OPTIMIZATION APPROACHES TO BIOLOGICAL INVASIONS AND CANCER

A Dissertation by
Eyyub Yunus Kibis

Master of Science, University of Houston Clear Lake, 2013
Master of Arts, University of Houston, 2011
Bachelor of Science, Bogazici University, 2009

Submitted to the Department of Industrial, Systems and Manufacturing Engineering
and the faculty the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2017
OPTIMIZATION APPROACHES TO BIOLOGICAL INVASIONS AND CANCER

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

______________________________________
İ. Esra Büyükahtakín, Committee Chair

______________________________________
Ehsan Salari, Committee Member

______________________________________
Abu Masud, Committee Member

______________________________________
Krishna Krishnan, Committee Member

______________________________________
Katie Mitchell-Koch, Committee Member

Accepted for the College of Engineering

______________________________________
Royce Bowden, Dean

Accepted for the Graduate School

______________________________________
Dennis Livesay, Dean
DEDICATION

To my beloved family
ACKNOWLEDGMENTS

I would like to thank all who encouraged me in many different ways during the course of my graduate study. I especially thank my doctoral advisor, Dr. Esra Büyüktahtakın to whom I am indebted for helping me enter the world of mathematical programming. Without her guidance and numerous discussions to improve the material in this dissertation, it would not have been possible.

I thank all my family, especially my wife, Zumrut, and my sisters for their endless support during my graduate study.

I am also grateful to my laboratory mate, Halil Cobuloglu, for sharing his practical knowledge and mentoring me at the beginning of my PhD study. His help was crucial to improving my thesis work. I would also like to thank Mostafa Hassan for his willingness to discuss cancer research.

I would like to acknowledge Dr. Greg Houseman from Biological Sciences Department at Wichita State University and Dr. Robert Haight, Dr. Kathleen Knight, and Charlie Flower from United States Department of Agriculture Forest Service for their help in providing data for my research. Finally, I gratefully acknowledge the support of the National Science Foundation CAREER Award, CBET #1554018 and the USDA Forest Service Grant.
ABSTRACT

Spatio-temporal models have been utilized in a wide range of disciplines to describe and predict spatially explicit processes that change over time. One of the various application areas of spatio-temporal models is ecological studies, specifically invasive species management (ISM) over large landscapes where scarce resources, such as budget, can be a limiting factor for controlling biological invasions. Another application area of spatio-temporal models is cancer treatment, where size, growth, and spread of cancer cells are tracked over time. However, the main challenge with spatio-temporal models is the high complexity of the problem in which model size expands exponentially as spatial and temporal dimensions are increased. Furthermore, incorporating growth and spread dynamics of invasive species, or cancer cells, significantly complicates the problem in terms of its solvability and solution time.

In this dissertation, we develop new spatio-temporal mathematical models and optimization-based solution algorithms for determining the optimal strategies to control invasive species and cancer growth. Specifically, we present nonlinear, mixed-integer, and stochastic programming models considering the detailed ecological characteristics of invasive species to analyze their economic impacts. In addition, we develop a spatio-temporal model to determine an optimal breast cancer treatment plan and sequence considering surgery, radiation therapy, and chemotherapy in combination with the optimal dose schedules for chemo- and radiotherapy treatments. In order to increase the solvability of the large-scale problems and reduce the solution time for instances that involve higher spatial and temporal dimensions, we develop linearization approaches and new cutting planes. The results of this dissertation provide practical insights into ISM and cancer treatment planning.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Motivation and Contributions</td>
<td>9</td>
</tr>
<tr>
<td>1.2. Summary of Chapters</td>
<td>11</td>
</tr>
<tr>
<td>1.3. Main Organization of Dissertation</td>
<td>14</td>
</tr>
<tr>
<td>1.4. References</td>
<td>15</td>
</tr>
<tr>
<td>2. REVIEW OF OPTIMIZATION MODELS ON INVASIVE SPECIES MANAGEMENT</td>
<td>19</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>19</td>
</tr>
<tr>
<td>2.2. Overview of the Literature</td>
<td>20</td>
</tr>
<tr>
<td>2.3. Literature Review</td>
<td>21</td>
</tr>
<tr>
<td>2.3.1. Studies on Invasive Species Control</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2. Studies Including Invasive Species Surveillance</td>
<td>25</td>
</tr>
<tr>
<td>2.4. Discussion and Conclusion</td>
<td>27</td>
</tr>
<tr>
<td>2.5. References</td>
<td>27</td>
</tr>
<tr>
<td>3. AGE-STRUCTURED BIO-ECONOMIC MODEL OF INVASIVE SPECIES MANAGEMENT: INSIGHTS AND STRATEGIES FOR OPTIMAL CONTROL</td>
<td>31</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>31</td>
</tr>
<tr>
<td>3.2. Mathematical Model</td>
<td>33</td>
</tr>
<tr>
<td>3.3. Case Study: Control of Sericea Invasion in the Great Plains</td>
<td>38</td>
</tr>
<tr>
<td>3.4. Results</td>
<td>40</td>
</tr>
<tr>
<td>3.4.1. Growth Behavior of Different Age Groups over 45 Years</td>
<td>40</td>
</tr>
<tr>
<td>3.4.2. Impact of Eradication Rate on Economic Damages</td>
<td>43</td>
</tr>
<tr>
<td>3.4.3. Impact of Dispersal Rate on Economic Damages</td>
<td>46</td>
</tr>
<tr>
<td>3.4.3. Impact of Different Treatment Strategies on Economic Damages</td>
<td>48</td>
</tr>
<tr>
<td>3.5. Discussion</td>
<td>50</td>
</tr>
<tr>
<td>3.6. References</td>
<td>54</td>
</tr>
<tr>
<td>4. OPTIMIZING INVASIVE SPECIES MANAGEMENT: A MIXED-INTEGER LINEAR PROGRAMMING APPROACH</td>
<td>58</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>58</td>
</tr>
<tr>
<td>4.1.1. Progress in Spatial Optimization Models for Invasive Species</td>
<td>62</td>
</tr>
<tr>
<td>4.2. Model</td>
<td>64</td>
</tr>
<tr>
<td>4.2.1. Mathematical Formulation</td>
<td>65</td>
</tr>
<tr>
<td>4.2.2. Dispersal Uncertainty</td>
<td>70</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>4.2.3. Linearization of Nonlinear Constraints</td>
<td>71</td>
</tr>
<tr>
<td>4.2.3.1. Proposed Linearization Method</td>
<td>72</td>
</tr>
<tr>
<td>4.2.3.2. Method 2: Piecewise Linearization Technique</td>
<td>73</td>
</tr>
<tr>
<td>4.2.3.3. Method 3 (Vielma &amp; Nemhauser [18])</td>
<td>74</td>
</tr>
<tr>
<td>4.2.3.4. Method 4 (Li &amp; Yu [38])</td>
<td>75</td>
</tr>
<tr>
<td>4.2.3.5. Method 5 (Williams [39])</td>
<td>75</td>
</tr>
<tr>
<td>4.3. Model Application</td>
<td>76</td>
</tr>
<tr>
<td>4.4. Results</td>
<td>79</td>
</tr>
<tr>
<td>4.4.1. Impact of Different Budget Allocations on Overall Damage</td>
<td>79</td>
</tr>
<tr>
<td>4.4.2. Verification of Model: Comparison of MIP, MINLP, and NLP Models</td>
<td>82</td>
</tr>
<tr>
<td>4.4.3. Treatment Decisions over Changing Budget Allocations and Initial Invasion Scenarios</td>
<td>85</td>
</tr>
<tr>
<td>4.5. Discussion and Conclusions</td>
<td>91</td>
</tr>
<tr>
<td>4.6. References</td>
<td>96</td>
</tr>
<tr>
<td>5. NEW MULTI-STAGE STOCHASTIC PROGRAMMING MODEL AND CUTTING PLANES FOR OPTIMAL SURVEILLANCE AND CONTROL OF EMERALD ASH BORER IN CITIES</td>
<td>101</td>
</tr>
<tr>
<td>5.1. Introduction</td>
<td>101</td>
</tr>
<tr>
<td>5.2. Model</td>
<td>106</td>
</tr>
<tr>
<td>5.2.1. Problem Definition</td>
<td>106</td>
</tr>
<tr>
<td>5.2.2. MSS-MIP Model</td>
<td>110</td>
</tr>
<tr>
<td>5.2.2.1. Scenario Tree</td>
<td>110</td>
</tr>
<tr>
<td>5.2.2.2. Notations</td>
<td>111</td>
</tr>
<tr>
<td>5.2.2.3. Mathematical Model</td>
<td>113</td>
</tr>
<tr>
<td>5.2.2.4. Cutting Planes for Linearization Variables</td>
<td>119</td>
</tr>
<tr>
<td>5.3. Data Calibration and Model Application</td>
<td>122</td>
</tr>
<tr>
<td>5.3.1. Data Calibration</td>
<td>122</td>
</tr>
<tr>
<td>5.3.2. Model Application</td>
<td>126</td>
</tr>
<tr>
<td>5.4. Results</td>
<td>129</td>
</tr>
<tr>
<td>5.4.1. Model Validation</td>
<td>129</td>
</tr>
<tr>
<td>5.4.2. Computational Results with Cuts</td>
<td>132</td>
</tr>
<tr>
<td>5.4.3. Impact of Surveillance on Net Benefits of Ash Trees</td>
<td>135</td>
</tr>
<tr>
<td>5.4.4. Rule of Thumb for Treatment and Removal</td>
<td>137</td>
</tr>
<tr>
<td>5.4.5. Cost vs Benefit Analysis</td>
<td>139</td>
</tr>
<tr>
<td>5.4.6. Sensitivity Analysis on Impact Rates</td>
<td>140</td>
</tr>
<tr>
<td>5.5. Discussion and Conclusion</td>
<td>141</td>
</tr>
<tr>
<td>5.6. References</td>
<td>144</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.</td>
<td>OPTIMIZING MULTIMODAL CANCER TREATMENT: MIXED-INTEGER PROGRAMMING</td>
<td></td>
</tr>
<tr>
<td></td>
<td>APPROACH</td>
<td>147</td>
</tr>
<tr>
<td>6.1.</td>
<td>Introduction</td>
<td>147</td>
</tr>
<tr>
<td>6.2.</td>
<td>MIP Models</td>
<td>153</td>
</tr>
<tr>
<td>6.2.1.</td>
<td>Notations</td>
<td>153</td>
</tr>
<tr>
<td>6.2.2.</td>
<td>Mathematical Model</td>
<td>155</td>
</tr>
<tr>
<td>6.2.2.1.</td>
<td>Treatment Method 1: Chemotherapy</td>
<td>157</td>
</tr>
<tr>
<td>6.2.2.2.</td>
<td>Treatment Method 2: Radiotherapy</td>
<td>161</td>
</tr>
<tr>
<td>6.2.2.3.</td>
<td>Treatment Method 3: Surgery</td>
<td>165</td>
</tr>
<tr>
<td>6.3.</td>
<td>Model Application</td>
<td>166</td>
</tr>
<tr>
<td>6.4.</td>
<td>Numerical Results</td>
<td>170</td>
</tr>
<tr>
<td>6.4.1.</td>
<td>Validation of Cancer Cell Growth: Comparison of Predicted PLA Values</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>and Simulated Gompertz Growth</td>
<td></td>
</tr>
<tr>
<td>6.4.2.</td>
<td>Optimal Sequence of Surgery, Radiotherapy, and Chemotherapy</td>
<td>173</td>
</tr>
<tr>
<td>6.4.3.</td>
<td>Alpha/Beta (α/β) Ratios</td>
<td>177</td>
</tr>
<tr>
<td>6.4.4.</td>
<td>Chemotherapy Treatment Efficiency and Dosage</td>
<td>178</td>
</tr>
<tr>
<td>6.5.</td>
<td>Discussion</td>
<td>180</td>
</tr>
<tr>
<td>6.6.</td>
<td>Conclusion and Future Work</td>
<td>183</td>
</tr>
<tr>
<td>6.7.</td>
<td>References</td>
<td>185</td>
</tr>
<tr>
<td>6.8.</td>
<td>Appendix</td>
<td>189</td>
</tr>
<tr>
<td>6.8.1.</td>
<td>Parameter Estimation: Threshold Rate for Cell Migration</td>
<td>189</td>
</tr>
<tr>
<td>6.8.2.</td>
<td>Parameter Estimation: Spread Rate</td>
<td>190</td>
</tr>
<tr>
<td>7.</td>
<td>CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS</td>
<td>192</td>
</tr>
<tr>
<td>7.1.</td>
<td>Summary of Contributions</td>
<td>192</td>
</tr>
<tr>
<td>7.2.</td>
<td>Future Research Directions</td>
<td>195</td>
</tr>
<tr>
<td>7.3.</td>
<td>References</td>
<td>196</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1</td>
<td>Classification of studies</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Initial population structure and parameters for sericea invasion</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Initial population structure and parameters</td>
<td>77</td>
</tr>
<tr>
<td>4.2</td>
<td>Total damage in seven-year period based on different scenarios and budget allocations, including corresponding M values, solution times, and optimality gaps</td>
<td>80</td>
</tr>
<tr>
<td>4.3</td>
<td>Total damage in seven-year period based on different budget and linearization methods for low invasion</td>
<td>85</td>
</tr>
<tr>
<td>5.1</td>
<td>Statistical analysis used to compare model predictions and actual data in site MB</td>
<td>126</td>
</tr>
<tr>
<td>5.2</td>
<td>Initial population structure and parameters</td>
<td>127</td>
</tr>
<tr>
<td>5.3</td>
<td>Paired-t-test results for comparison of real and predicted number of dead trees</td>
<td>131</td>
</tr>
<tr>
<td>5.4</td>
<td>Comparison of Solution Performances</td>
<td>134</td>
</tr>
<tr>
<td>5.5</td>
<td>Sensitivity Analysis on Impact Rates</td>
<td>141</td>
</tr>
<tr>
<td>6.1</td>
<td>Breast tumor and normal tissue parameters</td>
<td>167</td>
</tr>
<tr>
<td>6.2</td>
<td>Solution time and optimality gap for each treatment modality</td>
<td>171</td>
</tr>
<tr>
<td>6.3</td>
<td>ANOVA test for comparison of PLA equations and Gompertz growth</td>
<td>172</td>
</tr>
<tr>
<td>6.4</td>
<td>Linear regression results of Gompertz growth vs. PLA equations</td>
<td>173</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Representation of spatially explicit landscape</td>
<td>5</td>
</tr>
<tr>
<td>3.1</td>
<td>Field-level invasion by sericea for three age classes over 45 years in absence of control measures under initial conditions</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>Tradeoff between average yearly damages and budget allocation for different eradication rates over 15 years under initial conditions</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Tradeoff between average yearly damages and budget allocation for different dispersal rates over 15 years under initial conditions</td>
<td>47</td>
</tr>
<tr>
<td>3.4</td>
<td>Impact of different treatment frequency strategies on average yearly damage (presented on y-axis) and total (cumulative) damage</td>
<td>49</td>
</tr>
<tr>
<td>4.1</td>
<td>Representation of population dynamics and control of invasive species</td>
<td>65</td>
</tr>
<tr>
<td>4.2</td>
<td>Representation of piecewise linear function</td>
<td>74</td>
</tr>
<tr>
<td>4.3</td>
<td>Budget used and yearly damage with and without treatment each year for low initial invasion with respect to total allocated budget</td>
<td>86</td>
</tr>
<tr>
<td>4.4</td>
<td>Budget used and yearly damage with and without treatment each year for medium initial invasion with respect to total allocated budget</td>
<td>88</td>
</tr>
<tr>
<td>4.5</td>
<td>Budget used and yearly damage with and without treatment each year for high initial invasion with respect to total allocated budget</td>
<td>89</td>
</tr>
<tr>
<td>4.6</td>
<td>Treatment locations and sericea dispersal after treatment</td>
<td>90</td>
</tr>
<tr>
<td>5.1</td>
<td>Representation of EAB model with five time periods and four infestation levels</td>
<td>109</td>
</tr>
<tr>
<td>5.2</td>
<td>Representation of decision tree in multi-stage form</td>
<td>111</td>
</tr>
<tr>
<td>5.3</td>
<td>Maumee Bay State Park in Toledo, Ohio. Plots for impact and spread rate estimations</td>
<td>124</td>
</tr>
<tr>
<td>5.4</td>
<td>Representation of site MB with plots 1, 2, and 3</td>
<td>124</td>
</tr>
</tbody>
</table>
LIST OF FIGURES (continued)

Figure | Description |
--- | --- |
5.5 | Comparison of real data and predicted results for infestation level 4 over seven-year period for plots 1, 2, 3, and overall site MB, respectively |
5.6 | Comparison of real data and predicted results for infestation level 4 over seven-year period for sites 1 to 9 |
5.7 | Comparison of real data and predicted results for infestation level 4 over seven-year period for each plot in site 3 |
5.8a | Net benefits and costs for base case scenario |
5.8b | Marginal net benefits and marginal cost of surveillance for base case scenario |
5.8c | Net benefits and costs for base case scenario with ample budget |
5.9a | Total number of susceptible and infested trees over the years |
5.9b | Total number of treated and removed trees over the years |
5.10 | Net benefit and control costs adjustment for different budget allocations for base case scenario |
6.1 | Piecewise linear approximation of equation (6.1) |
6.2 | Piecewise linear approximation of equation (6.24) |
6.3 | Pointwise function |
6.4 | Three-dimensional representation of initial tumor population |
6.5 | Comparison of predicted and simulated Gompertz growth |
6.6 | Radiation fractionations over 25 days after surgery is performed with different effectiveness rates at the tumor margin |
6.7a | Optimal radiation fractionation after surgery in SRC regimen |
6.7b | Optimal chemotherapy drug concentration after surgery in SCR regimen
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.8a</td>
<td>Optimal radiation fractionation in RC regimen</td>
<td>176</td>
</tr>
<tr>
<td>6.8b</td>
<td>Optimal chemotherapy drug concentration</td>
<td>176</td>
</tr>
<tr>
<td>6.8c</td>
<td>Optimal chemotherapy drug concentration in CR regimen</td>
<td>176</td>
</tr>
<tr>
<td>6.8d</td>
<td>Radiation fractionation after chemotherapy treatment</td>
<td>178</td>
</tr>
<tr>
<td>6.9a</td>
<td>Comparison of radiation fractionations for the sensitivity values of $\alpha / \beta = [3;5]$ and $\alpha / \beta = [12;20]$</td>
<td>178</td>
</tr>
<tr>
<td>6.9b</td>
<td>Cancer cell population changes with respect to different sensitivity ratios</td>
<td>178</td>
</tr>
<tr>
<td>6.10</td>
<td>Cancer cell population change with respect to different chemotherapy drug efficiencies</td>
<td>179</td>
</tr>
<tr>
<td>6.11</td>
<td>Cancer cell population change with weekly (200 mg/m$^2$/week) and once in every three weeks (600 mg/m$^2$/three-weeks) chemotherapy regimens</td>
<td>180</td>
</tr>
<tr>
<td>6.A1</td>
<td>Representation of a voxel and tumor growth</td>
<td>189</td>
</tr>
<tr>
<td>6.A2</td>
<td>Migration of cancer cells to the six immediate neighbors</td>
<td>190</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>B&amp;B</td>
<td>Branch-and-Bound</td>
<td></td>
</tr>
<tr>
<td>BED</td>
<td>Biological Effective Dose</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Chemotherapy</td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>Dynamic Programming</td>
<td></td>
</tr>
<tr>
<td>EAB</td>
<td>Emerald Ash Borer</td>
<td></td>
</tr>
<tr>
<td>IP</td>
<td>Integer Programming</td>
<td></td>
</tr>
<tr>
<td>ISM</td>
<td>Invasive Species Management</td>
<td></td>
</tr>
<tr>
<td>LINAC</td>
<td>Linear Accelerator</td>
<td></td>
</tr>
<tr>
<td>LP</td>
<td>Linear Programming</td>
<td></td>
</tr>
<tr>
<td>MB</td>
<td>Maumee Bay State Park</td>
<td></td>
</tr>
<tr>
<td>MIP</td>
<td>Mixed-Integer Programming</td>
<td></td>
</tr>
<tr>
<td>MINLP</td>
<td>Mixed-Integer Nonlinear Programming</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>Mean Sum of Squares</td>
<td></td>
</tr>
<tr>
<td>MSS-MIP</td>
<td>Multi-Stage Stochastic Mixed-Integer Programming</td>
<td></td>
</tr>
<tr>
<td>MSSP</td>
<td>Multi-Stage Stochastic Programming</td>
<td></td>
</tr>
<tr>
<td>NLP</td>
<td>Nonlinear Programming</td>
<td></td>
</tr>
<tr>
<td>OAR</td>
<td>Organ-at-Risk</td>
<td></td>
</tr>
<tr>
<td>PLA</td>
<td>Piecewise Linear Approximation</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Radiation Therapy</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>Remaining Population Cut</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Surgery</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>SDP</td>
<td>Stochastic Dynamic Programming</td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>Stochastic Programming</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>Sum of Squares</td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Transition Population</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Invasive species are one of the world’s most serious threats to the environment. Although invasive species are non-native (or alien) to their new environments, not all non-native species can be called invasive [1]. While some non-native species die off naturally, some remain and adapt to the environment without causing any damage. On the other hand, some non-native species, (for example, sericea lespedeza, emerald ash borer (EAB), kudzu, and European starling) thrive in their new environment and damage ecosystems and biodiversity, consequently causing economic loss in different entities such as agriculture, international trade, forestry, fishery, power production, and human health [2].

Invasive species can be introduced to their new environments in several ways. While some invasive species such as snails, slugs, and beetles may be unintentionally transported through international trade, some species might be transported intentionally for animal food, forage, and soil-erosion prevention. Nonetheless, invasive species cause significant economic loss to the world economy. The Office of Technology Assessment of the U.S. Congress [3] reported that 59% of imported species between 1906 and 1991 has resulted in economic damage of $97 billion, with a worst-case scenario of $137 billion in total cost to agriculture, forestry, fishery, utilities, and natural areas. In a more recent study, Pimentel et al. [4] report that the cost of invasive species to the United States economy is around $138 billion every year and increasing steadily. In global scale, worldwide costs of invasive species such as plants, animals, and microbes have been estimated at US$1.4 trillion, or 5% of the global gross domestic product [5].
Eradicating invasion is the utmost purpose of the invasive species treatment. However, considering the scarce resources and the abundance of the invasion, a realistic approach would be to reduce the abundance of the invasion in order to allow space for native species to thrive. Treatment methods may vary depending on the type of the invasive species and its abundance. In this dissertation, we concentrate on terrestrial invasive species such as invasive weeds and insects. Therefore, the most widely used treatment methods for the aforementioned species include chemical, mechanical, and biological control as well as prescribed fire for invasive weeds [6].

Chemical control refers to the utilization of pesticides, which destroy or repel any pest including insects, weeds, and microorganisms [7]. The choice of pesticide depends on the abundance, type, and stage of growth of invasive species, as well as invasion distance to native species and close proximity to water resources [8,9]. Mechanical treatment includes techniques such as cutting, mowing, chopping, etc. to provide the immediate removal of invasive species. While mechanical treatment is an effective method, it can be labor intensive if the invasion is widely spread. Note that mechanical treatment should be conducted carefully to prevent the spread or reinvasion of weed species. Therefore, it is recommended to use pesticide followed by mechanical treatment to avert spread and reinvasion [10].

Biological control refers to the utilization of animals or specially bred species that could be the predators of the invader in order to control the invasion [11]. Biological treatment aims at targeting specific invasive species without damaging the ecosystem. For example, grazing has been used as a biological control of some invasive weeds such as kudzu and musk thistle. On the other hand, it is also argued that specially bred species can have negative impacts on the
environment by targeting native species. Therefore, biological treatment requires federal and state permits before implementation and constant surveillance afterwards [12].

Finally, prescribed fire is used as a key tool to reduce invasive species populations and stimulate the growth of native species, thereby restoring habitat [6]. Prescribed fire should be applied in small areas in order to control it. On the other hand, prescribed fire can also stimulate the growth of dormant invasive weed seeds by exposing them to the resources required for germination and growth [13]. Therefore, prescribed fire should be conducted several years in a row to minimize the risk of regeneration of invasive weeds.

Conservation of ecology and the ecosystem can highly benefit from the application of optimization methods such as linear programming (LP), nonlinear programming (NLP), mixed-integer programming (MIP), and stochastic programming (SP), which could integrate invasive species dynamics, landscape characteristics, costs, and management options in a mathematical model. Several optimization studies have been proposed to control, eradicate, or manage invasive species due to their high economic costs [14-21]. However, invasive species management (ISM) is a complex problem because various factors contribute to a species’ invasiveness. First, limited economic resources restrict management efforts to control invasive species. Second, each invasive species demonstrates distinct population and growth dynamics, and exhibits uncertainty in its growth and spatial dispersal characteristics rates as well as impacts on native species and the environment. The general characteristics of invasive species and their management can be summarized as follows:

- Age-dependent fecundity: Reproductive rate of each plant measured by the number of seeds produced, based on the age of the plant.
• Age-dependent mortality: Death rate of a given population class with respect to its age.
• Seed dispersal: Movement or spread of seeds away from the origin or parent plant.
• Seed bank: Natural storage of seeds under the soil. For example, sericea lespedeza seeds can remain viable in the seed bank up to 20 years until they find an opportunity to germinate [22].
• Density-dependent growth dynamics: Where density of a population is defined as the number of organisms living in a given area, density-dependent growth dynamics refers to the limitation of a population size, where the effects on the size or growth of a population vary with the density of the population itself. For example, when the density of a population is low, more seeds will germinate, and fewer individuals will die. On the other hand, when the density of a population is high, more individuals will die out, and fewer seeds will germinate [23].
• Carrying capacity: The maximum population size of the species that can live in an environment, given the necessary resources available.
• Logistic growth: Represents systems that grow exponentially until an upper limit or carrying capacity of the system is approached, at which point the growth rate slows and eventually saturates, forming the characteristic S-shape curve.
• Dispersal direction: Direction towards which seeds spread.
• Intraspecific competition: Refers to competition among members of the same species for limited resources in a given area.
• Steady-state condition of a population: Refers to a stable population level where the plant population does not have a net change over time.
Third, in addition to the specific biological traits of the species, landscape structure and characteristics of the potential habitat also impact the invasiveness of a species. In order to represent the landscape structure and heterogeneity of invasion, spatially explicit models are widely used in ISM studies (e.g., [24, 25]). Spatially explicit dynamics takes into account the spatial dependence among independent systems and enables the integration of these systems into the same model [26]. For example, Figure 1.1 demonstrates a spatially explicit landscape, which is divided into nine equal-size grid sites with a heterogeneous invader population. While each grid has its own carrying capacity, seed dispersal to the neighboring sites leads to spatial dependence among them. In addition, tracking seed production or population at each cell across time forms the problem’s temporal dimension. Therefore, considering the temporal dimension, spatially explicit models represent a system where data are collected across time and space.

![Figure 1.1. Representation of spatially explicit landscape](image)

While spatially explicit models are central to invasive species control model, integrating biological, spatial, and temporal dimensions of the problem into a mathematical model aggravates the computational complexity and solution time [21, 27]. In particular,
mathematical programming techniques, such as MIP, dynamic programming (DP), and non-linear optimization techniques [28-31], could provide a more efficient modeling alternative to simple ecological models or simulations. Considering spatial, temporal, and biological constraints complicates the models because they generate nonlinearities in the associated mathematical programs. Therefore, researchers develop heuristics such as the rolling horizon method [14, 19], recursive modelling [20], and genetic algorithm [15, 18] to overcome the adverse effects arising from the problem’s nonlinearity. Furthermore, in order to decrease the complexity, some researchers simplify the invasive species problem by relaxing spatial dispersal constraints [32-34]. Therefore, problem complexity has been tackled by using heuristic solution methods or reducing the size of the problem by ignoring some characteristics of invasive species.

In addition to spatio-temporal and biological aspects of ISM, it is critical to determine timely, cost-efficient, and dynamic detection strategies for the rapid control of invaders because surveillance of invasive species is crucial to provide timely treatments before the species invasion establishes and spreads to other landscapes. Early detection followed by rapid response, if not eradication, can keep invasive species under control, thus eliminating expensive control programs [35].

The aim of this dissertation is to develop spatio-temporal optimization models for the complex ISM problem to control the adverse impacts of the invasive species over a landscape under scarce economic resources by taking into account significant biological aspects such as age classes, spatial growth, and dispersal. Difficulties with the proposed models include the formulation of nonlinear biological intricacies, economic restrictions, and computational
burden caused by the complexity of modeling spatial heterogeneity and temporal dimensions in such a biological system. Therefore, the second aim of this dissertation is to develop solution methods such as linearization approaches and cutting planes as well as linearization techniques to facilitate optimal solutions of the complex mathematical models.

While spatially explicit models are widely used in the ISM literature, there are also other application areas where spatially explicit models incorporating temporal dimension are utilized. For example, spatio-temporal models have been used for cancer treatment, in particular radiation therapy \[36, 37\]. Here, a spatio-temporal model is utilized in external beam radiation therapy to shape the radiation beam as it exits the linear accelerator (LINAC) machine. Specifically, the beam sent from the LINAC is restricted by multileaf collimators, which are small rectangular leaves located inside the LINAC. Depending on the tumor size and structure, the leaves are opened and closed in order to shape the radiation beam. The beams pass through the patient and kill the cancer cells. After multiple beams, the impact of the radiation on different parts of tumor varies, ranging from high to low radiation delivery from the center to the margins of the tumor. It should be noted that while radiation kills cancer cells, it also passes through healthy tissues and organs, thus causing serious damage to surrounding healthy tissues. Therefore, radiation fractionations should be adjusted considering the vulnerability of healthy tissues and organs to radiation doses.

Chemotherapy is another method for treating cancer. In this systemic therapy, the chemo agent impacts the whole body by diffusing into the entire bloodstream. The chemo drug weakens and eradicates cancer cells at the original tumor site, as well as metastatic cells that spread to other parts of the body. Chemotherapy is generally used to clear the surgery site of
leftover cancer cells after surgical removal of the tumor, with the goal of reducing the risk of their recurrence. It can also be used prior to surgery to shrink the size of the tumor. Finally, it can be used as a palliative treatment method to destroy cancer cells as much as possible and to provide comfort to the patient [38]. Similar to the adverse impacts of radiation therapy, during chemotherapy treatment, the anti-cancer agent annihilates cancer cells but at the same time kills normal cells [39]. Therefore, chemotherapy should be limited and administered cautiously during the treatment period.

Surgery is a widely used method for breast cancer treatment. In order to ensure that the tumor is entirely resected from the cancer area, the tumor is removed along with the tumor margin, which is a layer of healthy tissue around the tumor [38]. Although the purpose of the surgery is to remove the entire tumor, it is possible that a portion of the cancer cells may remain in the breast area after surgery [40] because they are not visually detected.

Among the many forms of cancer, breast cancer is the second leading cause of death after lung cancer in women, with an estimation of 246,600 new cases and 40,450 deaths in 2016 [41]. Breast cancer usually begins in the cells of the lobules, which are the milk-producing glands, or the ducts, which are the passages that drain milk from the lobules to the nipple [38]. While non-invasive cancers stay within the milk ducts or lobules in the breast, invasive cancers grow into the normal, healthy tissue. Although early breast cancer can be cured and leads to a 100% five-year survival rate [42], postponing treatment could lead to an advanced stage cancer, metastasis, and death.

In this dissertation, in addition to studying ISM, we also develop spatio-temporal models for cancer treatment to determine an optimal treatment sequence considering surgery (S),
radiation therapy (R), and chemotherapy (C), in combination with the optimal dose schedules for chemo- and radiotherapy treatments, and we apply it to the case of breast cancer treatment.

1.1. Motivation and Contributions

Controlling biological invasions and diseases is a highly complex problem defined by the biological characteristics of the organisms, spatial context, and management objective of minimizing invasion damages given limited financial resources. While bio-economic optimization models provide a promising approach for invasive species control, current spatio-temporal optimization models ignore key ecological details such as seed dispersal and age structures, which could be essential to predict how populations grow and spread spatially over time and determine the most effective control strategies. Furthermore, the complexity of the problem is enhanced due to the exponential increase in number of decision variables while making decisions, such as when and where to apply surveillance and treatment over a spatial landscape. In addition, the ISM literature lacks studies that combine surveillance and control actions with biological characteristics of the species. On the other hand, taking into account surveillance, control, and biological characteristics of species in the same model intensifies the problem’s intricacy. Therefore, there is a research need to develop new models, specialized algorithms, or heuristics in order to solve the problem for practical size instances.

This dissertation fills the research gap in the ISM literature by developing optimization models that take into account the detailed biological characteristic of the species as well as economic aspects of invasive species management. Furthermore, a spatial problem structure is used to reflect the heterogeneity of the ecological system in consideration. Moreover,
proposed models solve the full dynamic problem and thus handle the damage of both current and forecasted invasions over the entire planning horizon at the same time. This feature enables policymakers to dynamically and optimally allocate the total budget over space and time. In addition, we contribute to the literature by proving general ISM models that can be specified for various invasive species by taking into account the invasion surveillance followed by control strategies. Finally, we develop unique cutting planes, and linearization and approximation solution methods to facilitate the solvability of the proposed models.

Although cancer demonstrates different and more complex biological characteristics, the growth of cancer cells shows similar traits to the invasive species in terms of demonstrating a spatio-temporal spread dynamic. Treatment options are determined by the stage and size of the cancer, and usually include surgery, radiation therapy, chemotherapy, or a combination of these treatment methods. Although both chemotherapy and radiotherapy are means of controlling the cancer population, they also have adverse effects on the human body. Therefore, the chemotherapy treatment dosage should be adjusted, based on the body’s drug tolerance. Furthermore, radiation damages healthy cells in the path of the radiation beam and should be adjusted based on the biological effective dose (BED) for both healthy and cancerous tissues. Moreover, while many studies concentrate on cancer treatment with surgery, radiation therapy, and chemotherapy, no MIP model combines all treatment strategies to analyze the impact of different treatment methods on tumor size, growth, and spread. In addition, treatment models include high nonlinearities, which could be handled by using proper linearization techniques. Furthermore, a three-dimensional spatial representation divided into
and growth combined with three main treatment options.

The proposed study fills the research gap in the cancer treatment literature by first providing a spatio-temporal MIP model that provides the optimal sequence of treatment modalities over the cancer-treatment period. Second, the spatial dimension assists in predicting the spread and regrowth of cancer cells during and after the treatment period, which would enable tracking of the size and extent of the growth and the spatial spread of cancer cells. New three-dimensional growth and spread models are developed for studying optimal cancer treatment. Finally, the proposed methodology differentiates the efficiency of surgery on the margin of the tumor compared to inner tumor structure.

1.2. Summary of Chapters

The key theme of this dissertation is the development of optimal resource allocation and treatment scheduling for controlling the rapid growth and spread of bio-invasions, including invasive species and cancer. This section provides an overview of the chapters presented in this dissertation.

Chapter 2 presents a literature review of the models developed for ISM from the perspective of surveillance and control actions. In order to cover the literature in a systematic way, the review on invasive pests and plants is broadened, and the literature is classified according to studies that include both “control” and “surveillance” efforts. Studies in these two research areas are further classified into two model groups: deterministic and stochastic. Deterministic models are clustered as MIP, integer programming (IP), and NLP models, while
stochastic models are classified as SP and stochastic dynamic programming (SDP) models. A total of 24 studies have been reviewed in this chapter.

Chapter 3 presents a novel age-structured optimization model as a spatial-dynamic decision framework for controlling invasive species. In particular, the proposed model allows for taking into account the biological competition among different age classes within the population. The potential use of the model is demonstrated on controlling the invasion of sericea (*Lespedeza cuneata*), a perennial legume threatening native grasslands in the Great Plains. Results show that incorporating an age structure into the model captures important biological characteristics of the species and leads to unexpected results, such as multi-logistic population growth with multiple, sequential, and overlapping phases of logistic form. These new findings can contribute to understanding time-lags and invasion growth dynamics. Additionally, given budget constraints, utilizing control measures every two to three years is found to be more effective than yearly control because of the time it takes for reproductive maturity. Results of the bio-economic optimization approach provide both ecological and economic insight into the control of invasive species. Furthermore, while the proposed model is specific enough to capture biological realism, it also has the potential to be generalized to a wide range of invasive plant and animal species under various management scenarios in order to identify the most-efficient control strategies for managing invasive species.

The complexity of the invasive species problem stems from the nonlinearity that is inherent in biological systems, consequently impeding researchers to obtain timely and cost-efficient treatment strategies over a planning horizon. Therefore, in Chapter 4, to cope with the complexity of the invasive species problem, an MIP model that handles the problem as a full
optimization model and solves it to optimality for the first time is developed. The applicability of the model is demonstrated on a case study of sericea (*Lespedeza cuneata*) infestation by optimizing a spatially explicit model on a heterogeneous 10-by-10 grid landscape for a seven-year time period. The solution quality of five different linearization methods used to obtain the MIP model are evaluated. The model is also compared with its mixed-integer nonlinear programming (MINLP) equivalent and NLP relaxation in terms of solution quality. The computational superiority and realism of the proposed MIP model demonstrate that this model constitutes the basis for future decision-support tools in invasive species management.

It is critical to determine timely, cost-efficient, and dynamic detection and treatment strategies for the rapid control of invaders. One of the most economically and environmentally damaging invasive species is the EAB, a pest of ash trees. While models and information are available to evaluate the economic impacts of an EAB invasion, there are no models to help evaluate the costs and benefits of surveillance of infested ash trees in combination with treatment and removal. In Chapter 5, a unique multistage stochastic mixed-integer programming model (MSS-MIP) is developed. The model takes into account the spatio-temporal stochastic growth of EAB infestation and allows for considering all possible scenarios for surveillance, treatment, and removal of ash trees over a planning horizon. The objective here is to maximize the net benefits of the ash trees on a given landscape by applying surveillance to the ash population, followed by treatment or removal of trees based on their infestation level. Due to the model’s complexity and state-of-the-art nature, novel cutting planes and a preprocessing algorithm to strengthen the MSS-MIP formulation and facilitate the corresponding optimal solution are developed. Results provide insights into surveillance and
control policies, and provide an optimal strategy to eradicate EAB infestation with scarce resources.

In Chapter 6, a new MIP model that explicitly formulates cancer growth and spread is developed, in order to determine the optimal sequence of the typically prescribed cancer treatment methods—surgery, chemotherapy, and radiotherapy—while minimizing the newly generated tumor cells in a three-dimensional spatio-temporal system. The quadratically constrained cancer growth dynamics and treatment impact formulations are linearized by using linearization as well as approximation techniques. Under supervision of medical doctors and utilizing several resources for the parameter values, the effectiveness of treatment combinations for breast cancer specified with different sequences (i.e., SRC, SCR, CR, RC) are compared by tracking the number of cancer cells at the end of each treatment method. Furthermore, the results provide the optimal chemotherapy and radiation dosage for each administration and fractionation to minimize the new growth of cancer cells.

1.3. Main Organization of Dissertation

This PhD dissertation is organized by chapters that correspond to one conference and four journal papers. Chapter 2 provides a literature review of the models developed for ISM from the perspective of surveillance and control actions. Chapter 3 provides a new age-structured and cost-effective nonlinear resource allocation model for invasive species management. Chapter 4 presents an MILP model for an problem to analyze the economic damages associated with invasive species under a limited treatment budget while taking into account its spatial dynamics. Chapter 5 proposes a MSSP model for ISM problem by integrating surveillance and control strategies and developing new cutting planes to substantially improve
the solution time of the proposed model. Chapter 6 proposes a spatio-temporal mathematical formulation that combines different treatment strategies for breast cancer treatment. Finally, Chapter 7 summarizes the contributions and future directions of the dissertation.

1.4. References


CHAPTER 2

REVIEW OF OPTIMIZATION MODELS ON INVASIVE SPECIES MANAGEMENT

2.1. Introduction

ISM has gained importance over the last few decades due to its adverse environmental and economic impacts on biodiversity, ecosystems, and economic enterprises such as agriculture, forestry, fisheries, power production, and international trade [1]. The economic costs of invasive species are similarly immense. Westphal et al. [2] link the rising number of biotic invasions all over the world to increased global trade. Furthermore, worldwide costs of invasive species such as plants, animals, and microbes have been estimated at US$1.4 trillion, or 5% of the global gross domestic product [3]. In addition, in a more recent study, the economic cost of invasive species on natural and agricultural systems within the U.S. has been estimated at $120 billion annually [4].

Conservation of ecology and the ecosystem can highly benefit from the application of optimization studies, which integrate invasive species dynamics, landscape characteristics, costs, and management options in a mathematical model. For example, each invasive species demonstrates distinct population and growth dynamics, and exhibits uncertainty in their growth rates as well as impacts on native vegetation and its dispersal. In addition to the specific biological traits of the species, landscape structure and characteristics of the potential habitat also impact their invasiveness. In addition, limited economic resources restrict management efforts. Therefore, explicit consideration of these components within a decision framework can have a dramatic effect on management strategies.
Operations research has been widely applied to study biological systems [5-7]. In this chapter, we review studies about ISM from the perspective of operations research applications. In section 2.2, we explain our approach and categorization of the reviewed papers. Here we provide a list of studies along with the models used and related areas. In section 2.3, we present papers on various optimization models that concentrate on control and surveillance efforts. Finally, section 2.4 provides a brief discussion on the main findings from the literature review and concluding remarks.

2.2. **Overview of the Literature**

Research conducted in ISM changes from field analyses to strategic decisions at the manager level and ranges from aquatic invasive species to pests and invasive plants. Furthermore, studies vary in terms of their concentration on prevention, surveillance, or control of invasive species. In addition, solution methodologies demonstrate differences depending on the invasive species studied and extent of the study. Therefore, in order to cover the literature in a systematic way, we broaden this review on invasive pests and plants, and classify the literature according to studies that include (1) “control”, and (2) “surveillance” efforts. Studies in these two research areas are further classified into two model groups as deterministic and stochastic. Deterministic models are clustered as MIP, IP, and NLP models, while stochastic models are classified as SP and SDP models. In this paper, we have reviewed a total of 27 studies. The reviewed articles are listed in Table 2.1 below.
### TABLE 2.1. CLASSIFICATION OF STUDIES

<table>
<thead>
<tr>
<th>Objective</th>
<th>Method</th>
<th>Application</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize economic costs</td>
<td>MINLP</td>
<td>Buffelgrass</td>
<td>[8]</td>
</tr>
<tr>
<td>Analysis of management implications</td>
<td>MINLP</td>
<td>Buffelgrass</td>
<td>[9]</td>
</tr>
<tr>
<td>Minimize damage to buildings, saguaros, and vegetation</td>
<td>Multi-objective MINLP</td>
<td>Buffelgrass</td>
<td>[10]</td>
</tr>
<tr>
<td>Determine stakeholder compensation</td>
<td>Game theoretical</td>
<td>Buffelgrass</td>
<td>[11]</td>
</tr>
<tr>
<td>Maximize healthy timber volume</td>
<td>MINLP</td>
<td>Pine beetle</td>
<td>[12]</td>
</tr>
<tr>
<td>Minimize economic costs</td>
<td>MINLP</td>
<td>Brown tree snake</td>
<td>[13]</td>
</tr>
<tr>
<td>Minimize invasion control efforts</td>
<td>NLP</td>
<td>Spartina alterniflora</td>
<td>[14]</td>
</tr>
<tr>
<td>Minimize economic and ecological loss</td>
<td>Diff**</td>
<td>Beaver</td>
<td>[15]</td>
</tr>
<tr>
<td>Maximize economic welfare</td>
<td>Diff</td>
<td>Hypothetical</td>
<td>[16]</td>
</tr>
<tr>
<td>Minimize control and invasion costs</td>
<td>IP</td>
<td>Hypothetical</td>
<td>[17]</td>
</tr>
<tr>
<td>Minimize invasion size</td>
<td>MIP</td>
<td>Pine beetle</td>
<td>[18]</td>
</tr>
<tr>
<td>Minimize economic costs</td>
<td>MIP</td>
<td>Lespedeza</td>
<td>[19]</td>
</tr>
<tr>
<td>Maximize economic welfare</td>
<td>SDP</td>
<td>Leafy spurge</td>
<td>[20]</td>
</tr>
<tr>
<td>Minimize control and invasion costs</td>
<td>SDP</td>
<td>Hypothetical</td>
<td>[21]</td>
</tr>
<tr>
<td>Minimize treatment cost</td>
<td>SDP</td>
<td>Hypothetical</td>
<td>[22]</td>
</tr>
<tr>
<td>Minimize economic costs</td>
<td>SDP</td>
<td>Hypothetical</td>
<td>[23]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective</th>
<th>Method</th>
<th>Application</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring and detection of invasion</td>
<td>Sim***</td>
<td>Hypothetical</td>
<td>[24]</td>
</tr>
<tr>
<td>Analysis of active and passive surveillance</td>
<td>IP</td>
<td>Hypothetical</td>
<td>[25]</td>
</tr>
<tr>
<td>Optimize budget among surveillance and treatment</td>
<td>SDP</td>
<td>Fire ant</td>
<td>[26]</td>
</tr>
<tr>
<td>Minimize detection and control costs</td>
<td>SDP</td>
<td>Gypsy moth</td>
<td>[27]</td>
</tr>
<tr>
<td>Minimize detection and control costs</td>
<td>Diff</td>
<td>Hypothetical</td>
<td>[28]</td>
</tr>
<tr>
<td>Minimize detection and control costs</td>
<td>LP</td>
<td>Gypsy moth</td>
<td>[29]</td>
</tr>
<tr>
<td>Optimize one-time surveillance effort</td>
<td>NLP</td>
<td>Orange hawkweed</td>
<td>[30]</td>
</tr>
<tr>
<td>Minimize monitoring and invasion costs</td>
<td>LP</td>
<td>Gypsy moth</td>
<td>[31]</td>
</tr>
</tbody>
</table>

*Left and right panels represent studies that involve control and surveillance efforts, respectively.  
**Differential equations model.  ***Simulation

#### 2.3. Literature Review

Due to the inherent complexity of biological dynamics of invasive species, researchers have been provided several methods to handle the intricacy of this problem. Studies provided in the following section have two common properties: (1) those based on spatially explicit models where the considered landscapes are divided into patches in order to take the heterogeneity of the landscape into account, and (2) those where invasive species spread to the entire landscapes through dispersal among the patches.
2.3.1. Studies on Invasive Species Control

Invasive species control has been a primary concern in many ISM studies. However, taking into account spatial constraints often complicates the models because they generate nonlinearities in the associated mathematical programs. Therefore, researchers develop heuristics to overcome the adverse effects arising from the problem’s nonlinearity. For example, Büyüktahtakın et al. [8] introduce a large-scale, nonlinear, zero-one IP model for the dynamic control of invasive weeds. They minimize economic damage caused by the buffelgrass weed to the environment and infrastructures. While the authors concentrate on model construction and solution algorithm concepts in [8], they focus heavily on the management implications of model results in [9]. Their results demonstrate the significance of the timing of the invasion and the frequency of follow-up treatments. In another study, Büyüktahtakın et al. [10] develop a multi-objective nonlinear MIP model that provides treatment strategies to buffelgrass invasion to minimize the economic damages on saguaros, buildings and vegetation. The consequences of various treatment strategies are evaluated based on the three resources. In [8-10], the complexity due to the nonlinearity is overcome by using a rolling-horizon solution in which they solve single-period problems one year at a time and feed the result into the next period’s problem. In a later study, Büyüktahtakın et al. [11] propose a game theoretical model to minimize total economic damages of invasion. The cooperation of decision makers is modeled by using solution concepts based on characteristic functions, where Shapley values of the cooperative game is determined. The goal is to find out the payoffs that will be awarded to each decision maker.
Aadland et al. [12] demonstrate that the full nonlinear spatial model even on a 2-by-2 grid cannot be solved, but instead they develop an approximately optimal recursive nonlinear model, where treatment strategies depend only on the current year’s host and insect stocks. Similar to a rolling horizon approach in [8-10], the recursive model enables authors to determine treatment rules for each year sequentially rather than jointly. Kaiser and Burnett [13] develop an MINLP model to address the effect of spatial variation in damage, cost, and biological growth on optimal treatment strategies over time using the case study of the Brown treesnake on the island of Oahu, Hawaii. The objective of their study is to minimize the present value of costs and damages for optimal removal of the snake invasion. They use a special genetic algorithm where the decision variable represents the number of spatial cells treated rather than the snake population in a given cell, which significantly reduces the problem’s complexity. Taylor and Hastings [14] develop a multi-objective nonlinear model to minimize the total effort to control the invasive grass spartina alterniflora and the risk of species colonization. Similar to [13], they use a genetic algorithm to address the question of whether it is more efficient to focus on the removal of young, low-density areas at the edge of an invasion or older, core population areas when using mechanical or chemical control methods. In another study, Bhat et al. [15] develop a nonlinear differential equations model to minimize the sum of economic and ecological loss due to beaver-inflicted timber damage and beaver-trapping cost. They develop a solution algorithm where the output in a preceding iteration is used as input in the following iteration until the solution converges based on a prespecified error criterion. Furthermore, Albers et al. [16] develop a nonlinear partial differential equation model to explore the costs and benefits of centralized and decentralized policies for invasive species.
control when two heterogeneous regions are connected through ecological and economic factors.

In order to decrease the complexity, some researchers simplify the invasive species problem by relaxing spatial dispersal constraints and modeling it as 0–1 linear programs, which can be solved efficiently by commercial solvers. For example, Epanchin-Niell and Wilen [17] consider the presence or absence of propagules on a cell rather than the abundance of propagules per cell in their model and minimize the control costs and invasion damages across space and time. They show that different shape and starting location of the invasion can result in dramatically different optimal control policies. There are also MIP problems that accounts for both biological dynamics of invasive species and spatial characteristics of the environment. For example, Hof et al. [18] minimize the total pest population over a planning horizon. Once infestation is removed, infestation will not be present again on a given site. In another study, Kibiş and Büyüktahtakın [19] minimize total economic costs due to the presence of invasive weeds by considering spatial heterogeneity, population dynamics and age-structures, stochastic dispersal dynamics, and a unique carrying-capacity constraint. Results demonstrate that treatment strategies vary depending on the abundance and frequency of the invasion.

Furthermore, some studies also consider stochasticity in some parameters such as growth or dispersal of invasive species, or both. For example, Hyder et al. [20] take into account stochastic population growth and spatial spread and incorporate it into a decision framework using SDP. Solutions from their SDP model determine the optimal strategy of whether to apply biological or herbicide control to leafy spurge to maximize the welfare due to the reduced amount of infestation. Furthermore, Haight and Polasky [21] use the Markov decision process
to minimize the sum of discounted costs associated with treatment and damage caused by invasive species. Based on the probability distribution, the decision maker chooses whether to treat the invasion or not. In a later study, Baxter et al. [22] use SDP to minimize the total expected cost of weed removal by taking into account the relation between invasion abundance and variation in the spread probabilities, which are affected by differences in the cost of removal, management time horizon, and population dynamics of the invasion. In another study, Olson and Roy [23] minimize the total cost and damage associated with invasive species in an SDP model to examine the economic conditions under which it is optimal to eradicate invasive species. Similar to [22], they demonstrate that the optimal time of the treatment is the time when the marginal cost of damage is higher than the marginal cost of treatment.

2.3.2. Studies Including Invasive Species Surveillance

In this section, we focus on studies that incorporate both surveillance and control actions into a decision framework. For example, Cacho et al. [24] develop a simulation model to analyze the significance of monitoring and detectability of invasion on the successful eradication of invasive species. They show that the treatment program significantly depends on the detection probability of the target plant, economic resources, and growth and spread rate of the invasive species. In a latter study, Cacho et al. [25] simulate the dispersal of invasive species, and incorporate both active and passive surveillance efforts (surveillance conducted by the professionals in charge of controlling the invasion, and surveillance by members of the public, respectively) into a decision framework. They show that increased passive surveillance reduces the control costs significantly and increases the probability of eradication as active
surveillance is enhanced. Baxter and Possingham [26] use an SDP model to optimize search-and-destroy strategies for the fire ant, Solenopsis invicta. They disregard the spatial spread since ants do not spread and colonize on a different nest. They demonstrate that search efforts are optimal if the pest is already widespread or if knowledge about the pest invasion is poor. Otherwise, complete eradication without surveillance is preferred. Furthermore, Mehta et al. [27] develop an SDP model to minimize the discounted expected total costs of detection and control. They analyze the relation between detection and subsequent control costs, and demonstrate that the cost of invasion significantly increases with delayed detection. Homans and Horie [28] use a stochastic differential equations model to minimize the total cost associated with detection, treatment, and economic damage caused by the invasive pest while choosing the optimal detection strategy. They demonstrate that a high dispersal range adversely impacts the detection of the dispersed population and results in a more aggressive approach to finding and suppressing the pest. Bogich et al. [29] develop a linear model to minimize the search and control costs of a gypsy moth infestation by optimizing detection and eradication efforts based on various growth rates. They demonstrate that optimal surveillance and eradication of the invasion depends significantly on biological characteristics such as growth rate, abundance and frequency of invasion, and economic parameters such as planning horizon length, eradication, and surveillance costs. While previous approaches have employed dynamic optimization methods, Hauser and McCarthy [30] develop a nonlinear model to minimize surveillance and treatment costs, and optimize a one-time surveillance effort when the presence of species is uncertain prior to detection. Results demonstrate that surveillance should be prioritized at sites where the invasive species will be relatively easy to detect.
Epanchin-Niell et al. [31] develop an LP model to minimize the total cost of surveillance and control. Different than previous models, their model enables multiple surveillance efforts over time. They show that a higher surveillance effort is required for invasive species that have higher growth and dispersal rates, and are costlier to eradicate.

2.4. Discussion and Conclusion

In this paper, we review articles in the literature regarding invasive pest and weed management from control and surveillance perspectives. We also cluster the articles in terms of their objectives, application areas, and solution methodologies. Although several studies concentrate on invasive species surveillance and control, these models ignore the biological dynamics in order to reduce problem complexity. Researchers should concentrate on optimization methods to deal with the complexity of solving these problems while taking into account biological characteristics and uncertainty [32, 33]. There is also a research need to study game theoretical models to analyze interactions among stakeholders [34]. Furthermore, the cost and benefits of surveillance on ISM in combination with treatment has not been well studied. Surveillance implies the eradication of invasion upon detection in previous studies, even though eradication may not be economically optimal at the point of surveillance. Therefore, there is a need for future studies that particularly optimize surveillance and subsequent control actions.

2.5. References


CHAPTER 3

AGE-STRUCTURED BIO-ECONOMIC MODEL OF INVASIVE SPECIES MANAGEMENT: INSIGHTS AND STRATEGIES FOR OPTIMAL CONTROL

3.1. Introduction

Controlling species invasions at the landscape scale is a highly complex problem. First, the rates of spread and impact on native communities are critically dependent on life span, growth rates, dormant stages, and dispersal, which may not be captured by simple population growth functions [1]. Second, landscapes are heterogeneous, and invasions often do not follow simple patterns of spread from a given introduction point [2]. Third, available resources (time and money) are almost always limited. When these factors are combined, intuitively determining the most efficacious control strategy quickly becomes intractable. For this reason, optimization models of invasion control that explicitly incorporate limited budgets can be useful decision tools to analyze the potential consequences of different control strategies (for a detailed review of these studies, see, e.g., [3-5]).

Although bio-economic optimization models for invasion control are not new [6], advances in optimization and computational power offer new opportunities to incorporate much greater ecological realism than previously possible. Several optimization models demonstrate the importance of spatio-temporal processes when controlling invaders [7-15]. While the progress in spatial-temporal modeling is encouraging, the real potential for such models remains under-utilized, because those models omit key ecological details such as age structures—which could be essential to forecast how populations grow spatially over time and determine the most effective control strategies. In this paper, we present a novel age-
structured optimization model as a spatial-dynamic decision framework for controlling invasive species, and demonstrate the potential use of the model for controlling the invasion of sericea (Lespedeza cuneata), a perennial legume threatening native grasslands in the Great Plains. In particular, we develop a new carrying capacity sub-model, which allows us to take into account the biological competition among different age classes within the population. The results demonstrate that incorporating age-structure into the model captures important biological characteristics of the species and lead to unexpected results such as multi-logistic population growth. These new findings can contribute to understanding time-lags and invasion growth dynamics thus provide new insight into controlling invaders.

We include the age structure of invasive species in the model because reproduction and survival vary with plant age. The simplest age-structured model is the Leslie Model [16], where population is divided into discrete age classes. Structured models include age-, weight-, stage- and size-structured models (see, e.g., [17-19]). Among all these possible structuring alternatives, we consider an age-structured model of the invasive species control because, for many species, reproduction and survival rate vary with age [20]. The seed stage is particularly important because seeds can either germinate quickly or form a long-term seed bank, which builds a reservoir of potential propagules that can increase future weed infestations [21]. Our model is unusual in that it accounts for density, frequency, age, dispersal, and seed bank dynamics of the invaders simultaneously in a spatio-temporal landscape to determine the optimal placement and timing of invasion control.

Here, population growth is derived from demographic information such as seed bank and dispersal, as opposed to the use of a logistic growth function, which is a central assumption
in previous invasion control models. The seedbank-based linear growth model contributes to the optimization of spatio-temporal population dynamic models by significantly improving its solvability compared to non-linear logistic growth counterparts while maintaining much greater biological complexity than other logistic or constant growth models. Furthermore, the model incorporates different seed production and loss rates by dividing the population into different age classes, and it tracks the growth of each age class over a multi-period time horizon. Incorporating seed bank growth and age structure into the model provides insight into population growth patterns, which is found to be more complex than the simple logistic growth [22] and has important implications for control strategies.

We also consider uncertainty in our deterministic model by performing sensitivity analysis of different uncertain parameters such as budget, eradication rate, and dispersal rate. In addition, we examine model solutions in order to provide the minimum necessary level of resources (labor and budget) that could efficiently control the invader under different initial population distribution scenarios. Finally, we evaluate the effectiveness of different treatment frequency strategies for controlling invasion damages.

3.2. Mathematical Model

The bio-economic model is formulated as follows: Let $T$ denote the time horizon, and let $t \in [0,T]$ be any year of the planning horizon. The area consists of rectangular cells with $I$ rows and $J$ columns. Any cell can be characterized by its coordinates $(i,j)$, where $i \in \{1,2,...,I\}$ and $j \in \{1,2,...,J\}$. The decision variable $x_{i,j}(t)$ is defined as the percent of area treated in cell $(i,j)$ in year $t$. 
In order to incorporate different seed production rates for different age groups, we define age groups (classes) \( k = 1, 2, 3, \ldots, n^* \), where each age group \( k \) defines a class of \( k \) year(s)-old species population, except that age group \( n^* \) includes the \( n \)-years-old and older population. Therefore, for individuals that reach maturity at the age of \( n \), where \( n \) can be any number depending on the species, transition population densities in cell \((i, j)\) are formulated as

\[
NP_{i,j}^k(t+1) = \delta \times \rho \times SB_{i,j}(t) \quad k = 1 \text{ and } \forall i, j, t \quad (3.1)
\]

\[
NP_{i,j}^k(t+1) = NA_{i,j}^{k-1}(t) \times (1 - \varphi_{k-1}) \quad k = 2, \ldots, n-1 \text{ and } \forall i, j, t \quad (3.2)
\]

\[
NP_{i,j}^k(t+1) = NA_{i,j}^{k-1}(t) \times (1 - \varphi_{k-1}) + NA_{i,j}^k(t) \times (1 - \varphi_k) \quad k = n \text{ and } \forall i, j, t \quad (3.3)
\]

where \( \delta \) is the seed germination rate, \( \rho \) is the survival rate of plants after becoming a seedling, \( SB_{i,j}(t) \) is a function representing seed bank population at time \( t \), \( \varphi_k \) is the loss rate of individuals when age class \( k \) grows into age class \( k+1 \), \( NA_{i,j}^k(t) \) is the population after treatment for the age class \( k \) at the beginning of time period \( t \), and \( NP_{i,j}^k(t+1) \) represents the potential population for age class \( k \) in cell \((i,j)\) at the beginning of time period \( t + 1 \) before carrying capacity is considered.

Equation (3.1) gives the number of one-year-old individuals at the beginning of period \( t + 1 \) that have germinated from the seed bank in time period \( t \) and become seedlings. Equation (3.2) denotes the transition population levels of individuals that are \( k = 2 \ldots n - 1 \) years old at the beginning of period \( t + 1 \) and were subject to individual losses at rate \( \varphi_{k-1} \) due to seasonal changes and ecological factors in period \( t \). Equation (3.3) provides the number of \( n^* \)-year-old individuals at the beginning of period \( t + 1 \), which are \( n - 1 \) and \( n^* \) years old and exposed to individual losses at rate \( \varphi_{k-1} \) and \( \varphi_k \), respectively, in period \( t \).
Here, we consider an invasive plant that disperses only through seeds. It is assumed that some of the seeds will be dispersing to eight adjacent cells, and some of them will remain within the cell. Define $M^{ij}$ as the set of eight adjacent cells of a cell $(i,j)$, where $M^{ij} = \{(i+1,j+1),(i+1,j),(i,j+1),(i-1,j-1),(i-1,j),(i-1,j+1),(i+1,j-1)\}$. We then formulate the seed dispersal to cell $(i,j)$ from its surrounding eight neighbors $(h,w) \in M^{ij}$ as

$$SD_{i,j}(t) = \lambda \sum_{k=1}^{n} \sum_{(h,w) \in M^{ij}} S(k) \ast NA_{h,w}^{k}(t) \quad \forall i, j, t$$

(3.4)

where $\lambda$ is the percentage of seeds dispersing from one of the eight neighboring cells to cell $(i,j)$ in period $t$, $S(k)$ is the number of seeds produced by one individual of age class $k$, and $NA_{h,w}^{k}(t)$ is the number of the individuals of age class $k$ in the surrounding cell $(h,w) \in M^{ij}$ following treatment. Equation (3.4) gives the total number of seeds dispersed from eight surrounding cells to cell $(i,j)$.

The number of seeds remaining in cell $(i,j)$ after dispersal is then given as

$$Seed_{i,j}(t) = \theta \sum_{k=1}^{n} NA_{i,j}^{k}(t) \ast S(k), \quad \forall i, j, t$$

(3.5)

where $\theta = 1 - 8 \ast \lambda$ is the percentage of remaining seeds after seeds are dispersed to the adjacent eight cells $(h,w) \in M^{ij}$. Equation (3.5) gives the total number of seeds produced in cell $(i,j)$, after dispersal.

After seed production by each individual, seeds can germinate, experience mortality from pathogens or seed predators, or become dormant, thus forming a seed bank. Hence, the longevity (viability) rate of seeds, which is defined by the percentage of seeds in the seed bank that remain viable over time, and the germination rate are important factors that must be
considered when modeling the seed bank population. Therefore, the number of seeds in the seed bank in cell \((i,j)\) at time \(t\), \(SB_{i,j}(t)\) is formulated as

\[
SB_{i,j}(t) = \sum_{s=0}^{t} \left( (\gamma - \delta)^{t-s} \times \right. \left[ SD_{i,j}(s) + Seed_{i,j}(s) \right] + \left. SB_{i,j}(0) \times (\gamma - \delta) \right) \quad \forall i, j, t
\]  

(3.6)

where \(\gamma\) and \(\delta\) represent the longevity rate and germination rate of seeds in the seed bank, respectively. Equation (3.6) indicates that the seed bank population in cell \((i,j)\) at time \(t\) includes seeds dispersing from the surrounding cells \((SD_{ij})\), seeds that are produced but not germinated within the cell \((Seed_{ij})\), and the initial seed bank population \((SB_{i,j}(0))\). Note that \(SD_{ij}(s)\) and \(Seed_{ij}(s)\) are given in equations (3.4) and (3.5), respectively. Also note that in equation (3.6), the seed bank population has a compound increasing rate depending on the longevity and germination rate of the seeds, which decays as time passes.

Although invasive species commonly produce many offspring, natural boundaries, soil characteristics, and ecological factors constitute barriers for total population in a given cell so that the population cannot exceed the carrying capacity of cell \((i,j)\), \(K_{i,j}\), which is the maximum density (number of individuals) in cell \((i,j)\). Therefore, the actual individual population before treatment, \(NB_{i,j}^{k}(t)\), is formulated as

\[
NB_{i,j}^{k}(t) = \min \{ NP_{i,j}^{k}(t), K_{i,j} \}, \quad k = n^+ \quad \text{and} \quad \forall i, j, t
\]  

(3.7)

\[
NB_{i,j}^{k+1}(t) = \begin{cases} 
0 \text{ if } K_{i,j} - \sum_{a=k+1}^{n^+} NB_{i,j}^{a}(t) \leq 0 \\
\min \left\{ \left( K_{i,j} - \sum_{a=k+1}^{n^+} NB_{i,j}^{a}(t) \right), NP_{i,j}^{k}(t) \right\} \text{ o.w.} 
\end{cases} \
\]  

(3.8)

In the model, the carrying capacity allows previously established plants to occupy cell \((i,j)\) before younger individuals do. If the cell population does not reach carrying capacity by the
individuals of age class $n^+$, the second-oldest class adds to the population up to the carrying capacity. This cycle continues until the population reaches the maximum population level in each cell.

In the case of treatment, the before-treatment population is multiplied by the factor $\left(1 - \omega \times x_{i,j}(t)\right)$, where $\omega$ is the eradication rate of the treatment, and $x_{i,j}(t) \in [0,1]$ is a decision variable representing the percentage of area treated in cell $(i,j)$ in year $t$. Therefore, the population after treatment for age class $k$, $NA_{i,j}^k(t)$, is calculated by

$$NA_{i,j}^k(t) = NB_{i,j}^k(t) \times \left(1 - \omega \times x_{i,j}(t)\right) \quad \forall i, j, t \quad (3.9)$$

The treatment in each time period $t$ is limited by available budget for treatment and labor. Therefore, the budget constraint becomes

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (C_{i,j} + H_{i,j}) \times x_{i,j}(t) \leq B(t) \quad \forall t \quad (3.10)$$

where $C_{i,j}$ is the treatment labor cost per cell $(i,j)$, $H_{i,j}$ is the specific (e.g., herbicide) cost of treatment per cell $(i,j)$, and $B(t)$ is the available budget for treatment and labor at time period $t$. Equation (3.10) ensures that the total amount of a budget spent for treatments in a period $t$ cannot exceed the available budget in period $t$.

The objective of the model is to minimize the total economic damages caused by the invasive species population in all cells and all periods of the planning horizon. The objective function is then formulated as

$$\text{Minimize } z = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{t=1}^{T} D_{i,j}(t) \quad (3.11)$$
where $D_{i,j}(t)$ represents the damage as an economic loss due to invasion in cell $(i,j)$ at the beginning of time $t$. The term $D_{i,j}(t)$ is given as

$$D_{i,j}(t) = E_{i,j}(t) \times \sum_{k=1}^{n} \frac{N_{i,j}^k(t)}{K_{i,j}} \quad \forall i, j, t$$

(3.12)

where $E_{i,j}(t)$ represents the economic value of cell $(i,j)$. The total sum of individuals at different age classes represents the total population in cell $(i,j)$ at the beginning of time period $t$. In equation (3.12), the loss of economic value in cell $(i,j)$ is proportional to the ratio of the total population with respect to carrying capacity, $K_{i,j}$, in that cell.

3.3. Case Study: Control of Sericea Invasion in the Great Plains

Sericea is a drought tolerant legume that can grow in a range of soil types, produce copious amounts of seeds, and has a long-lived seed bank [23]. Sericea was declared a noxious weed by the Kansas Department of Agriculture in 2000, has spread over 2,226,337 ha of the mid- to southern Great Plains [24], and led to $29$ million average annual forage loss in the Flint Hills region of Kansas [22]. Furthermore, this legume replaces the native species in grasslands and threatens biodiversity in the Great Plains. Although herbicides can effectively eradicate established plants, populations can quickly recover from control strategies by germination from the seed bank. Given the biological characteristics of sericea and limited financial resources, determining the most effective long-term control strategy is difficult without the use of complex response models. Our bio-economic optimization model provides decision strategies regarding where and when limited funds can be best allocated for effective control of invasions by applying a restricted budget across a 45-year planning horizon. The objective of the model is to minimize economic loss from haying and grazing due to sericea invasion. A gridded landscape
is utilized to represent the spatially heterogeneous growth, spread, damage, and control costs. In this case study, we represent the initial invasion on a 10 x 10 landscape (40 ha), where each cell represents 0.4 ha of land.

We examine responses using population maps reflecting three different frequency levels, representing the percentage of invaded areas of the gridded landscape—at 2% (low), 40% (medium), and 80% (high)—and three different abundance rates defining species population in each cell—such as U[1–20] (low), U[21–200] (medium), and U[201–2000] (high), where U[a–b] denotes an integer number drawn uniformly from the interval [a, b] (Table 3.1).

### Table 3.1 Initial Population Structure and Parameters for Sericea Invasion

<table>
<thead>
<tr>
<th>Description</th>
<th>Frequency</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of cells invaded in 10 x 10 landscape</td>
<td>2%</td>
<td>Low frequency</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>Medium frequency</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>High frequency</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description</th>
<th>Abundance</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial population of sericea ramets in one cell</td>
<td>U[1–20]</td>
<td>Low abundance</td>
</tr>
<tr>
<td></td>
<td>U[21–200]</td>
<td>Medium abundance</td>
</tr>
<tr>
<td></td>
<td>U[201–2000]</td>
<td>High abundance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Symbol</th>
<th>Units</th>
<th>Case study values</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss rate from age cluster k to k+1</td>
<td>( \varphi(k) )</td>
<td>—</td>
<td>4%, 9%, 22%</td>
<td>[25]</td>
</tr>
<tr>
<td>Number of seeds produced per ramet per age cluster k</td>
<td>( S(k) )</td>
<td>—</td>
<td>0.45, 900</td>
<td>[20]</td>
</tr>
<tr>
<td>Percentage of seed dispersal</td>
<td>( \lambda )</td>
<td>—</td>
<td>0.01%, 0.1%** 1%</td>
<td>*</td>
</tr>
<tr>
<td>Percentage of remaining seeds</td>
<td>( \theta )</td>
<td>—</td>
<td>99.92%, 99.2%, 92%</td>
<td>*</td>
</tr>
<tr>
<td>Longevity rate</td>
<td>( \gamma )</td>
<td>—</td>
<td>95%</td>
<td>*</td>
</tr>
<tr>
<td>Germination rate</td>
<td>( \delta )</td>
<td>seedlings/seeds</td>
<td>6.8%</td>
<td>[26]</td>
</tr>
<tr>
<td>Survival rate of seedlings</td>
<td>( \rho )</td>
<td>—</td>
<td>90%</td>
<td>[26]</td>
</tr>
<tr>
<td>Carrying capacity</td>
<td>( K_{ij} )</td>
<td>ramets/0.4 ha</td>
<td>1,936,000</td>
<td>[26]</td>
</tr>
<tr>
<td>Eradication rate of herbicides</td>
<td>( \omega )</td>
<td>—</td>
<td>90%, 95%, 99%</td>
<td>[27]</td>
</tr>
<tr>
<td>Labor cost</td>
<td>( C_{ij} )</td>
<td>$/0.4 ha</td>
<td>$3.25</td>
<td>*</td>
</tr>
<tr>
<td>Herbicide cost</td>
<td>( H_{ij} )</td>
<td>$/0.4 ha</td>
<td>$10.50</td>
<td>[27]</td>
</tr>
<tr>
<td>Budget allotted to control sericea in year t</td>
<td>( B(t) )</td>
<td>$</td>
<td>[0, 1400]</td>
<td></td>
</tr>
<tr>
<td>Revenue from hay</td>
<td>—</td>
<td>$/0.4 ha</td>
<td>$306</td>
<td>[28]</td>
</tr>
<tr>
<td>Revenue from forage</td>
<td>—</td>
<td>$/0.4 ha</td>
<td>$81.71</td>
<td>[28]</td>
</tr>
</tbody>
</table>

* Expert opinion. **Underlined parameter values correspond to percentage of seed dispersal, percentage of remaining seeds, and eradication rate of the herbicide represent reasonable baseline parameter values.
Therefore, nine different maps, each defined by a combination of three frequency and three abundance levels of the species, could be generated. However, for the sake of conciseness, the model is applied to the most representative five cases, which include extreme and average cases, and provide sufficient information regarding computational analysis: low frequency and low abundance (L-L), low frequency and high abundance (L-H), medium frequency and medium abundance (M-M), high frequency and low abundance (H-L), and high frequency and high abundance (H-H). Therefore, we utilize 50 maps and present the average results of ten maps for each case in each computational simulation (Figures 3.1–3.4).

Along with the initial population structure, we present model parameters with their symbols, units, and case study values (Table 3.1), in order to demonstrate the general behavior of the model. Sericea ramets generally start to produce seeds after two growing seasons, with the majority of ramets producing seed in year three. Therefore, we divide the sericea population into one-, two-, and three+-year-old age classes in order to incorporate different plant mortality and seed production rates for each age class.

3.4. Results

The proposed mathematical model is solved using the “CONOPT” in AMPL [29] with a nonlinear global optimization algorithm. In this section, we conduct four different computational simulations and provide the results and sensitivity analyses.

3.4.1. Growth Behavior of Different Age Groups over 45 Years

In the first computational simulation, we analyze yearly population changes of sericea without herbicide treatment by observing the growth of one-, two-, and three+-year-old ramets
on a 10 x 10 landscape for five different initial populations with different frequencies and abundances for 45 years (Figure 3.1).

In the M-M, H-L, and H-H cases (Figures 3.1(c)–(e)), the growth of sericea follows a bi-logistic growth form, where there are two distinct phases, each with a logistic pattern as proposed by Meyer [30]. On the other hand, in the L-L and L-H cases (Figures 3.1a, b), the growth of sericea follows a multi-logistic growth [31] with multiple, sequential and overlapping phases of simple logistic form. Here, multi-logistic growth represents a growth function that includes serial, overlapping logistic phases, in which a successive section of the multi-logistic curve shows a slowing rate of growth as the population approaches the carrying capacity and finally saturate when carrying capacity is reached (Figures 3.1(a), (b)). For example, in Figure 3.1(a) (L-L case), a logistic growth function or a growth phase is observed from year 9 to year 14, while another logistic function occurs from year 13 until year 18. Over a few decades, the multiple logistic growth functions dampen and show an asymptotic behavior. Note that we observe only two logistic growth phases in the M-M, H-L, and H-H cases, because carrying capacity is reached in these cases faster than the L-L and L-H cases.
Figure 3.1. Field-level invasion by sericea for three age classes over 45 years in the absence of control measures under initial conditions consisting of (a) low frequency and low abundance (L-L), (b) low frequency and high abundance (L-H), (c) medium frequency and medium abundance (M-M), (d) high frequency and low abundance (H-L), and (e) high frequency and high abundance (H-H).
In Figure 3.1, we also observe that in all cases, the population of one- and two-year-old ramets is oscillating in time due to higher loss rates than older stages and limitations imposed by the carrying capacity, whereas the population of three*-year-old ramets is increasing monotonically each year until carrying capacity is reached. This observed pattern of sericea growth can explain the multi (bi)-logistic behavior on a landscape with variation in the initial frequency and density of invasion among cells. Because reproduction is high and dispersal distance is limited, cells that have sericea plants quickly reach carrying capacity, while establishment into unoccupied cells is relatively slow. At the entire landscape scale, this translates into a multi-logistic rather than a smooth logistic pattern. This response is consistent with the responses of the different age classes. For example, at low-frequency invasion, the contribution of the first two age groups is high until newly occupied cells reach carrying capacity, at which point the three*-year-old age class becomes dominant, and the one- and two-year-old age classes start to diminish. Because new cells on the landscape are occupied, a spike in the one- and two-year-old age classes occurs followed by a shift to the three*-year-old age class forming a logistic growth phase. For landscapes with moderate- to high-frequency invasion (Figures 3.1(c), (e)), this spike in one- and two-year-old age classes is much higher, and we observe fewer growth phases than for landscapes with low-frequency invasion.

3.4.2. Impact of Eradication Rate on Economic Damages

Given the patterns of population growth without control measures, we examine the budget necessary to control invader growth. Figure 3.2 illustrates the tradeoff between the cost of control measures and economic damage. Because the effectiveness of the treatment is uncertain due to year-to-year variation or the care with which herbicide is applied, we perform
sensitivity analysis on different values of the eradication rate (90%, 95%, and 99%) to analyze their impact on the yearly economic damage for different budget levels over 15 years.

While a budget level of $0.1 thousand is sufficient for the L-L and L-H cases (Figures 3.2(a), (b)), $1 thousand, $0.8 thousand, and $0.6 thousand are required to completely eradicate the sericea population for 90%, 95%, and 99% eradication rates, respectively, in the M-M case. The necessary budget increases to $0.7 thousand for the 99% eradication rate in the H-L case (Figures 3.2(c), (d)).

In the H-H case (Figure 3.2(e)), sericea is eradicated with a budget allocation of $1 thousand for a 99% eradication rate but increases to $1.3 thousand for a 95% eradication rate. Here, a budget allocation of $1 thousand for a 90% eradication rate will lead to a damage level less than $0.05 thousand but is not sufficient to completely eradicate the sericea population due to the widespread seed bank and the remaining few three-year-old ramets, which potentially generate an enormous amount of seeds that will add to the population in the following years.
Figure 3.2. Tradeoff between average yearly damages and budget allocation for different eradication rates over 15 years under initial conditions consisting of (a) low frequency and low abundance (L-L), (b) low frequency and high abundance (L-H), (c) medium frequency and medium abundance (M-M), (d) high frequency and low abundance (H-L), and (e) high frequency and high abundance (H-H). Values are non-zero near/at high values on the x-axis. Note differences in y-axis scale.
3.4.3. Impact of Dispersal Rate on Economic Damages

Since the dispersal rate is potentially sensitive to variation in wind, animal, and human activity, we also perform sensitivity analysis on the impact of different spread rates (0.01%, 0.1%, and 1%) on the yearly damage for different budget levels over 15 years. For each of these dispersal scenarios, the control costs (yearly budget on x-axis in Figure 3.3) and the resulting economic damage are inversely related. When the dispersal rate level is increased from 0.1% to 1%, the increase in the average yearly damage is equal to $1.15 thousand for the L-L case, and $0.41 thousand for the H-L case under no treatment (budget allocation = 0 on x-axis in Figure 3.3). Furthermore, although the average yearly damage in the H-H case is more than the average yearly damage in the M-M case for all budget allocations, the impact of dispersal rates on the average yearly damage is more apparent in the M-M case. This occurs because seeds are more likely to spread to already-invaded cells in the H-H case than the M-M case, and the former is closer to the carrying capacity than the latter.
Figure 3.3. Tradeoff between average yearly damages and budget allocation for different dispersal rates over 15 years under initial conditions consisting of (a) low frequency and low abundance (L-L), (b) low frequency and high abundance (L-H), (c) medium frequency and medium abundance (M-M), (d) high frequency and low abundance (H-L), and (e) high frequency and high abundance (H-H). Values are non-zero near/at high values on x-axis. Note differences in y-axis scale.
3.4.4. Impact of Different Treatment Strategies on Economic Damages

We also examine the impact of three different treatment frequency strategies—every year (1-year), every two years (2-year), and every three years (3-year)—on the average yearly and total (cumulative) damages over 15 years (Figure 3.4). Computational tests are conducted by equally allocating a total treatment budget of $4.5 thousand for every year ($0.3 thousand), every two years ($0.56 thousand), and every three years ($0.9 thousand) of the fifteen-year period, starting with treatment from the beginning of year 1.

Without treatment, the total damage increases from $15.2 thousand to $206.6 thousand by year 15, and all strategies reduce damage considerably, compared to no treatment in all cases (Figure 3.4). In the L-L and L-H cases, the 1-year and 2-year strategies are the two best options, while the 3-year treatment strategy results in higher yearly and total damages (Figures 3.4(a), (b)). On the other hand, for the M-M case, the 2-year treatment strategy is more beneficial than the 1-year treatment (Figure 3.4(c)). Here, less frequent control measures allow more of the population to be treated than the 1-year approach, and—because the recovery of the invader in a cell requires at least two years—the benefit of treating larger areas less frequently exceeds that of smaller areas treated more frequently. In other words, the immediate and higher reduction in total population that is the result of using the 2-year treatment approach is more beneficial than the 1-year approach, even though the economic damage as a result of the 2-year strategy exceeds the economic damage of the 3-year strategy in some years.
Figure 3.4. Impact of different treatment frequency strategies on average yearly damage (presented on y-axis) and total (cumulative) damage (given above each subfigure) over 15 years under initial conditions consisting of (a) low frequency and low abundance (L-L), (b) low frequency and high abundance (L-H), (c) medium frequency and medium abundance (M-M), (d) high frequency and low abundance (H-L), and (e) high frequency and high abundance (H-H). Note differences in y-axis scale.
Results for the H-L and H-H cases suggest that the 3-year strategy will result in the lowest total damages costs over a 15-year period. The sericea population reaches carrying capacity earlier in the H-H case, compared to the other four cases, due to its high initial abundance. After the carrying capacity is reached, the $0.9 thousand budget allocation compensates for damage in the previous years in the 3-year treatment strategy and thus causes less total damage at the end of year 15. On the other hand, the 1-year treatment results in more consistent damage levels than the 2-year and 3-year treatment strategies, which have substantial year-to-year variation in damages.

3.5. Discussion

We present a novel spatio-temporal dynamic model, which integrates biological models into a decision theory framework, while incorporating seed bank and dispersal, different age classes, growth rates, treatment costs, budget, and relevant economic loss. Unlike previous spatial-temporal methods (for a detailed discussion of these methods, see, e.g., [5]) here the growth of the invasive population within each cell is modeled using the seed bank and influenced by the invasion state of neighboring cells, while the population is divided into classes of different age groups in order to reflect different seed production and loss rates of each age group into the model. Numerical results provide insights into biological growth and spread behavior of the species, in addition to strategies addressing relevant management questions.

The first key result is that the population growth response of sericea is more complex than simple logistic growth. In fact, the population follows a multi (bi)-logistic growth form, where there are multiple (two) distinct phases, each with a logistic pattern. Our results support Meyer’s contention [30] that the bi-logistic is useful in representing the growth of many
systems that contain complex growth processes that are not well modeled by the simple logistic function. Here, we observe logistic phases of growth where in the first half of each phase, the first two age groups are dominant, and in the second half, the three+ year-old age class becomes dominant, until carrying capacity is reached. Although not typically examined, logistic growth of a population may undergo oscillations of one type or another, for many reasons including frequency related to age structure and time-lag effects [32]. Introduced species commonly exhibit a lag-phase in which the non-native species remains at low abundance for an extended time before increasing exponentially [33]. Several proposed hypotheses suggest that this pattern results from genotypic, demographic, or extrinsic factors [34]. The multi (bi)-logistic response pattern exhibited by our model suggests that demographic factors may explain short-term lag-patterns and, when coupled with variation in extrinsic factors, may contribute to longer-term lag patterns. Such insight would not be possible with more simplistic models that ignore the biological detail included in our model and is likely to be relevant to other species with high seed production and persistent seed banks. Furthermore, the multi (bi)-logistic population growth pattern suggests that the timing of control measures may have stronger or weaker effects on the invader, depending on when treatment is applied.

Second, given a target goal, the model addresses efficient management strategies regarding the following: (1) how large the allocated yearly budget needs to be, (2) the size of the infestation and where it should be targeted, and (3) how often treatments should be applied to be effective. In this paper, computational results demonstrate growth responses for three age classes under no-treatment, tradeoffs between damage and budget levels, and the minimum required resources that must be allotted to alleviate the spread of sericea under
various treatment and management scenarios.

Third, we perform sensitivity analysis with respect to eradication rate and seed dispersal parameters to analyze the impact of uncertainty on the model outputs and observe the model behavior for extreme scenarios. Not surprisingly, the higher the eradication rate, the lower the damage levels, but interestingly, the impact of different eradication rates becomes clearer as the frequency levels increase. Such analyses illustrate the potential tradeoffs between the cost of treatment and its effectiveness for different initial population conditions. For example, by comparing the marginal (extra) cost with the marginal economic damage reduction benefit of using a more effective herbicide, managers decide whether to invest in herbicide or not. Note that while we consider the most effective herbicide treatment and the corresponding cost in the sericea treatment case, the bio-economic model can also be extended to include various herbicide types or control strategies with their related costs. The results of this experiment also suggest that the average yearly damage increases as the dispersal rate increases for all budget levels and cases, but the dispersal rate has a higher impact for low-frequency initial population distributions than high-frequency initial invasions. Thus, a key component of invader control is the prevention of seed dispersal by reducing human and animal interaction [35].

Next, we evaluate three treatment timing strategies—1-year, 2-year, and 3-year—that could be used by managers, and we compare them with each other as well as the no-treatment option. Results suggest that effectiveness of the control strategy is highly dependent on initial population levels. With a limited budget, it is better to treat yearly if the initial population abundance and frequency is low, while it may be better to apply treatment every second or
third year (with a higher per-treatment budget amount) when the frequency and abundance are high (Figure 3.4).

In this paper, we address uncertainty by performing sensitivity analyses on different stochastic parameters such as budget, eradication rate, and dispersal rate. However, if the probability distributions of uncertain parameters are known or can be estimated, those parameters could be directly incorporated into the optimization model by defining them as random variables. The resulting stochastic nonlinear model could then be solved using stochastic optimization algorithms or heuristic approaches. Furthermore, if difficulties, including the quantification and formulation of ecological damages and preferences of stakeholders, are solved, this research could be extended to considering multiple objectives of different stakeholders including economic and ecological damages.

Our spatial-temporal model can be applied to any species for which age structure is relevant such as fish, insects, mammals and plants (see, e.g., [36-39]). As an example, biofuel crops that pose a risk of invasiveness could be analyzed in combination with economic optimization models [40-43]. Furthermore, model equations that represent age structured growth (Equations (3.1–3.3)) and carrying capacity (Equations (3.7–3.8)) can be adapted to model the growth of stage- or size-structured species and estimate the population abundances of different stage and size groups given carrying capacity limitations, respectively while seed generation and seedbank based growth equations (Equations (3.4–3.6)) can be adapted to model dormancy and various offspring generation, accumulation and dispersal mechanisms.

The results of the bio-economic optimization approach illustrate the potential for new optimization approaches that incorporate demographic detail and spatio-temporal realism for
invasive species control into a single-decision framework. Furthermore, while the proposed model is specific enough to capture biological realism, it also has the potential to be generalized to a wide range of invasive plant and animal species under various management scenarios in order to identify the most efficient control strategy for managing invasive species over large, heterogeneous landscapes and long time periods.

3.6. References


CHAPTER 4

OPTIMIZING INVASIVE SPECIES MANAGEMENT: A MIXED-INTEGER LINEAR PROGRAMMING APPROACH

4.1. Introduction

Invasive species are one of the world’s most serious environmental problems. They pose significant damage to native vegetation by competing for shared resources and crowding them out [1], thus reducing the benefit from recreation areas [2], wiping out diversity and shelter for wildlife [3], and altering fire regimes and posing a fire threat to communities [4-5]. The economic costs associated with invasive species also escalate proportionally to the size of the invasion. Pimentel et al. [6] report that the cost of invasive species to the United States’ economy is around $120 billion every year and increasing steadily.

Due to the high economic costs of invasive species, several optimization models have been proposed to control them (see, e.g., [7-12]). Invasive species management (ISM) is a complex problem because various factors contribute to a species’ invasiveness. For example, each invasive species demonstrates distinct population dynamics, such as varying growth and dispersal rates, and impacts native vegetation in various ways. In addition to a species’ biological traits, landscape characteristics of the potential habitat also affect its growth and spread. Given different biological traits, the management problem is to control the adverse impacts of the invasive species over a landscape under scarce economic resources. In this study, our goal is to develop an MIP model to solve this highly complex ISM problem. The objective of the MIP model is to minimize the economic damage associated with invasive species under a limited treatment budget while taking into account its spatial dynamics. The
MIP model exploits significant biological aspects (age classes, survival, growth, and dispersal) of invasive species.

Our paper focuses on the computational difficulty of the ISM problem, which has remained the biggest challenge to implementing those models in computational software. Difficulties with the proposed model include the formulation of nonlinear biological intricacies, economic restrictions, as well as the computational burden caused by the complexity of modeling spatial heterogeneity and temporal dimensions in such a biological system. Even without considering the complicating biological factors, this is a spatio-temporal resource allocation problem, which is shown to be NP-complete in the strong sense [13]. In order to tackle the complexity of the model, we elegantly formulize the seedbank-based growth and linearize the nonlinear capacity and treatment constraints, thus leading to an MIP model. This linear model is preferable to nonlinear models because implementation of the former is easier and more straightforward for domain experts, thus providing significant computational advantage for stakeholders in ISM (land managers, government, invasive species control experts, and resource conservationists).

We demonstrate the use of our model for finding optimal treatment strategies to control an invasive weed sericea in the Great Plains of North America. Sericea is a perennial legume that is native to Asia. Once introduced into a landscape, it suppresses and crowds out native vegetation. Furthermore, it is very drought and soil tolerant, which enables the plant to thrive under conditions where other species struggle. A high seed-production rate per each stem and high dispersal potential accelerates the spread of sericea to surrounding sites. Once
established, sericea is very difficult to remove due to its seed bank, which may remain viable for decades.

In a previous study, Büyüktahtakın et al. [14] incorporate important biological characteristics of sericea, including age-dependent fecundity and mortality, seed dispersal, seed bank formation, and density-dependent growth dynamics, into a spatially explicit nonlinear model. Due to the complexity and nonlinearity of the model, the problem cannot be solved as a full dynamic model over a seven-year planning horizon. In order to handle this computational difficulty, the authors use a rolling horizon heuristic, where the NLP model is solved for each period, and the resulting population density at each cell is used as the next period’s initial condition. The rolling horizon method requires that a certain budget level is set in advance for each time period. This method may lead to a suboptimal solution because it does not handle current and forecasted damages at the same time.

Our approach significantly differentiates from Büyüktahtakın et al [14] by considering binary treatment decisions; incorporating dispersal direction, distance, and probabilities into the mathematical model; and linearizing the nonlinear model. In the proposed model, we first integrate the seed-dispersal probabilities by using a Cauchy dispersal kernel [15] and then incorporate seed dispersal as a function of distance between the centers of two neighboring cells and the direction of the invasion. Because we represent treatment decisions using binary variables, we obtain an MINLP model, which is then converted into an equivalent MIP model by using linearization techniques that exist in the literature. We show that this model obtained by the proposed linearization method provides higher-quality solutions compared to four other linearization methods. While the linearization methods applied to our model exist in the
literature, as far as we know, they have never been considered to enhance the computational solvability of such a model. Furthermore, previous computational analysis on the impacts of the big-M value is limited (see, e.g., [16-18]). To our knowledge, this is the first study that elaborates on the use of linearization in a biologically complex model and provides an extensive analysis of the value of big-M. Our detailed discussion on the computational impact of the big-M value could also offer insight into many other applications in spatial conservation and environmental resource allocation.

Our approach leads to a number of important results, as follows:

- First, our extensive computational experiments on the comparison of the MIP model solution with its MINLP counterpart and its NLP relaxation show that the proposed techniques are effective to solve a practical-size ISM problem. In particular, while NLP and MINLP models cannot even be solved as a full optimization problem for a 3-by-3 gridded landscape in a two-year time period, the MIP model outperforms NLP and MINLP models by optimally solving the problem for a 10-by-10 gridded landscape in a seven-year time period.

- Second, the fact that our linear model is solvable with available software such as CPLEX is a major advantage, providing a significant computational benefit. In particular, our MIP model could be easily incorporated into a decision-support tool and utilized by decision makers in ISM.

- Third, our model provides significant insights that would not be possible with existing models and methods in ISM. Specifically, the proposed MIP model solves the full dynamic problem and thus handles the damage of both current and forecasted
invasions over the entire planning horizon at the same time. This feature enables us to
dynamically allocate the total budget over space and time, as opposed to solving the
problem statically with fixed periodic budgets, which has been a common practice in the
existing literature (see e.g., [19-21]).

4.1.1. Progress in Spatial Optimization Models for Invasive Species

Spatially explicit mathematical models have been used in a variety of studies to handle
landscape heterogeneity and to account for the spatial spread of invasive species. This
approach is important because population growth and dynamics on each spatial patch affects
the extent of current and future invasions, and thus economic cost. For example, Bhat et al.
[22] develop a spatio-temporal trapping strategy to minimize the sum of economic and
ecological loss due to beaver-inflicted timber damage and beaver-trapping cost. They develop a
solution algorithm where the output in a preceding iteration is used as input in the following
iteration until the solution converges, based on a prespecified error criterion. Batabyal and
Beladi [23] develop a partial differential equations model by using a queueing theory
framework to prevent invasive species. They compare two Markovian processes to analyze
inspection methods for different shipment cases. In another study, Blackwood et al. [24]
develop a linear-quadratic control approach to minimize the total cost associated with the
presence of Spartina alterniflora and its harvesting cost by introducing a connected patch-based
model. Previous work also incorporates spatially explicit NLP models to represent the spread of
bio-invasions and explore control strategies [4, 11, 20, 25-30].

While spatially explicit models are central to invasive species control models, integrating
biological, spatial, and temporal dimensions of the problem into a mathematical model
aggravates the computational complexity and solution time [14, 31]. In order to deal with arising computational difficulties, researchers reduce the dimensions of the problem by limiting the number of variables, number of cells, or number of time periods, or by simplifying dispersal patterns and the interdependency of spatial cells. On the other hand, previous spatially explicit studies that incorporate biological complexity use heuristic strategies for solving the problem, which lead to suboptimal solutions. For example, Aadland et al [19] report that the full nonlinear spatial model is unsolvable even on a 2-by-2 grid, but instead they develop an approximately optimal recursive linear model, which is similar to a rolling horizon approach. Cacho et al [15] and Epanchin-Niell and Wilen [32] develop spatially explicit models for monitoring and controlling pest invasions, respectively. In order to simplify the problem, they consider the presence or absence of propagules on a cell rather than the abundance of propagules per cell in their model.

In this paper, we fill the gap in the literature by solving the invasive species control problem optimally as an MIP model for a 10-by-10 hypothetical landscape over a seven-year planning horizon. The proposed spatially explicit MIP model exploits the significant biological aspects of the case study species, sericea. Results demonstrate that the proposed MIP model outperforms the equivalent MINLP and NLP (MINLP relaxation) models in terms of solution quality and potential problem size. Furthermore, our work forecasts expected damage levels for a sericea invasion under various invasion frequency and abundance levels, and provides optimal management strategies accordingly. Unlike previous studies, optimal levels of budget allocations are found in each time period to efficiently control the sericea invasion on a gridded landscape with different invasion scenarios. Results suggest that there is no rule-of-thumb
control strategy that is optimal in each different case. Instead, treatment strategies demonstrate that optimal control methods vary depending on different budget allocations and initial invasion conditions.

The structure of this chapter is as follows: In Section 4.2, we present the MINLP and MIP models for ISM. In Section 4.3, we present the model application and data on the case of sericea invasion. Section 4.4 follows with computational results that illustrate the efficiency of the MIP model and provide related management implications. Section 3.5 discusses results and offers directions for future research.

4.2. Model

Regular rectangular or square grids are frequently used to study ecosystems, and for observations, experiments, and simulations [33]. Therefore, we build our model based on a map grid representation of an area with equal-size square cells to simulate the spatially explicit population dynamics of invasive species and to incorporate economic principles. The order of events for the biological growth and management of invasive species is shown in Figure 4.1.

Assuming that some events such as seed production, seed dispersal, seed bank generation, and age transitions occur simultaneously, the sequence of events can be described as follows: Each plant produces seeds in a given time period. After seed production, some seeds disperse to surrounding cells, and some remain within the seed originating cell. The remaining seeds and those dispersed from surrounding cells either germinate or remain in the soil, thus creating the seed bank (Figure 4.1(a)). These events are followed by seed germination. The density-dependent structure ensures that plants in the lower-age class move to the upper-age class based on the available space remaining from the established older-age groups in a given
cell (Figure 4.1(b)). At the beginning of the following time period, herbicide treatment will be applied, and the invasion population is reduced, based on the available budget (Figure 4.1(c)). Surviving plants of an appropriate age will produce seeds as described above. This cycle of events take place at each time period. Therefore, the management of invasive species is a dynamic problem, which is concerned with finding optimal spatial and temporal treatment strategies to minimize related damage over a multi-period timeframe.

Figure 4.1. Representation of population dynamics and control of invasive species

4.2.1. Mathematical Formulation

Following Büyükahtakın et al. [14], we describe the formulation of the MINLP model as follows: Let \( t \in [0, T] \) be any year of the planning horizon, where \( T \) represents the final time period. The area is divided into square cells with \( I \) rows and \( J \) columns, as shown on a 3-by-3 landscape in Figure 4.1. Any cell can be characterized by its coordinates \((i,j)\), where \( i \in \{1,2,\ldots,I\} \) and \( j \in \{1,2,\ldots,J\} \). We define a binary decision variable \( \chi_{ijt} \) that is equal to 1 if cell \((i,j)\) is treated in
year $t$, and 0 otherwise. Age clusters (age categories) are defined as $k = 1, 2, 3, ..., n^*$, where $k$ represents the age of each cluster, and $n^*$ defines the age cluster $n$ and older populations.

**Seed Dispersal Constraints:** Herbicide treatment is assumed to be applied just before the seed-production season in order to reduce the expected total seed production in cell $(i,j)$. Additionally, it is assumed that while some seeds remain within the cell, others will disperse to surrounding cells (Figure 4.1(a)). We assume that seed dispersal follows a Cauchy kernel [15]. The probability of seed dispersal depends on the direction of spread $\theta$ and distance between the center of the seed originating cell $(h,q) \in \Theta^i$ and center of the seed landing cell $(i,j)$, where $\Theta^i$ is the set of neighboring cells of cell $(i,j)$. Therefore, seed dispersal to cell $(i,j)$ from surrounding cells $(h,q) \in \Theta^i$ is formulated as

$$SD_{ijt} = \lambda \sum_{k=1}^{n^*} \sum_{(h,q) \in \Theta^i} \tau_{(h,q) \rightarrow (i,j)}^{\theta} P_{hqt}^k S_k^k \forall i, j, t$$

(4.1)

where $\lambda$ is the percentage of total seeds produced over all cells $(h,q) \in \Theta^i$ that disperse to cell $(i,j)$, $\tau_{(h,q) \rightarrow (i,j)}^{\theta}$ is the probability of seed dispersal from the seed originating cell $(h,q)$ to the landing cell $(i,j)$ in the direction of $\theta$, $P_{hqt}^k$ is the number of after-treatment populations of the invasive plants of age cluster $k$ in cell $(h,q) \in \Theta^i$, and $S_k^k$ is the number of seeds produced by each stem in each age cluster $k$. The number of seeds remaining in cell $(i,j)$ after dispersal at time $t$ is given as

$$SR_{ijt} = \omega \sum_{k=1}^{n^*} P_{ijt}^k S_k^k \forall i, j, t$$

(4.2)

where $\omega$ is the percentage of seeds that are produced in cell $(i,j)$ and remain after dispersal takes place to surrounding cells $(h,q) \in \Theta^i$.  

66
**Seed Bank Accumulation Constraint:** After seeds disperse to surrounding cells, the remaining seeds in cell \((i,j)\) and seeds dispersed from neighboring cells will form a seed bank (Figure 4.1(a)). It is assumed that only a portion of the seeds will remain viable due to natural disturbances such as scarce resources and seed predators. In addition, only viable seeds are able to germinate. With the given information, the seed bank population is formulated as

\[
SB_{ijt} = SB_{ij0} (\gamma - \alpha)^t + \sum_{s=0}^{t} ((\gamma - \alpha)^{s-t} (SD_{ij} + SR_{ij})) \quad \forall i, j, t
\]  

(4.3)

where \(SB_{ij0}\) is the initial seed bank population in cell \((i,j)\) at time 0, and \(\gamma\) and \(\alpha\) represent longevity and germination rate, respectively. Therefore, the seed bank equation includes the initial seed bank as well as all produced and dispersed seeds in cell \((i,j)\) from time zero until time \(t\), while reflecting the decay of seeds over time.

**Transition Population Constraints:** Germinated seeds from the seed bank become seedlings and turn into one-year-old plants with a given success rate (Figure 4.1(b)). In addition, each individual moves up one age class in a one-year period with a certain loss rate [34] and grows up until age \(n^+\), where it remains in that age cluster until it dies out [35]. Therefore, the transition population level for each age class is formulated as

\[
TP_{ij,t+1}^k = \alpha \rho SB_{ijt} \quad k = 1 \text{ and } \forall i, j, t
\]  

(4.4)

\[
TP_{ij,t+1}^k = P_{ij}^{k-1} (1 - \psi^{k-1}) \quad k = 2, ..., n-1 \text{ and } \forall i, j, t
\]  

(4.5)

\[
TP_{ij,t+1}^k = P_{ij}^{k-1} (1 - \psi^{k-1}) + P_{ij}^k (1 - \psi^k) \quad k = n^+ \text{ and } \forall i, j, t
\]  

(4.6)

where \(\alpha\) is the germination rate, \(\rho\) is the probability of becoming a one-year-old plant after germination, \(\psi^k\) is the loss rate of individuals while transitioning from age cluster \(k\) to \(k+1\), and
\( TP_{ij,t+1}^k \) is the transition population level of invasive plants for all age clusters before carrying capacity, \( K_{ij} \), is considered.

**Carrying Capacity and Actual Population Constraints:** Since a given cell \((i,j)\) accommodates individuals of different-age clusters, there will be intraspecific competition among individuals for the same resources. It is expected that younger individuals are vulnerable and weak against older ones. Therefore, as the abundance of a cell increases, older individuals will suppress and dominate younger ones. Furthermore, once the cell population reaches the carrying capacity, one-year-old individuals will only populate the available space when there is any loss from older individuals. Given this information, the actual individual before-treatment population, \( BP_{ij}^k \), is formulated as

\[
BP_{ij}^k = \min \{TP_{ij}^k, K_{ij}\}, \quad k = n^+ \text{ and } i,j,t
\]  

(4.7)

\[
BP_{ij}^k = \begin{cases} 
0 & \text{if } K_{ij} - \sum_{v=k+1}^{n^+} BP_{ij}^v \leq 0, \quad k = 1,\ldots,n-1 \text{ and } \forall i,j,t \\
\min \{(K_{ij} - \sum_{v=k+1}^{n^+} BP_{ij}^v), TP_{ij}^k\} & \text{otherwise},
\end{cases}
\]

(4.8)

Note that equation (4.7) allows the model to give priority to the oldest-age cluster, \( n^+ \), in a given cell \((i,j)\). If the transition population of oldest individuals in cell \((i,j)\) is more than the carrying capacity, then the before-treatment population (real population of oldest individuals before treatment) will be set to the carrying capacity; otherwise, the before-treatment population of oldest individuals will be equal to their transition population level. The first part of equation (4.8) states that if the carrying capacity in cell \((i,j)\) is reached by individuals at age clusters \( k+1, k+2,\ldots, n^+ \) (where \( k = 1\ldots n^+-1 \)), then \( k^{th} \) and younger-age clusters will not be able to
populate in cell \((i,j)\). On the other hand, if the carrying capacity is not reached by individuals at age clusters \(k+1, k+2, \ldots, n^+\), then the remaining carrying capacity will be occupied by age clusters \(k, k-1, \ldots, 1\), respectively, depending on the available remaining space determined by ecological limits (carrying capacity). If the remaining carrying capacity is more than the transition population of the \(k^{th}\)-age cluster, then the before-treatment population of the \(k^{th}\)-age cluster will be set as the transition population level of that age group. However, if the remaining carrying capacity is less than the transition population level of the \(k^{th}\)-age cluster, then the remaining carrying capacity will be assigned as the before-treatment population of the \(k^{th}\)-age cluster.

**Treatment Constraint:** In order to decrease the damage of invasive species, herbicide treatment is applied to invaded cells (Figure 4.1(c)). In the case of treatment, the before-treatment population is multiplied by \((1 - \phi \chi_{ijt})\), where \(\phi\) is the effectiveness rate of the herbicide treatment. Therefore, the after-treatment population is calculated as

\[
P_{ijt}^k = BP_{ijt}^k \times (1 - \phi \chi_{ijt}) \quad \forall i, j, k, t
\]

(4.9)

where \(\chi_{ijt}\) is a binary variable indicating where and when to apply herbicide treatment in a given time horizon under a given budget level. The binary variable \(\chi_{ijt}\) gets a value of 1 if cell \((i,j)\) receives herbicide treatment in period \(t\), and 0 otherwise.

**Budget Constraint.** Treatment is limited by the available budget, \(\Upsilon\), for the entire time horizon. Therefore, the budget constraint becomes

\[
\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} (L_{ij} + H_{ij}) \chi_{ijt} \leq \Upsilon
\]

(4.10)

where \(L_{ij}\) and \(H_{ij}\) are labor and herbicide costs per cell \((i,j)\), respectively.
**Objective Function:** The objective of the model is to minimize total damage caused by the invasive species population in all cells and all periods of the planning horizon. Therefore, the objective function is formulated as

\[
\text{Minimize } \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} \Gamma_{ijt} \tag{4.11}
\]

where \( \Gamma_{ijt} \) is the expected damage resulting from the invasive species in cell \((i,j)\) at time \(t\). The damage function is formulated as

\[
\Gamma_{ijt} = E_{ijt} \cdot \left( \sum_{k=1}^{n} P_{ijt}^k \right) / K_j \quad \forall i, j, t \tag{4.12}
\]

where \( E_{ijt} \) represents the expected revenue from cell \((i,j)\) at time \(t\). Therefore, revenue loss (damage) in a given area is proportional to the ratio of the total population over all ages with respect to the carrying capacity in that area.

4.2.2. Dispersal Uncertainty

As mentioned in the model, each age cluster \(k\) produces \(S^k\) seeds per time period, and it is assumed that seeds will disperse to neighboring cells. However, the amount of seed dispersal varies, depending on the distance between neighboring cells in spatially explicit models. Therefore, we consider a similar technique, presented in the work of Cacho et al. [15], to determine the probability of seed dispersal from cell \((i,j)\) to neighboring cells. A Cauchy dispersal kernel is used to find \(r_{(i,j)\rightarrow(h,q)}^\theta\), which defines the probability that a seed disperses from cell \((i,j)\) to cell \((h,q)\) in direction \(\theta\). In order to calculate \(r_{(i,j)\rightarrow(h,q)}^\theta\), we define the term \(d_{(i,j)\rightarrow(h,q)}\) as
\[
\delta_{(i,j)\rightarrow(h,q)} = \frac{1}{1 + \left( \frac{d_{(i,j)\rightarrow(h,q)} \sqrt{A}}{\varsigma} \right)^2}
\]  

(4.13)

where \( \varsigma \) is the dispersal constant, \( d_{(i,j)\rightarrow(h,q)} \) is the distance between cells \((i,j)\) and \((h,q)\) in meters, and \( A \) is the area of the cell in square meters.

Using \( \delta_{(i,j)\rightarrow(h,q)} \) values computed in equation (4.13), \( \tau_{(i,j)\rightarrow(h,q)}^\theta \) is calculated as

\[
\tau_{(i,j)\rightarrow(h,q)}^\theta = \frac{\delta_{(i,j)\rightarrow(h,q)}}{\sum_{(h,q)\in\Theta^\theta} \delta_{(i,j)\rightarrow(h,q)}}
\]  

(4.14)

A probability matrix \((H)\) of dimension \(mn\), where \(m\) and \(n\) represent the number of rows and columns of the extent of dispersal, is generated using \( \tau_{(i,j)\rightarrow(h,q)}^\theta \). The values of matrix \( H \) are normalized so that \( \sum_{(h,q)\in\Theta^\theta} \tau_{(i,j)\rightarrow(h,q)}^\theta = 1 \). Our formulation, equations (4.1)–(4.14), represents key spatial and temporal dynamics in an ecosystem, such as growth and dispersal among cells [36]. Therefore, our proposed model remains valid in situations where the grid cells are irregular polygons.

4.2.3. Linearization of Nonlinear Constraints

Note that nonlinearities arise in the model as the result of utilization of conditional functions in equations (4.7) and (4.8), and the multiplication of a binary and continuous variable in equation (4.9). The MINLP model, including equations (4.1)–(4.12), is converted into an equivalent MIP model by linearizing the nonlinear constraints (4.7)–(4.9).

In this section, we present five methods to linearize equations (4.7) and (4.8). Method 1 demonstrates the proposed methodology, while Methods 2–5 utilize piecewise linearization
methods that exist in the literature. In order to conserve space, we demonstrate the use of linearization techniques in equation (4.8) only; however, the same methods can be applied to linearize equation (4.7). Finally, we compare the results of each method with respect to the objective function value in order to validate the proposed linearization technique.

### 4.2.3.1. Proposed Linearization Method

Let \( \tilde{K}^{k}_{ijt} = K^{k}_{ij} - \sum_{v=k+1}^{n} BP^{v}_{ijt} \) for each \( i, j, t, \) and \( k \). Note that the value of \( \tilde{K}^{k}_{ijt} \) can never be negative because equations (4.7) and (4.8) ensure that the landscape is occupied with a priority from higher to lower age clusters while carrying capacity is respected. In order to deal with nonlinearities in equations (4.7) and (4.8), these equations are converted into the following analytical expressions, respectively:

\[
BP^{k}_{ijt} = TP^{k}_{ijt} z^{k}_{ijt} + K^{k}_{ij} (1 - z^{k}_{ijt}) \quad k = n^{+} \tag{4.15}
\]

\[
BP^{k}_{ijt} = \tilde{K}^{k}_{ijt} y^{k}_{ijt} + TP^{k}_{ijt} (1 - y^{k}_{ijt}) \quad k = 1, \ldots, n-1 \text{ and } \forall i, j, t \tag{4.16}
\]

where \( z^{k}_{ijt} \) and \( y^{k}_{ijt} \) are binary variables that are defined to determine the before-treatment population for age cluster \( k \) in cell \((i,j)\) at time \(t\). Since the problem is modeled to minimize damage associated with the presence of invasive plants, equations (4.15) and (4.16) behave similarly as equations (4.7) and (4.8), respectively.

Note that equations (4.15) and (4.16) include the products of continuous and binary variables. Therefore, the quadratic terms in equation (4.16) are linearized using a traditional linearization method as follows:

\[
LB_{1} y^{k}_{ijt} \leq G^{k}_{ijt} \leq UB_{1} y^{k}_{ijt} \tag{4.17}
\]
\[ \tilde{K}^k_{ijt} - UB_1 (1 - y^k_{ijt}) \leq G^k_{ijt} \leq \tilde{K}^k_{ijt} - LB_1 (1 - y^k_{ijt}) \] (4.18)

\[ LB_2 (1 - y^k_{ijt}) \leq Q^k_{ijt} \leq UB_2 (1 - y^k_{ijt}) \] (4.19)

\[ TP^k_{ijt} - UB_2 y^k_{ijt} \leq Q^k_{ijt} \leq TP^k_{ijt} - LB_2 y^k_{ijt} \] (4.20)

where \( G^k_{ijt}, Q^k_{ijt} \) are the substitute variables, and \( LB_1 = 0, UB_1 = K_{ijt} \) and \( LB_2 = 0, UB_2 = M \) (where \( M \) is a very large number), are the lower and upper bounds for terms \( \tilde{K}^k_{ijt} \) and \( TP^k_{ijt} \), respectively. Thus, equation (4.16) is equivalent to \( BP^k_{ijt} = G^k_{ijt} + Q^k_{ijt} \), where \( G^k_{ijt} \) and \( Q^k_{ijt} \) are restricted by equations (4.17)–(4.20). The linearization of terms \( TP^k_{ijt} y^k_{ijt} \) in equation (4.15) and \( BP^k_{ijt} \chi^k_{ijt} \) in equation (4.9) uses this same technique.

4.2.3.2. Method 2: Piecewise Linearization Technique

In this section, we present the linearization of equation (4.8) using the piecewise linearization technique discussed by Bazaraa et al. [37], as follows: Figures 4.2(a) and 4.2(b) represent the piecewise linear function for the before-treatment population of age groups \( k = n^+ \) (Equation 4.7) and \( k = 1, \ldots, n^* - 1 \) (Equation 4.8), respectively. The piecewise linear function \( BP^k_{ijt} \) has breakpoints 0, \( K_{ijt} \), and \( M \) for age clusters \( k = n^+ \); and 0, \( \tilde{K}^k_{ijt} \), and \( M \) for age clusters \( k = 1, \ldots, n^* - 1 \). Continuous variables \( 0 \leq v_{ijt}^{k0}, v_{ijt}^{k1}, v_{ijt}^{k2} \leq 1 \) are defined for each breakpoint, and binary variables \( z_{ijt}^{k0}, z_{ijt}^{k1} \in [0,1] \) are defined for each interval between breakpoints.
Figure 4.2. Representation of piecewise linear function: (a) age class \( k = n^+ \) (Equation 3.7); (b) age class \( k = 1, \ldots, n^*-1 \) (Equation 3.8).

Then, for age cluster \( k = 1, \ldots, n^*-1 \) (Equation 4.8), the piecewise functions can be written as

\[
TP^{k}_{ij} = \underbrace{0 \cdot v^{k0}_{ij}}_{\text{Quadratic term}} + \underbrace{\tilde{K}^{k}_{ij} \cdot v^{k1}_{ij}}_{\text{Quadratic term}} + M \cdot \underbrace{v^{k2}_{ij}}_{\text{Quadratic term}}
\]

(4.21)

\[
BP^{k}_{ij} = \underbrace{0 \cdot v^{k0}_{ij}}_{\text{Quadratic term}} + \underbrace{\tilde{K}^{k}_{ij} \cdot v^{k1}_{ij}}_{\text{Quadratic term}} + \underbrace{\tilde{K}^{k}_{ij} \cdot v^{k2}_{ij}}_{\text{Quadratic term}}
\]

(4.22)

\[
v^{k0}_{ij} \leq z^{k0}_{ij}, \quad v^{k1}_{ij} \leq z^{k0}_{ij} + z^{k1}_{ij}, \quad v^{k2}_{ij} \leq z^{k1}_{ij}
\]

(4.23)

\[
z^{k0}_{ij} + z^{k1}_{ij} = 1, \quad v^{k0}_{ij} + v^{k1}_{ij} + v^{k2}_{ij} = 1
\]

(4.24)

Equations (4.21)–(4.24) ensure that \( BP^{k}_{ij} \) receives the minimum value between \( \tilde{K}^{k}_{ij} \) and \( TP^{k}_{ij} \) for each age cluster \( k = 1, \ldots, n^*-1 \).

4.2.3.3. Method 3 (Vielma and Nemhauser [18])

Vielma and Nemhauser [18] developed a different piecewise linearization method with fewer variables and constraints compared to Method 2. While equations (4.21) and (4.22) are similar in this method, the following set of equations is added to the model to reformulate equation (4.8):
\[ \nu_{ij}^{k0} + \nu_{ij}^{k1} + \nu_{ij}^{k2} = 1, \quad \nu_{ij}^{k0} \leq z_{ij}^k, \quad \nu_{ij}^{k1} \text{ unrestricted}, \quad \nu_{ij}^{k2} \leq 1 - z_{ij}^k \tag{4.25} \]

### 4.2.3.4. Method 4 (Li and Yu [38])

Using the piecewise linearization technique proposed by Li and Yu [38], we formulate the corresponding functions for the age cluster \( k = 1, \ldots, n^+ - 1 \) for equation (4.8) as

\[
BP_{ij}^k = \tilde{K}_{ij}^k - \underbrace{\tilde{K}_{ij}^k u_{ij}^k}_{\text{Quadratic term}} + \delta_{ij}^k \tag{4.26}
\]

\[
TP_{ij}^k + M (u_{ij}^k - 1) \leq \delta_{ij}^k \quad \text{and} \quad \delta_{ij}^k \geq 0 \tag{4.27}
\]

where \( u_{ij}^k \) represents a binary variable that is used for linearization. Equations (4.26) and (4.27) ensure that \( BP_{ij}^k \) receives the minimum value between \( TP_{ij}^k \) and \( \tilde{K}_{ij}^k \). If \( TP_{ij}^k \leq \tilde{K}_{ij}^k \), then \( u_{ij}^k = 1 \), which implies that \( BP_{ij}^k = TP_{ij}^k \). Otherwise, if \( TP_{ij}^k \geq \tilde{K}_{ij}^k \), then \( u_{ij}^k = 0 \), which ensures that \( BP_{ij}^k = \tilde{K}_{ij}^k \).

Note that we have quadratic terms in Methods 2–4, which arise as a result of the non-stationary breakpoint, \( \tilde{K}_{ij}^k \) (Figure 4.2(b)) and are due to the dynamic structure of the carrying-capacity constraint in equation (4.8). To eliminate the nonlinearity, we fix \( \tilde{K}_{ij}^k \) to constant values for each age cluster \( k = 1, \ldots, n^+ - 1 \) (see section 3.3 for case study values of \( \tilde{K}_{ij}^k \)).

### 4.2.3.5. Method 5 (Williams [39])

In this method, we demonstrate the use of a different technique to linearize the piecewise linear function shown in Figure 4.2(b) (Equation 4.8). This technique avoids the use of quadratic functions; however, it requires an additional big-M value for our problem. Based on
this method, equation (4.8) can be represented by the following linear inequalities for age clusters \( k = 1, \ldots, n^+ - 1 \):

\[
z_{ijt}^0 + z_{ijt}^1 = 1 \tag{4.28}
\]

\[
0 - \bar{M} \cdot (1 - z_{ijt}^0) \leq T_{ijt}^k \leq \tilde{K}_{ijt}^k + \bar{M} \cdot (1 - z_{ijt}^0) \tag{4.29}
\]

\[
\tilde{K}_{ijt}^k - \bar{M} \cdot (1 - z_{ijt}^1) \leq T_{ijt}^k \leq M + \bar{M} \cdot (1 - z_{ijt}^1) \tag{4.30}
\]

\[
T_{ijt}^k - \bar{M} \cdot (1 - z_{ijt}^0) \leq B_{ijt}^k \tag{4.31}
\]

\[
\tilde{K}_{ijt}^k - \bar{M} \cdot (1 - z_{ijt}^1) \leq B_{ijt}^k \tag{4.32}
\]

where \( \bar{M} \) is a large constant, and \( M \) is the breakpoint. Equation (4.28) ensures that \( T_{ijt}^k \) must belong to exactly one interval. If \( z_{ijt}^0 = 1 \), then equation (4.29) becomes restrictive, and equation (4.30) becomes redundant. Similarly, if \( z_{ijt}^1 = 1 \), then equation (4.30) becomes restrictive, and equation (4.29) becomes redundant. Finally, since we are indirectly minimizing \( B_{ijt}^k \), equation (4.31) provides a lower bound on \( B_{ijt}^k \) if \( z_{ijt}^0 = 1 \); otherwise, equation (4.32) provides the related lower bound.

4.3. Model Application

Table 4.1 shows the initial population structure and parameters of the model. The collected data is specific to sericea, since each invasive species has unique characteristics. Most of the parameter values used in this paper are compiled from the work of Büyüktahtakin et al. [14] in which data are collected by utilizing several resources, including journal papers, government sources, and expert opinions.
Computational experiments are conducted by using a 10-by-10 landscape (40 ha), where each cell represents 0.4 ha (1 square acre) of land. The spatial growth behavior of sericea invasion is simulated based on three different hypothetical initial invasion frequencies. Each initial population is represented by the percentage of invaded areas of the gridded landscape with different abundance levels. Therefore, the initial population maps are defined as 2% (low), 20% (medium), and 40% (high) invasion with U[1–10], U[11–50], and U[51–250] abundance, respectively, where U[a,b] denotes an integer number drawn uniformly from the interval [a,b].

**TABLE 4.1. INITIAL POPULATION STRUCTURE AND PARAMETERS (ADJUSTED FROM BÜYÜKTAHTAKIN ET AL. [14])**

<table>
<thead>
<tr>
<th>Description</th>
<th>Frequency (%)</th>
<th>Category</th>
<th>Case study value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of cells invaded in 10-by-10 landscape</td>
<td>2</td>
<td>Low invasion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Medium invasion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>High invasion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Symbol</th>
<th>Unit</th>
<th>Case study value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of seeds produced per stem per age</td>
<td>$k$</td>
<td>—</td>
<td>0,45,900*</td>
<td>[40]</td>
</tr>
<tr>
<td>cluster $k$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of seed dispersal to all</td>
<td>$\lambda$</td>
<td>—</td>
<td>0.8%</td>
<td>*</td>
</tr>
<tr>
<td>neighboring 24 cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of remaining seeds</td>
<td>$\omega$</td>
<td>—</td>
<td>99.2%</td>
<td>*</td>
</tr>
<tr>
<td>Longevity rate</td>
<td>$\gamma$</td>
<td>—</td>
<td>95%</td>
<td>[41]</td>
</tr>
<tr>
<td>Germination rate</td>
<td>$a$</td>
<td>seedlings/seed</td>
<td>6.8%</td>
<td>[42]</td>
</tr>
<tr>
<td>Survival rate of seedlings</td>
<td>$\rho$</td>
<td>—</td>
<td>90%</td>
<td>*</td>
</tr>
<tr>
<td>Loss rate from age cluster $k$ to $k+1$</td>
<td>$\psi_k$</td>
<td>—</td>
<td>22%, 9%, 4%**</td>
<td>[47]</td>
</tr>
<tr>
<td>Carrying capacity</td>
<td>$K_{ij}$</td>
<td>stems/0.4 ha</td>
<td>2,000,000</td>
<td>[42]</td>
</tr>
<tr>
<td>Effectiveness rate of treatment</td>
<td>$\omega$</td>
<td>—</td>
<td>95%</td>
<td>[43]</td>
</tr>
<tr>
<td>Labor cost</td>
<td>$L_{ij}$</td>
<td>$/0.4$ ha</td>
<td>$3.25$</td>
<td>*</td>
</tr>
<tr>
<td>Herbicide cost</td>
<td>$H_{ij}$</td>
<td>$/0.4$ ha</td>
<td>$10.50$</td>
<td>[43]</td>
</tr>
<tr>
<td>Aggregate budget allotted to control sericea</td>
<td>$\gamma$</td>
<td>$/0.4$ ha</td>
<td>[0,8000]</td>
<td></td>
</tr>
<tr>
<td>Revenue from hay</td>
<td>—</td>
<td>$/0.4$ ha</td>
<td>$306$</td>
<td>[14]</td>
</tr>
<tr>
<td>Revenue from forage</td>
<td>—</td>
<td>$/0.4$ ha</td>
<td>$81.71$</td>
<td>[14]</td>
</tr>
</tbody>
</table>

*Expert opinion, **Values separated by commas for “number of seeds produced per stem per age cluster $k$” and “loss rate from age cluster $k$ to $k+1$” represent case study values for 1-, 2-, and 3-year-old age clusters.

The seed production amount of sericea changes depending on the age of its stems [40]. It is noted that although sericea stems cannot generate seeds in their first year, two- and three-
year-old (and older) stems can produce an average of 45 and 900 seeds, respectively [34, 40]. Therefore, we divide the sericea population into three age clusters to account for different seed production amounts and mortality rates. After seed production, it is assumed that while 99.2% of total seeds produced remain in the originating cell, 0.8% of the seeds disperse to the surrounding cells by natural disturbances such as wind, and human or animal interaction.

Sericea seeds can live more than 20 years under the soil and can produce a long-lived seed bank [41]. Here we assume a seed bank decay rate of 10%, which implies that the survival rate of seeds (longevity rate) in the seed bank is 90% from one year to the next. Seeds are assumed to germinate 6.8% each year from the seed bank. Therefore, if herbicide treatment is not applied, then the seed bank replenishes each year with an increasing rate due to newly produced and dispersed seeds.

After seeds germinate from the seed bank and become seedlings, some portion of them does not survive due to lack of resources and competition among sericea stems. Therefore, we assume that only 90% (survival rate) of the seedlings will become a one-year-old plant at the end of the first year. In addition, while one-year-old plants become two-year-old plants, and two-year-old plants turn into three-year-old plants, with a loss rate of 22% and 9%, respectively, three-year-old plants remain in the three+ year-old age cluster with a loss rate of 4% each year. The carrying capacity is set to 2,000,000 plants per cell [42]. Therefore, after reaching the carrying capacity in a given cell, a new plant can grow only if an older plant dies out.

In the case of herbicide treatment, the treatment-effectiveness rate depends on the type of chemical used and timing of the application. We select the treatment effectiveness rate
as 95% in the case study [43]. Moreover, the treatment cost of each cell depends on the cost of labor and herbicide, which are estimated to be $10.50 [43] and $3.25 (expert opinion), respectively. The objective function minimizes the expected average damage caused by the sericea population over all cells and time periods, where each cell has an expected economic value of $194 from hay and forage [14].

Finally, utilizing the results of Büyükahtaşkin et al. [14], we assume that the three-year-old age group can invade almost the entire region in a steady-state condition, while the one- and two-year-old populations can have, at most, 500,000 and 1,000,000 stems, respectively, before reaching their steady-state population levels. Therefore, $\tilde{K}_{ij}$ is fixed to 500,000 and 1,000,000 stems for age clusters $k = 1$ and $k = 2$, respectively, in the quadratic terms in Methods 2–4.

4.4. Results

The results here will demonstrate the computational feasibility of the proposed model using various linearization methods and to provide insight regarding the optimal strategies for controlling invasive species under various scenarios. A time limit of 3,600 CPU seconds is imposed for solving all test instances under various scenarios. Selected results chosen from interesting problem configurations are reported here.

4.4.1. Impact of Different Budget Allocations on Overall Damage

In this section, we examine the impact of different budget allocations on total damage in a seven-year time period for low, medium, and high initial invasions, as described in section 4.3. Table 4.2 presents details of the computational performance of the MIP model obtained with the proposed linearization method under various types of invasion and budget allocations.
TABLE 4.2. TOTAL DAMAGE IN SEVEN-YEAR PERIOD BASED ON DIFFERENT SCENARIOS AND BUDGET ALLOCATIONS

<table>
<thead>
<tr>
<th>Invasion type</th>
<th>Budget allocation ($)</th>
<th>Big-M value (^a)</th>
<th>Expected damage ($) (^b)</th>
<th>Solution time (^c)</th>
<th>Gap (%) (^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>0</td>
<td>100</td>
<td>94.47</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>25</td>
<td>6</td>
<td>47.34</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>50</td>
<td>4</td>
<td>2.42</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>75</td>
<td>4</td>
<td>0.26</td>
<td>218</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>100</td>
<td>4</td>
<td>0.02</td>
<td>192</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>125</td>
<td>1.5</td>
<td>0.01</td>
<td>121</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>150</td>
<td>1.5</td>
<td>0.01</td>
<td>89</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>175</td>
<td>1</td>
<td>0.01</td>
<td>86</td>
<td>0</td>
</tr>
<tr>
<td>low</td>
<td>200</td>
<td>1</td>
<td>0.01</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>medium</td>
<td>0</td>
<td>150</td>
<td>3,195.94</td>
<td>235</td>
<td>0</td>
</tr>
<tr>
<td>medium</td>
<td>700</td>
<td>2</td>
<td>22.7</td>
<td>3,600</td>
<td>0.99</td>
</tr>
<tr>
<td>medium</td>
<td>900</td>
<td>1.5</td>
<td>2.2</td>
<td>3,600</td>
<td>0.56</td>
</tr>
<tr>
<td>medium</td>
<td>1,100</td>
<td>1.5</td>
<td>1.36</td>
<td>3,600</td>
<td>0.50</td>
</tr>
<tr>
<td>medium</td>
<td>1,300</td>
<td>1.2</td>
<td>1.25</td>
<td>3,600</td>
<td>0.47</td>
</tr>
<tr>
<td>medium</td>
<td>1,500</td>
<td>0.9</td>
<td>0.78</td>
<td>3,600</td>
<td>0.45</td>
</tr>
<tr>
<td>medium</td>
<td>1,700</td>
<td>0.9</td>
<td>0.64</td>
<td>2,891</td>
<td>0.36</td>
</tr>
<tr>
<td>medium</td>
<td>1,900</td>
<td>0.9</td>
<td>0.48</td>
<td>1,774</td>
<td>0</td>
</tr>
<tr>
<td>medium</td>
<td>2,100</td>
<td>0.7</td>
<td>0.37</td>
<td>1,008</td>
<td>0</td>
</tr>
<tr>
<td>high</td>
<td>0</td>
<td>150</td>
<td>14,443.71</td>
<td>1,800</td>
<td>0</td>
</tr>
<tr>
<td>high</td>
<td>1,000</td>
<td>86</td>
<td>624</td>
<td>3,600</td>
<td>0.99</td>
</tr>
<tr>
<td>high</td>
<td>2,000</td>
<td>9</td>
<td>45.37</td>
<td>3,600</td>
<td>0.98</td>
</tr>
<tr>
<td>high</td>
<td>3,000</td>
<td>7</td>
<td>6.73</td>
<td>3,605</td>
<td>0.81</td>
</tr>
<tr>
<td>high</td>
<td>4,000</td>
<td>5.25</td>
<td>3.31</td>
<td>3,600</td>
<td>0.76</td>
</tr>
<tr>
<td>high</td>
<td>5,000</td>
<td>2</td>
<td>0.32</td>
<td>752</td>
<td>0</td>
</tr>
<tr>
<td>high</td>
<td>6,000</td>
<td>1.5</td>
<td>0.32</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>high</td>
<td>7,000</td>
<td>1</td>
<td>0.32</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>high</td>
<td>8,000</td>
<td>1</td>
<td>0.32</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\)In millions, \(^b\)Total economic damage in seven years, \(^c\)CPU seconds, \(^d\)\(|(\text{best integer solution–best lower bound})|/\text{best lower bound}\). The columns from left to right represent the initial invasion type, budget allocation, big-M value utilized during the computational experiment, expected damage (corresponding objective function value), solution time, and percentage optimality gap, respectively.

A few key observations can be made based on the results presented in Table 4.2. According to equations (4.19) and (4.20), the M value should be greater than or equal to the maximum transition population level for each year of the seven-year planning horizon. However, since the value changes depending on budget allocation and treatment decision, the M value is expected to decrease as a result of more treatment applications among the infested...
cells as more budget is allocated for herbicide treatment. Therefore, our results imply that determining the big-M value is highly dependent on the associated variables as well as specific conditions and constraints impacting the value of those variables. This insight is applicable to solving other problems that use the big-M in a linearization procedure.

Although there is no universal method to select an appropriate M value [16, 17, 44], M values should be chosen carefully to determine the proper bounds. When linearizing the product of a binary variable and a bounded continuous variable, the M value provides an upper bound to the continuous variable. The M value can be assigned a very large value; however, in this case, the feasible region unnecessarily becomes larger, thus leading to very weak relaxation bounds in the branch and bound (B&B) search tree. The loose relaxation bounds make it more difficult for the solver to prune nodes in the B&B search tree, thus leading to an exhaustive search for examining more nodes, which slows the solution process. For example, M values in equations (4.17) and (4.19) can be set to values that are derived from the natural growth of the invasion for each $k, i, j, \text{ and } t$. However, because the M values decrease as a result of treatment applications, the fixed values obtained by the invasion growth under no treatment provide unnecessarily large lower bounds for M values, which would further weaken the B&B algorithm. In contrast, if M is too small, then inequalities may cut off feasible regions by imposing lower bounds larger than they should be.

In order to find a proper M value for each scenario and prevent relevant problems, we run each scenario with different M values. We start with big values of M and reduce the value gradually at each run until the same results are reached within a one-hour time limit for two different M values in two consecutive runs. For example, in the medium-invasion scenario with
a $700 budget allocation, we start utilizing an $M$ value of 5,000,000 and reduce it by 250,000 in each run. We reach the same result at 2,000,000 and 2,250,000 within a one-hour time limit; thus, the $M$ value is chosen as 2,000,000 for the corresponding scenario.

Another important result of this experiment is that the solution time demonstrates first an increasing trend and then decreasing trend as the budget allocation increases for all scenarios. This is because when the budget is too low, the number of possible treatment locations is low as well. On the other hand, when the budget is increased to a medium level, the number of possible treatment locations increases, thus making it more difficult for the solver to decide on the optimal treatment locations among possible alternatives. Furthermore, if the budget is ample, then the solver simply applies the treatment budget to most of the invaded locations.

4.4.2. Verification of Model: Comparison of MIP, MINLP, and NLP Models

In this section, we compare the MIP model (presented in sections 4.2.1 and 4.2.3) with its MINLP equivalent (presented in section 4.2.1) as well as the NLP model, where binary treatment decisions of MINLP are relaxed to continuous decision variables. Both MINLP and MIP models are solved as full optimization models in order to compare their solution performances with each other. The NLP model is solved using a rolling horizon approach, as in the work of Büyüktahtakı̇n et al. [14], and its solution performance is compared with the solution of the proposed MIP model. The MINLP model is solved by using the KNITRO mixed-integer nonlinearly constrained optimization algorithm in AMPL, and the MIP model is solved by using CPLEX 12.6. Furthermore, the NLP model is solved by using the CONOPT nonlinear global optimization algorithm in AMPL [45].
We first compare the MIP model with its equivalent MINLP model. Due to the complexity of the nonlinear invasive species control problem, the MINLP model can only be solved for a 3-by-3 landscape with a two-cell initial invasion and a $50 budget allocation over a four-year period. For this scenario, equal damage levels ($0.039) are reached by solving both MINLP and MIP models in less than two CPU seconds with no optimality gap. Results indicate that although both models provide the same result for the particular small-invasion scenario, the MINLP model can only be solved for small problems with the current solver technology, whereas the MIP model can provide results for a 10-by-10 landscape over a seven-year time period.

Since the MINLP problem is a highly complex model and can only be solved for a small-size problem, in order to facilitate the solution capability, the MINLP is converted into an NLP model where binary treatment decisions are relaxed to continuous variables. Note that because the NLP problem is a relaxation of the MINLP model, it provides a lower bound on the optimal objective value of the MINLP problem. However, because of the complexity of the NLP model, it cannot be solved as a full dynamic model for a 3-by-3 gridded landscape and two time periods using state-of-the-art solvers [14]. Instead, researchers exploit a rolling horizon algorithm in order to solve the NLP model on a 10-by-10 landscape.

Table 4.3 presents results derived from the NLP model and the MIP model with the proposed linearization method (Method 1) and piecewise linearization techniques (Methods 2–5). Computational experiments show that the MIP model with Method 1 results in less damage in each scenario compared to the NLP model. This result is expected for a number of reasons. First, the population growth of sericea slows down in some years, followed by an exponential
growth in the following year [14]. Dividing the budget into equal amounts for herbicide treatment over the planning period provides suboptimal solutions because although a given yearly budget can control the invasion during a population growth slow-down, it will fall shorter than is necessary during a population growth burst. Therefore, the MIP model allocates budget efficiently throughout the seven-year period by taking into consideration population growth variability, providing improved results compared to applying treatment with an equal budget amount each year in the NLP model. Therefore, results of the MIP model always supersede NLP results in terms of solution quality for the instances presented in Table 4.3.

Finally, Method 1 is compared with Methods 2–5 to verify the validity of the proposed linearization method. The same big-M values are utilized for all methods to compare results. Table 4.3 shows that the objective values derived with Method 1 are superior to the results of Methods 2–5, in the majority of cases. The only two exceptions are the second and fourth instances, where Method 4 provides a slightly better objective value. In terms of the average solution time, Method 1 reaches optimality within 94.3 CPU seconds, Methods 2–5 provide results within 10.2 seconds on average, and Method 2 provides the best solution time with 3.4 seconds. The reason for this is that Method 1 provides an exact solution method, whereas the other methods are approximated because of their computational complexity. Results imply that although approximation of quadratic terms improves the solution time significantly, it may lead to sub-optimality. Furthermore, because Method 1 does not require an additional approximation as in the other linearization methods, we can claim that it is also the most practical application. We mainly present the results for a low-invasion scenario in Table 4.3.
because using a fixed big-M value for all instances leads to ill-conditioned results for medium- and high-invasion cases.

**TABLE 4.3. TOTAL DAMAGE IN SEVEN-YEAR PERIOD BASED ON DIFFERENT BUDGET AND LINEARIZATION METHODS FOR LOW INVASION**

<table>
<thead>
<tr>
<th>Budget ($)</th>
<th>NLP heuristic method</th>
<th>MIP model using different linearization techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Method 1</td>
</tr>
<tr>
<td>0</td>
<td>94.78</td>
<td>94.47</td>
</tr>
<tr>
<td>25</td>
<td>78.23</td>
<td>47.34</td>
</tr>
<tr>
<td>50</td>
<td>14.1</td>
<td>2.42</td>
</tr>
<tr>
<td>75</td>
<td>2.98</td>
<td>0.26</td>
</tr>
<tr>
<td>100</td>
<td>0.57</td>
<td>0.02</td>
</tr>
<tr>
<td>125</td>
<td>0.32</td>
<td>0.01</td>
</tr>
<tr>
<td>150</td>
<td>0.18</td>
<td>0.01</td>
</tr>
<tr>
<td>175</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>200</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**4.4.3. Treatment Decisions over Changing Budget Allocations and Initial Invasion Scenarios**

In this experiment, we allocate a specific budget throughout the seven-year period for each initial invasion scenario in order to observe the impact of yearly budget allocations on the average yearly damage. Results show that one specific treatment strategy is not optimal for all cases. Instead, each scenario requires different budget allocations and treatment locations in order to minimize total damage by the end of the seventh year.

Results shown in Figure 4.3 reveal that sericea damage levels are inversely proportional to budget allocations. Without any intervention, sericea demonstrates an exponentially increasing population level throughout the planning horizon for each invasion scenario (Figures 4.3(a), 4.4(a), and 4.5(a)). On the other hand, when the budget is allotted for treatment, we observe a significant reduction in damage level every year compared to damage levels under no
intervention. Furthermore, damage levels can increase or decrease based on the sufficiency of the budget level allocated in the corresponding year.

Figure 4.3. Budget used and yearly damage with and without treatment each year for low initial invasion with respect to total allocated budget: (a) $25, (b) $100, (c) $200, and (d) $400. Note differences in y-axis scale.

Although treatment strategies vary among different invasion scenarios, budget allocations demonstrate similar trends for low- and medium-invasion scenarios (Figures 4.3 and 4.4). Under these scenarios, results show that it is preferable to apply treatment in the first two years of the invasion in order to prevent the invasion from spreading to neighbor cells. As the budget allocation increases, the fourth year receives about the same or greater budget allocation compared to the first two years (Figures 4.3(b) and 4.4(c)) because two-year-old stems become three-year-old stems in the third year and produce copious seeds. Therefore, there will be significant seed dispersal to surrounding cells in year three, which will result in
excessive plant generation in the fourth year. Therefore, the MIP model foresees population bursts in the first and fourth periods, and allocates more resources to these periods.

As the invasion intensifies again in the following years, the outstanding budget is distributed over the remaining years to reduce the overall damage by deploying resources to those cells that lead the highest forecasted economic damage. Furthermore, it is observed that an almost equal budget is assigned to each year as the total budget reaches an ample level (Figures 4.3(d) and 4.4(d)). Under the ample-budget case, initially invaded cells receive treatment in all periods. Due to the 95% treatment efficiency rate, a very small amount of seed dispersal takes place from the initially invaded cells to the surrounding cells after each year’s treatment, which in turn causes a negligible amount of plant generation and corresponding damage in the surrounding cells over the seven-year planning horizon (Figures 4.3(d) and 4.4(d)). Therefore, while the initially invaded cells require repeated treatment due to the accumulated seed bank in these locations, the surrounding cells are not worth treating most of the time because of the insignificant amount of plant generation over the seven-year period. As a result, in the case of ample budget, an equal budget allocation is assigned each year to treat the initially invaded cells.
Figure 4.4. Budget used and yearly damage with and without treatment each year for medium initial invasion with respect to total allocated budget: (a) $700, (b) $1,000, (c) $1,400, and (d) $2,100. Note differences in y-axis scale.

In addition, it is efficient to postpone treatment to later years than applying treatment in the first year in the case of scarce resources for a high-invasion scenario (Figure 4.5(a)). Since the invasion is already spread in the high-invasion case, the priority here becomes to reduce the invasion abundance. Therefore, it is more cost efficient to postpone the treatment and allot more budget to the following years in order to apply treatment to more plants when resources are scarce. Similar to low and medium invasions, under high initial invasion, it is preferable to allocate budget equally each year when the budget is ample (Figure 4.5(d)).
Figure 4.5. Budget used and yearly damage with and without treatment each year for high initial invasion with respect to total allocated budget: (a) $1,000, (b) $2,000, (c) 3,000, and (d) $4,000. Note differences in y-axis scale.

Finally, identifying the optimal treatment location along with the budget allocation is also vital for effectively controlling invasions. Figure 4.6 demonstrates the treatment locations for low ($100), medium ($700), and high ($2,000) initial invasion (total budget) scenarios. Results support the idea that initially, invaded cells are treated primarily as long as available budget allows treatment during the seven-year period, in order to reduce the invasion abundance in each cell. Due to the limited budget, herbicide treatments are not observed in cells that surround the initially invaded cells in these three scenarios. Furthermore, among initially invaded cells, those that are close to the center are mostly treated because the MIP model predicts that central cells may lead to more economic damage in the long term. In summary, it can be inferred that treatment efforts mostly focus on initially invaded cells.
a. Initial population for low-invasion case (t=0); treated cells and sericea dispersal over seven years with $100 total allocated budget

b. Initial population for medium-invasion case (t = 0); treated cells and sericea dispersal over seven years with $700 total allocated budget

c. Initial population for high-invasion case (t = 0); treated cells and sericea dispersal over seven years with $2,000 total allocated budget

Figure 4.6. Treatment locations and sericea dispersal after treatment
4.5. Discussion and Conclusions

In this paper, we propose a unique linear ISM model to determine efficient treatment strategies for controlling invasive species, with an application to sericea invasion. We also compare five different linearization approaches to solve this realistic MIP model. Our elegant modeling and linearization techniques improve the computational solvability of the problem far beyond its existing boundaries. In that regard, our proposed model has the potential to constitute the basis for future decision-support tools in invasive species management.

Results demonstrate the strength of the proposed dynamic MIP model against the NLP model solved with a rolling horizon heuristic, as in the work of Büyüktahtakı̇n et al. [14]. To demonstrate the strength of the MIP model, we present comparisons of forecasted economic damage between NLP and MIP models for different initial invasions and budget scenarios. The MIP model provides better results than the NLP model in all scenarios in terms of solution quality. While there are effective approaches for quadratic problems, such as sequential quadratic programming or Newton’s methods, these approaches do not guarantee an optimal solution for a non-convex optimization problem similar to ours. On the other hand, the proposed linearization technique provides an MIP model, which is equivalent to the MINLP model, thus providing an exact optimal solution for the instances examined. The increased size of the MIP problem due to linearization, coupled with the non-stationary value of the big-M value, requires considerable computational time to reach the global optimum. Therefore, the solution time is reduced by choosing appropriate big-M values to shrink the size of the B&B search tree.
In addition, solution time as well as optimal solution can be affected by the integrality tolerance of the solver. Integrality tolerance indicates the closeness of a binary variable to an integer value before the solver accepts it as satisfying the integrality condition. Assume that the integrality tolerance setting is slightly greater than $\varepsilon > 0$, which means that any discrete variable that violates integrality by more than $\varepsilon$ will be further branched upon for resolution. As an example, assume that $Q_{ijt}^{k*} > 0$ is the optimal value of $Q_{ijt}^i$ given in equation (4.19). If $M \cdot \varepsilon \geq Q_{ijt}^{k*}$, then the solver will accept $1 - y_{ijt}^k = \varepsilon$ as part of an optimal solution. Therefore $Q_{ijt}^{k*}$ will be impacted by changes in the $\varepsilon$ value. For example, during the solution process, we reduce the default integrality tolerance value from 1e-05 to 1e-07 in order to obtain precise binary values. However, reducing the integrality tolerance, $\varepsilon$, without increasing the $M$ value forces the $M \cdot \varepsilon$ value to have an invalid upper bound on $Q_{ijt}^{k*}$, which would cut off the optimal solution. Because the $Q_{ijt}^{k*}$ value will be impacted by variations in integrality tolerance, the $M$ value should be adjusted in the model accordingly.

Furthermore, results of the MIP model with different linearization methods demonstrate the strength of the proposed linearization technique. While Methods 3 and 4 include fewer binary variables compared to other methods, the additional continuous variables in Methods 2-4 and the additional big-M value in Method 5 further increase the size of the MIP problem. Moreover, the dynamic nature of invasive species, particularly the carrying capacity, results in non-stationary breakpoints, which lead to quadratic terms in the piecewise functions. These quadratic terms are approximated, thus reducing the solution time. However, the proposed linearization method does not require the approximation of quadratic terms or an
additional big-$M$ value as in other linearization methods, thus providing a better objective function value in the majority of cases.

We demonstrate the applicability of the model on a case study of sericea (Lespedeza cuneata) infestation. While MINLP and NLP approaches could not solve the spatially explicit model for a 3-by-3 grid landscape with two time periods, we are able to solve our model on a heterogeneous 10-by-10 grid landscape with a seven-year time period. It is common to use time horizons between five and ten years in bio-conservation models [12], because most policymakers and municipals make their budget allocations for controlling invasive species based on a five-year projection, which aims to prevent further uncertainties occurring due to the longer time horizons [12]. Horie et al. [46] use 90 hexagonal 1.1 km$^2$ grid cells to model the oak wilt invasion, while Epanchin-Niell and Wilen [32] consider a 7-by-14 landscape to control a hypothetical invasion over a three-year period. Therefore, using a 10-by-10 or smaller gridded landscape is also a relevant application size for invasive species management.

Our approach and sensitivity analyses help to provide guidance about how to manage sericea invasion with different budget-allocation levels. Computational experiments provide expected damage levels and corresponding budget allocations for control strategies where both the budget falls short of what is needed and the budget is sufficient enough to control the invasion. Our results suggest that sericea damage is expected from the sericea invasion frequency and abundance impact expected, hence the reason for optimal treatment strategies over the seven-year time horizon. Results demonstrate that depending on the initial sericea frequency and budget, the decision maker needs to adopt different management strategies in order to minimize related damages.
In addition, if the initial invasion frequency is intensified by the time the decision maker recognizes the invasion, it is preferable to postpone treatment in order to efficiently use the limited resources. Therefore, when the budget falls short of the required amount in a high initial invasion case, deferring treatment in the initial period will enable farmers to eradicate more plants in the preceding year, despite the cost of damage to current-year forage. On the other hand, decision makers can allot an equal budget each year as the budget shifts to an ample level.

We also observe that optimal treatment efforts focus on the initially invaded cells for short planning horizons under limited or medium budget levels, in an effort to prevent substantial plant regeneration from the seed bank. Furthermore, among initially invaded cells, those that are close to the center are given treatment priority because the MIP model forecasts the fact that central cells may lead to more long-term economic damage. Results also indicate that one specific treatment strategy does not provide the best solution for all invasion scenarios. This shows the need and value of using state-of-the-art decision-making tools in order to employ efficient treatment approaches under different cases.

In this paper, we contribute to the literature by solving an invasive species control problem as a full optimization model. The biological dynamics of sericea invasion is incorporated into a spatially explicit MIP model in order to find optimum treatment strategies for various budget allocations. Solving the complex sericea invasion problem as a full dynamic model enables decision makers to make a series of interrelated decisions. This means that the model makes simultaneous treatment decisions by taking into account future potential damage. In summary, the MIP model allows land managers to make dynamic decisions about
where and when to apply herbicide treatments by deciding the budget allocation for sericea treatment over a seven-year time period.

Given input data and population growth specifics, our spatially explicit model can be applied to any species that follow a stage-dependent life cycle, including fish, insects, plants, or mammals, where the age, length, or weight of the species impacts its fecundity or mortality [46]. For instance, equations (4.1–4.3), which represent seed dispersal and seedbank generation, can be adapted to formulate offspring generation, dispersal mechanism, and dormancy. Furthermore, constraints that define age-structured dynamics, that is equations (4.4–4.6), can be adapted to provide state transitions among species' life stages. Finally, equations (4.7–4.8) can be used in the model to limit the population size of species within its natural boundaries.

Additional work will focus on extending the proposed MIP model to a full stochastic model by considering the uncertainty in growth and herbicide-effectiveness levels. Furthermore, the direction and intensity of wind could have a very significant impact on the seed dispersal rate and distance. Instead of using a radial dispersal mechanism, the impact of wind speed and direction could be modeled stochastically. In addition, MIP techniques help to solve the problem as a full dynamic model for most scenarios within 3,600 CPU seconds and enable an optimal solution to the ISM problem on a 10-by-10 gridded landscape over a seven-year period. Additional parameters and heuristics within CPLEX could be used to reduce the solution time at the expense of potentially reducing the solution quality. As a result, further research will focus on exact solution algorithms such as decomposition and cutting-plane
techniques [49] in order to increase the solvability of large-scale ISM models and reduce the solution time for instances that involve higher spatial and temporal dimensions.

4.6. References


CHAPTER 5

NEW MULTI-STAGE STOCHASTIC PROGRAMMING MODEL AND CUTTING PLANES FOR THE OPTIMAL SURVEILLANCE AND CONTROL OF EMERALD ASH BORER IN CITIES

5.1. Introduction

Invasive species cause significant environmental and economic damage. They disturb the ecosystem by suppressing and crowding out the native species [1], which eventually destroys biodiversity and shelter for wildlife. Invasive species also pose a fire threat by changing the fire regime [2] and they also alleviate the benefits from recreational areas [3]. Pimentel et al. [4] report that invasive species cost the United States around $120 billion every year in economic damages due to their continuous spatial spread. Economic damages along with the limited resources for controlling invaders highlight the vital importance for implementing optimal management strategies by government agencies.

Among more than five thousand invasive species in the U.S., the EAB is one of the most economically and environmentally damaging invasive species. The EAB is a wood-boring insect that is native to Asia. Since its discovery near Detroit in 2002, the EAB has spread to 25 states and two Canadian provinces and threatens to kill most, if not all, ash trees native to North America. In cities and communities, the EAB has already cost more than $60 billion and is projected to cost homeowners and local governments billions more in treatment or removal and replacement of landscape ash trees over the next decade [5]. In response, cities and communities are developing EAB management plans, including the application of systemic insecticides to kill EAB adults or larvae and the pre-emptive removal of infested trees before
larvae can complete development [6]. In addition, management plans may include surveillance to discover the location of infestations in early stages.

Surveillance is particularly difficult for invasive species like the EAB because it can infest ash trees for two or more years before damage is evident. Although there are several models and a significant amount of information available to evaluate the costs and benefits of invasive species control [7-12], studies that are concentrated on surveillance [13-16] do not differentiate surveillance and treatment decisions. In these studies, a surveillance decision implies treatment of invasion, i.e., surveillance is followed by treatment to immediately eradicate the invasion. However, surveillance and treatment decisions should be separated because treatment or removal immediately after surveillance might not be feasible due to a limited budget. Evaluating the costs and benefits of surveillance in combination with treatment and removal and taking a rapid action against the EAB invasion is a national priority. Consequently, cities and communities throughout the U.S. are developing strategies and programs to slow down EAB infestation.

This study addresses the problem of building a spatial-dynamic model to help communities develop cost-effective strategies for surveillance and control of invasive species. Our objective is to develop a novel MSS-MIP model that finds optimal strategies for surveillance of the ash population, followed by treatment or removal of trees based on their infestation level, while maximizing the net benefits of the ash trees on a given landscape by considering all possible scenarios. We provide the first MSS-MIP model to the invasive species literature by integrating invasive species population growth, dispersal, and impact on homeowners and local governments [6]. While the model has the potential to be generalized for various invasive
species, we focus on the population dynamics of the EAB by accounting for its dispersal and growth based on a stage-structured infestation formulation. Different than the existing literature (e.g., [5, 10, 15, 17]), trees are clustered into classes based on their infestation levels in which each infestation level poses a different extent of threat to the trees that are not infested. The MIP component of the model includes binary linearization variables for finding an infested population in each class along with surveillance, treatment, or removal of infested trees over a number of years under a limited budget. Unless an infested tree is treated, it moves to the upper infestation class in the following period until reaching a maximum infestation level, which requires removal.

The multi-stage stochastic program (MSSP) allows us to consider all possible scenarios for surveillance, treatment, and removal decisions over a planning horizon by considering uncertainty in the invasion growth. This information structure leads to a scenario tree where each stage corresponds to a time period while uncertain population growth is modelled using scenarios with certain probabilities. Furthermore, the surveillance decisions are integrated into uncertain population growth scenarios, forming a hybrid scenario tree with two types of decisions made at each node: surveillance decision for the landscape, and treatment or removal decision for the infested trees. We assume that at the beginning of each stage, the decision maker has a belief about the number of infested trees, which is only realized after surveillance is applied. Thus, treatment and removal decisions (i.e., number of treated and removed trees) can be adjusted at each stage and for each scenario, based on the true amount of infestation that is realized after surveillance. Therefore, the goal of the MSS-MIP model is to determine a plan that can be implemented for the treatment and removal of infested trees by taking into
account possible realization scenarios, and that can maximize the net benefits of ash trees over the planning horizon.

It should be noted that MSSP models are difficult to solve, and there are no general solution algorithms for this kind of problems. Most solution approaches are problem specific and usually exploit the scenario structures by decomposing the deterministic equivalent of the stochastic program. For example, Norkin et al. [18] describe a stochastic branch-and-bound (B&B) method where estimations of upper and lower values of the objective values are calculated using Monte Carlo simulation to converge the optimal solution. Furthermore, Kleywegt et al. [19] generate set of samples with randomly generated probability distributions, and sample average functions are used to estimate the expected value functions. The stochastic problem then is solved using deterministic algorithm using a stopping criteria. In another study, Shapiro [20] uses a conditional sampling procedure to guarantee that a consistent lower bound for the MSSP problem is reached by the sample average approximation. On the other hand, a dual-scenario decomposition is described in the work of Carøe and Schultz [21] where Lagrangian relaxation is applied to nonanticipativity constraints. The resulting dual problem reduces the problem into single-scenario size MIP problems which provides a lower bound for the primal problem. The upper bounds are obtained by utilizing heuristic methods.

In order to solve the proposed complex MSS-MIP model, we develop novel cutting planes. The solution algorithm has two parts: the first part represents a preprocess problem, and the second part represents the original model. In the first part, we receive the values of the complicating binary variables by solving the preprocess model under the worst-case scenario. In the second part, values of some binary variables are fed into the original MSS-MIP model by
following specific steps. Furthermore, additional cutting planes for complicating variables are defined to further facilitate the solution process. Our cuts significantly improve the solvability of the original MSS-MIP model.

The proposed model fills the research gap in the invasive species literature by first providing a novel MSS-MIP model for the first time. Second, we develop a novel solution methodology to solve the proposed model, which can be applied to other similar MSSP models. Third, we fill the research gap in the literature by integrating surveillance, treatment, and removal as separate decisions into an optimization model. In particular, surveillance is used as a method to track the true amount of infested trees at each infestation level on a given landscape followed by treatment and removal based on available resources. Finally, the proposed methodology will allow decision makers to examine all possible invasion scenarios based on distinct surveillance outcomes.

Our approach leads to a number of important results:

- First, our novel solution algorithm improves the solution time by an average of up to 44 times for 3-period instances, compared to solving the original MSS-MIP model. The solution time improvement becomes apparent as the size of the problem increases.
- Second, our computational experiments on the comparison of the original MSS-MIP model and the solution algorithm show that the proposed algorithm is effective to solve the proposed model. In particular, while the original model cannot be solved for a 6-by-6 gridded landscape in a five-year time period, the algorithm enables us to solve the model for a 10-by-10 gridded landscape in a five-year time period with an ample budget allocation.
• Third, the proposed model provides all possible surveillance, treatment, and removal decision approaches for decision makers under all possible scenarios over a planning horizon. Therefore, this model helps decision makers to optimally allocate resources at each stage by considering the current and forecasted damages of invasion.

• Fourth, model provides rule of thumb treatment and removal strategies for EAB infestation by targeting specific infestation classes in a gridded landscape over the planning horizon.

The structure of the remaining paper is as follows: In Section 5.2, we define the problem, and present the MSS-MIP model and cutting plane algorithms. In Section 5.3, we present the model application and data on the case of EAB invasion. Section 5.4 follows with computational results that illustrate the efficiency of the solution algorithm and provide related management implications. Section 5 presents discussion and offers directions for future research.

5.2. Model

5.2.1. Problem Definition

It is unlikely that the exact number of infested trees on a large landscape can be known in advance without controlling each tree, one by one, for infestation. Therefore, forest managers should form beliefs about the proportion of infestation on a spatially explicit landscape by inspecting a small sample of trees at each site, which in turn will provide a belief about or expectation of the number of infested trees in that location [17]. Although the sample provides some information about the number of infestations, it may not reflect the true
amount of infested trees. Therefore, managers should apply surveillance in order to determine the exact number of infested trees on a given landscape.

In order to better represent the infestation structure of the EAB, trees are divided into two categories: susceptible ash trees that are prone to infestation and infested trees. Infested trees are further clustered into $n$ classes based on their infestation levels. Each tree in an infestation level transitions to the upper infestation class in the following period until reaching a maximum infestation level. Furthermore, it is assumed that each infestation level poses a different extent of threat to the trees that are not infested.

In this model, we assume that managers have beliefs about the number of susceptible trees and also the number of trees at each infestation class at the beginning of the planning period. However, surveillance is needed to assist managers in realizing (determining) the true number of infested trees, which may be different than the belief of the infestation.

Therefore, the main structure of the problem is as follows: If surveillance is not applied to a class of infested trees, either in the current time period or before, then treatment will not be allowed, and consequently, transition among infestation classes occurs based on the initial belief in the following period. In this case, newly infested trees emerge in the following period based on the belief as well as the growth rate. However, if surveillance is applied in the current time period or before, then the true level of infestation is realized and treatment can be applied in the current period based on the available budget allocation.

Nevertheless, it can be the case where surveillance may not be applied in some years after the latest surveillance. In such a scenario, surveillance will not only target susceptible trees but should also be applied to all infestation levels except for the already realized ones.
This problem is represented as a network, which is demonstrated by an example shown in Figure 5.1, where surveillance is only applied in periods 2 and 5. In this example, infested trees are clustered into four levels. We assume that trees at infestation levels one and two can be freed from infestation with treatment. On the other hand, trees at infestation level three are irreversibly infested and, thus, should be removed from the population to prevent further damage. Finally, trees at infestation level four are assumed to be dead and can no longer pose threat to the population. Therefore, trees at level four should also be removed to avoid the risk of their falling. Finally, it is assumed that a new ash tree is planted for each removed tree, in order to sustain the environment. However, a newly planted tree is not considered a susceptible tree because EABs only invade mature trees with thick trunks.

We assume that the belief of infestation is set in period 1. However, since surveillance is not applied in period 1 in the example shown in Figure 5.1, the belief of infestation is not realized, and infestation classes transition to the upper level of infestation classes in period 2. Furthermore, some trees that are susceptible in period 1 become infested in period 2, due to the infestation spread from infected trees. Once surveillance is applied in period 2, infestation at each class is realized, and treatment may be applied to infested trees based on the available budget. In period 3, surveillance is not applied, and consequently, the number of newly infested trees (level one infestation) remains as a belief. Hence, treatment will not be applied to the newly infested trees in period 3 because they are not realized. On the other hand, treatment can be applied to level two, and removal can be applied to levels three and four infestations in period 3 because the number of trees in these classes has already been realized in the previous period when they were in levels one, two, and three infestation classes, respectively. Similarly,
removal can be applied to only level three and four infestations in the fourth year because corresponding infestation levels were realized in the second year. Note that unremoved trees at levels three and four transition to level four and remain in that infestation category, unless removal is applied. Finally, since surveillance is applied in period 5, the number of infested trees at all infestation classes are realized, and consequently, treatment or removal can be applied to all trees based on the available budget.

Figure 5.1. Representation of EAB model with five time periods and four infestation levels, where S represents susceptible trees, while 1, 2, 3, and 4 correspond to trees at infestation levels one, two, three, and four, respectively. Green nodes represent the real amount of susceptible trees; red nodes represent the real amount of infested trees in the corresponding infestation level for which treatment or removal can be applied; black nodes represent the level four infested trees, which are dead and do not impact susceptible trees; uncolored nodes represent the infestation classes in which the actual number of infested trees are not realized; black dashed lines represent the transition between infestation levels; black solid lines represent the removed trees; and green dotted lines represent treated trees that become susceptible two periods later.
In order to reduce the overall infestation and increase the monetary net benefits from susceptible ash trees, the model enforces surveillance and treatment decisions based on the available budget to prevent further spread and growth of infestation to susceptible trees. Furthermore, we assume that dead trees pose a threat to the environment and society. Therefore, the objective of this study is set to maximize the net benefit from susceptible ash trees while penalizing dead trees on a given landscape over the planning horizon. The details of the model are explained in the following section.

5.2.2. MSS-MIP Model

5.2.2.1. Scenario Tree

A scenario tree is a viable way of discretizing the dynamic stochastic data over time in a problem. We demonstrate our scenario tree in Figure 5.2. Black and white nodes represent surveillance and no-surveillance events (decisions) on a given stage, respectively. Outgoing arcs from circle nodes represent possible realizations of the beliefs. Each node $D$ represents treatment or removal decisions taken on the corresponding stage. Each black node yields three possible realizations, and each white node yields a belief of infestation. Therefore, we assume four possible outcomes (three realizations and one belief of infestation) from each decision node, based on the discretized surveillance decisions. Consequently, $4^t$ scenarios are generated by the end of stage $t$. 
Figure 5.2. Representation of decision tree in multi-stage form, where $U_i \{a\%, b\%\}$ represents uniform percentage change of infestation level (realizations) at each site at each stage.

5.2.2.2. Notations

Sets and Indices:

$\mathcal{I}$ Set of all sites ($i \in \mathcal{I}$ & $\mathcal{I} = \{1, \ldots, n\}$)

$\mathcal{K}$ Set of infestation levels ($k \in \mathcal{K}$ & $\mathcal{K} = \{1, 2, \ldots, n-1, n\}$)

$\mathcal{T}$ Set of time periods ($t \in \mathcal{T}$ & $\mathcal{T} = \{1, \ldots, T\}$)

$\Theta_i$ Set of neighboring sites of site $i$ ($j \in \Theta_i$)

$\Xi$ Set of scenarios in scenario tree ($s \in \Xi$)

Parameters:

$\pi_s$ Probability for scenario $s$
\( c_1 \)  Cost of surveillance
\( c_2 \)  Cost of treatment
\( c_3 \)  Cost of removal
\( \alpha \)  Monetary value of each susceptible tree
\( \vartheta_k \)  Penalty value assigned to each infested tree at infestation level \( k = n \) and \( n - 1 \)
\( r_k \)  Impact rate of each infested tree at infestation level \( k \) within a site \( i \), i.e., number of new infestations per infested tree at level \( k \)
\( \delta_t \)  Discount factor at time \( t \)
\( \Psi_s \)  Budget for scenario \( s \)
\( \theta_k \)  Infestation impact of \( k^{th} \)-level infested trees in neighboring site \( j \)
\( p_{j\rightarrow i} \)  Probability of infestation spread from site \( j \) to \( i \)
\( \beta_{\text{iks}} \)  Percentage change (realization) in belief of infestation for site \( i \), infestation level \( k \), at time \( t \), for scenario \( s \)
\( \overline{N}_i \)  Initial number of tree population at site \( i \)
\( \overline{T}_{ik} \)  Initial number of infested tree population at each infestation level \( k \), at site \( i \)

**Discretized Binary Parameters:**
\[
x'_s = \begin{cases} 
1 & \text{if surveillance is applied at time } t, \text{ for scenario } s \\
0 & \text{o.w} 
\end{cases}
\]

**Decision Variables:**
\( N'_{is} \)  Total number of trees at site \( i \), at time \( t \), for scenario \( s \)
\( S_{is}^t \) Number of susceptible trees at site \( i \), at time \( t \), for scenario \( s \)

\( \tilde{I}_{iks}^t \) Believed number of infested trees at site \( i \), at time \( t \), at infestation level \( k \), for scenario \( s \) before surveillance

\( \hat{I}_{iks}^t \) Transition number of infested trees at site \( i \), at time \( t \), at infestation level \( k \), for scenario \( s \) after surveillance without considering total tree population

\( I_{iks}^t \) Real number of infested trees at site \( i \), at time \( t \), at infestation level \( k \), for scenario \( s \) after surveillance with considering total tree population

\( V_{iks}^t \) Number of treated trees at site \( i \), at time \( t \), at infestation level \( k \), for scenario \( s \)

\( R_{iks}^t \) Number of removed trees at site \( i \), at time \( t \), at infestation level \( k \), for scenario \( s \)

\( H_{is}^t \) Number of trees surveyed at site \( i \), at time period \( t \), for scenario \( s \)

\( Q_{iks}^t \) Number of infested trees remaining after treatment and removal at site \( i \), at time \( t \), for scenario \( s \)

**Linearization Variables:**

\[ u_{iks}^t = \begin{cases} 1 & \text{if transition population is assigned to infestation level } k \\ 0 & \text{o.w} \end{cases} \]

5.2.2.3. **Mathematical Model**

A general MSSP model is formulated using the scenario tree shown in Figure 5.2. Based on the notation presented in section 5.2.1 and 5.2.2.1, the MSS-MIP model is described as follows:

**Budget Constraint.** It is assumed that while trees in the \( k = 1, \ldots, n - 2 \) infestation classes can be saved with treatment, the high-level infestation in \( k = n - 1 \) and \( n \) is irreversible, while \( k = n - 1 \)
still poses a threat to the environment, and $k = n$ represents dead trees. Therefore, while treatment could be applied to infestation levels $k = 1,\ldots,n - 2$, highly infested and dead trees should be removed to prevent further damage based on the available budget. Therefore, the budget constraint is formulated as

$$c_1 \sum_{t \in T} \sum_{i \in S} H_{is}^t + c_2 \sum_{t \in T} \sum_{i \in S} \sum_{k = 1}^{n-2} V_{iks}^t + c_3 \sum_{t \in T} \sum_{i \in S} \sum_{k = n-1}^{n} R_{iks}^t \leq \Psi_s \quad \forall s$$

(5.1)

which ensures that surveillance, treatment, and removal decisions are restricted by the available budget for each scenario $s \in \Xi$. In the budget constraint, $H_{is}^t$ represents the number of trees that are surveyed and is formulated as

$$H_{is}^t = \left( S_{is}^t + \sum_{k = 1}^{n-1} I_{iks}^t \right) x_{is}^t \quad \forall s, i \text{ and } t = 1$$

(5.2)

$$H_{is}^t = S_{is}^{t-1} x_{is}^t + \sum_{k \in K} \left[ \prod_{i = 1}^{k} \left( 1 - x_{is}^{\max[t-1,1]} \right) \right] x_{is}^{\max[t-1,1]} I_{iks}^{\max[t-1,1]} \quad \forall s, i \text{ and } t = 2,\ldots,T$$

(5.3)

Equation (5.2) ensures that the number of surveyed trees in the initial period is either zero or equals the number of susceptible plus infected trees based on surveillance decisions given in the initial period. Similarly, equation (5.3) determines the number of surveyed trees at period $t$ based on surveillance; however, this equation takes into account all surveillance efforts performed in previous $k$ time periods. Let $T = t \leq \overline{T}$ represent the number of periods until surveillance is applied in period $t$; then the number of susceptible trees and the real number of trees in the first $k = \overline{T}$ infestation level will be unknown and remain as a belief. If surveillance is performed in period $t$ after $\overline{T} \leq n$ periods (where $n$ is the highest infestation level), then trees
in the susceptible cluster and in the first \( k = \bar{t} \) infestation level will be surveyed in time period \( t \). Note that surveillance is not necessary for infestation levels \( k = \bar{t} + 1, \ldots, n \) because the number of trees at these infestation levels was realized \( \bar{t} \) periods previously, at the latest surveillance application. On the other hand, if \( \bar{t} \geq n \) and surveillance is applied, then susceptible trees and all infested trees will be surveyed in a given landscape.

**Total Population.** The tree population that can become infested at a given site \( i \) is reduced by the removal of highly infested trees. Furthermore, trees that are treated in a given period are safe from infestation for the following two years and can become infested two years later. Therefore, the total population that is under EAB threat in site \( i \) and scenario \( s \) is formulated as

\[
N_{is}^{t+1} = N_{is}^{t} - \sum_{k=1}^{n-2} V_{iks}^{t} - \sum_{k=n-1}^{n} R_{iks}^{t} \quad \forall s, i \text{ and } t = 1
\]

\[
N_{is}^{t+1} = N_{is}^{t} - \sum_{k=1}^{n-2} V_{iks}^{t} - \sum_{k=n-1}^{n} R_{iks}^{t} + \sum_{k=1}^{n-2} V_{iks}^{t-1} \quad \forall s, i \text{ and } t = 2, \ldots, \bar{t} - 1
\]

**Transition and Real Infestation Level.** Surveillance allows realization of the true number of trees at each infestation level. The believed (expected) number of infested trees for a particular infestation level \( k \), \( \tilde{I}_{iks}^{t} \), may change (increase, remain the same, or decrease) after the surveillance, or it may not be updated if surveillance is not applied. Therefore, the realization of infestation for each level \( k \) is formulated as

\[
\tilde{I}_{iks}^{t+1} = \tilde{I}_{iks}^{t} \cdot (1 + x_{s}^{t} \beta_{iks}^{t}) \quad \forall s, t, i, k
\]

where \( \tilde{I}_{iks}^{t} \) represents the transition number of infested trees at each infestation level \( k \) before considering the total tree population in a given site. We use a transition population level because the real number of infested trees at level \( k \) also cannot exceed the number of
remaining trees after \( k + 1, \ldots, n \) levels of infestation populate a region. As a result, if trees at level \( k + 1, \ldots, n \) infest the entire population, then (lower) infestation levels \( k, k - 1, \ldots, 1 \) do not exist at the given site. Therefore, the real number of infested trees in infestation level \( k \) after considering the total tree population is formulated as

\[
I'_{iks} = \min\left(1, \sum_{d = \min(k + 1, n)}^{n} I'_{ids} - \sum_{d = \min(k + 1, n)}^{n} I''_{iks} \right) \quad \forall s, t, i, k
\]  

(5.7)

where the real number of trees in an infestation level \( k \) is assigned the minimum value between the transition population level and the remaining population after \( k + 1, \ldots, n \) levels of infested trees populate a given site. Note that equation (5.7) is nonlinear, and can be replaced with the following linear inequalities:

\[
I'_{iks} \leq I''_{iks}
\]  

(5.8)

\[
I'_t ^{iks} \leq N'_t ^{iks} - \sum_{d = \min(k + 1, n)}^{n} I'^{t}_{ids}
\]  

(5.9)

\[
I''_{iks} - I'_t ^{iks} \leq N_i - (1 - u'_{iks})
\]  

(5.10)

\[
\left(N'_t ^{iks} - \sum_{d = \min(k + 1, n)}^{n} I'^{t}_{ids}\right) - I'_t ^{iks} \leq N_i \cdot u'_{iks}
\]  

(5.11)

where \( u'_{iks} \) is a binary variable that is defined to determine the \( k \)-th-level infested-tree population, and \( N_i \) is the initial total tree population in a given site. Note that while equations (5.8) and (5.9) provide an upper bound on the real number of infestation, equations (5.10) and (5.11) provide a lower bound by activating the binary variable \( u'_{iks} \). If the transition population level is less than the remaining capacity \((I''_{iks} \leq N'_t ^{iks} - \sum_{d = \min(k + 1, n)}^{n} I'^{t}_{ids})\), then \( u'_{iks} = 1 \), or vice versa.
Therefore, the linearization equations (5.8–5.11) ensure that the equality
\[ I_{iks}^t = \min \left( N_{is}^t - \sum_{d=\min(k+1,n)}^n I_{ids}^t + \tilde{I}_{iks}^t \right) \]
given in equation (5.7) holds.

**Susceptible Population.** Furthermore, the number of susceptible trees equals the total population less the total number of infested trees. Therefore, the susceptible tree population is formulated as
\[ S_{is}^t = N_{is}^t - \sum_{k=1}^n I_{iks}^t \quad \forall s, t, i \quad (5.12) \]

**Number of Treated and Removed Trees.** Treatment or removal can be applied to infestation level \( k \) if surveillance has been applied at least one time in the last \( k \) time periods including the current time period. Furthermore, the number of treated or removed trees cannot exceed the number of trees at the same infestation level. Therefore, the number of treated or removed trees is formulated as
\[ V_{iks}^t \leq I_{iks}^t - \sum_{a=\max(t-k+1,1)}^t x_a \quad \forall s, t, i \quad k = 1,\ldots,n-2 \quad (5.13) \]
\[ R_{iks}^t \leq I_{iks}^t - \sum_{a=\max(t-k+1,1)}^t x_a \quad \forall s, t, i \quad k = n-1 \text{ and } n \quad (5.14) \]

**Believed (Expected) Number of Infested Trees.** At each time period, susceptible trees in a given site can become infested by the impact of untreated but infested trees within the site and by the spread of infestation from surrounding sites \( j \in \Theta \). Therefore, the believed number of newly infested trees \((k = 1)\) at time \( t + 1 \) is formulated as
\[ \tilde{I}_{iks}^{t+1} = \sum_{g=1}^{n} Q'_{igsv} \cdot r_{g} + \sum_{g=1}^{n} \sum_{j \in \Theta} Q'_{jgs} \cdot \theta_{g} \cdot p_{j \rightarrow i} \quad \forall s, i \ & t = 1..\overline{t} - 1 \] (5.15)

where \( Q'_{igsv} \) represents infested but untreated or unremoved trees at time \( t \) and is formulated as

\[
Q'_{igsv} = \begin{cases} 
I'_{igs} - V'_{igs} & g = 1,\ldots,n-2 \\
I'_{igs} - R'_{igs} & g = n-1 \& n \quad \forall s, t, i
\end{cases}
\]

and \( p_{j \rightarrow i} \) represents the probability of infestation spread from site \( j \) to \( i \).

Furthermore, infested but untreated trees will transition to the upper infestation level as a belief in the following period. Therefore, the believed number of infested trees for infestation level \( k = 2,\ldots,n-1 \) is formulated as

\[
\tilde{I}_{iks}^{t+1} = I_{i(k-1)s}^{t} - V_{i(k-1)s}^{t} \quad \forall s, i \ & t = 1..\overline{t} - 1 \ & k = 2,\ldots,n-2
\] (5.16)

In addition, once the infested trees reach the highest infestation level \( k = n \), they remain in the \( n^{th} \) level unless they are removed from the population. Therefore, the believed number of infested trees at level \( n \) in a given period consists of the unremoved trees of level \( n-1 \) and \( n \) in the previous period. Therefore, the believed number of infested trees for level \( n \) is formulated as

\[
\tilde{I}_{iks}^{t+1} = (I_{i(k-1)s}^{t} - R_{i(k-1)s}^{t}) + (I_{iks}^{t} - R_{iks}^{t}) \quad \forall s, i \ & t = 1..\overline{t} - 1 \ & k = n-1,n
\] (5.17)

Finally, the initial total population and initial belief of infestation levels are defined as

\[
N_{iks}^{1} = \overline{N}_{i} \quad \forall s, i
\] (5.18)

\[
\tilde{I}_{iks}^{1} = \overline{I}_{ik} \quad \forall s, i,k
\] (5.19)

\[
N_{iks}^{t}, S_{iks}^{t}, I_{iks}^{t}, \tilde{I}_{iks}^{t}, \tilde{I}_{iks}^{t}, V_{iks}^{t}, R_{iks}^{t} \geq 0 \quad u_{iks}^{t} \in \{0,1\} \quad \forall s, t, i, k
\]
**Objective Function.** The objective function of this model is to maximize the net benefits of susceptible ash trees while penalizing those trees that are subject to removal over the entire landscape and planning horizon. Therefore, the objective is formulated as

\[
\max \sum_{s \in S} \pi_s \left( \sum_{t \in T} \delta_t s_t \left( \alpha S^i_{ts} - \sum_{k=n-1}^{n} \delta_k I^i_{ik} \right) \right)
\]

(5.20)

5.2.2.4. Cutting Planes for Linearization Variables

In this section, we discuss methods of generating cutting planes for the proposed model to facilitate the solution process. Although MSSP models are some of the most difficult to solve problems, there are no general algorithms to solve these problems. Most solution approaches are problem specific and common solution method is to decompose the deterministic equivalent of the stochastic problem. However, although decomposition algorithms are the widely used methodologies, decomposing the model does not always provide better solution process. Therefore, in this study, we develop a novel cutting plane algorithm and cutting planes in the form of complicating binary variable to facilitate solution process.

5.2.2.4.1. Cutting Planes for Remaining Capacity

As mentioned in section 5.2.2.3, the real number of infested trees in infestation level \(k\) receives the minimum value between the transition population and the remaining population after higher infestation levels populate a given site. Therefore, if infestation level \(k\) can only occupy the remaining population \(u_{iks}^i = 0\), then this implies that the given site has been totally infested by infestation levels \(k, k+1, ..., n\), and no space is left for infestation levels \(1, 2, ..., k-1\). Therefore, if \(u_{iks}^i\) is equal to zero, then the lower infestation levels \(k-1, ..., 1\) are assigned zero \(I_{iks}^i = 0\), implying that \(u_{iks}^{i(k-1)}\) is also equal to zero. On the other hand, if infestation level \(k\) is
assigned the transition population \( u_{ik}^t = 1 \), then lower infestation levels \( k - 1, \ldots, 1 \) can be assigned either the transition or remaining population. Thus, in order to represent this relationship between \( u_{ik}^t \) and \( u_{ik}^{t-1} \) variables, we derive remaining capacity cutting planes (RC cuts) as follows:

\[
\begin{align*}
u_{ik}^{t-1} & \leq u_{ik}^t & \forall s, t, i, k
\end{align*}
\]

We add the RC cuts given in equation (5.21) as valid inequalities to the model. Therefore, RC cuts strengthen the problem formulation without cutting off the optimal solution.

5.2.2.4.2. Preprocess Algorithm and Transition Population Cutting Planes

Our goal is to derive additional cuts that define transition population, so called transition population cutting planes (TP cuts). In order to generate TP cuts, we solve the problem for the worst-case scenario, in which we observe high realization in each stage. Once the problem is solved for the worst-case scenario, we can determine the first time period \( t' \), in which the infestation level \( k \) reaches the maximum allowable population. Under the worst case scenario, we can assign the transition population to infestation level \( k \), \( u_{ik}^t = 1 \), for each time period where \( t < t' \). We present the routine to generate the TP cuts in the following procedure:
The algorithm given in this procedure helps us to derive TP cutting planes in the form

\[ u'_{iks} = 1 \quad \forall s, i, k & t < t' \]  

(5.22)

TP cuts in equation (5.22) ensure that infestation level \( k \) is set to its transition population for all periods less than \( t' \). In order to generate the cuts of the form \( u'_{iks} = 1 \) for any \( s, i, k, \) and \( t < t' \) in the problem, we choose the scenario that leads to the minimum economic benefit from susceptible tree population (worst-case scenario) and solve the model for this specific scenario, namely \( s_{worst} \). Note that the worst-case scenario is the one where the realization of the
infestation is high in all periods, and no treatment or removal is applied to the infested trees. Once the problem is solved under the worst-case scenario, we can determine the first period $t'$ under which the remaining population is assigned to infestation level $k$, ($u'_{iksu_{worst}} = 0$). Because $t'$ is the earliest period when the population of infestation level $k$ reaches the maximum allowable population under the worst-case scenario, the transition population can be assigned to infestation level $k$, ($u'_{ik} = 1$), for all $s, i, k$, and each time period where $t < t'$.

On the other hand, the result $u'_{iksu_{worst}} = 0$ cannot be applicable for all $s, i, k$. The reason is that once treatment is applied to infested trees, they can be added to the total population two years later, which may lead to the assignment of a transition population to any infestation level $k$ in that time period. Therefore, if $u'_{iksu} = 0$ for an infestation level $k$ at a given time period $t$, then $u'_{ik}$ remains relaxed ($u'_{ik} \in \{0,1\}$) in the model for any $s, i, k$, and $t' \geq t$.

5.3. Data Calibration and Model Application

5.3.1. Data Calibration

The impact and spread rate of trees at each infestation level are estimated using real EAB infestation data. The data used for data calibration and model validation was collected by Knight et al. [22] and Flower et al. [23] in Toledo, Ohio, from 2005 to 2011 from ten different sites. Each site is a forested area with homogeneous tree species and includes three to six plots, where each plot represents a 400 m$^2$ circular landscape. Data includes the number of individual ash trees tracked at these sites and the canopy health of the trees, rated on a visual rating scale of 1 to 5, where 1 is a tree with a healthy canopy (healthy, full foliage in the canopy), 5 is a dead canopy (no foliage in the canopy), and 2–4 are stages of decline based on thinning of foliage.
and dieback of branches [24-25]. This rating system is rapid and user-friendly, and the ratings have been shown to correspond to EAB larval densities in the trees [26]. However, the rating system does not include destructive harvest of the trees to confirm infestation status.

EAB is very difficult to detect at low densities. Thus, ash canopy health class 1 trees may be infested by very low densities of EAB larvae [26]. The data were chosen for both data calibration and validation of the model because the five canopy health classes roughly correspond to the five infestation classes that were used in the model (susceptible, low, medium, high infestation, and dead trees). Although individual trees sometimes improved in health or declined by multiple categories within a year, the mean data at the plot level generally fit the model assumptions of a yearly decline by one health category.

In this study, the impact and spread rate parameters for each infestation level are calibrated using one site including three plots from Maumee Bay State Park (MB), as shown in Figure 5.3 (Plot 1: N41°40.847’ W83°21.181’, Plot 2: N41°40.821’ W83°21.204’, Plot 3: N41°40.860’ W83°21.216’).
The real infestation data consists of seven years of data from 2005 to 2011. Therefore, the impact and spread rate parameters are calibrated for a medium realization scenario over the seven-year period. The actual number of trees at the susceptible tree level and each infestation level in year 2005 is used and fed into the model as an initial tree infestation level in plots 1, 2, and 3. Fitted values of spread rates and impact rates for infestation levels 1, 2, and 3 are determined by trial and error. After several computational experiments, impact and spread rates are estimated as 0.18, 0.25, and 0.32 for infestation levels 1, 2, and 3, respectively. Finally,
the number of trees in infestation level 4 is used as a benchmark to compare the real data and predicted results after data calibration.

Computational experiments for data calibration are performed for the 2-by-2 gridded landscape shown in Figure 5.4. Figures 5.5(a)–(d) represent the comparison of the cumulative number of dead trees for real data and predicted results over the seven-year period for plots 1, 2, and 3 and the overall MB site.

Figure 5.5. Comparison of real data and predicted results for infestation level 4 over seven-year period for plots 1, 2, and 3 and overall MB site, respectively.

Figure 5.5 shows that although the cumulative number of dead trees is slightly overestimated, the predicted values of dead trees provide a good fit in all three plots (Figures
In order to provide a statistical analysis, the paired-t-test is performed for the entire site to compare the newly dead trees each year in predicted and actual data. Note that the paired-t-test assists in analyzing the differences between the pairs of yearly predicted and actual dead-tree populations. The results shown in Table 5.1 prove that there is no statistical difference between the predicted and actual data at the 5% significance level over the seven-year period.

### TABLE 5.1. STATISTICAL ANALYSIS USED TO COMPARE MODEL PREDICTIONS AND ACTUAL DATA AT MB SITE

<table>
<thead>
<tr>
<th></th>
<th>Real data</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9.33</td>
<td>8.96</td>
</tr>
<tr>
<td>Variance</td>
<td>33.86</td>
<td>104.08</td>
</tr>
<tr>
<td>Observations</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>t Stat</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.47</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.01</td>
<td></td>
</tr>
</tbody>
</table>

### 5.3.2. Model Application

Table 5.2 shows the initial population structure and parameters of the model. The parameter values are collected by utilizing several resources including journal papers, online resources, and expert opinions.

Computational experiments are conducted using a 4-by-4 landscape (100 ha), with each site representing 6.25 ha of land. Considering that 1 ha of landscape can include 50–70 trees [27], each site is assumed to contain 300–400 randomly generated trees. Invasion frequency $q \sim U[0,1]$ represents the percentage of initially infested trees at a site with uniform distribution,
where $U[a, b]$ denotes a number drawn uniformly from the interval $[a, b]$. The infested population is clustered into four levels. While trees at infestation levels one, two, and three pose a threat to susceptible trees, those trees in infestation level four are considered to be dead and do not spread infestation. Therefore, the initial number of infested trees at each infestation level one, two, and three at a given site is calculated by $q \cdot \nu_k$, where $\nu_k \sim U[0,1]$ is randomly generated with uniform distribution and $\sum_{k=1}^{3} \nu_k = 1$. Note that the number of trees in infestation level four is initially assumed to be zero at each site.

### TABLE 5.2. INITIAL POPULATION STRUCTURE AND PARAMETERS

<table>
<thead>
<tr>
<th>Description</th>
<th>Invasion frequency</th>
<th>Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of cells invaded in 4-by-4 landscape</td>
<td>$q \sim U[0,1]$</td>
<td>300–400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Symbol</th>
<th>Unit</th>
<th>Case study value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact rate of each infested tree at infestation level $k$</td>
<td>$r_k$</td>
<td>—</td>
<td>18%, 25%, 32%, 0%</td>
<td>1</td>
</tr>
<tr>
<td>Infestation impact of $k$th-level infested trees in neighboring site $j$</td>
<td>$\theta_k$</td>
<td>—</td>
<td>18%, 25%, 32%, 0%</td>
<td>1</td>
</tr>
<tr>
<td>Percentage change (realization) in belief of infestation</td>
<td>$\beta_k$</td>
<td>—</td>
<td>0%, 20%, 40%</td>
<td></td>
</tr>
<tr>
<td>Probability assigned to change in the belief</td>
<td>$\rho$</td>
<td>—</td>
<td>0.4, 0.3, 0.3</td>
<td></td>
</tr>
<tr>
<td>Probability for scenario $s$</td>
<td>$\pi_s$</td>
<td>—</td>
<td>(0.1)</td>
<td></td>
</tr>
<tr>
<td>Surveillance cost</td>
<td>$c_1$</td>
<td>$/tree$</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Treatment cost</td>
<td>$c_2$</td>
<td>$/tree$</td>
<td>120</td>
<td>[5]</td>
</tr>
<tr>
<td>Removal cost</td>
<td>$c_3$</td>
<td>$/tree$</td>
<td>888</td>
<td>1</td>
</tr>
<tr>
<td>Monetary value of each susceptible tree</td>
<td>$\alpha$</td>
<td>$/tree/year$</td>
<td>+54</td>
<td>1</td>
</tr>
<tr>
<td>Penalty value assigned to each highly infested tree</td>
<td>$\vartheta_k$</td>
<td>$/tree/year$</td>
<td>-50</td>
<td>1</td>
</tr>
<tr>
<td>Budget</td>
<td>$\Psi_s$</td>
<td>$/scenario$</td>
<td>$100,000–$450,000</td>
<td></td>
</tr>
<tr>
<td>Probability of infestation spread from site $j$ to $i$</td>
<td>$P_{j \rightarrow i}$</td>
<td>—</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>Discount factor</td>
<td>$\delta_i$</td>
<td>—</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>

1. Expert opinion
Each infestation level poses a different extent of threat to susceptible trees, except for those that are dead in the population. Therefore, the impact rate of each infested tree (number of newly infested trees due to infested tree) at infestation levels one, two, three, and four is estimated at 18%, 25%, 32%, and 0%, respectively. Furthermore, each tree in an infestation level \( k \) transitions to the upper infestation class in the following period until it reaches the dead-tree cluster.

Infestation size in a given landscape may change under changing weather conditions, soil nutrient levels, and transportation of woody materials. Although managers would have beliefs about infestation levels, the real number of infested trees at each infestation level might be different than their beliefs due to uncertainties in the aforementioned conditions. Therefore, surveillance is necessary to realizing (determining) the true number of infested trees, which may be different than the belief of the infestation (refer to Figure 5.1 in section 5.2.1 for details about surveillance). Therefore, we assume that the belief of infestation might change by 0%, +20%, and +40%, and (low, medium, and high realization, respectively) with 0.4, 0.3, and 0.3 probabilities after surveillance. If surveillance is not applied, then belief about the infestation does not change, with a probability of 1.

In the case of insecticide treatment, its effectiveness is assumed to be 100%. Treatment cost per one tree is estimated to be $120. If a tree with infestation level one or two is treated in a given period, then it is safe from a new infestation in the following two years of treatment. Furthermore, infestation levels three and four are subject to removal, at a cost of $888 including the replacement of an ash tree. The objective function maximizes the discounted net
benefits from susceptible ash trees by favoring each susceptible tree by $54$, and penalizing each third- and fourth-level-infested tree by $-50$ with a discount rate of $0.75\%$.

5.4. Results

Results demonstrate the computational effectiveness of the new cutting planes and provide insights regarding the optimal surveillance, treatment, and removal strategies for controlling EAB invasion under various scenarios. The multi-stage stochastic model presented in section 1 is solved by using CPLEX 12.6 [27] in a high-performance computer running the Linux Operating System with a 32 CPU and 64.0 GB of memory. A time limitation is not imposed for solving the test instances. Selected results chosen from interesting problem configurations are reported here.

5.4.1. Model Validation

After calibrating the impact and spread rate parameters, sites 1 to 9, as shown in Figure 5.6, are used for data and model validation. Similar to the MB site, each site has three plots; thus, a 2-by-2 gridded landscape is generated for each site to represent the EAB spread between plots in a given site. Visual comparisons of the predicted and real cumulative number of dead trees are shown for sites 1 to 9 in Figure 5.7. Furthermore, statistical analysis for comparing the predicted and real number of dead trees for the sites are presented in Table 5.3.

Figure 5.6 demonstrates that real and predicted data follow similar trends in the dead-tree population over the seven-year period. Furthermore, Table 5.3 represents the paired-t-test results for the comparison of the yearly real and predicted number of dead trees in the nine sites. The results in Table 5.3 demonstrate that there is no statistical difference between the predicted new dead-tree population and the real dead-tree data at each year of infestation.
at the 5% significance level for all sites tested. Based on the statistical results and visual model predictions compared with real data, it can be concluded that the proposed model provides a valid estimation for the number of dead trees for the considered seven-year time horizon.

Figure 5.6. Comparison of real data and predicted results for infestation level 4 over seven-year period for sites 1 to 9
Figure 5.7. Comparison of real data and predicted results for infestation level 4 over seven-year period for each plot in site 3

<table>
<thead>
<tr>
<th>TABLE 5.3. PAIRED-T-TEST RESULTS FOR COMPARISON OF REAL AND PREDICTED NUMBER OF DEAD TREES*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of trees</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>t Stat</td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
</tr>
<tr>
<td>t Critical one-tail</td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
</tr>
<tr>
<td>t Critical two-tail</td>
</tr>
</tbody>
</table>

*Hypothesized mean difference=0; Degrees of freedom=5

Although there is no statistical difference between the predicted new dead-tree population and the real dead-tree data for all sites, the predicted number of dead trees is slightly underestimated for sites 3, 6, and 8 (Figures 5.6(3), 5.6(6), and 5.6(8)) because there is a sudden increase in the number of dead trees in some plots of these sites. For example, Figure 5.7 shows a visual comparison real dead-tree data and predicted results for site 3.

Although the model provides statistically significant predictions for plots 1 and 3, it falls short of predicting the total number of dead trees in plot 2. As demonstrated in Figure 5.7 for
plot 2, there is a sudden increase in the number of dead trees from year 2006 to 2007. Although there was only one infested tree in infestation level one at the beginning of 2005 in real data, all trees became infested in 2006, and all of them died by 2008. The possible reason for the sudden and unexpected growth of EAB in that plot could be its spread from surrounding sites. Another reason might be the high uncertainty of detecting a visually healthy tree as a first-level infested tree in plot 2. Therefore, although there are several uncertainties that impact the EAB infestation, the validation results indicate that the proposed model successfully captures the EAB growth and spread dynamics, and is statistically successful in predicting spread and growth of the EAB invasion.

5.4.2. Computational Results with Cuts

In this section, we compare the results and solution performance of the original MSS-MIP model with the model including RC and TP cuts given in equations (5.21) and (5.22), respectively. Table 5.4 presents a comparison of the computational performance of the model with and without RC and TP cuts for different time horizon and landscape sizes with an ample budget allocation. The columns from left to right represent the number of time periods, size of the gridded landscape, and solution times and objective values for the original model and model with the preprocess algorithm, respectively.

Results show that RC and TP cuts improve the average CPU solution times by a factor of 44.75 for 3-period, 28.74 for 4-period, and 8.09 for 5-period instances. The improvement in solution times becomes more apparent as the difficulty of the problem increases temporarily and spatially. For example, while there is no significant difference in solution times between original model and problem with the algorithm for a 1-by-1 landscape size, solution times are
improved by a factor of 56.5 and 124 for a 10-by-10 landscape with three and four time periods, respectively. Solution times are also improved for five-period instances by a factor of 8.09 for a 5-by-5 landscape. Due to the complexity and size of the original model, instances with 6-by-6 and larger landscapes over five-period are not solvable with the current state of the art solver. However, larger instances with five periods is solvable within hours using the cutting planes. Furthermore, objective function values (net benefits) may slightly change as the landscape size increases when results of the original model and model with cuts are compared. The difference in the objective values may happen due to the default CPLEX integrality tolerance parameter, which specifies the amount in which an integer variable violates integrality requirement [23]. The tolerance of CPLEX results in epsilon deviations from integer 0 or 1 value. The impact of epsilon deviations on the objective value escalates as the number of variables and constraints are increased in the model which results in slight differences in the objective values. It can be concluded that the RC cuts and TP cuts obtained by the preprocess algorithm provides significant advantage to the model in terms of solution time.
### TABLE 5.4. COMPARISON OF SOLUTION PERFORMANCES

<table>
<thead>
<tr>
<th>No. of periods</th>
<th>Size</th>
<th>Original model</th>
<th>Original model + cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Real time</td>
<td>CPU time</td>
</tr>
<tr>
<td>3</td>
<td>1x1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2x2</td>
<td>5</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>3x3</td>
<td>5</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>4x4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5x5</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>6x6</td>
<td>19</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>7x7</td>
<td>47</td>
<td>454</td>
</tr>
<tr>
<td></td>
<td>8x8</td>
<td>61</td>
<td>513</td>
</tr>
<tr>
<td></td>
<td>9x9</td>
<td>69</td>
<td>482</td>
</tr>
<tr>
<td></td>
<td>10x10</td>
<td>146</td>
<td>1640</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>35.7</td>
<td>341.5</td>
</tr>
<tr>
<td>4</td>
<td>1x1</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2x2</td>
<td>71</td>
<td>1192</td>
</tr>
<tr>
<td></td>
<td>3x3</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>4x4</td>
<td>812</td>
<td>10012</td>
</tr>
<tr>
<td></td>
<td>5x5</td>
<td>987</td>
<td>9856</td>
</tr>
<tr>
<td></td>
<td>6x6</td>
<td>991</td>
<td>13412</td>
</tr>
<tr>
<td></td>
<td>7x7</td>
<td>998</td>
<td>13675</td>
</tr>
<tr>
<td></td>
<td>8x8</td>
<td>1812</td>
<td>23626</td>
</tr>
<tr>
<td></td>
<td>9x9</td>
<td>2111</td>
<td>19864</td>
</tr>
<tr>
<td></td>
<td>10x10</td>
<td>3042</td>
<td>33872</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1084.1</td>
<td>12554.5</td>
</tr>
<tr>
<td>5</td>
<td>1x1</td>
<td>3544</td>
<td>6779</td>
</tr>
<tr>
<td></td>
<td>2x2</td>
<td>8101</td>
<td>15992</td>
</tr>
<tr>
<td></td>
<td>3x3</td>
<td>11872</td>
<td>39613</td>
</tr>
<tr>
<td></td>
<td>4x4</td>
<td>21543</td>
<td>71305</td>
</tr>
<tr>
<td></td>
<td>5x5</td>
<td>23554</td>
<td>78401</td>
</tr>
<tr>
<td></td>
<td>6x6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>7x7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>8x8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9x9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>10x10</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>13722</td>
<td>42418</td>
</tr>
</tbody>
</table>

Average solution time (First five)
5.4.3. Impact of Surveillance on Net Benefits of Ash Trees

The timing and target population of the surveillance highly impacts the net benefits from the ash trees. Although a surveillance decision is given for the entire landscape, the surveyed population changes based on the number of infested tree population at each level and timing of the former surveillance (refer to Figure 5.1 in section 5.2.1 for detailed explanation). Here, the base case scenario represents the best scenario (scenario 704), which provides maximum net benefits for a 4-by-4 landscape over a five-year planning horizon with a budget allocation of $150,000. Furthermore, the best scenario also represents the lowest realization for each infestation level at each time period in which belief of the infested population does not change after surveillance.

In order to compare the impact of surveillance decisions, the net benefits and costs associated with the base case scenario is compared with scenarios that have the exact same realizations over time but differ in their surveillance regimen (768, 688, 684, and 683) in Figures 5.8(a)–5.8(c). Surveillance is applied in the first period, in the first two, three, four, and five periods for scenarios 768, 704, 688, 684, and 683, respectively.

Results shown in Figure 5.8(a) reveal that with the given budget allocation, surveillance should be applied in the first two years of the planning horizon. The reason is that surveillance in the first two years is sufficient to detect an infected population that would be treated or removed with the given budget allocation. Applying surveillance consecutively each year increases its marginal cost after the second year (Figure 5.8(b)), where applying surveillance again in the third year reduces the net benefits. The reason for this is that the budget necessary to control infestation is allocated for surveillance, which reduces the net benefits significantly.
As shown in Figure 5.8(a), tree removal is opted out, if surveillance is applied in the first three years consecutively. Furthermore, applying surveillance in the following years after the third-year results in a reduction in the number of treated trees due to limited budget, thus further reducing the net benefits.

![Figure 5.8(a)](image)

Figure 5.8(a). Net benefits and costs for base case scenario with $150,000 budget given surveillance decisions in consecutive years starting from year 1.

![Figure 5.8(b)](image)

Figure 5.8(b). Marginal net benefits and marginal cost of surveillance for base case scenario with $150,000 budget and surveillance decisions given in consecutive years starting from year 1.

However, as the budget allocation increases and becomes ample, applying surveillance in the first-time period is enough to control the infestation, as show in Figure 5.8(c). This is because once surveillance is applied, the true number of infestations at each level is realized in
the first period, and infestation is eradicated immediately. Therefore, since the region would be free of infestation in the first period, only susceptible trees will be targeted for surveillance after the first period, which only increases the cost of surveillance, treatment and removal without changing the net benefits.

Figure 5.8(c). Net benefits and costs for base case scenario with ample budget and surveillance decisions given in consecutive years starting from year 1

5.4.4. Rule of Thumb for Treatment and Removal

In this section, we analyze the changes in the number of trees with different infestation levels along with the number of treated and removed trees for the base case scenario. Results shown in Figures 5.9(a) and 5.9(b) reveal that insecticide treatment occurs primarily in periods one and two, in order to contain the infestation in the first and second infestation level. Treatment is applied to the entire first- and second-level infested trees after surveillance is applied in the initial period. Once all the first- and second-level populations are treated in the initial period, some budget is allocated to remove trees in the third-level infestation. Note that it is crucial to initially treat first, and especially second infestation levels because they will transition into upper infestation levels and pose more threat in the following period. Furthermore, although the third infestation level poses the highest threat to susceptible trees
in the first period, they will transition into a dead-tree cluster in the second period and will not spread infestation. Therefore, they are ignored for the sake of treating younger-age clusters until the second period when removal is applied.

In the second period, surveillance is applied to trees that were susceptible in the first period. This allows forest managers to detect newly infested trees due to the unremoved third-level infested trees in the first period. Once the newly infested trees are treated, the region becomes free of infestation, except for the dead trees, which do not spread infestation. A similar surveillance and control strategy is observed as the budget is increased to $200,000. However, increasing the budget to $250,000 enables managers to remove all highly infested third-level trees in the initial period. This results in eradication of all infested trees that pose a threat to the environment in the first period; thus, surveillance is not required in the following periods.

Therefore, to achieve the highest tree benefits, treatment occurs mostly on the second-level infestation followed by the first-level infestation in the initial periods to prevent the transition of second-level trees to the third level in the following period. Furthermore, removal mostly occurs after the second and third periods when there are no remaining trees in which to apply treatment.

![Figure 5.9(a). Total number of susceptible and infested trees over the years](image-url)
5.4.5. Cost vs Benefit Analysis

In this section, we analyze the changes in treatment, removal, and surveillance costs and net benefits by allocating different budget levels throughout the five-year period for the base case scenario. Results shown in Figure 5.10 reveal that treatment of first- and second-level infested trees is given more importance than removing the highly infested or dead trees, which supports the results in section 5.4.3. If there is tight budget ($100,000) for EAB control, surveillance is applied only once, so that the budget would be allocated more for treatment. Increasing the budget to $150,000 or $200,000 would allow managers to treat all trees in the first and second infestation level, but the budget would not be sufficient to remove all trees in the third infestation level. This causes new infestations in the second period, which requires surveillance of susceptible trees in the second period in order to detect newly infested trees. Therefore, surveillance cost and removal cost increase in the case of allocating $150,000 or $200,000, since surveillance is also applied in the second period and more budget is allocated for removal. The budget becomes sufficient to treat or remove all trees in the first three infestation levels in the initial period for the case of having a budget of $250,000 or more.
Therefore, surveillance is not necessary in the following periods since there are no infested trees remaining in the first, second, and third infestation levels after the first time period. Furthermore, the remaining budget after treating or removing the first three infestation levels is allocated for removing the dead trees. Therefore, surveillance and treatment cost does not change in the case of allocating a budget of $250,000 or more for EAB management. The model concentrates on removing highly infested trees after applying treatment, in order to alleviate the incurred penalties due to trees in the fourth infestation level.

Figure 5.10. Net benefit and control costs adjustment for different budget allocations for base case scenario

5.4.6. Sensitivity Analysis on Impact Rates

In this section, we analyze the effects of percentage changes in impact rates on the net benefits and show the results in Table 5.5. The columns from left to right represent the explanation for the alterations in impact rates, impact rates for the first, second, and third level of infestation, respectively, net benefits, and change in the net benefits with respect to the base case scenario. In order to see the effect of each infestation level’s impact on net benefits,
the total infested population in each cell is divided into three equal subpopulations, which represent the number of infested trees in infestation levels of one, two, and three.

Results show that changes in the impact rates significantly affect net benefits. A ±50% change in the impact rates results in a ±5.5% change in net benefits. Furthermore, increasing the impact rates of infestation levels reduces the net benefits proportionately. In addition, while a ±50% change in the impact rate of the third infestation level affects net benefits by approximately ±3%, changes in the impact rate of second- and first-level infestations would result in ±1.6% and ±0.8% variation, respectively. Therefore, results show that changes in the impact rate of third-level infestation causes more damage than other infestation levels.

<table>
<thead>
<tr>
<th>Explanation</th>
<th>Impact Rates</th>
<th>Net Benefits</th>
<th>Change in Net Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0.05 0.12 0.15</td>
<td>230,680</td>
<td>-</td>
</tr>
<tr>
<td>-50% on impact rates</td>
<td>0.025 0.06 0.075</td>
<td>243,342</td>
<td>0.055</td>
</tr>
<tr>
<td>+50% on impact rates</td>
<td>0.075 0.18 0.225</td>
<td>218,009</td>
<td>-0.055</td>
</tr>
<tr>
<td>+100% on impact rates</td>
<td>0.1 0.24 0.3</td>
<td>205,671</td>
<td>-0.108</td>
</tr>
<tr>
<td>+150% on impact rates</td>
<td>0.125 0.3 0.375</td>
<td>193,541</td>
<td>-0.161</td>
</tr>
<tr>
<td>+50% on 1st level impact rate</td>
<td>0.075 0.12 0.15</td>
<td>228,904</td>
<td>-0.008</td>
</tr>
<tr>
<td>+50% on 2nd level impact rate</td>
<td>0.05 0.24 0.15</td>
<td>227,102</td>
<td>-0.016</td>
</tr>
<tr>
<td>+50% on 3rd level impact rate</td>
<td>0.05 0.12 0.3</td>
<td>223,415</td>
<td>-0.031</td>
</tr>
<tr>
<td>-50% on 1st level impact rate</td>
<td>0.025 0.12 0.15</td>
<td>232,438</td>
<td>0.008</td>
</tr>
<tr>
<td>-50% on 2nd level impact rate</td>
<td>0.05 0.06 0.15</td>
<td>234,299</td>
<td>0.016</td>
</tr>
<tr>
<td>-50% on 3rd level impact rate</td>
<td>0.05 0.12 0.075</td>
<td>238,121</td>
<td>0.032</td>
</tr>
</tbody>
</table>

5.5. Discussion and Conclusion

In this paper, we propose a novel MSS-MIP model and solution algorithm to determine efficient management strategies for invasive species considering infestation uncertainty. The uncertain growth of invasion is realized by applying surveillance, which is followed by treatment.
or removal, based on the infestation level and available budget allocation. Surveillance, treatment, and removal decisions are integrated into a hybrid scenario tree. The proposed approach is applied for control of EAB invasion under uncertainty in the population growth. We present computational results using a realistically scaled spatio-temporal problem on a 5-by-5 gridded landscape over a five-year period.

Results indicate that surveillance would be effective if it is applied in the first two years of the planning horizon. Surveillance would be followed immediately by treatment without additional cost in the case of controlling perennial invasive plants such as sericea lespedeza, kudzu, or purple loosestrife. However, surveillance is a necessity for preemptive treatment and removal of infested ash trees because it allows decision makers to detect them in each infestation class with an additional cost and procedure, thus allowing treatment or removal of infested trees accordingly before all trees are infested or die. Once the true number of trees at each infestation level is detected, the optimal decision is to treat mid-level-infested trees, followed by low- and high-level-infested trees. This prevents mid-level-infested trees to become highly infested in the following period. Results indicate that if the budget is not sufficient, then decision makers may need to let some highly infested trees die in favor of treating low- and mid-level-infested trees.

The number of constraints and variables increase exponentially as the problem size increases spatially and temporarily. Therefore, we use a high-performance computer with a 32-core/64 GB ram/Linux Operating System to solve the proposed MSSP-MIP model. Even though a strong computer has been utilized, and the proposed model could be solved for a 10-by-10 gridded landscape over a five-year period (~50 million constraints) with ample budget
allocation, the problem size was reduced to a 5-by-5 gridded landscape (~12 million constraints) to solve the problem with a tight budget allocation. Therefore, although the solution time is highly improved with the proposed algorithm, different algorithms could be developed to solve the proposed model more efficiently.

The unique model structure with efficient constraints significantly reduces the size and complexity of the problem. We initially fixed the discretized binary variables to represent all possible surveillance decisions. Therefore, we prevent nonlinearities in the model, especially in equations (5.2), (5.3), and (5.6), thus facilitating the solution performance by considering all possible surveillance decisions. We further reduce the size of the problem by providing a surveillance decision for the entire landscape instead of for each site. Although the surveillance decision is given for the entire landscape in a given period, the elegant formulation of equations (5.2) and (5.3) by clustering infestation classes and tracking each infested tree prevents unnecessary surveillance of infested trees at each cell over a planning horizon. Therefore, discretizing the surveillance decisions makes the proposed model of practical importance.

Our numerical results indicate that the preprocess algorithm significantly improves the solution time. The integration of cuts obtained from the preprocess problem into the model along with the additional cutting planes reduce the size of the B&B tree and tighten the upper bound of the problem. This, in turn, improves the solution time by up to an average of 44 times, compared to solving original model. The improvement in the solution time is more apparent for larger instances, due to the increased number of sites in a given landscape and thus the increased number of constraints and variables.
Additional work could focus on developing new solution algorithms or decomposition methods to solve the proposed model for bigger instances. Furthermore, the discretized binary surveillance parameters could be replaced with binary decisions at the expense of increasing the problem complexity and thus the solution time.

5.6. References


CHAPTER 6

OPTIMIZING MULTIMODAL CANCER TREATMENT: MIXED-INTEGER PROGRAMMING APPROACH

6.1. Introduction

Cancer is defined as an imbalanced or uncontrolled cell proliferation of abnormal cells [1]. If left untreated, it spreads to different parts of the body (metastasis), which results in organ failure and thus leading to death. It is estimated that 1,685,210 new cases of cancer will be diagnosed and 595,690 patients would die due to different forms of cancer in 2016 [2].

Among the many forms of cancer, breast cancer is the second leading cause of death after lung cancer in women, with an estimation of 246,600 new cases and 40,450 deaths in 2016 [2]. Breast cancer usually begins either in the cells of the lobules, which are milk-producing glands, or the ducts, the passages that drain milk from the lobules to the nipple [3]. While non-invasive cancers stay within the milk ducts or lobules in the breast, invasive cancers grow into normal, healthy tissue. Although early breast cancer can be cured and leads to 100% five-year survival rate [4], postponing treatment could lead to an advanced stage cancer, metastasis, and death.

Surgery is a widely used method for early breast cancer treatment. In order to ensure that the tumor is entirely resected from the breast area, the tumor is removed along with the tumor margin, which is a layer of healthy tissue around the tumor [3]. Although the purpose of the surgery is to remove the entire tumor, it is possible that a portion of the cancer cells may remain in the breast area after surgery [5] because they are not visually detected.
Radiation therapy also plays an important role for the removal of cancer cells in all stages of breast cancer. Radiation annihilates cancer cells that may have been left after surgery [6]. It uses a special high-energy beam to damage the molecular structure of the cancer cells and consequently suppress them [7]. The goal of radiation treatment is to kill cancer cells using a uniform dose of radiation while limiting the potential damage to healthy organs and tissues. Therefore, the major challenge in radiation therapy is to limit the dose of radiation to prevent medical complications to the organs-at-risk (OARs), which are the structures or organs that are close to the tumor and in the region of radiation treatment [8].

Another widely used treatment method for breast cancer is chemotherapy, which targets the cancer cells both in the tumor area and in the other parts of the body (metastasis) considering the risk of cancer cell spread. During chemotherapy treatment, the anti-cancer agent annihilates the cancer cells while at the same time killing the normal cells [9]. This phenomenon leads to a tradeoff between toxicity, the amount of chemotherapy drug used, and cancer-cell reduction. Furthermore, during a long chemotherapy treatment protocol, some of the cancer cells may actually become resistant to the chemotherapy drug in use [10]. Therefore, chemotherapy treatment with higher initial doses at the beginning of the treatment is suggested to prevent resistance of cells to the drug [11]. However, due to the maximum allowable toxicity per treatment and acute toxicity level over the course of the treatment, the chemotherapy drug should be limited and administered cautiously. Therefore, the dependency and tradeoffs between the requirement to reduce the number of cancer cells, drug dosage limits, and its impact on normal cells are factors that complicate the development of the best chemotherapy treatment plan.
This study addresses the problem of developing a spatio-temporal model to determine an optimal breast cancer treatment sequence considering surgery (S), radiation therapy (R), and chemotherapy (C) in combination with the optimal dose schedules for chemo- and radiotherapy treatments. In the cancer treatment literature, mathematical models are typically defined as either pharmacokinetic or pharmacodynamic. While pharmacokinetic models concentrate on distribution and metabolism of the drug or radiation, pharmacodynamic models focus on the impacts of drugs and radiation on cancer cell population [12]. The objective in this study is to develop a pharmacodynamic MIP model that provides the optimal treatment sequence along with the optimal dose schedules by minimizing cancer population over the course of treatment.

The proposed multi-modal treatment problem consists of specific MIP models for surgery, radiotherapy and chemotherapy modalities. In this study, these three MIP models are solved sequentially for different combinations of treatment regimens, such as SRC (surgery-radiotherapy-chemotherapy), SCR, RC, and CR regimens. While each proposed model has distinct features, the objective function of minimizing cancer cell population and use of the Gompertz function [13-14] to track the spread and growth of cancer cell population are common aspects in all three models. In the Gompertz model, tumor growth is assumed to be exponential in the early stages of the cancer. However, the doubling time of the cancer only begins to increase as the tumor size gets larger. Eventually, the proliferation rate slows down as the cancer cell population asymptotically approaches the plateau population, and then the growth process stops due to lack of space and oxygen. The Gompertz function is a realistic representation of the tumor growth because it limits the population growth as the tumor size reaches the carrying capacity of the host organ. In all proposed MIP models, a three-
dimensional spatial structure is represented by using a grid of voxels. The surgery model involves the removal of voxels that are inside the tumor margin as well as the resection of cancerous cells on the tumor margin with a specific resection efficiency rate. Furthermore, the chemotherapy model takes into account the acute and per-treatment toxicity levels to find the optimal drug administration schedule. Finally, radiation treatment model incorporates optimal dose fractionations considering the vulnerability of the OARs to radiation. Therefore, this multimodal problem provides a comprehensive model for breast cancer treatment by considering the key dynamics of each treatment modality as well as the spatio-temporal tumor growth in three dimensions.

Although there are several pharmacokinetic models and a significant amount of information available to analyze radiation therapy, chemotherapy or combinations of these treatment methods [7, 8, 12, 15, 16, 17, 18], surgery is not included as a treatment method in the sequence of treatment modalities. However, surgery should be incorporated into the model because the optimal treatment dose and frequency of the radiation or chemotherapy treatment may change due to possible remaining cancer cells in the breast area after surgery. Considering the possible number of remaining cancer cells, chemo- and radiotherapy doses can be regulated. Consequently, the damages stemming from the low or high dose treatment can be eliminated over the course of the treatment.

Furthermore, while a three-dimensional structure is utilized in some studies [6-7], the goal in these studies, is to provide optimal radiation dose distribution among the voxels. In addition, while studies that concentrate on radiation therapy and chemotherapy treatment focus on developing DP models [15, 18, 19] or MIP models [6-7] to determine optimal
treatment regimens, spatio-temporal spread and growth have been avoided to reduce the complexity of the problem. Therefore, different than the literature, we propose an MIP model that incorporates spread and growth of cancer cells by integrating Gompertz growth dynamics into a dynamic spatial model. This allows tracking the population re-growth between and during each treatment modality, which would affect the radiation fractionation as well as chemotherapy drug administration schedules over the course of the treatment protocol.

In order to tackle the computational complexity of the model, we utilize piecewise approximation and linearization methods to linearize the highly complex quadratic equations and nonlinear constraints. The proposed three-dimensional spatio-temporal model is intractable to solve with DP or NLP methods, thus requiring the development of an equivalent or approximate linear model to solve the proposed multimodal problem. Furthermore, the MIP model could be solved using commercial MIP solvers such as CPLEX and Gurobi for reasonably sized instances.

The proposed model fills the research gap in the cancer treatment literature by first providing a spatio-temporal MIP model that provides the optimal sequence of treatment modalities over the cancer-treatment period. Second, the spatial dimension assists in predicting the spread and regrowth of cancer cells during and after the treatment period, which would enable tracking of the size and extent of the growth and spatial spread of the cancer cells. Finally, the proposed methodology differentiates between the efficiency of surgery on the margin of the tumor compared to surgery on the inner tumor structure.

The models here are applied to study the optimal treatment modalities for breast cancer treatment. The approach leads to several important results:
• First, the MIP model improves the solvability of the complex cancer treatment modality problem, while also considering the highly complex spatio-temporal cancer growth and spread.

• Second, linearizing the Gompertz growth function using a piecewise approximation approach results in a statistically similar growth pattern compared to the simulations using the non-linear Gompertz growth function.

• Third, it is optimal to apply surgery first before radiotherapy and chemotherapy modalities. In addition, although SRC and SCR treatment sequences eradicate the cancer population at the end of the treatment cycle, SRC provides better treatment effectiveness than SCR in terms of removing cancer cells in a shorter time period. Thus, it can be concluded that SRC is the best treatment sequence for the early stage breast cancer instance that is considered in this paper.

• Fourth, the proposed MIP model demonstrates that hypo-fractionation with initially high radiation doses would provide more effective results compared to hyper-fractionation in terms of annihilating cancer cells in a shorter time period.

• Fifth, a full-dose chemotherapy treatment with maximum allowable drug concentration at each drug administration may not be necessary unless chemotherapy is applied as an initial treatment before surgery and radiotherapy.

• Sixth, while a weekly chemotherapy schedule for 12 cycles leads to less cumulative toxicity, a once-every-three-weeks schedule results in faster cancer cell reduction compared to a weekly schedule.
The structure of this chapter is as follows: Section 6.2 defines the problem and presents the MIP model for different treatment modalities. Section 6.3 discusses an application of the model on the case of breast cancer and present related data. Section 6.4 follows with computational results. Section 6.5 presents discussion and offers directions for future research.

6.2. MIP Models

6.2.1. Notations

Sets

\( V \) \hspace{1cm} \text{Set of voxels} \ v \in 1.. \overline{V} \\
\( T \) \hspace{1cm} \text{Set of time periods} \ t \in 0.. \overline{T} \\
\( W \) \hspace{1cm} \text{Set of six neighbor sites} \ w \in 1.. \overline{W} \\
\( H \) \hspace{1cm} \text{Set of treatment types} \ h \in 1.. \overline{H} \ (1: \text{Chemotherapy}, 2: \text{Radiotherapy}, 3: \text{Surgery}) \\
\( OAR \) \hspace{1cm} \text{Set of organs at risk during radiotherapy treatment} \ i \in 1.. \overline{O}

Variables

\( P_v(t) \) \hspace{1cm} \text{Tumor population at voxel} \ v \ \text{at period} \ t \\
\( C(t) \) \hspace{1cm} \text{Chemo concentration at period} \ t \\
\( C_{eff}(t) \) \hspace{1cm} \text{Effective chemotherapy drug concentration at period} \ t \\
\( D(t) \) \hspace{1cm} \text{Chemo agent administered at period} \ t \\
\( R(t) \) \hspace{1cm} \text{Radiation dose for voxel} \ v \ \text{at period} \ t \\
\( \Theta_{w,v}(t) \) \hspace{1cm} \text{Cancerous cell spread to voxel} \ v \ \text{from neighbor voxel} \ w \ \text{at time period} \ t \\
\( M_{v,h}(t) \) \hspace{1cm} \text{Cancer cells annihilated at voxel} \ v \ \text{by treatment type} \ h \ \text{at period} \ t \\
\( r_v(t) \) \hspace{1cm} \text{Discrepancy between actual and prescribed radiation dose}
Binary Variables

\[ z_v(t) \begin{cases} 1 & \text{if prescribed irradiation dose is used for voxel } v, \\ 0 & \text{otherwise.} \end{cases} \]

Parameters

\( \tau \) Doubling time of tumor population

\( P_{\text{max}} \) Plateau population at voxel \( v \)

\( \phi \) Effectiveness rate of chemotherapy treatment

\( \lambda \) Spread rate of tumor

\( \xi \) Minimum occupancy rate for voxel before cancer cells start to migrate to immediate neighbors

\( C_{\text{max}} \) Maximum allowable drug concentration per treatment

\( C_{\text{cum}} \) Maximum allowable cumulative drug concentration

\( \zeta \) Decay rate of chemo concentration

\( C_{\text{th}} \) Threshold chemotherapy amount

\( PDose \) Total radiation dose over course of treatment

\( \varphi \) Percentage of voxels treated with prescribed dose

\( D^{OD} \) Maximum overdose per each radiation treatment

\( D^{UD} \) Maximum underdose per each radiation treatment

\( \gamma_i \) Maximum allowable total irradiation at organ at risk \( i \)

\( \alpha \) Patient specific radiation tolerance rate

\( \beta \) Tumor specific radiation tolerance rate

\( \psi \) Effectiveness rate of surgery at tumor margin
6.2.2. Mathematical Model

Cancer cells proliferate exponentially in the absence of drugs or surgery. Eventually the number of cancerous cells approaches a carrying-capacity population asymptotically, and then growth slows down. Furthermore, the spatio-temporal evolution of tumor cells results in the migration/spread of cancer cells to the surrounding areas. Finally, the impact of various treatment plans can be incorporated into the model describing the cell loss due to the selected treatment plan. Therefore, using the Gompertz equation [13], the cell population in voxel $v$, at time period $t$ is formulated as

$$
P_v(t+1) = P_v(t) + \sum_w \Theta_{w,v}(t) + \frac{1}{r} \ln \left( \frac{\ln(P_{\text{max}} / P_0)}{\ln(P_{\text{max}} / 2P_0)} \right) \left( P_v(t) + \sum_w \Theta_{w,v}(t) \right) \left( \ln \frac{P_{\text{max}}}{P_v(t) + \sum_w \Theta_{w,v}(t)} \right) - M_{v,h}(t)
$$

where $M_{v,h}(t)$ is the number of cancer cells annihilated with treatment type $h = 1, 2, \text{or } 3$, corresponding to chemotherapy, radiotherapy, and surgery, respectively.

Note that equation (6.1) is a quadratic constraint. A piecewise linear approximation (PLA) technique [20] is used to linearize the quadratic equation. In order to use this technique, the newly generated cancer cell population

$$
P_v(t) = \frac{1}{r} \ln \left( \frac{\ln(P_{\text{max}} / P_0)}{\ln(P_{\text{max}} / 2P_0)} \right) \left( P_v(t) + \sum_w \Theta_{w,v}(t) \right) \left( \ln \frac{P_{\text{max}}}{P_v(t) + \sum_w \Theta_{w,v}(t)} \right)
$$

is represented as a function of $P_v(t) + \sum_w \Theta_{w,v}(t)$, as shown in Figure 6.1. The reverse U-shaped function represents the fact that as the number of cancer cells increases in a voxel, the proliferation rate
increases with a decreasing rate, and after reaching the maximum proliferation rate, it decreases with an increasing rate until a plateau population is attained in a voxel. The piecewise linear function $P_v(t)$ has breakpoints $a_0, a_1, a_2, a_3,$ and $a_4$. Continuous variables $0 \leq y_{v,j}(t) \leq 1$ are defined for each breakpoint $j = 0, 1, \ldots, 4$, and binary variables $w_{v,l}(t) \in [0,1]$ are defined for each interval between breakpoints, $l = 1, 2, 3, 4$.

![Figure 6.1. Piecewise linear approximation of equation (6.1)](image)

Using the above notation, the PLA equations that are used to approximate equation (6.1) can be written as

$$P_v(t) + \sum_w \Theta_{w,v}(t) = \sum_j a_j \cdot y_{v,j}(t) \quad \forall v, t$$  \hspace{1cm} (6.2)

$$P_v(t) = \sum_j P_v(t,a_j) \cdot y_{v,j}(t) \quad \forall v, t$$ \hspace{1cm} (6.3)

$$y_{v,0}(t) \leq w_{v,1}(t), \quad y_{v,1}(t) \leq w_{v,1}(t) + w_{v,2}(t), \quad y_{v,2}(t) \leq w_{v,2}(t) + w_{v,3}(t), \quad y_{v,3}(t) \leq w_{v,3}(t) + w_{v,4}(t), \quad y_{v,4}(t) \leq w_{v,4}(t) \quad \forall v, t$$ \hspace{1cm} (6.4)

$$\sum_j y_{v,j}(t) = 1, \quad \sum_l w_{v,l}(t) = 1 \quad \forall v, t$$ \hspace{1cm} (6.5)
According to the available space, tumor cells can spread to neighbor cells. As the cancer cells proliferate, they tend to spread towards an attractor source such as a blood vessel [21]. However, for the sake of mathematical simplicity, spread in all directions is assumed, where cancer cells spread to surrounding voxels in equal amounts via a circular growth based on the available space in the neighbor voxels. Furthermore, it is assumed that cancer cells can spread to six immediate neighbor voxels at a rate of $\lambda$ once at least $100\%$ of the plateau population is reached in a given voxel (see Appendix A for further information). Without available space, i.e., none of the adjacent six neighbors on a three-dimensional grid has any vacancy, the cells are assumed to rest in a quiescent state until once again exposed to space. Therefore, the spread of cancer cells is formulated as

$$\Theta_{w,v}(t) = \left\{ \begin{array}{ll} 0 & \text{if } P_w(t) \leq \xi P_{\text{max}} \\ (P_w(t) - P_w(t-1)) \lambda & \text{o.w} \end{array} \right.$$  \hspace{1cm} (6.6)

In combination with equation (6.1), equation (6.6) ensures that the proliferated cancer cells migrate towards the neighbor voxels that are not at the carrying-capacity level.

In this proposed model, the objective is to minimize the new cancer cell population over the entire voxels and all time periods. Therefore, the objective function is described as

$$\min \sum_{v \in V} \sum_{t \in T} (P_v(t+1) - P_v(t))$$  \hspace{1cm} (6.7)

6.2.2.1. Treatment Method 1: Chemotherapy

The chemo concentration is detrimental to health and can also lead to life threatening side effects. Therefore, it should be controlled, and toxicity should be limited based on the drug
dosage in every treatment and over the course of the treatment. Therefore, periodic and cumulative chemical dose limit is formulated as

\[ C(t) \leq C_{\text{max}} \quad \forall t \]  \hspace{1cm} (6.8)

\[
\frac{1}{2} C(1) + \sum_{t=2}^{T-1} C(t) + \frac{1}{2} C(T) \leq C_{\text{cum}}
\]  \hspace{1cm} (6.9)

While equation (6.4) ensures that allowable drug dosage should be limited by \( C_{\text{max}} \), equation (6.5) guarantees that the cumulative acute toxicity during the treatment process cannot exceed an aggregate drug level \( C_{\text{cum}} \). Note that equation (6.9) provides an approximation of the area under the curve by using the trapezoid rule [22].

The drug concentration is reduced at every time step after the drug is administered with dose \( D(q) \) at period \( q \) [12]. Therefore, drug concentration change is formulated as

\[
C(t+1) = \begin{cases} 
\frac{1}{\varsigma} (e^{-\varsigma} - 1) \frac{D(q)}{h} + e^{-\varsigma} C(t) & \text{when } t = q \\
e^{-\varsigma} C(t) & \text{o.w.}
\end{cases}
\quad \forall t
\]  \hspace{1cm} (6.10)

where \( 1/\varsigma \) is the time constant. Equation (6.10) ensures that drug concentration decays exponentially with a constant rate \( e^{-\varsigma} \) after the drug is administered in period \( q \).

Furthermore, drugs may not be effective until a threshold level of the drug that is administered to the body or the drug concentration falls below the threshold level. Therefore, the effective drug concentration can be formulated as a piecewise function as follows [12]:

\[
C_{\text{eff}}(t) = \begin{cases} 
C(t) - C_{\text{th}} & C(t) > C_{\text{th}} \\
0 & \text{o.w.}
\end{cases}
\quad \forall t
\]  \hspace{1cm} (6.11)
While the drug is ineffective and does not kill the tumor cell population below the threshold, it still contributes to the toxicity of the body. Note that equation (6.11) creates a semi-continuous variable that results in a nonlinear function. Therefore, the nonlinear function given in equation (6.11) can be linearized as follows [23]:

\[ C_{\text{th}} - C(t) \leq C_{\text{th}} \cdot g(t) \]  
(6.12)
\[ C_{\text{eff}}(t) \leq (C_{\text{max}} - C_{\text{th}}) \cdot (1 - g(t)) \]  
(6.13)
\[ C(t) - C_{\text{th}} \leq (C_{\text{max}} - C_{\text{th}}) \cdot (1 - g(t)) \]  
(6.14)
\[ C_{\text{eff}}(t) - C(t) + C_{\text{th}} \leq C_{\text{th}} \cdot g(t) \]  
(6.15)
\[ 0 \leq C_{\text{eff}}(t) - C(t) + C_{\text{th}} \]  
(6.16)
\[ 0 \leq C_{\text{eff}}(t) \quad \forall t \]  
(6.17)

where $g(t)$ is a binary variable and receives a value of 1 if $C_{\text{th}} \geq C(t)$ and 0 otherwise. As an example, let the drug concentration in a given period be less than the threshold level ($C_{\text{th}} \geq C(t)$). In this case, the left-hand side of equation (6.12) is greater than 0, which forces $g(t)$ to be 1. Therefore, by equation (6.13), $C_{\text{eff}}(t)$ becomes zero, and equations (6.14) to (6.17) satisfy the condition. If $C_{\text{th}} \leq C(t)$, then the left-hand side of equation (6.14) becomes greater than zero, which ensures that $g(t)$ is zero. In this case, equation (6.13) and equation (6.16) provide an upper bound and lower bound for $C_{\text{eff}}(t)$, respectively, which are equal to each other. Thus, equations (6.12) to (6.17) ensure that $C_{\text{eff}}(t) = C(t) - C_{\text{th}}$ if $C(t) > C_{\text{th}}$, and 0 otherwise.
In the case of drug administration, the number of annihilated cells is a function of the effective dose of the chemotherapy agent, cancer population, and treatment efficiency of the drug [24]. Therefore, the number of annihilated cells in a given time period can be formulated as

\[ M_{r,h}(t) = C_{\text{eff}}(t)P_r(t)\phi \quad h = 1 \quad \forall v,t \]  

(6.18)

where \( \phi \) is the efficiency rate of the chemotherapy drug. Note that this is a quadratic function, which is difficult to linearize. Therefore, in order to formulate a linear constraint, the initial value of the tumor population at the beginning of the chemotherapy treatment, \( P_r(0) \), is used instead of \( P_r(t) \) at each time period, which leads to a linear function. However, this results in a constant number of cell annihilations at each period based on the initial population, which overestimates the population reduction. Therefore, in order to prevent an overestimation of the cell annihilation, the constant drug efficiency rate \( \phi \) is replaced with an efficiency function \( \hat{\phi}(t) \), which is strictly decreasing over time. This function is obtained using a number of simulations (see Appendix B for further information). Therefore, equation (6.18) is reformulated as

\[ M_{r,h}(t) = C_{\text{eff}}(t)P_r(0)\hat{\phi}(t) \quad h = 1 \quad \forall v,t \]  

(6.19)

In this way, equation (6.19) ensures that cells annihilated at each time period is a decreasing function of time, and thus equation (6.19) is analogous to the cell reduction function defined in equation (6.18).
6.2.2.2. Treatment Method 2: Radiotherapy

Similar to chemotherapy treatment, radiation cannot be prescribed more than an upper limit over the radiation treatment period. Furthermore, the radiation dose that is necessary to kill cancer cells cannot be given in one session. This can cause side effects such as damaging nearby healthy cells or organs, or radiation poisoning. Instead, the radiation dose is divided into smaller doses over a time period. In addition, the radiotherapy dose should exceed a threshold limit in order to provide a satisfactory treatment. Therefore, following the study of Lee et al. [7], the total number of voxels treated over a given period is formulated by using following set of constraints:

\[
\sum_r \left( R(t) - \sum_v r_v(t) \right) = PDose
\]

(6.20)

\[ r_v(t) \leq D_v^{\text{ref}} z_v(t) \quad \forall v, t \]  

(6.21)

\[ r_v(t) \geq D_v^{\text{ref}} (z_v(t) - 1) \quad \forall v, t \]  

(6.22)

\[ \sum_v z_v(t) \geq \varphi |V| \quad \forall t \]  

(6.23)

where \( r_v(t) \) is the gap between the actual administered dose over the treatment period and prescribed dose (e.g., a positive value of \( r_v(t) \) represents an overdose, while a negative value implies an underdose); \( z_v(t) \) is a binary variable that represents whether voxel \( v \) satisfies the prescribed dose bounds over the radiation treatment period or not; and \( \varphi \) is the minimum percentage of voxels that are supposed to be treated with the prescribed dose over the course of radiation treatment. Equations (6.20) to (6.23) ensure that at least \( \varphi \) percentage of the total number of voxels should be administered by the prescribed dose over the course of treatment.
It is known that there is a positive correlation between the level of irradiation of tumor cells and healthy tissue toxicity. Along with the dose limitation equations (6.20) to (6.23), organs-at-risk should also be taken into account because they are also exposed to radiation. Therefore, the concept of BED to limit the side effects of irradiation to the OARs. Following the study of Ulkenbach et al.[8], dose limitations for each OAR over the course of the treatment is formulated as

\[
\sum_i \gamma_i R(t) \left(1 + \frac{\gamma_i R(t)}{[\alpha / \beta]} \right) \leq Y_i \quad i \in OAR
\]  

(6.24)

where \([\alpha / \beta]\) is an OAR tissue sensitivity parameter, \(\gamma_i\) is the sparing factor, and \(Y_i\) is the upper bound for the radiation exposure for the \(i^{th}\) OAR. Equation (6.24) ensures that side effects of radiation at each fractionation are bounded by the tissue-specific exposure limit.

In the case of radiotherapy treatment, the number of cancer cells killed in each voxel is a fraction of the viable cancer cells following a single exposure to \(R\) Gy of radiation. Furthermore, radiation to the breast is delivered from two different angles with a difference of 180°, i.e., radiation comes from opposite directions facing each other [3]. This implies that each voxel \(v\) receives the same amount of radiation \(R\) at each fractionation. Therefore, the number of cancer cells killed with radiotherapy in each voxel is formulated as

\[
M_{r,h}(t) = P_v(t) \cdot \left(1 - e^{-\alpha R(t) - \beta R(t)^2} \right) \quad h = 2 \quad \forall v, t
\]  

(6.25)

where \(\alpha\) and \(\beta\) are the tumor sensitivity parameters for breast cancer.
Note that equations (6.24) and (6.25) are nonlinear. Similar to the linearization of equation (6.1), a PLA method is used to linearize equation (6.24). The constraint can be converted to equation (6.26) below:

\[
\frac{\gamma_i^2}{[\alpha / \beta]} \sum_t R(t)^2 + \gamma_i \sum_t R(t) \leq Y_i \quad (6.26)
\]

Therefore, \( R(t)^2 \) can be represented as a function of \( R(t) \), as shown in Figure 6.2.

![Figure 6.2. Piecewise linear approximation of equation (6.24)](image)

Thus, the PLA equations for equation (6.24) can be written as

\[
R(t) = \sum_j b_j \cdot x_j(t) \quad \forall t \quad (6.27)
\]

\[
R'(t) = \sum_j R(t, b_j) \cdot x_j(t) \quad \forall t \quad (6.28)
\]

\[
x_0(t) \leq \mu_1(t), \quad x_1(t) \leq \mu_1(t) + \mu_2(t), \quad x_2(t) \leq \mu_2(t) \quad \forall t \quad (6.29)
\]

\[
\sum_j x_j(t) = 1, \quad \sum_i \mu_i(t) = 1 \quad \forall t \quad (6.30)
\]
where \( b_0, b_1, \) and \( b_2 \) are the breakpoints; \( 0 \leq x_j(t) \leq 1 \) is defined for each breakpoint \( j = 0, 1, 2 \); and binary variables \( \mu_l(t) \in [0, 1] \) are defined for each interval between breakpoints \( l = 1, 2 \).

On the other hand, equation (6.25) is a highly nonlinear constraint, which can be linearized in two steps. Let \( r_{\text{eff}}(t) \) be the treatment efficiency per unit radiation dose; then

\[
r_{\text{eff}}(t) = \left(1 - e^{-\alpha_R(t) - \beta R(t)^2}\right).
\]

In the initial step, \( r_{\text{eff}}(t) \) can be defined as a function of \( R(t) \), which can be linearized using a piecewise linearization method. However, the resulting equality \( M_{v,h}(t) = P_v(t) \cdot r_{\text{eff}}(t) \) would become a multiplication of two continuous variables. In order to prevent such a nonlinear constraint, \( r_{\text{eff}}(t) = \left(1 - e^{-\alpha_R(t) - \beta R(t)^2}\right) \) is defined as a pointwise function (Figure 6.3), where \( R(t) \) receives a sequence of continuous values on a compact domain converging pointwise to \( r_{\text{eff}}(t) \).

![Figure 6.3. Pointwise function for \( r_{\text{eff}}(t) = \left(1 - e^{-\alpha_R(t) - \beta R(t)^2}\right) \)](image)

\[ r_{\text{eff}}(t) \]

\[ d_1 \]

\[ d_2 \]

\[ d_3 \]

\[ d_j \]
As shown in Figure 6.3, a pointwise function is generated by simulating \( r_{\text{eff}}(t) \) for a sequence of doses based on \( f_j = \left(1 - e^{-\alpha d_j - \beta d_j^2}\right) \) equation. After simulating \( r_{\text{eff}}(t) \) values (\( \forall t \)), the pointwise linear function for both \( R(t) = d_j \) and \( r_{\text{eff}}(t) \) can be formulated as

\[
R(t) = \sum_j d_j \cdot \eta_j(t) \quad \forall t
\]  
(6.31)

\[
r_{\text{eff}}(t) = \sum_j f_j \cdot \eta_j(t) \quad \forall t
\]  
(6.32)

\[
\sum_j \eta_j(t) = 1 \quad \forall t
\]  
(6.33)

where \( d_j \) are subsequent radiation doses, \( f_j \) is the efficiency rate of radiation dose \( d_j \), and \( \eta_j(t) \in [0,1] \) are binary variables defined for each radiation dose \( d_j \).

Note that \( M_v(t) = P_v(t) \cdot r_{\text{eff}}(t) \) becomes the product of the continuous and binary variables, which can be linearized using a traditional linearization method as follows:

\[
P_{\text{min}} \eta_{v,j}(t) \leq G_{v,j}(t) \leq P_{\text{max}} \eta_{v,j}(t)
\]  
(6.34)

\[
P_v(t) - P_{\text{max}}(1 - \eta_{v,j}(t)) \leq G_{v,j}(t) \leq P_v(t) - P_{\text{min}}(1 - \eta_{v,j}(t))
\]  
(6.35)

where \( G_{v,j}(t) \) is the substitute variable for \( P_v(t) \cdot r_{\text{eff}}(t) \), and \( P_{\text{min}} \) and \( P_{\text{max}} \) are the lower and upper bounds for \( P_v(t) \), respectively. Therefore, equation (6.25) can be approximated and linearized by utilizing equations (6.31) to (6.35).

6.2.2.3. Treatment Method 3: Surgery

Surgery is instantaneous and results in the resection of significant portion of the tumor structure. However, it is possible that some portion of the cells may remain undetected during X-ray surveillance or surgery, thus unremoved from the body. It is assumed that voxels at a
plateau population or inside the tumor margin can be detected and removed with 100% efficiency. On the other hand, a small percentage of cancer cells at the margin of the tumor structure may remain after surgery. Therefore, the number of resected cells at time $t$ is formulated as

$$M_{v,h}(t) = \begin{cases} P_v(t) & \text{if } P_v(t) = P_{\text{max}} \\ P_v(t) \cdot \psi & \text{o.w.} \end{cases}$$

where $\psi$ is the efficiency of surgery for the voxels at the margin.

6.3. Model Application

Table 6.1 shows the initial population structure and parameters of the model. The parameter values are collected by utilizing several resources including journal papers, online resources, and expert opinions.

Computational experiments are conducted for a stage-I breast cancer with a tumor size of 2 cm in diameter. It is assumed that the cancer has not spread outside the breast, and no lymph nodes are involved. In order to represent the 2-cm-diameter structure, a three-dimensional cuboid structure with dimensions of 5x10x5 mm layout is used in the experiment, as shown in Figure 6.4. The cuboid structure is divided into 250 equally shaped cubes (voxels) each with 8 mm$^3$ volume, where each voxel can contain 0 to 8x10$^6$ number of tumor cells. In the initial tumor population, while it is assumed that voxels in the inner site of the tumor structure are at their plateau population ($P_{\text{max}}$), the population of the voxels at the tumor margin, which is the border of the tumor tissue removed in cancer surgery, are randomly generated with uniform distribution $U[0, 8x10^6]$. Therefore, the initial cell population of the entire tumor is assumed to be a total of 7.44x10$^8$ cells at the beginning of the cancer treatment.
Depending on the available space, tumor cells can spread to neighbor cells. The tumor cell growth is assumed to be spherical and homogenous in a voxel. Once the cancer cells reach the facets in a given voxel, it is calculated that tumor cells occupy 52.3% (ξ) of the given voxel (see Appendix A). Once the tumor reaches the facets, 10.5% (λ) of the newly generated cancer cells migrate towards the immediate six neighbor voxels where the migration is limited by the number of voxels.

### TABLE 6.1. BREAST TUMOR AND NORMAL TISSUE PARAMETERS

<table>
<thead>
<tr>
<th>Structure</th>
<th>Model parameter</th>
<th>Symbol</th>
<th>Values</th>
<th>Units</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Doubling time of cancer cells</td>
<td>τ</td>
<td>14</td>
<td>days</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>Plateau population at each voxel</td>
<td>P_{max}</td>
<td>8x10^6</td>
<td>cells</td>
<td>[13]</td>
</tr>
<tr>
<td></td>
<td>Effectiveness rate of chemo drug</td>
<td>Φ</td>
<td>2.7x10^{-3}</td>
<td>1/days.C</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td>Spread rate of tumor</td>
<td>λ</td>
<td>0.105</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Minimum occupancy rate for a voxel before</td>
<td>ξ</td>
<td>0.523</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>cancer cells start to migrate to the</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>immediate neighbors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The maximum allowable drug concentration per</td>
<td>C_{max}</td>
<td>200</td>
<td>mg/m^2</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td>treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The maximum allowable cumulative drug</td>
<td>C_{cum}</td>
<td>16.9 x 10^3</td>
<td>mg</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td>concentration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decay rate of chemo concentration</td>
<td>Σ</td>
<td>0.27</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Prescribed total radiation dose</td>
<td>P_{D_{prescribed}}</td>
<td>50</td>
<td>Gy</td>
<td>[15]</td>
</tr>
<tr>
<td></td>
<td>Percentage of voxels treated with the</td>
<td>φ</td>
<td>80%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>prescribed dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum overdose for radiation treatment</td>
<td>D_{OD}</td>
<td>2</td>
<td>Gy</td>
<td>[15]</td>
</tr>
<tr>
<td></td>
<td>Maximum underdose for radiation treatment</td>
<td>D_{UD}</td>
<td>1.2</td>
<td>Gy</td>
<td>[15]</td>
</tr>
<tr>
<td></td>
<td>Patient related sensitivity</td>
<td>α</td>
<td>0.08</td>
<td>Gy^{-1}</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>Tumor related sensitivity</td>
<td>β</td>
<td>0.028</td>
<td>Gy^{-2}</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>Effectiveness rate of surgery at margin of</td>
<td>ψ</td>
<td>0.95</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>Initial tumor cell population</td>
<td>P_0</td>
<td>U[0, 8x10^6]</td>
<td>cells</td>
<td>-</td>
</tr>
<tr>
<td>tissue</td>
<td>Sensitivity ratio</td>
<td>α/β</td>
<td>5</td>
<td>Gy</td>
<td>[26]</td>
</tr>
<tr>
<td>Heart</td>
<td>Sparing factor</td>
<td>γ</td>
<td>0.083</td>
<td>-</td>
<td>[27]</td>
</tr>
<tr>
<td>tissue</td>
<td>Maximum allowable total irradiation</td>
<td>γ</td>
<td>4.29</td>
<td>Gy</td>
<td>[15]</td>
</tr>
<tr>
<td></td>
<td>Sensitivity ratio</td>
<td>α/β</td>
<td>3</td>
<td>Gy</td>
<td>[26]</td>
</tr>
<tr>
<td></td>
<td>Sparing factor</td>
<td>γ</td>
<td>0.0204</td>
<td>-</td>
<td>[27]</td>
</tr>
<tr>
<td></td>
<td>Maximum allowable total irradiation</td>
<td>γ</td>
<td>1.03</td>
<td>Gy</td>
<td>[15]</td>
</tr>
</tbody>
</table>
Although the solid structure of the tumor and the margin can be resected with surgery, there is a possibility that some cells may be undetected and left in the body during the surgery [28]. Therefore, it is assumed that while the voxels at the plateau population can be annihilated with 100% efficiency, cancer cells at the margins are resected with 98% efficiency. Hence, follow-up therapies, such as radiation therapy, chemotherapy, or both, are necessary to eradicate tumor cells in the breast area. Combinations of sequential radiotherapy and chemotherapy protocols to annihilate the tumor are analyzed after surgery, while the analysis of chemo-radiation therapy is left as a future work.

Figure 6.4. Three-dimensional representation of initial tumor population

For the radiation therapy, a standard radiation fractionation is applied for five consecutive days followed by a two-day rest period over a five-week period (25 days). A total of 50 Gy radiation is planned to be delivered over the course of radiation treatment. Furthermore, the maximum tolerance of organs-at-risk for the toxicity impact of radiation therapy is considered for the heart and lung using the BED equation (equation (6.13)). The total tolerated doses for heart and lung are 3 and 5 Gy, respectively, over the course of the radiation treatment.
In the case of chemotherapy treatment, a cyclophosphamide agent is used to reduce the risk of early-stage breast cancer recurrence. While the standard chemotherapy protocol for administering the drug consists of delivering a dose of 600 mg per square body-surface area (mg/m²) intravenously once every three weeks over a three-month period [29], it can also be administered once a week over a twelve-week period at a dose of 200 mg/m². Furthermore, by following the study of Sacco et al. [30], a body surface area of 1.7 m² is used to determine the weekly dose level. Therefore, the weekly drug concentration limit per treatment (C_max) is assumed to be 340 mg for a cancer patient. In addition, while the drug is administered once a week with dose D, the concentration of the dose in the body declines at a rate of e^{-0.27}. Note that a dose of 1020 (600x1.7) mg once in every three weeks for four cycles results in a higher acute toxicity than a dose of 340 (200 x 1.7) mg every week over 12 weeks. The reason is that while the 1020 mg drug concentration reduces to 4.2 mg at the end of 21 days, 340 mg reduces to 65.5 mg at the end of the first week. Since the drug concentration limit per day is 1020 mg and 340 mg, respectively, in these regimens, 1015.8 mg and 274.5 mg of chemotherapy agents will be administered in the following treatment, respectively, if full dose treatment is utilized. This will result in higher drug toxicity for the three-week regimen compared to one-week protocol. In the computational experiment, the performance of both two drug regimens is tested. By utilizing the 600 mg/m² dose protocol, acute toxicity C_cum can be calculated as

\[ C_{cum} = 4 \sum_{t=0}^{20} 600 \times 1.7 \times e^{-0.27t} \], which is 16,946.6 mg over a 12-week period.
6.4. **Numerical Results**

In this section, the proposed MIP models for the optimization of surgery, radiotherapy, and chemotherapy regimens for early breast cancer treatment are applied. The computational results for the optimal treatment sequence of these three treatment regimens in combination with the optimal radiation fractionations and drug dose protocols in order to minimize the new cancer cell population over the entire voxels and all time periods are presented. Furthermore, sensitivity analysis to evaluate optimal radiotherapy fractionation schedules for different radiation sensitivity values of the OARs including heart and lung is applied. In addition, hyper- and hypo-fractionation schedules are compared with each other in terms of radiation dose routines and resulting tumor cell populations. Finally, a sensitivity analysis on the effectiveness of the chemotherapy agent in chemotherapy treatment is performed. The three sequential models are solved by using CPLEX 12.6 in a high-performance computer running the Linux Operating System with a 32 CPU and 64.0 GB of memory. A time limitation of 7,200 CPU seconds is imposed for solving each treatment model. Selected results chosen from interesting problem configurations are reported here.

Table 6.2 presents details of the computational performance of the MIP models for surgery, radiotherapy and chemotherapy, separately for each treatment regimen, SRC, SCR, RC and CR. As shown in the first two rows of the computational results, applying surgery initially reduces the difficulty of the problem, and the MIP model for the treatment modality following surgery can be solved within given time limitation with no gap. This is an expected result because applying surgery leads to removal of the voxels inside the tumor margin, which reduces the size of the problem. Furthermore, without surgery, the initial treatment method
(radiotherapy or chemotherapy) becomes a difficult to solve problem within the time limitation. SRC is found to be the best treatment in terms of reducing the cancer population to zero in the shortest treatment regimen for the \((\alpha/\beta) = [3;5]\) case. Details of each regimen are discussed in the following sections.

**TABLE 6.2. SOLUTION TIME AND OPTIMALITY GAP FOR EACH TREATMENT MODALITY**

<table>
<thead>
<tr>
<th>Treatment regimen</th>
<th>Surgery</th>
<th>Radiation therapy</th>
<th>Chemotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC ((\alpha/\beta) = [3;5])</td>
<td>1 0 (4.8 \times 10^5)</td>
<td>5641 0 0 3 0 0</td>
<td>25</td>
</tr>
<tr>
<td>SCR ((\alpha/\beta) = [3;5])</td>
<td>1 0 (4.8 \times 10^5)</td>
<td>131 0 0 4377 0 0</td>
<td>85</td>
</tr>
<tr>
<td>RC, ((\alpha/\beta) = [3;5])</td>
<td>- - -</td>
<td>7600 29 0.18 25 0 -</td>
<td>110</td>
</tr>
<tr>
<td>RC, ((\alpha/\beta) = [12;20])</td>
<td>- - -</td>
<td>4212 0 0 12 0 -</td>
<td>11</td>
</tr>
<tr>
<td>CR ((\alpha/\beta) = [3;5])</td>
<td>- - -</td>
<td>227 0 1.12 7600 34 0</td>
<td>105</td>
</tr>
</tbody>
</table>

\(a\)-Clock time, \(b\)-\(|(\text{best integer solution}–\text{best lower bound})|/\text{best lower bound}\), \(c\)-Population after specific treatment, \(d\)-Number of treatment days to eradicate entire cancer population

### 6.4.1. Validation of Cancer Cell Growth: Comparison of Predicted PLA Values and Simulated Gompertz Growth

In this section, cancer growth predicted by the PLA equations, equations (6.2) to (6.5), is validated against the simulations obtained by the nonlinear Gompertz growth function [12]. Gompertz growth is simulated over a 360-day period using Microsoft Excel 2016 by taking into account the doubling time of cancer cells, plateau population of voxels, and cancer population at each period. ANOVA and linear regression methods are used at 5% significance level to compare the average cell population over a 360-day period and analyze how good the predicted values by the PLA method approximates the simulated cell population, respectively.

Table 6.3 presents the statistical analysis for the mean differences using the ANOVA test. The columns from left to right represent the source of variation in the data, sum of squares (SS), degrees of freedom, mean sum of squares (MS), calculated F statistic, p-value, and
F critical value. It can be seen from Table 6.3 that the p-value (0.943) is greater than the significance level (0.05). This implies failure to reject the null hypothesis in which the difference between the population means is zero. In other words, no statistically significant difference between the predicted and simulated values of cancer cell growth is detected.

**TABLE 6.3. ANOVA TEST FOR COMPARISON OF PLA EQUATIONS AND GOMPertz GROWTH**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>662294.5235</td>
<td>1</td>
<td>662294.5235</td>
<td>0.004987428</td>
<td>0.943718485</td>
<td>3.8544427</td>
</tr>
<tr>
<td>Within groups</td>
<td>95345228499</td>
<td>718</td>
<td>132792797.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>95345890794</td>
<td>719</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Furthermore, Table 6.4 presents linear regression results, which reveal the statistical relation between simulated Gompertz growth and the predicted values. The columns from left to right represent the source of interpretation, coefficients of the predicted variable and the intercept, t-statistics, p-value, and regression statistics. Simulated values represent the dependent variable, and predicted values represent the independent variable. In addition, the null hypothesis represents that predicted cell population does not explain the simulated cell population values, whereas the alternative hypothesis states that the predicted values explain the simulated values. Results shown in Table 6.4 demonstrate that the p-value is zero, thus rejecting the null hypothesis that there is no relation between the predicted and simulated cell population.
TABLE 6.4. LINEAR REGRESSION RESULTS OF GOMPERTZ GROWTH VS. PLA EQUATIONS

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>t Statistic</th>
<th>P-value</th>
<th>Regression statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>18.10106072</td>
<td>1.068496475</td>
<td>0.28601656</td>
<td></td>
</tr>
<tr>
<td>Predicted</td>
<td>1.005703271</td>
<td>812.5024932</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Multiple R</td>
<td></td>
<td></td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td>R Square</td>
<td></td>
<td></td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td>Adjusted R square</td>
<td></td>
<td></td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td></td>
<td></td>
<td>249.45</td>
<td></td>
</tr>
</tbody>
</table>

Based on these two different statistical tests, it can be concluded that PLA equations provide a cell growth that is statistically similar to the simulated Gompertz cell population. Finally, Figure 6.5 also provides a visual comparison for the fitted values of Gompertz growth and growth predicted by equations (6.2) to (6.5). Figure 6.5 shows that predicted values provide a very good approximation to the simulated cancer cell growth values.

![Figure 6.5. Comparison of predicted and simulated Gompertz growth. Note predicted growth function visually perfectly matches with simulated growth function.](image)

6.4.2. **Optimal Sequence of Surgery, Radiotherapy, and Chemotherapy**

This section analyzes the combinations of surgery, chemotherapy, and radiotherapy treatment sequences and compare these results in terms of dose regimes and number of treatment days to complete eradication of cancer population. Results show that implementing...
surgery first before chemotherapy and radiotherapy is more effective compared to having the surgery as a second or third treatment method. This is consistent with the idea that surgery provides an instantaneous treatment and as much cancer cell resection as possible [31]. On the other hand, as the effectiveness of surgery for the margin of the tumor is reduced from 98% to 82%, the total optimal radiation dose administration increases from 38 Gy to 44 Gy (by 6 Gy) over the course of the radiation treatment (Figure 6.6). Consistent with the increase in the radiation dose, radiation administration is given priority in the initial days to reduce the number of tumor cells immediately.

![Fig 6.6. Radiation fractionations over 25 days after surgery is performed with different effectiveness rates at tumor margin](image)

It is also shown that SRC provides better treatment effectiveness than SCR in terms of eradicating the cancer cells in a shorter time period. While cancer cells can be annihilated within 25 days radiation therapy (Figure 6.7(a)), it takes 85 days in chemotherapy treatment to completely kill all cancer cells (Figure 6.7(b)). It is shown that surgery followed by either chemotherapy or radiotherapy suffices to completely eradicate the stage-I breast cancer. This result is also consistent with the results of Kohandel et al. [31] showing that SRC results in better treatment efficiency than the SCR protocol as well as other protocols without surgery.
Figure 6.7: (a) Optimal radiation fractionation after surgery in SRC regimen; (b) Optimal chemotherapy drug concentration after surgery in SCR regimen.

It should be noted that surgery is the main treatment for the stage-I breast cancer. Furthermore, while radiation therapy is used after surgery, it is not a very common practice to utilize chemotherapy treatment for stage-I breast cancer before or right after surgery. Although it is not a common practice, we also compare RC and CR schedules without surgery because surgery may be avoided in some specific cases (e.g., age and comorbidity of the patient). Computational experiments show that both protocols eradicate cancer cells at the end of the treatment regimen. However, while RC regimen offers 47 radiation dose in combination with $2.1 \times 10^3$ acute toxicity, as shown in Figures 6.8(a) and 6.8(b), CR offers $14.5 \times 10^3$ acute toxicity followed by 35 Gy radiation, as shown in Figures 6.8(c) and 6.8(d). While chemotherapy annihilates cancer cells, radiation therapy has significantly less adverse effects on the body compared to chemotherapy in terms of poisoning the entire body [32]. Therefore, it is safe to claim that the RC regimen provides a more efficient treatment compared to the CR regimen because RC administers significantly less amount of chemotherapy drug compared to CR.
It should be noted that although a regular prescribed dose over the course of radiation treatment is 50 Gy, our results demonstrate that the optimal radiation dose administered should be limited to 47 Gy in the RC protocol, considering the sensitivity of the heart and lungs to radiation. Results imply that applying a full-dose radiation treatment would deteriorate the OARs at the cost of killing more cancer cells.
6.4.3. Alpha/Beta (\(\alpha/\beta\)) Ratios

Figure 6.9 presents the optimal radiation fractionation schedule for the RC treatment with different sensitivity values of \(\alpha/\beta\) ratios of the heart and lung. In this experiment, radiation fractionation for the base case sensitivity ratios of heart (\(\alpha/\beta = 3\)) and lung (\(\alpha/\beta = 5\)) is compared with the radiation fractionation for the ratios of \(\alpha/\beta = [12;20]\) (ratios for heart and lung, respectively). As shown in Figure 6.9(a), results demonstrate that radiation fractionation schedules for minimizing newly grown tumor cells results in a homogenous dose distribution over the radiation treatment for the base case sensitivity ratios. In addition, similar to results in the previous section, the total radiation dose administered is limited with 47 Gy over the course of the treatment. However, as the alpha/beta ratio increases \(\alpha/\beta = [12,20]\), consistent with the results of Badri et al. [15], it is observed that 50 Gy radiation is delivered homogenously within ten days of the treatment period with the hypo-fractionation schedule. Figure 6.9(b) demonstrates that hypo-fractionation with initially high doses reduces the cancer cell population faster than the hyper-fractionation schedule and annihilates the entire tumor structure. This is an intuitive result because as the \(\alpha/\beta\) increases, OARs become less sensitive (or more tolerant) to radiation, by equation (6.24), so OARs could be exposed to hypo-fractionation with higher doses.
6.4.4. Chemotherapy Treatment Efficiency and Dosage

Here, the optimal chemotherapy administration schedule and its impact on the number of cancer cells using the CR regimen is analyzed. Furthermore, sensitivity analysis is performed on the chemotherapy efficiency rates since cyclophosphamide (the chosen drug for the experiments) may have diverse impacts for different patients or cases. Finally, the effect of cyclophosphamide on the cancer population is analyzed by applying a 200 mg/m$^2$ weekly dose for 12 cycles and 600 mg/m$^2$ once every three weeks for four cycles.

Results indicate that a full-dose treatment (200 mg/m$^2$/week) is not necessary unless chemotherapy is applied as an initial treatment. The optimal dose regimen depends on the number of cancer cells in the tumor structure as well as efficiency of the drug concentration. Results reveal that the more efficient the drug, the shorter time it takes to eradicate the tumor cell population (Figure 6.10). Furthermore, it is shown that the cyclophosphamide agent is administered homogeneously and in equal amounts at each cycle over the course of chemotherapy treatment, irrespective of the initial number of cancer cells at the beginning of
chemotherapy treatment, as shown in Figure 6.8(b) and 6.8(c). This is an intuitive result because consistent and homogeneous drug administration can lead to a significant reduction in cancer population, compared to administering the same total dose less frequently at high doses [33]. Although the maximum allowable drug concentration can be administered at each cycle in real life, it is also important to determine the minimum total drug administration amount that would lead to the same final result, considering the adverse effects of chemotherapy drugs. Finally, consistent with the aforementioned results, Figure 6.11 demonstrates that, while administering cyclophosphamide in four cycles once every three weeks and in 12 cycles every week results in the eradication of cancer cells, the latter protocol results in less cumulative toxicity compared to the former method. Furthermore, the regimen of once every three weeks reduces the number of cancer cells faster than the other regimen.

![Figure 6.10: Cancer cell population change with respect to different chemotherapy drug efficiencies](image)
6.5. Discussion

This chapter presents three spatio-temporal MIP models, each model corresponding to surgery, chemotherapy, and radiation treatment regimens, respectively. In particular, the effects of various combinations of surgery, chemotherapy, and radiation therapy sequences on breast cancer treatment are studied using these three models. Each model incorporates tumor growth dynamics, migration of cancer cells among voxels, as well as the treatment impacts (e.g., impacts of radiation fractionation and chemotherapy drug administration with optimal dose schedules) in order to minimize the newly generated cancer cells and the tumor size. Furthermore, each mathematical model tracks the spatial and temporal evolution of the tumor by taking into account the cell population in a set of voxels over time to predict the growth of cancer. Several quadratic and nonlinear equations that have been proposed in the cancer treatment literature are linearized or approximated with linear functions and incorporated into the MIP model. Computational experiments are performed to analyze the optimal treatment
sequence and schedule using a realistically scaled spatial stage-I tumor structure on a three-dimensional system for the breast cancer case.

This study appears to be the first of its kind to fill the research gap in the literature by providing an MIP model to optimize the sequencing of surgery, chemotherapy, and radiotherapy as well as dose schedules in a spatially temporal tumor growth model. Furthermore, although surgery is incorporated into several cancer treatment models, this is the first study that differs efficiency of resection of the tumor margin and the inner structure inside the margin. Finally, the model tracks the spatial and temporal evolution of the tumor and considers the cell population in a set of voxels to predict spread and growth of cancer.

The proposed models use a mathematical approach supported by the cancer treatment literature. It is difficult to validate the predicted cancer cell population with or without treatment against clinical data using traditional statistical analysis methods. On the other hand, although the validity of the predicted number of cells cannot be compared with clinical data due to the lack of such information, the outcomes of the model, such as radiation fractionations, chemotherapy drug doses, and schedules, are compared with the results of studies in the literature.

Results indicate that PLA equations provide a good fit for the simulated Gompertz growth. This can help practitioners provide an optimal surveillance schedule based on the expected number of cancer cell populations after treatment. Results are also consistent with the cancer treatment literature in terms of providing similar treatment sequences and schedules for surgery, radiotherapy, and chemotherapy treatments. Results indicate that although SRC and SCR sequences provide the same results in terms of cancer population after
treatment, SRC provides an efficient treatment in terms of treatment length. This is consistent with the study of Beil and Wein [34], who claim that the SRC schedule is better than the SCR schedule when the patient suffers from early breast cancer. Furthermore, Miller et al. [35] also indicate that SRC and SR are practical applications for early breast cancer treatment and 51% of the stage-I breast cancer patients are treated using SRC or SR sequences. On the other hand, there is not a practical application for the SC schedule for stage-I breast cancer patients. In addition, while sequencing and timing of the treatments depend on the patient’s health condition, it is a common practice to have a lag time between treatments in order to allow healthy tissues to recover from the adverse impacts of treatment [19]. Contrary to this idea, Sertoli et al. [36] argue that delays between surgery and chemotherapy as much as four weeks do not have any significant impact on the control of the cancer population. Furthermore, De Ruysscher et al. [37] also report that the survival rate of the cancer patient increases as the lag time between chemotherapy and radiotherapy treatments is reduced. Therefore, lag time between treatment types is ignored in the proposed model.

Moreover, consistent with the study of Badri et al. [15], it is demonstrated that hypofractionated schedules provide better treatment than hyper-fractionation in terms of minimizing the cancer cell population at the end of the treatment. Different than the study of Badri et al. [15], the results from this study show that the prescribed 50 Gy dose is not fully administered by using the base case value of $\alpha / \beta$ ratio provided in Table 6.1 and is limited to 47 Gy over the course of radiation treatment. However, as the sensitivity ratios for heart and lung are increased, it is shown that hypofractionation is utilized with administering the prescribed dose in ten days. Results also reveal that chemotherapy drug administration
depends on the cancer cell population at the beginning of the treatment. This implies that drug concentration may vary over the course of treatment.

On the other hand, it can be argued that a full dose (340 mg) of cyclophosphamide can be administered each week for the case represented in Figure 6.7(b) to increase the chance of tumor cell annihilation, considering the stochastic biological characteristics of cancer as well as patient-related factors. While full-dose administration can be assumed by the time of drug administration [12], this may lead to unnecessary higher acute toxicity, which would be detrimental to the health of the patient. Therefore, regulating the drug concentration to maintain the white blood cell count while minimizing the size of the tumor is necessary in resolving chemotherapy treatment problems [38]. Furthermore, while administering cyclophosphamide both in four cycles every three weeks and in 12 cycles every week results in the eradication of cancer cells, the latter protocol results in less cumulative toxicity compared to the former method. Consistent with practical application, the computational results demonstrate that the schedule of once every three weeks reduces the number of cancer cells faster than an every-week regimen.

6.6. Conclusion and Future Work

This chapter explicitly formulates tumor growth and spread dynamics in a spatio-temporal MIP model that provides the optimal sequence of treatment modalities over treatment period. The spatial dimension assists in predicting the spread and regrowth of cancer cells during and after treatment, whereby the size and extent of the growth and spatial spread of cancer cells can be tracked. It was also demonstrated that linearizing Gompertz growth
function using a piecewise approximation approach results in a statistically similar growth pattern compared to the simulations using the non-linear Gompertz growth function.

The model here contributes to the cancer literature by proposing a new spatio-temporal MIP model for breast cancer treatment by combining surgery, chemotherapy, and radiotherapy modalities. However, the proposed model also has also some limitations. For example, both the tumor growth and spread are extremely complex processes that involve many patient- and tumor-specific factors, such as HER2 status, involvement of progesterone and receptors, and stage of the cancer [39-41]. Furthermore, using approximation methods can impact the robustness of the radiation therapy model compared to existing studies that use complex differential equations or quadratic programming models. For example, results demonstrate that the total radiation dose administered should be limited to 47 Gy over the course of the treatment while existing models allow administration of 50 Gy dose for the same case study values [15]. One possible explanation for restricting radiation dose to 47 Gy is the restriction of the hyper-fractionation radiation doses to integer values such as 0, 1, 2, and 3 Gy at each treatment in the MIP model in order to reduce the problem complexity stemming from equations (6.31) to (6.35), while fractional radiation doses any time during the treatment period (e.g., 2.5 Gy) could be administered in practice. Therefore, future work will focus on increasing the number of break radiation points including fractional values for equations (6.31) to 6.35) to provide a better approximation to equation (6.25).

As another future direction, the stochastic nature of growth and spread dynamics of breast cancer could be incorporated into the mathematical models. Furthermore, in radiation therapy treatment, beam angles and the impact of beams on the three-dimensional tumor
structure could be integrated into the model. Moreover, one of the objectives of chemotherapy is to administer a minimal amount of drug concentration while providing an effective treatment. Therefore, minimization the drug concentration as well as the cancer population could be handled in a multi-objective model. Incorporating any of the aforementioned future work will increase the complexity of the model. Therefore, specialized algorithms, cutting planes, or heuristics should be developed to increase the solvability of the complex MIP cancer treatment models.

6.7. References


6.8. Appendix

6.8.1 Parameter Estimation: Threshold Rate for Cell Migration

The volume of a sphere is calculated as \( \frac{4}{3} \pi r^3 \), where \( \pi \) is 3.14, and \( r \) is the radius of the sphere. Furthermore, the volume of a cube is calculated as \( a^3 \), where \( a \) is the length of an edge.

Let Figure 6.A1 represent the growth of cancer cells in a voxel where growth is assumed to occur radially from the center of the voxel (cube) and the length of an edge of the cube is 2 units. Therefore, considering spherical radial growth from the center of the voxel, the tumor will touch the faces of the voxel when the radius of the tumor equals 1. Therefore, the tumor is assumed to touch all six faces of the voxel when 100\( \xi \) percentage of the voxel is occupied, where \( \xi = \frac{3}{a^3} \) and given that \( r = 1 \), \( a = 2 \) and \( \pi = 3.14 \), \( \xi = 0.523 \). Therefore, it is assumed that cancer spread (e.g., radial growth beyond its initial voxel) will occur after 52.3% of the voxel is occupied.

Furthermore, as the tumor grows and reaches the faces of the voxel, a total of 100 \( \lambda \) percentage of the tumor will migrate to the immediate neighbor voxel (there are six immediate voxels), as shown in Figure 6.A2.
Therefore, \( \lambda \) the percentage of migration to an immediate neighbor voxel per a newly generated cancer cell after cancer cells occupy 52.3% of the host voxel is formulated as

\[
\lambda = \frac{\pi}{3} h^2 (3r' - h) \quad \frac{4}{3} \pi \left( r'^3 - r^3 \right)
\]

where \( h \) is the height of the spherical cap, and \( r' \) is the radius of the tumor after growing out of the voxel. For example, let the new radius be \( r' = 1.5 \) units with a height of \( h = 0.5 \) units of the spherical cap. Therefore, \( \lambda \) is calculated as 0.104. However, different \( \lambda \) values can be reached as \( r' \) changes from 1 to 1.73, which is the maximum possible value \( r' \) value (1.73 is the radius of the outer sphere for the cube with 2 unit edges). Therefore, after calculating \( \lambda \) with different \( r' \) values, the average spread rate is calculated as \( \lambda = 0.105 \).

### 6.8.2 Parameter Estimation: Spread Rate

Let the initial cancer cell population be \( P_v(0) = 8 \times 10^6 \) in the initial period in a given voxel. Furthermore, the maximum chemotherapy drug level per treatment is 340 mg and it decays exponentially with a constant rate \( e^{-\varsigma} \) after the drug is administered in a given period. Initially,
equation (6.18) is simulated using Microsoft Excel for one voxel assuming that full dose chemotherapy is applied over the course of treatment with 340 mg at each treatment.

\[ M_{v,h}(t) = C_{\text{eff}}(t)P_v(t)\phi \] (6.18)

Based on the simulation using equation (6.18), the population change of the given voxel for one-week period is shown in Table 6.B1.

<table>
<thead>
<tr>
<th>Period</th>
<th>Drug Concentration ((C_{\text{eff}}(t)))</th>
<th>Population at Beginning of (t) ((P_v(t)))</th>
<th>Efficiency ((\phi))</th>
<th>Population Killed at End of (t) ((M_v(t)))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>340</td>
<td>8000000</td>
<td>0.0015</td>
<td>4080000</td>
</tr>
<tr>
<td>2</td>
<td>258.4</td>
<td>3920000</td>
<td>0.0015</td>
<td>1519392</td>
</tr>
<tr>
<td>3</td>
<td>196.384</td>
<td>2400608</td>
<td>0.0015</td>
<td>707161.5022</td>
</tr>
<tr>
<td>4</td>
<td>149.25184</td>
<td>1693446.498</td>
<td>0.0015</td>
<td>379125.0086</td>
</tr>
<tr>
<td>5</td>
<td>113.431984</td>
<td>1314321.489</td>
<td>0.0015</td>
<td>223627.9867</td>
</tr>
<tr>
<td>6</td>
<td>86.20786278</td>
<td>1090693.502</td>
<td>0.0015</td>
<td>141039.5337</td>
</tr>
<tr>
<td>7</td>
<td>65.51797572</td>
<td>949653.9688</td>
<td>0.0015</td>
<td>93329.1085</td>
</tr>
</tbody>
</table>

Once the annihilated population is calculated for each period for the treatment period \(t\) \((M_v(t))\), the starting population level for the following period \(t+1\) is calculated by \(P_v(t+1) = P_v(t) - M_v(t)\). After simulating number of annihilated cells over the course of chemotherapy treatment, the efficiency rate for each period is calculated by the following equation:

\[ \hat{\phi}(t) = \frac{M_v(t)}{C_{\text{eff}}(t)P_v(0)} \]

Therefore, using the number of annihilated cells and drug concentration at each period, there is a vector of efficiency rates for the chemotherapy treatment \(\hat{\phi}(t)\) for each \(t\). Note that \(\hat{\phi}(t)\) is a strictly decreasing function of time.
CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

7.1. Summary of Contributions

In this dissertation, a collection of new mathematical models and solution approaches to find the optimal strategies for ISM and cancer treatment were developed.

Chapter 2, provided a review of articles in the literature regarding invasive pest and weed management from control and surveillance perspectives. The articles were clustered in terms of their objectives, application areas, and solution methodologies. Chapter 3 presented an NLP model to minimize the economic damages of the invasive species where the model incorporates different seed production and loss rates while dividing the population into different age classes and tracking the growth of each age class over a multi-period time horizon. Chapter 4 proposed an MIP model by considering the biological characteristics of invasive species and dispersal uncertainty to neighboring sites. This model outperforms the nonlinear model presented in the Chapter 3 in terms of optimal solution time. Chapter 5 presented a multi-stage stochastic optimization model along with unique cutting planes to find the optimal surveillance, treatment, and removal decision for all possible scenarios of invasion growth. Finally, Chapter 6, presented an MIP model for breast cancer treatment by combining surgery, radiation therapy, and chemotherapy to determine an optimal treatment scheme in order to minimize the size of the tumor.

One of the most important contributions of this dissertation regarding its applications is the utilization of seed dispersal dynamics of invasive species and incorporating biological characteristics into an optimization model. It is believed that this is the first study incorporating
carrying capacity of a landscape, stochastic dispersal, and age structure of the plants in the same optimization model. Furthermore, the research gap in the literature will be filled by integrating surveillance, treatment, and removal as separate decisions into an optimization model. Therefore, this dissertation helps decision makers to optimally allocate resources at each stage by considering the current and forecasted damages of invasion. In the cancer study, the best treatment strategy was determined by combining the surgery, radiation therapy, and chemotherapy. It is also believed that this is the first spatio-temporal MIP model developed for an integrated cancer treatment process. Finally, in order to provide better insights for practitioners, the computational experiments and sensitivity analysis are presented in the results sections of each chapter.

In terms of theoretical contributions, the nonlinear invasive species model is linearized by proposing a linearization technique that can be generalized to other studies. The model obtained by the proposed linearization method provides higher-quality solutions compared to four other linearization methods in the literature. While the linearization methods applied to the model here exist in the literature, it appears that they have never been considered to enhance the computational solvability of such a model in biological conservation. Furthermore, previous computational analysis on the impacts of the big-M value is limited. This is the first study that elaborates on the use of linearization in a biologically complex model and provides an extensive analysis of the value of big-M. The detailed discussion on the computational impact of the big-M value could also offer insight into many other applications in spatial conservation and environmental resource allocation.
Furthermore, in the EAB study, a new MSS-MIP model and unique cutting planes to solve the proposed model are developed, and they can also be applied to other similar MSS-MIP models. The unique cutting planes improve the solution time for three-stage instances by an average of up to 44 times compared to solving the original MSS-MIP model. The computational experiments on the comparison of the original MSS-MIP model and the solution algorithm reveal that the proposed algorithm is effective to solve the proposed model. In particular, while the original model cannot be solved for a 6-by-6 gridded landscape in a five-year time period, the algorithm enables the model to be solved for a 10-by-10 gridded landscape in a five-year time period with an ample budget allocation.

Finally, the proposed model fills the research gap in the cancer treatment literature by explicitly formulating tumor growth and spread dynamics in a spatio-temporal MIP model that provides the optimal sequence of treatment modalities over the cancer-treatment period. Linearization and approximation methods are used to linearize the nonlinear and highly quadratic equations. The MIP model improves the solvability of the complex cancer treatment modality problem, while also considering the highly complex spatio-temporal cancer growth. The proposed MIP model demonstrates that hypo-fractionation with initially high radiation doses would provide more effective results compared to hyper-fractionation in terms of annihilating cancer cells in a shorter time period. Furthermore, a full-dose chemotherapy treatment with maximum allowable drug concentration at each drug administration may not be necessary unless chemotherapy is applied as an initial treatment before surgery and radiotherapy. Additionally, an every-week chemotherapy schedule for 12 cycles leads to both less cumulative toxicity and faster cancer cell reduction compared to once-every-three-weeks
schedule for four cycles. Finally, linearizing the Gompertz growth function using a piecewise approximation approach results in a statistically similar growth pattern compared to simulations using the non-linear Gompertz growth function.

7.2. Future Research Directions

In the future, there are several directions in which each chapter of this dissertation can be further developed and extended. In its current form, this study provides resource allocation and treatment dose determination strategies for the treatment of invasive species and cancer. The models could be further extended by their scope, application, and methodologies.

One possible extension for the application area of the invasive species problem is to extend it to other species, such as fish, insects, mammals, and plants, for which age structure is relevant. Furthermore, age-structured growth and carrying capacity can be adapted to model the growth of stage- or size-structured species and estimate the population abundances of different stage and size groups given carrying capacity limitations, respectively. In addition, seed generation- and seedbank-based growth dynamics can be adapted to model dormancy and various offspring generation, accumulation, and dispersal mechanisms. Invasive species surveillance and control problem can be jointly studied with the biofuel production problem that involves non-native biomass that could pose a risk of invasion [1-3]. Finally, the proposed cancer treatment model can be adapted to solve other cancer treatment-planning problems such as for ovary, lung, brain cancer, etc.

The scope of the invasive species studies can also be improved by considering the uncertainty in growth and herbicide-effectiveness levels. Furthermore, the direction and intensity of wind could have a very significant impact on the seed dispersal rate and distance.
Instead of using a radial dispersal mechanism, the impact of wind speed and direction could be modeled stochastically. Although surveillance is utilized to reduce the risk of invasion, preventive decisions such as trapping methods and education of people could be integrated into the model to minimize risks associated with invasiveness.

The cancer study could be extended by incorporating beams of the LINAC and beam angles. Furthermore, the impact of radiation and chemotherapy on the three-dimensional shrinkage of tumor population could be incorporated into the model. In addition, impacts of multi-drug combinations on chemotherapy treatment could be analyzed in the model. Moreover, instead of providing sequential treatment, chemotherapy and radiation therapy model could be combined to incorporate chemo-radiation therapy into the model. Furthermore, a future study could consider costs of different treatment modalities.

In order to improve the solvability of the optimization models, additional parameters and heuristics within CPLEX could be used to reduce the solution time at the expense of potentially reducing the solution quality. As a result, further research would focus on exact solution algorithms such as decomposition and cutting-plane techniques (e.g., [4, 5]) in order to increase the solvability of the proposed ISM and cancer treatment models and reduce the solution time for instances that involve higher spatial and temporal dimensions.

7.3. References


