Growing confidence in forestry's future Research Programme

SCION forests·products·innovation

Date: June 2015
Reference: GCFF TN - 007

Technical Note

Predicting productivity using combinations of LiDAR, satellite imagery and environmental data

Author/s: Michael S. Watt, Jonathan P. Dash, Pete Watt, Santosh Bhandari

Corresponding author: michael.watt@scionresearch.com

Summary: An understanding of how plantation productivity varies spatially is important for forest planning, management and projection of future plantation yields and returns. Site Index and 300 Index are productivity indices widely used to determine height and volume production, respectively, in *Pinus radiata* D. Don plantations. Although Site Index and 300 Index are routinely characterised at the stand level, little research has investigated if remotely sensed data sources can be used in combination with environmental surfaces to predict these metrics at fine spatial resolution.

This study uses an extensive dataset obtained from *P. radiata* plantations in the central North Island, New Zealand. The objective of this research was to compare the precision of models of Site Index and 300 Index that included variables from aerially acquired Light Detection and Ranging (LiDAR), 5 m satellite imagery or environmental surfaces, and combinations of these three data sources. Models were constructed both with and without stand age as a variable as managers may not always have access to stand age. A total of 14 models were constructed for each productivity index using data from 433 plots. Precision and bias of these models was determined using an independent dataset of 60 plots.

LiDAR was the most useful data source for precise and unbiased predictions for both Site Index and 300 Index. In the absence of LiDAR data, models constructed using variables derived from environmental surfaces and satellite imagery were found to be most precise for both Site Index and 300 Index. Predictions made using only environmental surfaces were less precise for both productivity metrics while predictions made using only satellite imagery were generally the least precise of any of the information sources. Stand age was found to be a very useful determinant of 300 Index and Site Index.

Introduction

Forest managers want to understand and quantify how environmental factors influence tree growth and site quality. Site quality is an important baseline for forest-level planning and can play a major role in formulating silviculture strategies, projecting future yields and determining the economic returns from a forest. Site quality can be most accurately inferred from characteristics of the trees.

Site Index is the most common measure of forest site productivity. It expresses the height of dominant trees at a given age and is relatively unaffected by stand density. However, it does not accurately reflect volume productivity of a site.

A national growth model has recently been developed based on the volume productivity index – 300 Index – for *Pinus radiata* D Don within New Zealand**.** The 300 Index is calculated by adjusting field plot measurements for age, stand density, and silvicultural history to give the mean annual increment at age 30 for a reference silvicultural regime.

Stand level estimates of 300 Index and Site Index are made by averaging plot values. However, the aggregated estimate may not be adequate to detect the local growing variations within stands that result from interactions between environmental factors. Consequently, spatial models that provide greater

supported by forestgrowers commodity levy

FOREST OWNERS ASSOCIATION

This programme is also funded by Scion and the Ministry of Business, Innovation and Employment, New Zealand.

definition of Site index and 300 Index have been developed using geographic information systems (GIS).

These environmental layers (e.g. soil, climate and genetic factors) have been successfully used to develop spatial surfaces describing Site Index and 300 Index for *P. radiata* [1]. This approach has considerable merit as forest managers generally have access to GIS and a number of potentially useful predictive variables.

Another option is using LiDAR (Light detection and ranging) to characterise the forest canopy in three dimensions. LiDAR technology has been widely used to spatially quantify variation in tree height and crown dimensions at resolutions ranging from the stand level to individual tree level over the last 20 years. LiDAR has also been used to predict Site Index [2,3]. but the utility of LiDAR in predicting 300 Index has not been investigated.

Satellite imagery, although less detailed than LiDAR, is a more cost effective means of predicting stand attributes such as tree height and volume across forest resources. However, the utility of satellite data in predicting productivity indices such as Site Index and 300 Index has never been investigated.

Stand age –an important determinant of Site Index and 300 Index – is often not publicly available at a national scale. Consequently, there is considerable interest in determining the accuracy of productivity index models when stand age is not available.

This work has compared the accuracy of models of Site Index and 300 Index that included variables from LiDAR, satellite imagery or environmental surfaces and combinations of these three data sources. Models were constructed either with or without age as a variable.

Method

2.1 Study site

This study used data from the Kaingaroa Forest which is located in the Central North Island of New Zealand (Fig. 1). From a total of 493 plots, 60 plots were randomly selected and used for model validation. The remaining 433 plots were used for the model fitting process.

2.2. Field Measurement

Field plots (0.06 ha circles) were located within the forest at the intersections of a grid that had a randomised start point and orientation. Diameter at breast height (dbh) was measured for all trees in a plot. Tree height was measured for a subset of healthy, well-formed trees selected from across the dbh range.

2.3 Derivation of Site Index

Site Index for *P. radiata* in New Zealand is defined as the mean top height at age 20 years. Field data was used to fit a regression between dbh and height that was used to predict the heights of unmeasured trees within each plot. This information was used to calculate mean top height, defined as the average height of the primary leaders of the 100 largest diameter trees per hectare. Using an appropriate height-age function and the stand age, the mean top height was estimated at age 20 to estimate Site Index.

2.4 Derivation of 300 Index

The 300 index is estimated using numerous models including a stand-level basal area growth model, a height/age function, a mortality function, stand-volume function and a thinning function. The models were used to estimate the 300 Index value for each field plot.

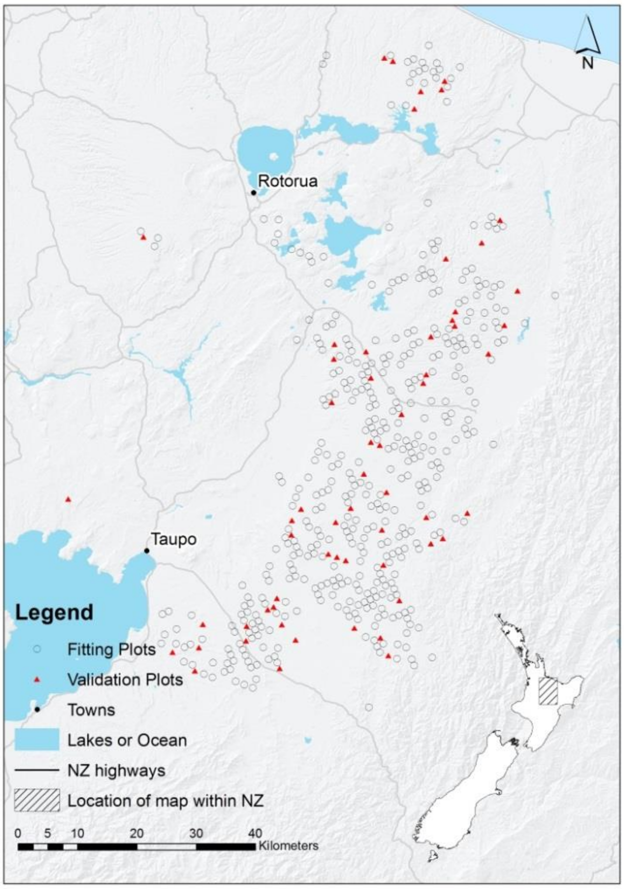

Figure 1. Map showing the location of the plots used for fitting and validation of the tested models.

2.5 Predictive variables used in the modelling

LiDAR data

The aerial LiDAR survey was completed in early 2014 using an Optech Pegasus scanner to collect a discrete, small footprint, dataset.

RapidEye (RE) imagery

The RE satellite system is a constellation of five satellites carrying identical sensors, all of which were launched at the end of 2008 [4]. RE imagery from January 16 and 28, 2014 were acquired over the study area. Spectral values, vegetation indices and texture metrics were extracted from the data for the plots used in this work.

Environmental surfaces

Environmental data for plot locations were obtained from biophysical GIS surfaces Key variables used in the analyses included mean annual and monthly air temperature, relative humidity, solar radiation, vapour pressure deficit and rainfall. Mean annual and seasonal root zone water storage, fractional available root zone water storage and the maximum available root zone water storage were determined for all plots. Soil fertility was represented by soil C:N ratio; a useful index of nitrogen mineralisation.

2.6 Analyses

Overview

Separate models of Site Index and 300 Index were created from the categories of data as listed in Table 1.

Table 1. The categories of data used to in creating Site Index and 300 Index models.

Data category	Age included	Age excluded
RE spectral values	✓	✓
RE vegetation indices (ratios)	✓	✓
RE textural metrics	✓	✓
All RE metrics	✓	✓
Environmental surface variables	✓	✓
All RE metrics and environmental surface variables	✓	✓
LiDAR variables	✓	*
LiDAR variables, all RE metrics and environmental surface variables	✓	*

*Predicting Site Index or 300 Index using LiDAR information without age is not practical.

Predictions of Site Index and 300 Index using LiDAR were made for each field plot

Modelling

All analyses were undertaken using the general linear model procedure (GLM) within SAS [5].

The precision of the models was assessed using the coefficient of determination (R^2) and root mean square error (RMSE). Bias was assessed through examination of plots between predicted and measured values.

3. Results

3.1 Model comparison – Site Index

The LiDAR models including age were the most precise (Fig 2). The R^2 for this model was 0.87 with RMSE of 1.66 m. Addition of environmental variables and metrics derived from satellite imagery to these models provided modest precision gains (Fig. 2).

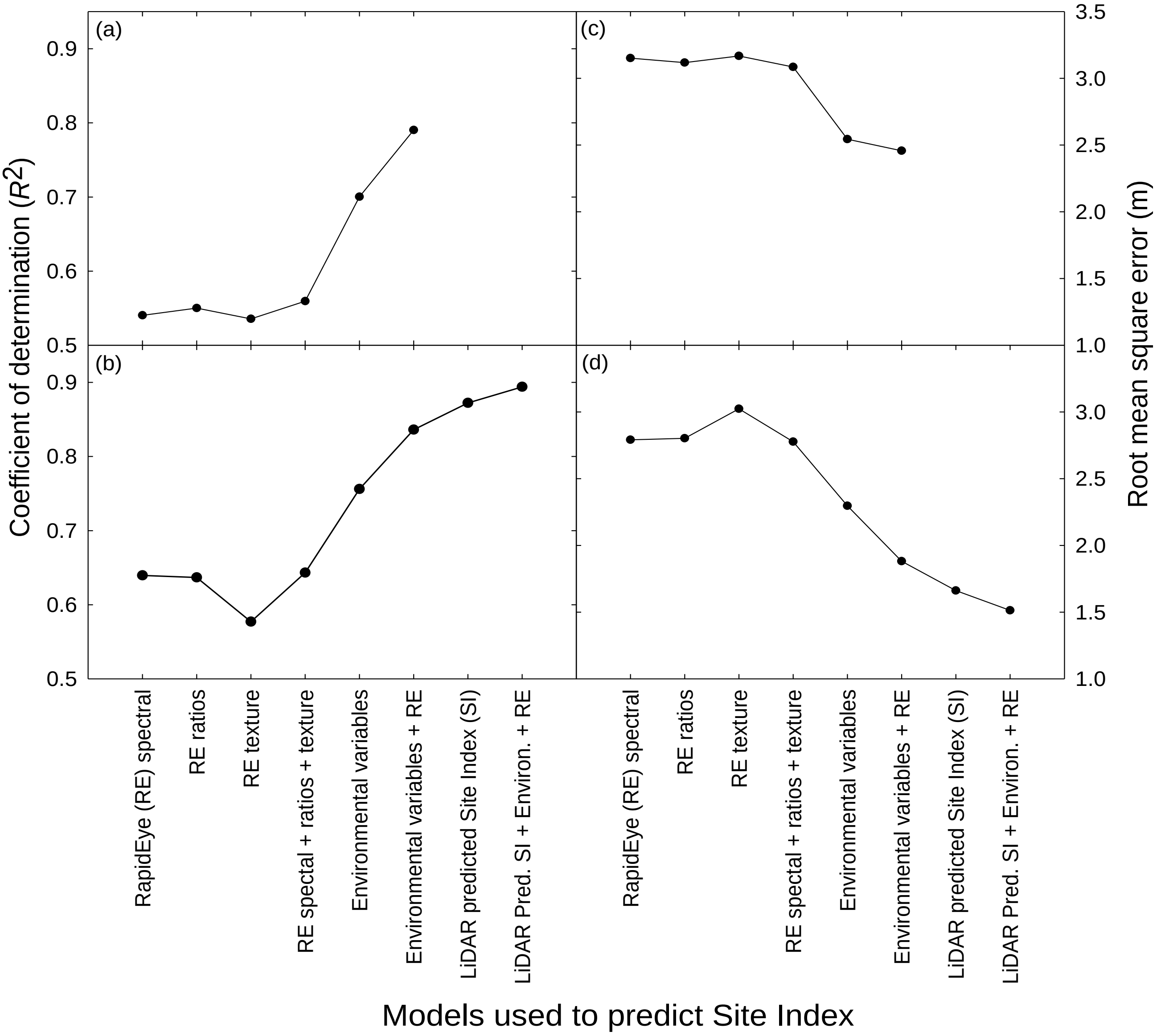

Figure 2. Variation in the (a, b) coefficient of determination (R^2) and (c, d) root mean square error (RMSE) of Site Index for models that (a, c) do not include age and (b, d) do include age as a predictive variable. Note that model precision is not given for use of LiDAR without age (panels a, c) as these models are not feasible.

Models that used environmental variables were not as precise as the LiDAR models (Fig. 2) but there were substantial gains in precision when data from satellite imagery was added to data from environmental surfaces (Fig. 2). Average spring air temperature was the environmental variable with most influence on Site Index.

The models created using metrics derived from satellite imagery had the lowest model precision (Fig. 2).

Inclusion of stand age as an explanatory variable improved the precision of all models.

3.2 Model comparison – 300 Index

The LiDAR models including age were the most precise (Fig. 3). The R^2 for this model was 0.77 while the RMSE was 2.55 m^3 ha^{-1} yr^{-1}. Addition of environmental variables and metrics derived from satellite imagery to these models provided modest precision gains (Fig. 3).

Models created from satellite imagery and environmental surfaces had similar precision and were less precise than the LiDAR models (Fig. 3). Average spring air temperature was the environmental variable with most influence on 300 Index.

Modest gains in precision occurred when metrics derived from satellite imagery were added to environmental variables (Fig. 3).

Addition of age improved the precision of all models.

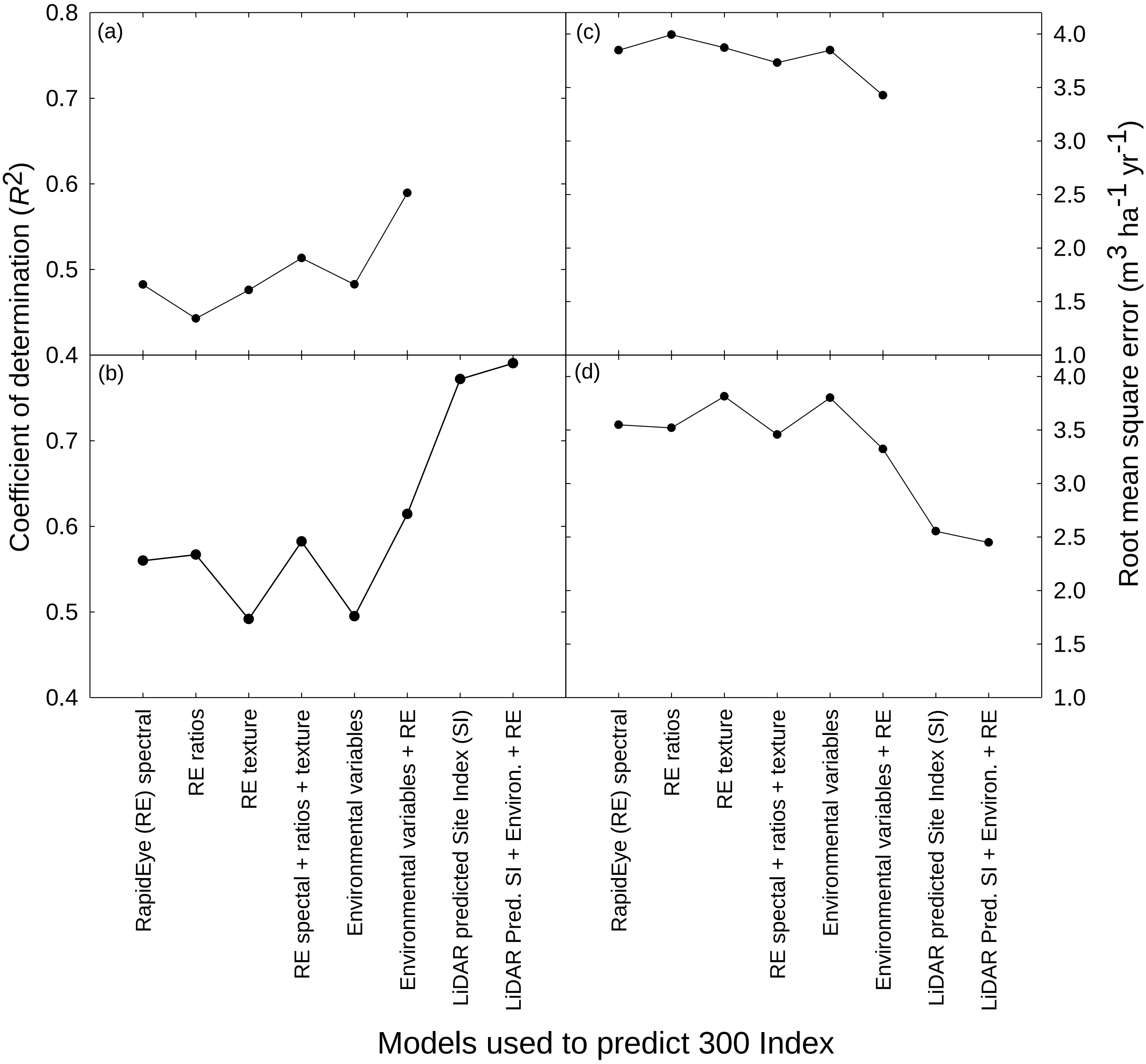

Figure 3. Variation in the (a, b) coefficient of determination (R^2) and (c, d) root mean square error (RMSE) for models of 300 Index that (a, c) do not explicitly include age and (b, d) do explicitly include age as a predictive variable. Note that model precision is not given for use of LiDAR without age (panels a, c) as these models are not feasible.

3.3 Model bias-Site Index and 300 Index

Plots of measured against predicted values for five models that included (i) all metrics derived from satellite imagery, (ii) environmental surface variables, (iii) variables derived from satellite imagery and environmental surfaces, (iv) LiDAR metrics and (v) all available variables) generally show the models to be relatively unbiased for both Site Index (Figs. 4a−e) and 300 Index (Figs. 4f−j).

Discussion

LiDAR provides the most precise predictions of Site Index and 300 Index. Models derived from variables extracted from environmental surfaces and satellite imagery were found to be of lower precision for both Site Index and 300 Index. The results clearly show the importance of age as a predictive variable for both Site Index and 300 Index.

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 [6,7,8,9,10].

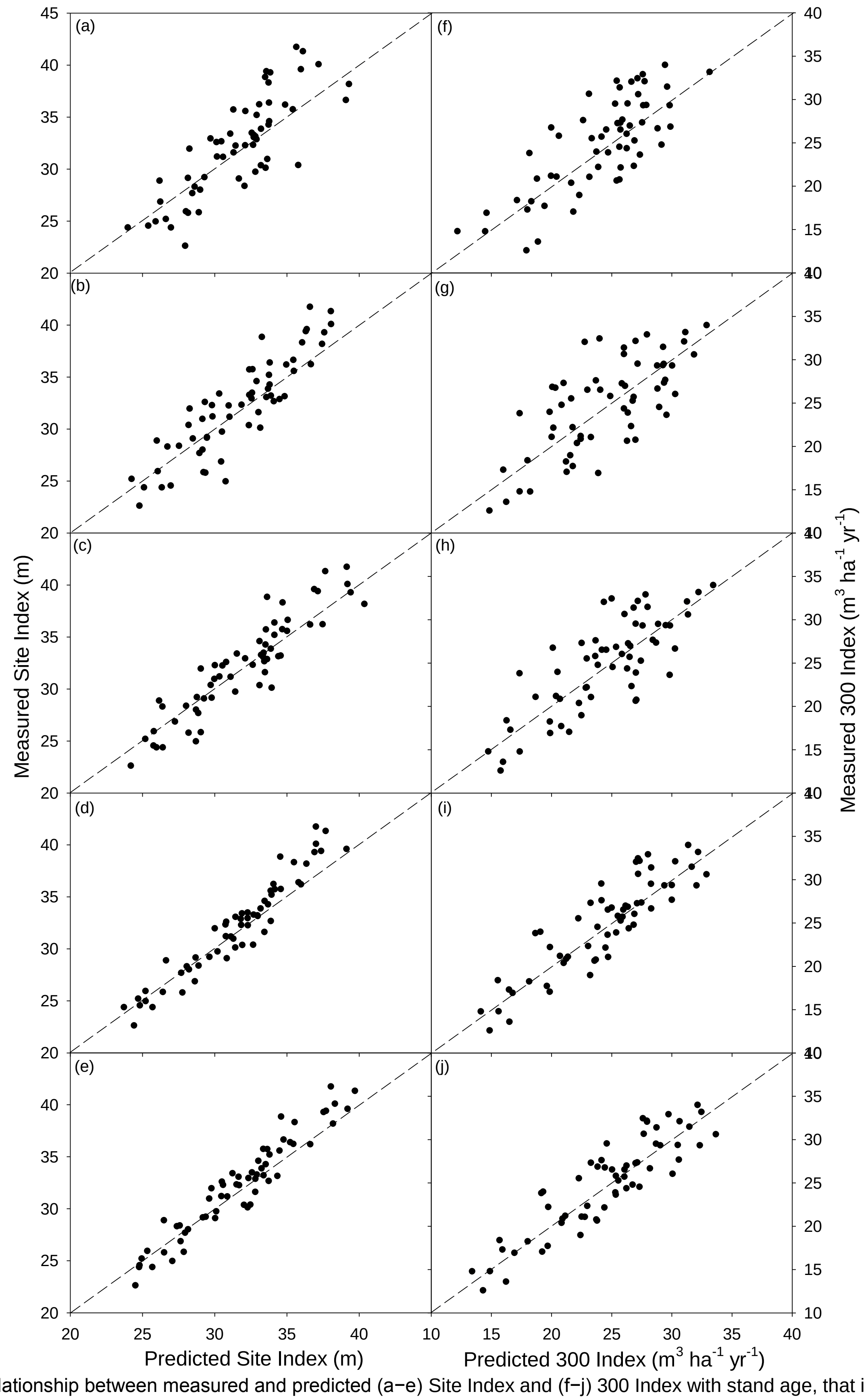

Figure 4. Relationship between measured and predicted (a−e) Site Index and (f−j) 300 Index with stand age, that include (a, f) variables derived from satellite imagery**,** (b, g) variables derived from environmental surfaces, (c, h) variables derived from satellite imagery and environmental surfaces, (d, i) Site Index predicted from LiDAR, (e, j) variables selected from all data sources. The 1:1 line is shown on all panels as a dashed line.

Given that canopy height is the dimension predicted with 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* [3 if stand age is available.

Our results extend this research by showing for the first time that LiDAR can be used to precisely predict 300 Index when stand age is available. Use of predicted Site Index, which was based on stand age and a LiDAR height based metric, was most useful for predictions. This result is consistent with previous research as most of the variation in volume in LiDAR based models is typically attributable to metrics describing LiDAR height percentiles [11].

Variables derived from satellite imagery provided moderately precise estimates of 300 Index and were the least precise for Site Index.

Models created using environmental surfaces were the least precise of those developed for 300 Index and of intermediate precision for Site Index. The precision range for these models was similar to that of previous national New Zealand models of *P. radiata* Site Index and 300 Index created from environmental surfaces where the most precise model had a respective RMSE of 2.70 m and 3.65 m^3 ha^{-1} yr^{-1} (Palmer *et al.*, 2009).

Stand age was found to be a very useful determinant of Site Index and 300 Index. Without stand age LiDAR data is of little use in predictions of these productivity metrics as height percentiles cannot be adjusted to the age at which Site Index and 300 Index are determined. Results show that, if age is not available, the combination of data derived from satellite imagery and environmental variables provides the most precise means of estimating both productivity indices.

Our results have clearly shown that LiDAR is the most useful data source for predicting Site Index and 300 Index when stand age is available. However, as LiDAR data is very expensive from an operational perspective, satellite imagery coupled with available environmental surfaces can be used as a cost-effective alternative for assessing the spatial variability of Site Index and 300 Index across planted forests. This methodology is likely to be particularly useful for regional or national scale predictions of Site Index or 300 Index or under circumstances when stand age is not readily available.

Acknowledgements

Funding for this research came from the “Growing Confidence in Forestry’s Future” research programme (C04X1306), which is jointly funded by the Ministry of Business Information and Employment (MBIE) and the Forest Growers Levy Trust, with the support of the NZ Forest Owners Association (FOA) and the NZ Farm Forestry Association (FFA). 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.

References

1. Palmer, D., Hock, B., Kimberley, M., Watt, M., Lowe, D., Payn, T., 2009. Comparison of spatial prediction techniques for developing Pinus radiata productivity surfaces across New Zealand. Forest Ecology and Management 258, 2046-2055.
2. Rombouts, J., Ferguson, I.S., Leech, J.W., 2010. Campaign and site effects in LiDAR prediction models for site-quality assessment of radiata pine plantations in South Australia. Int. J. Remote Sens. 31, 1155-1173.
3. 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, 1085-1092.
4. RapidEye AG, 2011. Satellite imagery product specifications. 3.2 edn. RapidEye AG.
5. SAS-Institute-Inc., 2008. SAS/STAT 9.2 User’s Guide. Cary NC: SAS Institute Inc.
6. Means, J.E., Acker, S.A., Harding, D.J., Blair, J.B., Lefsky, M.A., Cohen, W.B., Harmon, M.E., McKee, W.A., 1999. Use of large-footprint scanning airborne LiDAR to estimate forest stand characteristics in the Western Cascades of Oregon. Remote Sensing of Environment 67, 298-308.
7. Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning LiDAR. Photogrammetric Engineering and Remote Sensing 66, 1367-1371.
8. 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, 88-99.
9. Palmer, D., Hock, B., Kimberley, M., Watt, M., Lowe, D., Payn, T., 2009. Comparison of spatial prediction techniques for developing Pinus radiata productivity surfaces across New Zealand. Forest Ecology and Management 258, 2046-2055.