FRONT SHAPE SIMILARITY MEASURE FOR SHAPE-ORIENTED SENSITIVITY ANALYSIS AND DATA ASSIMILATION FOR EIKONAL EQUATION

M.C. Rochoux[1], A. Collin[2], C. Zhang[3], A. Trouvé[4], D. Lucor[5] and P. Moireau[6]

Abstract. We present a shape-oriented data assimilation strategy suitable for front-tracking problems through the example of wildfire. The concept of "front" is used to model, at regional scales, the burning area delimitation that moves, undergoes shape and topological changes under heterogeneous orography, biomass fuel and micrometeorology. The simulation-observation discrepancies are represented using a front shape similarity measure deriving from image processing and based on the Chan-Vese contour fitting functional. We show that consistent corrections of the front location and uncertain physical parameters can be obtained using this measure applied on a level-set fire growth model solving for an eikonal equation. This study involves a Luenberger observer for state estimation, including a topological gradient term to track multiple fronts, and of a reduced-order Kalman filter for joint parameter estimation. We also highlight the need – prior to parameter estimation – for sensitivity analysis based on the same discrepancy measure, and for instance using polynomial chaos metamodels, to ensure a meaningful inverse solution is achieved. The performance of the shape-oriented data assimilation strategy is assessed on a synthetic configuration subject to uncertainties in front initial position, near-surface wind magnitude and direction. The use of a robust front shape similarity measure paves the way toward the direct assimilation of infrared images and is a valuable asset in the perspective of data-driven wildfire modeling.

[1] CECI, Université de Toulouse, CNRS, CERFACS, 42 avenue Gaspard Coriolis, F-31057 Toulouse cedex 1, France.

[2] Inria, Université de Bordeaux (IMB), Bordeaux INP, France,

[3] Department of Fire Protection Engineering, University of Maryland, College Park, MD 20742, USA.

[4] Department of Fire Protection Engineering, University of Maryland, College Park, MD 20742, USA.

[5] LIMSI, CNRS, Université Paris-Saclay,

Campus universitaire bât 508, Rue John von Neumann, F-91405 Orsay cedex, France

[6] Inria – LMS, Ecole Polytechnique, CNRS – Université Paris-Saclay, France

Résumé. Nous présentons ici une stratégie d'assimilation de données de formes adaptée aux problèmes de suivi de front lors d'incendies. Le concept de "front" est utilisé pour modéliser, à l'échelle régionale, la limite de la surface brûlée qui se déplace, se déforme et change de topologie en raison des hétérogénéités de relief, combustible végétal et micrométéorologie. Les différences entre fronts simulé et observé sont représentées via une mesure de similitude de formes tirée du traitement d'images et fondée sur la fonctionnelle d'ajustement de contour de Chan-Vese. Nous montrons que cette mesure permet d'obtenir des corrections cohérentes sur la position du front et les paramètres physiques incertains dans le cadre d'un modèle d'ensembles de niveaux (*level set*) de la propagation d'incendie résolvant une équation eikonale. Cette étude repose sur un observateur de Luenberger pour l'estimation d'état (incluant un terme de gradient topologique pour capturer plusieurs fronts) et un filtre de Kalman d'ordre réduit pour l'estimation conjointe de paramètres. Nous mettons aussi en évidence l'importance de mener, au préalable de l'estimation de paramètres, une analyse de sensibilité s'appuyant sur la même mesure de similitude, et par exemple à l'aide de métamodèles de type chaos polynomiaux, pour s'assurer d'une solution physique au problème inverse. La performance de la stratégie d'assimilation de données de type objet est évaluée sur un cas synthétique sujet à des incertitudes sur la position initiale du front, l'intensité et la direction du vent proche du sol. L'utilisation d'une mesure robuste de similitude de formes de front ouvre la voie à l'assimilation d'images infra-rouge et est un atout en vue d'une modélisation des incendies recalée sur les données.

Introduction

Front-tracking problems arise in many applications, e.g. cardiac electrophysiology [15], tumor growth [36], seismic history matching [62], oil spill [34], precipitation forecasting [4], flaming combustion [44], wildfire behavior [52]. Such problems aim at tracking the contour (or "front") and the motion of an object, which is the indirect shape of, for instance, an electric field, a water saturation, a cloud, a temperature isovalue, etc. The object contour may be subject to strong shape deformations, its motion may be unsteady and present irregularities and topological changes can occur. Tracking these changes remains a challenging task from both modeling and observation viewpoint [42, 64]. Simulating the time-evolving front shape requires as inputs many uncertain quantities such as the initial front position and physical parameters that are usually difficult to measure. The model formulation itself cannot perfectly describe a physical system and its evolution, because of uncertainties on the physical and numerical parameters as well as uncertainties related to the simplification and discretization of the physics and to the numerical methods. Observations are also subject to instrumental and representativeness errors that are enhanced when the target object becomes partially occluded.

To overcome these limitations, data assimilation is considered as a powerful approach. Data assimilation aims at integrating available observations into the front-tracking model, while accounting for their respective uncertainties, in order to infer more accurate front position estimates and to produce more reliable forecasts [48]. The estimation process can target the model input parameters and/or the model prognostic variables [39, 57]; these estimated targets are referred to as the "control variables". The success of data assimilation lies in particular in the measure used to represent what we call the "innovation", i.e. the differences between the observations and the simulated counterparts from which the correction in the control space is derived. This measure shall be consistent with the physics of the problem to ensure optimal estimation. Classical data assimilation methods assume that errors are only of amplitude type by representing errors as additive Gaussian noise. In front-tracking problems, they perform poorly when the simulated contour is displaced from its observation. They compensate position errors by adjusting amplitudes, which can produce non-physical solutions and do not preserve coherent structures [4, 7, 21, 22, 46]. Our objective is therefore to design a shape-oriented data assimilation approach that addresses position errors, in which the innovation measures the observation-simulation distance in terms of shape discrepancies.

This is particularly relevant in the context of data-driven wildfire spread modeling. In previous works [51–54, 65] the observed fronts were treated as a discretized contour with a finite set of equally spaced markers. There,

a distance to the simulated fronts was computed by pairing each observed marker with its closest neighbor along the simulated front. Through this selection procedure, the discrepancies between the observations and the simulated fronts (i.e. the innovations) were formulated as Euclidean distances that have the advantage of featuring Gaussian error statistics. However, one limitation of such distance computation is that the selection procedure may be difficult to operate for large scale wildfires with complex front topology with significant temporal changes due to heterogeneous environment and fire behavior. That is why, we introduce in this paper a new distance metric to design a more robust data assimilation algorithm. This avoids specific treatment of the front markers [52]. This also avoids the use of morphing that involves a complex registration mapping [7] and provides valuable spatial information compared to usual similarity scores [6,23]. To this end, our goal is to rely on a front shape similarity measure introduced in the context of data assimilation for cardiac electrophysiology [14–16]. This measure – initially envisioned for object detection in image processing theory – formulates a shape similarity measure inspired from the Mumford-Shah functional, in particular its Chan-Vese version [10, 40]. In the present study, our first objective is to show that this front shape similarity measure can be generalized and successfully applied for level-set front-tracking simulator in the framework of data-driven wildfire spread modeling for both parameter estimation – using an Unscented Kalman Filter (UKF) – and state estimation – using a Luenberger observer (LO) [37, 38]. Our second objective is to highlight how this similarity measure can be used in an uncertainty quantification process – using a Polynomial Chaos (PC) strategy [33, 54, 55, 63] – to study the measurement sensitivity with respect to parameter changes, hence determine the influence of each control parameter on the variability of the front shape similarity measure. This is useful to ensure a meaningful feedback is achieved on the control parameters when performing joint state-parameter estimation. Even though different types of surrogates are reported in the literature [32], we focus here our attention on PC-based surrogate that is built accounting for the error statistics on the uncertain input variables and that is thus well adapted to the data assimilation framework associated with parameter estimation and already envisioned for interface variability studies [45].

1. FIRE FRONT-TRACKING PROBLEM: MODELS AND OBSERVATIONS

At regional scales (ranging from a few tens of meters up to several hectares), a wildfire is usually described as a 2-D front that self-propagates along its normal into the unburnt vegetation. The (fire) front separating the active flame areas from the unburnt vegetation propagates and deforms according to the environmental conditions such as local terrain orography, micrometeorology and biomass fuel conditions. The main quantity of interest is the propagating speed of the front commonly known as the "rate of spread" (ROS). This regional-scale viewpoint is adopted in current operational wildfire spread simulators, e.g. FARSITE [25]. These simulators require a parametric ROS formulation and a 2-D numerical scheme in order to account for wind and slope effects on fire growth. There is a wide variety of 2-D schemes, which can be sorted into two main types [9, 27, 47]: (1) Lagrangian type (i.e. tracking marker trajectories over time) using for instance asynchronous advection method based on discrete-event simulation [24] or Huygens-based ellipse expansion [49]; and (2) Eulerian type (i.e. tracking the burning area object over time) using for instance level-set methods [35, 41, 58, 59]. In the present work, we focus on Eulerian front-tracking simulators. While beyond the scope of this work, our data assimilation strategy could be easily extended to Lagrangian fire growth models.

1.1. Fire growth model

Several ROS submodels are reported in the literature [61]. The most widely known is the model due to Rothermel [3, 56] that evaluates the ROS in a pointwise manner given local orography, biomass and meteorological conditions:

$$\text{ROS}_{1D} \equiv \text{ROS}_{1D}\left([M_v, \delta_v, m_v'', \rho_v, \Sigma_v, \Delta h_v], d_{sl}, U_w\right), \tag{1}$$

with M_v [%] the fuel moisture (mass of water divided by mass of dry vegetation), δ_v [m] the fuel depth (vertical thickness of the vegetation layer), m_v'' [kg m^{-2}] the fuel surface loading, ρ_v [kg m^{-3}] the fuel mass density,

$\Sigma_{\rm v}$ [m^{-1}] the fuel particle surface-to-volume ratio, $\Delta h_{\rm v}$ [J kg^{-1}] the fuel heat of combustion, $d_{\rm sl}$ [rad] the terrain slope angle and $U_{\rm w}$ [m s^{-1}] the wind velocity vector at mid-flame height.

To represent the time-evolving burning active areas over the computational domain $\Omega \subset \mathbb{R}^2$, we introduce a regular progress variable $c \equiv c(\mathrm{x}, t)$ as the fire front marker. The fire front is identified as the contour line $c_{\rm fr} = 0.5$. Therefore $c < c_{\rm fr}$ corresponds to the unburnt vegetation, $c > c_{\rm fr}$ corresponds to the burnt vegetation. We thus denote the active burning areas $\mathcal{B}_c = \{\mathrm{x} = (x_1, x_2) \in \Omega \,|\, c(\mathrm{x}, t) > c_{\rm fr}\}$ and, when regular enough, the front line $\Gamma_c = \{\mathrm{x} = (x_1, x_2) \in \Omega \,|\, c(\mathrm{x}, t) = c_{\rm fr}\}$. The progress variable c is used as the prognostic variable and is calculated as a solution of the eikonal equation

$$\frac{\partial c}{\partial t}(\mathrm{x}, t) = \mathrm{ROS_{2D}} \, |\nabla c(\mathrm{x}, t)|\,, \quad c(\mathrm{x}, t_0) = c_0(\mathrm{x}), \quad \mathrm{x} \in \Omega, \ t \geq t_0, \tag{2}$$

with $c_0(\mathrm{x})$ the initial condition at time t_0. The quantity $\mathrm{ROS_{2D}}$ is an adequate extension [1, 12, 35] on the whole domain Ω of the Rothermel-based $\mathrm{ROS_{1D}}$ (Eq. 1) calculated on the front Γ_c. $\mathrm{ROS_{1D}}$ corresponds to a 1-D parametric function that is applied in 2-D cases by accounting for the local normal vector $\mathrm{n_{fr}} \equiv \mathrm{n_{fr}}(\mathrm{x})$. Note that for solving Eq. (2), we follow the choices made by [47] 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 [51].

1.2. Sources of model uncertainties

Despite the development of many wildfire-spread models, their use has been relatively limited operationally to respond to actively burning fires as they contain numerous uncertainties. Typically, the ROS submodel introduces epistemic uncertainties due to its limited domain of validity resulting from a calibration procedure based on laboratory-scale experiments [61]. The input parameters, especially the biomass moisture content $M_{\rm v}$ and the near-surface wind velocity $U_{\rm w}$ (see Eq. 1), are usually available with limited spatio-temporal resolution; they are subject to significant uncertainties that translate into uncertainties in the ROS predictions and thereby in the simulated front positions [17, 18, 31]. There is in general not much information on the fire ignition $c_0(\mathrm{x})$, so the initial position and shape of the fire are also part of the uncertainties in the wildfire problem.

Since this paper is methodology-oriented, we focus our attention on a test case with idealized conditions. We only consider the initial front condition c_0 and the near-surface wind velocity vector $U_{\rm w}$ as sources of uncertainties; they are known to have the most significant impact on the fire front shape and position. Note that in the present work, the wind velocity vector is specified using two scalar parameters, the wind magnitude $u_{\rm w}$ ($u_{\rm w} \equiv \|U_{\rm w}\|$) and the wind direction angle $d_{\rm w}$ such that $U_{\rm w} = (u_{\rm w} \sin d_{\rm w}, u_{\rm w} \cos d_{\rm w})^T$.

In the following, the front position uncertainties due to c_0 are addressed through state estimation. In practice, the uncertainties in the initial condition will be formulated in a parametric form with respect to the "center of mass" of the initial burning area, whose position is denoted by $\mathrm{x_{ign}} = (x_{1,\mathrm{ign}}, x_{2,\mathrm{ign}})$. The front position uncertainties due to $(u_{\rm w}, d_{\rm w})$ are addressed through parameter estimation. Control parameters are assumed uniform here. Our method extends easily to regular spatially distributed parameters – see for instance [65] – as soon as the spatial discretization of the parameters can be formulated with a reasonable number of degrees of freedom [11, 37]. The parameters $\mathrm{x_{ign}}$ and $(u_{\rm w}, d_{\rm w})$ are considered as independent random variables, each random variable being characterized by a marginal probability density function (PDF) that is either Gaussian (Kalman-based data assimilation, Sec. 2.2.3) or uniform (sensitivity analysis, Sec. 3).

1.3. Observations and Observing System Simulation Experiment (OSSE)

A wide inventory of wildfire observations is presented in [28]; the most valuable data for data-driven wildfire spread modeling are the time-evolving fire front locations that can be derived from infrared remote sensing images [13, 43, 50]. Several data structures can be envisioned. The first data structure is a map $\mathrm{z^{obs}} \equiv \mathrm{z^{obs}}(\mathrm{x}, t)$ of the burnt and unburnt regions (see Fig. 1). The second data structure is a close-contour delimiting the burnt regions. In this case, a binary map $\mathrm{z^{obs}} \equiv \mathrm{z^{obs}}(\mathrm{x}, t)$ corresponding to the interior points of this 2-D front could
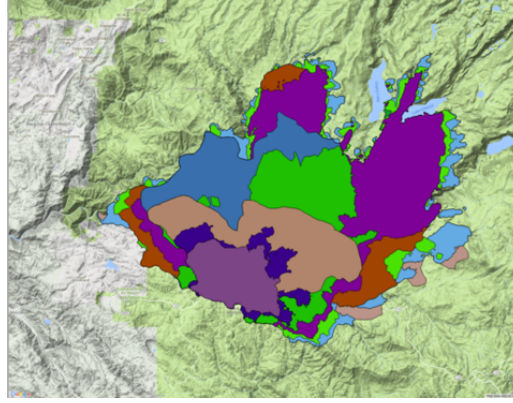
FIGURE 1. Typical burning area mapping from observations – Example of Rim fire (California, 20-25 August 2013), GeoMAC data (www.geomac.gov): each color represents the burning area at a given time from which a binary field (0 if unburnt, 1 if burnt) is reconstructed. Source: Evan Ellicott (Department of Geographical Sciences, University of Maryland).

be reconstructed. The third structure consists of a time-map $t_{\text{ign}} \equiv t_{\text{ign}}(\text{x})$ of the fire propagation. Then, a binary map can be defined by

$$z^{\text{obs}} \equiv z^{\text{obs}}(\text{x}, t) = \mathbb{1}_{t > t_{\text{ign}}}(\text{x}). \tag{3}$$

Moreover, we point out that our data $z^{\text{obs}}(\cdot, t)$ could also be defined as a continuous map $\mathcal{C}([0, 1])$ in order to represent the uncertain location of the front by a fuzzy set [8].

In the present observing system simulation experiment (OSSE) framework, observations are synthetically generated by integrating the fire growth model in Eq. (2) using reference (or "true") values of the uncertain parameters; the error in the observations is assumed to be small and the observation is then considered as the "target". The prior (or "background") values of the control variables are obtained by perturbing the true values. OSSE thus provides a verification framework to check the ability of the data assimilation algorithm to retrieve the target for varying uncertainty in the control variables and in the observations. In the present work, data assimilation aims at tracking the observed fire front when starting from a poor prior estimate by inferring wind parameters $(u_{\text{w}}, d_{\text{w}})$ that are more consistent with the reference values and/or by directly correcting the simulated field $c(\text{x}, t)$ to reduce uncertainties due to a wrong initial front position x_{ign}.

2. FRONT SHAPE SIMILARITY MEASURE ADAPTED TO FRONT DATA ASSIMILATION

2.1. Principles of sequential data assimilation

In essence, sequential data assimilation approaches aim at correcting a model dynamics by the use of the current available data. Let us consider that our forward model noted $\mathcal{M}$ is given by Eq. (2). The target trajectory cannot be simulated precisely since the initial condition c_0 and some parameters θ entering in ROS_{2D} are uncertain. We aim at estimating a trajectory characterized by the dynamics $\dot{y} = \mathcal{M}(y, t), y(0) = y_0$, with $y : t \mapsto (c(\cdot, t), \theta(t))$ the "extended state" at a function of the time t. The observations are cast in the observation vector $z^{\text{obs}} \in \mathcal{Z}$, with $\mathcal{Z}$ the observation space equipped with its norm $\| \cdot \|_{\mathcal{Z}}$. The observation operator $\mathcal{H}$ is defined as the map between the control space and the observation space such that $z = \mathcal{H}(y)$ denotes the simulated counterpart of the observations corresponding to the model state y, itself depending on the initial condition c_0 and on the parameters θ.

To begin with, let us consider a classical data assimilation approach. In this context, we typically assume that a similarity measure between the given observations z^{obs} and the simulated counterparts z can be computed

by a "least square" type data fitting functional of the form

$$\mathcal{J}_{\text{data}}(y) = \frac{1}{2}\|\mathcal{D}\|^2_{\mathcal{Z}},\tag{4}$$

with a discrepancy operator $\mathcal{D}$ of the form $\mathcal{D}(z^{\text{obs}}, y) = z^{\text{obs}} - z = z^{\text{obs}} - \mathcal{H}(y)$. Then a sequential estimator $\widehat{y}$ uses modified dynamics of the form

$$\begin{cases} \dot{\widehat{y}} = \mathcal{M}(\widehat{y}, t) + K\big(z^{\text{obs}} - \mathcal{H}(\widehat{y})\big) \\ \widehat{y}(0) = \widehat{y}_0 \end{cases}\tag{5}$$

where K is the sequential gain, in order to converge asymptotically to the target trajectory when assimilating noiseless data, namely $\widehat{y} \xrightarrow{t\to+\infty} y$; robustness property to noisy data is then obtained in a second step. Note that, as $\widehat{y}(t) = (\widehat{c}(\cdot, t), \widehat{\theta}(t))$ introduces dynamics for the estimated parameters $\widehat{\theta}(t)$, the property $\widehat{\theta} \xrightarrow{t\to+\infty} \theta$ allows to identify asymptotically the time invariant target parameters θ. The gain is in general of the form $K = P\, d\mathcal{H}^*$ where $d\mathcal{H}$ denotes the differential of $\mathcal{H}$ and P is a positive operator [11], hence the estimator correction of the dynamics is of the form

$$K(z^{\text{obs}} - \mathcal{H}(y)) = P\, d\mathcal{H}^*(z^{\text{obs}} - \mathcal{H}(y)) = -P\, \nabla\mathcal{J}_{\text{data}}(y).\tag{6}$$

This formulation is natural as the term $-\nabla\mathcal{J}_{\text{data}}(y)$ gives a direction of maximum decrease in the dissimilarity between the given observations and the simulation.

2.2. Discrepancy operator and data assimilation

2.2.1. *The Chan-Vese contour fitting functional*

The classical approach presented in Sec. 2.1 suffers from a strong limitation for front-tracking problems. Indeed, it is particularly difficult to define a data fitting term based on a simple Euclidean distance. A first approach was performed in [52] using a Lagrangian front description based on marker position comparison. This Euclidean computation was efficient but limited to a given front topology and suffers from potential mismatches during marker registration. To be more general, we investigate a more complex data fitting functional derived from image active contour segmentation approaches. Following the work of [15], an adequate similarity between the target front and the simulated one can be computed from the Chan-Vese contour fitting functional [10,40,41]. Let us define the 1-D Heaviside distribution H_v such that $H_v(\phi) = 0$ if $\phi < 0$, $H_v(\phi) = 1$ if $\phi > 0$. The Chan-Vese contour fitting functional can be written as

$$\mathcal{J}_{\text{data}}(\phi, z^{\text{obs}}) = \int_\Omega H_v(\phi)\left[z^{\text{obs}} - C_{\max}(z^{\text{obs}}, \phi)\right]^2 + (1 - H_v(\phi))\left[z^{\text{obs}} - C_{\min}(z^{\text{obs}}, \phi)\right]^2 dx,\tag{7}$$

where ϕ is a level set associated with c, for instance $\phi = c - c_{\text{fr}}$, and where $C_{\min}$ and $C_{\max}$ are scalar coefficients defined by

$$C_{\min} = \min(C_0, C_1), \quad C_{\max} = \max(C_0, C_1),\tag{8}$$

with

$$C_1(z^{\text{obs}}, \phi) = \frac{\int_\Omega H_v(\phi)\, z^{\text{obs}}\, dx}{\int_\Omega H_v(\phi)\, dx}, \quad C_0(z^{\text{obs}}, \phi) = \frac{\int_\Omega (1 - H_v(\phi))\, z^{\text{obs}}\, dx}{\int_\Omega (1 - H_v(\phi))\, dx}.\tag{9}$$

The constant C_1 then corresponds to the mean of z_{obs} on the region where $c > c_{\text{fr}}$ (i.e. $\phi > 0$), namely the simulated burnt region. Respectively C_0 corresponds to the mean of z_{obs} on the region where $c < c_{\text{fr}}$ (i.e. $\phi < 0$), namely the simulated unburnt region.

In the following, we provide some elements of understanding about this criteria through some illustrated examples displayed in Figure 2. For instance, we consider the case where the burnt region – corresponding to a

progress variable $c > 0$ – is included in the observed burnt region $z^{obs} = 1$, see Figure 2-C. The constant C_1 is the mean value of z_{obs} in the burnt region. Considering that $z_{obs} \equiv 1$ in this region, we thus have $C_1 = 1$. By contrast, we have $0 \leq C_0 \leq 1$ which is minimum when the contours $\Gamma = \{c = c_{fr}\}$ delimit exactly the burnt and unburnt regions observed through z^{obs}, see Figure 2-A. In this case, $C_{min} = C_0$, $C_{max} = C_1$ and $\mathcal{J}$ is null when the observed and simulated contours coincide. A second example – see Figure 2-E – comes when we consider a progress variable c with a burnt region completely disjoint from the observed burnt region. Then $C_1 = 0$ and $0 \leq C_0 \leq 1$ with a maximum value of 1 when the observed and simulated contours coincide. In this last case, ϕ and z^{obs} are in complete phase shift, hence the interest of introducing $C_{min} = C_1$ and $C_{max} = C_0$ in order for $\mathcal{J}$ to be maximum. We emphasize that introducing C_{min} and C_{max} is original here with respect to what is classically done with the Chan-Vese functional in image processing. Indeed, in its classical image segmentation use, there is no need to penalize a phase shift.
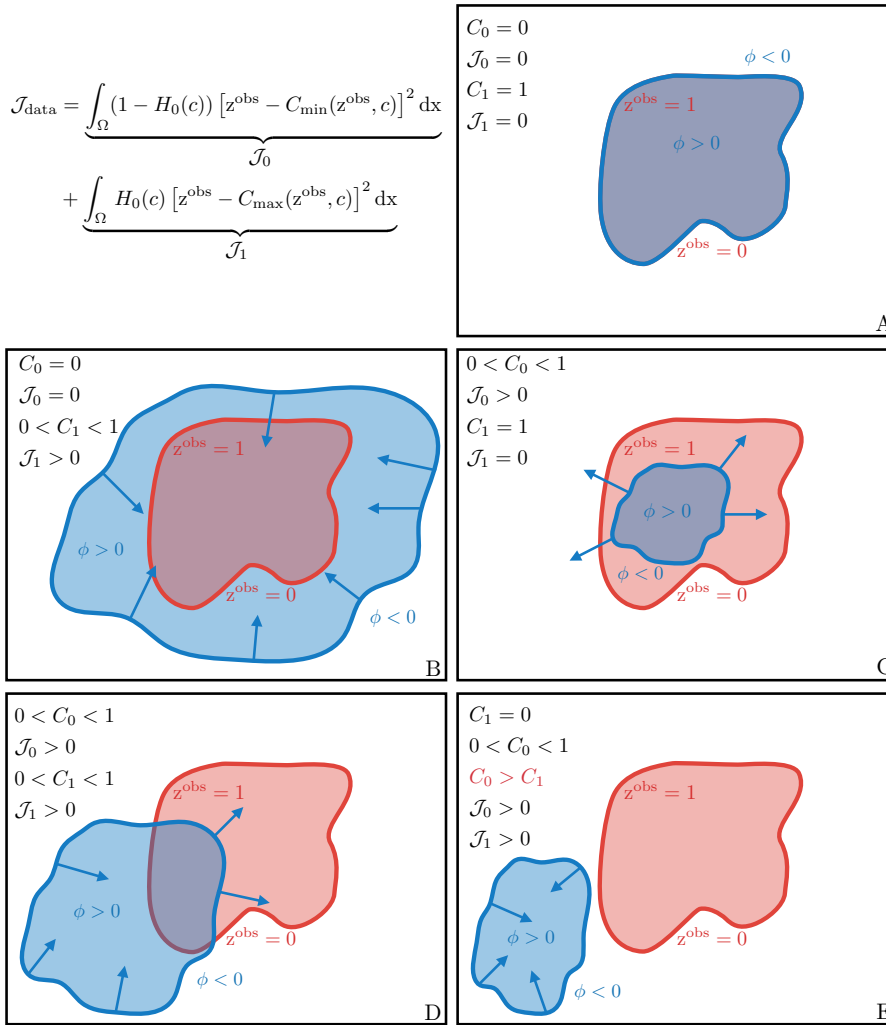

FIGURE 2. Behavior of Chan-Vese functional for some typical configurations of simulated and observed fronts. The arrows indicate how the shape-based state feedback acts on the contour of the observer simulated front.

Having defined the discrepancy functional $\mathcal{J}$, our objective is now to propose a sequential estimator based on a correction of the form (6) for the model dynamics and compatible with the data fitting functional of Eqs. (7)–(8)–(9).

2.2.2. A state observer

Let us first assume that all the parameters in θ are perfectly known. Only the initial condition c_0 is uncertain – due to uncertainty in x_{ign} – and we thus focus on state estimation only. In [15], the bidomain eikonal equation was corrected based on the Chan-Vese contour fitting functional in Eq. (7). To do so, as envisioned in Eq. (6), [15] computed the shape gradient of the similarity measure, namely

$$\nabla_{sh}\mathcal{J}_{data} = \delta(\phi) \left\{ \left[z^{obs} - C_{max}(z^{obs}, \phi) \right]^2 - \left[z^{obs} - C_{min}(z^{obs}, \phi) \right]^2 \right\}, \tag{10}$$

with δ the Dirac distribution playing the role of a localization operator, meaning that the gradient is concentrated on the front. This means that the variations of the criterion (7) can be diminished by acting on the front contour only. From the gradient computation in Eq. (10), [15] then proposed a correction of the form $-\lambda_{sh}\nabla_{sh}\mathcal{J}_{data}$ with a spatially distributed scalar gain $\lambda_{sh} \equiv \lambda_{sh}(x)$ (the subscript "sh" stands for "shape"). Following the same strategy, we propose as Luenberger state observer

$$\begin{cases} \dfrac{\partial\widehat{c}}{\partial t}(x,t) = ROS_{2D} \, |\nabla\widehat{c}(x,t)| \\ \qquad\quad -\lambda_{sh}(x)\,\delta(\phi) \left\{ \left[z^{obs}(x,t) - C_{max}(z^{obs}, \phi) \right]^2 - \left[z^{obs}(x,t) - C_{min}(z^{obs}, \phi) \right]^2 \right\}, \; t \geq t_0 \\ \widehat{c}(x,t_0) = \widehat{c}_0(x), \end{cases} \tag{11}$$

starting from a given inexact initial condition $\widehat{c}_0$. The gain λ_{sh} is a spatially distributed scalar parameter accounting for the local confidence in the observation z^{obs}; typically λ_{sh} is proportional to the inverse of the local observation error covariance. Figure 2 shows – through some illustrated examples – how the simulated front contour will evolve – see the blue arrows – to minimize the discrepancy.

The state observer given by Eq. (10) can deal with unobserved regions. Let us consider that z^{obs} is only defined on Ω^{obs} that is a restriction of Ω. We can replace Eq. (9) by

$$C_1(z^{obs}, \phi) = \frac{\int_{\Omega^{obs}} H_v(\phi)\, z^{obs}\, dx}{\int_{\Omega^{obs}} H_v(\phi)\, dx}, \quad C_0(z^{obs}, \phi) = \frac{\int_{\Omega^{obs}} (1 - H_v(\phi))\, z^{obs}\, dx}{\int_{\Omega^{obs}} (1 - H_v(\phi))\, dx}, \tag{12}$$

and $\lambda_{sh}(x) = \lambda_{sh}(x)\, \mathbb{1}_{\Omega^{obs}}(x)$ is then restricted to the observation region Ω^{obs}.

The state observer is devoted to pursue the target front from the discrepancy sensitivity $\nabla_{sh}\mathcal{J}_{data}$. However, this observer can only capture a number of fronts that is bounded by the number of initial fronts. However, we know that new ignitions can occur in wildfire problems, for instance by fire spotting [20], hence our need to capture new fronts during the estimation procedure. To address this difficulty, we propose to follow [16] and also compute the topological gradient of Eq. (7). A topological gradient, here given by

$$\nabla_{top}\mathcal{J}_{data} = \left\{ \left[z^{obs} - C_{max}(z^{obs}, \phi) \right]^2 - \left[z^{obs} - C_{min}(z^{obs}, \phi) \right]^2 \right\}, \tag{13}$$

typically computes the sensitivity of the discrepancy measure to new appearing fronts [29] (the subscript top stands for "topology").

From the sensitivity (13), we deduce – as in [16] – a data correction adapting the eikonal equation to new appearing fronts, see Figure 2-E for example. The resulting complete observer then reads

$$
\begin{cases}
\dfrac{\partial \widehat{c}}{\partial t}(\mathrm{x},t) = \mathrm{ROS_{2D}} \left| \nabla \widehat{c}(\mathrm{x},t) \right| \\
\quad - \lambda_{\mathrm{sh}}(\mathrm{x})\, \delta(\phi) \left\{ \left[\mathrm{z^{obs}}(\mathrm{x},t) - C_{\max}(\mathrm{z^{obs}},\phi) \right]^2 - \left[\mathrm{z^{obs}}(\mathrm{x},t) - C_{\min}(\mathrm{z^{obs}},\phi) \right]^2 \right\} \\
\quad - \lambda_{\mathrm{top}}(\mathrm{x})\, H_v(\nabla_{\mathrm{top}} \mathcal{J}_{\mathrm{data}} \times (c - c_{\mathrm{fr}})) \left\{ \left[\mathrm{z^{obs}}(\mathrm{x},t) - C_{\max}(\mathrm{z^{obs}},\phi) \right]^2 - \left[\mathrm{z^{obs}}(\mathrm{x},t) - C_{\min}(\mathrm{z^{obs}},\phi) \right]^2 \right\}, \\
\widehat{c}(\mathrm{x},t_0) = \widehat{c}_0(\mathrm{x}).
\end{cases}
$$
(14)

where $\lambda_{\mathrm{sh}} \equiv \lambda_{\mathrm{sh}}(\mathrm{x})$ is the gain associated with the shape gradient and where $\lambda_{\mathrm{top}} \equiv \lambda_{\mathrm{top}}(\mathrm{x})$ is the gain associated with the topological gradient. The impact of this additional topological feedback is typically illustrated in Figure 2-E. Indeed in this case, the simulated front tends to collapse due to the shape-based feedback term (see the blue arrows). However, the topological observer can introduce a new front in the domain $\{z^{\mathrm{obs}} = 1\}$. Note that the theoretical study – even for unnoisy data – of the asymptotic convergence $\widehat{y} \xrightarrow{t \to +\infty} y$ is beyond the scope of this methodology-oriented article. However, based on the results of [15] in a more complex context, we believe it is possible to demonstrate a weaker property, namely that the error $\|\widehat{y} - y\|$ decreases for small initial errors.

2.2.3. *Kalman-based filters and parameter estimation*

An alternative to propose a sequential estimator correction is to rely on a Kalman-based update. For simplicity purpose, we shall describe the Kalman correction after space and time discretization of the initial model and observations. We denote by $\mathcal{Y}_h$ and $\mathcal{Z}_h$ the corresponding finite dimensional state space and observation space. In the following, we use bold letters to refer to the vectors of degrees of freedom associated with the discretized solutions; n refers to the time index and $[n, n+1]$ refers to a given assimilation time window. Over the window $[n, n+1]$, the control augmented state vector is noted $\mathbf{y}_{n+1} = (\mathbf{c}_{n+1}, \boldsymbol{\theta}_{n+1})$. When considering a prior estimate of the state vector $\mathbf{y}^{\mathrm{b}}_{n+1}$ (the superscript b stands for "background") and when the similarity measure has a least square structure as in Eq. (4), the posterior estimate $\mathbf{y}^{\mathrm{a}}_{n+1}$ (the superscript a stands for "analysis") is obtained using the Kalman filter update equation

$$
\mathbf{y}^{\mathrm{a}}_{n+1} = \mathbf{y}^{\mathrm{b}}_{n+1} + \mathbf{K}^{\mathrm{al}}_{n+1} \left[\boldsymbol{D} \left(\mathbf{z}^{\mathrm{obs}}_{n+1}, \boldsymbol{H}(\mathbf{y}^{\mathrm{b}}_{n+1}) \right) \right].
$$
(15)

$\mathbf{K}^{\mathrm{al}}_{n+1}$ is the Kalman-type gain matrix such that $\mathbf{K}^{\mathrm{al}}_{n+1} = \mathbf{P}^{\mathrm{b}}_{\mathrm{yy}} \mathrm{d}\mathbf{H}^T \left(\mathrm{d}\mathbf{H}\, \mathbf{P}^{\mathrm{b}}_{\mathrm{yy}} \mathrm{d}\mathbf{H}^T + \mathbf{P}^{\mathrm{o}}_{\mathrm{zz}} \right)^{-1}$, where $\mathbf{P}^{\mathrm{b}}_{\mathrm{yy}}$ is the augmented state background error covariance matrix containing covariances between the errors in the state variable and the cross-covariances between errors in the parameters and in the state variable, where $\mathbf{P}^{\mathrm{o}}_{\mathrm{zz}}$ is the observation error covariance matrix, and where $\mathrm{d}\mathbf{H}$ is the derivative of the observation operator $\boldsymbol{H}$. Note that Eq. (15) is consistent with Eq. (6) as the time-discrete Kalman gain integrates the time step when accounting for discrete white noise [11].

In order to be compatible with a Kalman filter update computation of the form of (15), we thus need a discrepancy $\mathcal{J} = \frac{1}{2}\|\mathrm{D}\|^2_{\mathcal{Z}_h}$ of least square type associated with our front shape similarity measure. The Chan-Vese functional in Eq. (7) involves Heaviside distribution cannot be multiplied univocally [26], hence our functional is not a least square criterion. However, a pseudo least square strategy can be envisioned by decomposing the Chan-Vese functional into two functionals

$$
\mathcal{J}_{\mathrm{data}}(\mathrm{y}) = \mathcal{J}^+_{\mathrm{data}}(\mathrm{y}) + \mathcal{J}^-_{\mathrm{data}}(\mathrm{y}),
$$
(16)

with

$$
\mathcal{J}^+_{\mathrm{data}}(\mathrm{y}) = \int_\Omega H_v(\phi) \left[\mathrm{z^{obs}} - C_{\max}(\mathrm{z^{obs}},\phi) \right]^2 \mathrm{dx}, \quad \mathcal{J}^-_{\mathrm{data}}(\mathrm{y}) = \int_\Omega (1 - H_v(\phi)) \left[\mathrm{z^{obs}} - C_{\min}(\mathrm{z^{obs}},\phi) \right]^2 \mathrm{dx}. \quad (17)
$$

Indeed these two functionals can be approximated – see for instance [2] – by the least square functionals

$$
\begin{cases}
\mathcal{J}_{\mathrm{data}}^{\epsilon,+}(\mathrm{y}) = \dfrac{1}{2}\displaystyle\int_{\Omega} \left\{ \left[1 + \dfrac{2}{\pi}\arctan\left(\dfrac{\phi}{\epsilon}\right)\right]\left[z^{\mathrm{obs}} - C_{\max}(z^{\mathrm{obs}},\phi)\right] \right\}^2 \mathrm{dx}, \\[4mm]
\mathcal{J}_{\mathrm{data}}^{\epsilon,-}(\mathrm{y}) = \dfrac{1}{2}\displaystyle\int_{\Omega} \left\{ \left[1 - \dfrac{2}{\pi}\arctan\left(\dfrac{\phi}{\epsilon}\right)\right]\left[z^{\mathrm{obs}} - C_{\min}(z^{\mathrm{obs}},\phi)\right] \right\}^2 \mathrm{dx},
\end{cases}
\tag{18}
$$

where ϵ is here a small parameter defined with respect to the contour sharpness [19]. Integrating these two functionals into a Kalman-type filter algorithm is straightforward as it only consists in aggregating two sources of information, hence $\boldsymbol{D}\left(\mathbf{z}_{n+1}^{\mathrm{obs}}, \boldsymbol{H}(\mathbf{y}_{n+1}^{\mathrm{b}})\right) = (\boldsymbol{D}^{\epsilon,+}, \boldsymbol{D}^{\epsilon,-})^T$ with $\boldsymbol{D}^{\epsilon,+}$ and $\boldsymbol{D}^{\epsilon,-}$ discretizing the quantities

$$
\begin{cases}
\boldsymbol{D}^{\epsilon,+} = \left[1 + \frac{2}{\pi}\arctan\left(\frac{\phi}{\epsilon}\right)\right]\left[\mathbf{z}^{\mathrm{obs}} - C_{\max}(\mathbf{z}^{\mathrm{obs}}, \boldsymbol{\phi})\right], \\[3mm]
\boldsymbol{D}^{\epsilon,-} = \left[1 - \frac{2}{\pi}\arctan\left(\frac{\phi}{\epsilon}\right)\right]\left[\mathbf{z}^{\mathrm{obs}} - C_{\min}(\mathbf{z}^{\mathrm{obs}}, \boldsymbol{\phi})\right].
\end{cases}
\tag{19}
$$

As the state dynamics is already controlled by the LO (Eq. 14), we can restrict the definition of the Kalman filtering approach to the parameter space [37,54]. This is usually referred to as reduced-order Kalman filtering – ROKF – methods. Hence, our estimation strategy combines a Luenberger-type update that directly integrates the gradient of the Chan-Vese functional and a Kalman-type update based on a least square reformulation of the approximated Chan-Vese functional.

Note that the PDF associated with the control parameters $\boldsymbol{\theta}$ are assumed to be Gaussian in Kalman filtering. Note also that we consider an ensemble-based Kalman filter to account for some of the nonlinearities in the forward model. The key idea is to define a set of particles based on the prior PDF to estimate the Kalman gain $\mathbf{K}$ and more precisely, the error covariance statistics. Each particle represents one forward model integration and thus one realization of the state vector $\mathbf{y}$. In the present study, we use the reduced-order UKF algorithm noted ROUKF, meaning that the prior PDF is approximated by a deterministic sampling of particles (or "sigma-points") [37].

2.2.4. *Joint state-parameter estimation algorithm*

Eventually at each model time step, our joint state Luenberger Observer and parameter Reduced Order Unscented Kalman Filter (LO-ROUKF) data assimilation algorithm reads:

(1) sample sigma-points around the mean parameter based on the estimated parameter covariance and define corresponding state particles using the estimated parameter-state sensitivity;
(2) propagate the Luenberger state observer (LO) defined by the dynamics of Eq. (14);
(3) estimate the resulting model sensitivity with respect to the control parameters;
(4) compute independently the two data fitting terms $\boldsymbol{D}^{\epsilon,+}$ and $\boldsymbol{D}^{\epsilon,-}$ in Eq. (19);
(5) estimate parameter mean value and covariance based on sensitivity and data fitting terms;
(6) correct the state prediction by a ROUKF analysis step based on sensitivity and data fitting terms using Eq. (15).

3. Polynomial chaos (PC) based sensitivity analysis

3.1. **Polynomial chaos (PC) framework**

To make parameter estimation efficient, the choice of the control parameters $\boldsymbol{\theta}$ is of prime importance. They have to be actual sources of uncertainties, and the observed quantities have to be sensitive to them. We thus need to design a cost-effective approach that can measure this sensitivity, while being adapted to the Kalman-type front shape similarity measure presented in Eq. (19). For this purpose, we introduce a numerical strategy based on a PC expansion [54, 63]. The key idea is to replace the front shape similarity measure $\boldsymbol{D}^{\epsilon,\pm}$ obtained

with the fire growth model by a surrogate (or surface response) $\boldsymbol{D}_{\mathrm{pc}}^{\pm}$ of the form

$$\boldsymbol{D}_{\mathrm{pc}}^{\pm}(\boldsymbol{\theta}(\boldsymbol{\zeta})) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}^P} \gamma_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}). \tag{20}$$

$\boldsymbol{\theta} = (\theta_1, \theta_2)$ is defined in the input physical space and its counterpart in the standard probabilistic space is noted $\boldsymbol{\zeta} = (\zeta_1, \zeta_2)$ such that $\zeta_i = \frac{\theta_i - \mu_i}{\sigma_i}$ with μ_i the mean value and σ_i the standard deviation (STD) associated with the ith uncertain parameter θ_i in $\boldsymbol{\theta}$. $\boldsymbol{\theta}$ is thus rescaled in the standard probabilistic space to which the PC framework applies. It can be projected onto a stochastic space spanned by orthonormal polynomial functions $\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta})\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$. $\Psi_{\boldsymbol{\alpha}}$ is the $\boldsymbol{\alpha}$th multivariate basis function chosen in adequacy with the prior PDF $\rho_{\boldsymbol{\theta}}$ associated with the parameters $\boldsymbol{\theta} = (u_{\mathrm{w}}, d_{\mathrm{w}}) \in \mathbb{R}^2$. $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \{0, 1, \cdots, P\}^2$ is a multi-index of absolute value $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2$, which identifies the components of the multivariate polynomial $\Psi_{\boldsymbol{\alpha}}$ that shall be of total degree less or equal to P, i.e. $|\boldsymbol{\alpha}| \leq P$; the set of selected multi-indices is noted $\mathcal{A}^P$. $\gamma_{\boldsymbol{\alpha}}$ is the corresponding coefficient to determine to build the surrogate.

Since the fire growth model is of finite variance, the front shape similarity measure defined in Eq. (19) can be considered as a spatially distributed random vector for which there exists a convergence polynomial expansion of the form (20) representing how $\boldsymbol{D}^{\epsilon, \pm}$ varies according to changes in wind speed ($\theta_1 = u_{\mathrm{w}}$) and direction ($\theta_2 = d_{\mathrm{w}}$). Details of the PC surrogate construction are given in the following for a 2-D random vector $\boldsymbol{\theta}$ but could be easily generalized.

3.2. Construction of the response surface

3.2.1. Choice and truncation of polynomial basis

The basis functions are orthonormal with respect to the joint PDF $\rho_{\boldsymbol{\theta}}(\boldsymbol{\zeta})$, namely

$$\langle \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}), \Psi_{\boldsymbol{\beta}}(\boldsymbol{\zeta}) \rangle = \int_Z \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}) \, \Psi_{\boldsymbol{\beta}}(\boldsymbol{\zeta}) \, \rho_{\boldsymbol{\theta}}(\boldsymbol{\zeta}) \, \mathrm{d}\boldsymbol{\zeta} = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}, \tag{21}$$

with $\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ the Kronecker delta-function and $Z \subseteq \mathbb{R}^2$ the space in which $\boldsymbol{\zeta}$ evolves. In practice, the orthonormal basis is built using the tensor product of 1-D polynomial functions such that the $\boldsymbol{\alpha}$th multivariate basis function can be expressed as $\Psi_{\boldsymbol{\alpha}} = \phi_{\alpha_1} \phi_{\alpha_2}$ with ϕ_{α_i} the 1-D polynomial function and its index $\alpha_i \in \{0, 1, \cdots, P\}$. The choice for the basis functions depends on the probability measure of the random variables. For the present sensitivity analysis, we consider uniform PDFs: $\{\Psi_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ correspond to the Legendre polynomials [63].

There are several ways of choosing the number of terms r_{pc} in the expansion of form (20). Using a standard truncation strategy, r_{pc} is constrained by the number of random variables (the size of θ is 2) and by the total polynomial degree P as $r_{\mathrm{pc}} = (2+P)!/(2! \, P!)$, meaning that all polynomials involving the two random variables of total degree less or equal to P are retained in the PC expansion. Thus, the set of selected multi-indices for the multi-variate polynomials $\mathcal{A}^P$ is defined as

$$\mathcal{A}^P = \{\boldsymbol{\alpha} \in \mathbb{N}^2 : |\boldsymbol{\alpha}| \leq P\} \subset \mathbb{N}^2. \tag{22}$$

3.2.2. Computation of expansion coefficients

We focus on non-intrusive approaches to numerically compute the coefficients $\{\gamma_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ using a training set made of N_{e} simulations. Since we consider a 2-D random vector $\boldsymbol{\theta}$, Galerkin spectral projection is a suitable approach [54]. It relies on the orthonormality property of the polynomial basis. The $\boldsymbol{\alpha}$th coefficient $\gamma_{\boldsymbol{\alpha}}$ is computed using a tensor-based Gauss-quadrature as

$$\gamma_{\boldsymbol{\alpha}} = \langle \boldsymbol{D}^{\epsilon, \pm}, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}) \rangle \cong \sum_{k=1}^{N_{\mathrm{e}}} \boldsymbol{D}^{\epsilon, \pm}\left(\mathbf{z}^{\mathrm{obs}}, \boldsymbol{H}(\mathbf{y}^{(k)})\right) \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}^{(k)}) w^{(k)}, \tag{23}$$

where $\mathbf{z}^{(k)} = \boldsymbol{H}(\mathbf{y}^{(k)})$ with $\mathbf{y}^{(k)} = (\mathbf{c}^{(k)}, \boldsymbol{\theta}^{(k)})$ is the kth integration of the forward model corresponding to the kth quadrature root $\boldsymbol{\theta}^{(k)}$ (in the physical space) with its weight $w^{(k)}$, $k = 1, ..., N_{\mathrm{e}}$, and with $(P+1)$ the number of quadrature roots required per uncertain input random variable. In the present study, building the surrogate model $\boldsymbol{D}_{\mathrm{pc}}^{\pm}$ will require $N_{\mathrm{e}} = (P+1)^2$ fire growth simulations for a given total polynomial degree P, which is a lower bound that insures the orthonormality of the polynomial approximation basis. The N_{e} realizations form the training set.

3.2.3. *Polynomial Chaos (PC) algorithm*

Our algorithm to build the PC surrogate proceeds as follows:

(1) choose the polynomial basis functions $\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta})\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ assuming uniform marginal PDFs $(\rho_{\theta_1}, \rho_{\theta_2})$ on the 2-D random vector $\boldsymbol{\theta} = (\theta_1, \theta_2) = (u_{\mathrm{w}}, d_{\mathrm{w}})$ with $\boldsymbol{\rho_\theta} = \rho_{\theta_1}\rho_{\theta_2}$;
(2) choose the total polynomial degree P according to the complexity of the physical processes (e.g. non-linearity in the forward model and in the front similarity measure $\boldsymbol{D}^{\epsilon,\pm}$);
(3) truncate the expansion to r_{pc} terms using standard truncation strategy;
(4) apply Gauss-Legendre quadrature in Eq. (23) to compute the coefficients $\{\gamma_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ using $N_{\mathrm{e}} = (P+1)^2$ realizations of $\boldsymbol{\theta}$ and their corresponding N_{e} fire growth model snapshots (the front shape similarity $\boldsymbol{D}^{\epsilon,\pm}$ in Eq. (19) is computed at a given time for each snapshot);
(5) form the surrogate $\boldsymbol{D}_{\mathrm{pc}}^{\pm}$ in Eq. (20) that can be evaluated for any new set of parameters $\boldsymbol{\theta}^* = (u_{\mathrm{w}}^*, d_{\mathrm{w}}^*)$.

3.3. **Estimation of statistical moments and Sobol' sensitivity indices**

Once the surrogate $\boldsymbol{D}_{\mathrm{pc}}^{\pm}$ is built, we can directly derive from the coefficients $\{\gamma_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ the main statistical moments of the front shape similarity measure such as the mean μ_{pc} and the STD σ_{pc}. The coefficients also contain information on how the discrepancies between the observation and the simulated fire front position change according to variability in each uncertain input parameter. We use Sobol' indices [60] to measure this sensitivity between uncertain inputs and the front shape similarity measure. For the ith variable in the random vector $\boldsymbol{\theta}$, the first-order Sobol' index $\mathbb{S}_{\mathrm{pc},i}$ specifically evaluates the ratio of the front dissimilarity variance that is explained by the variance of the standalone parameter θ_i. The mean, variance and Sobol' indices are obtained through

$$\mu_{\mathrm{pc}} = \mathbb{E}[\boldsymbol{D}^{\epsilon,\pm}] = \gamma_0, \quad \sigma_{\mathrm{pc}}^2 = \mathbb{V}[\boldsymbol{D}^{\epsilon,\pm}] = \sum_{\substack{\boldsymbol{\alpha} \in \mathcal{A}^P \\ \boldsymbol{\alpha} \neq 0}} (\gamma_{\boldsymbol{\alpha}})^2, \quad \mathbb{S}_{\mathrm{pc},i} = \frac{\mathbb{V}_i[\boldsymbol{D}^{\epsilon,\pm}]}{\mathbb{V}[\boldsymbol{D}^{\epsilon,\pm}]} = \frac{1}{(\sigma_{\mathrm{pc}})^2} \sum_{\substack{\boldsymbol{\alpha} \in \mathcal{A}_i^P \\ \boldsymbol{\alpha} \neq 0}} (\gamma_{\boldsymbol{\alpha}})^2, \qquad (24)$$

with $\mathbb{E}[.]$ the expectation operator, $\mathbb{V}[.]$ the variance operator and $\mathcal{A}_i^P$ the set of multi-indices selected in $\mathcal{A}^P$ such that the computation of $\mathbb{S}_{\mathrm{pc},i}$ only includes terms that depends on the input variable θ_i, namely $\mathcal{A}_i^P = \{\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{N}^2, |\boldsymbol{\alpha}| \leq P \mid \alpha_i > 0, \alpha_{k \neq i} = 0\}$.

4. RESULTS AND DISCUSSIONS

4.1. **Numerical configurations**

We present results from an OSSE test case in which synthetic observations are generated using specified (or "true") values of ignition location $\mathrm{x}_{\mathrm{ign}}^{\mathrm{t}}$ and near-surface wind parameters $(u_{\mathrm{w}}^{\mathrm{t}}, d_{\mathrm{w}}^{\mathrm{t}})$ and in which there is the possibility to have a "spot" fire. The 2-D computational domain Ω is 180 m $\times$ 180 m (with a step size $\Delta x = 1$ m); the time window is 100 s (with a constant time step $\Delta t = 0.1$ s). The true "main" fire is ignited at time 0 s at $\mathrm{x}_{\mathrm{ign}}^{\mathrm{t}} = (90 \text{ m}, 90 \text{ m})$ as a circular front with radius $r_{\mathrm{ign}} = 10$ m. If spotting is considered, the true secondary fire ignites at time 40 s as a circular front centered at $\mathrm{x}_{\mathrm{ign}}^{*,\mathrm{t}} = (40 \text{ m}, 40 \text{ m})$ and of radius $r_{\mathrm{ign}}^* = 4$ m. The true fire front positions are mapped in Fig. 3 at 10-s time intervals from initial times (0 s for the main fire and 40 s for the spot fire) until 100 s. We consider flat terrain and heterogeneous biomass moisture content, $M_{\mathrm{v}}^{(1)} = 5$ % (yellow

areas in Fig. 3 – left panel) and $M_{\mathrm{v}}^{(2)} = 20~\%$ (blue areas in Fig. 3 – left panel). The rest of the input parameters (Sec. 1.1) is assumed uniform. The biomass fuel is characterized by $\delta_{\mathrm{v}} = 0.2$ m (layer thickness), $\rho_{\mathrm{v}} = 1~\mathrm{kg\,m^{-3}}$ (layer mass density), $\Sigma_{\mathrm{v}} = 9{,}000~\mathrm{m^{-1}}$ (particle surface-to-volume ratio), $\Delta h_{\mathrm{v}} = 18.6 \times 10^{6}~\mathrm{J\,kg^{-1}}$ (heat of combustion). Near-surface wind speed and directions are constant over time, with $(u_{\mathrm{w}}^{\mathrm{t}}, d_{\mathrm{w}}^{\mathrm{t}}) = (3~\mathrm{m\,s^{-1}}, 180°)$.
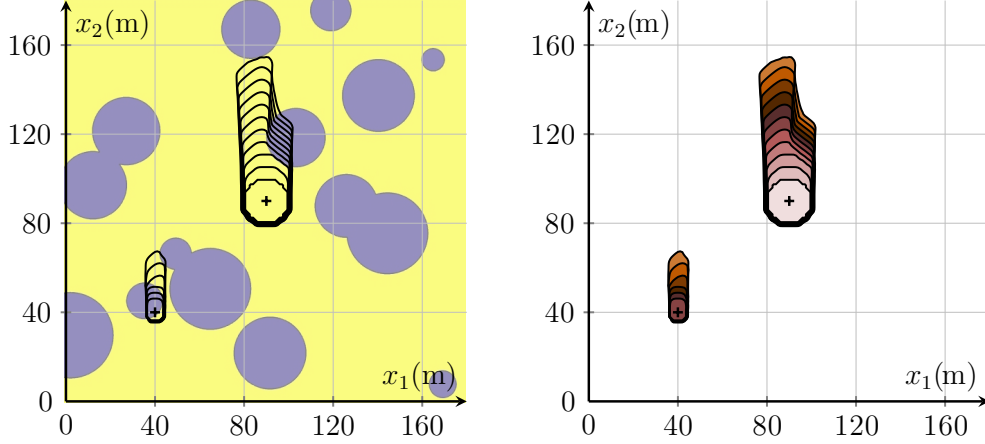
FIGURE 3. Left: Time-evolving location of the observed fire fronts from initial times t_0 until 100 s. The main fire ignites at $t_0 = 0$ s at $\mathrm{x}_{\mathrm{ign}}^{\mathrm{t}} = (90~\mathrm{m}, 90~\mathrm{m})$. The spot fire ignites at $t_0 = 40$ s at $\mathrm{x}_{\mathrm{ign}}^{*,\mathrm{t}} = (40~\mathrm{m}, 40~\mathrm{m})$. Black solid lines correspond to observations at 10-s time intervals; cross symbols indicate ignition position; and the colormap corresponds to the biomass fuel moisture content field: $M_{\mathrm{v}}^{(1)} = 5~\%$ in yellow areas and $M_{\mathrm{v}}^{(2)} = 20~\%$ in blue areas. Right: Corresponding superposed observation maps where the colormap corresponds to the arrival times (Eq. 3).

We consider here two test cases: the case with the standalone main fire (OMAIN) and the case with both main and spot fires (OSPOT). The two corresponding series of fire front positions are used as observations; confidence in the observations is high with the scalar shape weight equal to $\lambda_{\mathrm{sh}} = 0.04$ to evaluate the capability of the data assimilation algorithm to track observations in the "best-case" scenario (i.e. when observations are perfect). In practice, observations are not a time-continuous sequence of images, but a sequence of images acquired at distant times. In order to handle the data time-sampling, there are two possible strategies. We can either use the data only when they are available, or we can rely on time-interpolation. In the latter, as interpolating Heaviside functions is not recommended, we linearly interpolate shape observer corrections (first correction term in Eq. 14) and not the observations. Concerning the topological observer (second correction term in Eq. 14), as it is a Heaviside-type correction, we only apply it only when observations become available. In the present study, for both OMAIN and OSPOT test cases, the target fire front is observed every 10 s. Therefore, the data time-sampling is $\Delta T = 100\,\Delta t$ when compared to the model time-step Δt.

In the following, the main fire ignition location $\mathrm{x}_{\mathrm{ign}}$, the presence of a spot fire $\mathrm{x}_{\mathrm{ign}}^{*}$, the wind speed and direction $(u_{\mathrm{w}}, d_{\mathrm{w}})$ are considered as sources of uncertainties. We present different experiments with changing control variables to highlight the features of standalone LO-based state estimation (with or without activating the topological observer, step 2 of the algorithm presented in Sec. 2.2.4); standalone ROUKF-based parameter estimation (go directly from step 1 to step 4 in the algorithm presented in Sec. 2.2.4); and joint state-parameter estimation (combining LO and ROUKF through steps 1 to 5 in the aforementioned algorithm LO-ROUKF). The objective is to show that the new front shape similarity measure is applicable to any type of control variable and can retrieve the observed fire fronts considered here as the "targets".

4.2. Standalone state estimation

First, we consider uncertain fire ignition in terms of the main fire ignition location and the presence of a spot fire. The prior estimate of the front position (or "background") is obtained considering a single fire ignited at $x_{ign}^b = (94 \text{ m}, 94 \text{ m})$ (4 % error). We obtain a new estimate of the front position using the LO state estimator given in Eq. (11), with uniform gains λ_{sh} and λ_{top}. Figure 4 shows the result at times 10, 40, 60 and 100 s, when the topological gradient is not used ($\lambda_{top} = 0$ – left panel) and when it is used ($\lambda_{top} = 10$ – right panel). The main fire is accurately retrieved for both cases. However, the spot fire is only detected and tracked in the posterior estimate (or "analysis") obtained through the topological gradient. When $\lambda_{top} = 0$, the observer is limited in the sense that it can only correct the front shape that is already present in the simulation. On the contrary, when $\lambda_{top} = 10$, we show that the analysis accurately retrieves the spot fire rapidly after ignition at time 40 s.

If the prior wind $(u_w^b, d_w^b) = (4 \text{ m s}^{-1}, 225°)$ is subject to uncertainties but not corrected through parameter estimation, the LO state estimator is not able to track the observations since a wrong wind induces significant changes in the front shape as highlighted in Fig. 8 (left panel). This highlights the need for parameter estimation due to the parametric formulation of the fire growth model (Eq. 2) and its strong dependency on the quality of the inputs of the Rothermel-based ROS submodel.
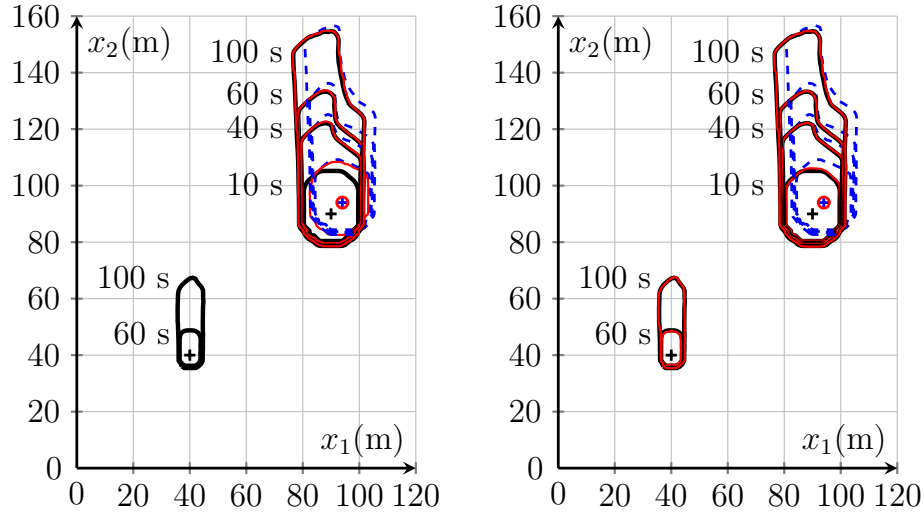

FIGURE 4. Mean front positions at times 10, 40, 60 and 100 s for the main fire and at times 60 and 100 s for the spot fire – OSPOT case. Comparison between observations (black solid lines), free run (blue dashed lines) and data-driven run (red solid lines). Symbols indicate ignition location. LO estimate (left) without topological gradient and (right) with topological gradient.

4.3. Sensitivity analysis and standalone parameter estimation

Second, we consider only the wind speed and direction as sources of uncertainties; the ignition position is assumed to be known. The prior estimate of the front position is obtained considering a single fire (OMAIN test case) ignited at the true position $x_{ign}^t = (90 \text{ m}, 90 \text{ m})$ but with prior wind values $(u_w, d_w)^b = (4 \text{ m s}^{-1}, 225°)$ that are significantly different from the true wind (25 % and 20 % errors, respectively).

Prior to parameter estimation, we carry out a sensitivity analysis of the discrepancy operator $(\boldsymbol{D}^{\epsilon,+}, \boldsymbol{D}^{\epsilon,-})^{\intercal}$ with respect to the uncertain wind magnitude and direction, $\boldsymbol{\theta} = (u_w, d_w)$. To limit the computational cost

of this process, we build a PC-surrogate considering a total polynomial order $P = 4$ and thus using a training set of $N_e = (P + 1)^2 = 25$ simulations. The 25 realizations of $\boldsymbol{\theta}$ are obtained considering uniform distributions centered on the prior control parameter vector $\boldsymbol{\theta}^b = (4 \text{ m s}^{-1}, 225°)$; the parameter realizations are plotted in Fig. 5 (right panel) and some of their corresponding simulated fronts are presented in Fig. 5 (left panel).

Figure 6 shows the statistics (mean and STD) of the discrepancy terms $\boldsymbol{D}^{\epsilon,+}$ (b–d) and $\boldsymbol{D}^{\epsilon,-}$ (c–e) that are directly derived from the coefficients $\{\gamma_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^P}$ ($P = 4$) of the surrogate $\boldsymbol{D}_{\text{pc}}^{\pm}$ (Eq. 20). It is found that the term $\boldsymbol{D}^{\epsilon,+}$ highlights the "hits" area (orange area where the observation and the simulated progress variable are equal) and penalizes the "false alarms" area (blue area where the simulated ensemble propagates and that is outside of the observed burnt area); the STD field is consistent with the patterns of the mean progress variable field. Moreover, the term $\boldsymbol{D}^{\epsilon,-}$ highlights the "misses" area (yellow area inside the observed burnt area but missed by the simulated ensemble members). The complete distinction between hits, false alarms and misses is not possible with the standard measure used to compute the observation-simulation differences (Eq. 4). This is one of the motivations to develop this new front shape similarity measure.

Figure 7a–b shows the 2-D map of the Sobol' indices $\mathbb{S}_{\text{pc},1}$ relating the variance of a) $\boldsymbol{D}^{\epsilon,+}$ and b) $\boldsymbol{D}^{\epsilon,-}$ with the variance of the wind speed $\theta_1 = u_w$; Fig. 7c–d shows the counterpart for the wind direction d_w, namely the 2-D map of the Sobol' indices $\mathbb{S}_{\text{pc},2}$ relating the variance of c) $\boldsymbol{D}^{\epsilon,+}$ and d) $\boldsymbol{D}^{\epsilon,-}$ with the variance of $\theta_2 = d_w$. Let us focus the discussion on the terms $\boldsymbol{D}^{\epsilon,+}$. They feature a high sensitivity to the wind direction d_w ($\mathbb{S}_{\text{pc},2} > 0.8$) on the flanks of the mean fire front. They also feature a significant sensitivity to the wind speed u_w ($\mathbb{S}_{\text{pc},1} > 0.5$) at the head of the fire, in the main direction of front propagation. These results are consistent with the fire growth model and the underlying ROS submodel (Eq. 2) since a change in the wind direction significantly modifies the main direction of front propagation, while a change in the wind speed mainly impacts the speed of the front propagation in the upwind direction. They highlight that only 25 fire growth simulations are sufficient to evaluate the contribution of each parameter to the behavior of the front shape similarity measure (via the PC surrogate) and therefore quantify the input-discrepancy sensitivity. The effects of the wind direction and wind speed on the front shape similarity measure are found to significantly differ, implying that no equifinality issue shall be encountered when performing parameter estimation.
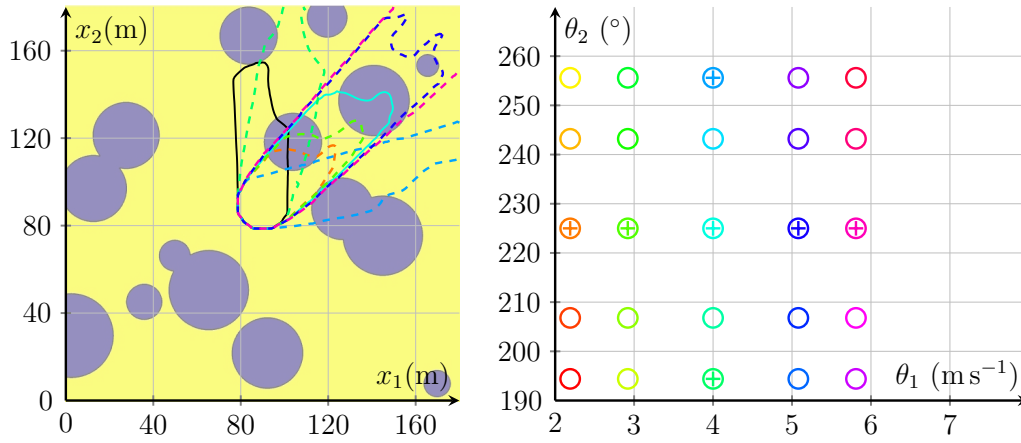

FIGURE 5. Input-output ensemble for the main fire – OMAIN case. Left: Comparison between observation (black solid line corresponding to $\boldsymbol{\theta}^t = (3 \text{ m s}^{-1}, 180°)$), prior mean (cyan solid line corresponding to $\boldsymbol{\theta}^b = (4 \text{ m s}^{-1}, 225°)$) and a subset of prior perturbations (colored dashed lines). The colormap corresponds to the biomass fuel moisture content field. Right: Ensemble of 25 $\boldsymbol{\theta}$-realizations. Each colored circle marker (right panel) represents a pair of wind magnitude $\theta_1 = u_w$ and wind direction $\theta_2 = d_w$; if there is a cross circle, the corresponding simulated front is represented on the left panel.
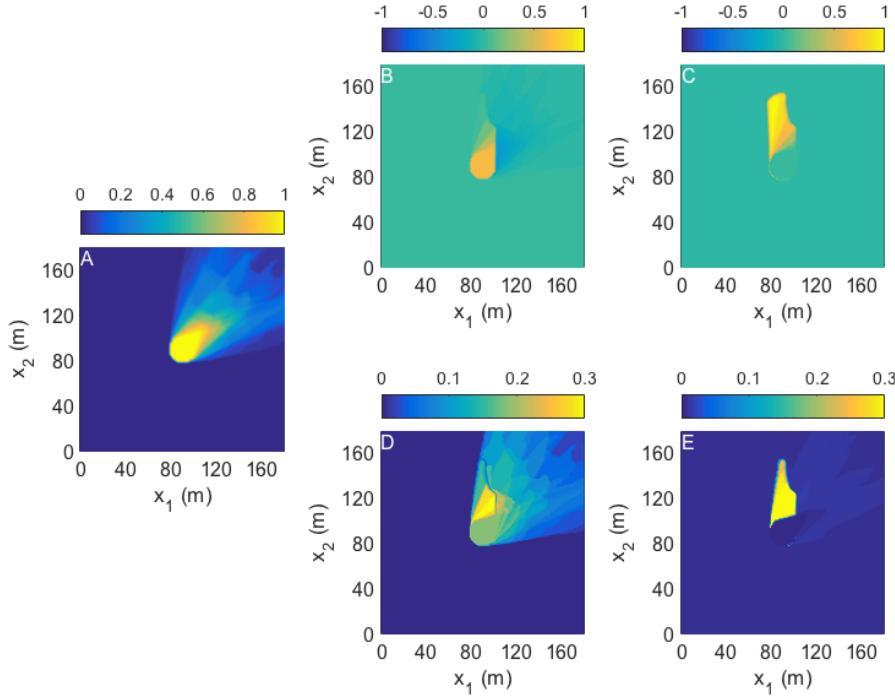
FIGURE 6. Uncertainty quantification – PC-based statistical fields (mean and STD – Eq. 24) of the quantity of interest derived from the PC surrogate built using the training set (see Fig. 5) – OMAIN case. a) Mean progress variable field $\mathbf{c} = \mathbf{c}(\mathbf{x}, t)$ (the progress variable takes values between 0 and 1). b) Mean discrepancy term $\boldsymbol{D}^{\epsilon,+}$ and d) associated STD. c) Mean discrepancy term $\boldsymbol{D}^{\epsilon,-}$ and e) associated STD.

Based on the sensitivity analysis results, we then carry out parameter estimation using the ROUKF. Figure 8 (right panel) presents results in terms of front position; Fig. 9 presents results in the parameter space. More accurate parameter values are found over the time period (the analysis converges to the true wind speed within 20 s and to the true wind direction within 40 s). These new values help to retrieve the main spreading direction and the correct ROS at the head of the fire. Still, there is a remaining gap between the analysis front estimates and the observations, especially on the flanks. There is no state estimation here and thereby not enough constraints from the wind parameters to perfectly track the observed fire fronts. We could extend the methodology to spatially-distributed parameter estimation to give the data assimilation algorithm more degrees of freedom to match the observations [65]. Note that in reality correcting parameters spatially does not ensure the retrieval of accurate parameter values. They shall then be viewed as effective values that incorporate multiple sources of uncertainties and that tend to be modified to reduce simulation-observation discrepancies, while they may only be partially responsible for them. This issue motivates the use of joint state-parameter estimation to obtain more realistic values of the parameters and thereby improve forecast performance.

4.4. **Joint state-parameter estimation**

We finally test the front shape similarity measure on the joint state-parameter estimation algorithm LO-ROUKF including a topological gradient for the state estimator [16] (Fig. 4 shows the influence of including a topological gradient in the LO formulation). We consider the ignition locations $(\mathbf{x}_{\mathrm{ign}}, \mathbf{x}_{\mathrm{ign}}^*)$ as well as the wind parameters $(u_{\mathrm{w}}, d_{\mathrm{w}})$ as sources of uncertainties. The prior values of these parameters are $\mathbf{x}_{\mathrm{ign}}^{\mathrm{b}} = (94 \text{ m}, 94 \text{ m})$ (same as in Sec. 4.2) and $(u_{\mathrm{w}}^{\mathrm{b}}, d_{\mathrm{w}}^{\mathrm{b}}) = (4 \text{ m s}^{-1}, 225°)$ (same as in Sec. 4.3). Their corresponding true values are
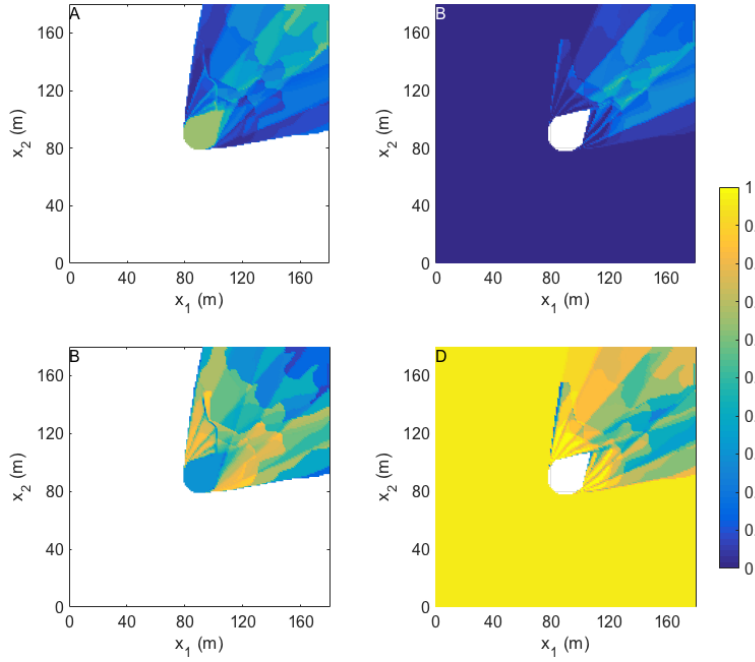
FIGURE 7. Sensitivity analysis – PC-based Sobol' indices (Eq. 24) derived from the PC surrogate built using the training set (see Fig. 5) – OMAIN case. They describe the contribution of each uncertain parameter on the variance of the discrepancy terms $\boldsymbol{D}^{\epsilon,+}$ and $\boldsymbol{D}^{\epsilon,-}$. Top row: Sobol' index associated with the wind speed u_w for a) $\boldsymbol{D}^{\epsilon,+}$ and b) $\boldsymbol{D}^{\epsilon,-}$. Bottom row: Sobol' index associated with the wind direction d_w for c) $\boldsymbol{D}^{\epsilon,+}$ and d) $\boldsymbol{D}^{\epsilon,-}$.
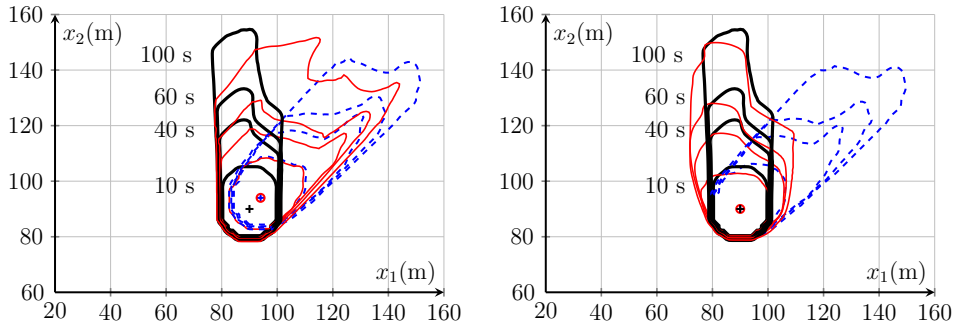
FIGURE 8. Mean front positions at times 10, 40, 60 and 100 s for the main fire – OMAIN case. Comparison between observations (black solid lines), free run (blue dashed lines) and data-driven run (red solid lines). Symbols indicate ignition location. Left: LO state estimation with wrong initial condition and wrong prior wind $(u_\mathrm{w}^\mathrm{b}, d_\mathrm{w}^\mathrm{b}) = (4~\mathrm{m\,s}^{-1}, 225°)$. Right: ROUKF parameter estimation with perfect initial condition and wrong prior wind $(u_\mathrm{w}^\mathrm{b}, d_\mathrm{w}^\mathrm{b}) = (4~\mathrm{m\,s}^{-1}, 225°)$.

$\mathrm{x}_\mathrm{ign}^\mathrm{t} = (90~\mathrm{m}, 90~\mathrm{m})$ and $(u_\mathrm{w}^\mathrm{t}, d_\mathrm{w}^\mathrm{t}) = (3~\mathrm{m\,s}^{-1}, 180°)$. Plus a spot fire starts at 40 s at $\mathrm{x}_\mathrm{ign}^{*,\mathrm{t}} = (40~\mathrm{m}, 40~\mathrm{m})$. Due to these combined uncertainties, the prior front estimate significantly deviates from the truth as shown in Fig. 10. The joint state-parameter estimator is found to accurately track the observed fronts (including the
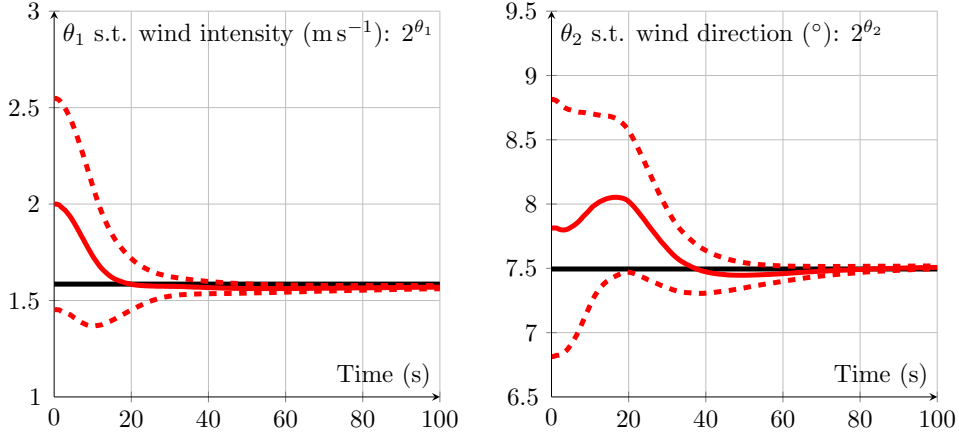
FIGURE 9. Time series of wind magnitude (left panel) and direction (right panel) estimates – OMAIN case; standalone ROUKF parameter estimation. Red solid lines correspond to the mean values equal to $(u_w^b = 4 \text{ m s}^{-1}, d_w^b = 225°)$ at initial time; red dashed lines correspond to the STD on both sides of the mean values; and black solid lines correspond to the true values $(u_w^t = 3 \text{ m s}^{-1}, d_w^t = 180°)$.

secondary fire) with a good balance between state estimation ($\lambda_{sh} = 0.04$, $\lambda_{top} = 10$) and parameter estimation: the true wind magnitude and direction are accurately identified as shown in Fig. 11. Note that the convergence to the true wind magnitude is slightly longer than in the standalone parameter estimation approach, while the convergence to the true wind direction takes less time (see Fig. 9). Moreover, we emphasize that here we focus on one iteration of the estimation procedure to illustrate the already very good performance of our method. Further studies could include multiple iterations of the procedure to help the parameter identification convergence in such a non-linear parameter dependency. Such strategies are known as iterative extended or unscented Kalman approaches [5, 30].

5. Conclusions and Perspectives

We presented a front shape similarity measure adapted to an eikonal equation solved using an Eulerian level-set front-tracking model. This measure derived from the Chan-Vese contour fitting functional in image segmentation theory is well suited to handle position errors that go beyond the problem of amplitude errors addressed by classical data assimilation framework. The present work shows that this formalism is valuable for wildfire applications (1) where uncertainties are both due to limited knowledge on the active fire location (location, spotting fires) and on the physical parameters required as inputs to the fire problem (e.g. near-surface wind speed and direction), (2) where observations typically correspond to mid-infrared images from which the position, shape and topology of the active burning areas shall be extracted to be compared to the fronts simulated by fire growth models. We showed on synthetic cases, the benefits of using a joint state-parameter estimation approach to address shape and topological errors: the Luenberger-type state estimation was useful to track the (potentially multiple) fire front shapes; and the Kalman-type reduced-order parameter estimation was useful to reduce model bias. We also provided a polynomial chaos framework to balance the weight of each approach through input-measure sensitivity analysis. Comparing Sobol' indices derived from the polynomial chaos surrogate is a robust method to classify by order of importance the parameters to estimate.

Future work includes the extensive evaluation of the front shape similarity measure in the context of real-world wildfire hazards. To address these problems, the framework could include more sources of uncertainties in
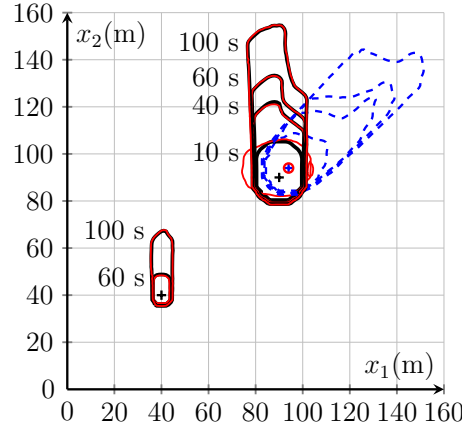
FIGURE 10. Mean front positions at times 10, 40, 60 and 100 s for the main fire and at times 60 and 100 s for the spot fire – OSPOT case. Comparison between observations (black solid lines), free run (blue dashed lines) and data-driven run (red solid lines). Symbols indicate ignition location. Joint LO/ROUKF state-parameter estimation with wrong initial condition ($x_{ign}^b = (94 \text{ m}, 94 \text{ m})$, absence of a spot fire) and wrong prior wind $(u_w^b, d_w^b) = (4 \text{ m s}^{-1}, 225°)$.
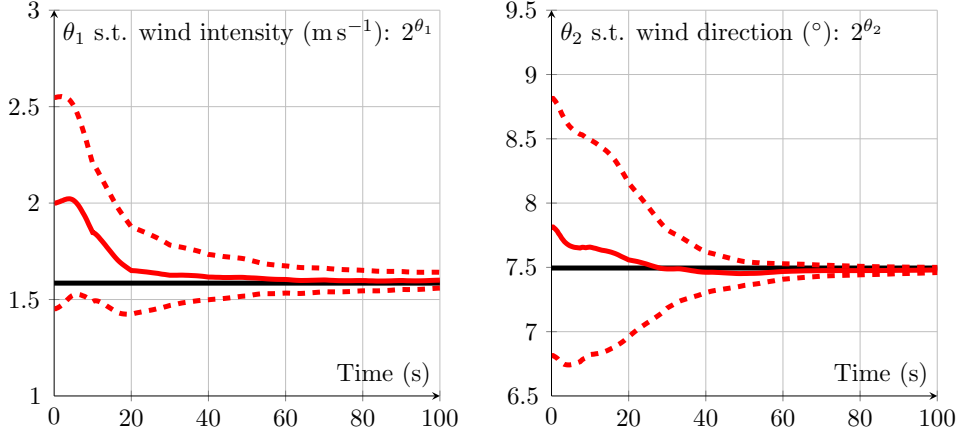

FIGURE 11. Same caption as in Fig. 9 but for OSPOT case and joint LO-ROUKF state-parameter estimation.

the control parameter vector (e.g. biomass fuel moisture content) and could be able to operate on Lagrangian-type fire growth models (e.g. FARSITE). We also need to evaluate, in practice, a spatially distributed gain in both state and parameter estimation, which may be hampered by partially observed fire fronts – as incomplete wildfire images may be the only source of information available due to the opacity of the fire-induced thermal plume and/or due to a limited monitoring. We may also require to define error covariance statistics that are well suited for the discrepancy operator, and to add a texture criterion – and not only shape and topological criteria – to provide more information on the active burning areas in the data assimilation process.

6. Acknowledgment

References

[1] D. Adalsteinsson and J. A. Sethian. The fast construction of extension velocities in level set methods. *J Comput Phys*, 148(1):2–22, 1999.

[2] J.-h. An and Y. Chen. Region based image segmentation using a modified mumford-shah algorithm. *Scale Space and Variational Methods in Computer Vision*, pages 733–742, 2007.

[3] P. L. Andrews, M. G. Cruz, and R. C. Rothermel. Examination of the wind speed limit function in the rothermel surface fire spread model. *International Journal of Wildland Fire*, 22:959–969, 2013.

[4] P. Arbogast, O. Pannekoucke, L. Raynaud, R. Lalanne, and E. Mémin. Object-oriented processing of crm precipitation forecasts by stochastic filtering. *Quarterly Journal of the Royal Meteorological Society*, 142(700):2827–2838, 2016.

[5] A. Armiento, P. Moireau, D. Martin, N. Lepejova, M. Doumic, and H. Rezaei. The mechanism of monomer transfer between two structurally distinct PrP oligomers. *PLoS ONE*, 12(7):e0180538, 2017.

[6] T. Artes, A. Cencerrado, A. Cortes, T. Margalef, D. Rodriguez-Aseretto, T. Petroliagkis, and J. San-Miguel-Ayanz. Towards a dynamic data driven wildfire behavior prediction system at european level. *Procedia Computer Science*, 29:1216 – 1226, 2014.

[7] J. Beezley and J. Mandel. Morphing ensemble Kalman filters. *Tellus*, 60A:131–140, 2008.

[8] I. Bloch. Fuzzy sets for image processing and understanding. *Fuzzy Sets and Systems*, 281(C):280–291, Dec. 2015.

[9] A. Bova, W. Mell, and C. Hoffman. A comparison of level set and marker methods for the simulation of wildland fire front propagation. *International Journal of Wildland Fire*, 25:229–241, 2015.

[10] T. F. Chan and L. A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*, VLSM '01, pages 161–, Washington, DC, USA, 2001. IEEE Computer Society.

[11] D. Chapelle, M. Fragu, V. Mallet, and P. Moireau. Fundamental principles of data assimilation underlying the Verdandi library: applications to biophysical model personalization within euHeart. *Medical and Biological Engineering and Computing*, 51:1221–1233, 2013.

[12] S. Chen, G. Charpiat, and R. J. Radke. Converting Level Set Gradients to Shape Gradients. In *Computer Vision – ECCV 2010*, pages 715–728. Springer, Berlin, Heidelberg, Berlin, Heidelberg, Sept. 2010.

[13] S. Clandillon and H. Yesou. GMES Emergency Response - Fire mapping experience within EO-based Rapid Mapping. In *EGU General Assembly 2011*, volume 13, 2011.

[14] A. Collin. *Analyse asymptotique en électrophysiologie cardiaque : applications à la modélisation et à l'assimilation de données*. PhD thesis, Université Pierre et Marie Curie - Paris VI, France, 2014.

[15] A. Collin, D. Chapelle, and P. Moireau. A Luenberger observer for reaction–diffusion models with front position data. *J Comput Phys*, 300(C):288–307, Nov. 2015.

[16] A. Collin, D. Chapelle, and P. Moireau. Sequential State Estimation for Electrophysiology Models with Front Level-Set Data Using Topological Gradient Derivations. In *Proceedings of the 8th International Conference FIMH*, volume LNCS 9126, pages 402–411, 2015.

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

[18] T. J. Duff, D. M. Chong, and K. G. Tolhurst. Quantifying spatio-temporal differences between fire shapes: Estimating fire travel paths for the improvement of dynamic spread models. *Environmental Modelling and Software*, 46:33 – 43, 2013.

[19] B. Engquist, A.-K. Tornberg, and R. Tsai. Discretization of Dirac delta functions in level set methods. *J Comput Phys*, 207(1):28–51, 2005.

[20] A. C. Fernandez-Pello. Wildland fire spot ignition by sparks and firebrands. *Fire Safety Journal*, pages –, 2017.

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

[22] N. Feyeux, M. Nodet, and A. Vidard. Optimal Transport for Data Assimilation. working paper or preprint, July 2016.

[23] J. B. Filippi, V. Mallet, and B. Nader. Evaluation of forest fire models on a large observation database. *Natural Hazards and Earth System Sciences*, 14(11):3077–3091, 2014.

---

[1]http://verdandi.sourceforge.net

[24] J.-B. Filippi, F. Morandini, J.-H. Balbi, and D. Hill. Discrete event front-tracking simulation of a physical fire-spread model. *Simulation, Transactions of the Society for Modeling and Simulation International*, 87:555–580, 2011.

[25] M. Finney. FARSITE: Fire Area Simulator - Model Development and Evaluation. Technical Report February, US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998.

[26] B. Fisher. The product of distributions. *The Quarterly Journal of Mathematics*, 1971.

[27] J. Glasa and L. Halada. On elliptical model for forest fire spread modeling and simulation. *Mathematics and Computers in Simulation*, 78(1):76–88, 2008.

[28] M. Gollner, A. Trouvé, I. Altintas, J. Block, R. De Callafon, C. Clements, A. Cortes, E. Ellicott, J.-B. Filippi, M. Finney, K. Ide, M.-A. Jenkins, D. Jimenez, C. Lautenberger, J. Mandel, M. Rochoux, and A. Simeoni. Towards data-driven operational wildfire spread modeling - report of the nsf-funded wifire workshop. Technical report, University of Maryland, 2015.

[29] L. He, C. Y. Kao, and S. Osher. Incorporating topological derivatives into shape derivatives based level set methods. *J Comput Phys*, 225(1):891–909, 2007.

[30] A. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[31] E. Jimenez, M. Hussaini, and S. Goodrick. Quantifying parametric uncertainty in the Rothermel model. *International Journal of Wildland Fire*, 17:638–649, 2008.

[32] L. Le Gratiet, S. Marelli, and B. Sudret. Metamodel-Based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes. In *Handbook of Uncertainty Quantification*, pages 1–37. Springer International Publishing, 2017.

[33] J. Li and D. Xiu. A generalized polynomial chaos based ensemble Kalman filter. *Journal of Computational Physics*, 228:5454–5469, 2009.

[34] L. Li, F.-X. Le Dimet, J. Ma, and A. Vidard. A level-set based image assimilation method: applications for predicting the movement of oil spills. Submitted to IEEE Transactions on Geoscience and Remote Sensing, Nov. 2016.

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

[36] M. Martins, S. Ferreira Jr, and M. Vilela. Multiscale models for the growth of avascular tumors. *Physics of Life Reviews*, 4(2):128 – 156, 2007.

[37] P. Moireau and D. Chapelle. Reduced-order unscented Kalman filtering with application to parameter identification in large-dimensional systems. *Control, Optimisation and Calculus of Variations*, 17(2):380–405, Mar. 2011.

[38] P. Moireau, D. Chapelle, and P. L. Tallec. Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 197(6–8):659 – 677, 2008.

[39] H. Moradkhani, S. Sorooshian, H. Gupta, and P. Houser. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28:135–147, 2005.

[40] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42(5):577–685, 1989.

[41] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Applied Mathematical Sciences - Springer, 2003.

[42] N. Papadakis and E. Mémin. A variational technique for time consistent tracking of curves and motion. *Journal of Mathematical Imaging and Vision*, 31(1):81–103, 2008.

[43] R. Paugam, M. Wooster, and G. Roberts. Use of handheld thermal imager data for airborne mapping of fire radiative power and energy and flame front rate of spread. *Geoscience and Remote Sensing*, 51:3385–3399, 2013.

[44] H. Pitsch. A consistent level set formulation for large-eddy simulation of premixed turbulent combustion. *Combustion and Flame*, 143(4):587 – 598, 2005.

[45] G. Poëtte, B. Després, and D. Lucor. Treatment of uncertain material interfaces in compressible flows. *Computer Methods in Applied Mechanics and Engineering*, 200(1):284 – 308, 2011.

[46] S. Ravela. Amplitude-position formulation of data assimilation. In *Proceedings of the 6th International Conference on Computational Science - Volume Part III*, ICCS'06, pages 497–505, Berlin, Heidelberg, 2006. Springer-Verlag.

[47] R. Rehm and R. McDermott. Fire front propagation using the level set method. Technical Report March, National Institute of Standards and Technology, 2009.

[48] R. Reichle. Data assimilation methods in the Earth sciences. *Advances in Water Resources*, 31:1411–1418, 2008.

[49] 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.

[50] P. Riggan and G. Robert. Airborne remote sensing of wildland fires. In A. Bytnerowicz, M. Arbaugh, C. Andersen, and A. Riebau, editors, *Wildland Fires and Air Pollution. Developments in Environmental Science 8*, pages 139–168. Elsevier, Amsterdam, The Netherlands, 2009.

[51] M. C. Rochoux. *Vers une meilleure prévision de la propagation d'incendies de forêt : évaluation de modèles et assimilation de données*. PhD thesis, Ecole Centrale Paris, France, 2014.

[52] 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*, 34:2641–2647, 2013.

[53] M. C. Rochoux, C. Emery, S. Ricci, B. Cuenot, and A. Trouvé. 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 Science*, 15(8):1721–1739, 2015.

[54] M. C. Rochoux, S. Ricci, D. Lucor, B. Cuenot, and A. Trouvé. Towards predictive data-driven simulations of wildfire spread – part i: Reduced-cost ensemble kalman filter based on a polynomial chaos surrogate model for parameter estimation. *Natural Hazards and Earth System Science*, 14(11):2951–2973, 2014.

[55] B. Rosic, A. Kucerov, J. Sykora, O. Pajonk, A. Litvinenko, and H. Matthies. Parameter identification in a probabilistic setting. *Engineering Structures*, 50:179–196, 2013.

[56] R. Rothermel. A mathematical model for predicting fire spread in wildland fuels. Technical report, U.S. Department of Agriculture, Forest Service, Intermountain forest and range experiment station, Ogden, Utah, USA, 1972.

[57] J. J. Ruiz, M. Pulido, and T. Miyoshi. Estimating model parameters with ensemble-based data assimilation: A review. *Journal of the Meteorological Society of Japan. Ser. II*, 91(2):79–99, 2013.

[58] J. Sethian. *Level set methods and fast marching methods*. Cambridge University Press, 1999.

[59] J. A. Sethian. Theory, algorithms, and applications of level set methods for propagating interfaces. *ANU*, 5(1):309–395, 1996.

[60] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliab Eng Syst Safety*, 93(7):964–979, 2008.

[61] A. Sullivan. Wildland surface fire spread modeling, 1990-2007. 2: Empirical and quasi-empirical models. *International Journal of Wildland Fire*, 18:369–386, 2009.

[62] M. Trani, R. Arts, and O. Leeuwenburgh. Seismic history matching of fluid fronts using the ensemble kalman filter. *SPE Journal*, 18(1):159–171, 2012.

[63] D. Xiu. *Numerical methods for stochastic computations. A spectral method approach*. Princeton University Press, 2010.

[64] R. Yildizoglu, J.-F. Aujol, and N. Papadakis. Active Contours without Level Sets. In *ICIP 2012 - IEEE International Conference on Image Processing*, pages 2549–2552, Orlando, Florida, United States, Sept. 2012. IEEE.

[65] C. Zhang, M. C. Rochoux, W. Tang, M. Gollner, J. B. Filippi, and A. Trouvé. Evaluation of a data-driven wildland fire spread forecast model with spatially-distributed parameter estimation in simulations of the fireflux i field-scale experiment. *Fire Safety Journal (in press)*.