CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Retro-Fitting Value-Creation Potential Indicators to Long-Term Supply Models

Gregory Paradis
Luc LeBel

May 2018

CIRRELT-2018-23

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Retro-Fitting Value-Creation Potential Indicators to Long-Term Supply Models

Gregory Paradis[1,2,*], Luc Lebel[2,3]

1. University of British Columbia, Department of Forest Resources Management, 2424 Main Mall, Forest Sciences Centre 2312, Vancouver, BC V6T 1Z4
2. Département des sciences du bois et de la forêt, Pavillon Abitibi-Price, 2405, rue de la Terrasse, Local 2121, Université Laval, Québec, Canada G1V 0A6
3. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

Abstract. It has been shown that credibility of wood supply optimization models can be improved by using a bilevel formulation that anticipates industrial fibre consumption. The upper level model corresponds to the standard long-term wood supply optimization model, and the lower level corresponds to a short-term network flow optimization model. The lower-level model maximizes profit from sale of primary forest products to exogenous markets. To compile such a model, we must be able to disaggregate species-wise timber volume output from the upper-level model into assortments of logs, and estimate value-creation potential of these assortments. Wood supply models used in many jurisdictions (including those used in Quebec, Canada) do not feature value-creation potential performance indicators. We describe a methodology for retro-fitting value-creation potential indicators to these wood supply models, based on exiting data sources and a previously-published volume disaggregation method. Our methodology greatly simplifies the otherwise onerous task of compiling value-creation potential indicators from available data. Although our method is specifically designed to be compatible with data and model structure used in Quebec, it could also be adapted to other contexts with relative ease, as a first step in implementing a value-driven forest planning process.

Keywords. Forest management, hierarchical planning, value-creation potential.

Acknowledgements. This study was supported by funding from the FORAC Research Consortium.

---

* Corresponding author: gregory.paradis@ubc.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2018

1 Introduction

Paradis et al. (2018) describe a bilevel wood supply model formulation that reduces risk of wood supply failures. They demonstrate a potential application of their bilevel model formulation using a realistic synthetic dataset. Linking upper- and lower-level models requires disaggregation of harvest volume output from the upper-level model into log assortments (by size, species, and quality) which are dispatched to different processing units in the lower-level model. The lower-level model simulates processing these logs into primary forest products, and sale of products to exogenous markets. The objective function of the lower-level model maximizes total network profit, which is estimated from the sum of revenue net of fibre procurement cost, processing cost, and transportation cost. Implementing such a bilevel model in practice presupposes availability of disaggregation coefficients for upper-level volume output, as well as unit value-creation potential (VCP) coefficients to estimate unit profit for all possible fibre flow paths through the lower-level network.

In practice, neither the disagregation coefficients nor the VCP coefficients have been compiled for wood supply models in Quebec, Canada. Furthermore, the task of compiling these coefficients is sufficiently complex and fraught with technical and methodological challenges so as to represent a substantial impediment to implementation of value-driven forest management planning. We endeavor to bridge this gap by developing reproducible and computationally tractable methodology, complete with purpose-built software implementation, which can be applied immediately using readily-available data.

This documents extends the methodology for compiling the disaggregation coeffients, described in Paradis and LeBel (2017a), by linking disaggregated volumes to value-creation data from an existing database and finally re-aggregating these into useful VCP coefficients that can be retro-fit to existing wood supply models. We present this method here, along with some sample output for management unit UA 064-51.

The remainder of this paper is organized as follows. We present some background information in §2 and describe our methodology in §3. Sample results from application of the method to management unit UA 064-51 are presented in §4, followed by discussion in §5 and concluding remarks in §6.

2 Background

Determining AAC in Quebec is the responsibility of the *Bureau du forestier en chef* (BFEC), an independent branch of the provincial government. BFEC analysts use the Woodstock modelling platform to model long-term wood supply, for each of 71 management units that constitute the public forest of Quebec. The Woodstock software does not feature a scripting interface, so any changes to the model input files must be applied manually (i.e., via keyboard and mouse). Furthermore, Woodstock models in Quebec are automatically generated by an in-house model-compiling interface (Horizon CPF), which results in relatively verbose models. For example, the Woodstock model for management unit UA 064-51, which we use as a test dataset to illustrate our methodology, contains over 600 000 lines of code. The task of retro-fitting value-creation-potential indicators (compiled using our methodology) to these models by manually editing the Woodstock code is too time-consuming to be practical. To make the retro-fitting task more tractable and results less error-prone, we use the `ws3` software library[1], which simplifies the tasks of importing and interpreting wood supply model input files, retro-fitting complex new indicators to these models, and interfacing with downstream software modules to create a modelling pipeline. We describe this workflow in more detail later in this document.

A branch of the Quebec government responsible for marketing fibre harvested from public forests (*Bureau de mise en marché des bois*, or BMMB) has published a simulation model (MERIS) which can be used to estimate value-creation potential of a stand, given a detailed stand table describing current inventory (in terms of the 45 standard species codes used in the Quebec forest inventory, and 26 2-cm-wide stem diameter size classes). BFEC Woodstock model output is aggregated in terms of 10 species group codes (with no stem diameter size class information). Paradis and LeBel (2017a) describe a methodology for compiling *disaggregation coefficients* that can be used to bridge the gap between these two models. We developped a methodology that uses these disaggregation coefficients to map wood supply model output to value data in the MERIS database and re-aggregate the value data to match the original wood supply model output aggregation level, effectively retro-fitting new value-creation potential indicators to the existing wood supply model. As mentionned earlier, the technical complexity of the process (combined with the large size of the original wood supply models) makes it virtually impossible to apply this method without a purpose-built intermediate software layer.

Our methodology can be applied with relative ease (compared to an *ad hoc* approach) to any forest management unit in Quebec, thereby making the bilevel wood supply modelling approach described by Paradis et al. (2018) much more accessible for researchers and forest practitioners. The method can be used to compile VCP coefficients for network flow optimization models, for example the LogiLab model described in Jerbi et al. (2012) and Bouchard et al. (2017), or other forest sector supply chain models. Alternatively, the method can be used to compile *a priori* value-creation indicators for long-term wood supply models (as opposed to *post hoc* injection of these indicators into the optimal solution, as we show here), in support of a potential shift towards value-driven wood supply planning.

Note that we developed this methodology in the context of a larger study, whose goal is to explore innovative business models to realize value-creation-

[1] See `http://ws3.readthedocs.io` for documentation of the `ws3` software library, which is freely downloadable from `http://github.com/gparadis/ws3`. The use-case described here was implemented using Jupyter Notebooks—the notebooks are available from the corresponding authour upon request. Please note that running the notebooks requires specific datasets—although these datasets are readily available upon request, terms of use of these datasets do not allow us to distribute the data directly.

potential from Quebec forests—our government and industry partners expect concrete, implementable solutions to relevant problems. We mention this to clarify why we tailored our methodology so specifically to the MERIS database and the Woodstock model format used by the BFEC in Quebec. Notwithstanding the Quebec-specific details and model design choices, we hope that the methodology presented here will be a useful framework for researchers and practitioners in other jurisdictions wishing to link long- and short-term models.

As an example of the application of our method and software framework, we compile value-creation potential of simulated harvest volume from a Woodstock model for management unit UA 064-51. We express output in terms of the 10 species groups, 3 cover types, 3 treatment types, and 26 stem size classes used in Paradis and LeBel (2017a).

3 Methods

The BFEC Woodstock models and the BMMB MERIS model were not designed to be compatible with each other. However, they are both designed to model forest management activities from public forests in Quebec, albeit at different scales. Thus, much of the information represented in these models is *conceptually* compatible, although the data used to represent this information is stored at different aggregation levels.

Documentation of the code structure of the BFEC Woodstock models and the MERIS database are limited, and both are quite complex. Not surprisingly, the methodology we developed to link these two models is also complex—a common consequence of working with real data and real models. We endeavoured to keep the description in this document as short as possible, while providing sufficient detail to facilitate replication of our methods.

The rest of this section is divided into two subsections. The first subsection provides an overview of the main steps in the methodology, and the second subsection describes each step in more detail.

3.1 Overview

In summary, our methodology extracts information from the MERIS database and *post hoc* injects VCP performance indicators into the development types harvested in the solution of a Woodstock wood supply model. This task (of retro-fitting VCP performance indicators into existing wood supply models) represents the culmination of the first phase in a larger modelling process—the second phase of this process, which we describe in Paradis and LeBel (2018), uses the VCP-augmented wood supply model as input for a network flow model that simulates fibre consumption behaviour of a network of profit-maximizing mills. Our methodology can be broken down into two steps.

In the first step, we compile 90 disaggregation coefficient vectors using the methodology described in Paradis and LeBel (2017a). These vectors of coefficients allow us to disaggregate harvest volume into the same 26 stem size class bins used to store VCP data in the MERIS database. Each of these vectors maps to one of the 90 cases of an intermediate aggregation scheme. This intermediate aggregation scheme was chosen such that *(a)* it is compatible with the Woodstock model format, *(b)* it is compatible with the MERIS database format, and *(c)* data is available to compile reasonable disaggregation coefficients for each case of this scheme. It is composed of 90 combinations of 10 species groups (documented in Table A1), 3 cover types (softwood, mixedwood, hardwood), and 3 harvest treatment types (clearcut, selection cut, commercial thinning).

In the second step, which is the focus of this document, we map value-creation potential data from the MERIS database onto our disaggregated volumes, and re-aggregate the data to produce the value-creation indicators we need for subsequent phases of our fibre supply modelling project. The MERIS database actually contains two distinct value-creation models. The first value model in MERIS represents *financial* VCP, from the perspective of an industrial facility that procures raw fibre from public forests, transforms this fibre into one or more primary forest products (and co-products, such as chips), and sells these products to end-customers in external markets at exogenously-defined prices. This is the model we use, when importing data from MERIS. We will be referring to this model in the remainder of this document, unless otherwise specified. The second value model in MERIS represents *economic* VCP, from the perspective of a government steward managing fibre flow from public forests for the benefit of society as a whole. Note that the first step of the methodology we describe here (i.e., disaggregation of wood supply model harvest volume) could potentially also be used to map wood supply models to the economic value model in MERIS, although we have not tested this.

The financial VCP model in MERIS is composed of six components: fixed cost, harvest cost, sylviculture credits, stumpage cost, transportation cost, and product values. The *fixed cost* component includes fixed costs related to fibre procurement (i.e., general administrative costs, access road planning and amortization costs)—it is expressed on an area (i.e., ha^{-1}) basis. The *harvest cost* component includes all variable costs associated with fibre extraction, including cost of loading logs onto trucks for transportation to processing facilities—it is expressed on a volume (i.e., m^{-3}) basis. In Quebec, cost of implementing prescribed non-commercial sylviculture treatments must be assumed by the licencee that harvests the fibre, however this cost is (mostly) offset by a credit applied to future stumpage fees. The *sylviculture credit* component this credit—it is expressed on an area (i.e., ha^{-1}) basis. In Quebec, a stumpage fee must be paid for each unit of fibre harvested from public forest. This corresponds to the *stumpage cost* component—it is expressed on a volume (i.e., m^{-3}) basis. The *transportation cost* component estimates cost of transporting fibre from the harvesting site to the processing facility (not including loading cost, but including unloading cost)—it is expressed on a volume (i.e., m^{-3}) basis. The *product value* component estimates revenue from sale of all primary products and co-products that will be produced from a given unit of fibre, net of processing cost and cost of transporting products to markets.

Figure 1 provides a schematic overview of the various steps in our modelling methodology.

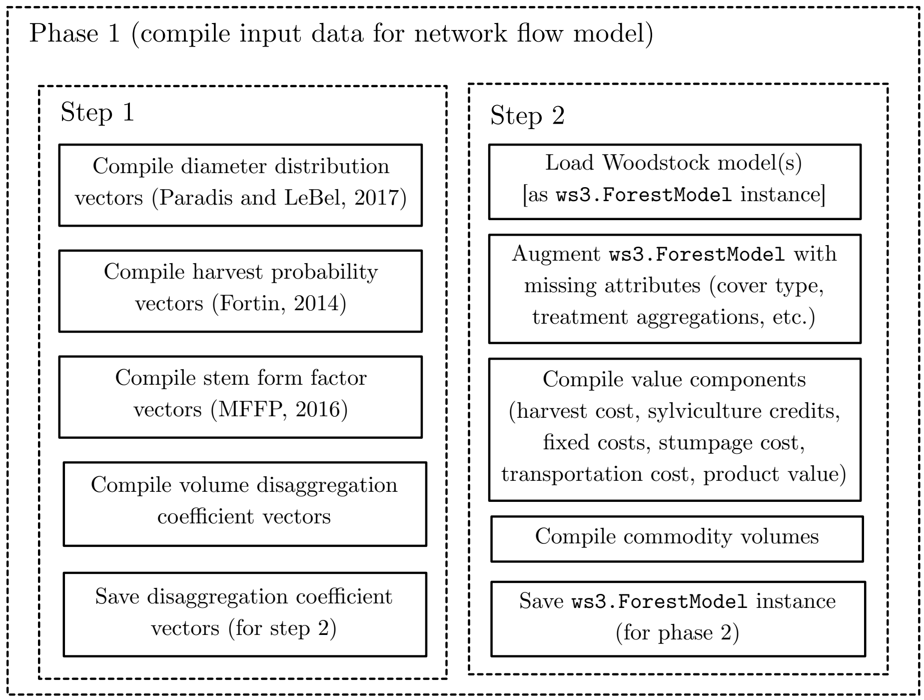

Figure 1: Schematic representation of modelling methodology for phase 1 (includes step 1, as described in Paradis and LeBel, 2017a).

The next subsection provides a more detailed description of our methodology.

3.2 Detailed Methods

As described earlier, the method is decomposed into two sequential steps, which we describe below.

The first step, *disaggregation of harvest volume*, is documented in Paradis and LeBel (2017a). An important part of this first step is modelling of stem diameter distribution, is documented in Paradis and LeBel (2017b)). The text below provides a high-level overview of this first step, and also documents a few of the key implementation details that we needed to address to apply the method to the UA 064-51 test case.

The second step, *compiling value-creation indicators*, is the main topic of this document, and is described in more detail.

3.2.1 Step 1: Disaggregating harvest volume

We start by importing a Woodstock model into `ws3`, loading the solution stored in the model (i.e., the solution used to determine AAC) into memory, and simulating its application sequentially for all planning periods.

In Woodstock lingo, a *development type* is equivalent to the combination of a forest stratum (i.e., unique combinations of stratification variables, or *themes*) and an ageclass. Woodstock solutions are composed of units of area of a given development type on which a given treatment is applied at a given period.

We want to aggregate our data by broad cover type (softwood, mixedwood, hardwood), but there is no attribute in the BFEC Woodstock models that directly maps to this aggregation level. We can, however, use yield components `yg_g_gf` (hardwood basal area) and `yg_g_gr` (softwood basal area) to deduce the cover type of a given development type—we can then inject a new cover type attribute (i.e., *yield component*) into each development type in the model. We define softwood cover type as having at least 75% softwood, hardwood cover type as having at least 75% hardwood, and the rest is mixedwood. We analyse basal area at optimal rotation age (i.e., maximum mean annual increment), so as to maximize classification accuracy of stands at maturity. This should correspond to the age at which the `yagemat` yield component equals 0 in our BFEC Woodstock models, but we found this attribute to be unreliable so instead we use `ws3` to compute maximum mean annual increment (MAI) age on the fly from the total volume yield component `yv_s`—this age corresponds to the optimal rotation age, which should be highly correlated with harvest age in even-aged stands in the harvest schedule.

We need to aggregate treatments in the Woodstock dataset into 3 types (1: final cut, 2: selection cutting, 3: commercial thinning) so we can disaggregate harvest volume using function $p_{cst}(x)$, as described in Paradis and LeBel (2017a). So, we add three aggregate actions to our model corresponding to these. Although some actions are difficult to classify, one must bear in mind that these values will be used to select the best volume disaggregation function for each treatment. The type 1 disaggregation profile assumes all stems are harvested, and so the stem size distribution matches standing inventory. Type 2 and 3 disaggregation profiles are based on a model published by Fortin (2014). Note that, for the softwood cover type, we restricted our stem size distribution analysis to high density (class A and B), high basal area (28+ $m^3 ha^{-1}$) stands. Applying commercial thinning treatments to mixedwood or hardwood cover types is just bad silviculture, so we are assuming that this will not come up in the model.

Next, we map the 45 species codes in the forest inventory to the 10 species group names used in the MERIS database. Fortunately (although not entirely coincidentally), there is a 1:1 match between the 10 species group definitions in MERIS and in the BFEC Woodstock models. See Table A1 for details.

Finally, we compile 90 vectors of 26 disaggregation coefficients (one per DBH stem size class), using the methodology described in Paradis and LeBel (2017a).

3.2.2 Step 2: Compiling value-creation indicators

We will now describe in more detail how data for each of the six value components are extracted from the database and compiled into a performance indicator that can be retro-fitted to the Woodstock model solution.

Although this step is somewhat complex, its execution is relatively straightforward using our scripted Python implementation. The first few steps involve importing product distribution, stumpage fee, transportation cost, and product value data from MERIS database. Once all the required data has been imported from the MERIS database and loaded into memory, we begin compiling this data to create the 6 value components described earlier, which we inject directly into a live `ws3` simulation.

The MERIS system defines two *profiles*, corresponding to hardwood and softwood sawmills. Some value component values vary depending on the profile selected by the user. We import fixed cost data from both hardwood and softwood sawmill profiles, assigning values from the hardwood sawmill profile to hardwood development types and values from the softwood sawmill profile to softwood and mixedwood development types.

Next, we import stumpage rate data from the MERIS database. Stumpage rates in Quebec vary by tarification zone. The geographical boundaries of these tarification zones do not always line up with management unit boundaries. Thus, each management unit may overlap several zones. The BFEC Woodstock models include a theme (i.e., a stratification variable), that specifies the tarification zone of each development type. Within a stumpage tarification zone, rates are specified in terms of species group and product. The MERIS database defines empirical product distributions for each combination of 45 species and 26 stem size classes. 8 merchantable products are defined in MERIS: veneer logs, 4 hardwood sawlog grades (F1, F2, F3, F4, according the classification scheme defined in Petro and Calvert, 1976), 2 softwood sawlog grades (small sawlogs correspond to DBH size classes 10 through 14, and large sawlogs correspond to DBH size classes 16 and up), and pulpwood. 3 unmerchantable products are defined in MERIS: rot, unutilized, and other. Note that stumpage is defined by species (rather than species group), so we must compile species-wise weight parameter vectors for each of the 30 combination of species group and cover type from the permanent sample plot data used in step 1 (see Paradis and LeBel, 2017a for details).

Next, we import transportation cost data from the MERIS database. Transportation cost coefficients in MERIS are compiled, for each stumpage zone, using 10 product-species groupings (henceforth referred to as *commodities*). Note that we split white and yellow birch pulpwood into two commodities (we will need this to map to timber licence contract aggregation level further downstream, when compiling the network flow optimization model), although both species are modeled as one commodity (birch pulpwood) in MERIS. Transportation cost coefficients in MERIS are compiled by stumpage tarification zone, using the mean transportation distance for the closest three processing facilities accepting a given commodity.

Next, we import product value data from the MERIS database. Product values are keyed on species and product class, with unit values for all combinations of species and product class specified for both hardwood and softwood sawmill profiles. The assumption is that a hardwood mill may accept softwood logs, but will pay a lower price for these logs than the softwood mill (to account for the trouble of having to store these logs until they can be dispatched to a softwood sawmill). The inverse goes for softwood sawmills. We will use these values to build a network flow optimization model (in a subsequent modelling phase, beyond the scope of the current document), which will only allow valid commodity-processor flows, so we import the higher of the profile-wise prices for each combination of species and product class.

Next, we compile harvest cost and sylviculture credit for each decision in the Woodstock optimal solution. Both harvest cost and sylviculture credit are derived from complex arithmetic functions (rather than being imported directly from the MERIS database). Independent input variables for these functions include harvested volume and mean piece size.

Harvest cost is estimated using a predictive model, compiled by BMMB staff for use in the MERIS system. The BMMB model is an amalgam of machine productivity functions, originally compiled from machine time study data, combined with several assumptions regarding frequency ratio of machines that compose different systems, relative proportions of system utilization, intensity of harvest prescriptions (final cut versus partial cut), mean skidding distances, roadside sorting complexity, etc. These productivity assumptions are combined with rental rate assumptions, and machine utilization ratio assumptions. The model mixes productivity functions for feller-bunchers, single-grip harvesters, grapple skidders, forwarders, delimbers, and slashers. The end result is a prediction of unit harvest cost on a volume basis, as a function of cover type, harvest intensity, and mean piece size. The harvest cost model is presented in an appendix (see §A1.2).

Sylviculture credit value is estimated for each component of the harvest schedule using arithmetic functions the published by government for the purpose of calculating silviculture credit. The sylviculture credit model is made up of 7 different functions. First, we classify (using expert judgement) each partial cut treatment in the wood supply model to one of three classes (used by government-defined criteria for selecting sylviculture credit formulas): progressive cut, selective cut, commercial thinning. We were able to simplify the model down to 4 functions (1, 2, 4, 7). For treatments classified as commercial thinning, we use function 1 for softwood and mixedwood cover types, and function 2 for hardwood cover types. For treatments classified as selection cuts, we use function 4. For treatment classified as progressive cuts, we use function 7 for softwood and mixedwood cover types, and function 4 for hardwood cover types. The formulas are a function of harvest volume, mean piece size of harvested stems, and mean piece size of standing inventory before harvest. The four retained formulas are presented in an appendix (see §A1.3).

Mean piece size is estimated from Woodstock model yield curve data (quotient of total volume and stem density curves). Although they are accurate (confirmed with David Pothier, personnal communication, October 2015), the stem density regressions are rather imprecise (i.e., we can expect a large random error, evenly distributed about the mean). The harvest cost prediction model has an overall inverse exponential shape (i.e., inverse J shape). Thus, underestimating mean stem size would tend to induce a relatively large increase in estimated harvest cost, whereas overestimating mean stem size would induce a relatively small increase in harvest cost. The BFEC Woodstock models use a combination of two different growth models (NATURA for even-aged stands, and ARTEMIS for uneven-aged stands). If we can estimate the error distributions of both total volume and stem density curves for NATURA and ARTEMIS models, then we can calculate the error distribution of the quotient of these random variates, which we can then use to calculate the expected value of the harvest cost function.

We implemented a help class in `ws3.common.rvquot_gen`, which encapsulates functions from the `pacal` library for calculating the quotient of two normally distributed random variates. `rvquot_gen` subclasses `scipy.stats.rv_continuous`, so we can simply call `rvquot.expect(...)` on an instance of our class (which has been instantiated with appropriate scale and location parameters for the numerator and denominator random variates) to output the expected value of the harvest cost function. The `pacal` library has functions for numerical integration of complex arithmetic functions of random variates. We tried solving the harvest cost and sylviculture credit functions (with random variates) using the `pacal` numerical integration functions to estimate the expected value of these functions, but the algorithm failed to converge after 24 hours of CPU time. In the end, we implemented a brute-force Monte Carlo algorithm within `ws3` to approximate the expected values of the functions, which converges in a reasonable time.

We contacted the developers NATURA model, and confirmed that the error terms for both total volume and stem density are normally distributed. The standard deviations for these error terms are documented in the NATURA documentation (Pothier and Auger, 2011)—Tables 8 and 11 of the NATURA documentation list standard deviations (REMQ) for stem density and total volume, by bio-climactic subdomain (i.e., *sous-domaine bioclimatique*, or SDB). UA 064-51 is in SDB `3ouest`, for 3 groups of strata (by simulation horizon length). We calculated the weighted-average standard deviation, using normalized strata counts in each group as weight coefficients. The mean standard deviations, expressed as a proportion of estimator value, are $\sigma = 0.386$ for total volume and $\sigma = 0.245$ stem density. Note that these error values seem to contractict annecdotal information we obtained from David Pothier (error on stem density estimate is higher than error on total volume). Unfortunately, these are not the errors we utimately seek, because the yield curves in our BFEC models are the result of aggregating several (NATURA or ARTEMIS, depending on the case) curves to form composite curves. Assuming that the error at all points on the BFEC composite curves is IID Gaussian distributed, and assuming that we treat each component curve as a single sample (i.e., use the value of component curves directly, ignoring that they are themselves IID Gaussian distributed, as discussed above), we can estimate the standard error σ_i of a given composite curve at any age class $i \in I$ from the the values of its component curves $j \in J$ at the same age

$$\sigma_{\hat{y}_i} = \sqrt{\frac{\sum_{j \in J} (y_{ij} - \hat{y}_i)^2}{|\hat{\mathbf{Y}}| - 1}} \tag{1}$$

where $\hat{y}_i$ represents the value of the BFEC composite curve at age class i, and y_{ij} is the value of component curve j at age class i.

NATURA and ARTEMIS component curves can allegedly be obtained from MFFP. However, we also need a mapping of component curves to composite curves—we were unsuccessful obtaining this information. Without these mappings, we cannot estimate $\sigma_{\hat{y}_i}$ as described above. For the sample results presented in this study, we use a conservative value of $\sigma = 0.5$ to model both total volume and stem density as random variates, for the purposes of estimating the expected values of harvest cost and sylviculture credit functions in the sample results presents later in this document.

Note that, although we were unable to obtain a copy of all the data required to correctly estimate the error distribution on mean piece size, this data is readily available in-house to government analysts. Thus, there should not be any problem, in practice, applying the methodology described here to correctly estimate expected value of harvest cost and sylviculture credit functions.

Finally, we use all the data we just assembled to compile VCP indicators that can be injected *post hoc* into the Woodstock model optimal solution. For each harvesting decision, we compile a total of 144 new indicators (6 value components plus net merchantable harvest volume, compiled at three aggregation levels [total, commodity-wise, and species-group–wise]). We compile net merchantable harvest volume indicators by multiplying total volume at harvest age by treatment-wise coefficients embedded in the Woodstock model code. At this point, we have not automated the process of extracting these coefficients from the Woodstock models (i.e., they must be manually extracted by an expert). These net merchantable harvest volume indicators will facilitate compilation of the network flow model in a subsequent phase.

A Jupyter Notebook containing Python code implementing our methodology for management unit UA 064-51 (the notebook includes detailed explanations of each step, mixed in with the blocks of code) is available from the corresponding author upon request.

In the following section, we show the result of applying our methodology to UA 064-51. As mentionned earlier, we designed our methodology to be easilty applicable to any one of the 71 management units that make up the public forest of Quebec, and have already collected all the data required to proceed with a province-wide deployment of our method.

4 Results

As an example, we present results of applying our methodology to management unit UA 064-51. Consistently with the methodology described in the previous section, model output is aggregated in terms of 3 cover types (softwood, mixedwood, hardwood), 3 harvest treatment types (clearcut, selection cut, and commercial thinning), and 10 species groups[2].

Figures 2 through 5 show results of applying our methodology to compile VCP for management unit UA 064-51 in Quebec, Canada. Figures 2 and 3 show unit VCP ($\$m^{-3}$), whereas Figures 4 and 5 show total VCP ($, i.e., product of unit VCP and harvest volume). Species group is fixed for a given row of subfigures, and cover type is fixed for a given column of subfigures. Treatment type 1 (circles) corresponds to clearcut harvesting, treatment type 2 (squares) corresponds to selection cut, and treatment type 3 (crosses) corresponds to commercial thinning.

[2]The original wood supply models we obtained from government analysts feature 11 species groups, however we merged the *red pine* group into the *white pine* group, resulting in 10 species groups. The merging was motivated by the relative scarcity of red pine observations in the PSP dataset, which made it difficult to reliably model stem diameter distributions to observed stem data.

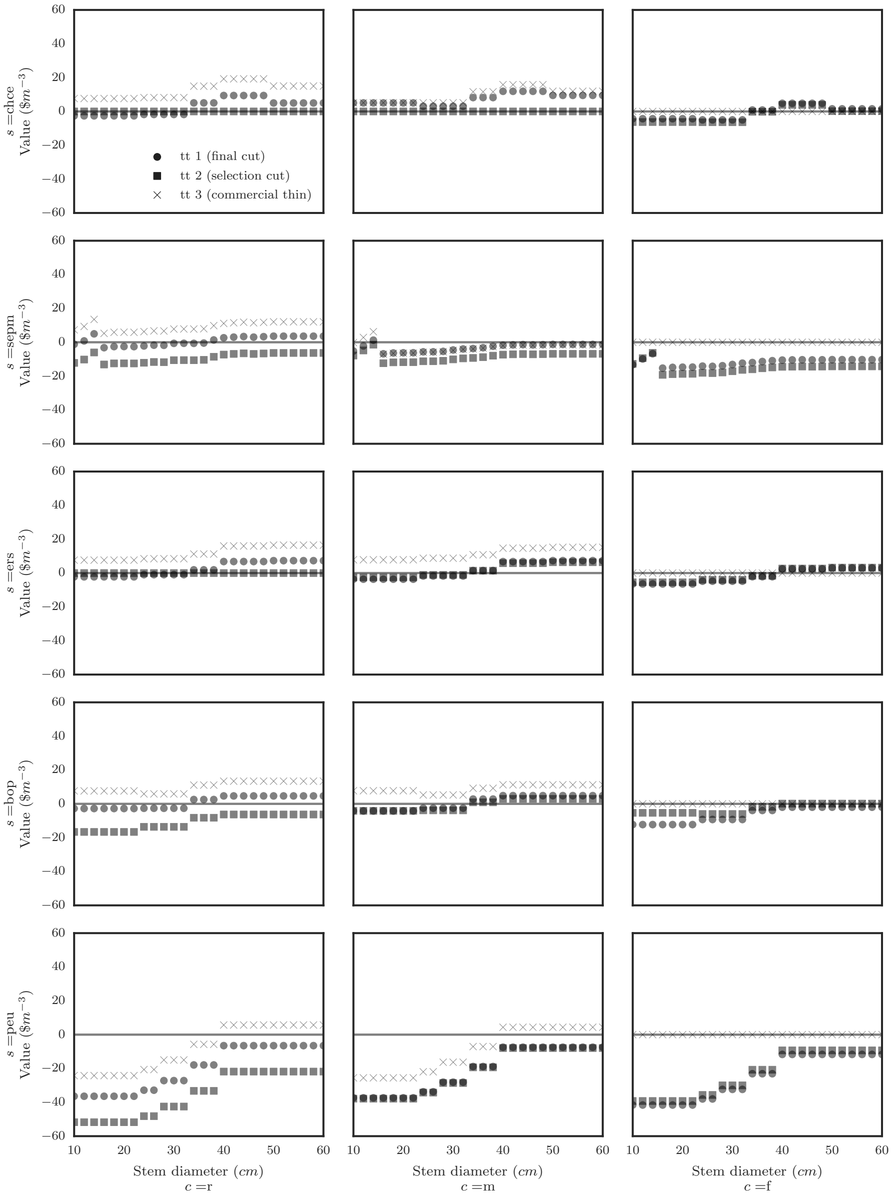

Figure 2: Unit value-creation–potential for management unit UA 064-51 in Quebec, Canada. Species group is fixed for a given row of subfigures, and cover type is fixed forn a given column of subfigures. Treatment type 1 (circles) corresponds to clearcut harvesting, treatment type 2 (squares) corresponds to selection cut, and treatment type 3 (crosses) corresponds to commercial thinning.

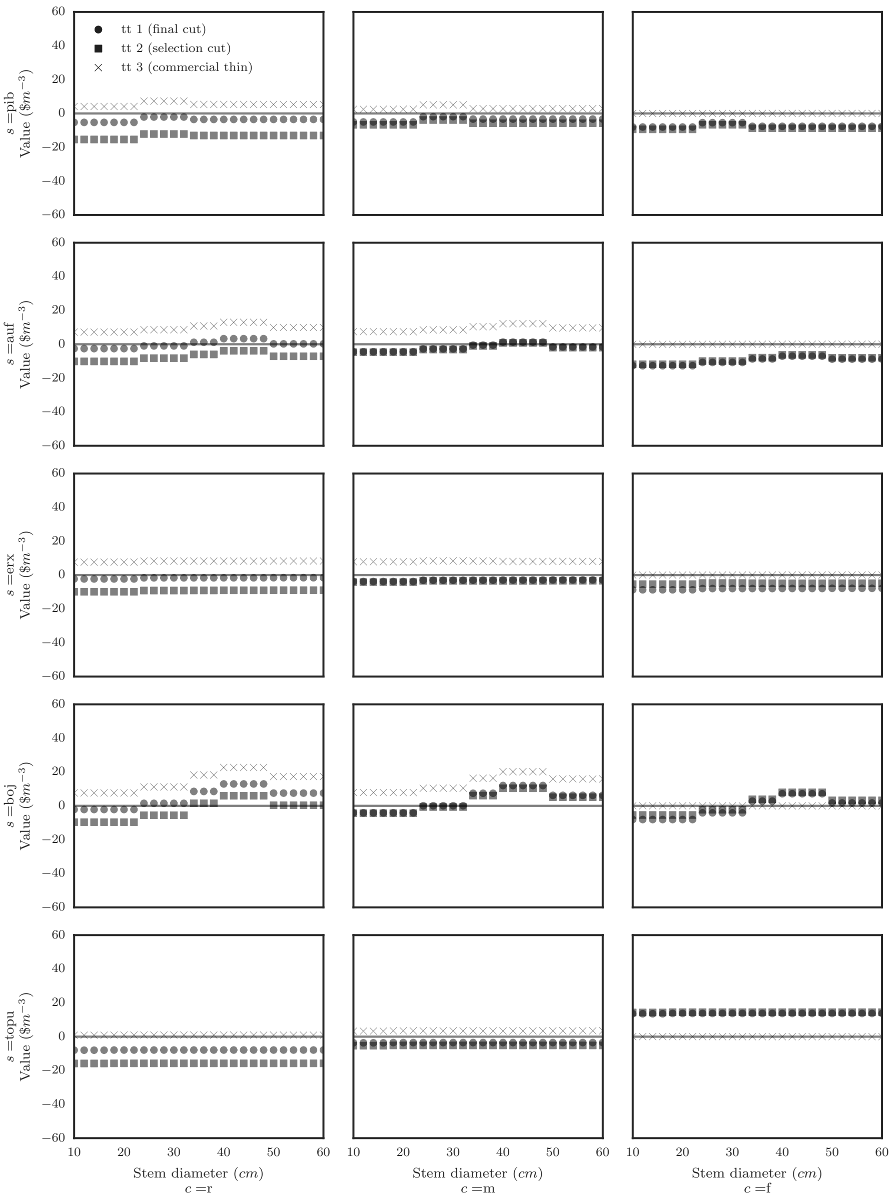

Figure 3: [Continued from Figure 2] Unit value-creation–potential for management unit UA 064-51 in Quebec, Canada. Species group is fixed for a given row of subfigures, and cover type is fixed forn a given column of subfigures. Treatment type 1 (circles) corresponds to clearcut harvesting, treatment type 2 (squares) corresponds to selection cut, and treatment type 3 (crosses) corresponds to commercial thinning.

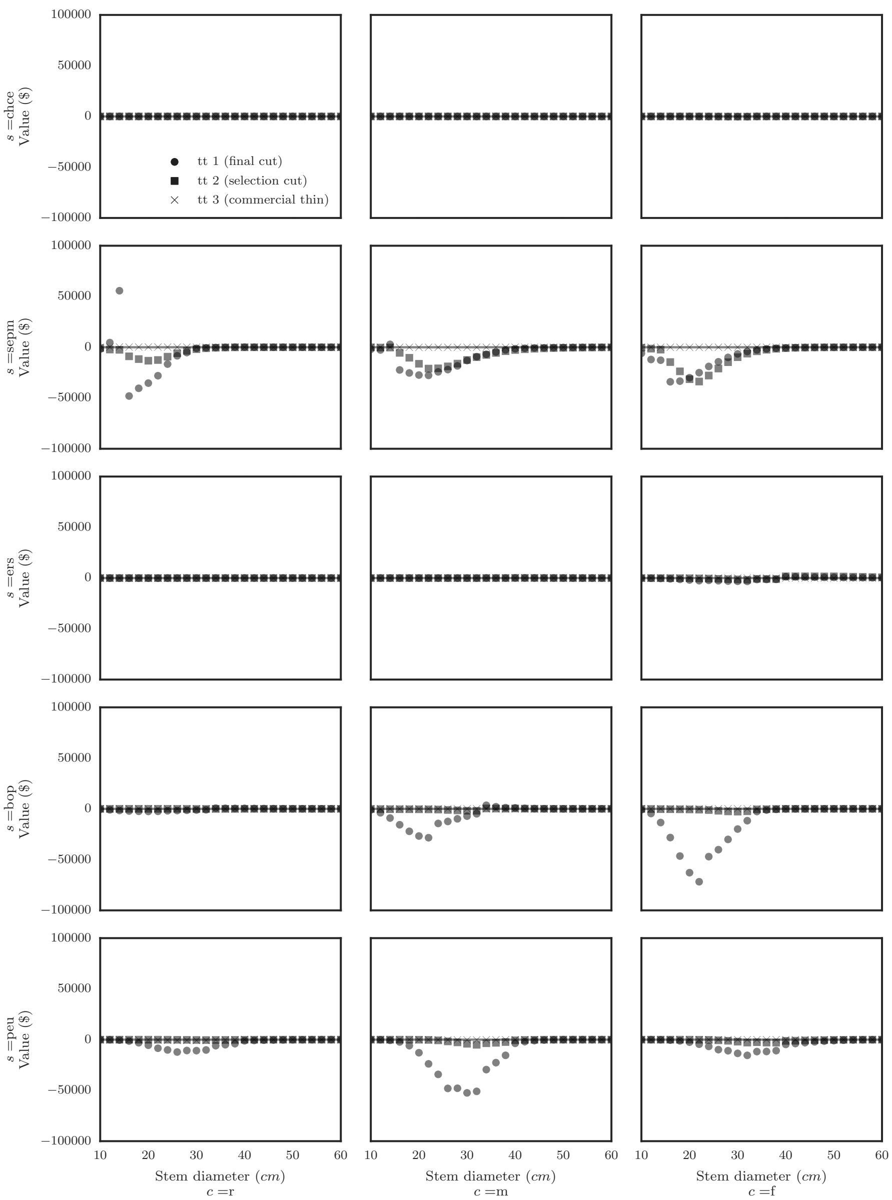

Figure 4: Total value-creation–potential compiled for management unit UA 064-51 in Quebec, Canada. Species group is fixed for a given row of subfigures, and cover type is fixed forn a given column of subfigures. Treatment type 1 (circles) corresponds to clearcut harvesting, treatment type 2 (squares) corresponds to selection cut, and treatment type 3 (crosses) corresponds to commercial thinning.

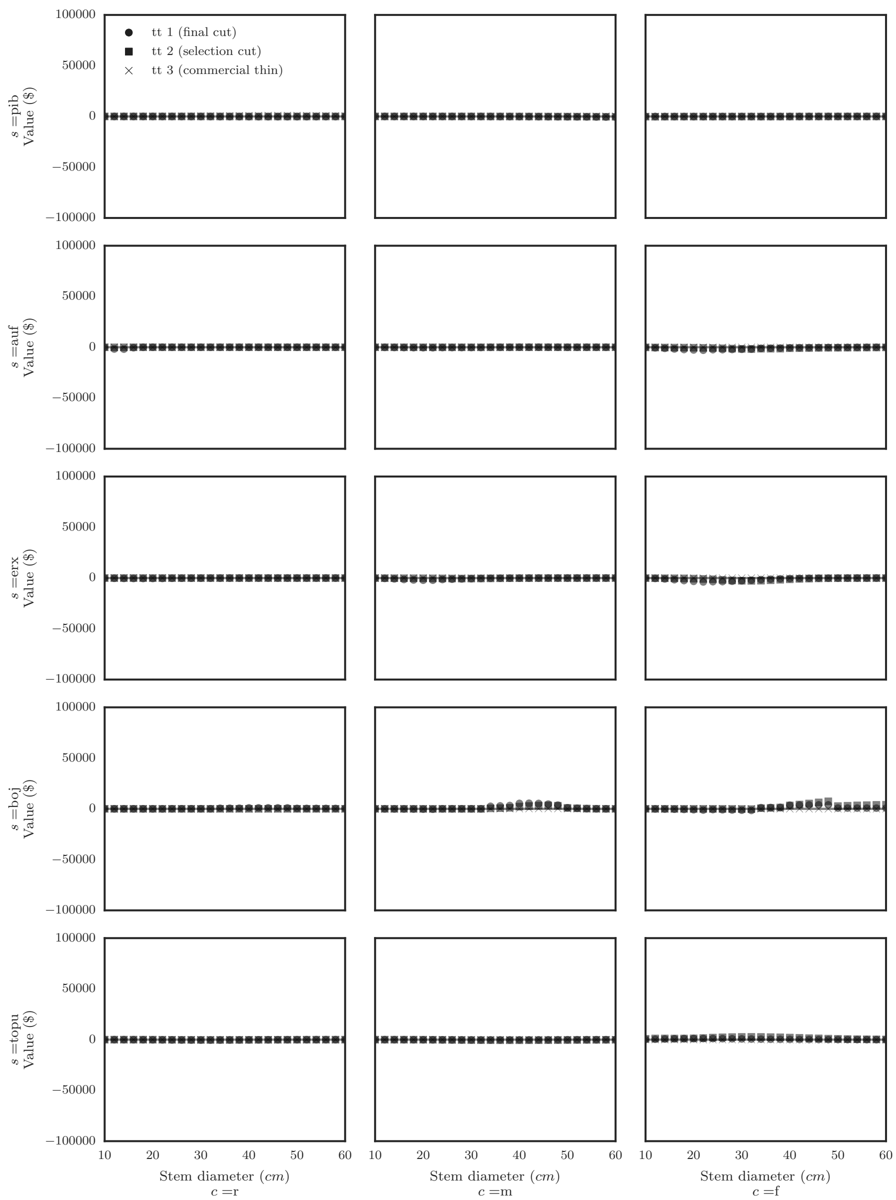

Figure 5: [Continued from Figure 4] Total value-creation–potential compiled for management unit UA 064-51 in Quebec, Canada. Species group is fixed for a given row of subfigures, and cover type is fixed forn a given column of subfigures. Treatment type 1 (circles) corresponds to clearcut harvesting, treatment type 2 (squares) corresponds to selection cut, and treatment type 3 (crosses) corresponds to commercial thinning.

Unit VCP data in the ordinate axis of all subfigures in Figures 2 and 3 is shown in the same range (i.e., -60 to +60). Similarly, total VCP data in the ordinate axis of all subfigures in Figures 4 and 5 is shown in the same range (i.e., -100000 to +100000).

Note that the data aggregation method we use to compile unit VCP results defaults to a null (0) value if no data is available to model a given combination of cover type, species group, treatment type, and stem size class. Also, some combinations of cover type and species group have little or no representation in the current inventory for our test case—total VCP value will show a null value (0) if no volume is harvested for a given combination of cover type, species group, treatment type, and stem size class.

5 Discussion

Our method simplifes the task of retro-fitting financial value performance indicators to existing wood supply models. The method works well, and was specifically designed to be applicable to all 71 mangement units that constitute the public forest of Quebec. However, the method relies on a custom-built software framework built on top of Jupyter Notebooks technology, requires a large amount of RAM to run (30 Gb or more in our tests), and requires a large and complex collection of input data to run.

Thus, application of our method, although a vast improvement over an *ad hoc* workflow, nonetheless requires advanced technical expertise to collect and assemble all the required data components and correctly connect these to the software prototype, run the model, and validate that output is reasonable at each (of many) steps in the data processing pipeline. Although we provide many helpful comments and explanations at each step in the notebook, and to document the method in a series of four technical report, applying our method remains a daunting task.

To mitigate the technical challenges associated with applying our method, we are currently working with government analysts to develop an additional software layer wrapping the existing prototype, which will further automate the process of importing and pre-processing input data, running the model, and automate quality assurance testing and reporting throughout the process. Also, the complexity of the processing requires several hours of CPU time to run the method on a single forest management unit. We are currently working on an enhanced prototype so that we can run multiple forest management units in parallel, as long as processing capacity and memory are available.

We tested our method on a real dataset, and compiled 90 distinct vectors of VCP values (one vector per combination of cover type, species group, and treatment type). As seen in Figures 2 and 3, several combinations of species group, cover type, treatment type, and stem diameter class yielded negative unit VCP values. If enough of these negatively-valued stems are present in a given stand, harvesting this stand will likely induce a financial loss. A profit-maximizing agent that is free to harvest any subset of AAC will therefore likely avoid harvesting these stands altogether. Some combinations of cover type, species group and treatment type yielded extemely negative unit VCP values (e.g., poplar

Generally, unit VSP values tend to increase as a function of stem diameter. This is consistent with the notion that larger stems contain larger proportions of veneer and sawlog product classes, which tend to have higher unit values. A step-wise pattern in unit value can be observed for most combinations of cover type, species group, and treatment type—this is an artifact of undocumented data aggregation in the MERIS database we used, and not a side-effect of our methodology.

There seems to be a pattern (for several combinations of cover type, species group, and treatment type) of step-down in unit value in the 50–60 cm stem size class range. Again, this is an artefact of financial data values we imported from the MERIS database, rather than a by-product of our method.

Unit VSP of stems harvested using a commercial thinning type treatment tends to be substantially higher than VSP of similar stems harvested using selection cut or final cut treatment types. Results presented here are derived from a complex disagregation-reaggregation process, including complex multivariate models used to approximate harvest cost and sylviculture credits, so identifying the cause of specific patterns such as this is not always easy. More analysis is required to identify the root causes of patterns such as this in VSP model output.

Note the discontinuity of unit VCP values for the *sepm* (i.e., spruce-pine-fir-larch) species group in the *r* (i.e., softwood) cover type between the 14 and 16 cm stem size classes—this discontinuity is attributable to the two-tier stumpage model for this species group and cover type. Indeed there is a sharp jump in unit stumpage price for 16-cm-and-up *sepm* stems. This discontinuity is amplified in the total VCP results, as the unimodal frequency distribution of harvested stem sizes happens to peak around the 15-cm threshold value (see Paradis and LeBel 2017a).

Figures 4 and 5 show that total VCP is null (or near null) for most combinations of cover type, species group, treatment type, and stem size class. We would not recommend spending too much time on further analysis of null or near-null VCP combination. Rather, we would recommend prioritizing further analysis and data refinement effort on dominant combinations, e.g., *sepm* (all cover types), *bop* (with particular attention to *m* and *f* cover types), *peu* (especially *m* cover type).

We would recommend that further validation and calibration efforts concentrate on auditing the data in the MERIS database (which we have not done, assuming that in this study that it could be used as-is) and testing additional sources of sample plot data for the diameter distribution modelling step (e.g., abundant temporary sample plot data is available, and might be integrated into the analysis). We recomment further validation and calibration of the method prior to deployment on all 71 management units in Quebec.

Finally, we would like to draw attention to the fact that these results are compiled using a complex harvest schedule, which represents only one of a virtually infinite number of possible harvest schedules that could have been output from the Woodstock wood supply model. Different schedules may induce very different output from our VCP estimation model. Furthermore, the Woodstock optimization model used to generate the harvest schedule does not feature any financial performance indicators (the objective function maximized even-flow harvest volume, subject to a complex set of constraints, none of which directly account for value of harvested fibre or mill demand). Our methodology could be used to compile value-driven performance indicators, which could be integrated into the Woodstock optimization model (either in the objective function or in constraints, or both), thereby potentially producing harvest solutions with improved value-creation potential.

In the next phase of this research project, we use the VCP coefficients compiled here as input to hybrid simulation-optimization model that emulates fibre-consumption behaviour of a network of profit-maximizing fibre-consuming mills, documented in Paradis and LeBel (2018).

6 Conclusion

We develop and present a method to simplify the task of retro-fitting VCP performance indicators to existing wood supply models, thereby improving the accessibility of the bilevel modelling approach described in Paradis et al. (2018). We designed and implemented a prototype software tool that automates the workflow, and test our method on real data from Quebec, Canada.

In a subsequent phase of this research project (beyond the scope of the current document), the VCP-indicator-enhanced wood supply models compiled in this phase are linked to a new network flow optimization model that will simulate profit-maximizing behaviour of a network of centrally-managed fibre-consuming mills.

Using only readily-available data, our methodology can be applied, as-is, to compile VCP indicators for any of the other 70 forest management units in Quebec. Our methodology could also adapted for use in other jurisdictions, assuming that input data is available to compile the disaggregations coefficients and to populate the MERIS database (or a similar database) with appropriate value data.

7 Acknowledgements

This study was supported by funding from the *FORAC Research Consortium*.

References

Bouchard, M., D'Amours, S., Rönnqvist, M., Azouzi, R., and Gunn, E. (2017). Integrated optimization of strategic and tactical planning decisions in forestry. *European Journal of Operational Research*, 259(3):1132–1143.

Bureau de mise en marché de bois (2014). Valeur des traitements sylvicoles commerciaux pour l'année financière 2014-2015. Technical report, Quebec.

Fortin, M. (2014). Using a segmented logistic model to predict trees to be harvested in forest growth forecasts. *Forest Systems*, 23(1):139–152.

Jerbi, W., Gaudreault, J., D'Amours, S., Nourelfath, M., Lemieux, S., Marier, P., and Bouchard, M. (2012). Optimization/simulation-based framework for the evaluation of supply chain management policies in the forest product industry. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2012)*. IEEE (accepted).

Paradis, G., Bouchard, M., LeBel, L., and D'Amours, S. (2018). A bi-level model formulation for the distributed wood supply planning problem. *Canadian Journal of Forest Research*, 48(2):160–171.

Paradis, G. and LeBel, L. (2017a). Compiling disaggregation coefficients to link long- and short-term planning models. Technical Report CIRRELT-2017-43, CIRRELT.

Paradis, G. and LeBel, L. (2017b). Diameter Distribution Models for Quebec, Canada. Technical Report CIRRELT-2017-34, CIRRELT.

Paradis, G. and LeBel, L. (2018). Estimating the value-creation potential of wood supply. Technical Report CIRRELT-2018-00, CIRRELT.

Petro, F. and Calvert, W. (1976). How to grade hardwood logs for factory lumber. Technical Report Forestry technical report no. 6, Canadian Forestry Service.

Pothier, D. and Auger, I. (2011). NATURA-2009 : un modèle de prévision de la croissance à l'échelle du peuplement pour les forêts du Québec. Technical Report (Mémoire de recherche forestière n° 163), Gouvernement du Québec, Direction de la recherche forestière.

Appendix

A1.1 Species aggregate mapping

Table A1 details mapping of species group names to species common and Latin names.

A1.2 Harvest cost model

The model used to estimate harvest cost is adapted from the function used in the MERIS system.

Table A1: Mapping of species group names to species common and Latin names. Alternate names are shown in parentheses.

Species Group	Common Name	Latin Name
Other Hardwoods	(white, American) ash	*Fraxinus americana*
	black ash	*Fraxinus nigra*
	(green, red) ash	*Fraxinus pennsylvanica*
	(North) American beech	*Fagus grandifolia*
	(American, white, water) elm	*Ulmus americana*
	slippery elm	*Ulmus rubra*
	(rock, cork) elm	*Ulmus thomasii*
	American hophornbeam	*Ostrya virginiana*
	American linden (basswood)	*Tilia americana*
White Birch	grey birch	*Betula populifolia*
	(white, paper) birch	*Beutul papyrifera*
Yellow Birch	yellow birch	*Betala alleghaniensis*
Oak-Hickory	(bitternut, swamp) hickory	*Carya cordiformis*
	shagbark hickory	*Carya ovata*
	([wild, mountain] black, rum) cherry	*Prunus serotina*
	white oak	*Quercus alba*
	swamp white oak	*Quercus bicolor*
	bur oak	*Quercus macrocarpa*
	(northern, eastern) red oak	*Quercus rubra*
	(butternut, white walnut)	*Juglans cinerea*
Spruce-Pine-Fir	white spruce	*Picea glauca*
	black spruce	*Picea mariana*
	Norway spruce	*Picea abies*
	red spruce	*Picea rubens*
	hybrid larch	*Larix X marschlinsii*
	Japanese larch	*Larix leptolepis*
	([eastern, American] larch, tamarack)	*Larix larincina*
	European larch	*Larix decidua*
	pitch pine	*Pinus rigida*
	([eastern, black] jack, grey, scrub) pine	*Pinus banksiana*
	Scots pine	*Pinus sylvestris*
	balsam fir	*Abies balsamea*
Other Maples	(silver, silverleaf) maple	*Acer saccharinum*
	black maple	*Acer nigrum*
	red maple	*Acer rubrum*
Sugar Maple	(sugar, rock) maple	*Acer saccharum*
Poplar	balsam poplar	*Populus balamifera*
	eastern cottonwood (poplar)	*Populus deltoides*
	(large-tooth, big-tooth) aspen	*Populus grandidentata*
	hybrid poplar	*Populus sp X P. sp.*
	([quaking, trembling] [aspen, poplar]	*Populus tremuloides*
Pine	white pine	*Pinus strobus*
	red pine	*Pinus resinosa*
Hemlock-Cedar	(eastern, Canadian) hemlock	*Tsuga canadensis*
	(eastern, northern) white-cedar	*Thuja occidentalis*

$$f_{HC}(p, s_1, s_2, s_3) = e^{A-(B \ln p)+(Cs_1)+(Ds_2)-(E(1-s_3))} + ((Fs_3) + (G(1 - s_3))) + K$$

where p represents piecesize (m^3 per stem). If p is a random variate, then we need to estimate the expected value of f_{HC}, given a distribution of p. The implementation of this function in the `ws3` package includes an optional switch to automate the process of estimating the expected value of the function.

s_1, s_2, and s_3 are binary switches that activate or deactivate different parts of the function, depending on the case. s_1 is set to 1 to model taking extra care during partial cutting (reduces productivity), and 0 otherwise. s_2 is set to 1 if the treatment is a final cut, and 0 otherwise. s_3 is set to 1 if partial cutting in a tolerant hardwood stand, and 0 otherwise.

Coefficents A, B, C, D, E, F, G, and K are given in Table A2.

Table A2: Coefficient values used in the harvest cost function.

Coefficient	Value
A	1.970
B	0.450
C	0.169
D	0.164
E	0.202
F	13.600
G	8.830
K	0.000

A1.3 Sylviculture credit model

Note that the sylviculture credit model used in Quebec is updated annually. The model presented here is the one currently implemented in *ws3* and used to generate the simulation results presented in this paper–it corresponds to the model published by the BMMB for the 2014-2015 period by Bureau de mise en marché de bois (2014).

The model is actually composed of seven different sub-models, corresponding to different combinations of treatment type, intensity, and cover type. For our analysis, we only retained four of these sub-models (1, 2, 4, 7). We manually map each treatment option in the wood supply model to a sub-model code, and the functions we built into `ws3` dispatch processing to the correct sub-model on the fly for each component of the simulated harvesting schedule.

Each submodel resturns a sylviculture credit value (expressed in $\$ha^{-1}$), given volume harvested (ha^{-1}) p, mean piece size of harvested stems (m^3) v_r, and mean piece size of standing inventory before harvesting (m^3) v_p.

If P, v_r and v_p are random variates, then we need to estimate the expected value of f, given a distribution of p. The implementation of this function in the `ws3` package includes an optional switch to automate the process of estimating the expected value of the function.

Submodel functions are presented below.

$$
\begin{aligned}
f_{SC1}(P, v_r, v_p) =& K_1 p(C_{1a} v_r^{C_{2a}} - e^{C_{7d} \ln v_p + C_{8d}} + C_{15h} e^{C_{16h} p} - C_{17i} p + C18j) + K2 \\
f_{SC2}(P, v_r, v_p) =& K_1 p(e^{C_{3b} \ln v_r + C_{4b}} - e^{C_{7d} \ln v_p + C_{8d}} + C_{11f} {v_r}^{-C_{12f}} - C_{13g} {v_p}^{-C_{14g}} + C_{15h} e^{C_{16h} p} \\
& - C_{17i} p + C_{18j}) + K2 \\
f_{SC4}(P, v_r, v_p) =& K_1 p(e^{C_{3b} \ln v_r + C_{4b}} - e^{C_{7d} \ln v_p + C_{8d}} + C_{11f} {v_r}^{-C_{12f}} - C_{13g} {v_p}^{-C_{14g}} + C_{15h} e^{C_{16h} p} \\
& - C_{17i} p + C_{18j}) + K2 \\
f_{SC7}(P, v_r, v_p) =& K_1 p(e^{C_{3b} \ln v_r + C_{4b}} - e^{C_{7d} \ln v_p + C_{8d}} + C_{15h} e^{C_{16h} p} - C_{17i} p + C_{18j}) + K2
\end{aligned}
$$

Coefficents values used in each submodel are given in Tables A3 through A6.

Table A3: Coefficient values used in sylviculture credit submodel 1.

Coefficient	Value
C_{1a}	4.5110
C_{2a}	-0.6280
C_{7d}	-0.3910
C_{8d}	1.9390
C_{15h}	3.9120
C_{16h}	-0.0094
C_{17i}	0.0698
C_{18j}	9.2529
K_1	1.0000
K_2	0.0000

Table A4: Coefficient values used in sylviculture credit submodel 2.

Coefficient	Value
C_{3b}	-0.2370
C_{4b}	2.5920
C_{7d}	-0.2370
C_{8d}	2.2470
C_{11f}	4.3546
C_{12f}	0.3400
C_{13g}	4.3543
C_{14g}	0.3400
C_{15h}	3.9120
C_{16h}	-0.0094
C_{17i}	0.0698
C_{18j}	7.1029
K_1	1.0000
K_2	0.0000

Table A5: Coefficient values used in sylviculture credit submodel 4.

Coefficient	Value
C_{3b}	-0.2370
C_{4b}	2.5920
C_{7d}	-0.2370
C_{8d}	2.2470
C_{11f}	4.3546
C_{12f}	0.3400
C_{13g}	4.3546
C_{14g}	0.3400
C_{15h}	3.9120
C_{16h}	-0.0069
C_{17i}	0.0517
C_{18j}	7.1029
K_1	1.0000
K_2	0.0000

Table A6: Coefficient values used in sylviculture credit submodel 7.

Coefficient	Value
C_{3b}	-0.3910
C_{4b}	2.2000
C_{7d}	-0.3910
C_{8d}	1.9390
C_{15h}	3.9120
C_{16h}	-0.0069
C_{17i}	0.0517
C_{18j}	7.1029
K_1	1.0000
K_2	0.0000