



FINAL REPORT

# Understanding Historical Old Forest Patch Size Patterns on the Edson Forest Products FMA Area in Alberta



**Final Report**

fRI Research Healthy Landscapes Program

April 30, 2021

David W. Andison, Bandaloop Landscape Ecosystem Services Ltd.

Alex M. Chubaty, FOR-CAST Research and Analytics

Eliot J.B. McIntire, Canadian Forestry Service



**fRI Research**  
Informing Land & Resource Management



## DISCLAIMER

Any opinions expressed in this report are those of the authors, and do not necessarily reflect those of the organizations for which they work or fRI Research.



## ACKNOWLEDGEMENTS

This report is part of the LandWeb project of the Healthy Landscapes Program (HLP) at fRI Research.

Many thanks to the LandWeb funding partners, including West Fraser Timber Co., Alberta Newsprint Company, Alberta Pacific Forest Industries Inc., Mercer Peace River Pulp Ltd., Vanderwell Ltd., Tolko Industries Ltd., Weyerhaeuser Company, Millar Western Forest Products, Canfor Company, Louisiana-Pacific Corporation, Mistik Management Ltd., Louisiana Pacific, Ducks Unlimited Canada, the governments of Alberta, Saskatchewan, the Northwest Territories. Thanks also to the Forest Resource Improvement Association of Alberta (FRIAA), for their generous support.

We would like to thank Ian Eddy, Tati Micheletti, Ceres Barros, and Yong Luo for useful modelling and SpaDES discussions, help describing and documenting each of the SpaDES modules, as well as their code contributions to the R packages and SpaDES modules used in this project.

We also thank CFS and Compute Canada for computing resources.

We also wish to acknowledge that the land on which this study is located is Treaty 8 territory and a traditional meeting ground and home for many Indigenous Peoples, including Dene Tha', Dene Chipewyan, Beaver and Metis Peoples.



## CONTENTS

Disclaimer .....	1
Acknowledgements .....	3
Executive Summary .....	5
1.0 Introduction .....	6
2.0 Goal .....	6
3.0 Desired Conditions and Outcomes .....	7
3.1 Indicators .....	7
3.2 Current Conditions as a Reference Point .....	7
3.3 Creating a Pre-Industrial Condition Baseline .....	8
4.0 Study Area .....	8
5.0 Methods: Choosing a Spatial Model.....	10
6.0 Methods: LandWeb and SpaDES .....	12
6.1 LandWeb Study area .....	12
6.2 SpaDES .....	12
6.3 Data Sources .....	13
6.4 Model Code.....	13
6.5 LandWeb simulation model .....	14
6.5.1 Overview .....	14
6.5.2 Data preparation.....	16
6.5.2.1 LandWeb_preamble Module .....	16
6.5.2.2 Biomass_speciesData Module.....	16
6.5.2.3 Biomass_borealDataPrep Module .....	16
6.5.2.4 Vegetation model (LandR Biomass) Module.....	19
6.5.2.5 Biomass_core Module.....	19
6.5.2.6 Biomass_regeneration Module.....	20
6.5.2.7 Fire model Module .....	20
6.5.2.8 LandWeb_output Module.....	22
6.5.2.9 timeSinceFire Module .....	22
6.5.2.10 Post-processing .....	22
6.5.3 Running the Model.....	23
6.6 Validation .....	23
6.6.1 Vegetation Dynamics .....	23
7.0 Results.....	26
8.0 Discussion .....	27
8.1 Overall Patterns .....	27
8.2 Possible Sources of Error in the Model.....	27
9.0 Recommendations.....	29
Literature Cited.....	30



## EXECUTIVE SUMMARY

This project was a spatial modelling exercise that created coarse-scale, pre-industrial landscape metrics for the Edson Forest Products FMA area in Alberta. The primary goal was to understand if, or in what ways the current conditions of the FMA area align with the historical, pre-industrial “natural” range. This report only covers the spatial results (i.e., patch sizes) of the analyses.

The results suggest that areas in old forest patches from 100 to 5000 ha on the study area are well within NRV.

Spatial patterns are far more challenging to interpret than non-spatial data. Definitions of what constitutes an “edge” are complicate further by the legitimacy of edges created by various anthropogenic activities. So while these results are good news for EFP, it would be advisable to better understand the nature and sources of how “patches” are defined with a more robust GIS exercise that includes a range of old forest edge definitions.



## 1.0 INTRODUCTION

Forest management in North America is evolving, with an inevitable move towards sustaining all forest values. Forest management is now expected to consider a wide range of biological values including water and soil conservation, toxin filtration, carbon cycling, fish and wildlife habitat, food, pharmaceuticals, and timber (Davis 1993).

Under the auspices of this evolution, the concept of the using (pre-industrial) forest patterns created by natural processes as management guides is gaining favour in North America (Franklin 1993), and is one of the foundations of an ecosystem-based management (EBM) approach (Booth et al. 1992, Grumbine 1994, Long 2009). The theory is attractive — by maintaining the type, frequency, and pattern of change on a given landscape, we are more likely to sustain historical levels of the various biological goods and services. So-called “coarse-filter” knowledge can also be applied directly to planning and management programs at all levels and scales. Thus, defining the historical range of various ecosystem patterns is a fundamental requirement of a natural pattern-based approach to forest management.

Developing coarse-filter, pre-industrial knowledge is perhaps most challenging at landscape scales. Reliable, pre-industrial landscape snapshots are rare to non-existent due to the combined impacts of fire control, cultural disturbance activities, and lack of historical records or data. What we do know about the disturbance history of Canadian boreal landscapes suggests that they are highly dynamic in time (Turner and Dale 1991, Payette 1993) and space (Andison and McCleary 2014). This means that historical levels of old forest are also likely to be both highly dynamic and spatially variable.

In the absence of detailed and repeated historical data and/or photos, the only means left to explore the dynamics of forest ecosystem patterns at the landscape scale is spatial simulation modelling. In its simplest form, spatial models allow us to explore how known (observed, recorded) probabilities of key variables intersect in time and space to create multiple possible landscape scenes or snapshots. When a sufficient number of landscape snapshots have been created by the model, each one is measured in a number of ways to capture the desired metrics, and then summarized to generate NRV.

This report describes a modelling process by which we generated multiple possible historical landscape scenes, summarized their patterns, and compared those to the current landscape condition for the Edson Forest Products FMA area managed by West Fraser Mills. The larger modelling project is LandWeb; *Landscape dynamics of Western Boreal Canada*. Note that this report only includes spatial modelling results. Please see Andison (2015) for the non-spatial NRV results.

## 2.0 GOAL

***D.W. Andison***

The goal of the LandWeb project is: ***to understand pre-industrial landscape-scale patch size patterns in the western boreal forest relative to the current condition***. Note that this goal is both narrow (it will



capture only landscape scale patterns) and humble (it will capture only a small number of simple metrics). This report includes the results for the Edson Forest Products FMA area.

## 3.0 DESIRED CONDITIONS AND OUTCOMES

*D.W. Andison*

### 3.1 INDICATORS

LandWeb project partners collectively identified two main classes of output/indicators as part of this project; 1) the area in each seral-stage × major vegetation types, and, 2) patch sizes of old forest × major vegetation types. Through a consultation process as part of this project, the LandWeb partners agreed to the following technical protocols:

- **Major vegetation types** were defined by polygons with at least 80% leading species of black spruce, white spruce, pine, deciduous, or fir (*Abies* spp.). All other forested areas that did not meet the 80% rule were classified as mixedwood.
- **Seral stages** were defined by the government of Alberta (GoA) provincial standard, and agreed to by everyone: young (<40 years), immature (40–80 years), mature (81–120 years), and old (>120 years).

In terms of old forest (i.e., >120 year old) patch sizes, the LandWeb partners also agreed that this project should report on the following patch sizes; >100 ha, >500 ha, >1000 ha, and >5000 ha. Patches should be reported by all forest types combined.

The LandWeb partners also asked to have NRV results summarized within several different geographic boundaries including a) jurisdiction (including the Edson Forest Products FMA area), b) ecological natural sub-regions (NSRs), and any c) existing caribou habitat range areas.

### 3.2 CURRENT CONDITIONS AS A REFERENCE POINT

The relevance of NRV modelling output is increased significantly when it is compared to the current condition since this provides a relevant reference point in time. These data must be provided in exactly the same format, using exactly the same rules as defined above.

In theory, current condition data exist in the form of inventories and updates. However, for the purposes of this project, the most recent data are notoriously challenging and time-consuming to a) acquire and b) summarize in a universal format. This is only magnified by the fact that the study area includes five different provincial / territorial jurisdictions, 15 different forest management areas, multiple provincial and federal parks, and provincially-managed areas. Moreover, the vintage of the most recent updates varies considerably across the study area. Acquiring and compiling these spatial data from scratch would have exceeded the entire budget of this project.

Instead, we took advantage of an existing initiative to compile forest inventory data from across Canada. The CASFRI (Common Attribute Schema for Forest Resource Inventories) is the first and only known



initiative to collect and standardize inventory data from multiple jurisdictions across Canada (Cosco 2011). Although this database was not 100% complete, and some of the data were outdated, it still saved us considerable time and costs. We acquired outstanding data directly from partners.

### 3.3 CREATING A PRE-INDUSTRIAL CONDITION BASELINE

Given that the goal of the modelling is to construct the NRV, the spatial data involved need to be free of all industrial human influence, including permanent and semi-permanent land use changes (e.g., infrastructure, agriculture), harvesting, and fire control. This can be done in two ways. Some NRV modelling exercises start with an existing landscape — complete with anthropogenic influences — and run the model forward hundreds to thousands of years to *fill in* the areas influenced by human activity. Alternatively, it is possible to re-create a single natural vegetation conditions on a single landscape scene *via* a GIS exercise that uses the following, hierarchical, rules: 1) historical (pre-disturbance) vegetation information in digital format, 2) historical (pre-disturbance) vegetation information from available maps, 3) rules and/or an algorithm that calculates the most likely vegetation type of missing polygons based on neighbours. For this project, we chose to go with the second option.

To create an initial, pre-industrial landscape, we first obtained the oldest digital version of the forest inventory (with the least amount of cultural disturbance). Then we used digital data, records, and maps to replace cultural features with pre-disturbed vegetation types. Any remaining culturally modified polygons were filled with the age and cover-type attributes of the adjacent polygon with the greatest length shared boundary. Thus, all towns, roads, cut blocks, mines, and other human developments were replaced by attributes of the last known or the most likely forest type. The “natural” pre-industrial snapshot created by this process still included biases and inaccuracies from a) fire control b) using data from different eras, and/or c) aging errors from forest inventories, all of which could influence the subsequent model output for centuries. To eliminate this risk, the model was run forward several thousands of years before landscape snapshots were collected and measured.

## 4.0 STUDY AREA

### ***D.W. Andison***

The area of concern for this report was the Edson Forest Products FMA area, covering a total of almost 266,000 ha in two pieces (Figure 1). The larger southern portion accounts for 226,000 ha, or 85% of the study area and the smaller northern piece the other 15% (Figure 1).

*Figure 1. Study area map showing the Edson Forest Products FMA*







Ecologically, most of the study area is in the Upper Foothills NSR (53%), followed closely by the Lower Foothills (43%). The other 4% includes relatively small areas of the Subalpine and Alpine NSRs (Table 1).

Table 1. Summary of Edson Forest Products FMA area by Natural Subregion (NSR)

NSR Name	FMA Area	
	Hectares	%
Alpine	434	0.2
Lower Foothills	114,284	43
Subalpine	9,106	3
Upper Foothills	141,948	53
<b>TOTAL</b>	<b>265,772</b>	<b>100</b>

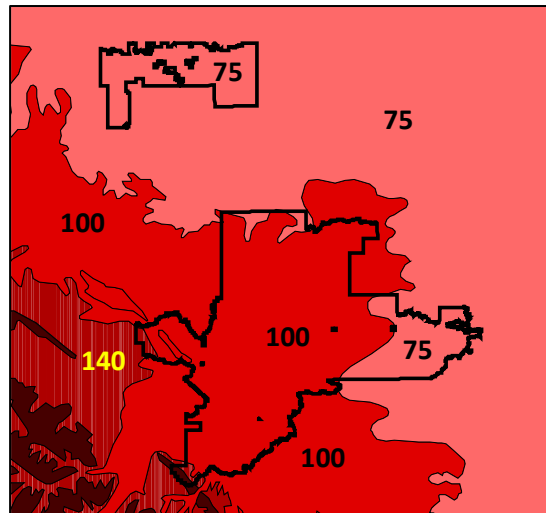
Ecological conditions vary significantly across the study area relative to its small size. Elevation ranges from 650 to 1750 m, topography from gently rolling to steep, and vegetation from dense mesic mixedwood forest to dense pure pine (Table 2).

Table 2. Summary of biotic and abiotic conditions across the study area.

Natural Region	Natural Subregion	Elevation	Topography	Climate	Vegetation	Soils	Growing Degree Days >5°C	Mean annual Precip (mm)	Relative Summer Moisture Index
Rocky Mountain	Subalpine	1300-2300m	Rolling to very steep	Very short cool wet summers, long snowy winters. Highly variable microclimate	Closed PI forest (low el) opening to mixed Se, L, and Abies forest & krummholz (high el). Wetlands and open water uncommon	Brunisols, with some regosols and non-soil	800	760	1.7
Foothills	Upper Foothills	950-1750m	Rolling to steeply sloped	Short wet summers, snowy cool winters	Dense PI forest (low el) to dense Sb, Sw forest (high el). Small area in wetlands.	Luvisols, with some brunisols	900	650	2
	Lower Foothills	650-1625m	Gently rolling with plateaus	Short summers with average precip, colder very snowy winters	Highly variable. Mostly mesic dense mixedwood forest (At, Pl, Sw, Pb, Ta, Fir, shrubs). Very little water or wetlands	Luvisols, with some brunisols	1,100	590	2.7

Estimates of the average pre-historical long-term fire cycle (LTFC) ranges from 75–100 years across the study area as one moves from east to west (Figure 2).

Figure 2. Long-term-fire-cycles for the study area (From Andison 2019).





## 5.0 METHODS: CHOOSING A SPATIAL MODEL

*D.W. Andison*

By definition, models are simple, incomplete representations of reality (Hammah and Curran 2009). There is also a key trade-off between complex models and simple ones. The “best” model is not necessarily the most complex or realistic one, but rather the model that best suits the purpose. The rule of parsimony for any modelling exercise *is as complex as necessary, but no more*. In other words, each modelling exercise should focus on achieving the desired objectives with the least possible number of explanations, equations, and assumptions (Hammah and Curran 2009). In this case, modelling objectives were very simple and general in nature; to define the natural, pre-industrial range of a) seral-stage levels and b) patch sizes by broad vegetation types, and broad geographic zones. This requires a model with the following attributes:

1. Fully spatial,
2. Fully stochastic,
3. Able to function at multiple scales,
4. Very good at capturing known fire patterns,
5. Able to accurately capture /represent known disturbance regime parameters (mostly frequency, size, and severity),
6. Able to generate results in a timely manner, and
7. Work at massive spatial scales (i.e., over 100 million ha).

These requirements were quite restrictive, and narrowed our options considerably since it meant the model must be a) raster-based at a scale of no larger than 10 ha, b) able to function across multiple fire regimes, c) able to handle and integrate multiple spatial data sources, and d) highly efficient in terms of language, memory and processing capacity.

At the outset of this project, there was no existing model that met all of these requirements. However, several were close enough that they could have been adapted with some effort (i.e., Landis, Bfolds, Landmine, Alces, and SELES). As part of the process for this project, the pros and cons of each model were researched and summarized, the likely costs associated with adapting each to suit the new parameters calculated with the help of local experts, and the risks of each not achieving the desired outcomes and objectives identified (e.g., what were the chances that scaling up model X to 100 million ha and adding component Y would even run on a computer, let alone produce output in a timely manner?). The cost and time estimates to upgrade any of the existing model options were considerable.

Another option presented itself at the same time. A CFS-Laval academic partnership (Drs. McIntire and Cumming respectively) were fleshing out the architecture of, and starting to write code for, an ensemble modelling framework called SpaDES (Spatially Discreet Event Simulator). Ensemble models are not models *per se*, but rather frameworks within which multiple models, and/or model components (i.e., modules) can interact (Krueger et al. 2012). In this case, the idea was to create a universal scheduling environment in R that would allow model modules (even ones from existing models) to communicate



and be interchangeable. For example, in Figure 3, there are four different spatial data modules, two fire spread modules, and three forest succession modules to choose from (see below).

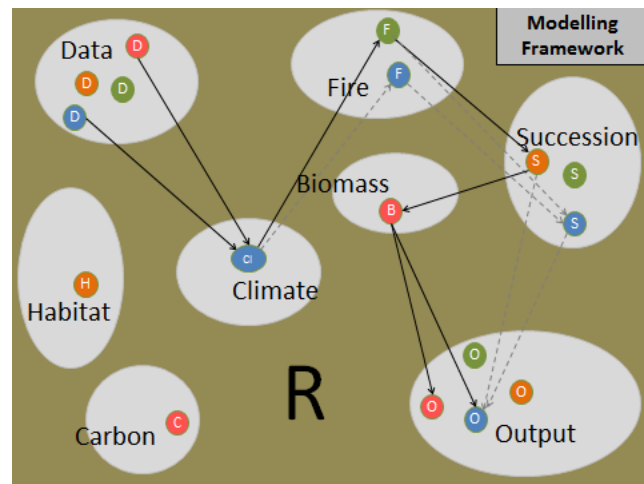
Thus, the alternative to investing in upgrading an existing model was to invest in the development of a new, potentially far more powerful modelling framework that is SpaDES, within which a specific module configuration would be developed to achieve the goals of this project.

There were several benefits of going with the SpaDES option. First, by design, the final product would be open source. This means the final product can be used, modified and shared openly and free of charge to anyone — as opposed to proprietary software, which is not only unavailable for independent review, but must be purchased. Second, because LandWeb would be associated with a larger, open source product it also creates a legacy. LandWeb partners are thus able to use the model for future, and different research and forecasting needs, as opposed to a one-off static model. Thus, the investment in the objectives of LandWeb could result in payoffs in terms of access to, and use of, a universal spatial model for multiple purposes. Third, the plan for LandWeb in SpaDES was to create a stand-alone app available (free of charge) online to anyone. Finally, the various modules necessary to fulfill the objectives of this project would be adapted from existing, proven models, as opposed to writing new modules from scratch.

The greatest risk of going with the SpaDES option was the unknown amount of time and effort required to not only design, build, test, and validate a new modelling framework, but to be the first to attempt to build a specific configuration and app within that framework. Writing, validating and error-checking code is notoriously challenging and time-consuming, and in this case there was no shortage of technical challenges to potentially overcome. So, although the original time and cost estimates from the modelling team were well within the timelines of the project, the resources to complete a LandWeb configuration within SpaDES could well have been significantly greater than we had. In the worst case scenario, resources would be depleted before the end of the project, and with no results to show for the effort. On the other hand, this same risk also existed for the existing model upgrade option. For example, model architecture aside, the sheer effort required to acquire, compile, validate, overlay, and access the massive spatial databases required is without precedent.

In the end, the HL Program Lead chose to support the work of the SpaDES modelling team to develop a needs-specific, LandWeb configuration.

*Figure 3. The SpaDES environment (brown shaded area) allows various modules to communicate and even be exchanged for other, parallel modules. The black lines represent one possible configuration of modules — out of dozens.*





## 6.0 METHODS: LANDWEB AND SPADES

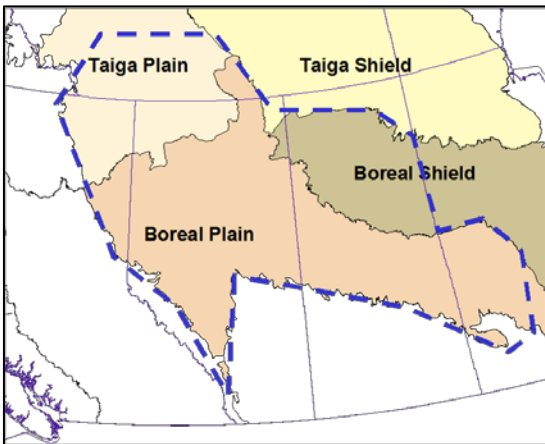
*A.M. Chubaty and E.J.B. McIntire*

### 6.1 LANDWEB STUDY AREA

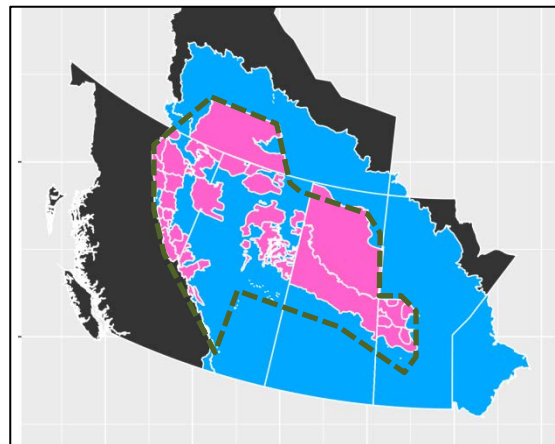
The study area for LandWeb includes the western-most 125 million ha of the Canadian boreal forest extending west from the Rocky Mountains to beyond the Manitoba border to the east, and from the southern boundary of the forest-grassland interface approximately to the 62<sup>nd</sup> parallel into the Northwest Territory. The area includes 73 million ha of Boreal Plain, 25 million ha of Taiga Plain, 20 million ha of Boreal Shield, and 7 million ha of transitional areas of Prairie, Montane Cordillera, Taiga Shield and Boreal Cordillera (Wilken 1986) (Figure 4).

The study area also includes several woodland caribou ranges (Figure 5). Note also that the area that was modelled extends well beyond the boundary of the study area. This is to avoid bias associated with edge effects, and common practice for spatial modelling (Figure 5).

*Figure 4. Map of the LandWeb study area ecozones.*



*Figure 5. Map of the LandWeb Study Area showing the modelling area (blue) and current caribou range (pink).*



### 6.2 SPADES

SpaDES is collection of packages for the R Statistical and Data Language used to develop and run spatially explicit simulation model (Chubaty and McIntire 2018; 2019a; McIntire and Chubaty 2019; Chubaty and McIntire 2019b). There are three key features of the SpaDES platform that make it an excellent choice for the implementation of the LandWeb model. The first is that SpaDES leverages the availability in R of a vast number of robust scientific computing and data visualization packages. Second, using R for data preparation, analysis, and simulation, provides a streamlined data-model pipeline and workflow. Finally, SpaDES is built with the explicit notion of model components that are interchangeable



and easily updatable (i.e., modular). In this sense, SpaDES simply schedules and run various model components (i.e., modules).

Although individual modules are designed to be standalone units, their design includes several features that facilitate use with other modules (i.e., module integration). Each module includes metadata that define its parameter values, as well as data inputs and outputs. These data dependencies are used by SpaDES to calculate module interconnectedness *via* the data objects shared among modules. The specific collection of modules (with their parameterizations and data dependencies) used by LandWeb (i.e., configuration) incorporate and build on models developed for and reusable in other research contexts. We describe each module used in LandWeb simulations in more detail below.

### 6.3 DATA SOURCES

Data used for the model are derived from multiple sources, and include both open (and freely available) data as well as proprietary partner-supplied data. Data sources for each module are identified in the module descriptions below (Table 3).

*Table 3. Summary of spatial data sources used*

Data product	Source URL
<a href="#">Pickell land cover and forest inventory data (Pickell and Coops 2016)</a>	N/A
<a href="#">"kNN data" (Beaudoin et al. 2014)</a>	<a href="http://tree.pfc.forestry.ca/">http://tree.pfc.forestry.ca/</a>
<a href="#">LCC2005 v1.4 (Latifovic and Pouliot 2005)</a>	<a href="ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCCLandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip">ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCCLandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip</a>
Forest Resource Inventory (LandWeb partners, prepared by Silvacom)	N/A
CASFRI v4 (2016) (described in Cosco 2011)	N/A

### 6.4 MODEL CODE

All modules are written in R and all model code was developed collaboratively using GitHub (<https://github.com>), with each module contained in its own (private) git repository (Table 4). Code that is shared among modules was bundled into R packages, and hosted in open git repositories. All package code is automatically and regularly tested using cross-platform continuous integration frameworks to ensure the code is reliable and free of errors.



Table 4. Module and package code repositories used for the LandWeb project. Module code repositories are currently private; package code repositories are open.

Code Repository	Description	URL
<b>Modules</b>		
LandMine	A reimplementation of Andison's fire model, simulating fire ignition and spread.	<a href="https://github.com/PredictiveEcology/LandMine">https://github.com/PredictiveEcology/LandMine</a>
LandR Biomass_speciesData	Prepares species input layers from multiple data sources.	<a href="https://github.com/PredictiveEcology/Biomass_speciesData">https://github.com/PredictiveEcology/Biomass_speciesData</a>
LandR Biomass_core	Simulates vegetation growth, mortality, aging, and dispersal. Updates biomass following other modules' events, and produces summary figures and tables.	<a href="https://github.com/PredictiveEcology/Biomass_core">https://github.com/PredictiveEcology/Biomass_core</a>
LandR Biomass_regeneration	Simulates post-disturbance (e.g. fire) biomass regeneration.	<a href="https://github.com/PredictiveEcology/Biomass_regeneration">https://github.com/PredictiveEcology/Biomass_regeneration</a>
LandR Biomass_borealDataPrep	Prepares multiple data objects used by Biomass_core; customized for Canadian Boreal Forests.	<a href="https://github.com/PredictiveEcology/eliotmcintire/Biomass_borealDataPrep">https://github.com/PredictiveEcology/eliotmcintire/Biomass_borealDataPrep</a>
LandWeb_output	Summarizes and prepares model outputs specifically for the LandWeb project.	<a href="https://github.com/fRI-Research/LandWeb_output">https://github.com/fRI-Research/LandWeb_output</a>
LandWeb_preamble	Creates study areas, including all FMA polygons, and prepares inputs for the main LandWeb simulation.	<a href="https://github.com/fRI-Research/LandWeb_preamble">https://github.com/fRI-Research/LandWeb_preamble</a>
timeSinceFire	Keeps track of forest pixel ages during the simulation.	<a href="https://github.com/fRI-Research/timeSinceFire">https://github.com/fRI-Research/timeSinceFire</a>
<b>Packages</b>		
LandR	Landscape Ecosystem Modelling in R	<a href="https://github.com/PredictiveEcology/LandR">https://github.com/PredictiveEcology/LandR</a>
LandWebUtils	Additional utilities for LandWeb analyses	<a href="https://github.com/PredictiveEcology/LandWebUtils">https://github.com/PredictiveEcology/LandWebUtils</a>
map	Defines a meta class of geographical objects, the 'map' class, which is a collection of map objects (sp, raster, sf), with a number of metadata additions to enable powerful methods (e.g., for leaflet, reproducible GIS, etc.)	<a href="https://github.com/PredictiveEcology/map">https://github.com/PredictiveEcology/map</a>
pemisc	Miscellaneous utilities developed by the Predictive Ecology Lab Group	<a href="https://github.com/PredictiveEcology/pemisc">https://github.com/PredictiveEcology/pemisc</a>

## 6.5 LANDWEB SIMULATION MODEL

### 6.5.1 OVERVIEW

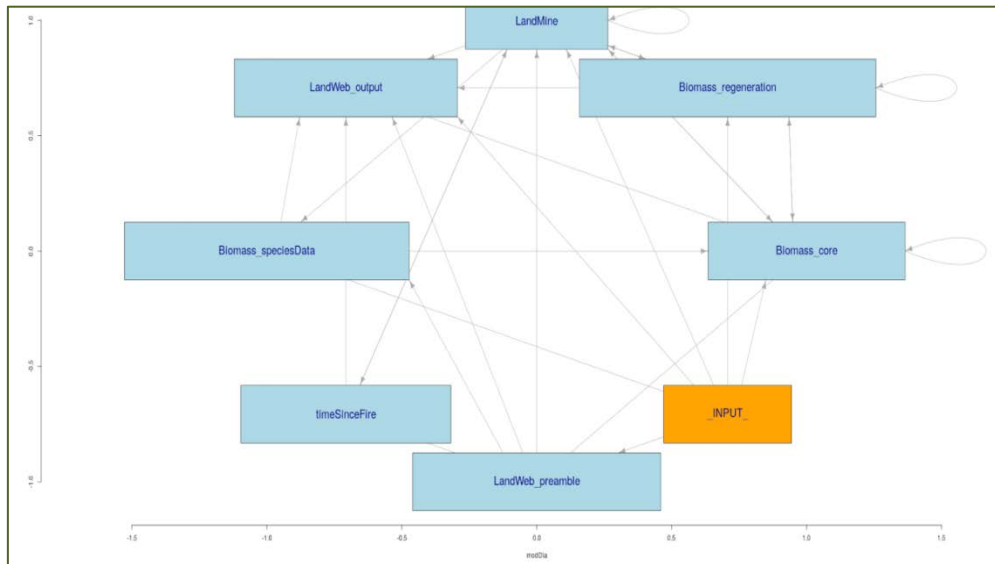
To our knowledge, LandWeb is the first large scale, data-driven approach to simulating historical NRV. In developing the model, analyses, as well as the infrastructure to host data, we strived to implement a single, reproducible workflow to facilitate running simulations, analyses, model reuse and future expansion. This tight linkage between data and simulation model was made possible *via* its implementation using the SpaDES<sup>1</sup> family of packages (Chubaty and McIntire 2018; 2019a; 2019b) within the R Statistical Language and Environment (R Core Team 2018). SpaDES facilitates the development of large-scale spatial simulation models.

<sup>1</sup> Packages used includes, SpaDES, SpaDES.core, SpaDES.tools, reproducible, quickPlot, LandR, LandWebUtils, amc, pemisc, map, raster, sp, sf, and data.table



The LandWeb model integrates two well-used models for forest stand succession and fire simulation, implemented in the SpaDES simulation platform as a collection of sub-models implemented as SpaDES modules. Each of these modules are generally categorized by their primary purpose, summarized in Figure 6 and are further described below.

Figure 6. Schematic diagram of the modules within the LandWeb model.



**Data preparation.** Simulations were run for the entire LandWeb study area, which spans most of the western Canadian boreal forest. Input data were derived from several publicly available, remote-sensed datasets (Beaudoin et al. 2014), as well as proprietary data compiled by Pickell and Coops (2016).

**Vegetation dynamics** were modeled using a re-implementation of the LANDIS-II Biomass model, a widely used and well-documented dynamic vegetation succession model (Scheller et al. 2007; Scheller and Mladenoff 2004; 2007). Our re-implemented model largely follows the original LANDIS-II source code (v 3.6.2), but with some modifications.

**Fire dynamics** were modeled using a re-implementation of the fire sub-model of Anderson’s (1996; 1998) Landmine model of landscape disturbance.

**Summary maps and statistics** were produced/calculated from simulation outputs, and consist of maps showing the time since fire as well as histogram summaries of 1) number and/or total area of large patches (i.e., patches above the number of hectares specified by the user) contained within the selected spatial area; and 2) the vegetation cover within the selected spatial area. Histograms are provided for each spatial area by polygon, age class, and species. Authorized users can additionally overlay current stand conditions onto these histograms. Simulation outputs were summarized for several publicly available reporting polygons (including Alberta Natural Ecoregions and woodland caribou ranges).



## 6.5.2 DATA PREPARATION

The following describe the modules used for LandWeb.

### 6.5.2.1 LANDWEB\_PREAMBLE MODULE

This module performs several GIS data preparation steps to 1) define the study area for LandWeb, and 2) to ensure that all downstream geospatial objects are converted to use the same geospatial geometries (e.g., projection, extent, resolution). Furthermore, this module implements several automated methods for ensuring the validity and the compatibility of input data layers with the downstream simulation components. In particular, it removes non-tree pixels from the Land Cover Classification 2005 and Forest Resource Inventory data sets, and overlays these inventory data into individual forest inventory (by species) and land cover layers (Table 5).

The module defaults to processing cover data for five species/genera: fir (*Abies spp.*), white spruce (*Picea glauca*), black spruce (*Picea mariana*), pine (*Pinus spp.*), and trembling aspen (*Populus tremuloides*).

Table 5. Data sources used by LandWeb\_preamble module

Forest Cover Layer(s)	Source URL
<a href="#">Pickell land cover and forest inventory data (Pickell and Coops 2016)</a>	N/A
<a href="#">“kNN data” (Beaudoin et al. 2014)</a>	<a href="http://tree.pfc.forestry.ca/">http://tree.pfc.forestry.ca/</a>
<a href="#">LCC2005 v1.4 (Latifovic and Pouliot 2005)</a>	<a href="ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCC/LandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip">ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCC/LandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip</a>
Forest Resource Inventory and Land Cover data (LandWeb partners, prepared by a.k.a. “Current Conditions” data)	N/A
<a href="#">CASFR1 v4 (2016); described in (Cosco 2011)</a>	N/A

### 6.5.2.2 BIOMASS\_SPECIESDATA MODULE

This module downloads and extracts several species cover data layers (Table 5) and overlays them to produce single cover layers by species. It also performs several data pre-processing steps to ensure 1) all data use the same geospatial geometries, 2) are cropped to the study area, and 3) attempts to correct or fill-in any inconsistent or missing data based on the data from the other layers. The details of how the layers used in this module were initially developed are reported in their respective reports and publications cited above (Table 5).

As above, this module defaults to processing cover data for five species/genera: fir (*Abies spp.*), white spruce (*Picea glauca*), black spruce (*Picea mariana*), pine (*Pinus spp.*), and trembling aspen (*Populus tremuloides*).

### 6.5.2.3 BIOMASS\_BOREALDATAPREP MODULE

This module converted open datasets that were available for all of Canada's forests into the input requirements for Biomass\_core, a forest landscape succession model derived from the Landis-II Biomass Succession Model (Scheller et al. 2007; Scheller and Mladenoff 2004). It was primarily used to estimate





vegetation growth parameters including maximum biomass, maximum aboveground net primary productivity (aNPP), and seedling establishment probability, and to simulate the tree cohorts necessary for Biomass core. This module also provided other parameters, such as species tolerances to shade, and other plant traits (e.g., longevity, ability to re-sprout, etc.). These traits are the same as those derived from LANDIS-II, though the specific values used in the LandWeb simulations were 1) selected to produce relative species abundances that resemble the initial conditions data (Table 6); and 2) others were determined using linear mixed effects models fit to the LandWeb study area (described below).

The module makes use of many datasets from the National Forest Inventory, including aboveground biomass, stand age, and species cover, (Beaudoin et al. 2014) as well as the 2005 National Land Cover of Canada (Latifovic and Pouliot 2005), and the Ecological Land Classification of Canada (LCC) (Statistics Canada 2018) (Table 7).

Table 6. Species traits values modified from LANDIS-II for LandWeb.

Species	Abie_sp	Pice_gla	Pice_mar	Pinu_sp	Popu_sp
Area	BSW	BP	BP	BP	BP
longevity	200	400	250	150	140
sexualmature	20	30	30	15	20
shadetolerance	3	2	3	1	1
firetolerance	1	2	2	2	1
seeddistance_eff	250	100	320	300	500
seeddistance_max	1250	1250	1250	3000	3000
resproutprob	1	1	1	1	1
resproutage_min	0	0	0	0	0
resproutage_max	400	400	400	400	400
postfireregen	resprout	resprout	resprout	resprout	resprout
leaflongevity	2	3	3	2	1
wooddecayrate	0.02	0.02	0.02	0.01	0.07
mortalityshape	15	15	15	15	25
growthcurve	0	1	1	0	0
leafLignin	0.2	0.2	0.2	0.2	0.1
hardsoft	soft	soft	soft	soft	hard

Table 7. Data sources used by Biomass borealDataPrep module.

Data Source	URL
<a href="#">Land cover and forest inventory data (Pickell and Coops 2016)</a>	N/A
"kNN data" (Beaudoin et al. 2014)	<a href="http://tree.pfc.forestry.ca/">http://tree.pfc.forestry.ca/</a>
<a href="#">LCC2005 v1.4 (Latifovic and Pouliot 2005)</a>	<a href="ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCC/LandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip">ftp://ftp.ccrs.nrcan.gc.ca/ad/NLCC/LandCover/LandcoverCanada2005_250m/LandCoverOfCanada2005_V1_4.zip</a>
Forest Resource Inventory and Land Cover data (LandWeb partners, prepared by Silvacom; 2016) a.k.a. "Current Conditions" data	N/A
<a href="#">CASFRI v4 (2016); described in (Cosco 2011)</a>	N/A
Initial communities (Landis-II)	<a href="https://github.com/LANDIS-II-Foundation/Extensions-Succession-Archive/master/biomass-succession-archive/trunk/tests/v6.0-2.0/">https://github.com/LANDIS-II-Foundation/Extensions-Succession-Archive/master/biomass-succession-archive/trunk/tests/v6.0-2.0/</a>
Species traits (Landis-II)	<a href="https://github.com/dcyr/LANDIS-II_IA_generalUseFiles">https://github.com/dcyr/LANDIS-II_IA_generalUseFiles</a>

A number of data cleaning operations were used to treat pixels with problematic sample sizes and logical inconsistencies. First, land cover classes (LCC) corresponding to recent burns, old burns, and cities were reclassified by searching the focal neighbourhood and using adjacent cover classes. These pixels were omitted from the subsequent fitting of statistical models, but were assigned predicted values from these models. Other situations arose where cover was 10% but biomass was zero, or biomass was 25 tons/ha but age was zero.



In these instances, tree species occupying fewer than 5 pixels (< 1 ha) were removed. Both age and biomass required fidelity to species cover, since cover was presumed to be the most accurately estimated variable. Species-specific above-ground biomass (AGB) was estimated for each tree species present in a given pixel by multiplying the relative cover of the tree by the total AGB of the pixel (this method assumed all tree species had identical cover/biomass relationships). Stand age also had to be corrected with respect to species longevity parameters. This was achieved by fitting a statistical model relating “correct” age observations (i.e., those already corrected for zero cover and with age estimates not exceeding longevity) against the interaction of observed biomass (totalB), species (speciesCode) and percent cover (cover), accounting for the random effect of combination of ecoregion and LCC (ecoregionCode):

$$age \sim totalB * speciesCode + cover + (1 | ecoregionCode) \text{ [Eq. 1]}$$

$$R^2 \text{ marginal} = 0.38, R^2 \text{ conditional} = 0.45$$

Predicted ages were subsequently bounded to zero on the lower limit. Parameters maxB and aNPP were then estimated from a linear mixed effects model reflecting the response of species-specific biomass (B) to the interaction between age (on the log scale, logAge) and species and % cover and species, accounting for the random effect of ecoregionGroup on the calculated slopes (per species) and intercepts:

$$B \sim logAge * speciesCode + cover * speciesCode + (logAge + cover + speciesCode | ecoregionGroup) \text{ [Eq. 2]}$$

The maximum aNPP was derived from the formula maximum aNPP = maximum AGB / 30, similar to LANDIS-II. Estimates of Species Establishment Priority were based on a generalized linear mixed effects model relating percent cover and species, accounting for the random effect of ecoregionGroup on the intercepts. In this case, species percent cover was treated as the number of times a species was observed (no. of pixels with cover > 0) per ecoregionGroup, thus following a binomial distribution that was accounted for in the model with a logit link function:

$$logit(\pi) \sim speciesCode + (1 | ecoregionGroup) \text{ [Eq. 3]}$$

where  $\pi$  is the probability of finding a species (cover > 0) in an ecoregionGroup, or, the proportion of pixels that it occupied.

For both models, coefficients were estimated by maximum likelihood and model fit was calculated as the proportion of explained variance explained by fixed effects only (marginal  $R^2$ ) and by the entire model (conditional  $R^2$ ). For the biomass model (Eq. 2), marginal and conditional  $R^2$  were 0.52 and 0.79, respectively; for the percent cover model (Eq. 3), they were 0.07 and 0.13. To estimate maxB we predicted biomass for unique combinations of species and ecoregion code assuming maximum age (i.e., longevity) and maximum cover (100%).



Parameters for the 'Recent burn' and 'Urban' LCC were input from the ecodistrict and LCC of neighbouring pixels using a focal window that iteratively expanded until a valid ecodistrict/LCC was returned.

One of the advantages of this module (and of using SpaDES/R more generally), is that the parameters used for the vegetation succession modules could also be directly estimated from data within the context of the simulation. This is achieved "automatically" should the data or study area change. As with any model, this means that model predictions need to be calibrated every time the study area changes.

#### 6.5.2.4 VEGETATION MODEL (LANDR BIOMASS) MODULE

LandR Biomass is a dynamic landscape vegetation model. As such, it simulated landscape-scale forest dynamics in a spatio-temporally explicit manner, using cohorts of tree species within each pixel. Multiple ecological processes were captured by the model, including vegetation growth, mortality, seed dispersal, and post-disturbance regeneration. These dynamics followed those of the LANDIS-II Biomass Succession module v3.2.1 (Scheller and Mladenoff 2004; Scheller and Miranda 2015), but were modified to improve general utility and computational performance (Barros et al. in prep). In brief, the LandR modules reproduced forest biomass dynamics in a spatially explicit manner at the landscape scale. They simulated biomass changes by cohort (species-age combinations) as a function of age, between-cohort competition for light resources, seed dispersal, germination, and regeneration following a disturbance, and background or fire-related mortality.

#### 6.5.2.5 BIOMASS\_CORE MODULE

This module provided the core vegetation dynamics, simulating vegetation growth and mortality processes. The functions that determine growth and mortality were unchanged from LANDIS-II. Growth and mortality dynamics were simulated in units of biomass ( $\text{g}/\text{m}^2$ ) for each cohort within a stand at an annual time step, regardless of the successional time step used for other processes, such as dispersal or regeneration. Growth was dependent upon the maximum annual primary productivity of a species, cohort age, and competition. Species-specific growth curves dictated the maximum growth for a cohort as it aged. Young cohorts had lower maximum growth, as small trees were not as productive as large, mature trees. Competition acted to reduce growth by limiting the available growing space, while recent disturbances (i.e., from the previous year) increased the available growing space. Competition occurred when a stand contained more than one species-age cohort.

Mortality was derived from two sources, senescence (age-related mortality) and development-related mortality due to the ongoing loss of individual trees and branches from a cohort (Scheller and Mladenoff, 2004). Mortality was dependent upon the living biomass of a cohort, while development-related mortality could not exceed aNPP. As cohorts near their longevity age, age-related mortality increased exponentially, eventually reaching the entirety of the cohort's biomass at the maximum lifespan of the cohort species. Age-related mortality was determined by pre-defined mortality curves that vary by species.



### 6.5.2.6 BIOMASS\_REGENERATION MODULE

This module simulated post-disturbance (in this case fire) regeneration, assuming fires were stand-replacing. In each burnt pixel, the module reset pixel biomass to zero and activated post-fire re-sprouting and/or serotiny depending on species' abilities to re-sprout, their seed establishment probabilities (SEP) in that pixel (i.e., the pixel's ecodistrict and land-cover classes), and their tolerance to shading conditions (which, in this case is zero given all biomass was totally removed after fire) (see Table 8 for species trait values). The module algorithm first determined for which species serotiny would be activated according to shading and SEP (light-loving species and higher SEP increased the probability of serotiny being activated). It then assessed which species rely on re-sprouting and will do so depending on their re-sprouting age limits, shading and re-sprouting probability (i.e., light-loving species and higher re-sprouting probability increased the probability of re-sprouting). For any given pixel, re-sprouting was limited to species that rely on re-sprouting for which serotiny was not activated. This provided an

*Table 8. Mean parameter values (and SE) for all geographically varying species inputs and map regions.*

Species	Species Establishment	Maximum ANPP	Maximum Biomass
BETU.PAP	0.78 (0.09)	478.76 (77.77)	3,655.17 (694.24)
LARI.LAR	0.60 (0.17)	260.48 (228.97)	1,004.48 (849.30)
PICE.GLA	0.68 (0.02)	929.87 (154.36)	10,559.91 (2,163.76)
PICE.MAR	0.37 (0.15)	551.85 (367.85)	3,816.86 (2,668.30)
PINU.BAN	0.78 (0.06)	1,129.29 (201.95)	12,177.80 (1,088.17)
POPU.BAL	0.82 (0.03)	988.64 (177.21)	7,843.75 (1,254.53)
POPU.TRE	0.82 (0.03)	988.64 (177.21)	7,843.75 (1,254.53)

advantage to serotinous species that would otherwise be out-competed by species that rely on re-sprouting.

Having insufficient data to draw from, we assumed that the overall proportion of each species in the landscape doesn't change much over the course of the simulation. Our previous simulation runs showed that stand regeneration — using the LANDIS-II defaults when coupled with the fire dynamics (described below) — was inadequate to ensure that the proportion of each species across the entire landscape remained consistent with current

condition data. Rather than re-engineer the underlying LANDIS-II approach to simulating these dynamics, we instead focussed on re-parameterizing the species traits that underlie these dynamics. In particular, we increased dispersal distances and regeneration rates for all species to ensure recolonization of burned pixels, resulting in a *de facto* state-transition model formulation, used successfully in ecological simulations.

### 6.5.2.7 FIRE MODEL MODULE

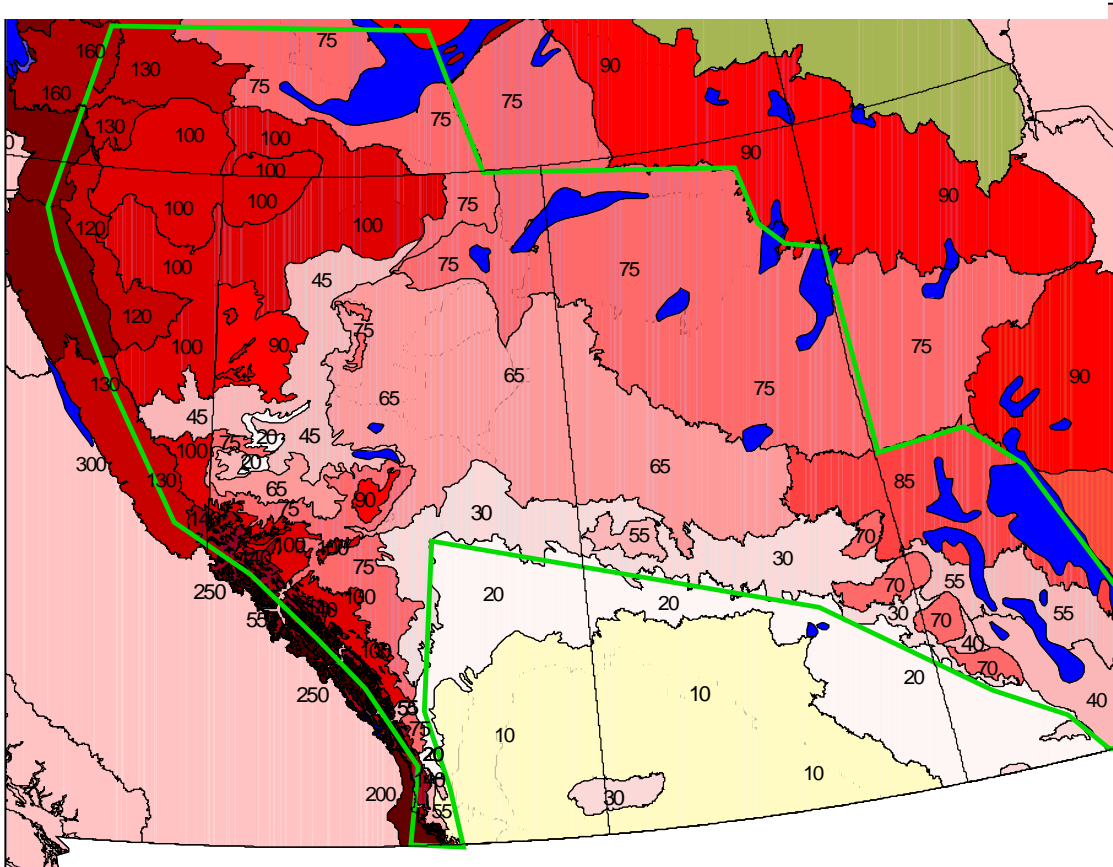
The LandR model has been designed to handle any number of generic disturbance events by accepting a disturbance layer and removing vegetation in those pixels. LandWeb considers fire as the only source of disturbance, as historically, fire is the dominant disturbance agent in boreal ecosystems.

LandWeb uses the fire initiation and spread module from the Landmine model. Landmine is a Monte Carlo based, spatially-explicit simulation model created for predicting the NRV for landscapes in the boreal forest (Andison 1996; 1998; Clarke et al. 1994) and has been widely used in various contexts both in the public and private sectors. It takes as an input a map of the Long-Term Fire Cycles (LTFC; Figure 7)



(Andison 2019) and simulates fire ignition and spread, and can be used to generate maps of forest disturbance (i.e., removes vegetation it burns). The LTFCs are used as fire return intervals in the simulations (Table 9).

Figure 7. Map of long-term fire cycles (in years) for the LandWeb study area (from Andison 2019).



For the LandWeb project, we re-implemented Landmine as a SpaDES module, with some modifications. Ignition is randomly assigned with a general area defined by fire return interval. Once a fire starts in a pixel its spread is affected by the vegetation type of neighbouring pixels (e.g., less likely to move into aspen). It “snakes” around searching neighbourhood for burnable pixels until it reaches its assigned fire size. If it gets stuck, it “jumps” to nearby pixels after a maximum number of tries. All burned pixels have their vegetation removed (i.e., all cohorts removed). The LandWeb implementation of Landmine differs slightly from the original in two ways: 1) fire sizes were drawn from a Truncated Pareto distribution (instead of a negative exponential), and 2) other parameters have not been fitted to the landscapes that are under study in the LandWeb project.

Table 9. Data sources used by Landmine fire module.

Data product	Source URL
<a href="#">Fire cycle map v6 (Andison 2019)</a>	N/A (fix)



We tracked proportion of area burned and compared against the area that was supposed to burn each year, noting that in the current version, we under-burn in many instances due to fires reaching the maximum number of “jumps” permitted. In other words, some fires simply cannot continue spreading/growing due to spatial restrictions imposed by neighbouring pixels that have inflammable cover classes or have already been burned. Even when only underburning by 1–2%, the area burned dictated by the fire return interval (LTFC) map is not achieved. Despite this, our earlier simulations showed very high disturbance causing excessive removal, coupled with insufficient regeneration of burned pixels. As mentioned above, these interactions required re-parameterization of the species traits to ensure sufficient regeneration post-fire.

#### 6.5.2.8 LANDWEB\_OUTPUT MODULE

This module produces raster maps of the leading vegetation types, as well as calculating the average time since fire over the course of the simulation.

#### 6.5.2.9 TIMESINCEFIRE MODULE

This module updates the pixel-level stand age (i.e., time since fire), by incrementing the age of unburned pixels, and resetting the ages of burned pixels to 0. It also produces raster maps of time since fire as outputs.

#### 6.5.2.10 POST-PROCESSING

Outputs from all simulation reps were used to calculate and report the NRV metrics identified by the partners, and generate custom maps for specific geographic areas (i.e., ‘reporting polygons’) within the study area. The collection of reporting polygons used in model post-processing reflects the principal considerations of forest managers and provincial government scientists, and can be classified into two main categories. First, there are reporting polygons corresponding to administrative boundaries such as provincial, park, and FMA boundaries. Second, there are reporting polygons that correspond to ecological boundaries such as ecological zones and caribou ranges. See Table 10 for a summary of reporting polygons used.

*Table 10. Summary of reporting polygons used in presenting LandWeb simulation model results.*

Reporting polygon	Source URL
<b>Administrative boundaries</b>	
Provincial boundaries	<a href="https://biogeo.ucdavis.edu/data/gadm3.6/Rsp/gadm36_CAN_0_sp.rds">https://biogeo.ucdavis.edu/data/gadm3.6/Rsp/gadm36_CAN_0_sp.rds</a> <a href="https://biogeo.ucdavis.edu/data/gadm3.6/Rsp/gadm36_CAN_1_sp.rds">https://biogeo.ucdavis.edu/data/gadm3.6/Rsp/gadm36_CAN_1_sp.rds</a>
Parks boundaries	<a href="https://www.altalis.com/map?id=117">https://www.altalis.com/map?id=117</a>
FMA area boundaries (2015)	<a href="https://www.albertaparks.ca/albertaparksca/library/downloadable-data-sets/">https://www.albertaparks.ca/albertaparksca/library/downloadable-data-sets/</a>
<b>Ecological boundaries</b>	
Ecological Land Classifications (Statistics Canada 2018)	<a href="http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/ecodistrict_shp.zip">http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/ecodistrict_shp.zip</a> <a href="http://sis.agr.gc.ca/cansis/nsdb/ecostrat/region/ecoregion_shp.zip">http://sis.agr.gc.ca/cansis/nsdb/ecostrat/region/ecoregion_shp.zip</a> <a href="http://sis.agr.gc.ca/cansis/nsdb/ecostrat/zone/ecozone_shp.zip">http://sis.agr.gc.ca/cansis/nsdb/ecostrat/zone/ecozone_shp.zip</a>
Alberta Natural Subregions (2005)	<a href="https://www.albertaparks.ca/media/429607/natural_regions_subregions_of_alberta.zip">https://www.albertaparks.ca/media/429607/natural_regions_subregions_of_alberta.zip</a>
<a href="#">Boreal Caribou Ranges (Environment Canada 2012)</a>	<a href="http://data.ec.gc.ca/data/species/protectrestore/boreal-caribou-ranges-in-canada/?lang=en">http://data.ec.gc.ca/data/species/protectrestore/boreal-caribou-ranges-in-canada/?lang=en</a>
Alberta Caribou Ranges	<a href="https://extranet.gov.ab.ca/srd/geodiscover/srd_pub/LAT/FWDSensitivity/CaribouRange.zip">https://extranet.gov.ab.ca/srd/geodiscover/srd_pub/LAT/FWDSensitivity/CaribouRange.zip</a>
British Columbia Caribou Ranges	<a href="https://catalogue.data.gov.bc.ca/dataset/caribou-herd-locations-for-bc">https://catalogue.data.gov.bc.ca/dataset/caribou-herd-locations-for-bc</a>



### 6.5.3 RUNNING THE MODEL

To ensure sample independence, the model was run for several thousand years, measuring snapshots every 100 years for a total of 60 snapshots.

## 6.6 VALIDATION

### 6.6.1 VEGETATION DYNAMICS

One of the ultimate measures of confidence in model output is the degree to which it compares to existing knowledge. One of the critical model assumptions imposed at the start of the project was that the current, existing proportion of vegetation types should reflect the average proportions from the modelling simulation runs. Although not a perfect assumption, it sufficiently captures reality notwithstanding climate change impacts. In this case, LandWeb created landscapes that shifted some vegetation types well beyond that which was expected. More specifically, the model was replacing conifer species with pioneer hardwood species and *Abies* at an unrealistic rate.

This suggested one or more model parameters, assumptions, or data inputs were not being accurately represented. This prompted a thorough and lengthy review of code and algorithms, input-data, parameters and other model assumptions. No major “bugs” were found in the code, although several data issues were identified. In the interests of time, the short-term fix was to ask the succession module to maintain (on average) the proportion of vegetation types observed on the landscape today.

After several months of attempting to reconcile this through error checking and manipulating parameters, the solution was to simplify the succession module from a vital attributes architecture (Noble and Slatyer 1980) to emulate a *de facto* state transition model (Stringham et al. 2003). However, this still created some unlikely vegetation type shifts.

There are several possible explanations for this inconsistency between actual and expected results.

- 1) The assumption that the average pre-industrial landscape conditions reflect current vegetation conditions was in error. Natural dynamics (such as fire frequency and severity) are constantly changing, and the model may in fact be accurately reflecting shifts in species based on the historical input assumptions.
- 2) The LTFC estimates (used as model inputs) were significantly wrong.
- 3) The model was under-estimating fire severity in the form of the amount and type of remnant vegetation. As the amount of unburned forest increases within individual fires, the lower the reliance on the youngest cohort to provide seed, and the greater the chances of later successional species such as white spruce to invade.
- 4) There are still un-discovered errors in the (one or more) model modules.
- 5) There are missing parameters in (one or more) of the modules that may be relevant.



- 6) The resolution (i.e., pixel size) of the model was too coarse to capture the scale at which the relevant dynamics (of mortality, forest dynamics, and succession) occur.
- 7) The succession module was not calibrated to properly reflect the ecological diversity across the larger LandWeb study area.

While some of these possibilities are more likely than others, there are arguments for and against each as follows (mirroring the same numbering reference as above):

- 1) There is merit to the possibility that vegetation types today do not reflect those of the past. However, the degree to which the model shifted vegetation types was well beyond anything expected, or likely, given our knowledge of boreal forest vegetation dynamics.
- 2) Long-term-fire-cycle is a highly influential model parameter influencing successional dynamics. The frequency and coverage of definitive, empirical studies across the LandWeb study area is highly variable. In an effort to address these gaps, a related but independent research project developed the LTFC map used here as input for the model using a combination of the available empirical evidence. The opinion of a large number of fire regime experts over four years of input was also solicited (Andison 2019). The quality of the evidence varies across the LandWeb study area. This study area is in an area with lower than average LTFC evidence quality, but the likelihood of being significantly wrong is relatively low.
- 3) The boreal forest has for many years assumed to be a “stand-replacing” ecosystem in which natural disturbances such as wildfire kill all or most of the trees resulting in single-aged forest (Johnson 1992). Most, or all, simulation models (including LandWeb) reflect this perception and a) kill 100% of the vegetation within any cell that is disturbed, and b) do not prioritize residual levels as either an input or output parameter. However, more recent evidence suggests that historical boreal wildfires are a mix of low, moderate and high severity fires (Andison and McCleary 2014). This is relevant to this study because as fire severity decreases, the amount of surviving forest increases, which changes the dynamics of regeneration, competition, and relative growth rates. For example, a fire in which only 20% of the trees survive will look very different than one in which 80% of the trees survive. It will also have very different species attributes as regards regeneration and growth.
- 4) It is not possible to be completely sure that there are not errors or logical inaccuracies. Case in point is that during the process of translating the succession module from LANDIS, the modelling team found a systematic error — in a model that has been used hundreds of times, with dozens of publications over the last 20+ years. As a reminder, models are representations of reality, and thus always wrong (to some degree). They are also notoriously under-tested against empirical data (Beverly and McLoughlin 2019). We use models because they are useful, not because they are perfect.
- 5) The possibility of the model not including key parameters is difficult to evaluate, which makes it a constant source of error of unknown influence. Just because a module is mechanistic (i.e., captures actual detailed functions) does not mean that the list of mechanisms is complete or the assumptions in terms of their influence to the output is accurate. In fact, more sophisticated





mechanistic models necessitate a significantly higher level of understanding of system processes, and thus a higher level of trust. What is the impact of parameter three (of 20+) on the outcome? What is the impact of not including parameter X, or getting it “right”? It is easier to be confident that individual model parameters are functioning as expected than it is to be confident that the various parameters fit together to create robust results.

- 6) One of the ways in which LandWeb is unique is that it attempted to blend fine-scale dynamics with coarse-scale ones. For example, the pixel size chosen for LandWeb was 6.25 ha — largely to accommodate computational efficiency. That corresponds to a square box with 250 m per side, and at least 125 m from the pixel centroid. In contrast, seeding distances for white spruce (for example) are 15–30 m. So the dispersal of white spruce seed is partly *within* pixels, and partly *between* pixels. How the model deals with such issues is critical. Similarly, the survival of individual (seed-bearing conifer) trees may not be accurately represented at a scale of 6.25 ha.
- 7) The succession module was calibrated to represent the entire LandWeb study area. In fact, the climatic, ecological, and wildfire dynamics conditions vary widely. So, while there may be places where the module performs very well, the LandWeb study area may require multiple, unique calibrations, particularly in the foothills.

As important as it is to find the source(s) of the inconsistencies described above, this issue will not impact the results in this case. Recall that the output metrics were both simple and broad. For example, when all vegetation types are combined (for both seral-stage levels and patch sizes) the results do not differ significantly from the vegetation type results. Thus, the LandWeb output will only marginally be affected by this unresolved problem. However, this issue may be more significant if/when the model is used for other purposes where the details of stand type parameters are important (e.g., habitat types, impact of climate change on species shifts, etc.).



## 7.0 RESULTS

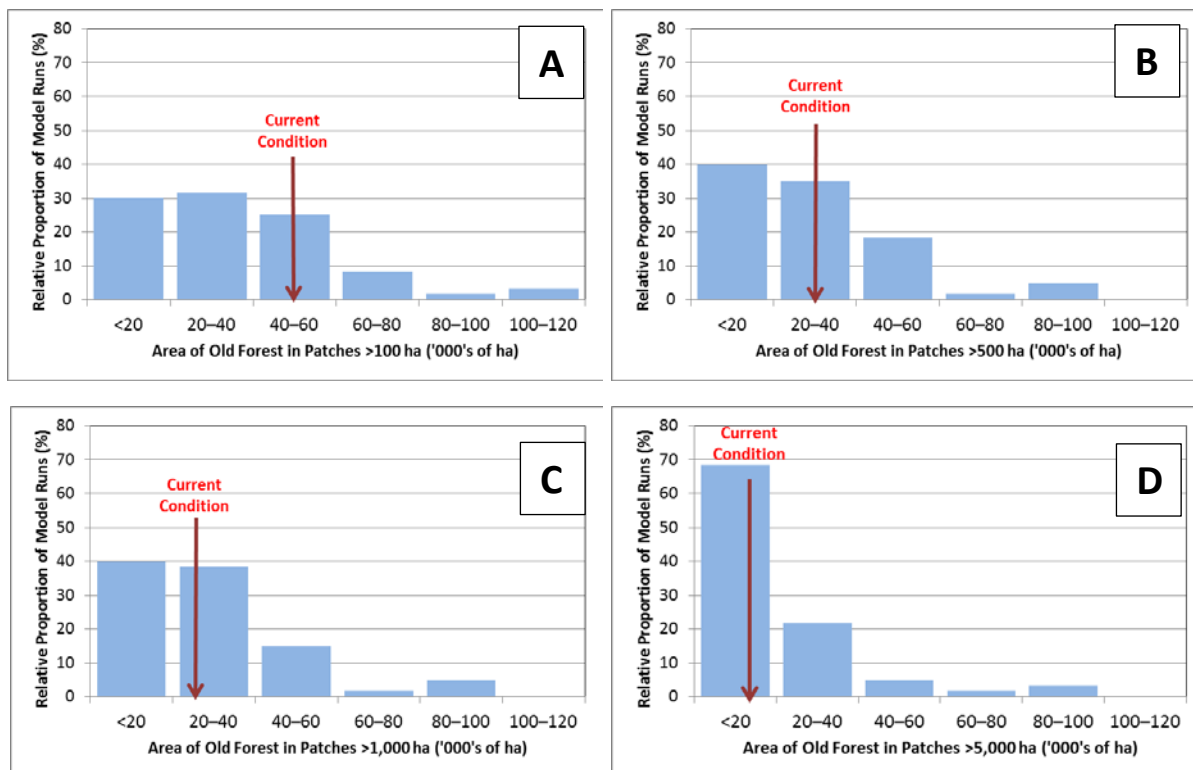
*A.M. Chubaty, E.J.B. McIntire, and D.W. Andison*

The results presented in this section include only spatial results on patch sizes. Note that the results have been adapted to align with the two pieces of the FMA area as appropriate. For non-spatial NRV results please see Andison (2015) developed for the adjacent Hinton Wood Products FMA area.

The results for four patch sizes of old forest are presented here as the number of patches >100 ha, >500 ha, >1000 ha, and >5000 ha. A “patch” is in this case defined by that portion of NRV and current condition that lies within the boundaries of the two parts of the Edson Forest Products FMA areas. Large forest patches that extend beyond the boundaries of the FMA area were captured by the model, but not reported here.

Current levels of old forest area in larger patches are all within NRV. The area of old forest patches on the study area >100 ha in size ranged between 6000 and 106,000 ha historically compared to the 50,000 ha observed today (Figure 8A). The area in old forest patches >500 ha historically ranged between <1000 and 97,000 ha compared to 31,000 ha today (Figure 8B). The pre-industrial area of old forest patches >1000 ha ranged from 0 to 94,000 ha, compared to 26,000 ha today (Figure 8C). Lastly, there were historically between 0 and 87,000 ha in large old forest patches >5000 ha on the Edson Forest Products FMA, compared to 17,500 ha today (Figure 8D).

*Figure 8. NRV (blue bars) and current condition (red arrow) of the area in old forest patch sizes on the Edson Forest Products FMA area. Top left (A) is all old forest patches >100 ha. Top right (B) is old forest patches >500 ha. Bottom left (C) is all old forest patches >1000 ha and lower right all old forest patches >5000 ha.*





## 8.0 DISCUSSION

*D.W. Andison*

### 8.1 OVERALL PATTERNS

The current area in old forest patches of various sizes are all well within NRV.

Spatial results are notoriously difficult to interpret. In this case, we defined any boundary between forest >120 years of age, and any other feature in the GIS database — natural or otherwise. However, one could argue that this is an oversimplification. Natural edges are not all equal. For example, should the boundary between 110 year and 130 year old stands be considered an edge (or in the same way) similar to that between a 20 year and 150 year old stand? Anthropogenic edges are even more challenging. Are the “edges” of highways, bush roads, and seismic lines equivalent? What about trails or 3-D seismic? What about a 20 year old seismic line versus a new one?

Edges created by human activity also come from various sources. Pickell et al. (2013, 2015) found that despite the fact that the disturbance footprint of the energy sector on this landscape was quite low, the impact on the resulting landscape patterns was far beyond that of forest management because it was so spatially ubiquitous. There is also considerable public infrastructure in the study area in the form of development, but also many types of linear features (e.g., roads, trails, rights of way).

The only way to understand the influence of each definition and sector on old forest patch size is to redo the calculation of current condition using a series of different baseline assumptions.

### 8.2 POSSIBLE SOURCES OF ERROR IN THE MODEL

One of the most widely known quotes about modelling is from Box (1979): “*all models are wrong, but some are useful*”. What he meant by that is, a) models are only representations of reality, b) every model (should) has a very specific purpose, and c) precise models are not necessarily “better” than accurate ones (Hammah and Curran 2009). This leads to the concept of parsimony: The best models should have the minimum number of parameters and assumptions necessary to address the objectives and explain the phenomenon, but no more (Haag and Kaupenjohann 2001). In other words, what is the bare minimum number of pieces moving parts to achieve the modelling goal? Parsimony also suggests that not all those parts or pieces influence the output equally.

Keeping in mind both Box’s advice and the concept of parsimony, recall that the purpose of this modelling exercise was to define some broad and simple landscape-scale, pre-industrial pattern metrics. And the reporting in this case was limited to patch sizes of old forest. Thus, the question is not so much whether the model simulated fire patterns, the probability of vegetative sprouting, or the distance of seed dispersal flawlessly, but rather which factors, parameters, or assumptions are mostly likely to *significantly* influence patch size. Thanks to the simplicity of the model — and its purpose — the range



of possibilities is limited. The most significant factor driving the area of different seral-stages is the frequency of disturbance (i.e., the LTFC). To illustrate, using a simple negative exponential mathematical model that is broadly associated with representing age-class distributions in the boreal forest (Johnson 1992), the average amount of forest older than 120 years with a 65 year long-term fire cycle (LTFC) is 16%, compared to 20% for a LTFC of 75 years, and 26% for a landscape with a LTFC of 90 years.

The process of identifying pre-industrial LTFCs in the study area was thorough and extensive, including a) an informal review of historical local records, b) a literature review, c) a two-day expert workshop, and, d) four iterations of a LTFC map from anonymous expert opinion over four years (see Andison 2019). In the end, the LTFC map represents the best available science; although the confidence level of the final LTFCs varies by region. The confidence levels of LTFCs in this particular part of the LandWeb study area were higher than average from the Andison (2019) study.

Another possible source of error could be the under-representation of low and moderate severity fires in the model. As with every other landscape-scale model today, the fire spread module in LandWeb captures and represents severity in a simplistic, binary fashion; either a pixel burns completely or not at all. However, evidence suggests that some percentage of historical fires left behind significant areas of partially burned forest (Andison 2004). This could influence succession dynamics in a number of ways. First, as residual forest levels increase, the “regeneration” components of the LANDIS succession model are less relevant, based on time-since-fire alone. For example, a 70 year old forest that experiences only 30% mortality from a fire will clearly be functioning as a sexually mature forest type, with a shade tolerant and re-sprouting understorey. Second, the introduction of low to moderate severity fires challenges, and suggests expanding on, simple definitions of a seral-stage to capture more complex forest age structures such as definitions of “old” forest (and thus patch sizes).

The last potential sources of error in the results are the current condition estimates. With reference to current condition (i.e., the red lines), data were taken from the most recent forest inventory. While AVI captures age data for every forest polygon, identifying the exact stand age is not a high priority for forest inventories. Comparisons suggest that accuracy is more of a concern than bias (Andison 1999a, 1999b).

There are two challenges inherent in the calculation of current condition for patch sizes. The first is tracking, classifying, and dating each disturbance feature. As with age data, AVI does not prioritize capturing details on all types of these data as part of its primary purpose. Fortunately, other agencies (e.g., the Alberta Biodiversity Monitoring Institute) have more recently been making significant progress on a province wide database of disturbance features that could be used to re-calculate current conditions for this project. The second challenge for the current condition estimate of patch sizes is more daunting: How to integrate and compare the impacts of forest edges of different sources and ages? For example, if/how do we differentiate edges along highways from a bush road, a large new seismic line, and a small old seismic line? For this study, any and all disturbance features were used, but this could easily be augmented by a sensitivity analysis that creates several alternative “edge” scenarios, perhaps using the ABMI data.



## 9.0 RECOMMENDATIONS

***D.W. Andison***

I would suggest two things, both intended to help clarify the results beyond that which this project was intended:

- 1) Confirm the date of, and if necessary update, the current condition red lines in the results graphs. This could be combined with the use of the new ABMI linear feature dataset.
- 2) If and when the inclusion of patch size metrics become important to track and manage for, the only way to better understand these results is to deconstruct current condition definitions using different criteria of what defines an “edge”, for how long, and from what source. This can be a simple GIS exercise, but it should also include some specific input scenarios that include explicit assumptions (e.g., “Seismic lines < X m wide that are more than Y years old will be ignored based on A and B”). This would make an excellent entry level grad student project, or a relatively quick and simple GIS contract.



## LITERATURE CITED

- Andison, D.W. 2019. Pre-industrial fire regimes of the western boreal forests of Canada. fRI Research, Hinton, Alberta. August 12, 2019. 49p.
- Andison, D.W. 2015. Using a Healthy Landscape Approach to Restore a Modified Landscape in Northeastern Alberta, fRI Research, Hinton Alberta. August 2019. 50p.
- Andison, D.W. 2004. Island remnants on foothills and mountain landscapes of Alberta: Part II on residuals. Foothills Model Forest, Hinton Alberta. 41p.
- Andison, D.W. 1999a. Assessing forest age data in foothills and mountain landscapes of Alberta: Laying the groundwork for natural pattern research. Disturbance Ecology Report Series #11 Foothills Model Forest, Hinton, Alberta. Oct. 1999.
- Andison, D.W. 1999b. Validating forest age data on the Mistik FMLA: Laying the groundwork for natural pattern research. Bandaloop Landscape Ecosystem Services. Coal Creek Canyon, Colorado. Dec. 1999.
- Andison, D.W. 1998. Temporal patterns of age-class distribution on the foothills landscapes in Alberta. *Ecography* 21: 543–550. <https://doi.org/10.1111/j.1600-0587.1998.tb00446.x>.
- Andison, D.W. 1996. Managing for landscape patterns in the sub-boreal forests of British Columbia. PhD thesis. UBC Forestry. 203p. <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0075275>.
- Andison, D.W. and K. McCleary. 2014. Detecting regional differences in within-wildfire burn patterns in western boreal Canada. *The Forestry Chronicle*. 90: 59–69. <https://doi.org/10.5558/tfc2014-011>
- Barros, C., Y. Luo, E.J.B. McIntire, A.M. Chubaty, D.W. Andison and S.G. Cumming (in prep). Land-R: a seamless union between landscape modelling and model parameterisation. To be submitted to *Methods in Ecology and Evolution*.
- Beaudoin, A., P.Y. Bernier, L. Guindon, P. Villemaire, X.J. Guo, G. Stinson, T. Bergeron, S. Magnussen, and R. J. Hall. 2014. Mapping Attributes of Canada's Forests at Moderate Resolution through KNN and MODIS Imagery. *Canadian Journal of Forest Research* 44: 521–532. <https://doi.org/10.1139/cjfr-2013-0401>.
- Beverly, J.L., and N. McLoughlin. 2019. Burn probability simulation and subsequent wildland fire activity in Alberta, Canada – Implications for risk assessment and strategic planning. *Forest Ecology and Management* 451: 117490. <https://doi.org/10.1016/j.foreco.2019.117490>.
- Booth, D.L., D.W.K. Boulter, D.J. Neave, A.A. Rotherham, D.A. Welsh. 1992. Natural forest landscape management: a strategy for Canada. (unpublished draft). 14 pp.
- Box, G. E. P. 1979. All models are wrong, but some are useful. *Robustness in Statistics* 202:549. <https://doi.org/10.1785/02201401213>
- Chubaty, A.M. and E.J.B. McIntire. 2018. SpaDES: Develop and Run Spatially Explicit Discrete Event Simulation Models. <http://cran.r-project.org/package=SpaDES>.
- Chubaty, A.M. 2019a. SpaDES.Core: Core Utilities for Developing and Running Spatially Explicit Discrete Event Simulation Models. <http://cran.r-project.org/package=SpaDES.core>.
- Chubaty, A.M. 2019b. SpaDES.Tools: Additional Tools for Developing Spatially Explicit Discrete Event Simulation (SpaDES) Models. <http://cran.r-project.org/package=SpaDES.tools>.



- Clarke, K.C., J.A. Brass, and P.J. Riggan. 1994. A cellular-automaton model of wildfire propagation and extinction. *Photogrammetric Engineering and Remote Sensing* 60: 1355–1367.
- Cosco, J.A. 2011. Common Attribute Schema (CAS) for Forest Inventories Across Canada. Timberline Natural Resource Group for Boreal Avian Modelling Project and Canadian BEACONS Project. [http://www.borealbirds.ca/files/CAS\\_Document\\_Final\\_Mar\\_2010\\_ALL\\_APPENDICES.pdf](http://www.borealbirds.ca/files/CAS_Document_Final_Mar_2010_ALL_APPENDICES.pdf).
- Davis, W. 1993. The global implications of biodiversity. M.A. Fenger et al. (eds.), *Our Living Legacy. Proceedings of a Symposium on Biological Diversity*. Victoria, BC. pp. 23–46.
- Franklin, J.F. 1993. Preserving biodiversity: species, ecosystems, or landscapes? *Ecological Applications* 3: 202–205. <https://doi.org/10.2307/1941820>
- Grumbine, E.R. 1994. What is ecosystem management? *Conservation Biology* 8: 27–38. <https://doi.org/10.1046/j.1523-1739.1994.08010027.x>
- Haag, D. and M. Kaupenjohann 2001. Parameters, predictions, post-normal science, and the precautionary principle—a roadmap for modelling decision-making. *Ecological Modelling* 144:45–60. [https://doi.org/10.1016/S0304-3800\(01\)00361-1](https://doi.org/10.1016/S0304-3800(01)00361-1)
- Hammah, R.E., and J.H., Curran. 2009. It is better to be approximately right than precisely wrong: Why simple models work in mining geomechanics. Presented at: the 43<sup>rd</sup> US Rock Mechanics Symposium and 4<sup>th</sup> US-Canada Rock Mechanics Symposium, Asheville, NC. June 28 – July 1, 2009.
- Johnson, E.A. 1992. *Fire and vegetation dynamics: Studies from the North American Boreal Forest*. Cambridge U. Press, Great Britain. 129pp.
- Krueger, T.T. Page, K. Hubacek, L. Smith, and K. Hiscock. 2012. The role of expert opinion in environmental modelling. *Environmental Modelling and Software*. 36: 4–18. <https://doi.org/10.1016/j.envsoft.2012.01.011>
- Latifovic, R., and D. Pouliot. 2005. Multi-temporal land cover mapping for Canada: methodology and products. *Canadian Journal of Remote Sensing* 31: 347–363. <https://doi.org/10.5589/m05-019>
- Long, J.N. 2009. Emulating natural disturbance regimes as a basis for forest management: A North American view. *Forest Ecology and Management* 257: 1868–1873. <https://doi.org/10.1016/j.foreco.2008.12.019>
- McIntire, E.J.B., and A.M. Chubaty. 2019. *Reproducible: A Set of Tools That Enhance Reproducibility Beyond Package Management*. Canadian Forestry Service, Victoria, BC.
- Noble, I.R., and R.O. Slatyer. 1980. The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbances. *Vegetatio* 43: 5–21. <https://doi.org/10.1007/BF00121013>
- Payette, S. 1993. Fire as a controlling process in North American boreal forest. In: West, D.C., H.H. Shugart, and D.B. Botkin (eds.), *Forest Succession: Concepts and Applications*. Springer-Verlag, New York. pp. 144–169.
- Pickell, P.D. D.W. Andison and N.C. Coops. 2013. Characterization of anthropogenic disturbance patterns in the mixedwood boreal forest of Alberta, Canada. *Forest Ecology and Management* 304: 243–253. <https://doi.org/10.1016/j.foreco.2013.04.031>
- Pickell, P.D., D.W. Andison, N.C. Coops, S.E. Gergel, and P.L. Marshall. 2015. The spatial pattern of anthropogenic disturbance in the western Canadian boreal forest following oil and gas development. *Canadian Journal of Forest Research* 45: 732–743. <https://doi.org/10.1139/cjfr-2014-0546>



Pickell, P.D, and N.C. Coops. 2016. Development of Historical Forest Attribute Layers Using Landsat Time Series and KNN Imputation for the Western Canadian Boreal Forest. University of British Columbia, Vancouver, BC, Canada.

R Core Team. 2018. R: A Language and Environment for Statistical Computing (version 3.5.2). Vienna, Austria: R Foundation for Statistical Computing. <http://www.r-project.org/>

Scheller, R.M., J.B. Domingo, B.R. Sturtevant, J.S. Williams, A. Rudy, E.J. Gustafson, and D.J. Mladenoff. 2007. Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling* 201: 409–19. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>

Scheller, R.M, and B. Miranda. 2015. LANDIS-II Biomass Succession v3.2 Extension User Guide. User Guide. <https://github.com/LANDIS-II-Foundation/Extension-Biomass-Succession/blob/master/docs/LANDIS-II%20Biomass%20Succession%20v3.2%20User%20Guide.docx>.

Scheller, R.M, and D.J. Mladenoff. 2004. A Forest Growth and Biomass Module for a Landscape Simulation Model, LANDIS: Design, Validation, and Application. *Ecological Modelling* 180: 211–229. <https://doi.org/10.1016/j.ecolmodel.2004.01.022>.

Scheller, R.M., and D.J. Mladenoff. 2007. An Ecological Classification of Forest Landscape Simulation Models: Tools and Strategies for Understanding Broad-Scale Forested Ecosystems. *Landscape Ecology* 22: 491–505. <https://doi.org/10.1007/s10980-006-9048-4>.

Statistics Canada. 2018. Ecological Land Classification, 2017. [http://publications.gc.ca/collections/collection\\_2018/statcan/12-607-x/12-607-x2018001-eng.pdf](http://publications.gc.ca/collections/collection_2018/statcan/12-607-x/12-607-x2018001-eng.pdf).

Stringham, T.K., W.C. Krueger, and P.L. Shaver. 2003. State and transition modelling: An ecological process approach. *Journal of Range Management* 56: 106–113. <https://doi.org/10.2307/4003893>.

Turner, M.G., and V.H. Dale. 1991. Modelling landscape disturbance. In: Turner, M.G. and R.H. Gardner (eds.), *Quantitative methods in landscape ecology*. *Ecological Studies* 82, Springer-Verlag. pp. 322–351.

Wilken, E.B. 1986. Terrestrial Ecozones of Canada. Ecological Land Classification No. 19. Environment Canada, Hull, Quebec. 26pp.