Chapter B4 - DESCRIPTION OF AGGREGATION OF TREE CROP AREAS

Standard for Description of Aggregation of Tree Crop Areas

Purpose

The purpose of this standard is to ensure that if the units of land area as recorded in the stand record system or generated spatially from base GIS data, are aggregated for the purposes of analysis or forest modelling, then the aggregation process is described and declared and the aggregation actually undertaken follows that process and any loss of information due to aggregation is noted.

STANDARD B4.1
Aggregation Procedure

For aggregation, the forest description shall:

- **state the purposes for the aggregation process;**
- **describe the aggregation process, including the rules for:**
  - **determining the group membership of each base land unit**
  - **combining the attributes of the base land units into group attributes;**
- **confirm that the membership rules are compatible with, and supported by, the attributes of the base land units;**
- **confirm that all the area is accounted for and no area is omitted or double counted on aggregation;**
- **ensure that the system that is used is capable of being audited to confirm the appropriate:**
  - **assignment of base land units to groups**
  - **combination of their attributes; and**
- **declare that the aggregation is fit for purpose without material loss of information and/or bias in the aggregation process.**

Guidance Notes on Aggregation

Formerly 'croptyping'

This chapter went under the title 'Croptyping' in the previous version of the standards. The term 'croptype' is used in the New Zealand forest industry to variously describe:

- the development of planning units, including aggregation;
- the identification of areas that are nominally uniform with respect to yield during the development of yield tables and sometimes to the development of the yield tables for those areas,
- the identification of areas where the crop is uniform with respect to any attribute, and
- inputs, and the development of those inputs, for some forest estate modelling systems; for example TigerMoth plunits and Woodstock development types.

So as to reduce the risk that the industry-specific, but overloaded, term "croptyping" might distract from the key valuation concerns of this chapter, which are potential loss of information and bias during aggregation, it has been replaced with the generic term "aggregation".

Background

This standard is about the aggregation of base units of land area (e.g. polygons or stands) for the specific purpose of reducing a forest estate model to a manageable size. Aggregation involves selectively losing non-critical information by grouping base land units and combining their attributes, such as area and yield, in a manner that serves the purpose for which the forest description is to be employed. Aggregation employs two distinct steps:

1. Assignment of each base land unit to a group. Examples include:

- Assigning parts of multiple stands to a harvest planning unit (coupe) which will, for the purposes of planning have a single yield table and harvest year.
- Assigning all intensively managed, radiata pine stands in the Far North District to a single National Exotic Forest Description croptype; and
- assigning all mature stands that are between 100 km and 150 km from a mill to a single planning supply source.

2. Calculation of group attributes by combination of the attributes of all the contributing land units, for example:
    - summing area;
    - averaging age or yield on an area-weighted basis;
    - averaging revenue weighted by volume and area; and
    - taking a modal category (e.g. using the clearfell year with the largest contribution to area).

The outcome of aggregation is a smaller number of land units than in the base data set, with each of the resulting land units considered to be uniform in its attributes (e.g. uniform age, yield, cost and planning intent).

Categorisation is a necessary prerequisite to aggregation, but is not the focus of this standard. Saying that two stands are in the same planning unit is categorisation. Averaging their yield tables so that the planning unit has one yield table instead of two is aggregation; it is the averaging that loses information.

In addition to resulting in lost information, the aggregation of attributes can also introduce bias, particularly when time is involved, because of the non-linear relationship between the discount factor[1] and time. The classic example of the potential for bias is when young stands with high expectations of mature yield are combined with older stands with a lower expectation of mature yield. This results in an average yield table that at clearfell age overstates the yield of the older stands and understates the yield of younger stands. The total expected volume over a long period of time may be correct in this example, but the discounted woodflows and cashflows will be overstated.

Purposes of aggregation

The term 'manageable size' covers a number of basic needs, including:

- ease of comprehension or communication;
- acceptable solve times for linear or integer programming solutions; and
- fitting within finite resources such as computer memory.

The business purpose for aggregation is usually associated with the development of planning units, often harvest-related and with greater spatial resolution for older stands. These planning units might not be suitable for valuation.

The reason for requiring disclosure of the purposes of aggregation in the forest description is that information loss that is immaterial in one context may be very material in another. For example, aggregating all of the younger stands into a single group may have no material effect on a short-term clearfell plan, but could be quite inappropriate for a valuation that targets a specific class of ownership in younger stands. A forest valuer may not have control over the level of aggregation, but they should understand its consequences.

It is useful to look at the process of getting from forest information through to forest estate model input as having four steps:

[1] The discount factor, where i is discount rate ($/$/year) and t is time (years), is not a linear function of time. Given a set of different time values (e.g. years until harvest for a group of stands), the discount factor calculated from the average time is not the same as the average of the discount factors calculated for each time.

1. **Disaggregation**. Disaggregate the forest into a complete set of mutually exclusive units of land area, each of which is uniform in the information that is necessary for modelling purposes, in this case valuation. It used to be safe to call these units stands because conventions and the technological limitations of the time meant that modelling information was stored against stands, which were defined in terms of adequate uniformity for most modelling purposes. Increasingly, the unit of land area that serves this purpose is the set of polygons that are the union of a number of GIS layers, each of which provides different information (e.g. ownership, site quality, harvest plan, inventory population etc). Attached to each land unit are categorical and continuous attributes representing the **base** information for the following steps.
2. **Derivation**. Attached to each land unit are attributes **derived** from the base information. For example, in a yield context one might determine that a distinct yield table should be used that reflects the existence of a past inventory operation, the intention to thin at some point in the future, and the location within a forest where a specific growth model and taper function are prescribed. The derived attribute in this case might be a yield table identifier and the base attributes might be obtained from many different base layers. Alternatively, one might derive a transport cost from the distance from wood catchment to market taken from a base layer.
3. **Subsetting**. Remove those land units that do not serve the purposes of this model (e.g. drop those that are not in the forest, ownership, productivity category or rotation that we are valuing based on base and/or derived information).
4. **Re-aggregation**. Aggregate the land units for the very specific purpose of reducing a forest estate model to a manageable size. The key points about this step are:
    a. we could, in theory, run a forest estate model against the entire set of land units because each has all of the necessary information, but we do not because we would run out of time or computer memory or would not be able to explain the results;
    b. in this step we are not attaching new information; but
    c. existing information might be lost as a result of aggregation (e.g. using the average yield or the average planned harvest year instead of different values for each polygon).

In practice, it is rare to observe a process that strictly follows the four steps as described. For example, it is common for efficiency reasons to carry out different parts of No. 2 at different levels of disaggregation. However, this does not remove the general usefulness of thinking in four steps because the effect of intersecting, then classifying, then intersecting again, should be identical to the effect of intersecting all the way to the lowest common denominator land unit, then classifying.

We can identify key concerns of the valuation standards for each of these steps and the transitions between them, for example:

1. Not losing or gaining total area other than as intended by **subsetting.**
2. **Derivation** according to well-defined, documented and plausible standards.
3. Not losing information in the **re-aggregation** step to the extent that we have a material effect on the outcome (i.e. the forest value).

Revision History

Original Standard	Released in May 1999
Revision in August 2020	Main changes are: • changing the emphasis from croptyping to aggregation. To reflect this, the title has been changed from Standard for Description of Croptyping to Standard for Description of Aggregation of Tree Crop Areas; • changing Standard B4.1 from Croptyping Procedure to Aggregation Procedure; and • deleting Standard B4.2 on Presentation of Croptyping.