Resilient Forests Programme

SCION
FORESTS ▪ PRODUCTS ▪ INNOVATION

Date: 20 March 2020
Reference: RFP-TN001

TECHNICAL NOTE

A data catalogue for developing prototype individual-tree growth simulator and specifications for future data collection

Summary:

A concise data catalogue prepared for developing a prototype of the new tree growth simulator has been described. This data catalogue created an initial inventory of data assets through the discovery, description and organization of distributed datasets, which provided a context to enable us to find and understand relevant datasets needed for developing this prototype simulator.

Author/s: Yue Lin, Alan Tan

Introduction

The objective of the sub-RA 2.2 is to understand the impact of inter-tree competition on tree growth and stand uniformity. A new individual-tree-based model will be developed to simulate the growth and competitive interactions among trees when there are different levels of stem and crown size inequality. The development of this model will based on fundamental ecological principles and informed by stem and crown information captured through Light Detection and Ranging (LiDAR). To develop a prototype of this simulator, we reviewed the state of the art in tree growth simulators and assembled available datasets needed to develop a prototype individual-tree growth simulator.

This technical note describes a concise data catalogue prepared for developing the simulator. This data catalogue creates and maintains the initial inventory of data assets through the discovery, description and organization of distributed datasets. It provides a context for finding and understanding relevant datasets needed for developing this prototype individual-tree growth simulator.

We focus on datasets that can potentially be used for modelling tree growth and tree-tree competition. Datasets which can be used for developing further functions or sub-models (e.g., wood quality and disease) are out of scope for this technical note, but were briefly reviewed, described and discussed with specifications for future data collection.

Method

Collecting and understanding datasets

The datasets discovered and assessed in our data catalogue can be categorized into seven groups (Table 1). All those datasets have different designs and structures suited to specific uses. There will be interpretations required to combine relevant and compatible fields in a manner suitable for constructing an overall summary of the data and it's suitability for modelling. As we found before (Lin et al., 2018), to combine those datasets and use for specific modelling requires a proper work of interpretation, so as to identify suitable linking fields and bring the data from each source to a compatible format for specific models.

There are several fundamental variables required for developing the prototype simulator: tree genetics/species, location of trees, age, diameter and height, with also a strong preference for plot spatial location with geographic coordinates and silvicultural and management information. In addition, there were a range of other variables with potential to support the growth modelling effort. Those key information and data can be extracted from forest sample plots datasets, e.g. from Scion permanent sample plots database (PSPs).

FOREST GROWERS RESEARCH

supported by forestgrowers commodity levy

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

Scion permanent sample plots database and other forest sample plots datasets

The PSPs database is managed by Scion for growth monitoring in commercial forests and prediction of forest growth by researchers (Hayes & Andersen, 2007). The Scion PSPs database captures data from plantation forest growth measurements. These growth measurements are used for growth monitoring by private timber management companies and for research into forest management. There are a series of trials and plots containing individual tree level information and location (e.g., FR308 trial) which can be used for developing the prototype individual-tree growth simulator. A joint of PSPs plot data with individual tree stem and crown information captured through LiDAR data of the available plot will be explored and used for modelling in next steps.

The Land Use Carbon Accounting System (LUCAS) database managed by Ministry for the Environment (MfE) (Ministry for the Environment, 2010) and the Field Measurement Approach (FMA) database managed by Ministry of Primary Industries (MPI) (Ministry for Primary Industries, 2017) are two other datasets can be useful for developing the simulator.

The LUCAS database is designed to store data measured specifically for national carbon reporting and analysis. The data is separated into pre-1990, post 1989 planted and post 1989 natural forests to align with international carbon reporting requirements. The planted forest data also comprise ground-based plot measurements and LiDAR data.

The FMA database is designed to store measurement following the FMA standard for property specific carbon modelling. Because of the property specific nature of the data, there is less focus on uniformity and compatibility of forest type and forest management descriptors used, particularly in character fields.

Although LUCAS and FMA database doesn't record information of individual tree location in each plot, they provide independent data for developing and validating tree growth models across New Zealand.

In addition, the National Vegetation Survey Databank (NVS) contains general vegetation survey data, from plots that are usually not permanently marked. These data include reconnaissance descriptions and Protected Natural Areas data. The NVS is suitable for vegetation description, studies of species distributions, and studies needing only coarse measurement of changes in vegetation. It can be used for modelling indigenous tree species growth and forest vegetation. Kunstler et al. (2013) used this dataset to calibrate and parameterize SORTIE, an individual tree-based forest dynamics model, for simulating how harvesting and earthquake disturbance affect the dynamics of a New Zealand podocarp–angiosperm forest.

NIWA virtual climate station network (VCSN) dataset

Climate (meteorological) datasets are essential to individual tree-based models, process-based models and non-parametric machine learning models. The NIWA's VCSN data are estimates of a set of fundamental climate variables interpolated on a regular 5 km grid covering the whole of New Zealand from 1972 to present. The original data was recorded in a daily form. This dataset can be restructured into monthly, seasonal and annual forms to meet different modelling requirements. It is suggested to use a seasonal/annual climate form which will be suitable for most modelling requirements.

There are other free climate datasets can be used for our modelling work. This data catalogue includes two datasets which are widely used and can be imported and analysed directly in R (Table 1).

LiDAR datasets

LiDAR is an airborne remote sensing technology that measures distance and surface properties of targets by illuminating them with a laser and recording the reflections. It has been shown to provide accurate information on individual-tree and stand-level forest structure, as well as on wood quality (Watt et al., 2013). Datasets comprise ground-based measurements of individual trees with stem and crown information captured through LiDAR will be prepared and used for the individual-tree level setup and initialization in the prototype simulator.

Scion's Forest Informatics team maintains a series of high quality LiDAR datasets. Some of these datasets were acquired by demand and were from forest trails recorded in PSPs database. It is important and useful to link the LiDAR data with PSPs database in the future, so the PSPs database can comprise ground-based plot measurements with LiDAR information. This can be achieved if location of individual trees and LiDAR data georeferenced into the New Zealand Transverse Mercator (NZTM) coordinate system.

In the LUCAS database, some planted forest data are complemented with LiDAR data captures. The LiDAR data, captured between 2006 and 2015, comprise swaths of approximately 70m-wide over various pre-1990 and post-1989 forest plantations. Metrics for multiple surfaces (tree-top to ground) can be extracted from this LiDAR data to characterise forest structure and ground elevation.

Future data collection and preferred datasets

There are other categories of datasets that can be used for improving sub-models and for developing additional functions of the new individual-tree growth simulator. For instance, a forest simulator with individual trees' wood quality features is appealing for industry stakeholders. A database comprises ground-based plot measurements, LiDAR information and wood quality records will be much preferred.

Soil information such as geochemical features, hydrological condition and fertility are essential component in mechanistic forest models and simulators. A collection of such datasets is necessary for developing and modelling underground processes.

Eco-physiological traits (e.g., photosynthetic capacity, specific leaf area, and leaf area index) of tree are important variables in process-based models and individual-tree based simulators. An important use of plant traits database is growth model calibration and parameterization (of different tree species), which is the core sub-model in process-based models and individual-tree growth simulators.

Datasets of indices and predictions of existing models, such as site index, 300-Index, 400-Index, 500-Index and 3-PG, are also important information for validating the new tree growth simulator.

Table 1. Data catalogue for developing a prototype individual-tree growth simulator

Category	Dataset	Accessibility	Description
Forest sample plots	PSPs	Scion has access to this dataset	Data from plantation forest growth measurements. The main uses of the data are for growth monitoring by private timber management companies and for research into forest management. Some plots data contains individual trees' location, DBH, height and breed/genotype information.
	LUCAS	Scion has access to this dataset	Designed to store data measured specifically for national carbon reporting and analysis. The data is separated into pre-1990, post 1989 planted and post 1989 natural forests to align with international carbon reporting requirements. No individual tree location information.
	FMA	Scion has limited access to this dataset	Designed to store measurement following the FMA standard for property specific carbon modelling. Because of the property specific nature of the data there is less focus on uniformity and compatibility of forest type and forest management descriptors used, particularly in character fields. No individual tree location information.
	NVS databank	Open access	A databank containing records of over 109,000 vegetation survey plots - including data from over 25,000 permanent plots. It is suitable for vegetation description, studies of species distributions, and studies needing only coarse measurement of changes in vegetation. These data can be used for modelling indigenous tree species growth and forest vegetation. Details of this database can be found on https://nvs.landcareresearch.co.nz/Home/Index
Climate	NIWA	Scion has access to this dataset	Daily records of climate variables such as temperature, precipitation, radiation hours, Potential evapotranspiration, Relative humidity, soil moisture, water vapour pressure, wind speed etc. Details of this dataset can be found on NIWA's website: https://cliflo.niwa.co.nz
	WorldClim	Open access	A dataset of global climate layers (gridded climate data in GeoTiff format) that can be used for mapping and spatial modelling. WordlClim contains average monthly climatic gridded data for the period 1970-2019 with different spatial resolutions, from 1 km2 to 340 km2. The dataset includes monthly climatic variables (e.g., minimum, mean and maximum temperature, precipitation, solar radiation, wind speed and water vapour pressure) and 19 derived bioclimatic variables (e.g., annual mean temperature, mean diurnal range, isothermality, temperature seasonality, max. temperature of warmest month, min. temperature of coldest month, temperature annual range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest month, precipitation of driest month, precipitation seasonality, precipitation of wettest quarter, precipitation of driest quarter, precipitation of warmest quarter, precipitation of coldest quarter). R package "raster". https://www.worldclim.org

	WBclimte	Open access	A climate dataset provided by the World Bank used on their Climate Change Knowledge Portal. All the climate data are derived from 15 global circulation models, the most comprehensive physically-based models of climate change available and used by the Intergovernmental Panel on Climate Change (IPCC) 4th Assessment Reports. The models simulate the response of the global climate system to increasing greenhouse gas concentrations. These data have been aggregated to both the country and basin levels. Note that these data are modelled estimates of temperature and precipitation changes in different time periods under different GCMs and scenarios. They include changes for future time periods and also as "backcasting" (model representations of the past) set for past time periods. The latter should not be confused with any instrumental or observed data. R package "rWBclimate". Details of this dataset can be found on https://datahelpdesk.worldbank.org/knowledgebase/articles/902061-climate-data-api
LiDAR data	Scion LiDAR datasets	Scion has access to the datasets	Scion maintains a series of high quality LiDAR datasets. It is important to link these LiDAR datasets with PSPs database, so the PSPs database can comprise ground-based plot measurements with LiDAR information.
	LUCAS NZ Forest LiDAR	Open access	The LiDAR data were captured between 2006 and 2015, and comprise swaths approximately 70m-wide over various pre-1990 and post-1989 forest plantations. Metrics for multiple surfaces (tree-top to ground) can be extracted from this LiDAR data to characterise forest structure and ground elevation. Locations of available LiDAR data and related orthophotography can be found from LUCAS NZ Forest LiDAR Footprints: https://data.mfe.govt.nz/layer/88106-lucas-nz-forest-lidar-footprints/
Wood quality (Preferred)	Scion integrated wood quality datasets	Developing Accessible	Integrated wood quality data from Discbot were collected at high resolution from trees' disc and core samples cross sections. This dataset is developing now within Resilient Forest project sub-RA2.1 and can be used for modelling together with other available wood quality datasets.
Soil (Preferred)	MfE soil dataset	Open access On request	The Ministry for the Environment can make available a variety of soil data sets obtained from both historical sources and ground-based field inventories. Requests for planted forest and soil data can be made to the Ministry for the Environment's LUCAS team.
	NSD	Open access	The National Soils Database (NSD) is a 'point' database containing descriptions of about 1,500 New Zealand soil profiles, together with analytical data on their chemical, physical, and mineralogical characteristics. The information is obtained from physically sampling and observing the soil on site. The NSD is available through the National Soils Database Viewer: https://viewer-nsdr.landcareresearch.co.nz/search, or by contacting the NSDR team of Landcare Research.
	LCDB	Open access	The New Zealand Land Cover Database (LCDB) is a multi-temporal, thematic classification of New Zealand's land cover. It identifies 33 mainland land cover classes. Details of LCDB can be found on Land Resource Information Systems: https://lris.scinfo.org.nz/
Ecological & Physiological data (Preferred)	TRY	Open access	TRY is the most comprehensive archive and database of global plant trait (morphological, anatomical, biochemical, physiological or phenological features of individuals or their component organs or tissues). The latest TRY (released in March 2019) contains almost 11 million trait records for 2100 traits of 4 million individual plants, representing 160,000 plant taxa (mostly species). About half of the data are geo-referenced, providing a global coverage of more than 15,000 measurement sites. The

			flora of New Zealand is well presented in TRY. Details of TRY can be found on its website: https://www.try-db.org/TryWeb/Home.php
	Ecological Traits of New Zealand Flora	Open access	This database holds a list of all native and all naturalised higher plant species in New Zealand. The database covers world and local distributional data, morphological, reproduction, and other attributes that influence species' environmental responses, and interactions with other plants and animals. Attributes relevant to weed management such as response to damage, reproduction capacity, seed longevity, and weed status elsewhere are included. The database does not capture information on weed control methods. Ecological traits for individual species are organised into six groups of attributes: Status in New Zealand, Distribution, Morphology, Reproduction, Flower and Fruit, Wetland and miscellaneous. Details of this dataset can be found on Landcare's website: https://ecotraits.landcareresearch.co.nz/

Acknowledgements

Funding for this research came from Scion and the Forest Growers Levy Trust, with the support of the NZ Forest Owners Association (FOA) and the NZ Farm Forestry Association (FFA).

References

Hayes, J., & Andersen, C. (2007). The Scion permanent sample plot (PSP) database system. *New Zealand Journal of Forestry, 52*, 31 - 33.

Kunstler, G., et al. (2013). Sustainable management, earthquake disturbances, and transient dynamics: modelling timber harvesting impacts in mixed-species forests. *Annals of Forest Science*, 70(3), 287-298.

Lin, Y., Dowling, L., & Meason, D.F. (2018). Data resource for modelling less common planted species - Interim report to MPI. Rotorua New Zealand Forest Research Institute. Ltd.

Ministry for Primary Industries. (2017). *A Guide to the Field Measurement Approach for Forestry in the Emissions Trading Scheme.* Retrieved from https://www.mpi.govt.nz/growing-and-producing/forestry/forestry-in-the-emissions-trading-scheme/using-the-field-measurementapproach/

Ministry for the Environment. (2010). *Measuring carbon emissions from land-use change and forestry: the New Zealand land-use and carbon analysis system*. Retrieved from Wellington:

MPI. (2017) Carbon Look-up Tables for Forestry in the Emissions Trading Scheme. In. Wellington: Ministry for Primary Industries.

Watt, M., Adams, T., Marshall, H., Pont, D., Lee, J., Crawley, D., & Watt, P. (2013). Modelling variation in Pinus radiata stem volume and outerwood stress-wave velocity from LiDAR metrics. *New Zealand Journal of Forestry Science*, 43, 1