Site productivity indices for native forests in southeast Queensland

V A Jay 1 and M Neumann 2,3

1 Bluechip Forest Services, Dubbo NSW Australia.
2 University of Natural Resources and Life Sciences, Vienna, Institute of Silviculture, Peter-Jordan-Str. 82, 1190 Vienna, Austria
3 Department of Chemistry and Biotechnology, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Hawthorn, VIC 3122, Australia

E-mail: alex@bluechipforestry.com.au  https://orcid.org/0000-0003-2472-943X (Neumann)

Abstract. Site productivity, or site quality, describes the potential biomass growth and yield of vegetation at a given location. Land managers have devised indices for site productivity using attributes related to plant yields or growth rates, and these have great utility when available spatially in maps. The main factors determining site productivity include climate, soil and terrain characteristics. Here we analysed four productivity indices (two based on remote sensing only, two based on modelling and algorithms using spatial datasets). The tested indices were available over a 150,000 km² area of southeast Queensland Australia, a region dominated by Eucalyptus and Acacia species. We were interested in comparing the indices regarding underlying drivers, effects on vegetation types and the general distribution of site productivity across our study region. Our methods included histograms of spatial attribute intersection, and multivariate linear regression. Remote sensing has clear advantages in capturing current conditions, which potential productivity algorithms cannot depict. On the other hand, maps with productivity algorithms provide large-scale robust information on biomass growth/yield that is sensitive to the main drivers of plant growth (e.g. climate, soil).

1. Introduction
Site productivity (also called site quality) describes the biomass production potential of vegetation limited by local site conditions, including climate, water availability, solar radiation, physical soil characteristics and nutrients supply. Robust information on site productivity is critical for land management, for grazing, agricultural and forested land [1–3]. Historically, one of the first sources for site productivity information were forest yield tables that are still used worldwide [4–7]. Yield tables often provided information on site class or bonity, that is, the wood volume production per unit area per year for a particular timeframe (e.g. m³ ha⁻¹ year⁻¹ at age 100). Other sources use maximum stand basal area (m² ha⁻¹), as a more easily measured metric [8] compared to increment or production, that require repeated measurements, increment cores or growth models [9, 10]. Basal area can be quickly measured with basal area sweeps, so called Bitterlich samples [11], and used as an allometric estimator for above ground biomass [12], as an input for forest growth models [13] and as guidance for thinning or selective harvesting [14].

Other site productivity information can be calculated from field measurements [15, 16] and estimated from remote sensing [17, 18]. Models are now increasingly used to obtain wall-to-wall maps of...
vegetation information [19], supported by the advent of consistent remote sensing data with high spatio-temporal resolution, including NASA’s Landsat and Terra and Aqua satellites and Sentinel missions by ESA [19–21]. The benefits of spatial site productivity data are the efficiency of data collection and consistent data quality. For sparsely populated areas with low site productivity, combined models with remote sensing is often the only viable source for productivity information. The Australian forest inventory is underpinned by a productivity map, obtained from climate data, remotely sensed vegetation cover and a biophysical model [23, 24]. The primary metric of this Australian productivity map is net primary production (NPP gC m⁻² year⁻¹), similar to the current global productivity metric based on temporal MODIS satellite data [25, 26].

We used here available datasets for site productivity for a study region in southeastern Queensland in Australia, including the Australian net primary productivity map, NPP [24], remotely sensed Foliage projected cover, FPC (i.e., percentage of ground area occupied by the vertical projection of foliage) using multispectral data [27, 28], remotely sensed 95th percentile of canopy height using LiDAR data, HT95 [29] and a new site quality index SQ [30], which follows concepts similar to those formulated by [24, 31, 32]. Despite the differences in origin and conceptual approaches, all these datasets (NPP, FPC, HT95 and SQ) should describe site productivity. To date, no comparison of these datasets have been made and it is unclear which dataset is more suited to evaluate plant production. The objectives of this study are, (1) to show the distribution of productivity index values within native forest vegetation types, (2) to analyse relationships between site quality indices and (3) to explore drivers and variable importance of environmental variables on the tested site quality indices.

2. Study Area
The study area is a 15 million ha region of southeast Queensland, Australia, situated south and east of latitude/longitude (-24.5°, 149.4°). Figure 1 shows the study area and main vegetation types. Appendix A contains maps depicting the variables used in the study. The climate is subtropical with summer rainfall, highest near the coast. The terrain is mostly low elevation <300 m above sea level except for a central range of hills at 400-800 m running N-S, highest in the south at 800-1000 m.

Figure 1. Study area and main pre-clearing vegetation types (see tables 1 and 2 for source and full description).
3. Materials and Methods

All data was analysed as raster values. Forest vegetation types (VEG) were converted from the polygon to the raster format on the equivalent scale. The spatial resolution for all rasters was 3 arc seconds latitude and longitude, or about 85 m between pixel centres in the central part of the study area. Some component layers and the index NPP were originally created at lower spatial resolutions, e.g. rainfall at 30 arc seconds, NPP at 9 arc seconds, and these were converted as required to match the standard 3 arc seconds. The HT95 map layer was clipped to match FPC, so these two layers contain year 2014 existing vegetation only. All other data in the study area includes both cleared land and vegetated land. We summarized the used data in Table 1. All indices and variables were modelled in some way, e.g. spatial interpolation from samples, rather than obtained by direct measure. FPC and HT95 were obtained by remote sensing, and adjusted for historical FPC series and some topographic anomalies (cliffs, landform shading, etc.). Appendix C contains details of the remote sensing measures HT95 and FPC.

Table 1. The map layers and indices referred to in this study. We show vegetation types (VEG), followed by the four site productivity indices (HT95, FPC, NPP and SQ). FPC represents conditions in 2014, HT95 for 2009. The lower 5 rows, (greyed) are environmental variables which were used in constructing the SQ map layer.

| Source | Abbreviation | Name             | Unit       | Description                                                                 |
|--------|--------------|------------------|------------|----------------------------------------------------------------------------|
| 1      | VEG          | Forest Vegetation Type |           | pre-clearing vegetation type, predicted based on remnant types and land units |
| 2      | HT95         | Site Height      | m          | Height where 95% of the plant cover has been intercepted (remote sensing)    |
| 3      | FPC          | Foliage Projected Cover | %         | surface area of leaves projected vertically downward                        |
| 4      | NPP          | Net Primary Productivity | t ha⁻¹ year⁻¹ | climate, soil, radiation                                                    |
| 5      | SQ           | Site Quality     | m² ha⁻¹    | climate, soil, radiation, topography sum of monthly values for an index combining rainfall, temperature and slope-adjusted solar radiation. |
| 6      | Climate      | Climate          |           | sum of monthly values of the ratio rainfall :pan evaporation                |
|        | PSMI         | Prescotts Soil Moisture Index |           | combination score from 6 variables, eg pH, clay content etc, each at different depths |
|        | Sfert        | Soil fertility index |           | combined depth in metres, as an indicator of volume available for tree roots |
|        | DES          | Depth of Soil A&B horizons |           | a measure of run-on and infiltration; low on ridgetops, high in valleys     |
|        | TWI          | Topographic Wetness Index |           |                                                                           |

1 https://www.data.qld.gov.au/dataset/broad-vegetation-groups-pre-clearing-and-2019-remnant-queensland-series
2 http://data.auscover.org.au/xwiki/bin/view/Product+pages/ICESat+Vegetation+Structure
3 https://www.data.qld.gov.au/dataset/landsat-foliage-projective-cover-queensland-2014
4 Reference [24]
5 Reference [30]
6 see [30] for methods. Data for Climate was obtained from sources a, b, c all others from source a
   a. CSIRO Soil and Landscape Grid for Australia http://www.clw.csiro.au/aclep/soilandlandscapegrid/GetData-GIS.html
   b. BOM Australian Bureau of Meteorology http://www.bom.gov.au/jsp/ncc/climate_averages/solarexposure/index.jsp?period=jan#maps
   c. WCLIM Temperature and rainfall http://www.worldclim.org/current
3.1. Vegetation types
We used a GIS map layer of pre-clearing native forest vegetation types ‘VEG’ [33] to assess relationships between vegetation type and site productivity. We focus on the major forest types of Rainforest, Eucalypt and Acacia and use here eleven VEG type categories (table 2). Minor forest types in the study area, including Melaleuca spp. and Callitris spp. woodlands, heath, and native grasslands dominated by Astrebla spp. and other species (altogether 7.6% of the total land area) are not included in the analyses. Much of the forest has been altered or cleared over the last century or more. We have not filtered the VEG mapping to remove cleared land. The reader may compare VEG with FPC or HT95 maps to see extent and density of year 2014 forest vegetation (see appendix A).

We describe and compare the distribution of index and variable values for each of the 11 VEG types in table 2 and figure 1. From the 11 considered VEG types, types 2, 209, and 210 have the highest yields and timber productivity (table 2).

Table 2. Vegetation types (“VEG”) used in this study, their abbreviated names and full description.

| Map Code | Code | Abbrev | Full name |
|----------|------|--------|-----------|
| 1        | Rf_scrub     | Complex to simple, semi-deciduous mesophyll to notophyll vine forest, sometimes with Araucaria cunninghamii. |
| 2        | WetE     | Notophyll to microphyll vine forests, frequently with Araucaria spp. or Agathis spp. |
|          |          | Semi-evergreen to deciduous microphyll vine thicket. |
| 209      | Coast | Moist to dry eucalypt open-forests to woodlands usually on coastal lowlands and ranges. |
| 210      | Corymbia | Corymbia citriodora dominated open-forests to woodlands on undulating to hilly terrain. |
| 211      | Basalt | Moist to dry eucalypt open-forests to woodlands mainly on basalt areas |
| 212      | Dry Hills | Dry eucalypt woodlands to open-woodlands, mostly on shallow soils in hilly terrain (mainly on sandstone and weathered rocks) |
| 213      | Mixed | Dry to moist eucalypt woodlands and open forests, mainly on undulating to hilly terrain of mainly metamorphic and acid igneous rocks |
| 215      | Temprt | Temperate eucalypt woodlands. |
| 4        | Alluv | Eucalyptus spp. dominated open-forest and woodlands drainage lines and alluvial plains. |
| 5        | Plains | Eucalyptus populnea or E. melanophloia (or E. whitei) dry woodlands to open-woodlands on sandplains or depositional plains. |
|          |          | Dry eucalypt woodlands to open-woodlands primarily on sandplains or depositional plains. |
| 10       | Brigalow | Acacia harpophylla sometimes with Casuarina cristata open-forests to woodlands on heavy clay soils. |
3.2. Site productivity indices and environmental variables.
The GIS rasters encompassed a gross area of 20.9 million pixels. We converted all raster layers (apart from VEG) to 7 classes in order to facilitate comparisons and analyses. Classes are based on the percentiles of the range of pixel values found in the original layer, and are numbered 1-7 in presumed ascending order of their effect on productivity. Note the class size intervals are not equal. These classes of variable are defined in table 3.

Table 3. Summary of site productivity variables and environmental indices, aggregated into 7 distributional classes. We show percentiles which are approximately the frequency-weighted centre for each class. Units for DES and HT95 are metres, FPC is percentage land cover, SQ and Climate are a maximum potential basal area measure, m² ha⁻¹, all others are indices with no units. Colour codes correspond to map colours in appendix A.

| Name   | Class | FPC | NPP | SQ   | Climate | PSMI | SFert | DES | TWI | HT95 |
|--------|-------|-----|-----|------|---------|------|-------|-----|-----|------|
| vv. High | 7     | 99  | 80  | 17,2 | 53,3    | 58,7 | 1,5   | 4,8 | 1,5 | 14,3 | 53,7 |
| v. High | 6     | 95  | 66  | 12,8 | 46,9    | 54,7 | 1,2   | 4,3 | 1,4 | 12,7 | 45,9 |
| High    | 5     | 90  | 59  | 11,1 | 43,2    | 51,7 | 1     | 4   | 1,3 | 11,9 | 36,6 |
| HighMed | 4     | 70  | 46  | 8,1  | 36,6    | 46,8 | 0,7   | 3,5 | 1,2 | 10,2 | 23   |
| Med     | 3     | 50  | 39  | 7    | 33,2    | 41,7 | 0,6   | 3,2 | 1,1 | 9    | 18,9 |
| LowMed  | 2     | 30  | 31  | 5,4  | 31,6    | 40,3 | 0,6   | 2,7 | 1   | 8    | 15,3 |
| Low     | 1     | 10  | 19  | 4,3  | 29,5    | 37,8 | 0,5   | 2,2 | 0,9 | 6,7  | 11,5 |

The gross study area of 20.9 million pixels is defined by the SQ map (appendix A). Eleven mapped VEG types cover 19.4 million pixels, including both cleared and currently forested land. The land with recent vegetation measures (year 2014 FPC and 2009 HT95 layers intersection) covers 11.8 million pixels. The FPC layer does not include cropland and water bodies, and 99% of FPC pixels have greater than 9% cover. We removed pixels in the HT95 layer with canopy heights smaller than 5 m, and then filtered the result to further remove any remaining pixels with zero FPC. Some minor areas (about 0.2% of pixel count) of low height and low FPC, eg regrowth clumps, are still included in the results.

We then intersected the four site productivity index layers (SQ, NPP, FPC, HT95) with a CASE layer of combined environmental variables (Climate, PSMI, SFert, DES and TWI) using the nominal classes 1-7. “CASE” means a 5 digit code from 11111 to 77777, where each digit represents the nominal classes for the five environmental variables (Climate, PSMI, SFert, DES, TWI) respectively. This analysis was done for every pixel in the study area. Pixels with a zero or missing value for any variable or index were excluded.

We then counted the instances of each CASE, which intersected with a pixel in each site productivity index class 1-7. We did this for all land combined, and then also for each VEG type separately. The sum of pixel counts for each productivity index class 1-7 were then used as weights to account for the number of samples per CASE in the regression model.

3.3. Analysis of drivers using linear multivariate regression
NPP [24] and SQ (appendix B) were derived using environmental variables, while FPC and HT95 are based solely on remote sensing data. We were interested in understanding the drivers of the analysed site productivity indices and used multivariate regression models to test whether environmental variables have linear relationships with site productivity. To this end, we used two model types, one with untransformed covariates (equation 1) and one with log-transformed covariates (equation 2).

\[ Y = k + a \times \text{Climate} + b \times \text{PSMI} + c \times \text{SFert} + d \times \text{DES} + e \times \text{TWI} \]  

(1)
\[ Y = k + a \cdot \log(\text{Climate}) + b \cdot \log(\text{PSMI}) + c \cdot \log(\text{SFERT}) + d \cdot \log(\text{DES}) + e \cdot \log(\text{TWI}) \] (2)

where \( Y \) is a class value 1-7 for SQ, NPP HT95 or FPC, and \( k, a, b, c, d, \) and \( e \) are fitted coefficients using weighted regression and R statistical software. The weighting factor was the number of pixels for each CASE within a given productivity class. For instance, for CASE 11111 and SQ = 1, the number of pixels for all land is \( n=1703 \). Similarly counts within VEG type were used as weights when analysing the productivity index within VEG. For example, CASE 11111 and SQ = 1 occurred in five VEG types, 210, 212, 213, 5 and 10, with pixel counts of 19, 334, 2, 498 and 452 respectively (sum= 1305). The difference 1703 - 1305 is land which is not one of the 11 mapped VEG types. We included HT95 and FPC for regression analysis, even though these metrics are affected by past history not just environmental factors; for example, forest regrowth may have considerably lower height and vegetation cover than the potential state for the site (as captured by SQ and NPP).

We also fitted linear models (equation 1) by vegetation type. With this analysis we can determine which variables are most important determinants for the tested index, and whether there are discernible and significant differences in site productivity drivers for the various VEG types.

We investigated the potential for the use of principal components analysis (PCA) for illustrating relationships between CASES and productivity indices for the VEG types. However, the matrix sizes proved too large for calculation with 10-20 million samples (pixels) with \( \sim 10,600 \) non-zero CASES for each VEG type and index combination. The histogram layout in the VEG results follows a manually sorted order of VEG groups, L->R with descending proportion of high site quality SQ. Since Climate and PSMI are covariates of SQ, the VEG group display order also follows an approximate descending order of pixel counts for these variables.

4. Results and Discussion

4.1. Vegetation types and site productivity

First we illustrate the distribution of the four tested productivity indices for each vegetation type (figure 2). The VEG type with highest SQ is the (current or former) Wet Eucalypt forest (VEG type 2), with the median SQ class 6 (about 46.9 m²/ha). Somewhat surprisingly, much of the Rainforest type (VEG 1) falls into SQ class 4 (SQ about 36.6 m²/ha), which is similar to the drier Eucalypt types (209, 4, 211, 210 and 213). This probably reflects the inclusion of the wide-spread inland “dry scrubs” in the rainforest category. The two productivity indices SQ and NPP had a rather similar distribution within VEG types, in particular for the more productive vegetation types.

The two remote sensing measures, FPC and HT95, which consider only current vegetation, have a similar pattern left to right with decreasing SQ. High SQ and high NPP is in general associated with high FPC and high HT95. The distributions of FPC and HT95 tend to match within VEG types, except for Plains type 5, which has some higher cover and low heights.

We noted that these maps are created from different data sources. Most vegetation types are found with a large range of FPC and HT95. Only the less productive vegetation types (Mixed, Temperate, Plains and Brigalow) are limited to classes 1-5. The two vegetation types with highest average SQ distributions within VEG types, for example types 1 and 2, have both high FPC and HT95, and the low SQ type 10 has both low FPC and HT95. VEG type 1 (Rainforests and scrubs), exists with a broad range of FPC and HT95, due to the variable natural conditions in this group and the historic significant clearing for agriculture and current regrowth.
Figure 2. Site productivity indices versus vegetation types. We show histograms (horizontal bars) of two productivity indices (SQ, NPP) and two remote sensing measures of vegetation structure (FPC, HT95) for the eleven vegetation types. We use relative frequency (% pixel counts) within the vegetation types. The vegetation types are arranged from left to right with decreasing SQ. The greyed line is class 3, containing the median values (40-60%). Class 5 starts at the 90th percentile (for details see table 3).

4.2. Vegetation types and environmental variables

Figure 3 shows the distribution of environmental variables within the 11 VEG types. The dendrogram at the base of figure 3 shows groups of VEG types with the most similar distribution patterns of pixel counts. Similarity was determined using hierarchical classification analysis (HCA) of the data in figure 3, after transposing the rows and columns (11 rows x 35 columns).

Climate and PSMI follow the same pattern as the site productivity indices shown in figure 2. Soil depth (DES), terrain wetness index (TWI) and soil fertility index (Sfert) do not fit this mould. For instance, the high SQ, high productivity wet eucalypt forests (type 2) are predominantly found on sites with low soil depth (median or below), on drier parts of the topography (ridges and slopes), and on less...
than median soil fertility. This confirms the primacy of Climate rather than landscape factors for forest productivity.

**Figure 3.** Environmental variables versus vegetation types. We show histograms (horizontal bars) of the five environmental variables values (Climate, PSMI, DES TWI, Sfert). We use relative frequency (% pixel counts) within the vegetation types. The vegetation types are arranged from left to right with decreasing SQ. The greyed line is class 3, containing the median values (40-60%). Class 5 starts at the 90th percentile (for details see table 3). From top to bottom the variables are arranged in approximate descending order of importance as determinants of SQ, (to be shown later using regression coefficients).

Only low productivity Plains (type 5) and Brigalow (type 10) vegetation have considerable number of pixels with DES greater than 60th percentile. These western areas have low rainfall, and thus a low Climate class, which means low productivity despite having deeper soils. The soil depth in these cases is a result of sandy depositions (Plains) or in situ formation (Brigalow) on clay soils which swell or
shrink with the seasonal wet/dry cycles (Vertosols), and result in the incorporation of organic matter to significant depth in deep soil cracks.

A similar pattern can be observed for TWI. The more productive forest types in the near coastal regions occur on hilly land with slopes and ridges, whereas the drier inland and relatively flat areas receiving run-on, do not support productive forest because rainfall is low and evaporation is high.

The highest site fertility values are found for vegetation types on Basalt, since the Sfert index reflects the importance of P as a scarce nutrient in most Australian soils. However due to the location away from the higher coastal rainfall zones (figure 1, appendix A), the Basalt vegetation types (type 211) have only average SQ compared to other vegetation in the study region. Some Wet Eucalypt and Rainforest types also occur on basaltic parent material, and will have some of the highest SQ in the study area.

Next we checked using equation 1 whether environmental variables had different importance in explaining different SQ within each of the eleven vegetation types (table 4).

Table 4. Multiple linear regression coefficients and model statistics for environmental variables in eleven vegetation types (VEG). We show intercepts and coefficients for the five environmental variables separately for each VEG. At the bottom of the table we show model statistics, p value and coefficient of determination (R\textsuperscript{2}). The length of the horizontal bars highlight differences in the relative importance of the covariates. Green bars indicate positive correlations e.g. higher climate class, higher SQ class for that VEG. The table is arranged left to right as per figures 2 and 3.

We found there were no substantial differences in the relative importance of the variables by vegetation type (table 4). Only the Temperate vegetation type 215 deviated, with more important TWI, while for all other vegetation Climate was most important. Figure 3 suggest that TWI has a wider variance and range than the other variables within Type 215, and we note above that Type 215 also has the lowest R\textsuperscript{2} and a large intercept value. This suggests that a linear model is not a good representation of the SQ algorithm for the temperate vegetation type, and/or that the vegetation type itself has high internal variability, which is not well described by the mapping.

Figure 2 showed that wet eucalypt and coastal eucalypt vegetation types (2 and 209 respectively) have a high proportion of high SQ and NPP classes. These vegetation types are also most strongly associated with high climate and high PSMI classes (figure 3). Table 4 however shows lower coefficient values for climate as a driver of SQ within vegetation types 2 and 209, and this arises because there is a relatively narrow range of climate and SQ in those types. Conversely, where climate variation is high within the vegetation type, for example for rainforest and Corymbia (types 1 and 210; see figure 3), then climate has the highest regression coefficient (table 4).
4.3. Site productivity and environmental variables

Now we illustrate the relationships between the indices and variables, for all land in the study area. Figure 4a shows that NPP and SQ are positively related, but not strongly. SQ and NPP both have weak positive relationship with HT95 and FPC. The strongest relationship in the matrix appears to be that sites with high cover tend to have tall trees, and vice versa.

Figure 4b shows the strongly linear association between climate and SQ, and to a lesser extent between SQ and PSMI. The associations with NPP are also linear, but not as strong as for SQ. Distribution of height and vegetation cover are also associated with Climate and PSMI but more weakly because of the inclusion of non-mature vegetation. The three landscape variables (DES, TWI and Sfert) do not appear to be strongly positively associated with any of the productivity indices.

4.4. Drivers for site productivity

Figures 2 and 3 indicated pronounced differences between the vegetation types in the study region regarding site productivity and the environmental variables. We were next interested in the underlying drivers of varying site productivity. We use equations 1 and 2 to compare the relative importance of environmental variables in the four tested site productivity indices, assuming linear (equation 1) and non-linear relationships (log-transformed, equation 2).

We have not shown standard error of the residuals (Residuals SE), because the software output using weighted samples does not produce meaningful results in this case. Instead we checked the residuals by applying one of SQ equations (linear model from table 5, equation 3) and found that 95% of residuals for SQ class were between -1 and +1, meaning that SE is <<1.

$$SQ = -2.046 + 0.7164 \times \text{Climate} + 0.3053 \times \text{PSMI} + 0.2815 \times \text{DES} + 0.2276 \times \text{TWI} + 0.1426 \times \text{Sfert} (3)$$

All models had low p-values (table 5). Our results further indicated, that log-transformed models (equation 2) were not more correlated with site productivity than the linear models (equation 1), as evident from general higher coefficients of determination (R²) for linear models. Highest R² was observed for SQ and NPP models. NPP and SQ are both expressed as single numerical index, and are the product of multiplicative algorithms and variables which follow that of [28] in calculating Net Photosynthetic Index, viz. NPI = FPC * Moisture * Fertility * Light * Temperature. It is then somewhat surprising, that equation 2 did not provide a better fit than linear formulation (equation 1), since the multiplicative (log) form more directly reflects the source algorithms (appendix B) [24]). We noted however in table 8, that the differences in R² for SQ using equation 1 and equation 2 are relatively minor, and that the order of importance of variables is largely unchanged.

Remotely sensed canopy height (HT95) was more correlated with the used environmental variables than was vegetation cover (FPC), although R² for both are quite low, (R²<0.33 for HT95 and <0.15 for FPC), because height and cover are an outcome of history not just environment.

Climate was by far the most important environmental variable for all of the used site productivity indices. NPP is less strongly related to Climate than is SQ (figure 4b, table 5). This was expected since Climate per se was a variable used directly in the SQ algorithm (see appendix B). NPP uses similar components to those in Climate, but with a different algorithm. Second most important driver of SQ and NPP was PSMI, the ratio of evaporation and precipitation. We expected that Climate and PSMI have some quite high autocorrelation since both include rainfall, evaporation, solar radiation and temperature as input variables, these being calculated monthly and summed to an annual value. When separated to 11 VEG types (table 4), the relative importance of these variables for productivity class remains similar to that for the study area taken as a whole (table 5).

TWI has almost no relationship with NPP (table 5), probably because the nearest analogous variable in NPP was soil water holding capacity value based on modelled soil depth and estimated plant water availability on a 250 m grid scale. At this scale and using very broad scale soil type mapping, the topographic influence on forest productivity would not be visible.
Figure 4. a) Correspondence matrix for the indices. We show histograms of relative Class frequency (% pixel counts as horizontal bars) for the intersecting layers in the study area; b) Environmental variables and productivity indices. We show histograms of relative Class frequency (% pixel counts as horizontal bars) for the intersecting layers in the study area. The variables are arranged top to bottom in descending order of variables importance as found in table 4.
Table 5. Multiple linear regression results for four site productivity indices and using two model types
linear models (left columns) and log-transformed models (right columns). We show intercepts and coefficients for the five environmental variables, separately for each of the four tested indices. At the bottom of the table we show model statistics, p value and coefficient of determination ($R^2$). The length of the horizontal bars highlight differences in the relative importance of the covariates. Green bars indicate positive correlations (e.g. higher climate, higher SQ) and red bars indicate negative correlations (e.g. lower site fertility, higher HT95).

|          | SQ    | NPP   | HT95  | FPC   |          | SQ    | NPP   | HT95  | FPC   |
|----------|-------|-------|-------|-------|----------|-------|-------|-------|-------|
| Intercept| -2.0460 | -0.2537 | 1.8884 | 2.1316 | (Intercept) | -1.3979 | 0.4885 | 2.0298 | 2.3494 |
| Climate  | 0.7164 | 0.4704 | 0.4470 | 0.3429 | log(Climate) | 1.8658 | 1.1415 | 1.1652 | 0.9218 |
| PSMI     | 0.3053 | 0.2858 | 0.0960 | 0.0517 | log(PSMI) | 0.8329 | 0.6978 | 0.2083 | 0.0321 |
| DES      | 0.2815 | 0.1086 | -0.0432 | 0.0791 | log(DES) | 1.0685 | 0.5001 | 0.1078 | -0.0205 |
| TWI      | 0.2276 | 0.0166 | 0.0751 | 0.0862 | log(TWI) | 0.3630 | 0.2152 | 0.3846 | 0.0047 |
| Sfert    | 0.1426 | 0.0598 | 0.1526 | 0.1646 | log(Sfert) | 0.3730 | 0.4428 | 0.3829 | 0.4233 |
| $p$ value| < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| $R^2$    | 0.844  | 0.519  | 0.334  | 0.155  | 0.800  | 0.427  | 0.2858 | 0.1267 |

Sfert has lowest coefficient values for SQ across the board, meaning that variation in Sfert is not strongly correlated with SQ at the landscape scale for any of the VEG types. Nevertheless, differences in soil fertility index will be associated with SQ changes at the localised or stand scale. The secondary variables are also important when Climate is not highly favourable; for example, it can be seen in Figure 4a that the medium SQ class 3, with mostly low Climate and PSMI classes, has a higher than average proportion of sites with class 4 and over for DES, TWI and Sfert.

While soil fertility index Sfert was weakly but positively correlated with SQ and NPP, we observed a negative correlation between Sfert and both HT95 and FPC. This suggests that fertile pixels have a higher SQ and NPP, but lower canopy height and vegetation cover (keeping all other variables constant). Negative or neutral coefficients also occurred for DES and TWI with height and cover. It appears sites with high soil fertility index, deep soils, and to a lesser extent the lower (wetter) points in local topography, are now associated with low height and cover, because they were preferentially cleared of mature native vegetation for agricultural use.

5. Conclusions
This study demonstrated the relative importance of environmental variables in site productivity indices, over a large region (150,000 km$^2$) in southeast Queensland containing diverse subtropical climates and vegetation types. We found that some forest vegetation types of the study area occur in distinct parts of the site productivity range. For forests, robust information on potential basal area as m$^2$ ha$^{-1}$, or carbon accumulation rate (growth) as tC ha$^{-1}$ year$^{-1}$ is especially important for assessing many of their diverse functions. Remote sensing can be a powerful tool to complement site productivity information; for example, we found that lower heights and lower vegetation cover based on remote sensing data were associated with deeper fertile soils, indicating that these land types have been relatively more cleared.

There are now several site productivity indices available, each with specific advantages. We found that the environmental variables which most strongly determined the productivity index values across all land, are ranked as follows: Climate > PSMI > DES > TWI > Sfert. Our observations illustrate that Climate and soil moisture are the predominant associates of forest productivity indices at a regional scale. Soil fertility and landscape factors do not have much effect where climate and water availability are favourable, but will nevertheless be important determinants of productivity at the localised scale of one hectare of less.
Appendix A. Additional maps. In SQ black lines are rainfall isohyet (west to east: 600, 800, 1000, 1500 mm year\(^{-1}\)). For data sources see table 1, table 3. For scale see figure 1.
Appendix B. Site quality algorithm.

\[
SQ = \left[ \sum_{month=1}^{12} \left( \frac{RainSol_{month}}{12} \times \sqrt{DayNightTemp_{month}} \right) \right] \\
\times \sqrt{Moisture \times Rooting \times Fertility \times Position}
\]
Appendix C. Details of remote sensing products and index sources.

HT 95 https://portal.tern.org.au/vegetation-height-structure-australia-coverage/21777

Peter Scarth 2013 Vegetation height and structure - derived from ALOS-1 PALSAR, Landsat and ICESat/GLAS, Australia coverage. Version 1. Terrestrial Ecosystem Research Network (TERN). Dataset. https://portal.tern.org.au/vegetation-height-structure-australia-coverage/21777

Temporal coverage from 2009-01-01 To 2009-12-31
Vertical plant profiles for the Australian continent were derived through integration of ICESat GLAS waveforms with ALOS PALSAR and Landsat data products. Co-registered Landsat Foliage Projected Cover (FPC) and ALOS PALSAR L-band HH and HV mosaics were segmented to generate objects with similar radar backscatter and cover characteristics. Within these, height, cover, age class and L-band backscatter characteristics were summarised based on the ICESat and Landsat time-series and ALOS PALSAR datasets.

ALOS PALSAR L-band Fine Beam Dual (HH and HV) polarisation data mosaics at 25 m spatial resolution, provided through the JAXA K&C Initiative, were combined with an Australian mosaic of persistent green vegetation cover derived from Landsat sensor time-series data [1]. These data were segmented to map clumps of contiguous vegetation structure and reduced to a maximum of 1000 unique classes using k-means based cluster analysis. These classes were then intersected with ICESat/GLAS L2 (Release 33) Global Land Surface Altimetry Data (GLA14) data acquired between 2003 and 2009 and obtained from the National Snow and Ice Data Centre (NSIDC). All GLAS laser shots were corrected for differences in footprint size, laser output power, vegetation/ground reflectance, and terrain slope. These laser shots were aggregated for each cluster and processed to mean vertical profiles of \( P_{\text{gap}}(\theta, z) \). Vertical plant profiles were then derived and height of peak foliage density (\( H_{\text{peak}} \)) and height percentiles (95th, 75th, 50th, and 25th) extracted. To produce a national map, each clump in the original segmentation was then assigned the mean vertical profile corresponding to the nearest cluster centroid.

FPC https://www.data.qld.gov.au/dataset/landsat-foliage-projective-cover-queensland-2014

Landsat Foliage Projective Cover - Queensland 2014

Foliage Projective Cover (FPC) is the percentage of ground area occupied by the vertical projection of foliage. The Remote Sensing Centre FPC mapping is based on regression models applied to dry season (May to October) Landsat-5 TM, Landsat-7 ETM and Landsat-8 OLI imagery for the period 1988-2014. An annual woody spectral index image is created for each year using a multiple regression model trained from field data collected mostly over the period 1996-1999. A robust regression of the time series of the annual woody spectral index is then performed. The estimated foliage projective cover is the prediction at the date of the selected dry season image for 2014. Where this deviates significantly from the woody spectral index for that date, further tests are undertaken before this estimate is accepted. In some cases, the final estimate is the woody spectral index value rather than the robust regression prediction. The product is further masked to remove areas classified as non-woody (see Landsat Woody Vegetation Extent - Queensland 2014). Corrections have been applied to remove errors due to topographic effects, cloud, cloud shadow, water, cropping, and regrowth following clearing. Some errors may remain. The product was generated from WRS-2 path/row scenes obtained from the United States Geological Survey (USGS).
SQ algorithm data sources
see Jay 2018 for methods. Data for Climate was obtained from sources a, b, c, all others from source a.

Jay 2018 methods; also Appendix B
https://www.researchgate.net/publication/344218529_Site_Quality_mapping_for_Private_Native_Forestation_in_southeast_Queensland

a. CSIRO Soil and Landscape Grid for Australia
https://www.clw.csiro.au/aclep/soilandlandscapegrid/ProductDetails-SoilAttributes.html

Soil Attribute Product Types
(i) National Soil Attribute Maps
The Soil and Landscape Grid of Australia's National Soil Attribute Maps were generated by combining Australia-wide digital soil attribute maps derived using consistent data mining-kriging models and regional maps for parts of Australia, derived using disaggregation and regression modeling. The combination of these maps in areas where the regional maps were available, should increase the evidence base of the National Soil Attribute Maps. Where the regional data was available, the combination was performed using a variance weighted mean at each pixel.

(ii) Australia-wide three dimensional digital soil attribute maps
The Soil and Landscape Grid of Australia's, Australia-wide Soil Attribute Maps were generated using measured soil attribute data from existing databases in the national soil site data collation and spectroscopic estimates made with the CSIRO's National spectroscopic database (Viscarra Rossel & Webster, 2012). The spatial modeling was performed using decision trees with piecewise linear models and kriging of residuals. Fifty environmental covariates that represent climate, biota, terrain, and soil and parent material were used in the modeling. Uncertainty was derived using a bootstrap (Monte Carlo-type) approach to derive for each pixel a probability density function (pdf), from which we derived 90% confidence limits. The approach is described in Viscarra Rossel et al. (2015a).

b. BOM Australian Bureau of Meteorology http://www.bom.gov.au/jsp/ncc/climate_averages/solar-exposure/index.jsp?period=jan#maps
Further details..
https://data.csiro.au/collections/collection/CI18852v002
Mean monthly total shortwave radiation on a sloping surface modelled using the 1" DEM-S - 3" mosaic
Start date 2000-02-10 End date 2000-02-21
Collection description.
Mean monthly solar radiation was modelled across Australia using topography from the 1 arcsecond resolution SRTM-derived DEM-S and climatic and land surface data. The SRAD model (Wilson and Gallant, 2000) was used to derive: • Incoming short-wave radiation on a sloping surface • Short-wave radiation ratio (shortwave on sloping surface / shortwave on horizontal surface) • Incoming long-wave radiation • Outgoing long-wave radiation • Net long-wave radiation • Net radiation • Sky view factor
All radiation values are in MJ/m2/day except for short-wave radiation ratio which has no units. The sky view factor is the fraction of the sky visible from a grid cell relative to a horizontal plane. The radiation values are determined for the middle day of each month (14th or 15th) using long-term average atmospheric conditions (such as cloudiness and atmospheric transmittance) and surface conditions (albedo and vegetation cover). They include the effect of terrain slope, aspect and shadowing (for sun positions at 5 minute intervals from sunrise to sunset), direct and diffuse radiation and sky view. The
monthly data in this collection are available at 3 arcsecond resolution as single (mosaicked) grids for Australia in TIFF format. The 3 arcsecond resolution versions of these radiation surfaces have been produced from the 1 arcsecond resolution surfaces, by aggregating the cells in a 3x3 window and taking the mean value. The 1 arcsecond tiled data can be found here: https://data.csiro.au/dap/landingpage?pid=csiro:9530. The 1 arcsecond mosaic data can be found here: https://data.csiro.au/dap/landingpage?pid=csiro:18851

c. WCLIM Temperature and rainfall http://www.worldclim.org/current
SQ data input included mean monthly precipitation evaporation, mean monthly temperature minima and maxima all at 30 arc seconds (30") gridded resolution (approx. 1 km grid). The GIS grids contained data with means 1950-2000, which comprised the antecedent version of this.

Fick S E and R J Hijmans 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37 (12): 4302-4315. https://doi.org/10.1002/joc.5086

We created a new global dataset of spatially interpolated monthly climate data at a 1 km² resolution, including monthly temperature, precipitation, solar radiation, vapor pressure and wind speed. Interpolations utilized MODIS satellite imagery, and instead of using a single model formulation for the entire world, we constructed the final product by selecting the best performing model for each region and variable. Satellite data improved prediction accuracy for temperature variables, particularly in areas with a low station density, but improvements were marginal for other variables, highlighting the importance of dense, high-quality climate station data.

Source data is available here https://worldclim.org/data/worldclim21.html
Worldclim.org is a site which compiles and distributes data from various sources.

Historical climate data.
This is WorldClim version 2.1 climate data for 1970-2000. This version was released in January 2020. There are monthly climate data for minimum, mean, and maximum temperature, precipitation, solar radiation, wind speed, water vapor pressure, and for total precipitation. There are also 19 “bioclimatic” variables. The data is available at the four spatial resolutions, between 30 seconds (~1 km²) to 10 minutes (~340 km²). Each download is a “zip” file containing 12 GeoTiff (.tif) files, one for each month of the year (January is 1; December is 12).
Variables: minimum temperature (°C), maximum temperature (°C), average temperature (°C), precipitation (mm), solar radiation (kJ m⁻² day⁻¹), wind speed (m s⁻¹), wind 10m, water vapor pressure (kPa)

Acknowledgements
We appreciate the spirit of sharing data that made this study possible. Specifically, we acknowledge Tom Lewis (Department of Agriculture and Fisheries) and Mark Stevenson (Department of Agriculture and Fisheries) who provided impetus for documenting the SQ algorithm and applying it to Queensland forests.
The GIS and Remote Sensing for Sustainable Forestry and Ecology (SUFOGIS) project has been funded with support from the EU ERASMUS+ program. The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.
References

[1] Neumann M and Smith P 2018 Sci. Total Environ. 643 902–11
[2] Daryanto S, Eldridge D J and Throop H L 2013 Agric. Ecosyst. Environ. 169 1–11
[3] He M, Kimball J, Maneta M, Maxwell B, Moreno A, Beguería S and Wu X 2018 Remote Sens. 10 372
[4] Curtin R A 1970 Aust. For. 34 97–106
[5] Egusa T, Kumagai T and Shiraishi N 2020 Sci. Rep. 10 1–9
[6] Ilvesalo Y 1927 Methods for preparing yield tables Silva Fenn. 5 1-23
[7] Vanclay J K 1994 Modelling forest growth and yield: applications to mixed tropical forests (CAB International, Wallingford, U)
[8] Jay V A and Dillon A 2016 Aust. For. 79 85–95
[9] Eastaugh C S and Hasenauer H 2013 For. Sci. 59 1–14
[10] Schmitt A, Trouvé R, Seynave I and Lebourgeois F 2020 Ann. For. Sci. 77 52
[11] Hasenauer H and Eastaugh C S 2012 Int. J. For. Res. 2012 961576
[12] Burrows W H, Hoffmann M B, Compton J F, Back P V and Tait L J 2000 Aust. J. Bot. 48 707–14
[13] Weiskittel A R, Hann D W, Kershaw J A and Vanclay J K 2011 Forest Growth and Yield Modeling (Chichester, UK: John Wiley & Sons, Ltd)
[14] Code of practice for native forest timber production on Queensland’s State forest estate 2020 (Queensland Parks and Wildlife Service) p 110
[15] Neumann M et al. 2016 Remote Sens. 8 1–18
[16] Antón-Fernández C, Mola-Yudego B, Dalsgaard L and Astrup R 2016 Can. J. For. Res. 46 794–803
[17] Li W, Niu Z, Shang R, Qin Y, Wang L and Chen H 2020 Int. J. Appl. Earth Obs. Geoinf. 92 102163
[18] Chen Y and Zhu X 2012 South. For. 74 217–27
[19] Moreno A, Neumann M and Hasenauer H 2017 Geosci. Data J. 4 17–28
[20] Immitzer M, Vuolo F and Atzberger C 2016 Remote Sens. 8 166
[21] Liu Z, Hu M, Hu Y and Wang G 2018 Int. J. Remote Sens. 39 1092–116
[22] Zhao M, Heinsch F A, Nemani R R and Running S W 2005 Remote Sens. Environ. 95 164–76
[23] Haverd V et al. 2013 The Australian terrestrial carbon budget Biogeosciences 10 851–69
[24] Kesteven J and Landsberg J 2004 Developing a national forest productivity model (National Carbon Accounting System Tech. Rep. 23) p 104
[25] Zhao M and Running S W 2010 Science 329 940–3
[26] Running S W and Zhao M 2015 User’s Guide Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Obser. System MODIS Land Algorithm (Version 3.0, 10/7/2015 28)
[27] Fisher A, Day M, Gill T, Roff A, Danaher T and Flood N 2016 Remote Sens. 8 515
[28] Read S and Howell C 2019 Aust. For. 82 53–5
[29] Scarth P, Armstrong J, Lucas R, Bunting P, Search P, Armstrong J, Lucas R and Bunting P 2019 Remote Sens. 11 147
[30] Jay A 2018 Site Quality mapping for Private Native Forestry in southeast Queensland. doi:10.13140/RG.2.2.34778.77760
[31] Specht A and Specht R L 1993 Biodivers. Conserv. 2 152–67
[32] Specht R L 1981 Oecologia 50 347–56
[33] Neldner V J, Niehus R E, Wilson B A, MacDonald W J F, Ford A J and Accad A 2017 Vegetation of Queensland. Descriptions of Broad Vegetation Groups. Version 3 (Queensland Herbarium, Department of Science, Information Technology and Innovation, Brisbane)