A learning heuristic for integrating spatial and temporal detail in forest planning

Eric B. Henderson1 | Howard M. Hoganson2

1Ecosystem Assessment and Planning, USDA Forest Service Northern Region, Missoula, Montana, USA
2Department of Forest Resources, North Central Research and Outreach Center, University of Minnesota, Grand Rapids, Minnesota, USA

Abstract
We present a learning heuristic using dynamic programming (DP) formulations to address both spatial and temporal detail in multiobjective forest management planning. The problem is decomposed into smaller problems to avoid the curse of dimensionality associated with DP. The heuristic learns from multiple decomposed problem formulations to identify stands assigned the same management option regardless of formulation. Consistently managed stands are recognized, and the problem is eventually distilled to the most difficult to solve portions of the forest. The heuristic is demonstrated on a forest management problem with short-lived core area wildlife habitat, where temporal detail is important. The problem is large due to the number of stand-level management timing options and associated interactions with the management timings of nearby stands. Results show solution improvement along with substantial time savings over previously used heuristics. The learning heuristic enables analysis of large problems that emerge with high levels of spatial and temporal detail.

Recommendations for Resource Managers:
Adding spatial and temporal detail to forest planning models can improve the utility of the solution for some applications such as wildlife habitat management.

- Decomposing a large optimization problem into linked, smaller subproblems can speed the solution search time.
- Learning from problem decomposition solutions to formulate subsequent search attempts can improve the quality of a solution.

**KEYWORDS**
core area, decomposition, dynamic programming, forest management, harvest scheduling, Kirtland’s Warbler, meta-heuristics, optimization

### 1 | INTRODUCTION

Forest management analyses have historically used coarse spatial depictions of a landscape coupled with aggregate time periods to derive strategic management plans for large forests (Johnson & Scheurman, 1977). Since then, advances in both computing technology and solution techniques have allowed for linkages between strategic and tactical-level modeling. Explicit spatial and temporal detail into forest planning models aimed at addressing tactical decisions is becoming more common. However, the resulting problems are often too complicated to fit into an exact optimization model structure of practical size. Metaheuristics are general algorithms for addressing complicated optimization problems. In forestry, metaheuristic applications include genetic algorithms, simulated annealing, and tabu search (Kangas et al., 2015).

As noted in a review of metaheuristics (Blum et al., 2011), emphasis has shifted to metaheuristics that are linked closely with optimization techniques with a more specific problem-oriented approach rather than algorithm oriented. Pellerin et al. (2019) describe this shift in a survey of hybrid metaheuristics for resource-constrained project scheduling problems. Mohamed-Mahmoud et al. (2015) in reviewing metaheuristics for agricultural land use optimization, conclude that success of meta-heuristics is problem dependent with need to escape from local optima. They suggest a future challenge is the use of parallelization techniques with hybrid metaheuristic techniques and involving stakeholders to help find key trade-offs between competing land uses. For forestry applications, Kangas et al. (2015, pp. 167–189) provide a recent review of heuristic optimization including meta-heuristics that have been linked with optimization models.

Temporal and spatial resolution of a planning problem affects its complexity, and therefore each is often considered as a possible simplification. Andersson and Eriksson (2007), in looking at impacts of temporal scale on long-term strategic planning for timber production, found that 5-year planning periods were generally adequate for representing financial returns from timber production. Forsell and Eriksson (2014) found that results were substantially more sensitive to temporal aggregation when also considering impacts of windfall damage. In a review of spatial modeling, DePellegrin Llorente et al. (2017) describe the importance of spatial and temporal interactions with stand-level harvest timing decisions which are also complicated by market...
destination options for multiple forest products. Bixby et al. (2019) address the complications of coordinating spatial planning across multiple forest ownerships. Then, not only are larger areas (forests) involved, but coordination is likely needed regarding time-sensitive value assumptions about spatially dependent nonmarket forest conditions.

Temporal and spatial simplification in planning problems is not always desirable, however. Rau et al. (2019) reviewed 295 studies on ecosystem services with only 2% of those studies considering changes in ecosystem services over time. They emphasize that future studies should: (1) more explicitly study temporal patterns including fine grain temporal patterns, (2) analyze trade-offs and synergies between services over time, and (3) integrate changes in supply and demand and involve and empower stakeholders in temporal ecosystem services research. Multiple ecosystem services are also typically involved making situations extremely complex.

Belavenutti et al. (2019) propose two goal programming model formulations for focusing on important time scale differences for better linking strategic and tactical forest management planning models. The modular nature of their formulation structure likely offers substantial flexibility for addressing a wide range of forestry problems. As the authors point out, the best approach for linking strategic and tactical models is problem dependent with improvements plausible with linkages to meta-heuristics to help overcome simplifying assumptions.

Dynamic programming (DP) is a problem structure from operations research that has been useful in solving some forestry problems. DP decomposes problems into a series of smaller, linked problems (Bellman, 1954). Hoganson et al. (2008) describe some of the history of DP in forestry. They outline early applications such as stand-level thinning. Applications of DP have involved solving problems for forest-wide spatial constraints such as adjacency (Hoganson & Borges, 1998), and patches of older forest core area (Hoganson et al., 2004; Wei & Hoganson, 2008). Core area forest is free from edge effects, assuming a certain buffer distance from an edge (Baskent & Keles, 2005; Ohman & Eriksson, 1998; Ohman, 2000). In another example, Wei and Hoganson (2005) set the shadow price of core area exogenously and allowed the DP to determine core area levels achieved at several prices.

A notable challenge with forestry-based DP formulations is they grow rapidly in size with moderate increases in the number of management options recognized for each stand, also known as the “curse of dimensionality” (Bellman, 1954). To address DP size concerns, Hoganson and Borges (1998) used a heuristic solution strategy involving a series of overlapping subproblems, or “moving windows,” to solve the adjacency problem for a forest comprised of uniform square stands. Borges et al. (1999) explored window design and DP formulation strategies with irregular polygons. Later, Hoganson et al. (2004) modified the heuristic to address core area management by using the concept of influence zones, or areas of the forest that are dependent on coordinated management of one or more stands to help achieve spatial objectives. The heuristic was used operationally as part of the USDA Forest Service planning process for the two National Forests in Minnesota (USDA Forest Service, 2004) involving ten planning periods and over 90,000 analysis areas. Since the size of the windows can influence both solution speed and solution quality, Wei and Hoganson (2008) investigated parameter settings of the heuristic solution strategy with the goal of increasing solution speed with minimal effects on solution quality. Their main tests specific to window design involved using different window widths, and different window overlap percentages. Generally, they showed that while smaller window sizes resulted in shorter solution times, they also yielded inferior solution values. To date, there have been few or no studies that have explored learning from common stand-level solutions between different window formulations.
2 | OBJECTIVES

The study objective is to develop a management strategy that includes tactical decisions associated with enhanced spatial and temporal consideration. The strategy is evaluated with a learning-based dynamic programming solution heuristic designed to quickly solve a spatial forest management problem with minimal loss in optimality. Interest in developing a better heuristic was rooted in a USDA Forest Service wildlife management problem (Sjogren, 2012) that sets explicit targets for the amount of Kirtland’s warbler (KW; *Setophaga kirtlandii*) core area habitat desired in each of 30 two-year planning periods (60 years total). An increased number of planning periods results in more management timing options per stand which can quickly complicate the problem with the curse of dimensionality. The method presented in this study is nested within a broader solution method that assumes explicit constraints on the amount of core area desired in each planning period. Here, we assume the value of core area is known in each planning period. In practice, the value of core area is determined through an iterative search that requires hundreds or thousands of DP solutions to search for shadow prices associated with the core-area constraints (Hoganson & Rose, 1984; Paredes & Brodie, 1988). Thus, it is advantageous to develop heuristics and exploit simplifications that allow the DP to be solved both quickly and accurately.

3 | METHODS

3.1 | Study problem

The data for the study comes from a forest and wildlife management problem faced by managers on the Hiawatha National Forest (HNF) in Michigan. The Forest’s land management plan has an objective to create and maintain 2711 hectares of suitable KW habitat consisting of young jack pine (*Pinus banksiana*) 6–16 years old (USDA Forest Service, 2006). This equates to a total of 13,557 hectares in a KW jack pine habitat system assuming this habitat area is regulated based on a 50-year rotation and 20% of this area between ages 6 and 16 years. Potential areas for KW habitat (sandy glacial outwash plains) on the HNF consist of approximately 70,620 hectares defined by 12,307 stands. The specific stands (13,557 ha) managed as suitable breeding habitat within the potential area have not been explicitly identified by the forest plan. Therefore, the forest management team has discretion in where it creates habitat and has the opportunity to design a management schedule that is both effective as habitat and financially efficient. The ideal spatial arrangement for KW habitat is in large, contiguous patches (Donner et al., 2010). Financial efficiency is measured by the present net value of the management strategy. It is important to coordinate management timing and spatial arrangement of habitat since it only exists for 10 years in young stands.

The planning horizon was 60 years, modeled as a series of 30 2-year time periods. Sixty years accommodates the average jack pine rotation length of 50 years and allows stands to convert to KW habitat later in the planning horizon for better spatial arrangement. Two-year time periods were chosen to address important detail in temporal fluctuations and the relatively short-lived nature of the suitable habitat (10 years). From a wildlife management perspective, a consistent abundance of habitat in each year is beneficial to wildlife recovery (Michigan Department of Natural Resources, 2015). Using short planning periods allows the model to evaluate the importance of different management timing decisions between adjacent stands in
consecutive planning periods. Habitat patches consisting of aggregated stands might slowly “walk” or shift on the landscape over time if adjacent timing options are staggered. Alternately, contiguous patches might emerge and desist at discrete times. With 2-year planning periods, each stand managed for KW produces habitat in five, 2-year planning periods. If adjacent stands are managed for KW habitat at similar times, each stand could be part of up to five different habitat patches over its lifetime if the habitat timing of those adjacent stands is staggered.

In addition to KW habitat, the problem was formulated to meet other desired conditions of the Hiawatha Forest Plan (USDA Forest Service, 2006), including covertype and size class minimum and maximum amounts. The model included management options resulting in KW habitat, as well as options to convert existing KW habitat to other covertypes. For the 12,307 stands recognized in the study, there were a total of 1.08 million management options considered, or an average of 88 per stand.

Rather than an explicit constraint on total KW habitat, the model specified desired amounts of habitat in KW core area each period, hereafter referred to as the spatial constraints of the problem. Core area is defined as the portion of a habitat patch sufficiently far from the nearest edge of a forest. Typically, edges of a patch are buffered inward to recognize the effect of the edge, and the portion of the patch not in the buffer, that is, far enough from the edge, is core area. In the case of KW habitat, an edge of larger trees is undesirable, as it allows the parasitic cowbird to more easily identify nest locations (Michigan Department of Natural Resources, 2015). To model core area, the landscape was simplified into a grid of 0.8 ha regular hexagons. Hexagons were then aggregated into stands that best represented the irregular polygons of the forest’s stand layer. Cells in a hexagonal grid have regular spatial interactions with all adjacent cells, which is not the case for stand designs such as squares or irregular polygons (Heinonen et al., 2007). The core area concept as it relates to hexagons is displayed in Figure 1. To produce core area, a cluster of at least three hexagons must meet the overall definition of KW habitat, meaning that core area exists approximately 56 m or more from the nearest edge. In Figure 1b, each triangle represents an area of potential core area production. Core area can be produced if

![Figure 1](image_url)
and only if the three hexagons in which it exists (e.g., stands \{1,2,3\} for triangle on the left) are managed as habitat in a specific time period. Efficiency can be gained if larger contiguous areas are managed as habitat. In the Figure 1 example, if three stands \{1,2,3\} are habitat in period 1, simply managing \textit{one more} stand \{4\} as KW habitat in period 1 would double the core area. Thus, the best solutions to problems that value core area consist of large, contiguous patches of adjacent stands in a compact arrangement.

The general DP formulation in this study is presented in Hoganson et al. (2004) and in Wei and Hoganson (2008). As a simplified example, consider four single-hexagon stands illustrated in Figure 1a. For each stand, assume there are two management options, \textit{a} (produce habitat in period 1) and \textit{b} (produce habitat in period 2). If core area produced in period 1 has a value of 2 and a unit of core area produced in period 2 has a value of 1, maximum value will result from managing all stands with option \textit{a}. The bottom portion of the figure (1b) is included to show a DP formulation of the problem, with the Stand 1 management decision as the “stage 1” decision, which is likely one of the simpler formulation options. Each stage of the DP represents the management decision for a single stand and each state at the start of a stage is defined by management decisions implied for stands associated with earlier stages in the network (e.g., state 1a is stand 1 option \textit{a}). These decisions must be remembered in the formulation so that core area conditions involving neighboring stands can be assessed (valued) in later stages of the network when decisions for all stands defining a triangle are implied. A stand can be dropped as a state descriptor only after all of its potential core area has been evaluated in the network. From Figure 1 it should be apparent how the number of states in each stage will increase if more management options are considered per stand. The DP network also grows substantially in size if the width of the area modeled is increased, resulting in more stands included as state variables at most stages of the formulation.

The problem in this study is too large to solve exactly with a single DP formulation. Therefore, we rely heavily on a moving windows DP heuristic first described by Hoganson and Borges (1998). A window is a subproblem consisting of a strip of forest stands spanning either the length or the width of the forest. The moving window heuristic utilizes three main parameters to define each subproblem: size, direction, and overlap. The size measures how many stands and resulting DP states span the window along the narrow axis. Direction describes the span of the forest in the long axis as well as the end from which the formulation starts. Finally, the overlap defines how many stands from the prior subproblem are included in the next subproblem. Window subproblems progress across the forest perpendicular to their direction. At each subproblem formulation, all of the stands behind the progression are considered as if their solutions are known and can be included without increasing the DP size. Window subproblem progression continues until all the stands in the forest have been evaluated.

### 3.2 A learning heuristic

This study seeks to quickly solve a large DP without compromising solution quality. Again, the DP is useful for this problem because it considers core area that results from coordinated management of adjacent stands. The study builds on work by Wei and Hoganson (2008) that showed smaller window subproblem formulations yield faster solutions, but often have inferior solution (objective function) values. Their study held direction constant, while incrementally increasing size and evaluating the trade-offs in the resulting objective
function value and solution time. Here, we evaluate similarities between solutions of many different problem formulation “trials” to inform subsequent formulations. In this way, the heuristic learns from known solutions to inform subsequent problem formulations with the intent of deriving a superior solution in less time.

The parameters of the learning heuristic include the window subproblem size and direction as well as the number of trial solutions that must be evaluated before simplifying the problem. A particular parameter set is called a “trial” and its solution is comprised of stand-level solutions, that is, chosen management options. The heuristic saves the solution to each trial to look for common stand-level management options between multiple trials. If a stand’s chosen management option is consistent between several trials, it supposed that global problem may be simplified by using only that management option for future trial formulations. Stand-level management options were classified as either part of the spatial solution ($\theta$), or not part of the spatial solution ($\delta$). If a stand had a specified number of consecutive trial solutions exceeding either the $\theta$ or $\delta$ threshold, that particular stand-level solution was accepted and used in subsequent trials. Theoretically, one could use the same value for both $\theta$ and $\delta$. However, we hypothesized that better results might be found with $\theta$ less than $\delta$; that is, take extra care in eliminating stands that are not part of the spatial solution too early. This hypothesis considers that stands with a nonspatial solution early in the search process may become part of the spatial solution if they are evaluated in a much larger context of the forest, made possible by accepting obvious spatial solutions first ($\theta$).

The heuristic is described in Figure 2. It initiates at the upper left with a set of management options $\{Rx\}$ available to each stand. The modeler then specifies both the trial direction and size parameters, and the problem is formulated and solved with the DP moving windows heuristic. If the DP solved exactly (single subproblem), the heuristic stops because by definition it is optimal. If, however, more than one window subproblem was required to solve the problem, each stand is evaluated for whether or not to accept that trial’s scheduled management option ($Rx$) for the stand by comparing it with the specified number of prior trial solutions. If a stand’s management option is accepted, it is the only management option included for that stand in subsequent trial formulations. This results in a simplified DP formulation, as stands with a single management option do not add to the DP size. Therefore more stands can be included in each subproblem of the next trial. Eventually, as more stands’ management options are accepted, the heuristic could result in an exact DP formulation and the heuristic will stop. In the case where an exact DP formulation is too large, the modeler may specify a maximum number of trials to evaluate and the best trial solution will be accepted (not shown).

Eight window subproblem directions were used for DP trials in this study. Directions included each end along the four sides of a forest (e.g., starting at the top end of a window subproblem that spans the north-to-south axis of a forest). Consider the illustration in Figure 3a in which the colored stands (“add to solution next” and “available to schedule next”) define a window across the forest. One window direction originates at the top of the window and works to the bottom and the other begins at the bottom and works to the top. When the DP for the trial is solved, the stand solutions closest to the edge of the forest are accepted for the remainder of the trial (“Add to assumed solution”) and the next window would add as many stands as possible (“Available to schedule next” and “Not yet included”) without violating the specified maximum window size. This overlap parameter was held constant at 40% for each trial.
3.3 Methods for evaluating the heuristic

The heuristic was tested at two scales to determine its performance in identifying a timely, accurate solution. The first scale (the baseline scenario) was small enough to arrive at an exact solution, yet large enough in size and complexity to be a meaningful test of a heuristic-derived solution. In the baseline scenario, the number of management options per stand was limited to six, including five options to create core area at different points in time and one that did not create core area, but contributed to other forest-wide desired conditions such as timber and other ecological conditions. Tests of the baseline scenario involved single trials of window direction and size combinations, as well as a test of the full heuristic. The second test scale (full scenario) included up to 28 management options per stand and was too large to solve with an exact DP formulation.

Results of each test show two financial metrics. First, the total objective function value is the discounted net value including spatial (KW) and nonspatial (timber) values using the shadow pricing approach as described and referenced earlier (Hoganson & Rose, 1984; Paredes & Brodie, 1988). Total objective function value can be used to assess how close the heuristic is to the true optimal solution. Second, its spatial value component is the net present value of core area based on the assumed per unit price (shadow price) for core area in each planning period. The summed spatial value is an indicator of the total amount and timing of the core area in the solution. All net present value estimates are based on a 4% annual discount rate in real dollars.
4.1 Baseline scenario: Tests for single window size and direction combinations

The first tests of the baseline scenario generated information to compare against the heuristic. The tests solved the problem with unique combinations of window size and direction. The maximum window size (number of states in a single stage of the DP) ranged from 1000 to 10 million, and all eight directions were tested for each size. Computing time ranged from less than a minute for the 1000 state limit to 82 min for the ten million state window size. As expected, larger windows yield better results (Table 1), and the same best solution (28608733) was found with the 10 million state window size at Directions 2, 3, 4, and 8 (bold font in Table 1). Although not absolute proof the solution is optimal, it seems unlikely that 4 of 8 DP solutions for substantially different combinations of overlapping windows would converge on anything but an optimal solution. The first part of Table 1 shows that even with a 1000 node limit, solution values were within 0.3% of the highest known objective function value. Increasing the node limit to 100,000 or 1,000,000 yields values to within thousandths of a percentage point of the maximum value. The lower section of the table shows the spatial value of the test as a percentage of the spatial value (8423033) associated with the best solution.
Again, increased maximum DP state-size limits generally identify solutions with higher spatial value. One important observation is that total objective function value and core area value are not necessarily coincident. An extreme case is seen with window direction 4 for the smallest window size, where the objective function value is 0.3% below the maximum, but the spatial value is 32% lower. Alternatively, consider window direction 5 solved with the largest window size. The core area value of this solution is measurably higher than the core area value of the best solution, yet the overall objective value is measurably lower than that of the best solution. These two outcomes suggest there are local optima for the problem close to the global optimum, and that the quality of the solution should be evaluated on not only the overall estimated value of the objective function, but also the estimated value of the spatial component. More important, and related to the first observation, is that the largest window sizes are needed to find solutions that are within 1%–2% of the spatial value associated with the best solution. By themselves, large windows result in long solution times (82 min average) that would potentially be quite limiting for iterative search methods that may require many intermediate DP solutions to help address additional forest-wide constraints.

| Percent objective function value | 1000 | 10,000 | 100,000 | 1,000,000 | 10,000,000 |
|---------------------------------|------|--------|---------|-----------|------------|
| 1                               | 99.80% | 99.94% | 99.99% | 99.999% | 99.9998% |
| 2                               | 99.80% | 99.97% | 99.98% | 99.990% | 100% |
| 3                               | 99.80% | 99.94% | 99.99% | 99.998% | 100% |
| 4                               | 99.70% | 99.96% | 99.97% | 99.998% | 100% |
| 5                               | 99.80% | 99.95% | 99.99% | 99.990% | 99.999% |
| 6                               | 99.90% | 99.96% | 99.97% | 99.990% | 99.9980% |
| 7                               | 99.80% | 99.95% | 99.99% | 99.990% | 99.9990% |
| 8                               | 99.70% | 99.94% | 99.99% | 99.990% | 100% |

| Percent maximum spatial value | 1000 | 10,000 | 100,000 | 1,000,000 | 10,000,000 |
|-------------------------------|------|--------|---------|-----------|------------|
| 1                             | 77.80% | 96.40% | 99.10% | 99.60% | 99.90% |
| 2                             | 80.70% | 94.30% | 96.30% | 96.30% | 100% |
| 3                             | 75.30% | 94.90% | 96.60% | 99.30% | 100% |
| 4                             | 68.00% | 94.20% | 98.50% | 98.90% | 100% |
| 5                             | 79.40% | 93.30% | 97.70% | 99.97% | 100.10% |
| 6                             | 81.50% | 93.60% | 97.00% | 96.50% | 99.60% |
| 7                             | 77.40% | 93.20% | 97.70% | 98.80% | 99.99% |
| 8                             | 70.20% | 92.10% | 98.50% | 98.90% | 100% |

| Solution times (minutes) | 1000 | 10,000 | 100,000 | 1,000,000 | 10,000,000 |
|--------------------------|------|--------|---------|-----------|------------|
| Average                  | 0.5  | 0.8    | 1       | 6.2       | 82         |
| Minimum                  | 0.5  | 0.8    | 1.0     | 5.0       | 15         |
| Maximum                  | 0.5  | 0.8    | 1.1     | 6.7       | 111        |

Note: Bold values indicate combinations that resulted in the maximum value. Italicized values indicate the lowest and highest financial values of the spatial component of the problem. Solution time comparisons are shown in the final portion of the table.
4.2 Baseline scenario: Testing the heuristic

The Baseline scenario was then used to test the learning heuristic. Tests used a series of trials that varied both window direction and window size as well as the number of trials evaluated before any part of a solution was accepted. With each successive trial, the window direction changed and the window size was increased. As more of the solution was accepted and the overall problem became smaller, the analysis was focused on the portions of the problem with inconsistent solutions, arguably those areas most difficult to solve. Additionally, as more of the solution was accepted, the most difficult areas could be evaluated in a larger spatial context, as windows could include much more of the forest in a feasible DP formulation.

There were two different rules used to increase the window size. First, a linear increase was used to adjust window size by a set amount every trial. For instance, one might initiate the search with a window size of 30,000 states, and increase by 10,000 states every trial. The second rule was to define thresholds according to the number of stands in the formulation. When there were only a few stands (such as 100) left in the problem, the window size was increased to a very large size such as 10 million. This potentially allowed the problem to solve with a single-window formulation and meet the stopping rule of the heuristic. The parameter test results are shown in Table 2. The first column indicates the initial window size in terms of DP states for any DP stage (window size) and the magnitude of the window size increase in each successive trial. Acceptance thresholds (θ and δ value) are shown in the second column, and indicate the number of window directions tried before accepting spatial and nonspatial management options, respectively. The third column indicates a threshold for the number of stands left in the DP before a large increase in maximum window size is triggered, and the fourth column indicates the size of increase for each successive trial. Columns five and six indicate the quality

| Initial size/ increase | Threshold stands | Size/ increase | Total loss | Spatial loss | Solution time (min.) | Trials to solve |
|------------------------|------------------|---------------|-----------|-------------|---------------------|-----------------|
| 1k/1k                  | 2/4              | 100           | 50k/10k   | 0.04%       | 3.24%               | 4.49            | 6               |
| 1k/1k                  | 4/6              | 100           | 50k/10k   | 0.02%       | 1.43%               | 5.67            | 9               |
| 1k/1k                  | 8/12             | 100           | 50k/10k   | 0.01%       | 0.26%               | 9.27            | 18              |
| 10k/10k                | 2/4              | 100           | 100k/50k  | 0.01%       | 1.23%               | 4.78            | 4               |
| 10k/10k                | 4/6              | 100           | 100k/50k  | 0.005%      | 0.89%               | 6.59            | 6               |
| 100k/20k               | 8/12             | 100           | 100k/50k  | 0           | 0                   | 11.82           | 12              |
| 30k/10k                | 2/4              | 100           | 1 M/500k  | 0.005%      | 0.89%               | 5.21            | 4               |
| 30k/10k                | 4/6              | 100           | 1 M/500k  | 0           | 0                   | 7.06            | 6               |
| 30k/10k                | 8/12             | 100           | 1 M/500k  | 0           | 0                   | 12.64           | 12              |
| 100k/20k               | 2/4              | 100           | 2 M/1 M   | 0           | 0                   | 6.29            | 4               |
| 100k/20k               | 4/6              | 100           | 2 M/1 M   | 0           | 0                   | 8.79            | 6               |
| 100k/20k               | 8/12             | 100           | 2 M/1 M   | 0           | 0                   | 18.2            | 12              |

Note: Bold values indicate tests that match the optimal solution. Italics indicate the test that found the optimal solution in the least amount of time.
of the solution respect to overall objective function loss and spatial loss. Loss is the value below the best solution; higher loss values indicate inferior solutions. Column six is the solution time in minutes, and the final column indicates the number of trials needed to find the solution.

The test that found the best solution in the least amount of time is indicated in italics in Table 2. It is not the test with the smallest problem formulation, but it accepts the spatial and nonspatial management options in the fewest number of trials. There are many potential parameter value tests that are not reflected in Table 2 that might reveal shorter solution times than those indicated. The point, however, is that the heuristic was generally consistent in identifying the best solution in measurably less time than the more complete DP formulations evaluated in Table 1 (6 vs. 82 min) for a range of parameter values tested.

4.3 Full scenario: Tests on larger problem sizes

One simplifying aspect of the baseline scenario is the maximum number of management options available to any stand was assumed to be six; five spatial and one nonspatial. Limiting the model to five spatial management options per stand speeds solution time, but potentially diminishes some of the benefits of using shorter planning periods. Therefore, the heuristic was tested on a larger problem that evaluated up to 28 management options per stand. Figure 4 shows a distribution of how many management options could be included in the DP formulation.

The learning heuristic was tested on the Full scenario by using a “trimming” algorithm to vary the maximum number of management options per stand. Briefly, the trimming was conducted with a stand-by-stand valuation of each management option that created core area in the context of the stand’s immediate neighbors, assuming the neighbors were assigned the management option that best aligned with the stand option in question. The stand’s management options with the highest values were then included in the DP formulation, up to the specified maximum. The rest were excluded, or trimmed, from the problem. Additionally, the

![Figure 4](image-url) Number of management options for stands used in the full analysis. The figure does not include the 8380 stands that had a mathematically optimal nonspatial solution and thus were excluded from the DP formulation.
management option with the highest noncore area value (“nonspatial”) was included in the DP formulation to allow scheduling to meet other forest conditions. Tests varied the number of management options per stand from 5 to 13. Additionally, the “All” test did not do any trimming (included up to 28 management options). This test was used as a control as it resulted in the highest objective function value.

In all, varying combinations of maximum management options per stand, window size, window increase, and \( \theta/\delta \) parameters were tested. Window sizes were evaluated at the 10,000, 30,000, 100,000, 200,000, 500,000, 1 million, and 2 million levels. Window size increases were set at 50% of the initial window size (i.e., a 30,000 window size was increased by 15,000 with each Trial). Theta/delta (\( \theta/\delta \); number of trials before solution acceptance) values were either 2/4, 4/6, or 8/12. This resulted in nearly 400 unique parameter combinations that could be tested.

Early tests indicated that certain parameter values were not likely to yield good results. Evaluating fewer than seven management options per stand failed to yield superior results even with large windows and large \( \theta/\delta \) values. Additionally, \( \theta/\delta \) values of 2/4 were too low to identify good solutions and \( \theta/\delta \) levels of 8/12, while yielding better solutions, also required additional solution time. Finally, initial window sizes greater than 100,000 seldom produced solutions with higher objective function values and lower solution times. Therefore, tests generally consisted of differing initial management option levels between 7 and 13, a \( \theta/\delta \) of 4/6, and initial window sizes of 30,000 and 100,000.

Test solutions were measured by total objective function value, spatial value, and solution time. Figure 5 displays 36 of the best solutions found. The solution with the greatest value is shown at (1,1) and is indicated as the “Best Solution Value.” The solution was found by using all potential management options and solved in 232 min. The \( x \)-axis shows how close the test objective function value is to the Best Solution Value (expressed as a percentage), and the \( y \)-axis shows how close the spatial value is to the spatial value of the Best Solution Value (expressed as a percentage). The size of each marker is inversely proportional to the amount of time it took to solve; that is, larger markers indicate shorter solution times. The shortest time is 6.72 min and is indicated by the largest point on the graph. All solutions were within 0.2% in the total value dimension and within 7% in the spatial dimension. One might conclude from these results that measurable time savings can be realized without compromising a large amount of overall objective function value or spatial production potential.

### 4.4 An unanticipated result

Thus far, solution quality has been discussed relative to overall solution value and spatial value. A closer examination of the results, however, highlights a concern not apparent by looking at values alone. Specifically, core area production through time can vary considerably between solutions with similar solution values. An example of these temporal variations is illustrated by Figure 6, which shows the final 23 periods of the planning horizon are shown for four of the tests included in Figure 5. The first several time periods of this problem are not particularly interesting since they are heavily influenced by existing and planned habitat and do not vary by test. In Figure 6, the thick solid line is the amount of core area associated with the best solution. The other three solutions are depicted with dotted or dashed lines, and the constraint level is the thin solid line. All nonoptimal solutions are within 0.1% the best solution value and within 3% of the spatial value of the best solution. Yet, in many periods, there are noticeable
differences in the estimated amount of core area between the solutions. For instance, in period 26, the best solution is within 14% of the desired constraint level, and the 11 Rx solution over-achieves the desired level by 43%. Similar differences can be detected in other periods as well, and this can complicate making a determination of whether habitat is valued correctly at each time step.

**FIGURE 5** Scatter plot summarizing the full forest tests. Spatial values and total values are relative to the best solution found. Point size is inverse to solution time; larger points represent shorter solution times.

**FIGURE 6** Core area production over time for select trials.
5 | DISCUSSION AND CONCLUSIONS

5.1 | Strengths of the learning heuristic

A strength of the learning heuristic is that it is based on a specific optimization model structure from operations research. The learning heuristic is distinct from random-search based heuristics. Some heuristic and meta-heuristic methods start with a randomly generated initial solution and seek ways to improve it (e.g., the cellular automaton meta-heuristic described by Heinonen & Pukkala, 2007, initiates with a random solution). Since each window formulation is unique, the learning heuristic does not start with a predetermined solution to move away from a local optimum. Rather, it seeks a global optimum with every formulation. The structure assumed is a DP formulation, yet this same model structure could also be described in terms of linear programming (Wei & Hoganson, 2007). Regardless of how this master problem is defined, the heuristic decomposes the problem into subproblems (windows) that include spatial overlap to help overcome spatial detail that is not truly separable in an exact sense.

The heuristic finds superior solutions by improving the methods of Wei and Hoganson (2008). The heuristic allows examination of larger window subproblems and multiple direction formulations. Larger window subproblems are likely to lead to better solutions because each window considers the spatial interdependencies of more stands. The learning heuristic allows larger window formulations (more included stands) by incrementally accepting common stand-level solutions from independently formulated problems. Searching through a variety of window directions also allows the heuristic to identify superior solutions. As a point of comparison, Wei and Hoganson (2008) conducted tests that corresponded with the first three window sizes of Direction 1 in the Table 1, which would not have identified the highest objective function value found here. Results of this study strongly suggest that distinguishing portions of the forest where one can likely accept intermediate solutions from portions of the forest where the problem is challenging to solve is an efficient use of time that yields better solutions. This strategy has practical appeal which this paper has demonstrated.

Another strength of the heuristic is the time savings that may be realized when parameter settings are set effectively. Figure 5 demonstrates substantial time savings with little compromise in solution quality; the shortest solution time is just under 7 min, compared to the best solution, which was found in 232 min. This translates into a 97% reduction in solution time associated with only a 0.14% reduction in overall value and 4% reduction in core area value. Arguably, this is a valuable trade-off to consider when solution time is limited.

While perhaps not unique to this study, the presence of other (nonspatial) constraints and the dynamics of the patch duration add substantial complexity. The approach presented in this study is similar to those such as Ohman and Lamas (2003) where harvest activities are clustered to realize economic efficiencies of scale from spatially coordinated harvesting activities. However, our study differs in several key areas that make the problem more complex. First, there are more nonspatial constraints to consider, such as cover type and size class constraints. Also, the clustering activity persists over several different time periods, which means that a single stand can conceivably be clustered with different sets of stands in different time periods. This phenomenon indicates the third key difference, which is that the habitat is allowed to move dynamically across the landscape through time. That is, so long as the core area constraint is satisfied, the design of the habitat might take the form of either a gradual shift, where a single patch grows and changes shape and location through time, or it might consist of spatially distinct patches that are created and then disappear at distinct points in time. In a
historical context, one might have expected habitat to mimic the spatially distinct patches that originate and disappear in large blocks since they were created by wildfire. In practice, the solution will likely be implemented with a combination of these two possibilities. The spatial context of these patches in relation to other patches is an aspect of this study that is not explicitly explored, however. Donner et al. (2010) found that larger, nonisolated patches were associated with earlier colonization and later abandonment, and birds may occupy relatively small patches if they are positioned in larger complexes of suitable habitat. Similar spatial concerns may arise when considering the year-over-year colonization patterns that require birds to locate suitable breeding habitat after seasonal migrations. It may be beneficial to design current habitat with good spatial proximity to suitable habitat from the year before.

Finally, comparing solutions based on objective function value alone may mask an important aspect of habitat management; namely, “pinch points” of low habitat availability that may affect species persistence (Marshall et al., 1998). The solutions depicted in Figure 6 show similar objective function values but measurable differences at each time step. In practice, managers will want to avoid deficits in habitat at any given time period and may evaluate results based on the worst-case scenario. Managers may benefit from having several near-optimal solutions from which to choose.

5.2 Future improvements

The learning heuristic allows for more exploration into what makes the “difficult areas” challenging to solve. One can easily pare down the problem to the most difficult spots by identifying those areas that do not solve exactly and that have differing solutions depending on the window size and direction (e.g., Figure 3d). It could be enlightening to study these areas in detail to look for patterns and develop improved solution tactics.

Additional study into the benefits of preprocessing could uncover greater efficiencies in the solution method. The management option trimming routine used to pare the management options included in the DP has a potential weakness in that it eliminates the management options for some stands that would be in the optimal solution (e.g., in Figure 5, even including the best 13 management options does not allow the DP to match the best solution). This outcome points to the concern that the DP or any other optimization approach can be limited by weaknesses in the preprocessing trimming routine. Stronger trimming rules that increase the likelihood of including the optimal management option and eliminate more nonoptimal options for each stand would allow the DP to find better solutions with shorter solution times.

ACKNOWLEDGMENTS

The USDA Forest Service, the University of Minnesota North Central Research and Outreach Center and the University of Minnesota Department of Forest Resources provided funding and support for this study project.

AUTHOR CONTRIBUTIONS

Eric Henderson: Learning Heuristic development, testing, documentation. Literature review. Manuscript writing. Manuscript review. Howard Hoganson: Learning Heuristic guidance and review. Literature review. Manuscript editing and review. Howard M Hoganson: conceptualization (supporting); formal analysis (supporting); methodology (equal); software (equal); writing review & editing (equal).
REFERENCES
Andersson, D., & Eriksson, L. O. (2007). Effects of temporal aggregation in integrated strategic/tactical forest planning. *Forest Policy and Economics*, 9(8), 965–981.
Baskent, E. Z., & Keles, S. (2005). Spatial forest planning: A review. *Ecological Modelling*, 188, 145–173.
Belavenutti, P., Romero, C., & Diaz-Balteiro, L. (2019). Integrating strategic and tactical forest management models within a multicriteria context. *Forest Science*, 65(2), 178–188.
Bellman, R. (1954). The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60, 503–515.
Bixby, J., Hoganson, H., & Wei, Y. (2019). Potential gains from spatially explicit coordinated planning between two large public ownerships in Minnesota. *Forest Ecosystems*, 6, 1–17.
Blum, C., Puchinger, J., Raidl, G., & Roli, A. (2011). Hybrid metaheuristics in combinatorial optimization: A survey. *Applied Soft Computing*, 11(6), 4135–4151.
Borges, J., Hoganson, H., & Rose, D. (1999). Combining a decomposition strategy with dynamic programming to solve spatially constrained forest management scheduling problems. *Forest Science*, 45(2), 201–212.
Donner, D. M., Ribic, C. A., & Probst, J. R. (2010). Patch dynamics and the timing of colonization-abandonment events by male Kirtland’s Warblers in an early succession habitat. *Biological Conservation*, 143, 1159–1167.
Forsell, N., & Eriksson, L. (2014). Influence of temporal aggregation on strategic forest management under risk of wind damage. *Annals of Operations Research*, 219, 397–414.
Heinonen, T., Kurrtila, M., & Pukkala, T. (2007). Possibilities to aggregate raster cells through spatial optimization in forest planning. *Silva Fennica*, 41(1), 89–103.
Heinonen, T., & Pukkala, T. (2007). The use of cellular automaton approach in forest planning. *Canadian Journal of Forest Research*, 37, 2188–2200.
Hoganson, H., Borges, J., & Wei, Y. (2008). Coordinating management decisions of neighboring stands with dynamic programming. In *Designing Green Landscapes* (pp. 187–213). Springer.
Hoganson, H. M., Bixby, J., Bergmann, S., & Borges, J. G. (2004). Large scale planning to address interior space production studies from northern Minnesota. *Silva Lusitana*, 12, 35–47.
Hoganson, H. M., & Borges, J. G. (1998). Using dynamic programming and overlapping subproblems to address adjacency in large harvest scheduling problems. *Forest Science*, 44(4), 526–538.
Hoganson, H. M., & Rose, D. W. (1984). A simulation approach for optimal timber management scheduling. *Forest Science*, 30(1), 220–238.
Johnson, K. N., & Scheurman, H. L. (1977). Techniques for prescribing optimal timber harvest and investment under different objectives. *Forest Science Monograph*, 18, 1–31.
Kangas, A., Kurrtila, M., Hujala, T., Eyvindson, & Kangas, J. (2015). *Decision support for forest management* (Vol. 30, 2nd ed., pp. 1–307). Cham: Springer Ingernational Publishing AG.
Marshall, E., Haight, R., & Homans, F. (1998). Incorporating environmental uncertainty into species management decisions: Kirtland’s Warbler habitat management as a case study. *Conservation Biology*, 76(3), 975–985.
Michigan Department of Natural Resources. (2015). US Fish and Wildlife Service; US Forest Service. Kirtland’s Warbler Breeding Range Conservation Plan.
Mohamed-Mahmoud, M., Lescourret, F., Xin, Y., & Lavigne, C. (2015). Metaheuristics for agricultural land use optimization: A review. *Agronomy for Sustainable and Development*, 35, 975–988.
Ohman, K. (2000). Creating continuous areas of old forest in long-term forest planning. *Canadian Journal of Forest Research*, 30, 1817–1823.
Ohman, K., & Eriksson, L. O. (1998). The core area concept in forming contiguous areas for long-term forest planning. *Canadian Journal of Forest Research*, 28, 1032–1039.
Ohman, K., & Lamas, T. (2003). Clustering of harvest activities in multi-objective long-term forest planning. *Forest Ecology and Management*, 176, 161–171.
Paredes, G., & Brodie, J. (1988). Activity analysis in forest planning. *Forest Science*, 34(1), 3–18.
De Pellegrin Llorente, I., Hoganson, H., Carson, M., & Windmuller-Campione, M. (2017). Recognizing spatial considerations in forest management planning. *Current Forestry Reports, 3*, 308–316.

Pellerin, R., Perrier, N., & Berthaut, F. (2019). A survey of hybrid metaheuristics for the resource-constrained project scheduling problem. *European Journal of Operational Research, 280*(2), 395–416.

Rau, A., Burkhardt, V., Dorninger, C., Hjort, C., Ibe, L., Kebler, L., Kristensen, J., McRobert, A., Sidemo-Holm, W., Zimmerman, H., Abson, D., von Wehrden, H., & Ekroos, J. (2019). Temporal patterns in ecosystem services research: A review and three recommendations. *Ambio*.

Sjogren, S. (2012). Technical guide Kirtland’s Warbler habitat management. Gladstone, MI: Hiawatha National Forest.

USDA Forest Service. (2004). Final environmental impact statement for forest plan revision. Duluth, MN: Chippewa National Forest Superior National Forest.

USDA Forest Service. (2006). Hiawatha National Forest 2006 Forest Plan. Author.

Wei, Y., & Hoganson, H. (2005). Landscape impacts from valuing core area in national forest planning. *Forest Ecology and Management, 218*, 89–106.

Wei, Y., & Hoganson, H. M. (2007). Scheduling forest core area production using mixed integer programming. *Canadian Journal of Forest Research, 37*(10), 1924–1932.

Wei, Y., & Hoganson, H. M. (2008). Tests of a dynamic programming-based heuristic for scheduling forest core area production over large landscapes. *Forest Science, 54*(3), 367–380.

**How to cite this article:** Henderson EB, Hoganson HM. A Learning Heuristic for Integrating Spatial and Temporal Detail in Forest Planning. *Natural Resource Modeling, 2021;34:e12299*. https://doi.org/10.1111/nrm.12299