INTEGRATING COMPLEXITY SCIENCE AND
ARTIFICIAL INTELLIGENCE:
GIS, AGENTS AND REINFORCEMENT LEARNING FOR
MODELING FOREST COVER CHANGE

by

Christopher Bone
Master of Science, Simon Fraser University
Bachelor of Arts, University of Toronto

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the
Department of Geography

© Christopher Bone 2009
SIMON FRASER UNIVERSITY
Summer 2009

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
Name: Christopher Bone
Degree: Doctor of Philosophy
Title of Thesis: Integrating Complexity Science and Artificial Intelligence: GIS, Agents and Reinforcement Learning for Modeling Forest Cover Change
Examining Committee:
  Chair: Dr. Nicholas Blomley
  Professor, Geography

Dr. Suzana Dragićević, Senior Supervisor
Associate Professor, Geography

Dr. Arthur Roberts, Supervisor
Professor, Geography

Dr. Roger White, Supervisor
Professor, Geography, Memorial University

Dr. Margaret Schmidt, Internal Examiner
Associate Professor, Geography

Dr. Arika Ligmann-Zielinska, External Examiner
Assistant Professor, Geography
Michigan State University

Date Defended/Approved: 20-July-2009
Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <http://ir.lib.sfu.ca/>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Last revision: Spring 09
ABSTRACT

Forest cover change is a complex spatially dynamic phenomenon involving the interaction of numerous processes leading to emerging forest patterns over time. This is especially true when considering forestry operations that attempt to harvest trees for maximizing short-term profits while contending with natural disturbances, fluctuating economies, and the need to conserve long-term ecosystem functions. Conventional computer models assist harvesting activities by generating forest cover strategies that satisfy both economic and ecological objectives. However, such models ignore the dynamic forces that govern the harvesting process, and as such produce strategies that can be in direct conflict with emerging patterns. The purpose of this dissertation is to enhance existing modeling approaches by bridging complex systems theory and artificial intelligence in order to incorporate spatial and temporal complexities of forest harvesting. Specifically, this dissertation introduces a novel approach for integrating geographic information systems (GIS), agent-based modeling (ABM) and reinforcement learning (RL) for developing intelligent agents that can represent stakeholder behaviours and their influence on forest cover change. Agents embedded with RL algorithms possess learning mechanisms that allow them to gain knowledge from their experiences in a dynamic forest environment as represented by GIS digital data structures. Agents learn where and when harvesting activities should take place in the forest in order to satisfy different and at times conflicting objectives that exist at varying spatial scales. These objectives are achieved amidst fluctuating timber prices, the presence of natural disturbances, and the
actions of other agents. Model results provide forest management with suitable harvesting strategies that satisfy conflicting objectives, information regarding the relationship between stakeholder interactions and emerging forest cover patterns, and the ability to evaluate the tradeoffs between different harvesting objectives. The developed approach is implemented in the context of forest management in British Columbia using datasets representing forest cover in the Chilliwack Forest District. This dissertation provides novel contributions to the fields of Geographical Information Science and Land Use/Cover Change by enhancing contemporary ABM approaches for simulating complex systems, and to the discipline of Forest Management by improving methods for understanding how to develop suitable forest cover patterns.

**Keywords:** Geographic information systems; complex systems theory; artificial intelligence; agent-based modeling; reinforcement learning; forest cover change
DEDICATION

For Beth, who continually reminds me where lie the true roots of learning.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Natural Sciences and Engineering Research Council of Canada (NSERC) for their generous funding through my Canadian Graduate Scholarship I received for three years, and for their funding and resources made available from the Discovery Grant awarded to Dr. Suzana Dragićević. My NSERC scholarship also led to the Gordon M. MacNabb Scholarship that was awarded by Precarn Inc. from which the funds were generously provided by MacDonald Dettwiler and Associates Ltd. I would also like to thank GEOIDE for their travel grant that allowed me to enjoy the Vespucci Geographical Information Science Workshop in Fiesole, Italy. In addition, I am grateful to Simon Fraser University for providing me with an entrance scholarship, a graduate fellowship, the President's PhD stipend, and several teaching opportunities through teaching assistantships and a sessional lecture position.

On a personal level, I am greatly indebted to my Senior Supervisor Dr. Suzana Dragićević for all her time, mentorship and support that I have received during both my Masters and PhD degrees at SFU. I consider myself fortunate to have had such strong guidance as she has taught me the virtue of hard work and dedication - lessons that will stick with me for a lifetime. I would like to express my thanks to my committee members, Dr. Arthur Roberts and Dr. Roger White, for their comments and advice they provided from the time of my comprehensive exams to the final defence. I thank my SAM Lab colleagues that I have had over the years, especially Shivinand Balram for all his help and guidance since I arrived at SFU in 2003, and my current lab mates Liliana
Perez and Anthony Jjumba for helping me get through the final years of my PhD. I also express my thanks for the help provided by the staff in the Geography Department, most notably the recently retired Marcia Crease, our department manager Liliana Hill, and to all others who have helped me along the way.

On a very personal level, I wouldn't be here if it wasn't for the two great people I call my parents. Their love and support (both financially and emotionally) has provided me with more opportunities than perhaps I deserve, and I only hope that they are proud of my accomplishments of which they are such a large part. I thank my two sisters Catharine and Karen, who likely don't realize how much they help me get through the tough times and enjoy the good ones. I thank the Richardson Clan who have become a very important part of my life over the past ten years; sweaters, marmalade and scotch are certainly the staples of a successful PhD. My hat goes off to Dave, who through thick and thin has always helped me to see the bright side of life, and to Tom for all our lengthy chats about every possible side of life. Finally and most importantly, I owe my accomplishments to my wonderful wife and my best friend Beth. While my PhD took four years, you have earned a lifetime of praise for putting up with me, my continual jousts with my computer, my empty gazes into nowhere, my bad back, and my sometimes less-than-tolerable moods after a bad day at work. I thank you for all your love, patience, encouragement, and for helping me appreciate what is truly important.
# TABLE OF CONTENTS

Approval .......................................................................................................................................... ii  
Abstract ........................................................................................................................................ iii  
Dedication ......................................................................................................................................... iii  
Acknowledgements ...................................................................................................................... vi  
Table of Contents ........................................................................................................................ viii  
List of Figures ............................................................................................................................... xi  
List of Tables ................................................................................................................................... xv  

1. **Introduction** ........................................................................................................................... 1  
   1.1 General Introduction ............................................................................................................. 1  
   1.2 Forest Use and Computer-based Models ............................................................................. 5  
      1.2.1 Early Computer-based Models ................................................................................... 6  
      1.2.2 Heuristic Models ....................................................................................................... 7  
      1.2.3 GIS Models .............................................................................................................. 8  
   1.3 Complex Systems Theory and Spatio-temporal Modeling ................................................ 9  
      1.3.1 Cellular Automata Models ....................................................................................... 10  
      1.3.2 Agent-based Modeling ............................................................................................. 11  
   1.4 Research Problem and Questions ....................................................................................... 14  
   1.5 Research Objectives .......................................................................................................... 17  
   1.6 Study Sites and Data ......................................................................................................... 19  
   1.7 Structure of Dissertation ................................................................................................. 20  
   1.8 References ...................................................................................................................... 23  

2. **Evaluating Spatio-temporal Complexities of Forest Management: An**  
   **Integrated Agent-based Modeling and GIS Approach** ...................................................... 32  
   2.1 Abstract ......................................................................................................................... 32  
   2.2 Introduction .................................................................................................................... 33  
   2.3 Methods ......................................................................................................................... 36  
      2.3.1 Forest Growth .......................................................................................................... 37  
      2.3.2 Agents ..................................................................................................................... 38  
      2.3.3 Economic conditions .............................................................................................. 41  
      2.3.4 AB-GIS Model ....................................................................................................... 42  
   2.4 Model Implementation and Simulation Results ............................................................... 45  
      2.4.1 Study Site and Data ................................................................................................. 45  
      2.4.2 Establishing Economic Conditions .......................................................................... 48  
      2.4.3 Forest Growth and Economic Value ...................................................................... 50  
      2.4.4 Harvesting Simulation Results .............................................................................. 50  
   2.5 Analysis ............................................................................................................................ 61  
   2.6 Conclusions ...................................................................................................................... 68
5.4 The Intelligent Agent Model ................................................................. 158
  5.4.1 Model Parameters ........................................................................ 159
  5.4.2 Agent Behaviour and Learning ..................................................... 163
5.5 Model Implementation and Results .................................................. 165
  5.5.1 Hypothetical Example ................................................................. 165
  5.5.2 South western British Columbia Example ................................. 175
5.6 Discussion and Conclusion ............................................................... 180
5.7 References ...................................................................................... 183

6. Simulation and Validation of a Reinforcement Learning Agent-based Model for Multi-Stakeholder Forest Management ................................................. 188
6.1 Abstract ......................................................................................... 188
6.2 Introduction .................................................................................... 190
6.3 The RL-ABM Multi-Stakeholder Framework .................................... 193
6.4 Methods .......................................................................................... 197
  6.4.1 Agents ....................................................................................... 197
  6.4.2 Landscape Dynamics ................................................................. 200
  6.4.3 Reinforcement Learning Algorithms ......................................... 201
  6.4.4 Validation ................................................................................ 205
6.5 Model Implementation ................................................................. 206
6.6 Results ............................................................................................ 211
  6.6.1 The Effects of Agent Behaviour ................................................ 211
  6.6.2 Validating Model Outcomes and Agent Behaviour .................... 217
6.7 Discussion ...................................................................................... 225
6.8 Conclusion ..................................................................................... 226
6.9 References ..................................................................................... 228

7. General Conclusions ......................................................................... 231
7.1 Summary of Findings ....................................................................... 233
7.2 Analyzing and Improving Model Sensitivity .................................... 242
7.3 Linking Current Work with Future Research ................................. 245
7.4 Research Contributions .................................................................... 247

Appendices ............................................................................................ 251
Appendix A: Co-authorship Statement ................................................ 252
Appendix B: Copyright Clearance from Springer .................................. 253
Appendix C: Copyright Clearance from Sage Publishing ..................... 255
Appendix D: Copyright Clearance from Wiley-Blackwell ..................... 257
LIST OF FIGURES

Figure 2-1. The process of the ABM-GIS model for a single time step representing 1 year of forest harvesting.................................................................44

Figure 2-2. Study site derived from forested region in southwestern British Columbia and spatially divided into five licensed areas.........................................................47

Figure 2-3. Twenty-five year simulation of forest growth in the absence of harvesting shown at time steps 1, 5, 15 and 25. At each successive time step the forest stands experience growth in volume as represented by the change from light to dark grey. ........................................................................................................51

Figure 2-4. Twenty-five year harvesting simulation in which harvesting costs are low, timber prices are high, and accessibility is minimal. Stands are either fully harvested or not harvested at all.................................................................54

Figure 2-5. Harvesting volume trajectories for low harvest cost, low accessibility and either (a) low timber prices or (b) high timber prices. The dashed line represents the Determined Cut Level as defined by the Government Agent every fifth time step. ..................................................................................55

Figure 2-6. Harvesting patterns generated from the twenty-five year model simulation under low harvesting costs and low timber access for A) high price, B) low price, C) cyclical price, and D) random price. ........................................................................56

Figure 2-7. Harvesting patterns generated from the twenty-five year simulation under low harvesting costs and high timber access for A) high price, B) low price, C) cyclical price, and D) random price. ..................................................57

Figure 2-8. Harvesting patterns generated from the twenty-five year simulation under high harvesting costs and low timber access for A) high price, B) Low Price, C) Cyclical Price, and D) Random Price..................................................59

Figure 2-9. Harvesting patterns generated from the twenty-five year simulation under high harvesting costs and high timber access for A) high price, B) low price, C) cyclical price, and D) random price.........................................................60

Figure 2-10. Total area harvested in m² under different economic and accessibility conditions at the end of the twenty-five year simulation. A = low timber price, B= moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price .................................................................64

Figure 2-11. Average size of patch harvested in m² under different economic and accessibility levels at the end of the twenty-five year simulation. A = low timber price, B= moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price. ........................................................................65
Figure 2-12. Number of patches under different economic and accessibility levels. A = low timber price, B = moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price. .................................................... 66

Figure 2-13. Time series graphs for volume harvested for each Timber Agent. The graphs compare harvesting under low, high, cyclical and random pricing scenarios coupled with either low accessibility (Access 1), moderate accessibility (Access 3) or high accessibility. .................................................................................. 67

Figure 3-1. The landscape for RL-LCC model is composed of the Economic Layer and Ecological Layer that are used to calculate if agents have achieved their goals, and the Value Layer that is assessed by agents to determine from which cells to extract resources. ...................................................................................................... 84

Figure 3-2. Pseudocode for describing the process of the RL-LCC. .......................................................................................... 90

Figure 3-3. The results from the original scenario in which the cell's values increase over time if they are beneficial for harvesting in order to achieve the defined objectives. Cell values are shown in relation to the cell's economic and ecological values. ......................................................................................................... 98

Figure 3-4. The learning curve of each agent with regards to obtaining high CD values. The dotted vertical line represents the moment when agents established transition rules that allow them to select specific stands continually to achieve their goals. ..................................................................................................................... 99

Figure 3-5. The average CD value for all agents for the different policies with harvesting levels of a) one stand per time step, b) five stands per time step, and c) ten stands per time step. ............................................................................................................... 102

Figure 3-6. The results from the Spatial Extent scenarios in which agents harvest ten stands per time step. Cell values are shown in relation to the cell's economic and ecological values. .................................................................................................... 103

Figure 3-7. The average CD value for all agents for the different policies using a) 20 x 20 cell spatial extent, b) 40 x 40 cell spatial extent, and c) 80 x 80 cell spatial extent. ....................................................................................................................... 104

Figure 3-8. The results from the Agent scenario in which six agents harvest ten stands per time step. Cell values are shown in relation to the cell's economic and ecological values. .............................................................................................. 106

Figure 3-9. The average CD value for all six agents for the different policies. .................................................. 107

Figure 4-1. Study site in southwestern British Columbia, Canada. ............................................................ 129

Figure 4-2. Pseudo code representing the process of IAM. .................................................................................. 133

Figure 4-3. Results from the initial IAM simulation. .......................................................................................... 134

Figure 4-4. Selected stands for each Forest Management Unit (FMU) are those with the highest stand quality that fall within the Allowable Cut Level. .................................................................................. 135

Figure 4-5. The process of agent learning is demonstrated in the improvement over time of (a) economic returns for each agent, (b) total area of harvested trees, and (c) average age of harvested trees. ................................................................. 137

Figure 4-6. Each point represents a solution generated by IAM in relation to the tradeoffs between (a) total economic return and total area harvested, (b) total economic
return and average age of harvested trees, and (c) total area harvested and average age of harvested trees. The location of the Quality Solution is highlighted.

Figure 4-7. The different disturbance simulations are defined by a combination of frequency and intensity values. The height of the bars depicts the coefficient of variation.

Figure 4-8. The difference in stand quality from different disturbance regimes defined by the frequency and intensity of disturbance.

Figure 5-1. Flowchart describing the process of the Intelligent Agent Model.

Figure 5-2. The hypothetical landscape of one hundred forest stands represented by either a low economic value (i.e. $2,000) or a high economic value ($20,000).

Figure 5-3. The sigmoidal curve depicting the decline in the step-size parameter and the probability of randomly selecting a forest stand over the course of the model.

Figure 5-4. The results for example 1, scenario \( P_{\infty} \) for five different time steps. The height of the bars represents the quality of each stand.

Figure 5-5. A comparison of the expected quality for the high value stands at time step one versus the estimated quality. Points that fall on the 1:1 diagonal line are accurately estimated.

Figure 5-6. The results at time step one for scenario \( P_{e} \), scenario \( P2 \) and scenario \( P1 \).

Figure 5-7. The study site in southwestern British Columbia. The forest stands are represented by the economic value of the timber they contain.

Figure 5-8. The results at the first time step for example 2, scenario \( P_{\infty} \). The optimal stands for harvesting are indicated.

Figure 5-9. The results at the first time step for example 2, scenario \( P50 \). The optimal stands for harvesting are indicated.

Figure 6-1. The RL-ABM framework for multi-stakeholder forest management.

Figure 6-2. The study site in southwestern British Columbia that is partitioned into four licensee areas.

Figure 6-3. The cognitive maps for the forest companies and the conservationist and the harvesting solution for Scenario 1.

Figure 6-4. The values for each objective resulting from Scenarios 1 to 5.

Figure 6-5. The episodes are separated into classes to demonstrate the change in cooperation levels as the number of episodes increases.

Figure 6-6. The difference between the total observed harvest for each forest company (FC) versus the expected total harvest as defined by the harvesting quotas. Positive differences represent that forest companies are harvesting above their quotas, while negative represent that forest companies are harvesting below their quotas.

Figure 6-7. The non-dominated solutions for Benchmark scenarios A (optimizing licensees’ objectives), B (optimizing the conservationist’s objectives), and C (equally optimizing all objectives). The final solution from each Benchmark
scenario is in bold, while the solutions for each multi-strategy scenario are represented by their respective numbers. .................................................................220

Figure 6-8. The learning rate defined by parameter \( \alpha \) that determines the amount of knowledge agents are able to acquire at each time step. .................................................................221

Figure 6-9. Total profit and the average distance to roads of all harvested stands at the end of each episode. The results are shown for scenarios 1, 3, 4 and 5 (scenario 2 was omitted as it presented the same findings as scenario 1). .................222

Figure 6-10. Average age and density of all harvested stands at the end of each episode. The results are shown for scenarios 1, 3, 4 and 5 (scenario 2 was omitted as it presented the same findings as scenario 1). .................................................................223
LIST OF TABLES

Table 2-1. Accessibility levels 1 to 5 as defined by the shortest distance of a stand to existing logging roads. .................................................................................................................................40

Table 2-2. Information on the total area, number of stands, total volume and the Determined Cut Level (DCL) of the forest licensed area for each Timber Agent........................................................................................................................................48

Table 2-3. The five timber pricing scenarios evaluated by the ABM-GIS model. Prices were defined by 1996 – 2007 log market reports provided by the Government of British Columbia..................................................................................................................................49

Table 2-4. The minimum value thresholds for the different harvesting cost scenarios. .................49

Table 4-1. Forest Management Unit information including the Allowable Cut Level (ACL). .........................................................................................................................................................130

Table 5-1. The temporal and spatial parameters of the intelligent agent model.........................160

Table 5-2. The calculation of the expected profits for all high value stands. HV = high value; LV = low value........................................................................................................................................171

Table 6-1. Attributes for each licensee's harvestable area................................................................209

Table 6-2. Timber pricing and timber price fluctuation data for each species. .........................210

Table 6-3. The percentage of stands that are harvested in each scenario presented in the vertical column that are also harvested in the scenarios presented in the horizontal column................................................................................................................213
1. INTRODUCTION

1.1 General Introduction

Land use and land cover change represent an explicit link between human activities and alterations to the earth’s terrestrial composition. Land cover is broadly defined by the attributes of the earth’s surface, which includes forest cover, agricultural and pastoral lands, and urban and rural areas. Land use refers to the exploitation of land cover through human activities such as forestry, agriculture, livestock herding and urban development (Lambin et al. 2006). Research on land-use/cover change (LUCC) informally emerged in the 1970s due to growing concerns over human impacts on climate change (Otterman 1974, Charney and Stone 1975), and has since become a branch of research as it broadened to investigate the complex relationships between a range of socio-economic interactions and various biophysical processes (Turner and Meyer 1994). While a definitive discipline of LUCC has yet to surface, its theoretical underpinnings have largely been supported by the development of complex systems theory, which attempts to understand how simple fundamental processes combine to form complex holistic systems (Gell-Mann 1994).

A component of the existing LUCC research has been dedicated to understanding how forestry practices have influenced the characteristic of forest cover change across the globe (Mapedza et al. 2003, Arima et al. 2005, Forrest et al. 2008, Liu et al. 2009, Siren and Brondizio 2009). Forests occupy over 2 billion ha of the earth’s surface (Kimmins 1997), and represent an invaluable resource for sustaining economies, communities and
living organisms. However, the use of forests as monetary-generating mechanisms through large-scale forestry operations has significantly altered natural forest cover and imposed negative societal and ecological impacts. Conventional methods for utilizing forest resources have been based on the assumption that forests can be controlled in isolation to produce maximum yields over long periods that will lead to sustained profits and economic growth. In order for this approach to be successful, system redundancies (i.e. those components of the system that are undesirable) are removed to improve the abundance and control of forest landscapes. However, research surrounding the use of forests and other natural resources has documented how removing these redundancies and focusing on only desirable components will eventually lead a system into a state of collapse that can take decades to recover (Walker and Salt 2007).

The failure of numerous enterprises dictating forest cover change in recent decades is often attributed to various problems stemming from advances in technology leading to exploitation of forests and the rise of multinational corporations resulting from the green revolution that began in the 1960s (Ponting 1991). However, it can also be argued that the roots of current failures are largely a result of our reliance on the scientific philosophy of reductionism. In a reductionist approach, the inexplicit network of interactions between components of a system is analyzed as a set of simple constituents that are linked together by simple rules (Cohen and Stewart 2000). A system is often viewed to be in a steady state of equilibrium that is void of fluxes and disturbances and closed off from external influences. As such, outcomes of a process can be explained in terms of the state of the system inputs, and any changes to these inputs should be explicitly manifest in the results. Reductionism has governed how science is conducted
through hypothesis testing and refuting implausible statements, which has no doubt led to scientific advances in numerous fields. However, the very foundations of this philosophy conflicts with the nature of natural resource management in which an intrinsic network of stakeholders interact within the complex constraints of the natural environment.

The detrimental consequences of traditional means for utilizing forest cover is gradually leading to a paradigm shift in the current science responsible for informing forest resource management. Many of the ideas regarding our ability to control forest processes in order to yield maximum returns is giving way to an acknowledgement of our uncertainty about forests due to their complex and adaptive nature. Stemming from the theory of complex adaptive systems (Holland 1992), this approach to science is a departure from reductionist thinking as it is based on the premise that a system cannot be understood solely as a sum of its constituent parts. All elements of a system have the potential of influencing the outcome of a process, as it is the interactions amongst the elements as well as their states that are crucial for understanding the system.

Geography, as a discipline, is in a beneficial position for investigating this paradigm shift in the context of forest cover change as geographic methods are gradually emerging that are focused the complex nature of spatial processes. Many of the quantitative methods developed in Geography during the twentieth century can be considered to be more in line with reductionist science, as they were based on estimates of linear spatial relationships between system elements in order to describe various patterns. This is evident in Geography’s only archived general law, the First Law of Geography (Tobler 1970), that states that all elements are related, but elements that are within closer proximity share more in common than those elements that are spatially
distant. This general notion of linear spatial relationships is the cornerstone of many statistically based geographic methods that followed. Methods for estimating spatial autocorrelation (i.e. the spatial distribution of objects), for example, describe a landscape in terms of the similarity between objects in a landscape and their distances. However, there is growing interest in Geography to provide insight into the relationship between system processes and the resulting spatial landscape patterns through linkages between real world systems and the components of complex systems theory (Hooke 2007, Bennett and McGinnis 2008, Lehotsky et al. 2008, Wilson 2009). This is evident when observing the emerging forest cover change research that utilizes geographical information systems (GIS) and spatio-temporal models based on complex systems theory. GIS are computer information systems that provide a medium for the storage, manipulation and display of geospatial data (Longley et al. 1999). GIS facilitate the digital classification and storage of forest cover attributes as individual spatial units, which can undergo spatial analysis for evaluating and determining suitable forest cover patterns. Employing spatio-temporal modeling provides a dynamic component to GIS by encoding human behaviour and natural processes into computer models that can access and alter data in order to simulate the dynamics that influence forest cover change. There now exists a substantial body of literature surrounding the use of GIS and spatio-temporal models for analyzing numerous LUCC and other geographically-related issues (Batty 1992, Evans and Kelley 2004, Brown et al. 2005b, Torrens and Benenson 2005), which has facilitated the creation of a discipline known formally as Geographical Information Science (GIScience) (Goodchild 1992). This has prompted research in LUCC to investigate the consequences of forest harvesting, yet there remains a need for enhancing methodological approaches that allow
us to develop forest cover patterns that are able to limit negative impacts by satisfying a
diverse range of objectives while understanding the complex nature of forests and timber
harvesting.

The following subsections provide a historical context to forest cover change
resulting from harvesting practices and discusses how conceptual approaches to forest
management have shaped the development of forest computer models over time.

1.2 Forest Use and Computer-based Models

The concept of forestry in North America surfaced in the late 1880s in order to
provide a framework for harvesting trees for profits (Miller 1997). In the following
decades, forest management practices were developed that explained how to convert
existing natural forests into landscapes that could be used to provide continual economic
returns over long periods of time (Thompson 1966). This is accomplished by changing
the composition of tree species, size and age structures in order to ensure that forest
stands containing highly profitable trees could be harvested each year. In 1934, this
forestry framework was coined ‘sustainable-yield forestry’ (Meyer et al. (1961), and
became enshrined as legislative policy in regions such as the U.S. and British Columbia
as a way to conduct forestry practices (Parminter 1999), Miller 1997).

Forest management under the sustainable yield framework is consistent with the
scientific philosophy of reductionism as tree growth is expected to occur in a predictable
manner, providing clear estimates of timber availability and profits that can be
anticipated at continual increments. Furthermore, reductionism views systems to be in a
steady state of equilibrium, void of fluxes or historical context as to how a system came
to exist in its current state (Blilie 2007). Similarly, successful implementation of the sustainable yield approach assumes that tree growth or the price of timber remains consistent. There is no consideration as to the interactions that shape forests over time, or the potential uncertainties of how external forces may alter the availability of timber. As a result, sustainable yield forestry as constructed through the principles of reductionist science has led to the perception that we can optimize forest operations in order to maximize our profits.

1.2.1 Early Computer-based Models

The concept of sustainable-yield forestry has its theoretical roots deep in the scientific philosophy of reductionism, as forests are understood simply by the states of their constituent parts. Sustainable-yield forestry assumes a linear and predictable relationship between current and future measurements of timber availability and timber prices. There is no consideration as to the interactions that shape forests over time, such as how different stakeholders’ actions or fluctuating timber prices influence harvesting practices. As a result, the development of the first computer forestry models was based on the notion that forest use could be optimized in a manner to obtain maximum profits. Early models were first based on linear programming procedures that were able to effectively estimate harvesting schedules at different increments of time (Wardle 1965, Littisch and Tcheng 1967, Weintraub and Navon 1976, Rorres 1978). In such models, the objectives of maximizing profits and maximizing total volume harvested are reduced to binary functions that the model learns to improve through simulation procedures. Linear programming provided a simplistic modeling approach as objectives were represented in binary form (i.e. 1’s and 0’s), which provided for fast computation for problems that have
clear economic objectives. However, the use of these methods became limited during the 1980s when forest management began to utilize the large amounts of spatial data that were becoming readily available as a result of enhanced data collection techniques.

1.2.2 Heuristic Models

Heuristic modeling techniques became popular in forestry in the 1990s for their ability to handle large spatial datasets and incorporate a number of diverse objectives. Heuristic techniques operate by self-educating the model to continuously improve a forest harvesting solution (Brack and Marshall 1992). Although the term ‘optimization’ is employed, heuristic methods focus on generating near-optimal solutions to a given problem because the complexity of integrating numerous objectives prevents us from knowing if the solutions are truly optimal. Simulated annealing (Boston and Bettinger 1999, Bettinger et al. 2002), tabu search (Gustafson and Crow 1998, Liu et al. 2006) and genetic algorithms (Venema et al. 2005, Ducheyne et al. 2006) present the three most common heuristic methods in the forest modeling literature today.

Simulated annealing and tabu search operate similarly as they both start with an initial forest harvesting solution that is iteratively changed during the model by using information regarding a solution’s ability to achieve different objectives. The solution is improved by removing those components of the solution that are considered disadvantageous and adding new components through random search processes (Liu et al. 2006). While an improved solution is developed over the course of the model, time is not taken into consideration as the solutions only dictate where harvesting should occur over the entire time period of interest. Genetic algorithms operate differently as they are developed to generate a population of solutions to a given forest harvesting problem and
improve through mechanisms mimicking biological evolution (Lu and Eriksson 2000). The algorithm selects the highest valued solutions from the solution population and breeds them through random crossovers. As a result, higher valued solutions survive while the low valued solutions perish.

1.2.3 GIS Models

In the 1990s, GIS began to play a prominent role in developing harvesting strategies due to their ability to integrate large spatial datasets that can be analyzed to determine suitable forest cover patterns that can meet different objectives (Naesset 1997, Gustafson and Crow 1998, Kangas et al. 2000, Varma et al. 2000, Store and Kangas 2001). Early attempts involved computer-assisted map overlays using Boolean operations and weighted linear combination for determining locations that satisfied a variety of objectives (Van Roessel 1986). The popularity of these methods was largely due to the fact that they are easy to implement and the results can be easily interpreted, but they are criticized for oversimplifying the complexity of the process upon which a forest cover decision is based (Malczewski 2004).

Multi-criteria evaluation (MCE) provides an improved alternative to map overlays (Carver 1991, Jankowski 1995). User preferences regarding the attributes of each objective as well as the relative importance between objectives can be implemented in the model (Mendoza and Prabhu 2000). Rather than overlaying datasets to determine homogenous descriptions of suitability, MCE provides results that depict a relative suitability of locations for meeting different objectives. Different harvesting objectives can be represented by a single data layer that can be weighted to represent the preference of one objective over another. The data layers are combined to produce a single output
map that depicts the best locations for harvesting in order to develop suitable forest cover patterns (Mendoza and Prabhu 2000, Kangas and Kangas 2005, Mendoza and Prabhu 2005)

1.3 Complex Systems Theory and Spatio-temporal Modeling

While sustainable-yield forestry has been responsible for shaping methods for determining optimal forest cover patterns from harvesting procedures, the science concerning forest management, and natural resource management in general, began to experience a paradigm shift towards the turn of the century as concepts such as sustainable-yields were challenged by theories that explore how systems are continually adapting to change (Holling et al. 2002). Complex systems theory lies at the headwaters of this shift, explaining how the multitude of local interactions between individuals in a system leads to emerging patterns over time (Holland 1992). To date, a single definition of a complex system does not exist, which is partly due to the fact that complex systems theory itself became formally recognized only as recent as the 1990s (Waldrop 1992). However, it is agreed that systems need to be analyzed in terms of the interactions between components, and not just the states of the components themselves (Manson 2001). Such interactions lead to and are governed by positive and negative feedbacks that influence the state of a system at different moments in time (Bennett and McGinnis 2008). Furthermore, complex systems theory considers how external influences and disturbances impact our certainty regarding system outcomes.

The very discourse of complex systems theory unsettles the foundation of traditional forest management that is heavily reliant on understanding the world through a
lens of predictability and certainty. This has lead to new theories about managing forests, such as “adaptive co-management” (Armitage et al. 2007), which embraces change and avoids the notion of optimization. As a result, complex systems theory is viewed as an emerging scientific philosophy that can replace traditional reductionist thinking when considering how to investigate the relationship between forest harvesting and resulting forest cover change (Colfer 2005). While complex systems theory is still considered to be in its early stages and has only really begun to influence our view of how we use and change forests, it has had an impact on the development of computer models that can address a host of LUCC issues. Most notably, cellular automata (CA) and agent-based models (ABM) have become the prominent complex systems approaches in this field.

1.3.1 Cellular Automata Models

CA models are considered a modeling approach consistent with the notions of complex systems theory (White and Engelen 2000). CA models typically consist of a regular grid of cells, each in one of a set of possible states. Each cell is surrounded by other cells in a defined neighbourhood, the shape of which is determined by the process being simulated. A set of transition rules determines if the state of a cell will change at the next time-step of the model based on its current state and the states of the cells in its neighbourhood.

CA modeling is consistent with complex systems theory because the simple rules that govern the relationships between cells in a neighbourhood produce emerging patterns over time that cannot be estimated by merely observing the initial cell states (Batty 2005). The transition of cell states leads towards positive or negative feedbacks in which the system’s existence along a specific trajectory is reinforced. The system can pass
through a threshold and enter a new trajectory in which the collective state of cells and
the set of available decision-making options are significantly different than those
available during the initial stages of the simulated process.

CA models are suitable for simulating forest cover change and other LUCC
phenomena, as landscapes can be classified into land use or cover categories and encoded
as individual cells that change states over time. Neighbourhoods can be constructed to
identify the spatial influence affecting land use and cover cells, and transition rules can
be developed to explicitly represent the processes that influence emerging spatial
patterns. CA modeling has received particular attention in the LUCC and GIScience
literature for simulating processes such as urban dynamics (White and Engelen 1993,
Batty and Xie 1994, Couclelis 1997, Clarke and Gaydos 1998) and rural residential
settlement patterns (Deadman et al. 1993). CA models have also been employed for
simulating issues surrounding forest use and forest cover change. These include modeling
deforestation for agricultural purposes (Menard and Marceau 2007), forest fires
(Alexandridis et al. 2008, Yassemi et al. 2008), insect outbreaks (Bone et al. 2006), and
for spatial allocation of forest management activities (Mathey et al. 2008)

1.3.2 Agent-based Modeling

While CA modeling is focused on how simple rules lead to land use and cover
transitions, ABM focuses on the interactions amongst computer-encoded agents that
represent the decision making behaviours of individuals, households and institutions.
Agents can either represent the autonomous decision-making behaviours of mobile
(Hoffmann et al. 2002, Lambin et al. 2003, Loibl and Toetzer 2003), and can possess multiple strategies for reacting to different states of the environment or to the actions of other agents (Zellner et al. 2009). ABM is considered a complex systems modeling approach because agents’ decisions are not merely influencing change to the environment/landscape or to the behaviour of other agents; their autonomous behaviour influences the overall behaviour of the system and the emerging patterns.

Agent autonomy requires agents to act in accordance with a cognitive model that relates their objectives to the environment through their behaviour (Wooldridge 1999). Cognition can be represented by reactionary behaviour to system changes, by rational choice-making, or by bounded rational behaviour. Although pure reactionary behaviour is simplistic in its ability to represent behaviours, it is still considered a form of cognition if agents possess the ability to react to change (Ruessell and Norvig 1997). Rational choice behaviour (derived from rational choice theory) represents agents that have the ability to solve mathematical optimization problems in order to balance long-term and short-term goals (Parker et al. 2003). This typically requires agents to possess omniscient knowledge of their environment and be able to anticipate the actions of other agents in order to maximize mathematical functions for solving problems. Without this type of knowledge, agents would not be able to make informed decisions that facilitate any form of goal achievement. Alternatively, bounded-rational behaviour represents inductive, discrete and evolving choices that lead agents towards satisfying their objectives (Manson 2006).

Parker et al. (2003) highlights some of the challenges facing the use and implementation of ABM in the context of LUCC. These include determining the balance between abstraction and the need to include critical system components; building an
adequate and acceptable experimental framework; model parameterization and validation; and determining the appropriate spatial and temporal resolutions for systems in which processes occur across different scales. As a result, several studies have responded to these challenges by examining how agent-based models can be developed in order to satisfy theoretical and implementation demands (Brown et al. 2005a, Manson 2008, Parker et al. 2008).

In addition to the aforementioned challenges, research in the GIScience literature has investigated the architectural requirements for integrating ABM with GIS. Brown et al. (2005b) explains that a primary concern with the GIS-ABM integration is defining four main relationships between the agents and their environment. These relationships are (1) identity relationships: the spatial attributes of the environment that are directly affected by agent behaviour; (2) causal relationships: the spatial attributes of the environment that are indirectly affected by agent behaviour; (3) temporal relationships: the timing of updating of the environment versus updating agents behaviour and knowledge, and (4) topological relationships: topological rules and spatial features of the environment determine whether agent actions are possible. How these relationships are formulated will influence model development and impact the resulting patterns emerging from the interactions within the agent-based model.

The benefits and challenges of developing and implementing ABM have assisted in understanding how to improve models of LUCC. Currently, ABM have been used for simulating processes such as urban migration (Benenson 1998), pedestrian movement (Batty 2001), animal foraging and competition (Ahearn et al. 2001, Bennett and Tang 2006), and urban utility use (Ducrot et al. 2004). Furthermore, the use of ABM for
simulating processes and patterns of forest cover change has been slowly emerging. Examples include modeling forest-agriculture land use transitions and consequential forest cover change (Deadman et al. 2004, Caplat et al. 2006, Evans and Kelley 2008, Bithell and Brasington 2009), co-management of tropical forests by different stakeholders (Purnomo et al. 2005, Purnomo and Guizol 2006), and for the development of spatial decision support systems for forest ecosystem management (Nute et al. 2004).

1.4 Research Problem and Questions

A review of the literature reveals an evident dichotomy between conventional methods for determining optimal patterns of forest cover and the potential for investigating forest cover processes and patterns using a complex systems approach. Linear programming, heuristic methods and GIS-based models using map overlay and MCE operations are 'top-down' models that assume forest harvesting processes will lead directly to the patterns generated by the model. These models do not consider the dynamic forces that shape forest cover over time, such as the interactions amongst different stakeholders, fluctuating economic markets, and natural disturbances. Such processes can have an impact on the availability of timber for harvesting and can dictate the emergence of forest cover patterns, which may significantly disagree with model solutions. As a result, the use of top-down models becomes limited when investigating the linkages between processes and patterns of forest cover change.

In order to address this problem, complex systems theory and complex systems models can provide a theoretical and methodological approach, respectively, for gaining insight into the relationships between forest cover processes and patterns. Complex
systems theory offers an epistemological perspective for inquiring how forest use coupled with socio-economic and biophysical processes influences forest cover dynamics. Modeling approaches such as ABM allow us to simulate how the interactions between stakeholders and the forest produce emerging patterns. Converse to ‘top-down' models, ABM are considered 'bottom-up' models because they generate results that are a consequence of the interactions amongst system components rather than overarching functions that attempt to produce optimal patterns.

While research using ABM for addressing forest cover issues is slowly emerging, there still remains a gap in the research literature regarding the ability of GIS and ABM to provide practical solutions for determining forest cover patterns that are able to achieve a diverse range of objectives. This is because ABM is widely considered an explorative method because its utility largely rests in simulating "what-if" scenarios that can inform decision makers about the likely outcomes given the inputs and constraints of a system. As such, conventional parameterization of agent cognition (whether it be reactive, rational or bounded-rational) does not provide agents with a means to learn about how their actions have influenced their ability to achieve their goals. Agents can possess different strategies that can adapt to system dynamics, but they do not contain methods that allow them to evaluate the linkages between their behaviour and emerging model results. For example, conventional models do not parameterize agent’s with the ability to estimate how landscape patterns resulting from harvesting activities measure against their objectives. This poses a limitation when attempting to generate forest cover patterns that are intended to satisfy a diverse range of objectives, and presents challenges for validating the model results.
Integrating ABM with artificial intelligence (AI) can alleviate this problem by providing agents with learning behaviour so that they can evaluate how their actions are able to satisfy their objectives in the presence of system complexity. AI is a branch of computer science that seeks to develop autonomous machines that can carry out complex tasks without human intervention (Brookshear 2006). AI originated in 1950 with the development of the Turing test (Turing 1950), and has since spawned a field that quickly infiltrated other disciplines by providing a range of techniques such as machine learning, search methods and pattern recognition. The use of AI has made its way into the GIScience literature, mostly for developing methods for spatial-allocation and multi-objective land use decision-making. However, integrating AI with complex systems theory and ABM for generating forest cover patterns has yet to receive attention.

Given the existing gap in the literature and potential for integrating complexity science and AI, this dissertation poses the following research questions:

1. How can GIS and ABM be integrated for simulating the emergence of forest cover patterns due to harvesting processes?

2. How can ABM be enhanced for improving forest cover patterns through the integration of AI?

3. Can intelligent agents effectively simulate forest cover patterns that satisfy objectives existing at different spatial scales in the presence of system complexity?

4. How can intelligent agents incorporate space and time in order to generate forest cover patterns that satisfy economic objectives under spatial constraints?

5. How can an intelligent agent approach be used to simulate forest cover patterns emerging from multi-stakeholder interactions? How can these results be used for informing forest management decision making?
1.5 Research Objectives

In order to address the proposed questions, the purpose of this dissertation is to integrate complex systems theory and artificial intelligence for generating forest cover patterns by developing intelligent agent models. Specifically, this dissertation proposes a novel approach to constructing agent cognition using reinforcement learning (RL) in order to allow agents to learn how to achieve forest cover objectives amongst the complex interactions of forest harvesting processes. RL is a heuristic based machine learning technique that originated in the 1950's with the inception of artificial intelligence. The theoretical foundations of RL evolved during the second half of the twentieth century, and the technical aspects were gradually used in applications in control theory (Waltz and Fu 1965, Mendel 1966), pattern recognition (Mendel and McLaren 1970), trial-and-error learning (Klopf 1975), and classifier systems (Holland 1986). The structure of RL as currently described in the literature was formalized by Barto et al. (1981) and Sutton (1988), and is now abundant in research on optimal control and game theory (Lee 2005, Tuyls and Nowe 2005, Gu and Yang 2007). While minimal research on the use of RL for spatial applications exists (Broggi and Cattani 2006, Marihiro et al. 2006, Park et al. 2007), the utility of RL for simulating forest cover change has yet to be explored. RL is a suitable technique for integrating learning into ABM as it reinforces the actions taken by individual agents by rewarding decisions that lead towards achieving objectives. Unlike contemporary agent-based models in the LUCC, agents in an RL model are not provided with a set of state-action behaviours that define the geographical movement or specified actions that they should take. Instead, it is assumed that agents have no a priori knowledge of their landscape. Through continual experience in their
environment, agents learn about economic fluctuations, natural disturbances, and the actions of other agents that characterize a forest harvesting process. This allows them to improve their decision making behaviour by influencing forest cover patterns in a manner to achieve their objectives.

The methods developed in this dissertation draw from the research literature on complex systems theory, GIScience, artificial intelligence, heuristic modeling, reinforcement learning, forest management and LUCC. This dissertation is not an effort to satisfy the expectations from all these fields, but instead represents an attempt to understand the necessary tradeoffs for the successful development and implementation of models for forest cover change. Given these intended pursuits, the main objectives of this dissertation are to:

1. Develop a GIS agent-based model for simulating forest cover change due to timber harvesting
2. Develop an intelligent agent approach for generating improved forest cover patterns by integrating the agent-based model with reinforcement learning
3. Implement the intelligent agent approach for evaluating how the interactions amongst multiple forest stakeholders lead to the emergence of forest cover patterns over time

To accomplish these objectives, this dissertation uses the context of forest harvesting processes in British Columbia, Canada. Forest companies represent the main actors in this region as their economic incentives are the main driving force behind emerging harvesting patterns. Government agencies also play a significant role because of their responsibility to regulate harvesting levels for maintaining economic growth while
avoiding exhaustion of available timber. Furthermore, ecological objectives and the behaviours of conservation groups are becoming significant contributors to the forest management process due to societal pressures for preserving forest services. These stakeholders and their objectives are represented as individual computer agents in the developed models in order to achieve the three objectives stated above.

1.6 Study Sites and Data

Different study sites are used throughout this dissertation for implementing the proposed methods. The data used in Chapter 2, pertains to forested areas that lie within the Chilliwack Forest district in southwestern British Columbia, which were provided by the Government of British Columbia. This location was selected because it is an area of immense forestry interest due to the existence of a significant logging road network and its proximity to timber processing and shipping resources (Pederson 2004). Furthermore, these forests contain some of the last remaining habitat for local endangered species, which raises concerns that a continuation of traditional forest management based on the desire for maximizing profits will inevitably lead to complete habitat loss and the demise of certain species. With regards to model parameterization, the study area is currently partitioned into different licensed areas that companies can lease for the purpose of harvesting trees. This information can be encoded in the GIS in order to represent where forest company agents are allowed to harvest. Furthermore, digital data exists that represent the extensive road network including highways and logging roads in and around the licensed areas.
Chapter 3 employs hypothetical datasets that were generated in ArcGIS (ESRI 2007). A cellular landscape was stylized in order to provide a representation of a forested area containing numerous forest stands. Each stand was given hypothetical economical and ecological values that influence their potential for harvesting by agents. It was deemed important to use a hypothetical dataset at this stage in the dissertation in order to test the implementation of RL for enhancing the agent-based model. For example, all forest stands in the dataset are identical in spatial dimension and contain the same volume of trees, which made it possible to effectively evaluate the explicit relationship between the parameters of the RL algorithms and the results as any influence from the stands’ spatial configurations or tree volume is negated. Furthermore, the use of a hypothetical dataset provided control over the distribution of economic and ecological attributes. This was useful during model development and calibration because the RL algorithms could be adjusted in order for the model to generate expected results.

Once the parameterization of the RL algorithms was deemed satisfactory, the integrated ABM-RL approach was implemented in Chapters 4 to 6 using datasets representing alternative forested areas in the Chilliwack Forest District. While the dataset in Chapter 4 consisted only of forest cover data, a digital road network was included in the models of Chapters 5 and 6 for representing the logging roads that forest company agents utilize when making their harvesting decisions.

1.7 Structure of Dissertation

The five chapters following the Introduction address the objectives of this dissertation through the development of models that integrate GIS, ABM and RL. Chapter 2 presents the development of a GIS agent-based model that simulates the
harvesting of forest stands by agents representing forest companies. Harvesting behaviour is influenced by timber pricing, the cost of conducting forestry operations, their ability to access timber, and by regulations set forth by a government agency that ensures harvesting levels remain at desirable levels. The results from this model are intended to reveal the types of forest cover patterns that emerge from varying economic influences such as timber pricing and harvesting costs, as well as landscape conditions such as the spatial arrangements of forest stands.

Agent cognition is explicitly represented by reaction behaviour in Chapter 2, and thus needs to be enhanced by embedding learning mechanisms. Therefore, Chapter 3 introduces the use and parameterization of RL for guiding agent learning in order for them to improve forest cover patterns by achieving economic and conservation objectives existing at different spatial scales. While the model is relatively limited in terms of including notions of complex systems, it demonstrates how different RL components influence agent learning, and evaluates if the ABM-RL approach can be easily transferred across spatial scales and with a different number of forest company agents.

In Chapter 4, this model is enhanced using a study site in southwestern British Columbia. The model is developed as a multi-objective decision making approach by formulating forest cover objectives to be maximized or minimized. This permits the results to be evaluated in terms of the non-dominated solutions that are generated by the model. Agent interaction occurs indirectly in this model, as the decisions of one agent to harvest in a particular area within its jurisdiction will have an impact on the decisions of other agents. Furthermore, complexity is introduced into the model by evaluating how agent behaviour is altered with the presence of different disturbance regimes.
Chapter 5 is a departure from the agent-based techniques employed in previous chapters. The model in this chapter is developed to evaluate how a single agent can navigate through a forest over time in order to achieve economic objectives given certain spatial constraints. The intent of this chapter is to enhance the RL algorithms developed in previous chapters to allow an agent to learn how early decisions in a forest harvesting process can influence later decisions based on the structure of the landscape and spatial harvesting limitations. A temporal weighting component is introduced to inform agents where it is best to begin harvesting operations in order to achieve forest cover objectives at each time-step of the model.

The methods developed in these four chapters are implemented in a multi-stakeholder RL-agent-based model in Chapter 6. The model contains RL-embedded agents equipped with multiple strategies and conflicting objectives that learn how to generate suitable forest cover patterns amidst different levels of altruistic behaviour. The model is validated based on the agents' ability to optimize their decision-making behaviour given the complexity of the system using a traditional analytical approach from the field of multi-objective decision making.

The concluding chapter of this research summarizes the general findings from this dissertation, discusses the potential for implementing RL and ABM for improving forest cover patterns resulting from harvesting practices, offers insight into potential future research, and discusses the contributions of this dissertation to the intended fields of research.
1.8 References


Siren, A. H. and E. S. Brondizio. 2009. Detecting subtle land use change in tropical forests. Applied Geography, 29, 201-211.


dynamics of urban and regional systems. Computers, Environments & Urban
Systems, 24, 383-400.


approach to distributed artificial intelligence (pp. 27-77). Cambridge,
Massachusetts: MIT Press.

Yassemi, S., S. Dragicevic, and M. Schmidt. 2008. Design and implementation of an
integrated GIS-based cellular automata model to characterize forest fire
behaviour. Ecological Modelling, 210, 71-84.

2009. The emergence of zoning policy games in exurban jurisdictions: Informing
collective action theory. Land Use Policy, 26, 356-367.
2. EVALUATING SPATIO-TEMPORAL COMPLEXITIES OF FOREST MANAGEMENT: AN INTEGRATED AGENT-BASED MODELING AND GIS APPROACH

2.1 Abstract

The objective of this study is to integrate agent-based modeling and geographic information systems for examining how interactions within forest management lead to patterns of land-cover change. Specifically, this study evaluates how management agents behave in the presence of variable timber prices, harvesting costs and accessibility to timber, and how their actions influence the spatial characteristics of the forest landscape over time. The GIS calculates the average harvested patch size, number of patches and total harvested area as measures of emergent patterns resulting from agent actions. The results from the agent-based GIS model reveal that good economic conditions lead to few but large harvested patches while deteriorating conditions will see more patches of smaller size if forest companies have access to high quality timber. This study emphasizes the need for a complex systems approach to forest management as the model illustrates how system elements interact in a manner to produce emergent spatial patterns over time.

2.2 Introduction

Forest management is a complex geographic process in which space and time play intrinsic roles in determining the location and timing of forestry operations. Landforms, for example, can present significant spatial constraints for enterprises with minimal harvesting technology, especially when high quality stands of timber are located at considerable distances from existing logging roads. Such problems are exacerbated during poor economic conditions as forest companies may not be able to access the quality of timber required to maintain profits. The presence of spatial constraints on harvesting practices coupled with fluctuating timber prices and harvesting costs can lead to the emergence of different landscape patterns that range from few but large harvested patches to several smaller patches dispersed across a region. The relationship between process and pattern has significant implications in forest management as it can interfere with government-based initiatives that impose top-down mechanisms for ensuring control over resource extraction. Thus, it is important that such initiatives reflect on the understanding of the influence of bottom-up spatial and temporal interactions in order for effective forest management to take place.

Complex systems theory offers an approach for evaluating the spatio-temporal relationships in forest management and other resource-based systems through modeling mechanisms such as agent-based models (ABMs). Modeling with agents can facilitate the simulation of how the behaviour of individuals and their interaction with each other and their environment produce emerging global patterns over time (Parker et al. 2003). While some models of geographic processes treat agents as interacting mobile individuals (Deadman and Gimblett 1994, Haklay et al. 2001, Batty et al. 2003, Itami et al. 2004,
Whalen et al. 2004, Torrens and Benenson 2005) and others as stationary units that make decisions affecting land-use and land-cover change (Lambin et al. 2003, Loibl and Toetzer 2003, Evans and Kelley 2004), there is a general consensus that agents are autonomous units that share an environment through an agent community and behave in a manner to reach their goals (Bousquet and Le Page 2004, Batty 2005). This has made ABMs appealing for simulating numerous resource management scenarios, including agricultural land use (Evans and Kelley 2004, Manson 2006), forest recreation (Deadman and Gimblett 1994, Itami et al. 2004), management of wildlife (An et al. 2006, Bennet and Tang 2006), fisheries (Whalen et al. 2004) and water (Ducrot et al. 2004), and other processes falling in the category of land-use/cover change (LUCC). With regards to forest management, ABMs can be linked with geographic information systems (GIS) in order to understand the spatial characteristics of harvesting patterns. While ABMs are responsible for system behaviour, GIS provide infrastructure for the storage, manipulation and display of spatial data (Burrough and McDoennel 1998). Thus, coupling ABMs and GIS provides a suitable approach for modeling the complex interactions of a system while capturing and measuring the emerging spatial properties that are important for understanding how such systems operate.

Traditionally, models for forest management were primarily concerned with timber growth and yield and were based on linear programming methods (Curtis 1962, Loucks 1962, Kidd et al. 1966, Thompson and Haynes 1971, Ware and Clutter 1971). These types of models offered top-down controlled solutions to a narrow spectrum of forestry problems. However, the formulation of complex systems theory and developments in computational efficiency and spatial data towards the end of the twentieth century
encouraged the understanding of the complexity of forest management. Models have become more sophisticated by incorporating social characteristics of management with ecological knowledge of the forest environment. ABMs have played a significant role in this progression for analyzing such things as how pastoral practices influence forest succession (Caplat et al. 2006), for developing frameworks for multiple-stakeholder forest management (Purnomo and Guizol 2006), and understanding the influence of human behaviour on wildlife habitat (An et al. 2006). Simultaneously, GIS has been employed in forest management due to the potential for integration with spatio-temporal modeling approaches such as cellular automata (Bone et al. 2007), incorporating multiple datasets with a broad range of information into the decision-making process (Guillermo and Prabhu 2000, Store and Kangas 2001, Bone et al. 2005, Kangas et al. 2005), and developing practical management policies (Naesset 1997, Varma et al. 2000).

While the use of ABMs and GIS is prominent in the field of forest management, there is limited research that integrates the temporal modeling potential of ABMs and the spatial analytical functionalities of GIS for evaluating the complexities of forest harvesting. This is especially true with regards to understanding the relationship between system dynamics and the emergence of spatial harvesting patterns. Harvesting practices create patches that have implications for the viability of wildlife populations (Hansson 1994, Williams et al. 2001), microclimatic variability (Chen et al. 1995), water balance (Hudson 2000, Murray and Buttle 2003), soil erosion (France 1997) and colonization of invasive species (De Grandpre et al. 2000). Such consequences emphasize the importance of understanding how different factors in the forest management process implicitly influence the patch characteristics.
The objective of this study is to use an integrated agent-based GIS (AB-GIS) approach for examining how local interactions within forest management lead to patterns of land-cover change. Specifically, this study evaluates how economic conditions and timber accessibility affect the patch size, number of patches, and total patch area. While research on harvested patches suggests limits to patch size and number of patches in order to preserve both ecological and geophysical conditions (Pawson et al. 2006), the complex systems approach used in this study attempts to determine if economic conditions and timber accessibility play implicit roles in the emergence of landscape patterns that can interfere with the implementation of such policies.

2.3 Methods

The model developed in this study simulates the process of forest harvesting under a timber licence-based system such as those established in British Columbia as explained by Kimmins (1997). Forest companies, who are represented as rational Timber Agents, are licensed to harvest timber from a specific area of a forested region. Their harvesting decisions are influenced by four factors: (1) a legislated harvesting level determined by a Government Agent, (2) the annual price of timber, (3) the cost of harvesting timber, and (4) accessibility to timber. Spatial landscape metrics such as average patch size, number of harvested patches and total area harvested are compared across various scenarios with varying economic conditions and accessibility constraints in order to evaluate the spatio-temporal relationships between forest management processes and harvesting patterns. The following sub-sections explain the natural growth of the forest from which Timber
Agents harvest, the role of the different agents in the model, and the economic and physical parameters that influence agent decision-making.

2.3.1 Forest Growth

Forest growth is measured by the increase in tree volume per area (m³/ha) on an annual basis. In order to estimate annual volume increase, this study employed a volume growth prediction model called the Variable Density Yield Projection (VDYP) (British Columbia Ministry of Forest 2007). Data such as stand age, species presence and species proportion of the total volume are used to calculate incremental volume increases for each year in the future for a specified timeframe. The VDYP model assumes that species composition and crown closure do not vary over time, nor does it account for changes in management activities. While this leads to rather generalized estimates of forest growth, the VDYP was calibrated from a large database of inventory sample plots and represents a rigorous approach when modeling the growth of many different stand types (Peterson et al. 1997). Thus, any error resulting from the use of the VDYP is assumed to be equal for all stands and should not impose a significant influence on the model developed in this study.

As a result of the VDYP, each stand will have an estimated volume every time step of the model. This information is then used to calculate the economic value of each stand at each time step of the model, which is used later when Timber Agents decide on which stands to harvest. The economic value in this study is represented in Canadian dollars due to the focus on forest management in Canada; however the economic value can represent any currency for the geographic location in which the model is applied. Calculating the economic value is performed using the equation
\[ Value = [v(P_{sp1}u_{sp1}) + [v(P_{sp2}u_{sp2})] + \ldots + [v(P_{sp3}u_{sp3})] \] (2-1)

where \( Value \) is the economic worth of the stand measured in dollars, \( v \) is the stand volume measured in m\(^3\)/ha, \( P \) is the proportion of total stand volume of species \( sp \), and \( u_p \) is the unit price of species \( sp \).

2.3.2 Agents

2.3.2.1 Timber Agents

_Timber Agents_ are the main agents in the model as they are licensed to harvest a specified timber volume each year and are primarily responsible for the change occurring to the landscape over time. Forest licences are used in situations when governments are responsible for allocating publicly owned land to companies or individuals interested in undertaking forestry operations. A forest licence defines the area in which harvesting can take place as well as a permitted volume of harvesting (Kimmins 1997), which is referred to in this study as the _Determined Cut Level_ (DCL). The process of defining the DCL, which is measured in m\(^3\), falls under the responsibility of the _Government Agent_ and is discussed in the next section. A _Timber Agent_ agrees to cut within some specified range of the DCL each year in order that the region can supply sufficient timber to local and foreign markets, and also to ensure economic stability within the region and continual local employment. In turn, the company or individual is guaranteed access to timber in their allocated area for a specified time period.

Experiences from real timber licensed-based systems have shown that providing some level of flexibility with regards to the DCL allows _Timber Agents_ to react appropriately to
fluctuating economic conditions. The model developed in this study follows the timber license system of British Columbia, which results in Timber Agents being permitted to harvest 50% above or below the DCL in any given year, however they must keep within 10% of the DCL over a five-year period. This allows Timber Agents to have some flexibility in order to react to fluctuating economic conditions, while remaining within acceptable harvesting limits.

In addition to economic conditions, Timber Agent behaviour is influenced by their accessibility to the timber in their licensed area, which is determined by the distance of a stand to existing logging roads. Timber Agents prefer to harvest stands that are closer to existing logging roads in order to avoid costs of establishing additional roads or harvesting by other methods (e.g. removing trees by helicopter). However, stands that are easily accessible may not be of significant value and can prevent Timber Agents from sustaining profits and thus maintaining viable operations. Accessibility is represented on a scale from 1 (low accessibility) to 5 (high accessibility), as defined by distance to existing logging roads (see table 2-1). The effect of accessibility on the spatial characteristics of harvesting is evaluated by altering accessibility within various economic conditions.
Table 2-1. Accessibility levels 1 to 5 as defined by the shortest distance of a stand to existing logging roads.

<table>
<thead>
<tr>
<th>Accessibility Level</th>
<th>Distance to Exiting Logging Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Less than 1,000 m</td>
</tr>
<tr>
<td>2</td>
<td>1,000 m - 2,000 m</td>
</tr>
<tr>
<td>3</td>
<td>2,000 m - 3,000 m</td>
</tr>
<tr>
<td>4</td>
<td>3,000 m - 4,000 m</td>
</tr>
<tr>
<td>5</td>
<td>Greater than 4,000 m</td>
</tr>
</tbody>
</table>

2.3.2.2 Government Agent

The Government Agent is responsible for establishing a regional DCL as well as an individual DCL for each Timber Agent every five years. The Government Agent calculates available timber volume across the entire region and sets the regional DCL to a level that will ensure long-term harvesting opportunities. Once the regional DCL is established, the Government Agent allocates a proportion of it to each Timber Agent based on the volume of timber contained in their licensed areas.

The Government Agent can adjust the DCL at every fifth time step if there is a significant change in harvesting by the Timber Agents. For example, during poor
economic conditions, *Timber Agents* have the flexibility to harvest below their *DCL*, which in turn will leave a greater proportion of the forest unharvested and destabilize the local employment situation. In reaction to this, the *Government Agent* increases the *DCL* to encourage large-sale harvesting in order for companies to offset the effects of poor economic conditions. Conversely, continual strong economic conditions would encourage *Timber Agents* to harvest above their *DCL* in order to take advantage of high timber prices. In this situation, the *Government Agent* would be forced to lower the regional *DCL* in order to ensure that the *Timber Agents* are not continually over-cutting the forest and depleting resources for future years. Altering the *DCL* to influence harvesting represents a feedback mechanism in the model as the *Government Agent* can exert pressure to increase or decrease harvesting at specified intervals of the simulation.

### 2.3.3 Economic conditions

The economic forces driving *Timber Agent* decisions are timber prices and harvesting costs. The model is tested using a variety of timber price and harvesting cost scenarios in order to provide a diversity of economic conditions for evaluating the spatial implications of forest management. Timber pricing scenarios were constructed using the information from timber market reports, and include (1) constant low price, (2) constant high price, (3) cyclically fluctuating prices, and (4) randomly fluctuating prices.

Harvesting costs incurred by forest companies are composed of the price of the license for harvesting, capital costs such as machinery and labour, transportation of timber, and any taxes imposed upon the company for selling the timber (Kimmins 1997). In order to incorporate all these factors into the model, a parameter called the *minimum value threshold* was developed that represents the minimum economic value of a stand
that should be harvested in the presence of current harvesting costs so that a forest company can maintain a profit. When harvesting costs are low, the \textit{minimum value threshold} will be set to a relatively low value as \textit{Timber Agents} can harvest low quality stands and still maintain profits. Conversely, high harvesting costs will force the \textit{minimum value threshold} to increase in order for \textit{Timber Agents} to receive profits. Due to the complexity of determining total harvesting costs, the model developed in this study evaluates management activities with regards to three general harvesting cost scenarios: (1) low cost, (2) moderate cost, and (3) high cost. The \textit{minimum value threshold} for each cost scenario is set to 25\%, 50\%, and 100\%, respectively, of the average stand economic value in the forested area. While these numbers are broad estimates of the cost of harvesting, they provide relative measures for evaluating how costs incurred by forestry companies affects their decisions of where and how much to harvest.

2.3.4 AB-GIS Model

The timber harvesting process described in the previous sub-sections is simulated in the AB-GIS model as shown in figure 2-1, in which one time step of the model represents a single year of timber harvesting. The time step begins with the growth of the forest, at which point the volume of each species in a stand is increased to meet the estimate of the VDYP model. The economic value of the stand is then calculated in order for the \textit{Timber Agents} to determine which stands are most beneficial to harvest. At every fifth time step (starting at time step 1) the \textit{Government Agent} determines the \textit{DCL} for the next five years for the entire region and for the areas belonging to each \textit{Timber Agent}. Next, the \textit{Timber Agents} register their \textit{DCL}, determine the current price of timber, and decide how much timber they should cut that year (referred to hereafter as the target cut). The agents then
evaluate all stands in their licensed area and determine those that are accessible; all accessible stands are placed in a queue based on economic value (i.e. in descending order, the stand with the highest value is placed at the front of the queue). Once the queue is established the *Timber Agents* evaluate the cost of harvesting and decide on a *minimum value threshold*. Next, the *Timber Agents* look at the first stand in the queue and determine if it is above the threshold. If so, the stand is harvested and the *Timber Agent* adds the volume of that stand to the current year’s total harvested volume. The agent will then compare the total harvested volume for that year to the target cut. If total harvested volume is less than the target cut, the agent will proceed to the next stand in the queue. If total volume is equal to the target cut, the *Timber Agent* will not harvest any further in that year. The time step is complete once every agent has either reached their determined cut or does not have access to any more stands above the minimum value. The model simulated a variety of scenarios in which timber prices, harvesting costs and accessibility to timber were varied, with each scenario simulated for a twenty-five year time period. The twenty-five year period was chosen in order to represent the typical length of major forest licenses given to forestry companies in British Columbia.
Figure 2-1. The process of the ABM-GIS model for a single time step representing 1 year of forest harvesting

TA = Timber Agent
GA = Government Agent
2.4 Model Implementation and Simulation Results

2.4.1 Study Site and Data

The dataset used for implementing the model was extracted from 1996 Forest Cover Data provided by the Government of British Columbia. A forest area of 70 km x 70 km from southwestern British Columbia was selected from the data as the study site because it contained a large enough area to represent a wide diversity of stands and could incorporate numerous Timber Agents. The study site consists of approximately 13,039 polygons representing homogenous forest areas ranging in size from approximately 0.1 ha to 2500 ha, and with an average size of 15.6 ha. Each polygon represents an individual stand containing information on species volume and the proportion of total stand volume that each species constitutes at the time the data were collected. The study site was divided spatially into five licensed areas as previously defined by the Government of British Columbia, each managed by a different Timber Agent. Figure 2-2 illustrates the location of the forested area and the spatial division into the licensed areas. The regional DCL for the study area was determined using the actual allowable cut level in 2007 for the forest region in which the study site is located. This value was scaled-down in order to make it appropriate for the size of the study area, which resulted in a regional DCL volume of 202,356 m³. The regional DCL was then apportioned to each licensed area based on the proportion of timber volume it contained. Table 2-2 provides information on the size, number of stands, total tree volume, and the DCL of each licensed area. The DCL was adjusted by the Government Agent at every fifth time step in order to ensure that the long-term harvesting strategy is met. The long-term harvesting strategy represents the total expected harvest by the Government Agent for the simulation period.
(i.e. annual $DCL$ multiplied by the total number of time steps of the simulation), which is 5,058,900 m$^3$. Thus, the *Government Agent* evaluates the total harvested volume at every fifth time step and adjusts the $DCL$ so that each agent will harvest an amount that collectively leads to meeting the long-term strategy.

The AB-GIS model was implemented by developing simulation routines in Agent Analyst (Argonne National Laboratory 2006), a software for integrating the spatial data storage and display capabilities of ArcGIS with an agent modeling toolkit called The Recursive Porous Agent Simulation Toolkit (Repast) (North et al. 2006). The programming language Python was used in Agent Analyst to simulate natural forest growth, and to formalize the responsive behaviour of *Timber Agents* to varying economic and accessibility conditions and the decision behaviour of the *Government Agent*. 
Figure 2-2. Study site derived from forested region in southwest British Columbia and spatially divided into five licensed areas.
Table 2-2. Information on the total area, number of stands, total volume and the Determined Cut Level (DCL) of the forest licensed area for each Timber Agent.

<table>
<thead>
<tr>
<th>Timber Agent</th>
<th>Total Area</th>
<th>Number of Stands</th>
<th>Total Volume</th>
<th>DCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36,465.8 ha</td>
<td>2,487</td>
<td>10,939,794 m³</td>
<td>36,425 m³</td>
</tr>
<tr>
<td>2</td>
<td>39,044.4 ha</td>
<td>2,310</td>
<td>12,507,899 m³</td>
<td>40,471 m³</td>
</tr>
<tr>
<td>3</td>
<td>31,635.0 ha</td>
<td>1,900</td>
<td>8,900,078 m³</td>
<td>28,330 m³</td>
</tr>
<tr>
<td>4</td>
<td>49,243.8 ha</td>
<td>3,057</td>
<td>17,230,266 m³</td>
<td>56,660 m³</td>
</tr>
<tr>
<td>5</td>
<td>50,619.6 ha</td>
<td>3,285</td>
<td>11,838,452 m³</td>
<td>38,448 m³</td>
</tr>
</tbody>
</table>

2.4.2 Establishing Economic Conditions

Timber pricing scenarios were constructed using the information provided by British Columbia timber market reports (Government of British Columbia 2007), which specify the annual price for each species of timber per cubic meter. Data were available for the period between 1996 and 2007, which were analyzed to establish an initial price for each species present in the dataset and to determine the pattern of pricing fluctuations. From 1996 to 2007, the annual price of timber fluctuated ±30% around 2007 levels. This information was used to develop the pricing schematic presented in table 2-3. All scenarios using randomly fluctuating prices were simulated 100 times; results reported for the random scenarios are an average of these simulations. Low, moderate and high harvesting costs were represented by setting the minimum value threshold for each scenario as presented in table 2-4.
Table 2-3. The five timber pricing scenarios evaluated by the ABM-GIS model. Prices were defined by 1996 – 2007 log market reports provided by the Government of British Columbia.

<table>
<thead>
<tr>
<th>Timber Pricing Scenario</th>
<th>Pricing Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Low Price</td>
<td>Set to 30% below 2007 prices</td>
</tr>
<tr>
<td>Constant Moderate Price</td>
<td>Set at 2007 Prices</td>
</tr>
<tr>
<td>Constant High Price</td>
<td>Set 30% above 2007 prices</td>
</tr>
<tr>
<td>Cyclically Fluctuating Price</td>
<td>Fluctuates systematically between -30% and +30% around 2007 prices</td>
</tr>
<tr>
<td>Randomly Fluctuating Price</td>
<td>Fluctuates randomly between -30% and +30% around 2007 prices</td>
</tr>
</tbody>
</table>

Table 2-4. The minimum value thresholds for the different harvesting cost scenarios.

<table>
<thead>
<tr>
<th>Harvesting Cost Scenario</th>
<th>Minimum Threshold Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Harvesting Costs</td>
<td>$10,000</td>
</tr>
<tr>
<td>Moderate Harvesting Costs</td>
<td>$25,000</td>
</tr>
<tr>
<td>High Harvesting Costs</td>
<td>$50,000</td>
</tr>
</tbody>
</table>
2.4.3 Forest Growth and Economic Value

Volume growth for each forest stand was determined by the Variable Density Yield Predictor (VDYP) software. Data for stand age, species type and the proportion of total volume of each species were entered into the software for calculating annual volume increases. This information was used by the AB-GIS model to increase the volume of each stand accordingly. Figure 2-3 shows the results from a twenty-five year simulation of forest growth in the absence of harvesting. The results demonstrate that volume per hectare increases to a certain level, at which point growth ceases due to the fact that space nutrient resources availability are increasing limited over time. Subsequent to volume update, the model calculates the new economic value of each stand by combining information on the updated volume and the price of each species present in the stand (equation 2-1).

2.4.4 Harvesting Simulation Results

A harvesting scenario is presented in figure 2-4 in which harvesting costs are low, timber prices are high, and Timber Agents have minimal accessibility to timber. The figure illustrates the pattern of harvesting by the Timber Agents over time in which stands are represented as either fully harvested or non-harvested. Due to the minimal access to timber, most of the harvesting is observably clustered in adjacent or nearby stands for each Timber Agent. It should be noted that the observable horizontal pattern of forest growth depicted in figure 2.3 is an artefact of previous harvesting within the study area. An examination of the dataset and study area reveal that previous harvesting predominantly occurred in large patches within valleys that were oriented along a west-to-east gradient. This created a mosaic of forest stands defined by different volume
Figure 2-3. Twenty-five year simulation of forest growth in the absence of harvesting shown at time steps 1, 5, 15 and 25. At each successive time step the forest stands experience growth in volume as represented by the change from light to dark grey.
classes that are evident in figure 2.3, which consequentially presents spatial constraints for agents determining where to harvest.

The total combined annual harvest volume for the five Timber Agents over the twenty-five year simulation is provided in figure 2-5a along with the regional DCL. The graph depicts a gradual increase in harvest leading to a spike every five years. This harvesting pattern is due to the fact that demand for timber is constantly low, therefore the Timber Agents are minimizing losses by harvesting below the DCL in hopes that the timber prices will increase. However, since they are contracted to harvest a specific volume of timber every five years, they must eventually cut an amount of trees that meets the five-year target – hence the gradual increase in harvest levels. The inverse occurs when timber prices are constantly high, as Timber Agents harvest above the DCL in order to take advantage of high prices, but they must eventually reduce harvest volumes in order to stay within their DCL limits (figure 2-5b). The DCL in figure 2-5a is increased by the Government Agent at the fifth time step in order to ensure that the agents will meet the long-term harvesting strategy as harvesting in the first five years was below anticipated levels. Conversely, the DCL was lowered in figure 2-5b to enforce a negative feedback mechanism that slows down the rate of harvesting. Further adjustments occur to the DCL, as observed to changes in the dashed line in figure 2-5b, in order for the Government Agent to regulate harvesting.

The difference in spatial harvesting patterns caused by economic conditions and accessibility can be observed in the results from the AB-GIS model simulation as presented in figures 2-6 to 2-9. Each figure depicts the results from the four timber price scenarios under different harvesting cost and accessibility conditions. The four timber
price scenarios are as follows: high timber prices (Scenario A), low timber prices (Scenario B), cyclical timber prices (Scenario C), and random timber prices (Scenario D). Forest stands are represented as either completely harvested or not harvested. In order to emphasize the influence of accessibility on the decisions of which stands to harvest, figures 2-6 to 2-9 focus on the difference between accessibility level 1 (i.e. low accessibility) and accessibility level 5 (i.e. high accessibility).

Figure 2-6 displays the results for different timber pricing scenarios under conditions of low cost and low accessibility. Each map shows there is minimal difference in the location and amount of harvested stands between the four pricing scenarios as harvesting costs are low, which means that Timber Agents are not concerned with finding high quality stands in order to maintain operations. Furthermore, the patterns remain relatively constant because accessibility is minimal, which means that Timber Agents can only access timber within close proximity to existing logging roads. Figure 2-7 provides results from timber pricing scenarios under low cost and high accessibility. The results are identical to figure 2-6 for Scenarios A-C even though accessibility to timber is maximized. The reason is that, following the logic of the previous examples, low harvesting costs do not require Timber Agents to seek high quality timber in order to maintain operations. Instead, they will continue to harvest stands that are most accessible as long as they are above the minimum value threshold. Thus, increasing accessibility does not influence patterns during these economic periods because no mechanism is present that forces Timber Agents to harvest in different locations. The reason for the differences between Scenario D in figures 2-6 and 2-7 are strictly due to the random fluctuations of timber pricing.
Figure 2-4. Twenty-five year harvesting simulation in which harvesting costs are low, timber prices are high, and accessibility is minimal. Stands are either fully harvested or not harvested at all.
Figure 2-5. Harvesting volume trajectories for low harvest cost, low accessibility and either (a) low timber prices or (b) high timber prices. The dashed line represents the Determined Cut Level as defined by the Government Agent every fifth time step.
Figure 2-6. Harvesting patterns generated from the twenty-five year model simulation under low harvesting costs and low timber access for A) high price, B) low price, C) cyclical price, and D) random price.
Figure 2-7. Harvesting patterns generated from the twenty-five year simulation under low harvesting costs and high timber access for A) high price, B) low price, C) cyclical price, and D) random price.
Simulation results under high cost – low access are presented in figure 2-8. Scenario A shows that high harvesting costs are offset by high timber prices as there is only a minimal decrease in the harvest when comparing Scenario A from figures 2-6 and 2-7. Scenario B indicates a significant decrease in harvest compared to the other scenarios, which explains that there are minimal high quality stands in close proximity to logging roads that can help agents to increase profits and compensate for poorer economic conditions. Therefore, when economic conditions are poor and accessibility is minimal, *Timber Agents* can not expect to sustain operations for long periods, especially if there is minimal high quality timber available. Scenarios C and D depicts intermediate levels of harvesting, which is intuitive as the cyclical and random pricing fluctuations will ensure that at times timber pricing will be low while at other times it will be high. Therefore, levels of harvesting should fall between low and high pricing scenarios.

The evidence for accessibility affecting the harvesting patterns is clearly visible in figure 2-9 in which harvest costs are high and accessibility maximized. While Scenario A remains relatively the same as in figures 2-6 to 2-8, Scenarios B, C, and D indicate that increasing accessibility creates a more dispersed pattern of patches across the landscape.
Figure 2-8. Harvesting patterns generated from the twenty-five year simulation under high harvesting costs and low timber access for A) high price, B) Low Price, C) Cyclical Price, and D) Random Price.
Figure 2-9. Harvesting patterns generated from the twenty-five year simulation under high harvesting costs and high timber access for A) high price, B) low price, C) cyclical price, and D) random price.
2.5 Analysis

The findings from the simulation results were analyzed using spatial metrics in order to further explain how economic conditions and timber accessibility lead to varying forest spatial landscape patterns over time. Spatial metrics are quantitative methods for measuring the spatial structure and patterns (Herold et al. 2005, Ji et al. 2006). These measures are obtained through digital analysis of thematic maps that contain spatial heterogeneity at a specific scale. While several measures exist, the analysis in this study uses patch size, number of patches and total patch area in order to evaluate harvesting patterns. Figure 2-10 illustrates that total area harvested remains relatively stable under all low and most medium cost scenarios, but decreases significantly under some high cost scenarios (i.e. low, cyclical and random prices). The reason for these results is that poorer economic conditions will cause Timber Agents to reduce the amount that they harvest because they cannot maintain profits with the available timber. However, there is a notable increase for the low, cyclical and random price scenarios under high cost as accessibility is increased. As explained by the simulation results, this is due to the fact that improving accessibility to timber during poor economic conditions allows the Timber Agents access to high quality timber stands that help them offset high costs and low prices.

Figure 2-11 illustrates that average patch size remains relatively stable under low and medium costs, but is significantly smaller for the low, cyclical and random price scenarios when harvesting costs are high. This is because total harvests decline under poor economic conditions which in turn results in smaller overall patches. However, there exists minimal difference between these values when improving accessibility to timber.
This is counter-intuitive as it could be expected that an increase in accessibility (and hence higher harvesting levels) would directly lead to larger patches. The reason for patches remaining the same size, as supported by the simulation results in figures 2-6 to 2-9, is because the high quality stands in less accessible areas that are needed to offset poor economic conditions are not all located adjacent to each other. This is further supported by figure 2-12 which indicates that the total number of patches increases significantly when accessibility is improved for most high cost scenarios. More patches under poorer economic conditions reveals that *Timber Agents* are selecting stands that are dispersed across the landscape as they provide them with enough revenue to sustain operations. The highly dispersed pattern under poor economic conditions is depicted in the simulation results in Scenario D of figure 2-9.

To summarize the results from figures 2-10 to 2-12, harvesting costs have a direct impact on all three spatial metrics as high costs lead to a decline in total harvesting, and consequentially the number of sizes and average patch size. Timber prices also exert influence on patterns as the low, cyclical and random scenarios generally produced lower values of all three spatial measures, however there were no significant changes under the medium and high timber price scenarios. While accessibility was not an important factor in determining patch size, increasing accessibility did lead to a rise in the number of patches and total area harvested under high cost scenarios.

The findings from the simulation results are further reinforced in the time series graphs presented in figure 2-13. The graphs depict the annual harvest for each agent under high harvesting costs, and emphasize the influence of timber prices and accessibility in such conditions. The high price graphs show that the majority of *Timber*
Agents are able to maintain operations for the entire simulation period, with the exception of Timber Agent 5. However, this agent is able to harvest over the entire twenty-five year period as accessibility increases. This is likely due to the fact that Timber Agent 5, which has the third highest volume amongst all the agents, does not have sufficient high quality timber in low accessibility areas. This finding stresses that the location of quality timber is a prominent factor in the ability to maintain harvesting. The low price scenarios repeat the previous finding by showing that coupling low prices and high costs reduces the likelihood that agents will be able to harvest successfully for long periods. Timber Agent 4 is the only agent that, when accessibility is maximized, is able to maintain operations. This is due to Timber Agent 4 having the highest volume of timber in its licensed area, which increases the probability of finding high quality stands, especially as accessibility increases. Thus, large licensed areas may be better at withstanding poor economic conditions as timber availability is diversified. The cyclical and random pricing scenarios show harvested volumes that are between the high and low price scenarios. As mentioned above, this is due to prices being periodically either high or low, which leads the Timber Agents to experience the effects of both high and low pricing scenarios.
Figure 2-10. Total area harvested in m² under different economic and accessibility conditions at the end of the twenty-five year simulation. A = low timber price, B = moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price.
Figure 2-11. Average size of patch harvested in m² under different economic and accessibility levels at the end of the twenty-five year simulation. A = low timber price, B = moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price.
Figure 2-12. Number of patches under different economic and accessibility levels. A = low timber price, B = moderate timber price, C = high timber price, D = cyclical timber price, and E = random timber price.
Figure 2-13. Time series graphs for volume harvested for each Timber Agent. The graphs compare harvesting under low, high, cyclical and random pricing scenarios coupled with either low accessibility (Access 1), moderate accessibility (Access 3) or high accessibility.
2.6 Conclusions

The simulation results in this study reveal that economic conditions and timber accessibility impose significant influences on the spatial patterns of timber harvesting. Patch sizes are likely to be largest during good economic conditions as forest companies will focus harvesting efforts on stands that are most accessible. Patch sizes decrease as costs go up and companies are forced to look for high quality stands that can exist in dispersed locations. As a result, the number of patches will also increase as the total harvest is divided into numerous smaller patches. However, this only occurs when accessibility increases as the high quality stands would otherwise be unavailable. The fact that the spatial measures do not significantly change between low and moderate costs suggests that there exists a harvesting cost threshold that determines the ability of Timber Agents to maintain harvesting operations. A similar threshold may also exist for timber prices as moderate and high timber prices do not exhibit significant change, even under high cost scenarios, but low timber prices cause a significant decrease in all three spatial measures. Furthermore, observations from the time-series graphs explain that the size of licensed areas and the abundance of high quality stands also play a role in the emerging spatial patterns from harvesting as larger areas have a greater chance of containing the quality of timber necessary for compensating losses during poor economic conditions.

The findings from this study can be attributed to the design of coupling agent-based modeling and GIS, as the influences of both space and time on forest land-cover change were available for analysis. One of the main advantages of the AB-GIS approach is the flexibility of the model to be used in a variety of forestry applications. A model such as the one presented in this study can be applied on different scales, from small
independently-owned forests to large regional forests. With regards to the Timber Agents, the model can operate on few agents or a few hundred agents, but the number of possible outcomes becomes enormous when agents are abundant. Furthermore, the model is based on the behaviour of agents and as a result is not dependent on a particular dataset, which enhances its flexibility by allowing it to be implemented on different study sites. Thus, different areas can easily be compared in order to determine how specific management strategies perform in a diversity of ecological and economic environments. The model could also incorporate real-time timber pricing and harvesting costs so that possible future scenarios can be provided based on current conditions.

Agent-based modeling provides a useful approach for analyzing and understanding the complexities of forest management; however future research in this field requires a focus on the issue of model testing and validation. This involves comparing the modeled results to reality, the latter of which is typically represented in a dataset of the same geographic area as the model input, but from a later period in time. Validation presents a challenge as forest landscapes that are often used in modeling experiments stretch across vast areas and contain considerable information so that it is extremely expensive and time consuming to collect data of the same area for different periods in time. Furthermore, even if data could be collected at different time intervals, complex geographic processes are subject to feedbacks and random events that can lead to several different outcomes. Thus, a dataset used for validation may represent only one of a set of plausible outcomes. Therefore, these special research efforts are important but beyond the scope of this study. They are future steps that are necessary in order to insure
the model generates emergent patterns that can be used in real forest management applications.

Forest management can integrate knowledge gained from the AB-GIS model in policy development in order to represent the implicit nature of interacting factors in harvesting operations. A variety of economic and accessibility scenarios can be simulated that are specific to local forestry operations. This can provide management with an understanding of the challenges of implementing top-down mechanisms for dictating how forest resources are extracted (e.g. specifying maximum number of harvest patches) in the presence of real economic and accessibility constraints implicit in forestry operations. The ability to use this approach for development of practical policies will prove beneficial for linking processes in forestry, as well as other process of LUCC, to the emergence of spatial landscape patterns over time.
2.7 References


3. DEFINING TRANSITION RULES WITH REINFORCEMENT LEARNING FOR MODELING LAND COVER CHANGE

3.1 Abstract

Spatio-temporal modeling provides the opportunity to simulate geographic processes of land use and land cover change (LUCC) by integrating geographic information systems (GIS) with various machine learning approaches to computing. Contemporary models are often developed using a training dataset to define a set of probabilistic transition rules that govern how a landscape changes over time. However, the use of training datasets can be problematic for spatio-temporal modeling as they can limit the ability to incorporate system complexity and hinder the transferability of the model to different datasets. The purpose of this study is to evaluate a machine learning approach called reinforcement learning (RL) for defining transition rules for GIS-based models of land cover change due to natural resource extraction. Specifically, RL is evaluated based on its potential for constructing models independent of training datasets that can handle different levels of complexity and be transferred across different spatial extents. An RL model for Land Cover Change (RL-LCC) is developed for considering economic and ecological goals involved in natural resource management, and implemented using a hypothetical forest management scenario. Simulation results reveal

---

that agents in the RL-LCC model are able to develop transition rules from their experience in their landscape in a variety of simulation scenarios that allows them to achieve their goals. This study demonstrates the benefits of integrating RL and GIS in order to address important issues of space, time and complexity.
3.2 Introduction

Spatio-temporal modeling offers the ability to explore the characteristics of complex geographic processes in which landscapes evolve through the interactions of natural and human entities. The advancement of spatial models in recent years has been facilitated by coupling geographic information systems (GIS) with existing dynamic modeling approaches from a variety of disciplines. GIS are computer-based systems that provide the ability to input, store, manage, analyze and display data of geographic nature (Burrough and McDonnell 1998). As these systems typically provide static geographic representations, approaches for modeling dynamic phenomena can be integrated with GIS in order to provide the dimensions of both space and time when analyzing geographic processes. Landscapes can thus be represented with dynamic entities whose change can be simulated using a set of transition rules that explain how such entities change over time.

While the broad range of spatio-temporal modeling techniques is often characterized under different contextual frameworks, such as geocomputation (Openshaw 2000), land-use/cover change (LUCC) (Parker et al. 2003) and geosimulation (Beneson and Torrens 2004), most methods are concerned with exploring and explaining the relationship between geographic processes and their resulting spatial patterns. Modeling approaches such as cellular automata (CA) and agent-based modeling (ABM) abound in the spatial modeling literature for GIS-based simulation of a variety of geographic phenomena, including resources management (Deadman and Gimblett 1994, Janssen 2003, Whalen et al. 2004), urban dynamics (White and Engelen 1993, Clarke et al. 1997, Batty et al. 1999, O'Sullivan and Torrens 2001), and ecological process of land cover
change (Darwen and Green 1996, Baltzer et al. 1998, Matsinos and Troumbis 2002, Feagin et al. 2005). The majority of such applications are developed by explicitly defining land use/cover transition rules and performing a number of simulations from which a response to environmental constraints, stochastic events and varying parameters can be observed. Such observations are useful for exploring the nature of geographic processes and for providing practical knowledge for solving complex spatial problems.

More recently, spatio-temporal modeling applications have incorporated the use of machine learning methods in order to enhance the way in which transition rules are defined. As explicitly-stated user-defined transition rules can be subjective and may lack knowledge of the uncertainty in the system, machine learning offers numerous approaches for implementing artificial intelligence within the transition rule definition process (Diplock 2000).

The most common forms of machine learning for modeling geographic processes are those that implement supervised learning algorithms, which includes approaches such as artificial neural networks (ANNs) (Li and Yeh 2002, Dai et al. 2005, Pijanowski et al. 2005), Bayesian networks (Lei et al. 2005, Kocabas and Dragićević 2007, Ma et al. 2007) and decision tree learning (Aalders and Aitkenhead 2006, McDonald and Urban 2006). In a spatio-temporal modeling context, these supervised learning procedures typically involve determining transition rules of a system by constructing a relationship between the inputs and outputs of the system. The input is most commonly a dataset of a geographic area representing one moment in time, while the output is a dataset of the same area representing a later period. Through numerous iterations the algorithms continually adjust estimates of how the land use variables in the inputs change to become
those observed in the output data. Once the transition rules are determined they can be applied for extrapolating future LUCC scenarios.

While the inception of supervised learning algorithms in spatio-temporal modeling has enhanced the ability to explore geographic processes, the fact that learning is based on the relationship between existing input and output datasets limits the extent to which they can be applied to simulate systems of increasing complexity. This is because complex geographic processes are often characterized as having numerous entities interacting with each other at different spatial and temporal scales, which in turn has an effect on the emerging spatial patterns of a landscape (Batty and Torrens 2005). The generated patterns are subject to different inherent forces such as stochastic processes and feedbacks that result in non-linear relationships between system inputs and outputs. These present considerable challenges to supervised learning methods as the specific output dataset used for training the model is potentially only one in a set of plausible outcomes. Furthermore, the need for an output dataset for model training can pose a challenge as spatial datasets are expensive to collect and not always readily available for the area or time of interest. As a result of these challenges, there is a need for spatio-temporal modeling approaches in which transition rules are developed independently of the causal structure of system inputs and outputs.

The objective of this research is to provide an alternative machine learning approach for developing transition rules in the simulation of complex geographic processes. In particular, reinforcement learning (RL) is evaluated for its ability to provide a flexible model framework for simulating land cover change dynamics as a result of natural resource management. RL is a machine learning method in which transition rules
are learned through the experience of system agents in their landscape rather than information retrieved from the relationship between inputs and outputs of a system. RL models are simulated using agents that are not directed on how to behave based on user-defined rules or training data, but instead they make independent decisions by selecting areas within the landscape to extract resources without knowledge about the attributes of the selected areas. After agents make their decisions, they receive either a positive or negative reinforcement depending on whether their decision has led them closer to reaching their goal. As a result, RL-embedded agents learn over time how to define a set of transition rules that brings them towards goal achievement.

While the history of RL is one that involves the confluence of ideas from many fields, its foundations were formed in the mid-twentieth century in the initial speculations on computers and the use of artificial intelligence, such as those put forth by Turing (1950). The term reinforcement learning was not widely used until the 1960’s when applications in engineering applied RL to control systems (Waltz and Fu 1965, Mendel 1966) and pattern recognition (Mendel and McLaren 1970). Its popularity was fostered by to RL-related developments in computer science such as classifier systems (Holland 1986)) and trial-and-error learning investigated by Klopf (1975). However, contemporary formulizations of RL only date back to the 1980’s with the work by Barto et al (1981) and Sutton (1988). While the popularity of RL has reached numerous fields in which complex computer models and artificial intelligence play significant roles, it has received minimal attention in the GIS and spatio-temporal modeling literature despite its potential for overcoming the limitations of other machine learning approaches based on supervised learning. Thus, this study examines the utility of RL for developing transition rules for
modeling LUCC processes. Specifically, an RL model for land cover change (RL-LCC) is developed to evaluate the economic and ecological goals involved in natural resource decision-making that results in changes to landscape cover. The model is implemented on a hypothetical forest management scenario in which forest management agents make decisions to maximize profits while minimizing ecological degradation. The model is evaluated using different RL learning parameters, different spatial extents, and varying levels of system complexity.

### 3.3 Methods

The RL-LCC is designed to suit a variety of processes involving the management of natural resources. The objective of the model is to simulate the extraction of resources by computer agents representing resource managers in a manner that balances economic and ecological goals. Agents, which are a central component to RL models, are programmed with a set of behaviours that allows them to evaluate their surrounding landscape and the actions of other agents in the model in order to make decisions that help them achieve their goals.

A comprehensive review of agent-based models of LUCC is provided by Parker et al. (2003) and Bosquet and Le Page (2004). The implementation of RL algorithms is intended to facilitate agent learning with regards to meeting these goals in a complex landscape with multiple agents. The RL-LCC model is composed of five main components: (1) RL-embedded agents and related goals, (2) the landscape in which the agents interact and make decisions, (3) the rewards that are assigned if agents are able to reach their goals, (4) the value function that uses the rewards to estimate the transition
rules, and (5) the *policy* that explains how agents will determine suitable transition rules. The following sections explain these components and the development of the RL-LCC model.

### 3.3.1 RL Agents and Goals

Agents in the RL-LCC model represent resource managers who are responsible for extracting an available resource over time from the land while abiding by certain principles or guidelines for how they can undertake management activities. Thus, each agent seeks to achieve a balance between two goals: (1) receiving high economic returns (i.e. *economic goal*), and (2) preserving specific ecological qualities of the landscape, such as habitat or a specific species of plant or animal (i.e. *ecological goal*). Any type of *ecological goal* is suitable as long as it can be formalized as a numeric *reward* (i.e. a reinforcement) signal that is passed from the environment to the agent. The *economic goal* of an agent is reached if, at the end of the simulation, the agent selects resources in a manner that allows it to meet or surpass a desired profit level. The *ecological goal* is achieved at the end of a simulation if a specific area of a given quality of land is preserved. Defining both goals will depend on how much of the resource the agents are permitted to extract each time step of the simulation and on the total number of time steps.

### 3.3.2 Landscape

The landscape represents the space in which the resource management agents interact with each other and make decisions. The landscape in the RL-LCC model is stored in a GIS database and is composed of multiple digital data layers, each containing
different types of information about the landscape referred to as attributes. Examples of attributes are the economic value of ecological sensitivity of the landscape. Each layer is a GIS raster data file which is composed of a rectangular matrix of cells. Individual cells contain a value expressing the quality of the attribute existing at that specific location. However, these layers are not directly used by agents when determining areas from which to extract resources. Instead, agents look at a *Value Layer* in which each cell contains a *value* that estimates the benefit of extracting the resource from a given location. Initially, the *values* of all cells in the *Value Layer* are set to 0.0 to indicate that agents do not possess a knowledge of the landscape. As a result agents randomly decide which cells to select.

If an agent is successful during a simulation in achieving its pre-defined goals, those cells that were selected for resource extraction will experience a positive reinforcement resulting in an increase in *value* in order to indicate to the agent that the resource at that location has positive qualities. Conversely, those cells selected during a simulation in which the goals are not achieved will experience a negative reinforcement and hence a decrease in their *value*, which will signify that such cells should be avoided. The cell *values* are updated at the end of each simulation based on information provided from the attribute layers. Figure 3-1 illustrates the landscape framework for the RL-LCC model. The landscape is composed of two layers – an *Economic Grid* representing the economic worth of the resource at each location, and an *Ecological Layer* explaining the sensitivity of the landscape. The agents observe the *Value Layer* when making decisions, in which the *values* are updated using the information from the other two GIS layers. Figure 3-1 also depicts how the landscape can be divided in order that different resource
management agents have access to a defined area. Landscape 1 is divided to accommodate the presence of three agents, while Landscape 2 is divided in order to accommodate six agents. The influence of the different landscapes on the RL-LCC is evaluated as described below.

Figure 3-1. The landscape for RL-LCC model is composed of the Economic Layer and Ecological Layer that are used to calculate if agents have achieved their goals, and the Value Layer that is assessed by agents to determine from which cells to extract resources.
3.3.3 Rewards

Rewards are numeric values that are used to determine by how much a cell value will increase when an agent achieves its goal. While there are numerous methods for determining rewards, the RL-LCC model uses a delayed reward (Sutton and Barto 2000) approach in which the goals are not evaluated until the final time step of each simulation. The main challenge with defining rewards for spatio-temporal models is dealing with the fact that there are a number of cells that can be selected during each simulation, thus each cell should receive a reward depending on its ability to help the agent achieve its goal. For example, an agent will receive a positive reward if it achieves both economic and ecological goals, and as a result the values of each of the selected cells will increase. However, it is logical that those cells with higher economic values and lower ecological sensitivity values should receive higher rewards. Conversely, if one or both goals are not met, those cells with the lowest economic value and the highest sensitivity values should receive the lowest negative rewards as they had a greater influence in preventing goal achievement. Therefore, instead of providing a single reward for all cells, formalizing the reward with spatio-temporal modeling requires the ability to provide scalar rewards based on the contribution to goal achievement or prevention for each cell.

The notion of relative rewards was implemented in the RL-LCC model using a three-step process in which the difference between the economic and ecological values for each selected cell is first measured, and then the Z-scores of the difference value is calculated in order to obtain a relative notion of how each cell influences goal achievement. A cell’s difference value, \( dv \), was calculated using the equation
\[ dv_c = a_c - b_c \]  \hspace{1cm} (3-1)

where \( a_c \) represents the economic value of a selected cell \( c \), and \( b_c \) represents the ecological value of cell \( c \). Thus, cells with higher \( dv \) values are better at facilitating goal achievement as they represent cells with high economic and low ecological values. The next step for formalizing the reward is calculating the Z-scores for each \( dv \) of the selected cells. This is performed using the standard Z-score equation

\[
Z = \frac{dv_c - \overline{dv}}{\sigma_{dv}}
\]  \hspace{1cm} (3-2)

The Z-score provides a standardized measure of the difference between an observation in a population and that particular population's mean. In the case of the RL-LCC model, the Z-score is the difference between a cell's difference value and the mean difference value of all selected cells. The objective of using the Z-score as part of the RL-LCC is to give each cell a value that depicts its benefit of being selected for extracting a resource relative to all other selected cells. That is, those cells with high positive Z-scores exhibit a greater relative benefit, while those cells with low negative Z-scores have a much lower relative benefit. Furthermore, the use of the Z-scores assumes that the values used to calculate the rewards at each time step are normally distributed. This assumption is considered valid when either the economic or ecological values of the cells constitute a distribution that is normal. The assumption is also considered true due to the search
procedures of the agents that involve selecting cells that agents know are most beneficial for achieving their goals as well as alternative cells that agents explore to determine if they can improve their goal-seeking decision making. Such searching behaviour results in a distribution of stands that contains the majority of economic and ecological values close to the mean and fewer values as the distance to the mean increases.

The final step was to calculate the reward values using the equations

\[
\begin{align*}
\text{if goal achieved} & \quad r = \frac{Z_{\text{min}} + Z_i}{|Z_{\text{min}}|} \quad (3-3) \\
\text{if goal not achieved} & \quad r = -1 \left( \frac{Z_{\text{max}} + Z_i}{Z_{\text{max}}} \right) \quad (3-4)
\end{align*}
\]

where \(Z_{\text{min}}\) and \(Z_{\text{max}}\) represent the minimum and maximum Z-scores, respectively, from that simulation. Equations 3-3 and 3-4 ensure that the reward is positive for all cells when the goals are met and negative for all cells when the goals are not achieved.

3.3.4 Value Function

The value function is an equation that uses the rewards to estimate the increase or decrease in cell value depending on whether the agents’ goals were achieved. The equation of the value function is

\[
V(c_t) \leftarrow V(c_t) + \alpha \left[ r + V(c_{t+1}) - V(c_t) \right] \quad (3-5)
\]
where \( V(c) \) and \( V(c_{t+1}) \) represent the value \( V \) of cell \( c \) at times \( t \) and \( t+1 \), respectively, and \( r \) represents the reward. The value function is based on the difference in a cell's value from successive time steps, which is termed the \textit{temporal difference} (TD) in the RL literature (Sutton 1988). This is calculated by updating the value of cell \( c \) at time \( t \) by subtracting it from the value of cell \( c \) at time \( t + 1 \). The method of using future values to update current values in a model is referred to as bootstrapping, which assumes that if cell \( c \) is a suitable cell to select at time \( t+1 \), then it is quite probable that it is also suitable to select cell \( c \) at time \( t \). However, this assumption is tested by repeating each episode a number of times that is deemed sufficient for allowing estimated values to converge towards true values. Without the incorporation of TD learning in the RL-LCC model, decisions made by agents to extract resources would only be evaluated based on the success of those decisions at specific moments in time. The link between current and future decisions would be removed, and agents would thus concern themselves with achieving goals only in the present instead of in the long term. Therefore, incorporating TD learning ensures that agents’ decisions at each time step are directed towards achieving goals both in the present and at the final time step of the model.

The parameter \( \alpha \) in equation 3-5 is the step-size parameter, which is a small positive fraction that influences the rate of learning. It is usually reduced over time so that changes to the value function are initially more significant than in later episodes (Sutton and Barto 2000). This represents the learning curve of the agents, as knowledge gained early has a significant impact, while knowledge gained after many episodes will provide only small additions to the agents’ overall knowledge base.
3.3.5 Policy

The policy in RL models defines the method for how agents will go about determining the transition rules that are most suitable for obtaining their goals. Specifically, it explains how agents will select cells based on their values in order to determine if they are beneficial or not with regards to extracting resources. It is logical for agents to choose cells that have the highest value as these will help them achieve their goals. This way the agents are exploiting their learned knowledge of the landscape. However, it is advantageous for agents to occasionally choose cells with values that are not the highest in order to explore different cells to see if agents can determine alternative sets of cells that are better for achieving their objectives. Thus, there is an important balance between exploitation and exploration.

Constant exploitation by selecting cells with the highest values is called a greedy policy, as the agent always takes the greedy action. This policy ensures that agents are doing their best at each time step to achieve their goals. However, it does not acknowledge the potential of other cells being able to facilitate goal achievement. Policies that explore every so often are called $\varepsilon$-greedy policies. In such policies the agents act greedy most of the time, but explore based on a random probability $\varepsilon$. The $\varepsilon$-greedy policies allow agents to search other cells that may eventually lead to higher rewards. With regards to spatio-temporal models where numerous potential outcomes exist, it is intuitive that agents should initially explore often in order to experience the many different cell combinations, and decrease exploration over time once cell values become reinforced. As a result, agents will learn which cells they should continually choose in order to best achieve their goals, thus establishing the transition rules of the system.
The pseudo code for explaining the process of the RL-LCC model is provided in figure 3-2. At each time step agents select a number of cells from which to extract resources. The number of cells selected must fall within a pre-determined limit. Once each agent has reached the allowable limit the cells' values are updated using the value function. At the final time step the reward is calculated and the values are again updated. After the final time step, the landscape is reset back to its initial state in which no resources are extracted; however the updated cell values remain in order to help the agents learn which cells to select in future simulations.

```
Set V(c) to 0.0 for all forest stands
Repeat for each simulation
  Repeat for n time steps
    for all agents
      while harvested resource < allowable harvested resource
        select cell based on \( \epsilon \)
        update value of \( V(c) \) for selected cells
  At time step = n
    for all agents
      if economic goal achieved and ecological goal achieved
        goals = achieved
      else
        goals = not achieved
    calculate \( dv \)
    calculate Z-scores
    calculate r
    update value of \( V(c) \) for selected cells
```

Figure 3-2. Pseudocode for describing the process of the RL-LCC.
3.4 Model Implementation: RL for Simulating Forest Management

The RL-LCC model was implemented using a hypothetical forest management scenario in which individual agents representing forest managers select forest areas to harvest based on their economic worth and ecological sensitivity. The model was tested using different search parameters across three spatial extents and with different numbers of agents in order to evaluate the flexibility of RL for defining system transition rules in a variety of scenarios. The following sections explain the model parameters and the specifics of each modeled scenario.

3.4.1 Agents and Goals

All agents in the RL-LCC model for forest management are assigned a designated Tree License Area (TLA) from which they can harvest trees. Each TLA contains a number of cells with a diversity of economic and ecological values. Each cell represents a forested area of similar trees, referred to as a forest stand.

The overall goal of each agent is to harvest a specified number of forest stands every time step in order to reach its economic goal of maximizing economic profit while reaching its ecological goal by minimizing the harvest of ecologically sensitive land. The economic goals are locally defined as they are independent for each agent. Conversely, the ecological goal is regionally defined, meaning that all forest management agents must collectively minimize the harvest of ecologically sensitive land for the entire forest region. This adds an additional element of complexity to the modeled scenario as decisions by one agent to harvest certain forest stands will have direct consequences on the decisions of the other agents. For example, if one agent harvests stands that are deemed highly ecologically sensitive, then the other agents must have harvested stands
that are relatively low in sensitivity in order to meet the regional *ecological goal*. As a result, the desire to achieve the *ecological goal* imposes a global constraint in the outcome of the model.

Goal achievement is evaluated at the end of each simulation based on agents’ ability to maintain or improve their stand selections from previous simulations. This was accomplished by measuring the summed difference of the *economic* and *ecological values* of all selected stands at the end of each simulation. A greater collective summed difference reveals that agents have been effective at selecting stands of higher economic worth and lower ecological sensitivity. The *collective difference*, $CD$, was calculated using the following equation

$$CD = \frac{\Sigma a_e}{k} - \Sigma B_e$$  \hspace{1cm} (3-6)

where $\Sigma a_e$ is the sum of *economic values* for stands harvested by a single agent, $\Sigma B_e$ is the sum of *ecological values* for stands harvested by all agents in the region, and $k$ represents a constant for standardizing the *economic* and *ecological values* to ensure that they are comparable. As a result of this process, the agents want to achieve a $CD$ value that is equal to or greater than previous simulations.

### 3.4.2 Forest Landscape

The artificial forest landscape is represented by two GIS data layers in which the smallest spatial unit is a forest stand covering an area of 1 ha. The first layer describes the *economic value* of each stand in dollars that describes the quality of timber that is available for harvest. The *economic value* of each stand in the dataset was defined by
randomly assigning a dollar amount from a normal distribution ranging between $0 and $100,000. The second layer represents the ecological value of each stand. For explanatory purposes, the values in this layer range from 1 to 5, where 1 indicates a stand that can be harvested without any concerns towards ecological sensitivity, while 5 represents a stand that should be of high priority for preservation because it contains important ecological characteristics. An ecological value was randomly selected for each stand by the model from a uniform distribution of values from 1 to 5; thus each stand had the same probability of receiving any of the five ecological values.

### 3.4.3 Rewards and Value Function

The rewards were assigned at the end of each simulation depending on whether an agent had achieved its goals. Rewards were calculated using the three step process explained above in which Z-scores were used. An agent only receives a positive reward if its calculated CD value is equal to or larger than the value of CD from previous simulations, referred to as \( CD_{\text{max}} \). Conversely, an agent receives a negative value if its calculated value of CD is lower than previous simulations. The value function presented in equation 3-5 was used to update the values of harvested stands at the end of each simulation. The step-size parameter \( \alpha \) in equation 3-5, which controls the rate of learning, was defined by

\[
\alpha = \frac{0.1}{S} \quad \text{(3-7)}
\]
where $S$ represents the current simulation number. The numerator value of 0.1 ensures that updates to stand values remain relatively small compared to the stand values themselves. The presence of $S$ ensures the successive updates are incrementally decreased over time, demonstrating that the agent is learning more information during the initial stages of the model.

### 3.5 Results

#### 3.5.1 RL-LCC Algorithm

The algorithms governing the RL-LCC model were implemented using Python programming language (Python Software Foundation 2007) in a middleware application called Agent Analyst (Argonne National Laboratory 2006). This program provides the ability to program dynamic modeling procedures using agent-based simulation libraries from REPAST (North et al. 2006), an agent-based modeling package. These procedures are then linked to ArcGIS 9.0 (ESRI 2004), which provides the storage and retrieval capabilities for spatial datasets. At each time step of the model, the procedures written in Agent Analyst access the spatial datasets, perform calculations based on the RL algorithms, and then write the information to the spatial datasets in order to update the stand values for each cell.

#### 3.5.2 Initial Scenario

The initial scenario modeled by the RL-LCC model used a 20 x 20 cell GIS layer representing an artificial landscape of forest stands which was divided into three separate Tree License Areas (TLAs), each managed by a separate forest management agent (see
Landscape 1 in figure 3-1). The agents were responsible for selecting a single 1ha forest stand at each time step of the model for a total of ten time steps in order to represent a suitable time frame for forest management. The selection of forest stands was not constrained by any spatial criteria such as the need to harvest adjacent stands or stands that are a specified distance from each other. The initial scenario employed a greedy policy, which ensures that agents are continually choosing forest stands with the highest cell value. The model was run for 1,000 simulations in order for the agents to gain substantial experience so they are able to learn to develop transition rules that will lead them towards goal achievement by using the RL algorithms presented in the methods section. To ensure the model was enabling agents to correctly evaluate and select appropriate forest stands, the model was calibrated using a simplistic representation of the forest landscape in which the most beneficial stands for achieving agent goals were known a priori. This was accomplished by randomly selecting ten stands from each TLA and assigning maximum economic values and minimum ecological values, while the remaining stands were given less desirable values. Thus, the model was considered acceptable when the three agents all selected the ten stands in their TLA that were predetermined as most beneficial. Model calibration was performed by altering the number of previous simulations over which an agent had to maintain or surpass the $CD_{max}$ parameter in order to receive a positive reinforcement. The number of previous simulations for maintaining or surpassing $CD_{max}$ was altered using values from 1 to 10. The calibration revealed that a value of 5 allowed the agents to select the most beneficial stands in the shortest number of simulations.
Figure 3-3 presents the results from the original scenario with regards to how cell values change over numerous simulations as the agents learn which stands are better to select. The figure also explains the relationship between stand values and the economic and ecological values. Initially, all stand values are set to 0.0 as the agents have yet to experience the benefit of selecting them. However, over numerous simulations, those stands that are beneficial for goal achievement are continually selected and receive positive rewards, which in turn increases the stand’s value. As this simulation employs a greedy policy, there are minimal stands that are selected as beneficial for goal achievement because agents are not exploring alternative stands. The learning curve of each agent with regards to their ability to obtain high $CD$ values is presented in figure 3-4. The graph illustrates that agents initially make poor decisions, but over numerous simulations they learn which stands to select that will provide them with higher values of $CD$. At the completion of 1000 simulations, the agents have developed a set of transition rules that utilize the stand values to determine which stands are harvested at each time step.

3.5.3 Sensitivity to Agent Policies

Three different levels of policies were tested in order to evaluate the sensitivity of the model to various searching parameters. The greedy policy was evaluated in order to determine how agents were able to develop effective transitions rules while only selecting those stands with the highest values. In addition, two $\epsilon$-greedy policies were implemented to evaluate how exploration facilitates transition rule development. Exploration was constantly performed for both policies for a specific period of time and then reduced during each simulation. This ensures that agents initially always select random stands in
order to fully explore different stand combinations, and then over time focus on those stands that are most beneficial.

For the first $\varepsilon$-greedy policy (i.e. low exploration policy), exploration was constantly performed at each time step for the first ten simulations in order to represent a short period of constant exploration, then reduced to occur with probability

$$\varepsilon = \frac{1}{S - 10} \quad (3-8)$$

where $S$ is the current number of simulations performed. Exploration with the second $\varepsilon$-greedy policy (i.e. high exploration policy) was constantly performed at each time step for the first hundred simulations in order to represent a long period of exploration, then reduced to occur with probability

$$\varepsilon = \frac{1}{S - 100} \quad (3-9)$$
Figure 3-3. The results from the original scenario in which the cell's values increase over time if they are beneficial for harvesting in order to achieve the defined objectives. Cell values are shown in relation to the cell’s economic and ecological values.
Figure 3-4. The learning curve of each agent with regards to obtaining high CD values. The dotted vertical line represents the moment when agents established transition rules that allow them to select specific stands continually to achieve their goals.

The policies were evaluated based on how they facilitated goal achievement for three different harvesting levels: (1) one stand per time step, (2) five stands per time step, and (3) ten stands per time step. Figure 3-5 illustrates these results by displaying the average CD value of all agents at each simulation for the different policies. The graph in figure 3-5a explains that the greedy policy was more effective for defining transition rules leading to the highest CD values. However, as the harvesting level increases to five (figure 3-5b) and ten (figure 3-5c), the greedy policy initially does better, but eventually becomes the least effective of the three policies for increasing CV values. Conversely, the ε-greedy policies start off poorly but eventually lead to higher returns.
3.5.4 Sensitivity to Spatial Extent

Three levels of spatial extent of the forest landscape were chosen by doubling the numbers of rows and columns of cells. Thus, the initial layer of 20 x 20 cells (i.e. the small spatial extent) was doubled to form the medium spatial extent (40 x 40 cells), which was doubled resulting in the large spatial extent (80 x 80 cells).

The results from transferring the RL-LCC model to different spatial extents is presented in figure 3-6. These results depict simulations in which the harvesting level was set to ten stands per time step and using the high exploration policy. Figure 3-6 reveals that the RL-LCC model is easily transferable to different spatial extents as agents are still able to learn how to select those stands with high economic values and low ecological values. Furthermore, as the spatial extent increases, agents have a greater selection of stands from which to choose, which in turn decreases the number of stands with intermediate economic and ecological values compared with what is observed from the smaller spatial extents. As a result of greater selection, agents can expect to have higher returns of CD values.

The observation of higher returns from larger spatial extents is emphasized in the graphs in figure 3-7, which depict the average CD values of all agents for different policies using the various spatial extents. Figure 3-7a reveals that there is minimal difference in the CD values between the different policies when the spatial extent is relatively small. However, as the spatial extent increases from the medium (figure 3-7b) and large (figure 3-7c), it is apparent that policies that invoke greater periods of exploration are required for achieving maximum returns. This is logical as datasets with a greater number of stands will require agents to search for longer periods in order to
determine which stands are most beneficial to harvest due to the increased number of potential stand combinations.
Figure 3-5. The average CD value for all agents for the different policies with harvesting levels of a) one stand per time step, b) five stands per time step, and c) ten stands per time step.
Figure 3-6. The results from the Spatial Extent scenarios in which agents harvest ten stands per time step. Cell values are shown in relation to the cell's economic and ecological values.
Figure 3-7. The average CD value for all agents for the different policies using a) 20 x 20 cell spatial extent, b) 40 x 40 cell spatial extent, and c) 80 x 80 cell spatial extent.
3.5.5 Sensitivity to Number of Agents

Simulations were performed with six agents in order to evaluate the RL-LCC model’s ability to incorporate additional complexity. The landscape, which is divided into six Tree License Areas (TLAs), is presented in figure 3-1 (Landscape 2). The simulations were performed using the 80 x 80 cell GIS layer and the policy with the high exploration policy, and each agent was responsible for harvesting ten forest stands at each time step. The largest spatial extent was chosen in order to ensure that the evaluation of increasing agents in the model was not affected by the availability of suitable stands to harvest.

The results from the six-agent simulations are presented in figure 3-8, which depicts the stand values in relation to the economic and ecological values. The figure reveals that a greater amount of stands are now being considered beneficial for harvesting, and the majority of select stands have intermediate to high economic values or intermediate to low ecological values. These two observations are due to the fact that there are now six agents selecting ten stands during each time step. Therefore, more stands are being selected in order for all agents to achieve their goals, including those with relatively lower economic or higher ecological values. This leads to an overall decline in the average CD values of each agent at the end of the simulations because each agent has fewer stands from which to select. This can be observed from the average CD values presented in figure 3-9 for the different policies, which are lower than the three-agent simulation using the same policy and harvesting level (figure 3-7c). Figure 3-8 also reveals that policy selection is more important as the complexity of the scenario is increased. The greedy policy, which initially gives higher returns, fails to improve, and as a result performs poorly compared to the two ε-greedy policies. The high exploration
policy takes the longest to obtain its maximum $CD$ values, but the increased exploration eventually leads to the highest returns.

Figure 3-8. The results from the Agent scenario in which six agents harvest ten stands per time step. Cell values are shown in relation to the cell's economic and ecological values.
3.6 Conclusion

The results from the RL-LCC model reveal how reinforcement learning can help construct spatio-temporal models of land cover change, especially with regards to forest management. The RL algorithms established transition rules that lead to goal achievement by clearly indicating which stands are most beneficial for harvesting with regards to their economic and ecological characteristics. The output maps depicting stand values across the forest landscape are linked to attribute tables in the GIS that provide management with relative indicators of how each stand represented as cells in the GIS database contributes towards goal achievement.

The ability to integrate multiple landscape attributes existing at different scales (i.e. locally-defined economic goals and regionally-defined ecological goals) is a prominent advantage of the RL-LCC in contemporary management scenarios in which...
harvesting is constrained by many factors. In addition to the economic and ecological requirements, agent goals could also focus on spatial constraints for decision-making such as harvesting adjacent stands or only harvesting stands that are a given distance from each other. In fact, any goal that can be formalized as a numeric reward is suitable for an RL model.

The fact that RL develops transition rules based on the experience of agents in a landscape rather than on a set of training data or explicitly-defined rules allows it to be more flexible for simulating LUCC. This aspect of RL facilitated an increase in system complexity through the addition of agents by simply altering the policy to one that provides for further exploration. Furthermore, the ability to transfer the model seamlessly across different spatial extents suggests that RL can adequately address issues of scale and resolution that often impinge on spatio-temporal modeling approaches.

The integration of GIS and RL proves to be beneficial as RL algorithms can easily access information stored in spatial datasets in order to determine how to simulate LUCC over time, while GIS can provide a variety of spatial analytical procedures for use in developing RL parameters such as goals and rewards. The ability of RL and GIS to facilitate in the simulation of complex spatial processes is of particular importance to contemporary forest management around the world as demands for forest products increase with a growing global population coupled with dwindling resources. Different types of objectives that are specific to various study sites can be implemented in RL models to determine how stakeholders should make decisions in order to achieve their goals that may be in conflict with each other and perhaps exist at varying spatial scales. Furthermore, the potential for RL-GIS integration extends beyond forest management, as
modeling procedures can adapt to evaluate the use of agricultural, aquatic and mineral resources. In fact, the RL methods presented in this study can be applied to any system in which resource managers are required to balance multiple objectives when determining how to allocate elements of the landscape. The development of an RL model for land cover change only requires that decision making behaviour can be encoded into computer agents, and that agent actions can be rewarded. As a result, RL has the potential for becoming and instrumental approach in helping to determine how to integrate the many needs of resource management by providing an understanding of how different goals existing across space can be integrated to provide suitable management decisions over time.
3.7 References


4. GIS AND INTELLIGENT AGENTS FOR MULTIOBJECTIVE NATURAL RESOURCE ALLOCATION: A REINFORCEMENT LEARNING APPROACH

4.1 Abstract

An important component of natural resource management is determining how to allocate resources within a landscape to different stakeholders in a manner that satisfies multiple objectives. Developing decision making tools for assisting natural resource allocation is a challenging endeavor as stakeholders’ objectives typically exist at varying spatial scales, their actions are defined by the spatial constraints in which they operate, and the spatial distribution of resources can be altered due to system disturbances. The nature of such challenges suggests the need for a geographic approach that can investigate these spatial complexities in order to generate a suitable set of solutions.

The objective of this study is to develop and evaluate an Intelligent Agent Model for multiobjective natural resource allocation. The model integrates agent-based modeling in a GIS environment with reinforcement learning – a heuristic method for generating, evaluating, and improving solutions for multiple objective decision making.

The model is implemented by simulating a forest management scenario in which agents

---

that represent forest companies learn how to harvest trees in a manner that maximizes economic return while minimizing the adverse ecological impact to the surrounding landscape. In addition, the model simulates forest disturbances of varying frequencies and intensities to determine how disturbance events affect the decision-making ability of agents. The model is validated to demonstrate that it can provide practical solutions to natural resource decision making.
4.2 Introduction

Natural resource allocation is a process in which decision makers develop solutions that incorporate multiple objectives for determining how resources should be allocated to stakeholders across space and time. Geography plays a central role in the development of such solutions as it is imperative to understand the spatial distribution of objects that are of importance to management initiatives and the spatial scale at which different objectives exist. A geographic perspective reveals how landscapes are formed by complex histories that lead to non-linear, dynamic spatial relationships between natural resource entities, how stakeholders' objectives are often conflicting in intent and scale, and how their actions are influenced by the spatial structure of the landscape. However, incorporating these spatial complexities into the decision-making process presents a significant challenge that requires sophisticated methods for providing practical solutions.

GIScience offers a wide breadth of methods for addressing the needs of natural resource allocation. Geocomputation through the coupling of GIS and agent-based modeling (ABM), for example, is a suitable approach for modeling complex geographic processes in order to inform spatial decision-making (Malczewski 2004). GIS lends a database structure and analytical tools required for spatial analysis, while ABM provides a dynamic component to the otherwise static nature of GIS (Brown et al. 2005) by encoding computer agents with the ability to capture the spatial decision-making behaviours of individuals, households or institutions (Parker et al. 2003). Agents can represent either interacting mobile individuals (Deadman and Gimblett 1994, Batty 2001, Torrens and Benenson 2005) or stationary individuals that make decisions affecting land
use (Hoffmann et al. 2002, Lambin et al. 2003, Evans and Kelley 2004), but they should exhibit autonomous behaviour, share an environment through an agent community and behave in a manner that will enable them to reach their objectives (Bousquet and Le Page 2004). Due to their goal-oriented structure, ABMs provide considerable potential for informing natural resource decision making as multiple objectives of different agents can be represented at different spatial scales (Ducrot et al. 2004, Castella et al. 2005).

However, the strength of ABM for geocomputation is primarily focused on understanding the emerging patterns that arise from the interactions of agents in a landscape. Agents are typically provided with a set of behaviours that are implemented in response to different states of the landscape or in response to the actions of other agents. As such, traditional ABM for geocomputation can inform decision makers about the types of results that can be expected from certain behaviours and interactions, but lack a substantive ability to evaluate and improve agent behaviour in order to address multiobjective decision making.

Multiobjective decision making has received much attention in the GIScience literature through the use of heuristic algorithms (Diamond and Wright 1988, Brookes 1997, Cromley and Hanink 1999, Cova and Church 2000a, Cova and Church 2000b). Among of the most prominent heuristic approaches employed are those belonging to the class of evolutionary algorithms (EA), which, as Diplock (2000) explains, includes the broad categories of genetic programming (Koza 1992) and genetic algorithms (Holland 1975). Following Darwinian evolution principles, solutions in an EA-based model are considered individuals in a solution population, each having a fitness score depicting its suitability for solving a particular problem (Krzanowski and Raper 2001). A solution
consists of a sample of areas within the landscape that is selected for some particular purpose (e.g. resource extraction). Solutions evolve over numerous runs of the model until a near-optimal solution is reached. The fitness scores of the evolved solutions provide a ranking of relative optimality that indicates how suitable a solution is for solving the problem of interest. While EAs present significant potential due to their ability to integrate multiple objectives to determine optimal land use patterns (Zhou and Civco 1996, Brookes 2001, Xiao et al. 2002, Ducheyne et al. 2006, Mooney and Winstanley 2006), an underlying problem exists with EA approaches and other heuristic algorithms in that solutions are typically generated and evaluated at the level of the entire landscape. Such methods do not consider that different stakeholders could govern different jurisdictions within the landscape, and that these jurisdictions possess their own spatial and non-spatial characteristics that shape stakeholder actions and, consequentially, the effects these actions have on other stakeholders. Furthermore, all areas that belong to a solution are evaluated equally with no attention given to how each individual area contributes towards the achievement of different objectives. This leads to a loss of detail as there is no information provided regarding the relative benefit of performing activities in one area versus another. However, these challenges can be addressed by incorporating elements of independent decision making such as those found in ABM in the solutions derived from heuristic algorithms.

The objective of this study is to present and evaluate an approach to multiobjective decision making for natural resource allocation that integrates concepts from ABM and heuristic algorithms. A model is developed in which agents representing individual stakeholders have their actions evaluated by algorithms based on reinforcement learning
(RL) - a heuristic method that is used to reward decisions made by individual agents that leads towards achieving specific objectives. Each agent explores its jurisdiction within the landscape and takes an action to extract a resource. The evaluation of actions is recorded as an attribute of the landscape, and is used to help agents learn over time the most beneficial locations for extracting resources. While the notion of "intelligent agents" has previously been proposed in the GIScience literature (Bennett et al. 1996, Manson 2005, Bennett and Tang 2006), this study extends this idea for the explicit use of multiobjective decision making for natural resource allocation. Intelligent agents for multiobjective decision making can potentially incorporate numerous objectives and implement a diverse range of agent behaviour, but the model developed in this study is focused on agents acting to maximize their own economic returns, and interacting in an indirect manner to achieve ecological objectives that exist over larger geographical areas (i.e. at the global scale). The model is evaluated on its ability to provide solutions for achieving multiple objectives and for the ability of agents to incorporate and adapt to random system disturbances. The model, called the Intelligent Agent Model (IAM), is implemented in the context of forest management on a study site in southwestern British Columbia, Canada.

4.3 Intelligent Agent Model for Multiobjective Forest Management

The theoretical foundations of RL evolved during the second half of the twentieth century as various elements were implemented in control theory, pattern recognition, trial-and-error learning, and classifier systems. However, the structure of RL as currently described in the literature was formalized by Barto et al. (1981) and Sutton (1988). While
minimal research on the use of RL for spatial applications exists, the utility of RL for natural resource allocation has yet to be explored. In the development of IAM for forest management, RL algorithms are used to determine when and where numerous forest management agents should harvest trees to meet their economic expectations while retaining certain ecological aspects of the forests. Agents in IAM do not possess \textit{a priori} information regarding the landscape, nor are they provided with information concerning rational behaviour that states how to act when presented with a given state of the landscape. Instead, agent actions are governed by a balance between exploring the landscape to find improved solutions and exploiting learned information for ensuring that beneficial actions are maintained. While stakeholders in the forest management process do not possess this type of cognition, the agents in the developed heuristic model are embedded with this particular learning behavior as a means to explore extensively large spatial datasets in which enumerable potential harvesting solutions exist.

Forests are typically classified into forest stands that are defined as an area of trees of similar age and species. As such, agents are concerned with harvesting a set of forest stands that will help them achieve their objectives. The agents select a number of stands in which to harvest a pre-determined volume of trees at each time step of the model for a number of time steps that represents a given period. Once a stand has been selected, the trees are harvested and the stand is no longer available for harvesting for the remainder of the period. At the end of the period, the model evaluates the set of harvested stands based on how close it satisfies the multiple objectives. Each stand in the set receives a positive reward if the solution to the multiobjective problem is maintained or improved, which is then used to update the quality of the stand. Quality is a stand
attribute that informs the agent how beneficial it is to harvest the trees in the stand in order to satisfy the objectives. The *quality* of each stand is initially 0.0 to represent that the agents do not possess any prior knowledge of the landscape. Lacking any knowledge of the landscape means that agents will initially explore the landscape by making random stand selection. The rewards are calculated at the end of the process and the *quality* of harvested stands is updated to reflect the outcome of the agents’ decisions. The process is repeated numerous times; each time the landscape is updated to better reflect the true *quality* of the stands and to improve the ability of the agent to make informed decisions.

The time period representing a single simulation of the process (i.e. from time step 1 to the terminal time step) is referred to as an episode. Through the simulation of numerous episodes, agents rely less on exploring the landscape through random stand selection and begin to exploit their knowledge of the landscape by selecting those stands that possess a relatively high *quality*. By the final episodes the agents are able to select those sets of stands that are most beneficial for achieving their objectives. Furthermore, the balance between exploration and exploitation facilitates a substantive search through the list of potential solutions, which can consequentially provide important information regarding the tradeoffs between the different objectives.

Similar to other RL models described in the literature (Sutton and Barto 2000), IAM consists of four main components: (1) the agents and their objectives, (2) the numeric *rewards* used to reinforce positive actions that lead to objective achievement or demerit actions that lead to undesirable outcomes, (3) the *quality function* that updates stand *quality*, which is used to facilitate agent learning, and (4) the *search function* that
defines how agents are to search through the database. Each component is described in
detail below.

4.3.1 Agents and Their Objectives

Each agent in the model represents a single forestry company that harvests trees
with the intent of sustaining profits. The agents are required to manage a defined area of
forest called a Forest Management Unit (FMU) by extracting a volume of timber on an
annual basis. The amount of timber that agents are permitted to harvest at each time step
of the model is specified by an Allowable Cut Level (ACL), which is determined by the
total volume of timber that is available in the agent’s FMU. As some jurisdictions impose
an ACL for a given time period, a single time step in the model represents that defined by
the ACL. The primary objective of each agent in IAM is to receive maximum economic
returns from harvesting, referred to hereafter as their economic objective. Forest stands
are usually associated with an economic value that represents the financial worth of the
timber in the stand. In order to receive positive reward, the agents must balance their
economic objective with minimizing the adverse ecological impact on the landscape.
IAM focuses on harvesting stands with minimal tree age and minimizing the area over
which timber is harvested. These are represented by an age objective and an area
objective, respectively. While the economic objective is locally defined, the age and area
objectives are defined at the regional level. This means that agent rewards are also
dependent on their ability to minimize the average age of all harvested stands and the
sum of area that is harvested. The existence of the age and area objectives represents a
global pursuit in which one agent’s decisions has a direct consequence on all other agents.
While traditional agent models typically involve explicit agent interaction that influences
behaviour, the regional-level objectives in IAM ensure that agents indirectly interact with each other.

4.3.2 Defining Rewards

At the end of each episode, the model evaluates the stands harvested by all agents (i.e. the solution) in comparison to all previous solutions. The model assigns a positive reward to all harvested stands if the solution is at least as good as the previous best solution. That is, a positive reward is assigned if the solution meets the following reward conditions: (1) the sum of economic values of all harvested stands in each agent's FMU is greater than or equal to sum of economic values of all harvested stands in their FMU from previous best solution, (2) the harvested area for the entire study site is less than or equal to the harvested area for the entire study site from previous best solution, and (3) the average age of harvested trees for the entire study site is less than or equal to the average age of harvested trees for the entire study site from previous best solution. A negative reward is assigned if at least one of the conditions is not satisfied.

As harvested stands contribute differently towards objective achievement or prevention, each stand receives a positive or negative reward that is a function of this contribution. Variable rewards were implemented by developing a three-step process in which the three variables that define each objective (i.e. economic value, age and area) are standardized, combined and converted into a numeric reward. First, in order to standardize the variables for each harvested stand, the z-scores were calculated for the economic value \( v \), age \( a \), and area \( b \) using the equations
where $\mu_v$ is the mean economic value and $\sigma_v$ is the standard deviation of all harvested stands in the FMU from which the stand resides;

$$Z_v = \frac{v - \mu_v}{\sigma_v}$$  \hspace{1cm} (4-1)

where $\mu_a$ is the mean age and $\sigma_a$ is the standard deviation of all harvested stands in the entire study site;

$$Z_a = \frac{a - \mu_a}{\sigma_a}$$  \hspace{1cm} (4-2)

where $\mu_b$ is the mean area and $\sigma_b$ is the standard deviation of all harvested stands in the entire study site. The second step is to combine these z-scores to form a single value $z_t$, for each stand, which is calculated by

$$Z_t = Z_v - Z_a - Z_b$$  \hspace{1cm} (4-4)
The justification behind this formula is that a harvested stand with the highest positive z-score for economic value \( z_v \), and the lowest negative z-scores for age \( z_a \) and area \( z_b \) is the most beneficial in the group of harvested stands for facilitating objective achievement. Thus, the equation ensures that such harvested stands have relatively higher values of \( z_t \). The equation also ensures that all three objectives are weighted equally. That is, no objective is considered more important than the others when estimating stand quality. Weights could potentially be assigned by decision makers either prior to running the model or intermittently during model simulation. However, these two methods, referred to as prior articulation and interactive articulation, are problematic in that it can be challenging to agree upon a reasonable weighting schematic, and weighting the objectives may overlook some solutions that are successful in satisfying the objectives (Xiao et al. 2007). Instead, the model employs equal weighting to each objective that should be observable in the results. Furthermore, the model provides information that allows decision makers to view the tradeoffs between the different objectives in order that they can determine appropriate weighting for future runs of the model.

The final step uses \( z_t \) to assign either a positive reward (i.e. when all reward conditions are satisfied) or negative reward (i.e. when at least one reward condition is not satisfied), which is performed by the equations

\[
\text{if positive reward, then } r = \frac{\min z_t + z_t}{\min z_t} \quad (4-5)
\]

\[
\text{if negative reward, then } r = -1.0 \left( \frac{\max z_t - z_t}{\max z_t} \right) \quad (4-6)
\]
where min$z_t$ and max$z_t$ are the minimum and maximum $z_t$ values in the set of all harvested stands.

### 4.3.3 Updating Stand Quality with the Quality Function

As mentioned above, the *quality* of each stand in the forest is initially set to 0.0 to represent the fact that agents do not possess knowledge of which stands are most beneficial for harvesting. Over the course of numerous episodes, the *quality* $Q$ of stands is updated using the *quality function*, which has been derived from Sutton and Barto (2000). For this function, let $Q$ represent the *quality* of stand $s$ when selected at time step $t$, and let $k-1$, $k$, and $k+1$ represent the previous, current and next episode that stand $s$ was or will be chosen at time $t$, respectively. The *quality function* is thus expressed as

$$Q(s_{t,k+1}) \leftarrow Q(s_{t,k-1}) + \alpha \left[ r + Q(s_{t,k}) - Q(s_{t,k-1}) \right]$$ (4-7)

where $r$ is the reward and $\alpha$ is a positive fraction called the *step-size parameter* that controls the rate of agent learning. The *quality function* specifies that the *quality* of stand $s$ at time $t$ in the next episode, $k + 1$, is based on the temporal difference between the *quality* of stand $s$ at time $t$ from the current episode, $k$, and previous episode, $k-1$. As such, the temporal difference represents a measure of the error in the estimation of stand *quality*, which should decrease over the course of the model as the *quality* of stands becomes more representative of its true value. The *step-size parameter* is used to influence agent learning as it is typically represented by a small fraction that decreases as the number of episodes increases. This will cause the initial updates of a stand's *quality* to
be relatively large as the agent is reaching its full learning potential due to the fact that it has no prior knowledge. As the number of episodes increases, the agents potential for learning new information about the landscape diminishes, therefore the updates to stand quality is decreased.

4.3.4 Search Function

The knowledge learned from each episode increasingly provides agents with information in the form of stand quality that facilitates stand selection. If agents constantly select those stands with the highest quality, they are said to be exploiting their knowledge, or implementing a greedy search function. While this type of search function is logical, it may actually prevent agents from finding better solutions to their problems as they may continually select a small sample of stands and as a result converge towards poor quality solutions. To solve this problem, agents should also explore stands that do not have the highest quality, as such actions could potentially lead them down a path with more beneficial outcomes. In order to implement exploration, agents can possess an $\epsilon$-greedy search function in which stands that do not have the highest quality are selected with probability $\epsilon$. An $\epsilon$-greedy search function is implicitly linked to the step-size parameter because of the relationship between exploration and learning. That is, agents should be learning more information about the landscape during periods of exploration, and learning will diminish as agents begin to exploit their knowledge. As a result, it is logical that the $\epsilon$-greedy search function and the step-size parameter diminish in the same fashion over the course of the model. Furthermore, searching and learning is dependent on the size of the dataset and the complexity of scenario. This is because larger datasets and more complex scenarios will inevitably lead to more potential solutions. Therefore,
more search time is required and as well as a longer period of learning. Therefore, the \( \epsilon \)-
greedy search function and the step-size parameter should be defined through model
calibration to ensure that each stand in the landscape is sampled a significant number of
times in order to determine if it is beneficial for harvesting with regards to achieving the
different objectives.

4.3.5 Incorporating System Disturbances

Disturbance events such as those caused by forest fires or insect outbreaks were
simulated in IAM at random time steps to determine how the agents respond to novel
situations in which all trees in different stands are eliminated. The strength of the
disturbance is determined by frequency, \( f \), and intensity, \( i \), which are kept constant during
each simulation in order to evaluate agent learning under different disturbance regimes.
Frequency defines how often a disturbance occurs, which can range from as low as one
disturbance per episode to as much as one disturbance per time step. Intensity determines
the number of forest stands that are affected by a disturbance. The probability of a
specific stand being randomly selected to experience a disturbance is a function of its
economic value. This represents the notion that higher value trees are either older or more
susceptible to forest disturbances such as wind or insects. For explanatory purposes, the
probability of stands experiencing a disturbance was defined by a linear function based
on the average price of a stand. A probability of 0.0 is associated with the minimum stand
economic value, and a probability of 0.9 is associated with the maximum stand economic
value. The susceptibility value of 0.9 was chosen to ensure that the same stands were not
continually selected for disturbance. The model will select stands to experience
disturbance until the number selected is equal to \( i \). Agents will then have to learn which
remaining stands are most beneficial for harvesting to meet their objectives. The goal of incorporating system disturbances here is not to recreate a specific forest disturbance, but rather to determine if the agents are able to provide logical solutions in the presence of random disturbances, and, if so, how these solutions differ from the solutions in which disturbances are absent.

### 4.4 Model Implementation

IAM was implemented on a study site located in the Chilliwack Forest District in southwestern British Columbia, Canada. The district covers an area of 1.4 million ha in which the Government of British Columbia determined an Allowable Cut Level (ACL) at 1,270,000 m³/ha. The study site covers an area of 127,927.1 ha, which was previously divided by government legislation into three Forest Management Units (FMUs) as shown in figure 4-1. The area is represented by a GIS vector data structure in which forest stands are encoded as individual polygons. Each forest stand contains information on tree volume, average tree age, and the total area of the stand as summarized in table 4-1. This information was used in conjunction with the region’s ACL to calculate the initial ACL for the FMUs. The price value of each stand was calculated by combining information on the species composition and volume within the stand with information on the current pricing of timber as provided by British Columbia timber market reports (Government of British Columbia, 2007). The volume and price of each forest stand was increased on an annual basis to represent forest growth. This was implemented by increasing volume by 1% each time step, followed by a recalculation of the economic value of the forest stand.
Figure 4-1. Study site in southwest British Columbia, Canada.
A sigmoid curve was selected to govern the rate of decline in the $\epsilon$-greedy search function and the step-size parameter in order that agents initially explore for an extended period of time, followed by a gradual decline towards full exploitation. Altering the slope of the curve was used as a method for calibrating the model. A more gradual slope ensured longer periods of exploration, but also resulted in longer periods of learning that were not useful in distinguishing between beneficial and non-beneficial stands as stand quality was relatively homogenous. Conversely, a relatively steep curve ensured distinctive values of stand quality, but often lead the model to converge towards less desirable solutions. As a result, a moderate slope was selected to govern the $\epsilon$-greedy search function and the step-size parameter, which was defined by the equation

$$\epsilon, \alpha = 1 - \left[ 1 + e^{-\frac{1}{\alpha 0.001 (k - 10,000)^2}} \right]$$  \hspace{1cm} (4-8)$$

where $k$ is the current episode in the model. This equation, coupled with a 10,000 episode simulation, ensured that each stand was sampled at least fifty times through the course of
the model. While it is difficult to estimate a sufficient range for sampling, the results generated from equation 4-8 proved to be suitable for providing adequate solutions.

4.5 Simulation Results

Agent decision-making and the calculation of spatial forest characteristics were encoded in procedures developed in the Python Programming Language (Python Software Foundation 2007). Agent behaviour was programmed to allow agents to search a GIS database and make decisions regarding harvesting forest stands based on their objectives. The Python procedures were implemented in Agent Analyst (Argonne National Laboratory 2006), which is a middleware that links the spatial database structure of ArcGIS (ESRI 9.1 2004) with the agent-based modeling package REPAST (North et al. 2006). The temporal modeling libraries of REPAST were used to update the spatial data stored in ArcGIS’s vector shapefile format. Figure 4-2 represents the procedural code that was used for developing IAM. For the initial forest management scenario, IAM was performed in the absence of disturbance to evaluate the ability of RL-embedded agents to adequately determine a solution for harvesting trees. Additional scenarios were then introduced in which the frequency and intensity of disturbances were incrementally increased, and the resulting change in the quality of each stand was recorded.

4.5.1 Initial Forest Management Scenario

The results from the initial scenario without disturbances are presented in figure 4-3. The figure shows the map of the study site with stands represented by their quality for year 10 of the final episode of the model. The quality of forest stands is distributed
between values of -10.0 to 10.0, where a quality closer to -10.0 demonstrates that a stand is an extremely poor selection for achieving the three objectives, while quality closer to 10.0 represents stands that are extremely beneficial. The stands with the highest quality for each time step that were within the limits of the Allowable Cut Level were selected to form the solution for maximizing economic returns and minimizing the average age and total area of harvested trees. These stands constitute the Quality Solution, which is demonstrated for each agent in figure 4-4, provides a collective economic return of $112,249,427, an average harvested tree age of 77 years, and a total harvested area of 2,185 ha.
Set $V(c)$ to 0.0 for all forest stands
Repeat for each episode
    Repeat for 10 time steps
        for all agents
            while harvested volume < ACL
                select forest stand based on $\varepsilon$
                update value of $V(c)$ for harvested stands at time step -1
    At time step = 10
        for all agents
            if economic return > economic goal
                if harvest average age < goal age
                    if harvested total area < goal area
                        goal = achieved
                    else
                        goal = not achieved
            calculate Z-scores
            calculate $Z_t$
            calculate $r$
            update value of $V(c)$ for harvested stands at time step = 10

Figure 4-2. Pseudo code representing the process of IAM.
Figure 4-3. Results from the initial IAM simulation.
Figure 4-4. Selected stands for each Forest Management Unit (FMU) are those with the highest stand quality that fall within the Allowable Cut Level.
The process of agent learning is presented in figure 4-5, which depicts how agents improved their decision-making over time to maximize economic returns while minimizing the adverse ecological impact on the region. The graphs show a distinctive improvement in learning after approximately 5000 episodes. This is the period during the simulation of the model that agents begin to engage more in exploiting their knowledge rather than exploring forest stands in order to improve their decision making behaviour. Prior to this, agents are still learning which stands are relatively more advantageous than others to harvest, but their reliance on exploration does not result in any significant improvement in achieving their objectives. Conversely, the increase in exploitation after 5000 episodes results in agents increasingly selecting more beneficial stands, and thus improving their economic returns and decreasing the adverse negative ecological impacts on the forest. The agents gain experience in each simulation of the ten-year harvesting scenario with regards to which stands lead them to or prevent them from achieving their objectives. As a result, the economic returns begin to increase and the total area harvested and average harvested tree age decrease to improved levels. The variation existing during later episodes is due to the small probability of exploration, and due to the agents continually attempting to improve their solutions at the expense of one of the objectives or of another agent. However, the structure of IAM ensures that one agent cannot continually increase its economic returns as the objectives of all three agents must be met.
Figure 4-5. The process of agent learning is demonstrated in the improvement over time of (a) economic returns for each agent, (b) total area of harvested trees, and (c) average age of harvested trees.
4.5.2 Model Validation

The model was validated by determining if the Quality Solution is a non-dominated solution and if the tradeoffs between the objectives are representative of the equal weighting that was assigned to the objectives in equation 4-4. In multiobjective decision making, a non-dominated solution is one that is not dominated by another solution. As explained by Xiao et al. (2007), solution \( x' \) dominates solution \( x'' \) if, for one objective, a value of \( x' \) is no worse than \( x'' \), and, for another objective, the value of \( x' \) is better than \( x'' \). This formulation can be evaluated through an observation of the tradeoffs between objectives as presented in figure 4-6. Each point in the graphs of figure 4-6 represents a single solution generated by IAM. It is likely that IAM did not generate all potential solutions as the range of dominated solutions appears to be limited in the graphs. However, for each set of tradeoffs, the figure demonstrates that the Quality Solution belongs to the set of non-dominated solutions. Furthermore, the location of the Quality Solution within this set represents a relatively even compromise between the different objectives, which demonstrates that the equal weighting of objectives in equation 4-4 is manifest in the Quality Solution. That is, the location of the Quality Solution in the non-dominated set is such that satisfying one objective does not significantly compromise another objective Cohon (1978). These results demonstrate that IAM adequately generates a Quality Solution given the specific weighting scheme.
Figure 4-6. Each point represents a solution generated by IAM in relation to the tradeoffs between (a) total economic return and total area harvested, (b) total economic return and average age of harvested trees, and (c) total area harvested and average age of harvested trees. The location of the Quality Solution is highlighted.
4.5.3 Evaluating Response to Forest Disturbances

The results from the disturbance scenarios are intended to demonstrate how agents in IAM respond to random events and how the resulting solutions differ from the simulations in which disturbances are absent. The coefficient of variation (CV) of the quality of all stands was used as a measure for explaining the degree of confidence in determining which stands are most beneficial to harvest for achieving the objectives. Forest disturbance scenarios that result in high CV values indicate that there are few stands with relatively high quality, while the majority of stands possess low quality. This suggests a high confidence in a relatively small number of stands. Low CV values reflect results where the majority of stands have similar quality, thus indicating a lower confidence in determining which stands are most beneficial.

The different forest disturbance scenarios and the resulting CV values are provided in figure 4-7, from which two important observations can be made. First, there is an evident trend of declining CV values as disturbances become more frequent and intensified. This should be expected because agents can be more certain about where to harvest when the landscape exists without significant perturbations. However, as disturbances increase, the ability of agents to learn which stands are most beneficial diminishes due to the fact that all stands are not always available for harvesting - especially those that are of high economic value. The second observation is that the CV declines in a non-linear fashion. The CV remains relatively high for a set of disturbance scenarios, then significantly declines for the remaining scenarios. This suggests that some threshold could exist that determines the level of system perturbation over which agents can adequately determine the sets of stands that will lead to beneficial results. This observation is reiterated in figure 4-8, which provides a visual comparison between three
disturbance scenarios. The map depicting the lowest disturbance (i.e. \( f = 1 \) and intensity \( i = 10 \)) demonstrates that there are few stands considered to be very beneficial for harvesting compared to the total number of stands. As such, there are only a select few stands that are able to provide the highest level of economic return while minimizing the average age of harvested trees and total area harvested. The second scenario in which frequency \( f = 2 \) and intensity \( i = 20 \) shows that there are still numerous stands that have a relatively high quality, while the majority of stands in the landscape are still below 0.0. As disturbance is considerably intensified in the third scenario (i.e. \( f = 6 \) and \( i = 30 \)), there is no indication of stands with a high quality, but there is a much greater presence of stands with values above 0.0. This is because the increased disturbances are affecting those stands with higher economic values, thus decreasing the likelihood that harvesting them will lead to objective achievement. Instead, the stands with lower quality that are not as likely to be affected by disturbance are being selected more often to ensure that the agents receive some economic return.
Figure 4-7. The different disturbance simulations are defined by a combination of frequency and intensity values. The height of the bars depicts the coefficient of variation.
Figure 4-8. The difference in stand quality from different disturbance regimes defined by the frequency and intensity of disturbance.
4.6 Conclusion

The Intelligent Agent Model developed in this study introduces an approach for integrating agent-based modeling and reinforcement learning for multiobjective natural resource extraction. The use of independent agent decision making for developing solutions for the entire landscape is important for recognizing that the areas operated by each agent possess different spatial and non-spatial characteristics that define the types of decisions that will be made. It is important to note that the simplistic behaviour of the agents in IAM was intended to show the merits of the approach rather than simulate the complexities of the system. The agents in IAM are primarily focused on harvesting a specific volume of timber within a given timeframe, and interaction is limited to altering stand selection in order to minimize the global ecological consequences of harvesting. Agent behaviour can be altered in order to simulate responses to timber market fluctuations, introduce actions of cooperation or greed, and implement activities that help improve the productivity of the forest. Such behaviour could lead to an evolving relationship between agents and the manner in which they manage the landscape, which would potentially provide a different set of solutions. However, the level of complexity should not interfere with the ability of the RL algorithms to recognize objectives and improve solutions. Therefore, it is important in this approach to strike a balance between complexity and optimization.

While the issue of time received minimal attention in this study, it is important to note that IAM offers an advantage over previous work on spatial optimization as it provides information on the degree to which each stand contributes to objective achievement in different time steps. The nature of the quality function that was developed
in this study ensures that each stand has a *quality* for every time step in an episode. This information can be used to help agents learn when is the best time to harvest stands in addition to which stands are best to harvest. Knowledge of the appropriate timing of harvesting is particularly important if natural disturbances can be realistically simulated in the model. That is, knowing when a disturbance is more likely to occur in certain parts of the landscape will allow agents to adapt to ensure that objectives are still achieved, and the results can provide decision makers with the ability to effectively to manage resources in the presence of uncertain landscape dynamics.

The focus of this study was to generate a solution given an equal weighting scheme to all objectives. However, providing decision makers with a single optimal solution may prove to be impractical as spatial constraints in the landscape or existing policies may prevent the solution from being implemented. Therefore, there is a significant utility in providing decision makers with a set of non-dominated solutions and the ability to evaluate the tradeoffs between the objectives to determine how increasing the importance of one objective leads to a compromise in another. As such, future work with intelligent agent modeling for multiobjective natural resource allocation should consider how the entire set of potential solutions, rather than a single solution, can be generated and utilized for transferring practical information to the decision maker.
4.7 References


5. INCORPORATING PATH DEPENDENT KNOWLEDGE IN AN INTELLIGENT AGENT MODEL FOR NATURAL RESOURCE MANAGEMENT

5.1 Abstract

Space and time are intrinsic components of the decision-making process in natural resource management. Decisions to extract resources from a specific location have consequences for all future decisions as they may lead to profitable opportunities or, conversely, down a path of unfavorable outcomes. As such, the path dependent nature of decision making should be acknowledged and incorporated into models developed to assist the management of natural resources. The objective of this research is to develop an Intelligent Agent Model that is able to learn through repetitive simulation how to make optimal decisions regarding natural resource extraction. Specifically, an agent is guided by heuristic algorithms to search a natural landscape and learn which locations hold the highest profits and when it is best to extract the resource in order to improve the potential of future opportunities. The model is implemented using hypothetical and real datasets to emulate the process of harvesting trees in natural forests in order to maximize profits while respecting spatial constraints that are imposed in order to conserve various aspects of the forest. Simulation results reveal the ability of the Intelligent Agent Model to utilize path dependent knowledge in order to learn how to devise optimal solutions in a variety

---

4 A version of this chapter has been submitted to the journal *Landscape and Urban Planning* under the co-authorship of Suzana Dragićević.
of scenarios. Furthermore, the model demonstrates how the timing of decisions is linked to the spatial constraints imposed on the operation. The findings from this research can be used to inform natural resource management about the importance of the relationship between the location and timing of resource-based activities.
5.2 Introduction

Human decision-making has a significant impact on the transformation of natural landscapes (Antrop 1998). Consider the consequences of natural resource extraction, whereby decisions regarding the timing and location of events dictate the spatial patterns that emerge in landscapes over time, and the consequential effects imposed on various attributes of the landscape (Wimberly and Ohmann 2004). Decisions governing natural resource management are implicitly linked to various feedbacks that influence the balance between opportunities and constraints (Holling et al. 2002). The development of infrastructure, for example, provides access to resources and allows companies to receive economic returns from their operations. A positive feedback forms as profits are used for further development of infrastructure that leads to access to more resources. Conversely, negative feedbacks appear in the form of constraints that limit resource extraction due to the finite availability of resources, the geographical structure of a landscape, and as limits imposed as a form of conservation-related efforts (Colfer 2005). The potential for future resource extraction declines as a company draws closer to these constraints. As such, both positive and negative feedbacks are formed by previous actions and play a central role in structuring the availability of future opportunities (Brown et al. 2005, Castaldi and Dosi 2006). The intrinsic link between feedbacks and decision-making reveals the presence of path dependency, a term used to explain how decisions one is faced with in the present are dependent on those made in the past (Gerrits and Marks 2008). With regards to natural resource management, path dependency emerges when a decision to extract a resource from a spatial location at a particular moment in time is governed by the existence of feedbacks and dictates the potential for where future resource extraction
can occur. It is thus important to understand how the spatial and temporal characteristics of path dependency can be utilized for dictating landscape change through natural resource extraction.

Incorporating notions of path dependency requires the ability to examine the non-linearity of the system in both space and time. Space can be considered non-linear due to the asymmetric associations between objects in a landscape (Phillips 2003, Verburg et al. 2006). Principles of spatial relationships such as the First Law of Geography formulated by Tobler (1970) have been developed to explain how objects are related to each other in space. While such principles are likely too general to provide meaningful solutions to scenarios in which natural landscapes display significant heterogeneity, their utility has come in the form of assisting in the development of numerous spatial statistics and analytical techniques that are equipped to capture some form of spatial patterns. Examples of these techniques include landscape metrics (Venema et al. 2005, Brehmer et al. 2007, Miyamoto and Sano 2008), cluster analysis (Castro-Ortiz and Lluch-Belda 2007, Stefenon et al. 2008) and spatial autocorrelation (Perkins and Matlack 2002, Peterson et al. 2009). However, as these methods are focused on describing the overall spatial distribution of entities in a natural landscape, they are prone to overlook local variations where non-linearity may manifest – with the exception of statistics that are primarily based on local variation.

A suitable alternative to statistics for capturing spatial non-linearity is the growing field of spatial simulation. Several simulation techniques such as linear programming and heuristic search algorithms are useful for exploring datasets of landscape attributes in order to extract information and provide optimal solutions (Russell and Norvig 2003,
Malczewski 2004, Gustafson et al. 2006). Heuristic algorithms, in landscape planning context, typically operate by searching through a dataset representing a natural landscape to determine how the landscape can be spatially optimized for achieving one or more objectives (Westphal et al. 2007). Algorithms generate and improve solutions to multi-objective problems by evaluating the spatial patterns that are a consequence of management decisions. The main benefit of heuristic algorithms from a resource management perspective is that large and detailed datasets can be sufficiently explored to satisfy a diverse range of objectives. Limited assumptions need to be made regarding the spatial association between objects as local variations are explicitly investigated by the algorithm. Furthermore, these methods can be coupled with spatial statistics in order to combine local and global spatial information. This has lead to the successful employment of strategies such as hill-climbing (Liu et al. 2006, Mouton et al. 2008), simulated annealing (Boyland et al. 2004, Nalle and Arthur 2005, Pukkala and Heinonen 2006, Hynes et al. 2008) and genetic algorithms (Janssen et al. 2004, Ducheyne et al. 2006, Chamberlain and Meitner 2009) for generating and evaluating landscape scenarios for resource extraction. Furthermore, heuristic techniques can be implemented with spatio-temporal simulation models that incorporate both space and time in order to evaluate the process of landscape change as well as the resulting patterns. This is commonly performed by utilizing an agent modeling approach in which agents represent the decision making behaviour of individuals (Gimblett et al. 2001, Parker et al. 2003, Veldkamp 2004), while the heuristic algorithms help the agents to learn about optimal decision making behaviour (Manson 2005, Bennett and Tang 2006).
The literature concerning the integration of heuristic algorithms and agent models typically treats time as a measurement of change rather than a component of the model that can be used for formulating solutions. Treating time solely as a measurement of change has two drawbacks for using such approaches for incorporating path dependency. First, much detail of the process can be lost depending on the coarseness of the temporal resolution. As most heuristic models operate in discrete time, there exists the possibility that the period represented by a single time step is too long to capture sufficient information to adequately describe the process being modeled. This has adverse consequences for understanding the role of path dependency as it is not always possible to know the precise sequencing of events. As a result, all events that take place within a single time step have the same influence on future decisions regardless of the order in which they occurred. While the challenges of spatial resolution are well-documented in the literature (Benenson 1998, Menard and Marceau 2005, Kocabas and Dragićević 2006), little has been documented regarding sensitivity to time. The second drawback is that at each time step the information used to make a decision is based on the information made available from past decisions. By restricting the model to information from the past, it ignores the tradeoff between extracting a resource in the present versus extracting the resource at some future moment in time. The decision to extract a resource in the present may lead to significant short-term economic returns, but it may restrict future decisions towards a path of low profits in the long-term. Therefore, it would be advantageous if heuristic models could acquire information regarding the path dependent nature resulting from extracting resources in different locations within the landscape.
The importance of path dependency in natural resource management promotes the need for methods that can effectively evaluate and utilize the dimensions of space and time. To address these needs, this research integrates an agent modeling approach with a heuristic algorithm called reinforcement learning (Barto et al. 1981) for acquiring knowledge about the path dependent nature of a process in order to provide solutions to natural resource management problems. Specifically, an Intelligent Agent Model is developed in which an agent representing a forest company makes decisions to harvest trees in a forested landscape over time. Through repetitive simulation of a single forest harvesting process, the agent's actions are evaluated by a reinforcement learning algorithm that reward's the agent for actions that allow it to maintain acceptable profits while abiding by the spatial constraints that are imposed as a form of conservation-related efforts.

During initial simulations of the process, the agent explores the landscape to gather information as to which areas are most profitable for harvesting and the potential paths that are available when initiating forest harvesting in different areas across the landscape under various spatial constraints. After repeating the process numerous times, the agent exploits its knowledge of the landscape by examining the long-term benefits of extracting resources from specific locations and evaluating if it is better to extract the resource at the present time or at some time in the future. The agent thus learns where to harvest at each time step in order to maximize profits, minimize travel distance while abiding by spatial constraints. The Intelligent Agent Model is similar to a host of existing forest harvesting models that evaluate where and when to harvest trees with regards to conservation constraints (Csuti et al. 1997, Marinov et al. 2004) and the need to construct
roads (ReVelle and Snyder 1996, Hruza 2003, Gumus et al. 2008). However, this research focuses on the explicit use of space and time in the decision making process for improving forest harvesting (Spring et al. 2008) in order that it can meet both conservation and economic objectives. While contemporary forest models also focus on space and time for determining how to meet different objectives, this research provides a novel approach by representing the forest harvesting process through agent actions that have the potential to respond to dynamic landscapes and varying model parameters.

The model is first implemented using a simplified forest management scenario for explanatory purposes, and then implemented using a study site from southwestern British Columbia, Canada. In order to better understand the utility of the model, the next section offers a description of the role of path dependency in the context of forestry operations.

5.3 Forestry Operations as a Path Dependent Process

Forestry operations worldwide are tied to path dependency as the decisions to harvest trees in a particular area at a specific moment in time will have consequences for future decisions. In British Columbia, a forestry company or individual wishing to harvest trees must apply and be granted a license that permits them to cut a pre-determined volume of timber on an annual basis, referred to as the annual allowable cut (AAC) (Kimmins 1997). A company constructs a management plan that allows it to harvest trees within the constraints of the AAC while attempting to maximize profits. The company, therefore, engages in balancing the tension between negative and positive feedbacks.
While a company may determine that there are numerous areas that are highly profitable due to the species and age of trees, it typically begins its operations at one specific location. The decision of where to begin can have a profound effect on the decisions made thereafter as the company will need to construct roads in order to move from one area to another. The construction of roads is a costly procedure because numerous trees and other obstacles need to be cleared and the roads need to withstand the transport of large and heavy forestry equipment. It is therefore in the best interest of the company to decide on a location to start that is profitable and close to other profitable areas in order to minimize the length of road needed to maintain operations. Making good decisions allows a company to spend less of its return on infrastructure and enjoy higher profits.

The decisions a forestry company faces are further complicated by spatial constraints that can be imposed by the government or agreed upon during the development of management plans. Spatial constraints usually, but not always, are developed in order to limit the scope of negative effects resulting from harvesting operations. Constraints can be placed on specific species of trees that provide habitat for wildlife or are of some form of social importance. Constraints can also exist in the form of spatial measurements that are intended to protect the overall health of the forest. As a result of such constraints, companies need to consider when and where to harvest in order to maximize returns while minimizing the negative impacts on the forests. These various considerations dictate the location of harvesting at each moment in time and influence the availability of future opportunities.
5.4 The Intelligent Agent Model

The model developed in this study is based on an intelligent agent approach, as described by Russell and Norvig (2003). A forest company engaging in harvesting trees is represented in the model by an agent that learns by way of RL algorithms how to make beneficial decisions through the repetitive simulation of decision-making events. RL algorithms enable the agent to improve its decisions over time not by explicitly defining agent actions, but rather by allowing an agent to make its own decisions and receive a positive reinforcement in the form of a numeric reward if the decision or set of decisions leads to a favorable outcome (Abdulhai and Kattan 2003). Over numerous simulations of a given scenario, an agent receives substantial reinforcement that allows it to learn which set of decisions are optimal or near-optimal (Sutton and Barto 2000).

RL is a suitable choice for understanding and incorporating path dependence in models of forest harvesting. As a heuristic function, RL provides an agent with the ability to search extensively through large datasets to determine which areas in a forest are most suitable for harvesting. Theoretically, RL models should be able to evaluate a dataset of any size by altering the amount of exploring required by the agent. Similarly, scenarios that are more complex (i.e. have a greater number of potential solutions) can also be handled by increasing agent exploration. While RL has been previously employed for addressing forest management decision making, what clearly separates this model is that time plays a central role in the decision making process. As the agent makes decisions, the model keeps record of the profit received when harvesting an area at a specific time step. As a result, the agent can determine the benefit of harvesting trees in one area versus another, the long-term benefits of extracting resources from specific locations, and can
evaluate if it is better to extract the resource at the present time or at some time in the future. Furthermore, the model keeps track of the sequence of when each area is harvested regardless of the temporal resolution. This information is useful for assigning greater significance to earlier decisions which are likely to have a greater impact than later decisions on the outcome of the process (Mahoney 2006).

5.4.1 Model Parameters

The forested landscape searched by the agent is divided into areas referred to as forest stands that are composed of trees that are similar in species and age. At each time step of the model an agent will harvest the trees in a stand until it meets a defined level of allowable harvest. A stand is classified as either one that was harvested during the course of the episode, stand \( h \); one that was most recently harvested, stand \( i \); or one that is currently being examined by the agent for harvesting, stand \( j \). The remaining parameters can be divided into two classes: temporal parameters and spatial parameters (table 5-1).

The temporal parameters are responsible for keeping track of time, recording when harvesting activities take place, and governing the manner in which the agent searches and learns. A time step represents the period of time that is defined by the allowable harvesting level. For example, the AAC used in British Columbia limits the volume of timber that can be harvested in a single year; therefore, the time step in the model would represent one year. The model simulates forest harvesting for a defined number of time steps referred to as an episode (i.e. the period of time from the first time step to the terminal time step). Thousands of episodes are simulated in order for the agent to gain enough experience in the landscape so that it can learn which areas are the most
suitable for harvesting. A greater number of episodes is required for larger datasets and for scenarios that have a higher number of potential solutions.

Table 5-1. The temporal and spatial parameters of the intelligent agent model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>temporal</td>
<td>Time step</td>
</tr>
<tr>
<td>$R$</td>
<td>temporal</td>
<td>Total number of episodes</td>
</tr>
<tr>
<td>$r$</td>
<td>temporal</td>
<td>Current episode</td>
</tr>
<tr>
<td>$w_j$</td>
<td>temporal</td>
<td>Sequential weight parameter</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>temporal</td>
<td>Probability of exploration</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>temporal</td>
<td>Step-size function</td>
</tr>
<tr>
<td>$V_j$</td>
<td>spatial</td>
<td>Economic value of stand $j$</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>spatial</td>
<td>Distance between stand $j$ and stand $i$</td>
</tr>
<tr>
<td>$dij_{MAX}$</td>
<td>spatial</td>
<td>Maximum search distance from stand $i$ for stand $j$</td>
</tr>
<tr>
<td>$H_{MAX}$</td>
<td>spatial</td>
<td>Maximum harvested timber volume permitted per time step</td>
</tr>
<tr>
<td>$H$</td>
<td>spatial</td>
<td>Current harvested timber volume</td>
</tr>
<tr>
<td>$P_{MAX}$</td>
<td>spatial</td>
<td>Spatial constraint</td>
</tr>
<tr>
<td>$P_i$</td>
<td>spatial</td>
<td>Measure of landscape with regards to spatial constraint at time $t$</td>
</tr>
<tr>
<td>$Q_{j,t}$</td>
<td>spatial</td>
<td>Quality of stand $j$ that is being considered for harvesting at time $t$</td>
</tr>
<tr>
<td>$Q_{h,t}$</td>
<td>spatial</td>
<td>Quality of stand $h$ that was harvested at time $t$</td>
</tr>
</tbody>
</table>

The sequential weight parameter, $w_j$, is the weight applied to a decision based on when the decision was made. Higher weights are placed on earlier decisions to emphasize their greater significance in the overall outcome of the harvesting process. The sequential weight parameter $w_j$ is calculated using the equation
\[ w_h = \frac{n - s_h + 1}{n} \] (5-1)

where \( s_h \) is the sequence in which stand \( h \) was harvested, and \( n \) is the total number of harvested stands.

The search parameter \( \varepsilon \) is the probability of an agent exploring by randomly selecting a stand to harvest rather than exploiting its knowledge by selecting a stand that it knows is beneficial for harvesting. A value of 1.0 for \( \varepsilon \) means that the agent is randomly selecting all stands, while a value of 0.0 results in the agent only selecting the most beneficial stands. The value of \( \varepsilon \) is initially high as the agent does not have sufficient knowledge of the landscape to make informed decisions. The value diminishes over numerous episodes as the agent’s knowledge of the landscape is increased and the agent can exploit what it has learned. As the model draws closer to the final episodes, \( \varepsilon \) declines close to 0.0 and the agent selects only those stands that are most beneficial for achieving its objectives. The manner in which \( \varepsilon \) declines over the course of the model is important for ensuring that the agent does not settle on a local optimum, as it is possible for the agent to get trapped into believing that a sub-optimal area within the landscape is the best place to harvest trees if it is not given sufficient opportunity to explore.

The step-size parameter \( \alpha \) represents the rate of agent learning over the course of the model as the agent discovers the characteristics of the landscape. The value of \( \alpha \) is initially high (i.e. close to 1.0) to represent that an agent is close to reaching its full learning potential due to the fact that it does not possess prior information about the forest stands. As the number of episodes increases, \( \alpha \) declines and eventually nears 0.0 to represent how the amount of knowledge that an agent can potentially learn diminishes with each episode. The step-size parameter is linked with the search parameter because
learning decreases as the agent exploits its knowledge and selects stands that are most beneficial. Therefore, it is logical to diminish $\alpha$ and $\varepsilon$ at the same rate over the course of the model. The rate at which these two parameters diminish is assumed to follow a sigmoid curve represented by the equation

$$
\varepsilon, \alpha = 1 - \frac{1}{1 + e^{-\beta \left( \frac{r - R}{R/2} \right)}}
$$

where $r$ is the current episode number, $R$ is the total number of episodes, and $\beta$ determines the slope of the curve and is dependent on the number of episodes used to simulate the model. Altering the value of $\beta$ changes how the agent explores the forested landscape, and can thus be optimized to ensure that each stand is selected a sufficient number of times in order for the agent to know which stands are most beneficial for harvesting.

The spatial parameters of the model provide information on the characteristics of the forest stands. The economic value of a stand is calculated using the volume and market price of each species present in the stand. It is typical for different stands to have different values due to the heterogeneity of forests. The distance $d_{ij}$ between the currently harvested stand $j$ and the previously harvested stand $i$ is measured as the Euclidean distance between the coordinates of stand centers. The maximum search distance $d_{ij,MAX}$ is the maximum distance from stand $i$ that the agent will search in order to find a suitable stand. The maximum search distance for each decision is initially set as the immediate neighborhood as the agent will want to minimize the distance of road needed to be built.
However, \(d_{ij,\text{MAX}}\) is iteratively increased if no suitable stand can be located. The maximum value of timber to harvest \(H_{\text{MAX}}\) relates to the AAC described above. This value is dependent on the size of the forest area, the type and volume of trees, and the limits set by a regulatory body. In this model, the agent is allowed to harvest permitted that \(H < H_{\text{MAX}}\). The spatial constraint \(P_{\text{MAX}}\) represents the spatial constraint that is intended to protect one or more characteristics of the forest. The parameter \(P_t\) is simply the current spatial measurement of harvesting with regards to \(P_{\text{MAX}}\). For example, if \(P_{\text{MAX}}\) represents the maximum allowable harvested patch size, the value for \(P_t\) would represent the size of a harvested patch at time \(t\). Thus, stand \(j\) can be harvested if \(P_t < P_{\text{MAX}}\). Finally, \(Q_{h,t}\) represents the quality of stand \(h\) at time step \(t\). Quality \(Q_{h,t}\) is calculated as the expected economic return an agent can receive when harvesting stand \(h\) at time \(t\), which is further discussed in the next section.

### 5.4.2 Agent Behaviour and Learning

The model flowchart describing the decision making behaviour of the agent is presented in Figure 5-1. The model commences by initializing the quality of each stand to 0.0 to represent that the agent has yet to experience the consequences of harvesting any stands within the landscape. If \(H < H_{\text{MAX}}\), the agent selects a stand in which \(d_{ij} \leq d_{ij,\text{MAX}}\) based on probability \(\varepsilon\). The criteria of \(d_{ij} \leq d_{ij,\text{MAX}}\) is ignored if it is the first stand to be harvested in an episode due to the fact that no previous stands were harvested. The selected stand is harvested if \(Q_{j,t} > Q_{j,t+n}\) and \(P \leq P_{\text{MAX}}\), where \(n\) represents time. If no stands in the immediate neighborhood meet these criteria, \(d_{ij,\text{MAX}}\) is iteratively increased in order to expand the spatial extent of the search until the criteria are satisfied. The agent continues to harvest stands in that time step until \(H = H_{\text{MAX}}\). At the final time step of an
episode, the quality of each stand that was harvested during the episode is updated using the formula

\[
Q_{h,t,r+1} \leftarrow Q_{h,t} + \alpha \left( \frac{\sum V_h - c \left[ \sum_{y} d_y \right]}{w_h} - Q_{h,t} \right) \tag{5-3}
\]

where \( c \) is the cost of constructing a unit of road. This formula ensures that the quality \( Q_{h,t} \) of stand \( h \) at time \( t \) is updated each time the stand is selected to be harvested at time \( t \). The quality of a stand changes based on the difference between its returns as measured by economic returns and its existing quality. This difference is multiplied by the step size parameter \( \alpha \) in order to ensure that changes to \( Q_{h,t} \) are diminished over time. As a result, forest stand \( h \) at time \( t \) receives a positive reinforcement if it belongs to a set of harvested stands that provide a profit that is greater than the current quality of stand \( h \) at time \( t \).

That is, a positive reward is assigned when the agent is able to improve upon its previous decisions.

Once \( Q_{h,t} \) is updated for all harvested stands, a new episode begins with \( t \) reset to 1, and the agent begins its search for an initial stand to harvest. The entire process repeats itself while \( r < R \). During each episode, the agent improves its knowledge of which stands should be harvested at specific time steps in order to maximize its economic returns while adhering to the spatial constraints. Those stands that are more beneficial for harvesting will eventually be selected more often as their quality increases over numerous episodes and the search parameter \( \varepsilon \) declines. If the approximate number of episodes is appropriately selected, the quality for those stands that are most beneficial will converge.
towards their true profit values. That is, the stand’s quality will converge towards the profit the agent can expect to receive if it harvests stand \( j \) at time \( t \).

5.5 Model Implementation and Results

The model was encoded using Python programming script in Agent Analyst (Argonne National Laboratory 2006), a middleware that integrates an agent-based modeling toolkit REPAST (North et al. 2006) with the spatial database structure of ArcGIS (ESRI 9.1 2004). REPAST contains a set of libraries for assisting in the development of dynamic simulation through the actions of agents encoded with specific behaviours. Agents access and alter landscape attributes that are stored as fields in vector data structures.

5.5.1 Hypothetical Example

The model was first implemented using a hypothetical forest harvesting scenario in which one hundred stands are positioned in a 10 x 10 cellular grid as shown in figure 5-2. Each stand represents an area that is approximately 1 ha. The economic value of the stands and the distance between stands is simplified for explanatory and evaluative purposes. Each stand is represented by either a low economic value of $2,000.00 or a high economic value of $20,000.00. The cost for building a logging road is $2,000.00 per distance between two stands.
Figure 5-1. Flowchart describing the process of the Intelligent Agent Model.
The hypothetical example was simulated as a ten-year harvesting process, where each year is represented by a single time step in which $H_{MAX}$ equals one forest stand. A ten-year harvesting period was selected in this study to represent the timeframe over which several forest licenses are negotiated in British Columbia (Kimmins 1997). The model was calibrated to ensure that each stand was selected at least 30 times at every time step in order to provide adequate exploration of the landscape. A minimum sampling
of 30 times for each forest stand proved to be sufficient for the agent to learn the true quality of each stand. If each stand is sufficiently sampled, the model should be able to determine which stands are most beneficial for harvesting. The calibration resulted in $R = 10,000$, $\beta = 0.001$, and the step-size/exploration curve presented in figure 5-3. The simplistic nature of this scenario allows the model to be evaluated against a calculated expected outcome. Three different spatial constraint scenarios were separately simulated by varying the allowable harvested patch size. The first scenario, $P_\infty$, involved an infinite patch size. As such, the agent was permitted to harvest anywhere it decided. The second and third scenarios, $P_2$ and $P_1$, involve a maximum harvest patch size of two stands and one stand, respectively.

Figure 5-3. The sigmoidal curve depicting the decline in the step-size parameter and the probability of randomly selecting a forest stand over the course of the model.
The results for scenario $P_\infty$ are presented in figure 5-4. The height of the bars represents the stand quality calculated by the model when harvesting stand $h$ at time $t$. The quality shown at each time step is that observed by the agent. Thus, later time steps will exhibit lower quality due to the use of the sequential weight parameter $w_h$ in equation 5-2 in order to place greater weight on earlier decisions. An evident trend of which stands are most profitable can be visualized for the different time steps as presented in figure 5-4. This trend indicates that, when no spatial constraints are imposed, the agent should harvest adjacent stands from one corner of the landscape to the other (i.e. from stand in location 0,0 to stand in location 9,9, or vice versa). The results also illustrate that it is more beneficial to begin harvesting low value stands that are in close proximity to stands with the highest quality than it is to begin harvesting high value stands towards the centre of the landscape. While these results are intuitive, when applied to more complex datasets they can reveal that a decision that is less profitable in the short term may be more beneficial in the overall process.

The results for scenarios $P_2$ and $P_1$ are presented in figure 5-5. The same trend exists whereby stands at locations $x,y$ of 0,0 and 9,9 exhibit the highest quality. However, the overall quality is lower and the variation in quality of all stands in the landscape has diminished. Furthermore, the results from scenario $P_1$ display an increased spatial complexity as low value stands adjacent to high value stands are lower in quality than their low value neighbors. This is due to the agent learning that it cannot harvest adjacent stands, and should therefore avoid those low value stands that are adjacent to the high value stands.
Figure 5-4. The results for example 1, scenario $P_\infty$ for five different time steps. The height of the bars represents the quality of each stand.
Table 5-2. The calculation of the expected profits for all high value stands. HV = high value; LV = low value.

<table>
<thead>
<tr>
<th>Stand Location</th>
<th>Number of Harvested HV Stands</th>
<th>Harvested HV Stands Total Value</th>
<th>Number of Harvested LV Stands</th>
<th>Harvested LV Stands Total Value</th>
<th>Total Cost of Road Construction</th>
<th>Expected Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario (P_\infty)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0,0 / 9,9</td>
<td>10</td>
<td>$200,000</td>
<td>0</td>
<td>$0</td>
<td>$20,000</td>
<td>$180,000</td>
</tr>
<tr>
<td>1,1 / 8,8</td>
<td>9</td>
<td>$180,000</td>
<td>1</td>
<td>$2,000</td>
<td>$20,000</td>
<td>$162,000</td>
</tr>
<tr>
<td>2,2 / 7,7</td>
<td>8</td>
<td>$160,000</td>
<td>2</td>
<td>$4,000</td>
<td>$20,000</td>
<td>$144,000</td>
</tr>
<tr>
<td>3,3 / 6,6</td>
<td>7</td>
<td>$140,000</td>
<td>3</td>
<td>$6,000</td>
<td>$20,000</td>
<td>$126,000</td>
</tr>
<tr>
<td>4,4 / 5,5</td>
<td>6</td>
<td>$120,000</td>
<td>4</td>
<td>$8,000</td>
<td>$20,000</td>
<td>$118,000</td>
</tr>
<tr>
<td>Scenario (P_2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0,0 / 9,9</td>
<td>7</td>
<td>$140,000</td>
<td>3</td>
<td>$6,000</td>
<td>$28,000</td>
<td>$118,000</td>
</tr>
<tr>
<td>1,1 / 8,8</td>
<td>6</td>
<td>$120,000</td>
<td>4</td>
<td>$8,000</td>
<td>$28,000</td>
<td>$100,000</td>
</tr>
<tr>
<td>2,2 / 7,7</td>
<td>6</td>
<td>$120,000</td>
<td>4</td>
<td>$8,000</td>
<td>$28,000</td>
<td>$100,000</td>
</tr>
<tr>
<td>3,3 / 6,6</td>
<td>5</td>
<td>$100,000</td>
<td>5</td>
<td>$10,000</td>
<td>$28,000</td>
<td>$82,000</td>
</tr>
<tr>
<td>4,4 / 5,5</td>
<td>5</td>
<td>$100,000</td>
<td>5</td>
<td>$10,000</td>
<td>$28,000</td>
<td>$82,000</td>
</tr>
<tr>
<td>Scenario (P_1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0,0 / 9,9</td>
<td>5</td>
<td>$100,000</td>
<td>5</td>
<td>$10,000</td>
<td>$40,000</td>
<td>$70,000</td>
</tr>
<tr>
<td>1,1 / 8,8</td>
<td>5</td>
<td>$100,000</td>
<td>5</td>
<td>$10,000</td>
<td>$40,000</td>
<td>$70,000</td>
</tr>
<tr>
<td>2,2 / 7,7</td>
<td>4</td>
<td>$80,000</td>
<td>6</td>
<td>$12,000</td>
<td>$40,000</td>
<td>$52,000</td>
</tr>
<tr>
<td>3,3 / 6,6</td>
<td>4</td>
<td>$80,000</td>
<td>6</td>
<td>$12,000</td>
<td>$40,000</td>
<td>$52,000</td>
</tr>
<tr>
<td>4,4 / 5,5</td>
<td>3</td>
<td>$60,000</td>
<td>7</td>
<td>$14,000</td>
<td>$40,000</td>
<td>$34,000</td>
</tr>
</tbody>
</table>
Figure 5-5. A comparison of the expected quality for the high value stands at time step one versus the estimated quality. Points that fall on the 1:1 diagonal line are accurately estimated.
The results from scenario $P_\infty$, $P_2$ and $P_1$ are evaluated using the expected quality of high value stands presented in table 5-2. The information presented in this table states the expected maximum profit that can be earned when selecting each of the high value stands to begin the harvesting process. A comparison of the expected maximum profits and the estimated maximum profits that was generated by the model results are displayed in figure 5-6 for the high value stands. The results indicate that the model is able to adequately estimate the profit when selecting each of the high value stands to begin the harvesting process for scenario $P_\infty$. With regards to scenarios $P_2$ and $P_1$, the expected quality of the high value stands is significantly lower than scenario $P_\infty$ due to the presence of spatial constraints. Although the complexity of decisions has increased in scenarios $P_1$ and $P_2$, figure 5-6 demonstrates that the agent is able to adequately estimate the true quality of the high value stands.
Figure 5-6. The results at time step one for scenario $P_\infty$, scenario $P2$ and scenario $P1$. 
5.5.2 South western British Columbia Example

The model was next implemented using the dataset pertaining to a forest landscape in southwestern British Columbia in order to evaluate how the model performs under different spatial constraints in a realistic simulation context. The study site, shown in figure 5-7, is an area covering approximately 29,000 ha of forest in the Chilliwack Forest District. The diversity of the forest results in a wide distribution of stand economic values displayed in figure 5-7, which were calculated using data containing the volume of each species in a stand and the economic value of that species as stated in British Columbia timber market reports (Government of British Columbia 2008). An AAC of 30,000 m$^3$ was determined for the study site using information gathered from previous timber supply reviews of the region (Pederson 2004). Information from logging company reports was used to establish a cost for road construction at $100,000 per kilometer. The model simulates a ten year harvesting process in which each year is represented by a single time step. The model was calibrated to meet the sampling requirements, resulting in $R = 100,000$ and $\beta = 0.0001$. The increased number of episodes is necessary due to the significant increase in forest stands in the study site compared to the previous example. Two scenarios were simulated to evaluate how the agent responds to different spatial constraints. The first was the infinite patch size Scenario $P_{\infty}$, and the second was a maximum patch size of 50 ha, Scenario $P_{50}$.

As this study is focused on understanding the consequences of spatial constraints on decision making, the maximum patch size was based on an area deemed large enough to alter forest harvesting behaviour. However, future implementation of the model needs to determine a maximum patch size that is based on the need to conserve contiguous stands of forest for the protection of ecological functions.
Figure 5-7. The study site in southwestern British Columbia. The forest stands are represented by the economic value of the timber they contain.
The quality (i.e. total expected profit) of forest stands at $t = 1$ for scenarios $P_\infty$ and $P_{50}$ is presented in figure 5-8 and figure 5-9, respectively. The stands highlighted were determined to be the most profitable for harvesting over the ten year period. Stands are numbered to demonstrate the initial stand harvested during the first time step (1) and the last stand harvested during the final time step (10). The estimated economic return from harvesting the optimal stands was $31,489,523$ for scenario $P_\infty$ and $27,587,349$ for scenario $P_{50}$. The results demonstrate that there exists a different area of highest profit for the different scenarios, which demonstrates the agent’s response to the spatial constraints. In scenario $P_\infty$, several high value stands are clustered together, making them more beneficial to harvest when no spatial constraints exist. However, under the spatial constraints of scenario $P_{50}$, this area is no longer as profitable as high value stands exist at such a distance that they cannot all be harvested under the 50 ha maximum patch size. Instead, the agent learned to harvest high quality stands in another area that are of sufficient distance apart so as not to conflict with the allowable patch size.
Figure 5-8. The results at the first time step for example 2, scenario $P_\infty$. The optimal stands for harvesting are indicated.
Figure 5-9. The results at the first time step for example 2, scenario P50. The optimal stands for harvesting are indicated.
5.6 Discussion and Conclusion

The intelligent agent model developed in this study offers an approach that builds on the capabilities of heuristic algorithms and the dynamic simulation aspects of modeling that takes space and time into consideration. The intelligent agent model is able to explore a landscape sufficiently to learn which set of forest stands is most beneficial for harvesting, and utilize information about the timing of decisions in order to learn if it is more beneficial to harvest trees in a particular location at the current time step or at some time in the future. The scenarios from the hypothetical example demonstrate how the intelligent agent can construct a detailed knowledge of the landscape for each time step that relates specifically to the objectives of the operation without the agent having direct knowledge of that objective. That is, instead of concerning itself explicitly with the value of each forest stand or the distances between harvested stands, the agent focuses on the quality of the forest stands that not only incorporate these objectives but also integrate the temporal significance of their decisions. Each stand in the forest contains a value that serves as an indicator of its potential as a factor of time. That is, knowledge of time becomes embedded as an attribute of the landscape. This has benefits in addition to learning which stands are the most beneficial for harvesting because information is provided regarding the potential profits for all areas within the landscape. This can be particularly useful if harvesting the most optimal stands becomes impractical in reality due to unforeseen constraints, as it can help avoid selecting alternative stands that are appealing due to their short term benefits, but may lead to adverse long term consequences.
The results from the British Columbia scenarios revealed a rather intuitive observation that the potential profit from harvesting is likely to decline when spatial constraints are added due to the need for additional roads. More importantly, the results demonstrate that the addition of spatial constraints can completely alter where harvesting should take place in order to maximize returns. Areas that contain stands with the highest economic values may not be the most profitable due to how these stands are spatially arranged and how they can be accessed over time. Compared to the British Columbia example, the simplicity of the hypothetical example allowed for a practical validation of the model by comparing expected and estimated values of stand quality for those stands that would return the highest profit. This task becomes increasingly difficult when the number of forest stands in the dataset significantly increases due to the large rise in the number of potential outcomes. Traditional validation techniques for spatio-temporal modeling are inappropriate for measuring the success of the results in this research due to the absence of a dataset with which to compare the predicted results. Instead, it is practical to develop evaluation techniques that analyze the model to ensure that it is performing acceptably and not selecting sub-optimal solutions. The development of such techniques is beyond the scope of this research in which the intention is to test the inclusion of path dependent knowledge into agent decision-making. However, future work in intelligent agent modeling for natural resource management should examine how these models can be adequately evaluated.

The agent presented in this model had one objective that it attempted to achieve under a single spatial constraint. However, the intelligent agent approach has significant potential for incorporating multiple objectives that vary across spatial and temporal
scales. Agent behaviour can also be made more complex by allowing it to change its decision making to understand a dynamic landscape in which resources are altered due to growth and disturbances, and a dynamic financial system that influences changes to resource pricing. Such endeavours will provide further insight on how changes in spatial constraints influence natural resource management and how acquiring knowledge on path dependency can assist in the planning of how we utilize landscapes.
5.7 References


6. SIMULATION AND VALIDATION OF A REINFORCEMENT LEARNING AGENT-BASED MODEL FOR MULTI-STAKEHOLDER FOREST MANAGEMENT

6.1 Abstract

Spatial optimization and agent-based modeling present two distinct approaches that have been implemented in forest management research for incorporating the objectives of multiple stakeholders. However, challenges arise in their implementation as optimization procedures do not consider the interactions amongst stakeholders, and agent-based models generate results from which it is difficult to determine if objectives have been successfully achieved. The purpose of this research is to overcome these limitations by improving the ability of an agent-based model to achieve optimal forest harvesting strategies through the integration of reinforcement learning (RL). A simulation model is developed in which forest company agents harvest trees in order to maximize their profits while considering the potential to cooperate with a conservationist agent whose objectives are based on protecting species habitat. RL algorithms are implemented to allow the forest company agents and the conservationist agent to learn where harvesting should occur in order to achieve their objectives. The model is validated by

---

5 A version of this chapter has been submitted to the journal *Computers, Environment and Urban Systems* under the co-authorship of Suzana Dragićević.
determining if generated solutions can be considered optimal given system constraints, and by comparing observed agent behaviour against learning functions as defined by the RL algorithms. The results obtained demonstrate a non-linear relationship between different levels of cooperation and the ability of agents to achieve their objectives. The model also provides outputs that depict the relative quality of forest areas and the tradeoffs between objectives for different optimal solutions.
6.2 Introduction

Forest management requires the ability to integrate numerous objectives in order to satisfy the goals of different stakeholders (Bettinger et al. 2003, Seely et al. 2004, Kangas et al. 2005, Shao et al. 2005). Forest companies, for example, are typically driven by economic incentives that involve harvesting high quality timber and minimizing the construction of logging roads, while conservation-minded groups are interested in preserving the long term ecological functions of the forest. In addition, government agencies are motivated by the need to spur economic growth while avoiding the exhaustion of available timber.

A broad range of spatial optimization procedures exist that have the ability to integrate multiple and often-times conflicting objectives that can potentially exist across spatial scales. Heuristic modeling techniques represent the most recognized collection of approaches because of their ability to produce feasible solutions to large-scale spatial problems (Baskent and Keles 2005). Such methods include simulated annealing (Baskent and Jordon 2002, Ohman and Lamas 2003), tabu search (Caro et al. 2003, Richards and Gunn 2003) and genetic algorithms (Venema et al. 2005, Ducheyne et al. 2006), all of which can evaluate sets of spatial patterns with the aim of improving forest harvesting strategies in order to meet different objectives. However, the implementation of spatial optimization procedures can prove to be challenging when patterns resulting from harvesting processes are largely dictated by the complex interactions between stakeholders and various system components (Cerda and Mitchell 2004). Economic markets, ecological processes, political and social dynamics lead to uncertainty in our ability to implement strategic plans (Nelson 2003). As such, the dynamics that shape the
forest harvesting process can be in direct conflict with spatially-optimal patterns derived from heuristic modeling techniques.

In order to investigate the dynamics of forest management, agent-based modeling (ABM) offers a simulation approach in which computer agents represent the decision making behaviours of individual entities that influence their surrounding environment (Parker et al. 2003, Brown et al. 2004, Manson 2006). Agents can possess different strategies for responding to the actions of other agents and to the dynamic components of the landscape (Zellner et al. 2009). Positive feedbacks are formed as agent actions cause a system to become entrenched along a specific trajectory, while negative feedbacks can exist due to constraints that prevent a system from entering certain states (Manson 2006). As a consequence, the results of an agent-based model are viewed as patterns that emerge from the various dynamics of the system (Li et al. 2008). ABM is suitable for simulating forest management as the behaviour of interested stakeholders such as forest companies, conservation groups and government agencies as well as natural forest processes can be explicitly encoded into a spatial model in order to evaluate how the interactions amongst these elements lead to the emergence of different forest harvesting patterns (Purnomo et al. 2005).

The utility of ABM is evident in the number of applications for simulating forest-related processes that have surfaced in the literature in recent years. Examples include the use of ABM for simulating the emerging patterns of forest-agriculture transition (Deadman et al. 2004, Castella et al. 2005, Caplat et al. 2006, Evans and Kelley 2008, Bithell and Brasington 2009), multi-stakeholder management of tropical forests (Purnomo et al. 2005, Purnomo and Guizol 2006), and as well for developing forest
management spatial decision support systems (Purnomo et al. 2005, Purnomo and Guizol 2006). Yet, there still remain challenges when considering how to implement the information gained from an agent-based model into practical forest management strategies due to the perception that ABM remains largely a mechanism for exploring system dynamics rather than providing predictive results (Brown et al. 2006). This perception stems from the fact that simulating system complexity often involves the inclusion of various stochastic components that can lead a single model to produce numerous results that are all plausible outcomes of a system process (Batty and Torrens 2005). It is difficult to determine which outcome is most representative of the process given what the agents are trying to achieve, and how, if at all, the generated harvesting patterns satisfy the objectives of the different agents. As such, a paradox exists within the attempt to develop computational models for assisting forest management because of our conflicting desires to generate optimal spatial patterns while acknowledging the spatial and temporal complexities of the system.

The objective of this study is to address this optimization-complexity paradox through the development and validation of a model for multi-stakeholder forest management that integrates ABM and reinforcement learning (RL). RL is a computation approach stemming from the literature on machine learning and artificial intelligence that is used to improve model outcomes by providing numeric reinforcing rewards to those actions in a system that lead towards the achievement of a set of defined objectives (Barto et al. 1981, Sutton 1988). In this study, RL provides a means to incorporate optimization procedures into an agent-based model that allows agents to interact with each other and their environment while learning how to improve their decision-making
behaviour. RL algorithms evaluate landscape patterns and relay information to the agents that describes where and when forest harvesting strategies should take place in order for them to achieve their objectives. Furthermore, the RL agent-based model is parameterized as a multi-objective optimization model, which facilitates the use of traditional multi-objective evaluation methods for validating the ability of the model to produce optimal results given the complexity of the system.

The model is implemented using a scenario in which forest companies, a conservation group, and a government agency interact to influence harvesting patterns emerging from a dynamic forest landscape. Forest companies largely dictate harvesting patterns by attempting to maximize their economic returns, but they can also consider cooperating with the conservationist in order to incorporate objectives that protect species habitat. A study site in south western British Columbia, Canada, provides a context of conflicting objectives between harvesting timber for profit and conserving specific areas of the forest for species habitat. The conflict is also influenced by the need to maintain a level of harvesting that can support a local economy while avoiding resource exhaustion. The simulation model is specifically used in this study to evaluate how different levels of cooperation between forest companies and the conservationist influence the outcome of the harvesting process and to investigate reasons leading to varying results.

6.3 The RL-ABM Multi-Stakeholder Framework

The framework for integrating RL and ABM for multi-stakeholder forest management is presented in figure 6-1. Agents representing forest companies (F) harvest trees in the landscape based on the availability and price of timber for a specified number
of time steps. The period from the first to the final time step is referred to as an episode; the forest harvesting pattern resulting from an episode represents a single harvesting solution. For computational purposes, the forest company agents have no knowledge about the landscape at the onset of the model. They experience the landscape for the first time and make decisions that are only reactionary with respect to landscape dynamics. An agent that represents a government agency (G) estimates if the current amount of harvested timber is acceptable. If it is not acceptable, the government agency agent can either increase or decrease the harvesting quota of the forest companies.

At the end of the episode, RL algorithms are employed to evaluate whether the resulting solution satisfies the objectives of the forest companies agents. A positive reinforcing numeric reward is assigned to those areas in the landscape that constitute a beneficial solution, which informs the forest company agents that these areas are of substantial quality for achieving their objectives. Conversely, negative reinforcing numeric rewards are assigned when objectives are not satisfied in order to signal that these areas should be avoided. At the same time, an agent representing a conservationist (C) also evaluates the harvesting solution with a different set of RL algorithms to determine if the harvesting pattern is consistent with its own objectives. The rewards are encoded into a cognitive map for each agent, which represents the knowledge the agents have gained about the forest regarding where harvesting should occur in order to ensure their objectives are satisfied. This term was derived from Bennet and Tang (2006), who used the idea of cognitive landscapes to store the memory of mobile agents. With the first episode complete, the landscape is reset to its initial conditions, and the process is simulated again to produce another solution that is evaluated by the RL algorithms. The
scenario is repeated for a number of episodes that is dependent on the number of potential solutions.

Figure 6-1. The RL-ABM framework for multi-stakeholder forest management.
During initial episodes, the forest company agents continue to make reactionary decisions with no knowledge of the quality of locations and as such are not behaving in a manner to achieve their objectives. However, as the number of episodes increases, the forest company agents gradually access their own cognitive maps in order improve to their actions so that they can begin to achieve their objectives. Furthermore, forest company agents develop a memory of the total profits they received from previous solutions. At each time step of the model, they calculate the difference between their current profits and the maximum obtained profit from previously remembered episodes. If this difference is considered acceptable, the forest company agent switches its strategy to cooperate with the conservationist agent by accessing the conservationist agents’ cognitive map in order to incorporate its objectives. Whether or not the forest company agents decide to cooperate or act in their own interest, it is important that the rate at which they access the cognitive maps increases in small increments to avoid making decisions that are based on limited knowledge. However, by the final episodes, the forest company agents are accessing the cognitive maps for making the majority of their decisions as it is assumed they have gained sufficient knowledge for achieving the objectives.

As the purpose of developing the RL agent-based model is to provide optimal solutions to multi-stakeholder forest management given the complexity of the system, the model is validated by determining if agents’ decision making leads to results that can be considered optimal. This is accomplished by examining the tradeoffs between the different objectives in each solution using a method from the multi-objective decision making literature (Xiao et al. 2007). In addition, the nature in which the agents achieve
the different objectives over the course of the model is compared to learning functions that are parameterized within the RL algorithms. This assists in validating the model by determining if the generation of solutions is consistent with agent learning.

6.4 Methods

The model developed in this study consists of three types of stakeholder agents that interact in a forest landscape. Forest dynamics are represented by tree growth and fluctuating timber prices. The RL algorithms ensure that the stakeholder agents learn to contend with forest changes and the actions of other stakeholders when attempting to achieve their objectives.

6.4.1 Agents

The methods describing the model are discussed in the context of forest management in British Columbia. The main agents in the model are the forest companies that have been granted a license for harvesting a specified volume of timber within a designated jurisdiction at each time step of the model. The forest companies’ jurisdictions are composed of forest stands, which are spatially defined units of trees that are relatively homogenous in species and age. Harvesting thus takes place at the stand level, where all trees in the stand are removed by the agent who receives a profit based on the value of the timber located in the stand. The main goal of a forest company is to receive the highest attainable profit from its harvesting operations over the time period simulated by the model. As such, the forest company has two objectives: (1) harvest stands that are most
profitable, and (2) harvest stands that are closest to existing logging roads or to previously harvested stands in order to minimize the construction of new logging roads. Profit is calculated as the economic worth of harvested timber divided by the volume of timber harvested (i.e. dollars per volume harvested).

During early episodes, each forest company selects a number of random stands that fulfils its harvesting quota. They increase their access to their cognitive maps (the rate at which they are so is defined below) as the number of episodes increase. When they access their cognitive maps, the forest companies are presented with the value of each stand in their jurisdiction for the current time step and all future time steps. This allows the agents to determine which stands are most advantageous to select in the current time step in order to achieve their objectives, but it also allows them to evaluate if it is better to harvest the stand now or at some time in the future. The forest company is also given the option to harvest an amount of timber that is either 50% below or above its quota for each time step, but it must harvest within 10% above or below its quota for a five year period. This decision, which is adapted from the forest management principals in British Columbia, is based on the forest company’s memory of its profit potential from previous episodes. A forest company that is meeting or surpassing the profit from all previous episodes in its memory will decide to increase its harvest proportionally as it will have more expendable income to spend on constructing logging roads and acquiring or maintaining capital. Conversely, a forest company who is not meeting desired profits will reduce its harvest proportionally to its decline in profit due to the lack of expendable income.
Finally, the forest company has a choice between two strategies: to act in its own interest by harvesting with the intention of maximizing profits, or to act cooperatively with the conservationist by harvesting stands that abide by specific conservation principals. This decision is based on the cooperative nature of the forest company and the amount of profit it has currently received compared to its memory of maximum profits from previous episodes. That is, a very cooperative forest company will forgo a greater profit to cooperate with the conservationist, while a minimally cooperative agent will only cooperate if it is close to achieving maximum profits.

The conservationist agent represents an institution whose goal is to protect local wildlife habitat. Their objectives are thus: (1) minimize the overall area of harvested trees, and (2) minimize the harvesting of trees with characteristics that are of importance to a specific wildlife species in terms of habitat. The conservationist is unable to apply pressure on the forest companies, but instead it evaluates the solution generated from each episode and develops its own cognitive map of where the forest companies should harvest in order to meet the conservation objectives. The conservationist provides information regarding the best stands to harvest to any agent that is willing to cooperate with it at any time during the model. Therefore, the conservationist should be most effective at achieving its objectives when forest companies are more cooperative.

The remaining agent in the model represents a government agency whose objective it is to regulate the amount of timber harvested for the entire study area. The government agency does not possess learning mechanisms or memory, but instead exists to ensure that the forest companies are harvesting timber that is in line with the productivity of the forest. Following harvesting guidelines as set forth in British
Columbia, the government agency establishes how much timber can be harvested at each time step over the entire study area, and then apportions this amount to each forest company based on the available timber in each forest company’s jurisdiction. At five year intervals, the government agency evaluates if the amount of timber harvested has surpassed or fallen short of the amount initially established for the given time period. If the agents have collectively harvested above their combined quota, the government agency will reduce their quotas proportionally based on the amount of overharvesting and the availability of timber within each forest company’s jurisdiction.

### 6.4.2 Landscape Dynamics

Two attributes of the forest change each time step of the model independent of agent actions. The first dynamic component is forest growth, which is implemented in the model by increasing the age and volume of each stand in the forest. While age is synchronously increased for all stands at each time step, volume is increased at an amount that is dependent on the species composition of the stands. Stands that contain all fast growing species will experience greater increases in volume at each time step, as will stands that are relatively younger. The exact growth rate for different stands is estimated using tree growth projections based on the study area.

The second dynamic component is the economic value of each species present in the forest. Tree values in real timber markets fluctuate over time based on the demand for certain wood products. Therefore, it is important to present this fluctuation as it has a direct effect on forest companies’ decisions of which stands to harvest. Information from timber market reports are used to estimate the probable annual fluctuation of the price of
each species, as well as an estimate of the likely maximum and minimum price of each species.

6.4.3 Reinforcement Learning Algorithms

The reinforcement learning algorithms are employed at the completion of each episode to update the cognitive maps of each agent. The first step in this process is to evaluate the resulting landscape harvesting pattern to determine if it produces a solution that is more optimal for meeting objectives than all previously generated solutions (the method for which is described below). An objective function, \( f(x_F) \), is calculated independently for each forest company agent in which they are to maximize profits and minimize the distance to previously harvested stands is calculated using the equation

\[
f(x_F) = \min \left[ 1 - \left( \frac{x_1 - x_{1\text{min}} + 1}{x_{1\text{max}} - x_{1\text{min}}} \right) \right] + \left( \frac{x_2 - x_{2\text{min}} + 1}{x_{2\text{max}} - x_{2\text{min}}} \right)
\]

(6-1)

where \( x_1 \) equals the total profit acquired from all harvested stands from the episode and \( x_2 \) equals the length of new logging roads required to reach all harvested stands. The minimum and maximum profits generated from all previous solutions is represented by \( x_{1\text{min}} \) and \( x_{1\text{max}} \), while \( x_{2\text{min}} \) and \( x_{2\text{max}} \) represent the minimum and maximum average road length. Similarly, equation 6-2 is used to calculate the conservationist agent's objective function, \( f(x_C) \), in which \( y_1 \) equals the average density of harvest stands and \( y_2 \) equals an attribute of the forest that is used as a measure for conserving species habitat. Density is used as a surrogate measure for area as the harvesting of denser stands will lead to less overall area from which trees are removed.
A positive reinforcement is awarded to a solution if the objective function calculated from the current episode is less than or equal to the lowest objective function previously generated that is retained in the agent’s memory. Each stand in the solution receives a reinforcement that is a function of its contribution to achieving the objectives. The reinforcement, \( r \), for each stand, \( s \), is calculated for the forest companies using the equation

\[
r_s = \left( \frac{v_s}{v_{s\text{max}}} + \frac{d_s}{d_{s\text{max}}} \right)
\]

(6-3)

where \( v \) is the economic value of stand \( s \), and \( d \) is the distance of stand \( s \) to the nearest logging road, while \( v_{s\text{max}} \) and \( d_{s\text{max}} \) are maximum respective values from the set of harvested stands. The justification behind the use of minimum and maximum values in the objective and reward functions is to provide a measure of standardization that allows the two objectives to be combined into a single value. The reinforcement for the conservationist is calculated as

\[
r_s = \left( \frac{q_s}{q_{s\text{max}}} + \frac{m_s}{m_{s\text{max}}} \right)
\]

(6-4)
where $q$ is the tree density of stand $s$, and $m$ is the measure relating to species habitat.

The reinforcements are used in a formula called the *Quality Function* that updates the *Quality* of each stand in the solution. Stand *Quality* is the attribute represented in the *cognitive maps* that forest companies access in order to learn where harvesting should take place in order to achieve their objectives. The *Quality* encompasses the different objectives as well as the appropriate timing of when specific decisions regarding harvesting should be made. The *Quality Function* for updating the forest companies’ *cognitive maps* is expressed as

$$Q_{s,k+1} \leftarrow Q_{s,k} + \alpha \left( \frac{r_{s,k}}{w_{s,k}} - Q_{s,k} \right)$$

(6-5)

and the *Quality Function* for updating the conservationist’s *cognitive map* is defined by the equation

$$Q_{s,k+1} \leftarrow Q_{s,k} + \alpha \left( r_{s,k} - Q_{s,k} \right)$$

(6-6)

where $Q_{s,k}$ and $Q_{s,k+1}$ are the *Quality* of stand $s$ at the current and next episode, respectively. The parameter $w_{s,k}$ in the forest companies’ *Quality Function* represents a temporal weighting function that assigns a greater weight to those stands harvested earlier in the current episode. Stands harvested earlier are considered to be of higher importance because they have a greater influence on where forest companies decide to harvest in

203
subsequent time steps. The value for \( w_{s,k} \) is expressed as a real number between 0.0 and 1.0, where stands harvested earlier in the episode are closer to 1.0, and the weight diminishes as stands are harvested later in the episode.

The parameter \( \alpha \) in the Quality Function is the agent learning rate, which is set to 1.0 at the beginning of the model and declines to 0.0 based on the equation

\[
\alpha = 1 - \left[ \frac{1}{\beta \left( \frac{k - \frac{K}{2}}{2} \right)} \right]
\]

(6-7)

where \( k \) is the current episode, \( K \) is the total number of episodes for which the model is run, and \( \beta \) is a real number less than 1.0 that governs the slope of the curve. The equation ensures that the learning rate – or the rate at which the potential to learn about the landscape declines – decreases according to a sigmoid curve to represent that the agents are learning to their full potential in early episodes, and then learning gradually diminishes through the course of the model as agents develop a better understanding of their landscape with regards to their objectives. Scenarios with a higher number of potential solutions will require a lower value of \( \beta \) which leads to a more gradual slope and a slower rate of learning. The sigmoid curve is also employed to determine when forest companies access the cognitive maps. In order for the forest companies not to converge on a set of sub-optimal solutions, it is important for forest companies not to access these maps during early episodes as the maps will only represent partial knowledge of the landscape. This is a realistic representation of the relationship between forest company knowledge and their actions as it demonstrates that they will make bad
decisions if they do not provide enough opportunity for learning about how to achieve their objectives. However, as the agents gain experience through repetitive simulation, the Quality values in the cognitive maps begin to converge towards more representative values, and as such the forest companies increase their access to the maps to improve their decision-making behaviour. At each time step, the probability of a forest company accessing the cognitive maps is defined by equation 6-7. The model is simulated for a number of episodes that allows each stand to be sampled a sufficient number of times in order for the agents to learn how to satisfy their objectives.

6.4.4 Validation

The validation of the model takes place in two stages to determine if the heuristic optimization employed by the reinforcement learning algorithm produces solutions that can be considered optimal for a given set of objectives, and if agents are acquiring knowledge at a rate that is consistent with the learning rate. For the first stage, three “Benchmark” optimization scenarios are simulated to evaluate if the final solution generated by the model belongs to the set of the most optimal solutions generated over the course of the model. The three Benchmark simulations are as follows: (1) forest companies harvest stands in order to satisfy only their own objectives by only accessing their own cognitive maps; (2) forest companies harvest stands in order to satisfy only the conservationist’s objectives by only accessing the conservationist's cognitive map; and (3) the forest companies harvest stands in order to equally satisfy both their objectives and the conservationist’s objectives by using the sum of quality values from their cognitive map and that of the conservationist. The set of non-dominated solutions from each scenario is identified according to the formulation by Xiao et al. (2007), in which
solution $x'$ dominates solution $x''$ if and only if $\forall i \ f_i(x') \leq f_i(x'')$ and $\exists i \ f_i(x') < f_i(x'')$.

Using the forest companies’ objective function as an example, solution $x'$ dominates solution $x''$ if either the total profit or average road distance of $x'$ are no worse than $x''$ while the other objective of $x'$ is better than $x''$. The optimization component of the model is considered validated if the final solution resulting from each Benchmark scenario belongs to the set of non-dominated solutions for that scenario. The Benchmark solutions are then compared to the solutions from the multi-strategy scenarios to evaluate how objectives are achieved given the complexity of agent interactions.

The second validation stage is based on determining if the observed rate of agent knowledge acquisition is in agreement with the learning rate as defined by the parameter $\alpha$ in the Quality Function. This is accomplished by a visual comparison between the manner in which solutions are improved over the course of the model versus the defined learning rate. The two should ideally be consistent with each other when agents are able to sufficiently achieve their objectives. A failure of the two rates to coincide reveals that the agents are not learning in a manner as defined by model parameters, or that the complex nature of the system is preventing agents from achieving their objectives. If the former is found to be true, the model is considered invalid based on the inability of agents to accurately learn about their landscape.

### 6.5 Model Implementation

The model is implemented through the simulation of forest harvesting in a study site located in the Chilliwack Forest District in south western British Columbia. The area provides opportunities for harvesting due to the availability of desired timber and
proximity to timber processing and shipping locations. However, the area also lies within the habitat of the Northern Spotted Owl, a species that has been placed on Canada’s Endangered Species list due to declining populations as a result of habitat loss.

The study area is partitioned into four forest company jurisdictions as previously defined by the Government of British Columbia. The study site depicting the four forest company areas is shown in figure 6-2. A timber supply review (Pederson 2004) was used to define the annual harvesting quota for each forest company. The quota measures the volume of timber that can be harvested on an annual basis, which is measured in cubic meters (m$^3$). Harvesting quotas and other information concerning each forest company area are provided in table 6-1. The conservationist’s objectives are defined by their desire to protect the habitat of the locally endangered species, the Northern Spotted Owl. As this species needs large areas of old growth forest (i.e. stands over 140 years) to survive, the second objective of the conservationist is to minimize the average age of stands harvested by the forest company.
Figure 6-2. The study site in southwestern British Columbia that is partitioned into four licensee areas.
Table 6-1. Attributes for each licensee's harvestable area.

<table>
<thead>
<tr>
<th>Licensee</th>
<th>Area (ha)</th>
<th>Total Volume (m³)</th>
<th>Average Age (years)</th>
<th>Average Stand Distance to Logging Roads (m)</th>
<th>Total Value of Timber (CDN Dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7662.01</td>
<td>5,980,513.94</td>
<td>145.33</td>
<td>1042.40</td>
<td>550,043,571.05</td>
</tr>
<tr>
<td>2</td>
<td>4793.10</td>
<td>3,868,193.20</td>
<td>119.45</td>
<td>702.54</td>
<td>360,913,924.78</td>
</tr>
<tr>
<td>3</td>
<td>10020.25</td>
<td>5,668,080.50</td>
<td>82.63</td>
<td>289.55</td>
<td>534,036,498.59</td>
</tr>
<tr>
<td>4</td>
<td>6781.27</td>
<td>4,393,509.49</td>
<td>73.85</td>
<td>495.09</td>
<td>381,772,161.88</td>
</tr>
<tr>
<td>Total</td>
<td>29256.66</td>
<td>19,910,297.13</td>
<td>104.40</td>
<td>608.15</td>
<td>1,825,996,156.30</td>
</tr>
</tbody>
</table>

The information from table 6-1 reveals a north-to-south gradient between the different harvesting jurisdictions with regards to the average age of trees and the accessibility to forest stands. This gradient likely exists due to the fact that Jurisdiction 4 is located within the closest proximity to major transportation routes and is thus most easily accessible. As a result, Jurisdiction 4 has the lowest average stand age and a relatively low distance to between stands and logging roads because it has experienced the most harvesting. Conversely, Jurisdiction 1 is the furthest from major transportation routes and has thus maintained relatively older trees and fewer logging roads. This gradient will impose some influence on the forest company agents for being able to achieve their objectives, but it is likely not significant as those agents with younger, less profitable stands also have to spend less money on logging roads while those agents with older trees and higher valued trees will have to spend more money in order to access such trees. As a result, it can be assumed that no forest company agent has a significant advantage over any other.
Timber prices and timber price fluctuation (table 2) were determined by information from British Columbia timber reports (Government of British Columbia 2007). The rate of annual volume growth for each stand was generalized based on a classification of species within the stand. Using a volume growth prediction toolkit called the Variable Density Yield Prediction (VDYP) software provided by the Government of British Columbia (British Columbia Ministry of Forest 2007), volume growth was estimated for stands dominated by coniferous species (i.e. cedar, cypress, fir, and hemlock,), stands dominated by deciduous species (alder, cottonwood and maple), and stands that contain a relatively even composition of coniferous and deciduous species.

The model was encoded in Python within an agent-based modeling application called Agent Analyst that links the agent simulation procedures of the Recursive Porous Agent Simulation Toolkit with ArcGIS 9.2. Each episode in the model represents a ten-year harvesting time frame in which a single year is represented by one time step. This temporal resolution coincides with the timber quotas that are defined on an annual basis, and with the estimated rate of timber growth as provided by VDYP.

Table 6-2. Timber pricing and timber price fluctuation data for each species.

<table>
<thead>
<tr>
<th></th>
<th>Alder</th>
<th>Cedar</th>
<th>Cottonwood</th>
<th>Cypress</th>
<th>Fir</th>
<th>Hemlock</th>
<th>Maple</th>
<th>Spruce</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>49.0</td>
<td>99.0</td>
<td>33.0</td>
<td>79.0</td>
<td>64.9</td>
<td>49.0</td>
<td>40.0</td>
<td>73.0</td>
</tr>
<tr>
<td>Maximum</td>
<td>80.0</td>
<td>164.0</td>
<td>41.0</td>
<td>189.0</td>
<td>110.0</td>
<td>70.0</td>
<td>66.0</td>
<td>137.0</td>
</tr>
<tr>
<td>Average</td>
<td>66.1</td>
<td>129.6</td>
<td>35.8</td>
<td>135.5</td>
<td>101.9</td>
<td>59.4</td>
<td>50.6</td>
<td>108.1</td>
</tr>
<tr>
<td>Maximum Fluctuation</td>
<td>31.0</td>
<td>43.2</td>
<td>5.0</td>
<td>33.0</td>
<td>35.1</td>
<td>13.0</td>
<td>26.0</td>
<td>40.0</td>
</tr>
</tbody>
</table>
In order to evaluate how different levels of cooperation between forest companies and the conservationist influence the outcome of the harvesting process, five different multi-strategy scenarios were simulated in which the forest companies’ willingness to cooperate is varied. For the five scenarios, agents will cooperate if the difference between current profits and the maximum known profit from previous episodes stored in the forest companies’ memory is (1) 0%; (2) no less than 5%; (2) no less than 10%; (3) no less than 25%; and (5) no less than 50%. It is expected that increases in forest company cooperation will lead to greater knowledge acquisition by the conservationist, which will in turn allow the conservationist to be more successful in meeting its objectives.

6.6 Results

For each scenario, the forest harvesting process represents a ten-year period, where each year is represented by a single time step. The model was run for 10,000 episodes as this lead to each stand being selected at least 50 times, which was found to provide a sufficient level of sampling and this avoided simulating too many episodes at the expense of no significant changes to the results.

6.6.1 The Effects of Agent Behaviour

The results from the multi-strategy Scenario 1 are presented in figure 6-3. The cognitive maps for the forest companies and the conservationist are presented for the final episode of the scenario, as is the harvesting solution that is based on those stands harvested in the final episode. The cognitive maps depict the Quality of each forest stand, which the forest companies utilize for determining where and when to harvest. This
solution is compared to the final solution generated by the other four multi-strategy scenarios with regards to the proportion of stands that are harvested in both scenarios (table 3). While it is expected that scenarios with similar levels of cooperation would produce more similar solutions, the results reveal that the solution from Scenario 3 is overall least similar to the solutions from other scenarios. The amount of dissimilarity does not appear to be significantly different for all the scenarios, but Scenario 3 does exhibit dissimilarity that is disproportionate to the other solutions.

The results depicting the values for the objectives for each scenario are displayed in figure 6-4. Conversely to the expectation that the conservationist will better achieve its objectives with increasing cooperation, the results demonstrate that all agents are able to best achieve their objectives at an intermediate level of cooperation. Another contradiction can be observed in which the forest companies collectively are able to reduce the average distance of additional logging roads and increase their profits when agreeing to provide an increased level of cooperation instead of only cooperating when they are certain that they are maximizing their profits. However, as with the conservationist’s objectives, the forest companies are only able to improve their ability to achieve objectives with a certain level of cooperation.
Table 6-3. The percentage of stands that are harvested in each scenario presented in the vertical column that are also harvested in the scenarios presented in the horizontal column.

<table>
<thead>
<tr>
<th>Percentage of Overlapping Harvested Stands From Different</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100.0</td>
<td>46.22</td>
<td>41.51</td>
<td>40.00</td>
<td>37.74</td>
</tr>
<tr>
<td>2</td>
<td>42.98</td>
<td>100.0</td>
<td>41.23</td>
<td>33.33</td>
<td>35.09</td>
</tr>
<tr>
<td>3</td>
<td>36.67</td>
<td>39.17</td>
<td>100.0</td>
<td>23.33</td>
<td>30.93</td>
</tr>
<tr>
<td>4</td>
<td>38.39</td>
<td>33.93</td>
<td>26.92</td>
<td>100.0</td>
<td>32.14</td>
</tr>
<tr>
<td>5</td>
<td>38.46</td>
<td>38.46</td>
<td>35.58</td>
<td>36.62</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Figure 6-3. The cognitive maps for the forest companies and the conservationist and the harvesting solution for Scenario 1.
Figure 6-4. The values for each objective resulting from Scenarios 1 to 5.
Figure 6-5 provides information regarding the percentage of harvesting events in which forest company agents decide to cooperate with the conservationist agent. This information is divided into four classes to depict how cooperation changes in the different scenarios depending on the episode of the model. The results demonstrate that cooperation increases most significantly for Scenarios 1 and 2 over the course of the simulation, while Scenario 3 observes a moderate increase in cooperation. This is contrary to the expectation that forest companies would be cooperating more often in Scenario 3 because their behaviour is defined to do so. Scenarios 4 and 5 show no observable pattern of change as the forest companies cooperate the majority of the time throughout the course of the model.

Figure 6-5. The episodes are separated into classes to demonstrate the change in cooperation levels as the number of episodes increases.
Another set of important findings relates to the total volume of trees harvested by the forest company agents in the final solution of the model. Figure 6-6 demonstrates that all forest companies in Scenarios 1 and 2 harvested at least 10% over their harvesting quotas as defined by the government agent. This reveals that, in these scenarios, forest companies were able to reach the maximum profits as recorded in their memory, and as such harvested beyond their quotas due to the availability of expendable income. All forest companies in Scenario 3 also harvest above their defined quotas, but to a noticeably lesser degree than previous scenarios. Furthermore, all agents in the final two scenarios harvest well below their quotas, demonstrating that they were not able to achieve the maximum profits.

6.6.2 Validating Model Outcomes and Agent Behaviour

The non-dominated solutions from the three Benchmark optimization scenarios are shown in figure 6-7 with regards to the tradeoffs between the different objectives. The most important tradeoffs to consider are those displayed between (a) the average stand age and average stand density, and (f) between the average stand distance and total profit, as these represent the conservationist’s objectives and forest companies’ objectives, respectively. The results clearly indicate that each Benchmark scenario generates a distinct set of non-dominated solutions. That is, Benchmark A provides a set of non-dominated solutions that are best for achieving the forest companies’ objectives; Benchmark B provides a set of non-dominated solutions that are best for achieving the conservationist’s objectives; and the non-dominated solutions from Benchmark C lie somewhere between the other two sets.
Figure 6-6. The difference between the total observed harvest for each forest company (FC) versus the expected total harvest as defined by the harvesting quotas. Positive differences represent that forest companies are harvesting above their quotas, while negative represent that forest companies are harvesting below their quotas.

In addition, the final solution from each Benchmark scenario is highlighted in figure 6-7, demonstrating that the final harvesting solution derived from the model is within the set of non-dominated solutions. Thus, it can be concluded that the optimization component of the model adequately generates solutions based on specified objectives. Figure 6-7 also displays the final solution generated from each multi-strategy scenario. These results indicate that, similar to the results presented in figure 6-4, an intermediate level of cooperation is most suitable for the agents to achieve their objectives. In fact, the solution generated from Scenario 3 is at least as good as the non-dominated solution set.
generated from Benchmark B in which the forest companies equally consider all objectives.

The second stage of model validation compares the sigmoid curve that defines the learning rate (figure 6-8) with the manner in which agents improve their ability to achieve their objectives (the inverse of the curve is used to compare those objectives in which the values are to be maximized – i.e. the average stand density and total value). The values representing the forest companies' objectives at each episode are presented in figure 6-9, while the conservationist's objectives are presented in figure 6-10 (Scenario 2 was excluded due to its similarity to Scenario 1). The shape of the curve can be observed for each objective, particularly for Scenarios 1 and 3, illustrating that agents are initially learning about their landscape and not making the best decisions, but the experience they gain over numerous episodes is used to improve their knowledge of which stands should be harvested in order to meet their objectives. The curve is most evident for each objective in Scenario 3, which demonstrates that the observed rate of agent knowledge acquisition is most consistent with the learning rate when agent behaviour best facilitates objective achievement. Furthermore, the graph depicting the total profit from harvested stands for Scenario 3 reveals that the forest companies increased their exploration of alternative solutions as there is a greater range of profit values generated within the scenario. The fact that the presence of the sigmoid curve begins to deteriorate through Scenarios 4 and 5 suggests that the behaviour of the agents defined in these scenarios may be preventing the forest companies and the conservationist from fully exploring their landscape and consequentially learning how to improve their objectives.
Figure 6-7. The non-dominated solutions for Benchmark scenarios A (optimizing licensees' objectives), B (optimizing the conservationist's objectives), and C (equally optimizing all objectives). The final solution from each Benchmark scenario is in bold, while the solutions for each multi-strategy scenario are represented by their respective numbers.
Figure 6-8. The learning rate defined by parameter $\alpha$ that determines the amount of knowledge agents are able to acquire at each time step.
Figure 6-9. Total profit and the average distance to roads of all harvested stands at the end of each episode. The results are shown for scenarios 1, 3, 4 and 5 (scenario 2 was omitted as it presented the same findings as scenario 1).
Figure 6-10. Average age and density of all harvested stands at the end of each episode. The results are shown for scenarios 1, 3, 4 and 5 (scenario 2 was omitted as it presented the same findings as scenario 1).
The observation that all agents better achieve their objectives in Scenario 3 than in the first two scenarios is contrary to initial expectations, but the results reported in figure 6-5 and figure 6-6 reveal how system feedbacks played a role in this outcome. Forest companies in the final episode of Scenarios 1 and 2 must be reaching maximum profit expectations because the figures illustrate that they are harvesting well above their quotas. However, the positive feedback that allows them to increase their harvest also ensures a greater level of cooperation with the conservationist agent. As a result, the majority of harvesting that takes place above the quota is based on meeting the conservationist’s objectives, which does not necessarily help to improve the forest companies’ profits or ability to reduce the need for new logging roads. Furthermore, the fact that minimal cooperation occurred in earlier episodes limits the knowledge that the conservationist could gain regarding its own objectives, which compromises its decisions in later episodes when it is often included in the decision making process. As a result, the positive feedbacks that cause forest companies to increase harvesting levels lead to a decrease in profits because excess harvesting is based on the limited knowledge of the conservationist. The observation that forest companies in Scenario 3 harvest above their quotas to a lesser degree than in the first two scenarios implies that they are not as successful at achieving maximum profit expectations. This leads to fewer instances of cooperation in the final episodes of the model as shown in figure 6-5, which results in a higher achievement of the forest companies' objectives. In addition, figure 6-5 demonstrates that the conservationist agent was incorporated more often in the decision making process during earlier episodes of the model, which allowed it to sufficiently learn where the forest companies should harvest in order to achieve the conservation
objectives once the model reached its final episode. The implications of these findings are
that increasing levels of cooperation do not lead to expected changes in the ability to
achieve objectives. The non-linear relationship between agent behaviour and model
outcomes is driven by system feedbacks that demonstrate how moderate levels of
cooperation between stakeholders is beneficial in order for a variety of objectives to be
achieved.

6.7 Discussion

The results from this study reveal that agent behaviour has an influence on the
ability to learn about forest harvesting patterns that are beneficial for achieving their
objectives. However, it can be safely concluded that specific changes in agent behaviour
are non-linearly related to the simulated results as the forest companies’ willingness to
cooperate does not provide predictable outcomes. Information extracted from the
comparison of solutions, the outcomes of the objectives, and the rates of learning
suggests that – given the modeled scenario of forest harvesting and agent-agent and
agent-landscape interaction in this study – the forest companies and the conservationist
are able to best meet their objectives at some intermediate level of cooperation. Beyond
this level, the system appears to enter into a state of decline as the forest companies
perhaps relinquish too much of their profits to appease the conservationist. Such
behaviour is detrimental to the forest companies because they are spending more time
focusing on achieving the conservationist’s objectives rather than their own, and as a
result they do not improve their decision-making. This can be observed in the difference
in the overall learning rate of the forest companies for Scenarios 4 and Scenario 5 relative
to other scenarios. When the forest companies’ are uncertain about how to achieve their
objectives, they are always selecting a variety of different stands in order to determine which are best for achieving their objectives. This type of behaviour also prevents the conservationist for achieving its objectives because it is not provided ample opportunity to have its objectives considered in order for it to improve its cognitive maps.

The validation procedures applied in this study are unlike commonly employed statistical metrics that compare the similarity of the model outputs and a dataset that is assumed to be a true representation of reality at a given moment in time. It is unlikely that such a comparison will yield useful information for this study as any real outcome will likely have unfolded differently than the model results. Instead, this study demonstrated that the model can be validated based on the ability to achieve a non-dominated set of solutions for Benchmark scenarios to which the outcomes of complex processes can be compared. This validation technique is appropriate for the complexity-optimization framework because it compares the outcomes of agent interactions with regard to achieving different objectives against outcomes that are considered optimal for a given set of conditions. The ability of agents to achieve their objectives is thus evaluated based on the complexity of the system.

6.8 Conclusion

The model developed in this study provides three outputs that are of important use to forest management. The first is the harvesting solution that is generated in the final episode of the model, which depicts the optimal decision-making behaviour of the forest company agents given their interactions with the conservationist and the government agents. Decision makers can utilize such information for determining if the resulting
spatial patterns conflict with management policies that dictate the spatial constraints of harvesting. Second, the *cognitive maps* provide useful information regarding the relative benefit of harvesting trees in each area of the forest with regards to achieving different objectives. These maps differ from those produced by alternative optimization procedures because they are based on agents’ experience with the dynamics of the landscape and the actions of other agents. The third output of use are the graphs representing the tradeoffs between the different objectives. This information demonstrates to stakeholders to what level their objectives can be achieved given the complexity of the system, and how the achievement of one objective comes at a cost of other objectives. Together, these outputs can be used to improve the ability of forest management to understand how the complex interactions amongst multiple stakeholders influences the ability to achieve different objectives.

This study presents a method for integrating agent-based modeling and reinforcement learning in order to simulate the complex interactions in multi-stakeholder forest management and generate solutions for improving forest harvesting strategies. The model was presented in the context of forest management, but can be extended to analyze other management applications that involve the interactions of different stakeholders for dictating resource outcomes. Reinforcement learning was considered a suitable method for implementing optimization as individual agent behaviour could be sufficiently influenced and analyzed. However, it is unlikely that the integration of agent-based modeling and optimization will be able to fully satisfy the expectations of each approach as defined by their respective fields. It is thus important to understand the necessary tradeoffs between complexity and optimization in order to address specific questions.
6.9 References


7. GENERAL CONCLUSIONS

Forest harvesting is an essential practice in today’s society for generating economic wealth, sustaining levels of employment, and for producing goods that support our livelihood. Since it is not possible to avoid harvesting of the world’s forests, it is necessary to develop methods that improve forest cover patterns in order to achieve economic, ecological and other important objectives. Complex systems theory provides an epistemological approach to understanding the important components of harvesting that influence the emergence of forest cover patterns. Integrating complex systems models with artificial intelligence offers a mechanism for simulating the interactions between such components in order to learn how beneficial forest cover patterns can be achieved.

The purpose of this dissertation was to enhance existing modeling approaches by developing and evaluating the integration of GIS, agent-based modeling and reinforcement learning. The fields of complex systems theory, GIScience, artificial intelligence, heuristic modeling and LUCC offer a broad range of methods that can address these needs. However, ABM and RL were selected due to their immense potential for representing the learning behaviour of stakeholders and their response to economic and environmental variables. Encoding stakeholders as agents in a model produces patterns that are dictated by interactions. Time thus becomes a central element of the model as feedbacks and system perturbations influence emerging forest cover.
Shaping agent cognition with reinforcement learning proves to be advantageous as the actions of individual agents can be reinforced based on the ability to achieve different and at times conflicting objectives. Furthermore, integrating ABM and RL also provides a means for understanding how the results generated from stakeholder interactions compare to optimal forest cover solutions.

The methods developed and presented in this dissertation represent a paradigm shift in the scientific theories and methodologies that are used for informing decision making on forest cover change. Traditional models that build on reductionist assumptions are limited when attempting to understand how various system components interact over time to produce emerging forest cover patterns. Such methods provide top-down strategies that insinuate an ability to optimize our actions in a static world. The implementation of these traditional methods requires the elimination of redundancies in order to focus on the components of the system will lead towards optimal results. Conversely, developing models with agents provides an entirely different insight into how individual behaviours and reactions to various events can produce unexpected results. The attention is shifted towards bottom-up solutions that take into consideration the inevitable uncertainties involved in any management endeavour when different interests exist across varying spatial scales. Constructing agent cognition with artificial intelligence represents how individuals react and remember the consequences of their actions, which can serve to enhance the realism of emerging forest cover patterns. Furthermore, encompassing agents and agent learning in a GIS environment as accomplished in this dissertation provides an explicitly spatial context to understanding the processes and patterns of forest cover change.
7.1 Summary of Findings

The nature in which this dissertation was conducted provided successive building blocks for the integration of ABM and reinforcement learning. The model developed in Chapter 2, based on conventional agent cognition (i.e. reactionary behaviour, demonstrates that emerging forest cover patterns are largely dictated by agent responses to various economic conditions and forest characteristics. Forest company agents were observed to be most successful in maintaining harvesting quotas when timber pricing is high and costs are low, and when they have ample capital to access a variety of timber. While such findings can be considered trivial, the utility of the model is that it demonstrates how different spatial patterns unfold during the simulating of forest harvesting activities in response to varying economic situations. As timber pricing and the cost of harvesting change, so too does the behaviour of forest harvesting companies, which can lead to results that cannot be predicted using conventional modelling methods. Statistical methods, for example, may determine where harvesting should take place over a given period in order for forest companies to meet their objectives by performing a series of procedures that estimate the spatial relationships among the attributes of forest stands. However, circumstances can potentially change over the given period to the extent where those predictions are no longer a viable option. However, the use of the ABM approach demonstrates the potential outcomes given what is known about the dynamics of the system. The dichotomy between statistical and ABM approaches suggests that shaping forest management policy with traditional models that provide top-down solutions may be directly in conflict with how harvesting patterns actually emerge in reality. This finding emphasizes the need to engage in understanding how to think
about forest cover change from harvesting practices as a bottom-up processes where individual behaviours, and not just spatial relationships between components in the forest, are responsible for shaping forest cover patterns.

The results from Chapter 2 also revealed that the size of the licensed areas in which forest companies operate has an influence on their ability to maintain operations during periods when harvesting costs are high and timber prices are low. Forest company agents that have larger license areas were able to withstand unfavourable economic conditions better than those with smaller areas, which is likely due to having a greater number of profitable stands that are more accessible. The importance of this finding is that space plays a role in determining how forest companies can respond to system dynamics as the spatial extent, and likely the spatial configuration, of a license area can aid a company in meeting objectives during varying economic situations. ABM can thus play a significant role in developing forest management policy by determining precisely how license areas should be defined in order to provide all companies with some means of resilience during unfavourable periods.

The agent-based model for simulating forest harvesting patterns was improved upon in Chapter 3 by implementing RL for developing intelligent agents that learn where to harvest trees in order to achieve different objectives at varying scales. In order to focus on RL parameterization, the initial model integration is a departure from typical agent-based models as system complexity was not a priority. As such, the agents in this model are considered a mechanism for implementing optimization behaviour. The issue of exploitation versus exploration in agent searching was addressed by determining agent policies that are suitable for sufficient sampling of the landscape that leads to achieving
objectives. The results also demonstrated that the relatively simple overhead involved with implementing RL facilitates the ability of a single model to be transferred to different spatial scales and to include varying numbers of agents.

The findings from Chapter 3 reveal how agent-based models can be improved upon by incorporating learning behaviour through means of artificial intelligences. Instead of developing agents with only reactionary behaviour, the model was developed to include learning and memory mechanisms that enhance the ability of agents to achieve their objectives. While reactionary agent behaviour is suitable for analyzing how individuals behave in given circumstances, it does not represent the knowledge that individuals gain over time that helps them to improve their decision making skills. Furthermore, the use of learning algorithms also provides a means to create a distinction between each agent’s local objectives and the collective global objectives of all agents. This is an improvement over traditional spatial methods that incorporate all objectives at one scale, and is particularly important in a process such as forest harvesting where ecological and economic goals exist can exist at very different scales.

The intelligent agent approach was further enhanced in Chapter 4 by contextualizing the model as a multi-objective decision-making approach for generating forest cover patterns. While the initial attempts at integrating RL and ABM combined all objectives into a single function for implementing agent learning behaviour, the methods employed in the enhanced model investigate the ability of the agents representing forest companies to achieve each objective independently. The results demonstrate how agents initially make poor decisions regarding their objectives, but over the course of the model run they are able to utilize the knowledge gained in their landscape in order to improve
their behaviour. Furthermore, the parameterization of the model as a multi-objective decision making approach facilitated the analysis of the results with regards to the entire set of solutions that were generated by the model. This demonstrated that the solution deemed most suitable for achieving all objectives belongs to the set of non-dominated solutions, indicating that the RL algorithms can equip agents with the means to achieve improved forest cover strategies.

Complexity was introduced into the model by analyzing how agents’ abilities to derive distinct forest cover solutions is affected by the presence of simulated forest disturbances. The results from this analysis reveal that increasing intensities of system disturbances leads to a decline in the ability of agents to provide distinctive beneficial solutions, but in a non-linear fashion. That is, agent learning can withstand some level of disturbance by learning alternative locations for harvesting in order to achieve their objectives, but beyond some apparent threshold this ability is greatly reduced. This finding sheds light on the fact that tradeoffs exist between complexity and optimization as it becomes increasingly difficult to improve solutions when simulating scenarios with high degrees of system uncertainty. However, the resulting forest cover output maps from the disturbance scenarios still yield useful information to forest managers as they depict the change in benefit of harvesting trees from specific areas given our uncertainty about the system.

The model developed in Chapter 4 provides a novel approach to forest management decision making and to the field of GIScience as it begins to bridge the rigorous elements of optimization with complex systems theory through the explorative capabilities of ABM. As discussed in Chapter 1, optimization models are the founding
computer techniques for informing forest management, and have maintained their relevance due to the potential of heuristic methods that can provide solutions to various forestry-related problems. The success of these models is due to the ability to integrate several objectives together for producing one or more solutions that can be easily interpreted by decision makers. The model developed in Chapter 3 presents an example of an optimization model as three objectives were combined into a single function upon which the model attempts to improve. However, the success of optimization models is based on the assumption that the world is in a steady state, void of any change that can impact how decisions are made. As such, the inclusion of agents and dynamic variables into models of optimization as presented in Chapter 4 is a significant contribution because it introduces a method to investigate the uncertainty in spatial decision making, and also parameterizes optimization based on the objectives of individual agents rather than on the overall goal of the model.

From a GIScience perspective, the bridging of optimization and complexity as accomplished in Chapter 4 is important because it provides a means to improve the information derived from spatio-temporal modelling techniques such as ABM. In the GIScience literature, ABM applications are abound for simulating the complex interactions of spatial decision making behaviours that lead to emerging patterns over time. However, as mentioned in Chapter 1, it has been repeatedly argued that ABM is limited to explorative capabilities that can only investigate ‘what-if’ scenarios because of concerns surrounding the ability to adequately validate results. As such, the inclusion of learning mechanism provides a means for agents to improve upon their decision making abilities in order to draw closer towards optimizing their objectives. This is not to
say that agents behave optimally each time they are faced with a decision. Instead, the bridging of optimization and complexity improves ABM by allowing agents to make the best decisions possible given their knowledge of their surrounding environment and the dynamics of the system.

The potential of the RL-ABM approach to incorporate both space and time was investigated in a single intelligent agent application in Chapter 5 in which a forest company agent navigates through a forest in order to maximize profits under different spatial constraints. Specifically, the concept of path dependency was considered by simulating an agent’s ability to learn where in the landscape to initiate harvesting activities, and where to harvest at each subsequent time step in order to maximize economic returns and minimize the length of roads required to access forest stands. This represents the notion of path dependency because the decision to harvest in a specific location in each time step has a direct consequence on the availability of potential harvestable stands given the objectives of the agent and the defined forest cover constraints. A temporal weighting function was developed that assigns a weight to each stand based on the sequence in which it was harvested, which was integrated into the quality function for updating the quality of each forest stand. The results indicate that RL algorithms can be used for identifying where and when it is most beneficial to harvest trees in order to meet forest cover objectives that are affected by the temporal ordering of events and the spatial constraints of the system.

By incorporating path-dependent knowledge into agent cognition, the model developed in Chapter 5 directly addresses the issue of time in spatial decision making. Time is often treated as only a record of when specific events take place in contemporary
agent-based models, and is not utilized for determining how agents make decisions.

Ignoring the importance of time limits agent cognition to the spatial dimension because there is no consideration of how the timing of decisions influences the outcome of the model. The development of the temporal weighting function in Chapter 5 presents the notion of temporal cognition, whereby agents learn the benefit or detriment of making all decisions at specific moments in time. The temporal weighting function enhances the representation of individuals with computer agents because it incorporates the human ability to understand when certain actions should be taken over others. The model only consisted of a single agent in order to effectively evaluate if time was adequately considered in agent decision making. However, considering the success of the agent to make beneficial decisions, the temporal weighting function can easily be implemented into applications of the model with numerous agents, as is demonstrated in Chapter 6.

The knowledge gained from integrating RL and ABM was utilized in Chapter 6 to develop a multi-stakeholder model in which the complex interactions between different forest company agents with multiple strategies, a conservation agent and a government agent are governed by system dynamics and learning mechanisms as encoded through RL. A conceptual framework for integrating ABM and RL was developed to demonstrate how the complexity of harvesting can be simulated in a manner that allows agents to evolve their understanding of how their actions influence forest cover patterns. The concept of cognitive maps was introduced as a representation of the knowledge that different agents are able to gain within specific system constraints. These maps are accessed at moments during the model when agent input is needed to determine where and when to harvest in order to meet their specific objectives. The results reveal that
certain levels of altruism were found to be more beneficial for agents to gain sufficient experience for improving their decision making behaviour.

The model developed in Chapter 6 provides contributions to research on forest cover change by simulating the interactions amongst different stakeholders and the way in which their goal-seeking actions lead to different landscape patterns. However, rather than simple goal-seeking behaviours that are governed by reactionary cognition, agents encoded with learning mechanisms attempt to improve upon their decisions while contending with the fact that other agents are also improving their decisions. As the results from Chapter 6 demonstrate, the timing of agent actions governed by learning mechanisms has significant implications for the outcome of the process. For example, when the actions of forest company agents are dictated by greedy behaviour (i.e. they only cooperate with the conservationist agent when they are reaching maximum profit potential), the conservationist agent is not included in the process until late in the model simulation. At this point, it is too late for the conservationist agent to gain enough knowledge to be able to reach its objectives to an acceptable level. Conversely, if the forest company agents are highly cooperative, they will never learn how to achieve substantial profits. This also jeopardizes the ability of the conservationist agent to achieve its objectives because it is inconsistently involved in the decision making process. However, if the forest company agents agree to cooperate when profit levels are at an intermediate level, they are receiving enough financial returns to allow them to achieve their objectives while also ensuring that the conservationist agent is consistently included in the decision making process so that it is able to learn how to achieve its own objectives. These finding, as stated above, reveal how cooperative behaviour can be
beneficial for different stakeholders to simultaneously achieve their objectives. The findings also emphasize the utility of the model developed in Chapter 6 for informing policies regarding how stakeholders should engage in negotiations regarding the use of forest resources.

A validation procedure was also introduced in Chapter 6 to evaluate agent behaviour and the model outcomes in comparison to optimal expectations. This is a departure from traditional ABM validation procedures that compare simulated patterns with a dataset representing a known reality. These procedures rely on methods such as the Kappa statistic or landscape metrics to measure the agreement between two datasets, and determine if a model can produce results that are consistent with expectations. However, these methods are inappropriate for validating the RL models developed in this dissertation for two main reasons. First, the datasets used in this dissertation depict relatively large areas at a fine spatial resolution as the forests are represented at the stand level. As such, while the datasets provide a large amount of information, they are very expensive to develop and update. For this reason, the Government of British Columbia, who provided much of the data used within this dissertation, has last updated the forest cover datasets in 1996, and will likely not update them again for some time. Therefore, it is not possible to obtain a dataset to compare against the simulated results. Second, and perhaps more significantly, the purpose of the RL models developed in this study was not to use an agent-based model to simulate a single process of forest harvesting and to validate the results against an expected or observed patterns. The purpose was to integrate artificial intelligence into agent decision making in order improve forest cover solutions and to incorporate complexity into forest cover models to determine how individual
stakeholder behaviour influences the ability of all agents to achieve their objectives. As such, it was deemed more appropriate to improve upon existing validation techniques from the multi-objective decision making literature for evaluating the ability of the developed model for producing acceptable results. The validation method presented in Chapter 6 determines if agents can achieve their objectives given the complexity of the system. Assumptions are made regarding the relative degree to which agents will satisfy their objectives considering the behaviour of all agents and the dynamics of the system, and the model outputs are compared to these assumptions to determine the validity of the model.

7.2 Analyzing and Improving Model Sensitivity

The models developed in Chapters 3 through 6 demonstrate how RL can be utilized with ABM for determining where forest companies should harvest over time and for simulating how multi-stakeholder interactions lead to different forest cover outcomes over time. For either purpose, the development of agent-based models with RL involves the specification of model parameters and equations that are specific to the simulated application. The manner in which the various parameters and equations are defined will influence the type of results that the model produces, and a model that is oversensitive to changes in parameters will most likely not produce consistent findings. As such, the modifications to various RL components through Chapters 3 to 6 were undertaken not only to adapt to different applications, but also in attempt to minimize the sensitivity of the model. In particular, special attention was paid when defining the reward functions, quality functions, the step-size parameter and the search function.
The reward functions developed in Chapters 3 and 4 were based on calculating Z-scores in order to standardize the different objectives before they are combined to produce a single value for the reward. This was considered an acceptable method for calculating rewards because the landscape attributes and the agent search procedures were deemed suitable for meeting the assumptions of a normal distribution. However, standardizing the objectives in this manner resulted in the rewards being sensitive to the number of forest stands that were selected for harvesting. That is, harvesting a greater number of stands with less tree volume as oppose to fewer stands with greater tree volume appeared to provide a better estimate of the true rewards, which is likely due to having a larger sample size. While the exact influence that the Z-score calculations has on the results of the model is not considered to be significant, the reward function was altered in Chapters 5 and 6 by standardizing the objectives using maximum or minimum values derived from the different objectives. This method avoids having the results of the model being influenced by the sample size of harvested stands.

The quality function developed in Chapter 3 was also altered in subsequent chapters in order to reduce the sensitivity of the model. In Chapter 3, the quality values of the harvested stands were updated at each time step of the model. Although updating at each time step is a common procedure in the RL literature, this method means that the quality value of each stand to be somewhat sensitive to the specific time step in which cell was selected. An informal analysis of the results demonstrated that the estimated quality value of earlier harvested stands was closer to their true quality value than those stands selected during later time steps. This was likely a consequence of stands selected at earlier time steps receiving more updates to their value. In order to solve this problem,
the value function in subsequent chapters updated the value of each harvested stand only at the end of each episode to ensure that all harvested stands were updated equally.

Regarding the RL search function, a sensitivity analysis was performed in Chapter 3 to determine the ability of the model to provide acceptable solutions given varying search parameters. The results from this analysis demonstrated that agents are more successful at achieving their objectives if they engage in higher levels of exploration. This means that agents spend a greater period of time randomly selecting forest stands to harvest (i.e. exploration) rather than only utilizing the quality value of each stand in order to make their decisions (i.e. exploitation). This finding was especially true when the RL model was implemented on a relatively large dataset due to the fact that agents required more time to search the significantly larger solution space, which reflects the need to adjust the search function based on the number of potential solutions. The knowledge gained from this sensitivity analysis was utilized for developing a flexible search function that could be easily adjusted in order to ensure that each forest stand in the dataset is sampled a sufficient number of times. This required determining how agents should make the transition from exploration to exploitation, which is challenging due to the fact that an abrupt transition will lead the agents towards local optima (i.e. solutions that are considered significantly less than optimal), while a slow transition will never allow agents to fully learn which stands are most beneficial to harvest. As such, a sigmoid curve was selected to govern the transition from exploration to exploitation because it allowed agents to explore for a prolonged period of time, then gradually move towards increasingly higher levels of exploitation. Furthermore, the equation of the sigmoid curve allowed for it to be easily adjusted in order to satisfy specific sampling requirements.
The final RL parameter that was considered in terms of its influence on model sensitivity was the step-size parameter, also referred to as the rate of learning. This parameter represents a real number between 0.0 and 1.0 that is utilized in the quality function in order to weight the temporal difference in quality values between successive time steps. An informal analysis revealed that the model is sensitive to the step-size parameter in a similar manner as it is to the search function. That is, the model converges towards local optima when the step-size value is relatively low (i.e. close to 0.0), while a value close to 1.0 does not permit agents to fully learn which stands are most beneficial for harvesting. Due to the step-size parameter and search function having similar influences on the model outcomes, the sigmoid function was also utilized for governing how the step-size parameter will change over the course of the model. This entailed having values close to 1.0 during early episodes and then gradually declining the step-size value towards 0.0. Defining the step-size parameter as such results in agents initially gaining knowledge to their full potential as they lack any a priori information regarding the forest stands. As the model progresses, the step-size parameter slowly declines as governed by the sigmoid function to represent that there is increasingly less knowledge to be gained over time.

7.3 Linking Current Work with Future Research

While the modeling approach developed in this dissertation has significant potential for informing forest management, it is important to consider how it can be enhanced through further research. The imminent focus for ensuring successful implementation should be to develop an adequate spatial decision support system in
which interested stakeholders can easily access the information derived from the model. The RL-ABM approach developed in this dissertation produces three outputs that are of significant use to management: (1) the final solution indicating where harvesting activities should take place in order to satisfy the set of defined forest cover objectives; (2) the cognitive maps that demonstrate the relative quality in the forest for achieving the set of objectives; and (3) the non-dominated solutions that reveal the tradeoffs between objectives for the beneficial solutions. It would thus be advantageous to develop an interactive application that allows stakeholders to examine these different outputs in coordination to gain a full understanding of the relationship between specific harvesting strategies and the ability to achieve their objectives.

More generally, future research is also needed for evaluating and improving upon some of the existing challenges with integrating complexity and AI. Specifically, it is important to learn how the use of heuristic algorithms such as RL imposes restrictions on agent behaviour and the emergence of the model outcome. For example, the use of the function that governs the relationship between agent exploitation and exploration (i.e. the rate at which agents access their cognitive maps) assumes that agents gain sufficient knowledge during the model in order for them to optimize their objectives during the final model episodes. However, constructing the model as such challenges the ability of agents to react to novel landscape circumstances that can emerge during the modeled process. This challenge can be addressed in future implementation by exploring how agent cognition can be formulated to respond to varying levels of complexity. Conversely, incorporating system complexity into a heuristic modeling approach limits the potential for generating a range of solutions that can be useful for informing forest
management. Heuristic methods now exist that are able to produce a full set of plausible non-dominated solutions to a given multi-objective problems that provide stakeholders with the necessary means to evaluate the full tradeoffs existing between different objectives. However, the presence of complexity demonstrated that only a sub-set of the non-dominated solutions could be produced because model outcomes were dictated by agent interactions rather than the need to generate a diverse range of solutions.

Rather than viewing these issues as hindrances to the successful implementation of the methods developed in this dissertation, it is more appropriate to acknowledge the challenges of integrating two distinctively different methods for generating forest cover patterns. It is not possible to adhere to all expectations set forth by both complex systems theory and AI because certain concessions need to be made when attempting to harmonize notions of emergence and optimization within the same model. As such, future research in integrating complex systems theory and AI should examine the inevitability of tradeoffs with regards to the overall purpose of the model, the process being simulated, and the intended use of the results. Such considerations will facilitate the useful implementation of agent-based modeling and reinforcement learning or other forms of AI for addressing issues of forest cover change and LUCC in general.

7.4 Research Contributions

This dissertation investigates the integration of complex systems theory and AI for improving models of forest cover change through the use of GIS, ABM and RL, and thus offers contributions to different scientific fields. First, this dissertation adds to the existing research in GIScience by improving agent-based models for understanding
spatial dynamic phenomena. As conventional agent cognition relies on reactionary, rational and bounded-rational behaviour, the research conducted herein provides the additional ability for agents to learn about the relationship between their actions and how their objectives are achieved. Encoding learning procedures enhances how individual behaviour is represented by computers, and provides a means for generating solutions that are more advanced at achieving objectives. Furthermore, RL was utilized in this dissertation by allowing agents to improve their decision making behaviour in repetitive simulations of processes. However, RL can also be used for agents to learn within a single simulation of a LUCC process by reinforcing their actions at each time step of the model in order to improve decisions made at later time steps.

This dissertation also provides contributions to research in forest management as it presents a novel approach for deriving forest cover patterns for achieving economic and ecological objectives. While conventional forest models have evolved to generate patterns that can achieve stakeholder objectives, it is difficult to understand if such patterns are likely to manifest without incorporating stakeholder behaviour and acknowledging the dynamics of the system. The top-down models of contemporary forest management are ill-equipped to understand the presence of uncertainty as a result of social or natural processes that can influence the availability and accessibility of timber as well the behaviours of different stakeholders. That is, contemporary models fail to acknowledge the coupling of human and ecological processes in order to understand forest management as a complex system. Furthermore, the methods developed in this dissertation enhance forest management decision making by presenting the model results in different ways. Rather than simply providing suggested forest cover strategies, the use
of RL also produces the cognitive maps that agents use to guide their actions. These maps can be utilized by management to learn the relative benefit of harvesting trees in different areas of the forest for meeting specific objectives. In addition, implementing a heuristic modeling approach allows the results to be depicted as the tradeoffs between different objectives resulting from system dynamics. Forest management or interested stakeholders can utilize such information for determining how certain objectives are achieved at the expense of others, and for evaluating how different behaviours, such as varying levels of altruism, influence the ability to optimize objectives.

Finally, this dissertation adds to the growing body of research concerning LUCC that is concerned with understanding the linkages between human activities and resulting land cover change. In this sense, the methods presented can provide a computer simulation laboratory for observing how different behaviours interact with socio-economic and biophysical processes involved in forest harvesting. The implementation of RL for enhancing agent cognition will lead to more realistic representations of stakeholder behaviour that can improve how research understands the relationship between how forests are used and resulting forest cover patterns.

The world’s forests provide an important resource upon which the livelihood of humans and countless other organisms depend. Harvesting these forests is essential for sustaining economies and for providing us with products that help to maintain and improve our standards of living. However, harvesting also has negative societal and ecological impacts as the removal of forests can threaten the existence of communities and ecosystems. As such, it is important to understand how we can maintain forest harvesting while acknowledging harmful consequences. This dissertation has attempted
to address these issues by providing a novel approach to developing forest cover patterns that incorporate different objectives while understanding amongst the dynamic forces that influence management decision-making and impact forest cover change over time.
APPENDICES
Appendix A: Co-authorship Statement

Chapter 2
Title: Evaluating Spatio-temporal Complexities of Forest Management: An Integrated Agent-based Modeling and GIS Approach
Co-author: Suzana Dragićević
Role of co-author: Manuscript preparation

Chapter 3
Title: Defining Transition Rules with Reinforcement Learning for Modeling Land Cover Change
Co-author: Suzana Dragićević
Role of co-author: Manuscript preparation

Chapter 4
Title: GIS and Intelligent Agents for Multiobjective Natural Resource Allocation: A Reinforcement Learning Approach
Co-author: Suzana Dragićević
Role of co-author: Manuscript preparation

Chapter 5
Title: Incorporating Path Dependent Knowledge in an Intelligent Agent Model for Natural Resource Management
Co-author: Suzana Dragićević
Role of co-author: Manuscript preparation

Chapter 6
Title: Simulation and Validation of a Reinforcement Learning Agent-based Model for Multi-Stakeholder Forest Management
Co-author: Suzana Dragićević
Role of co-author: Manuscript preparation
Appendix B: Copyright Clearance from Springer

Copyright clearance for the use of the paper "Evaluating Spatio-temporal Complexities of Forest Management: An Integrated Agent-based Modeling and GIS Approach" as published in the journal *Environmental Modeling and Assessment*:

----- Original Message ----- From: "Permissions Europe/NL" <Permissions.Dordrecht@springer.com> To: cbone@sfu.ca Sent: Monday, 18 May, 2009 03:04:32 GMT -08:00 US/Canada Pacific Subject: RE: Customer question from springer.com

Dear Mr. Bone,

With reference to your request (copy herewith) to reprint material on which Springer Science and Business Media controls the copyright, our permission is granted, free of charge, for the use indicated in your enquiry.

This permission
- allows you non-exclusive reproduction rights throughout the World.
- permission includes use in an electronic form, provided that content is
  * password protected;
  * at intranet;
- excludes use in any other electronic form. Should you have a specific project in mind, please reapply for permission.
- requires a full credit (Springer/Kluwer Academic Publishers book/journal title, volume, year of publication, page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) to the publication in which the material was originally published, by adding: with kind permission of Springer Science and Business Media.

The material can only be used for the purpose of defending your dissertation, and with a maximum of 40 extra copies in paper.

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Kind regards,
Estella Jap A Joe
Springer
Rights & Permissions
__
Van Godewijckstraat 30 | 3311 GX
P.O. Box 990 | 3300 AZ
Dordrecht | The Netherlands
fax +31 (0) 78 657 6377
estella.japajoe@springer.com
**************************************************
To whom it may concern,

I am writing to ask permission to publish a paper that has been accepted by Environmental Modeling And Assessment in my Ph.D. dissertation. The details of the paper are as follows:


Thank you for your time, and I look forward to hearing from you.

Sincerely,
Christopher Bone
Appendix C: Copyright Clearance from Sage Publishing

Copyright clearance for the use of the paper "Defining Transition Rules with Reinforcement Learning for Modeling Land Cover Change" as published in the journal Simulation:

Under the terms of your contributor agreement, without seeking permission, you may:

- At any time, distribute on a not-for-profit basis photocopies of the published article for your own teaching needs or to supply on an individual basis to research colleagues.
- At any time, circulate or post on any repository or website the version of the article that you submitted to the journal (i.e. the version before peer-review) or an abstract of the article.
- At least 12 months after publication, post on any non-commercial repository or website the version of your article that was accepted for publication.
- At least 12 months after publication, re-publish the whole or any part of the Contribution in a printed work written, edited or compiled by you provided reference is made to first publication by SAGE/SOCIETY.

When posting or re-using the article, please provide a link/URL from the article posted to the SAGE Journals Online where the article is published: http://online.sagepub.com and please make the following acknowledgment ‘The final, definitive version of this paper has been published in <journal>, Vol/Issue, Month/Year by <<SAGE Publications Ltd.>>/<<SAGE Publications, Inc.>>, All rights reserved. © [as appropriate]

The licenses granted above in this paragraph are expressly made subject to and limited by the following restrictions:

- The SAGE-created PDF of the published Contribution may not be posted at any time.
- In each instance of use of the Contribution, or any part of it, must include the copyright notice that appears on the issue of the Journal in which the Contribution is first published and a full bibliographic citation to the Journal as published by SAGE;
- Copies of the Contribution, or any part of it, shall not be sold, distributed, or reproduced for commercial purposes (i.e., for monetary gain on Contributor’s own account or on that of a third party, or for indirect financial gain by a commercial entity);
• The Contribution, or any part of it, shall not be used for any systematic external
distribution by a third party (e.g., a listserve or database connected to a public
access server).
Appendix D: Copyright Clearance from Wiley-Blackwell

Copyright clearance for the use of the paper "GIS and Intelligent Agents for Multiobjective Natural Resource Allocation: A Reinforcement Learning Approach" as published in the journal Transactions in GIS:

----- Original Message ----- From: "Permission Requests - UK" <permissionsuk@wiley.com> To: "Chris Bone" <cbone@sfu.ca> Sent: Friday, 5 June, 2009 06:12:44 GMT -08:00 US/Canada Pacific Subject: RE: Permission to Publish Paper

Dear Christopher Bone,

Thank you for your email request. Permission is granted for you to use the material below for your thesis/dissertation subject to the usual acknowledgements and on the understanding that you will reapply for permission if you wish to distribute or publish your thesis/dissertation commercially.

Best wishes,

Laura Mothersole

Permissions Assistant
Wiley-Blackwell
9600 Garsington Road
Oxford OX4 2DQ
UK
Tel: +44 (0) 1865 476160
Fax: +44 (0) 1865 471158
Email: laura.mothersole@wiley.com
-----Original Message-----
From: Chris Bone [mailto:cbone@sfu.ca]
Sent: 07 May 2009 18:50
To: Permission Requests - UK
Subject: Permission to Publish Paper

To whom it may concern,

I am writing to ask permission to publish a paper of mine in its entirety in my Ph.D. dissertation. The details of the paper are as follows:

Bone, C., Dragićević, S., Accepted . GIS and intelligent agents for natural resource allocation: a reinforcement learning approach. Transactions in GIS.

Thank you for your time, and I look forward to hearing from you.

Sincerely,
Christopher Bone