



**ADVANCES IN
FOREST FIRE
RESEARCH**

DOMINGOS XAVIER VIEGAS

EDITOR

2014

A Comparative Study of Parameter Estimation and State Estimation Approaches in Data-Driven Wildfire Spread Modeling

M.C. Rochoux^a, C. Emery^a, S. Ricci^a, B. Cuenot^a, A. Trouvé^b

^a CERFACS, SUC-URA1875, CNRS, , 42 Av. G. Coriolis, 31057 Toulouse (France);

melanie.rochoux@graduates.centraliens.net, sophie.ricci@cerfacs.fr, benedicte.cuenot@cerfacs.fr

^b Department of Fire Protection Engineering, University of Maryland, College Park, MD 20742 (USA); atrouve@umd.edu

Abstract

The objective of this work is to demonstrate the ability of the FIREFLY data-driven wildfire simulator to forecast the fire spread behavior over complex terrain topography. The prototype simulator features the following main components: a level-set-based fire propagation solver that adopts a regional-scale viewpoint, treats wildfires as propagating fronts, and uses a description of the local rate of spread (ROS) of the fire as a function of vegetation, topographical and meteorological properties based on Rothermel's model; a series of observations of the fire front location; and a data assimilation algorithm based on an ensemble Kalman filter (EnKF). The data assimilation algorithm also features a choice between a parameter estimation (PE) approach in which the estimation targets are the input parameters of the ROS model, and a state estimation (SE) approach in which the estimation targets are the spatial coordinates of the discretized fire front. This study shows the extension of the FIREFLY data-driven simulator to complex terrain topography. The fire propagation is represented by time-evolving two-dimensional fronts along the horizontal plane in order to remain consistent with the formulation of the PE-/SE-based EnKF algorithms that were initially developed for flat terrain configuration. While evaluated on synthetic cases in this study, the performance of the EnKF algorithm is shown to be preserved in the case of complex terrain. Thus, this study emphasizes the potential of data assimilation to dramatically increase wildfire simulation accuracy in real-world wildfire events.

Keywords: *Data assimilation, Ensemble Kalman Filter, Wildfire, Fire modeling, Terrain topography*

1. Introduction

The challenges found on the route to developing quantitative fire models are two-fold. First, there is the classical modeling challenge associated with providing accurate mathematical representations of the multi-physics multi-scale processes that involve biomass pyrolysis, combustion, flow dynamics as well as atmospheric dynamics and chemistry (Sullivan 2009a, 2009b). Second, there is the less common data challenge associated with providing accurate estimates of the input parameters required by the models (Viegas 2011). Current regional-scale fire models are limited in scope because of the large uncertainties associated with the accuracy of physics-based models, because also of the large uncertainties associated with many of the environmental conditions that are required as input parameters to the fire problem (Rochoux 2014). A possible approach to overcome the limitations found in numerical simulations of wildfire spread as well as to develop predictive models that are compatible with operational applications is data assimilation (DA). This approach takes advantage of recent progress made in airborne remote sensing that is promising for real-time monitoring of the fire front location (Paugam *et al.* 2013). DA offers a convenient framework for integrating fire sensor observations into a computer model along with the estimated modeling/observation uncertainties with the goal to provide optimal estimates of the control variables and thereby to improve predictions of fire spread behavior. The application of DA and inverse modeling has recently been explored in the area of wildfire research (e.g. Mandel and Beezley 2007; Lautenberger 2013; Rochoux *et al.* 2013). The key idea is that, when used alone, neither measurements nor computer models can provide a reliable and complete description of the real state of the physical system.

The present study is an extension of our previous work in Rochoux *et al* (2013, 2014a, 2014b) and Rochoux (2014), in which a prototype data-driven wildfire spread simulator capable of forecasting the wildfire spread behavior was developed. The prototype simulator features the following main components: a level-set-based fire propagation solver (Rehm and McDermott 2009) that treats wildfires as propagating fronts and uses a description of the local rate of fire spread (ROS) as a function of vegetation, topographical and meteorological conditions based on the model due to Rothermel (1972); a series of observations of the fire front location; and a DA algorithm based on the ensemble Kalman filter (EnKF) to account for some of the non-linearities in the mapping between control space and observation space. The DA algorithm also features a choice between a parameter estimation (PE) approach in which the control variables are the input parameters of the ROS model (Rochoux *et al* 2014a) and a state estimation (SE) approach in which the control variables are the spatial coordinates of the fire front (Rochoux *et al* 2014b). This study presents the extension of the FIREFLY-based data-driven wildfire spread simulator to problems with complex terrain topography; the formulation of the PE-/SE-based DA strategies is not affected by the increasing complexity of the wildfire spread model. The paper is organized as follows. The data-driven wildfire spread model (including the fire spread model with complex terrain topography and the DA algorithms) is presented in Section 2; examples for each component of the data-driven simulator are also provided. Its performance is evaluated for synthetic DA experiments in Section 3.

2. Methodology: Data-driven wildfire spread model

As any current operational fire spread simulator, FIREFLY adopts a regional-scale perspective. This implies that wildfires are assumed to feature a front-like geometry at scales ranging from a few tens of meters up to several kilometers and may be described as a line that propagates normal to itself into unburnt vegetation and over complex topography, at a ROS denoted by Γ_{3D} .

2.1. Forward model: Fire spread model

2.1.1. Geometrical considerations

A rectangular Cartesian coordinate system (x_0, y_0, z_0) is introduced to geolocalize the propagation of the fire front over complex terrain topography. (x_0, y_0) represents the horizontal plane, with the x_0 -axis pointing towards the East direction and the y_0 -axis pointing towards the North direction; the z_0 -axis is the vertical direction ($z_0 = 0$ m represents the horizontal plane, where the terrain elevation h is zero). The reference frame is illustrated in Figure 1a.

Based on Sharples (2008) and Lautenberger (2013), the terrain topography is locally characterized in FIREFLY by a pair of angles noted (α_a, α_{sl}) , see Figure 1a. α_a [°] is the terrain aspect angle indicating the downslope direction and is defined in a clockwise representation, where 0° indicates the North direction (i.e. the y_0 -direction). α_{sl} [°] is the terrain slope angle, taking values between 0° (flat terrain) and 90° (vertical wall). The direction from which the wind blows is noted α_w [°] and is defined in a clockwise representation, where 0° indicates the North direction, see Figure 1b. In this framework, the direction of fire propagation is noted α_{fr} [°].

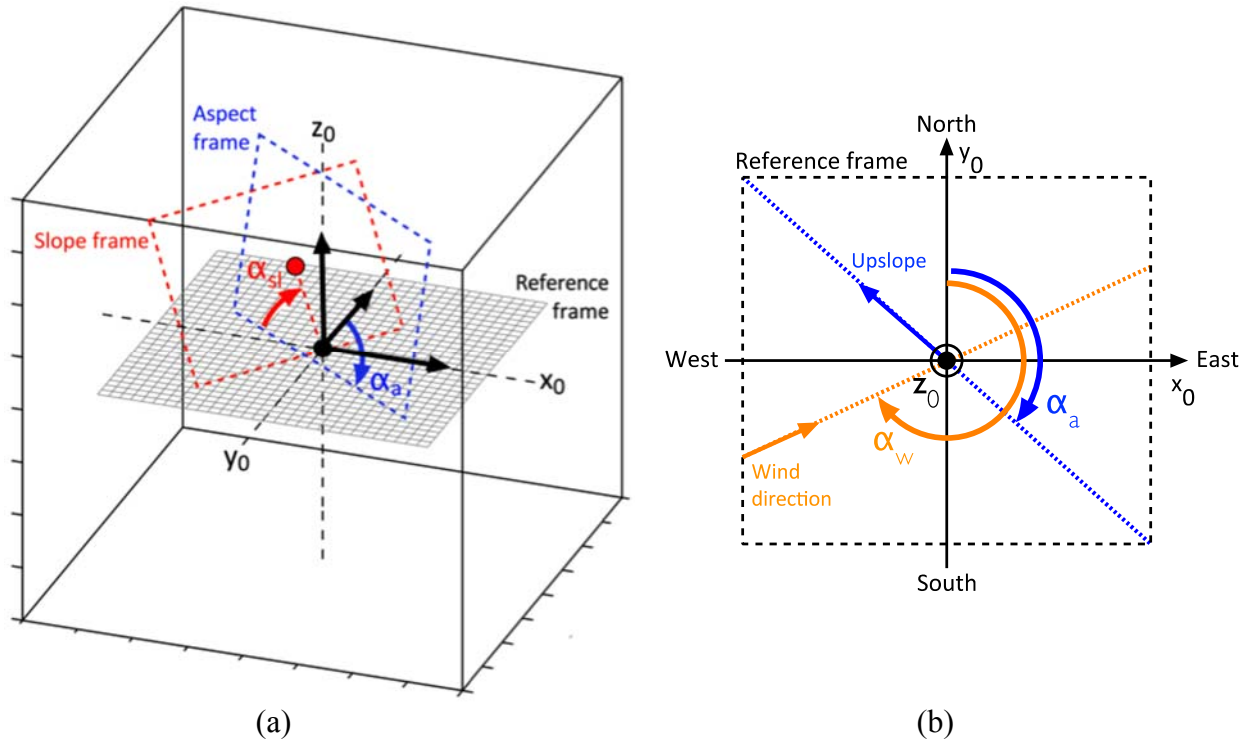


Figure 1. Formalism to include terrain topography in FIREFLY. (a) Three-dimensional reference frames, where (x_0, y_0) represents the horizontal plane, where α_a represents the aspect angle direction (i.e. the downslope direction) and where α_{sl} represents the terrain slope angle; the blue and red planes are perpendicular. (b) Two-dimensional horizontal plane (x_0, y_0) with the aspect angle direction α_a and the wind direction angle α_w .

The pair of local parameters (α_a, α_{sl}) can be retrieved from altimetry elevation data $h \equiv h(x_0, y_0)$ as follows (Vico and Porporato 2009, Emery *et al* 2013):

$$\begin{pmatrix} \sin \alpha_a \\ \cos \alpha_a \end{pmatrix} = -\frac{1}{\tan \alpha_{sl}} \begin{pmatrix} \partial h / \partial x \\ \partial h / \partial y \end{pmatrix}, \quad (1)$$

with:

$$\tan \alpha_{sl} = \sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}. \quad (2)$$

In FIREFLY, the terrain elevation h corresponds to input data that are interpolated at each grid point of the computational domain. The calculation of its gradient is performed through a classical centered finite difference scheme.

2.1.2. Extension of the Rothermel's model to wind and slope effects

The original Rothermel's ROS model (noted Γ_{1D}) that was developed for one-dimensional upslope and/or upwind fire propagation is of the following form:

$$\Gamma_{1D} = \Gamma_0 (1 + \Phi_w^* + \Phi_{sl}^*), \quad (3)$$

where Γ_0 [$m s^{-1}$] is the no-wind no-slope ROS, and where the 1-D coefficients $\Phi_w^* \equiv \Phi_w^*(\mathbf{u}_w)$, $\Phi_{sl}^* \equiv \Phi_{sl}^*(\alpha_{sl})$ represent the wind and slope correction coefficients to account for the additional effects of

wind and slope on the propagating speed of the head fire. Note that the wind correction coefficient depends on the surface wind velocity vector noted \mathbf{u}_w [m s^{-1}], which can be provided for instance by meteorological forecasts.

Following choices made by Lautenberger (2013), a modification of the wind and slope contributions to the Rothermel-based ROS is introduced to account for wildfire spread in other directions than the uphill/upwind direction. The 2-D wind and slope correction coefficients in the modified Rothermel's formulation are noted Φ_w and Φ_{sl} , respectively. The modification of the slope correction coefficient Φ_{sl} relies on the following assumptions: (1) when the fire propagates in the upslope direction (i.e. $\alpha_{fr} = \alpha_a + 180^\circ$), the slope contribution to the ROS is maximum (i.e. $\Phi_{sl} = \Phi_{sl}^*$); (2) if the fire propagation occurs in the normal direction to upslope or downslope (i.e. $\alpha_{fr} = \alpha_a \pm 90^\circ$), the slope does not contribute to the fire front propagation; and (3) if the fire propagates in the downslope direction such that $\alpha_{fr} \in [\alpha_a - 90^\circ, \alpha_a + 90^\circ]$, the ROS is forced to Γ_0 , meaning that the fire cannot propagate at a lower ROS than the no-slope no-wind Γ_0 . The modification of the wind correction coefficient Φ_w relies on similar arguments. Thus, the Rothermel's formulation of the wind- and slope-aided ROS (noted Γ_{3D}) becomes:

$$\Gamma_{3D} = \Gamma_0 \max(1, 1 + \Phi_w + \Phi_{sl}), \quad (4)$$

with the following expressions for $\alpha_{fr} \notin [\alpha_a - 90^\circ, \alpha_a + 90^\circ]$:

$$\Phi_{sl} = \cos(\alpha_{fr} - (\alpha_a + \pi)) \Phi_{sl}^*, \quad (5)$$

$$\Phi_w = \cos(\alpha_{fr} - (\alpha_w + \pi)) \Phi_w^*. \quad (6)$$

Γ_{3D} corresponds here to the evaluation of the ROS in the slope frame (see Figure 1a); this value is projected onto the reference frame (x_0, y_0, z_0) to obtain the ROS Γ_{2D} required by the FIREFLY level-set solver to propagate the fire front on the horizontal plane (x_0, y_0) along its normal direction \mathbf{n}_{fr} . Based on geometrical considerations, the projected ROS Γ_{2D} reads:

$$\Gamma_{2D} = \Gamma_{3D} \left(1 + \tan^2(\alpha_{sl}) \cos^2(\alpha_a - \alpha_{fr}) \right)^{\frac{1}{2}}. \quad (7)$$

The derivation of the formulation in Eq. (7) can be found in Emery *et al* (2013). Figure 2 shows the variations of the slope-aided ROS Γ_{2D} evaluated using Eq. (7), with respect to the slope angle α_{sl} (horizontal axis) for different aspect angles α_a (colorbar) varying between 0° (dark-blue-plain line) and 90° (brown-plain line). For $\alpha_a = 0^\circ$, the fire spreads in the upslope direction with a ROS reaching up to 1 m/s for a slope angle α_{sl} above 65° ; the effect of the slope is considerable since the ROS Γ_{2D} can be multiplied by a factor up to 25 compared to the no-slope no-wind ROS $\Gamma_0 = 0.048$ m/s. On the contrary, for $\alpha_a = 90^\circ$, the fire propagates in the transverse direction to the slope, implying that the slope does not modify the ROS and $\Gamma_{2D} = \Gamma_0 = 0.048$ m/s.

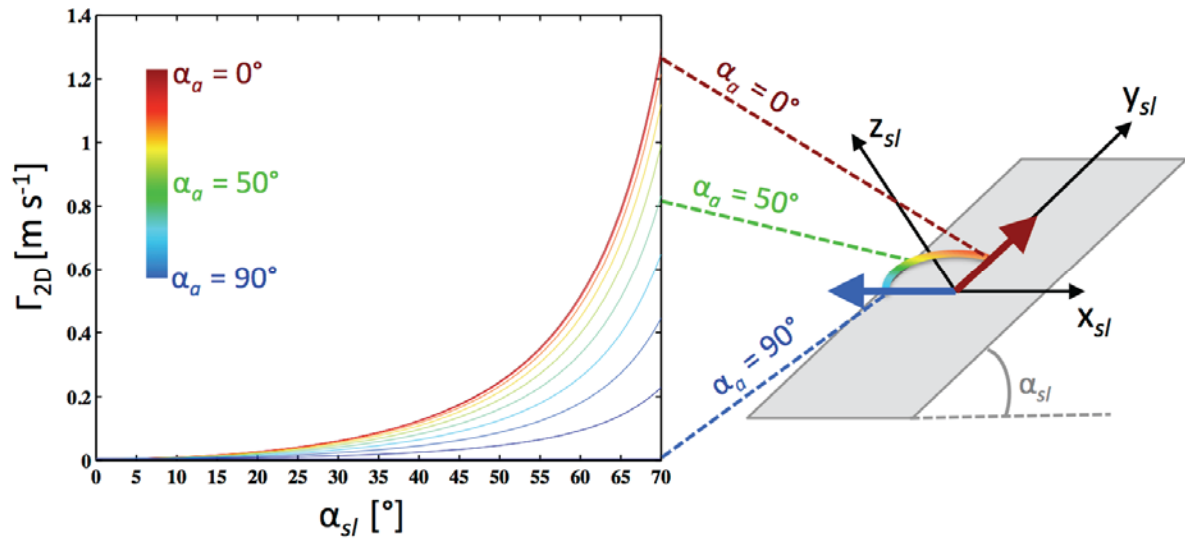


Figure 2. Slope-aided ROS Γ_{2D} with respect to the terrain slope angle α_{sl} (horizontal axis) for different values of the terrain aspect angle α_a (represented by the colorbar) for a no-wind plane configuration and $\Gamma_0 = 0.048$ m/s.

2.1.3. Level-set solver

FIREFLY tracks the time-evolving fire front location on the horizontal plane (x_0, y_0) using a level-set-based solver, see Figure 3. A progress variable $c \equiv c(x, y, t)$ is introduced as a flame marker: $c = 0$ in unburnt vegetation, $c = 1$ in burnt vegetation, and the contour line $c_{fr} = 0.5$ is identified as the flame front.

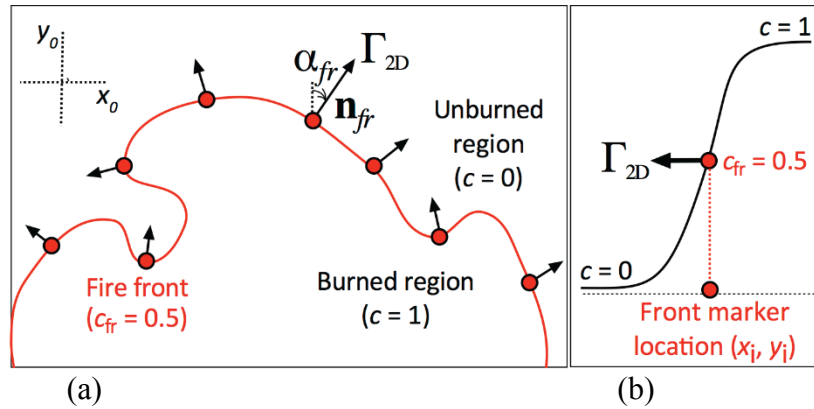


Figure 3. Schematic of FIREFLY. (a) Propagation of the contour line $c_{fr} = 0.5$ onto the horizontal plane (x_0, y_0) , Γ_{2D} measures the local ROS of the fire along the normal direction to the front \mathbf{n}_{fr} (defined by the direction angle of fire propagation α_{fr}). (b) Profile of the spatial variations of the progress variable c across the fire front, (x_i, y_i) representing the location of the i th fire front marker.

The progress variable field c is calculated as a solution of the propagation equation:

$$\frac{\partial c}{\partial t} = \Gamma_{2D} |\nabla c|, \quad (8)$$

where $\Gamma_{2D} \equiv \Gamma_{2D}(\mathbf{u}_w, \alpha_{sl}, M_v, \Sigma_v, \beta_v, \delta_v, \dots)$ calculated by Eq. (7) is defined along the local normal direction to the fire front $\mathbf{n}_{fr} = -\nabla c / |\nabla c|$ (see Figure 3) and depends on the local environmental conditions such as the biomass fuel moisture content M_v [-], the biomass fuel particle surface-to-

volume ratio Σ_v [m^{-1}], the biomass fuel packing ratio β_v [-] and the biomass fuel layer thickness δ_v [m]. Equation (8) is solved by using a second-order Runge-Kutta scheme for time-integration and a second-order total variation diminishing scheme with a Superbee slope limiter for spatial discretization, following choices made by Rehm and McDermott (2009). Convergence of the numerical method was demonstrated in Rochoux (2014).

2.1.4. Reconstruction of the simulated fire front

The instantaneous location of the fire front at time t is extracted by geolocating the contour line $c_{fr} = 0.5$ and satisfies: $c(x_i, y_i, t) = c_{fr}$ for every marker i of the front, with $1 \leq i \leq N_{fr}$ (N_{fr} being the total number of simulated markers along the fireline). Two steps are required for this reconstruction of the simulated fire front. First, the algorithm extracts the contour line $c_{fr} = 0.5$ from the two-dimensional progress variable c with respect to the computational grid resolution on the horizontal plane (x_0, y_0) in FIREFLY. Second, this algorithm discretizes the contour line $c_{fr} = 0.5$ with a fixed number (N_{fr}) of equally-spaced markers. Thus, the outputs of FIREFLY are the two-dimensional coordinates of the N_{fr} front markers. Further technical details on the contour line algorithm are provided in Rochoux (2014).

2.1.5. Examples of fire propagation over terrain topography

Two synthetic test cases are presented to illustrate the effects of non-uniform terrain topography on the fire front location as simulated by FIREFLY. This terrain topography corresponds to first, an inclined plane and second, a terrain elevation $h(x_0, y_0)$ typical of mountainous regions.

Inclined plane. The fire spread is simulated for a uniform biomass fuel characterized by $\delta_v = 1$ m, $M_v = 15$ % and $\Sigma_v = 11500$ m^{-1} . The terrain is a uniform inclined plane, tilted by $\alpha_{sl} = 15^\circ$ with respect to the horizontal plane, whose aspect angle is $\alpha_a = 225^\circ$ and whose dimensions are 600 m \times 600 m (with a mesh step size $\Delta x = \Delta y = 1$ m). The initial condition is described by a circular front centered at ($x_0 = 300$ m, $y_0 = 300$ m) and of radius $r_0 = 5$ m; there is no external flow ($u_w = 0$). FIREFLY is integrated during 1000 s (with a time step $\Delta t = 0.5$ s). Results presented in Figure 4 show that the slope induces a constant propagation in the upslope direction, while the spread at the back of the fire remains very limited. The effective simulated ROS of the head of the fire is equal to 0.264 m s^{-1} , which is consistent with the theoretical value 0.261 m s^{-1} provided by the 0-D Rothermel's model Γ_{1D} (see Eq. 3). The no-slope ROS Γ_0 is equal to 0.068 m s^{-1} ; the slope induces a propagation that is 4 times faster than in a no-slope configuration.

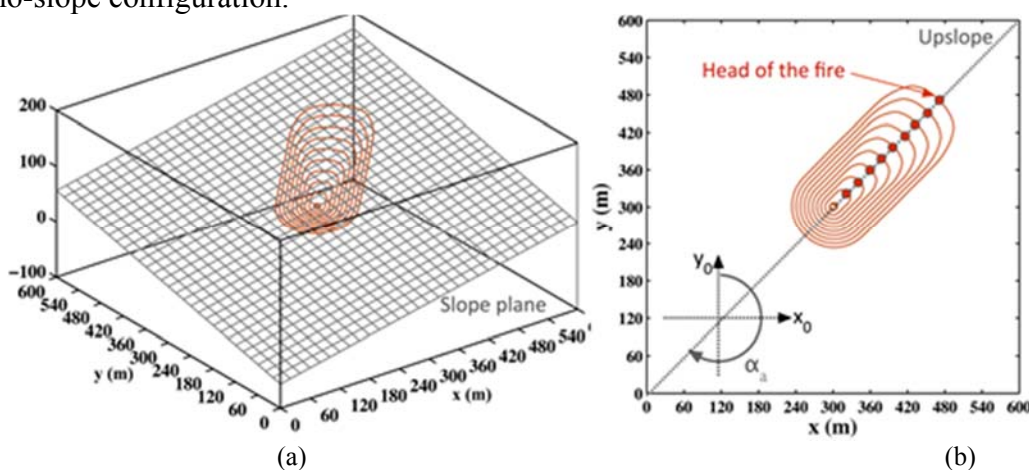


Figure 4. Time-evolving location of the simulated fire fronts at 100 s intervals on a slope plane with the aspect angle $\alpha_a = 225^\circ$ (a) Three-dimensional representation. (b) Projected representation onto the horizontal plane (x_0, y_0).

Mountainous region. A simulation of wildfire spread induced by complex terrain topography is performed for moderate wind conditions (0.75 m s^{-1} , 315°); the biomass fuel is uniformly-distributed

over the 200 m x 200 m field, with $\delta_v = 1$ m, $M_v = 20$ % and $\Sigma_v = 10000$ m⁻¹. The initial condition is described by a circular front centered at ($x_0 = 100$ m, $y_0 = 100$ m) and of radius $r_0 = 5$ m. FIREFLY is integrated during 1500 s. Figure 5 illustrates the growth of the burnt area over time (at 300 s time intervals). Even though the validation of FIREFLY for the treatment of the wind and slope effects remains to be performed against real-world wildfire events, this simulation is consistent with the main features of wildfire spread: the largest ROS is obtained in the upslope direction (the effect of terrain elevation on the fireline behavior is high compared to the effect of moderate wind, partly due to the high terrain elevation in one corner of the domain).

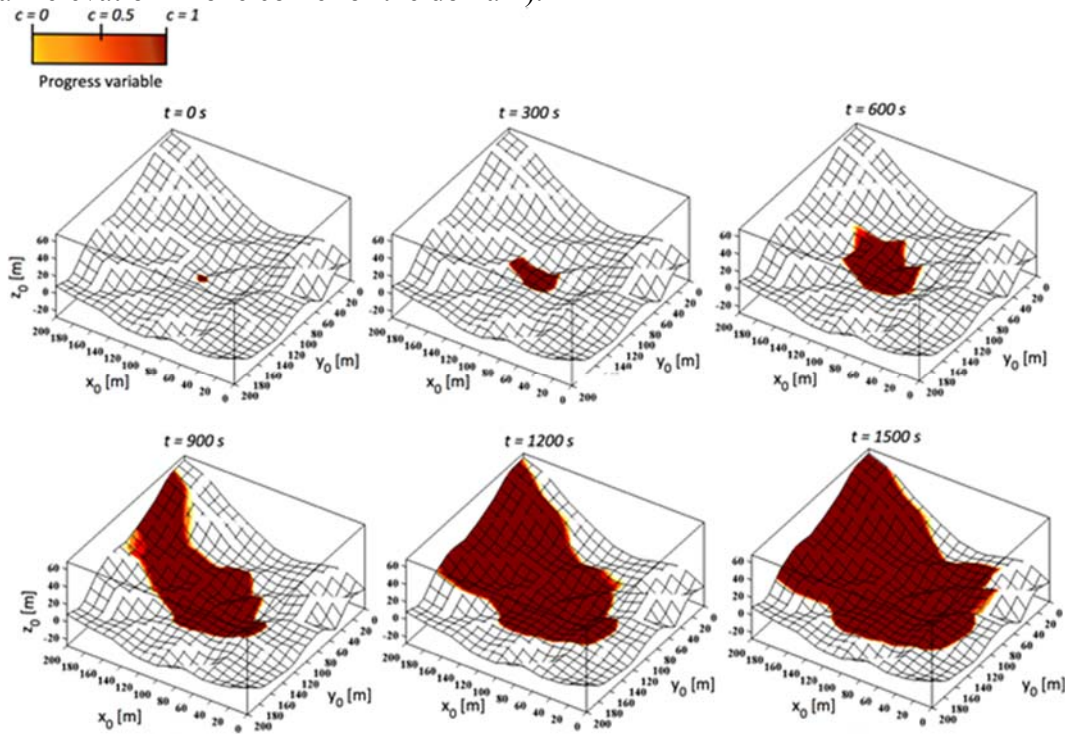


Figure 5. Growth of the burnt area from $t = 0$ s to $t = 1500$ s, simulated with FIREFLY over a complex terrain topography; moderate horizontal wind conditions (0.75 m s⁻¹, 315°);

2.2. Data assimilation algorithms: parameter estimation versus state estimation

2.2.1. Model counterparts of the observations

The correction provided by the DA algorithms relies on a comparison between the (FIREFLY-based) predicted and observed fire front locations at time t . The predicted fire front is a high-resolution discretized line described by a set of N_{fr} markers. Since observations of the fire front position are likely to be provided with a much coarser resolution and since they may cover only a fraction of the fire front perimeter, the observed fire front is a discretized line with a set of N_{fr}^o markers, with N_{fr}^o much lower than N_{fr} . In the following, we assume for simplicity that $N_{fr}^o = (N_{fr} / r)$ where r is an integer taking values much larger than 1. A selection operator is applied to the simulated fire front in order to pair a subset of N_{fr}^o markers along the simulated fire front with the N_{fr}^o markers along the observation fire front, associating each marker of the observation fire front with its closest neighbor along the simulated fire front.

2.2.2. Ensemble Kalman filter

For both parameter estimation (PE) and state estimation (SE) approaches, the EnKF algorithm is sequentially applied; each assimilation cycle consists of two successive steps (see Figure 6): (1) a prediction step (forecast) in which the system is evolved from time $(t-1)$ until time t (the next observation time) through an integration of the FIREFLY fire spread model for a large sample N_e of

control variables, each realization leading to a new prediction of the fire front location; (2) an update step (analysis) in which new observations are considered at the analysis time t and the ensemble of forecast realizations is modified consistently with the observations in order to reduce uncertainties in FIREFLY predictions. The new ensemble of N_e realizations is obtained by the application of the classical Kalman filter equation, assuming that the errors on the control variables and on the observations follow Gaussian probability density functions. One advantage of the EnKF algorithm is that the mapping between the control variables and the observation space is estimated stochastically. Thus, the EnKF algorithm is able to handle model non-linearities.

In the PE approach, the control vector includes the n input parameters of the Rothermel-based ROS model that are subject to uncertainties and to which FIREFLY is sensitive (Rochoux *et al.* 2014a). In the SE approach, the control vector is made of the coordinates of the N_{fr} front markers along the simulated fire front such that $n = 2N_{fr}$ (Rochoux *et al.* 2014b). The main difference between PE and SE approaches relies on the definition of the observation operator. In the SE approach, the observation operator is limited to the selection operator introduced in Section 2.2.1. Since in the PE approach the observation operator maps the environmental parameters onto the model counterparts of the observations, it also includes the integration of the FIREFLY fire spread model (this aspect emphasizes the need for a DA approach that accounts for model non-linearities).

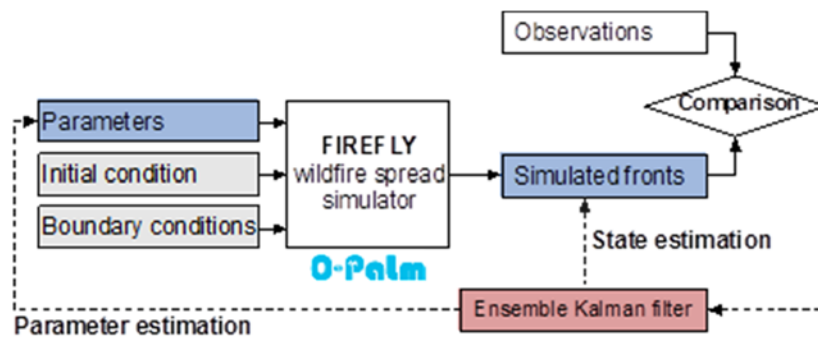


Figure 6. DA flowchart for PE and SE approaches over one assimilation cycle based on the OpenPALM dynamic coupler (www.cerfacs.fr/globc/PALM_WEB/); control variables are colored in blue.

2.2.3. Validation of the ensemble Kalman filter

We present results from a validation test in which observations are taken from an experimental database corresponding to a controlled grassland fire (personal communication and unpublished data). This validation test corresponds to a real reduced-scale (4 m x 4 m), flat, open-field grassland lot burning in which the ROS takes values on the order of 1 cm s^{-1} under moderate wind conditions. The properties of the grass are (approximately) known: $\delta_v = 0.08 \text{ m}$, $M_v = 22 \%$ and $\Sigma_v = 11500 \text{ m}^{-1}$. The mean wind conditions are also approximately known: $u_w = 1 \text{ m/s}$ and $\alpha_w = 307$ degrees. The fire spread is recorded during 350 s using a thermal-infrared camera; the fire front (represented by $N_{fr}^o = 40$ markers) is defined at the 600 K iso-temperature contour at 14 s intervals, from $t_0 = 50 \text{ s}$ to $t_4 = 106 \text{ s}$. In the PE approach, 4 parameters are used as control variables: the fuel moisture content and particle surface-to-volume ratio as well as the wind magnitude and direction angle. These parameters are varied around mean values and with prescribed uncertainties; the EnKF ensemble contains $N_e = 1000$ members. In the SE approach, the control variables are the spatial coordinates of the discretized fire front. An ensemble of $N_e = 100$ members is produced based on assumed uncertainties in the input parameters of the ROS model (i.e., the fuel depth, moisture content and particle surface-to-volume ratio as well as the wind magnitude and direction angle) as well as assumed uncertainties in the initial fire location at t_0 .

Figure 7a presents the averaged fire front position at time $t_4 = 106 \text{ s}$, using the PE and SE ensembles as predicted from a numerical integration of FIREFLY, and with an EnKF update performed at time $t_3 = 92 \text{ s}$. The numerical integration is performed N_e times for all members of the EnKF ensemble but

only the mean predictions are plotted in Figure 7a. It is shown that the mean forecast significantly underestimates the observed ROS. In contrast, Figure 7b shows that the mean analysis (i.e. a simulation with an EnKF update at current time t_4) successfully reduces the distance between model predictions and observations; this result is achieved by an adjustment of the ROS model parameters in the PE approach or directly of the fire front position in the SE approach.

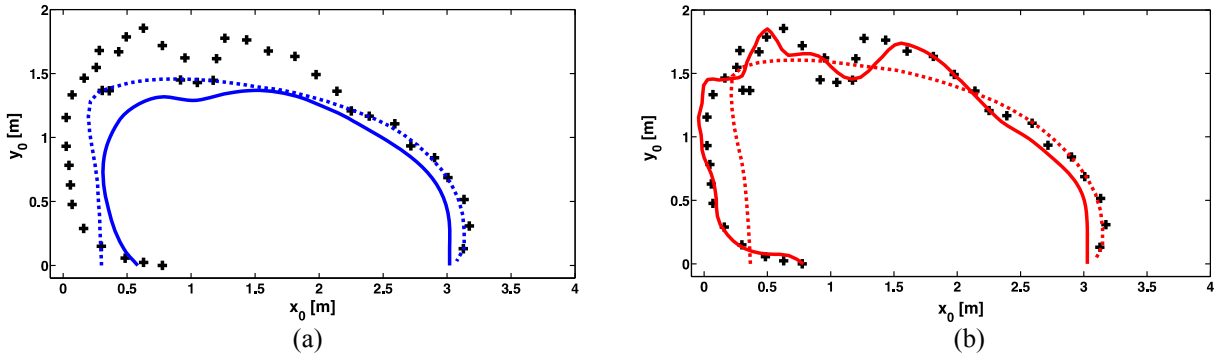


Figure 7. Comparison between simulated (lines) and observed (symbols) front positions at time $t_4 = 106$ s on the horizontal plane (x_0, y_0) . The simulated front position is the mean position calculated as the average of the EnKF ensemble; dashed lines (solid lines) correspond to the PE-based (SE-based) simulations. (a) Forecast (with an EnKF update at $t_3 = 92$ s). (b) Analysis (with an EnKF update at $t_4 = 106$ s).

In spite of the quality of the PE-based or SE-based correction, the mean forecast, while still significantly more accurate than the simulations without DA, remains limited. The mean predictions presented in Figure 7a have the benefits of an EnKF update performed at earlier time t_3 ; however, some of these benefits are lost at time t_4 due to the limited persistence of the initial condition for the SE-based forecast and due to the temporal variability of the errors in the environmental conditions for the PE-based forecast. We find that in this configuration, the PE algorithm provides better forecasts than the SE algorithm but this ranking is problem-dependent. While a DA approach provides excellent forecasting performance at short lead-times, the assimilation frequency (i.e., frequency at which the analysis is renewed by observations) needs to be adjusted to the persistence of the initial condition and/or to the temporal variability of the errors in the environmental conditions.

3. Evaluation of the performance of the data-driven simulator for complex terrain topography

An anisotropic case of wildfire spread subject to moderate wind conditions is considered for the complex terrain topography illustrated in Figure 5 for a deterministic run. In the present DA test, observations are synthetically generated using a similar deterministic run over the time window $[0; 750$ s], called *true solution* of the FIREFLY fire spread model, with specified values of the environmental conditions (referred to as the *true parameters* in Table 1); random errors characterized by a 1-m standard deviation (STD) are then added to the true solution to obtain the location of the observation markers ($N_{fr}^o = N_{fr} = 100$) at time 750 s (there is a single observation time over $[0; 750$ s]). Since this observation error STD is low relatively to the size of the fireline perimeter, the objective here is to evaluate the ability of the EnKF algorithm to retrieve the true location of the fire front.

An ensemble of $N_e = 320$ forecasts is produced over the time window $[0; 750$ s], based on assumed uncertainties in a subset of ROS model parameters, specifically in the fuel layer depth δ_v , the fuel moisture content M_v , the fuel particle surface-to-volume ratio Σ_v and the wind properties (u_w, α_w) . Thus, uncertainties are due to variations in 5 parameters, whose mean and STD are presented in Table 1. Note that the perturbed parameters remain spatially-uniform for each ensemble member. Note also that this large ensemble is necessary to accurately describe the spatial variability in the errors in the SE approach as well as to dissociate the effects of each control parameter on the fireline position in the PE approach. Figure 8a presents a comparison between the true and forecast fire front locations

at time $t = 750$ s (the forecast estimates between PE and SE approaches are equivalent). Due to uncertainties in the ROS model parameters and the presence of heterogeneous terrain topography, the shape of the simulated fire fronts significantly varies between the members (with a burnt area varying between 500 and 10000 m²).

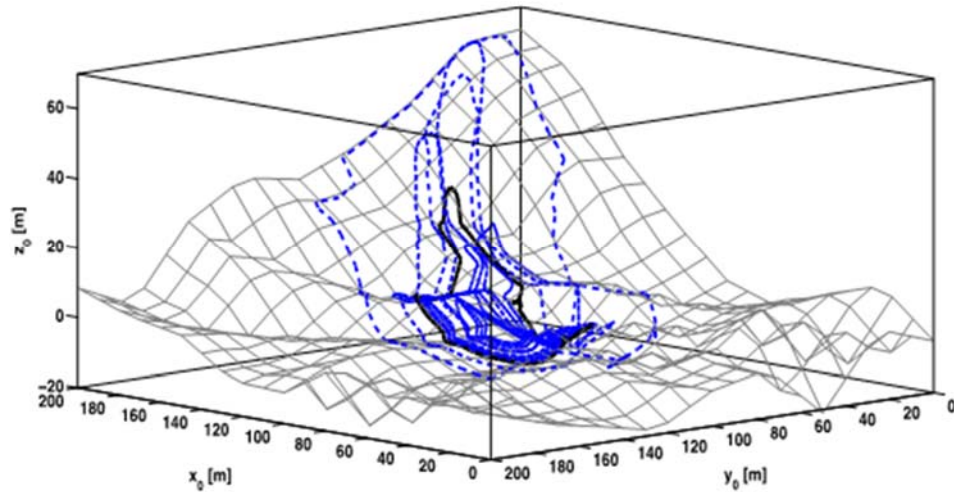
The EnKF update is performed at time 750 s, either for the PE approach or for the SE approach. While the SE approach leads to a direct adjustment of the fire front location, the PE approach works by an adjustment of the ROS model parameters; the 5 perturbed parameters (δ_v , M_v , Σ_v , u_w , α_w) are used as control variables. Figure 8b compares the mean analysis estimate obtained at the same time 750 s using a PE-based EnKF update (red dashed line) and a SE-based EnKF update (red solid line). These analysis estimates are compared to observations (black crosses) and to the mean forecast estimate (blue solid line). For both PE and SE approaches, the analysis estimates feature a much reduced scatter around the true location of the fire front than the ensemble of forecast estimates. At time 750 s, the mean distance between the (PE-/SE-based) mean ensemble estimate and the observations is reduced by a factor of at least 20. Thus, both PE and SE approaches are able to retrieve an accurate estimation of the fire front location, even though the prior information is subject to high levels of uncertainties and the terrain topography is complex. Note that as for the real case presented in Figure 7, the SE approach is more effective at retrieving the topology of the front at the analysis time than the PE approach; the latter provides valuable information on the input parameters to forecast the front at future lead-times (see Rochoux *et al.* 2014a, 2014b for further details).

Table 1. Properties of the ensemble forecast in the control space; anisotropic OSSE test

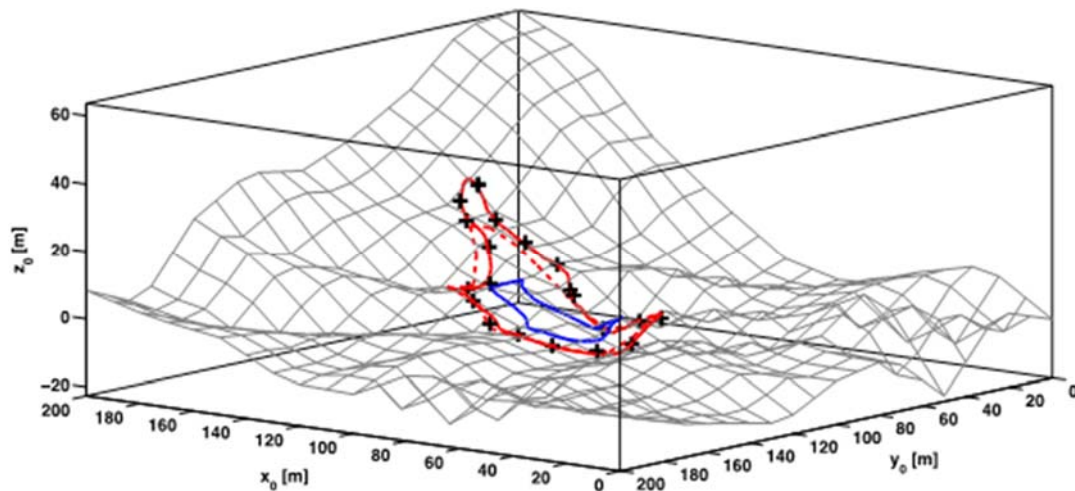
	True	Forecast (mean/STD)		PE-based analysis (mean/STD)	
δ_v	0.30 m	0.20 m	0.10 m	0.26 m	0.06 m
M_v	10 %	14.0 %	4.0 %	11.7 %	3.6 %
Σ_v	11500 m ⁻¹	10000 m ⁻¹	4000 m ⁻¹	13078 m ⁻¹	2436 m ⁻¹
u_w	0.75 m s ⁻¹	0.65 m s ⁻¹	0.20 m s ⁻¹	0.83 m s ⁻¹	0.15 m s ⁻¹
α_w	315°	315°	45°	288°	28°

4. Conclusions

This study presents a prototype data-driven wildfire simulator capable of correcting inaccurate predictions of the fire front position and of subsequently providing an optimized forecast of the wildfire behavior over complex terrain topography. The simulator features a regional-scale wildfire spread model FIREFLY coupled with an EnKF-based data assimilation algorithm and a parameter estimation (PE) approach or a state estimation (SE) approach. The study assumes that observations of the fire front location are available at frequent times but possibly provide an inaccurate and incomplete description of the fire front. The results obtained for a controlled grassland fire indicate that the forecasting performance of a PE approach or a SE approach may be limited to short lead-times. Future plans include the development of a dual SE/PE approach that could partly overcome these limitations. The PE approach could be extended to the case of weak spatial variations of the spread-rate model parameters. Assuming that the errors on the parameters vary slowly in time, the PE-based correction can reasonably be used for forecast, thus allowing for mid- to long-term forecast. In addition, the SE correction could be used for short-term forecast in order to locally correct the shape of the fire front. Future plans also include the evaluation of the PE/SE-based estimation strategy for real-world regional-scale wildfire hazards, in order to design the observation requirements (in terms of spatial and temporal resolutions) to obtain reliable forecasts of wildfire spread.



(a) Forecast step at time 750 s



(b) Analysis step at time 750 s.

Figure 8. Comparison between simulated (lines) and observed (symbols) front positions at time 750 s; single assimilation cycle. (a) Forecast step: comparison between the true front location (black solid line) and the ensemble of forecast estimates (blue dashed lines). Only a subset of the ensemble is shown here for clarity purposes (b). Analysis step: comparison between observations (black crosses) and the mean analysis estimate: the dashed line (the solid line) corresponds to the PE-based (SE-based) analysis (with an EnKF update at 750 s and starting from the same initial condition as the forecast at time 0). Only a subset of observation front markers is shown here for clarity purposes.

5. Acknowledgments

The financial support provided by the Agence Nationale de la Recherche under the IDEA project grant ANR-09-COSI-006 (2010-2013) and by the LEFE-MANU grant (INSU-CNRS program, 2011-2013) was greatly appreciated. The authors also gratefully acknowledge Ronan Paugam and Martin Wooster (King's College London) for sharing the experimental data, Florent Duchaine and Thierry Morel (CERFACS) for support on Open-PALM.

6. References

- Emery C, Rochoux MC, Ricci S, Trouvé A (2013) State estimation using data assimilation for simulation of regional-scale wildfire spread with complex topography. Technical Report TR-CMGC-13-63, CERFACS (France) - University of Maryland (USA).
- Lautenberger C (2013) Wildland fire modeling with an Eulerian level-set method and automated calibration, *Fire Safety Journal* 62, 289–298.
- Mandel J, Beezley JD (2007), Morphing Ensemble Kalman filter, *Tellus A* 60.
- Paugam R, Wooster M, Roberts G (2013), Use of handheld thermal imager data for airborne mapping of Fire Radiative Power and Energy and flame front rate of spread, *IEEE Transactions on Geoscience and Remote Sensing* 51, 3385-3399.
- Rehm RG, McDermott RJ (2009) Fire front propagation using the level-set method. NIST, Technical Report 1611.
- Rochoux MC, Delmotte B, Cuenot B, Ricci S, Trouvé A (2013) Regional-scale simulations of wildland fire spread informed by real-time flame front observations, *Proceedings of the Combustion Institute* 34, 2641–2647.
- Rochoux MC (2014) Vers une meilleure prévision de la propagation d'incendies de forêt: évaluation de modèles et assimilation de données (PhD Thesis - in English, Ecole Centrale Paris: France)
- Rochoux MC, Ricci S, Lucor D, Cuenot B, Trouvé A (2014a) Towards predictive data-driven simulations of wildfire spread – Part I: Reduced-cost ensemble Kalman filter based on polynomial chaos surrogate model for parameter estimation, *Natural Hazards and Earth System Sciences Discussion* 2, 3289–3349.
- Rochoux MC, Emery C, Ricci S, Cuenot B, Trouvé A (2014b) Towards predictive data-driven simulations of wildfire spread – Part II: Ensemble Kalman filter for the state estimation of a front-tracking simulator of wildfire spread, *Natural Hazards and Earth System Sciences Discussion* 2, 3769-3820.
- Rothermel RC (1972) A mathematical model for predicting fire spread in wildland fuels. USDA Forest Service, Research Paper INT-115.
- Sharples J (2008) Review of formal methodologies for wind-slope correction of wildfire rate of spread. *International Journal of Wildland Fire* 17, 179–193.
- Sullivan AL (2009a) Wildland surface fire spread modeling, 1990-2007. 1: Physical and quasi-physical models, *International Journal of Wildland Fire* 18, 349-368.
- Sullivan AL (2009b) Wildland surface fire spread modeling, 1990-2007. 2: Empirical and quasi-empirical models, *International Journal of Wildland Fire* 18, 369-386.
- Vico G, Porporato A (2009) Probabilistic description of topographic slope and aspect. *Journal of Geophysical Research* 114, 1–13.
- Viegas DX (2011) Overview of forest fire propagation research, in *Fire Safety Science - Proceedings of the Tenth International Symposium*, International Association for Fire Safety Science 10, 95-108.