**Using hybrid physiological/mensurational modelling to predict site index of *Pinus sylvestris* L. in Sweden: A pilot study**

Euan G. Masona\*, Emma Holmströmb & Urban Nilssonb

aSchool of Forestry, University of Canterbury, Private Bag 4800, Christchurch. euan.mason@canterbury.ac.nz, corresponding author

bSouthern Swedish Forest Research Centre, Sveriges Lantbruksuniversitet, P.O Box 49, SE-23053 Alnarp, Sweden

**Abstract**

Precision and bias of a model designed to predict site index of Scots pine (*P. sylvestris* L.) from site variables in Sweden were tested using data from 1985 inventory plots. The model was biased and relatively imprecise with a standard error of 3.7 m. A new model was constructed using a fitting subset of the data, employing the sum of mean monthly estimates of photosynthetically active radiation modified by local monthly climatic conditions as a primary independent variable. The best model used a day-time temperature modifier to calculate potential radiation use efficiency. Modifiers for vapour pressure deficit and soil water did not add significantly to the model. Elevation and distance to the sea added small but significant improvements to the predictions. Phytometer indicators of nutritional fertility also slightly improved the fit. The final model had a standard error of 2.06 m for predictions of site index that ranged from 18-30 m at age 100. When applied to a validation subset of plots the model displayed a standard error of 2.09 m and very similar residual patterns to those observed during fitting. The new model represents a significant improvement over the older model, and further improvements may be feasible when historical climatic estimates and a higher resolution digital elevation model become available.

**Key words:** Site index, Mensuration, Tree physiology, Ecophysiology, Scots pine

**Word count:** 3467

**Introduction**

Foresters need predictions of site quality in order to plan silvicultural strategies, run growth and yield models in the absence of tree measurements, and value land. Most managers use site index, the mean top height of a stand at a standard, “index” age, as an estimate of quality. Mean top height is less dependent on stocking than measures of diameter (Assman 1970), although it has been found not entirely independent of stocking, particularly when selection ratios of alternative stockings are equivalent (Maclaren et al. 1995). Site index is usually species-specific, with different species often having different site indices on the same sites.

Site index can be predicted from mean top height at any given age if appropriate species- and region-specific difference models of mean top height growth and yield have been developed (Burkhart and Tennent 1977), and these often have standard errors as low as 0.5 m. However, in the absence of tree measurement, site variables can be used to predict site index (Hunter and Gibson 1984), albeit with larger standard errors (Ekö et al. 2008).

Estimations of site productivity have a major influence in forestry planning, stand optimization and regulatory control through legislation in Swedish forestry. In the regeneration phase, site properties and site index are used for selection of tree species and the need for various regeneration treatments. Also regeneration density in new stands is regulated based on site index. Thinning operations are, most commonly, enforced due to basal area growth in the stand, which in the planning phase is estimated based on site index. Stand rotation length is determined by several objectives, due to a forest owner´s preferences, but site index in combination with stand age is an important tool in the planning of rotation length. Furthermore, in the Swedish Forestry Act the lowest cutting age of a forest is dependent on site productivity, measured as site index. Constant evaluation and development of the measures and modelling of site productivity is therefore of utmost importance, considering the influence that site productivity estimations has on how individual forest stands are managed.

Swedish foresters have site index models for their principal species. The site index used in Swedish forest practice is species-dependent and estimates height of dominant trees at an age of 100 years (Elfving and Kiviste 1997). Site indices predicted from mean top height measurements are generally given the acronym SIh, while those predicted from site variables (Hägglund and Lundmark 1977) are labelled SIs. SIh is subject to tree measurement error, and in practice its precision probably improves as plot size increases and as proximity to index age diminishes. However, as SIs models are created using plot estimates of SIh, SIh should be more precise than SIs. In addition, ideally SIs should be an unbiased predictor of SIh despite its lower precision, but SIs estimated with a model developed during the 1970s (Hägglund and Lundmark 1977) is often below actual values since new, planted forests grow more rapidly than the old-growth that the SIs-functions were based on Nilsson et al. (2012). New forests are often planted rather than naturally regenerated and this may partially explain growth differences between old and new forests, but also fertilisation through deposition of nitrogen and longer growing seasons due to climate change may also cause newer forests to be more productive.

Recent progress with high-resolution geographical information system (GIS) and radiation-use efficiency modelling (Landsberg and Waring 1997) offers an opportunity to try to improve models of SIs. Mason et al. (2007a) have shown how simple radiation use efficiency modelling, with potential radiation use by trees (determined by environmental modifiers similar to those in the 3-PG model (Landsberg and Waring 1997)) summed across years, can be employed to predict tree development on alternative sites.

The objectives of the study described here were to:

1. Determine the precision and bias of estimating SIh from SIs for Scots pine (*Pinus sylvestris*) in Sweden; and
2. Conduct a pilot study to determine whether or not potentially useable radiation sums can improve predictions of SIh from site variables.

**Materials and Methods**

Data from 1985 plots from the Swedish National Forest Inventory were used during this study. These data came from recent, planted stands of Scots pine (*Pinus sylvestris* L.) in which pine formed more than 80% of the forest canopy. A summary of the data is shown in table 1 and the locations of plots are shown in Figure 1.

 [Table 1 about here]

 [Figure 1 about here]

Residuals were calculated as observed (SIh) minus predicted (SIs) values from Hägglund & Lundmark’s (1977) model. Hägglund & Lundmark’s model predicts SIs from latitude, altitude, distance from the coast, soil depth class, sub-surface water flow frequency, soil texture, and herbaceous vegetation type. Residuals were plotted against SIs values and the standard error of SIs prediction of SIh was calculated.

Monthly mean rainfall, mean daily maximum temperature and mean daily minimum temperature estimates averaged between 1960 and 1990 for Sweden were downloaded from the Bioclim website[[1]](#footnote-1). These were at 0.5’ resolution (Figures 2-4) and comprised monthly averages for one year, i.e.: 12 values for each variable at each grid point across Sweden.

[Figures 2-4 about here]

The Sveriges Meteorologiska och Hydrologiska Institute (SMHI) provided a series of monthly estimates of photosynthetically active radiation between 2007 and 2012 at a resolution of 16.9’ latitude and 8.05’ in longitude (Formetta et al. 2013). These were averaged to provide mean monthly estimates for a single year (Figure 5).

[Figure 5 about here]

Potential radiation use efficiency was simulated for each sample plot using object-oriented R coding developed by Euan Mason (Mason et al. 2011, Mason et al. 2007b). All calculations and analyses were coded using R statistical software (R Development Core Team 2004). An outline of the functions used is given below.

PAR radiation sums for each month were multiplied by a temperature modifier, and sums from January to December in an average year were created. The radiation sum was therefore:

 (1)

where in (1)

 RTot= modified PAR sum over one average year for the sample plot

 Rt = the total PAR radiation sum for month t

 ƒ = a soil water balance modifier from 0-1, reflecting restriction of water deficits on photosynthesis

 ƒD = a vapour pressure deficit (VPD) modifier (equation 2) with values between 0-1, reflecting the influence of VPD on stomatal conductance

 ƒT = a temperature modifier (equation 3) for month t, with values between 0 and 1, representing temperature limitations on photosynthesis.

A water balance model was created for each plot that was identical to that used in 3-PG (Landsberg and Waring 1997) and was run on a monthly time step. Leaf area index was assumed to be 4, because values above this tend to result in little change to the water balance model. Soil rooting depth was available from the inventory database as three categories, 100 mm, 450 mm and 800 mm. Soil type was classified in the inventory database as either coarse sand, sand, silt or clay. Where the soil was peat this was recorded in the dataset and the effect of peat was tested as a dummy variable during modelling. Maximum stomatal conductance was assumed to be 0.02 m/sec and maximum boundary layer conductance was assumed to be 0.2 m/sec. The water balance model was run on a monthly time step over a single year in each plot, using climate variables from the Bioclim GIS layers. The soil water modifier (see equation 1) used for PAR sums was also identical to that used in the 3-PG model.

Vapour pressure deficit (VPD) was estimated from Bioclim’s mean daily maximum and minimum temperatures by assuming that vapour pressure deficit was 0.5 times the saturated vapour pressure at the maximum temperature minus saturated vapour pressure at the minimum temperature. The average VPD should logically be related to the difference between minimum and maximum temperature in this way because water is lost from the air at minimum temperature as dew.

A VPD modifier on radiation use was used that is identical to that used in the current version of the 3-PG model. It was represented as:

  (2)

where in (2) VPD=vapour pressure deficit in kPa and e is the base of the natural logarithm. This modifier was also used to calculate stomatal conductance from maximum stomatal conductance in the water balance model assuming that stomatal conductance was directly related to fD as a fraction between 1 and 0.

The effect of temperature on radiation use was identical to that used in the 3-PG model was based on the minimum, optimum and maximum temperatures for photosynthesis as:

  (3)

where in (2) *fT* = 0 if *Ta* ≤ *Tmin* or *Tmax* ≤ *Ta*, and *Tmin*, *Topt* and *Tmax* were the minimum, optimum and maximum air temperatures for net photosynthetic production in degrees Celsius. *Ta*was the mean temperature for each month. The minimum, optimum, and maximum temperatures for photosynthesis were assumed to be -2, 15 and 25 degrees respectively (Kolari et al. 2007). Each plot had an elevation in plot records that was determined on-site and that was likely to be more accurate than an elevation estimate from the digital elevation model, and so a lapse rate was calculated to account for differences between these elevations and elevations of the GIS temperature raster points using Equation 4.

$Lapse=(E\_{R}-E\_{P})\*0.0098$ (4)

Where Lapse=adjustment to monthly temperature values, ER=elevation from temperature raster, and EP= elevation inplot records. This adjustment was necessary because the raster grid points were at different elevations from the plot points. Temperatures were adjusted by adding the total lapse rate calculated for any given point to the monthly temperature values from the Bioclim raster layer.

PAR sums were also accumulated for unmodified radiation, radiation only modified by temperature, and radiation modified by temperature and vapour pressure deficit but not by the soil water modifier. These different PAR sums were all tested as candidates for predictions of SIh.

Both “field” and “bottom” vegetation descriptions were available in each plot. These have been used as indicators of nutritional status when predicting SIs (Hägglund and Lundmark 1977) and are well known to foresters in Sweden. They comprise lists of species that are present on sites that indicate different levels of nutritional fertility. They were coded with numbers scaled to their levels of influence on productivity observed in Hägglund & Lundmark’s study. Phytometers have been used as indices of fertility in many other studies, for instance see the study by Ure (1950).

Multi-linear models often require transformations to stabilise variance and make relationships linear. In the study reported here scaled power transformations were used (Equation 5).







(5)

Where x is the variable being transformed, and  is a parameter that varies usually between -2 and 5, providing a range of transformed shapes.  values were chosen that made frequency distributions of variables as normal as possible.

The sample plot data were separated into a fitting (1361 plots) and validation (624 plots) datasets for the purposes of multi-linear modelling. A model was first constructed using the fitting dataset, then its predictions were tested using the validation dataset to ensure that bias and precision observed during fitting were the same as when the model was applied to the validation set.

Multicollinearity was evident between the modified radiation sum, elevation and distance to coast, and so the R procedure “anova” for comparing models was employed to determine whether or not these latter two variables added statistically significant extra information. This procedure applies the same rationale as the protocol outlined by Cook & Weisberg (1999) for dealing with multicollinearity.

**Results**

Residuals of an SIs prediction of SIh had a mean bias of 2.9 m with a standard error of 3.7 m (Figure 6). When the bias was removed by a linear model adjustment the standard error was 2.3 m.

[Figure 6 about here]

The best model of SIh using radiation-use efficiency had a standard error of 2.06 m and was unbiased across the range of predictions. Other terms in the model added a small amount compared to radiation-use efficiency, but they were nonetheless significant (Table 2 shows the final fit with all data). Both coast distance and elevation were transformed using equation 5, with  values of 0.57 and 0.61 respectively.

[Table 2 about here]

When this model was applied to the validation dataset the standard error was 2.09 (compared with 2.06 during fitting), and plots of residuals were very similar to those observed during fitting (Figure 7). A final model was then fitted using all data.

[Figure 7 about here]

R procedure “anova” comparing alternative models revealed that both elevation and distance to coast added small but statistically significant (P=7.327e-12) extra information, despite their correlations with modified radiation sum, the dominant independent variable. Note that this P value refers to the differences between models, and not individual lines in the table of coefficients. The influence of elevation, for instance was significant as an interactive effect and so the main effect was retained in order to satisfy the principle of marginality.

**Discussion**

Bias in estimation of SIh using SIs has very likely arisen due to both improved forest management practices and climate change since the SIs model was created in 1977. Large areas of Sweden are now planted with improved breeds and also with the benefit of site preparation. Temperatures in Sweden have been higher in recent epochs, and this may also be affecting tree productivity. It is therefore timely to begin creating a new estimation procedure, particularly one that can be used as the climate changes.

Biased and imprecise estimates of SIs produced by the old model would lead to incorrect estimates of future yields, applications of incorrect thinning schedules, and sub-optimal estimates of harvest age. The new, less biased and more precise model therefore will lead to improvements silvicultural and harvest planning.

An advantage of the SI estimation procedure demonstrated here is that the scale can be adjusted from old stand inventory polygons to a chosen pixel size or raster resolution. Increasing the resolution of SI estimates from mean values of entire stands of several hectares to pixels of 20 x 20 m (roughly the scale of plots used for calibration) will increase the precision of estimates of stand productivity. Increased resolution will also enable a more accurate choice of management methods. Small scale biotope identification might increase production and economy and allow targeting of nature conservation measures to areas where they are most efficient.

Radiation was best modified by temperature only. It is likely that with better localisation of inputs (Mason et al. 2017) soil water modifiers and vapour pressure deficit modifiers may become useful. Estimates of soil depth may also have to be improved, as these were given simply by categories in the database and the middle of each category was used in the water balance calculation. It is possible that better estimation of photosynthetic responses of Scots pine to temperature variation will eliminate elevation and distance to coast from the model.

The pilot study reported here is a very promising interim result. Modern GIS technology and hybrid physiological/mensurational modelling has allowed us to create a less biased and more precise estimation procedure, but it has some drawbacks that will be addressed in a more thorough study. These drawbacks and solutions to them are outlined below.

Firstly, we used long run averages of climatic variables to estimate SI, when estimates of weather during periods when trees were growing will very likely provide more precise models. Estimates as far back as 1960 on a 4 x 4 km grid system are being sourced. Localising the estimates (Mason, Salekin and Morgenroth 2017) will improve our models. Moreover, the improved models will be capable of projecting “what-if” projections of changes in productivity across Sweden if climate changes.

Secondly, the digital elevation model (DEM) we used had a 50 x 50 m resolution, and recent advances in DEM resolution in Sweden will allow us to test for effects of aspect and slope on productivity as detected in a similar analysis in New Zealand (Mason et al. in prep).

Thirdly, ground vegetation can be used to assess soil fertility so long as observers are skilled and it is the right time of year. However, vegetation has not been recorded everywhere, and it would be better if fertility could be assessed using data from GIS layers so that maps of site index could be constructed at high resolution for all of Sweden. Geological GIS layers may offer some opportunities to avoid the use of phytometers.

Fourthly, this pilot study was conducted for only one species, and it is anticipated that models will be constructed for several tree species commonly employed in Sweden.

Finally, estimation of site index and then running conventional growth and yield models driven by site index builds extra error sources into growth and yield estimation when compared to models that have physiology fully integrated into the growth and yield modelling framework (Mason, Methol and Cochrane 2011). We anticipate that fully coherent hybrid physiological/mensurational growth and yield models will be built for Swedish foresters in future.

The second stage of this study will address the shortcomings of our pilot study.

SIs was found to be a biased and imprecise estimator of modern SIh estimates of Scots pine in Sweden. Mean bias was 2.9 m and the standard error of estimation was 3.7 m.

Cumulative modified radiation sum was the most powerful variable in a model designed to estimate SIh from site variables. Other variables in the model included field vegetation, peat, elevation and coast distance. The standard error of the model was 2.07 m and it was relatively unbiased. This model represents a significant improvement when compared to the older site index estimation model.

**Acknowledgements**

The authors wish to thank the University of Canterbury for providing study leave for Professor Euan Mason, and the Southern Swedish Forest Research Centre for travel and accommodation funding. We also wish to acknowledge the Sveriges Meteorologiska och Hydrologiska Institute (SMHI) for access to monthly estimates of photosynthetically active radiation across Sweden.

**Disclosure**

The authors declare that they have no conflicts of interest.

**References**

Assman E. 1970. The principles of forest yield study. SH Gardiner, translator Oxford, England: Pargamon Press.

Burkhart HE, Tennent RB. 1977. Site index equations for radiata pine in New Zealand.

Cook RD, Weisberg S. 1999. Applied Regression Including Computing and Graphics. John Wiley and Sons New York.

Ekö PM, Johansson U, Petersson N, Bergqvist J, Elfving B, Frisk J. 2008. Current growth differences of Norway spruce (Picea abies), Scots pine (Pinus sylvestris) and birch (Betula pendula and Betula pubescens) in different regions in Sweden. Scand J Forest Res.23:307-318.

Elfving B, Kiviste A. 1997. Construction of site index equations for Pinus sylvestris L. using permanent plot data in Sweden. Forest Ecology and Management. Nov;98:125-134.

Formetta G, Rigon R, Chavez JL, David O. 2013. Modeling shortwave solar radiation using the JGrass-NewAge system. Geoscientific Model Development. 2013;6:915-928.

Hägglund B, Lundmark J. 1977. Site index estimation by means of site properties : Scots pine and Norway spruce in Sweden (Skattning av höjdboniteten med ståndortsfaktorer : tall och gran i Sverige). Studia Forestalia Suecica.138.

Hunter IR, Gibson AR. 1984. Predicting site index of radiata pine from environmental variables. New Zealand Journal of Forestry Science.14:53-64.

Kolari P, Lappalainen HK, Hanninen H, Hari P. 2007. Relationship between temperature and the seasonal course of photosynthesis in Scots pine at northern timberline and in southern boreal zone. Tellus.542-552.

Landsberg JJ, Waring RH. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecology and Management.95:209-228.

Maclaren JP, Grace JC, Kimberley MO, Knowles RL, West GG. 1995. Height growth of Pinus radiata as affected by stocking. New Zealand Journal of Forestry Science.25:73-90.

Mason EG, Methol R, Cochrane H. 2011. Hybrid mensurational and physiological modelling of growth and yield of Pinus radiata D.Don. using potentially useable radiation sums. Forestry. Apr;84:99-108.

Mason EG, Morgenroth JA, Bown HE. in prep. Precision forestry: What does hybrid physiological/mensurational modelling offer?

Mason EG, Rose RW, Rosner LS. 2007a. Time vs light: A potentially-useable-light-sum hybrid model to represent the juvenile growth of Douglas-fir subject to varying levels of competition. Canadian Journal of Forest Research.37:795-805.

Mason EG, Rose RW, Rosner LS. 2007b. Time vs. light: a potentially useable light sum hybrid model to represent the juvenile growth of Douglas-fir subject to varying levels of competition. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere. Apr;37:795-805.

Mason EG, Salekin S, Morgenroth JA. 2017. Comparison between meteorological data from the New Zealand National Institute of Water and Atmospheric Research (NIWA) and data from independent meteorological stations. New Zealand Journal of Forestry Science.47.

Nilsson U, Elfving B, Karlsson K. 2012. Productivity of Norway Spruce Compared to Scots Pine in the Interior of Northern Sweden. Silva Fennica. 2012;46:197-209.

R Development Core Team. 2004. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Ure J. 1950. The natural vegetation of the Kaingaroa Plains as an indicator of site quality for exotic conifers. New Zealand Journal of Forestry Science.6:112-123.

Table 1 – Summary of inventory dataset used for the study

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | SIh(m) | SIs(m) | Elevation (m) | Age(years) | Year measured |
| Minimum | 17 | 15 | 230 | 31 | 1983 |
| Mean | 22 | 19 | 371 | 46 | 1994 |
| Maximum | 27 | 22 | 500 | 58 | 2002 |

Table 2 – Model analysis of variance table and coefficients

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | df | SS | F value | P |
| T mod. rad. sum | 1 | 3722.5 | 868.7 | < 2.2e-16 |
| Field vegetation | 1 | 782.4 | 182.6 | < 2.2e-16 |
| Coast distance  | 1 | 539.6 | 125.9 | < 2.2e-16 |
| Peat | 1 | 847.4 | 197.8 | < 2.2e-16 |
| Bottom vegetation | 1 | 413.8 | 96.6 | < 2.2e-16 |
| Elevation  | 1 | 0.9 | 0.2 | 0.64 |
| Rad\*BottomVeg\*Elev | 1 | 221.7 | 51.7 | 8.98e-13 |
| Residuals | 1977 | 8471.8 |  |  |

 Adjusted r2 0.54 SEres 2.07

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source |  value | Coefficient | Standard error | t value |
| Intercept |  | 14.81 | 1.302 | 11.368 |
| T mod. rad. sum |  | 0.002612 | 0.0001512 | 17.279 |
| Field vegetation |  | -0.2196 | 0.01625 | -13.513 |
| Coast distance  | 0.57 | 0.0843 | 0.007512 | 11.221 |
| Peat |  | -4.535 | 0.3225 | -14.062 |
| Bottom vegetation |  | -0.2908 | 0.1041 | -2.794 |
| Elevation  | 0.61 | -0.09082 | 0.01374 | -6.609 |
| Rad\*BottomVeg\*Elev | 0.61 (Elev) | 0.000003593 | 4.995E-07 | 7.193 |

1. http://www.worldclim.org/bioclim [↑](#footnote-ref-1)