Front shape similarity measure for data-driven simulations of wildland fire spread based on state estimation: Application to the RxCADRE field-scale experiment

C. Zhang [a], A. Collin [b], P. Moireau [c], A. Trouvé [a], M.C. Rochoux [d,*]

[a] *Department of Fire Protection Engineering, University of Maryland, College Park MD, 20742, USA*
[b] *Inria, Université de Bordeaux, Bordeaux INP, 351 cours de la Libération, Talence cedex 33405, France*
[c] *Inria – LMS, Ecole Polytechnique, CNRS – Université Paris-Saclay, 1 rue Honoré d'Estienne d'Orves, Campus de l'Ecole Polytechnique, Palaiseau 91120, France*
[d] *CECI, Université de Toulouse, CNRS, CERFACS, 42 Avenue Gaspard Coriolis, Toulouse cedex 01 31057, France*

Abstract

Data-driven wildfire spread modeling is emerging as a cornerstone for forecasting real-time fire behavior using thermal-infrared imaging data. One key challenge in data assimilation lies in the design of an adequate measure to represent the discrepancies between observed and simulated firelines (or "fronts"). A first approach consists in adopting a Lagrangian description of the flame front and in computing a Euclidean distance between simulated and observed fronts by pairing each observed marker with its closest neighbor along the simulated front. However, this front marker registration approach is difficult to generalize to complex front topology that can occur when fire propagation conditions are highly heterogeneous due to topography, biomass fuel and micrometeorology. To overcome this issue, we investigate in this paper an object-oriented approach derived from the Chan–Vese contour fitting functional used in image processing. The burning area is treated as a moving object that can undergo shape deformations and topological changes. We combine this non-Euclidean measure with a state estimation approach (a Luenberger observer) to perform simulations of the time-evolving fire front location driven by discrete observations of the fireline. We apply this object-oriented data assimilation method to the three-hectare RxCADRE S5 field-scale experiment. We demonstrate that this method provides more accurate forecast of fireline propagation than if either the fire spread model or the observations were taken separately. Results show that when the observation frequency becomes lower than $1/60 \text{ s}^{-1}$, the forecast performance of data assimilation is improved compared to simply extrapolating observation data. This highlights the need of a physics-based forward model to forecast flame front propagation. We also demonstrate that the front shape similarity measure is suitable for both Eulerian

---

* Corresponding author.
  E-mail address: melanie.rochoux@cerfacs.fr (M.C. Rochoux).

and Lagrangian-type front-tracking solvers and thereby can provide a unified framework to track moving structures such as flame front position and topology in combustion problems.

1. Introduction

At regional scales (i.e., at scales ranging from tens of meters up to several kilometers), wildland fires are usually described as fronts that self-propagate normal to themselves into unburnt vegetation; the local speed of the propagating front is referred to as the rate of spread (ROS). Current operational fire spread simulators adopt this regional-scale perspective using Eulerian [1,2] or Lagrangian [3,4] front-tracking solvers. The accuracy of the model predictions is limited due to both errors made in the ROS model and uncertainties in the input parameters of the ROS model representing heterogenous environmental and meteorological conditions.

A possible approach to overcome these limitations is data assimilation (DA) [1,5–10]. DA offers a valuable framework to integrate fire sensor observations into a computer model, with the goal to find optimal estimates of the targets or "control variables" (e.g., physical parameters, external forcing, initial conditions) and thereby to improve wildfire spread forecast. The control variables can be of different nature: in a parameter estimation approach, they are the parameters of the ROS model, which allows a correction of the fire spread model [8,10]; in a state estimation approach, they are the coordinates of the fireline, which allows a correction of the initial conditions of the fire spread model [9]. DA fits into the wider category of data-driven systems, where sensor is used to inform a model of the system, leading to the reduction of uncertainties on the model and its outputs. In wildland fire applications, the main challenge is to design a DA framework that takes advantage of remote sensing technologies [11,12] and assimilate observations of active burning areas [13].

The goal of DA is to infer control variables that minimize the discrepancies between observations and simulations by a certain metric or "measure". The choice of this measure strongly depends on the nature of the observations and influences the resulting optimal values for the control variables. In wildland fire applications, the observations correspond to time-evolving firelines. It is thus of prime importance to have an adequate measure to represent differences in front position. The fire front has a coherent structure: it is easy to identify such a structure by eye but a complete mathematical description remains a challenge [14,15]. Standard point-wise local metrics suffer for instance from the double penalty effect, i.e., a misplaced structure is predicted where it should not be and is not predicted where it should be [16]. We thus need to design a measure that characterizes position shift, shape deformation and topological change of the fire front.

In previous work, a data-driven wildfire modeling system was developed and successfully evaluated against reduced-scale and field-scale (FIREFLUX I) experiments [8–10]. In Refs. [8–10], the observed fire contours were treated as discretized contours with a finite set of markers. The distance between the simulated and observed fronts was then computed by pairing each observed marker with its closest neighbor along the simulated front. Such marker registration was considered as a temporary solution to the problem of representing position errors because it relied on some arbitrary choices made in the pairing process, required user-input and was therefore neither robust nor automatic. This solution is no longer adequate when addressing regional-scale wildland fires in highly heterogeneous conditions that feature complex front shapes and potentially multiple fire fronts due to spotting. To overcome this issue, we propose in the following a front shape similarity measure derived from image segmentation theory [17] and previously applied in cardiac electrophysiology [18]. This measure considers the burning area as a moving object that can deform under heterogeneous conditions and quantifies the match between observed and simulated burning areas.

The present work aims at demonstrating the relevance and performance of the new front shape similarity measure for wildland fire problems through an application to the 2012 RxCADRE S5 [19] field-scale experiment. The study considers a state estimation approach and both Eulerian and Lagrangian formulations of the fire spread model. The study is a continuation of preliminary work [20]: the focus of Ref. [20] was on mathematical considerations and verification tests against synthetic observations; the present focus is on a validation test against RxCADRE dataset with the objective of showing the merits of DA over standalone model predictions and observation extrapolation. The fire spread models and the RxCADRE dataset are presented in Section 2. The new front shape similarity measure and state

estimation approach are presented in Section 3. The application to RxCADRE is presented in Section 4.

2. Models and data

2.1. Fire spread models

Regional-scale fire spread models feature two main components: (1) a rate-of-spread model that gives values of ROS as a function of local topography, biomass and meteorological conditions (we use the Rothermel model [21]); and (2) a front-tracking algorithm that provides a solution of the front propagation equation on the two-dimensional horizontal plane (x, y).

The fire spread model used in our previous work [8–10] is an Eulerian front-tracking simulator. A progress variable $c \equiv c(x, y, t)$ is first introduced as a marker of the fireline; the fireline is then defined as the isocontour $c_{\text{fr}} = 0.5$; $c < c_{\text{fr}}$ ($c > c_{\text{fr}}$) represents unburnt (burnt) vegetation. In this framework, c is calculated as a solution of the propagation equation for $(x, y) \in \Omega$ and $t \geq t_0$

$$\frac{\partial c}{\partial t} = \text{ROS} \, |\nabla c|, \; c(x, y, t_0) = c_0(x, y), \quad (1)$$

where $c_0(x, y)$ is the initial condition at time t_0 over the computational domain Ω. The level-set function $\phi_c(x, y, t) = (c(x, y, t) - c_{\text{fr}})$ also satisfies Eq. (1); the fire front is represented by $\phi_c(x, y, t) = 0$. To solve Eq. (1), we follow the choices made in Ref. [22] and use a second-order Runge–Kutta scheme for time-integration and a second-order total variation diminishing (TVD) scheme with a Superbee slope limiter for spatial discretization.

In the present work, we also consider a Lagrangian front-tracking solver similar to the model adopted in FARSITE [3]. FARSITE is based on Huygens' principle and assumes that the fire front features a local elliptical shape. The front is parameterized by the closed curve $(x(s, t), y(s, t))$ with $0 < s < 2\pi$. The Lagrangian equations are for $(x, y) \in \Omega$ and $t \geq t_0$

$$\begin{cases} \partial_t x = F(s, t, \partial_s x, \partial_s y) \\ \partial_t y = G(s, t, \partial_s x, \partial_s y) \end{cases}, \quad \begin{cases} x(s, t_0) = x_0(s) \\ y(s, t_0) = y_0(s) \end{cases}, \quad (2)$$

where $(x_0(s), y_0(s))$ is the initial condition at time t_0, and where F and G are functions of the local topography, biomass and meteorological conditions (the shape and orientation of the local ellipse are determined by the wind and slope conditions, while its size is determined by the ROS). In this framework, the fire front is represented by the envelope of growing individual ellipses. To solve Eq. (2), we follow the choices made in Ref. [23].

Note that in the Lagrangian model, the Rothermel-based ROS is used to describe fire spread in the wind and/or upslope direction, i.e., in the head fire region, while spread in other directions is described by calibrated properties of the assumed local elliptical shape. In contrast, in the Eulerian model, the Rothermel-based ROS is used to describe fire spread along the entire fire perimeter, not only at the head but also on the flanks and in the rear (this modeling choice is inconsistent with the original intent of Rothermel's model to focus on the fire head [21] but this is a usual assumption).

2.2. RxCADRE S5 experiment

The Prescribed Fire Combustion and Atmospheric Dynamics Research Experiment (RxCADRE) corresponds to a series of field experiments aimed at model development and validation. In the present work, we consider the 2012 S5 experiment [19]. The S5 experiment was a 15-min-long prescribed fire conducted on a flat terrain characterized by a surface area of approximately 180m × 180m and a mixed grass and shrub vegetation. During the fire, the surface wind blew mainly from the North direction. In the simulations, we assume uniformly-distributed vegetation fuel as well as uniform and constant wind. The input parameters to the Rothermel model are based on experimental measurements: the fuel depth is 0.2 m; the fuel surface loading is 0.28 kg m^{-2}; the fuel particle surface-to-volume ratio is 9,000 m^{-1}; the fuel moisture content is 10 %; the wind velocity at mid-flameheight is 2 m s^{-1} and the wind direction is 345° (corresponding to a northwest wind).

The fire was ignited on the North side of the lot and propagated into the southern direction. Fire propagation was recorded through a series of temperature maps starting at time $t = 34$ s after ignition and recorded at 1-Hz frequency using a long-wave thermal infrared imaging system [19]. Since the initial fire only covers a very small area, we use the observed fire at time $t = 60$ s as initial conditions for spread simulations. So in the following, time $t = 0$ s corresponds to time $t = 60$ s in the RxCADRE dataset. Figure 1a presents the map of flame arrival times, from $t = 0$ to 480 s (corresponding to the time interval [60 s; 540 s] in the RxCADRE dataset), showing the time at which the fireline arrives at a given pixel of the S5 burn lot. In the assimilation procedure, we assume that observations are available at 60-s time intervals. For example, Fig. 1b shows the observed unburnt/burnt binary field that is assimilated at time $t = 480$ s. This binary field is obtained using the corresponding map of flame arrival times and after filtering to remove small-scale holes and outliers.

3. Front shape similarity measure in data assimilation

We now present a front shape similarity measure derived from image segmentation theory to go beyond the limitations of Euclidean-type distance and reduce position errors through DA [18,20].
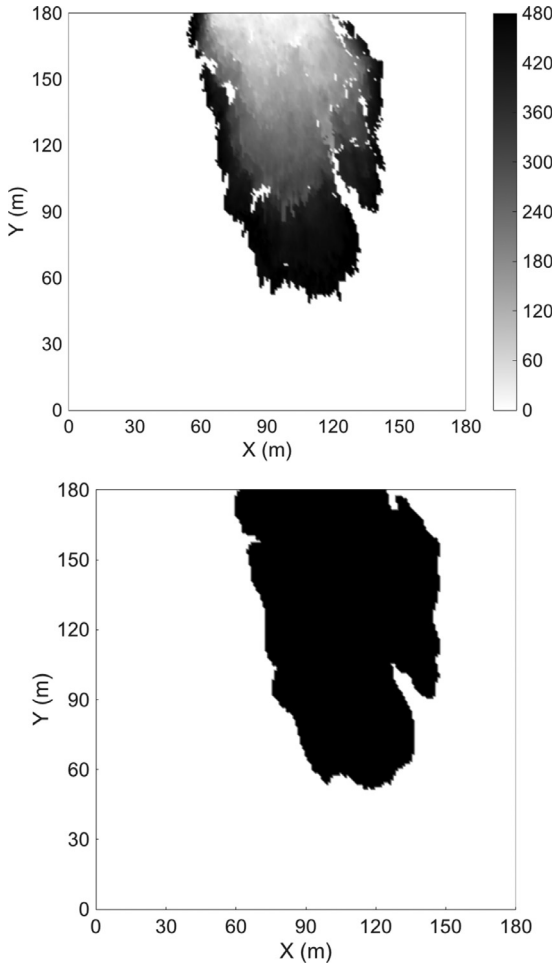
Fig. 1. Dataset of the RxCADRE S5 fire. (a) Map of flame arrival times (0-480 s). (b) Binary image showing unburnt (white) and burnt (black) vegetation at time $t = 480$ s.

3.1. Chan–Vese Contour Fitting Functional

We can define a similarity measure between a target front and a simulated front from the Chan–Vese contour fitting functional [17,24]. In a level-set formalism, this measure can be written as

$$
\begin{cases}
\mathcal{J}(y^\circ, \phi_c) = \int_\Omega H_v(\phi_c) [y^\circ - C_{\max}(y^\circ, \phi_c)]^2 \\
\quad + (1 - H_v(\phi_c))[y^\circ - C_{\min}(y^\circ, \phi_c)]^2 \, dx \, dy,
\end{cases}
$$
(3)

where y° is the observed binary field, ϕ_c the level-set function ($\phi_c = c - c_{fr}$), H_v the Heaviside function ($H_v(\phi_c) = 0$ if $\phi_c < 0$; $H_v(\phi_c) = 1$ if $\phi_c > 0$); and where $C_{\min}$ and $C_{\max}$ are scalar coefficients defined by $C_{\min} = \min(C_0, C_1)$ and $C_{\max} = \max(C_0, C_1)$

with

$$
\begin{cases}
C_0(y^\circ, \phi_c) = \dfrac{\int_\Omega (1 - H_v(\phi_c)) \, y^\circ \, dx \, dy}{\int_\Omega (1 - H_v(\phi_c)) \, dx \, dy}, & 0 \le C_0 \le 1, \\[2ex]
C_1(y^\circ, \phi_c) = \dfrac{\int_\Omega H_v(\phi_c) \, y^\circ \, dx \, dy}{\int_\Omega H_v(\phi_c) \, dx \, dy}, & 0 \le C_1 \le 1.
\end{cases}
$$
(4)

C_1 corresponds to the mean of y° across the simulated burnt region ($H_v = 1$) and measures the match between the observed and simulated burnt areas. C_0 corresponds to the mean of y° across the simulated unburnt region ($H_v = 0$) and measures the mismatch between the observed and simulated unburnt areas. If the observed and simulated fronts coincide, $C_1 = 1$ and $C_0 = 0$.

3.2. State estimation methods

We now present Eulerian- and Lagrangian-type state estimators, which are derived from the data fitting functional presented in Eq. (3) and introduce new relaxation terms in the propagation Eqs. (1) and (2). This type of state estimator is known as a Luenberger observer (LO). The gradient of the similarity measure in Eq. (3) reads

$$\nabla \mathcal{J} = \delta(\phi_c)\, \mathcal{D}(y^o, c), \tag{5}$$

where δ is the Dirac delta-function, and where $\mathcal{D}(y^o, c)$ is the discrepancy term defined as

$$\mathcal{D}(y^o, c) = [y^o - C_{max}(y^o, \phi_c)]^2 - [y^o - C_{min}(y^o, \phi_c)]^2. \tag{6}$$

The function δ plays the role of a localization operator (the correction associated with state estimators is only active along the simulated fire front) [18].

For the Eulerian fire spread model, Eq. (1) is modified as follows:

$$\begin{cases} \dfrac{\partial c}{\partial t} = \text{ROS}\, |\nabla c| - \lambda \delta(\phi_c)\{[y^o - C_{max}(y^o, \phi_c)]^2 \\ \quad - [y^o - C_{min}(y^o, \phi_c)]^2\}, \end{cases} \tag{7}$$

where the feedback term is expressed as $-\lambda \nabla \mathcal{J}$, and where $\lambda \equiv \lambda(x, y)$ is the gain that describes confidence in the observations (λ corresponds to the inverse of the observation error variance). Starting from given (possibly incorrect) initial conditions c_0 at time t_0, λ controls the rate at which the simulated progress variable c converges towards the observations y^o. Note that in the present study, the Dirac delta-function is numerically approximated as $\delta(\phi_c) = |\nabla c|$ since the initial fire front is thin. Equation (7) is referred to as the "Eulerian Luenberger observer" (LO–EUL).

The "Lagrangian Luenberger observer" (LO–LAG) can be derived by analogy to the LO–EUL. If $X(s, t) = (x(s, t), y(s, t))^T$ denotes the front parameterization and $V = (F, G)^T$ the associated velocity vector, then Eq. (2) is modified as follows:

$$\begin{cases} \partial_t x = V(s, t, X(s, t)) - \lambda \{[y^o - C_{max}(y^o, c)]^2 \\ \quad - [y^o - C_{min}(y^o, c)]^2\}\, n_{fr}. \end{cases} \tag{8}$$

where n_{fr} is the normal vector to the fire front and $\lambda \equiv \lambda(s)$. We use a ray casting algorithm [25] to construct the binary burnt/unburnt fields from the front marker positions given by Eq. (8); these fields are required in the evaluation of the discrepancy term $\mathcal{D}(y^o, c)$.

The feedback term features a similar formulation in the Lagrangian and Eulerian models. In principle, $\text{ROS} = (V \cdot n_{fr})$ provides the equivalence between the Eulerian and Lagrangian formulations of a front propagation problem. However, we adopt here the usual formulations of the wildland fire research field so that the way to handle fire propagation on the flanks and at the rear of the fire is different in the present Eulerian and Lagrangian models [3,22] (see discussion in Section 2.1).

3.3. Illustration of the effect of the Luenberger observer corrections

We illustrate the behavior of the front shape similarity measure by considering the Eulerian framework (LO–EUL) and a case for which the simulated fireline is enclosed by, and initially quite different from the observed fireline (see Fig. 2).

We consider a computational domain of $200\,\text{m} \times 200\,\text{m}$; the grid resolution is $\Delta x = \Delta y = 1\,\text{m}$ and the temporal resolution is $\Delta t = 0.1$ s. The gain is $\lambda = 4$. For illustration purposes, the propagation equation in Eq. (7) is solved without the ROS term, implying that the λ-parameter only features the speed of the correction process; a smaller λ-value would provide the same correction but at a slower rate. Coefficient C_1 corresponds to a perfect match for the simulated burnt area, $C_1 = 1$; the coefficient C_0 corresponds to some level of mismatch for the simulated unburnt area, $0 \leq C_0 \leq 1$; therefore, $C_{min} = C_0$, which gives $0 \leq C_{min} \leq 1$, and $C_{max} = C_1 = 1$. In the vicinity of the simulated fireline (at $c \approx c_{fr}$), the discrepancy term is negative, $\mathcal{D}(y^o, c_{fr}) = \{[y^o - C_{max}]^2 - [y^o - C_{min}]^2\} = -[1 - C_{min}]^2 < 0$, which corresponds to a positive term on the right-hand-side of Eq. (7) and as seen in Fig. 2a, to outward propagation of the simulated fireline. Figure 2a shows how the simulated fireline is progressively modified to match the observed front shape; Fig. 2b presents the discrepancy term $\mathcal{D}(y^o, c)$ at time $t = 5$ s; negative (positive) values of $\mathcal{D}(y^o, c)$ correspond to outward (inward) propagation; this term is multiplied by $\delta(\phi_c)$ so that the correction in Eq. (7) is only active at $c \approx c_{fr}$.

4. Data-driven simulation results

We now apply the LO–EUL and LO–LAG state estimators to data-driven simulations of the RxCADRE S5 experiment. Input parameters are selected using available experimental data (see Section 2). The computational domain is $180\,\text{m} \times 180\,\text{m}$; the total simulation time is 480 s with a fixed temporal resolution Δt. For LO–EUL, $\Delta x = \Delta y = 1$ m and $\Delta t = 0.05$ s. For LO–LAG, $\Delta t = 0.5$ s and the initial number of front markers is 40 (the number of front markers increases during simulation due to the increasing length of the fireline). The initial condition c_0 (indicating the location of the fireline at ignition time) is supposed to be unknown. The objective is to demonstrate that our state estimation algorithm is able to overcome an imperfect knowledge of the fire situation
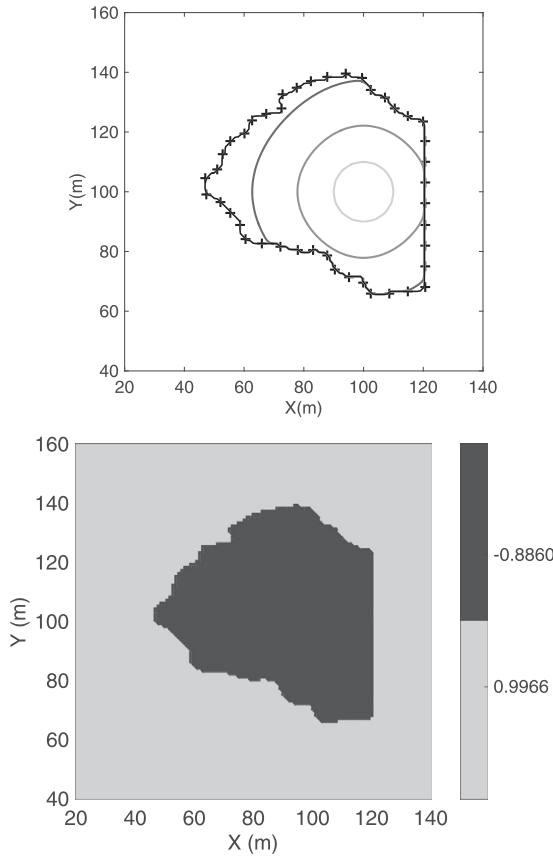
Fig. 2. Simulation of a representative test case. (a) Comparison of simulated (lines) and observed (symbols) firelines at $t = 0, 5, 10$ and 15 s. (b) Discrepancy term $\mathcal{D}(y^o, c)$ (Eq. (5)) at $t = 5$ s.

at initial time. To initialize the fire spread model, we thus consider a simple, approximate fireline location, i.e., a semi-circular front with a 15-m radius and a center located at $(x, y) = (90$ m$, 180$ m$)$. Observations are assimilated at 60-s time intervals. Between two observations made at time $t = t_n$ and $t = t_{n+1}$ (i.e., during an analysis), the model state is continuously steered towards the observations made at time t_{n+1}. The intensity of the steering process is controlled by the gain λ (the higher the value of λ, the higher the level of confidence in the observations and the lower the value of the observation error standard deviation). Beyond the last observation (i.e., during a forecast), the relaxation terms are de-activated and the model solves the original propagation Eqs. (1) and (2).

We consider simulations with a large prescribed value of λ for both LO–EUL and LO–LAG models, in order to push the data-driven strategy to its limits in a situation where observations are considered accurate, the standalone model prediction (or "free run") is far away from the observations and the correction is significant. In practice, the λ-value should be set according to the information

we have on the observations. Here $\lambda = 1$, which corresponds to a very large weight on the observation data. Figure 3 compares the observed and simulated firelines during the first 8-min of the S5 experiment. The simulated firelines include free runs (that are not informed by the observations) and data-driven simulations using LO–EUL in Fig. 3a and LO–LAG in Fig. 3b. It is seen that the free runs underestimate the ROS in both the headfire and flank regions indicating the presence of significant model errors for both Eulerian and Lagrangian models. In contrast, the data-driven simulations successfully reproduce the location and shape of the fire perimeters. This shows the capability of the data-driven model to accurately retrieve the shape of the observed firelines.

We also examine the forecast performance at 60-s lead time of the LO–EUL and LO–LAG data-driven runs in terms of distance error. The distance error between the predicted and observed fireline positions is evaluated using the Hausdorff distance [26] (in meter), which corresponds to the largest of all the distances from a point in one set to the closest point in the other set. Note that we choose
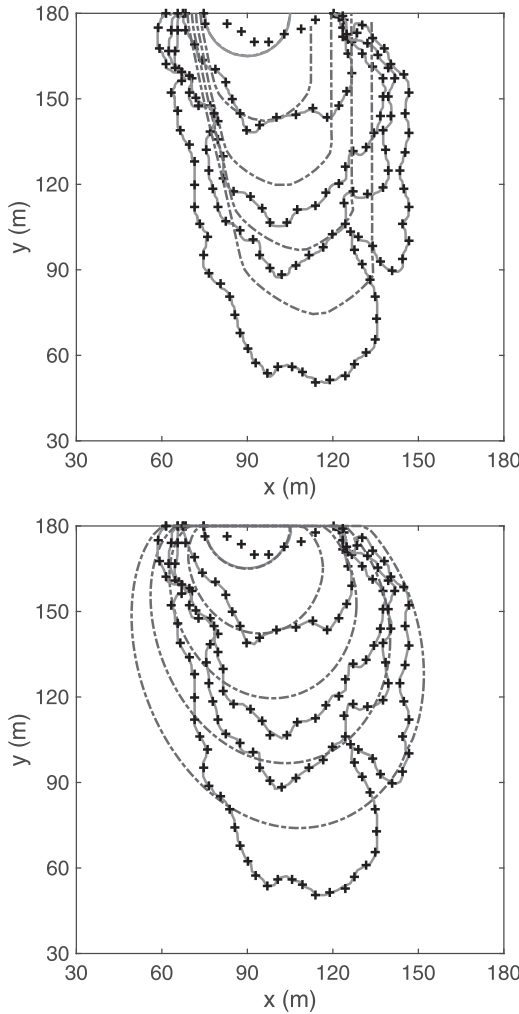
Fig. 3. Comparison of simulated (lines) and observed (symbols) firelines at 120-s time intervals during $0 \leq t \leq 480$ s; $\lambda = 1$. Dashed (solid) lines correspond to free runs (data-driven runs) using (a) the Eulerian model and (b) the Lagrangian model.
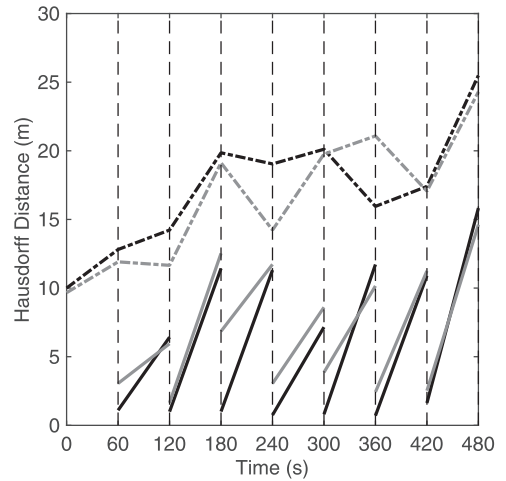

Fig. 4. Time variations of the Hausdorff distance between the simulated and observed firelines. Dashed (solid) lines correspond to free runs (data-driven runs) using LO–EUL (black) and LO–LAG (gray).

the Hausdorff distance as a diagnostic tool since it is easier to interpret to an end-user than the front shape similarity measure (the latter is still used in the DA algorithm to calculate the discrepancy term $\mathcal{D}(y^o, c)$) and since it is a more conservative estimate than the mean distance error (the Hausdorff distance has the meaning of maximum distance error between observation and simulation). Figure 4 compares the temporal evolution of the Hausdorff distance in the free runs (dashed lines) and data-driven runs (solid lines). The evolution of the error in free forecast mode is a continuously increasing function of time. This error takes moderate values (approximately 25 m at time $t = 480$ s) primarily due to the limited scale and duration of the S5 ex-

periment but the error is unbounded and keeps increasing in time. In contrast, the evolution of the error in data-driven mode is a discontinuous function: deviations of model predictions from observations are periodically reduced (to less than 5 m) during the analysis events (when integrating new observations). After each analysis event, the error increases but remains bounded and takes small-to-moderate values (on the order of 10 m). LO–EUL and LO–LAG models provide similar results. The rapid increase of the error seen in Fig. 4 after each analysis indicates that the ROS model has limited accuracy and that the benefits of assimilating new observations have limited persistence in time. Assimilation at $1/60$-s^{-1} frequency results in a distance error of 10 m; assimilation at lower (higher) frequencies would result in larger (lower) errors.

We finally evaluate the contribution of the ROS model in the quality of the short-term forecast by comparing, in forecast mode, the DA strategy with a simpler methodology that is exclusively based on extrapolation of observations. For this purpose, we test a simple forecast method (referred to as the "extrapolation method") based on extrapolation of fireline data at the two most recent observation times and the sensitivity of the method to the observation frequency (we use three different RxCADRE datasets with data available at 30-s, 60-s and 120-s time intervals). Consider that two observations are available at times t_1 and t_2; each observed fireline is represented using 100 front markers. The extrapolation method computes the trajectory of each marker between t_1 and t_2 and extrapolates its propagation until time t_3 assuming the same ROS and direction between t_2 and t_3 (in the current work, the forecast time window $[t_2; t_3]$ has
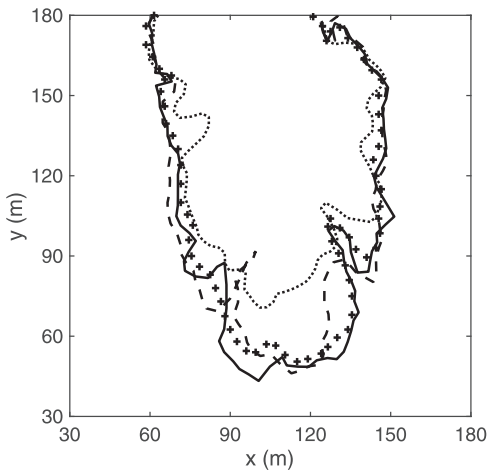
Fig. 5. Comparison of predicted (lines) and observed (symbols) firelines at time $t = 480$ s. Predictions are obtained with the extrapolation method for varying observation frequency: 1/30 s^{-1} (solid line); 1/60 s^{-1} (dashed line); 1/120 s^{-1} (dotted line).

the same length as the observation time window [t_1; t_2], which corresponds to the assimilation time period when using state estimation). Figure 5 presents the fireline position obtained at time 480 s using the extrapolation method for varying observation frequency. Results indicate that when the observation frequency is high (i.e., when the time between successive observations is 30 s or less), the extrapolation method is able to track the actual observation at time 480 s; the forecast performance of the extrapolation method is then comparable to that obtained with the LO-EUL and LO-LAG state estimators. However, the forecast performance of the extrapolation method degrades when the observation frequency is 1/60 s^{-1} and 1/120 s^{-1}. In the first case, there is a significant deviation in the position of the flanks of the fire. In the second case, the headfire propagates too slowly compared to the actual observation. Thus, when the observation frequency is low (i.e., when the time between successive observations is 60 s or more), the LO-EUL and LO-LAG state estimators yield a better forecast performance than the extrapolation method. The LO-EUL and LO-LAG state estimators reduce the Hausdorff distance error by a factor of approximately two compared to the extrapolation method. These results show the ability of DA to steer inaccurate fire spread models towards observed firelines and to provide an improved forecast of the fire behavior compared to standalone model prediction or observation extrapolation by combining observations with a physics-based forward model when observations become less frequent.

5. Conclusions

This study presents an application of a data-driven wildland fire spread simulator to the Rx-CADRE S5 field-scale controlled fire experiment. The simulator features a front-tracking solver based on a ROS formulation and a DA algorithm assimilating fireline perimeters and based on state estimation. State estimation allows for regular corrections of the initial conditions of the forward model at times of new observations. Compared to previous work based on Euclidean distance, this study presents an original measure derived from image segmentation theory to quantify discrepancies between simulated and observed firelines in the DA algorithm. The performance of the new front shape similarity measure combined with state estimation is evaluated in the three-hectare Rx-CADRE burn, in both analysis and forecast modes, and using an Eulerian or Lagrangian fire spread model. The results demonstrate the ability of the proposed data-driven models to reduce uncertainties and to provide an improved short-term forecast. In particular, the results show that combining observations with a ROS model in a DA framework provides a better forecast than simply extrapolating information from available observations.

Longer-term forecast performance requires a more accurate forward model; one solution is to correct the fire spread model through parameter estimation in addition to state estimation. Future work will therefore include the development of a joint state-parameter estimation approach and investigate the impact of the DA frequency on the forecast performance. It is important to evaluate the required assimilation frequency to obtain a good forecast and how this frequency upscales with the fire size. Future work will include validation tests against a series of past wildfires with spatial variability of biomass, wind and topography, in which the burnt areas may feature complex front topology. DA provides an efficient framework to account for observation uncertainties, which is expected to be a useful capability to monitor wildfire hazards where observation uncertainties are large. More generally, the proposed methodology is believed to be a powerful approach to track moving structures and patterns in combustion problems.

References

[1] J. Mandel, J.D. Beezley, A.K. Kochanski, V.Y. Kondratenko, M. Kim, *Proc. Comput. Sci.* 9 (Supplement C) (2012) 1100–1109, doi:10.1016/j.procs.2012.04.119.

[2] C. Lautenberger, *Fire Saf. J.* 62, Part C (2013) 289–298, doi:10.1016/j.firesaf.2013.08.014.

[3] M. Finney, FARSITE: Fire Area Simulator - Model Development and Evaluation, *Technical Report*, USDA Forest Service, Rocky Mountain Research Station, 1998.

[4] J.-B. Filippi, X. Pialat, C. Clements, *Proc. Combust. Inst.* 34 (2013) 2633–2640.

[5] M. Denham, K. Wendt, G. Bianchini, A. Cortés, T. Margalef, *J. Comput. Sci.* 3 (2012) 398–404.

[6] M.C. Rochoux, B. Delmotte, B. Cuenot, S. Ricci, A. Trouvé, *Proc. Combust. Inst.* 34 (2013) 2641–2647, doi:10.1016/j.proci.2012.06.090.

[7] O. Rios, E. Pastor, M. Valero, E. Planas, *Intl. J. Wildland Fire* 25 (2016) 1033–1047, doi:10.1071/WF16031.

[8] M.C. Rochoux, S. Ricci, D. Lucor, B. Cuenot, A. Trouvé, *Nat. Hazards Earth Syst. Sci.* 14 (11) (2014) 2951–2973, doi:10.5194/nhess-14-2951-2014.

[9] M.C. Rochoux, C. Emery, S. Ricci, B. Cuenot, A. Trouvé, *Nat. Hazards Earth Syst. Sci.* 15 (8) (2015) 1721–1739, doi:10.5194/nhess-15-1721-2015.

[10] C. Zhang, M.C. Rochoux, W. Tang, M. Gollner, J.B. Filippi, A. Trouvé, *Fire Saf. J.* 91 (2017) 758–767, doi:10.1016/j.firesaf.2017.03.057.

[11] NIROPS, US Forest Service National Infrared Operations Unit, Available at https://fsapps.nwcg.gov/nirops/.

[12] MODIS, Active Fire and Burned Area Products, Available at http://modis-fire.umd.edu/.

[13] R. Paugam, M. Wooster, G. Roberts, *Geosci. Remote Sens.* 51 (2013) 3385–3399.

[14] P. Arbogast, O. Pannekoucke, L. Raynaud, R. Lalanne, E. Mémin, *Q. J. R. Meteorol. Soc.* 142 (700) (2016) 2827–2838, doi:10.1002/qj.2871.

[15] N. Feyeux, *Transport optimal pour l'assimilation de données images*, Communauté Université Grenoble Alpes, 2016 Ph.D. thesis.

[16] J. Chang, S. Hanna, *Meteorol. Atmos. Phys.* 87 (1) (2004) 167–196, doi:10.1007/s00703-003-0070-7.

[17] T.F. Chan, L.A. Vese, in: Proceedings of the IEEE Workshop on Variational and Level Set Methods, in: VLSM, IEEE Computer Society, Washington, DC, USA, 2001, pp. 161–168.

[18] A. Collin, D. Chapelle, P. Moireau, *J. Comput. Phys.* 300 (C) (2015) 288–307.

[19] J. O'Brien, E. Loudermilk, B. Hornsby, et al., *Intl. J. Wildland Fire* 25 (1) (2016) 62–75.

[20] M. Rochoux, A. Collin, C. Zhang, A. Trouvé, D. Lucor, P. Moireau, *ESAIM Proc. Surv.* (2017) 1–22.

[21] R. Rothermel, A Mathematical Model for Predicting Fire Spread in Wildland Fuels, *Technical Report*, USDA Forest Service, Intermountain Forest and Range Experiment station, 1972.

[22] R. Rehm, R. McDermott, Fire Front Propagation Using the Level Set Method, *Technical Report*, National Institute of Standards and Technology, 2009.

[23] G.D. Richards, *Int. J. Numer. Methods Eng.* 30 (6) (1990) 1163–1179, doi:10.1002/nme.1620300606.

[24] S. Osher, R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, 153, Applied Mathematical Sciences - Springer, 2003.

[25] S.D. Roth, *Comput. Graph. Image Process.* 18 (2) (1982) 109–144, doi:10.1016/0146-664X(82)90169-1.

[26] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (9) (1993) 850–863.