

**Development of a Regeneration Model  
for Prognosis<sup>BC</sup> – Design Document**

Prepared for

Abdel Azim Zumrawi  
Ministry of Forests, Research Branch  
Victoria

Prepared by

Donald C.E. Robinson  
ESSA Technologies Ltd.  
300 – 1765 West 8th Avenue  
Vancouver, BC V6J 5C6

March 31, 2003

Citation: **Robinson, D.C.E.** 2003. Development of a regeneration model for Prognosis<sup>BC</sup> – design document. (Draft). Prepared by ESSA Technologies Ltd. Vancouver, BC for BC Ministry of Forests, Victoria, BC. 9 pp.

## Table of Contents

<b>List of Figures</b> .....	<b>iii</b>
<b>Acknowledgements</b> .....	<b>iv</b>
<b>1.0 Introduction</b> .....	<b>1</b>
<b>2.0 System Design</b> .....	<b>3</b>
2.1 Component Software.....	3
Operating System .....	3
Graphical Interface.....	3
Regeneration Database .....	4
Imputation Methods .....	4
Constraints on Regeneration Data.....	4
Growth Model.....	5
2.2 Regeneration Database .....	5
Plot Information.....	6
Regeneration Information .....	7
<b>3.0 Next Steps</b> .....	<b>8</b>
<b>4.0 References</b> .....	<b>9</b>

## List of Figures

**Figure 2.1:** The regeneration system will be linked to the existing Prognosis<sup>BC</sup> growth model using the programme flow shown here. .... 2

## Acknowledgements

We express our appreciation to the individuals who provided ideas and information at a workshop held at UBC, December 17, 2002, organized by Drs. Valerie LeMay and Abdel-Azim Zumrawi. We also thank Cornel Lencar for providing us with early versions of the regeneration database, and Nicholas Crookston, for providing ideas and assistance during the development of this document.

Forestry Innovation Investment provided funding for this work, through a contract with the Research Branch of the Ministry of Forests.

This report was formatted and produced by Kelly Robson.

## 1.0 Introduction

The regeneration system links together a stand growth model and its initial inventory, a collection of observations of regeneration found at many sites, and statistical methods to infer the specific regeneration that is most likely to be found on a simulated site. When complete, it will provide plausible regeneration estimates over the course of stand succession and in the face of stand management activities.

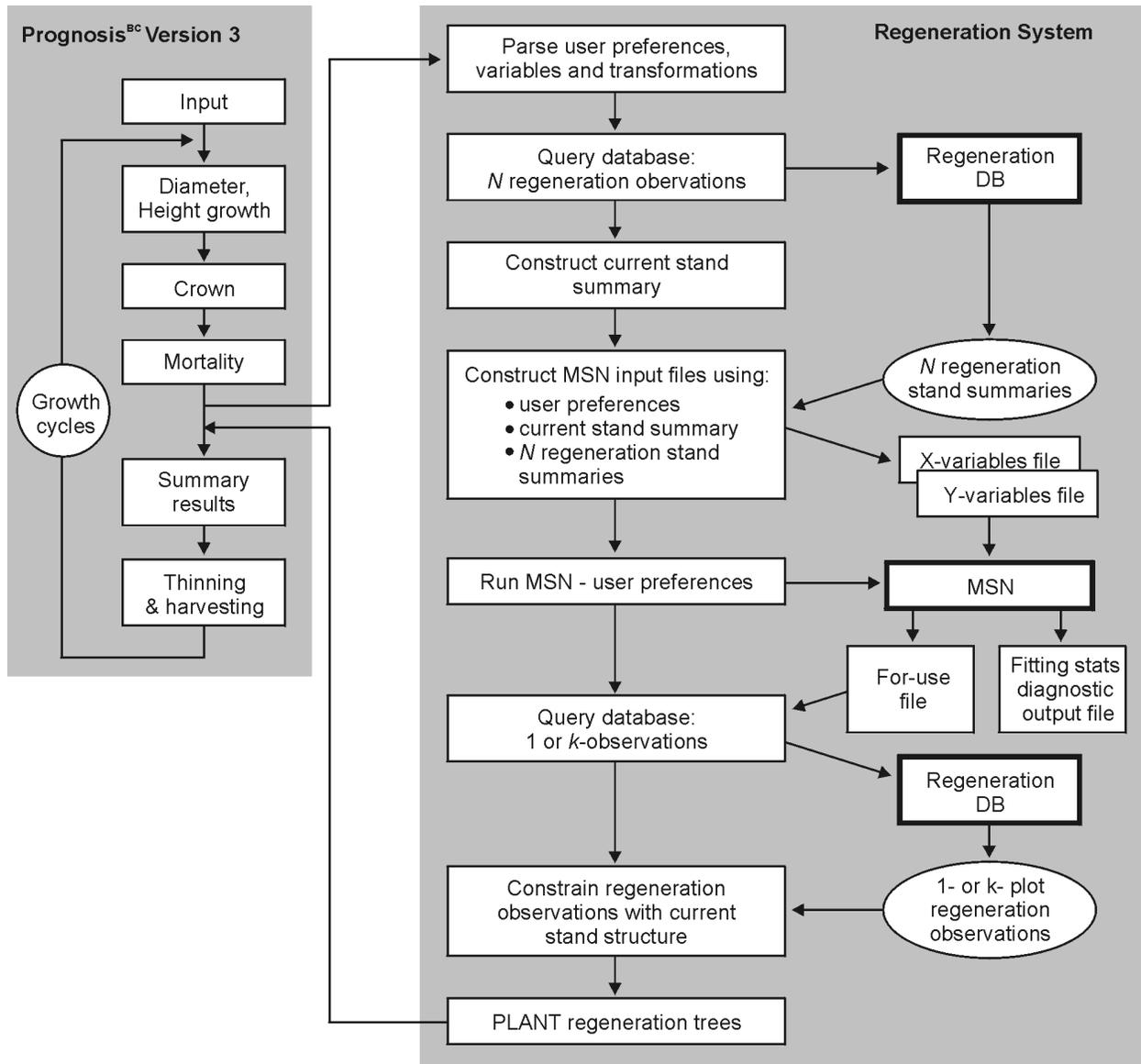
The system combines these key elements:

- a growth model;
- a statistical programme that uses a database of multivariate regeneration observations to match an empirical observation from the database to the stand where regeneration is being simulated;
- methods for model users to control the variables that are included in the similarity matching;
- methods to constrain the data that are chosen by the similarity matching; and
- an interface to simplify its use.

The first two components of this system are already available: the project brings together Prognosis<sup>BC</sup> as the growth model and the Most Similar Neighbor (MSN; Moeur and Stage, 1995; Crookston *et al.* 2002) approach to calculating similarity. A regeneration database is also being developed for this and other applications, and will be incorporated into the system. The principal task of this project is to develop ways to automate the user's preferences and the activities and interactions among the components over the course of a stand simulation.

The project has resulted in the creation of a prototype system in which user preferences are 'hard coded' into the regeneration logic, and result in the creation of intermediate files that are used as input to MSN. The MSN software is automatically invoked and the output is further manipulated so that the regeneration is 'planted' in the stand. Not all details of the system design are yet implemented: users do not yet have the flexibility to prescribe the variables and transformations they wish to use, nor do they have the ability to constrain the regeneration that is predicted.

A graphical view of the system's operation is in Figure 2.1.



**Figure 2.1:** The regeneration system will be linked to the existing Prognosis<sup>BC</sup> growth model using the programme flow shown here.

## 2.0 System Design

### 2.1 Component Software

The proposed regeneration system consists of 6 components:

1. a host operating system;
2. a graphical interface to simplify use;
3. a regeneration database;
4. methods for selecting the most appropriate regeneration data;
5. methods for constraining the inclusion of selected regeneration trees; and
6. a growth model.

#### Operating System

The following four Microsoft Windows operating systems will be supported:

- Windows 98;
- Windows NT 4.0;
- Windows 2000; and
- Windows XP

With the exception of the Windows-based Prognosis<sup>BC</sup> graphical interface (described below), the regeneration system should work on most other operating systems (OS) with additional development work. The key OS-dependent elements of the regeneration system are:

- an ODBC connection to the database, whereby SQL queries can be performed;
- a “System” call within the Fortran compiler; and
- syntax differences among operating systems (e.g. the use of “\” to separate sub-directories on a DOS-derived OS; and “/” on a Unix-derived OS).

#### Graphical Interface

The Windows-based Prognosis<sup>BC</sup> interface (a set of 3 linked VisualBasic 6 applications: DataProg, SimProg and ViewProg) provides model users with a simplified visual environment for making stand simulations. It is not supported on non-Windows operating systems. The SimProg interface will be modified to simplify the automatic selection of regeneration data. Design and implementation of the interface will follow the development of the core functionality, and will include these elements:

- the ability to schedule regeneration automatically or manually;
- the ability to select categorical and continuous variables from drop-down lists, for inclusion and exclusion in the data selection process;
- the ability to perform transformations on variables, as appropriate (e.g. slope<sup>2</sup>; log((1-p)/p));

- live linkage to the database, making it possible to customize the variables that are presented to the user;
- automatic generation of keyword sets based on the user's selections.

## Regeneration Database

The prototype database will be based on Microsoft Access 97. The choice of this software and architecture is based on its simplicity for development purposes and its ubiquity for prospective users. With minor changes, Access 2000 and Access XP (2002) could also be employed. Because the internal architecture of Access 2000/XP differs from Access 97, a decision to use one rules out using the other. If the regeneration database grows to a large size (exceeding 1 GB), alternative database systems could be employed with modest effort. The key linkage to a standalone database residing on the user's computer is the existence of an ODBC driver to link the growth model with the database, and the use of SQL queries that work in all operating environments. The ODBC connection is not strictly required, and could probably be implemented with direct file access, with a possible loss of portability to other operating systems. With additional work, the database could also be ported to an Internet-based SQL server, removing the need to issue database updates.

## Imputation Methods

The regeneration system will use the MSN software (Most Similar Neighbor Imputation Program Version 2 (Crookston et al. 2002)) to select most appropriate regeneration data. The method is based on using canonical correlation to calculate a distance metric used to find the best match (the minimum of a multivariate distance measure) between a collection of multivariate observations made from many stands, to a target stand. MSN provides five methods for computing the match. Users will initially have the ability to employ the default method only; links to other methods will be developed if resources permit. The MSN software allows categorical and continuous variables to be used in the distance calculation. During initial development, the system will make use of two hardwired variables: one categorical and one continuous. During subsequent development, users will be able to provide lists of variables to be included, as well as a number of transformations on those variables.

Besides MSN's canonical correlation approach to obtain a metric indicating similarity, other approaches are also possible. The possibility of using alternative approaches (e.g. tabular methods) will not be considered in this project, but the changes to the growth model code (described below) will be made in such a way that alternative approaches will be simple to add. Although it is simple to design the system so that other approaches easy to add, the actual implementation of any new method will likely require a substantial amount of work.

## Constraints on Regeneration Data

During the development of the modelled stand it is easy to imagine mismatches between current stand structure and the stand structures encountered regeneration data. These mismatches would result in "out-of-context" new trees 'planted' in the stand. In a separate white paper, Nicholas Crookston (pers. comm.) has identified a number of possible mismatch cases, particularly where the regeneration data themselves and time lags produce differences in the diameter distributions of the two stands. The general solution to these realistic but complex situations is to devise methods to constrain the regeneration data so that new trees are not out-of-context to the stand in which they are placed.

It is also possible to imagine stands in which  $k$  sample plots are present in the inventory. In such cases it may be appropriate to allow the user the option to select the  $k$  most similar matches, using those matches to impute different regeneration on the sample plots.

The initial prototype will be unconstrained and will return a single most similar match. Constraint-methods will be studied and tested over the course of the project. As resources permit, these constraint methods will be implemented.

## **Growth Model**

The regeneration system will be linked to the Southern Interior variant of Prognosis<sup>BC</sup> (SIBC Version 3), and will be applicable to current and future versions of the model. The existing ‘STRP’ (“stripped-down” Regeneration Model) suite of regeneration subroutines will be the basis for code enhancement. Changes to this code will be the means whereby these steps are followed:

1. regeneration is scheduled with a NEWTREES keyword invoked explicitly or through the Event Monitor;
2. user-selections are parsed, including variable-selections and transformations;
3. the current stand structure is summarized based on user-specifications;
4. the database is queried for regeneration plot summaries;
5. MSN input files are constructed based on user-specifications, employing information from the current stand structure and from summaries extracted from the database;
6. the MSN programme is run with user-supplied options, optionally generating debug information;
7. MSN output is used to select the appropriate regeneration plot data;
8. the database is again queried to retrieve regeneration data;
9. data are constrained so that mismatches are reconciled; and
10. new trees are ‘planted’ in the simulation.

Since the STRP routines are used by all but two of the Prognosis variants (the exceptions are the Northern Idaho and Eastern Montana variants, which use Ferguson and Crookston’s (1991) regeneration model). The program enhancements made to accomplish this linkage will be immediately transferable to all other variants of the Prognosis/FVS family that make use of the STRP routines.

## **2.2 Regeneration Database**

The regeneration database is being developed by Cornel Lencar, Department of Forest Resources Management, UBC. The (March 6, 2003) database exists as a number of Microsoft Excel spreadsheets, but will be exported to a Microsoft Access database and reformatted to accommodate efficient queries. The Access database will be linked to the regeneration system using an ODBC connection.

The proposed structure of the database tables and fields is shown below. Secondary tables are shown following each of the main tables.

## Plot Information

The PlotLocation Table provides a global key to information about each sample plot. The LocationID, BlockID and PlotID fields will uniquely define the topography of each sample point.

---

### PlotLocation Table

LocationID – key field  
 BlockID – key field  
 PlotID – key field  
 DB\_PlotID

### Location Table

LocationID  
 LocationName

### Block Table

BlockID  
 BlockName

### Plot Table

PlotID  
 PlotName

---

The PlotInfo Table contains topographical information and other information related to the location and history of the sample plots.

---

### PlotInfo Table

DB\_PlotID – key field  
 RegionID  
 MapsheetID  
 MapUnitID  
 BECID  
 SiteSeriesID  
 SlopePositionID  
 DisturbanceID  
 Sample Year  
 YearsSinceDisturbance  
 Slope  
 Aspect  
 Elevation

### Region Table

RegionID  
 RegionName

### Mapsheet Table

MapsheetID  
 MapsheetName

### MapUnit Table

MapUnitID  
 MapUnitName

### BEC Table

BECID  
 BECName

### SiteSeries Table

SiteSeriesID  
 SiteSeriesName

### SlopePosition Table

SlopePositionID  
 SlopePositionName

### Disturbance Table

DisturbanceID  
 DisturbanceName

---

## Regeneration Information

There are currently three spreadsheets of regeneration data: ‘Best,’ ‘Random’ and ‘all’ trees. These will be integrated into the single entity shown below.

---

### Regen Table

DB\_PlotID – key field  
 SubPlotID  
 SampleTypeID  
 SpeciesID  
 Age  
 HeightClassID  
 Height  
 DBH

### SubPlot Table

SubPlotID  
 SubPlotName centre, north, south, east, west

### SampleType Table

SampleTypeID  
 SampleTypeName best, random

### Species Table

SpeciesID  
 SpeciesCode 1, 2, 3 ,...  
 SpeciesName western white pine, western larch, Douglas-fir, ...

### HeightClass Table

HeightClassID  
 HeightClassName  
 AvgHeight  
 AvgDBH

---

### 3.0 Next Steps

The proposed sequence of development steps are described below:

1. Other members of the project team review the design document, with written suggestions considered and incorporated into a final design document.
2. Regeneration database structure is finalized, based on work and input by other team members. Currently there are some important gaps in current database:
  - There is not yet any key that unambiguously locates regeneration data to specific plots and sites, however, polygon ID may be used.
  - The content is not yet consistent: in some cases there are multiple height classes in record, with and without count data.
3. Develop hardwired code to allow generation of MSN distance based on one categorical and one continuous variable.
4. Develop code to allow keyword control over included and excluded attributes.
5. Develop code to allow keyword control for constraining new trees, removing regeneration trees that are out-of-context for the current stand.
6. Develop code to allow transformations of input variables, possibly through an Event Monitor-like command compiler.
7. Develop SimProg Interface to allow users access to the keyword functions.

## 4.0 References

**Crookston, N.L., M. Moeur and D. Renner.** 2002. User's guide to the Most Similar Neighbor Imputation Program Version 2. Gen. Tech. Rep. RMRS-GTR-96. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rock Mountain Research Station. 35 p.

**Ferguson, D.E. and N.L. Crookston.** 1991. User's guide to version 2 of the Regeneration Establishment Model: part of the Prognosis Model. Gen. Tech. Rep. INT-279. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 34 p.

**Moeur, M. and A.R. Stage.** 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science* 4:337-359.