SYSTEMS ANALYSIS IN FOREST RESOURCES

For. Sci. 66(4):443–456
doi: 10.1093/forsci/fxz012
Copyright © 2019 Society of American Foresters

economics

Managing Fragmented Fire-Threatened Landscapes with Spatial Externalities

Christopher J. Lauer, Claire A. Montgomery, and Thomas G. Dietterich

Accounting for externalities generated by fire spread is necessary for managing fire risk on landscapes with multiple owners. In this paper, we determine the optimal management of a synthetic landscape parameterized to represent the ecological conditions of Douglas-fir (*Pseudotsuga menziesii*) plantations in southwest Oregon. The problem is formulated as a dynamic game, where each agent maximizes their own objective without considering the welfare of the other agents. We demonstrate a method for incorporating spatial information and externalities into a dynamic optimization process. A machine-learning technique, approximate dynamic programming, is applied to determine the optimal timing and location of fuel treatments and timber harvests for each agent. The value functions we estimate explicitly account for the spatial interactions that generate fire risk. They provide a way to model the expected benefits, costs, and externalities associated with management actions that have uncertain consequences in multiple locations. The method we demonstrate is applied to analyze the effect of landscape fragmentation on landowner welfare and ecological outcomes.

Study Implications: This research builds on several important ideas for forest management. Fire risk for any particular stand on a landscape is a function of vegetation conditions across the entire landscape. Landowners who wish to achieve a management objective that is affected by fire risk need to account for the risk generated by broader landscape conditions. This work expands on a tractable model to account for the spatial interactions generated by fire spread that affect the optimal timing and spatial location of timber harvest and fuel treatments. In this paper, we demonstrate that optimal behavior changes when there are multiple landowners. On a sufficiently fragmented landscape, one landowner's actions can create additional risk for their neighbors. This work suggests that policy interventions to incentivize risk reducing behavior may be appropriate on sufficiently fragmented landscapes.

Keywords: wildland fire, stochastic dynamic games, spatial, ecological disturbance, risk, approximate dynamic programming, multiagent reinforcement learning

Ecological disturbances such as wildfire can create spatial interactions that complicate forest management. Value on a forest landscape can be diminished, or destroyed, by wildfire, and the extent of the damage is at least partially outside manager control. Since fire spreads across property boundaries, conditions on one part of a landscape can create externalities for other parts of a landscape. Landowners can take action to increase their welfare by modifying landscape conditions that give rise to fire risk (Spies et al. 2014). However, they are only able to address conditions on their own land. A landowner's willingness to engage in management that reduces fire risk will primarily depend on how management affects value at risk on their own land, not on the overall benefit for the landscape. Several authors have identified transboundary risks as an important consideration for managing wildfire, including Zaimes et al. (2016) and Ager et al. (2018).

At the individual stand level, Routledge (1980) and Reed (1984) demonstrated that the optimal rotation age for timber harvest determined by Faustmann (1849) can be adjusted to account for the possibility of an unpredictable natural disaster damaging or destroying stand value. They showed that the optimal rotation age is shorter as the probability of stand destruction increases. Reed built upon this model in subsequent papers to demonstrate how optimal management is affected when fire arrival rate is a function of stand characteristics. He also explored optimal scheduling of investment in fire protection, such as fuel treatment or fire-fighting infrastructure (Reed 1989, 1993). Studies that built upon Reed's insights include Amacher et al. (2005) and Garcia-Gonzalo et al. (2014) who looked at how fuel treatment and silvicultural interventions affect optimal rotation age for fire-threatened forest stands. Daigneault et al.

Manuscript received December 31, 2017; accepted February 26, 2019; published online May 10, 2019.

Affiliations: *Christopher J. Lauer (christopher.lauer@noaa.gov), National Oceanic and Atmospheric Administration, 1315 East-West Highway, Silver Spring, MD 20910. Claire A. Montgomery (claire.montgomery@oregonstate.edu), Oregon State University, Forest Engineering, Resources & Management, 280 Peavy Hall, Corvallis, OR 97331. Thomas G. Dietterich (tgd@oregonstate.edu), Oregon State University, Electrical Engineering and Computer Science, 1148 Kelley Engineering Center, Corvallis, OR 97331.*

Acknowledgments: *This material is based upon work supported by the National Science Foundation under CyberSEES Computing and Visualizing Optimal Policies for Ecosystem Management (grant no. 1331932) and the USDA Forest Service Western Wildland Environmental Threat Assessment Center (joint venture agreement 12-CR-11261907-100). This work does not necessarily reflect the views of the funding agency, and no official endorsement should be inferred.*

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

(2010) examined how carbon sequestration is affected by manager response to fire risk.

These stand-level analyses do not consider spatial fire spread. Fire can travel large distances across a landscape, and hence fire risk on any individual stand is a function of the condition of the entire landscape. Therefore, the value of fuel treatment in one location depends partly on its effect on fire risk elsewhere. Incorporating spatial interactions into the problem adds greatly to its complexity. Studies that incorporate spatial interactions and model a single decisionmaker who places fuel treatments on a real landscape typically involve static analyses that evaluate the effect of fuel treatment on expected loss to the next big fire. For example, Wei et al. (2008) separated fire arrival probability into ignition and spread probabilities; the latter can be influenced by management activities. They modeled placement of fuel treatments on the landscape to minimize expected value at risk, formulating and solving it as a mixed integer programming problem. Ager et al. (2010) did not optimize, but used repeated fire spread simulations to compare the damage probabilities for structures in the wildland–urban interface under different fuel-management strategies. Chung et al. (2013) developed a model that places fuel treatments and timber harvest over two time periods to minimize expected loss. They incorporated a spatially explicit fire spread model directly into the algorithm and solved it using simulated annealing. Although the Chung et al. model is intertemporal, in that vegetation evolves over the treatment periods, all of these models are static because they do not account for how optimal management will adjust on a post-fire landscape, should fire actually occur.

Fire, when it does occur, can change landscape conditions dramatically, and land managers should respond by adapting their management to the new conditions. Therefore, the problem of optimal management of a fire-threatened landscape is inherently dynamic. Konoshima et al. (2008, 2010) moved from static to dynamic spatial stochastic optimization by formulating the model as a stochastic dynamic programming problem and solving it by complete enumeration. Their model suggests that landowners will try to protect on-site timber values by shortening rotations, as suggested by Reed, but will try to protect adjacent stand timber values by postponing harvest in order to avoid high spread rates associated with young stands. Although inferences about general behavior were obtained, the practical usefulness of Konoshima et al.'s approach for forest planning and policy analysis is limited by the need to greatly simplify the problem for the sake of tractability. Lauer et al. (2017) employed approximate dynamic programming methods to determine an optimal management policy in a setting where the size of the landscape and the complexity of the ecological models precluded the use of exact dynamic programming methods because of the curse of dimensionality. The policies found in this paper corroborated the Konoshima results in a more realistic setting.

The single-agent approach may be useful for federal agencies that manage large expanses of forest land, but some of the most challenging and intriguing policy questions involve fire that crosses ownership boundaries. There are several examples of analyses of fire on multiple ownerships. Yoder et al. (2003) and Yoder (2004) examined how liability rules and negligence standards affect precautionary effort on the part of landowners who use controlled burning as a management tool and adjacent residential property owners who might be harmed by escaped fire. Crowley et al. (2009) used the Reed model to determine optimal rotation age for two adjacent stands with interdependent fire risk. Busby et al. (2012) developed a game-theoretic model to demonstrate how the pattern of public and private ownership might affect fuel management by residential property owners in the wildland–urban interface in a dynamic context. Again, the practical usefulness of these papers for planning and policy analysis is limited by the need to greatly simplify the landscape, fuel models, fire behavior, and weather for the sake of tractability.

In this study, we extend the spatial stochastic approximate dynamic programming problem developed in Lauer et al. (2017) by incorporating a game-theoretic model to explore how the configuration of multiple ownerships on a landscape affects optimal fuel treatment and harvest choices relative to that of a single owner. To solve this game, we use approximate dynamic programming (ADP), a form of value function iteration used to solve high-dimension dynamic optimization problems (Sutton and Barto 1998, Powell 2007, 2009). We applied our method to model optimal timing and placement of timber harvest and fuel treatment for two agents with interdependent fire risk on a landscape that we parameterized to represent the ecological conditions of southwest Oregon. The value functions we estimate provide a way to model the expected benefits, costs, and externalities associated with different management actions that have uncertain consequences in multiple locations on the landscape. We use this optimization framework to explore how ownership fragmentation affects landowner welfare. One example of a highly fragmented landscape is the Oregon and California Railroad Revested Lands (O&C lands) of western Oregon. As ownership on a forest landscape becomes increasingly fragmented, each individual landowner has less ability to modify the conditions that affect their fire risk. Landowners must compensate for this increased risk by modifying their management strategies.

This paper is organized as follows. The economic model is described, and then the solution methods and the implementation of the ADP algorithm used to estimate value functions that determine optimal actions are described. The parameterization of the model landscape is then described, and a description of the scenarios associated with ownership fragmentation on the landscape is provided. Finally, results are presented, followed by concluding remarks.

Model

A bio-economic model provides insight into how landscape conditions that give rise to fire risk affect landowner behavior. Although our model accounts for the financial incentives faced by agents, it could also be specified to include nonmarket incentives. Faced with these incentives, the economic agent makes decisions in the context of the ecological processes that determine the evolution of the landscape. Vegetation evolves, stochastic disturbances (in this case, fire) occur over time so that future costs and revenues depend on actions taken today, and, likewise, the present value depends on those future rewards. Because we account for the effect of fire spread across the landscape on each stand's value, our model is explicitly spatial.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

A single agent who controls the entire landscape has the incentive to account for all of these spatial interactions. However, when there are multiple agents controlling and deriving benefit from only their own piece of the landscape, management strategies employed by one agent may lead to externalities for the others.

Markov Decision Process and Bellman's Equation

The economic and ecological components of this problem can be integrated by representing the problem as a Markov Decision Process (Puterman 1994), which has five different components.

1. A set of possible *States* that depend on the attributes of the individual stands contained by the landscape. The state S_t describes the conditions of the landscape at time t.
2. A set of *Actions* that describes what a land manager can do. In our setting, the overall action at time t is a vector x_t of the management activities applied to each stand in the landscape. For each stand, there are four possible management activities: harvest timber (clearcut), treat fuel to reduce fire risk, implement both activities, or do nothing; in each time step, one of these four options must be chosen for each stand.
3. A *Reward Function*, $C(S_t, x_t)$, that describes the immediate (i.e., current period) costs or revenues associated with a particular action for a particular state.
4. A *State Transition Model*, $S_{t+1} = S'(S_t, x_t, W_t)$, that describes how the state evolves over time. The transition from S_t to S_{t+1}, is a function of the current state, S_t, the current action, x_t, and a vector of stochastic events, W_t, which includes fire arrival and weather.
5. A *Discount Factor*, δ, that determines how current rewards are valued relative to future rewards.

To analyze this Markov Decision Process, Bellman's equation (Bellman 1957), a recursive equation that assigns a value to a particular state, is used (Equation 1). This equation represents a landowner's decisionmaking process and can be broken down into two parts. A manager will consider the immediate cost or benefit of a management action, which is captured by the reward function. The decisionmaker will also consider the implications of this action for the future by forming an expectation about the value of the next period's state. We employ the so-called action-value representation, also known as the "Q-value" (Watkins and Dayan 1992). The quantity $Q(S_t, x_t)$ (Equation 1b) is the expected return of taking action x_t in state S_t and then behaving optimally thereafter. In approximate dynamic programming, it is often useful to first compute $Q(S_t, x_t)$ for each possible action x_t and then choose the action that has the highest Q value. This is the optimal action in state S_t, and it defines the value $V(S_t)$:

$$V(S_t) = \max_{x_t} Q(S_t, x_t) \tag{1a}$$

$$Q(S_t, x_t) = C(S_t, x_t) + \mathbb{E}_{S_{t+1}}[\delta V(S'(S_t, x_t, W_t))] \tag{1b}$$

Spatial Interactions

To incorporate spatial interactions, we decompose $Q(S_t, x_t)$ into a separate action-value function for each stand in the landscape. Let $Q_j(S_t, x_t)$ denote the contribution of stand $j = 1, 2, \ldots, J$ to the value of the overall landscape. Note that the state S_t, action x_t, and stochastic event W_t are not indexed by j because they represent the entire landscape.

$$Q_j(S_t, x_t) = C_j(S_t, x_t) + E_W\left[\delta V_j(S'(S_t, x_t, W_t))\right] \tag{2}$$

$V_j(S_t)$ is stand j's contribution to the value of landscape S given that the landscape is managed optimally in future periods, i.e., $\sum_j V_j(S_t) = V(S_t)$. The vector of actions, x_t, is chosen to maximize the overall Q-value for the landscape, which is the sum of the contributions from each stand:

$$V(S_t) = \max_{x_t} Q(S_t, x_t) = \max_{x_t} \sum_{j=1}^{J} Q_j(S_t, x_t) \tag{3}$$

The individual stand value function depends not only on that stand's vector of attributes, but also on the attributes of the entire landscape (as indicated by S_t) to account for potential fire spread. Therefore, the actions that maximize the value of the landscape may not maximize the value of each individual stand.

Multiple Agents

When there are multiple agents on a landscape, each agent will maximize the value of the stands they control. We assume these agents will not consider the implications of their actions for the welfare of the other owners. Let $n = 1, \ldots N$ index the owners/agents on the landscape and $j_n = 1, \ldots J_n$ index the set of stands owned by agent n. $Q^n_{j_n}\left(S_t, x^n_t | \boldsymbol{x_t^{-n}}\right)$ represents the action-value function for stand j_n owned by agent n, x^n_t is the vector of actions taken by agent n, and $\boldsymbol{x_t^{-n}}$ denotes the vectors of actions taken by the other agents. Because of the spatial interactions created by fire spread, the value of each agent's Q-functions depends on the actions of every agent on the landscape.

$$Q^n_{j_n}\left(S_t, x^n_t | \boldsymbol{x_t^{-n}}\right) = C^n_{j_n}(S_t, x^n_t) + E_W \delta\ V^n_{j_n}\left(S'\left(S_t, x^n_t, \boldsymbol{x_t^{-n}}, W_t\right)\right) \tag{4}$$

$$V^n(S_t) = \max_{x^n_t} Q^n\left(S_t, x^n_t | \boldsymbol{x_t^{-n}}\right) = \max_{x^n_t} \sum_{j_n=1}^{J_n} Q^n_{j_n}\left(S_t, x^n_t | \boldsymbol{x_t^{-n}}\right) \tag{5}$$

Each agent will try to maximize the value of the stands they control; however, agents cannot take actions on parts of the landscape they do not own, even though conditions on these parts of the landscape may also contribute to their fire risk. The actions of an agent will be determined in part by their reaction to the conditions created by other agents. In order to account for this, we can model the interaction between agents as a game, where each agent will adjust their actions to account for the actions taken by other landowners. The optimal set of actions for each agent is found by solving for the set of actions implied by the value function (Equation 5) for all N agents simultaneously (Equation 6). This solution is the Nash equilibrium, where no agent has anything to gain by changing their vector of actions as long as the actions of the other agents also remain unchanged.

$$\begin{aligned} X^1(S) &= \arg\max_{x^1_t} Q^1\left(S_t, x^1_t | \boldsymbol{x_t^{-1}}\right) \\ &\vdots \\ X^N(S) &= \arg\max_{x^N_t} Q^N\left(S_t, x^N_t | \boldsymbol{x_t^{-N}}\right) \end{aligned} \tag{6}$$

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

Solution Method

Determining the optimal action at each state requires solving for the actions of all agents simultaneously (Equation 6). This problem is difficult because the value function is unknown, and it is intractable to enumerate all the states, actions, or stochastic events that could occur on the landscape—this is the curse of dimensionality, which precludes the use of exact dynamic programming methods. Multiple agents compound this problem because the actions of each agent will affect the welfare of the others. We address this problem through a combination of two techniques described by Powell (2007): postdecision states and approximate dynamic programming (ADP).

Postdecision State

The postdecision state technique is a method for simplifying the transition model and the action-value functions; it was first described by Van Roy et al. (1997). In the context of multiple agents, a postdecision state can be defined as the state of the landscape after an agent has taken an action, but before the stochastic events (in this case fire arrival), *or the actions of the other agents*, are observed. This formulation is similar to the multiagent Q-learning techniques described by Littman (2001) and Busoniu et al. (2008). Instead of representing the transition model as a function $S_{t+1} = S'(S_t, x_t^n, \boldsymbol{x_t^{-n}}, W_t)$, which depends on the current state S_t, the management actions $x_t^n, \boldsymbol{x_t^{-n}}$, and the stochastic realization W_t, we can divide the transition into two steps: (1) the effect of the agent's own management action and (2) the stochastic transition to the subsequent state, which is also a function of the actions of the other agents. Each agent faces the same starting state but controls different portions of the landscape and takes separate actions, which leads to a different postdecision states for each agent in each period. Agent n's postdecision state is written as $S_t^{x^n} = g(S_t, x_t^n)$. The most important property of the postdecision state is that it can be anticipated accurately by the decisionmaking agent; it depends only on the current state, which is observed by the agent, and the actions of that agent. The transition to the next period state is a function of the actions of all the agents as well as the stochastic events $S_{t+1} = S'(S_t, x_t^n, \boldsymbol{x_t^{-n}}, W_t)$. With this change, we can rewrite the Q and V functions in terms of the postdecision state S^{x^n}; the Q-function is only a function of agent n's actions:

$$Q_{j_n}^n\left(S_t, x_t^n\right) = C_{j_n}^n\left(S_t, x_t^n\right) + V_{j_n}^n\left(g(S_t, x_t^n)\right) \;\; = C_{j_n}^n\left(S_t, x_t^n\right) + V_{j_n}^n\left(S_t^{x^n}\right) \tag{7}$$

We can rewrite the Bellman equation in terms of the postdecision state as:

$$V^n\left(S_t^{x^n}\right) = E_{S_{t+1}}\left[\delta \max_{x_{t+1}^n}\left\{C^n\left(S_{t+1}, x_{t+1}^n\right) + V^n\left(g\left(S_{t+1}, x_{t+1}^n\right)\right)\right\}\right] \tag{8}$$

where $C^n = \sum_{j_n}^{J_n} C_{j_n}^n$ and $V^n = \sum_{j_n}^{J_n} V_{j_n}^n$. This version of the Bellman equation relates the value of the postdecision state for agent n, $S_t^{x^n}$, to the expected value of the next period reward and the value of the next period postdecision state. Hence, this form of the Bellman equation goes from postdecision state to postdecision state, whereas the standard Bellman equation goes from (predecision) state to (predecision) state. The advantage of this is that the expectation in Equation 8 is outside the max operation, whereas in Equation 5, we must compute a separate expectation for each possible combination of actions before we can find the maximum. The value function for the postdecision state is the value given that the other agents follow the Nash equilibrium strategy.

Value Function Approximation

After reformulating the Bellman equation using the postdecision state, we solve it using value function approximation. Each stand's contribution to the value of the postdecision state $V_{j_n}^n(S^{x^n})$ is represented by a linear model defined with respect to a set of basis functions $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_I)$, and coefficients $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ as shown in Equation 9. A separate approximation is created for each agent:

$$V_{j_n}^n\left(S_t^{x^n}\right) \approx \bar{V}_{j_n}^n\left(S_t^{x^n}\right) = \sum_{i=1}^{I} \theta_i^n \times \phi_i^n\left(S_t^{x^n}\right), \tag{9}$$

Basis functions $\boldsymbol{\phi} = (\phi_1, \ \ldots, \ \phi_I)$ are a numeric representation of observable attributes for the stand and surrounding landscape that influence the net present value of the stand. Knowing values for $\boldsymbol{\theta}$ allows us to compute the value of any state by computing the basis function values in that state, multiplying by the coefficients, and summing the results. The value function approximation (Equation 9) is conceptually similar to a hedonic model for determining the value of a forest stand. Hedonic models are used to estimate how attributes of a composite good contribute to its value. In this case, the composite good is the stand, which depends on the attributes of the landscape. Each agent's value function for the landscape and the optimal policy it determines are now defined using Q-values for each stand that include an approximation of the stand's contribution to the value of the postdecision state, and can be defined based on the actions of agent n only:

$$Q_{j_n}^n\left(S_t, x_t^n\right) = C_{j_n}^n\left(S_t, x_t^n\right) + \bar{V}_{j_n}^n\left(S_t^{x^n}\right) \tag{10a}$$

$$V^n\left(S\right) = \max_{x^n} Q^n\left(S_t, x_t^n\right) = \max_{x^n} \sum_{j_n=1}^{J_n} Q_{j_n}^n\left(S_t, x_t^n\right) \tag{10b}$$

Value Iteration Algorithm

Here we give a brief description of the value iteration algorithm employed to estimate the unknown coefficients $\boldsymbol{\theta}$; we provide details in the Appendix. To compute the optimal values of the coefficients, we initialize $\boldsymbol{\theta^n} = \boldsymbol{\theta^{n,0}}$ for $n = 1 \ \ldots \ N$ agents and then update the coefficients for each agent simultaneously through a series of $k = 1 \ldots K$ iterations of Equation 10b.

The iterative updates to the value function coefficients occur during a number of simulation cycles. In each cycle, we start in a chosen starting state, S_0, and simulate management for several years into the future. In each time step t, simulation outcomes are used to update the estimate of the value function coefficients. The most recent estimate of the value function $\bar{V}_{j_n}^{n,k}$ is used to find the optimal action x_t^n for each agent, given the current state S_t. We employ a heuristic optimization method that combines simulated annealing (Kirkpatrick 1984) with tabu search (Reeves 1993). This vector

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

of actions x_t is applied to the landscape to find the postdecision state. We then simulate fire events (which could be "no fire") on the postdecision landscape by drawing random ignition and weather events from the distributions described below. The optimal action and the value for each of the states resulting from the simulated fires are found for each agent to calculate the realized contribution each stand makes to the value of the postdecision state for that agent. In the next step, the coefficients $\boldsymbol{\theta}$ are updated to reduce the difference between our current estimate of stand value and the realized values of the states resulting from the simulations. This completes one "Bellman backup" step for value iteration. Then, we choose one of the simulated resulting states S_{t+1} as the next starting state. If there are no time steps left in the current cycle, a new cycle is started with the starting state S_0. Additional steps are necessary to adequately explore the state space and avoid local optima; these are also described in the Appendix.

Data and Parameters

The underlying ecology is a crucial driver of agent behavior. To demonstrate this optimization framework, a representative landscape for SW Oregon was created using pre-existing ecological models and parameters to characterize the state variables, transition functions, and reward functions. The action variables we model are fuel treatment and timber harvest. Although there are many possible objectives for forest management, in this application we assumed that each landowner's goal is to maximize the expected net present value of harvested timber on the landscape. This objective is easy to define and characterizes the objectives of some state and private forest landowners. Our financial parameters include a real discount rate of 4 percent ($\delta = 0.96$; Row et al. 1981), log prices obtained from the Oregon Department of Forestry (2016), and harvest/haul costs estimated based on a harvest cost model developed by Fight et al. (1984).

Landscape Parameters

We modeled a representative forest landscape as an 8 × 8 grid consisting of 64 forty-acre square stands that are flat (no elevation change) with the same soil conditions, climate, and weather. The defining feature of a forest stand is that its vegetation is relatively homogeneous and can be treated in a uniform manner (Tappeiner 2007). Smaller stand sizes increase the resolution and landscape heterogeneity, especially for determining the effect of fire on the landscape; larger stand sizes decrease modeling complexity. We selected 40 acres as a reasonable minimum size for a timber harvest unit. In order to account for edge effects, and because our model landscape is relatively small, we model the landscape as a torus that wraps on itself. This construction eliminates the need to model the costs associated with fire that spreads from outside the landscape, or fire that spreads off the edge of the landscape. It ensures that all effects of the fire are captured in the model, because a fire that spreads to the Eastern boundary (for example) wraps around and continues spreading inward from the Western boundary. It works as long as we assume that the surrounding landscape is similar to the model landscape in terms of vegetation, fire behavior, weather, and management options and objectives.

The initial landscape is created by randomly assigning an age class, with an associated vector of attributes, to each stand in the landscape. Each stand evolves over time independently of the other stands. We tracked the evolution of stand characteristics over time using a transition table—attributes we track are stand age, total cubic feet of biomass per acre, merchantable cubic feet per acre, merchantable board feet per acre, quadratic mean diameter, crown base height, tree height, and fuel model. These characteristics were used to drive simulations of fire events and to compute the reward function resulting from landowner actions.

Stands transition into the next state as a result of vegetation growth, fire, harvest, and fuel treatment as follows:

- *Vegetation growth* was simulated using the Inland CA/Southern Cascades variant of Forest Vegetation Simulator (FVS) (Dixon 2002). In the vegetation simulations, bare ground is prepared for planting by piling and burning surface fuel and planting 500 Douglas-fir trees (*Pseudotsuga menziesii*) per acre. At age 15, the stands are thinned from below to a density of 300 trees per acre. After this time, trees are allowed to grow until they are harvested or destroyed in a fire, at which time the stand is re-planted. *Harvest* age for each stand is determined by the optimization algorithm. This approximates typical even-age stand management that would occur in this type of forest (Hobbs et al. 1992, Tappeiner et al. 2007). Surface fuel models classify a wide number of vegetative covers for the purpose of modeling fire spread (Anderson 1982). We used the Fire and Fuel Extension to the Forest Vegetation Simulator (FFE-FVS; Reinhardt and Crookston 2003) to assign fuel models to each stand in each time step as it grows, receives silvicultural treatments, is harvested, and has *fuel treatments* applied (simulated in FVS as piling and burning surface fuel).
- For *weather*, we used FireFamily Plus (Bradshaw & McCormick 2000), a software tool that analyzes weather observations and computes fire danger indices, to analyze Remote Automatic Weather Station (RAWS) data for several weather stations in SW Oregon in order to determine mean wind speed and fuel moisture conditions for four different fire danger categories: low fire danger was the mean conditions of the 0th–15th percentile of the fire danger index, moderate was the mean conditions of the 16th–89th percentile, high was the mean of the 90th–97th percentile, and extreme was the mean of the 98th–100th percentile. Weather was drawn according to the following discrete distribution: lower fire danger, probability .15; moderate fire danger, .65; high fire danger; .07; extreme danger .03; this fire danger controls how quickly fire spreads through the landscape. Wind is not equally probable from all directions; it is more likely to come from some directions than others and the level of fire danger may be correlated with the wind direction. We assumed that each wind direction had the same distribution of fire danger. However, we modeled a prevailing wind direction by averaging each weather station's ranked wind direction probability and rounding to the nearest whole number. Since our representative landscape is symmetrical, it does not matter which direction the prevailing wind comes from. On our landscape, wind probabilities were specified as 30 percent from the west, 15 percent each from NW and SW, 10 percent each from north and south, 8 percent each from NE and SE and 4 percent from the east.
- *Fire* occurrence is characterized by ignition, spread rate, and duration. Because we assumed that fire arrival leads to stand destruction and complete value loss for the standing timber, we only modeled fire spread and not also fire intensity. *Ignition*

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

probability was determined using statistics from the SW District of the Oregon Department of Forestry (Thorpe 2011). All stands on the representative landscape had an equal probability of ignition. We used the BEHAVE fire modeling system (Andrews et al. 2003) to determine the fire *spread rates* associated with each fuel model/weather combination. The extent of fire spread for each ignition is controlled by the fire weather danger (described in the previous paragraph) and fire *duration*. Duration was a randomly drawn number between 24 and 96 h, with longer durations more likely under more dangerous weather conditions. Fire duration is not based on empirical data; instead, repeated simulations on randomly generated landscapes using the parameters described above were performed to find a distribution of durations that led to a fire size distribution similar to that which has historically occurred in SW Oregon (Thorpe 2011). Fuel-treatment costs were determined based on a study by Calkin and Gebert (2006).

Ownership Configurations

To explore the effect of ownership fragmentation, we examine two different cases. In each scenario, the landscape was populated with $N = 2$ agents. In the *low-fragmentation* scenario, the landscape was divided between the two agents by drawing a line down the middle. Agent 1 controlled the west half of the landscape, and agent 2 controlled the east half (Figure 1). In the *high-fragmentation* scenario, the same landscape was divided into four-stand blocks, and alternating blocks were assigned to each agent creating a checkerboard pattern with many more between-owner adjacencies (Figure 2). Fragmentation on a landscape may not always be this severe or neatly segmented; however, the O&C lands that occupy approximately 2.5 million acres of Western OR provide an example of this pattern of fragmentation. A scenario that represents the optimal management of the landscape under a single owner (i.e., $N = 1$) was included for comparison, and this scenario will be referred to as the *social planner* scenario. This case represents the value of the landscape managed for maximum social welfare.

Value Function Specification

Several different specifications for the basis functions $\phi = (\phi_1, \ldots, \phi_I)$ were attempted in preliminary experiments. The specification that performed best for our model landscape is described below. A value function was specified for each agent, and the coefficients of these value functions were estimated for each scenario. Since both agents face the same set of rewards and underlying ecological processes, we are able to specify the same value function

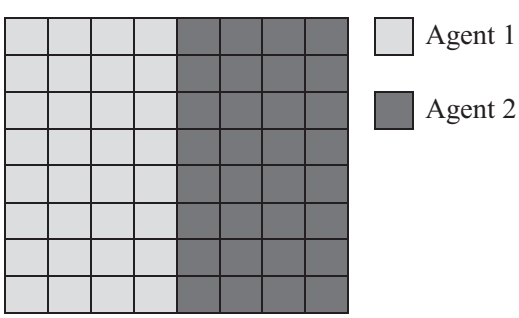

Figure 1. Low-fragmentation ownership configuration.

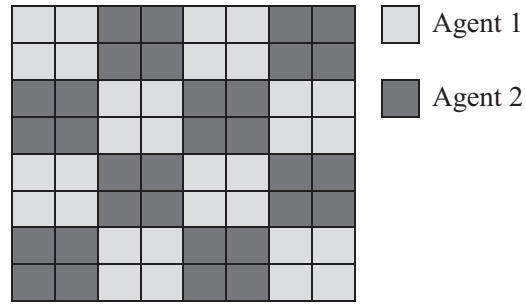

Figure 2. High-fragmentation ownership configuration.

for stands owned by each agent. The same basis functions (Equation 11) were used for all of the scenarios:

$$\begin{aligned}\bar{V}^n_{j_n}\left(S^{x^n}|\theta^n\right) = {} & \theta^n_1 + \theta^n_2 \mathrm{LTV}_{j_n} + \theta^n_3 \mathrm{LTV}_{j_n}{}^2 + \theta^n_4 \mathrm{SR}_{j_n} \\ & + \theta^n_5 \mathrm{SR}_{j_n}{}^2 + \theta^n_6 \mathrm{LTV}_{j_n} \mathrm{SR}_{j_n} \\ & + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{7_a} \mathrm{SR}_{a \in \mathrm{ADJ}_{j_n}} + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{8_a} \mathrm{SR}^2_{a \in \mathrm{ADJ}_{j_n}} \\ & + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{9_a} \mathrm{SR}_{j_n} \mathrm{SR}_{a \in \mathrm{ADJ}_{j_n}} + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{10_a} A_{j_n} \mathrm{SR}_{a \in ADJ_{j_n}} \\ & + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{11_a} D_{a \in \mathrm{ADJ}_{j_n}} + \sum_{a \in \mathrm{ADJ}_{j_n}} \theta^n_{12_a} D_{a \in \mathrm{ADJ}_{j_n}} \mathrm{SR}_{a \in \mathrm{ADJ}_{j_n}}\end{aligned} \tag{11}$$

In this formula we include the following:

- $j_n = 1, \ldots J_n$ indexes the stands owned by agent n.
- $a \epsilon \mathrm{ADJ}_{j_n}$ is the set of stands that are directly adjacent (i.e., share an edge or corner) with stand j_n (Figure 3).
- LTV_{j_n} is the land and timber value of stand j_n according to the Faustmann rotation, which does not account for fire risk. It is calculated using the vegetation-growth simulations and cost equations described in the previous section.
- A_{j_n} is the age of stand j_n.
- SR_{j_n} is the predicted down-wind spread rate under extreme weather conditions. It is difficult to directly compare the likely effect of one fuel condition to another. This variable helps capture the risk created by fuel conditions on the landscape that facilitate fire spread.
- $D_{a \epsilon \mathrm{ADJ}_{j_n}}$ is a dummy variable describing whether an adjacent stand is owned by another agent.

This specification of the value function approximation includes fire risk from the first ring of adjacent stands (Figure 3). It does not incorporate all of the spatial information about potential fire risk; it is possible for a fire to travel from any stand on the landscape to any other stand if the fuel and weather conditions are right. However, nearby stands have a greater impact on fire risk than distant stands. On the representative landscape we created, preliminary experiments showed that the inclusion of more distant stand attributes did not improve the ability of the value function parameters to converge, or predict the expected value of the target stand.

The specification of the value function and the variables included will depend on the state space where it is applied. It will also depend on the objectives of the agent. A different set of landscape parameters, different models of fire behavior, or different management objectives will require a different value function specification. In cases where agents are asymmetric (e.g., industrial forest land and federal forest land), a different value function will need to be

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

$a_8 \in ADJ_{j_n}$	$a_1 \in ADJ_{j_n}$	$a_2 \in ADJ_{j_n}$
$a_7 \in ADJ_{j_n}$	Target stand j_n	$a_3 \in ADJ_{j_n}$
$a_6 \in ADJ_{j_n}$	$a_5 \in ADJ_{j_n}$	$a_4 \in ADJ_{j_n}$

Figure 3. Depicts adjacent stands whose attribute are included in the value function approximation.

specified for each agent based on the landscape characteristics that contribute to each agent's objective.

The coefficients for this value function were estimated using the algorithm described earlier and the Appendix; the resulting approximation for the value function, $\bar{V}^*_{j_n}\left(S^{x^n}_t|\theta^*\right)$, should accurately approximate the expected value of being in state $S^{x^n}_t$. To initialize the algorithm, we set $\theta^0_2 = 1$ and all other coefficients to 0 for our initial guess θ^0. This seemed to be a good first approximation because the Faustmann rotation has long been recognized as optimal in the absence of complicating factors such as fire risk (Samuelson 1976). A reasonable initial estimation of the value function is important for ensuring the convergence of the value function iteration algorithm to a good optimum. This is especially important in the context of timber harvest because most stand value depends on anticipated future value and not the current reward; there is a long interval between positive rewards that occur when trees are harvested.

Results

Evaluating the Value Function Approximation

Each value function should converge to a value/policy that is optimal given the other agents value/policy. The value function iteration algorithm used to generate each agent's value function coefficients does not provably converge to the true coefficient values except in very specific conditions, which are not met for this case (Powell 2007). Hence, it is important to assess whether the chosen specification of the value function accurately estimates the expected landscape value for each agent. We evaluated the value function approximation by addressing two questions: (1) Does each coefficient estimate converge to a stable value? (2) Are the landscape values returned by the value function approximation close to the landscape values generated by Monte Carlo simulations?

To answer the first question, we tracked the evolution of the value function coefficients. Coefficients should move rapidly from their initial value and then make small oscillations around a stable value. The coefficients will never completely converge because of stochastic ignition and weather events. The coefficients did evolve as we expected: in the early iterations, they make wide oscillations that trend toward their true value; in the later iterations, the coefficients make smaller oscillations around a fixed value with no apparent trend in at least the last thousand iterations. For a more complete discussion of the convergence of the value function coefficients, see Lauer et al. (2017). Each scenario leads to different landscape values and different value function coefficients. A table of the final values for these coefficients is included in the Appendix (Table A1).

Each agent operates on the same landscape facing the same incentives and underlying ecological processes, so the value function coefficients for each agent should converge to the same values. In practice, the coefficients converge to similar values, but they are not exactly the same. This happens because each agent controls a different part of the landscape, and they are affected by the stochastic events differently.

Each agent's value function is a prediction of the discounted stream of rewards that will be generated for the agent over an infinite time horizon from that starting state. To determine whether the value function approximations are accurate predictions, we created a set of 500 simulations of fire events drawn from distributions of ignition location, fire duration, and weather as described earlier for the same starting landscape, each simulation being 150 years. Timber harvest and fuel treatment were optimally chosen to solve Equation 10 for each agent in each time step. This set of simulations was completed for each scenario using the same set of fire events. Each individual simulation led to different states and different reward streams, because of the different weather events. We tracked the rewards earned in each period on each stand for every simulation and compared the discounted sum of the rewards for the portion of the landscape owned by each agent (Equation 12—V^n_r is the realized value for simulation $r = 1, \ldots 500$), to the value predicted by each agent's value function approximation, $\bar{V}^n$(Equation 13).

$$V^n_r = \sum_{t=0}^{150}\left[\delta^t \sum_{j_n=1}^{J_n} C_{j_n}\left(S_t, x^n_t\right)\right] + \delta^{150}\sum_{j_n=1}^{J_n}\bar{V}^n_{j_n}\left(S^{x^n}_{150}\right) \quad (12)$$

$$\bar{V}^n = \max_{x^n}\sum_{j_n=1}^{J_n} C_{j_n}\left(S_0, x^n\right) + \bar{V}^n_{j_n}\left(S^{x^n}_0|\theta\right) \quad (13)$$

The ending value, $\bar{V}^n_{j_n}\left(S^x_{150}\right)$, was computed using the value function approximation and the state of the ending landscape. $\bar{V}^n\left(S|\theta\right)$ cannot perfectly predict the value of a specific stream of rewards for the agent because of stochasticity, but it should be very close to the mean V^n_r generated by the simulations.

The landscape values predicted by the value function approximations, and the mean realized values for the landscape under each scenario (social planner, low-fragmentation, and high-fragmentation), are listed in Table 1. In every case, the difference between the predicted value of the landscape and the mean realized value is 3.14 percent or less. There will always be some difference between the predicted and mean realized values because of stochasticity. Other factors that may contribute to a larger difference are an imperfectly specified set of basis functions (several different specifications were tried in preliminary experiments) and inadequate exploration of the state space (see Appendix).

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

Fragmentation Effects

As the level of fragmentation increases, the mean realized value of the landscape decreases. Actions by one agent are more likely to create externalities for the other. The benefits realized by one agent for a particular action will not offset the externalities generated for the other agent, leading to a lower landscape value. The mean realized values for each scenario are listed in Table 1. We also conducted several pairwise statistical tests to determine whether the different outcomes for each scenario were statistically significant. We used Welch's *t*-test to assess whether the mean difference for each scenario was significant. Welch's *t*-test allows us to compare scenarios with unequal variances but assumes that data are normally distributed. We confirm the results of this test using a Wilcoxon signed rank test that does not rely on the assumption of normal distributions. We also report the results of a Kolmogorov–Smirnov (K-S) test for determining the likelihood that the data in each scenario are from the same distribution, and we report an effect size that determines the number of standard deviations between each scenario's mean. The results of these tests are reported in Table 2.

In the low-fragmentation configuration, there is only 0.2 percent loss of landscape value on average compared to the outcomes under the social planner, a difference that is not statistically significant according to the *t*-test. Each agent is able to manage their land without being significantly impacted by their neighbor. We also compared the outcome distributions for these two scenarios using the K-S test, which resulted in a *P*-value of .663, suggesting that the outcome under the social planner does not stochastically dominate the outcome on the low-fragmentation landscape.

In the high-fragmentation case, there is a 4 percent decrease in the mean overall landscape value compared to the social planner scenario. This difference is statistically significant (P = .00056 for the *t*-test). The K-S test suggests that the distribution of outcomes under the social planner is stochastically greater than the distribution of outcomes under the high-fragmentation scenario. We calculate an effect size of 0.2066 meaning that the mean outcome under the Social Planner is 0.2 standard deviation larger than the mean outcome on the highly fragmented landscape. Each individual agent is worse off in the high-fragmentation case because they must adjust their management to account for externalities generated by the other agent. The empirical cumulative distribution functions for the 500 simulation outcomes, as measured by the landscape net present value, are depicted for all three scenarios in Figure 4. These cumulative distribution functions show that the social planner and low-fragmentation scenarios stochastically dominate the high-fragmentation scenario.

Increasing fragmentation decreases a landowner's ability to manage the fuel conditions that generate fire risk. The agents respond to this increasing fire risk by harvesting earlier—a result predicted by Reed (1984). Fire arrival probability depends on the current state of the landscape, which leads to a distribution of harvest ages caused by each agent modifying their harvest activity based on the current landscape conditions. The distributions of harvest ages for the social planner compared to the low-fragmentation and high-fragmentation scenarios are shown in Figures 5 and 6 respectively. The mean harvest age for the social planner is 39.05 years; it is 38.04 years for the low-fragmentation scenario and 36.86 years for the high-fragmentation scenario.

Agents also modify their fuel-treatment behavior depending on the level of fragmentation. Agents incur the cost of fuel treatment for two primary reasons, to protect timber value on the stand, and to prevent fire from spreading to nearby stands. Under the social planner, fuel treatment occurs for all stands at age 20 when stands are beginning to produce merchantable timber, and the fuel conditions can be changed from moderate risk to low risk. Treatments also occur in young stands with high-spread-rate fuel models when they are adjacent stands with valuable timber. In the low-fragmentation scenario, fuel-treatment activity is similar to the social planner with slightly more fuel treatment in the young high-spread-rate fuel models. In the high-fragmentation scenario, there is a substantial reduction in fuel-treatment activity for both the young stands with high-spread-rate fuel models and the 20- to 25-year-old stands with moderate spread rate fuel models. Ownership fragmentation prevents the agent from capturing the increased value for adjacent stands created by fuel treatment, thus decreasing the marginal value of this action.

The level of fragmentation also has implications for ecological outcomes on the landscape. As an example, we compared the distribution of fires sizes that occur in the simulations. For the low-fragmentation scenario and the social planner case, the distribution of fire sizes is almost indistinguishable. This is confirmed using a K-S test; the *P*-value was equal to .95, which does not allow us to reject the null hypothesis that these distributions are equal. However, in the high-fragmentation case, the distribution of fire sizes shifted toward larger fires, as shown in Figure 7. This result is also confirmed using a K-S test which had a *P*-value of .000 when the high-fragmentation scenario was compared to the social planner case. In the high-fragmentation scenario, there is less incentive to treat fuel and manage the landscape for small fires, since a spreading fire may damage your neighbor instead of you. Additionally, earlier

Table 1. Mean net present value of simulation outcomes for the whole landscape.

	Mean realized value (V_r) ($ million)	Predicted value ($\bar{V}$) ($ million)	Percentage difference between predicted and mean realized value percentage Δ
Social Planner	15.075	15.137	0.41
Low-fragmentation	15.045	15.517	3.14
High-fragmentation	14.517	14.610	0.64

Table 2. Pairwise comparisons for fragmentation scenarios.

	Mean difference ($ million)	*P*-value of one-tail Welch's *t*-test	*P*-value of one-tail Wilcoxon signed rank test	*P*-value of one-tail Kolmogorov–Smirnov test	Effect size Cohen's *d*
Social Planner v. Low-fragmentation	0.030	.4287	.0707	.663	0.0114
Social Planner v. High-fragmentation	0.558	.0006	.0000	.0001	0.2066
Low-fragmentation v. High-fragmentation	0.526	.0010	.0000	.0000	0.1965

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

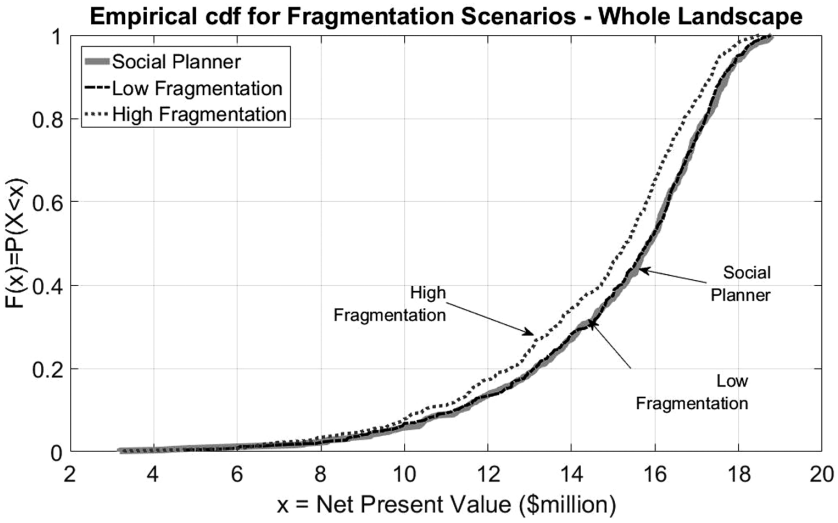

Figure 4. Figure depicts the empirical cdfs for the social planner case and the two fragmentation cases.

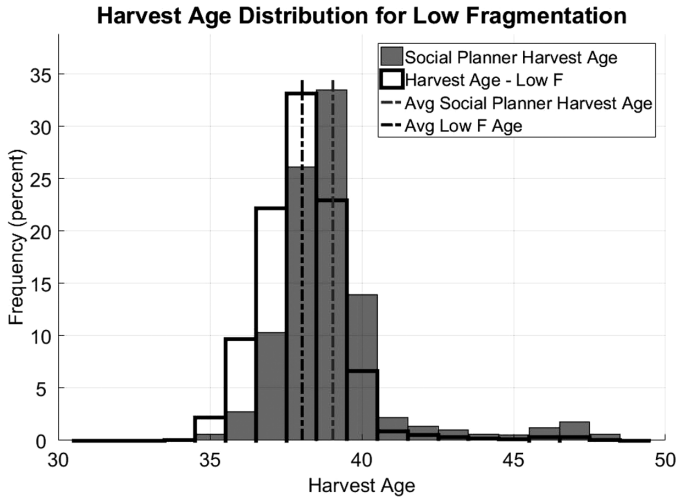

Figure 5. Distribution of harvest ages social planner vs. low-fragmentation landscape.

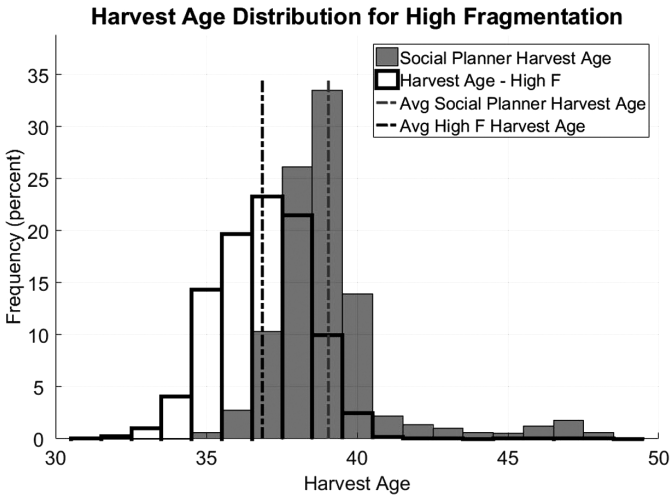

Figure 6. Distribution of harvest ages social planner vs. high-fragmentation landscape.

harvest is incentivized on the fragmented landscape, which leads to an increased occurrence of fuel conditions that generate high spread rates.

Discussion and Conclusions

Spreading fire creates spatial interactions that generate externalities affecting the welfare of forest landowners. The effects of these spatial interactions can be amplified or mitigated by human management. Property boundaries that fragment the landscape are one example of an institution that has important implications for landowner welfare and ecological outcomes. In this paper, we develop a method to analyze the interaction of multiple agents in a dynamic and spatial context. We use our method to demonstrate that ownership fragmentation can impede landowners' ability to manage their fire risk. We show that owning larger spatially continuous tracts of land will make it easier for landowners to mitigate fire risk created by externalities and achieve maximum landscape value. In cases where fragmentation cannot be avoided (e.g., O&C lands), cooperation between landowners could help to minimize externalities and lead to welfare improvements. Fischer and Charnley (2012) explore barriers to cooperation among nonindustrial private forest owners and highlight collective actions that could decrease fire risk.

The effect of ownership fragmentation will depend on ecological conditions specific to the site in question. Incorporation of more detailed, accurate, and site-specific ecological models is an obvious opportunity for future work. In this analysis, we modeled optimal management for timber as a financial asset under risk of stand-destroying fire. Fire is one example of a stochastic ecological process that creates spatial interactions—windthrow, the spread of invasive species, and the movement of wildlife are others. Likewise, maximum financial asset value of the forest is just one possible objective for forest management. For example, a landowner might instead want to manage fire to minimize threat to habitat connectivity or to meet other restoration objectives. The optimization framework demonstrated here is flexible enough to be adapted for these and other problems related to the management landscapes with complex ecological processes.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

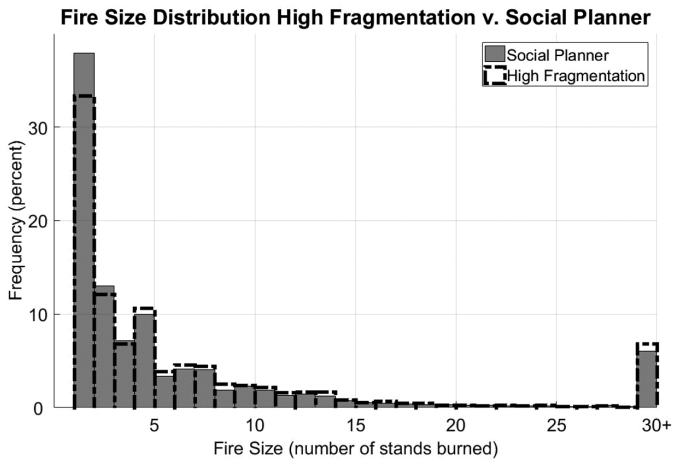

Figure 7. Fire size distribution social planner vs. high ownership fragmentation.

For the ecological conditions we modeled, ownership fragmentation led to larger fires on the landscape. This happened for a couple of reasons. First, as fragmentation increases, the marginal value of fuel treatment decreases, because landowners are unable to capture the value of reduced fire risk for all of the neighboring stands. Additionally, landowners are incentivized to harvest trees earlier because of the increased fire risk and their reduced ability to protect timber through fuel treatments in neighboring stands. More harvest leads to younger trees that spread fire more effectively.

The scale of ownership fragmentation relative to fire size is an important consideration. If the majority of damaging fires are larger than the ownership units, then fragmentation will likely have important ecological and welfare effects. Landowners will alter their management in reaction to the possibility of spreading fire from a neighboring landowner. Conversely, if most damaging fires are smaller than the units of ownership, fire will be less likely to cross property boundaries, and landowners' behavior will not be influenced by fire spread externalities. This result reinforces the assertion by Ager et al. (2017) that scale mismatches can result in less effective wildfire governance.

We modeled a risk-neutral wealth-maximizing landowner, which is appropriate if we assume large-scale corporate landowners for whom risk is spread over time and space. However, in some areas, the landscape is dominated by small woodland owners. This is more likely to be true on a fragmented landscape. These owners may be relatively risk-averse because their forest may represent a large portion of their wealth. A method for incorporating risk preferences into the objective function will be necessary for landowners who are not risk-neutral. A risk-averse agent may not pursue a Nash equilibrium strategy; instead, for example, they may choose a max–min strategy that maximizes the value of the worst possible outcome. In these cases, the optimization framework demonstrated in this paper would not be appropriate.

We demonstrated a tractable method for modeling spatial interactions in the framework of a dynamic game between multiple landowners who can create externalities for each other. The value functions we estimate can provide a measure of the cost of fire risk created by altering landscape conditions without damage needing to be explicitly realized. It allows the effect of management actions that change the state of the landscape to be characterized, and it provides a way to compare the known cost of an action to the expected benefits of that action on the landscape. The landscape and ecological process models we used, while simple, are far more realistic than in previous studies (e.g., Konoshima et al. 2008, 2010, Busby et al. 2012). The optimization platform we developed allows for introduction of more detail where it might be pertinent to the policy scenario under analysis. This method can help policymakers understand and predict how human institutions that interact with the landscape can affect landowner welfare in ecological systems where disturbance is a factor.

Appendix

Value Function Iteration Algorithm

The following is a detailed description of the algorithm used to estimate the coefficients (θ^n) for each agent's value function approximation. Figure A1 shows a visual depiction of the process, and a written description of each step is also provided.

Step 1

Define the landscape parameters, the initial state, and the initial estimation of the value function. The parameters and starting state are described in detail in the Data and Parameters section. The initial estimation of the value function $\bar{V}_{j_n}^{n,0}(S|\theta^{n,0})$ for each agent $n = 1 \ldots N$ is described in the Solution Method section.

The following series of steps (steps 2–8) is repeated until the maximum number of cycles (indexed by $h = 1 \ldots H$) is reached, or the stopping criteria are met.

Step 2

The current period is set to $t = 0$. The starting state for the cycle is determined; with some probability, it is the initial state defined in step 1, otherwise it is a randomly generated starting state. Choosing a different starting state allows the algorithm to learn from states it may not otherwise see, and is one strategy for ensuring adequate exploration of the state space. The number of periods T (years) in the cycle is defined. For early cycles, the number of periods is relatively small. This is because the landscape evolves over time based in part on actions chosen by the agent. These actions are determined based on the current estimation of the value function. We know that the initial approximation of the value function is likely to be wrong, and we expect it to improve as the approximation process progresses. It is important to expose the algorithm to states that are similar to those that are likely to occur under the optimal policy in order to accurately approximate the value function. In the early stage of the value function estimation process, going too many time steps into the future may lead to states that are unlikely to occur under the optimal policy. In later cycles, the number of periods is increased, i.e., $T(h)$, $\frac{dT}{dh} > 0$.

Step 3

In this step, the optimal action for the current state is determined for each agent using each agent's current value function approximation. Determining the optimal action in this case is a complex combinatorial problem. For each stand in the landscape, the agent can harvest timber, treat fuel, do both, or do neither. However, actions taken on a particular stand affect not only the value of that stand, but also the value of other stands on the landscape. On a 64-stand landscape, with two agents who each control 32 stands, 4^{32} different

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

possible combinations are possible for each agent. It is not computationally practical to evaluate each one; instead, a hybrid simulated annealing/tabu search algorithm (Kirkpatrick 1984, Reeves 1993) is employed to rapidly find a near-optimal solution.

Step 4

To avoid local optimum in determining the value function approximation, we explore states that might not otherwise be found by an agent using the current estimate of the value function. This is especially important in early iterations because early on, the value function approximation is not likely to be accurate and may incentivize actions that are not optimal. In each time step, there is some probability (P) that a small change will be made to the optimal action chosen for the agent in step 3. This probability decreases as the number of iterations increases, and the approximation of the value function improves $P(k)$, $\frac{dP}{dk} < 0$. Each agent's action (either the optimal action or modified action) is used to calculate their postdecision state. The postdecision states for each agent are used together in fire simulations and state transitions, and used for each agent individually in updating the value function approximation.

Step 5

In this step, we advance to the next time period (i.e., $t = t + 1$) and fire events are simulated. Multiple simulations ($m = 1 \ldots M$) are created, and each resulting state, S_t^m, is saved. A vector random ignition and weather variables (W^m) are drawn to determine whether and where fire occurs. Next, each agent's postdecision states $S_{t-1}^{x^n}$ are used along with the possible fire occurrence, W^m, and vegetation-growth models to generate a new state $S_t^m = S'(S_{t-1}^{x^1}, \ldots, S_{t-1}^{x^N}, W^m)$. The value of the resulting state for each agent will be used to update each agent's value function coefficients in the next steps. To prevent the value function update from overreacting to low probability events that are drastically different from the expected outcome, the number of simulations is higher in early cycles. After a number of cycles have been completed, the value function approximation should move toward a more accurate estimate, so the number of simulations per time step is decreased to get more variance in the observed outcomes and fine-tune the value approximation $M(h)$, $\frac{dM}{dh} < 0$.

Step 6

In this step, each agent's optimal action for each of the M states resulting from the fire simulations is determined using the same process described in step 3. This action is used to calculate the value of being in state S_t^m and determine each stand's contribution to that value $Q_{j_n}^n\left(S_t^m, X_t^{n,m^*} \mid \bar{V}_{j_n}^{n,k}\right)$. The value of being in the current state is also a realized value of being in the last period's postdecision state.

Step 7

In this step, ($k = k + 1$), and the value function approximation for each agent $\bar{V}_{j_n}^n$ is updated. This process uses the agent's postdecision state variable from previous period, the current approximation of the agent's value function, and the mean of the realized values of the previous postdecision state $Q_{j_n}^{n,avg} = \sum_M Q_{j_n}^n\left(S_t^m, X_t^{n,m^*}\right)/M$. The method used for updating the value function estimate is a stochastic gradient, which works as follows:

$$\bar{\theta}^{n,k} = \bar{\theta}^{n-1,k} - \alpha_{k-1}\left(\bar{V}_{j_n}^{n,k-1}(S_{t-1}^{x^n}) - Q_{j_n}^{n,avg}\right)\frac{\partial \bar{V}_{j_n}^{n,k-1}}{\partial \theta}$$

$\bar{\theta}^{n,k}$ is the vector of new parameter estimates for agent n's value function. These new parameters are determined by adjusting the current parameter value using the difference between the predicted value of the stand in the postdecision state landscape $\bar{V}_{j_n}^{n,k-1}(S_{t-1}^{x^n})$, and the mean of the realized or observed values of the stand in the postdecision state $Q_{j_n}^{n,avg}$, multiplied by the change in the value function for a change in the parameter. A learning rate α_{k-1}, a number between 0 and 1, also referred to as a step size, determines how much weight is placed on the most recent observation versus how much weight is placed on the current parameter value.

The value function we estimate, $\bar{V}_{j_n}^n$, is an approximation of the contribution a stand makes to the value of all stands owned by agent n. Since all stands on the landscape are homogeneous in our model landscape and use the same basis functions, it is possible to use every stand in the landscape to update the estimate of the coefficients. Since each agent on our landscape controls 32 stands, each stand's new coefficients are given a weight of 1/32 to determine the coefficients for the next iteration's value function. On a landscape with heterogeneous stands, the attributes that create heterogeneity would need to be included in the value function, or a separate value function approximation would need to be estimated for each stand.

Since the value function estimate $\bar{V}_{j_n}^n$is linear in θ^n, the derivative of the value function estimate with respect to the coefficient parameters is the vector of basis functions:

$$\bar{\theta}^{n,k} = \bar{\theta}^{n-1,k} - \alpha_{k-1}\left(\bar{V}_{j_n}^{n,k-1} - Q_{j_n}^{n,avg}\right)\begin{pmatrix}\phi^1\left(S^{x^n}\right)\\ \phi^2\left(S^{x^n}\right)\\ \ldots\\ \phi^I\left(S^{x^n}\right)\end{pmatrix}$$

The learning rate or step size is a very important factor in determining the ability of this algorithm to accurately approximate the value function. The step size can take many different forms, but it must have certain properties to guarantee convergence; these properties are met with a step size of $\alpha_n = 1/n$ (Powell 2007). However, the step size $1/n$goes to 0 too fast to get convergence in practice. This is because updates to the value function coefficients are initially made using incorrect estimates of the true value of the state. Because of this, it is important to weight later observations more heavily than early observations. Another important factor that leads to slower convergence is variance in the outcomes caused by the stochastic process. Events that deviate significantly from the true expected value may send the wrong signal if too much weight is placed on that observation. Therefore, it is important to carefully choose a step size rule. There are several different ways for choosing a step size, including constant step sizes or step size rules that decrease over a number of iterations to a target step size.

For this particular problem, we used the bias-adjusted Kalman filter rule outlined by Powell (2007). This rule chooses a step size by estimating the bias and variance of the value function approximation after k iterations. It increases the step size if bias is large and decreases the step size if variance is large. By increasing the step size when bias is large, the parameters move more quickly toward its true value. By decreasing the step size when the variance is large, the parameter updates are prevented from overreacting to any particular stochastic realization. Accounting for bias and variance, and adjusting the step size accordingly, leads to faster convergence.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

Step 8

If there are more time periods in the current cycle, one of the states created in step 5 is randomly chosen as the current state and the algorithm returns to step 3. If no time steps are left in the current cycle, and the stopping criteria have not been achieved, the algorithm returns to step 2. If the stopping criteria have been met (measured by the size of the change in the coefficient estimated for the value function approximation from the previous cycle), or if the specified number of cycle has been completed, the algorithm is stopped.

Table A1. Final value function coefficients.

Coefficients	Social planner	Low-fragmentation Agent 1	Low-fragmentation Agent 2	High-fragmentation Agent 1	High-fragmentation Agent 2
θ_1	–0.05178	–0.04450	–0.04830	–0.02315	–0.02297
θ_2	0.97401	0.98037	0.97876	0.99271	0.99392
θ_3	–0.01329	–0.01057	–0.00997	–0.00212	–0.00087
θ_4	–0.00622	–0.00761	–0.00636	–0.00468	–0.00322
θ_5	0.00080	0.00069	0.00097	0.00058	0.00144
θ_6	–0.01195	–0.00976	–0.01029	–0.01183	–0.00961
θ_{71}	–0.00512	–0.00375	–0.00475	–0.00209	–0.00262
θ_{72}	–0.00458	–0.00394	–0.00393	–0.00265	–0.00296
θ_{73}	–0.00502	–0.00439	–0.00347	–0.00204	–0.00357
θ_{74}	–0.00458	–0.00394	–0.00398	–0.00270	–0.00301
θ_{75}	–0.00512	–0.00376	–0.00483	–0.00220	–0.00263
θ_{76}	–0.00458	–0.00425	–0.00501	–0.00358	–0.00327
θ_{77}	–0.00571	–0.00487	–0.00553	–0.00329	–0.00241
θ_{78}	–0.00458	–0.00425	–0.00502	–0.00357	–0.00317
θ_{81}	0.00362	0.00264	0.00373	0.00229	0.00276
θ_{82}	0.00251	0.00291	0.00184	0.00201	0.00253
θ_{83}	0.00337	0.00305	0.00237	0.00239	0.00280
θ_{84}	0.00251	0.00292	0.00180	0.00201	0.00245
θ_{85}	0.00362	0.00264	0.00371	0.00216	0.00280
θ_{86}	0.00304	0.00294	0.00410	0.00241	0.00144
θ_{87}	0.00360	0.00406	0.00394	0.00224	0.00096
θ_{88}	0.00304	0.00292	0.00403	0.00250	0.00150
θ_{91}	–0.00060	0.00006	–0.00044	–0.00013	0.00063
θ_{92}	0.00000	–0.00115	–0.00055	–0.00071	–0.00065
θ_{93}	–0.00078	–0.00103	–0.00034	–0.00027	–0.00118
θ_{94}	0.00000	–0.00116	–0.00056	–0.00072	–0.00066
θ_{95}	–0.00060	0.00006	–0.00045	–0.00015	0.00063
θ_{96}	–0.00012	–0.00119	–0.00176	–0.00147	–0.00018
θ_{97}	–0.00073	–0.00101	–0.00175	–0.00045	0.00064
θ_{98}	–0.00012	–0.00119	–0.00178	–0.00146	–0.00016
θ_{101}	–0.00142	–0.00141	–0.00148	–0.00266	–0.00337
θ_{102}	–0.00064	–0.00081	–0.00032	–0.00160	–0.00212
θ_{103}	–0.00064	–0.00106	–0.00165	–0.00215	–0.00067
θ_{104}	–0.00064	–0.00077	–0.00043	–0.00175	–0.00224
θ_{105}	–0.00142	–0.00145	–0.00173	–0.00288	–0.00335
θ_{106}	–0.00167	–0.00153	–0.00224	–0.00192	–0.00139
θ_{107}	–0.00206	–0.00266	–0.00218	–0.00253	–0.00154
θ_{108}	–0.00167	–0.00157	–0.00225	–0.00194	–0.00115
θ_{111}	0.00000	0.00000	0.00000	–0.00009	–0.00079
θ_{112}	0.00000	0.00097	0.00280	0.00114	–0.00020
θ_{113}	0.00000	0.00151	0.00251	0.00049	–0.00136
θ_{114}	0.00000	0.00098	0.00276	0.00110	–0.00020
θ_{115}	0.00000	0.00000	0.00000	–0.00012	–0.00081
θ_{116}	0.00000	0.00238	0.00381	0.00036	0.00015
θ_{117}	0.00000	0.00287	0.00353	0.00062	–0.00090
θ_{118}	0.00000	0.00239	0.00379	0.00042	0.00020
θ_{121}	0.00000	0.00000	0.00000	–0.01166	–0.01147
θ_{122}	0.00000	–0.00087	–0.00280	–0.01158	–0.01142
θ_{123}	0.00000	–0.00087	–0.00280	–0.01285	–0.01230
θ_{124}	0.00000	–0.00087	–0.00280	–0.01157	–0.01155
θ_{125}	0.00000	0.00000	0.00000	–0.01149	–0.01149
θ_{126}	0.00000	–0.00250	–0.00229	–0.01158	–0.01142
θ_{127}	0.00000	–0.00250	–0.00229	–0.01030	–0.01067
θ_{128}	0.00000	–0.00250	–0.00229	–0.01157	–0.01155

*Agent 1 and Agent 2 had symmetrical objectives, but their coefficients do not converge to exactly the same values for two reasons: (1) each agent had a different starting landscape, and (2) there are multicollinearity issues in the value function model; age and timber value are highly correlated with fuel conditions. Multicollinearity does not reduce the predictive power of the model as a whole, but it can affect the estimation of individual coefficients within the model.

In cases where symmetrical agents exist, faster convergence could be achieved by incorporating information from both agents in the value coefficient updates, but on most landscapes, agents are not symmetric, and separate value functions and reward functions will need to be specified for each agent.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

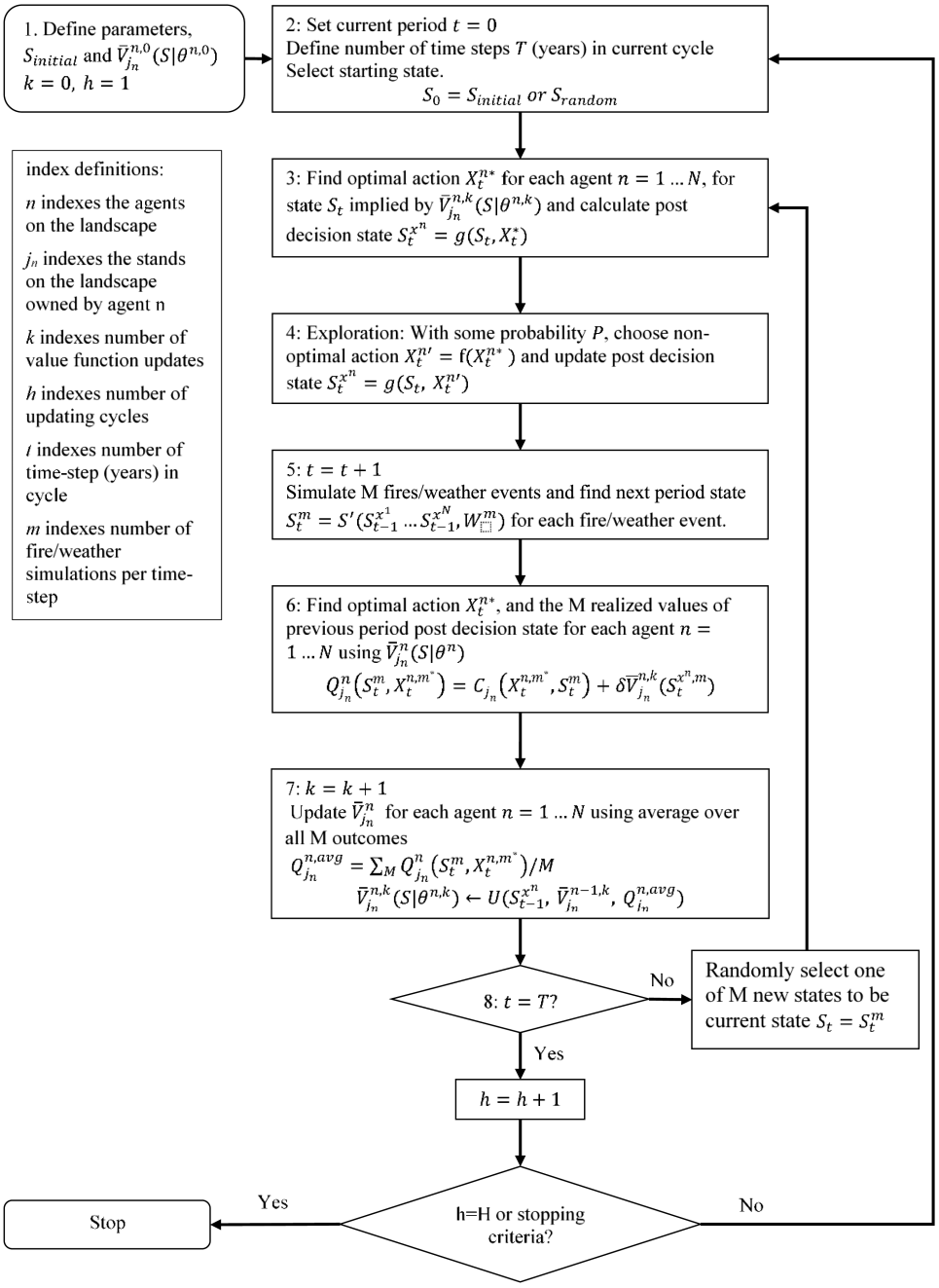

Figure A1. Value function iteration algorithm.

Literature Cited

Ager, A.A., N.M. Vaillant, and M.A. Finney. 2010. A comparison of landscape fuel treatment strategies to mitigate wildland fire risk in the urban interface and preserve old forest structure. *For. Ecol. Manage.* 259(8):1556–1570.

Ager, A.A., C.R. Evers, M.A. Day, H.K. Preisler, A.M.G. Barros, and M. Nielsen-Pincus. 2017. Network analysis of wildfire transmission and implications for risk governance. *PLoS One* 12(3):e0172867.

Ager, A.A., P. Palaiologou, C.R. Evers, M.A. Day, and A.M. Barros. 2018. Assessing transboundary wildfire exposure in the southwestern United States. *Risk Anal.* 38:2105–2127.

Amacher, G.S., A.S. Malik, and R.G. Haight. 2005. Not getting burned: The importance of fire prevention in forest management. *Land Econ.* 81(2):284–302.

Anderson, H.E. 1982. *Aids to determining fuel models for estimating fire behavior.* USDA Forest Service Gen. Tech. Rep. INT-122, Intermountain Forest and Range Experiment Station, Ogden, UT. 22 p.

Andrews, P.L., C.D. Bevins, and R.C. Seli. 2003. *BehavePlus fire modeling system: Version 2.0: User's guide.* USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Bellman, R.E. 1957. *Dynamic programming.* Princeton University Press, Princeton, NJ. 151 p.

Bradshaw, L.S., and E. McCormick. 2000. *FireFamily Plus user's guide.* USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Busby, G.M., H.J. Albers, and C.A. Montgomery. 2012. Wildfire risk management in a landscape with fragmented ownership and spatial interactions. *Land Econ.* 88(3):496–517.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020

Busoniu, L., R. Babuska, and B. De Schutter. 2008. A comprehensive survey of multiagent reinforcement learning. *IEEE Trans. Syst. Man Cybern. C Appl. Rev.* 38(2):156–172.

Calkin, D., and K. Gebert. 2006. Modeling fuel treatment costs on forest service lands in the Western United States. *West. J. Appl. For.* 21(4):217–221.

Chung, W., G. Jones, K. Krueger, J. Bramel, and M. Contreras. 2013. Optimizing fuel treatments over time and space. *Int. J. Wildland Fire* 22(8):1118–1133.

Crowley, C.S.L., A.S. Malik, G.S. Amacher, and R.G. Haight. 2009. Adjacency externalities and forest fire prevention. *Land Econ.* 85(1):162–185.

Daigneault, A., M. Miranda, and B. Sohngen. 2010. Optimal forest management with carbon sequestration credits and endogenous fire risk. *Land Econ.* 86(1):155–172.

Dixon, G.E. 2002. *Essential FVS: A user's guide to the forest vegetation simulator.* USDA Forest Service, Forest Management Service Center, Fort Collins, CO.

Faustmann, M. 1849. Calculation of the value which forest land and immature stands possess for forestry, W. Linnard (trans.). In *Martin Faustmann and the evolution of discounted cash flow*, M. Gane (ed.). University of Oxford, Oxford.

Fight, R.D., C.B. LeDoux, and T.L. Ortman. 1984. *Logging costs for management planning for young-growth coast Douglas-fir.* USDA Forest Service Gen. Tech. Rep. PNW-GTR-176 Pacific Northwest Forest and Range Experiment Station, Portland, OR. 10 p.

Fischer, A.P., and S. Charnley. 2012. Risk and cooperation: Managing hazardous fuel in mixed ownership landscapes. *Environ. Manage.* 49(6):1192–1207.

Garcia-Gonzalo, J., T. Pukkala, and J.G. Borges. 2014. Integrating fire risk in stand management scheduling. An application to maritime pine stands in Portugal. *Ann. Oper. Res.* 219(1):379–395.

Hobbs, S.D., S.D. Tesch, P.W. Owston, R.E. Stewart, J.C. Tappeiner II, and G.E. Wells. 1992. *Reforestation practices in Southwestern Oregon and Northern California.* Forest Research Laboratory, Oregon State University, Corvallis, OR.

Kirkpatrick, S. 1984. Optimization by simulated annealing: Quantitative studies. *J. Stat. Phys.* 34(5–6):975–986.

Konoshima, M., H.J. Albers, C.A. Montgomery, and J.L. Arthur. 2010. Optimal spatial patterns of fuel management and timber harvest with fire risk. *Can. J. For. Res.* 40(1):95–108.

Konoshima, M., C.A. Montgomery, H.J. Albers, and J.L. Arthur. 2008. Spatial-endogenous fire risk and efficient fuel management and timber harvest. *Land Econ.* 84(3):449–468.

Lauer, C.J., C.A. Montgomery, and T.G. Dietterich. 2017. Spatial interactions and optimal forest management on a fire-threatened landscape. *Forest Policy Econ*. 83:107–120.

Littman, M.L. 2001. Value-function reinforcement learning in markov games. *Cogn. Syst. Res.* 2(1):55–66.

Oregon Department of Forestry. 2016. *Timber sales – log prices. Available online at* https://www.oregon.gov/ODF/Working/pages/TimberSales.aspx; last accessed May 2016.

Powell, W.B. 2007. *Approximate dynamic programming: Solving the curses of dimensionality*, Vol. 703. Wiley, New York.

Powell, W.B. 2009. What you should know about approximate dynamic programming. *Nav. Res. Logist.* 56(3):239–249.

Puterman, M.L. 1994. *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, New York.

Reed, W.J. 1984. The effects of the risk of fire on the optimal rotation of a forest. *J. Environ. Econ. Manage.* 11(2):180–190.

Reed, W.J. 1989. Optimal investment in the protection of a vulnerable biological resource. *Nat. Resour. Model.* 3(4):463–480.

Reed, W.J. 1993. The decision to conserve or harvest old-growth forest. *Ecol. Econ.* 8(1):45–69.

Reeves, C.R. 1993. *Modern heuristic techniques for combinatorial problems.* John Wiley, New York.

Reinhardt, E., and N.L. Crookston. 2003. *The fire and fuels extension to the forest vegetation simulator.* USDA Forest Service General Technical Report, RMRS-GTR-116, Rocky Mountain Research Station, Ogden, UT. 209 p.

Routledge, R.D. 1980. The effect of potential catastrophic mortality and others unpredictable events on optimal forest rotation policy. *For. Sci.* 26(3):389–399.

Row, C., H. Fred Kaiser, and J. Sessions. 1981. Discount rate for long-term forest service investments. *J. For.* 79(6):367–376.

Samuelson, P.A. 1976. Economics of forestry in an evolving society. *Econ. Inq.* 14:466–492.

Spies, T.A., E.M. White, J.D. Kline, A.P. Fischer, A. Ager, J. Bailey, J. Bolte, et al. 2014. Examining fire-prone forest landscapes as coupled human and natural systems. *Ecol. Soc.* 19(3):9.

Sutton, R.S., and A.G. Barto. 1998. *Introduction to reinforcement learning.* 1st ed. MIT Press, Cambridge, MA.

Tappeiner, J.C., D.A. Maguire II, and T.B. Harrington. 2007. *Silviculture and ecology of western US forests.* Oregon State University Press, Corvallis, OR.

Thorpe, D. 2011. *Boot prints: A centennial summary of activities and events of Oregon Department of Forestry in Jackson and Josephine Counties.* Southwest Oregon District, Oregon Department of Forestry, Central Point, OR.

Van Roy, B., D.P. Bertsekas, Y. Lee, and J.N. Tsitsiklis. 1997. A neuro-dynamic programming approach to retailer inventory management. P. 4052–4057 in *Proceedings of the 36th IEEE Conference on Decision and Control*. Vol. 4.

Watkins, C.J., and P. Dayan 1992. Q-learning. *Mach. Learn*. 8(3-4):279–292.

Wei, Y., D. Rideout, and A. Kirsch. 2008. An optimization model for locating fuel treatments across a landscape to reduce expected fire losses. *Can. J. For. Res.* 38(4):868–877.

Yoder, J. 2004. Playing with fire: Endogenous risk in resource management. *Am. J. Agric. Econ.* 86(4):933–948.

Yoder, J., M. Tilley, D. Engle, and S. Fuhlendorf. 2003. Economics and prescribed fire law in the United States. *Rev. Agric. Econ.* 25(1):218–233.

Zaimes, G.N., M. Tufekcioglu, A. Tufekcioglu, S. Zibtsev, R. Corobov, D. Emmanouloudis, R. Uratu, A. Ghulijanyan, A. Borsuk, and I. Trombitsky. 2016. Transboundary collaborations to enhance wildfire suppression in protected areas of the Black Sea Region. *J. Eng. Sci. Technol. Rev.* 9(2):108–114.

Downloaded from https://academic.oup.com/forestscience/article/66/4/443/5487863 by Julie Blankenburg user on 30 September 2020