UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ADVANCED TECHNIQUES IN FOREST MANAGEMENT UNDER CONDITIONS OF FIRE UNCERTAINTY

TESIS PARA OPTAR AL GRADO DE
DOCTOR EN SISTEMAS DE INGENIERÍA

JAIME ADRIÁN CARRASCO BARRA

PROFESOR GUÍA:
ANDRÉS WEINTRAUB POHORILLE

MIEMBROS DE LA COMISIÓN:
MARCELO OLIVARES ACUÑA
VLADIMIR MARIANOV KLUGE
ANDREAS WIESE

SANTIAGO DE CHILE
2019

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTOR EN SISTEMAS DE INGENIERÍA
POR: JAIME ADRIÁN CARRASCO BARRA
FECHA: 2019
PROF. GUÍA: ANDRÉS WEINTRAUB POHORILLE

ADVANCED TECHNIQUES IN FOREST MANAGEMENT UNDER CONDITIONS OF FIRE UNCERTAINTY

The general objective of this thesis is to address the problem of forest fires from a preventive perspective, mainly motivated by the increase of this phenomenon worldwide and exacerbated by climate change. The current incidents indicate that preventive policy measures must be taken to reduce the risk of fire occurrence, managing the land in an effective way to protect natural forests, agricultural areas, and human lives. However, the problem of deciding where, when, and how to perform fuel treatments is a very complex problem and still remains open.

To address this problem, in the first two chapters (Chapter 2 and 3), we will develop a Cell-based Fire Growth Model, suitable for decision making that includes different sources of fire uncertainty such as: (1) ignition point(s) selected via a user-defined probability distribution; (2) a coefficient of variation (cv_{ROS}) capturing the stochastic aspects of Rate of Spread(ROS); and (3) a set of user-generated weather stream files (scenarios) that can be provided to the simulator. Using the tools mentioned in the previous paragraph, in Chapter 4 a Decision Support System (DSS) is developed that integrates fire simulation and fuel treatment decision making to minimize wildfire losses. We introduce the Downstream-Protection-Value (DPV) concept, an adaptable metric that ranks units of the landscape via their impact on fire propagation by modeling a forest as a network and the fire as multiple tree graphs. Our DPV metric stands out in all experiments, rapidly detecting fire ignition areas, preventing fire occurrence and decreasing the expected area burned by more than half compared to state-of-the-art alternatives. We test our methodology using public data of real forest patches from Alberta and British Columbia provinces, Canada.

In Chapter 5, we provide a device to improve the performance of our simulator by adjusting its internal input parameters in order to obtain better and more reliable fire scars, through a robust methodology and easy-and-fast implementation based in Derivative-Free Optimization Algorithms. Also, we discuss how this approach could be used to develop simulators that learn from the data provided in real time from an ongoing fire; and also from historical fire scars.

Finally in Chapter 6, we developed a methodology based on fire risk, to delimit the Wildland Urban Interface for Chile. To meet this objective, we first built a Bagged Decision Tree (BDT) model to quantify the risk of fire occurrence from different variables, grouped into: Human Activity, Geographic and Topographic, and Land Cover. Subsequently, using the mathematical elements underlying the BDT models, we determine individual relationships between the risk and the variables considered to address the problem of delineating a WUI-map.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTOR EN SISTEMAS DE INGENIERÍA
POR: JAIME ADRIÁN CARRASCO BARRA
FECHA: 2019
PROF. GUÍA: ANDRÉS WEINTRAUB POHORILLE

ADVANCED TECHNIQUES IN FOREST MANAGEMENT UNDER CONDITIONS OF FIRE UNCERTAINTY

El objetivo general de esta tesis es abordar el problema de los incendios forestales desde una perspectiva preventiva, motivado principalmente por el aumento de este fenómeno en todo el mundo y exacerbado por el cambio climático. Los incidentes actuales indican que se deben tomar medidas preventivas para reducir el riesgo de incendio, gestionando la tierra de forma efectiva para proteger los bosques naturales, las áreas agrícolas y las vidas humanas. Sin embargo, el problema de decidir dónde, cuándo y cómo realizar los tratamientos de combustible forestal es un problema muy complejo y aún permanece abierto.

Para abordar este problema, en los primeros dos capítulos (Capítulo 2 y 3), desarrollaremos un modelo propagación espacial para simular los incendios forestales, adecuado para la toma de decisiones que incluye diferentes fuentes de incertidumbre tal como: (1) punto(s) de ignición aleatorios; (2) un coeficiente de variación (cv_{ROS}) que captura los aspectos estocásticos de la tasa de propagación (ROS); y (3) un conjunto de escenarios climáticos, los cuales se pueden proporcionar al simulador para obtener distintas métricas de riesgo.

En el Capítulo 4, se desarrolla un Sistema de Apoyo a la toma de Decisiones (DSS) que integra la simulación de incendios y el tratamiento de combustibles para minimizar las pérdidas por incendios forestales. Introducimos el concepto Downstream-Protection-Value (DPV), una métrica adaptable que clasifica las unidades del paisaje a través de su impacto en la propagación del fuego al modelar un bosque como una red, y la propagación del fuego como múltiples grafos dirigidos. DPV se destaca en todos los experimentos con respecto a las alternativas de vanguardia. Probamos nuestra metodología utilizando datos públicos forestales reales de las provincias de Alberta y British Columbia, Canadá.

En el Capítulo 5, proporcionamos un dispositivo para mejorar el rendimiento de nuestro simulador ajustando sus parámetros de entrada internos para obtener cicatrices de incendio mejores y más confiables, a través de una metodología robusta y una implementación fácil y rápida basada en Derivative-Free Optimization algorithms. Además, discutimos cómo este enfoque podría usarse para desarrollar simuladores que aprendan de los datos proporcionados en tiempo real de un incendio; y también de cicatrices históricas de incendios.

Finalmente en el Capítulo 6, desarrollamos una metodología basada en el riesgo de incendio, para delimitar la Inferfaz urbano-rural para Chile. Para cumplir con este objetivo, primero creamos un Bagged Decision Tree (BDT) model para cuantificar el riesgo de ocurrencia de incendios a partir de diferentes variables, agrupadas en: Actividad humana, geográfica y topográfica, y cobertura del suelo. Posteriormente, utilizando los elementos matemáticos subyacentes a los modelos BDT, determinamos las relaciones individuales entre el riesgo y las variables consideradas para abordar el problema de delinear un mapa WUI.

A mi hija Amanda Carrasco Strehlow, soporte indispensable en todo este proceso.

Acknowledgements

Agradezco en especial a mi profesor guía Andrés Weintraub. Y agrego que más que un profesor, ha sido un maestro y amigo en todo este proceso. Definitivamente es un premio haberlo conocido y haber tenido la oportunidad de ser su alumno. Muchas gracias.

Agradezco a mi familia en el amplio sentido de la palabra: sanguínea y elegida por la vida, por la paciencia y el apoyo incondicional, que me dio la fortaleza en varios momentos que me acechaba el pesimismo y flaqueaban mis fuerzas. Gracias a ti mamá y a ti papá, gracias hermano Esteban.

Agradezco también a Linda Valdez, Fernanda Melis, siempre dándome ánimo y ayudándome en todo, son excelentes profesionales. Doy gracias al ISCI, en especial a Karla Jaramillo, muchas gracias por tu apoyo y tu tremenda amabilidad.

Contents

List of Tables

List of Figures

Chapter 1

General Introduction

The relationship between increased fires and climate change intensifies the need for a paradigm shift in our relationship with forest fires (Moritz et al., 2014). From the view that forest fires are an inevitable and natural process, how do we mitigate social and ecological damage? Should we focus our efforts on suppression? Or should we take preventive measures? or both, in a combined effort? Looking at recent world events, everything indicates that measures must be taken beforehand. From the point of view of prevention policies, we could take different measures, such as reducing the risk of fire — see for example (Curt et al., 2016a) — as well as helping in the management of fuel in the Urban-Forest Interface (WUI) and/or in large areas of forest (Calviño-Cancela et al., 2016).

One way to mitigate the uncontrolled effect of fires and at the same time protect our communities and natural resources is through *forest fuel management*. Activities such as cutting and cleaning wood, prescribed burns, commercial harvesting and thinning can promote the reduction of the danger of fire (Agee and Skinner, 2005; Loehle, 2004). In addition, these actions could also benefit the diversity of species (Shinneman et al., 2019), improve the health of the forest ecosystem and help in the extinction of fires, among many others benefits (Finney, 2001; Hirsch et al., 2001b; Moghaddas and Craggs, 2008). However, land planners continually face difficult questions as what, where, when and how to act (Chung, 2015).

In practice, for a given planning horizon, depending on the state of the forest, the planners decide where to locate a treatment. The condition of the forest depends on the previous efforts of the land managers (from the previous year for example) and possible modifications of the vegetation cover due to the fires. Now, where and when the fires will occur and what will be the severity, these are stochastic concatenated events. Therefore, the question we can aim to answer is in expected terms: what are the best decisions so that the loss due to fires is minimal? However, regardless of who or how we make the decisions, can we test their effectiveness in the newly exposed framework?.

In this thesis, we address the problem of locating harvest decisions under fire risk, so that these actions minimize the damage of possible future fires. The studies and tools incorporating fire risk into the selection of an spatially explicit optimal plan are still insufficiently addressed, mainly because they: i) usually deal with very simple forest landscapes, or ii)

they are based on the simulation of a limited number of fires to adjust the fire risk during the planning period, or iii) they use simplified fire spread simulators that do not allow running multiple simulations over large landscapes, among others. — see Bettinger (2009); Kim (2009); Konoshima et al. (2010); González-Olabarria and Pukkala (2011)

Given this context, in Chapter 2 we develop a fire growth simulator suitable for this purpose, using a modeling of fire behavior that includes the most important variables of a landscape: topography, forest fuel and climate. In this study we rely on the Canadian Fire Behavior Prediction (FBP) System. However, we are currently extending our tools to manage forests in Chile and Spain.

Subsequently, in Chapter 3, we incorporate three sources of uncertainty to develop an efficient and more realistic fire growth simulator. This extension is also aimed to provide us with different fire risk metrics, based on the fire behavior from multiple simulations. In particular, we include stochasticity by: ignition point(s) selected via a user-defined probability distribution; a perturbation level representing the stochastic aspects of the predicted *Rate of Spread* (ROS) of fire, accounting its inherent approximation error as well as allowing the user to obtain different (and average) fire scars including uncertainty in the fire dynamic; and a set of user-generated weather stream files (scenarios) with specific probabilities can be provided to our simulator.

In Chapter 4, we address the problem of evaluating the effectiveness of different metrics to prioritize the location of harvest decisions. For this, we propose an integrated simulation and optimization system that reports at the end of the process, the expected loss due to fires in each period of our planning horizon. In addition, the system could evaluate any empirical decision, for example, a "myopic" decision maker that does not take fire risk into account. Our methodology proposes a friendly framework with projections to be applicable to real situations. Comparisons for different levels of demand are made in order to show the mean as well as the standard deviation for each criterion.

In Chapter 5, not far from the topics of the first three chapters, we propose a new methodology to improve the performance of forest fire simulators. In practical applications, it is common that wildfire simulators do not correctly predict the evolution of the fire scar. This is due to multiple factors including inaccuracy in the input data such as land cover classification, moisture, improperly represented local winds, among many others. We address this problem using Derivative-Free Optimization (DFO) algorithms for adjusting the Rate of Spread (ROS) factors. To achieve this, we solve an error minimization optimization problem that captures the difference between the simulated and observed fire, which involves the evaluation of the simulator output in each iteration as part of a DFO framework, allowing us to find the best possible factors for each fuel present on the landscape.

Finally in Chapter 6, we develop a methodology based on fire risk, to delimit the Wildland Urban Interface for Chile. To meet this objective, we first built a Bagged Decision Tree (BDT) model to quantify the risk of fire occurrence from different variables, grouped into: Human Activity, Geographic and Topographic, and Land Cover. Subsequently, using the mathematical elements underlying the BDT models, we determine thresholds to map the WUI at national and subnational scales, based on historical fire occurrence data. We finish this chapter, presenting a first WUI-map for Chile.

Chapter 2

Cell2Fire: A Cell Based Forest Fire Growth Model

Authors

Jaime Carrasco and Andrés Weintraub
University of Chile, Industrial Engineering Department

Cristobal Pais
University of California Berkeley, IEOR Department

David L. Martell
University of Toronto, Faculty of Forestry

David L. Woodruff
University of California Davis, Graduate School of Management

Abstract

Cell2Fire is a new cell-based forest and wildland landscape fire growth simulator that is open-source and exploits parallelism to support the modelling of fire growth cross large spatial and temporal scales in a timely manner. The fire environment is characterized by partitioning the landscape into a large number of cells each of which has specified fuel, weather, fuel moisture and topography attributes. Fire spread within each cell is assumed to be elliptical and governed by spread rates predicted by a fire spread model such as the Canadian Forest Fire Behavior Prediction (FBP) System. The simulator includes powerful statistical and graphical output and spatial analysis features to facilitate the display and analysis of projected fire growth.

In this chapter, we validated Cell2Fire by using it to predict the growth of real and realistic hypothetical fires, comparing our fire growth predictions with those produced by the state-of-the-art Prometheus fire growth simulator. Cell2Fire is structured to facilitate its use for predicting the growth of individual fires or embedding it in landscape management simulation models. It can be used to produce probabilistic fire scar predictions by allowing for uncertainty concerning the basic spread rate predictions and uncertain weather scenarios that might drive their growth.

2.1 Introduction

The effects of global warming on temperature, precipitation levels, soil moisture and other forest and wildland fire regime drivers have increased and are expected to continue to increase both the number of and area burned by wildfires around the globe (Westerling, 2016). Wildfires have burned large areas and important infrastructure, thousands of homes and forest resources have been destroyed and many lives have been lost in recent years. Recent examples include catastrophic incidents in the United States, Canada, Chile, Portugal and southwestern Australia in the years 2016-2018. That has also resulted in increases in expenditures by forest and wildland fire management agencies (see, e.g., Stocks and Martell, 2016b). Despite concerted efforts, wildfire growth remains a complex and very difficult to model process.

Two of the most important characteristics of a wildfire are its rate of spread (ROS) and intensity which are influenced by fuel type, fuel moisture, wind velocity, and slope. The Canadian Forest Fire Behavior Prediction (FBP) System includes empirical fire spread rate models that can be used to predict the rate of spread and the intensity of wildfires based on weather, fuel moisture, time of year and topographical variables for specified fuel types; e.g., for individual grid cells that contain homogeneous fuel types (Forestry-Canada, 1992).

However, the FBP System alone cannot be used to predict how a fire will grow across a heterogeneous landscape/grid over time. Spatial fire growth models like Prometheus, a deterministic fire growth simulator, are designed to use FBP spread rates to do so (Tymstra et al., 2010b). Prometheus is a vector-based fire growth simulation model that is based on

an adaptation of Huygens' principle of wave propagation, i.e., the propagation of the fire front is modelled in a fashion similar to a wave, shifting and moving forward continuously in time and space. It uses spatially explicit fire environment input data concerning topography (slope, aspect and elevation) and FBP fuel types along with a weather stream and fire danger rating codes and indices,

FARSITE Finney (2005) is another widely used fire growth simulator. It is based on the U.S. Forest Service's BEHAVE fire behavior prediction system and is also a vector-based Huygens' type model. A review of twenty-three simulators that can be used to predict forest fire growth can be found in Papadopoulos and Pavlidou (2011a). The two models found to best simulate historical fires were FARSITE in the United States and Prometheus in Canada.

Recent years have witnessed growing interest in the development of detailed cell-based deterministic/stochastic fire simulators and some of the modelling assumptions that have been adopted include, for example the use of memoryless distributions (Markovian processes) to model the fire spread dynamics (Boychuk et al., 2009a), homogeneous forests (cells' characteristics are identical), reductions in the number of adjacent cells to which a fire can spread (e.g., from 8 to 6 or 4), no spotting, and no stochasticity is included.

Our goal was to include many realistic aspects of fire behaviour in Cell2Fire — an open-source (`https://github.com/cell2fire/Cell2Fire`) cell-based fire growth simulator, with a view to achieving high computational performance via parallelism when simulating large-scale fire instances to provide valuable insight to inform both fire and forest management. We therefore use both real and realistic hypothetical fire instances to validate our simulator and assess its computational performance.

FireSmart (Hirsch et al., 2001b) forest and fuel management calls for landscape management and fire growth simulation models that have well-structured interfaces that facilitate the exchange of data between them to inform the iterative re-planning that takes place when strategic plans are modified in response to what fires materialize and what areas actually burn over long planning horizons.

Acuna et al. (2010) demonstrated the importance of integrating fire management with strategic forest management planning models to develop and evaluate FireSmart forest management plans. Although stand-alone fire growth models can be used to evaluate specified forest and fuel management plans, the fire growth simulators described above were designed to simulate fire growth and cannot readily be incorporated strategic planning frameworks that can be used to develop good or optimal strategic FireSmart landscape management plans.

Cell2Fire is designed specifically for use in a fuel-management framework with the intent of mitigating the detrimental impact of large fires efficiently. Cell2Fire can therefore be used as a pure simulation tool to model the growth of specific fires over a short time frame and/or embedded in a landscape management framework to evaluate fuel-management strategies or linked with optimization software to develop "optimal" fuel management strategies over long planning horizons.

The primary objective of our research was to develop an efficient and realistic fire growth

simulator that enables simulation of the dynamics of fire growth across a grid representation of a real or hypothetical forest using cell attributes such as fuel type, elevation and weather, given an ignition point or initial fire perimeter, to inform FireSmart forest management. Our ultimate goal is to develop a tractable methodology that can be used to generate realistic spatial fire scar scenarios that can be used to support fuel management and harvesting planning.

The chapter is organized as follows: In section 2, the two main fire growth simulation paradigms are described and a review of state-of-the-art simulators is included. Section 3 describes the Cell2Fire simulator structure, the main simulation steps, the fire growth dynamics model and the computational implementation. The results of a case study based on a real forest landscape in Canada and several test instances that were used to validate the simulator output and compare its computational performance with state-of-the-art simulators are discussed in section 4. Finally, section 5 contains our conclusions and thoughts concerning future research needs.

2.2 Description of fire growth dynamics

The two methods that have most often been used to simulate fire spread rates and fire growth across heterogeneous landscapes are the wave propagation approach and the cellular automata approach. We therefore begin by providing a brief overview of those two approaches.

2.2.1 Wave-propagation model: Huygens

Huygens considered every point on a wavefront of light as a source of individual wavelets and described the new wavefront as the surface tangential to the circumferences of the secondary waves. The use of Huygens' principle to simulate fire growth is based on the assumption that the shape of a fire can be represented by a polygon, a plane figure composed of a sequence of straight-line segments forming a closed path, whose vertices are a tangential envelope of the elliptical "firelets". Huygens' principle was first applied to the simulation of fire spread by Sanderlin and Sunderson (1975). Anderson et al. (1982) later developed a simple elliptical model based on Huygens' principle of wave propagation to simulate the spread of grass fires. Richards (1990) then extended the Anderson et al (1982) model by deriving a set of partial differential equations to model the growth of fires across a heterogeneous landscape.

Both the FARSITE (Finney, 2004) and Prometheus models use Richards (1990) partial differential equations to propagate each vertex on a fire's perimeter. However, the models differ in the fire danger rating system components and fuel models used to model fire spread rates. FARSITE uses the US National Fire Danger Rating System and fire behavior prediction fuel models developed by Rothermel (1972a) and extended by Anderson (1982) and Scott and Burgan (2005), whereas Prometheus uses the Canadian Forest Fire Danger Rating System.

2.2.2 Cell-based fire growth models

Cellular automata models that employ a raster-grid of square or hexagonal cells, are widely used to model wildfire spread. Fuel and terrain conditions are usually assumed to be homogeneous within each cell in order to simplify basic fire spread rate calculations. The fire propagates through the grid-cells basis, typically from a cell's center to the center of an adjacent cell. Each ignited cell behaves as an ignition source that is independent of any adjacent burning cells. To spread the fire from one cell to another, a search mechanism such as an adjacency or spread template is required.

Kourtz and O'Regan (1971) developed the first computer simulation model to spatially simulate the growth of a small fire. Their model was based on a heterogeneous and discontinuous fuel-type grid but did not account for the effects of terrain and wind. This deterministic model predicted how long it would take a fire to burn through one square area or cell within a fuel grid when the location of the fire, the starting time, and the grid resolution were known. Travel times were calculated using fixed rates of spread (based on the fuel type and the spread index for the day) and fixed spread directions from the burning cell. Later, O'Regan et al. (1973) developed a method for using directional rates of spread to predict fire growth. They also rewrote the original model for use on what was then a large computer to simulate fires of up to 15,000 ha. in size.

O'Regan et al. (1976) calculated average directional rates of spread using the equation:

$$ROS(\phi) = \begin{cases} \dfrac{a\,(1-\mathrm{e}^2)}{1-\mathrm{e}\cdot cos\phi}, & 270^{\mathrm{o}} < \phi < 90^{\mathrm{o}}, \\ a\,(1-\mathrm{e}^2)\,, & 90^{\mathrm{o}} \leq \phi \leq 270^{\mathrm{o}}, \end{cases} \tag{2.1}$$

where a and b are major axes, the quotient c/a is the eccentricity e and ϕ is the focal angle. Kourtz et al. (1977) then further modified this model to accommodate variation hourly wind conditions. Todd (1999) adapted the Kourtz et al. (1977) model to create an eight-point symmetric fire growth model called Wildfire, which incorporates FBP System spread rates. The features and functionalities of the Wildfire model were assessed and considered during the design of the Prometheus model (Tymstra et al., 2010b).

Boychuk et al. (2009a) developed a stochastic model of fire spread using a lattice Markov chain model, in which they associated probabilistic transition functions with each cell. Each of these cells interacts with its four nearest neighbors and a cell transitions from unburned to burning depending on the state of the neighboring cells. The use of a simplified cellular automata model describing the dynamics of fire spread on a heterogeneous landscape accounting for weather factors (wind speed and direction) as well as the type and density of vegetation was used to successfully model the Spetses Island fire, Greece was after tuning the main parameters of the simulator (Alexandridis et al., 2008a).

Recently, a fire growth simulator (Arca et al., 2019) was released with the goal of assisting civil protection and fire management agencies. In order to provide short term forecasts of how the fire will grow, they based their simulator on models of wind field progression. The fire growth model uses the level set technique (see Ghisu et al. (2015) and the Rothermel (1972a) fire behavior model.) The goal of our study, in contrast, is to not to support large fire

management operations but rather to inform strategic harvest planning and fuel management. We are interested in producing final fire scars that are reasonable. To do that, we need to have a reasonable evolution of the fire, but approximations with respect to minute-by-minute status are acceptable in the interest of computational efficiency because many scars may need to be generated to explore methods for strategic timber harvest scheduling and forest management planning or to validate the use of simpler methods for creating fire scar scenarios such as, e.g., Kuhlmann et al. (2014).

2.2.3 Canadian Fuel Behavior Prediction System: FBP

The Canadian FBP System is a set of empirical models that can be used to predict fire spread rate, fuel consumption and fire intensity in homogeneous fuel types as functions of fuel type, fuel moisture and current weather expressed in terms of the Canadian Forest Fire Weather Index System (FWI) codes and indices (Forestry-Canada, 1992). It includes fuel models that are used to classify forest types into 17 fuel types that collectively represent most of the major forest cover types in Canada. Figure 2.1 is based on a figure in Tymstra et al. (2010b) and shows the main inputs and outputs of the FBP System.

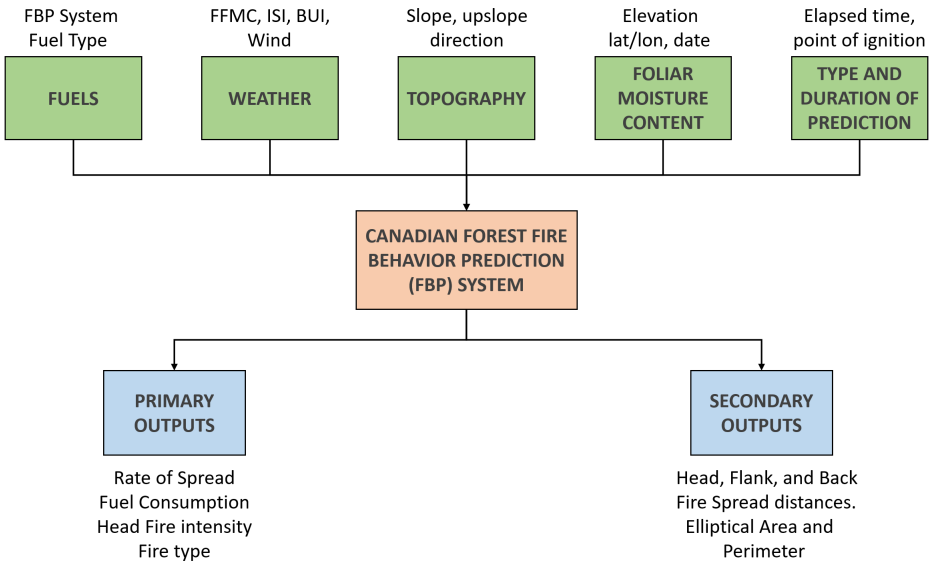


Figure 2.1: A conceptual diagram of the Canadian Forest Fire Behaviour Prediction (FBP) System.

2.3 Cell2Fire growth simulator

2.3.1 Description

Cell2Fire is an open-source cell-based fire growth simulator developed in Python and C++ for use on laptop or desktop computers as well as on High-Performance Computer (HPC)

systems. It allows a user to simulate fire growth across a grid that represents a real forest landscape using fire environment variables such as the fuel type, elevation (topographic/terrain components) of each cell, fire ignition points and weather.

A forest landscape is mapped into a rectangular region comprised of n rows and m columns partitioned into a series of identical area square cells to produce a grid where the cell size depends on the desired spatial resolution that will of course, be influenced by the spatial scale for which cell attribute data is available. Each cell represents a specific portion of the landscape and has two information layers that pertain to its topographic and fuel characteristics. Those layers define the characteristics of each independent cell, allowing the simulator to treat them as individual objects that can interact with other cells on the forest landscape to model fire growth.

Algorithmically, Cell2Fire simulates the growth of fire by tracking the state of all cells as the model progresses through discrete equally-spaced time steps. The status of the fire and all the cells on the landscape are updated at the end of every time step and smaller time steps lead to better fire growth precision.

A cell can be in one of five states: *"Available"*, *"Burning"* , *"Burned"*, *"Harvested"*, or *"Non-Fuel"* where the label *"Available"* indicates that the cell contains a flammable fuel type; *"Burning"* represents that the cell contains an active fire, *"Burned"* indicates that the fire has passed through the cell; *"Non-Fuel"* is a non-flammable fuel type such as rivers, lakes, or rocky. The state *"Harvest"* is provided so that the simulator can be embedded in a strategic harvest planning system. The harvest planning module would be responsible for labeling the cells that are harvested and provide the appropriate post-harvest fuel type.

Once an ignition point has been specified, the fire is ignited. During each time step the fire may spread along the axes emanating from the centre of each burning cell to its neighboring cells. The predicted FBP system Head Rate of Spread (HROS), Flank Rate of Spread (FROS) and Back Rate of Spread (BROS) are used to model elliptical fire growth within each cell with the focus of the ellipse at or near the center of the cell (see this process described in more detail in 2.3.7). The geometry of the ellipse is then used to predict the fire spread rates along the axes emanating from the centre of each cell.

A signal/message is then initiated to any adjacent cells whose center is reached by the fire. In the present implementation, it is assumed that each cell has at most 8 adjacent cells (see Figure 2.2) because the grid is assumed to be rectangular. These are the only neighbors considered because the simulation time step is assumed to be small enough to ensure that the fire can not spread beyond adjacent cells in one time step. For the examples we tested using a 100m by 100m cell size, simulation time steps below one minute resulted in almost no change in final fire scar compared to one minute. Shorter simulation time steps result in longer run times for the simulation.

Note that there is also an approximation because the fire enters a cell a from its neighbor and the neighbor's cell characteristics are used to model fire spread within that cell until the fire reaches the center of the destination cell. At that point,the characteristics of the destination cell take over.

When the fire spreads from the edge of a cell and reaches its center, the cell receives a message. When the cell receives a message it calculates what its ROS based on its characteristics and the current weather and if that ROS is greater than a user-specified parameter (which was zero in the experiments reported here), its state is labeled as *"Burning"*. At this point, the main Rate of Spread values are calculated by the FBP System module and fire progress begins to be calculated along with its available axes. Predicted ROS values for each axis are based on the assumption that fire spreads in the shape of an ellipse in each cell and the geometry of the ellipse and its orientation are used to predict spread rates along the axes emanating from the cell as described later in Section 2.3.7.

The fire's progress is updated at each fire time period by examining the state of all *"Burning"* cells. Once no adjacent cells are available or a burn-out criterion (See assumption (A5) in 2.3.2) has been satisfied the cell changes its status to *"Burned"* and is omitted from the simulation in further steps.

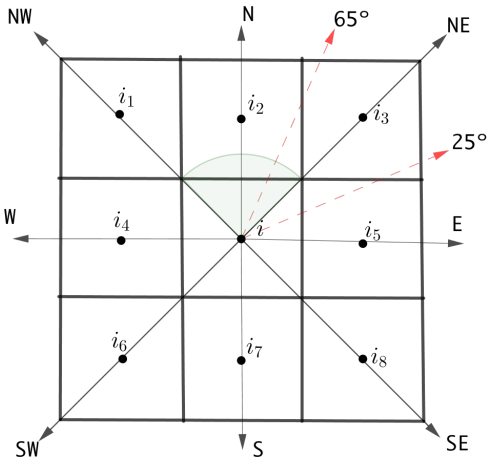


Figure 2.2: Neighbors cells.

This process is repeated until a fire ending event state is reached, i.e., the total number of fire burning periods or hours have passed or there are no more cells available to burn.

Three different sources of uncertainty can be included in our model to account for the most important perturbations that can significantly affect the fire growth dynamics: (1) ignition point(s) selected via a user-defined spatial probability distribution or else simply with equal probability for each cell, (2) a perturbation level the stochastic aspects of the basic FBP predicted Rates of Spread via its coefficient of variation, accounting for its inherent approximation error, allowing the user to obtain different fire scars taking into account uncertainty in the fire spread rates, and (3) a set of user-generated weather stream files (fire weather scenarios that can occur with specified probabilities can be provided to Cell2Fire, performing a series of simulations with different weather scenarios thereby allowing the user to study the behavior of the fire under different weather conditions for a specified forest/instance.

A series of relevant outputs are generated depending on the user needs: fire evolution maps (scars) at different time-precision resolutions, burn grids that indicate the final status

of each cell for statistical analysis as well as fire probability maps if there are probabilistic inputs, and a simulation log that documents all the interactions between cells and statistics of the final state of the forest.

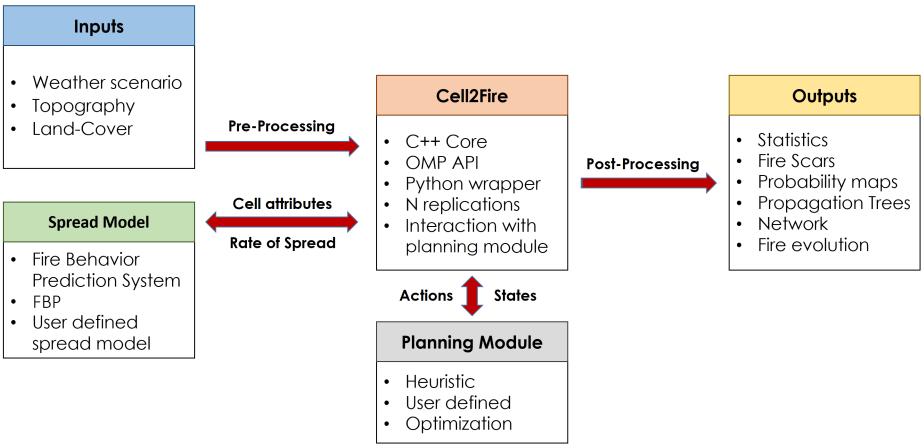


Figure 2.3: Simulation framework. (1) Raw data is pre-processed into Cell2Fire's format, (2) Cell2Fire calls an independent fire spread model (e.g., FBP), performing the simulations including harvesting plans provided by the user (if needed), and (3) finally, outputs are generated and returned to the user.

2.3.2 Modelling assumptions

Based on the previous description description of our model we can summarize the main simplifying assumptions upon which Cell2Fire is based as follows:

(A1) The growth of the fire depends on the "Rate of Spread (ROS)" of the fire from a burning cell towards its neighboring cells. We assume that a cell is ignited when the fire reaches its center and conditions for burning are met (see A5). Each cell has at most 8 neighbors.

(A2) The Rates of Spread along the 4 principal orthogonal axis (which are functions of the weather, fuel moisture conditions and characteristics of each cell) are calculated using the Canadian FBP System. Note that a the main axis of the ellipse that burns in each cell is aligned in the direction that the wind is blowing (the HROS direction) for that time step and the BROS is aligned in the opposite direction. The FROS is perpendicular to the wind direction (the two flanks). Note that other fire spread models could be used in lieu of the Canadian FBP system, a point we discuss in our Conclusions section.

(A3) Each cell that burns serves as a new source of fire. Fire spread occurs at the cellular level and cell size depends on the spatial resolution and corresponding data availability.

(A4) The effect of fire suppression action is not modelled. It is beyond the scope of this study and will be addressed in a future research project.

(A5) There are two sets of conditions for modeling the termination of fire growth (a fire ending event) in Cell2Fire: the cellular level and general fire evolution conditions. At a cellular level, each cell becomes unavailable (burned) for future fire dynamics if (i) the

ROS along the main available axes is less than some empirical threshold $\delta > 0$, (ii) the cell does not have any adjacent cells that are available to burn, (iii) the residual fuel available in a cell is not sufficient to support fire spread (implicit in the FBP system), or (iv) a user-defined head fire intensity (HFI) threshold is provided and the HFI is below that threshold.

Regarding the general fire dynamic, the total duration of the simulated wildfire event is determined by (1) maximum number of hours of burning per day — a season-dependent constant Parisien et al. (2005c), or drawn from a probability distribution – and (2) the total fuel remaining in the forest (available cells).

2.3.3 Cells

Cells that contain the information concerning the structure of the forest as well as the past and current state of the simulation are the main processing units in Cell2Fire. The main fields that define the state of a cell $i \in \mathcal{N}$ are the following:

i) Fuel type: following the classification criteria of the FBP System (Forestry-Canada, 1992), each cell is assigned a specific fuel type (e.g., conifer, grass, non-fuel) represented by an FBP code. This will be used for selecting the specific fire spread models and coefficients defined by the FBP System in order to predict the rate of rate of spread (ROS) in that cell.

ii) Slope: the slope % in terms of the vertical rise over the horizontal run and adjacent cells, and the upslope direction (radians) have a significant impact on the predicted ROS.

iii) Elevation: altitude in meters of the current cell with respect to the sea level.

iv) Location: latitude and longitude coordinates are provided with the instance.

v) Status: cells are classified into five different states during the simulation time steps. Cells that have not been harvested/burned and represent valid fuel types are labeled as *"Available"* cells. Cells that are actively burning are classified as *"Burning"* cells, updating the fire progress along each of its 8 axes every time-step. When a burned-out condition (see Section 3.2) is reached, burning cells are turned into *"Burned"* cells, meaning that they can no longer contribute to the propagation of the fire across the forest and can be omitted in further simulation steps. Finally, a cell can be labeled as *"Harvested"* or *"Non-Fuel"* and therefore, not available for the fire spread.

In addition, each cell contains a series of secondary parameters that allow Cell2Fire to track the evolution of the fire within the forest and change the state of the current simulation run. The Fire dynamics group includes the fire ignition date and time of each cell, the fire's progress along each axis, and the effective $ROS(t, \theta)$ values (per period and axis angle). The average age of trees inside the cell, approximate volume of wood/products available, ID label, perimeter, area, adjacent cells list, and distance to adjacent cells centers can be found in the *Demographic* category.

Due to this independent structure, cells can be treated as individual units allowing an efficient parallel computing approach for each iteration. We can therefore update their status

and generate the relevant fire messages to model the fire dynamics of each burning cell in parallel using independent threads and update the global status of the forest at the end of each fire time period and thereby obtain significant improvements in execution times (from hours in serial mode to minutes in parallel mode) when dealing with large fires that have many simultaneously active cells.

2.3.4 Fire propagation dynamics

As described above, the fire growth model is simple but powerful: every time a cell is ignited by an adjacent cell it acts as a new source of potential ignition for neighboring cells in the forest, updating the progress of the fire for each available axis (center-to-center directions).

Following an object-oriented programming paradigm in Python and C++, a series of classes for the most relevant components of the problem were developed: Forest, Cells, Weather, Ignitions, FBP System methods and Input/Output formatting. Then, the main program instantiates the different objects and applies the pertinent methods required to simulate fire growth. Once the fire instance data has been read and the forest has been initialized inside the simulator engine, the main simulation steps are as follows:

i) Relevant fire parameters are calculated by performing calls to the FBP System module to determine the Rate of Spread (ROS) for each available fire spread axis of the burning cell: based on fuel characteristics, topography, and weather different $ROS(t, \theta)$ are obtained where t is the current fire time period and θ is the angle with respect to the center of the adjacent cells ($0°$= East, increasing counter-clockwise) based on the procedure described in Section 2.3.7. Following a discrete time simulation approach, the internal simulator clock advances one unit of time — a user-input precision parameter — and the fire's progress is updated for each axis.

ii) Fire spread between cells is modeled by using a sending/receiving message approach (which enables parallelization) based on the computed ROS along each axis. If the fire reaches the center of a cell during the simulated time step, a message is sent. Checking environmental and cells characteristics, a the cell begins burning (or not). This is the core of the simulator and thus, the critical performance bottleneck that comes into play when simulating fire spread across large lanscapes. However, we designed it to maximizing the parallel performance of the code, obtaining a large percentage of naturally parallelizable code, representing around 80% of its structure.

iii) The previous steps are repeated until some specified ending criterion is satisfied: e.g., the maximum number of weather periods, the maximum simulation time, and/or some fire ending event condition. Statistics regarding the status of the forest as well as plots and other outputs of the fire scar evolution are produced.

Figure 2.4 illustrates a forest with 9 cells in which a fire ignites in cell 4, after a harvesting period (e.g., years) from which it can spread to other cells or forest stands for two more fire spread time periods (minutes). If no messages are sent to neighboring cells based on

the current environmental conditions (burned-out conditions) or the maximum simulation time for the current fire has stopped growing, go to the time the next fire ignites (randomly generated or user-provided) or stop the simulation. A general pseudo-code of the simulation steps can be seen in Algorithm 2.

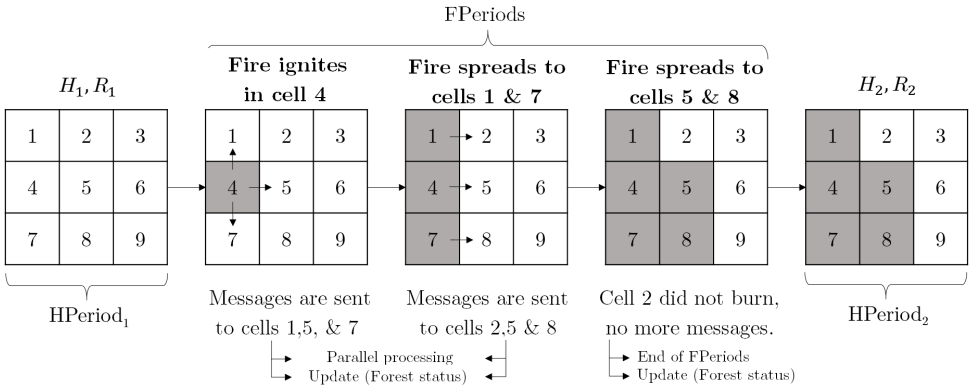


Figure 2.4: Simulation scheme: Send/Receive messages structure allows natural parallelization. Messages are sent when the fire reaches the center of an available adjacent cell.

We use the following notation to describe the main steps of the simulation Algorithm 2:

- $\mathcal{N}$ denotes the set of cells in the forest and $i \in \mathcal{N}$ represents cell $i \in \mathcal{N}$.
- $Adj\,(i)$ denotes the set of available cells that re adjacent to cell i (at most 8).
- $BurningCells$ is the set of actively burning cells in the simulation process.
- $i \rightarrow_t j$ indicates that cell i "sends a message" to cell j at time t

Algorithm 1 Cell2Fire Pseudo-code

procedure SIM($ForestData, FTypes, Ignitions, Weather, TMax, Options$)
    Step 0: Initialize Cell2Fire
        Parse inputs, options, read data, initialize objects
        $FPeriod \leftarrow 0, nsim \leftarrow 0$
    Step 1: Ignition
        $ic \leftarrow$ *Choice(Ignitions)*
        $BurningCells \cup \{ic\}$
        $FPeriod \leftarrow 1$
    Step 2: Fire Dynamics (Send-Receive)
        Let $i \in BurningCells$, ***if*** $i \rightarrow_{FPeriod} j$, *where* $j \in Adj(i)$:
            $BurningCells \cup \{j\}$
        $FPeriod \leftarrow FPeriod + 1$, *Update Forest, Weather*
        Repeat, until $FPeriod < TMax$
    Step 3: Results and Outputs generation
        Generate Grids, Statistics, Output plots
end procedure

2.3.5 Main Inputs

As is the case with other state-of-the-art fire growth simulators, the Cell2Fire model requires a number of inputs including a minimum set of data layers that define an instance/forest to start the simulations. The relevant inputs needed to simulate the growth of a fire are as follows:

i) Forest raster data: ASCII grid forest files that specify the number of cells in the forest, their geographical coordinates and information concerning each cell including as its fuel type, elevation, slope (% and azimuth), degree of curing. Files can be in either .csv or .asc format. If topographic data is not fully available, default input dummy values for parameters such as elevation or slope can be provided by the user (null by default).

ii) Fuel type dictionary: Fuel type codes and descriptions that match the Canadian FBP System fuel types (Table 2.1) are included inside a .csv file. Future implementations will allow the use of a custom dictionary file that includes fuel types not currently included in the set of Canadian FBP fuel types.

grid_value	export_value	descriptive_name	fuel_type
1	1	Spruce-Lichen Woodland	C-1
2	2	Boreal Spruce	C-2
3	3	Mature Jack or Lodgepole Pine	C-3

Table 2.1: FBP fuel type dictionary sample. The grid_value field refers to the encoding of the forest inside the ASCII files and the fuel_type column contains its translation into the FBP code.

iii) Ignition points: An optional file that that specifies the cell(s) in which fires are to be ignited during the simulation, paired with their corresponding ignition time periods.

iv) Weather stream: Hourly weather records for one or more fire weather stations located near the area of interest that include the date-time, precipitation, temperature, wind speed/direction, relative humidity, scenario ID, as well as the daily fire danger rating codes and indices (FFMC, DMC, DC, ISI, BUI, and FWI) of the Canadian Forest Fire Danger Rating System (Alexander et al., 1996), registered by the stations and equations inside the FBP System module (Table 2.2). Data from the nearest weather station is used for each cell.

Sample files are included with the distribution of Cell2Fire for the publicly available Dogrib instance (`http://www.firegrowthmodel.ca/prometheus/software_e.php`) as well as simple generated testing instances.

Besides the main input files, a set of options and user-provided parameters for exploiting all the flexibility of the simulation engine including tuning options (see Section 2.3.7) are available as secondary inputs when running Cell2Fire.

Scenario	datetime	APCP [mm]	TMP [C°]	RH [%]	WS [m/s]	WD[°]
JCB	2001-10-16 13:00	0.0	17.7	20	21	225
JCB	2001-10-16 14:00	0.6	16.9	18	25	205
JCB	2001-10-16 15:00	1.2	16.1	20	27	190
JCB	2001-10-16 16:00	10.0	15.8	20	37	232
JCB	2001-10-16 17:00	5.3	13.9	25	43	225
JCB	2001-10-16 18:00	2.1	12.1	35	45	222
JCB	2001-10-16 19:00	0.9	10.6	41	46	241
JCB	2001-10-16 20:00	0.0	11.3	39	18	248

Table 2.2: Extract of an hourly weather stream file. Average precipitation (APCP), temperature (TMP), relative humidity (RH), wind speed (WS) and wind direction (WD).

2.3.6 Main Outputs

Once a simulation run has been completed, the following outputs are available:

i) Burn-Grids: Files in which 1s indicate burned cells and 0s indicate those that are available to burn. Useful for statistical comparisons with other simulators as well as to generate burn probability maps.

ii) Plots: Initial forest state, fire scar evolution, and message sending/receiving can be visualized by a series of plots (Figure 2.5) generated after the simulation run has been completed.

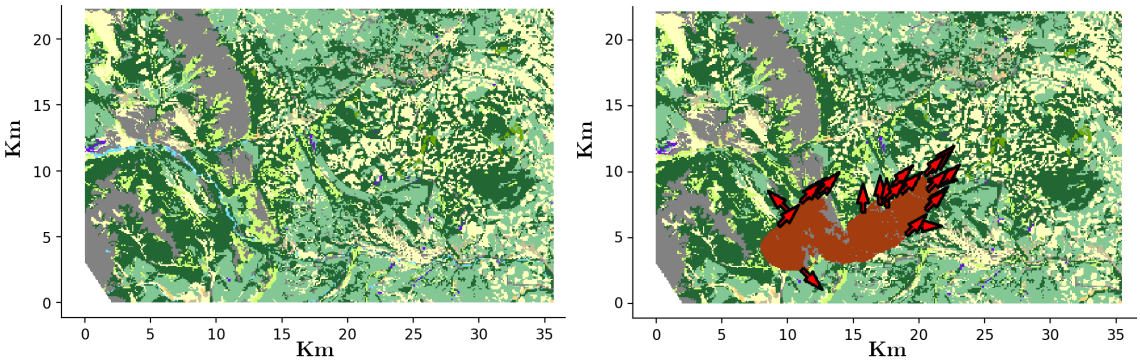


Figure 2.5: Initial forest state (left side) plot example for Dogrib instance. Each color represents a cell with a specified fuel type, encoded according to the FBP system. Fire scar evolution example (right side) including burning cells (orange) and new fire messages directions.

iii) Statistics: Final status of the forest with relevant information regarding the messages, fire ignition dates, fire evolution (distance per axis and ROS values), average number of burned/available cells, the average percentage of the forest available for all the tested scenarios (including multiple ignition points, weathers).

2.3.7 Rate of Spread from cell to cell

We use the approach proposed by the Canadian FBP System Tymstra et al. (2010b) where an elliptical distribution scheme is applied. This method does not require the ignition point or the point of propagation to coincide with either of the two foci of the ellipse, although the authors indicate that "*small differences between the point of ignition and the focus of the ellipse do not change the results*", i.e., the elliptical propagation shape/size.

- An elliptical fire has a HROS (head), a BROS (backing) and a FROS (flank). Let a, b, and e be the semi-major, semi-minor, and eccentricity of the ellipse, respectively.
- The FBP system predicts the HROS, the BROS and the length to breadth ratio (LB) which is 2a/2b or a/b.
- During the first time interval t, the fire will spread from its ignition point to the center of the ellipse and then from the center of the ellipse to the farthest edge of it.
- At time t, we have:

$$a = \frac{HROS + BROS}{2} \times t, \tag{2.2}$$

$$b = \frac{2 \times FROS}{2} \times t, \tag{2.3}$$

  to expand the ellipse generated by the propagation of fire at time t on the two main axes.
- Noting that $LB = \frac{a}{b}$ we have $FROS = \frac{HROS + BROS}{2LB}$. Therefore, the eccentricity is:

$$\mathrm{e} = \sqrt{1 - \left(\frac{(FROS \times t)^2}{\frac{(HROS + BROS) \times t^2}{2}} \right)} \tag{2.4}$$

Using these equations and the procedure described in Tymstra et al. (2010b), we can estimate the Rate of Spread from the center of a cell to the center of any adjacent cell as in equation (1).

2.3.8 Computational details

We chose C++ for the parallel implementation because the objects we created in Python are naturally mapped to C++ objects, obtaining at least 15-20x average speedups depending on the forest characteristics. As previously discussed, our algorithm contains 3 sections during each time step: (1) checking for new lightning fire ignitions (igniting), (2) updating the intensity of already-burned cells and analyzing newly burned cells (sending messages), and (3) marking newly burned cells as burning (receiving messages).

Analyzing Cell2Fire running times, the ignition stage is very quick, with most simulations only igniting a single time at the first time step of the simulation. The sending-messages

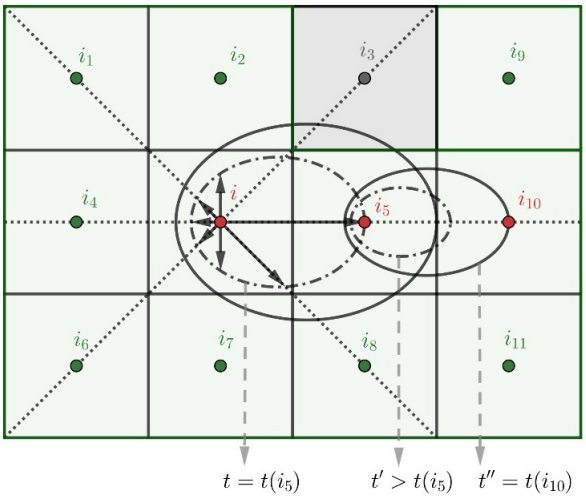


Figure 2.6: Elliptical Rate of Spread distribution scheme using the ellipses defined by the Canadian FBP System. At any time t, the backfire will be $BROS \times t$ behind the point of ignition and the head fire will be $HROS \times t$ ahead of the point of ignition of the fire, expanding the ellipse. Then, if the fire emitted by cell i reaches the center of an adjacent cell j at time $t(j)$, a new ellipse is generated at time $t' > t(j)$, calculating the corresponding rate of spreads, and so on.

stage updates the fire progress in every burning cell. Because we can have a large number of cells burning at once and there are no direct dependencies on neighboring cells, this part is easily parallelizable. Each cell, in addition to updating its current status, can also "send a burning message" to an adjacent cell. In the receiving messages stage, we analyze the "burn messages" sent to non-burning cells and mark them as burned if the conditions are met. This part is also potentially parallelizable, but because the number of newly burned cells at a single time-step is dwarfed by the number of currently burning cells, we found that a speedup here is of lower priority.

Due to the easily parallelizable structure of our code, the most suitable approach for parallelizing its execution consists of a shared-memory approach using the well-known OpenMP API (Dagum and Menon, 1998). This is an advantage of Cell2Fire since the code will be also optimized for its execution in daily-use computers, without needing a multi-node architecture for exploiting its parallelism. Using it, we were able to obtain an average of 15%-20% extra performance for the parallel region.

2.4 Results and Discussion

In this section, we compare the predicted burn scars produced Cell2Fire and Prometheus for several instances created for this purpose as described in Section 2.4.2. We did not compare either simulator with the realized fire scars in our study because it is difficult to determine the extent to which the final shape of real fires that were fought was or was not influenced by suppression action.

2.4.1 Comparison Methodology

The comparison methodology we used consists of measuring the difference between fire scars generated by Prometheus and Cell2Fire simulations using three metrics. The simplest and most widely used full-reference quality metric is the mean squared error (MSE), which objectively quantifies the strength of the error signal. However, two distorted images with the same MSE may have very different types of errors, some of which are much more visible than others. Thus, as we are interested not only in the difference of the marks but also in structural information, we use a measure of similarity suggested in Zhou et al. (2004) and denoted by SSIM (structural similarity index). Finally, we also include the Frobenius norm of the difference between two scars X, Y, $\delta_{norm} = ||X - Y||$.

First, we compare the evolution of Cell2Fire and Prometheus fire scars on a period-to-period basis (where a period represents one hour) in order to measure the differences in the fire propagation. We denote by $PromGrid^t$ a 0-1 Matrix at time t, which represents the fire scar obtained with Prometheus, where $PromGrid^t_{ij}$ is equal to 1 if the cell (i, j) was burned and 0 otherwise. Analogously, we define the fire scar obtained by Cell2Fire as $Cell2Grid^t$.

Below, μ_{X^t}, μ_{Y^t}, σ_{X^t}, σ_{Y^t} and $\sigma_{X^tY^t}$ represents the means, standard deviations, and cross-covariance for scars X^t and Y^t respectively, and C_1, C_2, and C_3, are internal parameters of the metric Zhou et al. (2004). The methodology is as follows:

1. Choose a ignition point for each instance and run Prometheus for T time periods. Thus, we obtain $PromGrid^t$, $t = 1, ..., T$ (0-1 Cell Matrices).
2. Choose the same ignition point as above and then run Cell2Fire for T time periods from this point. Outputs: $Cell2Grid^t$, $t = 1, ..., T$.
3. Set $X^t = PromGrid^t$ and $Y^t = Cell2Grid^t$ and calculate for all t:
   (a) Mean Squared Error:
   $$MSE\left(X^t, Y^t\right) = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left|X^t_{ij} - Y^t_{ij}\right|,$$
   to measure average of the squares of the pixel differences of the fire scars, and
   (b) Structural similarity Measure:
   $$SSIM\left(X^t, Y^t\right) = \frac{\left(2\mu_{X^t}\mu_{Y^t} + C_1\right)\left(2\sigma_{X^tY^t} + C_2\right)}{\left(\mu^2_{X^t} + \mu^2_{Y^t} + C_1\right)\left(\mu^2_{X^t} + \mu^2_{Y^t} + C_2\right)},$$
   to measure the change in structural information between the fire scars obtained from the two simulators: Cell2Fire and Prometheus.
4. Measures analysis: MSE and SSIM evolution for all time t, δ_{norm} for the final scar.

2.4.2 Instances

We used three sets of fire instances to compare the performance of Cell2Fire with Prometheus: (1) Sub-instances set, (2) British Columbia province real landscapes with simulated wildfires and (3) Case study: Dogrib fire landscape.

Sub-instances

We used portions of the Dogrib landscape data (see Section 2.4.2), we generate two sub-instances that we label Sub-1 and Sub-2 with a cell resolution of 100 × 100 meters. The first one represents a sub-forest from the Dogrib landscape that is 20 × 20 cells (400 ha.) and the second a 40 × 40 cell (1600 ha.) instance. Both consist of heterogeneous landscapes that include different fuel types as well as non-flammable cells (such as mountains or rivers). An ignition point was selected for each instance as a starting point for the fire simulation. Three weather stream files: Weather-1, Weather-2, and Weather-3 of 6, 14 and 22 hours respectively, were used as main inputs: the first one contains data for the 6 hours during which the real Dogrib fire made a run (extreme weather conditions) while the second and third contain additional meteorological measurements from the same day of the fire, before and after that spread event, thus, being an extension of the Dogrib's weather scenario.

After the ignition point was fixed for both instances, we proceeded to run the simulation in Prometheus and the deterministic version of Cell2Fire, generated the hourly fire scars and calculated the performance indicators (1-MSE and SSIM).

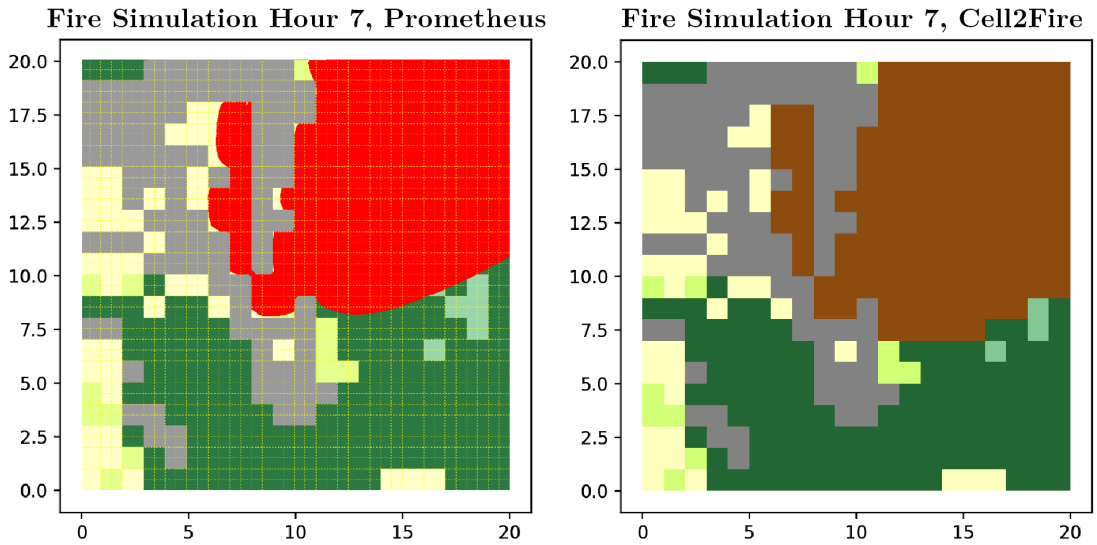


Figure 2.7: Sub-Instance 2 Fire evolution visualization in Prometheus and Cell2Fire including a heterogeneous landscape with non-flammable cells (mountains, gray cells) and different fuel-types (green and yellow cells).

Based on the results shown in Table 2.3 and Figure 2.7 we can see that Cell2Fire produced results that are similar with respect to the hourly fire growth and final fire scar to those produced by Prometheus. When testing the 6 critical hours of the Dogrib fire, the level of difference (MSE) is less than 5% for the first sub-instance and less than 6% for the sub-instance 2. For the full day simulation, a slight increment of the differences is registered, reaching average MSE levels values close to 6% and 7 % for Sub-2 and Dogrib instance respectively.

On the other hand, it is possible to observe how the structural similarity measurement has a decreasing pattern (see Table 2.3), a situation that is even more explicit when comparing

the 22 hour fire weather stream. This indicates that the fire growth predicted by Cell2Fire differs from the one predicted by Prometheus due to: (1) as expected, the approximation of the elliptical model into an adapted cellular-automata approach implies a different fire dynamic, increasing the differences on every time step when comparing with the wavefront model and (2) differences in the calculations/approximations of the effective $ROS(t, \theta)$ values.

The wave-propagation model based on the Huygens' principle implemented in Prometheus performs a series of approximations with respect to the burning area (ellipse) such that a cell is classified as a *Burned* cell in the Burn Grid output (0-1 Matrix) only if more than 50% of its area has been covered by fire (belongs to the interior of the approximated ellipse) while in Cell2Fire a cell is always completely available or burned. In addition, an ignition point represents a complete cell in Cell2Fire (an area) while it is just a vertex/point for Prometheus, defining two different (but consistent) starting points for the fire spread evolution. This approximation based on discrete cells improves as the cell size decreases.

	Sub-1		**Sub-2**	
Hour	**1 - MSE [%]**	**SSIM [%]**	**1 - MSE [%]**	**SSIM [%]**
1	99.75	93.59	99.98	95.44
2	96.75	77.40	99.94	95.01
3	97.75	74.03	95.56	85.53
4	97.50	78.70	95.38	84.41
5	96.75	79.61	96.50	82.25
6	96.00	73.96	94.03	75.01
AVG [%]	97.42	79.55	96.90	86.25

Table 2.3: Sub-Instances accuracy and structural similarity index measure values per hour (6 hours evolution)

The two fire growth models are consistent and correlated in that they produce very similar fire scars for the test instances as seen in Figure 2.7.

British Columbia wildfire set

The British Columbia instances set contains five different regions — ArrowHead (265,536 ha.), Revelstoke (391,314 ha.), Mica Creek (348,404 ha.), Glacier Natural Park (559,746 ha.), and Central Kootenay (494,665 ha.) — of the province. For each area, two fires with random ignition points and 24-hours stream weather conditions based on the historical weather dataset from the Climate Information Section of the Agriculture and Forestry site of Alberta, Canada, and data from the Yaha Tinda Auto station (coordinates: 51.6547 , -115.3617) are compared. These instances are provided with BurnP3 http://www.firegrowthmodel.ca/burnp3/software_e.php. We generated subsets of the large forests, simulated fires in Prometheus using 24 hours weather scenarios — using historical data from the zone — and then we compare the fire scars from Prometheus with the ones obtained with Cell2Fire.

The final fire scars and performance metrics — focusing in the affected area of the instance for easier visualization — obtained for the 10 simulated wildfires in both Prometheus (columns 1 and 3) and Cell2Fire (columns 2 and 4) can be seen in Figure 2.8. Results indi-

cate the high similarity between the scars, obtaining good performance across the main three metrics (Table 2.4) for all forests. This validates Cell2Fire for different fuel types, landscapes, and weather scenarios, being able to approximate the results of a state-of-the-art simulator like Prometheus.

Different ignition points and weather scenarios were tested on these landscapes, obtaining similar results in terms of the main performance metrics. A similar pattern was observed with respect to the hourly evolution of the scars.

	MSE	$SSIM$	δ_{norm}
Mean	0.09	0.68	27.36
Std	0.04	0.09	8.88
Max	0.18	0.85	42.64
Min	0.03	0.46	10.10

Table 2.4: British Columbia simulations summary statistics when comparing the simulated final fire scars from Prometheus and Cell2Fire. $\delta_{norm} = ||X - Y||$ where X and Y are the binary BurnGrids matrices obtained from both simulators.

Dogrib fire instance

The Dogrib fire (Tymstra et al., 2010b) started on September 25, 2001 the province of Alberta. The fire was detected late in the afternoon on September 29 and assessed early the next day at 70 ha. in size. Suppression action began early on October 1. The fire was 828 ha. and out of control on October 15. A wind event resulted in a major fire run on October 16. Local terrain funneled wind flow along the Red Deer River and through a gap in the surrounding mountains. This pushed the fire east along the river valley. The fire jumped the Red Deer River and a road and then resumed spreading in a northeast direction. The final fire size was 10,216 ha. The October 16 fire run accounted for ninety percent of the total area burned and resulted in high to very high burn severities. In Figure 2.9 we can see the Dogrib fire perimeter and burn severity as detected by Landsat.

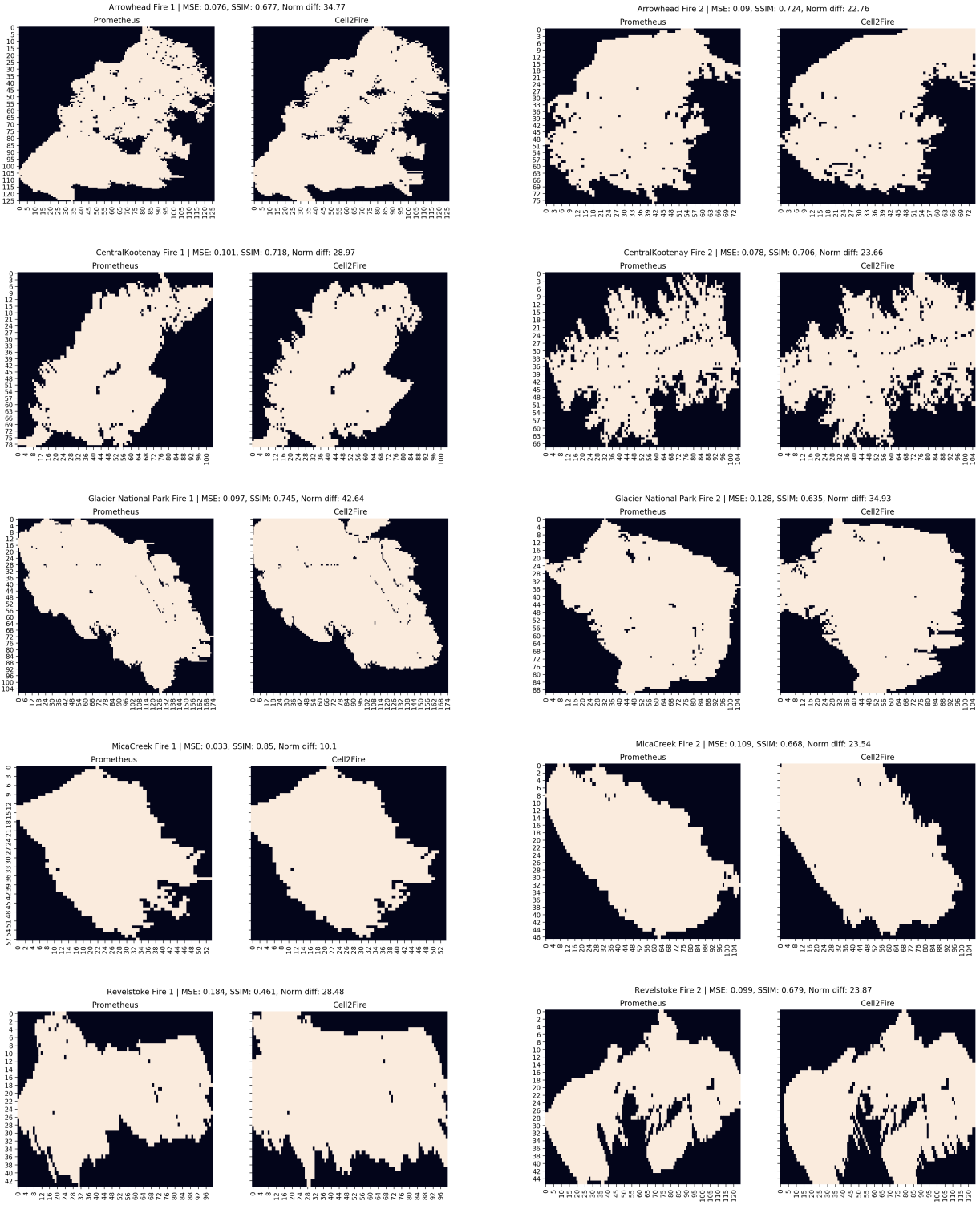


Figure 2.8: British Columbia wildfire instances. From the final scars and statistics (left side Prometheus, right side Cell2Fire), it can be seen how accurate is Cell2Fire w.r.t. simulated scars from Prometheus, reaching $\overline{MSE} = 0.0995$, $\overline{SSIM} = 0.6863$, $\overline{\delta}_{norm} = 16.347$.

We chose to model this particular fire due to the large amount of documentation and real data available — weather conditions recorded from The Yaha Tinda Automatic station, and demographic/topographic data collected from the area — as well as the fact that it contained a representative set of different fuel-types documented in the Canadian FBP system. We divided the landscape into 79,611 100×100; $[m^2]$ cells, used the Dogrib fire's ignition point located at (51.652876°, -115.477908°) and started spreading the fire on October 16, 2001, 13:00 hrs. As discussed in the previous section, the ignition point is translated into an ignition area (cell) in Cell2Fire containing its coordinates.

The comparison of the performance metrics after 22 hours of fire growth is shown in Figure 2.10 . Here, it can be seen that Cell2Fire has a very similar evolution with respect to the wave-front approach, obtaining good performance when compared with Prometheus, not exceeding a 20% difference in both measurements. An average of 87.91% of structural similarity and a global average of 91.82% of accuracy (1-MSE) are obtained during the 22 hours of active fire growth. A clear pattern can be seen in the graph where both performance metrics start very high and remain stable during the first 4 hours, then they show a significant negative slope between hours 4 to 11, and finally reach a steady state for the rest of the simulation. The explanation behind these results is clear: during the initial 4 hours of the fire, similar fire growth occurs due to weather conditions that are not extreme, resulting in diminished model differences (wave-front and cellular-automata), however, weather factors between hours 4 to 11 contain the most extreme conditions (strong wind speed, high temperature, low precipitation levels, etc.) magnifying the fire growth differences/approximations between both approaches in terms of the number of burning cells per hour (fire scar). After hour 11, differences between fire scars tend to be very stable due to the lack of new extreme weather episodes, but keeping the structural differences — mainly in terms of the current perimeter of the fire — in the fire scars obtained in the previous hours.

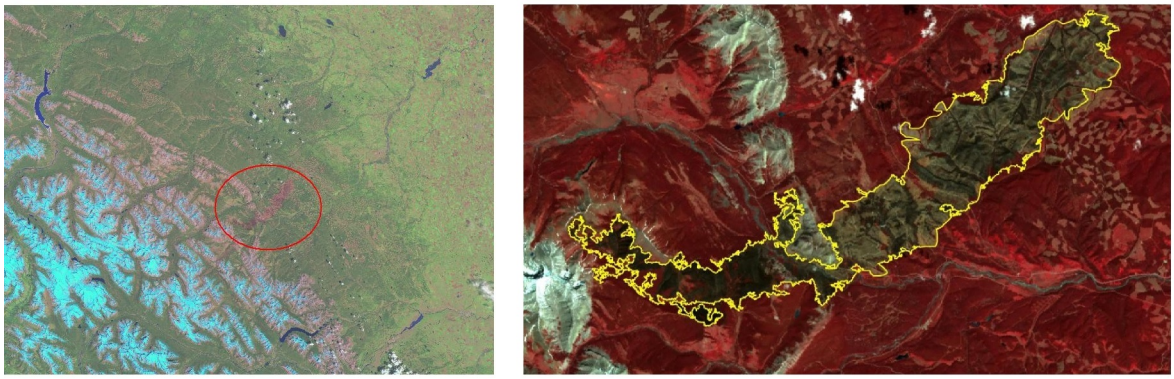


Figure 2.9: Dogrib fire perimeter registered on June 22, 2002 from Landsat.

In addition, some of the structural differences between the generated fire scars can be explained in part by the extra modeling features included in Prometheus but not in Cell2Fire such as *Breaching* where non-fuel grid cells or linear fuel breaks fail to stop an advancing fire front, a feature that is not currently included in Cell2Fire.

Detailed results are presented in Table 2.5 where both $(1 - MSE)$ and $SSIM$ values can be seen for the full 22 simulated hours. We conclude that Cell2Fire produces results that are similar to those produced by Prometheus, obtaining similar final fire scars as seen in Figure

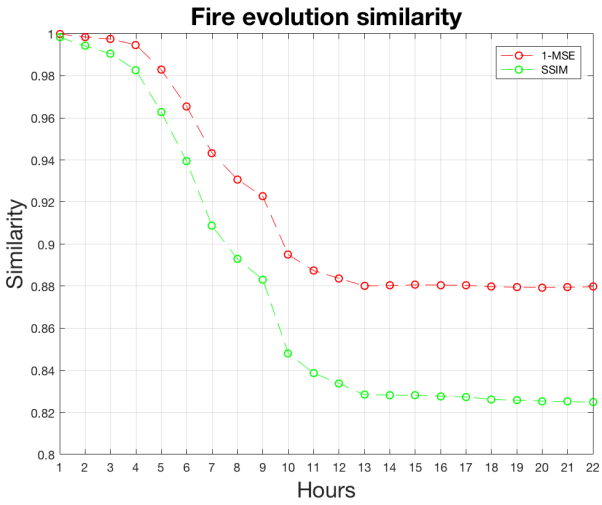
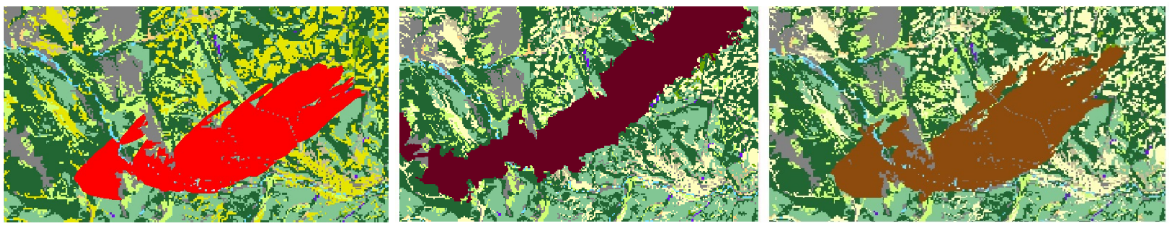


Figure 2.10: Dogrib MSE and SSIM hourly evolution (22 hours).

2.11 where the simulated fires and the real satellite images are shown.



(a) (b) (c)

Figure 2.11: a) Prometheus fire scar obtained for the region of Dogrib, Canada, comparison with the b) real fire projected into grid format in 2002 and c) Cell2Fire final output.

Hour	1-MSE [%]	SSIM [%]	Hour	1-MSE [%]	SSIM [%]
1	99.98	99.83	**12**	88.37	83.38
2	99.85	99.42	**13**	88.01	82.84
3	99.74	99.05	**14**	88.03	82.83
4	99.47	98.28	**15**	88.06	82.81
5	98.29	96.27	**16**	88.06	82.77
6	96.54	93.94	**17**	88.04	82.74
7	94.33	90.86	**18**	87.98	82.62
8	93.08	89.30	**19**	87.96	82.58
9	92.26	88.32	**20**	87.94	82.54
10	89.50	84.81	**21**	87.95	82.51
11	88.74	83.87	**22**	87.97	82.49
AVG [%]	95.62	93.08	**AVG [%]**	88.03	82.74

Table 2.5: Dogrib accuracy and structural similarity index measure values per hour (22 hours evolution).

2.5 Conclusions

Cell2Fire provides numerous opportunities for researchers interested in incorporating fire growth in their models for strategic harvest planning and fuel management. We are currently using it in ongoing projects. Because it is open-source and modular it lends itself to customization as needed. The simulator is fast and scales well in parallel computing environments so it is well-suited for use with large forests and large studies that may require many simulation runs.

Using the FBP fire spread model, we have compared the simulated fire scars with those produced by the state-of-the-art simulator, Prometheus. Other fire spread models can be employed instead, which extends the range of environments where Cell2Fire can be used and also allows for comparison of fire spread models when used in a growth simulator. In addition to supporting stochastic ignition and weather, the simulator also supports modeling of the uncertainty in the rate of spread.

By adding a highly parallelizable, open-source fire growth simulator to the tool set available, we hope to provide transparent support for ongoing research.

Acknowledgements

We thank B.M. Wotton and D. Boychuck for their helpful comments and suggestions. This research has been supported by Complex Engineering Systems Institute (CONICYT-PIA-FB0816) and FONDECYT 1191531.

Chapter 3

On the possible practical applications of the inclusion of different sources of uncertainty in Cell2Fire simulator

Authors

Jaime Carrasco and Andrés Weintraub
University of Chile, Industrial Engineering Department.

Cristobal Pais
University of California Berkeley, IEOR Department.

David L. Martell
University of Toronto, Faculty of Forestry.

David L. Woodruff
University of California Davis, Graduate School of Management.

Abstract

Several authors have mentioned the need to include uncertainty in the modeling of the spread of fire. The reasons are many: to capture the intrinsic nature of fire; there is limited accuracy of the input parameters, such as the speed of propagation and climatic conditions that can affect the predictions; decision makers need to consider the uncertainty of fire, for evaluation of harvest management plans at operational, tactical and/or strategic level, among many others. In this study, we incorporate three random components to a new cell-based forest fire spread simulator recently developed, called Cell2Fire. This simulator incorporates fuel and fire models from the Canadian Forest Fire Behavior Prediction (FBP) System such as fuel moisture, fuel type, surface fire, and crown fire, and has been designed to naturally assess the impact and expected losses due to fire propagation on heterogeneous landscapes. An analysis of the main sources of uncertainty and their impact in the simulations is shown.

3.1 Introduction

Large and destructive megafires are becoming more frequent worldwide (Maditinos and Vassiliadis, 2011; Tedim et al., 2013; Attiwill and Adams, 2013; Ferreira-Leite et al., 2015; de la Barrera et al., 2018a), experts have attributed this to the climate change, which is causing hotter and drier conditions as well as to a century of fire suppression policies that have left forests with more vegetation to fuel the flames than in the past (Moritz et al., 2012; Dillon et al., 2011). Several efforts have been made to understand this new phenomenon, but since scientists cannot experiment with large and destructive wildfires, so they have fallen back on examining statistical correlations and/or simulate the phenomenon to try to tease out the key factors associated with megafires.

Traditionally fire regimes, fire danger, and fire behavior have been predicted using a series of rules and empirical models, being the most used the fire weather index to estimate daily fire danger (Van Wagner et al., 1987; Stocks and Martell, 2016a) and fire behavior models to predict how fire will spread once occurs (Van Wagner et al., 1987; Rothermel, 1972b), which have led to more sophisticated computing tools of simulation such as FARSITE (Finney, 1998), Prometheus (Tymstra et al., 2010a) and others (Papadopoulos and Pavlidou, 2011b; Sullivan, 2009). These models and subsequent modifications are the basis of most fire research studies dealing with fire occurrence and fire behavior. They have shown their usefulness over the years, but still, they present certain limitations that new approaches can overcome. (1) Either are highly demanding in data in the case of fire behavior models, being difficult to apply over large areas due to inventory costs, (2) not accounting for uncharacteristic variations on interacting conditions or able to identify shifting points in climatic/environmental variables that may significantly perturb the accuracy of the predictions, or (3) being deterministic on the way fire interacts (occurs, spreads) with different ecosystems, adapting well to those conditions to whom the models were developed for, but not accounting for uncharacteristic variations on unseen environments. Even more, fire simulators and fire danger indexes highly rely on rules applied to fine fuels and how they dry according to weather conditions, sensible to changes on climate, but in the case of fire danger, these systems do not account for many

other factors influencing the occurrence and their interactions.

Several authors have pointed the need to include uncertainty in the modeling of the propagation of fire for different reasons; the intrinsic nature of fire behavior is better modeled (Kourtz, 1972; Boychuk et al., 2009b; Cruz and Alexander, 2013) and/or to obtain a fire risk measure (Parisien et al., 2005a; Finney et al., 2011b). However, despite the existence of several deterministic and stochastic fire simulators, fire-smart forest and fuel management requires that both the decision support system and simulation models have a well-structured interface,facilitating the exchange of data between them (Pais et al., 2019). The above mentioned simulators were not designed for purposes other than fire spread simulation and were not incorporated in a land management framework. For this reason that Cell2Fire was developed, specifically for used in a fuel-management context with the intent of mitigating the impact of large detrimental fires efficiently, including uncertainty. Thus, Cell2Fire could work as a pure simulation tool and/or as a tool that can be used to evaluate fuel-management strategies or linked with optimization software to develop "optimal" fuel management strategies. A review of recent Forest and Wildland Fire Management Decision Support Systems can be found in Martell (2015).

The main objective of our research was to incorporate three sources of uncertainty to develop an efficient and more realistic fire spread simulator, that will allow one to simulate the dynamics of wildfire across a grid representation of a real or hypothetical forest, and in this way, obtain different fire risk measures that could support the decision of fuel management or suppression. In particular, we include stochasticity by: (i) ignition point(s) selected via a user-defined probability distribution or a simple uniform approach for each period of the planning horizon, (ii) a perturbation level representing the stochastic aspects of the predicted Rate of Spread (ROS) of fire via a coefficient of variation, accounting its inherent approximation error as well as allowing the user to obtain different (and average) fire scars including uncertainty in the fire dynamic, and (iii) a set of user-generated weather stream files (scenarios) with specific probabilities can be provided to Cell2Fire, performing a series of simulations with different weather conditions such as wind direction/speed, precipitation, etc. allowing the user to analyze the fire propagation under different scenarios for a certain forest/instance. We emphasize that these uncertainty components can work separately or in combination, thus providing greater flexibility for the fire manager.

3.2 Incorporating uncertainty into fire spread model

Cell2Fire (Pais et al., 2019) is a cell-based fire growth model on a raster-grid containing a series of square cells (same area). Fuel and terrain conditions are assumed to be homogeneous within each cell in order to simplify basic fire spread rate calculations. The fire propagation is modeled from a cell's center to the center of an adjacent cell. Each ignited cell behaves as an ignition source and is independent of any adjacent burning cell. We refer to this approach as: "Cellular Huygen's Principle". To spread the fire from one cell to another, a pre-defined spread template is required such that a raster grid of cells and others data layers.

However, as in any modeling of some phenomenon of nature, the scientist faces an inher-

ent problem such as the calibration of the input parameters and/or the intrinsic randomness of some of them that can lead to erroneous and unrealistic results. Zhang et al. (2016) pointed out that: "*an inherent error in the input data to a prediction model negatively influences overall performances*". Cruz (2010) points out that incorporating stochastic analysis in fire spread models would be more valuable in obtaining error bounds and probability-based outcomes in the prediction models. Hajian et al. (2016) agree that stochastic analysis has advantages over the deterministic approach, while Han and John Braun (2014) go as far as stating that a realistic fire spread model, based on the rate of spread, have an intrinsic noise component which could be used to model the uncertainty in ROS.

In order to address these aspects, several efforts have been proposed to include uncertainty in fire growth models: (i) some focused on the intrinsic randomness in input parameters, for example considering random location ignition point where an empirical distribution is based on historical data and/or considering random weather scenarios; others (ii) modifying the transition/spread probabilities, e.g. using a continuous-time Markov chain on a lattice or a cellular approach; and (iii) through directly including a random noise in the Rate of Spread (ROS).

With respect to the first approach i), an excellent tool was developed in Canada based on Prometheus simulator and historical data: Burn-P3 (Probability, Prediction, and Planning) (Parisien et al., 2005a). It is a simulation model that evaluates the fire likelihood or burn probability (BP) of a large fire-prone landscape. The model combines deterministic fire growth based on the FBP System and spatial data for forest fuels and topography with probabilistic fire ignitions and spread events derived from historical fire and weather data. Model components include the location and frequency of ignitions, the rate at which fires escape initial attack and become large, the number of days on which each fire achieves significant spread, the fire weather conditions associated with these spread event days, and the deterministic fire spread. Other tools with similar characteristic but not identical are FlamMap developed by Finney (2006) and Fire Spread Probability (FSPro), implemented in the Wildland Fire Decision Support System (WFDSS, `http://wfdss.usgs.gov`) and used as a strategic decision aid tool — looking at fire risk as it is determined by uncertainty in the weather; see Finney et al. (2011a).

For examples of ii), we refer to Boychuk et al. (2009b), which develops a stochastic fire spread model using a lattice Markov chain approach, in which they associated transition functions with each cell. Each of these cells interacts with its four nearest neighbors and a cell burns depending on the state of the neighboring cells. Other simplified cellular automata models describing fire spread on a heterogeneous landscape including weather factors (wind speed and direction) as well as the type and density of vegetation has been modeled by Alexandridis et al. (2008b) where the Spetses Island fire, Greece, was successfully modeled using this approach after tuning the main parameters of the simulator. Almeida and Macau (2011) models the dynamics of fire spread as a stochastic event with an effective fire spread probability which is a function of three probabilities that characterize: the proportion of vegetation cells across the lattice, the probability of a burning cell becomes burnt, and the probability of the fire spread from a burning cell to a neighboring vegetation cell.

Finally, in case iii), we mentioned to Dionysus, a simulator recently developed by Han

and John Braun (2014). This is a project that tries to randomize Prometheus using the observations of Cruz and Alexander (2013), who claim that the empirical relation between ROS and ISI (the initial spread index, see Hirsch et al. (1996)) has a noise component which could be used to model the uncertainty in ROS. Then, Dionysus can be consider as a statistical method for modeling fire growth prediction uncertainty based on residual variation in the ROS model whose results are to obtain a distribution of fire front predictions and/or accurate probability contours.

In this study, we include aspects of the i) and iii). From i) we use the ignition locations as a user-defined parameter, giving control over the pattern of ignitions on the landscape. These points can be selected randomly or deterministically before the simulation begins. The latter models the fall of lightning over the forest. When a lightning falls on a cell i at time t, there is a "ignition probability" that the fire will start or not. On the other hand, Cell2Fire contain a *stochastic weather module* that draws a weather scenario from a set provided by the user. In this point, the user must previously construct the scenarios using, for example, historical data to produce representative-different variations of the weather, where the latter has the same format as those of Prometheus simulator and Burn-P3. In addition, users can assign a probability to each scenario in order to generate relevant statistical outputs while providing more flexibility when analyzing the impact of potentially extreme but rare (with low probability) events.

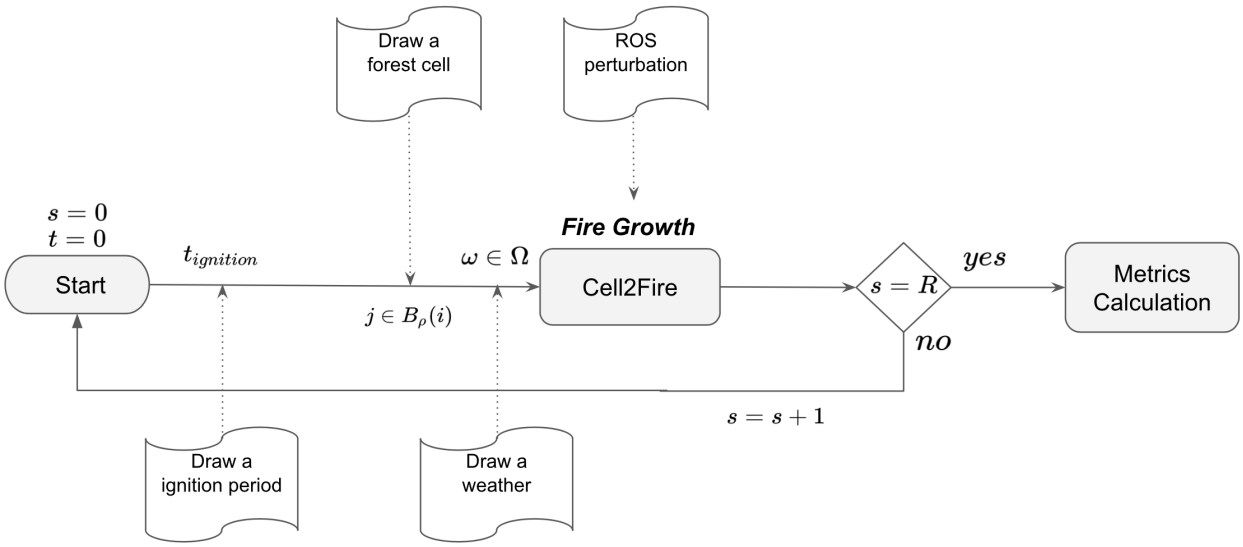


Figure 3.1: Framework of the inclusion of the different sources of uncertainty in the new system.

In this version of our system, fires begin with randomly occurring lightning strikes, igniting a cell. Also, an ignition point can be chosen deterministically as discussed in Pais et al. (2019). Cell2Fire simulates the spread of fire by tracking the state of all cells as the model progresses through discrete time steps on a raster grid. A cell can be in one of five states: *"Available"*, *"Burning"* , *"Burned"*, *"Harvested"*, or *"Non-Fuel"* where the term *"Available"* indicates the cell in question contains a flammable fuel type. Once the cell where the fire starts is selected, a weather scenario is drawn. At each time step, fire may spread along specified axes emanating from the center of the cell to adjacent cells based on the predicted Head Rate of Spread

(HROS), Flank Rate of Spread (FROS), and Back Rate of Spread (BROS) from the fire spread model and eventually perturbed by a random factor. A signal/message dynamic is then initiated between adjacent cells (see Figure 3.1).

In the present implementation, it is assumed that each cell has at most 8 adjacent cells. In this point, is important to notice that the concept of adjacent cells can be naturally extended in Cell2Fire in order to capture potentially more complex fire spread dynamics since the ROS calculation and distribution along the relevant axes are performed in a continuous space (see next section), allowing to easily include as many layers of adjacent cells (e.g. from 8 to 24 by adding the second layer of 16 cells surrounding a cell) as needed. Continuing, a cell starts to burn when the fire reaches its center, and then, each cell that burns serves as a new source of fire ("Cellular Huygens principle"); noting that the effect of fire suppression action is beyond the scope of this study and will be addressed in a future research project. The total duration of the simulated wildfire event is determined by (1) maximum number of burning hours per day — constant, season-dependent, or drawn from a probability distribution — as described in Parisien et al. (2005a) and (2) the total fuel remaining inside the forest (available cells).

3.3 Uncertainty sources and implementation

In this section, we define and provide a detailed description of the main sources of uncertainty included in our system (see Figure 3.1) and the importance of their inclusion. Then, we discuss their impact on the simulation dynamic when applying them as isolated or combined effects by showing examples of their usage.

Three different sources of uncertainty are included in our model to account for the most important perturbations that can significantly affect the fire growth dynamics: (1) ignition point(s) selected via a user-defined probability distribution or a simple uniform approach for each period of the planning horizon, (2) a coefficient of variation (cv_{ROS}) capturing the stochastic aspects of ROS predicted by the fire spread model, accounting its inherent approximation error as well as allowing the user to obtain different (and average) fire scars based on potential wildfires including uncertainty in the fire dynamic, and (3) a set of user-generated weather stream files (scenarios) with specific probabilities that can be provided to Cell2Fire, running simulations with different weather conditions such as wind direction/speed, precipitation, etc.

In most cases, we show the effect of the disturbances through R replications of the simulator, following the scheme in Figure 3.1.

3.3.1 Ignition points/area

In order to incorporate uncertainty when selecting the starting ignition point(s) of a fire, we include three main approaches:

I) If no ignition points (or probability distribution) are provided by the user, a simple

random selection can be performed by using a random variable $X \sim U(0, N \times M)$ where N is equal to the number of rows and M the number of columns representing the forest grid. With this scheme, a fire simulation will start at the selected cell i at time $t = 0$. Same random sampling will be performed at every time step t in case of simulating multiple fire periods for a planning horizon T, sampling independent random numbers $X_t \sim U(0, N \times M)\ \forall t \in T$.

II) An ignition probability distribution can be provided by the user, indicating the probability of ignition associated with each cell in an auxiliary file following two main formats:

   i) Each line contains a pair $(i, \mathbb{P}_i(ignition))$ where i corresponds to the cell ID number (starting from the top-left corner) and $\mathbb{P}_i(ignition)$ is equal to its ignition probability (for $1 \geq \mathbb{P}_i(ignition) > 0$).

   ii) An $N \times M$ comma/space separated grid file containing the ignition probabilities of each cell can be provided, representing a numerical heat-map where each entry $p(n, m)$ corresponds to $\mathbb{P}_{n,m}(ignition)$.

   After processing the provided distribution(s), independent random numbers $X_{(n,m)} \sim U(0, 1)$ are generated in a Monte-Carlo scheme for each cell with a positive ignition probability, triggering the corresponding ignition events, and thus, starting the fire spread dynamic. Future releases of Cell2Fire will include support to basic discreet/-continuous distributions such that the user could simply provide the relevant set of cells and the main parameters of the distribution.

III) A random sampling area can be generated from a list of ignition points provided by the user. This is defined by a radius ρ that represents the number of adjacent layers to include — using the provided cells as centers — defining a series of square-shaped sampling neighborhoods $\mathbf{B}_{\rho,t}(i)$ (see Figure 3.2) for each pair (i, t) of ignition points and time periods. Notice that this approach can be naturally extended when providing multiple ignition points for the same period t by simply including a vector of ignition points $\vec{y}$ such that the user provides $(\vec{y}, t)$ instead of (i, t).

   Formally, we define the adjacent layers as the number of "adjacents-to-adjacent cells" to be included during the random sampling procedure surrounding the specified ignition points. Thus, the first adjacent layer for a cell i simply corresponds to its adjacent cells set $Adj_1(i) := Adj(i)$, the second layer to $Adj_2(i) := Adj_1(i) \cup \{l \mid l \in Adj(j)\ \forall j \in Adj_1(i)\}$, and the k-th layer can be defined recursively as $Adj_k(i) := \cup_{r=1}^{k-1} Adj_r(i) \cup \{l \mid l \in Adj(j)\ \forall j \in Adj_{k-1}(i)\}$. Therefore, based on a list of ignition points (i, t) provided by the user for each period $t \in T$ and a radius $\rho \geq 0$ representing the number of adjacent layers to be included in the random sampling procedure, an initial ignition cell for period t is drawn (following a uniform distribution or one specified by the user) from the neighborhood centered in the provided ignition point i $\mathbf{B}_{\rho,t}(i) := Adj_\rho \cup \{i\}$. An example for a homogeneous forest of 20×20 cells can be seen in Figure 3.3 where different fire scar distributions are shown for several ρ values when sampling 10 ignition points.

   Thanks to this approach, the user can test the impact of defining different ignition areas, potentially including his/her own probability distributions to represent specific (and potentially more complex) phenomena. For example, the provided ignition point could represent the most likely area for a lightning strike while the cells inside its neighborhood $\mathbf{B}_\rho(i)$ can be associated with a certain ignition probability as a function

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

$i = 15, \rho = 2$
$Adj_1(15) := \{8,9,10,14,16,20,21,22\}$
$Adj_2(15) := Adj_1(15) \cup \{1,2,3,4,5,7,11,13,17,19,23,25,26,27,28,29\}$
$\mathbf{B}_\rho(15) := Adj_2(15) \cup \{15\}$

Figure 3.2: Random sampling area example. Here, cell 15 is provided as an initial ignition point and a radius $\rho = 2$ is selected. First adjacent layer $Adj_1(15)$ is shown in gray and the second layer $Adj_2(15)$ contains both the light and dark gray cells. The $\rho = 2$ neighborhood of cell 15 $\mathbf{B}_\rho(15)$is then defined as the union between $Adj_2(15)$ and the cell 15. If $\rho = 3$, the $Adj_3(15)$ set will be the union between $Adj_2(15)$ and the white cells.

of their distance to i, among several other possibilities.

In addition to the previous features, Cell2Fire includes a simple lightning-strike model for representing the temporal stochasticity of the ignition phenomenon by selecting the initial ignition period $t_{ignition}$ from a probabilistic model, as an optional feature. Two main models are implemented based on the Poisson distribution: (1) Homogeneous and (2) heterogeneous. In the first one, the user provides an initial lightning-strike rate λ_0 [Fires/t] and Cell2Fire determines if an ignition event occurs at period t, if true, an ignition is triggered (or not) based on one of the previously described methods, otherwise, the simulation advances to the next time step until an ignition occurs.

When using the heterogeneous version, a time-dependent rate $\lambda(t)$ [Fires/t] is used, allowing to modify the lightning-strike probability with respect to the time period t as well as include more complex phenomena (if needed) such as incorporating a peak lightning season reflected in a larger rate λ at time t^*, that later decreases with respect to t. As an example, a lightning-strike model could follow a Poisson distribution with rate $\lambda(t)$ [Fires/t]. Here, $\lambda(t)$ can be an increasing function on t, having a higher probability of ignition in a period t if no ignitions occurred before. We could define:

$$\lambda(t) = \lambda_0(1 + \alpha(t-1)) \qquad (t \geq 1, \alpha \in (0,1)) \tag{3.1}$$

$$\Lambda_{(t_a,t_b)} = \int_{t_a}^{t_b} \lambda_t \mathrm{d}_t = \frac{\lambda_0}{2}(2(t_b - t_a) + \alpha(t_b^2 - t_a^2 - 2)) \tag{3.2}$$

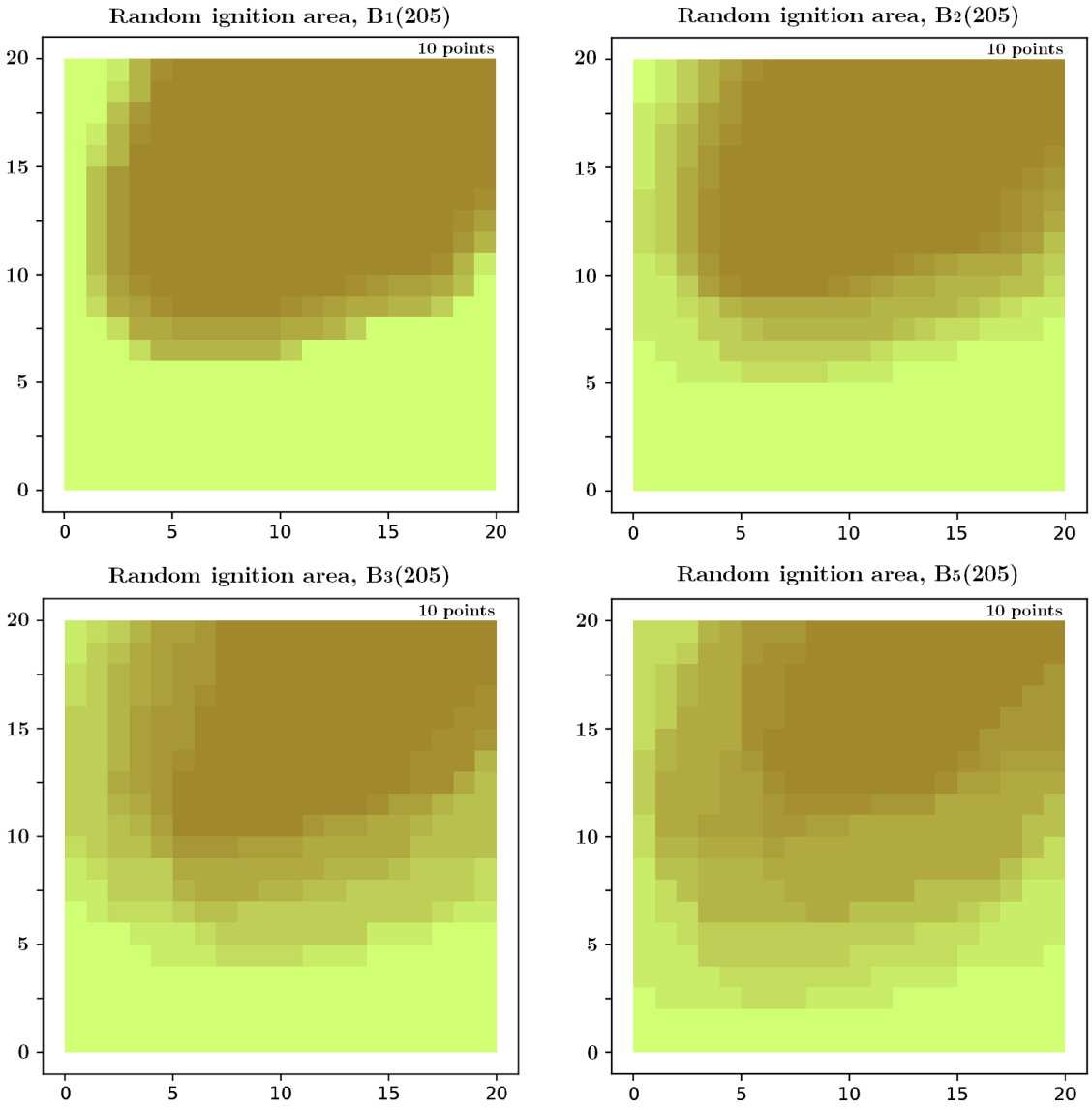


Figure 3.3: Random sampling area example when using $\rho \in \{1, 2, 3, 5\}$ in a 20×20 C1 fuel homogeneous instance, centering the neighborhood in the cell 205 ($\mathbf{B}_\rho(205)$). Ten different random ignition points are generated for each ρ value. Darker colors represent the intersection of more fire scars (more likely fire evolution given random ignition area) while lighter colors represent less likely fire spread dynamics.

Since we are dealing with a non-homogeneous Poisson process with $N(t) :=$ a Poisson counting process of the number of fires up to period t, we have that the probability of having k ignitions until period t is equal to:

$$\mathbb{P}(N(t) = k) = \frac{(\mathrm{e}^{-\Lambda_{(0,t)}t})\Lambda_{(0,t)}^k}{k!} \tag{3.3}$$

Thus, the probability of not having an ignition in period $t = 1$ is:

$$\mathbb{P}(N(1) = 0) = \frac{(\mathrm{e}^{-\Lambda_{(0,1)}})\Lambda_{(0,1)}^{0}}{0!} = \mathrm{e}^{-\Lambda_{(0,t)}} \tag{3.4}$$

Finally, the probability of not having an ignition in period t, if no ignition happened in the previous period is:

$$\mathbb{P}(N(t) = 0/N(t-1) = 0) = \mathbb{P}(N(t) - N(t-1) = 0) = \mathrm{e}^{-\Lambda_{(t-1,t)}} \tag{3.5}$$

Using this approach, ignition probabilities could be related to previous events instead of following the memoryless property present in a homogeneous process. Another example of this approach consists of including a $\lambda(t)$ function with a peak in t^*, the middle of a certain horizon T, and then decrease it until the end of the season:

$$\lambda(t) = \begin{cases} \lambda_0 t, & \text{if } t < t^* \\ \lambda_0(\phi_t - t), & \text{otherwise} \end{cases}$$

Where ϕ_t is a parameter that can change through time (or not), allowing the user to test different probability distributions: symmetric, higher probability in extreme periods, etc.

Using all the previous elements, Cell2Fire allows the user to test and incorporate, in a simple and flexible way, different models of uncertainty for the fire ignition process.

3.3.2 ROS: approximation error

Following a simplification of the proposal in Han and Braun (2014), a random term is included that represents the inherent uncertainty when calculating the ROS that does not depend on weather conditions. Here, we include randomness in Cell2Fire as follows: if the user-supplied parameter cv_{ROS} (representing the ROS coefficient of variation) is greater than zero, random numbers χ are generated from a $Normal(0, 1)$ distribution, perturbing the values of the main ROS components by a factor of $(1 + cv_{ROS}\chi_{i,t})$, for $i \in \{HROS, BROS, FROS\}$, $t \leq T$. Therefore, the user can incorporate the inherent stochasticity of the calculated Rate of Spread, testing different levels of error/approximation via the provided cv_{ROS} parameter.

For example, we can see in Figure 3.4 how different fire scars are obtained when perturbing the original ROS using a coefficient of variation equals to 0.5 on 10 simulations. From the results, we can see that some fires are not able to spread from the ignition point to the rest of the forest, obtaining a darker area in that particular cell. On the other hand, we can notice fuzzy borders on the ellipse including less intense cells, representing other possible (and less likely) fire scars. Clearly, increasing the cv_{ROS} leads to a greater fire evolution variance, obtaining a wider variety of final fire scars.

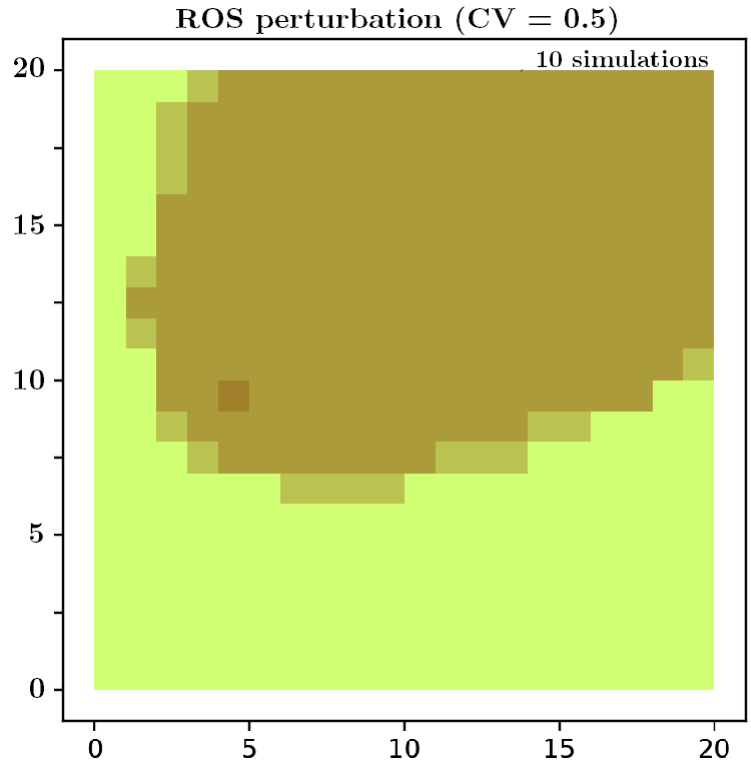


Figure 3.4: ROS perturbation example using $cv_{ROS} = 0.5$ for 10 simulations. As expected, the fire scar is concentrated in the ignition point and then, some variations are obtained as can be seen in the borders of the ellipse: depending on the random perturbation, the total magnitude of the fire can be slightly different. Notice that in this example, some fires are not able to propagate after the initial ignition, giving more weight to the ignition point.

3.3.3 Weather scenarios

Users can provide sets Ω of weather stream files defining a series of "weather scenarios" $\omega \in \Omega$ for the simulations by varying several conditions such as wind speed/direction, temperature, relative humidity, radiation, among other characteristics supported by the FBP system. Different scenarios ω will be reflected in different rate of spread values (ROS_ω), leading to different fire spread evolution depending on the conditions provided. In addition, a probability distribution Φ_Ω can be associated with the scenario set Ω by including the probability of occurrence $p_\omega \sim \Phi_\Omega$ of each scenario $\omega \in \Omega$. — e.g., giving flexibility for modeling rare events/black swans — allowing Cell2Fire to include this information and provide insightful statistics of the most likely fire evolution as well as the distribution of the fire scar for each simulated period $t \in T$ for a particular set of scenarios.

As an example, we can see in Figure 3.5 how different fire scars can be obtained when performing a random weather scenario selection from a set Ω containing all possible wind directions ($\theta \in [0, 360[$), allowing the user to easily visualize the most likely fire scar given the selected scenarios.

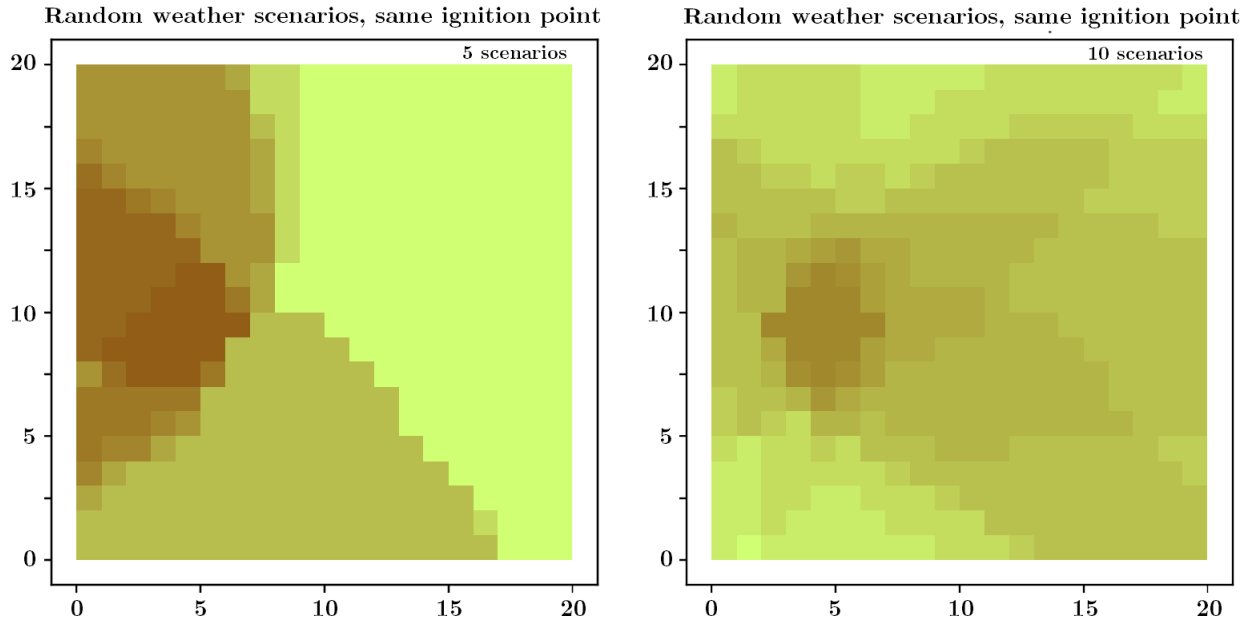


Figure 3.5: Multiple random weather scenarios — with different wind direction — output in a 20×20 homogeneous instance (fuel type C1). Darker zones represent more likely fire spread evolution/scars than lighter zones. If a probability distribution Φ_Ω is provided, colors are proportional to each weather scenario ω probability p_ω.

Therefore, fire managers will be able to assess and understand the impact of different weather conditions on the wildfire dynamic in a heterogeneous landscape context, using these insights as inputs for decision support systems.

3.3.4 Cell2Fire Pseudocode

Important is to highlight that all the previously discussed sources of uncertainty — ignition points, ROS approximation error, and weather scenarios — can be easily combined in any way for the ongoing research, giving even more flexibility to the user when studying the fire evolution in a stochastic framework.

We use the following notation to describe the main steps of the simulation Algorithm 2:

- $\mathcal{N}$ denotes the set of cells in the forest and $i \in \mathcal{N}$ represents any cell.
- $Adj\,(i)$ denotes the set of available adjacent cells of i (at most 8).
- $BurningCells$ is the set of burning/active cells in the simulation process.
- $i \rightarrow_{(\chi,\omega,t)} j$ indicates that cell i "sends a message" to cell j at time t for a weather scenario $\omega \in \Omega$ and possibly with a disturbance $(1 + cv_{ROS}\chi)$ if $cv_{ROS} > 0$. It is important to note that the message that travels from i to j depend on the weather ω and the variation $(1 + cv_{ROS}\chi)$ because the Rate of Spread is a direct function on the spreading axis, i.e., $ROS_{ij}(t) = f(\omega_t, \chi_t)$ for some function f.
- $Stochastic \in \{0, 1\}$ is a binary parameter equals to 0 if the simulator model is run in

its deterministic form, and 1 for its stochastic version.

- $Choice\,(\Delta, Stochastic)$ represent a random draw using the Inverse Transform Method for the corresponding random variable Δ. It is activated if $Stochastic = 1$.

Algorithm 2 Cell2Fire Pseudo-code

procedure SIM($ForestData, FTypes, Ignitions, Weather, TMax, Options$)
  Step 0: Initialize Cell2Fire
    Parse inputs, options, read data, initialize objects
    $FPeriod \leftarrow 0, nsim \leftarrow 0, Stochastic \in \{0, 1\}$, Set radius $\rho \geq 0$ and $cv_{ROS} \geq 0$
  Step 1: Ignition and Weather
    $ic \leftarrow$ *Choice(Ignitions, Stochastic)*
    $\omega \leftarrow$ *Choice(Ω, Stochastic)*
    $\chi \leftarrow$ *Choice($N(0,1)$, Stochastic)*
    $BurningCells \cup \{ic\}$
    $FPeriod \leftarrow 1$
  Step 2: Fire Dynamics (Send-Receive)
    Let $i \in BurningCells$, ***if*** $i \rightarrow_{(\chi,\omega,FPeriod)} j$, *where* $j \in Adj(i)$:
      $BurningCells \cup \{j\}$
    $FPeriod \leftarrow FPeriod + 1$, *Update Forest, Weather*
    Repeat, until $FPeriod < TMax$
  Step 3: Results and Outputs generation
    Generate Grids, Statistics, Output plots
end procedure

The main outputs of our system refer to fire risk assessment metrics. The most used is called: Map of probability of burns. This is a value assigned to a cell, and is defined as the percentage of times a cell is burned in simulation R replicas. However, for the evaluation to be reliable, the user must be sure to properly model the variability in the system inputs and specify a large-enough number of iterations.

*In the cell selection process, it is assumed that there is a selection of the ignition time; not explicitly mentioned.

3.4 On Practical Applications

In this section we address the inclusion of uncertainty through some practical applications that a decision maker could use. By including uncertainty, Cell2Fire can capture and model the real fire spread stochasticity, being able to analyze the most likely fire dynamics and fire scar evolution under certain environmental conditions as well as detect the potential risks of extreme scenarios (Black Swans) with low probability of occurrence but with significant impact in the decision making process. This methodology allows the decision maker to test different forest management plans including the inherent stochasticity (and risk) of the fire spread dynamic by defining a series of potential weather/ignition scenarios as well as accounting for the approximation error when calculating the rate of spread, being able to design strategies for decisions under uncertainty as well as incorporating relevant risk measurements.

For our applications in this section, we use real landscapes. One is located in the region of Alberta, Canada (Dogrib-instance) and the other in British Columbia (Arrowhead-instance). The first is a patch that has a surface of 79,611 ha, divided into 100×100 $[m^2]$ cells. It can be downloaded from http://www.firegrowthmodel.ca/prometheus/software_e.php. The second, that we denoted as Arrowhead, has an area of 159,963 ha. The full British Columbia instance is part of the Burn-P3 documentation and can be downloaded from http://www.firegrowthmodel.ca/burnp3/software_e.php.

The module of selection of weather scenarios, we feed it with historical data, obtained from the Climate Information Section of the Agriculture and Forestry site of Alberta, Canada, and data from the Yaha Tinda Auto station (coordinates: 51.6547°, -115.3617°). This is the station closest to the coordinates of the forests used for simulation. First, we extracted the data of the daily averages of the most important climatic indicators (Temperature, precipitation, wind speed, and direction) of the last 5 calendar years (from 01/01/2014 to 12/31/2018). From this data, we chose the 40 days with the worst weather conditions, which favor the propagation of fire and proceeded to obtain 100 scenarios of 24 hours, including the indexes of the Canadian Forest Fire Weather Index (FWI) System. Due to their dimensions (small), fires are simulated during 12 hours in the Alberta province, while the full day hourly stream scenarios are used in the British Columbia instances.

3.4.1 Application 1

As mentioned in Section 3.3, relevant visualization outputs are generated in order to easily understand the effect of activating different uncertainty sources. For example, in Figure 3.6, we can see and compare the deterministic fire scar (left side) with potential fire scars obtained when activating and combining different uncertainty sources. From this, we can identify potential fire dynamics that represent a real threat to a significant area of the forest that were not captured by a deterministic simulation, not being able to represent and evaluate the real risk of having a fire with similar characteristics. Thanks to this information, fire managers could take into account the variability of the fire evolution, being able to protect other sections of the forest based on a risk analysis: e.g., by identifying how likely a cell can

get burned, leading to larger fire scars.

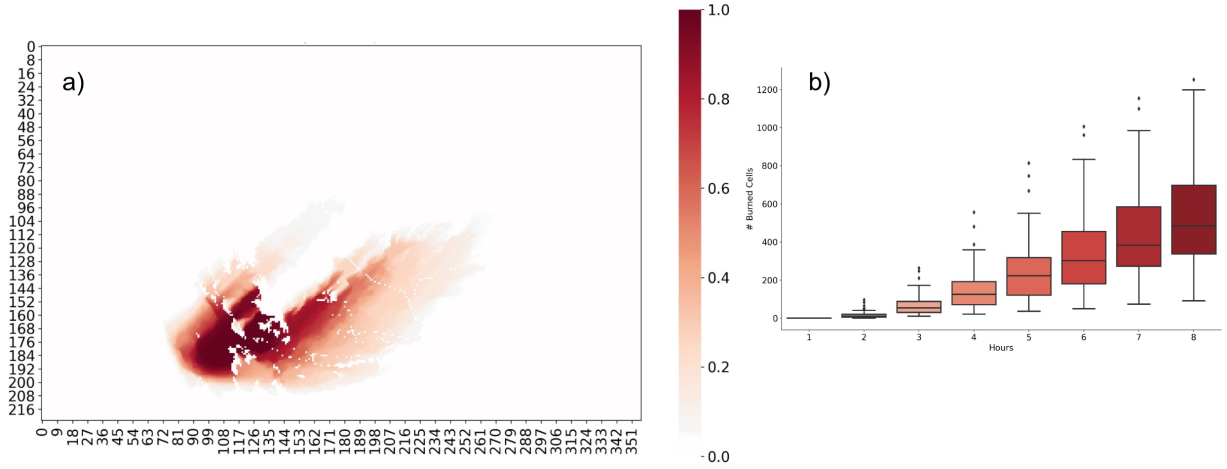


Figure 3.6: Burn Probability Map for Dogrib-instance, using $R = 100$ replications, with $cv_{ROS} = 1.0$. and $\rho = 2$.

Following this idea, Cell2Fire generates probability maps and provides an analysis of the evolution and distribution of the number of burned/available cells per time step via plots and statistical sheets. In Figure 3.6, we can see a probability map and the hourly distribution of the number of burned cells for the Dogrib instance, when simulating 100 different scenarios including the main three sources of uncertainty discussed in the previous sections. From them, we can observe the most likely fire scar (darker red with probability close to 1) given the initial conditions provided, and identify potential risks such as the probability of having a fire that is able to reach the top-right corner of the instance (with an average probability of 0.4), as well as the hourly variability and distribution of the number of burned cells for different scenarios. In this example we can observe an increasing pattern: a small variation is obtained during the initial hours since the main source of uncertainty consists of the initial ignition point (including two adjacent layers, $\rho = 2$), and then, the rest of the uncertainties start to play a significant role as advance the simulation-time, significantly impacting the fire spread evolution for certain scenarios increasing the variability in the distribution of burned cells as seen in the box-plot (see Figure 3.6, right side).

The combination of these outputs represents and summarize a powerful and useful set of statistical tools included in the simulator that can be applied when using Cell2Fire as part of a decision support system, being able to analyze the effect and impact of each uncertainty source (with different levels and combinations) on the landscape under study. A detailed analysis of the statistical tools included in Cell2Fire — their interpretation, potential uses, and outputs —will be included in a future research.

3.4.2 Application 2

We model a forest as a two-dimensional lattice with identical cell areas. The attributes and spatial relation between these cells are modeled using an undirected planar graph $G = (\mathcal{N}, \mathcal{E})$

where the set $\mathcal{N}$ represent the forest cells (graph nodes) and $\mathcal{E}$ the edges between neighboring cells, determined by the structure of the forest.

From the cellular structure of Cell2Fire, when a fire occurs, a messaging process is generated between the nodes of G during the simulation, obtaining a directed graph $G_D = (\mathcal{N}_D, \mathcal{E}_D)$ where $\mathcal{N}_D \subseteq \mathcal{N}$ is the set containing all the cells burned during the replication and $\mathcal{E}_D$ is built from these signals, representing the fire propagation between adjacent cells.

In this dynamic between the cells, fire propagation lines are generated, clearly identified as shown in Figure 3.7. On the left side, a fire is shown in the Arrowhead instance. On the right side two graphs are shown: the one above shows the fire propagation lines ("Propagation Tree"), and on the bottom one the Rate of Spread (ROS) of fire when the contagion occurs, represented on a heatmap.

The propagation lines and the values of the ROSs could help firefighters away from those dangerous areas or eventually protect areas where fire will arrive faster.

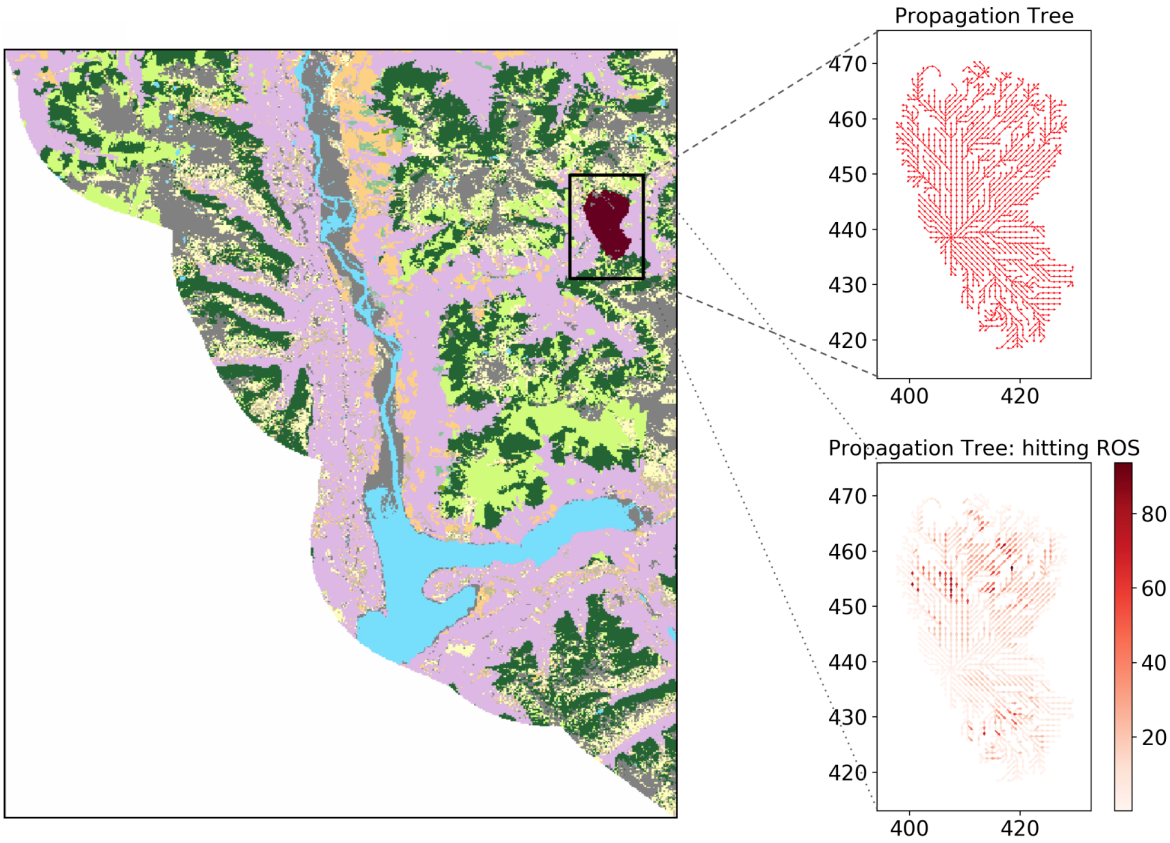


Figure 3.7: This figure shows a fire in the Arrowhead region (left side) and its representation via propagation lines and a heatmap for the hitting ROS (right side), $R = 1$.

3.4.3 Application 3

The concept of burn probability maps (BP-maps) is commonly used to assess the likelihood of burning (Ager et al., 2007, 2010; Finney et al., 2011b; Moritz et al., 2005; Parisien et al., 2005b). Several software such as Burn-P3, FSPro, and FlamMap (Parisien et al., 2005b, 2011; Finney et al., 2011b) can calculate these values using sophisticated fire growth algorithms to produce high-resolution spatial estimates. Fuel treatment applications using this approach can be reviewed in Liu et al. (2013); Parisien et al. (2006); Salis et al. (2018)

Although the methodologies differ somewhat (in the aforementioned softwares), they all incorporate the main components included in our study: random selection of ignition point, selection of a climate scenario, recording burned cells and repeating the process a number of times (see Figure 3.1). The main difference with our approach is that they do not consider a disturbance in the ROS, therefore, in the event of randomly selecting the same ignition point and the same climate scenario, their systems get the same fire scar.

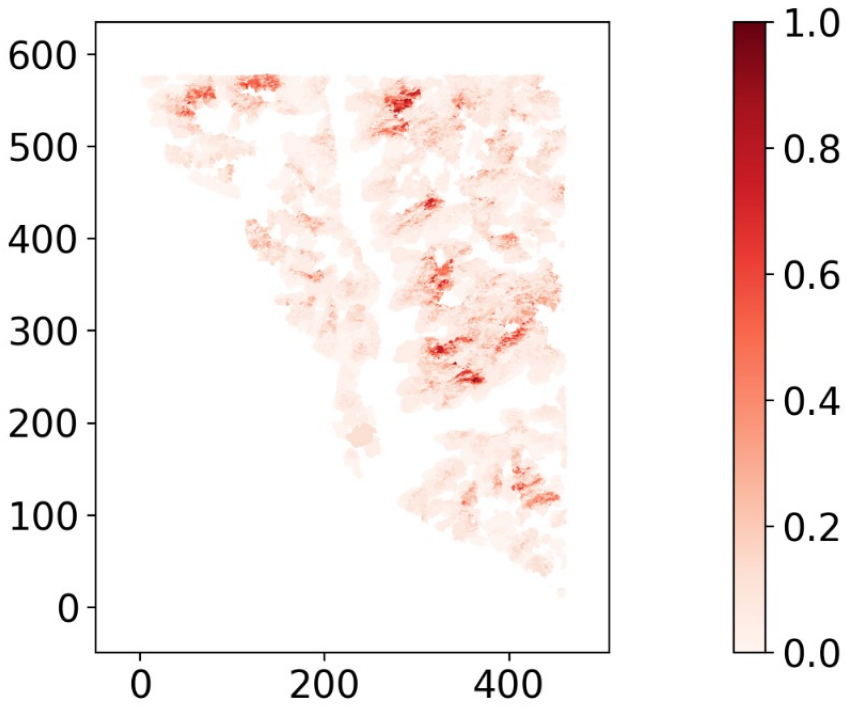


Figure 3.8: This figure shows Burn Probability Map for Arrowhead region, for $R = 100$ replications and $|\Omega| = 100$ weathers scenarios of 24 hours.

3.5 Conclusions

Cell2Fire can work as a deterministic simulator (Pais et al., 2019), or with stochastic components as introduced in this chapter. In the present study, we have shown how to use and exploit Cell2Fire when including uncertainty. As a kick start, we have considered randomness based on three sources: ignition points, through weather scenarios, and via a controlled

random error in the rate of spread. These — combined or isolated — affect the shape of fire front-line over time by changing the fire ROS.

Although Cell2Fire is quite robust, some sources of uncertainty are left out. Examples of them: duration of fire, wind direction, wind speed from a probability distribution (as in Hajian et al. (2016)), temperature, and relative humidity among others. Further elements such as the inclusion of the spotting phenomenon and its probabilities will be discussed in a future research as well as the calibration and application of these new components to a real case.

A stochastic fire spread simulator helps to predict the wildfire spread more accurately and realistically and also as a support for the decision maker, helping her to select the best actions in firefighting and/or Forest Management by providing a certain level of confidence of the predictions (Brandes, 2001; Palma et al., 2007b; Parisien et al., 2006). This allows for better evaluation of the risk involved in the decision-making process. For example, Burnt Probability maps have enormous potential to support strategic fire and fuel management planning decisions at a landscape level (Martell, 2015; Thompson and Calkin, 2011). Alongside this, the analysis of "stochastic fire contours" — a precise estimate of the fire perimeter at a certain time — could improve the determination of escape routes or support operational decisions such as where to locate the firefighters/drones/helicopters (suppression forces) depending on the observed fire scenario.

Chapter 4

Evaluating the effectiveness of Fire-Smart Forest Management plans under wildfire risk

Authors

Jaime Carrasco and Andrés Weintraub
University of Chile, Industrial Engineering Department

Cristobal Pais and Max Shen
University of California Berkeley, IEOR Department

Abstract

The destructive potential of wildfires has been exacerbated by climate change, causing their frequencies and intensities to continuously increase globally. Generating fire-resistant terrains via efficient and calculated fuel-treatment plans is critical to protecting relevant areas including native forests, agricultural resources, biodiversity, and human populations. To address this challenge, we propose a framework that integrates fire spread, optimization, and simulation models. We introduce the concept of *Downstream Protection Value* (DPV), a flexible metric that assays and ranks the impact of treating a unit of the landscape, by modeling a forest as a network and the fire propagation as a tree graph. Using our open-source decision support system, the decision maker can optimize custom performance metrics to minimize wildfire losses, obtaining an effective treatment plan. Experiments with real forests show that our model is able to consistently outperform traditional methods and accurately detect high-risk and potential ignition areas, focusing the treatment on the most critical zones. Results indicate that our methodology is able to decrease the expected area burned and fire propagation speed by more than half in comparison to alternative methods under ignition and climatic uncertainty.

4.1 Introduction

The Amazon forest, the largest expanse of rain forest in the world has experienced over 74,000 fires this year. This represents an 84% increase in fires observed by the same time with respect to the previous year (Dunne and Mcsweeney, 2019; INPE, 2019). This explosion of fire occurrences has decreased its ability to perform key tasks for the environment such as carbon sequestration and the accommodation of a large number and variety of species (Balch et al., 2015; Gibbens, 2019; Rappaport et al., 2018). Sadly, this phenomenon is not limited to a particular region, for example, The Congo forest, the second-largest rain forest in the world located in Central Africa has experienced similar if not worse fire trends this year (Turkewitz, 2019). In the United States, California experienced the largest recorded fire in its history in 2018 where a single fire destroyed more structures than any other in modern history (Edwards, 2019). Research indicates that these fire events will increase due to the effects of climate change on temperature, precipitation levels, and soil moisture, having increased the number of fires and area they consume around the globe, generating larger events than before (Running, 2006; Westerling, 2016; Westerling et al., 2006).

The current incidents indicate that preemptive policy measures must be taken to reduce the risk of fire occurrence (Curt et al., 2016b), managing the land in an effective way (Calviño-Cancela et al., 2016) to protect natural forests, agricultural areas, and human lives. These concepts are included in what is known as FireSmart Forest Management (FSFM). As discussed in Hirsch et al. (2001a), this paradigm considers opportunities in three dimensions: i) decrease of the fire behavior potential of the landscape, ii) reduction of the potential for fire ignitions, and iii) increase in the fire suppression capability.

Several researchers have reported that the intensity and severity of wildfires can be reduced

through fuel management activities. Targeted fuel treatments and fuel-breaks are used to alter the composition of wild land vegetation and forest in order to modify the behavior of future fires. There are multiple fuel treatments activities to delay the spread of a fire such as cutting and clearing wood, prescribed burns, and thinning (Agee and Skinner, 2005). These actions are mainly used to decrease the intensity and size of potential wildfires, but they could also affect species diversity (Shinneman et al., 2019), improve the health of ecosystems, and suppress fires, among many others (Finney, 2001; Hirsch et al., 2001b; Moghaddas and Craggs, 2008).

Trailblazing research by Finney (2001, 2004), indicates that it is possible to modify fire behavior and progress across landscapes through strategic placement of treatments. In Salis et al. (2016), the authors discuss strategies to define treatments using burning probability maps, the area burned or the flame length. Other researchers have focused their attention on finding the optimal spatial allocation of prescribed burning activities (Alcasena et al., 2018; Matsypura et al., 2018) and the design of fire breaks to control of fire spreading (Russo et al., 2015).

In Parsons and van Wagtendonk (1996), authors report that the systematic segmentation of fuel can influence fire size and fire growth. However, the empirical evaluation of the efficiency of a fuel-break and fuel treatment is difficult, if not impossible. The climatic conditions, as well as the state of a forest (e.g., land cover) at any given moment, are non-reproducible. This unreliability in physical conditions is one of the advantages of deterministic fire simulators such as FARSITE, Prometheus and Wildfire Analyst (Finney, 2004; Tymstra et al., 2010a; Ramírez et al., 2011), which can reproduce fires with and without treatment activities. However, these simulators do not have a well-structured interface to facilitate data exchange during multi-period interactions – e.g., to systematically evaluate fuel management in operational or tactical planning – and do not consider temporal changes of vegetable fuels. Additionally, evidence of the high stochasticity of this problem and inconsistency of forestry tools are highlighted in Barnett et al. (2016) where the authors show that, historically, the majority of treatments rarely intersect with wildfire occurrences.

Therefore, finding an effective treatment layout across a landscape – even for a single time period – is not a simple task. Recent work in Chung (2015) discuss challenges such as timing, selection of an adequate fuel-treatment method, and the impact of high uncertainty levels over multiple time periods (e.g., climate and ignition). To fully address the optimal treatment problem, the development and systematic evaluation of effective fire risk metrics in an integrated, robust, and flexible decision support system is required, a problem that remains open.

In this chapter, we develop and introduce the theory and concept of downstream protection value (DPV), a flexible metric that measures and ranks the impact of treating a unit of the landscape to minimize wildfire losses. Our objective is to select the subset of cells, defined as homogeneous sections of the forest with similar fuel type, age, and topographic conditions, with the highest impact on interrupting the fire propagation by incorporating DPV into an open-source, integrated Decision Support System (DSS), providing fire managers with optimal landscape treatment plans under wildfire risk. To assess the performance of our new methodology, we generate and apply the fuel treatment plans in two Canadian regions, using

traditional fire risk metrics (Brandes, 2001; Freeman, 1977; Palma et al., 2007a; Parisien et al., 2005c) and our DPV metric to select the treated cells. We discuss the sensitivity of traditional metrics and DPV with respect to their learning capacity and performance under multiple sources of uncertainty (e.g., climatic conditions). Their performance is measured by estimating (i) expected wildfire losses, (2) average fire propagation speed, and (3) the fire prevention rate.

4.2 Methods

4.2.1 Study area

In this work, we use public data of real forest patches from Alberta and British Columbia provinces, Canada (Table 4.1). The Alberta instance is a sample data set of the Rocky Mountain subregion for the 2001 Dogrib fire[1]. This patch has a surface of 79,611 ha, divided into 100×100 $[m^2]$ cells. British Columbia instance is a very large section (1,854,838 ha) and it is characterized by mountainous surfaces and discontinuous fuel levels. In order to analyze the possible effects of topography on risk measures and accelerate our computational capacity, we built 6 sub-instances from it. We denoted them as ArrowHead (AH - 159,963 ha.), Revelstoke (RT - 350,956 ha.), Mica Creek (MC - 304,781 ha.), Glacier Natural Park (GNP - 464,664 ha.), Central Kootenay (CK - 399,401 ha.), and Neptune Peak (NP - 350,956 ha). We picked this naming convention following the names of the mountains within each patch. The full British Columbia instance is part of the Burn-P3 documentation[2]. Each instance contains all data layers (fuel composition, topographic characteristics, and climatic conditions) for each cell.

Instance	Area [h]	Elev. range [m]	Mean elev. [m]	Slope range [°]	Mean slope [°]	Dominant Fuel	Total fuels
Dogrib	79,611	[1299, 2825]	1693	[0, 189]	21.04	Boreal Spruce	11
AH	159,963	[430, 2894]	1329	[0, 220]	42.96	Red/White Pine	11
RT	350,956	[429, 2997]	1494	[0, 251]	45.91	Red/White Pine	11
MC	304,781	[516, 2895]	1523	[0, 230]	48.41	Non-fuel	11
GNP	464,664	[300, 3000]	1701	[0, 149]	45.79	Non-fuel	11
CK	399,401	[300, 3000]	1811	[0, 144]	51.64	Non-fuel	11
NP	350,956	[429, 2997]	1494	[0, 251]	45.91	Red/White Pine	12

Table 4.1: Basic description of the instances considered in this study including total area, topographic characteristics, and land-cover.

4.2.2 Uncertainty

The location of an ignition point can be configured deterministically (FI) or randomly (RI). In a stochastic setting, ignition locations can be drawn from a given probability distribution,

[1] http://www.firegrowthmodel.ca/prometheus/software_e.php
[2] http://www.firegrowthmodel.ca/burnp3/software_e.php

otherwise, a uniform distribution is set by default. The number of ignitions is set to be one per replication.

Weather scenarios include hourly observations of temperature, relative humidity, wind speed, wind direction, as well as their associated Fire Weather Index (FWI) System (Van Wagner and Pickett, 1987) fuel moisture codes and fire behavior indexes from the FBP System (Hirsch et al., 2001a). The selection of weather scenarios can be deterministic (FW) or random (RW). In the first case, a pre-defined scenario is drawn. Otherwise, simulations are performed using a set of weather scenarios Φ, selected to represent multiple contexts in which a fire can spread, capturing the variability of this phenomenon. However, it is necessary to include only scenarios that boost the propagation of fire. As mentioned in Finney (2005); Finney et al. (2011b); Parisien et al. (2005b) large fires are more useful to risk assessment. Historical data are obtained from the Climate Information Section of the Agriculture and Forestry site of Alberta, Canada, and data from the Yaha Tinda Auto station. This is the closest station to the coordinates of the forests used for simulation. We extract data of the daily averages of the climatic parameters of the last 5 calendar years (2014 to 2018). From this data, we chose the 40 days with the worst weather conditions, which favor the propagation of fire and proceeded to obtain $|\Phi| = 100$ scenarios of 24 hours.

4.2.3 Fire modeling

In this study we use Cell2Fire, a new fire growth model, presented in Pais et al. (2019). This fire simulator has a similar performance to Prometheus (the most prominent simulator in Canada, see more details in Pais et al. (2019)). Both consider the same aspects of a landscape such as: slope, aspect, and elevation; fuel types from the FBP System (Forestry-Canada, 1992); and weather to simulate fire growth. Cell2Fire is a cellular automata – fire spreads from cell to cell – unlike Prometheus which is based on wave propagation equations (Tymstra et al., 2010a). This is an advantage of Cell2Fire, because it can save and change the status of the cells in simulation time. During a simulation, a cell can be in one of five states: *"Available"*, *"Burning"*, *"Burned"*, *"Harvested"*, or *"Non-Fuel"* where the label *"Available"* indicates that the cell contains a flammable fuel type; *"Burning"* represents cells containing an active fire; *"Burned"* indicates that the fire has consumed the fuel available and passed through the cell; and *"Non-Fuel"* is a non-flammable fuel type such as rivers, lakes, or rocks. The *"Harvested"* state is provided so that the simulator can be embedded in a strategic harvesting planning system, as we will see in section 4.2.5. The harvesting planning module will be responsible for labeling the cells that are harvested and provide the appropriate post-harvest fuel type.

4.2.4 Measuring Wildfire Risk

We experiment and compare our new metric with three established approaches: *Burn Probability Maps*, *Betweenness Centrality*, and *Fire Protection Value (FPV)*. The concept of burn probability maps (BPM) is commonly used to assess the risk of fire (Ager et al., 2007, 2010; Finney et al., 2011b; Moritz et al., 2005; Parisien et al., 2005b). Several software such as

Burn-P3, FSPro, and FlamMap (Parisien et al., 2005b, 2011; Finney et al., 2011b) can calculate these metrics using different methodologies. Fuel treatment applications using this approach can be reviewed in Liu et al. (2013); Parisien et al. (2006); Salis et al. (2018). Another metric applied in the context of fuel management is the betweenness centrality (BC) value (Brandes, 2001). It aims to identify and rank those areas that act as a link between sections of the forest. For example, in Gray and Dickson (2016), the authors use this metric for the strategic placement of fuel treatments to mitigate the spread of fire. Other applications of this method in this context can be seen in Russo et al. (2016). In Palma et al. (2007a), the authors developed the fire protection value (FPV). Using this metric, the decision maker can identify sections of the forest where the fire tends to travel faster, focusing the treatment in these critical locations. The main difference with our approach is that the model is static in the sense that it relies on unique weather conditions failing to capture temporal interactions.

4.2.5 Downstream Protection Value

We model a forest as a two-dimensional lattice with identical cell areas. The attributes and spatial relation between these cells are modeled using an undirected planar graph $G = (\mathcal{N}, \mathcal{E})$ where the set $\mathcal{N}$ represent the forest cells (graph nodes) and $\mathcal{E}$ the edges between neighboring cells (see Figures 4.2 and 4.1). Nodes $i \in \mathcal{N}$ are associated with relevant attributes $\vec{\nu_i}$ such as total volume available, selling price per cubic meter, treatment costs per area, fuel type, and any other relevant data provided by the user (see Pais et al. (2019)). Specific attributes can be considered for edges such as distances between cell centers, slope, transportation costs, among others.

When a fire occurs, a messaging process is generated between the nodes of G during the simulation, obtaining a directed graph $G_D = (\mathcal{N}_D, \mathcal{E}_D)$ where $\mathcal{N}_D \subseteq \mathcal{N}$ is the set containing all the cells burned during the replication and $\mathcal{E}_D$ is built from these signals, representing the fire propagation between adjacent cells. Based on the network structure and the characteristics of the fire-growth model, we introduce the following notations/concepts:

- $Adj\,(i)$: denotes the set of available cells adjacent to $i \in \mathcal{N}$.
- $t\,(i)$: represents the time at which the fire reaches the center of cell i and starts propagating to its neighbors.
- We denote $i \rightarrow_t j$ if cell i is **"sending a message"** to cell j at time-step t, simulating and updating the fire evolution along the minimum distance between the center of both cells. Here, we assume that $t \leq t\,(j)$, the time to reach the center of cell j. When $t = t\,(j)$, the cell j changes its state to *"burning"* and we establish that **"i sent a message to j"** ($i \rightarrow_{t(j)} j$). Following this notation, we can represent the fire traveling time from the center of cell i to the center of cell j. Namely, $t_{(i,j)} = t\,(j) - t\,(i)$, and thus if $i \rightarrow_t j$ then $t\,(i) < t \leq t\,(j)$.
- For notation simplicity, we denote by $i \rightarrow j$ when the fire propagating from cell i reaches the center of cell j, i.e. $i \rightarrow_{t(j)} j$ occurs. Only the **shortest paths** are registered during the execution of the simulator.
- A cell i becomes inactive when no more messages can be sent from i. This occurs when no adjacent cells are available, $Adj(i) = \{\emptyset\}$. (e.g., surrounded by non-flammable or

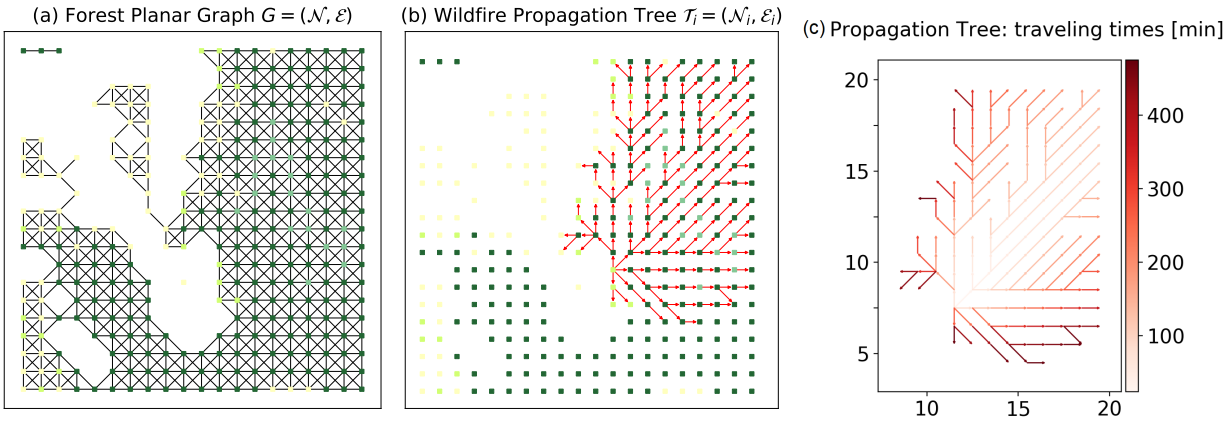


Figure 4.1: (a) Planar graph, (b) simulated "Propagation Tree" from a heterogeneous instance, and (c) fire shortest traveling times. Red arrows indicate edges where fire was propagated. The root i of the tree $\mathcal{T}_i$ corresponds to the ignition point of the fire.

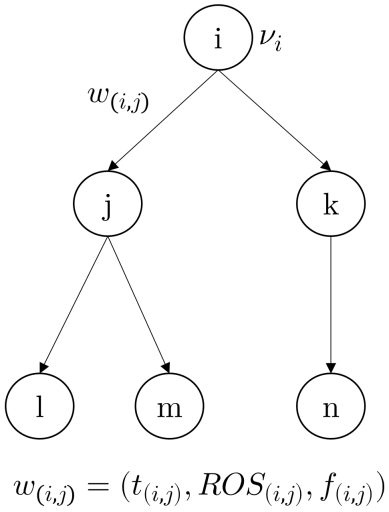


Figure 4.2: Example of a "Propagation Tree" graph generated from Cell2Fire. Cells are represented by nodes and the edges indicate that fire was propagated from one cell to another during the simulation. Weights $w_{(i,j)}$ include information w.r.t. the traveling times, average rate of spread (ROS), and usage frequency.

already burned cells).

We emphasize that the notation "$\rightarrow$" is only defined for adjacent cells on graph G, so in what follows, we extend its definition.

Definition 1 We say that "i **reaches** j", and we denote it by $i \rightarrow j$ if there are $i_1, i_2, ..., i_p \in \mathcal{N}$ such that:

$$i \rightarrow_{t(i_1)} i_1 \rightarrow_{t(i_2)} i_2 \rightarrow \cdots \rightarrow i_p \rightarrow_{t(j)} j.$$

Note that the fire traveling time from i to j is given by:

$$(t(i_1) - t(i)) + \cdots + (t(j) - t(i_k)) = t(j) - t(i).$$

Definition 2 We say that the graph $\mathcal{T}_i = (\mathcal{N}_i, \mathcal{E}_i)$ is the **"Propagation Tree"** of the cell i if:

- $i \in \mathcal{N}_i$ is the root of the tree $\mathcal{T}_i$, where $\mathcal{N}_i = \{j : i \to j\} \cup \{i\}$.
- $\mathcal{E}_i = \{(k, j) : k, j \in N_i, k \leftrightarrow j\}$ where $\leftrightarrow$ represents the existence of an edge between cells k, j due to the propagation of a fire message during the simulated wildfire.
- Each edge $e = (k, j) \in \mathcal{E}_i$ is associated with a weight $w_{(k,j)} = (t_{(k,j)}, ROS_{(k,j)}, f_{(k,j)})$ including the traveling time between cells, average ROS during the new ignition, and usage frequency.

The concept of "Propagation Tree", although simple, is very powerful and useful. Given any cell inside the forest, it summarizes the fire propagation dynamics during a simulation, starting from this particular root cell. Intuitively, we would like to decrease the size of each "Propagation Tree", meaning that the number of cells burned after the wildfire season is minimized. We can easily obtain useful insights and assess the risk of a cell during a fire season by noting the size of its tree. The bigger $\mathcal{T}$, the larger the number of cells burned because of the presence of the root cell, acting as the shortest link between two different areas.

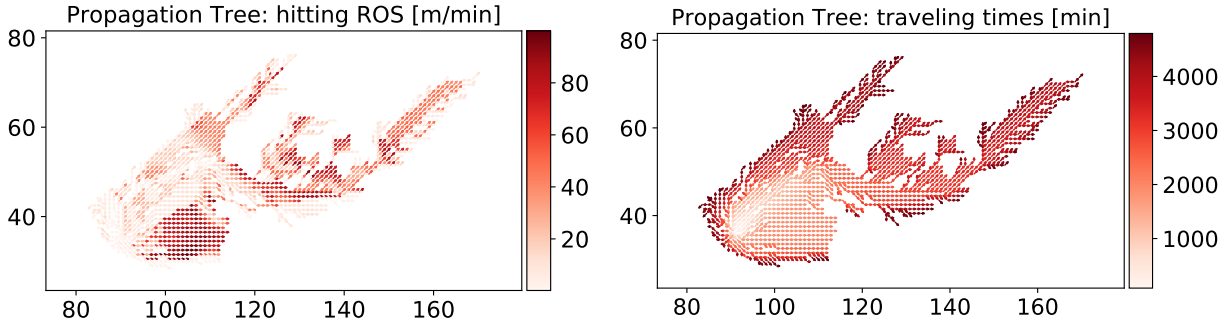


Figure 4.3: Propagation Tree example from Dogrib instance. Useful information regarding the ROS [m/min] and traveling times [min] are registered at the spreading times between cells.

Traveling times recorded inside the "Propagation Trees" provide the researcher with useful information regarding the shortest propagation paths inside the forest. From the discussion above, each edge (i, j) represents the shortest path of the fire from the ignition point to cell i and from i to j. Furthermore, average propagation time and ROS can be calculated for a tree $\mathcal{T}_i$, providing practical insights for the decision maker when deciding the fuel treatment plan by also incorporating relevant elements for operational fire suppression (e.g., average propagation speed) that can be critical in the event of a future wildfire. From these outputs, interesting conclusions can be extracted with respect to the role of different fuel types inside the propagation dynamic, as well as climatic conditions that affect the effective ROS between cells, among others (see Figure 4.3).

We introduce the concept of "Global Propagation Tree", the tree generated from the superposition of individual trees (Figure 4.4). In the context of multiple replications – wildfire simulations under different conditions – this graph allows us to calculate all the required parameters for our optimization model such as ignition probabilities, expected area burned, average rate of spread, and risk metrics.

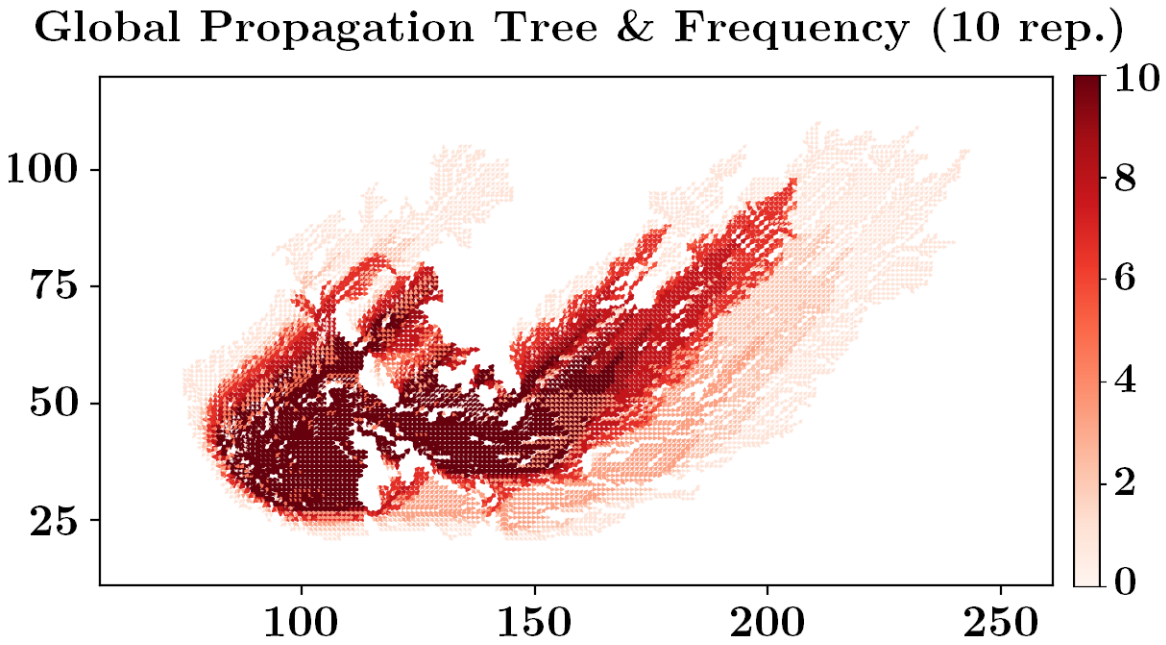


Figure 4.4: Global Propagation Tree obtained from 10 replications starting from the same ignition point including uncertainty in the weather scenario $\phi \in \Phi$ in Cell2Fire. Edges frequency usage is indicated with colors.

Definition 3 We define the graph $\mathcal{GT}_{\mathcal{R}} = (\mathcal{N}_R, \mathcal{E}_R)$ as the **"Global Propagation Tree"** of the landscape after R replications where we have:

- The set $\mathcal{N}_R := \cup_{r \in R} \mathcal{N}_r$ contains all the ignited/burned nodes among all replications R.
- $\mathcal{E}_R := \cup_{r \in R} \mathcal{E}_r$, the set of all edges traversed by fire during the R replications.
- The weights $w_{(i,j),R} := \left(t_{(i,j),R}, ROS_{(i,j),R}, f_{(i,j),R}\right)$ associated with any edge $(i,j) \in \mathcal{E}_R$ where:
  - $t_{(i,j),R} = \sum_{r \in R} \frac{\mathbf{1}_{(i,j) \in \mathcal{E}_r} t_{(i,j),r}}{|R|}$ the average traveling time between node i to j.
  - $ROS_{(i,j),R} = \sum_{r \in R} \frac{\mathbf{1}_{(i,j) \in \mathcal{E}_r} ROS_{(i,j),r}}{|R|}$ the average ROS when fire hits node j from i.
  - $f_{(i,j),R} = \sum_{r \in R} \mathbf{1}_{(i,j) \in \mathcal{E}_r}$, the number of times an edge (i,j) is traversed by fire among the $|R|$ replications.

Mathematical Model

Following the idea of evaluating the level of protection given by treating a certain area, we introduce the downstream protection value (DPV), a new fire risk model that aims to measure the value of what is burned downstream, starting from a burning cell i. Thanks to this metric, we can quantify the impact of treating cell i in terms of an optimization problem defined by the decision maker, minimizing the expected losses when planning under wildfire uncertainty.

Definition 4 Let $i \in \mathcal{N}$, and $\mathcal{T}_i = (\mathcal{N}_i, \mathcal{E}_i)$ its propagation tree with root node i. Then, the Downstream Protection Value, $DPV\,(i)$ is defined as

$$DPV\,(i) = \sum_{j \in \mathcal{N}_i} VaR_j$$

where VaR_j is an appropriate value at risk.

From the previous definition, DPV is the summation of Value-at-Risk (VaR) factors. For example, volume arises as a consistent and simple VaR variable. We can easily extend it and include more information by weighting values by probabilities and other characteristics of the nodes/edges. In addition, due to the additive property of the model, we notice its natural extensions to work with larger units (cluster of cells) and to calculate the total risk per cell from multiple replications R, using the global propagation tree $\mathcal{GT}_{\mathcal{R}}$ instead of $\mathcal{T}$. In order to illustrate potential applications and limitations depending on the selected VaR, we proceed to analyze and test different possibilities.

Suppose that we have three different "Propagation trees" (Figure 4.5) and we want to calculate the DPV using just the volume as our VaR. For simplicity, we assume identical cells (1 unit of volume). In this case, the DPV is exactly the same for cell 1 in all graphs, but their structures, traveling times, and ROS are significantly different, therefore, this VaR may not be appropriate for the situation.

We tested different extensions and modifications to our initial DPV: (1) Adding a decay factor $\alpha(j)$ depending on the depth inside the tree, (2) weighting the VaR by average traveling times $t(j)$ or ROS, and (3) multiplying the original summation by the degree $\mathrm{d}(i)$ of the node. In Table 4.2 and Figure 4.6, we can see how different values can be obtained depending on the approach used. Weighting the volume by the total degree $\mathrm{d}(i)$ of the node allows the DPV to better capture this information, increasing the risk value when cells tend to be more connected inside the network, and thus, tend to spread fire to more sections of the generated tree.

4.2.6 Optimal fuel-treatment plans

Formulating an explicit optimization model, we seek to solve the problem of finding the adjacent cells that maximize the total summation of the fire risk value considered, subject

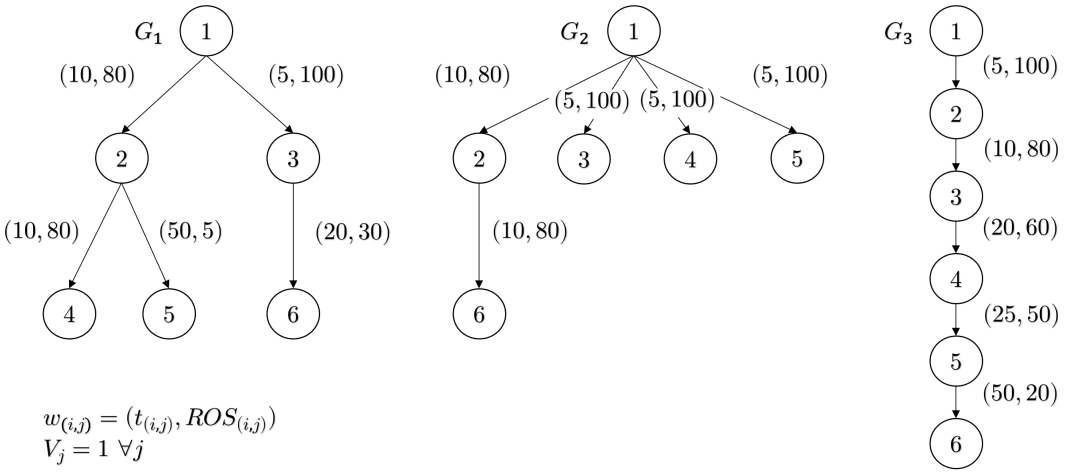
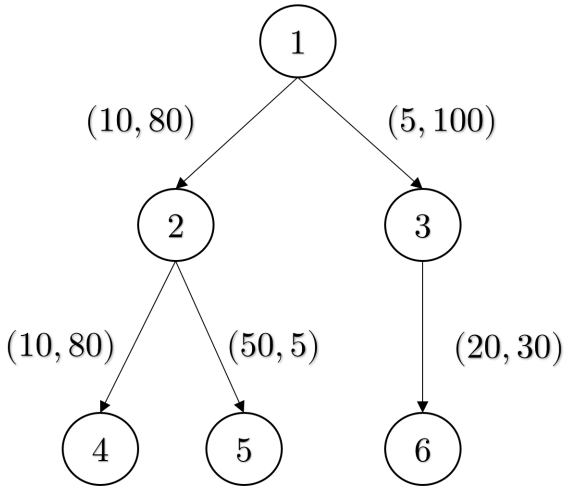


Figure 4.5: DPV comparison for three different "Propagation Tree" graph structures.

DPV formula	G_1	G_2	G_3
$\sum_{j \in N_i} V_j$	6	6	6
$\sum_{j \in N_i} \mathrm{d}(i) V_j$	12	24	6

Table 4.2: DPV comparison example.



DPV version	**Value**
$\sum_{j \in N_i} V_j$	6
$d(i) \sum_{j \in N_i} V_j$	12
$d(i) \sum_{j \in N_i} \alpha(j) V_j$	9
$d(i) \sum_{j \in N_i} t(j) V_j$	24.4
$d(i) \sum_{j \in N_i} \alpha(j) d(j) V_j$	13

Figure 4.6: DPV versions example. Several combinations of relevant variables of the graph can be used for calculating the Value-at-Risk of the landscape. Its components will depend on the objective and planning horizon of the study.

to specific constraints provided by the decision maker such as resources allocation (e.g., total budget), demand of forestry products, ecologic/demographic restrictions (e.g., protected areas), among others. Adjacency constraints are included with the purpose of obtaining more realistic and easily adaptable solutions by forestry experts in practice (Carvajal et al., 2013), where the fuel treatment is carried out such that units are adjacent to each other. Formally, we solve the following general optimization problem:

Definition 5 Given an undirected graph $G = (\mathcal{N}, \mathcal{E})$ and a weight function $\nu : \mathcal{N} \to \mathbb{R}_0^+$, we seek to identify the subset of nodes S where $\sum_{i \in S} \nu_i$ is maximized and the induced sub-graph G_S is connected. A feasible region $\mathbf{\Omega}$ is defined by a set of operational constraints (e.g., maximum budget, max capacity). Thus, we formulate the Protection Value Problem (PVP) as:

$$(PVP) \quad \begin{array}{ll} \max_{S \in \Lambda_{\mathcal{N}}} \sum_{i \in S} \nu_i & \\ s.t. & S \text{ connected} \\ & S \in \Omega \end{array} \tag{4.1}$$

where $\Lambda_{\mathcal{N}}$ is the set of all subsets of nodes in $\mathcal{N}$, and ν_i is some fire risk metric (BP, BC, FPV, DPV) for each cell $i \in \mathcal{N}$.

Therefore, for a given undirected and node-weighted graph $G = (\mathcal{N}, \mathcal{E})$, we want to find the subset of nodes with maximal sum of DPV (or another relevant metric), inducing a connected subgraph. This is a variant of the maximum-weight connected subgraph problem (MWCSP) a known NP-hard combinatorial problem (Álvarez-Miranda et al., 2013; Johnson, 1985), where extra constraints depending on the fire managers needs are included.

In this work, we solve this optimization problem using a greedy heuristic that selects which cells to harvest, starting from the one at the top of the ranking and adding the adjacent cells with the highest value until a stop condition (e.g., maximum number of treated cells) is satisfied. A detailed discussion of alternative techniques to solve this problem, such as Mixed Integer Programming (MIP) or meta-heuristics will be included in a future study.

4.3 Results and Discussion

4.3.1 Design of simulation experiments

The proposed Decision Support System (DSS) is divided into two main stages: simulation and optimization phases. In the first one, a set of R replications are performed, simulating multiple fires within the landscape(s) of interest in order to capture the uncertainty of the fire ignition and fire growth models as well as the weather stochasticity. Fire occurrence/probability maps, relevant statistics (e.g. average number of burned cells, shortest paths), useful visualizations, and the underlying graph structure of the fire are the main outputs of this stage (see Section 4.2). In the second phase, fires are processed in order to calculate the different risk metrics (or any variation proposed by the user) per cell. Then, an optimization problem is solved (see PVP problem in Eq. 4.1 and Figure 7.1), generating the optimal fuel treatment plan to minimize the impact of future fires given the conditions/constraints provided by the decision-maker.

Following this scheme, we perform a series of experiments using the data described in previous sections, assessing and comparing the performance of the proposed methods when measuring the fire risk of different sections of the forest. Multiple sources of uncertainty are included, aiming to generate robust fuel treatment plans to decrease the losses generated by future wildfires.

STAGE I

In this stage, the different outputs to calculate the fire risk values mentioned in section 4.2 are generated. For this, two modules are integrated: Fire Ignition Probability Map and Fire Growth Model (Cell2Fire) – see Figure 7.1. In the first module, a cell is selected for ignition. This probability map can be built simply by using a uniform distribution in the number of landscape cells, a probability function provided by the user, or through advanced methods of Machine Learning that capture the risk of occurrence based on variables that take into account the spatial distribution of forest fuel, distance to roads, demography and other anthropogenic characteristics that influence the risk of fire (see Chapter 6 in this thesis). In the selected regions, ignitions are usually caused by lightning. Then, we consider that the location is uniformly distributed. Once a cell is ignited, a weather scenario is randomly selected. Fire is generated in the second module, whose duration depends on the temporal information of the selected climate. Each fire spreads through the forest cells, and a directed graph G_D is produced, modeling the fire on the graph associated with the forest. By repeating this procedure R times, R graphs representing the R fires are generated, obtaining a $MultiDiGraph$ ($\mathcal{GT}_{\mathcal{R}}$). Using this information, different risk maps are produced using the metrics mentioned in section 4.2.

This offers an advantage over other approaches such as in Palma et al. (2007a) or Gray and Dickson (2016), who use FBP System and FlamMap outputs, respectively. Unlike our approach, where the ROS between two forest cells will depend on the evolution of the fire for a given simulation and a specific climatic scenario, these outputs are isolated representations of the potential fire behavior at each landscape grid cell and can be used to effectively compare *relative* rates of spread across the landscape, without actually simulating fire spread (Finney, 2006; Hirsch, 1996). In opposition, we can use several approaches to obtain a representative ROS between two cells, like the average of the ROS obtained from simulations where the fire went through such cells or other alternatives, providing more flexibility to our system.

STAGE II

The values obtained by the BP, BC, DPV from STAGE I or some of their variations are calculated for each cell/stand of the forest using the previous stage outputs. Now, our main hypothesis in this study is that managing the cells/stand with greater value (e.g. higher ignition probability) will play a fundamental role in the spread of fire, and therefore, on the protection of the forest in terms of expected losses from wildfires. Then, with the purpose of obtaining more realistic and easily adaptable solutions by forestry experts in practice, the fuel treatment of several cells/stands is carried out in such a way that units are adjacent to each other (a continuous patch of the original forest). These restrictions are known as "adjacency constraints" – see for example Carvajal et al. (2013).

Formulating an explicit optimization model, this module solves the PVP problem (Eq. 4.1) of finding the adjacent cells/stands that maximize the total summation of the fire risk value considered, subject to specific constraints provided by the decision-maker such as resources allocation (e.g. budget), demand of forestry products, ecologic/demographic restrictions (e.g.

protected areas), among others.

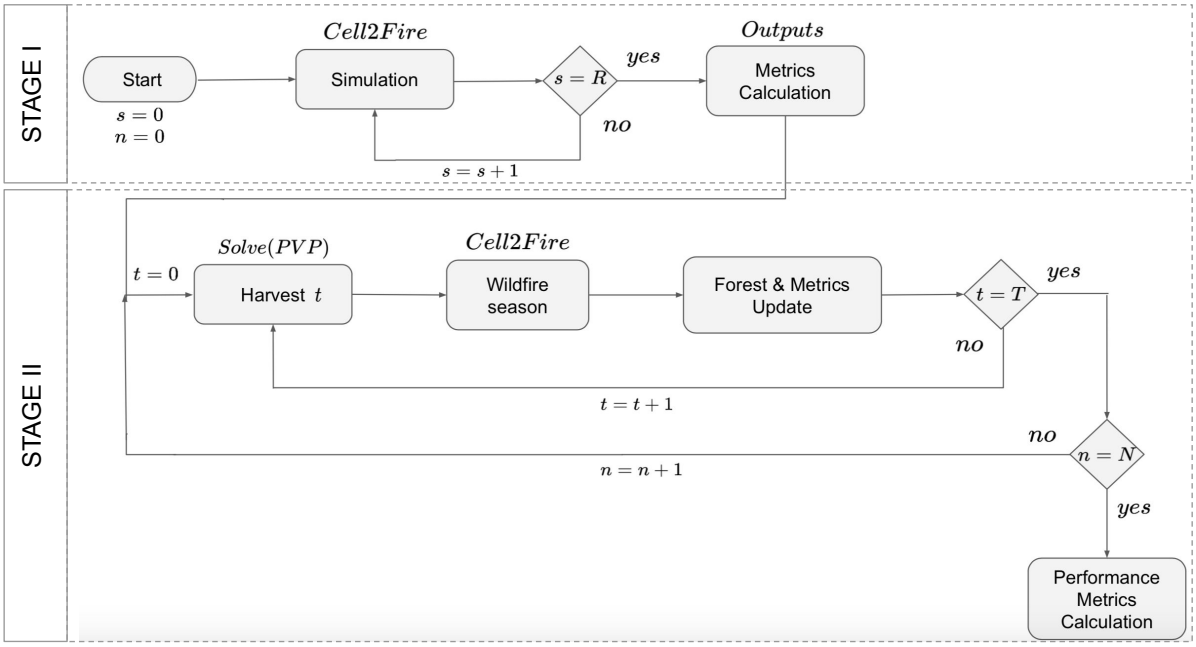


Figure 4.7: Decision support system schematic. STAGE I: in this stage, different outputs used to calculate the fire risk values are generated. For this calculation, two modules are integrated: fire ignition probability map and the fire growth model (Cell2Fire). A cell i is selected for ignition using a uniform distribution, a probability function provided by the user. Once a cell is ignited, a weather scenario is selected, and a wildfire is generated. Each fire spreads through the cells following a provided fire spread model, and a *propagation tree* $\mathcal{T}_i$ is obtained. After R replications, a multidigraph $\mathcal{GT}_{\mathcal{R}}$ is obtained. Using this information, different risk maps are produced using the discussed metrics. STAGE II: formulating an explicit optimization model (*PVP problem*), this module solves the problem.

4.3.2 One-stage framework to evaluate the effectiveness of fuel treatment

A total of 65,500 experiments, involving testing the different metrics with (i) heterogeneous landscapes of varying sizes – 79,611 to 559,746 cells – and land cover compositions, (ii) different numbers of replications $R \in \{10, 50, 100, 1000\}$, (iii) a random or deterministic selection of the weather scenarios and ignition point per replication, (iv) In the STAGE II, we consider a period for evaluation: $|T| = 1$ and (v) varying the total fraction tf representing the maximum area of the landscape to be treated given the decision maker's resources are conducted. Due to size limitations, fires are simulated for 12 hour windows in the Alberta province, while full day scenarios are used in the British Columbia instances.

For any given wildfire, using our model decreases the average rate of spread and expected

area burned by 53% and 57%, respectively, compared to its closest competitor (BC) thus creating effective fire-resistant terrains. These improvements are observed across all instances and weather scenarios, reaching differences above 60% of the total area protected with respect to the second-best metric when treating 5%, 10%, and 15% of the landscape. The results are useful for decision makers facing budget and operational limitations, where resources must be spent in the most economical manner. The DPV performs best in all experiments, rapidly extracting the core fire propagation patterns within the landscape. It identifies high risk regions, focusing the treatment in those zones, with its effectiveness elevated as the number of replications is increased. We measure this amelioration by noting the significant decrease in the number of wildfires and the expected losses – with differences up to 67% of the total area protected – when compared with the second-best metric (BC) across all treatment levels.

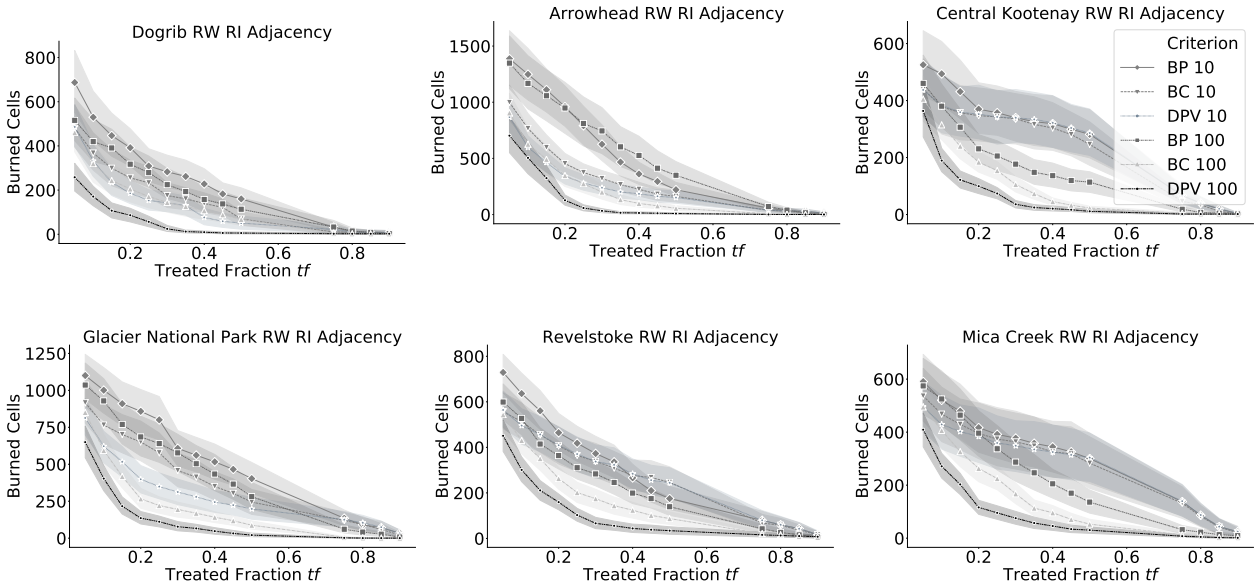


Figure 4.8: Adding new simulations from potential wildfires significantly improves the performance of the tested metrics under our framework, which is critical to generate effective treatment plans. Focusing on the Arrowhead instance $tf = 0.25$, we observe that using the DPV model, we decrease the average area burned from 281.1 ± 348.0 (DPV 10) to 58.4 ± 139.1 ha. (DPV 100). Here, we notice that even training the DPV model with 10 samples is almost as good as those of BC 100, showing the potential of our method. Minor performance differences can be seen between the BC 10 and DPV 10 models in the Revelstoke instance, indicating that the complexity of the landscape, more fragmented than the Arrowhead area, is relevant in determining the number of replications R to obtain robust results. Increasing R, DPV is able to significantly improve its performance, decreasing the number of burned hectares from 370.5 ± 382.2 to 102.6 ± 146.6, an improvement of 72.3%, learning faster than any other metric when more data are provided.

In general, we observe that increasing the number of replications positively influences the performance of all methods by decreasing (i) the volatility of the results and (ii) the total expected area burned after the application of the fuel-treatment plan. These effects are evident in the RW-RI scenarios, where uncertainty is a crucial part of simulated fires. DPV

not only surpasses the front line alternatives but it has the fastest learning capacity with an improvement up to 65% in the protected area when increasing R from 10 to 100.

However, this positive relation is not true for all metrics. This can be seen in the Arrowhead instance (Figure 4.8, top-center), where including 100 replications does not improve BP treatment plan's effectiveness, yielding a global average for $tf \leq 25\%$ of 1066.7 ± 1191.7 burned cells when training the metric with 100 replications (BP 100) versus 1099.6 ± 1090.3 when including only 10 training samples (BP 10). We omit the results of the FPV model since it is consistently outperformed by the DPV/BC both in the protection of the landscape and in the computational performance.

We can illustrate the previous discussion by observing the "Global Propagation Trees" $\mathcal{GT}_R$ obtained from a section of the Dogrib instance for a different number of replications R (Figure 4.9). From the plots, the number of potential trajectories captured when including more simulations is clear: $\mathcal{GT}_{10}$ contains limited information for all nodes in the graph, leading to classify several cells with null DPV (or alternative metric) since no observation include them into the wildfire propagation dynamic. On the other hand, the $\mathcal{GT}_{100}$ graph provides a representative and exhaustive picture of the potential fire propagation risk in different sections of the landscape, allowing the decision maker to easily identify which sections should be treated, as well as gathering relevant information regarding the expected ROS and traveling times of the fire within the forest.

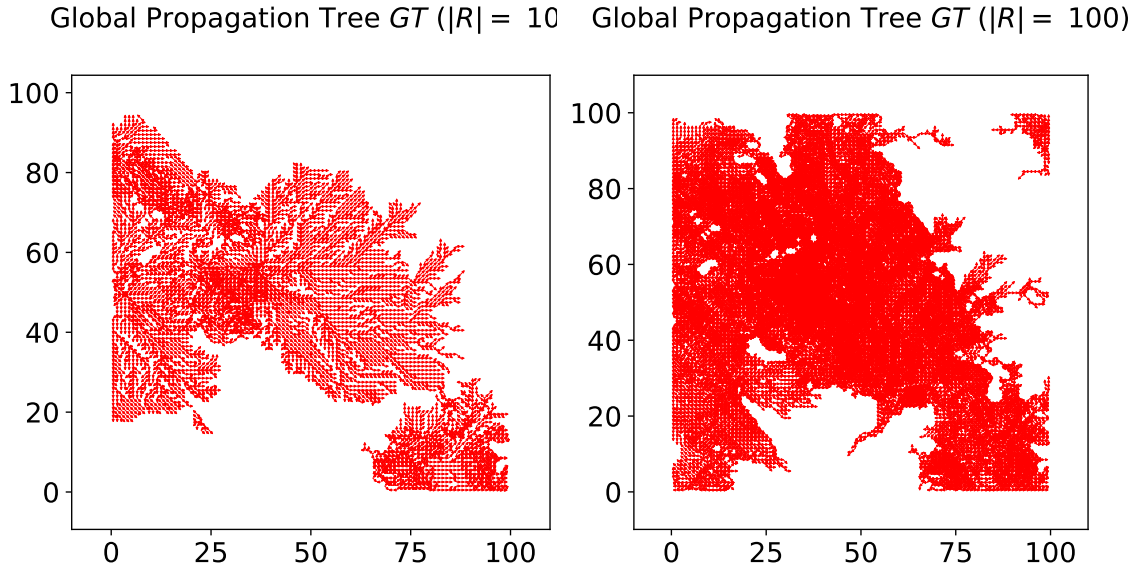


Figure 4.9: "Global Propagation Trees" for $|R| \in \{10, 100\}$ replications of the Dogrib RW-RI instance. More complex and general wildfire patterns are captured when increasing the number of replications, allowing the metrics to exploit this information when ranking the cells for the treatment plan.

Comparing the impact of the different sources of uncertainty included in the experiments (Figure 4.10), we observe similar trends: the DPV is able to converge faster toward a full protection of the landscape when fixing the ignition area (FI), preventing approximately 90%

of the fires, and generating the most prepared landscapes for addressing fires under weather and ignition (RW-RI) uncertainty across all tested instances.

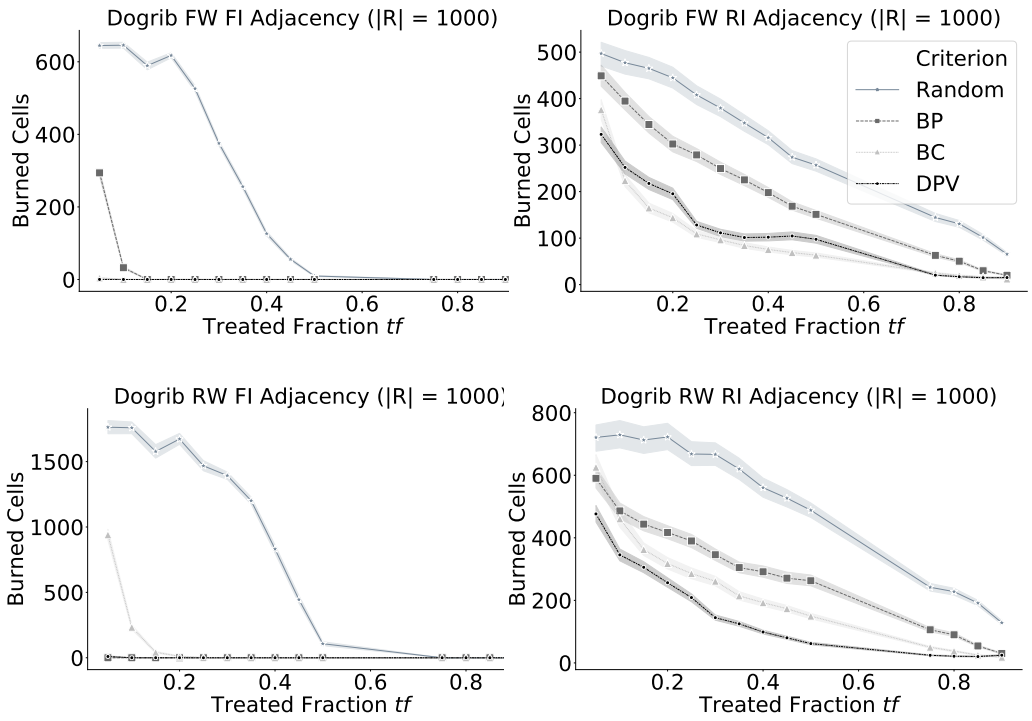


Figure 4.10: Treatment plan effectiveness depending on the simulated scenario for Dogrib instance. Fixing the ignition point to a bounded area (left side) significantly simplifies the problem, obtaining fast convergence toward no losses due to wildfire with all metrics.

Our DSS can be implemented in several practical applications. As an example, fire managers could capture the fire propagation patterns for a particular forest by replicating historical wildfires and their climatic conditions, obtaining a fully-adapted and trained fire risk model. Moreover, its data-driven approach allows decision makers to continuously update the parameters of our optimization model, allowing them to add new information from the landscape and incorporate it into the optimal fuel-treatment plan. Different subsets of the forest could have different values for the decision maker. Factors such as native forest areas, highly population-dense sectors, or regions naturally predisposed for hosting certain species should be accounted for by the the decision maker. These can be easily incorporated in our framework by modifying the VaR included in the DPV model as well as adding specific constraints to the PVP model. Thus, certain sections of the forest could have different ν_i weights when optimizing the plan. This difference is translated into multiple DPV heatmaps, and thus, different treatment plans are obtained, showing the flexibility of the framework.

A series of valuable outputs are obtained from the execution of the proposed methodology, allowing the decision maker to modify her relevant Values-at-Risk based on the information gathered from the experiments. For example: (1) ROS heatmaps projected on the "Propagation Trees" can be analyzed to determine which sections of the landscape tend to rapidly spread the fire under certain weather scenarios. Thus, helping the decision maker to identify

highly flammable and risky land covers within the forest as well as the most frequent fire propagation lines. (2) Shortest traveling times between all nodes inside the "Global Propagation Tree" could provide critical information when planning the suppression activities and resources allocation. Knowing the frequency and propagation patterns of the most likely potential fires within the landscape could be used to generate efficient prevention plans and to estimate the number of units of different suppression resources (e.g. helicopters, drones, firefighters) needed to act under several scenarios, in order to suppress a detected ignition. (3) Custom metrics based on the DPV model can be easily tested and adapted by the user. After developing a series of experiments for a landscape, the decision maker could customize the metric by adjusting or even proposing a completely new VaR, exploiting knowledge gathered from multiple replications and the performance achieved by the proposed metrics during the experiments. This way, more general and complex metrics – even including elements such as lightning strikes probabilities – can be easily tested in our open-source DSS, giving the user full flexibility and expanding the potential of the system.

The proposed model can take advantage of historical data to develop more efficient fuel-treatment plans with our DSS, obtaining landscapes that are better prepared for dealing with wildfires in comparison to state-of-the-art metrics. Overall, DPV emerges as the most effective option for minimizing large wildfires with significant room for improvement thanks to its flexible and expandable model. Fuel-treatment plans generated from our model are able to detect high-risk areas, focusing the treatment on the most critical sections of the forest. Therefore, the treatment plans generated result in optimized forest structures that minimize wildfire impact.

4.3.3 Multi-stage framework using a rolling horizon approach

We naturally extend our methodology to a multi-stage setting where we solve a series of iterative problems $t \in T$ with T the set containing all the periods up to the planning horizon. Using $|T| = 4$, we simulate the problem faced by a decision maker when deciding which portions of the landscape to treat in between fire seasons. For simplicity of the analysis, we assume that only one fire occurs per season. However, our DSS allows the user to include as many fires as wanted using a fixed number or following a historical distribution/density function. In addition, tf levels are calculated with respect to the remaining forest hectares after applying the fuel-treatment plan and passing the fire season, for every time-step t.

In this section, for our numerical results, we extract a patch of 10,000 hectares for testing the proposed methods (see Table 6.1) from Dogrib instance. This instance is referred as: Sub100.

Solving the multi-stage problem with $|T| = 4$ in the Sub100 instance, we observe (Figure 4.12) that the DPV model outperforms the alternative metrics for different tf levels. From the experiments, differences in performance tend to be even greater than in the single stage problem. This can be explained by the fact that treatment plans based on DPV are able to systematically modify the landscape, and transform it into a fire-resilient terrain since it consistently minimizes the expected losses due to wildfires, capturing the most frequent propagation patterns, their impact on the forest ecosystem, and translating this information

into a robust classification ranking for prioritizing the treatment of the cells. A future research study will be focused in the systematic evaluation of this challenging setting.

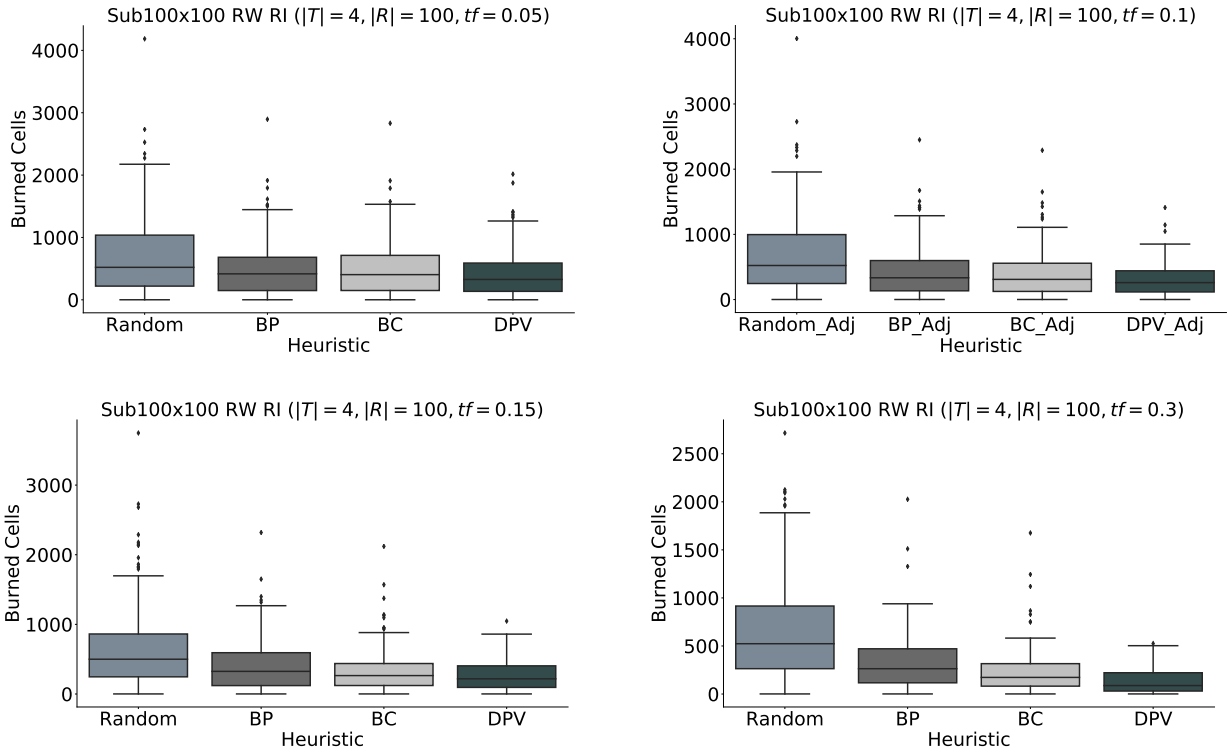


Figure 4.12: Multi-stage Sub100 instance algorithms performance comparison for multiple tf levels. DPV arises as the most effective and robust approach in the experiments developed.

This scheme can be naturally extended to any planning horizon T, even with the possibility of using different VaRs depending on the period t, providing great flexibility to the decision maker. Important is to notice that in this first study, we are using the information captured by the "Propagation Trees" obtained during the training stage to calculate the relevant VaR of the DPV. After every fire season, the original graph is updated with the new state of the forest and the decisions performed in the previous harvesting plan. Therefore, we do not perform new simulations between periods for updating the VaR inside the graph. Although the second approach could be more effective since the new graph is calculated with respect to simulated trajectories accounting for the new state of the landscape, it requires a significant amount of time to be evaluated since $|T| \times |R|$ simulations would be performed. This version and alternative/complementary approaches are being studied and will be covered in future studies.

4.4 Conclusions

The proposed methodology has been evaluated using realistic instances and exhaustive experimental settings. Results indicate that it significantly improves the performance obtained by traditional/alternative approaches when measuring the wildfire risk of different sections of

a landscape, allowing the decision planner to develop more effective fuel-management plans. It allows the decision maker to explicitly account and incorporate the wildfire risk as part of the decision making process. Using our proposed DSS, landscape managers will be able to include their relevant VaRs as part of the DPV calculations, obtaining quantitative information regarding the importance of each cell/stand inside the landscape to minimize the impact and losses associated with wildfires. Solving an optimization problem — including extra constraints and objectives as needed — we provide a flexible and robust methodology and an open-source DSS for performing a systematic, effective, and efficient tactical fuel-management.

The learning ability of the proposed DSS when increasing the size of the training set, as well as its flexibility and natural expansion/customization, make it a powerful system not only useful for landscape managers but also to a broader research community involved in topics such as animal conservation (e.g including connectivity corridors inside the landscape) under wildfire risk, modeling plague dynamics, and forestry industry in general, among other possibilities. Our work is the first open-source DSS for systematic fuel-treatment under wildfire risk, fully-integrated with a state-of-the-art fire growth simulator.

Several future research challenges remain open. In the current version, an individual VaR (DPV) is calculated so each cell is associated with a value. However, different configurations of harvested cells could lead to better/worse performance, even when the individual or total sum of DPVs is identical. Future research could explore the value of the interaction between cells in a simple but effective way: given their individual VaR, location in the graph, degree, and other characteristics; calculate the total DPV weighted by a factor capturing the interaction between the cells. An evaluation approach based on simulation could be applied, testing different combinations/shapes of cells and registering the most effective ones in order to understand the underlying patterns that make them more effective. However, the problem is highly combinatorial, exploding with the size of the instance, and thus, being extremely demanding from a computational perspective.

The current multi-stage approach is myopic, in the sense that decisions are taken based on the actual state, without including expectations about future periods of the planning horizon. This is repeated for each decision period. An alternative is to model the full tactical/strategical problem as a multi-stage stochastic problem. However, it triggers several challenges such as model size/complexity or scenario definition (content, number, associated probabilities). Therefore, it remains an open challenge.

Once we take decisions that affect the forest/landscape such as harvesting certain cells, the VaR of the full forest is modified. How can we update it without running new simulations is an interesting research question. In the current version, we eliminate the values of the already harvested cells and recalculate the remaining VaRs, however, that value was calculated after simulating fires with the original configuration of the forest, and thus, are not completely representative of the new situation. In order to tackle this limitation, we are currently exploring the use of supervised learning techniques for training a model as part of a future research study as follows: (1) Generate a large training dataset containing fire simulations of different classes of forests (homogeneous, heterogeneous, certain fuel types/weathers, etc.), (2) the relevant VaR is calculated using the exposed method, (3) using the instance data as

inputs, we train a deep learning model to estimate the VaR value of the instance – acting as a label for the sample, (4) once the model attains a good performance, we use it as a proxy for calculating the VaR of an instance "on-the-fly", updating the original values every new period. The idea is to increase the size of the dataset with every new fire simulated, improving the performance of the model through time.

Similarly, the action space will be extended in future research from a pure/binary harvesting setting to a flexible version including multiple management actions such as pruning, thinning, or application of different treatments (e.g. use of fire retardants) with the aim of developing an even more realistic decision model.

Chapter 5

Adjusting Rate of Spread Factors through Derivative-Free Optimization: A New Methodology to Improve the Performance of Forest Fire Simulators

Authors

Jaime Carrasco and Andrés Weintraub
University of Chile, Industrial Engineering Department

Cristobal Pais and Zuo-Jun Max Shen
University of California Berkeley, IEOR Department

Abstract

In practical applications, it is common that wildfire simulators do not correctly predict the evolution of the fire scar. Usually, this is caused due to multiple factors including inaccuracy in the input data such as land cover classification, moisture, improperly represented local winds, cumulative errors in the fire growth simulation model, high level of discontinuity/heterogeneity within the landscape, among many others. Therefore in practice, it is necessary to adjust the propagation of the fire to obtain better results, either to support suppression activities or to improve the performance of the simulator considering new default parameters for future events, best representing the current fire spread growth phenomenon. In this chapter, we address this problem through a new methodology using Derivative-Free Optimization (DFO) algorithms for adjusting the Rate of Spread (ROS) factors in a fire simulation growth model called Cell2Fire. To achieve this, we solve an error minimization optimization problem that captures the difference between the simulated and observed fire, which involves the evaluation of the simulator output in each iteration as part of a DFO framework, allowing us to find the best possible factors for each fuel present on the landscape. Numerical results for different objective functions are shown and discussed, including a performance comparison of alternative DFO algorithms.

5.1 Introduction

Considerable efforts have been made in recent decades to simulate fire through a heterogeneous forest landscape due to the increase of such events as a result of global warming and human carelessness (Running, 2006; Westerling et al., 2006; Westerling, 2016). To date, a wide set of fire growth simulators are available, which utilize a range of different modeling approaches and underlying fire behavior prediction systems to simulate the fire spread dynamics based on demographic, topographic, and environmental conditions. An excellent review of these can be found in Papadopoulos and Pavlidou (2011a).

One of these computational tools is Prometheus, a deterministic fire growth simulator released in May 2009 (Tymstra et al., 2010b). Using spatial fire behavior input data on topography (slope, aspect, and elevation) and the Canadian Forest Fire Behavior Prediction (FBP) System fuel types along with an hourly weather stream, it simulates fire growth based on the Huygens' principle of wave propagation. Another simulator with similar characteristics is FARSITE (Finney, 2005) based on the American BEHAVE System. As indicated in Papadopoulos and Pavlidou (2011a), these two models obtain the best simulations of historical fires: FARSITE in the United States and Prometheus in Canada.

Although Prometheus and FARSITE have excellent performance in general, they are not suitable for fire-smart forest management that requires that both the harvest and simulation models have a well-structured and natural interface for easily exchanging data at each iteration of the process, in order to incorporate decision-making modules. This becomes even more relevant in a multistage framework in which multiple events, whether harvest decisions and/or fires may occur each year for a given planning horizon. For this reason, a

new cellular-based fire growth simulator has been recently developed called Cell2Fire (Pais et al., 2019). Developed in C++, it was designed to run both on daily-user machines and High-Performance Computer (HPC) systems. In its current version, it incorporates fuel and fire models from the Canadian Forest Fire Behavior Prediction (FBP) System, allowing the user to simulate the fire dynamics across a grid that represents a real forest landscape.

However, the performance of any simulator and/or computer programs depends to a large extent on its input parameters, calculated as a function of the data provided by the user or following theoretical bases that support their values. This way, some performance problems may be a result of inaccurate data on fuel moistures, fuel descriptions, weather, or improperly represented local winds, among other sources of errors (Finney, 1998). As is indicated in Pais et al. (2019), the ability to build a realistic fire scar with Cell2Fire depends on the Rate of Spread (ROS), which causes the fire to advance faster or slower through the cells of the forest. A similar approach is followed by simulators such as FARSITE and Prometheus (Finney, 1998; Tymstra et al., 2010b). Modeling the landscape as a continuous surface, they base their propagation model through the forest on the assumption that under homogeneous conditions, the burned area has an elliptical shape whose deformation mainly depends on the wind speed and its axes are constructed from the main rate of spread values: Head ROS (HROS), Back ROS (BROS) and Flank ROS (FROS). In the case of FARSITE, these parameters are provided by BEHAVE, and obtained from the Canadian FBP System for Prometheus and Cell2Fire.

The noise of the data and different approximation inaccuracies directly affect the propagation model by perturbing the calculation of the different ROS values mentioned above, and thus, this can lead to an over/underestimation of the ROS magnitude, in disagreement with the observed fire. For these reasons, systems like FARSITE, BEHAVE, WildFire Analyst, and others (Rothermel and Rinehart, 1983; Finney, 1998; Ramírez et al., 2011), have introduced a module of adjustment factors directly on the ROS, allowing the user to use expert judgment or local data to tune the simulation parameters to observed or actual fire spread patterns. These factors are fuel model specific, multiplied by the rate of spread to achieve the specified adjustment (Finney, 1998; Ramírez et al., 2011). The latter is done manually in FARSITE, without any guarantee that the new factors configuration will improve the prediction. Moreover, the complexity of searching for factors increases along with the heterogeneity of the forest, directly depending on the number of fuels inside it. A recent article (Srivas et al., 2017) addresses this problem in FARSITE, introducing an automatic calculation using an Ensemble Kalman Filter that exploits the uncertainty information on the simulated fire perimeter, fuel adjustment factors, and measured fire perimeters. On the other hand, Cardil et al. (2019) proposes a method to determine the optimal adjustment factors by fuel model from the fire observed in real-time in order to minimize the arrival time error with respect to the simulated fire.

The main objective of this study is to show how a novel methodology based on Derivative-Free Optimization can serve to improve the performance of fire simulators in both real-time and future events simulations through an automatic adjustment of the Rate of Spread factors, obtaining more accurate fire scars. For this, we use Prometheus simulator as a proxy because it provides us the evolution of fire scars hour-by-hour with the purpose of imitating the behavior of a fire in real-time. Using this information, we minimize an objective function

that represents the error in the fire spread evolution, having the ROS factors as the decision variables of the optimization model. Also, in order to show the adjustment approach for future events, we analyze a case study in Canada where a landscape located in the Alberta Region suffered a large fire on September 25, 2001.

The chapter is organized as follows: In section 2, the main fire spread simulation logic of Cell2Fire is described and a brief description of Derivative-Free Optimization algorithms is given. Section 3 introduces the Derivative-Free optimization framework and the methodology approach to adjust the ROS factors in Cell2Fire, analyzing the fire spread and growth dynamics of each fuel type included in the FBP system, as well as describing the main algorithms implemented. In section 4, optimization results are discussed for a case study based on a real forest in Canada and a comparison of the performance with different sets of parameters is included for several (non)-homogeneous test instances. Finally, section 5 contains the conclusions and future work of the project.

5.2 Methods

5.2.1 Cell2Fire Simulator

Cell2Fire is a fire growth cell-based simulator developed in C++. It allows the user to simulate the fire dynamics inside a grid instance that can represent a real forest based on variables such as fuel type of each cell, weather, elevation (topographic/terrain components), ignition points, and all the aspects considered by the FBP System – see more detail in (Pais et al., 2019).

In algorithmic terms, Cell2Fire simulates the growth of fire by tracking the state of all cells as the model progresses through discrete time steps on a raster grid where a cell can be in one state *"Available"*, *"Burning"*, *"Burned"*, *"Harvested"*, or *"Non-Fuel"*. Fires begin with randomly occurring lightning strikes that will ignite a cell or an ignition point can be chosen deterministically to reproduce, for example, a historical event. At each time step, the fire may spread to neighboring cells based on calculations of the Head Rate of Spread (HROS), Flank Rate of Spread (FROS, identical to both sides) and Back Rate of Spread (BROS) obtained from the Canadian FBP System. Subsequently, a messaging signal process is initiated between the neighboring cells. In the current implementation, each cell has (at most) 8 neighbors. Fire progresses through each available axis of the current burning cell towards the center of the adjacent cells, starting a new fire when it reaches another cell's center. Thus, a higher rate of spread (ROS) entails that there is a greater chance of fire spreading to a neighboring cell.

As mentioned in Pais et al. (2019), Cell2Fire is based on an adaptation of the elliptical model proposed in O'Regan et al. (1973, 1976) using the main Rate of Spread components of the FBP System (HROS, FROS, and BROS). It considers the center of a cell as a focus of an ellipse and an adjustment of the ROS to the generated ellipse is performed as shown in Figure 5.1. The implemented ROS distribution scheme is as follows:

i. Parameterize the ellipse from one of its foci, using polar coordinates (r, ϕ) where r is the length of the radius vector and ϕ represents its angle with respect to the horizontal line (0°East).

ii. Set the main parameters of the ellipse (axes and eccentricity) at time t:

$$a = \frac{HROS + BROS}{2} \times t, \tag{5.1}$$

$$b = \frac{2 \times FROS}{2} \times t, \tag{5.2}$$

where $FROS = \dfrac{HROS + BROS}{2LB}$ and $LB = \dfrac{a}{b}$, the length-to-breadth ratio. Therefore, the eccentricity is calculated by:

$$\mathrm{e} = \sqrt{1 - \left(\frac{(FROS \times t)^2}{\frac{(HROS + BROS) \times t}{2}^2} \right)} \tag{5.3}$$

to expand the ellipse generated by the fire propagation at time t on the two main axes.

iii. Calculate the Rate of Spread r as a function of the angle ϕ:

$$r = \frac{a\,(1 - \mathrm{e}^2)}{(1 - \mathrm{e} \cdot cos\phi)}.$$

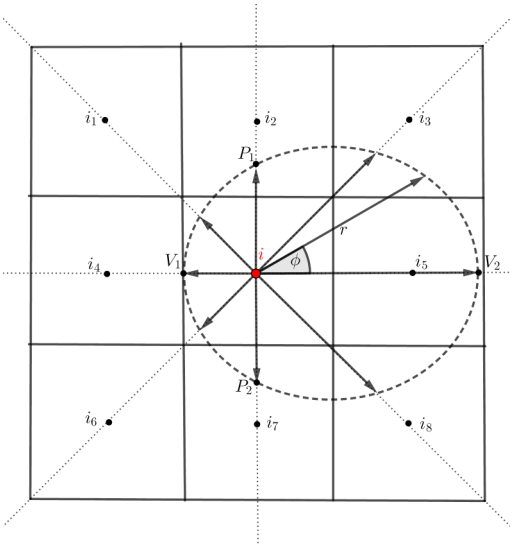


Figure 5.1: Elliptic approach for $WD = 0°$. ROS (r) is calculated for each angle ϕ.

Therefore, the full dynamic of the fire is determined by these three magnitudes: HROS, BROS, and FROS plus a fourth related to the eccentricity of the ellipse. Despite its practical usage, this elliptical model may have its limitations for practical situations as pointed out by Richards (1993): local wind fluctuations may decrease the LB value and then overestimate the heading spread of a fire at the expense of flanking spread. For this reason, Finney (1998)

considers pertinent to introduce some type of compensation through the use of spread rate adjustment factors. The above justifies the introduction of four ROS adjustment factors: x_1, x_2, x_3, x_4. These variables internally multiply the rate of spreads: $HROS$ (x_1), $FROS$ (x_2), and $BROS$ (x_3) obtained by Canadian FBP module; and a additional factor ECC (x_4) respective to adjust the ellipse eccentricity.

5.2.2 Derivative-Free Optimization

Derivative-free optimization (DFO) is an area of nonlinear optimization that deals with problems where the derivatives of the objective function (and potentially, constraints) are not available. Due to a growing number of applications in science and engineering, the development of DFO algorithms has increased and given greater attention in recent decades. Some applications using DFO algorithms can be found in Alarie et al. (2013); Alexandridis et al. (2008b); Begin et al. (2010); Hare (2010).

There are different situations where this methodology is appropriate: i) the functions defining the problem are provided through a computer simulation that cannot be easily subjected to automatic differentiation (see Figure 5.2); ii) the optimization problem involves conducting a laboratory experiment, with no explicit mathematical expressions; iii) the objective function is noisy and the gradient estimation may be completely useless; iv) when the evaluation of the functions require a significant amount of computational power, it may be prohibitive to perform the necessary number of function evaluations – normally no less than the number of variables plus one – to provide a single gradient estimation.

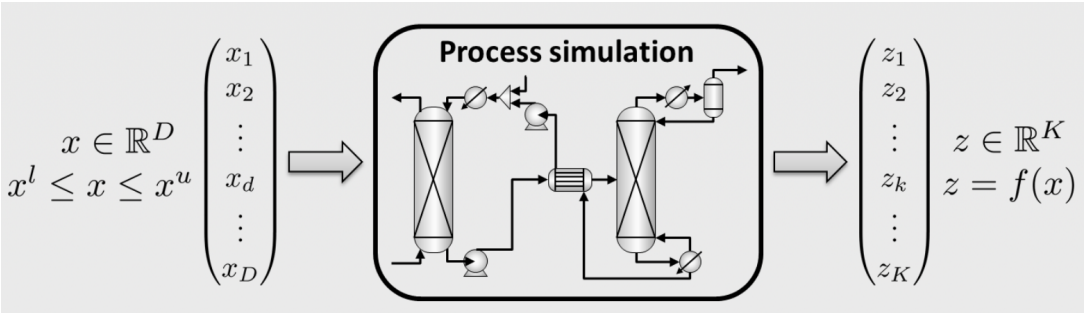


Figure 5.2: A "Black-Box System" assumes that the black box can be queried through a simulation or experimental measurements that provide a system output for specific system input values $x \in \mathbb{R}^D$. A principal challenge in practical optimization is how to optimize an objective function $z = f(x)$ that depends on this process in the absence of an algebraic model.

The diversity of applications includes problems in engineering, mathematics, physics, chemistry, economics, finance, medicine, transportation, computer science, business, and operations research (see e.g. Conn et al. (2009); Audet and Hare (2017)). Some examples of them are: Tuning of algorithmic parameters (Audet and Orban, 2006; Begin et al., 2010); Engineering design (Booker et al., 1998a,b); Molecular geometry (Alberto et al., 2004); Medical image registration (Oeuvray, 2005); and dynamic pricing (Levina et al., 2009).

In our research, we will follow the ideas of "Parameter Fit" presented in Audet and Orban (2006) and Alexandridis et al. (2008b). As we pointed out in the Introduction, most numerical codes (for simulation, optimization, estimation, etc) depend on a number of internal parameters. Researchers implementing numerical algorithms know how critical the choices of these parameters are and how much they influence the performance of solvers. Typically, these parameters are set to values that either have some mathematical justification or satisfactory empirical results. One way to automate the choice of the parameters — in order to find possibly optimal values — is to consider an optimization problem whose variables are the parameters and whose objective function measures the performance of the solver for a given set of parameters, measured by CPU time or by some other indicator such as the number of iterations taken by the solver (see Conn et al. (2009)). However, in our study, we are not interested in the CPU time or the number of iterations that Cell2Fire makes to get a more accurate fire scar. Since Cell2Fire simulated scars depend dynamically – in simulation time – on the fuel type (ROS obtained from the FBP System), our parameters to adjust/re-scale the fire spread model will be such that they change the magnitude of the ROS among the main propagation axes. This way, our main performance measurement will be the adjustment error with respect to a real/historical fire scar observed, a real-time fire scar provided to predict the most likely evolution of an on-going fire, or a scar simulated by an already calibrated software.

In Figure 5.3, we show a diagram of the internal dynamics of Cell2Fire as a function of the ROS factors $\vec{x} = (x_1, x_2, x_3, x_4)$. Note that for $\vec{x} = (1, 1, 1, 1)$, Cell2Fire behaves normally (by default). Let $Cell2Fire(\vec{x})$ be the "simulator function". Since we want to compare simulated and observed fire scars for different ROS factors, the $Cell2Fire(\vec{x})$ function maps the $\vec{x}$ vector to burn-grids $S_t(\vec{x})$, for all time steps $t = 1, ..., T$, where T is the total simulation time and $S_t(\vec{x})$ a 0-1 matrix with the same dimensions as the landscape where components $s^t_{(i,j)} = 1$ when a cell – located at coordinates (i, j) – was burned before time t and 0 otherwise (see Figure 5.4). Thus, $\sum_{i,j} s^{t_1}_{i,j}(\vec{x}) \leq \sum_{i,j} s^{t_2}_{i,j}(\vec{x})$ if $t_1 \leq t_2$. This function represents the core of the DFO framework and will be explained in detail in the next section.

Finally, we mention that, since the magnitude of the ROS depends on the characteristics of the type of forest fuel being burned at time t (Rothermel, 1972a; Van Wagner and Pickett, 1987), the adjustment factors may have a dependence by type of fuel (Rothermel and Rinehart, 1983; Finney, 1998). In order to address this problem, we naturally extend the initial framework by introducing a vector $\vec{x}_\omega = (x_{\omega 1}, x_{\omega 2}, x_{\omega 3}, x_{\omega 4})$ for each $\omega \in \Omega$, where Ω denotes the set of all fuel types existing within the landscape (see Section 5.3.2 for more details).

5.3 Rate of Spread adjustment factors

In this section, we describe the theoretical background of the DFO algorithms, their practical advantages/limitations, as well as how we can exploit their characteristics for modeling our fitting problem as a simple error-minimization problem. Theoretical and computational implementation details are discussed for the selected fitting strategies.

Based on the discussion from Section 5.2, the main tunable parameters present in Cell2Fire

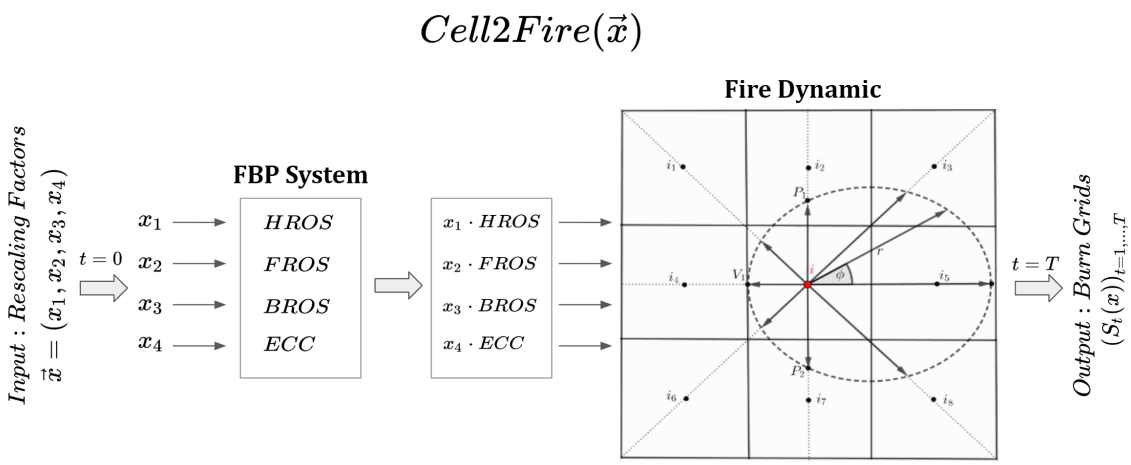


Figure 5.3: $Cell2Fire(\vec{x})$ Function. $S_t(\vec{x})$ is defined as a function that depends on the variables x_1, x_2, x_3 which modify the magnitude of the parameters HROS, FROS, and BROS obtained by the FBP System, and a fourth variable x_4 that adjusts the eccentricity of the ellipse, called ECC factor. Its outputs $S_1(\vec{x})$, $S_2(\vec{x})$,...,$S_T(\vec{x})$ are simulated fire scars for $t = 1, t = 2, ..., t = T$ hours using the parameters $\vec{x} \in \mathbb{R}^4$.

consist of the adjustment factors for the main ROS components (Head, Back, and Flank ROS) that define the shape and magnitude of the fitted-ellipses, associated with each burning cell. As mentioned in Pais et al. (2019), this step is critical to reproduce valid and realistic fire scars as well as represent the fire spread dynamic evolution with minimum error. Inappropriate ROS factors values can lead to poor results in both the evolution and final fire scar obtained, specially depending on the structure and characteristics of the forest: (1) different fuel types obtain different accuracy performance — associated with different ROS estimations, (2) heterogeneous landscapes tend to increase the structural differences between the wave-propagation and the cellular-automata models, and (3) extreme weather conditions significantly affect the behavior of the propagation dynamic due to the differences in the conditions under which fire models algorithms were developed (Duff et al., 2018).

Two main optimization approaches are tested in order to find the best ROS adjustment factors $\vec{x}$ and improve the performance of Cell2Fire with respect to an observed fire through the formulation of an optimization problem that minimizes the fire scar evolution error over time. The first one seeks to find four general factors, which we will denote by $\vec{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4_+$, independent of the type of fuel inside the landscape. Thus, each factor is an "average factor" that re-scales all ROS independent of the fuel type in order to improve the estimate of the observed scar. This method will be faster in computational terms, but less precise than a fuel-dependent model. The extension is covered in our second model, where we try to address the problem of factors adjustment empirically proposed in Finney (1998), however, we go a little further: we seek not only to adjust the corresponding factor to the HROS, but we also introduce factors for the BROS, FROS, and for the eccentricity of the elliptical model. This way, this is a fuel model specific approach where the factors depend on the type of fuel. Let Ω be the set of all fuel types, then $\vec{x}_\omega = (x_{1\omega}, ..., x_{4\omega})$ denotes the ROS factors associated with a specific fuel type $\omega \in \Omega$. Therefore, the new decision vector

in this model is $\vec{x} = (\vec{x}_\omega)_{\omega \in \Omega} \in \mathbb{R}_+^{4|\Omega|}$.

5.3.1 Global approach

The simplest tuning approach consists of finding the $\vec{x}^*$ vector that minimizes the global fire scar evolution error for a heterogeneous landscape. In this case, a unique $\vec{x}(\vec{\mu}) \in \mathbb{R}_+^4$ vector containing the relevant re-scaling parameters is optimized by solving the ROS Adjustment Factor *(RAF)* problem, using a set of desired weights $\vec{\mu} \in \mathbb{R}_+^{|T|}$ depending on the relevance of each period t accuracy for the research purpose: e.g. get the best accuracy during the first hours after the fire ignition or obtain the most likely final fire scar (see Figure 5.4 and Definition 6). This is equivalent to equally modify the shape and magnitude of the fire ellipses generated by each burning cell inside Cell2Fire.

Based on the previous discussion, we introduce the following notation and definitions:

- Let the subscript $t \in T$ be the simulation time measured in hours and T the discrete set containing the hours to simulate (simulation horizon). Our methodology allows other time intervals depending on the evolution of the observed fire. For simplicity, we use one hour in this study as the main time-step unit.
- A vector $x \in \mathbb{R}^4$ is denoted by $\vec{x}$ or $\boldsymbol{x}$, representing the four ROS adjustment factors.
- Let $S_t : \mathbb{R}^4 \to \{0,1\}^{m \times n}$ be a function whose outputs are the burn-grids obtained from the $Cell2Fire(\vec{x})$ "simulator function", dependent on the parameters $\vec{x} = (x_1, x_2, x_3, x_4)$, where x_1, x_2, x_3, and x_4 are the re-scaling factors that multiply the main rate of spread values for the ellipse-fitting procedure of Cell2Fire: *HROS*, *FROS*, and *BROS*, and an additional parameter for *ECC* factor, respectively (see section 5.3).
- The $S_t(\vec{x})$ function maps the four main ROS factors into a series of Burn-Grids, each one associated with a certain period t containing the simulated fire scar up to that hour. For a certain period t and a forest containing $m \times n$ cells, each Burn-Grid consists of a $m \times n$ binary matrix where entries containing $1s$ represent burned cells while the ones with $0s$ represent available cells. Therefore, the output of $S_t\,(1,1,1,1)$ is equal to the binary matrix obtained without any adjustment on the ROS values, i.e. using the direct output from the FBP system.
- Let Π_t be the Burn-Grid binary matrix generated by Prometheus after t hours of simulation or a historical wildfire scar. Similar to $S_t(\vec{x})$, this matrix represents the burned cells (with value 1) up to time t.
- Objective Function: we seek to minimize the hourly fire scar evolution error, ideally using observed fire data (when available). In our study, we use Prometheus' outputs as a proxy.

Using the above notations and observations, we define the following optimization problem which seeks to find the optimal ROS adjustment factors $\vec{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*)$.

Definition 6 Let $\varepsilon_T(\vec{x}, \vec{\mu})$ be the black-box error function for a simulation horizon T, $S_t(\vec{x})$ the simulator function at time-step t and a ROS adjustment factor vector $\vec{x}$, Π_t the expected binary matrix at period t, $||.||$ a matrix norm function (e.g. Frobenius norm), and

$\vec{\mu} = (\mu_1, ..., \mu_T)$ the vector of weights associated with each time-step. We define the **ROS adjustment factor (RAF) problem** for fitting the values of the $\vec{x}$ vector as:

$$(RAF) : \min_{\vec{x} \in \mathbb{R}^4} \varepsilon_T(\vec{x}, \vec{\mu}) := \sum_{t=1}^{T} \mu_t \left\| S_t(\vec{x}) - \Pi_t \right\|.$$

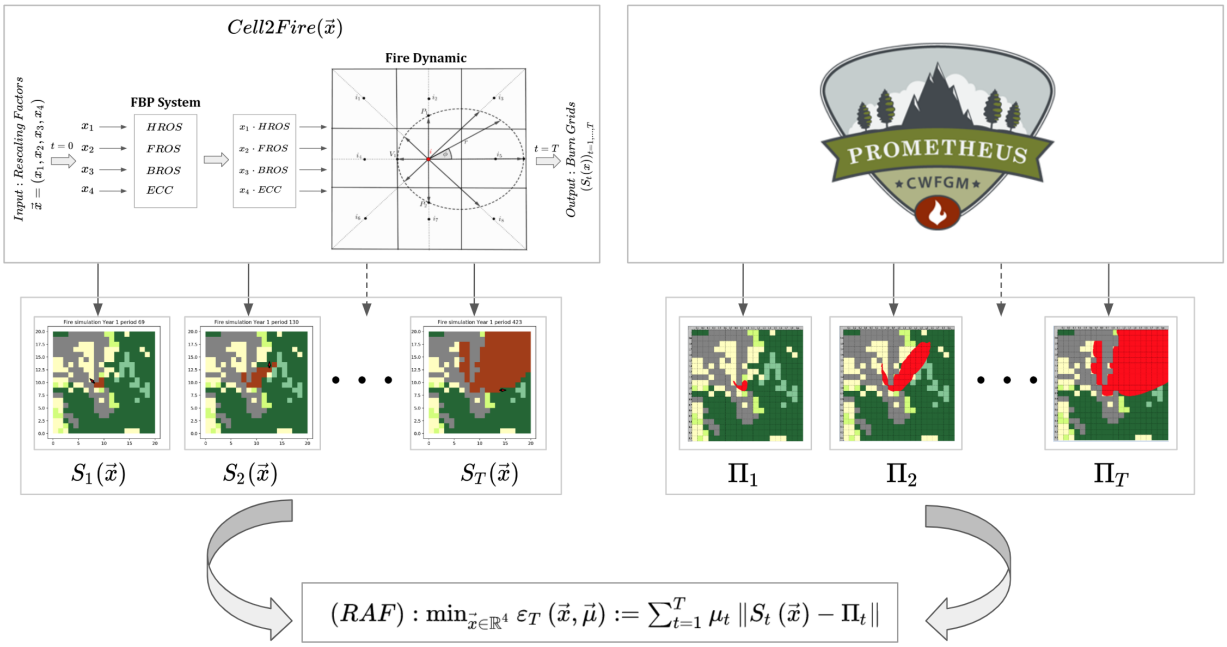


Figure 5.4: Derivative-Free Optimization framework: This figure represents the scheme of how the (RAF) problem is built. On the left side, Cell2Fire generates the fire scars $S_1(\vec{x})$,..., $S_T(\vec{x})$ which are dependent of the ROS adjustment factor vector $\vec{x}$ and time $t \in T$. On the right side, the observed/simulated fire scars Π_1,..., Π_T are provided. From this, the Black-Box function $\varepsilon_T(\vec{x}, \vec{\mu})$ is formulated as the weighted sum of the errors made using parameters $\vec{x}$ at each time t. Finally, the (RAF) problem aims to find those parameters that minimize the total error.

Observations:

- The term $\|S_t(\vec{x}) - \Pi_t\|$ measures the error between Cell2Fire and the provided fire scars at time t hours using the $\vec{x}$ ROS parameters. Important is to emphasize at this point that in case of having the evolution of the observed fire scar, it would not be necessary to use simulated ones (e.g. from Prometheus), which would improve the accuracy of the adjustment for practical cases.
- $\varepsilon_T : \mathbb{R}^4 \to \mathbb{R}_+$ is a function that measures the error between the simulated and observed fire scar evolution. In order to obtain it, Cell2Fire must be executed for the corresponding parameters $\vec{x} \in \mathbb{R}^4$.
- Different weight vectors $\vec{\mu}$ can be selected for different fitting objectives, e.g. minimize global hourly error or give specific emphasis on final scar differences, adding flexibility for analyzing different types of decisions – operational, tactical, and strategical.
- The RAF problem is an optimization problem that cannot be solved with conventional methods because we do not have any information about the derivatives or the algebraic structure of the function $\varepsilon_T(\vec{x}, \vec{\mu})$.

Despite the explicit computational advantage of this scheme with respect to the individual optimization of each fuel type factor, one of the main limitations of this approach consists of the fact that the optimal $\vec{x}^*$ vector is instance-dependent and thus, it is not likely useful for different forest compositions and fuel type distributions.

5.3.2 Fuel Model Specific (FMS) Approach

In this section, we denote the ROS adjustment factors vector by $\vec{x} = (x_{1\omega}, ..., x_{4\omega})_{\omega\in\Omega}$ where Ω denote the set of all fuel types, i.e. we suppose that the number of variables in the RAF problem increase to $4 \cdot |\Omega|$. We proceed to find the optimal vector $\vec{x}^*$ such that the combined fuel types dynamic is adjusted by these parameters, minimizing the error in the fire evolution of a specific heterogeneous landscape instance. Then, $\vec{x}^*$ is the solution of

$$(RAF_{FMS}) : \min_{\vec{x}\in\mathbb{R}_+^{4|\Omega|}} \varepsilon_T(\vec{x}, \vec{\mu}) := \sum_{t=1}^{T} \mu_t \|S_t(\vec{x}) - \Pi_t\|.$$

Thanks to this approach, we obtain the most accurate parameters for each fuel type, minimizing the simulation error of the internal fire spread model.

This approach requires more computational resources than the global approach presented in section 5.3.1 in order to obtain the optimal $\vec{x}^*(\boldsymbol{\mu})$ vectors since we are increasing the number of adjustable parameters by a factor of $|\Omega| - 1$. However, it arises as one of the most precise approaches for very heterogeneous forest during our experiments, as we will discuss when applying it to our case study (see Section 6.4).

5.3.3 DFO algorithms

In order to solve the (RAF) problem, we apply a series of powerful and easy-to-implement DFO algorithms following the techniques and recommendations from Conn et al. (2009) and Audet and Hare (2017). Based on the characteristics of our problem and the expected performance of the different algorithms (convergence to the global optimum is not guaranteed), we implement, test, and compare the following algorithms in order to find the optimal $\vec{x}$ parameters:

- **Nelder-Mead**: an algorithm introduced in Nelder and Mead (1965), it starts with a set of points that form a simplex – a generalization of the notion of a triangle or tetrahedron to arbitrary dimensions. On each iteration, the objective function values at the corner points of the simplex determine the worst corner point. The algorithm attempts to replace the worst point by introducing a new vertex in a way that results in a new simplex. Candidate replacement points are obtained by transforming the worst vertex through a number of operations around the centroid of the current simplex: reflection, expansion, inside, and outside contractions.
- **COBYLA** (Powell, 1994): was developed to solve non-linearly constrained optimization problems. This algorithm follows an approach similar to the DFO method (Conn and Toint, 1996; Conn et al., 1997), but it uses a linear model approximation for the objective function and constraints, interpolating at the vertices that form a simplex and where a trust-region bound restricts the variables perturbation. Thus, a new vector of variables is calculated which may replace one of the current vertices, either to improve the shape of the simplex or because it is the best vector that has been found so far according to a merit function that gives attention to the greatest constraint violation. The trust-region radius is never increased, and it is reduced when the approximations of a well-conditioned simplex fail to yield improvement to the variables until the radius reaches a prescribed tolerance value that controls the final accuracy.
- **NEWUOA**: is an unconstrained optimization method using a quadratic interpolation approximation. Like the DFO method, it seeks to calculate the least value of an objective function by applying the trust-region iteration for adjusting the variables. Now, as we mentioned above, all n-dimensional quadratic models have $(n+1)(n+2)/2$ parameters. This means that, unless other conditions are imposed, we require this number of interpolation points to build them. However, in NEWUOA this is an input parameter denoted by m. In Powell (2006), the author proposed to use a quadratic model relying on fewer than $(n+1)(n+2)/2$ interpolation points. The remaining degrees of freedom in the interpolation are determined by minimizing the change to the Hessian of the surrogate model between two consecutive iterations. The latter is an advantage since a DFO algorithm aims to use fewer evaluations of the objective function.
- **BOBYQA**: is an iterative algorithm for finding a minimum of an n-d*imensional* function subject to box-constraint. BOBYQA is a extension of NEWUOA, based on a quadratic interpolation approximation — see Powell (2009).

The development of derivative-free algorithms dates back to the works of Spendley et al. (1962) and Nelder and Mead (1965) with their simplex-based algorithms. An excellent review and numerical comparisons of state-of-the-art algorithms can be found in Papadopoulos and

Pavlidou (2011a).

5.4 Results and Discussion

In this section, we analyze and discuss the main results obtained following the two adjustment schemes described in Section 6.3. First, we report and compare the performance of Cell2Fire with and without tuning the $\vec{x}$ vector when dealing with our case study instance, but using the fire scars obtained from Prometheus hourly simulations for finding $\vec{x}^*$ both in a full information (after a wildfire) and partial information (real-time on-going wildfire) approaches. Subsequently, we will review how the adjustment behaves compared to the real historical fire scar. In this case, we do not have the evolution of the fire, so we only adjust the parameters using the final scar.

Finally, a detailed analysis and discussion of the simulation results obtained for Dogrib's 2002 fire in Canada when applying the described tuning approaches with different weight vectors $\vec{\mu}$ is performed and a benchmark of the DFO algorithms applied is shown.

All experiments have been conducted in a daily-use laptop with a 4th generation I7 CPU (1.9 GHz, 2 cores), 8 GB of RAM, and Ubuntu 14.0 OS.

5.4.1 ROS adjustment factors via Prometheus proxy

In these experiments, the *RAF* problem is solved with full information of the wildfire of interest in order to adjust the $\vec{x}$ vector. This way, the full evolution of the wildfire is provided with the aim of adjusting the simulator parameters with historical data. Notice that the optimization framework can be easily extended to include multiple historical fires by modifying the objective function in order to obtain a representative set of optimal parameters $\vec{x}^*$ for a certain region/area, training and evaluating the model performance with both a training and testing set of fires to avoid over-fitting, following a machine learning model scheme.

Global approach results: Dogrib-North

Starting with $\vec{x}_0 = (1, 1, 1, 1)$, we use the fire scars generated with Prometheus simulator using the Dogrib instance with the same weather conditions as in the real fire but using a constant wind direction (North) for visualization purposes, recording the hourly fire scars obtained 7 hours after its ignition (full information). We set $\mu_i = 1/7, i = 1, ..., 7$ to represent the average error between the observed scars and the ones simulated by Cell2Fire. We solve the *RAF* problem using the algorithms presented in section 5.3.3, focusing the discussion on the results obtained using the BOBYQA algorithm. Later, we will perform a detailed comparison with the other algorithms mentioned in the previous section.

Because Cell2Fire and Prometheus both rely on the FBP System (Pais et al., 2019), the

initial error value $\varepsilon_{T=7}(\vec{x}_0, \vec{\mu})$ is very low (approx. 36.86 using the euclidean norm). This represents the average difference in the number of cells after 7 hours of fire evolution. Since the total burned area is 3,777 cells (or hectares), then, the percentage difference is 0.97%. Subsequently, after 107 evaluations of the objective function, a minimum error value of 29.10 (0.77%) is reached. In this case, the best algorithm (BOBYQA) converged after 9 minutes.

The optimal values are very close to the starting point $\vec{x}_0$, which was to be expected since both Cell2Fire and Prometheus obtain the ROS values from the FBP System. The best parameters $\vec{x}^*$ are: $HFactor = 1.41$, $FFactor = 1.16$, $BFactor = 1.85$, $EFactor = 1.16$. Similar results in both computational time and accuracy are obtained for multiple wind directions (8 main axes plus the original weather of the case study).

FMS approach results: Dogrib-North

In this section, we address the RAF_{FMS} problem with the Fuel Model Specific (FMS) approach. Following the discussion of section 6.3, each fuel model $\omega \in \Omega$ of the Dogrib instance corresponds to a ROS factor 4-tuple. This way, since the Dogrib forest contains 8 different fuel types, the RAF_{FMS} problem has 32 variables, in comparison to the 4 variables of the global approach. The optimal factors are shown in Table 6.1 and Figure 5.5 shows the fire scar generated by Prometheus (a), Cell2Fire without adjustment (b), the global (c), and the FMS (d) tuning approaches. Graphically, figures (a), (c), and (d) are very similar, however, without adjustment (b) the fire in the last hour of simulation lacks intensity (strength). This occurs just as the fire crosses the river and is later found in a zone of fuel type C-1. To this respect, important is to mention that Prometheus includes both breaching and spotting phenomena when simulating wildfires while Cell2Fire does not include them in its first version. Therefore, the FMS approach determines that an increase in the magnitude of the C-1 *HFactor* from 1 to 1.495 is a better approximation (strategy) to decrease the error between both fire scars.

Fuel	**HFactor** (x_1^*)	**FFactor** (x_2^*)	**BFactor** (x_3^*)	**EFactor** (x_4^*)
C-1	1.495	0.786	0.995	1.424
C-2	1.281	1.064	1.047	1.238
C-3	1.287	1.324	1.000	1.722
C-4	1.002	1.000	0.998	1.002
C-5	0.998	1.003	1.001	1.001
D-1	0.998	1.010	0.992	0.998
O-1a	1.219	1.055	0.995	1.125
M-1	1.445	1.449	1.004	1.560

Table 5.1: Optimized *ROS* adjustment factors (*HROS, FROS, BROS,* and *ECC*) for each fuel type available in the Dogrib instance using Prometheus proxy and BOBYQA algorithm.

In contrast, we observe that parameters associated with certain fuels like C-4, C-5, and D-1 do not suffer significant modifications — they are almost identical to the default parameters $\vec{x} = (1, 1, 1, 1)$. This pattern is due to the fact that Cell2Fire is able to capture

similar propagation patterns as the ones modeled by the wave-front propagation approach of Prometheus, as mentioned in Pais et al. (2019).

Lastly, we observe that the performance in terms of accuracy is superior to the global approach — as expected — since each fuel has its own adjustment factors adding more degrees of freedom to the spread model. The minimum value reached using BOBYQA as the main algorithm for this approach is 25.55, i.e. an error of 0.67%. However, in this case, the number of evaluations of the objective function is 392, which increases the computation time by a factor of 4.

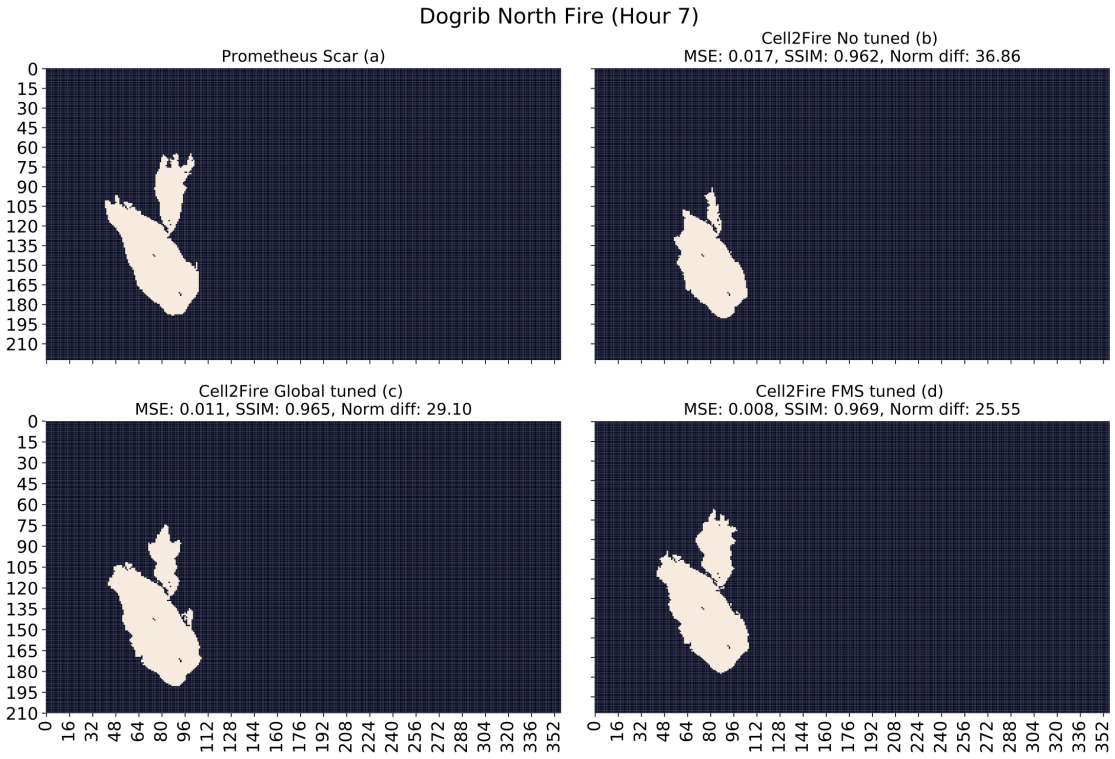


Figure 5.5: Dogrib fire scar evolution comparison using a constant wind (North) generated by Prometheus (a), Cell2Fire without adjustment (b), Cell2Fire using global tuning (c), and Cell2Fire using a FMS approach (d).

5.4.2 Real-Time adjustment

In this second set of experiments, only partial information is provided when solving the *RAF* problem. Assuming the presence of an on-going fire, we aim to train and evaluate the performance of the framework in an iterative process where the optimal parameters are updated with new information (fire scars), collected every time-step t. This way, we (1) train the model with the information available up to time t obtaining $\vec{x}_t^*$, (2) simulate the fire scar for $t+1$ using $\vec{x}_t^*$, (3) compare its performance with the observed data at $t+1$, and (4) obtain the $\vec{x}_{t+1}^*$ using $\vec{x}_t^*$ as the starting point for the optimization procedure. This process is repeated until a convergence criterion is achieved or no more data is provided.

This setting is significantly relevant in practical applications. As an example, firefighters and decision-makers need to establish their strategy in order to contain an on-going fire. Having an automatic system that learns from the currently available data will improve and support the simulator outputs, becoming a fundamental input for their action plan as more accurate predictions are achieved.

Global approach results

As in the previous section, we start from the default parameters $\vec{x}_0 = (1, 1, 1, 1)$, but in this case, we use the 7 hours fire scars generated with Prometheus simulator for the Dogrib North instance to update the optimal parameters every time-step t – in an iterative process – using the current optimal parameters as a starting point when solving the RAF problem. This way, we initially update the parameters $\vec{x}_0$ providing the first-hour fire scar Π_1. Once the optimal vector $\vec{x}_1^*$ is obtained, we proceed to the next hour ($t = 2$) providing Π_2 and starting the optimization procedure using $\vec{x}_1^*$, and so on. Therefore, for adjusting the parameters at time t, we use the optimal vector $\vec{x}_{t-1}$ as the starting point when solving the optimization problem, including the last observed scar Π_t.

At time t, we set $\mu_t = 1/t, i = 1, ..., t$ to represent the average error between the observed scars by time t and the ones simulated by Cell2Fire. Again, we focus our analysis using the BOBYQA algorithm.

Hour	HFactor	FFactor	BFactor	EFactor	Initial Error	Final Error
1	0.94	1.03	0.99	0.62	8.18	3.87
2	0.90	0.89	0.99	1.01	13.45	7.74
3	0.99	0.99	0.13	1.12	16.03	12.24
4	1.20	1.00	0.12	1.25	19.08	13.60
5	1.20	1.00	0.12	1.26	23.62	23.60
6	1.46	1.04	0.19	1.39	28.21	25.90
7	1.66	1.02	0.35	1.49	27.69	26.77
AVG	1.19	1.00	0.41	1.16	19.47	16.25

Table 5.2: Global approach results for the real-time adjustment experiment. All factors evolution and a comparison between the initial and final objective value (error) per hour are included.

From the results presented in Table 5.2, we observe that the most significant adjustment occurs in the BFactor, in special after the second hour of the ignition time. The perturbation along the larger axis of the ellipse is boosted by the increasing pattern on the HFactor parameter, allowing the fire growth model to reach and surpass the river at the north of the ignition area. This is consistent with the results obtained in the experiments discussed in the previous sections. In addition, we observe how both the initial and final error values reduce their gap (maximum of 52.69% at hour 1, minimum of 0.08% at hour 5) when increasing the number of simulated hours, indicating that the performance of the $\vec{x}$ vectors starts to converge. The total run-time needed for solving the whole set of scars was 41.6 [min] – an

average of 6 minutes per time-step tuning – requiring an average of 74 evaluations of the objective function per $t \in T$ to reach convergence (BOBYQA).

In order to account for potential estimation errors in the input data of the model, several replications with multiple weather scenarios and other potential uncertainty sources (e.g. ignition point) should be performed. This way, confidence intervals could be calculated for each parameter of the optimization problem, allowing the researcher to capture and assess the risk of different potential outcomes.

FMS approach results

Following the procedure described above, we solve the real-time RAF_{FMS} problem, updating the optimal values of the previous time step vector $\vec{x}^*_{t-1}$ once a new scar is available. The average and standard deviation of the optimal factors obtained during the simulation of the 7 hours are shown in Table 5.3 and both Prometheus and Cell2Fire final scars can be seen in Figure 5.6, showing the high-precision of the adjustment method. Graphically, both figures are very similar, reaching an objective value of 23.37, a Mean Squared Error (MSE) of 0.7% and a very high Structural Similarity Index (SSIM) of 97.4%.

Looking at Table 5.3, we observe again that the most significant changes occur on C-1 (AVG $HFactor = 1.77 \pm 0.87$, AVG $FFactor = 1.56 \pm 0.55$) and O1-a (AVG $HFactor = 0.69 \pm 0.47$) fuel types, being consistent with the results of the previous sections and showing how the model is able to balance the effect of a higher HFactor on one fuel by decreasing it on another (C-1 vs O1-a). In addition, an average running time of 17 minutes and 340 evaluations per hour simulated are needed by BOBYQA to reach convergence.

Fuel Type	**HFactor**		**FFactor**		**BFactor**		**EFactor**	
	AVG	**STD**	**AVG**	**STD**	**AVG**	**STD**	**AVG**	**STD**
C-1	1.77	0.87	1.56	0.55	1.05	0.02	0.90	0.16
C-2	0.98	0.35	0.81	0.06	1.21	0.24	1.40	0.28
C-3	1.37	0.29	1.15	0.29	0.98	0.02	1.20	0.22
C-4	1.00	0.01	1.02	0.02	1.01	0.03	1.01	0.01
C-7	1.00	0.01	1.00	0.01	0.98	0.04	1.00	0.01
D-1	1.03	0.06	1.00	0.05	1.00	0.01	1.05	0.15
O-1a	0.69	0.47	0.99	0.09	0.99	0.01	1.29	0.68
M-1	1.02	0.04	1.00	0.01	1.00	0.01	1.03	0.06

Table 5.3: Summary results for real-time Dogrib's adjustment. Average and standard deviation for all factors and fuel types obtained from the 7 hours simulation are shown.

Previous results are complemented by Figure 5.7, where the initial and final per-hour errors can be seen. Notice that since the RAF_{FMS} problem has multiple local minima, we can observe situations like in the fourth-hour optimization, where the optimal vector $\vec{x}^*_4$ reaches a better (lower) objective value than $\vec{x}^*_3$. In addition, we observe how the method

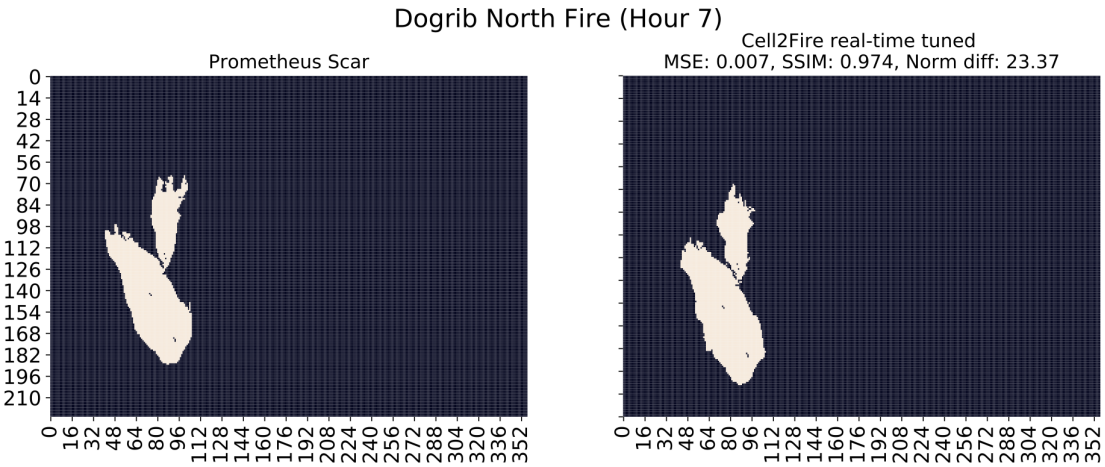


Figure 5.6: Comparison of Dogrib North's final scars obtained from Prometheus and Cell2Fire real-time FMS adjusted.

reaches a steady-state in the last 3 hours, being able to stabilize the final error achieved as well as decreasing the initial error towards this value, indicating that the succession of $\vec{x}_t^*$ vectors start to converge.

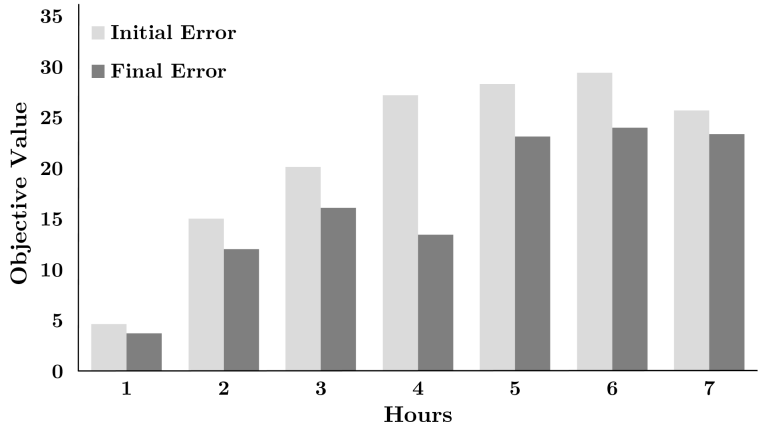


Figure 5.7: Hourly error evolution using the optimal vector $\vec{x}_{t-1}^*$ as the starting point for the time-step t during the real-time tuning FMS approach. Both initial and final errors show a convergence pattern by the end of hour 7.

5.4.3 Case Study: Dogrib Fire

The Dogrib fire (Tymstra et al., 2010b) started on September 25, 2001, in the Rocky Mountain front ranges of southwest Alberta. The fire was detected late in the afternoon on September 29 and assessed early the next day at 70 ha. in size. Firefighting started early on October 1. The fire was 828 ha. and out of control on October 15. A wind event resulted in a major fire run on October 16. Local terrain funneled wind flow along the Red Deer River and through a gap in the surrounding mountains. This pushed the fire east along the river valley. The fire jumped the Red Deer River and a road and then resumed a northeast spread direction.

The final fire size was 10,216 ha. The October 16 fire run accounted for 90% of the total area burned and resulted in high to very high burn severities.

We chose to model this particular fire due to the large documentation and real data availability — weather conditions recorded from The Yaha Tinda Automatic station, and demographic/topographic data collected from the area — as well as containing a representative set of different fuel-types documented in the Canadian FBP system. Divided into 79,611 100×100 $[m^2]$ cells, we replicate the original fire's ignition point located at (51.652876°, -115.477908°) starting the fire spreading dynamic on October 16, 2001, 13:00 hrs. For this, the ignition point is translated into an ignition area (cell) in Cell2Fire containing its coordinates.

Global approach results

Starting with $\vec{x}_0 = (1, 1, 1, 1)$ — default parameters — we use the real final fire scar from the Dogrib instance obtained 7 hours after its ignition. Therefore, we set $\{\mu_i\}_{i=1}^{6} = 0$ and $\mu_7 = 1$. We solve the *RAF* problem using BOBYQA as the main algorithm and compare its performance with COBYLA, NEWUOA, and NELDER-MEAD in section 5.4.4. From the experiments performed, convergence is achieved after 96 evaluations (467.78 [s] = 7.79 [min]) decreasing the objective value from 96.98 to 81.94 (15.5% improvement), obtaining an MSE = 7.9% and an SSIM = 0.818. Notice that since we are optimizing a non-convex function, the solution can converge to multiple local optimum values, and thus, it is recommended to perform a series of re-optimizations starting from the best $\vec{x}$ (incumbent) obtained or simply starting from different initial points $\vec{x}$ to explore more potential solutions. This is the inherent exploration versus exploitation trade-off, a challenge that is out of the scope of this article and will be covered in future research.

FMS approach results

Below, we show the results obtained from the application of the BOBYQA algorithm in our FMS scheme. Again, we use the starting point $\vec{x_0} = (1, 1, 1, 1)$ and we obtained relevant results in both computation time and final precision with respect to the observed fire. The algorithm converges after 37.58 minutes and the objective function reaches a minimum value equal to 77.53 — an improvement of 5% with respect to the global approach. The optimal values obtained for the ROS adjustment factor are shown in Table 6.4 and a comparison between the real and simulated scars can be seen in Figure 5.8.

Analyzing the results in Table 6.4, it is interesting to notice that certain fuels are significantly affected by the optimal adjustable parameters of $\vec{x}^*$. For example, both the HFactor and EFactor of the O-1a fuel type are significantly modified (1 to 2.066 and 1 to 10.82, respectively) in order to capture its spread pattern based on the real fire scar. Similarly, C-1, C-2, and C-3 fuels are also impacted, mainly in their HFactor, indicating that the default parameters obtained from the FBP system — given the instance data provided — are not able to capture the strength of the fire in the main wind direction, and thus, the optimal vector $\vec{x}^*$ tries to adjust the shape of the ellipses, expanding them into the north-east direction.

Fuel	**HFactor** (x_1^*)	**FFactor** (x_2^*)	**BFactor** (x_3^*)	**EFactor** (x_4^*)
C-1	1.623	1.243	1.023	2.479
C-2	2.462	0.955	1.388	2.425
C-3	2.289	0.996	1.188	1.93
C-4	1.158	1.138	0.287	4.91
C-5	1.018	1.068	1.05	3.17
D-1	0.263	2.569	0.825	1.625
O-1a	2.066	1.675	0.269	10.82
M-1	0.987	0.962	0.95	1.004

Table 5.4: Optimized *ROS* adjustment factors (*HROS, FROS, BROS, and ECC)* for each fuel type available in the Dogrib instance from the Canadian FBP system. The tuning has been performed using the real fire scar of the Dogrib instance.

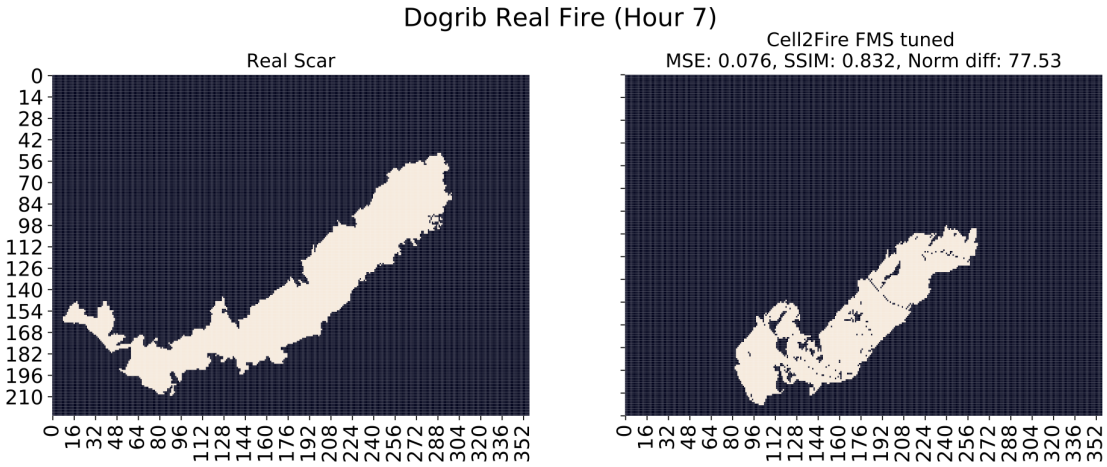


Figure 5.8: Dogrib real fire scar (left) and Cell2Fire simulated scar using the $\vec{x}^*$ obtained from the RAF_{FMS} problem. A Mean Squared Error value of 0.076 and a Structural Similarity Index value of 0.832 are achieved. The current fire spread and growth model is able to capture the expanding pattern of the real fire in the main direction (north east). However, more detailed information (such as the effect of local wind conditions) is needed in order to capture the back-propagation pattern of the original wildfire.

From the results, we are able to identify significant differences — as expected — when using Prometheus or real fire scars. As mentioned, both Prometheus and Cell2Fire use the FBP System as their main fire spread model to determine the relevant ROS values, and differences between both approaches are mainly associated with their fire growth models logics: wave-propagation and cellular automata. However, the initial error with respect to a real fire scar is significantly higher and requires a deeper level of adjustment (more computational time) to find sets of parameters $\vec{x}$ that are able to better capture the real fire propagation dynamics. However, we also remember the fact that historical scars tend to include the effect of suppression activities that are not modeled or included in any of the mentioned simulators, and thus, it plays against the accuracy of the propagation model. In order to account for this bias, suppression efforts should be explicitly modeled by (1) including it as part of the simulation tool or (2) by modifying the data structure of the instance to reflect the

position/effect of different activities such as the use of retardant or the location of firewalls, among others.

5.4.4 Benchmarking DFO Algorithms

In this section, we show and discuss the performance of the implemented DFO algorithms, namely BOBYQA, NEWUOA, NELDER-MEAD, and COBYLA. These algorithms have been widely used in different applications – a review can be found in Papadopoulos and Pavlidou (2011a) and Rios and Sahinidis (2013). However, to our knowledge, few comparisons in real problems have been reported and even less in ROS adjustment factors in the field of forest fire simulation.

Our methodology consists of solving the RAF problem for the two approaches, Global and FMS, and record the number of evaluations ($NEVAL$) of the objective function until a specified tolerance (error) is reached, which we will denote by $xtol\_abs$. Then, $xtol\_abs$ is a stopping criterion of the algorithm. Δx is usually a measure of how much x changes by from one iteration to the next, or the diameter of a search region. This way, the algorithm stops when $|\Delta x| < xtol\_abs$ or a maximum running time (1 hour Global, 3 hours FMS) is reached. This last notation comes from the NLOPT package – see (Johnson, 2014) – a free/open-source library for nonlinear optimization, providing a common interface for a number of different free optimization routines available online as well as original implementations of various other algorithms, in particular: BOBYQA, NEWUOA, COBYLA, and NELDER-MEAD.

As pointed out in Conn et al. (2009); Audet and Hare (2017), in the Derivative-Free Optimization context, CPU time is irrelevant since we have the number of evaluations as a perfect machine-independent criterion. However, with the propose of supporting decision making in fire fighting, which is usually carried out in real-time, we want this methodology to respond to us as quickly as possible. Therefore, we use two performance indicators: i) the number of Black-Box evaluations ($NEVAL$) and ii) run-time ($RUNTIME$) measured in minutes [min]. In all cases, we execute the algorithms taken as starting point $\vec{x_0} = 1_{\mathbb{R}^4}$ in the Global approach or $\vec{x_0} = 1_{\mathbb{R}^{4|\Omega|}}$ in the FMS approach with $xtol\_abs = 1\text{e}^{-16}$.

Dogrib-North results

In Table 6.3, we observe the performance results of each DFO algorithm selected in our study for the Dogrib-North instance, using their default configuration – internal parameters – provided by the NLOPT package and our Global tuning approach. In general, the algorithms that stand out most are BOBYQA and NEWUOA in both $NEVAL$ and $RUNTIME$ metrics. Although the one with the best/lowest objective value is NELDER-MEAD, it is the algorithm with the largest run-time (more than one hour) and number of evaluations before achieving convergence, being a critical bottleneck for situations where fast and high-quality responses are required.

From the results, the NEWUOA algorithm arises as the most suitable method for this par-

Algorithm	$NEVAL$	$RUNTIME$ [min]	$MinValue$
BOBYQA	**107**	18.03	29.10
NEWUOA	125	**16.33**	26.41
NELDER-MEAD	343	67.60	**25.59**
COBYLA	224	40.60	29.13

Table 5.5: A summary of the performance of each algorithm is presented in this table for Dogrib-North instance using Prometheus proxy. Best (minimum) results are highlighted.

ticular instance. Several experiments with multiple randomly generated weather conditions were performed, obtaining similar results to the ones already indicated. However, in practice, further improvements can be found by testing multiple configurations of the algorithms, modifying parameters like their tolerance, convergence rate, bounds of the constraints/trust region, among others. This experimentation remains open as part of future research projects.

Dogrib real fire results

Similarly, we show in Table 5.6 the performance results of each DFO algorithm selected in our study for the Dogrib real fire, using the Global approach. Again, the algorithms that stand out are both BOBYQA and NEWUOA in terms of the number of evaluations and the total run-time needed to achieve convergence towards a better objective function. As in the previous section, NELDER-MEAD is the algorithm that achieves the best objective value in the instance, mainly due to its default configuration in the NLOPT package, including a larger exploration phase than exploitation, in contrast to BOBYQA and NEWUOA which tend to be more focused on exploitation than exploration. However, both the run-time and the number of evaluations needed by NELDER-MEAD are again very high for time-constrained situations.

Algorithm	$NEVAL$	$RUNTIME$ [min]	$MinValue$
BOBYQA	**96**	**7.78**	81.94
NEWUOA	104	12.23	85.56
NELDER-MEAD	332	45.33	**80.54**
COBYLA	217	25.56	86.58

Table 5.6: A summary of the performance of each algorithm is presented in this table. Best (minimum) results are highlighted.

Note that the final errors reached by the algorithms – the minimum value of the RAF objective function – are similar in magnitude for both instances. However, the error in the real Dogrib fire instance is quite higher than in Dogrib-North. As we already mentioned, this is explained because the fire evolution data is obtained with Prometheus which simulates a fire in its natural form, unlike Dogrib fire's scar which includes the effects of suppression.

5.5 Conclusions

The adjustment of parameters is a transversal problem in all the disciplines that usually involve the development of software or computer code. In this chapter, we have presented a methodology for adjusting the parameters of fire spread models, critical for the accuracy of fire simulators using Derivative-Free Optimization methods that respond to this problem in an efficient way. The DFO-algorithms allowed us to find optimal parameters in the sense that they minimize the evolution error of the hourly fire scars obtained by Cell2Fire and Prometheus. We note at this point that, if real data is available, it would not be necessary to use simulated scars because historical fire data could give us a better fit to the underlying characteristics of the forest under study. However, this is not entirely clear and may produce misleading results because a real scar could be influenced by firefighting actions and thus, the adjusted parameters would contain this latter information (bias) and not just the natural evolution of the fire. In this case, we do not recommend using the real scar for tuning purposes, unless the researchers explicitly include the suppression actions effects, either via extra parameters in the simulation model or by modifying the instance data structure (e.g. introducing new fuel types to model the effect of these activities).

As mentioned, Cell2Fire is a cell-based fire spread simulator that models the fire spread phenomenon via an adaptation of the elliptical model proposed in O'Regan et al. (1973, 1976) using the main four ROS components of the FBP System. Nowadays, Prometheus, FARSITE and other simulators widely used are wave-propagation based, and thus, the fire dynamic is modeled in a different approach. Therefore, this methodology can serve as a bridge between the wave propagation approach and the proposed cellular approach in Pais et al. (2019).

The proposed tuning framework is able to capture both the local and global factors that affect the fire spread dynamics in a heterogeneous landscape: (1) local effects due to discontinuities of fuel types in the forest are translated to the cell-based model by re-scaling and modifying the shape of the individual ellipses generated for each fuel type and (2) the general structure (spatial disposition, topography, etc.) of the forest is captured by a general re-scaling vector, modifying the strength and magnitude of the fire evolution depending on the characteristics of the heterogeneous landscape. It worth mention that the proposed framework is general enough such that it could be easily adapted to similar simulation tools in different contexts.

Finally, it is important to note that though we have formulated the (RAF) problem using four parameters relative to the rate of spread, we could have used others. In future works, we will focus our research in this direction. It is possible that some parameters such as the slope effect, curing degree, and resilience time are not well adjusted and therefore we could include them as decision variables in the (RAF) problem as well as include new fuels, among other possibilities, in order to improve and extend Cell2Fire to new applications. As a future step, the project will address the development and adaptation of Fuel Prediction Behavior systems for the U.S., Spanish, and Chilean forests based on national field data (work in progress).

Chapter 6

Multiscale definition of the wildland-urban interface to mitigate fire risk: an evidence-based approach

Authors

Jaime Carrasco and Andrés Weintraub
University of Chile, Industrial Engineering Department, Santiago, Chile.

Alejandro Miranda
Universidad de La Frontera, Departamento de Ciencias Forestales, Temuco, Chile.

Alexandra Syphard
Conservation Biology Institute, Corvallis, Oregon, USA

Abstract

Include the fire risk in building WUI-maps is key to protecting our human communities. This importance has become greater due mainly to the climatic change that has exacerbated these incidents, both in occurrence as well as in magnitude and damage to public and private property. However, quantifying the fire risk relationship with the different human, topographic and climatic drivers is not an easy task. In this study we address both problems. For them, we first built a Bagged Decision Tree (BDT) model to quantify the risk of fire occurrence from different variables, grouped into: Human Activity, Geographic and Topographic, and Land Cover. Subsequently, using the mathematical elements underlying the BDT models, we determine individual relationships between the risk and the variables considered to address the problem of delineating a WUI-map. Our approach is applied in Chile, but could be replicated in any region of the world.

6.1 Introduction

The Wildland-Urban Interface (WUI) is the spatial manifestation of the coupling of human communities and ecosystems, where interactions can take many forms (Bar-Massada et al., 2014). One of the result of this interaction is the increased fire risk (Radeloff et al., 2005), for this reason in many parts of the globe, the WUI accounts for large percentages of fire control expenditures because it is where most human fatalities and structure losses occur (Kramer et al., 2018; Moritz et al., 2014). Wildfires at the WUI constitute a socio-environmental and security problem, as well as an economic. There is a need to define where to concentrate management actions of our landscapes, to obtain the maximum social return (Stein et al., 2013). Knowing which areas of the WUI are most fire-prone, would help to protect people and surrounding ecosystems, through fire risk mitigation and prevention, e.g., managing the flammability and continuity of landscape fuels (Spyratos et al., 2007) or diminishing human-started wildfire (Balch et al., 2017). Therefore, a fire-risk based definition of the spatial distribution of the WUI , it is of the utmost importance to prioritize areas and actions for properly distributing prevention action and management investments in rural and urban areas near wildlands (Calkin et al., 2014; Tonini et al., 2018).

Fire risk can be understood as the chance that a fire might start and then respond to multi-scale spatial and temporal interactions among vegetation, or fuel type, with topography and the human footprint and behavior (Hardy, 2005). Since we have to consider that there are many changes by human communities, as population grows, that have side effects in fire-prone ecosystems; the risk may expand due to land abandonment, shifting vegetation dynamics, or other types of land-use change (Badia et al., 2019; McWethy et al., 2018). Shifting demographics and human population growth ,can lead to changing development patterns either via urban sprawl, the intensification of the road network or by rural-urban migration (Tonini et al., 2018; Blackhall and Raffaele, 2019) could alter the distribution of human-caused ignitions (Ganteaume and Syphard, 2018). Also considering that the climate change has the potential to dramatically increase human and natural communities' vulnerability to wildfire.

As explained above, it should be a priority delineating the WUI worldwide, would allow us to elaborate informed maps to assist the community and plan the use of the land, as well as the adoption of homeowner mitigation practices to minimize fire risk (Moritz et al., 2014; Calkin et al., 2014; Keeley and Syphard, 2019). However, most of the efforts in the development of WUI-maps have focused on delineate the values at stake (human lives and integrity, communities, infrastructure) without explicitly including the fire risk (Johnston and Flannigan, 2018). WUI commonly used for mapping three types of data (Radeloff et al., 2005; Platt, 2010), residential or population density and their development, ii) vegetation cover and type, and iii) minimum distance between them (Johnston and Flannigan, 2018). The WUI is then delineated using an arbitrary threshold, that quantifies the degree of spatial overlap between these variables and the distance among them. This is a clear and quick method for mapping purposes, but a significant challenge still remains, is identify appropriate thresholds for different socio-ecological contexts as the local scale and also the national (Amato et al., 2018), that imply the maximum social return diminishing fire risk where higher potential damage can cause.

Accounting for a full range of effects and interactions among environmental and anthropogenic drivers is difficult using traditional statistical modeling approaches (Breiman et al., 2001; Elith et al., 2008). Artificial Intelligence (AI) algorithms provide a compelling alternative to evaluate risk based on multiple drivers because they can capture complex relationships among predictors and response variables (Elith et al., 2008). We used these tools to model fire risk in the landscape and subsequently define thresholds to map the WUI at national and subnational scales. In this chapter, we address the problem of including fire risk to define the WUI in Chile. For them, we propose a methodology based on machine learning methods, which provides us with the mathematical elements to determine the critical thresholds based on historical fire occurrence data. Our methodology is robust and replicable, and could possibly be used elsewhere in the world.

6.2 Material and Methods

6.2.1 Study area

The study area includes eight administrative regions (∼212.000 km2), which constitute 28% of Chilean national territory. It covers from 32° to 43.8°S, accounting for 86% of the country's human population and 98.5% of its historical national fire occurrence (González et al., 2018) (see Figure 6.1). The northern part of the study area has a semi-arid and Mediterranean-type climate with mean annual rainfall ranging from 100 to 500 mm, with precipitation concentrated during the austral winter and a long dry season with peak temperatures over 30° (Viale and Garreaud, 2015). The southern part (∼38°S) is dominated by a temperate climatic regime with a mean annual rainfall along the central valley between 1200–2000 mm with higher values toward the Andean range. Farther south, precipitation can reach a yearly mean between 1500 mm and over 3500 mm.

In the last five decades, central and southern Chile have experienced highly dynamic

land use and land cover change (LULCC) with different temporal/geographic variability and patterns (Heilmayr et al., 2016; Miranda et al., 2017). In central Chile (33°–34°S) and in southern Chile (40°–42°S) the main pattern has been the conversion of native forest to shrublands, with 45% of total change to date. From El Maule (35°) to Los Ríos regions (40°S), the expansion of dense exotic tree plantations for pulp and paper production was the main contributor of native forest loss, creating highly homogeneous (Miranda et al., 2015) and flammable landscape (McWethy et al., 2018). According to these climatic, land cover, and LULCC latitudinal variations, we divided the area into five sub-national zones (SNZ), which vary between 1.3°and 2.9°degrees along this latitudinal gradient (see Figure 6.1).

6.2.2 Wildfire data

We use a subset of the public database of wildfires, provided by Chilean Forest Service (CONAF, `www.conaf.cl/conaf/seccion-stadisticas-historicas.html`) that gives us the spatial coordinates of ignition points for every fire within 0.01 ha mapping units. From these data, we created a temporal subset, that includes all fires from January 2013 to December 2015. We used this temporal window, because CONAF is registering the coordinates of the nearest ignition point from 2013, and most accurate land cover map for the country was developed in 2014 (Zhao et al., 2016). The final database included 19,413 fires, covering a total area burned of 245,815 ha spanning all land cover classes, where 70% of fires were less than 1 ha, 23% from 1 to 10 ha, 6% from 10 to 200 ha, and only 1% over 200 ha but accounting for 71% of burned area.

6.2.3 Wildfire risk drivers

We selected potential variables to use for our risk model based on factors found to be important in driving fire occurrence both in Chile as well as other regions (Bar-Massada et al., 2014; Curt et al., 2016b; Syphard et al., 2019). We considered three different groups of drivers: i) human activity, ii) geographical and topographical, and iii) land cover (González et al., 2018; McWethy et al., 2018; Urrutia-Jalabert et al., 2018; Gómez-González et al., 2019). For each group, we prepared a set of spatially explicit variables (Table 6.1), which were assigned to each ignition point (Figure 6.1, right panel). We calculated the drivers for each ignition point in two different ways (1) extracting the value of the pixel in the precise ignition point coordinates and (2) by zonal statistics of 500 m diameter buffer around the ignition point (see details in Table 6.1). For land cover drivers we re-categorized the (Zhao et al., 2016) map in nine classes: croplands, native forest, forest plantation, pasturelands, shrublands, wetlands, water bodies, impervious surfaces, and bare lands.

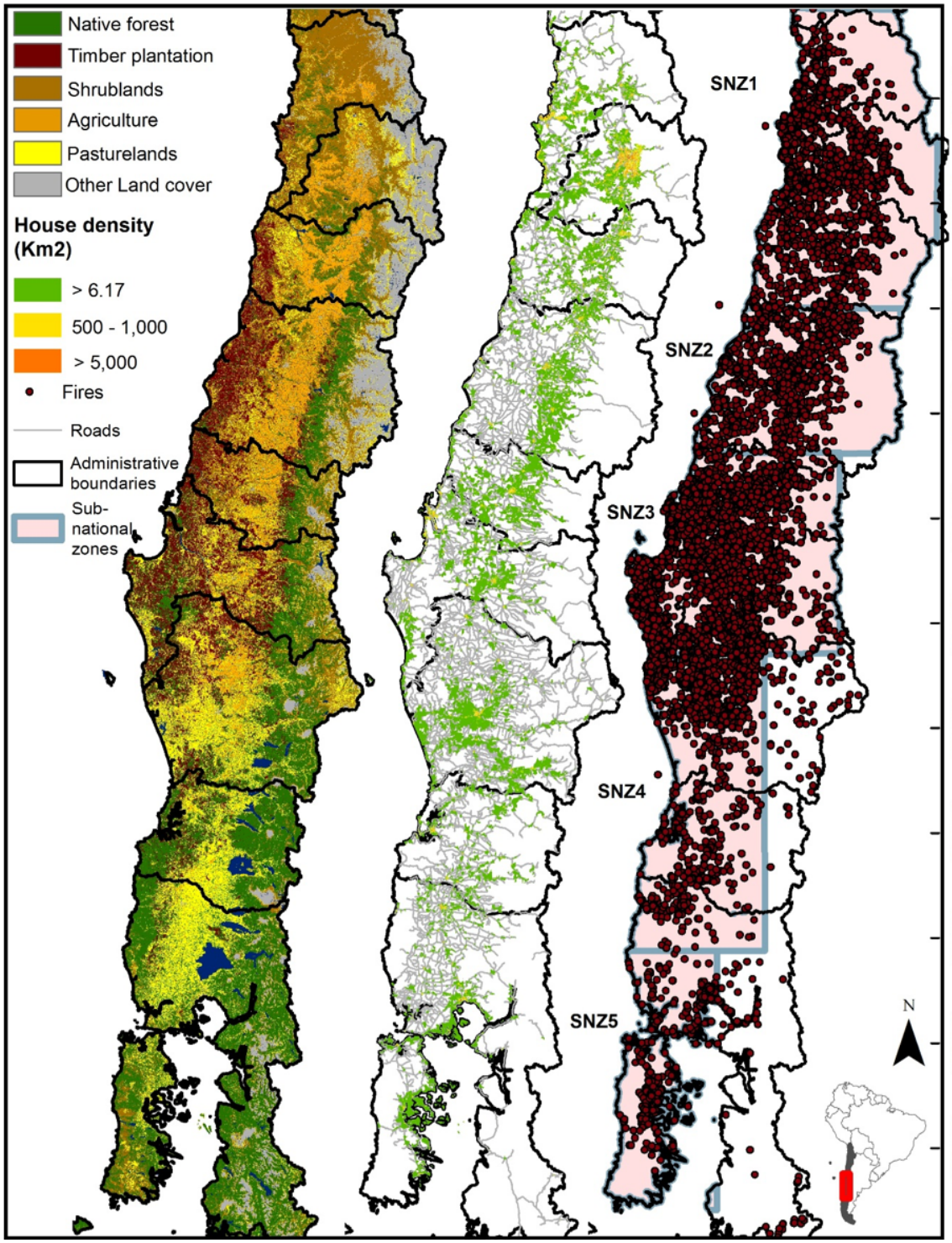


Figure 6.1: Study area. Left panel: land cover map for 2014 from Zhao et al. (2016). Centre panel: Housing density (housing units/km2) in national block census (INE 2018) with main roads (MOP 2018). Right panel: fire ignition occurrence locations from the CONAF (2018) database and subnational zones from SNZ1 to SNZ5.

Drivers	Variable	Description	Units	Data source
Human Activity	Road density	Calculated as the mean distance to the nearest road in a 500 m radius buffer around the ignition point	meters	MOP (2018)
	Road distance	The distance of ignition point to the nearest road	meters	MOP (2018)
	City distance	Distance of ignition point to the nearest city	meters	INE (2018)
	Population density	Number of houses divided by census block area	Population/km2	INE (2018)
	Housing density	Houses density	housing units/km2	INE (2018)
Geographic and Topographic	Elevation	Mean elevation in a 500 m radius buffer	m.a.s.l.[1]	SRTM
	Slope	The mean slope in 500 m radius buffer	Percentage	SRTM
	Longitude	Longitude of the ignition point	decimal degrees	CONAF (2018)
	Latitude	Latitude of ignition point	decimal degrees	CONAF (2018)
Land cover[2]	Crops	Proportion of crops in the landscape	Percentage	Zhao et al. (2016)
	Native forest	Proportion of native forest in ∼80 ha	Percentage	Zhao et al. (2016)
	Forest plantation	Proportion of forest plantation in ∼ 80 ha	Percentage	Zhao et al. (2016)
	Impervious[3]	Proportion of impervious land in ∼80 ha	Percentage	Zhao et al. (2016)
	Shrublands	Proportion of shrublands land in ∼80 ha	Percentage	Zhao et al. (2016)
	Barelands	Proportion of barelands in ∼80 ha	Percentage	Zhao et al. (2016)
	Pasturelands	Proportion of pasturelands in ∼80 ha	Percentage	Zhao et al. (2016)
	Land cover	Land cover in ignition point	Categorical	Zhao et al. (2016)

[1] Meters above sea level; [2] Proportion of the landscape cover by it in a 500 m radius buffer; [3] Artificial structures.

Table 6.1: These are all the variables to build our machine learning models in this study to evaluate the probability of fire occurrence (fire risk). SRTM data was downloaded from Earth Resources Observation and Science Center (EROS, https://www.usgs.gov/centers/eros/data-tools).

6.2.4 Modelling approach

Background

Artificial Intelligence (AI) methods, particularly those of Machine Learning (ML), are useful in finding complex interactions between the study variables directly from the data (Olden et al., 2008). This method is very beneficial in practical applications, where there are no

parametric models that explain the phenomenon in question. Moreover, some authors have pointed out that traditional statistical methods are not suitable for problems that come from ecology or related fields due to their low predictive power (Breiman et al., 2001; Elith et al., 2008). We can find several applications in the field of modeling fire risk assessment (Oliveira et al., 2012; McWethy et al., 2018). Their results show that the decision tree models reach a higher predictive ability than the commonly used Multiple Linear Regression model. However, most of the studies in this field use traditional Logistic Regression Models (González et al., 2006; Wotton et al., 2010).

Additional data

The main objective in conceptual and computational terms, was to understand which factors were related to fire ignitions. The database contains the ignition points, but to understand what drivers lead to "non-fire ignition", 20,000 random points were generated over the study area. We eliminated those points at less than 2 km from any ignition point to avoid overlap and bias, thus obtaining a total of 14,360 points. Following this procedure, we built a full database including 33,773 points (19,413 + 14,360) where each data is marked using a binary codification with value "1" if there was a fire or "0" otherwise.

Bagged Decision Tree

We implemented Bagged Decision Tree (BDT) method (Breiman, 1996). BDT is used in classification and regression models (Opitz and Maclin, 1999). It is a type of ensemble method[1], meaning that different classification models (called "weak learners") — that we will denote by $\mathcal{T}_k$ — are combined to obtain a more robust prediction. In BDT, each *weak learner* is built by drawing random training sets from the given sample. This re-sampling is usually done following a bootstrapping technique, where numerous versions of the training set are created using sampling with replacement and each of those versions is used to build a different learning model.

In mathematical terms, let $\mathcal{D}_1 := \{(\vec{x}^i, y^i)\,,\ i = 1, ..., N_1\}$ be the training set, where $\vec{x}^i$ is the vector of variables (whose components are the variables described in Table 6.1) labeled by $y^i \in \{0, 1\}$, where $y^i = 1$ if it is an ignition point and $y^i = 0$ if it is not. The bootstrap aggregating method consists of generating subsets $\mathcal{D}^k \subset \mathcal{D}_1$ of size $M < N_1$ by randomly drawing – with replacement – samples from the original training set (see Figure 6.2 for an example of decision tree). Then, each decision tree $\mathcal{T}_k$ is built from $\mathcal{D}^k$, for $k = 1, ..., K$ (see Figure 6.3). Later, classifiers $\mathcal{T}_k$ are used in parallel, each one offering an opinion $\mathcal{T}_k(\vec{x})$ as to which class the example should be labeled with, in order to predict most likely class for an unobserved vector $\vec{x}$. A "master classifier" $\mathcal{T}$ collects this information and then chooses the label y that has received more votes as follows:

[1]In the words of Dietterich (2000): "Ensemble methods are learning algorithms that construct a set of classifiers and then classify new data points by taking a (weighted) vote of their predictions"

$$\mathcal{T}(\vec{x}) = y \quad \Leftrightarrow \quad y = \begin{cases} 1, & \tilde{y} > 0.5, \\ 0 & otherwise, \end{cases} \qquad \text{where } \tilde{y} := \frac{1}{K} \sum_{k=1}^{K} \mathcal{T}_k(\vec{x}). \tag{6.1}$$

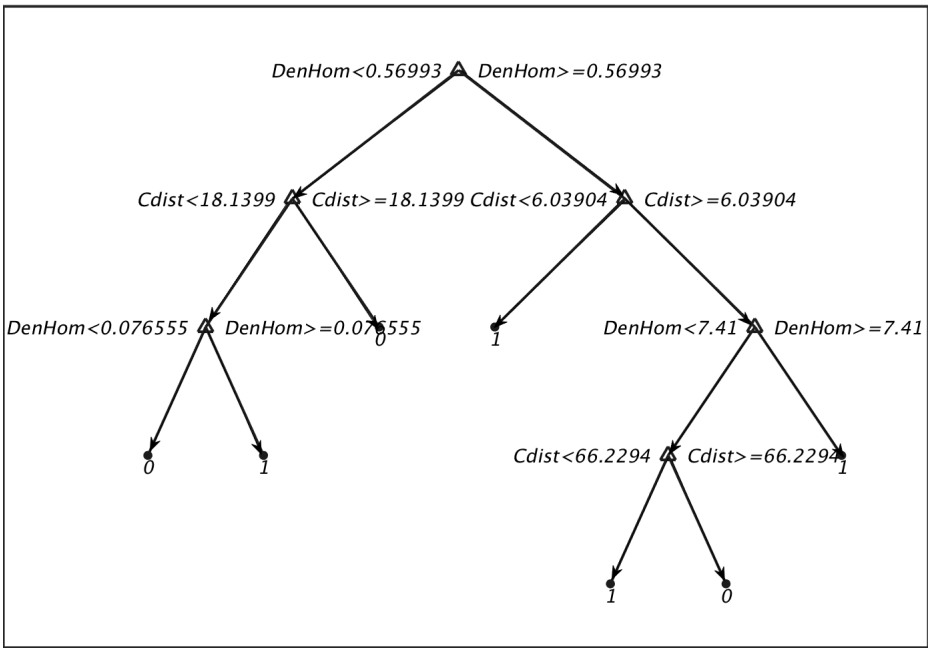


Figure 6.2: Example of decision tree for two variables: City distance ($Cdist$) and Population density ($DenHom$).

We notice that higher values of $\tilde{y}_i$ indicate that a greater number of $\mathcal{T}_k$ gave their vote to classify $\vec{x}$ as an occurrence of type "one point" instead of a "zero point", and vice versa. Based on this, we define the *fire risk* for a point $\vec{x}$ and denoted it by $\mathcal{R}(\vec{x})$ using the value of $\tilde{y}_i$. In addition, this risk measure increases (or decreases) when there are more positive (negative) points occurring in similar places. In other words, points that share the same geographical area have the same attributes, and therefore – in the process of building the $\mathcal{D}_k$ – they will be more likely to be selected to train the decision trees $\mathcal{T}_k$ during the execution of the bootstrapping technique. This way, for any new point $\vec{x}$, there will be more voters with similar responses, increasing/decreasing the value of $\mathcal{R}(\vec{x}) := \tilde{y}$ depending on the type and density of different events surrounding it.

Voting schemes can be seen as approximations under a Bayesian framework. In fact, informally denoting $\mathbb{P}(\mathcal{T}_k) = 1/K$ as the uniform selection of decision trees (K is the number of decision trees in the BDT method; see Figure 6.3); Y the random variable which represents the choice of the occurrence label; and replacing $\mathcal{T}_k(\vec{x})$ in the Eq. 6.1 with the class posterior probabilities $\mathbb{P}(Y \mid \vec{x}, \mathcal{T}_k)$ – relative frequency of classes, before the tree leaf – the Eq. 6.1 (right side) can be rewritten as follows:

$$\mathbb{P}(Y \mid \vec{x}) = \sum_{k=1}^{K} \underbrace{P(Y \mid \vec{x}, \mathcal{T}_k)}_{T_k(\vec{x})} \cdot \underbrace{\mathbb{P}(\mathcal{T}_k)}_{1/K} \tag{6.2}$$

Equation 6.2 helps us to understand the intuition behind the previously defined fire risk (Eq. 6.1) as a probability that is conditional on the attributes (variables) of the point, obtained as a weighted summation of the class posterior and decision trees selection probabilities. Therefore, in what follows, we use the concepts of risk and probability of occurrence without any distinction.

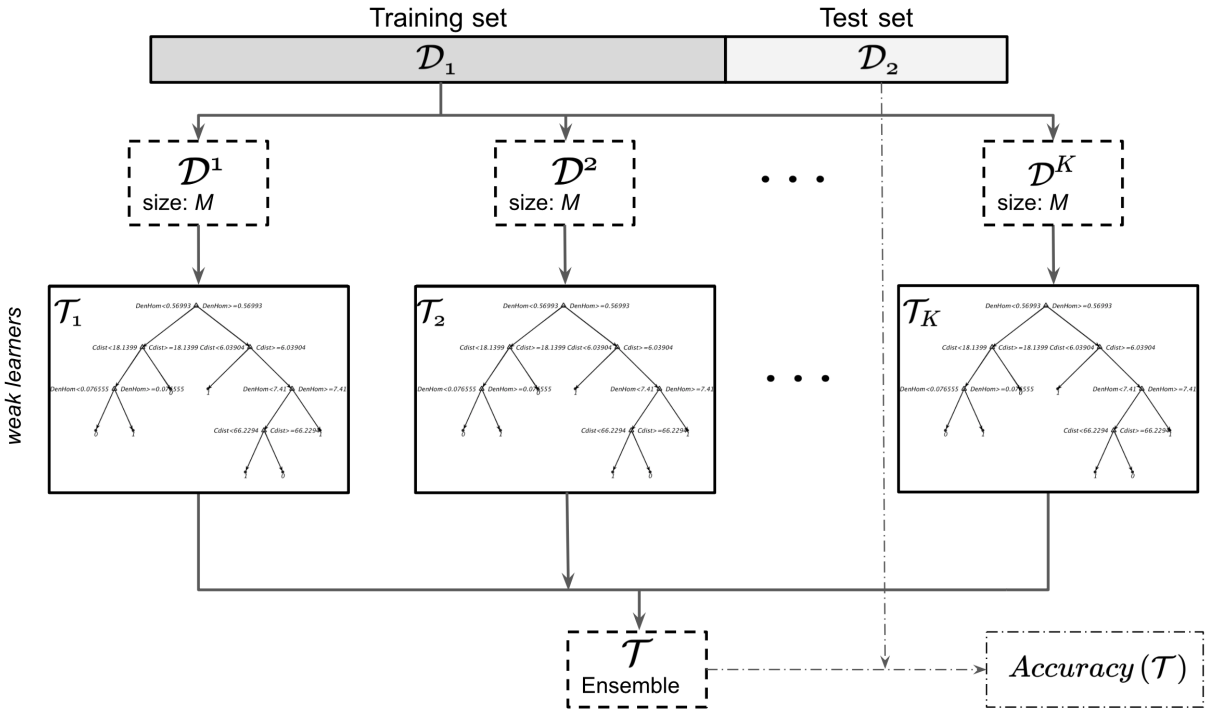


Figure 6.3: BDT scheme: reading from top to bottom, first K subsets of data are created from training sample $\mathcal{D}_1$ chosen randomly with replacement (bootstrap technique). Then, each subset data $\mathcal{D}^k$ is used to train their decision trees $\mathcal{T}_k$; and in this way, a set of different models is obtained. Finally, decision trees are combined into a "stronger ensemble" $\mathcal{T}$. The output response of $\mathcal{T}$ is calculated by the Eq. 6.1 (right side). Using the data set $\mathcal{D}_2$ and Eq. 6.3 the accuracy level of $\mathcal{T}$ is calculated.

National and subnational models

We built national and subnational models for fire risk evaluation, the national model considers all fires in the study area while the subnational models were defined following five bioclimatic zones for local fire risk evaluation which in sum contain more than 99% of fires considered in national model (see Figure 6.1).

For the national fire risk model, the full database was considered for the training and testing sets. In the case of the SNZ's models, we only considered the subsets of the original database that corresponded with each region (Figure 6.1). In the national model, we retained 75% of the ignition data for training and 25% for testing. In the case of SNZ's models were trained with 85% of the data and tested on the remaining 15% — we chose this data

set distribution for SNZ's models to retain more data for training. The final datasets was balanced between non-ignitions (0s) and ignitions (1s) observations in order to have similar observations of each category (Salas-Eljatib et al., 2018).

Zone	Latitudinal range	Number of fires	Percentage of fires (%)	Native forest (%)	Forest plantation (%)	Crops (%)	Pastures (%)	Shrubland (%)
SNZ1	-32.0°, -34.9°	4,343	22.2	6.7	4.9	13.4	11.3	49.1
SNZ2	-34.9°, -36.2°	1,858	9.5	12.5	15.5	20.1	10.0	28.9
SNZ3	36.2°, -38.2°	10,290	52.6	18.7	31.8	5.5	20.8	17.1
SNZ4	38.2°, -41.0°	2,305	11.8	27.4	15.1	8.2	36.6	8.4
SNZ5	-41.0°, -42.5°	617	3.2	37.9	0.8	0.5	35.1	18.4

Table 6.2: Description of subnational zones for fire risk modeling. Land cover was calculated based on the zones' total area and not only within the vicinity of fires.

Validation accuracy

After training the model, a validation accuracy score estimates the performance of the model on an independent data set $\mathcal{D}_2 := \{(\vec{x}^i, y^i), i = 1, ..., N_2\}$ (see Figure 6.3). This procedure is known as *holdout validation* (Kim, 2009). It is worth mentioning that the full data set is such that $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, $\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$.

Let $\mathcal{T}$ be the ensemble model built following the procedure described in the previous section, using the training set $\mathcal{D}_1$. Then, the accuracy of the method is defined by:

$$Accuracy\left(\mathcal{T}\right) = 100 \times \frac{1}{N_2} \sum_{i=1}^{N_2} \delta\left(y^i - \mathcal{T}(\vec{x}^i)\right), \text{ where } \delta\left(z\right) := \begin{cases} 1, & z = 0, \\ 0, & otherwise; \end{cases} \tag{6.3}$$

which represents the percentage of successes obtained by the model on the validation data set $\mathcal{D}_2$.

Variable importance measure

We used "Predictor Importance (PI)" (Breiman, 1996), a measure that assigns a score to each variable based on its contribution to the model learning process. With PI, candidate predictor variables can be compared according to their impact in predicting the response, or even their causal effect (Strobl et al., 2008). Because we used a random process to build the model ensemble from a bootstrap technique, it was possible that choice of different seeds for the generation of random numbers could lead to different relative importance values. To assure the more robust information, we generated 100 models with multiple random seeds and report the distribution of these measures — via boxplots — for each variable. Our results are discussed in Section 6.3 (see Figure 6.4).

At the beginning of our study, PI measure also helped us to dismiss variables with low predictive power and thus obtain a simpler BDT model to reduce computational cost and to

obtain a model that is less likely to overfitting (Olden et al., 2008). Only variables with a relative importance value greater than 2% were retained.

National and subnational thresholds definition

We used Partial Dependence Plots (PDPs) (Elith et al., 2008), to analyze relationships between features and predicted responses. In particular, this consists of a line plot of occurrence probability, estimated by the BDT model against a single predictor variable. For example, if we separate the attribute vector $\vec{x}$ in a subset of attributes $\vec{x}_S$ and $\vec{x}_C$, where the latter vector represents the complement over $\vec{x}$, we get $\vec{x} = (\vec{x}_S, \vec{x}_C)$ and the expectation of predicted answers regarding $\vec{x}_C$ is naturally given by:

$$\mathcal{T}_{\vec{x}_S}(\vec{x}_S) = \mathbb{E}_{\vec{x}_C}\left[\mathcal{T}(\vec{x}_S, \vec{x}_C)\right] \approx \frac{1}{N}\sum_{i=1}^{N}\mathcal{T}(\vec{x}_S, \vec{x}_C^i) \tag{6.4}$$

where $N = |\mathcal{D}|$ and $\vec{x}^i = (\vec{x}_S^i, \vec{x}_C^i)$ is the i-th observation. The last sum in Eq. 6.4 represents an approximation, assuming that each observation is equally likely. The PDP is defined as the graph of the function $\mathcal{T}_{\vec{x}_S}$, which we will denote by $\mathcal{G}\left[\mathcal{T}_{\vec{x}_S}\right]$. A PDP can show whether the relationship between the fire risk and a specific variable is linear, monotonic or more complex (see Figure 6.5 of the next section). We assumed that inflexion point or abrupt changes in the slope of the curve showed us thresholds for specific variables where a management action could potentially attain the best social return for decreasing fire risk. In general, these functions can be very irregular since they depend directly on the observations, making it difficult to accurately determine these threshold points. For this reason, we adjusted different algebraic curves using nonlinear least squares (Marquardt, 1963). With these curves, we obtained more robust and reliable thresholds.

We found three kinds of relationships between individual variables and fire risk; i) monotonically decreasing effect curve (MD) (Figure 6.5a,b), ii) monotonically increasing effect curve (MI) (Figure 6.5c,e) and iii) plateau effect curve (Pl) (Figure 6.5d,f). These patterns of interaction determined the kind of thresholds proposed: MD must have had a continuous decreasing trend, however could have a maximum value but decrease from that point. MI had to show a positive relationship with the individual variable increase; and Pl had an initial increasing trend in fire risk that remained relatively constant until a new critical value at which fire risk fell.

For MD we estimated an upper threshold of variables over which risk was defined. For MI, we estimated a bottom threshold above which variables with this type of curve were at were defined as high risk. For the Pl case, the threshold represented an interval between which the highest risk was concentrated.

6.3 Results

6.3.1 National scale fire risk model

The best model included variables that reached ∼90% of global accuracy at national scale; this result was obtained using Eq. 6.3, the main variables related to fire risk were: elevation and latitude (geographic variables), road density and road distance (human activity), and forest plantation and native forest (land cover variables) (Figure 6.4). The national model shows that the relative importance of the variables is dominated by human activity accounting for 40% of relative importance. Latitude was the most important variable potentially related to fire occurrence, highly concentrated between 35°S and 38.5°S (Table 6.2).

6.3.2 Subnational scale fire risk model

Bagged Decision Trees reached high accuracy in each SNZ (range: 83-90%, Table 6.4). The best models were for SNZ1 and SNZ3, with accuracy as high as the national model: 90% and 88% respectively. At sub-national level, accuracy decreases compared to the national model, because there is less data available for training. Fire risk models showed latitudinal variation in relative importance of variables (Figure 6.4; Table 6.4). Land cover variables had higher differences in relative importance between subnational zones (CV = 55.5%), followed by geographical (CV = 48.1%), and human activity variables (CV = 34.4%). Land cover group variable importance varied between SNZs from almost no differences in adjacent SNZ's (e.g. SZN1 and SZN2) and doubled the importance between others (SNZ3 and SZN4). In SZN4 land cover variables reached 50.5% of the importance in the model with highest values in proportion of shrublands, native forest and forest plantation (Table 6.3). The minimum relative importance for tree plantation was related to its lower percentage of cover in the landscape in SNZ5 (Table 6.3). However, the proportion of land cover in the landscape was not always correlated with the importance for prediction (Table 6.3). Human activity variables had the mean highest value of importance for SNZ1, SNZ2, and SNZ3 where the main cities and human population in the country are most concentrated. For this group, the most important factor was distance to cities, where SNZ5 showed the maximum value in an area with low human and roads density, revealing the impact of human influence in fire risk.

6.3.3 Wildland urban interface fire risk thresholds

We found three kinds of relationships between individual variables and fire risk; i) monotonically decreasing effect curve (MD) (Figure 6.5a,b), ii) monotonically increasing effect curve (MI) (Figure 6.5c,e) and iii) plateau effect curve (Pl) (Figure 6.5d,f). These patterns of interaction determined the kind of thresholds proposed: MD must have had a continuous decreasing trend, however could have a maximum value but decrease from that point. MI had to show a positive relationship with the individual variable increase; and Pl had an initial increasing trend in fire risk that remained relatively constant until a new critical value at

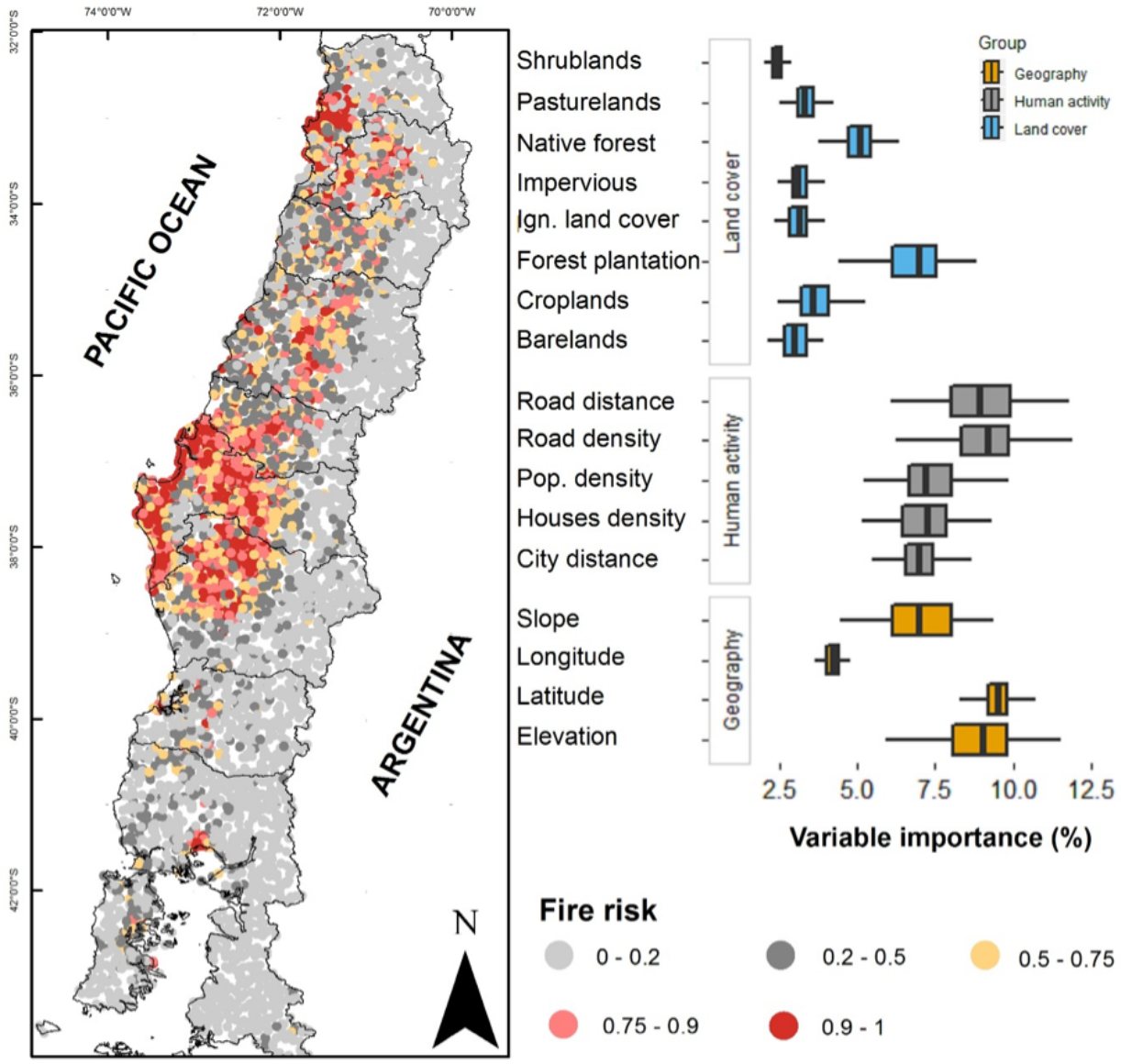


Figure 6.4: Left panel: predicted national fire risk, Low: 0-0.5 prob., Medium: 0.5-0.75 prob., Medium-high: 0.75-0.9 prob. and High: 0.9-1 prob. Right panel: variables relative importance for national model.

which fire risk fell.

National fire risk thresholds

At the national scale for the human activity variables group, we found a positive effect on fire risk for the three most important variables, with the main individual effects on increasing ignition probability. Distance to cities and roads had a local effect in increasing or maintaining fire risk, especially for the first 1200 m and 1500 m respectively (Figure 6.5a, b). However, they had different effects. Fire risk had a continuously decreasing trend (MD) from roads (Figure 6.5a), while distance from cities maintained a high constant probability over that

Variable/Zone	National		SNZ1		SNZ2		SNZ3		SNZ4		SNZ5	
Land cover	mean	sd	mean	sd	mean	sd	mean	sd	mean	Sd	mean	sd
Shrublands	2.4	0.2	5.0	0.7	3.8	0.5	2.0	0.1	**11.3**	0.7	4.0	0.3
Pasturelands	3.3	0.4	5.0	0.7	2.8	0.3	1.9	0.1	6.2	0.5	3.4	0.3
Native forest	5.1	0.6	1.8	0.2	3.5	0.5	5.2	0.6	9.8	0.7	3.8	0.5
Impervious	3.1	0.3	4.5	0.8	4.7	0.7	1.7	0.2	3.1	0.3	**10.5**	1.9
Ignition land cover	3.1	0.4	4.2	0.7	3.4	0.5	2.5	0.3	5.6	0.5	1.7	0.2
Forest plantation	**6.9**	1.0	1.5	0.3	5.0	0.6	**6.3**	0.8	8.4	0.4	1.3	0.1
Croplands	3.6	0.6	**9.8**	1.1	**7.4**	1.0	2.0	0.1	4.9	0.4	7.6	1.4
Barelands	3.0	0.5	4.6	0.7	4.4	0.8	2.9	0.5	1.2	0.2	8.5	1.8
sum	30.6	–	36.4	–	35.0	–	24.5	–	50.5	–	41.0	–
mean	3.80	–	4.60	–	4.40	–	3.10	–	6.30	–	5.10	–
Human activity												
Road distance	9.0	1.3	**9.0**	1.4	9.2	1.3	9.9	1.5	6.0	0.6	2.9	0.3
Road density	**9.1**	1.4	8.9	1.2	**9.6**	1.4	**10.3**	1.5	6.9	0.6	3.1	0.3
Population density	7.4	1.1	8.3	1.2	8.2	1.2	2.6	0.2	7.1	0.5	9.5	1.3
Housing density	7.2	0.9	8.3	1.2	8.2	1.1	2.6	0.3	6.8	0.6	9.7	1.2
City distance	7.0	0.7	5.2	0.6	7.6	1.0	9.6	1.1	**7.7**	0.3	**14.3**	2.0
sum	39.7	–	39.6	–	42.8	–	35.1	–	34.5	–	39.5	–
mean	7.90	–	7.90	–	8.60	–	7.0	–	6.90	–	7.90	–
Geographical and Topographical												
Slope	7.0	1.1	8.4	1.3	7.2	1.0	9.0	1.4	3.9	0.2	3.0	0.3
Elevation	9.0	1.2	**8.7**	1.2	**8.1**	1.1	**17.2**	1.9	5.2	0.3	2.9	0.3
Longitude	4.2	0.3	6.8	0.9	6.9	1.0	14.2	1.4	**6.0**	0.3	**13.7**	1.8
Latitude	**9.5**	–	0.6	–								
sum	29.8	–	24.0	–	22.2	–	40.5	–	15.1	–	19.5	–
mean	7.40	–	8.0	–	7.4	–	13.5	–	5.0	–	6.50	–

Table 6.3: Variables relative importance for national and subnational models. The highest values are highlighted in bold-text. We observe that theses values varies according to subnational zones. This is explained due to changes in land use throughout the national territory.

range with the probability reaching under 0.6 at 8 km (Figure 6.5b). For housing density we found an exponential trend of ignition probability from zero to 5.6 housing units/km2, increasing fire risk more than 10% in that range. From that housing density, we found a lower increasing rate of 3.5% of fire risk until 35 housing units/km2 where it stabilized, maintaining a high risk (Figure 6.5c).

We found differing relationships between individual land cover variables and fire risk (Figure 6.5, Table 6.4). Forest plantation had the main individual effect, with risk increasingly sharply with more cover. When forest plantation cover increased from 0 to 20%, the fire risk probability hit 0.65, but the highest rate of increase in fire risk was when cover increased from 0 to 4.5%, which thereby defined the thresholds at which fire risk probability stabilized (Figure 6.5e). Conversely, native forest cover proportion has a rapidly increasing effect, with a maximum fire risk at 18% cover that remained high up to 50%, after which there was a steep decreasing trend (Figure 6.5d). Croplands also showed an increasing effect but with a

stabilization threshold at 20% which fell after 60% (Figure 6.5f).

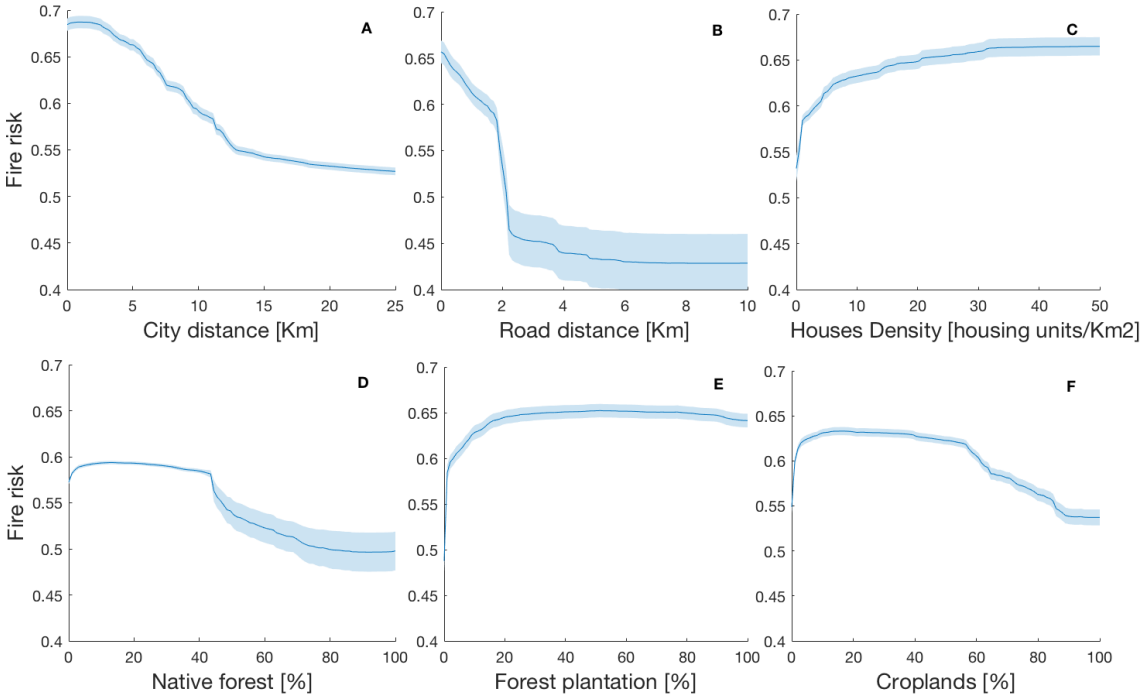


Figure 6.5: One-dimensional effect of human activities on fire risk at the national scale, a) distance to cities (km), b) distance to roads (km), c) housing density (housing units/km2), d) native forest proportion (%), e) forest plantation proportion and f) croplands proportion (%). The light blue band represents the standard deviation of the fire risk after training 100 models with multiple random seeds. The dark line represents the average risk.

Subnational fire risk thresholds

We found the same types of relationships between variables and fire risk in the subnational models as we did in the national scale model, MD, MI, and Pl; however, in some SNZs, the effect of individual variables was negligible and did not exceed 0.5 of fire risk (Table 6.4). Human activity variables, in general, showed higher individual effects on increasing fire risk for all SNZs except in SNZ1 and SNZ4 where the maximum value was reached by Croplands and Forest plantations respectively. For every SNZ, distance to cities had a strong and consistently increasing effect on fire occurrence probability from at least 0.8 km to a maximum of 1.71 km, in SNZ3. The stabilization of fire risk in terms of housing density varied from 3.3 to 5.7 housing units/km2, and this variable did not show a significant effect on southern SNZs (Table 6.4).

For land cover variables, native forest reached a minimum increasing effect until 11% of cover in SNZ1, then subsequently decreased. In SNZ2 the risk increased until complete cover of native forest, but the highest risk was between 19% and 29% cover (Table 6.4). For all zones, we found positive and monotonically increasing fire risk relative to forest plantation

cover, with the main increase in SNZ3, with a 16% increase in fire occurrence probability with 10% of tree plantation cover that continued to increase until 67% cover, with a total gain of 23% in fire risk. Croplands for different SNZs had all three types of relationships between individual variables and fire risk, therefore, different kinds of thresholds were proposed (Table 6.4).

Subnational Zone	Accuracy Assessment (%)	City distance (Km)	Road distance (Km)	Housing density (housing units /km2)	Forest plantation (%)	Native forest (%)	Croplands (%)
SNZ1	90	1.3 ↓	1.7 ↓	4.9 ↑	5.0 ↑	11-13	20-60
SNZ2	83	1.1 ↓	1.3 ↓	3.3 ↑	2.2 ↑	19-29	40-60
SNZ3	88	1.7 ↓	1.4 ↓	5.7 ↑	9.8 ↑	13-28	9.4 ↓
SNZ4	83	1.1 ↓	0.9 ↓	(*)	5.4 ↑	17.3 ↓	40 ↑
SNZ5	86	0.8 ↓	(*)	(*)	(*)	(*)	(*)
National	89	1.2 ↓	1.5 ↓	5.6 ↑	4.5 ↑	18-50	20-60

Table 6.4: Defined threshold for national and subnational models. The upper threshold is represented by ↓, ↑ indicates a bottom threshold and (*) denote a non-relevant threshold with under 0.5 fire risk for the variable.

6.3.4 First WUI-map for Chile

We conclude this section showing a WUI-map in Figure 6.6, delineated from the thresholds obtained at national level. In particular, for its construction, the areas that satisfy the following relationships (all at once) are identified: i) areas that are below (<) 1.5 km of road distance; ii) areas that are below (<) 1.2 km away from the cities; iii) areas where the proportion of native forest less (<) than 18%; iv) areas where the density of houses is greater (>) than 5.6 houses/km2; and the v) proportion of plantations is greater (>) than 4.5%.

6.4 Discussion

6.4.1 National fire risk model

The strong human influence identified in the national model is consistent with the fact that more than 95% of fires in Chile are human-caused (González et al., 2018). The dominance of human-caused fires is common in other regions with Mediterranean-type climates (Manzello et al., 2018); and human influence on fire occurrence is increasing globally with population growth and urban expansion into wildland vegetation (Balch et al., 2017). Consistent with other studies, the national model identified most ignitions start near the place where people live, work, travel, recreate or do other activities in areas with human-dominated fire regimes (Fusco et al., 2016) and is consistent with other studies in Chile (McWethy et al., 2018). Even elevation could be considered as human activity proxy because the main roads and major cities are in the coastal and central valley of Chile (lower altitude) (Figure 6.1). In addition,

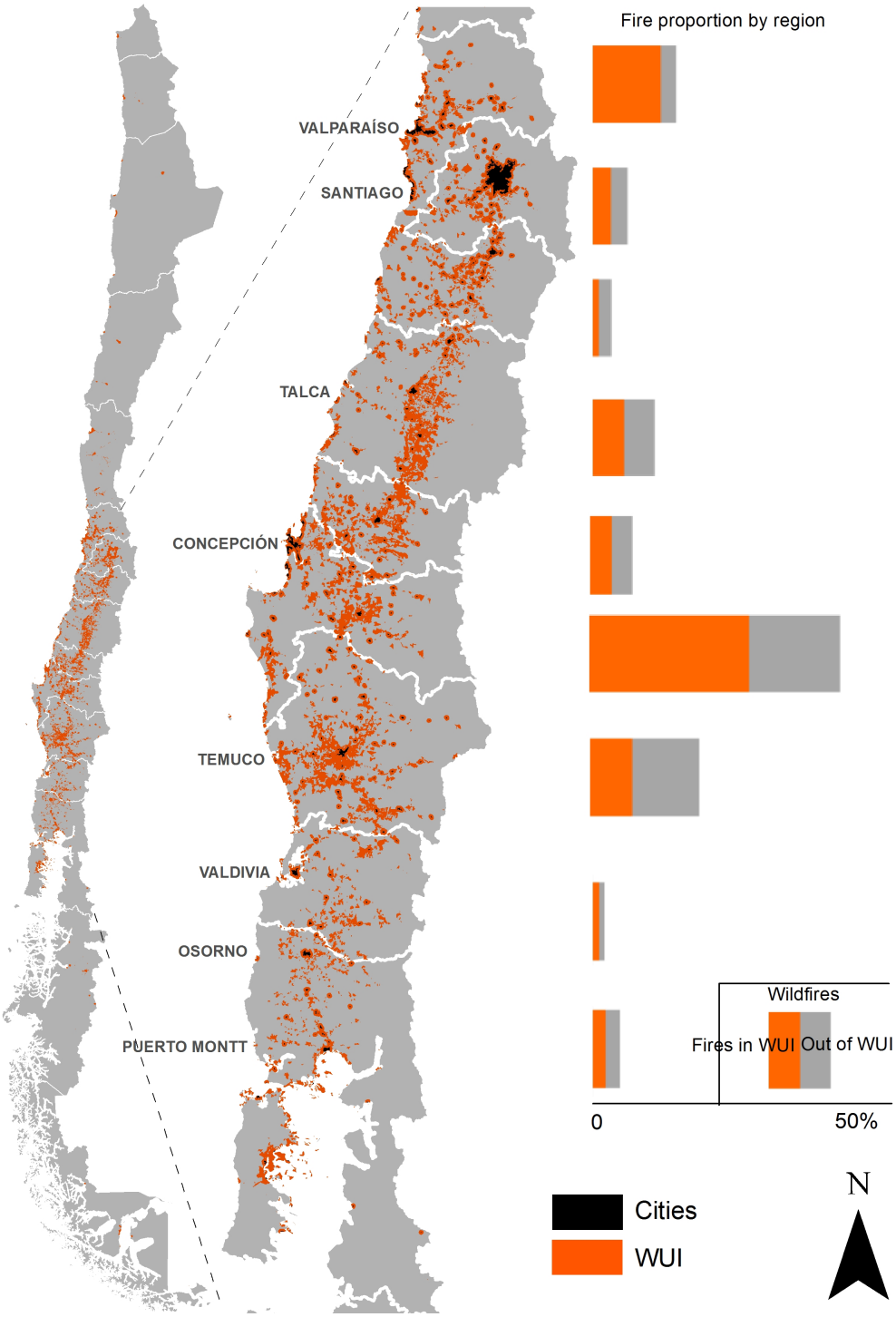


Figure 6.6: Wildland-Urban Interface (WUI) areas in Chile delineated through the thresholds defined at national level. The relative proportion of fires by region and the proportion of fires inside and outside the proposed WUI are shown on the right side.

we identified (and quantify) a non-linear relationship between population density and fire risk which has also been shown globally (Syphard et al., 2019), and these also strongly interacted with land cover. At a national scale, more than half of fires occurrence concentrates between

-36.2°S and -38.2°S (SNZ3), but the higher population and roads density is further north this area (SNZ1). This nonlinear relation between subnational zones is consistent with the nonlinear individual effect of human activities variables observed in Figure 6.5. This result supports our approach in terms of artificial intelligence based analysis to take into account variable interaction of different nature.

We did not differentiate if fire ignition was intentional or accidental, which could potentially change the spatial pattern of risk (Syphard and Keeley, 2015). However, by including all ignition sources together, we nevertheless were able to identify the overall spatial pattern of human-caused ignitions.

Human activities variable is the primary driver of fire risk for Chile. In SNZ1, SNZ2 and SNZ3, where most of the fires concentrate (Table 6.2), cities and road area mainly in the coastal range and in the central valley at low elevation. These areas also focus other economics activities as industrial forest plantation, agriculture, and productive pastures, while in the pre-Andean and Andean range is the significant proportion of native forest with least LULCC dynamic (Miranda et al., 2015, 2017). The area where most of the fires occur in Chile is also where the higher substitution of native forest for tree plantation was concentrated in the past decades (Heilmayr et al., 2016; Miranda et al., 2017) and now a homogeneous forest plantation dominate the landscapes ((Zhao et al., 2016), Figure 6.1). The proportion of tree plantation, croplands and native forest in the landscape were the most important landscape variables for our fire risk model, and they're almost double in importance to proportion of pasture or shrublands cover. Tree plantation industry also spread over vast agricultural fields, especially in the central valley and coastal range (Miranda et al., 2015). A ten-fold increase of the nationwide exotic tree plantation area from 300.000 ha in 1974 to three million in 2018 has occurred. All these changes make that currently, central and southern Chile concentrate the extensive anthropogenic land uses (Miranda et al., 2017).

This unbalanced geographic distribution of fires justifies the multi-scale models approach to explore fire risk in each subnational zone without the weight effect of latitude in model fitting. The use of only national-scale model could cause a misinterpretation of variable importance at local scale and with a strong effect on fire geographic concentration.

6.4.2 Subnational fire risk model

The subnational models illustrated how drivers of fire risk can interact differently across bioclimatic zones (Fusco et al., 2016). That is, different combinations and interactions among climate, land cover, and human land use produced unique local effects in different regions. Although we did not explore the influence of climate variables explicitly, our sub nationals areas represented a climatic latitudinal north to south gradient from arid-semiarid to humid ombrotypes, which indirectly accounted for climatic variation (Luebert and Pliscoff, 2006). Although climatic variability is more related to total burned area than with ignition probability generally (Syphard et al., 2019), and in Chile (Gómez-González et al., 2019), the variability of importance in landscape variables here revealed an interaction between ignition and land cover and landscape composition. The importance of land cover in our model suggests that, in addition to influencing fire spread, it also affects ignition probability (Oliveira

et al., 2012), which could be a function of flammability juxtaposed with the distribution of human presence. Even patterns of lightning ignitions have been shown to be sensitive to fuels, climate, and topography (Narayanaraj and Wimberly, 2012). This may be because some land covers produce more flammable material, likes small branches and dry leaves, or are associated with different fuel conditions governing flammability (Blackhall and Raffaele, 2019). The unique effects of different spatial contexts on different fire patterns may help to better delineate the most vulnerable portions of the WUI.

6.4.3 WUI fire risk thresholds

A precise definition of WUI or the management of its components may have a strong effect on hazard mitigation (Spyratos et al., 2007). The WUI is commonly mapped by the establishment of certain thresholds of housing density, vegetation cover, and distance between prone ecosystem and people, whose intersection generates value-at-risk areas for human development (Radeloff et al., 2005; Glickman and Babbitt, 2001). This threshold is usually predefined as an adaptation of standard definition — e.g. (Radeloff et al., 2005) — and independently from fire risk, resulting in an inconsistent description through literature (Johnston 2016). Our approach to delineating thresholds based on fire risk advances the potential for a WUI definition that includes the combination of fire ignition and spread with vulnerability to fire (Chuvieco et al., 2014).

Our machine learning modeling approach enabled us to identify thresholds for common variables that influence fire risk at the WUI. In particular, these proposed thresholds indicated the value of each variable where the fire risk, defined by Eq. 6.2; increased, stabilized, or decreased (see Figure 6.5). Although these thresholds are currently independent of the value at risk to be protected, they could be applied to define the WUI in different scenarios depending on prevention, management, or landscape planning strategy. For example, the national model determined that the main increase in fire probability occurred between zero and 5.5 housing units/km2 and then stabilized (see Figure 6.5c). Then, a land planner who considers the risk of fire and faces the problem where to build (or relocate homes), could discriminate his decision through this threshold.

National and subnational fire risk thresholds

We found a positive effect on fire probability occurrence when tree plantation, agricultural, pastureland and shrublands increase, and a decreasing probability with a rising native forest cover proportion. One of the most common metrics used to define the WUI is the distance to large patches of vegetation. However, our results suggest that that distance is not necessarily applicable for all kind of vegetation because we found different directions and magnitudes of effect depending on the type of land cover (Johnston and Flannigan, 2018). This is consistent with the fact that different vegetation types tend to be associated with differences in fire regimes (Wells et al., 2004) and that the patterns of fire vary across different resource gradients (Bradstock, 2010). Different land cover types may reflect those gradients differently. For instance, tree plantations are associated with hazardous fuel conditions highly associated

with wildfire activity (de la Barrera et al., 2018b; Bowman et al., 2019; Gómez-González et al., 2019; McWethy et al., 2018).

From 20% of the land cover by planting trees, the risk of fire does not present significant changes in the slope of the curve (see Figure 6.5c). This result is consistent with the effect of large patches on risk. For agricultural fields and shrublands these effects is not continually increasing with cover. The opposite case is native forest that with cover increment, fire probability decrease, which was also observed by McWethy et al. (2018). This result should not be seen as that native forest do not burn; however the effect of big size continuous patches of fast-growing, high-density trees planted for commercial purposes generate environmental condition that promotes fire ignition and spread (Gómez-González et al., 2018). These patterns are consistent across subnational zones for tree plantation, but for native forest we can see a low fraction cover effect on increasing fire probability (see Figure 6.5d), suggesting that fragmented, isolated or degraded patches of native forest, could be prone to fire than large continuous patches, especially in the drier areas of the north.

At a national scale, we observed a high fire risk probability up to a threshold of 1.2 km, which is half of the 2.4 km threshold used by Radeloff et al. (2005). The difference is that they proposed a distance to where a firebrand can travel from a wildfire and threaten human-populated areas (Glickman and Babbitt, 2001), but our threshold may better differentiate the areas where fire risk is highest . Also at the national scale we found a stabilizing threshold at 5.6 housing unit/km2, which is similar to the 6.17 housing unit/km2 proposed by Glickman and Babbitt (2001) as the value-at-risk threshold. At a broad scale, structure loss has been positively associated with areas mapped at the WUI, with most other structure loss occurring at densities too low to make it into the definition of WUI (Kramer et al., 2018). However, at the subnational level this threshold could be variable depending on population density areas. Syphard et al. (2019) also found varying housing densities where structure loss was highest depending on the region. This pattern reveals the need for local definition of values at risk considering local density population and how it interacts with other factors such as land cover.

Until now, there has been little effort at defining an empirical definition of the WUI based on different fire risk thresholds (Johnston and Flannigan, 2018), and little progress in exploring multi-dimensional definitions of thresholds based on the spatially explicit interaction among drivers. There have been recommendations for thresholds to be flexible or probabilistic to for planning purposes (Amato et al., 2018).

The results of this work revealed a potentially dangerous combination of human infrastructure and forest plantation cover. In light of other evidence of fire spread in this type of land cover (Gómez-González et al., 2018), particularly under extreme conditions like those seen in 2017 (de la Barrera et al., 2018b; Bowman et al., 2019), these areas may be particularly vulnerable in the future, especially if population increases in areas with continuous forest plantation cover. This combined effect on fire risk needs further exploration to obtain the n-dimensional optimal territorial thresholds.

6.5 Conclusions

We present the first methodological approach that can be used to delineate the WUI based on a fire risk assessment. To achieve this, we developed a geographical framework to model the most prominent drivers of fire risk and their interactions. Through this, we defined spatial explicit thresholds for mapping the WUI.

Overall, fire risk was higher near human infrastructure (cities, roads, and houses), at both the national and sub-national scales, which supports the general influence of human activities on the fire regime. However, there was also significant local variation in fire risk depending on land cover, the latter being consistent, with geographical variability through a latitudinal gradient in the national territory.

Our approach was tested using south-central Chile as a study model since fire activity and subsequent human impacts have increased in this area in recent years (González et al., 2018). In 2017, Chile experienced the most active wildfire season in history with ∼530,000 ha burned, representing ten-fold the national average (55,000 ha/yr). As a result, different environmental (e.g., air pollution, endemic forest loss) and economic losses were reported, including eleven fatalities and 3,000 houses lost (de la Barrera et al., 2018b; Bowman et al., 2019) associated to WUI. This unprecedented event demonstrates that wildfires are becoming an essential Chilean citizen security issue, which will require a better understanding of fire risk and an evidence-based definition of WUI.

We have presented a methodology that can be used to identify both general thresholds for large-scale applications as well as local thresholds for defining the Wildland Urban Interface (WUI) both operationally and empirically level. AI algorithms, in particular Bagged Decision Tree, support fire risk modelling using a high number of variables, capable of detecting different types of interactions,, for defining thresholds that may allow for a realistic interpretation of spatially complex processes.

Chapter 7

General conclusions

Developing optimal adaptive management plans for landscapes is still an open field of research. Two questions remain, how to address climate change and the spatial integration of the spatially explicit nature of fire risk. Uncertainty and risk play an important role in the development of forest management planning (Yousefpour et al., 2012). However, nowadays few tackle the difficult land-use issue, of where and how humans choose to build their communities in the first place. The relationship between the increase in fires and climate change intensifies the need for a paradigm shift (Moritz et al., 2014). Unless people understand and plan for fire as an inevitable and natural process, it will continue to have serious consequences for both social and ecological systems.

In this thesis we try to contribute to the aforementioned. We have addressed the problem of forest fires in two ways: i) through *Forest Fuel Management*, and ii) through the inclusion of fire risk to define the *Wildland-Urban Interface (WUI)*. In the first case, we developed an integrated simulation and optimization system to locate fuel treatments and evaluate their effectiveness in order to minimize the expected losses from future fires. In the second, from the quantification of the fire risk based on anthropogenic and geographical variables; through machine learning methods, we determine critical thresholds in order to adequately delineate the high risk areas.

Our two approaches, although developed separately, could be integrated as shown in Figure 7.1. In STAGE I, there is a first module where a cell is selected with the purpose of starting a fire in the study landscape. However, this selection could eventually be made using the probability of fire occurrence, using the method proposed in Chapter 6. The latter would give us new opportunities for analysis inside and outside the WUI, such as the determination of conditional probabilities: What is the probability that a cell will burn in the landscape, given that there are such probabilities of fire occurrences?, or what is the expected fire damage, conditioned to these probabilities? This is an interesting question because places with a high risk of occurrence, does not imply that large damages are expected (referring only to economic damage), mainly because it depends on the continuity of the fuel, topography, access to fire fighting, among many other factors.

Figure 7.1 shows our vision of the system we want to develop. In STAGE I, an ignition

module, a module for generating climate scenarios and the fire growth model (Cell2Fire) are integrated, to obtain risk measures; and this way, to address the location of the fuel treatment in the next stage. In the STAGE II: formulating an explicit optimization model (*PVP problem*) to locate harvest decisions and evaluate its effectiveness. Now, this is a simplified way of dealing with the problem, because in complex landscapes, where there are not only forest plantations, other actions can be taken to reduce fuel and therefore minimize damage from future fires. In addition, for a planning horizon of more than one year, we assumed that the vegetation does not grow. Therefore, this thesis proposes new challenges to face: How to include other fuel treatment? How to integrate growth models for vegetation to our DSS? How do we generate representative fire scenarios to include them in stochastic optimization models?.

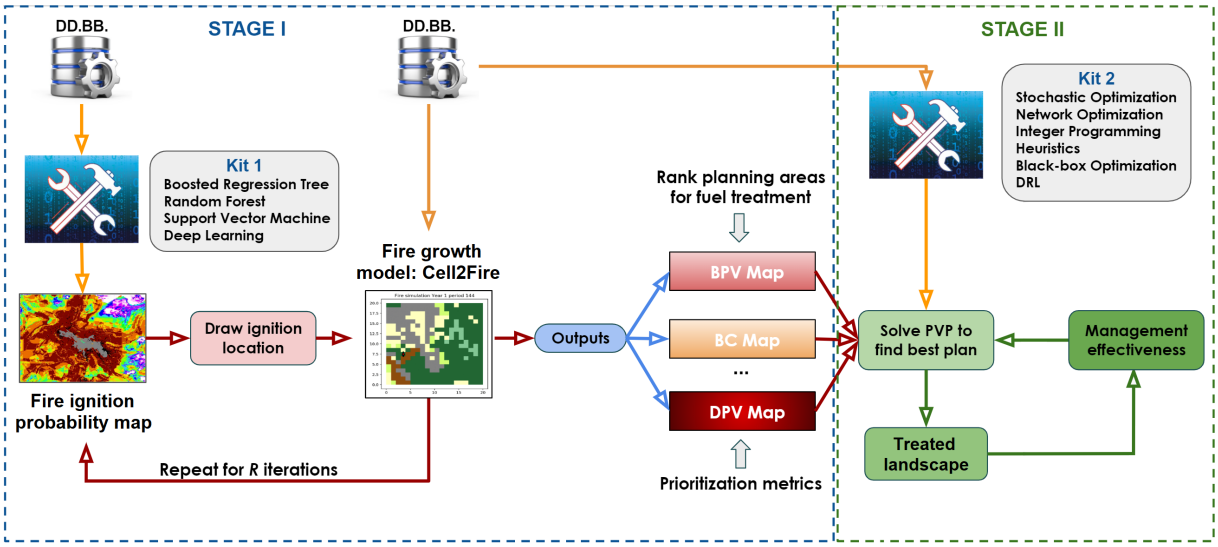


Figure 7.1: Decision support system schematic (vision)

Bibliography

M. A. Acuna, C. D. Palma, W. Cui, D. L. Martell, and A. Weintraub. Integrated spatial fire and forest management planning. *Canadian Journal of Forest Research*, 40(12):2370–2383, 2010. doi: 10.1139/X10-151. URL https://doi.org/10.1139/X10-151.

J. K. Agee and C. N. Skinner. Basic principles of forest fuel reduction treatments. *Forest ecology and management*, 211(1-2):83–96, 2005.

A. A. Ager, M. A. Finney, B. K. Kerns, and H. Maffei. Modeling wildfire risk to northern spotted owl (strix occidentalis caurina) habitat in central oregon, usa. *Forest Ecology and Management*, 246(1):45–56, 2007.

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

S. Alarie, C. Audet, V. Garnier, S. Le Digabel, and L.-A. Leclaire. Snow water equivalent estimation using blackbox optimization. *Pac J Optim*, 9(1):1–21, 2013.

P. Alberto, F. Nogueira, H. Rocha, and L. N. Vicente. Pattern search methods for user-provided points: Application to molecular geometry problems. *SIAM Journal on Optimization*, 14(4):1216–1236, 2004.

F. J. Alcasena, A. A. Ager, M. Salis, M. A. Day, and C. Vega-Garcia. Optimizing prescribed fire allocation for managing fire risk in central catalonia. *Science of the total environment*, 621:872–885, 2018.

M. E. Alexander, B. J. Stocks, and B. D. Lawson. The Canadian forest fire danger rating system. *Initial Attack*, 1996:6–9, 1996.

A. Alexandridis, D. Vakalis, C. I. Siettos, and G. V. Bafas. A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990. *Applied Mathematics and Computation*, 204(1):191–201, 2008a.

A. Alexandridis, D. Vakalis, C. I. Siettos, and G. V. Bafas. A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990. *Applied Mathematics and Computation*, 204(1):191–201, 2008b.

R. M. Almeida and E. E. Macau. Stochastic cellular automata model for wildland fire

spread dynamics. In *Journal of Physics: Conference Series*, volume 285, page 012038. IOP Publishing, 2011.

E. Álvarez-Miranda, I. Ljubić, and P. Mutzel. The maximum weight connected subgraph problem. In *Facets of Combinatorial Optimization*, pages 245–270. Springer, 2013.

F. Amato, M. Tonini, B. Murgante, and M. Kanevski. Fuzzy definition of rural urban interface: An application based on land use change scenarios in portugal. *Environmental Modelling & Software*, 104:171–187, 2018.

D. H. Anderson, E. A. Catchpole, N. J. De Mestre, and T. Parkes. Modelling the spread of grass fires. *The ANZIAM Journal*, 23(4):451–466, 1982.

H. E. Anderson. Aids to determining fuel models for estimating fire behavior. *US Dep. Agric., For. Serv., Intermt. For. Range Exp. Stn., Ogden, UT Gen Tech. Rep. INT*, 122: 22, 1982.

B. Arca, T. Ghisu, M. Casula, M. Salis, and P. Duce. A web-based wildfire simulator for operational applications. *International Journal of Wildland Fire*, 28(2):99–112, 2019.

P. M. Attiwill and M. A. Adams. Mega-fires, inquiries and politics in the eucalypt forests of victoria, south-eastern australia. *Forest Ecology and Management*, 294:45–53, 2013.

C. Audet and W. Hare. *Derivative-free and blackbox optimization*. Springer, 2017.

C. Audet and D. Orban. Finding optimal algorithmic parameters using derivative-free optimization. *SIAM Journal on Optimization*, 17(3):642–664, 2006.

A. Badia, M. Pallares-Barbera, N. Valldeperas, and M. Gisbert. Wildfires in the wildland-urban interface in catalonia: Vulnerability analysis based on land use and land cover change. *Science of the total environment*, 673:184–196, 2019.

J. K. Balch, P. M. Brando, D. C. Nepstad, M. T. Coe, D. Silvério, T. J. Massad, E. A. Davidson, P. Lefebvre, C. Oliveira-Santos, W. Rocha, et al. The susceptibility of south-eastern amazon forests to fire: insights from a large-scale burn experiment. *Bioscience*, 65 (9):893–905, 2015.

J. K. Balch, B. A. Bradley, J. T. Abatzoglou, R. C. Nagy, E. J. Fusco, and A. L. Mahood. Human-started wildfires expand the fire niche across the united states. *Proceedings of the National Academy of Sciences*, 114(11):2946–2951, 2017.

A. Bar-Massada, V. C. Radeloff, and S. I. Stewart. Biotic and abiotic effects of human settlements in the wildland–urban interface. *Bioscience*, 64(5):429–437, 2014.

K. Barnett, S. Parks, C. Miller, and H. Naughton. Beyond fuel treatment effectiveness: Characterizing interactions between fire and treatments in the us. *Forests*, 7(10):237, 2016.

T. Begin, B. Baynat, F. Sourd, and A. Brandwajn. A dfo technique to calibrate queueing

models. *Computers & Operations Research*, 37(2):273–281, 2010.

P. Bettinger. prototype method fur integrating spatially-reft rented wildfires into a tactical forest planning model. *Research Journal of Forestry*, 3(1):8–22, 2009.

M. Blackhall and E. Raffaele. Flammability of patagonian invaders and natives: When exotic plant species affect live fine fuel ignitability in wildland-urban interfaces. *Landscape and Urban Planning*, 189:1–10, 2019.

A. Booker, P. Frank, J. Dennis, Jr, D. Moore, and D. Serafini. Managing surrogate objectives to optimize a helicopter rotor design-further experiments. In *7th AIAA/USAF/-NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, page 4717, 1998a.

A. J. Booker, J. Dennis, P. D. Frank, D. B. Serafini, and V. Torczon. Optimization using surrogate objectives on a helicopter test example. In *Computational Methods for Optimal Design and Control*, pages 49–58. Springer, 1998b.

D. M. Bowman, A. Moreira-Muñoz, C. A. Kolden, R. O. Chávez, A. A. Muñoz, F. Salinas, Á. González-Reyes, R. Rocco, F. de la Barrera, G. J. Williamson, et al. Human–environmental drivers and impacts of the globally extreme 2017 chilean fires. *Ambio*, 48 (4):350–362, 2019.

D. Boychuk, W. J. Braun, R. J. Kulperger, Z. L. Kroughly, and D. A. Stanford. A stochastic fire growth model. *Environmental and Ecological Statistics*, 16:133 – 151, 2009a.

D. Boychuk, W. J. Braun, R. J. Kulperger, Z. L. Krougly, and D. A. Stanford. A stochastic forest fire growth model. *Environmental and Ecological Statistics*, 16(2):133–151, 2009b.

R. A. Bradstock. A biogeographic model of fire regimes in australia: current and future implications. *Global Ecology and Biogeography*, 19(2):145–158, 2010.

U. Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

L. Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

L. Breiman et al. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16(3):199–231, 2001.

D. E. Calkin, J. D. Cohen, M. A. Finney, and M. P. Thompson. How risk management can prevent future wildfire disasters in the wildland-urban interface. *Proceedings of the National Academy of Sciences*, 111(2):746–751, 2014.

M. Calviño-Cancela, M. L. Chas-Amil, E. D. García-Martínez, and J. Touza. Wildfire risk associated with different vegetation types within and outside wildland-urban interfaces. *Forest Ecology and Management*, 372:1–9, 2016.

A. Cardil, S. Monedero, C. A. Silva, and J. Ramirez. Adjusting the rate of spread of fire

simulations in real-time. *Ecological Modelling*, 395:39–44, 2019.

R. Carvajal, M. Constantino, M. Goycoolea, J. P. Vielma, and A. Weintraub. Imposing connectivity constraints in forest planning models. *Operations Research*, 61(4):824–836, 2013.

W. Chung. Optimizing fuel treatments to reduce wildland fire risk. *Current Forestry Reports*, 1(1):44–51, 2015.

E. Chuvieco, I. Aguado, S. Jurdao, M. L. Pettinari, M. Yebra, J. Salas, S. Hantson, J. de la Riva, P. Ibarra, M. Rodrigues, et al. Integrating geospatial information into fire risk assessment. *International journal of wildland fire*, 23(5):606–619, 2014.

A. R. Conn and P. L. Toint. An algorithm using quadratic interpolation for unconstrained derivative free optimization. In *Nonlinear optimization and applications*, pages 27–47. Springer, 1996.

A. R. Conn, K. Scheinberg, and P. L. Toint. Recent progress in unconstrained nonlinear optimization without derivatives. *Mathematical programming*, 79(1-3):397, 1997.

A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to derivative-free optimization*, volume 8. Siam, 2009.

M. G. Cruz. Monte carlo-based ensemble method for prediction of grassland fire spread. *International Journal of Wildland Fire*, 19(4):521–530, 2010.

M. G. Cruz and M. E. Alexander. Uncertainty associated with model predictions of surface and crown fire rates of spread. *Environmental Modelling & Software*, 47:16–28, 2013.

T. Curt, T. Fréjaville, and S. Lahaye. Modelling the spatial patterns of ignition causes and fire regime features in southern france: implications for fire prevention policy. *International Journal of Wildland Fire*, 25(7):785–796, 2016a.

T. Curt, T. Fréjaville, and S. Lahaye. Modelling the spatial patterns of ignition causes and fire regime features in southern france: implications for fire prevention policy. *International Journal of Wildland Fire*, 25(7):785–796, 2016b.

L. Dagum and R. Menon. Openmp: an industry standard api for shared-memory programming. *IEEE computational science and engineering*, 5(1):46–55, 1998.

F. de la Barrera, F. Barraza, P. Favier, V. Ruiz, and J. Quense. Megafires in chile 2017: Monitoring multiscale environmental impacts of burned ecosystems. *Science of the total environment*, 637:1526–1536, 2018a.

F. de la Barrera, F. Barraza, P. Favier, V. Ruiz, and J. Quense. Megafires in chile 2017: Monitoring multiscale environmental impacts of burned ecosystems. *Science of the total environment*, 637:1526–1536, 2018b.

T. Dietterich. Ensemble methods in machine learning. in 'multiple classifier systems'.(eds, j.

kittler and f. roli.) pp. 1–15, 2000.

G. K. Dillon, Z. A. Holden, P. Morgan, M. A. Crimmins, E. K. Heyerdahl, and C. H. Luce. Both topography and climate affected forest and woodland burn severity in two regions of the western us, 1984 to 2006. *Ecosphere*, 2(12):1–33, 2011.

T. Duff, J. Cawson, B. Cirulis, P. Nyman, G. Sheridan, and K. Tolhurst. Conditional performance evaluation: Using wildfire observations for systematic fire simulator development. *Forests*, 9(4):189, 2018.

D. Dunne and R. Mcsweeney. Media reaction: Amazon fires and climate change (accessed 4 october 2019, Aug 2019. URL https://www.carbonbrief.org/media-reaction-amazon-fires-and-climate-change.

W. P. Edwards. The new normal: Living with wildland fire. *Natural Resources & Environment*, 33(3):30–33, 2019.

J. Elith, J. R. Leathwick, and T. Hastie. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813, 2008.

F. Ferreira-Leite, A. Bento-Gonçalves, A. Vieira, and L. da Vinha. Mega-fires around the world: A literature review. *Wildland fires: a worldwide reality'.(Eds AJ Bento Gonçalves, AA Batista Vieira) pp*, pages 15–33, 2015.

M. A. Finney. Farsite: Fire area simulator-model development and evaluation. *Res. Pap. RMRS-RP-4, Revised 2004. Ogden, UT: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 47 p.*, 4, 1998.

M. A. Finney. Design of regular landscape fuel treatment patterns for modifying fire growth and behavior. *Forest Science*, 47(2):219–228, 2001.

M. A. Finney. Farsite: fire area simulator — model development and evaluation. *Rev. ed. US Dept Agric., ForServ. Res. Pap. RMRS-RP-4*, page 48, 2004.

M. A. Finney. The challenge of quantitative risk analysis for wildland fire. *Forest Ecology and Management*, 211:97 – 108, 2005.

M. A. Finney. An overview of flammap fire modeling capabilities. In *In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28-30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 213-220*, volume 41, 2006.

M. A. Finney, I. C. Grenfell, C. W. McHugh, R. C. Seli, D. Trethewey, R. D. Stratton, and S. Brittain. A method for ensemble wildland fire simulation. *Environmental Modeling & Assessment*, 16(2):153–167, 2011a.

M. A. Finney, C. W. McHugh, I. C. Grenfell, K. L. Riley, and K. C. Short. A simulation of probabilistic wildfire risk components for the continental united states. *Stochastic*

Environmental Research and Risk Assessment, 25(7):973–1000, 2011b.

Forestry-Canada. Development and structure of the Canadian forest fire behavior prediction system. Technical report, Forestry Canada, Headquarters, Fire Danger Group and Science and Sustainable Development Directorate, Ottawa. Information Report ST-X-3. 64 p., 1992.

L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

E. J. Fusco, J. T. Abatzoglou, J. K. Balch, J. T. Finn, and B. A. Bradley. Quantifying the human influence on fire ignition across the western usa. *Ecological applications*, 26(8): 2390–2401, 2016.

A. Ganteaume and A. D. Syphard. Ignition sources. *Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires*, pages 1–17, 2018.

T. Ghisu, B. Arca, G. Pellizzaro, and P. Duce. An optimal cellular automata algorithm for simulating wildfire spread. *Environmental Modelling & Software*, 71:1–14, 2015.

S. Gibbens. The amazon is burning at record rates-and deforestation is to blame (accessed 4 october 2019), Aug 2019. URL https://www.nationalgeographic.com/environment/2019/08/wildfires-in-amazon-caused-by-deforestation/.

D. Glickman and B. Babbitt. Urban wildland interface communities within the vicinity of federal lands that are at high risk from wildfire. *Federal Register*, 66(3):751–777, 2001.

S. Gómez-González, F. Ojeda, and P. M. Fernandes. Portugal and chile: Longing for sustainable forestry while rising from the ashes. *Environmental Science & Policy*, 81:104–107, 2018.

S. Gómez-González, M. E. González, S. Paula, I. Díaz-Hormazábal, A. Lara, and M. Delgado-Baquerizo. Temperature and agriculture are largely associated with fire activity in central chile across different temporal periods. *Forest Ecology and Management*, 433:535–543, 2019.

J. R. González, M. Palahí, A. Trasobares, and T. Pukkala. A fire probability model for forest stands in catalonia (north-east spain). *Annals of Forest Science*, 63(2):169–176, 2006.

M. E. González, S. Gómez-González, A. Lara, R. Garreaud, and I. Díaz-Hormazábal. The 2010–2015 megadrought and its influence on the fire regime in central and south-central chile. *Ecosphere*, 9(8):e02300, 2018.

J.-R. González-Olabarria and T. Pukkala. Integrating fire risk considerations in landscape-level forest planning. *Forest Ecology and Management*, 261(2):278–287, 2011.

M. E. Gray and B. G. Dickson. Applying fire connectivity and centrality measures to mitigate the cheatgrass-fire cycle in the arid west, usa. *Landscape ecology*, 31(8):1681–1696, 2016.

M. Hajian, E. Melachrinoudis, and P. Kubat. Modeling wildfire propagation with the stochastic shortest path: A fast simulation approach. *Environmental modelling & software*, 82: 73–88, 2016.

L. Han and W. Braun. Dionysus: a stochastic fire growth scenario generator. *Environmetrics*, 25(6):431–442, 2014. doi: 10.1002/env.2236. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2236.

L. Han and W. John Braun. Dionysus: a stochastic fire growth scenario generator. *Environmetrics*, 25(6):431–442, 2014.

C. C. Hardy. Wildland fire hazard and risk: problems, definitions, and context. *Forest ecology and management*, 211(1-2):73–82, 2005.

W. Hare. Using derivative free optimization for constrained parameter selection in a home and community care forecasting model. In *International perspectives on operations research and health care, Proceedings of the 34th meeting of the EURO working group on operational research applied to health sciences*, pages 61–73, 2010.

R. Heilmayr, C. Echeverría, R. Fuentes, and E. F. Lambin. A plantation-dominated forest transition in chile. *Applied Geography*, 75:71–82, 2016.

K. Hirsch, V. Kafka, C. Tymstra, R. McAlpine, B. Hawkes, H. Stegehuis, S. Quintilio, S. Gauthier, and K. Peck. Fire-smart forest management: a pragmatic approach to sustainable forest management in fire-dominated ecosystems. *The Forestry Chronicle*, 77(2):357–363, 2001a.

K. Hirsch, V. Kafka, C. Tymstra, R. McAlpine, B. Hawkes, H. Stegehuis, S. Quintilio, S. Gauthier, and K. Peck. Fire-smart forest management: A pragmatic approach to sustainable forest management in fire-dominated ecosystems. *The Forestry Chronicle*, 77(2):357–363, 2001b.

K. G. Hirsch. Canadian forest fire behavior prediction (fbp) system: user's guide. nat. resour. can., can. for. serv., northwest reg., north. for. cent., edmonton, alberta. *Spec. Rep*, 7:121, 1996.

K. G. Hirsch et al. *Canadian forest fire behavior prediction (FBP) system: user's guide*, volume 7. 1996.

INPE. Área queimada - programa queimadas (accessed 4 october 2019), Aug 2019. URL http://queimadas.dgi.inpe.br/queimadas/portal-static/situacao-atual/.

D. S. Johnson. The np-completeness column: an ongoing guide. *Journal of Algorithms*, 6(3): 434–451, 1985.

S. G. Johnson. The nlopt nonlinear-optimization package, 2014.

L. M. Johnston and M. D. Flannigan. Mapping canadian wildland fire interface areas. *International journal of wildland fire*, 27(1):1–14, 2018.

J. E. Keeley and A. D. Syphard. Twenty-first century california, usa, wildfires: fuel-dominated vs. wind-dominated fires. *Fire Ecology*, 15(1):24, 2019.

J.-H. Kim. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational statistics & data analysis*, 53(11):3735–3745, 2009.

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

P. Kourtz. Probability makes fire danger index more reliable. *Fire Contr Notes Washington*, 1972.

P. Kourtz, S. Nozaki, and W. G. O'Regan. Forest fires in the computer — a model to predict the perimeter location of a forest fire. *Fish Environ. Can Ottawa, ON. Inf. Rep. FF-X-*, 65, 1977.

P. H. Kourtz and W. G. O'Regan. A model for a small forest fire ... to simulate burned and burning areas for use in a detection model. *For Sci*, 17(2):163–169, 1971.

H. A. Kramer, M. H. Mockrin, P. M. Alexandre, S. I. Stewart, and V. C. Radeloff. Where wildfires destroy buildings in the us relative to the wildland–urban interface and national fire outreach programs. *International journal of wildland fire*, 27(5):329–341, 2018.

C. A. Kuhlmann, D. L. Martell, R. J.-B. Wets, and D. L. Woodruff. Generating Stochastic Ellipsoidal Forest and Wildland Fire Scar Scenarios for Strategic Forest Management Planning under Uncertainty. *Forest Science*, 61(3):494–508, 12 2014.

T. Levina, Y. Levin, J. McGill, and M. Nediak. Dynamic pricing with online learning and strategic consumers: an application of the aggregating algorithm. *Operations Research*, 57 (2):327–341, 2009.

Z. Liu, J. Yang, and H. S. He. Studying the effects of fuel treatment based on burn probability on a boreal forest landscape. *Journal of environmental management*, 115:42–52, 2013.

C. Loehle. Applying landscape principles to fire hazard reduction. *Forest Ecology and management*, 198(1-3):261–267, 2004.

F. Luebert and P. Pliscoff. *Sinopsis bioclimática y vegetacional de Chile*. Editorial Universitaria, 2006.

Z. Maditinos and C. Vassiliadis. Mega fires: can they be managed effectively? *Disaster Prevention and Management: An International Journal*, 20(1):41–52, 2011.

S. L. Manzello, K. Almand, E. Guillaume, S. Vallerent, S. Hameury, and T. Hakkarainen. Forum position paper the growing global wildland urban interface (wui) fire dilemma: Priority needs for research. *Fire safety journal*, 100, 2018.

D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal*

of the society for Industrial and Applied Mathematics, 11(2):431–441, 1963.

D. L. Martell. A review of recent forest and wildland fire management decision support systems research. *Current Forestry Reports*, 1(2):128–137, 2015.

D. Matsypura, O. A. Prokopyev, and A. Zahar. Wildfire fuel management: network-based models and optimization of prescribed burning. *European Journal of Operational Research*, 264(2):774–796, 2018.

D. B. McWethy, A. Pauchard, R. A. García, A. Holz, M. E. González, T. T. Veblen, J. Stahl, and B. Currey. Landscape drivers of recent fire activity (2001-2017) in south-central chile. *PLoS one*, 13(8):e0201195, 2018.

A. Miranda, A. Altamirano, L. Cayuela, F. Pincheira, and A. Lara. Different times, same story: Native forest loss and landscape homogenization in three physiographical areas of south-central of chile. *Applied Geography*, 60:20–28, 2015.

A. Miranda, A. Altamirano, L. Cayuela, A. Lara, and M. González. Native forest loss in the chilean biodiversity hotspot: revealing the evidence. *Regional Environmental Change*, 17 (1):285–297, 2017.

J. J. Moghaddas and L. Craggs. A fuel treatment reduces fire severity and increases suppression efficiency in a mixed conifer forest. *International Journal of Wildland Fire*, 16(6): 673–678, 2008.

M. A. Moritz, M. E. Morais, L. A. Summerell, J. Carlson, and J. Doyle. Wildfires, complexity, and highly optimized tolerance. *Proceedings of the National Academy of Sciences*, 102(50): 17912–17917, 2005.

M. A. Moritz, M.-A. Parisien, E. Batllori, M. A. Krawchuk, J. Van Dorn, D. J. Ganz, and K. Hayhoe. Climate change and disruptions to global fire activity. *Ecosphere*, 3(6):1–22, 2012.

M. A. Moritz, E. Batllori, R. A. Bradstock, A. M. Gill, J. Handmer, P. F. Hessburg, J. Leonard, S. McCaffrey, D. C. Odion, T. Schoennagel, et al. Learning to coexist with wildfire. *Nature*, 515(7525):58, 2014.

G. Narayanaraj and M. C. Wimberly. Influences of forest roads on the spatial patterns of human-and lightning-caused wildfire ignitions. *Applied Geography*, 32(2):878–888, 2012.

J. A. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

R. Oeuvray. Trust-region methods based on radial basis functions with application to biomedical imaging. Technical report, EPFL, 2005.

J. D. Olden, J. J. Lawler, and N. L. Poff. Machine learning methods without tears: a primer for ecologists. *The Quarterly review of biology*, 83(2):171–193, 2008.

S. Oliveira, F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J. M. Pereira. Modeling spatial patterns of fire occurrence in mediterranean europe using multiple regression and random forest. *Forest Ecology and Management*, 275:117–129, 2012.

D. Opitz and R. Maclin. Popular ensemble methods: An empirical study. *Journal of artificial intelligence research*, 11:169–198, 1999.

W. G. O'Regan, S. Nozaki, and P. Kourtz. A method for using directional rates of spread to predict forest fire configurations. western states section. *The Combustion Institute*, WSS/CI Paper 73-17, 1973.

W. G. O'Regan, P. Kourtz, and S. Nozaki. Bias in the contagion analog to fire spread. *For Sci*, 2(1):61–68, 1976.

C. Pais, J. Carrasco, D. L. Martell, A. Weintraub, and D. L. Woodruff. Cell2fire: A cell based forest fire growth model. *arXiv preprint arXiv:1905.09317*, 2019.

C. D. Palma, W. Cui, D. L. Martell, D. Robak, and A. Weintraub. Assessing the impact of stand-level harvests on the flammability of forest landscapes. *International Journal of Wildland Fire*, 16(5):584–592, 2007a.

C. D. Palma, W. Cui, D. L. Martell, D. Robak, and A. Weintraub. Assessing the impact of stand-level harvests on the flammability of forest landscapes. *International Journal of Wildland Fire*, 16(5):584–592, 2007b.

G. D. Papadopoulos and F.-N. Pavlidou. A comparative review on wild fire simulators. *IEEE Syst. J.*, 5:233–243, 2011a.

G. D. Papadopoulos and F.-N. Pavlidou. A comparative review on wildfire simulators. *IEEE systems Journal*, 5(2):233–243, 2011b.

M. Parisien, V. Kafka, K. Hirsch, B. Todd, S. Lavoie, and P. Maczek. Using the burn-p3 simulation model to map wildfire susceptibility. *Canadian Forest Service Report NOR-X-405*, 2005a.

M.-A. Parisien, V. Kafka, K. Hirsch, J. Todd, S. Lavoie, P. Maczek, et al. Mapping wildfire susceptibility with the burn-p3 simulation model. *Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Information report NOR-X-405.(Edmonton, AB)*, 2005b.

M. A. Parisien, V. G. Kafka, K. G. Hirsch, J. B. Todd, S. G. Lavoie, and M. PD. Using the burn-p3 simulation model to map wildfire susceptibility. *Information Report NOR-X-Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, AB*, 405, 2005c.

M.-A. Parisien, D. R. Junor, and V. G. Kafka. Using landscape-based decision rules to prioritize locations of fuel treatments in the boreal mixedwood of western canada. In *In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28-30 March 2006; Portland, OR. Proceedings RMRS-*

P-41. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 221-236, volume 41, 2006.

M.-A. Parisien, S. A. Parks, C. Miller, M. A. Krawchuk, M. Heathcott, and M. A. Moritz. Contributions of ignitions, fuels, and weather to the spatial patterns of burn probability of a boreal landscape. *Ecosystems*, 14(7):1141–1155, 2011.

D. J. Parsons and J. W. van Wagtendonk. Fire research and management in the sierra nevada national parks. *Science and ecosystem management in the national parks. University of Arizona Press, Tucson, Arizona, USA*, pages 25–48, 1996.

R. V. Platt. The wildland–urban interface: evaluating the definition effect. *Journal of Forestry*, 108(1):9–15, 2010.

M. J. Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in optimization and numerical analysis*, pages 51–67. Springer, 1994.

M. J. Powell. The newuoa software for unconstrained optimization without derivatives. In *Large-scale nonlinear optimization*, pages 255–297. Springer, 2006.

M. J. Powell. The bobyqa algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, pages 26–46, 2009.

V. C. Radeloff, R. B. Hammer, S. I. Stewart, J. S. Fried, S. S. Holcomb, and J. F. McKeefry. The wildland–urban interface in the united states. *Ecological applications*, 15(3):799–805, 2005.

J. Ramírez, S. Monedero, and D. Buckley. New approaches in fire simulations analysis with wildfire analyst. In *7th International Conference on Forest Fire Research*, 2011.

D. I. Rappaport, D. C. Morton, M. Longo, M. Keller, R. Dubayah, and M. N. dos Santos. Quantifying long-term changes in carbon stocks and forest structure from amazon forest degradation. *Environmental Research Letters*, 13(6):065013, 2018.

G. D. Richards. An elliptical growth model of forest fire fronts and its numerical solution. *Int. J. Numer. Math Eng*, 30:1133–1149, 1990.

G. D. Richards. The properties of elliptical wildfire growth for time dependent fuel and meteorological conditions. *Combustion science and technology*, 95(1-6):357–383, 1993.

L. M. Rios and N. V. Sahinidis. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293, 2013.

R. C. Rothermel. A mathematical model for predicting fire spread in wildland fires. *US Dep. Agric., For. Serv., Intermt. For. Range Exp Stn Ogden, UT. Res. Pap. INT*, 115:52, 1972a.

R. C. Rothermel. A mathematical model for predicting fire spread in wildland fuels. *Res. Pap. INT-115. Ogden, UT: US Department of Agriculture, Intermountain Forest and Range Experiment Station. 40 p.*, 115, 1972b.

R. C. Rothermel and G. C. Rinehart. Field procedures for verification and adjustment of fire behavior predictions. *Gen. Tech. Rep. INT-142. Ogden, UT: US Department of Agriculture, Forest Service, lntermountain Forest and Range Experiment Station. 25 p.*, 142, 1983.

S. M. Running. Is global warming causing more large wildfires? *Science*, 313:927–928, 2006. doi: 10.1126/science.

L. Russo, P. Russo, I. Evaggelidis, and C. Siettos. Complex network statistics to the design of fire breaks for the control of fire spreading. *Chemical Engineering Transactions*, 2015.

L. Russo, P. Russo, and C. I. Siettos. A complex network theory approach for the spatial distribution of fire breaks in heterogeneous forest landscapes for the control of wildland fires. *PloS one*, 11(10):e0163226, 2016.

C. Salas-Eljatib, A. Fuentes-Ramirez, T. G. Gregoire, A. Altamirano, and V. Yaitul. A study on the effects of unbalanced data when fitting logistic regression models in ecology. *Ecological Indicators*, 85:502–508, 2018.

M. Salis, M. Laconi, A. A. Ager, F. J. Alcasena, B. Arca, O. Lozano, A. F. de Oliveira, and D. Spano. Evaluating alternative fuel treatment strategies to reduce wildfire losses in a mediterranean area. *Forest Ecology and Management*, 368:207–221, 2016.

M. Salis, L. Del Giudice, B. Arca, A. A. Ager, F. Alcasena-Urdiroz, O. Lozano, V. Bacciu, D. Spano, and P. Duce. Modeling the effects of different fuel treatment mosaics on wildfire spread and behavior in a mediterranean agro-pastoral area. *Journal of environmental management*, 212:490–505, 2018.

J. C. Sanderlin and J. M. Sunderson. A simulation for wildland fire management planning support (fireman). vol. 2. *Prototype models for FIREMAN (Part II): Campaign fire evaluation Mission Research Corp. Contract 21-343, Spec. 222*, 2:249, 1975.

J. H. Scott and R. E. Burgan. Standard fire behavior fuel models: a comprehensive set for use with rothermel's surface fire spread model. *US Dep. Agric., For. Serv., Rocky Mtn. Res. Stn., Fort Collins, CO Tech. Rep. RMRS-GTR-153*, page 72, 2005.

D. J. Shinneman, M. J. Germino, D. S. Pilliod, C. L. Aldridge, N. M. Vaillant, and P. S. Coates. The ecological uncertainty of wildfire fuel breaks: examples from the sagebrush steppe. *Frontiers in Ecology and the Environment*, 2019.

W. Spendley, G. R. Hext, and F. R. Himsworth. Sequential application of simplex designs in optimisation and evolutionary operation. *Technometrics*, 4(4):441–461, 1962.

V. Spyratos, P. S. Bourgeron, and M. Ghil. Development at the wildland–urban interface and the mitigation of forest-fire risk. *Proceedings of the National Academy of Sciences*, 104

(36):14272–14276, 2007.

T. Srivas, R. A. de Callafon, D. Crawl, and I. Altintas. Data assimilation of wildfires with fuel adjustment factors in farsite using ensemble kalman filtering. *Procedia Computer Science*, 108:1572–1581, 2017.

S. M. Stein, J. Menakis, M. Carr, S. Comas, S. Stewart, H. Cleveland, L. Bramwell, and V. Radeloff. Wildfire, wildlands, and people: understanding and preparing for wildfire in the wildland-urban interface-a forests on the edge report. *Gen. Tech. Rep. RMRS-GTR-299. Fort Collins, CO. US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 36 p.*, 299, 2013.

B. Stocks and D. L. Martell. Forest fire management expenditures in canada: 1970–2013. *The Forestry Chronicle*, 92(3):298–306, 2016a.

B. J. Stocks and D. L. Martell. Forest fire management expenditures in Canada: 1970–2013. *The Forestry Chronicle*, 92(3):298–306, 2016b.

C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. Conditional variable importance for random forests. *BMC bioinformatics*, 9(1):307, 2008.

A. L. Sullivan. Wildland surface fire spread modelling, 1990–2007. 1: Physical and quasi-physical models. *International Journal of Wildland Fire*, 18(4):349–368, 2009.

A. D. Syphard and J. E. Keeley. Location, timing and extent of wildfire vary by cause of ignition. *International Journal of Wildland Fire*, 24(1):37–47, 2015.

A. D. Syphard, H. Rustigian-Romsos, M. Mann, E. Conlisk, M. A. Moritz, and D. Ackerly. The relative influence of climate and housing development on current and projected future fire patterns and structure loss across three california landscapes. *Global Environmental Change*, 56:41–55, 2019.

F. Tedim, R. Remelgado, C. Borges, S. Carvalho, and J. Martins. Exploring the occurrence of mega-fires in portugal. *Forest Ecology and Management*, 294:86–96, 2013.

M. P. Thompson and D. E. Calkin. Uncertainty and risk in wildland fire management: a review. *Journal of environmental management*, 92(8):1895–1909, 2011.

B. Todd. User documentation for the wildland fire growth model and the wildfire display program. can. for. serv., north for. *Cent., Edmonton, AB*, 37 p. Draft rep, 1999.

M. Tonini, J. Parente, and M. G. Pereira. Global assessment of rural–urban interface in portugal related to land cover changes. *Natural Hazards and Earth System Sciences*, 18 (6):1647, 2018.

J. Turkewitz. The amazon is on fire. so is central africa., Aug 2019. URL https://www.nytimes.com/2019/08/27/world/africa/congo-angola-rainforest-fires.html.

C. Tymstra, R. Bryce, B. Wotton, S. Taylor, O. Armitage, et al. Development and structure

of prometheus: the canadian wildland fire growth simulation model. *Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Information Report NOR-X-417.(Edmonton, AB)*, 2010a.

C. Tymstra, R. W. Bryce, B. M. Wotton, S. W. Taylor, and A. OB. Development and structure of Prometheus: the Canadian wildland fire growth simulation model. *Information Report NOR-X-Edmonton (AB): Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre*, 417:102, 2010b.

R. Urrutia-Jalabert, M. E. González, Á. González-Reyes, A. Lara, and R. Garreaud. Climate variability and forest fires in central and south-central chile. *Ecosphere*, 9(4):e02171, 2018.

C. Van Wagner, P. Forest, et al. Development and structure of the canadian forest fireweather index system. In *Can. For. Serv., Forestry Tech. Rep.* Citeseer, 1987.

C. E. Van Wagner and T. L. Pickett. Development and structure of the Canadian forest fire weather index system. Technical Report 35, Canadian Forest Service, Forestry (Ottawa, ON), 1987.

M. Viale and R. Garreaud. Orographic effects of the subtropical and extratropical andes on upwind precipitating clouds. *Journal of Geophysical Research: Atmospheres*, 120(10): 4962–4974, 2015.

M. L. Wells, J. F. O'leary, J. Franklin, J. Michaelsen, and D. E. McKinsey. Variations in a regional fire regime related to vegetation type in san diego county, california (usa). *Landscape Ecology*, 19(2):139–152, 2004.

A. L. Westerling. Increasing western us forest wildfire activity: sensitivity to changes in the timing of spring. *Phil. Trans. R. Soc. B*, 371(20150178), 2016.

A. L. Westerling, H. G. Hidalgo, D. R. Cayan, and S. TW. Warming and earlier spring increase western us forest wildfire activity. *Science*, 313:940–943, 2006. doi: 10.1126/science.1128834).

B. M. Wotton, C. A. Nock, and M. D. Flannigan. Forest fire occurrence and climate change in canada. *International Journal of Wildland Fire*, 19(3):253–271, 2010.

R. Yousefpour, J. B. Jacobsen, B. J. Thorsen, H. Meilby, M. Hanewinkel, and K. Oehler. A review of decision-making approaches to handle uncertainty and risk in adaptive forest management under climate change. *Annals of forest science*, 69(1):1–15, 2012.

Y. Zhang, S. Lim, and J. J. Sharples. Modelling spatial patterns of wildfire occurrence in south-eastern australia. *Geomatics, Natural Hazards and Risk*, 7(6):1800–1815, 2016.

Y. Zhao, D. Feng, L. Yu, X. Wang, Y. Chen, Y. Bai, H. J. Hernández, M. Galleguillos, C. Estades, G. S. Biging, et al. Detailed dynamic land cover mapping of chile: Accuracy improvement by integrating multi-temporal data. *Remote Sensing of Environment*, 183: 170–185, 2016.

W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image qualifty assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, April 2004.