Forecasting Wind-Driven Wildfires Using An Inverse Modelling Approach

Oriol Rios i Rubiras

Promoter: Dr. Guillermo Rein
Imperial College London

[*They see only their own shadows or the shadows of one another, which the fire throws on the opposite wall of the cave*]

Book VII of *The Republic*,
The Allegory of the Cave
Plato

Summary/Abstract

Wildfires are complex multiscale phenomena involving a series of nonlinear and transient processes. They are governed by coupled physicochemical interactions that are not yet known which pose the main obstacle when trying to model their behaviour. Thus, a technology able to forecast wildfire dynamics would lead to a paradigm shift in the response to emergencies, providing the Fire Services with essential information about the on-going fire.

The thesis at hand presents and explores a novel methodology to forecast wildfire dynamics using real time data assimilation and solving an inverse problem for the invariants- governing parameters that remain constant for a significant amount of time- with a tangent linear approach and a forward automatic differentiation. Rothermel's rate of spread theory with a perimeter expansion model based on Huygens principle are used to build up the forecasting algorithm. Its potential is investigated using synthetic data and evaluated in different wind-driven wildfire scenarios. The results show the high capacity of the method to quickly predict the location of the fire front with a positive lead time (ahead of the event). This work opens the door to a further advanced framework with more sophisticated models (e.g. pyrolysis models, CFD) keeping the computational time suitable for operativeness.

El foc forestal és un fenomen complex que involucra processos no lineals i avarca múltiples escales. El conjunt complert de les interaccions fisico-químiques que el governen resten encara desconegut, fet que suposa un obvi obstacle a l'hora de simular el seu comportament. Una tecnologia capaç de predir la dinàmica d'un incendi forestal suposaria, per tant, un canvi de paradigma per els agents involucrats en combatre'l.

La present tesis presenta i explora una peculiar metodologia per a predir el comportament d'incendis forestals utilitzant informació proporcionada en temps real i resolent el problema invers per als invariants del sistema -paràmetres que el governen i resten constants durant un determinat interval de temps- mitjançant una aproximació de tangència lineal (Tangent Linear Appoach) i una diferenciació automàtica directa. La teoria de propagació de Rothermel, conjuntament amb un model d'expansió basat en el principi de Huygens, constitueixen el nucli de l'algoritme predictiu. El seu potencial és avaluat utilitzant dades sintètiques i en diferents escenaris de foc forestals conduits pel vent. Els resultats mostren una gran capacitat del mètode per a predir ràpidament i amb prou antel·lació, la posició del front de foc i el potencial per realitzar una predicció millorada amb informació addicional.
Aquest treball obre la porta per al desenvolupament d'aquesta eina amb models de propagació més sofisticats (i.e. models de piròlisis, CFD) tot mantenint el requeriment computacional prou baix perquè esdevingui una eina operativa.

Contents

List of Figures

List of Abbreviations

- **CA**: Cellular Automata

- **CD**: Compact Disc

- **CFD**: Computational Fluid Dynamics

- **FBP**: Canadian Fire Behaviour Prediction System

- **GIS**: Geographic Information System

- **GPS**: Global Positioning System

- **LSM**: Level Set Method

- **PDE**: Partial Differential Equation

- **RANS**: Reynolds-averaged Navier–Stokes equations

- **RoS**: Rate of Spread

- **SI**: International System of Units (abbreviated SI from French: *Le Système international d'unités*)

- **TLM**: Tangent Linear Model

- **WFDS**: Wildland-urban interface Fire Dynamics Simulator

- **WUI**: Wildland-Urban Interface

- **WW**: Window Width

Chapter 1

Introduction to Wildfire Forecasting

> *An operational model is validated through its utility for application. An operationally valid model may not correctly describe the processes or phenomena; it need only identify a pattern of behaviour that is useful in some way.*
>
> – Mark Finney[17]

Wildfires are not a recent phenomenon. Their first evidence is a charcoal fossil dated to 420 million years ago[51], much time before humanity existed. From the beginning of human times, the wildfire phenomenon was considered part of the forest natural circle of life as the ignition sources were normally human independent. Rural societies learnt to manage the forest in such a way that wildfires did not reach massive dimensions and were not perceived as an actual thread. However, the wildfire perspective changed when towns and cities expanded towards the forest and the traditional forest management was abandoned. The spectacular growth of the wildland urban interface (WUI) entailed a growing wildfire hazard and a shift in the wildfire perception from a harmless natural phenomenon to a threatening large scale situation that requires large suppression tasks.

Wildfires have recently shown the potential to reach colossal dimensions that were never seen before. Fires in the Victoria region in Australia, 2009[11], or Colorado, USA, 2012, as examples, have led to a new scientific term; *mega-fires*[75], used to describe the large, highly-destructive wildfire that grows beyond any currently existing extinguishing capacity.

On top of it, the global warming worsens the future perspective of wildfire regime, increasing its trend in the number of fires and the burnt surface per year[47]. Liu et al.[34] used global models to predict that, in the near future, the forest fire potential will increase worldwide, and the period susceptible to high fire risk will enlarge from 2 to 8 months in all the regions studied. In Australia, as an example, the Intergovernmental Panel on Climate Change estimated that the likelihood of serious forest fires will increase by up to 25 percent by 2020, and 70 percent by 2050[35].

The natural question then arises; why are forest fires still beyond our control?

The answer to this question is twofold: On one hand, the scientific community has been unable to fully understand the processes involved in a wildfire and create an exhaustive theory to be the basis of the discipline (Finney et al.[17]), on the other hand, the community has not yet engineered a reliable predicting tool to radically change the way of tackling wildfires. Insight into wildfire must be considered to understand the reasons behind those deficiencies.

The wildfire phenomenon involves several processes coupled together. Starting from the fuel, it is a porous organic material which undergoes phase and radical thermal properties change during combustion. Its shape usually has thermally thick (trunk) and thermally thin (leaves) regions which implies specific heat transfer mechanisms for each part. Wildfire spreads following wind and slope directions as those enhance convection and radiation by moving hot gases downwind (increasing convective heat transfer) and tilting the flames towards adjacent fuel (increasing radiation). On a ground level, conduction also plays a role in the spread of the fire independently of wind and slope. And even advection of pines and firebrands (hot wood embers) takes part in fire spread by spotting the fire beyond conventional heat transfer distances.

Wildfire is thus a complex multi-scale phenomenon governed by non-linear scale-dependant interactions. The complexity stems from the large amount of chemical and physical processes taking place together and can produce unexpected and radical changes in the fire behaviour (as the eruptive fire phenomenon discussed by Viegas et al. 2010[73] and Thomas et al. 2012[70]). The multi-scale processes involved in a forest fire range from the millimetre scale of chemical reactions related to the pyrolysis up to a 10-km fire front that couples to the local meteorology. Heat transfer phenomena are present and coupled in all the range and the dynamics can only be accurately captured if all the scales are resolved. Thus, the scaling of the problem is extremely complicated and cannot be exclusively studied in a bench-scale experiments. Real scale data is essential to understand all the phenomena involved.

The most common way to gather data from a real wildfire is doing post fire measurements which provide information about the final fire behaviour and its effects but do not supply data from the fire development itself. In a real wildfire situation, proper measurements are hindered by the unfolding emergency, limiting the amount of available data to fire front locations (if airborne images -either in visible or infra-red range- are accessible) or qualitative information on flame height or front velocity given by fire fighting crews. To obtain more data in a real fire situation prescribed burnings have been used in the literature by Anderson et al. 1982[6], Morandini et al. 2006[43] or Cruz et al. 2009[10] together with a sensor setup. Prescribe burnings are fires intentionally ignited under tightly controlled conditions normally used in agriculture to manage forest or to prevent fires by creating fire breaks. Prescribe burnings performed for scientific purposes are the principle source of reliable data of full-scale wildfires.

Nevertheless, the scientific community posed itself the problem and started to study forest fires in 1920s when Hawley et al.[24] carried out the first work to understand the physical factors that govern them. Later, in 1949, Fons et al.[18] performed rigorous approaches on the measuring and modelling of fire. Since then, many different approaches have been explored to understand the forest fire behaviour and, ultimately, create an operational forecasting tool.

1.1 Wildfire modelling and simulation

Wildfire discipline utilises multiple approaches and focuses. It spans from modelling the long term ecological impact of a wildfire to predicting the amount and type of pollutants emitted. In this thesis we focus on the fire front locations and we revise the models that deal with front position of the fire spread. There are several ways to classify the models to structure a literature review; by the nature of the underlying equations, the variables studied, the system studied (surface fires, crown fires, spotting...), the mathematical methodology used, the principal assumptions, weather it is based on empirical correlations or its applicability. It is also necessary differentiating models from simulations; while models come up with a set of equations to capture the behaviour of the relevant processes, simulations normally use and merge parts of the pre-existing models and implement them to simulate the fire spread over a surface as well as other fire properties (fire-front intensity, flame hight, temperature...). Aside from simulators, we also consider the so-called mathematical analogous models which use a variety of mathematical strategies or concepts that have no real-world connection to the fire phenomenon itself but exhibit great resemblance and potential as a predicting tool.

For the purpose of this introduction and the relevance to this thesis, two criteria will be used to briefly review some of the existing studies and point out their principal weaknesses and strengths. The first will be based on the nature of the model underlying equations (as established by Pastor et al.[50]) whereas the second focuses on the mathematical technique used in simulators (explored by Sullivan[68]).

1.1.1 Theoretical vs empirical models

If we consider the underlying equations in the model we can identify two main groups. The theoretical models that exclusively use physically derived equations and the empirical and semi-empirical models that use experimentally or statistically correlated equations.

Theoretical models

Fons[18] obtained the first wildfire spread equation in 1946 by stating the energy conservation to a system of uniform solid particles: He found a logarithmic relation between fire spread and fuel bed temperature that was successfully validated with small-scale experiments despite the fact that he linearized the radiation contribution. After Fons' work, several eminent researchers in the combustion field such as Emmons[14] (1964), Hottel[27] (1965), Thomas[69] (1967) made theoretical contributions to the basis of the fire spread in forest fire and its phenomenology. They brought their knowledge of combustion into the field and characterised the fuel bed by constant assuming properties; its moisture content, its packing ratio and the surface area to volume ratio of its constituent particles. Thomas also explored gradient temperatures inside the fuel particles. In spite of their effort, these theoretical models, in a form of partial differential equations, were not able to capture some complicate effects as the convective heat transfer from the hot gases produced by combustion, or the geometry of the flame, and needed to be complemented with empirical correlations in order to be be compared with experiments.
The scientific contribution to physical models diminished after 1970s due to the loss of interest (and support) of the governmental organisations and thus, the models which appeared during this period are scarce and only Williams, 1982[74] contributed to the field with a relevant compendium on wildfires and laid the foundation for wildland-urban interface (WUI).

The increase in computational capacity together with the development of geographical sensing (GPS) and terrain mapping (GIS) during 1990s enabled new theoretical approaches that brought together the combustion (i.e pyrolysis), the heat transfer, the flow and turbulence (i.e Navier Stokes with either RANS or LES implementations) and in some cases even meteorological data. Examples of these theoretical models are FIRETEC[33] and WDS[42] in U.S.A. and FIRESTAR[44] in France. These news models solve the governing coupled equations using a grid or a finite-volume scheme and can cope with complex geometries and fuel and terrain heterogeneities as long as they are properly resolved by the grid size. A key aspect in differentiating these fire spread models from a generic flow model is the presence of combustion chemistry that inputs the heat into the system.

Although these models are still in the validation phase, they have been capable of simulating highly detailed fire behaviours in a complex fuel and terrain and showed great capacity for becoming a useful research tool. However, physical models have two main downsides that make them unsuitable for real time forecasting. They require an extremely large amount of detailed data (fuel properties, forest layout, initial conditions for wind and meteorological quantities) to be input as initial and boundary conditions, and they also require long computational times. As an illustrative example, Running FIRETEC in a 360mx360m domain with a spatial accuracy of 2m using a reasonable large supercomputer (64 processors) will take approximately 5 hours to give a 200s forecast[66]. It is worth mentioning that 200s is not an operative forecasting time and that 64 processors supercomputers are restricted to research centres and are beyond the current computational capacity of emergency responders.
It was exactly the need of seeking faster and more accessible models to serve as a tool for day to day forest mangers that boosted the empirical and semi-empirical models

Empirical and semi-empirical models

In 1972 Rothermel[60] developed the first semi-empirical model (see section 2.1.1 for an in-depth discussion) with the intent of providing a useful tool for land managers. Rothermel merged a simple source-sink model with laboratory-scale experiments. The model rapidly gained popularity in the U.S.A and became the basis of the National Fire Danger System as well as part of the fire behaviour simulator BEHAVE[26]. Due to its success, the model was exported to the Mediterranean forests with unsatisfactory results because the empirical correlations used did not capture the Mediterranean fuels[50].
At the same time in Australia, McArthur[41] used statistical correlations to developed fire-danger meters (in a form of circular charts) that determine main surface fire parameters. The model showed high agreement with the Eucalypt fires and was adopted as a common fire predicting and decision support tool and embedded in the SiroFire[9] simulator.
In Canada, 500 fires were studied to create the Fire Behaviour Prediction System (FBP)[65], based on statistical analysis and correlations. This model is extensively used by Canadian forest agencies becom-

ing an essential tool for forest management.

Although some of these models can be used as an operational tool, their empirical nature restricts their applicability to the specific conditions used to extract the correlations. As an example of its low versatility, figures 1.1 and 1.2 show the wind and the moisture content contribution to the rate of spread (*RoS*) for 19 different semi-empirical models developed during the period from 1990 to 2007. The large scatter and disagreement between models is due to the experimental conditions used to calibrate them and illustrates the inaccuracies of using any model out of its range.
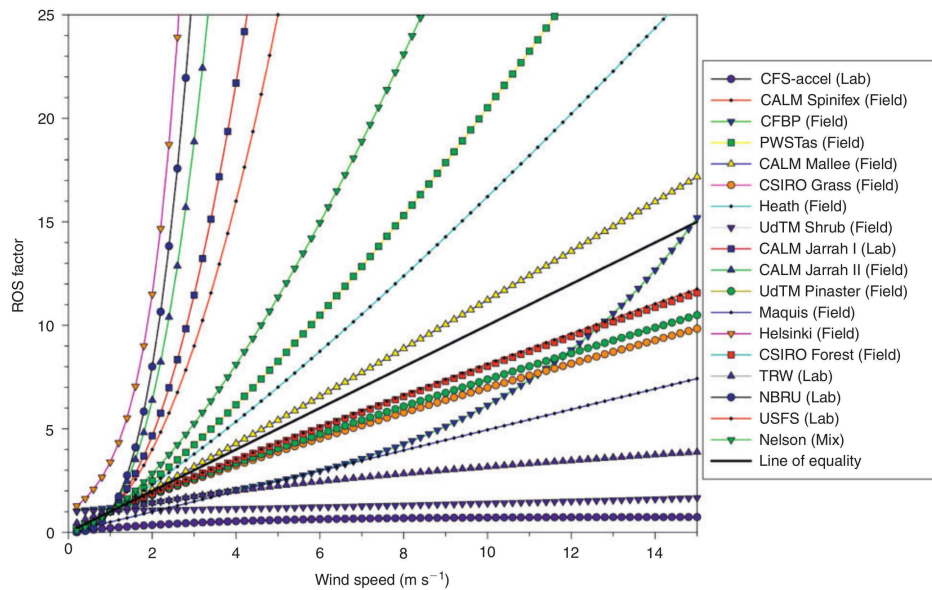
Figure 1.1: Wind factor correlations used in different empirical models. Note the large discrepancy between them. From Sullivan 2009[67]
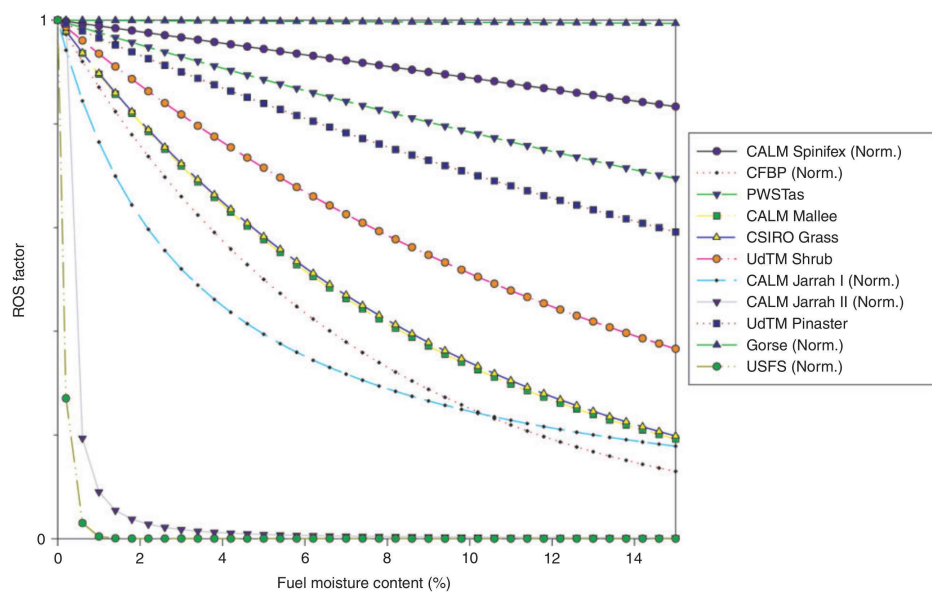
Figure 1.2: Moisture content functions used in different empirical models. From Sullivan 2009[67]

1.1.2 Wildfire spread simulators

With the increasing accessibility to computational capacity, simulation of forest fire flourished to create a workaround some of the limitations of the approaches above. Simulation models focus on the practical and the operational capability to simulate real wildfire spread (in a domain from 10x10m to 1000x1000m) and, thus, do not try to solve the intermediate interactions.

Following this idea, simulators focus on two problems. First, representing the fire perimeter in a mathematical way. Second, coming up with an algorithm to expand this perimeter. There are mainly two strategies to deal with the first; representing the fire front by adjacent cells (known in the literature as a raster implementation) or by defining it with connected points (referred to as a vector-implementation). Regarding the second -perimeter expansion-, Huygens principle, Eulerian level set equations and the near-neighbour proximity, are the principal techniques used. In the following sections, we explore the pros and cons of some of the simulators that use the aforementioned approaches.

Huygens' wavelet principle

This principle is named after Christiaan Huygens, who originally postulated it in 1920 to explain light wavefront propagation[28]. The principle puts forward the idea that any point reached by the light acts as a new source of spherical wave called wavelets. The sum of these wavelets determines the new light wavefront. By means of this principle, Huygens successfully derived the light's reaction and refraction laws.

Extrapolating it in forest fire, this principle considers every point in the fire perimeter at time t as an new ignition source that spreads during a time dt following a template shape (known sometimes as *firelet*). The corresponding fire front line at time $t + dt$ is the outer curve that envelopes the wavelets (see figure 1.4).
Sanderlin[61] introduced this concept in forest fire simulation in 1977 in his attempt to develop a fire management support system called FIREIFGTHER . However, Anderson et al.[5] were the first that actually called it Huygens' principle and started exploiting all its potential with elliptical *firelets* -fire wavelet-. Since then, it has been been extensively and successfully used to reproduce the shape of the fire front [4],[9],[15].
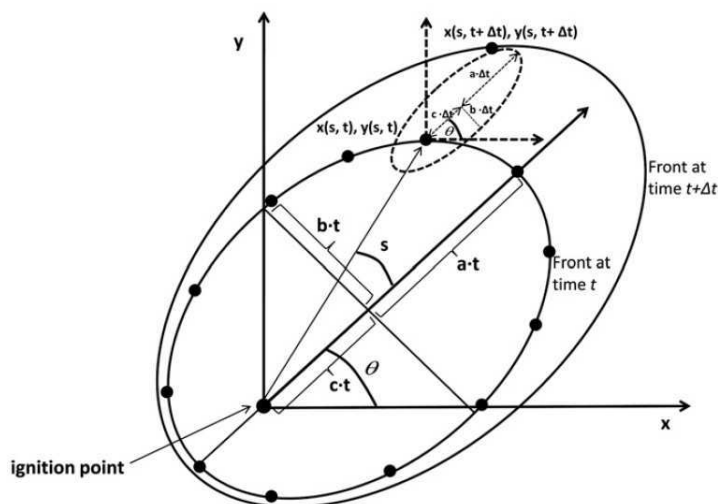
Figure 1.3: Geometrical characteristics of Huygens expansion principle using elliptical firelets. Based on this, Richard (1990)[54] derived mathematical expressions to calculate the fire front at $t + dt$. After S.L MacPhee[36].
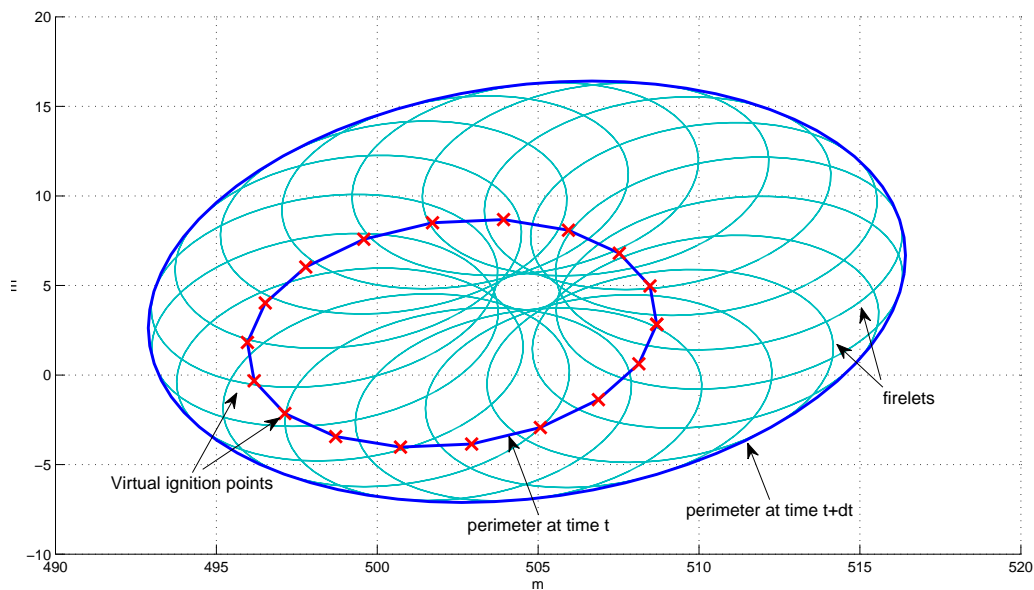
Figure 1.4: Illustration of the Huygens' principle applied to fire spread. The fire front at time t is discretised into vertices (red crosses). Each of this vertex become the point source for a local elliptical expansions (*firelets*). The curve that envelopes all the *firelets* is the new fire perimeter at time $t + dt$.

Other template shapes like double ellipses or ovals were initially explored by Peet, 1967[52] and Albini, 1976[2] but according to Alexander, 1985[3], the ellipse is the most reported shape when the fire reaches steady state spread regime. Other authors such as Anderson[5] and Green[21],[20] also studied elliptical firelets and they agreed that the shape of the front tended to tunr into an ellipse as fire grows. Green also implemented it in a simulator called IGNITE using McArthur[41] flame spread correlations.

In 1990 Richards[54] mathematically derived the set of partial differential equations needed to implement Huygens principle using ellipses. In 1993, after stating that the ellipse is the mathematically most appropriate shape to be implemented in a model[56], he revisited his equations and extended them to cope with wind direction probability and, thus, propose a set of probabilistic equations for the fire front position. In his first study Richard already encountered and pointed out one of the main problems when applying the Huygens' principle. If the fuel abruptly changes, the fire front can turn itself inside-out enclosing unburned fuel and reach physically impossible shapes (see fig. 1.5). Facing this problem, he proposed an algorithm to correct it based on the maximum curvature between adjacent fire front vertex. His model was incorporated in two of the most popular forest fire simulating tools; FARSITE[15] and PROMETHEUS[72]. The first was developed in the U.S whereas the second is part of the Canadian Fire Behaviour Prediction System (FBP), the principal tool used in this country.

Knight et al.[31] developed an alternative implementation of Huygens principle that prevents loops overlaps and high concave shapes by adding various algorithms that detect and correct the new perimeter points that have a high curvature (i.e sharply complex) or overlap previous *burnt* areas (shown in figure 1.5). His method was embedded in SIROFIRE[9], a fire simulator developed and tested in Australia which is now part of PHOENIX, a forest fire managing support program. Finney, in 2002[16], used elliptical expansion in a slightly different approach based on minimum travel time between two points in a grid (i.e. Fermat's principle used for rays propagation). The elliptical expansion was also embedded and, therefore, the enveloping fire front matches Huygens-based simulations exactly. His alternative approach, however, prevents the model from overlapping and closed loops and enables parallel computation.
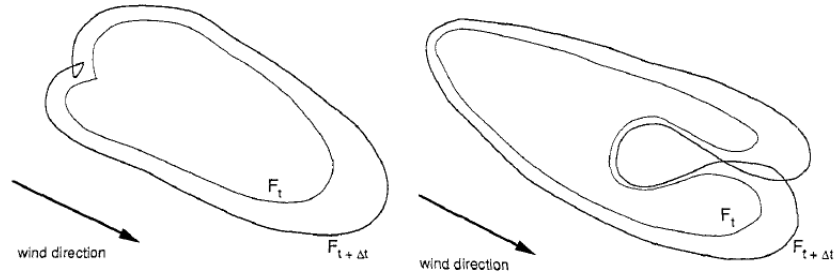
Figure 1.5: Perimeters expansions based on Huygens's principle might form loops and overlaps when the spread rate is suddenly lowered (by a change of fuel or a fire break). Knight et al.[31] proposed an algorithm to correct this effect and compute the corresponding enveloping curve. Figure from Knight et al.[31]

Eulerian Level Set Method (LSM)

The level set method (LSM) is a mathematical technique that tracks interfaces, or fronts, between separate regions. Several authors such as Mandel et al. 2009[38], Mallet et al. 2009[37], and Lautenberger, 2012[19], have recently explored its potential to generate wildfires fronts.
Given the fire front curve $\Gamma(t)$ evolving with the speed $S(x, y, t)$, the level-set method represents this curve in terms of a level-set function φ as:

$$\Gamma(t) = \{(x, y) \in \mathbb{R}^2 : \varphi(x, y, t) = 0\}$$

Then, the evolution of the curve $\Gamma(t)$ is converted into a partial differential equation for the level-set function of the form:

$$\frac{\partial \Gamma}{\partial t} + S \cdot \|\nabla \varphi\| = 0 \tag{1.1}$$

Where $\|\nabla \varphi\|$ is the Euclidean norm of the level-set function's gradient .

The geometrical interpretation of Euclidean level-set method is illustrated in figure 1.6 . Level-set function φ can be seen as the surface that generates the required curve $\Gamma(t)$ when intersected at zero level. Despite $S(x, y, t)$ being measured in the intersected plane it is related to the vertical evolution of surface φ.
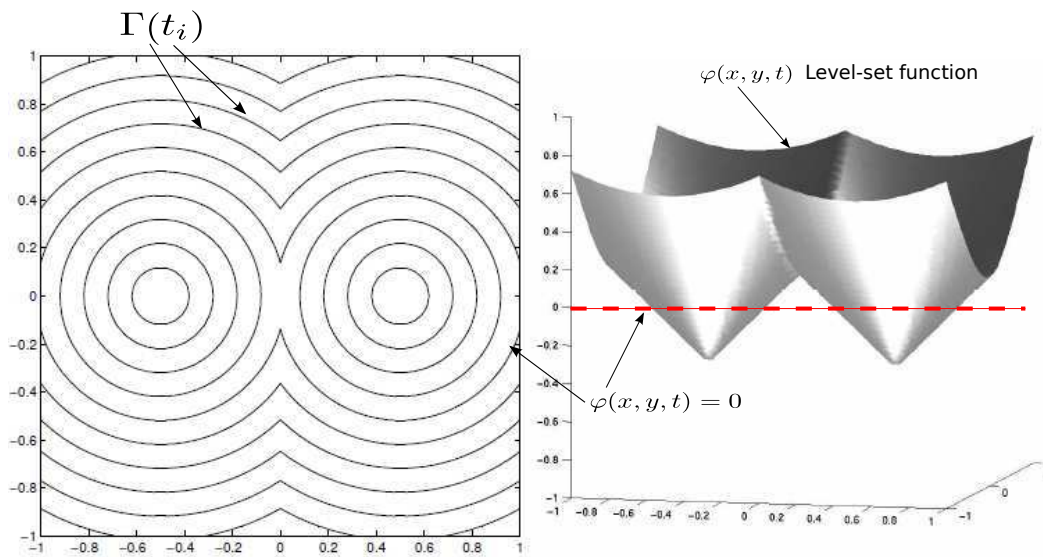
Figure 1.6: Geometrical interpretation of the level set method. On the right is the level-set function $\varphi(x, y, t)$ that moves vertically and the zero intersection line (dashed). On the left is the corresponding generated curves at different times $\Gamma(t_i)$. After Stanley Osher, 2003[71], with modifications.

With this method, Lautengerger simulated six hours of hypothetical fire fronts locations using Rothermel's rate of spread RoS as $S(x, y, t)$. He also incorporated weather data (2m height temperatures, wind speed and humidity) and topology information to enhance the reliability of the fronts. The fire fronts simulations were credible although were not compared with real data. According to the same author this Eulerian approach is more robust than the Lagrangian alternative (that tracks fire front particles instead of fire front curves) since spotting and inactive enclaves (due to fire breaks or fuel discontinuities for example) can be handled without complex logic. However, Rochoux[57] found that it is computationally more expensive in some cases. In contrast to the previously discussed Huygens' method, Eulerian approach has no direct connection to the physics and must be seen as a mathematical tool that is still under exploration and in development in the wildfire modelling field.

Raster Implementation and Cellular Automata

In those models the simulated domain is represented by a grid of cells that can be in three states: burnt, burning or unburnt. The difference between a raster-based model and purely cellular automata (CA) is that while the first uses a physical or empirically-based model for rate of spread, the second only uses mathematical rules to spread the fire cells (called autonomous agents).

Raster-based models use a diverse set of spread equations; from Rothermel's equations to least-time-to ignition formulas. Using a model based on heat balance and flame contact spread, Green[21] found that the fire fronts tend to become an ellipse (see figure 1.7). They are particularly suitable for heterogeneous fuel and weather conditions since the variables can be locally stored in each cell.

Belcher et al.[7] applied the same technique to successfully model the invasion of smouldering fire into homogeneous fuel. However, handling this amount of data creates a tread-off between fire front resolution and memory and computational requirements[68].
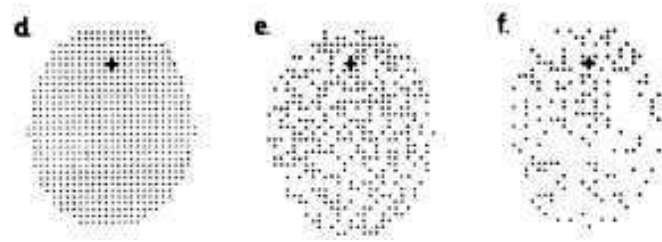
Figure 1.7: Raster-based model developed by Green[20] in 1983. Black cells represent burnt fuel whereas white cells are unburned areas (either because there is no fuel or the fire has not spread yet). The fuel cover decreases from left to right from 100% to 25%. Despite this, the actual fire front resembles elliptical shapes. Figure from Green[20]

By contrast, cellular automata (CA) are based on the assumption that complicated interaction equations can be simulated with a fairly simple set of rules that define the state of each cell on the next time step. CA has been applied in many different fields, from crowd simulations to laws of the universe. One of its principal developers and supporters, Stephen Wolfram, claims that CA will create a paradigm shift in the science by unveiling the rules that create complicated systems[76].

In one of the most recent applications of CA in forest fires, G. Achtemeier, 2013[1], validates a free-agent cellular automata with a real-scale field experiment. He uses a set of rules called "Rabbit Rules" to account for fire and weather interaction. The name is an analogy between fire consuming fuel, spreading and spotting, and rabbits jumping, eating and reproducing. His model predicts non-linear interactions that cannot be simulated with empirical-based models. Despite performing much faster than physical model -eventually reaching positive lead times- the author admits that the accuracy of the model needs to be improved for it to be able to serve as a forecasting tool.
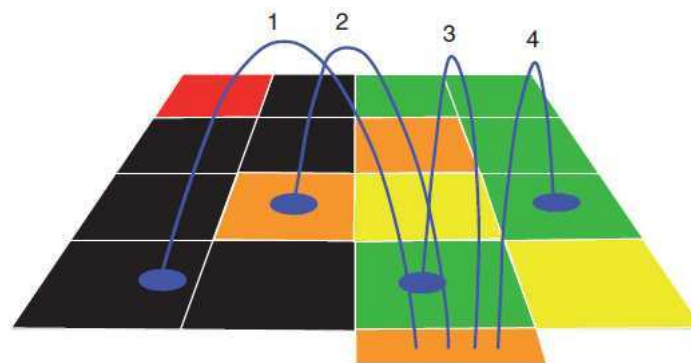
Figure 1.8: Possible "Rabbit" paths used to steer a cellular automata model to reproduce fire spread accounting and weather interactions. Black and green cells represent eaten and uneaten food. Red, orange and yellow cells are occupied by different kinds of rabbits. After Achtemeier, 2013[1]

Bond Percolation and Fractals

Percolation is the physical process of transporting fluid through porous materials. Thus, this approach sees fire as a flow that is percolating through the porous interface created by trees. Percolation has been used in physics to study problems such as conductivity of solids or epidemic spreading.

Since 1990 some authors have tried to apply it in forests using theoretical (Drossel et al. 1992[12]) and experimental (Nhamias et al. 2000[46]) approaches. They represented the forest domain as a lattice of different elemental shapes connected by bonds. Then, they defined the probability of a fire spread through each bond based on neighbour's interactions and experimental correlations.

An important characteristic of bond percolation problems is the critical probability p_c that allows the cluster (i.e. fire spread in the case at hand) to grow infinitely. This quantity is related to the fractal dimension and, according to the universality principle, must be invariant and independent of the dynamics of the system itself.

Although bond percolation studies showed some capacity for qualitatively simulating fire spread, all the attempts to match percolations model with experiments concluded that the model was not adequate. Nevertheless, zooming out the problem, and analysing satellite images of fires (i.e. fire scars) some authors manage to estimate the percolating critical exponent to be between $D = [1.15-1.9]$. This exponent describes the fractal properties of the front and gives information for the scaling. Despite the inherent beauty in the fractal study of fire fronts and their eventual potential to unveil a universal parameter for all wildfire (i.e. critical exponent), no successfully predicting tools have been created by following this direction.

Reaction-Diffusion

Reaction-diffusion equations are normally used to explain how substance concentration ($q(x,t)$) varies when it spreads out (diffuses) and transforms into another substance (reacts). Mathematically, reaction-diffusion processes are parabolic partial differential equations. In one dimension it takes the form of eq. 1.2.

$$\frac{\partial q(x,t)}{\partial t} = \mathbf{D}\frac{\partial^2 q(x,t)}{\partial x^2} + R(q) \tag{1.2}$$

Where $\mathbf{D}$ is the diffusion coefficient that multiplies the energy diffusion term, and $R(q)$ is the reaction term that acts as a source.

Reaction-diffusion equations with different reaction terms can be utilised to explain a wide range of phenomena, from zebra stripes pattern formation to tumour growth.

Combustion discipline brought them into fire to study the reaction-diffusion processes of the flames. In the forest fire discipline Margarit and Sero-Guillaume, 2002, developed and investigated several 2D and 3D models based on irreversible thermodynamics and making different heat transfer assumptions[63],[40]. More relevant to the thesis in hand, at 2008 they demonstrated that the Rothermel theory coupled together with the Huygens principle can be explained, in fact, with a set of reaction-diffusion equations if the thermal equilibrium in gas and solid phase is assumed[64]. To conclude this, they reformulated the problem using Hamilton-Jacobi equations and field theory concepts.

1.2 Towards an operative forecasting tool

Current forest fire fighting methods can be classified into two groups; the ones that are based on fuel removal (hand tools and suppression fire[1]) and the ones based on water suppression (ground water lines and aerial fire fighting). Suppression capacity was enormously improved when the aerial fighting came into the field shortly after the Second World War and some bombers were tested to drop water into wildfire. However, those direct extinguishing methods are insufficient to tackle the blaze[32] especially for *megafires*. This can be illustrated with the following rough calculations.

Hansen, 2012[23] estimated that the heat release rate of hight-intensity wildfire could be taken to be about $2000kW/m^2$. Then, by using the water cooling capacity he estimated that the flow rate required to extinguish such a fire would be $0.1kg \cdot s^{-1} \cdot m^{-1}$. Knowing that each hose line can deliver about 150 l/min (assuming that there is endless water supply, which is not usually the case), then at least 800 water hose lines will be required to extinguish a 5km long fire with a 4m high flames (not even in a megafire scale), or equivalently, 120000 l/min must be evenly delivered along the fire-front. This, is clearly beyond the manpower capacity not even with the assistance of aerial means.

Therefore, a technology able to forecast wildland fire dynamics would lead to a paradigm shift in the response to emergencies, providing the Fire Services with essential information about the on-going fire. The main objective is to enhance the protection of citizens, property and the environment by providing a key tactical tool to fight fires. The cornerstone is to find a computational algorithm that combines a physical model with data assimilation that reliably delivers an accurate forecast within a positive lead time, (i.e. enough time before the event) and enables emergency responders to make tactic decisions and

---

[1]Suppression Fire is the term used to refer to intentional fire lit to fight a blaze. Despite being a delicate technique fire fighting experts around the world have been successfully using it for last decades[53].

deploy resources ahead of it. But the state-of-the-art of computational wildfire dynamics is not fast or accurate enough to provide valid predictions on time (as pointed out in section 1.1). Current wildfire models do not achieve lead times mainly because they take too long to provide a solution (negative lead times) or require data that is impossible to sense (for example, flame temperature or detailed fuel and flow characterisation).

Such an operational forecasting tool needs to meet some characteristics. It has to strike the balance between accuracy and execution time; as the accuracy is kept high enough for operational use, the model has to deliver the forecast faster than real time (positive lead time). At the same time, it has to be versatile enough to be adapted in different fire situations (wide range of fuels, complex topography, weather conditions). Furthermore, it has to be able to incorporate fire fighting interaction (as water lines, fire breaks) and weather forecast, and it must not require high computational capacity (i.e. super computers) as then it will be useless for the forest fire responders.

A potential solution is the assimilation of real-time sensor data, which has been shown to greatly accelerate fire simulations without loss of forecast accuracy (Mandel et al. 2008[39]; Jahn et al. 2011[29]; Jahn et al. 2012[30]; Rochoux al. 2012[58]).

1.2.1 Data Assimilation and Inverse Modelling

Data assimilation has been historically developed and applied in the weather forecasting discipline to combine observations with the models and create operational tools that deliver reliable forecasts. One of the most common strategies used to merge assimilated data with the model is by means of inverse modelling.

Inverse modelling consists in studying the observed measurements to gain information about the physical system behind them using a large variety of mathematical tools. In other words, instead of looking at the outputs of the model, it takes those outputs to unveil the physical interactions and the inputs behind.

One main difference between fire forecasting and weather forecasting problems is that the weather governing equations are already known and, despite using several estimated parameters, the assimilation problem is normally used as a method to solve an initial condition problem and estimate the initial values to run an accurate forecast. By contrast, in the forest fire discipline, the set of equations are not fully developed yet. In this case data assimilation has to estimate the initial conditions for the model as well as adjust the model itself so it fits the observations and enables a reliable forecast.

Nevertheless, inverse modelling is particularly appropriate for forest fire modelling due to large amount of unknowns in real-life fires. The fuel properties and its location, the distance between fuel particles, the shape and the area covered by the branches, the moisture content... are necessary variables to initialise a physical classical forecasting model and hardly ever can be measured. By contrast, the inverse approach can use any kind of available data if the forward model is tweaked accordingly.

Despite its great capacity for coping with complex problems with a large number of variables, few authors have recently tried to apply it to the fire field. Among these, Jahn et al.[29],[30] successfully pioneered this approach to forecast fires in enclosures using both simple and complex models (two zones model and CFD).

In the field of forest fire, Mandel et al.[38] explored this technique to predict time-temperature curve of a sensor placed in the way of an advancing wildfire. They examined a reaction-diffusion equation (see section 1.1.2) and a semi-empirical fire line propagation model coupled with a Eulerian level-set-based equation (see section 1.1.2). Additionally, they coupled weather forecast information to the model demonstrating the powerful potential of data assimilation. Despite its potential, their implementation was found to be unstable due to the ill-definition of temperature measurements in a wildfire.

Following this work Rochoux et al.[58] applied real-time data assimilation to predict the wildfire front location using infra-red observations. Data was also assimilated with a Kalman filter to balance computational and observational errors, -central part of their model-. Rochoux's model assimilates perimeter locations in different times and uses the fuel depth (height of the combustible, $\delta(x,y)$) as an input. The propagating model used two components: the rate of spread (RoS) represented by a product between the fuel depth (δ) and a parameter to be identified ($RoS = \Gamma \cdot \delta$), and a level-set-based equation to cast the fire perimeter. Rochoux tested the model in a controlled small scale fire but only assimilating one fire front and delivering one 30sec forecast. The fact that their model deals with errors using Kalman

Filters implies the assumption that the errors are normally distributed (i.e. can be represented by the mean and the covariance). However, in fire, errors do not follow a Gaussian distribution[39] as fire itself creates discontinuities.

Our forecasting algorithm

In this thesis, we propose and explore a novel approach based on real-time data assimilation. We formulate the inverse problem based on the premise that some invariant exists by following the contributions of Jahn et al.[29],[30] on forecasting fire dynamics in enclosures. Invariants are the set of governing parameters that are mutually independent and constant for a significant amount of time. If the physicochemical parameters cast into each invariant remain constant for a significant period of time, so does its effect on the fire front. Therefore, our implementation relies on the assumption that some properties of the system remain constant during a period of time. The invariants are not restricted to physical observables and can represent their effect to the system. For instance, if the wind speed changes but its effect to the fire remains constant (due to topography or fuel distribution) the proper invariant will be its effect to the rate of spread and not the wind speed itself.

Thus, we propose a forecasting algorithm that relies on the existence of such fundamental invariants and calculates them by assimilating the position of the wildfire front in real time. After an assimilation window of observations, the invariants converge to a value that is used to forecast the perimeter expansion in the subsequent time interval. This forecast is then accurate until any of the invariants change significantly, which would be detected by the continuous feed from sensor data. The observational errors embedded in the assimilated data are considered to be smaller than the forward model accuracy and therefore their influence is not directly considered. To suit our model to a larger range and more complex situations we will expand the forward model and explore different ways of adapting it to specific situations depending on available data.

Regarding data inputs, in this thesis we only considered fire fronts positions hypothetically supplied by air-born observations, or ground crew. However, additional data such as heat release rate, flame height or spreading rate (recently measured by infra-red images and stereo vision[59]) should be considered in future developments.

To fully exploit the potential of this tool with the time-constraints of a master's thesis, we narrowed down the scenarios and focused on wind driven fires leaving the effect of the terrain slope for further development. The reason why we chose to simplify the problem by leaving the slope out off the picture is because it complicates the representation of the model outputs. The perimeters must be projected to the horizontal plane while the spread is taking place in the local plan tilted by the local slope. Therefore it requires some more scripting, although the methodology, and particularly its potential, remains unaffected.

Chapter 2

Methodology: Data assimilation and inverse modelling

Each problem that I solved became a rule, which served afterwards to solve other problems.

– René Descartes

2.1 Building up the Forward Model. Casting the invariants

The initial step when posing an inverse modelling problem is to determine the forward model and its invariants (see section 1.2.1). The forward model is the set of equations that relates the invariants to the observables. In the forecasting model it is necessary for two different purposes. First, it is iteratively used to identify the invariants and then to deliver a forecast valid until the invariants change or the next assimilation process is started.

To build up the forward model we combined Rothermels[60] and Richards[54] models. The first model estimates the Rate of Spread (RoS) of any point in the fire front whereas the second model, uses this RoS to generate the firelets that expand the fire front and computes its location at any time. Following the ideas presented in section 1.2.1, and after some exploration of the resultant model we cast the invariants and reduced the forward model down to its essentials.

2.1.1 Recasting the spread model

This model was developed by Rothermel and the US Forest Service in 1972. Although it relies on the previous work carried out by Fons and Byram[8] it is the first quasi-empirical model that is used as a functional tool being the basis of the National Fire Danger Rating System[67] and the prediction tool Behave[26]. The model is based on an energy balance equation where the heat sources and sinks are identified to estimate the RoS (equation 2.1) of a surface fire. The original formulation uses several experimental correlations deduced from data obtained in tunnel experiments using artificial fuel beds of varying conditions. The model was derived for fires that are in a quasi-steady spread phase. This means that the initial acceleration of the fire is not considered. The shape of the fire front is assumed to have no influence on the RoS and as such, RoS changes only if the underlying variables change. Despite this assumption, Rothermel's model is suitable for our forward model since the data assimilation copes with this limitation as new data is fed to the system (see section 1.2.1).

The original Rothermel's formula for RoS is an energy balance equation with a heat source and sink corrected by wind and slope factors (see appendix A for the full set of underlying correlations):

$$RoS = \frac{\gamma_r \xi}{\rho_b \varepsilon Q_{ig}} (1 + \Phi_w + \Phi_s) \tag{2.1}$$

Where:

- RoS = Rate of spread $[m/s]$: Speed of propagation of the fire front.

- ξ = Propagating flux ratio: Ratio between propagating flux (front energy density that goes to drive the fire) and reaction intensity.

- γ_r = Reaction intensity $[kJ/m^2]$: The heat source. It accounts for the heat that is being generated in the fire front per unit area.

- Φ_w = Wind factor: Experimental correlation to caption the boosting influence of the mid-flame wind speed on the front spread.

- Φ_s = Slope factor: Experimental correlations to estimate the effect of the positive slope[1].

- ρ_b = Ovendry bulk density $[kg/m^3]$: Quantity of fuel per unit area (function of fuel depth δ).

- ε = Effective heating number: Correcting factor for the amount of fuel that is actually involved in the combustion process.

- Q_{ig} = Heat of preignition: Experimental correlation with moisture content (Mf) to account for heat of gasification [2] (pyrolysis).

When it comes to identify the proper invariants some considerations need to be taken. The invariants should be as physically independent as possible. So if one physical condition change only the correspondent invariant needs to be identified again while the others can still remain constant. Moreover, they should be functionally related to a physical variable so their value can be initially estimated.

After considering several configurations the equation 2.1 is recast with three invariants that can be defined using Rothermel model:

$$RoS = I_{mf}(1 + I_u \cdot I_w) \tag{2.2}$$

All the fuel properties are captured by I_{mf}; bulk density (Wo), surface-area to volume ratio (σ), moisture content (M_f), moisture of extinction (M_x) and fuel depth (δ). Those properties cannot be cast with further invariants on account of multiplicity problems due to functional dependency (see section 3.2.1).

$$I_{mf} = \frac{I_r \xi}{\rho_b \varepsilon Q_{ig}} = \mathcal{F}(\sigma, Wo, \delta, M_f, M_x) \tag{2.3}$$

The wind speed is directly represented by I_u. This invariant will be also used in the second part of the forward model (Huygens, section 2.1.2).

$$I_u = U \tag{2.4}$$

But the effect of the wind speed on the fire spread depends on fuel properties such as layout, bulk density, surface-area to volume ratio and fuel depth. Its effect is embedded in I_w that can be formulated by using the Rothermel's functional relation for wind factor Φ_w:

$$\Phi_w = CU^B \left(\frac{\beta}{\beta_0}\right)^{-E} = \mathcal{K} \cdot U^{B-1} \cdot U = I_w \cdot U \tag{2.5}$$

$$I_w = \mathcal{K}(\sigma, Wo, \delta) \cdot U^{B-1} \tag{2.6}$$

Where C, B, E, β, β_0 are experimental coefficients determined by Rothermel and presented in appendix A.
The fact that I_w is analytically dependent on I_u seems to contradict the casting criteria. However, as it will be extensively discussed in the following sections, this theoretical dependency does not play a role in the forecasting algorithm.

---

[1]Note that the slope factor is presented for completeness and, as argued in section 1.2.1, will not be considered in this thesis.
[2]According to Dougal Drysdale[13] nomenclature

2.1.2 Huygens principle; firelet expansion

Although Rothermel's model can estimate the RoS of any point, it is an appraised mean value for the heading portion of the fire[60] and does not give information about fire spreading off the main direction. Therefore, it is not sufficient in predicting the fire front shape and location. Thus, in parallel to RoS estimation, some other model must be used to represent the fire perimeter expansion.

We chose Huygen's principle as described in section 2.1.2 due to the physical interpretation of the elliptical firelet expansion as virtual ignition sources. Despite this election, the forecasting algorithm at hand can be also explored with the other models presented in section 1.1.2.

An elliptical shape can be defined by the eccentricity (that is directly related to the length-to-breadth ratio, equation 2.14) and an axis (either the major or the minor) with its direction. The first depends on the wind speed whereas the second is determined by the RoS and wind vector direction by aligning the major axis of the ellipse to it.
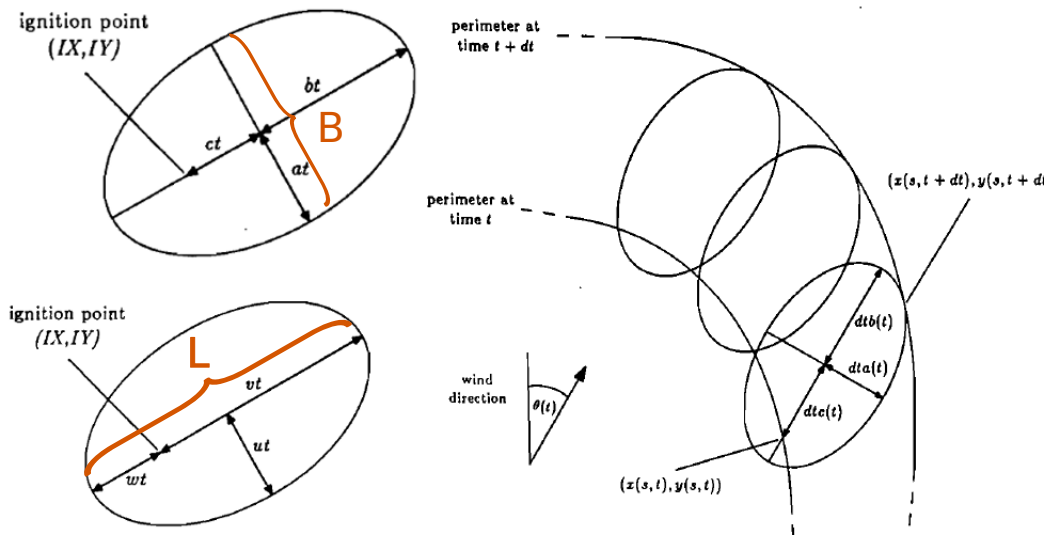
Figure 2.1: Fire perimeter at time t and $t+dt$ and firelets geometrical properties (Length (L) and Breadth (B)). The spread velocities in the bottom-left (u, v, w) are the physical velocities whereas the variables in the top-left (a, b, c) are the ones used in the formulation. The major axis is aligned with the wind (or slope-wind) vector. From Richards1993[55]

Following this idea, Richards[54] developed a set of partial differential equations to calculate the $\{(x_i(t), y_i(t)\}$ coordinates of fire front vertexes. Considering the initial fire-front coordinates $\{\hat{x_0}(s), \hat{y_0}(s)\}$, parameterised with $s \in [0 - 2\pi]$, they can be integrated as:

$$x(s, \hat{t}) = \hat{x_0}(s) + \int_0^{\hat{t}} \big(a(t,s)cos\theta(t)cos\phi(s,t) + b(t,s)sin\theta(t)sin\phi(s,t) + c(t,s)sin\theta(t)\big)dt \qquad (2.7)$$

$$y(s, \hat{t}) = \hat{y_0}(s) + \int_0^{\hat{t}} \big(-a(t)sin\theta(t,s)cos\phi(s,t) + b(t,s)cos\theta(t)sin\phi(s,t) + c(t,s)cos\theta(t)\big)dt \qquad (2.8)$$

Where θ is the slope-wind vector direction, t is the time variable used for integration (to distinguish between $\hat{t}$ which is the time where the function is to be evaluated), $a(r,s), b(r,s)$ and $c(r,s)$ are the lateral, forward and backward spread velocities (as shown in figure 2.1) that depend on time and (or) position, and ϕ is defined as:

$$cos\phi = \frac{A(s,t)}{\sqrt{(A^2(s,t) + B^2(s,t))}}$$

$$sin\phi = \frac{B(s,t)}{\sqrt{(A^2(s,t) + B^2(s,t))}}$$

with:

$$A(s,t) = a(t)\big(\hat{x}_s(s)sin\theta(t) + \hat{y}_s(s)cos\theta(t)\big) \tag{2.9}$$

$$B(s,t) = b(t)\big(-\hat{x}_s(s)cos\theta(t) + \hat{y}_s(s)sin\theta(t)\big) \tag{2.10}$$

Where the subscript s denotes the partial derivative with respect to s ($\frac{\partial}{\partial s}$). Richard demonstrated that this derivative is time-independent and therefore can be determined with the initial conditions[55].

If the fire is ignited in a single point $\{X_0, Y_0\})$, the initial perimeter could be approximated to a small circle of radius dt:

$$\hat{x}(s) = X_0 + dt\,cos(s)$$
$$\hat{y}(s) = Y_0 + dt\,sin(s)$$

Substituting these in eq. (2.9) and eq. (2.10) and collapsing the circle radius dt to zero leads to a simplification of eq. (2.7) and eq. (2.8):

$$x(s,\hat{t}) = X_0 + \int_0^{\hat{t}} \left(\frac{a^2(t)cos\theta(t)cos(\theta+s) + b^2(t)sin\theta(t)sin(\theta+s)}{\sqrt{a^2(t)cos^2(\theta+s) + b^2(t)sin^2(\theta+s)}} \cdot c(t)sin\theta(t) \right) dt \tag{2.11}$$

$$y(s,\hat{t}) = Y_0 + \int_0^{\hat{t}} \left(\frac{a^2(t)sin\theta(t)cos(\theta+s) + b^2(t)cos\theta(t)sin(\theta+s)}{\sqrt{a^2(t)cos^2(\theta+s) + b^2(t)sin^2(\theta+s)}} \cdot c(t)cos\theta(t) \right) dt \tag{2.12}$$

The b, c velocities can be calculated by imposing Rothermel's fire head rate of spread from the *new* ignition point situated on the rear focus (see fig. 2.1):

$$b + c \equiv RoS \tag{2.13}$$

Lateral front velocity a, however, is directly related to the eccentricity of the firelet (see figure 2.1). It was originally estimated using an experimental correlation found by Anderson[5] that relates the ratio between the major and the minor firelet's axis (length-to-breath ratio LB depicted in figure 2.1), and thus, the ratio between b and a (independent of the time step Δt used). Its value depends on the wind speed ($U[m/s]$) in accordance with the equation:

$$\frac{a}{b} = 0.936e^{0.2566U} + 0.461e^{-0.1548U} - 0.397 \equiv LB \tag{2.14}$$

The constant 0.397 is a modification of the Anderson's original formula to ensure that the fire expands circularly ($LB = 1$) under no-wind conditions ($U = 0$). According to Alexander[3] the above correlation is valid for average wind speed below $30km/h$.

Once the length-to-breadth ratio is known, a, b, c velocities can be calculated using eq. (2.13) and the elliptical geometry relationship between b/a and v/u ratios:

$$\frac{v}{u} = \frac{LB + \sqrt{LB^2 - 1}}{LB - \sqrt{LB^2 - 1}} \equiv HB$$

$$a(s,t) = RoS(s,t)\frac{1 + 1/HB}{2LB} \tag{2.15}$$

$$b(s,t) = RoS(s,t)\frac{1 + 1/HB}{2} \tag{2.16}$$

$$c(s,t) = b(s,t) - \frac{RoS(s,t)}{HB} \tag{2.17}$$

If the invariant I_u (eq. (2.4)) introduced in Rothereml's model is reused in Huygens' firelets expansion, only one additional invariant is required to account for the principal direction of spreading determined by the direction of the blowing wind:

$$I_\theta = \theta \tag{2.18}$$

Then, the firelet expansion equations for fire front positions become:

$$x(s,\hat{t}) = X_0 + \int_0^{\hat{t}} \left(\frac{a^2(t)cos(I_\theta)cos(I_\theta + s) + b^2(t)sin(I_\theta)sin(I_\theta + s)}{\sqrt{a^2(t)cos^2(I_\theta + s) + b^2(t)sin^2(I_\theta + s)}} \cdot c(t)sin(I_\theta) \right) dt \qquad (2.19)$$

$$y(s,\hat{t}) = Y_0 + \int_0^{\hat{t}} \left(\frac{a^2(t)sin(I_\theta)cos(I_\theta + s) + b^2(t)cos(I_\theta)sin(I_\theta + s)}{\sqrt{a^2(t)cos^2(I_\theta + s) + b^2(t)sin^2(I_\theta + s)}} \cdot c(t)cos(I_\theta) \right) dt \qquad (2.20)$$

$$LB = 0.936e^{0.2566I_u} + 0.461e^{-0.1548I_u} - 0.397 \qquad (2.21)$$

$$a(s,t) = RoS(s,t)\frac{1 + 1/HB}{2LB} \qquad (2.22)$$

$$b(s,t) = RoS(s,t)\frac{1 + 1/HB}{2} \qquad (2.23)$$

$$c(s,t) = b(s,t) - \frac{RoS(s,t)}{HB} \qquad (2.24)$$

The forward model is then a function of four invariants:

Forward model equations:

$$\mathcal{M}(I_u, I_w, I_{mf}, I_\theta, T) = \begin{cases} RoS = \mathcal{R}_t(I_u, I_w, I_{mf}) \\ \{x, y\} = \mathcal{H}(RoS, I_u, I_\theta, T) \end{cases} \qquad (2.25)$$

Where T is the time when the fire front is estimated, $\mathcal{R}$ represents Rothermel's model with cast invariants (eq. 2.2) and $\mathcal{H}_t$ firelet expansion (from eq. 2.20 to eq. 2.24).

2.2 Sensitivity analysis

The behaviour of the invariants can be studied by means of a sensitivity analysis which enables us to better understand the effect of the underlying variables and to establish their physical possible values. Invariants I_θ and I_u directly refer to a physical observable quantity, thus, their range can be already estimated. I_θ must obviously cover all the directions whereas I_u is limited by eq 2.21:

$$I_\theta \in [0 - 2\pi](rad)$$

$$I_u \in [0 - 8.3](m/s)$$

Although $8.3m/s$ seems to be easily exceeded, this refers to mean wind speed and not gust.
For the other two invariants, a Monte-Carlo sensitivity analysis is performed by randomly creating 20000 sets of Rothermel's original variables (see eq. (A.1)). The limits for those variables are displayed in table 2.2 and are taken from the tabulated data used in current models. The invariant values obtained are depicted in a histogram plot. Besides establishing the range, this representation allows for its distribution which will help us when analysing the model's performance.

Variable	Units	Base Value	Range
Fuel depth, δ	$[m]$	0.6	[0.03-1.5]
Moisture content, M_f	[%]	10	[3-20]
Wind speed, U	$[m/s]$	2	[0-7]
Surface-area-to-volume ratio (SAV) σ	$[1/m]$	5900	[2200-7200]
Ovendry fuel loading W_o	$[kg/m^2]$	0.7	[0.022-1.5]
Moisture content of extinction, M_x	[%]	20	-

Table 2.1: Base values and range for the sensitivity analysis and synthetic data generation. M_x has not been modified since the dependency in the model is always $\frac{M_f}{M_x}$. Values range taken from the fuel models described by Scott and Burgan 2005[62] and more recently by Overholt 2012[49]
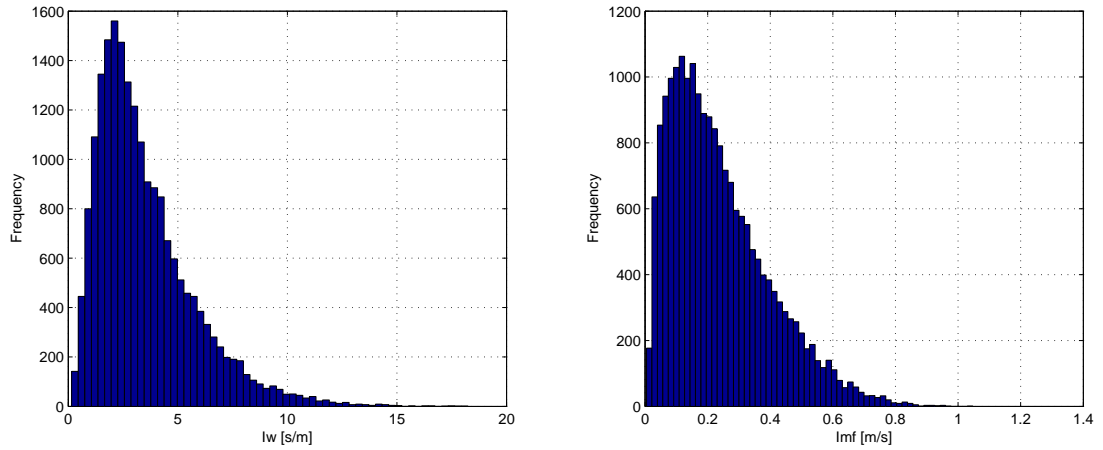
Figure 2.2: Histograms with I_w (on the left) and I_θ (on the right) values for a sample of 20000 randomly generated sets of Rothermel's variables

Observing figure 2.2 the ranges for I_w and I_{mf} can be easily estimated:

$$I_w = [0 - 100](s/m)$$

$$I_{mf} = [0 - 1.2](m/s)$$

Both invariants have a similar distribution. Instead of a uniform distribution (as the underlying values), they rather follow a poisson distribution. Note that I_w multiplies wind speed (I_u) and therefore its units are s/m to become a dimensionless product.

To better understand the model it is important to explore the sensitivity of each invariant. Again, due to their nature, I_w and I_θ are directly related to their background variable (if wind speed or its direction change a certain amount, so do their invariants).
However, I_w and I_{mf} depend on many quantities and thus, we want to know how they change when the quantities are modified. To do so, those two invariants are calculated with the quantities in the table 2.2 perturbed $\pm 20\%$ from its base value. The response of the invariants is represented in figure 2.3.
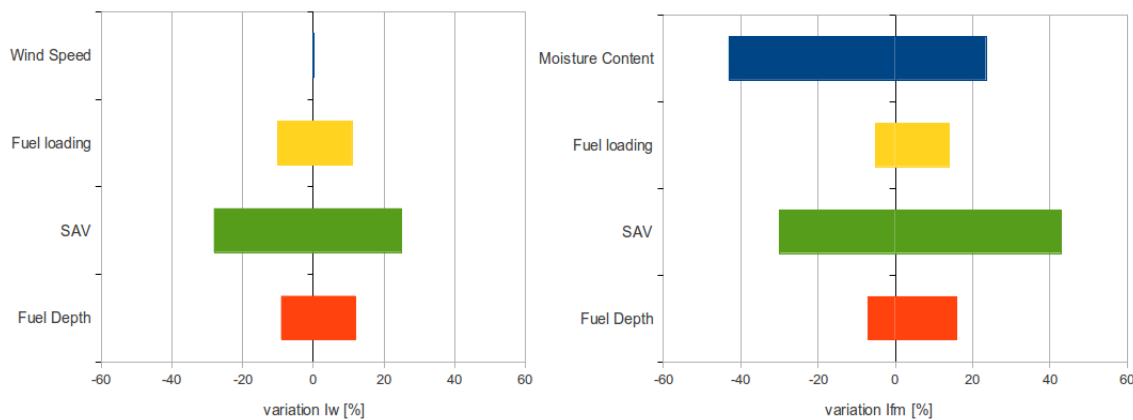
Figure 2.3: Sensitivity analysis of the invariants I_w and I_{mf}

Although I_w analytically depends on the wind speed (see eq. (2.5)) the sensitivity analysis reveals that its effect is negligible. Regarding I_{mf}, the moisture content together with the surface-area-to-volume ratio (SAV) play the the principal role steering this invariant.
Following the same philosophy, the same sensitivity analysis is performed regarding the final output of

the model (i.e. fire front positions). Now the invariants are calculated according to the values in table 2.2 and perturbed ±20%. The output of the model is not a single value but a set of coordinates. Thus, the Euclid norm between the fire front vertices is used to account for the difference between invariants vector (set of invariants) as depicted in figure 2.4. The integration time is set to one minute and three fire fronts are considered.
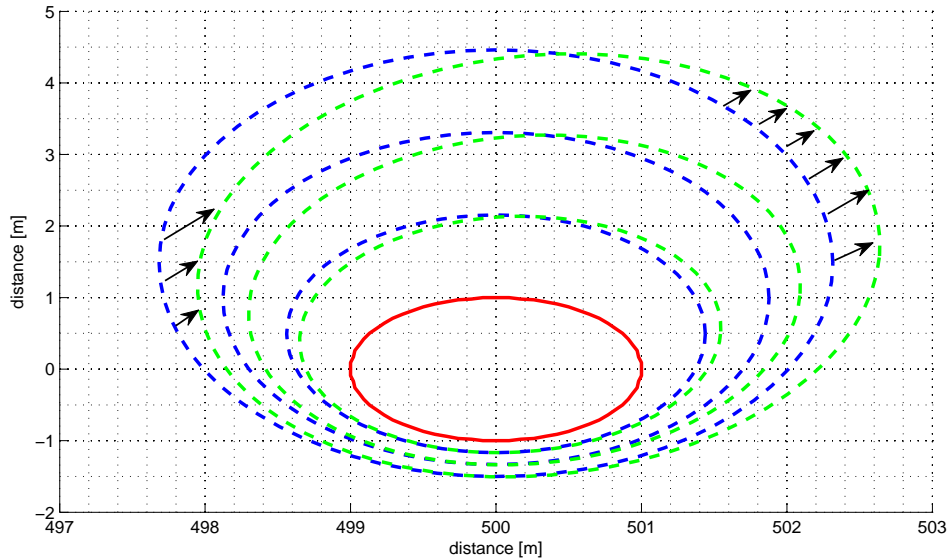
Figure 2.4: Fire front representation for two invariants vector. Blue dashed lines correspond to the base value whereas for green dashed lines the invariant vector is perturbed with +20% in I_θ invariant. The red solid line represents the initial ignition perimeter. The black arrows show some of the vertexes used to calculate the Euclidean norm that differentiates one perimeter form the other.
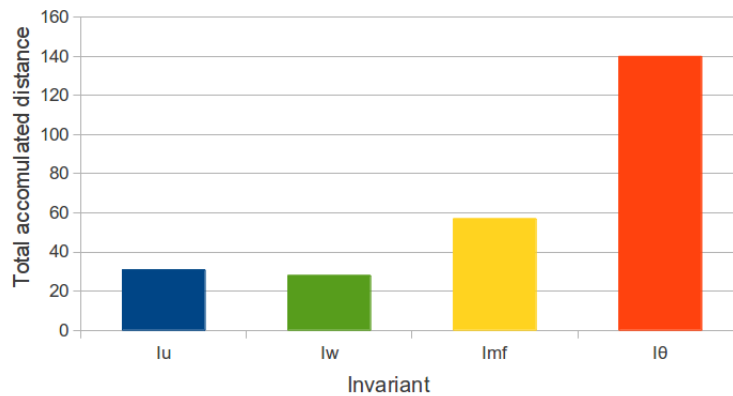
Figure 2.5: Bar plot with the accumulated Euclidean distance between fire fronts generated with base values and when they are perturbed ±20%

Figure 2.5 shows that the wind direction has the maximum influence on the proper resolution of the fire perimeter. A 20% change in the direction created a massive change of the fire-line. This points to the importance of an accurate estimation of I_θ since a slight discrepancy would lead to an erratic forecast. The moisture-fuel related invariant I_{mf} turns out to have a larger influence to the perimeter than the wind speed. This is because its range is smaller, therefore, it must be treated carefully. Those findings, together with the validity range of the invariants and the discovery that wind speed does not play a role

in I_w and is thus independent of I_u, will be essential for implementing and understanding the forecasting algorithm.

2.3   Reversibility of the invariants

Depending on the acquirable sensor data, the invariants can be turned into input data for the problem. For example, if reliable wind speed data arrive, there is no need to solve for it but instead it is directly used as input in the forward model. This particular feature makes this methodology suitable for a large range of different wildfire situations and for an ever-evolving technological support. The assimilating data algorithm actually imbibes both, measured and forecasted data, to deliver the most accurate forecast.

In the present study, the strategy of tweaking the invariants when additional data is available -and eventually turning them into input parameters-, relies on the empirical correlations developed by Rothermel[60] to close his model(see section 2.1.1). They allow us to calculate or redefine the correspondent invariant when the needed data is provided. To better understand the concept we shall consider the following credible case:

In an given forest fire, the fuel properties (dimensions, depth, density...) are unknown. However, the moisture content (m_f) might be estimated and reliably predicted according to the weather forecast (or day-night variations).
Then, the fuel-moisture invariant I_{mf} can be redefined as:

$$I_{mf} = \widetilde{I}_{mf} \cdot \mathcal{F}(m_f)$$

Where $\mathcal{F}(m_f)$ is a known function -or a correlated equation- and $\widetilde{I}_{mf}$ becomes the new invariant. Note that now I_{mf} does not depend on m_f.
The algorithm is now run swapping I_{mf} with $\widetilde{I}_{mf}$. The moisture content information feeds the model during the assimilation phase and, once the invariants are identified, it also feeds the forward model to create a more accurate forecast (see figure 3.6).

The same strategy showed here with m_f can be applied to any of the eleven physical variables underlying the model at hand (see section 2.1.1). The only difference is that if the new input data affects more than one invariant then the invariant cast needs to be extended to all the invariants involved. In section **??** this case is also explored as wind speed and fuel depth, respectively, are considered to be input data.

2.4   Cost Function

The invariants are calculated by minimizing a cost function J that measures the difference between the model output and the observations. The cost function proposed is based on the Euclidean norm summed over the different assimilation times.

$$\mathcal{J}(\mathbf{p}) = \sum_{t=t_i}^{t_f} \sqrt{[\mathbf{x}_i - \hat{\mathbf{x}}_i(\mathbf{p})]^T \mathbf{W}_i [\mathbf{x}_i - \hat{\mathbf{x}}_i(\mathbf{p})] + [\mathbf{y}_i - \hat{\mathbf{y}}_i(\mathbf{p})]^T \mathbf{W}_i [\mathbf{y}_i - \hat{\mathbf{y}}_i(\mathbf{p})]} \tag{2.26}$$

Where $\{\mathbf{x}_i, \mathbf{y}_i\} \in \mathbb{R}^2$ are the N-coordinates set of the observed fire front position in a given time i and $\hat{\mathbf{x}}_i(p) = \mathcal{M}_x(\mathbf{p}))$ are the corresponding model output position for a set of invariants ($\mathbf{p}$).
$\mathbf{W}_i$ is a weigh function that could be used to give more value to the latest assimilation on time in order to better capture the invariants to create the forecasting. However, in the present work no weighting function is used ($\mathbf{W}_i = \mathbb{I}$) but the door is left open for further exploration in this direction.
Equation 2.26 can be simplified if the $\mathbf{x} - \mathbf{y}$ coordinates (both, observables and model output) are concatenated as 1-row vector $\bar{\mathbf{y}}_i$ and $\tilde{\mathbf{y}}_i = \mathcal{M}_i(\mathbf{p})$:

$$\mathcal{J}(\mathbf{p}) = \sum_{t=t_i}^{t_f} \sqrt{[\bar{\mathbf{y}}_i - \tilde{\mathbf{y}}_i(\mathbf{p})]^T [\bar{\mathbf{y}}_i - \tilde{\mathbf{y}}_i(\mathbf{p})]} \tag{2.27}$$

Although the square rood gives the correct Euclidean norm, it does not affect the minimisation as can be easily shown:
The relative extreme (i.e. also minima) condition is:

$$min. \quad \mathcal{F}(x) \Rightarrow \nabla\mathcal{F}(x^*) = 0 \tag{2.28}$$

Where x^* are the minima coordinates. Then if:

$$\mathcal{J}_1(x) = \sqrt{\mathcal{F}(x)} \rightarrow \nabla\mathcal{J}_1(x^*) = \frac{\nabla\mathcal{F}(x^*)}{2\sqrt{\mathcal{F}(x^*)}} = 0$$

And thus:

$$\Rightarrow \nabla\mathcal{F}(x^*) = 0 \qquad \text{if} \qquad \mathcal{F}(x^*) \neq 0$$

Which is exactly the same condition as stated in eq. (2.28).
Therefore, due to computational reasons, the chosen cost function is:

$$\mathcal{J}(\mathbf{p}) = \sum_{t=t_i}^{t_f} [\bar{\mathbf{y}}_i - \mathcal{M}_i(\mathbf{p})]^T [\bar{\mathbf{y}}_i - \mathcal{M}_i(\mathbf{p})] \tag{2.29}$$

2.5 optimisation

They are two main approaches to minimize eq. (2.29); the gradient-free and the gradient-based[48]. The first group are stochastic algorithms that evaluate the cost function $\mathcal{J}(\mathbf{p})$ at many random points to find the absolute minimum, whereas the second group use an initial guess (called background vector, $\mathbf{p}^b$) and follow the gradient direction towards the closest minimum. Although gradient-free algorithms ensure that the absolute minimum will be found, they have to evaluate the cost function multiple times which might be computationally costly if the forward model $\mathcal{M}(\mathbf{p})$ is complex. On the other hand, when the cost function is continuous and the possible domain of the invariants ($\mathbf{p}$) is known and delimited (as it is in our problem, see section 2.2), the gradient-based algorithms are more suitable and efficient. It is true that gradient-based algorithms can converge to a local minima instead of a global one. However, the extended sensitivity analysis performed in appendix **??** showed that the system is benign in the sense that all the functions involved behave smoothly and therefore the convergence to the absolute minima will depend on the initial guess. This is a key point of the forecasting algorithm and is discussed in section 3.1.
If the forward model $\mathcal{J}(\mathbf{p})$ is lineal then the cost function is quadratic and can be minimised by easily solving a system of linear equations (as will be shown in the following sections). For forward models that are not linear (as it is the case, see section 2.1) the Tangent Lineal Model (TLM) has been widely used to tackle the problem[22].

2.6 Tangent Linear Model

TLM consist in linearising the forward model $\mathcal{M}(\mathbf{p})$ in the vicinity of an initial guess $\mathbf{p}^b$. This linearisation can be done if the model is weakly non linear as the one at hand (see appendix **??**). The viability of the TLM relies on the initial guess and the fact that the procedure is iterated until convergence. To calculate the TLM we use Taylor's series expansion as follows:

$$\mathcal{M}(\mathbf{p}) = \mathcal{M}(\mathbf{p}^b) + \nabla\mathcal{M}(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b) + O^2(\mathbf{p} - \mathbf{p}^b) \approx \mathcal{M}(\mathbf{p}^b) + \nabla\mathcal{M}(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b)$$

Where ∇ is the gradient operator and $O^2(\mathbf{p} - \mathbf{p}^b)$ includes the second and higher order terms that are neglected.

By inserting this in eq. (2.29) we get:

$$\mathcal{J}(\mathbf{p}) = \sum_{t=t_i}^{t_f} \left[\bar{\mathbf{y}}_i - \left(\mathcal{M}(\mathbf{p}^b) + \nabla\mathcal{M}(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b)\right)\right]^T$$

$$\left[\bar{\mathbf{y}}_i - \left(\mathcal{M}_i(\mathbf{p}^b) + \nabla \mathcal{M}_i(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b) \right) \right] \tag{2.30}$$

The gradient of the linearized function is now:

$$\nabla_{\mathbf{p}} \mathcal{J}(\mathbf{p}) = 2 \sum_{t=t_i}^{t_f} \left[\left(\nabla_{\mathbf{p}} \mathcal{M}_i(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b) \right) \right]^T \left[\bar{\mathbf{y}}_i - \left(\mathcal{M}_i(\mathbf{p}^b) + \nabla_{\mathbf{p}} \mathcal{M}_i(\mathbf{p}^b)(\mathbf{p} - \mathbf{p}^b) \right) \right] \tag{2.31}$$

Applying the first order condition for minimisation (eq. (2.28)) and introducing the following notation:

$$\mathbf{M}_i = \mathcal{M}_i(\mathbf{p}^b)$$

$$\mathbf{H}_i = \nabla_{\mathbf{p}} \mathcal{M}_i(\mathbf{p}^b)$$

$$\bar{\mathbf{p}}_i = (\mathbf{p} - \mathbf{p}^b)$$

eq. (2.31) becomes:

$$\sum_{t=t_i}^{t_f} \mathbf{H}_i^T \left(\bar{\mathbf{y}}_i - [\mathbf{M}_i + \mathbf{H}_i \bar{\mathbf{p}}] \right) = 0$$

And rearranging,

$$\sum_{t=t_i}^{t_f} \mathbf{H}_i^T \mathbf{H}_i \bar{\mathbf{p}} = \sum_{t=t_i}^{t_f} \mathbf{H}_i^T (\bar{\mathbf{y}}_i - \mathbf{M}_i) \tag{2.32}$$

Then, defining the terms:

$$\mathbf{A} = \sum_{t=t_i}^{t_f} \mathbf{H}_i^T \mathbf{H}_i$$

$$\mathbf{b} = \sum_{t=t_i}^{t_f} \mathbf{H}_i^T (\bar{\mathbf{y}}_i - \mathbf{M}_i)$$

eq. (2.32) becomes a linear system of equations:

$$\mathbf{A}\bar{\mathbf{p}} = \mathbf{b} \tag{2.33}$$

which can be easily solved by using a QR factorisation[3] with column pivoting[48].

2.7   Automatic Differentiation

Calculating the jacobian multiplication term $\mathbf{H}_i^T \mathbf{H}_i$ in eq. (2.32) requires partially differentiating the model with respect to the different invariants. This has to be done $p \times 2n \times t$ times, where p is the number of invariants used, $2n$ the coordinates of the fire front and t the assimilating time.
The simplest way to numerically evaluate the jacobian is by finite centred differences:

$$\mathbf{H}_{k,i}^j = \frac{\partial \mathcal{M}_i^j(\mathbf{p}^b)}{\partial p_k} \simeq \frac{\mathcal{M}_i^j(\mathbf{p}^b + \epsilon_k) - \mathcal{M}_i(\mathbf{p}^b)}{\| \epsilon_k \|}$$

where $\epsilon_k \in \mathbb{R}^p = \{0, 0, ..\epsilon, ..0\}$ has a perturbation ϵ in the position k.

Despite its simplicity, this approach has multiple downsides: The forward model has to be evaluated twice each time, and ϵ should be reduced as much as possible which introduces numerical truncation errors when ϵ is too small [22]. For these reasons we chose an automatic differentiation approach.

---

[3]QR factorisation is based on the decomposition of any matrix $\mathbf{A}$ into an orthogonal matrix $\mathbf{Q}$ and a upper triangular matrix $\mathbf{R}$.

Automatic differentiation allows us to directly calculate the jacobian matrix $\mathbf{H}_i$ (normally called *Tangent Linear* or *Forward*) or $\mathbf{H}_i^T$ (called *Adjoint*). It consists in iteratively applying the *chain rule* used in differentiation calculus to script the differentiated model that gives all the partial derivative. An illustrative example is shown below:

Forward and adjoint automatic differentiation example

Let us consider the function:

$$y = sin(\frac{x_1}{x_2})$$

The statements in that should be implemented in a code to compute y when $x_1 = 1$ and $x_2 = 2$

```
Program
x1=1
x2=2
x3=x1/x2
y= sin(x3)
```

The analytical partial derivative with respect to x_2 is:

$$\frac{\partial y}{\partial x_2} = cos(\frac{x_1}{x_2})(-\frac{x_1}{x_2^2}) = -0.23$$

Another way to calculate this value is by scripting the forward differentiation of the `Program`. The `dProgram` computes the partial derivative with respect to x_2 without directly knowing the analytical derivative of the function $y(x_1, x_2)$:

```
dProgram
x1=1
dx1=0
x2=2
dx2=1
x3=x1/x2
dx3=-x1/(x2*x2)*dx2
y= sin(x3)
dyx2= cos(x3)*dx3
```

Which gives `dyx2=-0.23`.
The adjoint differentiation (sometimes called backwards differentiation) works as follows:

```
aProgram
x1=1
x2=2
x3=x1/x2
y= sin(x3)
dy=1
ax3=cos(x3)*dy
ax2=ax3*(-1/(x2*x2))
ax1= ax2* x1
```

Which also gives `ax1=-0.23`.

Note that the number of statements in both `dProgram` and `aProgram` are twice the initial number and thus, the time to run it could be expected to double. However, just by running the `dProgram` once, the Jacobian matrix can be computed. This elemental example shows how this algorithm can be applied to differentiate any kind of code. It is even suitable to differentiate numerical integral evaluations (as is our case, see fig. 2.1) since all the statements can be split down to elemental mathematical operations.

The tangent linear differentiation is indicated when the number of observables (i.e. model outputs) is much larger than the quantity of independent variables (i.e. the number of invariants in our case). By contrast, calculating the Adjoint Differentiation is more convenient and efficient when there is a large number of independent variables.

Therefore, in the present work, we use forward differentiation approach. The differentiation of the code was generated with the help of `ADiMat`[25], a set of functions and libraries developed by Scientific Computing centre in Aachen University.

Nevertheless, in the future expansion of the methodology, if the number of degrees of freedom of the forward model (i.e. number of invariants) is increased, the Adjoint Automatic Differentiation should be explored to keep the computational efficiency and maximise the lead time.

2.8 Forecasting program structure

Having stated the mathematical framework, all the steps are linked together in a program scripted in MATLAB® language. Despite the fact that some high-level languages might suit the computational requirements better, MATLAB® was chosen due to its versatility and capacity to handle and plot data. The following diagram (figure 2.6) summarises the principal parts of the assimilating and inverse modelling program:
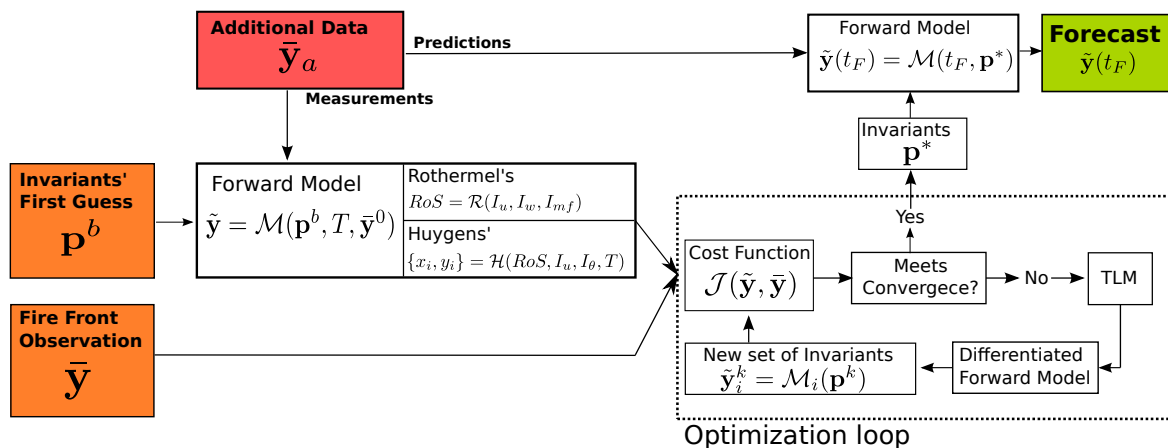

Figure 2.6: Program structure flow diagram. Orange boxes are the required inputs, green box is the output and red box shows additional inputs.

First, fire front positions $\bar{\mathbf{y}}$ are assimilated during a specific period of time (called assimilation windows). Meanwhile, an educated guess estimates the first set of invariants $\mathbf{p}^b$. This first guess is based on roughly estimated data and the influence of the model is explored in section 3.1. This invariants guess is input into the forward model together with the assimilated time information and one fire front position (or the initial ignition point) $\mathcal{M}(\mathbf{p}^b, T, \mathbf{y}^0)$. The consequent first prediction set of fronts $\widetilde{\mathbf{y}}_i$ is compared with the assimilated data by means of the cost function $\mathcal{J}(\widetilde{\mathbf{y}} - \bar{\mathbf{y}})$ (see eq. (2.27)). If the fronts do not match (i.e. the cost function is not zero)the program starts the optimisation iteration loop.

The first statement in the loop is to run the differentiated forward model to calculate the jacobian terms in eq2.33. The solution to this equation gives a new set of invariants $\mathbf{p}^k$ that is input to the forward model again to get a new estimated set of fire fronts. If the convergence criteria are reached by comparing them again with the observation then the best estimated invariants vector is found ($\mathbf{p}^*$) and thus the forecast is delibered by running the forward model with the forecasting time. Otherwise, the loop is iterated again.

The fact that a loop is needed to estimate the invariants reduces the inaccuracy added by applying a tangent linear approach to a non-linear model since in every new iteration the model is linearized (i.e.

the differentiated forward model is run) in a new state point ($\mathbf{p}^{k+1}$). In addition, if any of the new invariant values in the vector $\mathbf{p}^{k+1}$ exceeds the physical range, established in section 2.2, its value is set back to the initial guess to prevent non-physical results.

Note that every time that the differentiated forward model is run, the forward model is also evaluated (see `dProgram` in section 2.7). Thus the forward model is always evaluated at the same time as the differentiated model speeding up the simulations and enabling the use of complex forwards models.

Regarding convergence, two criteria can be requested. The first is to state the maximum discrepancy between simulated and observed fire perimeters (i.e. the value of the cost function) whereas the second imposes convergence between consecutive estimations of the invariant vector. While the first ensures the simulation matches the observation, it might not always happen. By contrast, the second condition will always be fulfilled (if there is any local minimum) although the perimeter might not be always fully simulated. In the following sections, both criteria will be explored and compared.

2.9   Synthetic data

The forecasting model is validated with synthetic data generated by the same tool as the forward model, that is a Rothermel-Huygens firelet expansion model with all the correlated formulas and experimental values embedded. All those expressions can be found in appendix A. The synthetic data generator was scripted in MATLAB® as well, and was initialised with the fuels properties provided by Scott and Burgan[62]. Wind speed and direction were chosen from within a reliable range as listed in table 2.2. The synthetically generated data was stored in a file and inputted into the forecasting algorithm in due time trying to mimic the data acquisition in a real situation.

It is important to point out that the Huygens firelet expansion used in both the forecasting algorithm (as a part of the forward model) and in the model to generate synthetic data are actually the same implementation of equations 2.7 and 2.8.

2.9.1   Verification

To check that the Rothermel-Huygens model was correctly scripted and behaves as reported in the literature, a verification analysis was performed to check that no human-induced errors are present in the code and that it correctly performs the implemented calculations.

Firelets

The first test consisted in calculating and depicting all the firelets and checking that the integrated fire front envelopes them all. Figure 2.7 shows that, indeed, the calculated front is the composition of the firelets and it allows us to understand how the backfire -fire spreading upwind- is simulated with Huygens. The white ellipse in the center that mimics the ignition ellipse (red solid line) is an unexpected side effect from the geometrics.
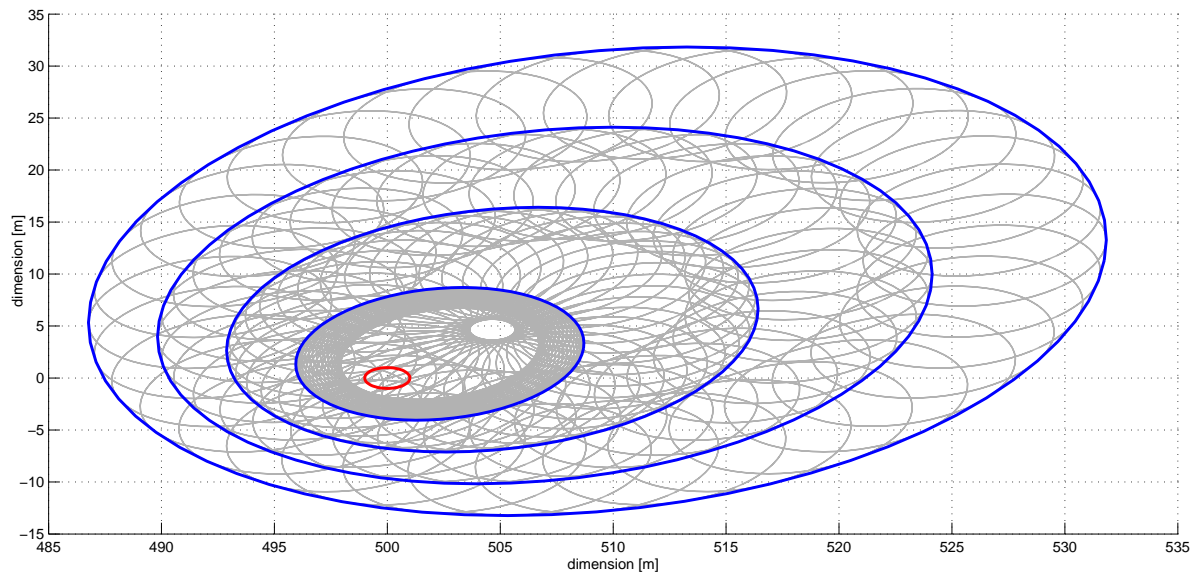
Figure 2.7: A simple case with constant wind speed and direction to illustrate the Huygens principle. The red line is the initial ignition perimeter. The spreading ellipses are depicted in grey and the enveloping curve (new fire-front) is coloured in blue.

Wind response

The second verification deals with the influence of wind direction since it is involved in the differential equations to be integrated, which correspond with the critical part of the model. The wind direction is swept over the 360° keeping the other quantities constant. Figure 2.8 shows the symmetry effect of the wind blowing in both diametrically opposed directions. The cases where the wind is blowing from north and south (first and fourth case) seem to have a different shape than the other cases but that is because the x-axis scale changes to better display those configurations.
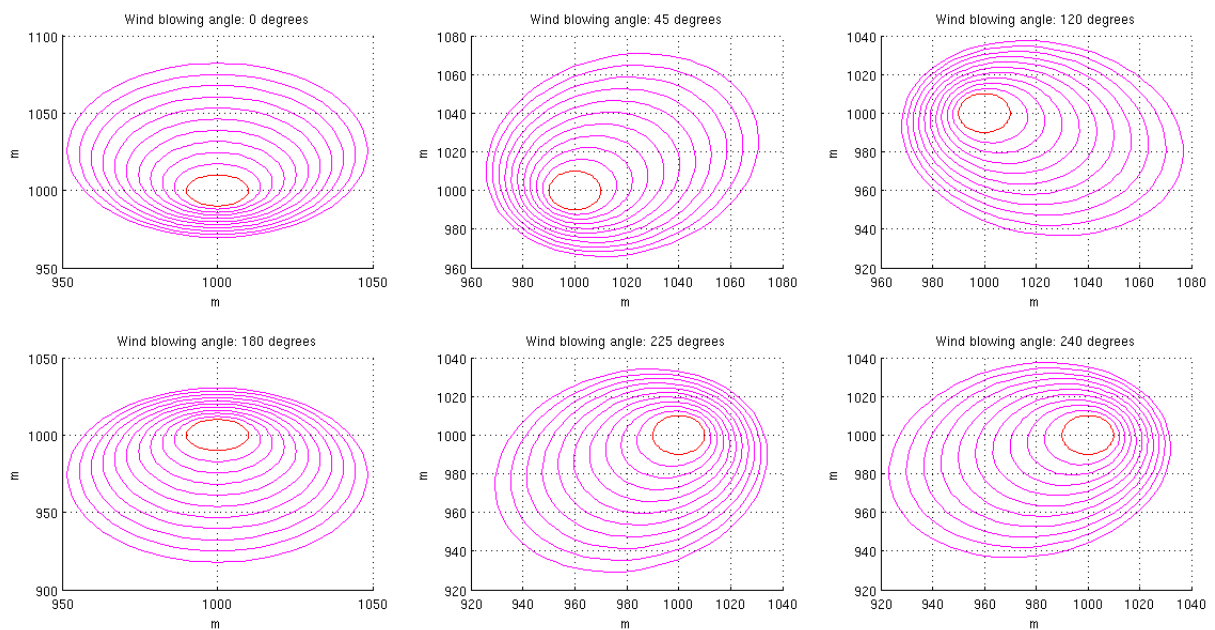

Figure 2.8: Wind direction influence on front spread with a homogeneous fuel and wind speed. Six different blowing directions from 0 to 250 degrees shown with. All figures have the same ignition perimeter source (red line).

Commutativity in wind directions

This verification aims to check a particular and curious property of Richard's equations (equations 2.11, 2.12); the commutativity of wind speed. If the wind speed is assumed to vary with each time step and the respective values at each instant are permuted to create a new wind-time vector that contains exactly the same wind speed values but in a different order, then, the last fire front from both wind-time vectors must match each other.

Figure 2.9 illustrates and demonstrates this property. Magenta and blue lines come from different runs of the model where the wind speed has been permuted. Despite creating different shapes throughout the integration time, both last fronts match perfectly. This property is peculiar since each fire front is calculated from the previous front location and no commutativity seems possible. The fact that Richard[55] mentioned it and we found it with our model, indicates that it is properly scripted.
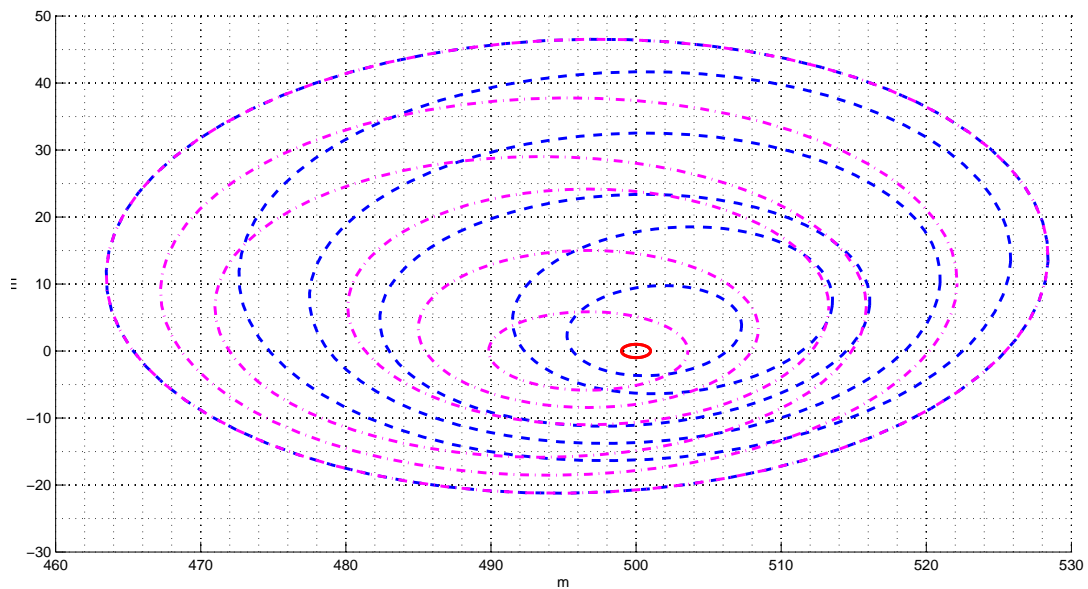
Figure 2.9: Commutativity of wind directions. The wind direction vector is permuted to generate two different fire front expansions (blue and magenta dashed lines). Due to the model's particular commutativity in wind direction, both cases match on the final front.

Effect of fuel heterogeneity

The last tests is to quantitatively check if the forward model can simulate more realistic fire front shapes. To do so, we created an hypothetical situation where wind speed and direction change and the fuel depth is randomly perturbed by ±20% of its base value (table 2.2). Ten fire fronts are generated within an interval of one minute.
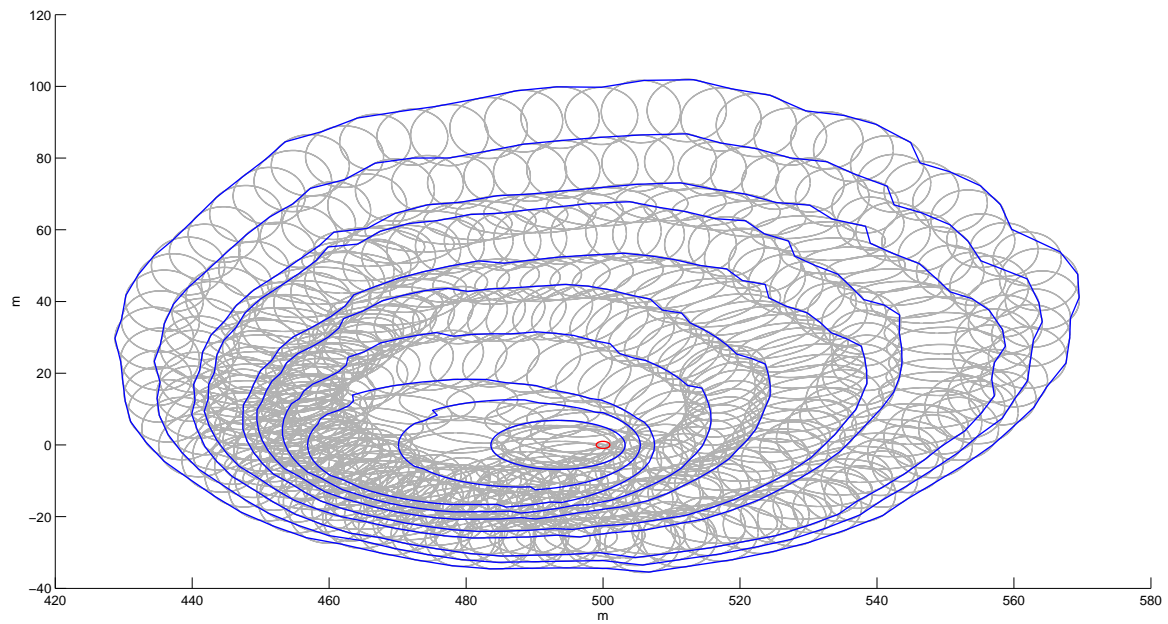
Figure 2.10: Ten fire fronts with randomly generated heterogeneous fuel depth and changing wind speed and direction at every time step. The interval between correlative fronts is one minute

Indeed, although no real data is available for comparison, fire fronts in figure 2.10 resemble real fire fronts and prove that despite the elliptical nature of the firelets this Rothermel-Huygens firelet expansion model can output shapes that are not elliptically alike.

Chapter 3

Results and Discussions

Nothing great
was ever achieved
without enthusiasm.

– Ralph Waldo Emerson

The forecasting algorithm developed in this thesis is investigated in different situations where synthetic data simulates the observations to assimilate. The tests are performed varying different parameters like the assimilation window, assimilated data (fire fronts locations and feeding frequency) and initial guess. Then we look at several features like convergence of the invariant, minimisation of the cost function, effect of the initial guess, effect of the assimilating window width, the computing time and the leading times obtained.
These analyses are performed with the forward model discussed in the previous section 2.1. However, the same methodology is applied and explored with alternative invariants to handle more realistic situations where some of the physical quantities assumed as constant are now allowed to vary.

The fact that the synthetic data is generated with Rothermel's-Huygens model implies that it exists at least one true invariant vector that generates the observed fronts exactly. However, this is not the case in reality since the forward model used is only an approximation of the real fire dynamics. Thus, to explore the case where there is no unique and exact solution to generate the fire fronts, the synthetic data will be randomly perturbed before feeding the forecasting algorithm.

In all of the following test, punctual ignition source is considered as the initial integration point for the fire front expansion. This ignition point source is depicted as a red spot in all the following plots and is a required piece of information to run the forecasting algorithm. In a real wildfire situation, it could be identified as the first reported location of the fire -the bottom of the smoke plume, as an example-. If the fire has spread out before the first bit of information arrives and it is no longer a point source, the first assimilated fire front can be also used as a virtual ignition perimeter by considering the whole fire front as a set of initial integration points. Then, equations 2.7 and 2.8 have to be used to calculate the fire fronts.

3.1 Initial educated guess

Apart from the assimilated data, the forecasting algorithm needs an initial guess of the invariant's value where the first Tangent Linear Approximation (TLM) is performed. This first educated guess can be directly generated by considering the range of validity of each invariant only and directly estimate its value without considering any hint from the actual wildfire. However, some of the invariants (such as I_w or I_{mf} for example) are hard to estimating because they depend on several physical quantities. Thus, Rothermel equivalent equations for the invariants (equations 2.3, 2.4 and 2.6, 2.18) can be used to estimate the initial guess by guesstimating the six physical underlying quantities δ, M_f, M_x, σ, W_0 and θ which can be roughly done by observing the fuel and wind (see section 2.1.1 and appendix A). Although both strategies generate a valid guess, the latter one justifies it better and will be employed in

the following section.

The initial guess values together with the true value of the invariant used to generate the synthetic data are displayed in the bottom part of every plot of the following sections. The INFO string also provides information about the number of assimilated fire fronts and the assimilating period (minutes between observations).

It is worth mentioning that sometimes the differences between the true invariants value and the first educated guess seem to be small. That is because of two reasons. First, the absolute value and range of the invariants are small and thus the percentage difference must be considered when comparing them. Second, despite the possibility of wrongly guesstimating the underlying variables, their divergence can be balanced in the final invariant value.

To perform a rigorous study of the effect of the initial guess, the six initialising variables were swept according to operational-based considerations. For instance, the fuel depth δ was considered to be an easily estimable variable in a real wildfire (it can be easily distinguished between half metre tall grass or 3m tall forest) and thus, its offset in the initial guess is considered to be lower than 1.50m. In contrast, some other variables such as moisture content (M_{mf}) or ovendry fuel loading (W_0), cannot be estimated with such accuracy and therefore the educated guess exploration covers a larger differing range.

3.2 Identifying the invariants

The first important output to check is the convergence of the cost function (equation 2.27). If the cost function converges towards zero it indicates that the implemented algorithm correctly solves the optimisation problem.

The first scenario investigated here assimilates 15 fire fronts with a window width of 15 minutes (i.e. one fire front per minute). The invariants converge within 3 iterations (i.e. 3 runs of TLM and thus, 3 estimations of the invariants). Figure 3.1 shows the observed data (black triangles), the fronts generated with the initial invariants guess (red dashed lines) and the respective fire fronts after each iteration (dashed lines) until convergence is reached (green solid line). The invariants and cost function convergence is depicted in figure 3.2. Its left plot shows the rapid decrease of the cost function towards zero. The slope quantifies the converging rate. At the first iteration the slope is steep which indicates that the algorithm quickly corrects the large discrepancies. As the cost decreases so does the steepness indicating that convergence is achieved since in a minimum the steepness flattens.

As mentioned in section 2.1, some invariants are not completely independent. I_w (eq. 2.4)and I_{mf} (eq. 2.6) receive contribution from the same physical magnitudes (δ, σ and W_0). Thus it is possible that despite the cost function converging towards zero, the invariants do not converge to true value. Without any additional assimilated data or information about the system, there is no solution for this and it could well take place during any forecast. In the case shown in figure 3.2 (right) this has not taken place (all invariants converge to true value within 2% of percentage difference). Cases where multiplicity of solutions appear are discussed in the following sections. Figure 3.2 also shows that with the first iteration (orange dashed lines) the direction I_θ is already identified (below 10% accuracy) although the very last fire front is not matched because the other invariants are not yet resolved.

INFO: Assimilated times: 15 // Itrue:[1.00 2.27 2.19 0.87] // Iguess:[1.30 2.25 2.03 0.52]
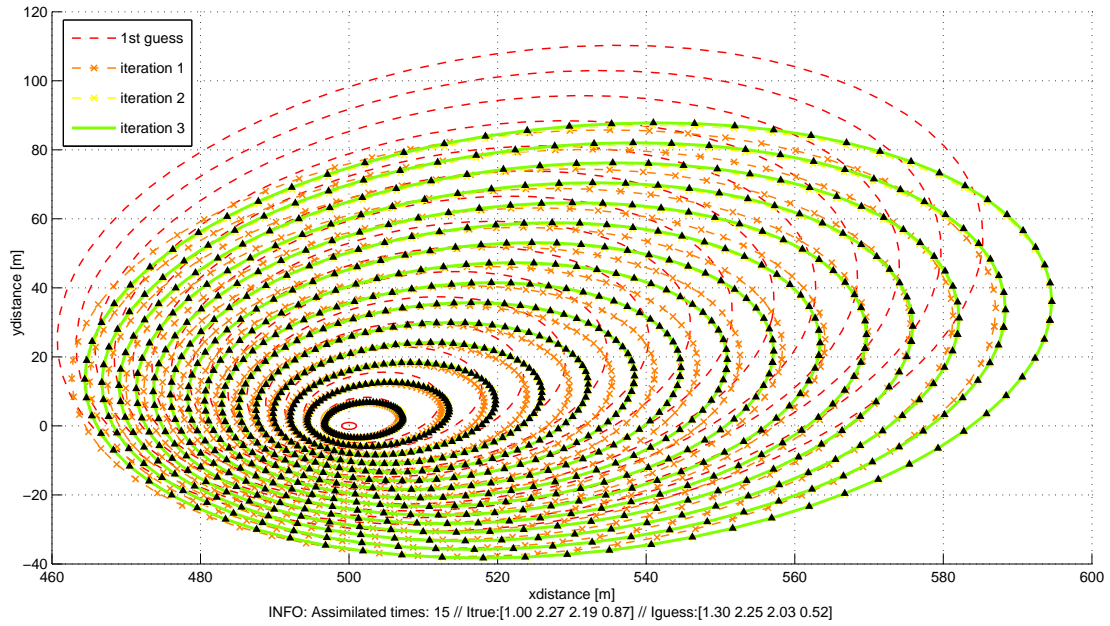
Figure 3.1: Guess, observation, and optimisation iterations of fire fronts in an x-y plane (plan view of a wildfire). The black triangles are the 15 observed positions. The red dashed lines are the fire fronts generated with the first guess and the dashed lines are the following iterations of invariants. The last iteration is depicted with green solid lines.
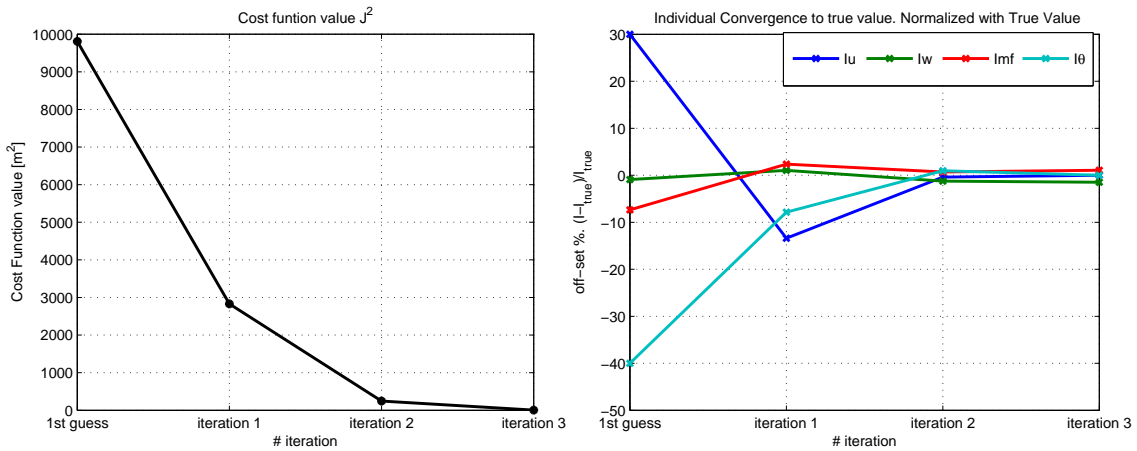

Figure 3.2: Convergence of cost function (left) and individual convergence of each invariant to the true value (right) as a percentage difference. Assimilation windows = 15 min (1 assimilation/min)

In all the scenarios investigated, varying the first guess and the assimilation window (more than 50), the number of iterations needed for convergence was always less than 12. In some cases, however, the convergence was not reached due to the wrong estimation of one of the invariants. The reason is that when the rate of spread (RoS) calculated with the initial guess differs more than about five times from the true RoS value, the geometrical relationship between observed and guessed fire fronts become ill-defined. Figure 3.3 illustrates this aspect. The cost function, calculated as the difference between corresponding fire perimeters at same time step (i.e. front observed at time t with front predicted at time t), is no longer a proper quantity to be minimised. Its value is so large that the forecasting algorithm predicts non-physical invariants values (i.e. out of the range stated in section 2.2). This motivates the implementation of a correction algorithm to automatically reset negative invariants values. Nevertheless, there is a second reason that prevents the algorithm from correctly minimising the cost function in those

situations. Because the invariant vector is too far from the true value the linearisation of the forward
model leads to a false local minimum where the TLM algorithm gets trapped, providing poor solutions.
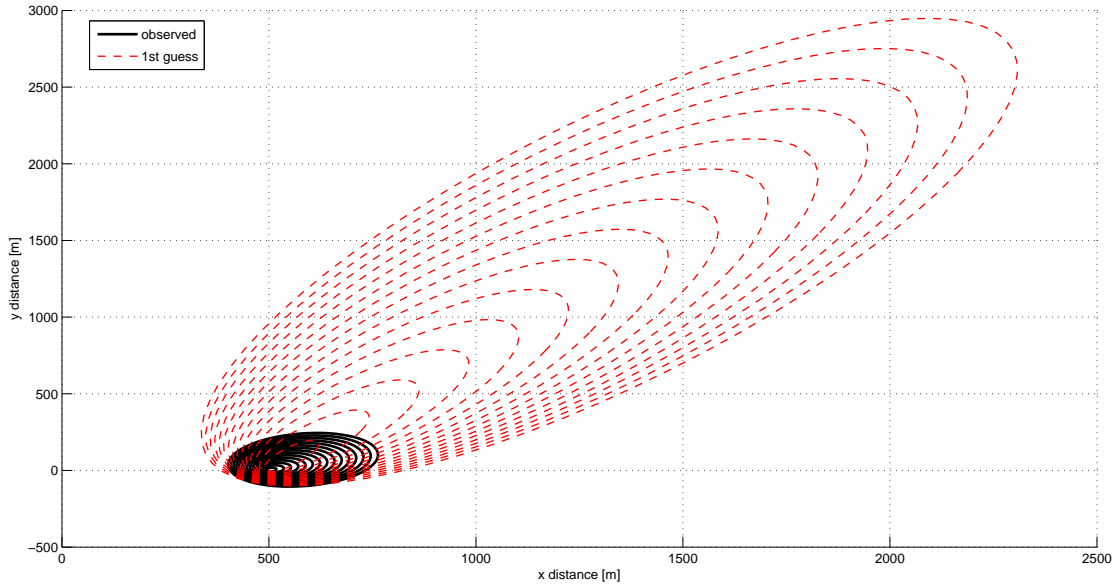

Figure 3.3: Illustrative figure of a case where the TLM approximation and the cost function chosen are
no longer valid to perform a proper optimisation and identify the invariants.

3.2.1   Influence of the assimilation window width

The window width (WW) is the amount of time while the forecasting algorithm is being fed data (i.e.
fire front location in the case at hand). The time between consecutive fire front observations is called
assimilation period (T) and can be directly related to the assimilating frequency $(F = 1/T)$.

The main effect of the number of assimilated fronts (WW/T) is the resolution of multiplicity. The
value of the cost function tends to increase when the assimilation windows increase and more fronts are
assimilated. The error of the initial guess amplifies with the propagation (the previous fire front position
is required to calculate the new one, see Huygens principle in section 2.1.2) and therefore the forecasting
algorithm is more sensitive to the wrong identification of invariants. This is shown in figure 3.4.
In figure 3.4, instead of assimilating 15 min (and 15 fire fronts) -as in the converging example above-
we assimilate front positions during 3 min (i.e. 3 front positions). The cost functions rapidly drops to
zero but in this case the value estimated for both I_{mf} and I_w differs from the true value by 10%. The
reason is that now the initial cost function has a lower absolute value since the propagation of an inac-
curate estimation is truncated in time and therefore the effects of an incorrect assimilation are hidden.
It is worth mentioning that despite the possibility of I_{mf} and I_w misconverging, RoS is always correctly
estimated as it has no multiplicity in the forward model and only one value can match (or approximate)
the observations.

INFO: Assimilated times: 3 // Itrue:[1.00 2.27 2.19 0.87] // Iguess:[1.30 2.25 2.03 0.52]
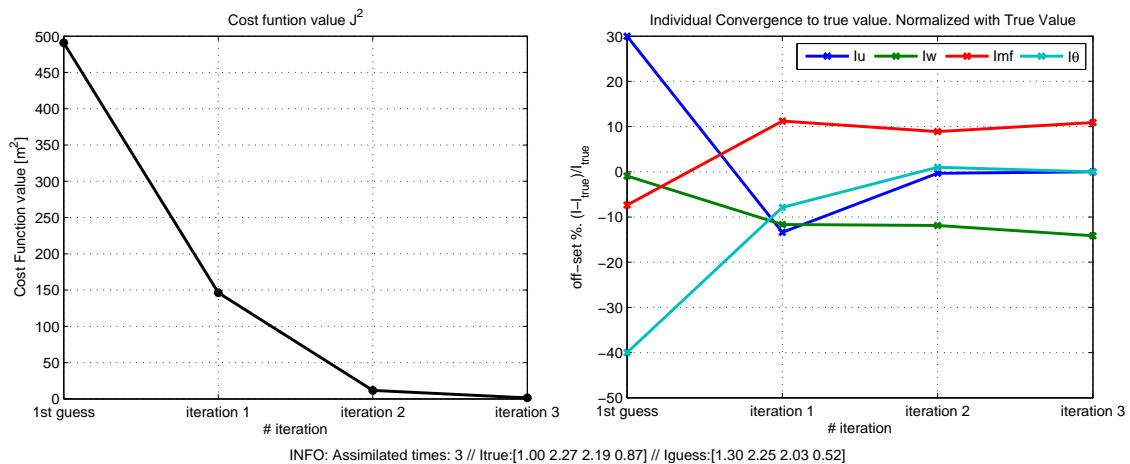
Figure 3.4: Convergence of cost function (left) and individual convergence of each invariant to true value (right). Assimilation windows = 3 min (1 assimilation / min).

It is clear that increasing the assimilation window helps to correctly estimate the dependent invariants, as demonstrated above, yet it is not always sufficient to ensure the proper identification of I_w and I_{mf}. The fire variables (either the fire is driven by the wind or the RoS is mainly determined by the fuel) play an important role in it and it is hard to say in advance if the situation is prone to misconvergence.

Dealing with Multiplicity

One way to deal with multiplicity is by defining only one invariant for the Rate of Spread. This approach, however, does not allow for the forecasting algorithm to be ameliorated if any extra data is available (as will be done in section 3.3) since no information about particular contributions is achieved. Thus, a more interesting way to diminish multiplicity is recasting the invariants and inputting extra data in a way that the invariants become functionally independent. For instance, if the fuel-moisture invariant is somehow multiplied by any measurable quantity as fuel depth or moisture content that varies spatially or over time, then, its value is no longer exchangeable with the wind factor. The same strategy could be used for this second invariant if wind speed is known. This approach is successfully explored in the following sections.

The third way to deal with multiplicity is by assimilating additional quantities that are predicted by the forward model. It is worth pointing out the difference between inputting additional values and assimilating more information. The first consists of extra inputs to run the forward model and allows it to handle more complex situation. Examples of this could be information of moisture content, fuel properties or wind speed. Data assimilation, in contrast, requires the quantifiable information to be the output of the forward model. Thus, in our case, only the positions of the fronts can be assimilated but the forward model can be complemented so it delivers additional characteristics such as flame height or fire intensity. By assimilating this additional information the invariant multiplicity is narrowed down since each invariant is then part of different equations and they are no longer dependent.

3.3 Forecasting the fire spread

Once the invariants are identified the forecasting algorithm predicts the location of the fire by running the forward model again with the correct invariants. The forecast will be valid as long as the conditions present when assimilating the data will remain constant.

In order for it to be an operative tool, the forecasting algorithm must deliver the forecast ahead of the event, thus, any forecast must meet the positive lead time requirement. The lead time is defined as the amount of time between the delivery of the forecast and the predicted event. If the forecasting algorithm needs 1 min (computing time) to deliver a 20 min forecast then the lead time is 19 min.

The lead time principally depends on the number of assimilated fronts and the initial guess (i.e. iterations

required for convergence). The forecasting time (either we ask for a 15 min or 45 min forecast) also plays a role when the forward model is computationally demanding. However, due to the simplicity of the forward model used in the case at hand, its contribution is limited as shown in figure 3.5.
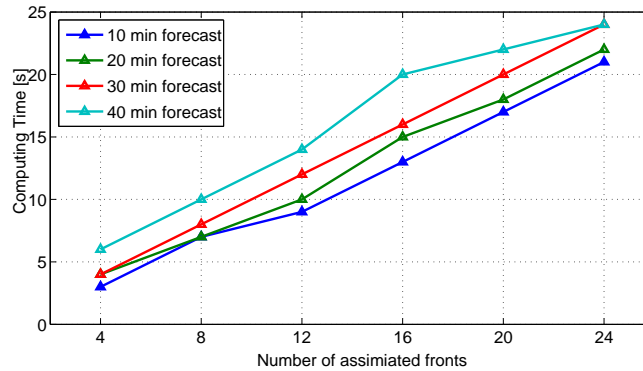

Figure 3.5: Computing time required for 4 different forecasting windows (10, 20, 30 and 40min) versus the number of assimilated fire fronts

They keystone of the forecasting algorithm at hand is that if any condition changes when the invariants are already captured, it can be also used to create the forecast. For instance, in the case shown in figure 3.2 the wind direction is identified to be 50°. If we get the information -weather forecast in this case- that the wind direction will suddenly change and start blowing towards the north (such a wind pattern is characteristic and well known in some areas) then, the wind direction invariant can be updated to the new value $I_u = 0$ and the forecast can be run with the other identified invariants that are still valid (i.e. fuel has not changed for example). Those two different forecasting ways are shown in figure 3.6.
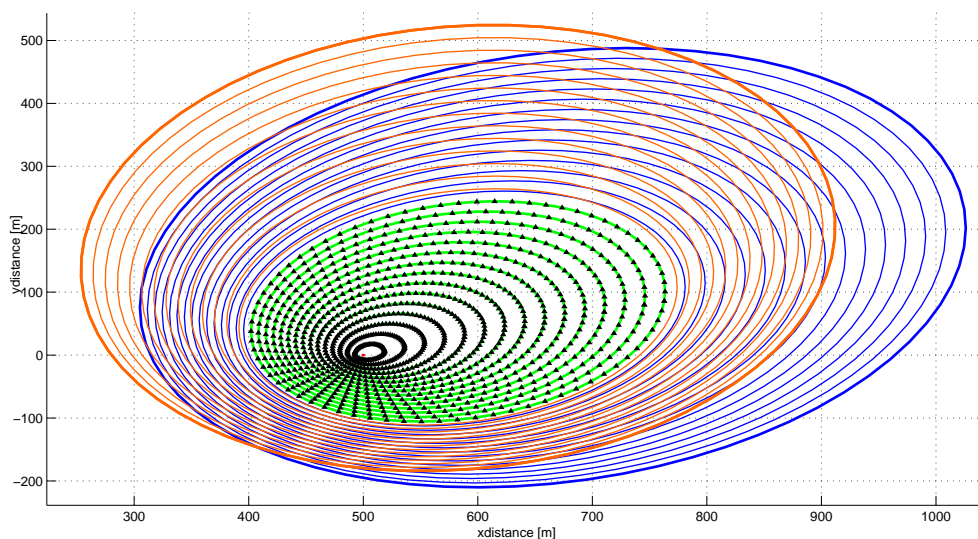
Figure 3.6: Two different forecasts with the invariants identified in figure 3.1. Black triangles are observed data fitted with the invariants of the case above (green solid fronts). Blue perimeters forecast is calculated assuming constant wind direction during assimilation and forecast, whereas orange fronts are based on the available data that inform of a change of wind direction towards north. The other invariants are kept constant since they are sill valid. The time between fire front is 1 min and the last forecast (wider lines) is for 15 min from the last observation.

3.4 Applying the forecasting algorithm to different data contexts

As explained in section 2.3 the invariants can be adapted to different data situations. To show the versatility of our model three different cases with different available data are explored; wind speed, wind speed and direction, and fuel depth.

In the first and second case, wind speed and direction are assumed to be spatially independent and the main difference is the amount of available information and the degree of freedom (i.e. the number of invariants) of the final invariant cast. By contrast, in the third case, the fuel depth δ is allowed to vary spatially which increases the validity of the model for heterogeneous situations. All three cases rely on realistic measuring capabilities in a real fire. Wind speed and direction can be gathered from deployed units as well as from weather stations spread over the fire area. Regarding the information about fuel, forest managers usually map forest areas in advance to list their spatially distributed characteristics. New techniques recently brought into the field such as the use of LIDAR -Light Detection and Ranging- (Mutlu et al. 2008[45]), potentially increase the accuracy and availability of this information and opens the door for preparing operative measuring systems for the situations when these data are not known.

3.4.1 Wind speed as input data

The first step is to recast the invariants related to wind speed. I_u is now directly the wind speed and I_w is redefined using the wind factor functional relation (eq. 2.5):

$$\Phi_w = C U^B \left(\frac{\beta}{\beta_0} \right)^{-E} = \mathcal{P}(\sigma, \beta, w_0, \delta) \cdot U^B = I_{w1} \cdot U^{I_{w2}} \tag{3.1}$$

Thus,

$$I_{w1} = \mathcal{P}(\sigma, w_0, \delta) = C \left(\frac{\beta}{\beta_0} \right)^{-E} \tag{3.2}$$

$$I_{w2} = \mathcal{F}(\sigma) = B \tag{3.3}$$

Where C and B are calculated with experimental correlations derived by Rothermel and β, β_0 are the nominal and the optimal packing ratio respectively. The analytical formulas are presented in appendix A).

The other two invariants I_{mf} and I_θ remain the same and, thus, the model is still described by four invariants.

$$\mathcal{M}(I_{w1}, I_{w2}, I_{mf}, I_\theta, T) = \begin{cases} RoS = \mathcal{R}_t(I_{w1}, I_{w2}, I_{mf}) \\ \{x, y\} = \mathcal{H}(RoS, I_u, I_\theta, T) \end{cases} \tag{3.4}$$

The reason why four degrees of freedom are needed despite the new assimilated data is because the effect of the wind in the RoS and the firelets shape is still unknown as it depends on several fuel parameters as such the packing ratio. However, the important difference is that now the the wind is allowed to change (is not an invariant any more) and, therefore, the forecasting algorithm can deal with more complicated -less idealised- situations.

Despite this recast being, to some extent, more complicated than the previous one, it allows to identify the invariants more accurately than the previous recast. Nevertheless, on average, more iterations are needed to reach the required convergence which slightly increases the computing time.

Figure 3.7 shows the perfect convergence of the invariants to true factor with this recast. The first peak in the I_{mf} invariant is due to the algorithm used which keep the invariant estimations within the physically meaningful range. After seven iterations, the invariants converge to true value with an accuracy lower than 1%.
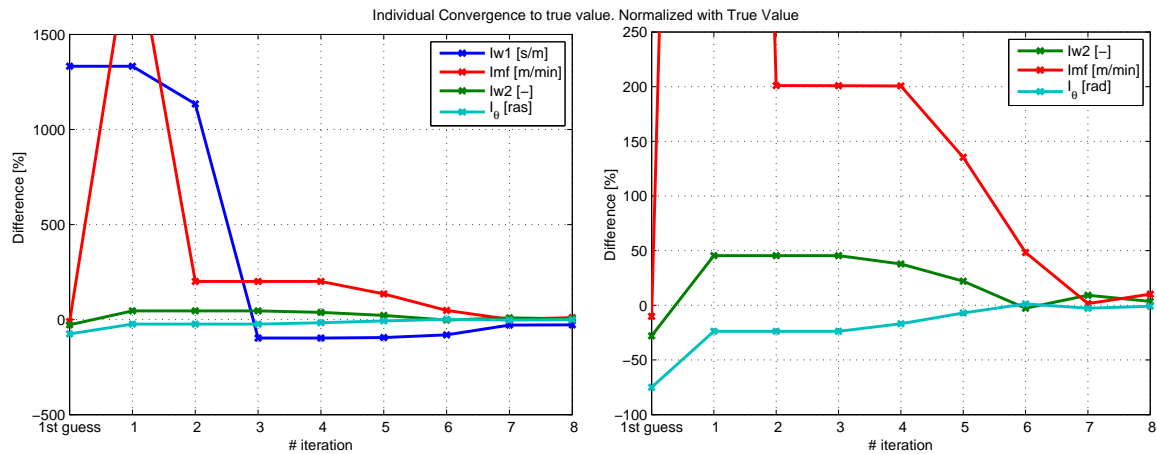
Figure 3.7: Invariants convergence with an assimilation window of 10min with a sampling interval of 2min (i.e. 5 assimilated fronts). The educated guess of Rothermel's variables gives a high perturbation in I_{w1} invariant (left). The figure on the right zooms in to better appreciate the initial discrepancy of the other invariants. All of them converge to true value.

Apart from assimilating fronts with variable wind speed, this implementation of the forecasting algorithm allows to run the forecast while also considering variable wind speed. In this case we assume a wind speed forecast of $U = [1.3, 1, 0, 1.2, 2](m/s)$ for the next 10 minutes. The forecasted fire front locations are displayed in figure 3.8. The effect of no wind can be clearly observed in the second and third forecasted fire fronts.
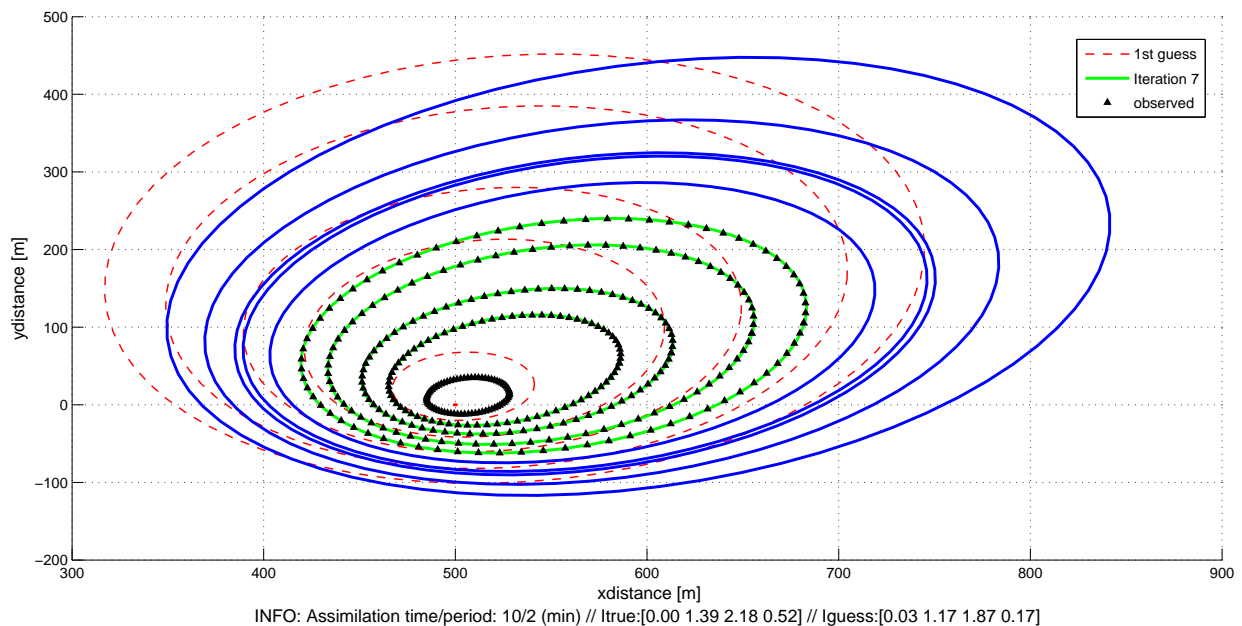
Figure 3.8: Plot of the initial guess (red dashed lines), assimilated data (black triangles), final optimisation (green solid line) and 10 min forecast with a varying wind speed for the same set of invariants as in figure 3.7 (blue solid lines)

3.4.2 Wind speed and direction as input data

If both wind speed and direction are measurable we can consider the same invariant cast as in the previous section (equation 3.4) and directly identify the direction invariant with the wind direction data $I_\theta = \theta$. Several tests are performed to explore this situation and show perfect capacity to correctly

identify the three invariants left. The reduction in the number of invariants speeds up the forecasting algorithm while the iterations needed are on average similar to the previous cases.

The forecast algorithm can also consider wind speed and direction predictions to deliver a more accurate forecast when these quantities vary. In figure 3.9, five fire fronts are assimilated during 25 min (at a frequency of 1 fire front every five min). The invariants are perfectly identified with six iterations. Then, a forecast in launched for the next 25 min with a synthetic prediction of wind speed and direction.

INFO: Assimilation time/period: 25/5 (min) // Itrue:[0.00 1.39 1.95 0.02] // Iguess:[1.17 2.59
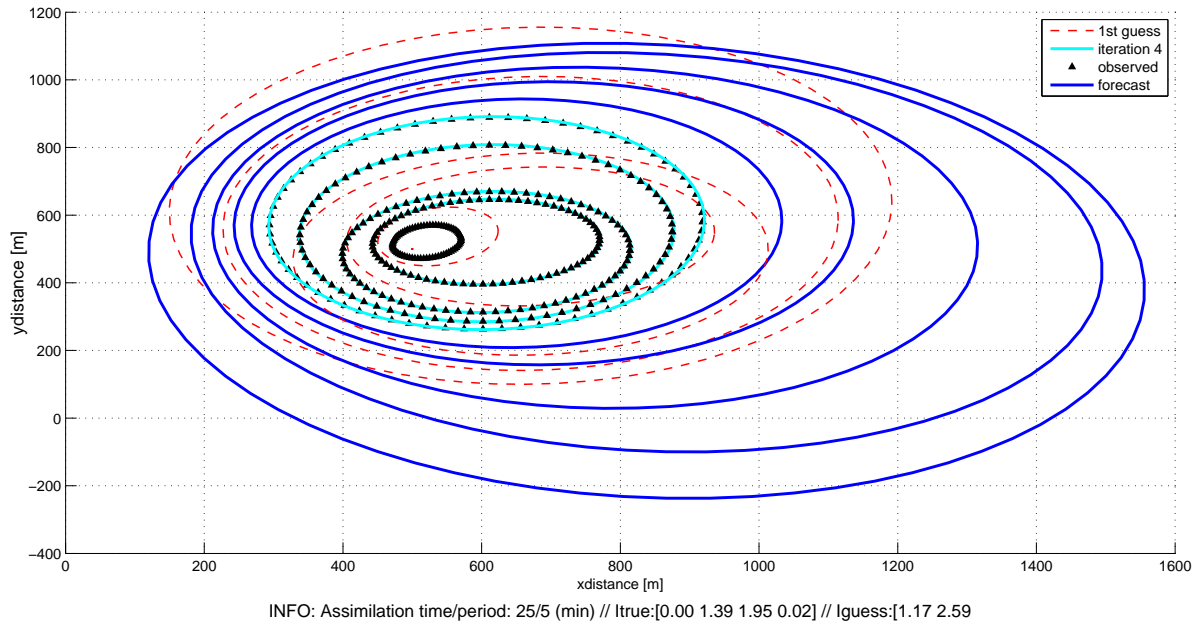
Figure 3.9: Assimilation case where the wind speed and direction change and are inputted as a known value (black triangles). The three remaining invariants (see equation **??**) are identified with four iterations (light blue solid lines). The assimilation window is 25 min with a period of 5min between assimilations. A 25 min forecast is delivered (blue solid lines) by using wind speed and direction predictions.

Figure 3.10: Convergence of the cost function and the invariants when wind speed and direction are used as an input. The peak in the third iteration of the cost function is due to the correcting algorithm that resets negative values.

3.4.3   Fuel depth as input data

We consider now the case where fuel depth information is available and varies spatially. To cast the new invariants we use the informati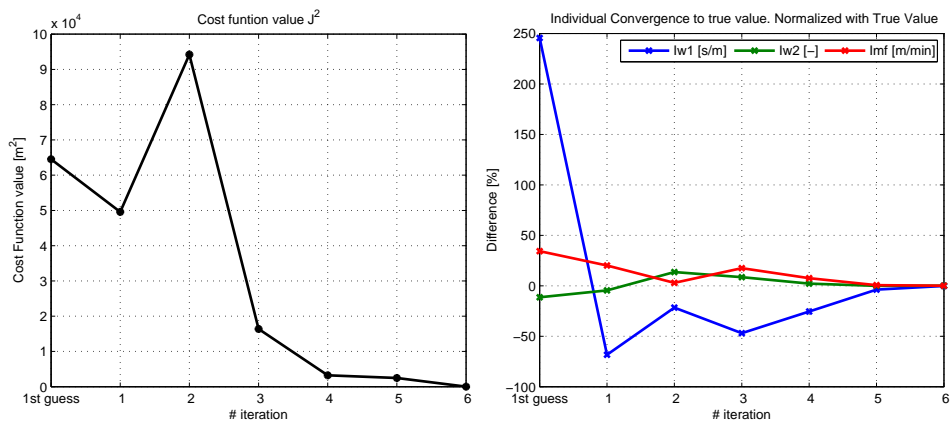on obtained with the sensitivity analysis performed on Rothermel's model (see section **??**). The analysis reveals that RoS is linearly related to fuel depth δ as first approximation. Thus, the RoS can be written now as:

$$RoS = I_{mfw} \cdot \delta(x, y)$$

Where $\delta(x, y)$ varies spatially.

The wind contribution is now included in $RoS = I_{mfw}$ and therefore we have to create a new parameter that accounts for the shape of the elliptical firelets (i.e. the eccentricity): I_{LB}. Where LB stands for length to breadth ratio. This invariant also depends on wind speed and, thus, it is not independent of I_{mfw}. This does not affect the capacity of our forecasting model since I_{LB} could be interpreted as a shaping factor and the way it is used in the forward model (only in the Huygens expansion part) prevents it from being mixed with I_{mfw}. As in the previous cases the wind direction invariant I_θ is required to close the invariants cast.

The influence of assimilating space dependent variable is that RoS now also depends on the location. This adds an extra non-linear behaviour to the model, since now when the fire front location changes, the RoS changes as well. Despite this higher complexity, our algorithm handles it in the optimisation loop and correctly identifies the invariants (figure 3.13).
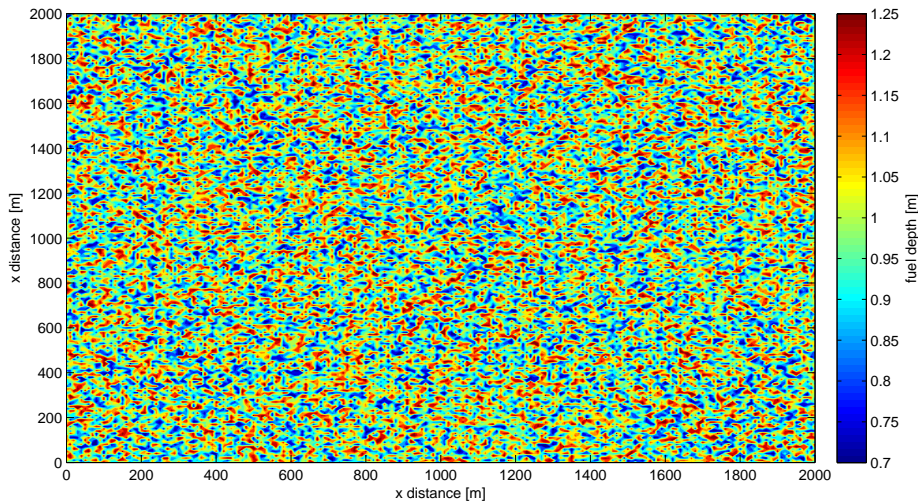

Figure 3.11: Ilustration of the fuel depth used as input. The forest is simulated by randomly varying a ±30% base fuel depth value (1m in this case).
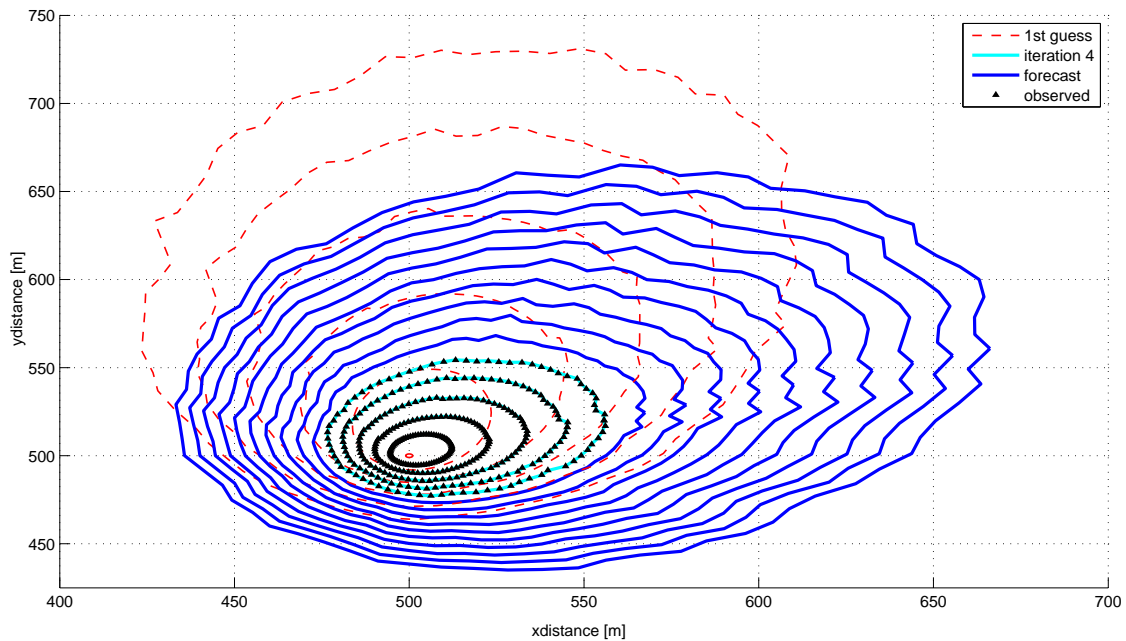
Figure 3.12: Five assimilated fire fronts with 1 minute intervals (black solid lines). The first guess (red dashed line) is taken to be far from the true invariants vector to check the algorithm capability to converge to the true invariants value. A ten min forecast (blue solid lines) is also calculated using fuel depth as an input.
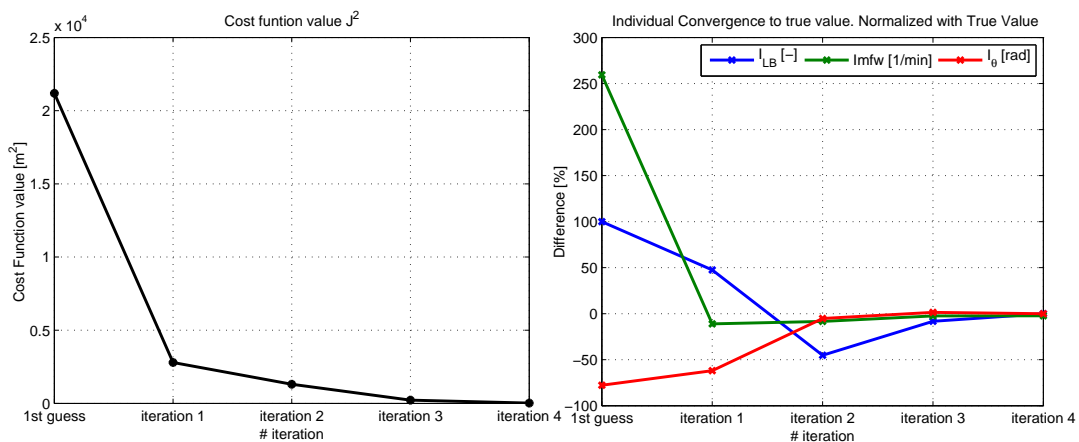
Figure 3.13: Cost function and invariants convergence when fuel depth is input

3.5 Lead time

The lead time for the four different implementations discussed above is investigated by feeding them different number of fire fronts and recording the computing time to deliver a 30 min forecast. It is worth mentioning that the number of assimilated fronts does not necessarily have to be the total assimilating time since it depends on the assimilation frequency (i.e. the number of assimilations per unit of time). Changing this frequency has a minor influence on the computing time since its contribution is linear in our forward model but might be important if more complex forward models are to be used (such as CFD based, for example). The Rothermel's variables that generate the synthetic data and the educated guess were kept constant for all the scenarios when they were not an input parameter (as wind speed, wind direction or fuel depth).

Figure 3.14 depicts the computing time time versus the number of assimilated fronts. The invariant cast for the situation when wind speed and direction are known parameters (red solid line) turns out to be the faster case. As expected, decreasing the number of invariants to be identified, speeds up the model since the dimension of the the matrices involved in the optimisation process decreases. The exception is when fuel information is inputted (light blue lines). The spatial dependency of the fuel depth and the fact that RoS has to be recalculated in every node raises the computing time and thus, this case is the slower one. The effect of feeding the algorithm with wind speed becomes noticeable above 16 assimilated fronts when the complexity of the fire fronts shapes increases the number of iterations required to reach convergence.

Despite these significant differences, when eight fronts are assimilated the forecast is delivered in less than a min and even when 24 fronts are assimilated the lead time is well above 25 min for a 30 minutes forecast.

A laptop with dual processor core of 2.2GHz is used as a computational tool since (as stated in the initial requirements) the forecasting algorithm must be suitable for normal computers.
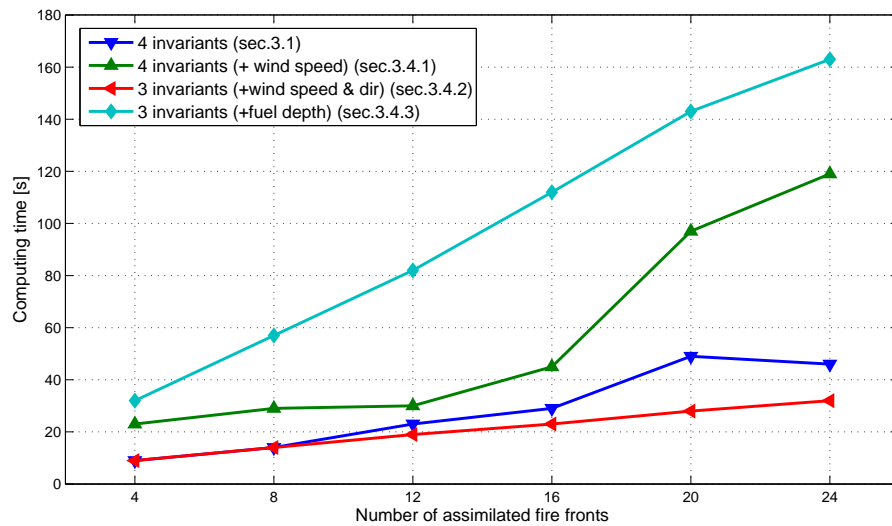

Figure 3.14: Computing time for all the implementations of the forecasting algorithm studied (referencing section displayed in the legend).

3.6 Effect of Perturbed Data

In all the scenarios investigated so far the cost function decreased until zero because the synthetic data was generated using similar froward models and, thus, there existed a true value that generates the observed fire fronts. To test the forecasting algorithm in a situation where such a true vector does not exist any more (and perfect convergence is then impossible), the synthetic data used in the fuel depth input data case (section 3.4.3) has been randomly perturbed with an error uniformly distributed in the range of $[0 \pm 10]$m. Apart from exploring the responds of the forecasting algorithm in a case where the forward model cannot properly describe the fire locations, this test can be seen as a sensitivity check of the errors and accuracy involved in data acquisition.

As expected, the best optimisation does not match now the observations perfectly and, thus, the cost function converges to a value of $2500m^2$ instead of zero (see figure 3.16). Despite this fact the convergence of the invariants -towards the values used to generate the unperturbed synthetic data- is still reached with an error lower than 5%.

Figure 3.15 shows the observed fronts and the corresponding optimisation after four iterations. The sharp corners in the observed perimeters are due to both the random distribution of the fuel depth (see figure 3.11) and the added error.
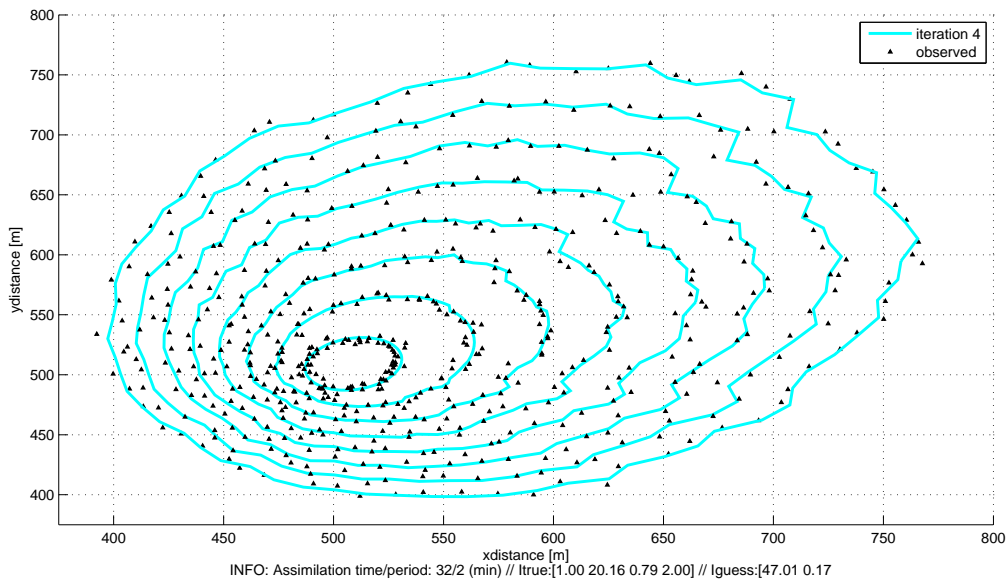
Figure 3.15: Perturbed fire fronts (black lines) correctly assimilated after 4 integrations. The invariants cast used is this of section 3.4.3 where fuel depth is used as an input and three invariants are identified.
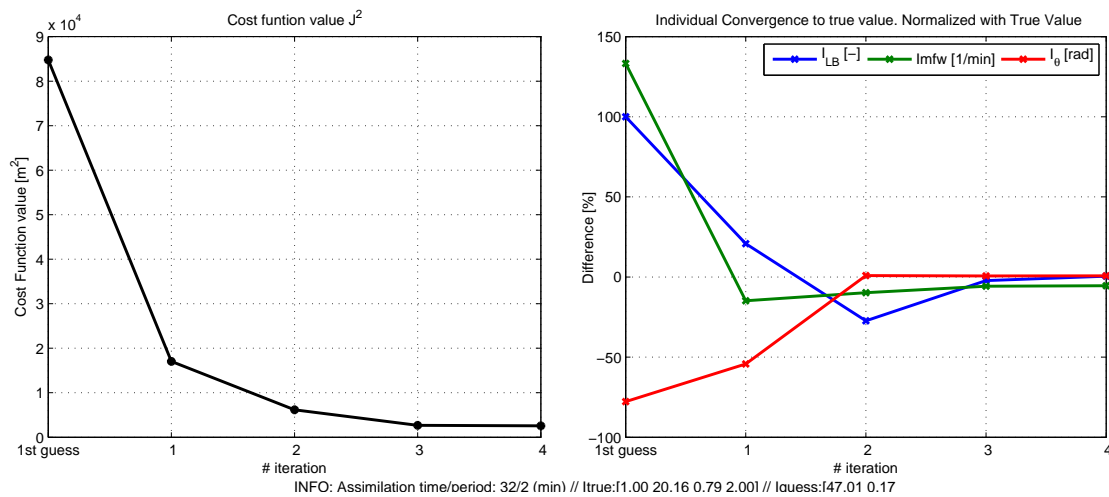
Figure 3.16: Convergence of cost function and invariants when perturbed data is assimilated

More tests performed while extending the error added to observation demonstrate that the algorithm manages to assimilate the observations with up to 20m of perturbations. The result that are the optimised front are smoother since the invariants are constant along the front as depicted in figure 3.17. The invariants also converged in this case which demonstrates the potential of this forecasting algorithm even when scarce and inaccurate data is available. These results open the door to further development the algorithm in these circumstances.
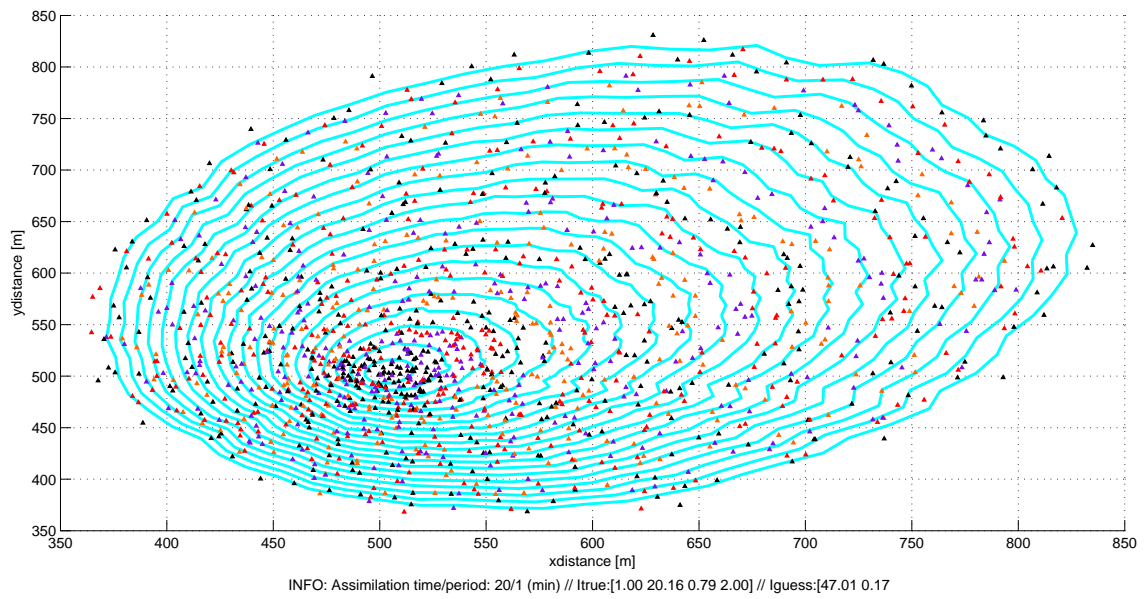
Figure 3.17: Assimilation of 20 perturbed fire fronts. The assimilated data are displayed with coloured triangles (black-red-purple-orange) to distinguish between observing times. The light blue lines are the fire fronts generated with the identified invariants.

Chapter 4

Conclusions and Future Work

In this thesis we set up a novel methodology to forecast wildfire dynamics based on data assimilation. We implemented it and explored its performance by focusing on wind-driven wildfires and assimilating synthetically generated front locations at different times. Rothermel's spread theory and Huygens expansion principle were chosen as the basis for casting the invariants of the system.

The forecasting algorithm uses direct automatic differentiation and a tangent linearization of the forward model to solve the optimisation problem. This strategy showed great efficiency finding the invariants within less than a minute and 10 iterations (runs of TLM).

The first attempt to identify 4 invariants encountered with a problem due to multiplicity in the determination of the RoS. In some scenarios, and depending on the initial guess and the assimilation window, the invariants I_{mf} and I_w could not be correctly identified and they converged to a wrong value. The multiplicity reduces when more fronts were assimilated although a minimum number of assimilations could not be determined because the misconvergence kept present for some first guess configurations. The fact that two invariant could be misidentified, prevented the forecasting algorithm to reach its potential since both invariants needed to be kept constant to deliver a correct forecast. Three ways to overcome the multiplicity problem were identified; casting only one invariant for RoS, extending the forward model so it predicts extra parameters (as flame hight or heat release rate) and assimilate them, or including extra information about the system to break the multiplicity. The later was implemented and evaluated in 3 scenarios: wind speed, wind speed and wind direction and fuel depth.

For all the scenarios where additional information was used, all the invariants were correctly identified, even when the first guess greatly differed from the true value. In those cases the forecast could be improved with additional data (fuel depth or wind forecast) to deliver a more reliable forecast. The computing time was of the order of one minute (less than three in the longest simulation studied) and thus, all the implementations had a positive lead time (time ahead of the event). The most computing expensive implementation is the one that uses fuel depth since the RoS varies in each node of the front.

The forecasting algorithm implementation that uses wind speed information, was evaluated with altered synthetic data. This recreates the situation where a set of invariants that can match the observations does not exist. In this case the forecasting algorithm successfully estimates the invariants closely fitting the fire front. A random deviation of up to 20m from each perimeter node was used to alter the synthetic data.

After all the evaluations performed to the different implementation of the forecasting algorithm, we can conclude that we set up the framework for a promising forecasting tool and proposed first simple yet powerful implementations.

Future work

The next required step to make this algorithm operational is to check it with real data. This could not be done during this thesis due to temporal constraints. To keep developing the methodology some identified limitations should be tackled as spotting fires -which does not follow classical (i.e. Rothermal's) fire spread and the capacity of the forecasting algorithm to deal with uncertainties caused by the lack of reliable data, and deliver probabilistic values as outputs. To pursue this, we propose to explore more sophisticated forward models (e.g pyrolisis, CFD) and greatly increase the number of invariants to several dozen. Then the automatic direct differentiation should be switched to adjoint differentiation (adjoint modelling approach) to keep the low computational cost requirement.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Guillermo Rein for always encouraging me with his undying enthusiasm that overcame any distance or obfuscation. Thank you as well for giving me the opportunity to work at Imperial College London as a Visiting Scholar in the course of researching and writing this thesis.
I am also thankful to Dr. Wolfram Jahn for wisely guiding me through the inverse world.

A very special thanks goes out to Professor Drysdale who contributed his wisdom and charm, and the Fire Group people for their motivation and inspiring environment that undoubtedly enriched my stay in Edinburgh. Agustín opened the Fire Group doors from the very first day and made it possible, *gracias*!

Thanks are also due to Critián Maluk for the several inspiriting philosophical talks on fire science and the long working days (and nights) spent together.

Whenever MATLAB or LATEX defeated my patience, Xavier Clotet gave me a helping hand. Thanks! I am also grateful to Emilia Wrobel, who put in time to correct this thesis. All remaining mistakes are entirely mine.

This thesis is a culmination of a two year long journey during which I walked with many people who encouraged me along the way: IMFSE mates, Maria, Jonathan, thanks for sharing it with me.

Bibliography

[1] G L. Achtemeier. Field validation of a free-agent cellular automata model of fire spread with fire - atmosphere coupling. *International Journal of Wildland Fire*, 22(2):148, 2013.

[2] F A Albini. *Estimating wildfire behavior and effects.* Intermountain Forest and Range Experiment Station, Forest Service, US Department of Agriculture, 1976.

[3] M E Alexander. Estimating the length-to-breadth ratio of elliptical forest fire patterns. In *Proceedings of the eighth conference on fire and forest meteorology*, volume 29, pages 4–85, 1985.

[4] D H Anderson, E A Catchpole, N J De Mestre, and T Parkes. Modelling the spread of grass fires. *Journal of the Australian Mathematical Society*, 23:451–466, 1982.

[5] Hal E Anderson and Intermountain Forest. *Predicting wind-driven wild land fire size and shape.* US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 1983.

[6] PARKES T. ANDERSON D H, CATCHPOLE E. A., DE MESTRE N. J. Modelling the spread of grass fires, 1982.

[7] CM Belcher, Yearsleya M.J, RM. Hadden, J C. McElwaina, and G Rein. Baseline intrinsic flammability of Earth's ecosystems estimated from paleoatmospheric oxygen over the past 350 million years. *Proceedings of the . . .*, 2010.

[8] GM Byram and WL Fons. Thermal properties of forest fuels. *US Forest Service, Department of Agriculture*, 1952.

[9] JR Coleman and A L Sullivan. A real-time computer application for the prediction of fire spread across the Australian landscape. *Simulation*, 1996.

[10] MG Cruz and Jim Gould. Field-based fire behaviour research: past and future roles. *Proceedings of the 18th World IMACS Congress*, (July):247–253, 2009.

[11] MG Cruz, A L Sullivan, and Jim Gould. Anatomy of a catastrophic wildfire: the Black Saturday Kilmore East fire in Victoria, Australia. *Forest Ecology and . . .*, 2012.

[12] B Drossel and F Schwabl. Self-organized critical forest-fire model. *Physical Review Letters*, 69(11):1629–1632, 1992.

[13] Dougal Drysdale. *An Introduction to Fire Dynamics.* John Wiley & Sons, 2011.

[14] HW Emmons. Fire in the forest. *Fire Research Abstracts and Reviews*, 1964.

[15] Mark Finney. *FARSITE, Fire Area Simulator–model development and evaluation*, volume 3. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998.

[16] Mark Finney. Fire growth using minimum travel time methods. *Canadian Journal of Forest Research*, 32(8):1420–1424, 2002.

[17] Mark Finney, Jack D. Cohen, Sara S. McAllister, and W. Matt Jolly. On the need for a theory of wildland fire spread. *International Journal of Wildland Fire*, 22(1):25, 2013.

[18] WL Fons. Analysis of fire spread in light forest fuels. *Journal of Agricultural Research*, 1946.

[19] GIDAI, editor. *Modeling Wildland Fire Spread Using an Eulerian Level Set Merhod and High Resolution Numerical Weather Prediction*. INTERNATIONAL CONGRESS Fire Computer Modeling, 2012.

[20] David G. Green. Shapes of simulated fires in discrete fuels. *Ecological Modelling*, 20(1):21–32, October 1983.

[21] D.G. Green, a.M. Gill, and I.R. Noble. Fire shapes and the adequacy of fire-spread models. *Ecological Modelling*, 20(1):33–45, October 1983.

[22] Andreas Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Number 19 in Frontiers in Appl. Math. SIAM, Philadelphia, PA, 2000.

[23] Rickard Hansen. Estimating the amount of water required to extinguish wildfires under different conditions and in various fuel types. *International Journal of Wildland Fire*, 21(6):778, 2012.

[24] LF Hawley. Theoretical considerations regarding factors which influence forest fires. *Journal of Forestry*, 1926.

[25] Christian H.Bischof, H Martin Bücker, Bruno Lang, A Rasch, and Andre Vehreschild. Combining Source Transformation and Operator Overloading Techniques to Compute Derivatives for {MATLAB} Programs. In *Proceedings of the Second {IEEE} International Workshop on Source Code Analysis and Manipulation ({SCAM} 2002)*, pages 65–72, Los Alamitos, CA, USA, 2002. IEEE Computer Society.

[26] F A Heinsch and P L Andrews. *BehavePlus fire modeling system, version 5.0: design and features*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 2010.

[27] HC Hottel. Modelling of fire spread trhough a fuel bed. In *Tenth Symposium (International) on Combustion, The Combustion Institute*, pages 997–1007, 1965.

[28] C Huygens. Traité de la lumière. 1678.

[29] W Jahn, G Rein, and J L Torero. Forecasting fire growth using an inverse zone modelling approach. *Fire Safety Journal*, 46(3):81–88, 2011.

[30] W Jahn, G Rein, and J L Torero. Forecasting fire dynamics using inverse computational fluid dynamics and tangent linearisation. *Advances in Engineering Software*, 47(1):114–126, 2012.

[31] I Knight and J Coleman. A Fire Perimeter Expansion Algorithm-Based on Huygens Wavelet Propagation. *International Journal of Wildland Fire*, 3(2):73, 1993.

[32] Marc Lacey. Intuition, Fighting Wildfires With Computers and.

[33] RR Linn and FH Harlow. FIRETEC: a transport description of wildfire behavior. 1997.

[34] Yongqiang Liu, John Stanturf, and Scott Goodrick. Trends in global wildfire potential in a changing climate. *Forest Ecology and Management*, 259(4):685–697, February 2010.

[35] C Lucas, K Hennessy, G Mills, and J Bathols. Bushfire Weather in Southeast Australia: Recent Trends and Projected Climate Change Impacts. 2007.

[36] Stephanie L. MacPhee, Jason I. Gerhard, and G Rein. A novel method for simulating smoldering propagation and its application to STAR (Self-sustaining Treatment for Active Remediation). *Environmental Modelling & Software*, 31:84–98, May 2012.

[37] V Mallet, D E Keyes, and F E Fendell. Modeling wildland fire propagation with level set methods. *Computers & Mathematics with Applications*, 57(7):1089–1101, 2009.

[38] J Mandel, J D Beezley, J L Coen, and M Kim. Data assimilation for wildland fires. *Control Systems, IEEE*, 29(3):47–65, 2009.

[39] J Mandel, L S Bennethum, J D Beezley, J L Coen, C C Douglas, M Kim, and A Vodacek. A wildland fire model with data assimilation. *Mathematics and Computers in Simulation*, 79(3):584–606, 2008.

[40] J. Margerit and O. Séro-Guillaume. Modelling forest fires. Part II: reduction to two-dimensional models and simulation of propagation. *International Journal of Heat and Mass Transfer*, 45(8):1723–1737, April 2002.

[41] AG McArthur. Weather and grassland fire behaviour. 1966.

[42] W Mell, MA Jenkins, Jim Gould, and Phil Cheney. A physics-based approach to modelling grassland fires. *International Journal of Wildland Fire*, 16(1):1, 2007.

[43] F Morandini, X Silvani, JL Rossi, and PA Santoni. Fire spread experiment across Mediterranean shrub: Influence of wind on flame front properties. *Fire Safety Journal*, 2006.

[44] D. Morvan, S. Méradji, and G. Accary. Physical modelling of fire spread in Grasslands. *Fire Safety Journal*, 44(1):50–61, January 2009.

[45] M Mutlu, SC Popescu, C Stripling, and T Spencer. Mapping surface fuel models using lidar and multispectral data fusion for fire behavior. *Remote Sensing of . . .*, 2008.

[46] J Nahmias and H Téphany. Fire spreading experiments on heterogeneous fuel beds. Applications of percolation theory. *Canadian Journal of . . .*, 2000.

[47] Michelle Nijhuis. Forest fires: Burn out. *Nature*, 489(7416):352–4, September 2012.

[48] Jorge Nocedal and Stephen J Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 1999.

[49] K J Overholt, J Cabrera, A Kurzawski, M Koopersmith, and O A Ezekoye. Characterization of Fuel Properties and Fire Spread Rates for Little Bluestem Grass. *Fire Technology*, pages 1–30, 2012.

[50] E Pastor, L Zarate, E Planas, and J Arnaldos. Mathematical models and calculation systems for the study of wildland fire behaviour. *Progress in Energy and Combustion Science*, 29(2):139–153, 2003.

[51] Juli G. Pausas and Jon E. Keeley. A Burning Story: The Role of Fire in the History of Life. *BioScience*, 59(7):593–601, July 2009.

[52] GB Peet. The shape of mild fires in Jarrah forest. *Australian Forestry*, 1967.

[53] D Perroy and M H Fernandez. fireparadox.org, 2010.

[54] G D Richards. An elliptical growth model of forest fire fronts and its numerical solution. *International Journal for Numerical Methods in Engineering*, 30(6):1163–1179, 1990.

[55] 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.

[56] G D Richards and RW Bryce. A Computer Algorithm for Simulating the Spread of Wildland Fire Perimeters for Heterogeneous Fuel and Meteorological Conditions. *International Journal of Wildland Fire*, 5(2):73, 1995.

[57] M Rochoux, A Trouvé, and S Ricci. Forest Fire Propagation using Data Assimilation. Master's thesis, Dept. of Engineering Mathematics, Numerical Modeling. INSA, Toulouse, 2010.

[58] M C Rochoux, B Delmotte, B Cuenot, S Ricci, and A Trouvé. Regional-scale simulations of wildland fire spread informed by real-time flame front observations. *Proceedings of the Combustion Institute*, 2012.

[59] JL Rossi, T. Molinier, M. Akhloufi, A. Pieri, and Y. Tison. Advanced stereovision system for fire spreading study. *Fire Safety Journal*, pages 1–9, December 2012.

[60] R C Rothermel and I Forest. *A mathematical model for predicting fire spread in wildland fuels*. Intermountain Forest & Range Experiment Station, Forest Service, US Department of Agriculture, 1972.

[61] JC Sanderlin and RJ Van Gelder. A simulation of fire behavior and suppression effectiveness for operation support in wildland fire management. *Proceedings of the 1st International Conference on ...*, 1977.

[62] Joe H Scott and Robert E Burgan. Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model. *The Bark Beetles, Fuels, and Fire Bibliography*, page 66, 2005.

[63] O. Séro-Guillaume and J. Margerit. Modelling forest fires. Part I: a complete set of equations derived by extended irreversible thermodynamics. *International Journal of Heat and Mass Transfer*, 45(8):1705–1722, April 2002.

[64] O. Séro-Guillaume, S. Ramezani, J. Margerit, and D. Calogine. On large scale forest fires propagation models. *International Journal of Thermal Sciences*, 47(6):680–694, June 2008.

[65] B. J. Stocks, T. J. Lynham, B. D Lawson, M E Alexander, C. E. Van Wagner, R. S. McAlpine, and D. E. Dubé. Canadian Forest Fire Danger Rating System: An Overview. *The Forestry Chronicle*, 1989.

[66] 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.

[67] A L Sullivan. Wildland surface fire spread modelling, 1990–2007. 2: Empirical and quasi-empirical models. *International Journal of Wildland Fire*, 18(4):369–386, 2009.

[68] A L Sullivan. Wildland surface fire spread modelling, 1990–2007. 3: Simulation and mathematical analogue models. *International Journal of Wildland Fire*, 18(4):387–403, 2009.

[69] P.H. Thomas. Some Aspects of the Growth and Spread of Fire in the Open. *Forestry*, 40(2):139–164, 1967.

[70] P.H. Thomas and Dougal Drysdale. Letter to the Editor. *Fire Safety Journal*, pages 1–2, November 2012.

[71] R Tsai and S Osher. Review article: Level set methods and their applications in image science. *Communications in Mathematical Sciences*, 2003.

[72] C Tymstra, RW Bryce, and BM Wotton. *Development and structure of Prometheus: the Canadian wildland fire growth simulation model*. 2010.

[73] D. X. Viegas and A Simeoni. Eruptive Behaviour of Forest Fires. *Fire Technology*, 47(2):303–320, October 2010.

[74] FA Williams. Urban and wildland fire phenomenology. *Progress in Energy and Combustion Science*, 1982.

[75] Jerry Williams. Exploring the onset of high-impact mega-fires through a forest land management prism. *Forest Ecology and Management*, 294(null):4–10, April 2013.

[76] S Wolfram. *A new kind of science*. Wolfram media Champaign, 2002.

Appendices

Appendix A

Rothermel's Model Equation and Correlations

The original Rothermel's formula for RoS needs 11 input variables:

$$RoS = f(W_o, \delta, \sigma, Mf, Mx, U, \phi_s, h, \rho_p, S_t, S_e) \tag{A.1}$$

Where:

- W_o = Ovendry fuel loading $[kg/m^2]$: Quantity of fuel per unit area.

- δ = Fuel depth $[m]$: Height of the surface fuel canopy.

- σ = Fuel particle surface-area-to-volume ratio $[1/m]$

- Mf = Moisture content[%]: Moisture content of fuel.

- Mx = Moisture content of extinction[%]: Experimental minimum value from which the combustion is no longer taking place

- U = Wind velocity $[m/s]$: Wind velocity at midflame height

- ϕ_s = Slope steepness

- h = Fuel's heat of combustion $[kJ/kg]$: Nominal[†1] value= $18.6MJ/kg$

- ρ_p = Ovendry particle density $[kg/m^3]$: Nominal[†] value= $512kg/m^3$

- S_t = Fuel particle mineral content: Content of incombustible material in the fuel: Nominal[†] value= 5.55%

- S_t = Fuel particle effective mineral content: Content of incombustible material without considering silica since it was prove that it does not effect the pyrolization rate: Nominal[†] value= 5.55%

The explicit equations and correlations are (original unites are not in SI):
Rate of Spread:

$$RoS = \frac{\gamma_r \xi}{\rho_b \varepsilon Q_{ig}} (1 + \Phi_w + \Phi_s)$$

Reaction intensity B.t.u/ft^2 min

$$\gamma_r = \bar{\Gamma} w_n h \eta_m \eta_s$$

---

[1][†]The values corresponding to the last 4 variables were estimated by Albini[2] in 1976 and are used in all Rothermel's implementations since then.

where:

$\bar{\Gamma} = \bar{\Gamma}_{max}(\beta/\beta_{op})^A exp[A(1 - \beta/\beta_{op})]$ Optimum reaction velocity (min^{-1})

$\bar{\Gamma}_{max} = \sigma^1.5(495 + 0.494\sigma^1.5)^{-1}$ Maximum reaction velocity (min^{-1})

$\beta_{op}3.348\sigma^{-0.818}$ Optimum packing ratio

$A = 1/(4.774\sigma^0.1 - 7.27)$

$\eta_M = 1 - 2.59\frac{M_f}{M_x} + 5.11(\frac{M_f}{M_x})^2 - 3.52(\frac{M_f}{M_x})^3$ Moisture damping coefficient

$\eta_s = 0.174S_e^{-0.19}$ Mineral damping coefficient

$\xi = (192 + 0.259\sigma)^{-1}exp[(0.792 + 0.681\sigma^0.5)(\beta + 0.1)]$ Propagating flux ratio

$\Phi_w = CU^B \left(\frac{\beta}{\beta_0}\right)^{-E}$ Wind coefficient

$C = 07.47exp(-0.133\sigma^0.55)$

$B = 0.02526\sigma^0.54$

$E = 0.715exp(-3.59 \cdot 10^{-10}\sigma)$

$w_n = \frac{w_0}{1+S_T}$ Net fuel loading (lb/ft^2)

$\Phi_s = 5.275\beta^{-0.3}(tan\Phi)^2$ Slope factor

$\rho_b = w_0\ \delta$ Ovendry bulk density (lb/ft^3)

$\epsilon = exp(-138/\sigma)$ Effective heating number

$Q_{ig} = 250 + 1.116M_f$ Heat of preignition btu/lb

$\beta = \frac{\rho_b}{\rho_p}$ Packing ratio

Appendix B

Sensitivity Analysis of Rothermel's Model

When it comes to casting the invariants and linearise the forward model it is important to know how it behaves. To check this, RoS response to Rothermel's inputs is studied and displayed in figure B.1. All six functions are well-behaved since they are smooth and no mathematical singularities are present. This is of importance when performing the TLM.
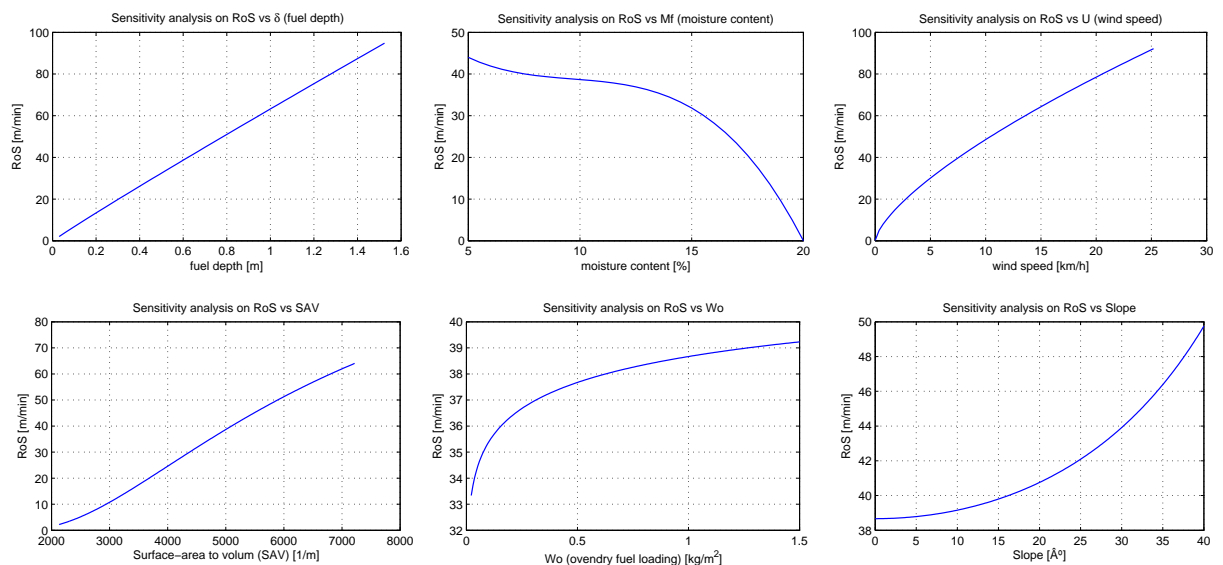The base values and the sweeping range are taken from table 2.2.

Figure B.1: Sensitivity analysis on RoS depending on Rothermel's model variables

Appendix C

Software CD

A CD is attached to this thesis with the MATLAB® programs scripted to implement the forecasting method and evaluate the different scenarios presented above.
The folders are named as follow.

Folder name	Folder information
`4-invariants`	4 invariants without extra data. Section 3.2
`wind-input`	3 invariants with wind speed information. Section 3.4.1
`wind-dir-input`	3 invariants with wind speed and direction information. Section 3.4.2
`fuel-input`	3 invariants with fuel depth information. Section 3.4.3
`Synthetic`	Synthetic data program
`Extras`	Additional programs scripted

Table C.1: Information of the attached folders

Each of the 4 forecasting algorithm implementations needs 11 extra MATLAB functions included in the respective folders. The folder called `synthetic` contains the program to generate the synthetic data used. Additional functions, required to perform the forward automatic differentiation (`Adimat`), are located in `Extras` folder together with a sensitivity analysis and plot functions used to study and represent the different forecasting algorithms implemented.

Appendix D

Conference abstract

The following paper was submitted and accepted for oral presentation at the *Numerical Wildfires 2013: Numerical simulation of Forest Fires, from combustion to emissions workshop*. It will take place in Corsica from 13 to 18 May 2013 at IESC Cargèse, and is organised by CNRS SPE, University of Corsica. The author of this thesis will present it on Thursday, May 17.

Forecasting wildfire dynamics for wind-driven conditions using Huygens' front expansion and tangent linearisation

Oriol Rios[1], Wolfram Jahn[2] and Guillermo Rein[*2]

[1]University of Edinburgh
[2]Imperial College London

March 2013

Wildland fires are complex multiscale phenomena involving a series of nonlinear and transient processes. They are governed by coupled physicochemical interactions that are not yet known which pose the main obstacle when trying to model their behaviour. It is not only about solving a set of equations that do not have an analytical solution or known initial values (as it is in weather forecasting) but solving a problem without even having the complete set of equations (Finley et al. 2013).

A technology able to forecast wildland fire dynamics would lead to a paradigm shift in the response to emergencies, providing the Fire Services with essential information about the on-going fire. The main objective is to enhance the protection of citizens, property and the environment by providing a key tactical tool to fight fires. This paper presents a methodology to forecast wildfire dynamics posed as an inverse problem to solve for the invariants governing the spread of wildfire and using a tangent linear approach for quick optimisation.

The key is to find a computational algorithm and combines a physical model with data assimilation that reliably delivers an accurate forecast within a positive lead time, (i.e. enough time before the event) so that emergency responders can make tactic decisions and deploy resources ahead of it. However, the state-of-the-art of computational wildfire dynamics is not fast or accurate enough to provide valid predictions on time. Current wildfire models do not achieve lead times mainly because they take too long to provide a solution (negative lead times) or require data that is impossible to sense (e.g. flame temperature or detailed fuel and flow characterisation). The solution is the assimilation of real-time sensor data, which has been shown to greatly accelerate fire simulations without loss of forecast accuracy (Mandel et al 2008; Jahn et al. 2011; Jahn et al. 2012; Rochoux al, 2012).

Data assimilation of wildfire was first studied by Mandel et al 2008 who reproduced the time-temperature curve of a sensor placed in the way of an advancing wildfire by coupling a reactive-diffusion system of equations with a Kalman filter. Following this work, Rochoux et al., 2012 applied real-time data assimilation to predict the wildfire front location at the local scale using infrared observations. They used Rothermel's theory for the rate of spread with one parameter and coupled it with an Eulerian level set method to simulate the front spread. Data assimilation was also done with a Kalman filter to balance computational and observational errors.

---

[*]g.rein@imperial.ac.uk

By means of data assimilation and solving an inverse problem, we propose a new methodology to forecast the fire front. Our forward model is simple but powerful and captures the fundamentals of wildlife dynamics. It couples the full theory of Rothermel's rate of spread with a perimeter expansion model based on Huygens principle (Richards 1990). Huygens principle assumes that the spread of wildfire is given by combination of multiple ellipses propagating from the previous front position as illustrated in Figure 1. This is the same principle used in FARSITE.
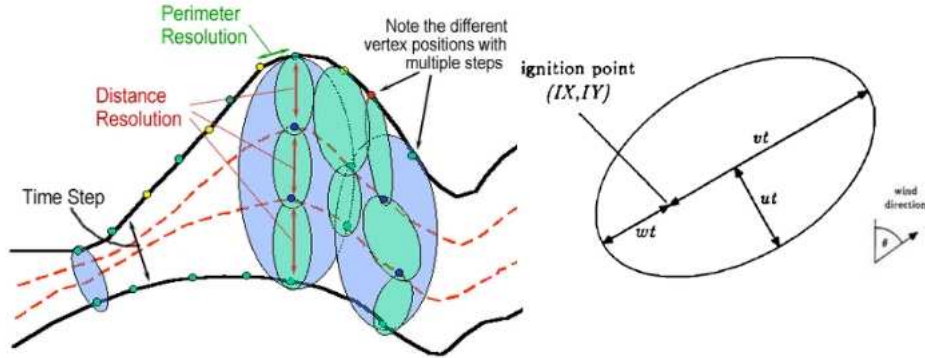

Figure 1: Illustration of the geometric principle in Huygen's expansion of ellipses used in Richards 1990 model for wildfire spread. From Farsite and Richards

We formulate the problem based on the premise that some invariant exists following the contributions of Jahn et al. (2011 and 2012) on forecasting fire dynamics in enclosures. The invariants are the set of governing parameters that are mutually independent and constant for a significant amount of time. If the physicochemical parameters cast into each invariant remain constant for a significant period of time, so does its effect on the fire front. Therefore, this methodology formulates the existence of fundamental invariants and calculates their contribution in the forward model by assimilating the position of the wildfire front in real time. After an assimilation window of observations, the invariant converges to a value that is used to forecast the perimeter expansion in the subsequent time interval. This forecast is then accurate until any of the invariants change significantly, which would be detected by the continuous feed of sensor data. For our forward model, we identified a minimum of six invariants: moisture damping (I1), fuel factor (I2), wind speed factor (I4), wind direction (I5), breadth-to-length spread ratio (I3) and slope (I6). More invariants can be formulated adding complexity to the forecast, but what we explore here is the first the simplest yet still meaningful set.

Rothermel's Rate of Spread:

$$RoS = I_{mf}(1 + I_w \cdot I_u)$$

Huygen's perimeter expansion:

$$I_\theta = \theta$$

Where θ is the wind direction.

The formulation of the problem is general enough that it is suitable to work with airborne observations from drones, helicopters or satellites, as well as ground observations from permanent sensors and deployed units. It can also integrate information typically gathered by forestry agencies (e.g., topography GIS, fuel distribution maps, weather conditions, location of transport route and urban interfaces). On top of this, we formulate the problem in such a way that depending on the available sensor data, the invariants can be turned into input data for increased accuracy and speed. For example, if reliable wind speed data arrives, there is no need to solve

for it but instead it is directly used as input to the spread model. This makes the methodology suitable for a large range of different wildfire situations and for an ever-evolving technological support.

In this paper we present the general formulation of the problem and then explore the results in a series of large-scale wildland fire cases with synthetic data. The sensibility of the forecast to data from different sources (i.e. the range of frequencies and spatial accuracies) is also explored. Figure 2 shows an example. We focus on wind-driven fires, where the slope is not a parameter, and hence solving for five invariants. The results show the high capacity of the method to predict the location of the fire front quickly.
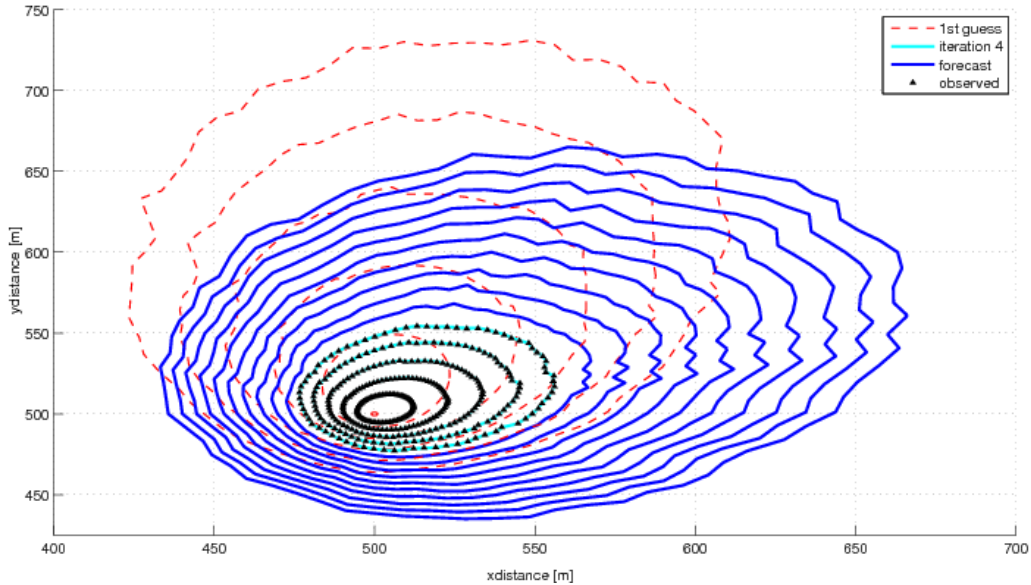

Figure 2: Example of predicted wildfire front spread using synthetic data (black triangles) for the case of short grass fuel under a constant wind speed and using synthetic fuel depth information.

This work opens the door to a further advanced framework with more sophisticated forward models (i.e. pyrolysis models, CFD) and greatly increases the number of invariants to several dozen by means of an adjoint modelling approach.

References

Finney, M., Cohen, J. D., McAllister, S. S., & Jolly, W. M. (2013). On the need for a theory of wildland fire spread. International Journal of Wildland Fire, 22(1), 25.

Jahn, W., Rein, G., & Torero, J. L. (2012). Forecasting fire dynamics using inverse computational fluid dynamics and tangent linearisation. Advances in Engineering Software, 47(1), 114–126.
Jahn, W, G Rein, JL Torero, Forecasting Fire Growth using an Inverse Zone Modelling Approach, Fire Safety Journal 46, pp. 81–88, 2011

Mandel, J., Beezley, J. D., Coen, J. L., & Kim, M. (2009). Data assimilation for wildland fires. Control Systems, IEEE, 29(3), 47–65.

Richards, G. D. (1990). An elliptical growth model of forest fire fronts and its numerical solution. International Journal for Numerical Methods in Engineering, 30(6), 1163–1179.

Rochoux, M. C., Delmotte, B., Cuenot, B., Ricci, S., & Trouvé, A. (2012). Regional-scale simulations of wildland fire spread informed by real-time flame front observations. Proceedings of the Combustion Institute.

Rothermel, R. C., & Forest, I. (1972). A mathematical model for predicting fire spread in wildland fuels. Intermountain Forest & Range Experiment Station, Forest Service, US Department of Agriculture.