use of LiDAR without age (panel **a**) as these models are not feasible

that age was the eighth most important predictor for parametric models and fourth most important for non-parametric models (Fig. 3).

The *R*² and RMSE (in brackets) for various parametric or non-parametric models that included stand age were, respectively, (i) 0.582 (3.46 m³ ha⁻¹ year⁻¹) or 0.506 (3.81 m³ ha⁻¹ year⁻¹) for all metrics derived from satellite imagery; (ii) 0.495 (3.80 m³ ha⁻¹ year⁻¹) or 0.569 (3.55 m³ ha⁻¹ year⁻¹) for environmental layer variables; (iii) 0.614 (3.32 m³ ha⁻¹ year⁻¹) or 0.586 (3.48 m³ ha⁻¹ year⁻¹) for variables derived from a combination of satellite imagery and environmental layers; (iv) 0.772 (2.55 m³ ha⁻¹ year⁻¹) or 0.789 (2.48 m³ ha⁻¹ year⁻¹) for LiDAR metrics; and (v) 0.790

($2.45 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$) or $0.786 (2.50 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1})$ for all available variables. Plots of measured against predicted values for these ten models generally show parametric models to be relatively unbiased although there was slight underprediction of the 300 Index at higher values for the five models (Fig. 5a–e). This underprediction was also evident for the non-parametric models, and bias for these models was greater than that for parametric models (Fig. 5).

Discussion

Of the three data sources considered in this study, results clearly show that LiDAR provides the most precise predictions of the 300 Index. The use of variables extracted from environmental layers was found to provide the least precise estimates of the 300 Index while the use of variables derived from satellite imagery was of intermediate precision. Combining different data sources only resulted in modest gains in predictive precision of

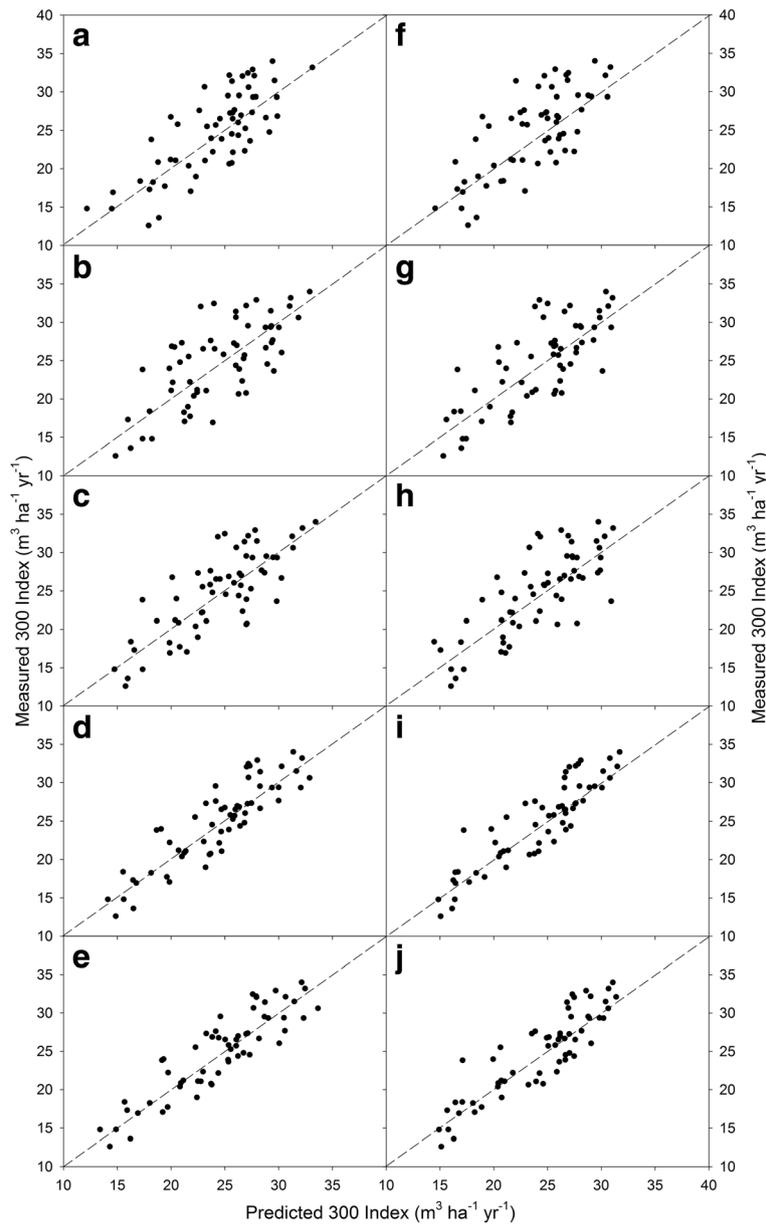


Fig. 5 Relationship between measured and predicted 300 Index for parametric models (a–e) and non-parametric k -NN (with optimised values of k) models (f–j) with stand age that include a, f variables derived from satellite imagery; b, g variables derived from environmental layers; c, h variables derived from satellite imagery and environmental layers; d, i Site Index predicted from LiDAR; and e, j variables selected from all data sources. The 1:1 line is shown on all panels as a dashed line

the 300 Index. Results clearly show the importance of age as a predictive variable and demonstrate non-parametric models to be more precise than parametric models if age is not available as an independent variable.

The use of LiDAR as a technology for predicting stand attributes is widely accepted within forestry. Since the first application of LiDAR in forestry almost three decades ago, LiDAR data have been used to accurately predict stand height and volume (Coops et al. 2007; Means et al. 2000; Næsset 2002; Means et al. 1999; Watt et al. 2014; Dash et al. 2016; Melville et al. 2015). Correlations of moderate to high strength have been found between LiDAR metrics and basal area (Næsset 2002, 2004, 2005; Nord-Larsen and Schumacher 2012; Means et al. 1999; Means et al. 2000), diameter (Næsset 2002) and green crown height (Næsset and Økland 2002). Given that canopy height is the dimension predicted with the most precision by LiDAR, it is not surprising that recent research has shown that Site Index can be predicted with high precision by LiDAR in plantation species such as *Eucalyptus urograndis* (Packalén et al. 2011) if stand age is available.

Our results extend this research by showing that LiDAR can be used to precisely ($R^2 = 0.79$) predict the 300 Index when stand age is available. The use of predicted Site Index, which was based on stand age and a LiDAR height-based metric, was most useful for predictions of the 300 Index. This result is consistent with previous research as most of the variation in tree volume in LiDAR-based models is typically attributable to metrics describing tree-height percentiles using LiDAR (Watt and Watt 2013).

Variables derived from satellite imagery provided moderately precise estimates of the 300 Index. Although we are unaware of any research that has used satellite imagery to predict the 300 Index, previous research does show a moderate correlation between *P. radiata* stand volume and textural attributes ($R^2 = 0.64$) derived from high-resolution satellite imagery (Shamsoddini et al. 2013). The utility of models that use solely spectral information, or vegetation ratios, is often limited because as the canopy reaches closure, the spectral response flattens and is unresponsive to continued forest growth. This reduces the predictive ability of satellite-based models to estimate changes in stand volume (Donoghue and Watt 2006) and is particularly problematic in highly productive plantations.

Models generated using environmental layers were the least precise of those developed. However, the precision range for these models (RMSE range of 3.55–3.85 m³ ha⁻¹ year⁻¹) was similar to that of previous national New Zealand models of *P. radiata* 300 Index created from environmental layers (Palmer et al. 2009b) where the most precise model had

RMSE of 3.65 m³ ha⁻¹ year⁻¹. This consistency does highlight the utility of satellite imagery and LiDAR as an alternative more precise means of predicting the 300 Index.

Air temperature was the environmental variable with the greatest influence on the 300 Index. This is consistent with previous research that showed air temperature to be the most important determinant of *P. radiata* growth in New Zealand (Jackson and Gifford 1974; Hunter and Gibson 1984; Watt et al. 2010), with air temperature in most locations in New Zealand, including the central North Island, sub-optimal for growth under current climatic conditions (Kirschbaum and Watt 2011). Models of Site Index for plantation species growing outside of New Zealand have also frequently found air temperature to be an important determinant of Site Index (Sharma et al. 2012), but rainfall is at least as important as air temperature in many drier regions (Mohamed et al. 2014; Sabatia and Burkhart 2014).

Although non-parametric models were slightly more biased, our results show these models to be more precise than parametric models when stand age is not available. This is consistent with previous models developed for predicting Site Index. Such models demonstrate the utility of non-parametric methods particularly when there are a large number of predictors with complex model forms (Aertsen et al. 2010). Prediction of other stand metrics, such as height, biomass, volume and diameter distribution from remotely sensed data, also often show non-parametric methods such as *k*-NN to outperform parametric regression (Packalén and Maltamo 2006; Zhou et al. 2011; Maltamo et al. 2009; Bollandsås et al. 2013) but not always (Mora et al. 2013; Penner et al. 2013). Pierce et al. (2009) found that a variant of *k*-NN, the gradient nearest neighbour, generally performed better than linear regression models or classification trees in temperate marine forests but that regression performed better in the more complex forests of adjacent temperate steppe and Mediterranean forests. This result suggests that the best technique may be reliant on the characteristics of the particular forest or landscape. In a recent review, Brosofske et al. (2014) outline a number of important considerations when selecting the most appropriate model type for prediction of forest inventory attributes from remotely sensed data.

These results confirm that the selection of an appropriate value of *k* is an important decision made by the analyst, which is dependent on the end uses of the model outputs. Small values of *k* are more susceptible to noise in the training data. The error rate decreases as the value of *k* increases, but both the likelihood of distant observations being included and the over-smoothing of model outputs increase. Finding an appropriate balance between these two is a recurring issue

and the focus of substantial interest (Dash et al. 2015; Ghosh 2006; McRoberts 2012; McRoberts et al. 2015). Numerous algorithms for selecting suitable values of k are available, but selecting k to optimise a model criterion such as RMSE is a logical approach.

Stand age was a useful determinant of the 300 Index. Without stand age, LiDAR data is of little use in predictions of the 300 Index as height percentiles cannot be adjusted to the age at which the 300 Index is determined. Results show that if age is not available, then the combination of data derived from satellite imagery and environmental variables provides a more precise means of estimating the 300 Index ($R^2 = 0.65$ and $RMSE = 3.21 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$ for the non-parametric model).

Acquiring LiDAR data is very expensive from an operational perspective. Results suggest that a combination of satellite imagery and available auxiliary environmental data may provide a cost-effective alternative for assessing the spatial variability of the 300 Index across planted forests. This methodology is likely to be particularly useful for regional- or national-scale predictions of the 300 Index or under circumstances when stand age is not readily available.

Conclusions

In conclusion, our results demonstrate that LiDAR is the most useful data source for predicting the 300 Index when stand age is available. Using these data, the most accurate model had an R^2 of 0.789 and RMSE of $2.48 \text{ m}^3 \text{ ha}^{-1} \text{ year}^{-1}$. In the absence of stand age, non-parametric models that used variables derived from both satellite imagery and auxiliary environmental data were found to be the most precise.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MSW was the primary author and undertook all the parametric analyses. JD was the secondary author and undertook all the non-parametric analyses. PW and SB prepared the satellite imagery for analysis and contributed to the writing. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by the 'Growing Confidence in Forestry's Future' programme which is jointly funded by the New Zealand Ministry of Business, Innovation and Employment and the New Zealand Forest Growers Levy Trust. Duncan Harrison of Scion is acknowledged for producing Fig. 1. We are grateful to Timberlands Ltd. for supplying the plot information and LiDAR dataset. The authors also gratefully acknowledge the diligent and professional contribution of all field teams involved. All research carried out in this manuscript follows the local, national or international guidelines and legislation and the required or appropriate permissions and/or licences for the study.

Author details

¹Scion, PO Box 29237, Fendalton, Christchurch, New Zealand. ²Scion, PO Box 3020, Rotorua, New Zealand. ³Indufor Asia-Pacific Ltd, PO Box 105039, Auckland, New Zealand.

Received: 3 November 2015 Accepted: 19 April 2016

Published online: 06 May 2016

References

- Aertsen, W, Kint, V, van Orshoven, J, Ozkan, K, & Muys, B (2010). Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological Modelling*, 221(8), 1119–1130. doi:10.1016/j.ecolmodel.2010.01.007.
- Baraldi, A, & Parmiggiani, F (1995). An investigation of the textural characteristics associated with gray-level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 293–304. doi:10.1109/36.377929.
- Bollandsås, OM, Maltamo, M, Gobakken, T, & Næsset, E. (2013). Comparing parametric and non-parametric modelling of diameter distributions on independent data using airborne laser scanning in a boreal conifer forest. *Forestry*, 86(4), 493–501.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Brosfokske, KD, Froese, RE, Falkowski, MJ, & Banskota, A (2014). A review of methods for mapping and prediction of inventory attributes for operational forest management. *Forest Science*, 60(4), 733–756. doi:10.5849/forsci.12-134.
- Coops, NC, Hilker, T, Wulder, MA, St-Onge, B, Newnham, G, Siggins, A, et al. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees-Structure and Function*, 21, 295–310.
- Crookston, NL, & Finley, AO. (2008). yalmpute: an R package for kNN imputation. *Journal of Statistical Software*, 23(10), 1–16.
- Dash, JP, Marshall, HM, & Rawley, B (2015). Methods for estimating multivariate stand yields and errors using k-NN and aerial laser scanning. *Forestry*, 88(2), 237–247. doi:10.1093/forestry/cpu054.
- Dash, JP, Watt, MS, Bhandari, S, & Watt, P. (2016). Characterising forest structure using combinations of airborne laser scanning data, RapidEye satellite imagery, and environmental variables. *Forestry*, 89(2), 159–169. doi:10.1093/forestry/cpv048.
- Donoghue, DNM, & Watt, PJ. (2006). Using LiDAR to compare forest height estimates from IKONOS and Landsat ETM+ data in Sitka spruce plantation forests. *International Journal of Remote Sensing*, 27(11), 2161–2175.
- Ghosh, AK. (2006). On optimum choice of k in nearest neighbor classification. *Computational Statistics and Data Analysis*, 50(11), 3113–3123.
- Haralick, RM, Shanugam, K, Dinstein, I. (1973) Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics* (Volume: SMC-3, Issue 6), pages 610–621 doi:10.1109/TSMC.1973.4309314.
- Hewitt, AE. (1993). *New Zealand Soil Classification (Landcare Research Science Series 1): Manaaki Whenua-Landcare Research*.
- Holmgren, J, Nilsson, M, & Olsson, H. (2003). Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science*, 49, 419–428.
- Hunter, IR, & Gibson, AR. (1984). Predicting *Pinus radiata* site index from environmental variables. *New Zealand Journal of Forestry Science*, 14(1), 53–64.
- Hyypä, J, Kelle, O, Lehtikoinen, M, & Inkinen, M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 969–975.
- Jackson, DS, & Gifford, HH. (1974). Environmental variables influencing the increment of radiata pine. (1) Periodic volume increment. *New Zealand Journal of Forestry Science*, 4(1), 3–26.
- Kayitakire, F, Hamel, C, & Defourny, P (2006). Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, 102(3–4), 390–401. doi:10.1016/j.rse.2006.02.022.
- Kimberley, MO, West, G, Dean, M, & Knowles, L. (2005). The 300 Index—a volume productivity index for radiata pine. *New Zealand Journal of Forestry*, 50, 13–18.
- Kirschbaum, MUF, & Watt, MS (2011). Use of a process-based model to describe spatial variation in *Pinus radiata* productivity in New Zealand. *Forest Ecology and Management*, 262(6), 1008–1019. doi:10.1016/j.foreco.2011.05.036.
- Leathwick, JR, & Stephens, RTT. (1998). *Climate surfaces for New Zealand* (Landcare Res. Contract Report LC9798/126, p. 19). Lincoln: Landcare Research.
- Leathwick, J, Morgan, F, Wilson, G, Rutledge, D, McLeod, M, & Johnston, K. (2002). *Land environments of New Zealand: a technical guide* (p. 184). Hamilton: Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research.
- Leathwick, J, Wilson, G, Rutledge, D, Wardle, P, Morgan, F, Johnston, K, et al. (2003). *Land environments of New Zealand*. Hamilton: Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research.
- Liaw, A, & Wiener, M. (2012). *Breiman and Cutler's random forests for classification and regression* (46-7th ed.).

- Lim, KS, & Treitz, PM. (2004). Estimation of aboveground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, 19, 558–570.
- Lu, D. (2005). Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *International Journal of Remote Sensing*, 26(12), 2509–2525. doi:10.1080/01431160500142145.
- Magnussen, S, & Tomppo, E. (2014). The k-nearest neighbor technique with local linear regression. *Scandinavian Journal of Forest Research*, 29(2), 120–131.
- Maltamo, M, Peuhkurinen, J, Malinen, J, Vauhkonen, J, Packalén, P, & Tokola, T. (2009). Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. *Silva Fennica*, 43, 507–521.
- McRoberts, RE. (2012). Estimating forest attribute parameters for small areas using nearest neighbors techniques. *Forest Ecology and Management*, 272, 3–12.
- McRoberts, RE, & Tomppo, E. (2007). Remote sensing support for national forest inventories. *Remote Sensing of Environment*, 110, 412–419.
- McRoberts, RE, Næsset, E, & Gobakken, T. (2015). Optimizing the k-Nearest Neighbors technique for estimating forest aboveground biomass using airborne laser scanning data. *Remote Sensing of Environment*, 163, 13–22.
- Means, JE, Acker, SA, Harding, DJ, Blair, JB, Lefsky, MA, Cohen, WB, et al. (1999). Use of large-footprint scanning airborne LiDAR to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment*, 67(3), 298–308.
- Means, JE, Acker, SA, Fitt, BJ, Renslow, M, Emerson, L, & Hendrix, CJ. (2000). Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering and Remote Sensing*, 66(11), 1367–1371.
- Melville, G, Stone, C, & Turner, R. (2015). Application of LiDAR data to maximise the efficiency of inventory plots in softwood plantations. *New Zealand Journal of Forestry Science*, 45, 9.
- Mitchell, ND. (1991). The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). *Journal of the Royal Society of New Zealand*, 21, 13–24.
- Moer, M, & Stage, A. (1995). Most similar neighbour: an improved sampling inference procedure for natural resource planning. *Forest Science*, 41, 337–359.
- Mohamed, A, Reich, RM, Khosla, R, Aguirre-Bravo, C, & Briseno, MM (2014). Influence of climatic conditions, topography and soil attributes on the spatial distribution of site productivity index of the species rich forests of Jalisco, Mexico. *Journal of Forestry Research*, 25(1), 87–95. doi:10.1007/s11676-014-0434-5.
- Mora, B, Wulder, MA, White, JC, & Hobart, G (2013). Modeling stand height, volume, and biomass from very high spatial resolution satellite imagery and samples of airborne LiDAR. *Remote Sensing*, 5(5), 2308–2326. doi:10.3390/Rs5052308.
- Næsset, E (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1), 88–99. doi:10.1016/S0034-4257(01)00290-5.
- Næsset, E (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19(2), 164–179. doi:10.1080/02827580310019257.
- Næsset, E (2005). Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 98(2–3), 356–370. doi:10.1016/j.rse.2005.07.012.
- Næsset, E, & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, 79(1), 105–115.
- Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne LiDAR system. *Remote Sensing of Environment*, 56(1), 1–7.
- Nord-Larsen, T, & Schumacher, J. (2012). Estimation of forest resources from a country wide laser scanning survey and national forest inventory data. *Remote Sensing of Environment*, 119, 148–157. doi:10.1016/j.rse.2011.12.022.
- Ozdemir, I (2008). Estimating stem volume by tree crown area and tree shadow area extracted from pan-sharpened Quickbird imagery in open Crimean juniper forests. *International Journal of Remote Sensing*, 29(19), 5643–5655. doi:10.1080/01431160802082155.
- Packalén, P, & Maltamo, M. (2006). Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. *Forest Science*, 52(6), 611–622.
- Packalén, P, Mehtätalo, L, & Maltamo, M. (2011). ALS-based estimation of plot volume and site index in a eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. *Annals of Forest Science*, 68(6), 1085–1092.
- Palmer, DJ. (2008). *Development of national extent terrain attributes (TANZ), soil water balance surfaces (SWatBal), and environmental surfaces, and their application for spatial modelling of Pinus radiata productivity across New Zealand* (PhD thesis, p. 410). Hamilton: University of Waikato.
- Palmer, D, Höck, B, Kimberley, M, Watt, M, Lowe, D, & Payn, T (2009a). Comparison of spatial prediction techniques for developing *Pinus radiata* productivity surfaces across New Zealand. *Forest Ecology and Management*, 258(9), 2046–2055. doi:10.1016/j.foreco.2009.07.057.
- Palmer, DJ, Höck, B, Kimberley, MO, Watt, MS, Lowe, DJ, & Payn, TW. (2009b). Comparison of spatial prediction techniques for developing *Pinus radiata* productivity surfaces across New Zealand. *Forest Ecology and Management*, 258(9), 2046–2055.
- Palmer, DJ, Watt, MS, Höck, BK, & Lowe, DJ. (2009). A dynamic framework for spatial modelling *Pinus radiata* soil water balance (SWatBal) across New Zealand. *Scion Bulletin*, 234, 93.
- Penner, M, Pitt, D, & Woods, M. (2013). Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. *Canadian Journal of Remote Sensing*, 39(5), 426–443.
- Pierce, KB, Ohmann, JL, Wimberly, MC, Gregory, MJ, & Fried, JS. (2009). Mapping wildland fuels and forest structure for land management: a comparison of nearest neighbor imputation and other methods. *Canadian Journal of Forest Research*, 39, 1901–1916.
- Popescu, SC, Wynne, RH, & Scriveri, JA. (2004). Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. *Forest Science*, 50, 551–565.
- R Development Core Team. (2014). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- RapidEye AG. (2011). Satellite imagery product specifications. 3.2 edn. RapidEye AG. http://www.planet.com/assets/themes/planet/pdf/1601.RapidEye.Image.Product.Specs_Jan16_V6.1_ENG.pdf.
- Rombouts, J, Ferguson, IS, & Leech, JW. (2010). Campaign and site effects in LiDAR prediction models for site-quality assessment of radiata pine plantations in South Australia. *International Journal of Remote Sensing*, 31(5), 1155–1173.
- Sabatia, CO, & Burkhart, HE (2014). Predicting site index of plantation loblolly pine from biophysical variables. *Forest Ecology and Management*, 326, 142–156. doi:10.1016/j.foreco.2014.04.019.
- Saremi, H, Kumar, L, Stone, C, Melville, G, & Turner, R. (2014). Sub-compartment variation in tree height, stem diameter and stocking in a *Pinus radiata* D.Don plantation examined using airborne LiDAR data. *Remote Sensing*, 6(8), 7592–7609.
- SAS-Institute-Inc. (2008). *SAS/STAT 9.2 User's Guide*. Cary: SAS Institute Inc.
- Shamsoddini, A, Trinder, JC, & Turner, R (2013). Pine plantation structure mapping using WorldView-2 multispectral image. *International Journal of Remote Sensing*, 34(11), 3986–4007. doi:10.1080/01431161.2013.772308.
- Sharma, RP, Brunner, A, & Eid, T (2012). Site index prediction from site and climate variables for Norway spruce and Scots pine in Norway. *Scandinavian Journal of Forest Research*, 27(7), 619–636. doi:10.1080/02827581.2012.685749.
- Skovsgaard, JP, & Vanclay, JP. (2008). Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. *Forestry*, 81, 13–31.
- Solberg, AHS (1999). Contextual data fusion applied to forest map revision. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1234–1243. doi:10.1109/36.763280.
- Tait, A, Henderson, R, Turner, R, & Zheng, Z. (2006). Thin plate smoothing interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *International Journal of Climatology*, 26, 2097–2115.
- Tomppo, E, & Katila, M. (1991). Satellite image-based national forest inventory of finland for publication in the igarss' 91 digest. In *Geoscience and Remote Sensing Symposium*, 1991. IGARSS'91. Remote Sensing: Global Monitoring for Earth Management, International (Vol. 3, pp. 1141–1144). IEEE.
- Tuominen, S, & Pekkarinen, A (2005). Performance of different spectral and textural aerial photograph features in multi-source forest inventory. *Remote Sensing of Environment*, 94(2), 256–268. doi:10.1016/j.rse.2004.10.001.
- Véga, C, & St-Onge, B. (2009). Mapping site index and age by linking a time series of canopy height models with growth curves. *Forest Ecology and Management*, 257(3), 951–959.
- Watt, MS, & Palmer, DJ (2012). Use of regression kriging to develop a Carbon: Nitrogen ratio surface for New Zealand. *Geoderma*, 183, 49–57. doi:10.1016/j.geoderma.2012.03.013.
- Watt, PJ, & Watt, MS. (2013). Development of a national model of tree volume from LiDAR metrics for New Zealand. *Remote Sensing of Environment*, 34, 5892–5904.
- Watt, M, Palmer, D, Kimberley, M, Hock, B, Payn, T, & Lowe, D (2010). Development of models to predict *Pinus radiata* productivity throughout New Zealand. *Canadian Journal of Forest Research*, 40(3), 488–499. doi:10.1139/X09-207.

- Watt, MS, Meredith, A, Watt, P, & Gunn, A. (2013). Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations. *New Zealand Journal of Forestry Science*, 43(1), 18.
- Watt, P, Meredith, A, Yang, C, & Watt, MS. (2013). Development of regional models of *Pinus radiata* height from GIS spatial data supported with supplementary satellite imagery. *New Zealand Journal of Forestry Science*, 43(1), 11.
- Watt, MS, Meredith, A, Watt, P, & Gunn, A. (2014). The influence of LiDAR pulse density on the precision of inventory metrics in young unthinned Douglas-fir stands during initial and subsequent LiDAR acquisitions. *New Zealand Journal of Forestry Science*, 44, 18.
- Watt, MS, Dash, JP, Bhandari, S, & Watt, P (2015). Comparing parametric and non-parametric methods of predicting Site Index for radiata pine using combinations of data derived from environmental surfaces, satellite imagery and airborne laser scanning. *Forest Ecology and Management*, 357, 1–9. doi:10.1016/j.foreco.2015.08.001.
- West, G, Moore, JR, Shula, RG, Harrington, JJ, Snook, J, Gordon, JA, et al. (2013). *Forest management DSS development in New Zealand. Paper presented at the Implementation of DSS Tools into the Forestry Practice*. Slovakia: Technical University of Zvolen.
- Zhou, GM, Xu, XJ, Du, HQ, Ge, HL, Shi, YJ, & Zhou, YF. (2011). Estimating aboveground carbon of Moso bamboo forests using the k nearest neighbors technique and satellite imagery. *Photogrammetric Engineering and Remote Sensing*, 77(11), 1123–1131. Accessed 4 February 2015
- Zvoleff, A. (2014). glcm: Calculate textures from grey-level co-occurrence matrices GLCMs) in R. R package version 1.0. <http://CRAN.R-project.org/package=glcm>. Accessed 4 February 2015.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
