Towards predictive simulation of wildfire spread using a reduced-cost Ensemble Kalman Filter based on Polynomial Chaos approximation.

By M.C. Rochoux †‖‡, S. Ricci ‡††, D. Lucor ††, B. Cuenot †, A. Trouvé ¶ AND J.-M. Bart †¶

The sequential correction of a fire spread model parameters is performed via the assimilation of airborne-like fire front observations, in order to improve the simulation and forecast of the fire propagation. An Ensemble Kalman Filter (EnKF) is applied to reduce the uncertainties in the atmospheric and vegetation parameters for the Rate Of Spread (ROS) model. The non-linear relation between the parameters and the fire front position induced by the non-linearities of the fire spread is described stochastically over the EnKF members. In order to reduce the computational cost of the data assimilation algorithm, a surrogate model based on a Polynomial Chaos (PC) approximation is used in place of the forward propagation model. The merits of using the EnKF algorithm based on the PC approximation are highlighted on experiments using synthetical and real measurements.

1. Introduction

Because wildfire spread involves both multi-physics and multi-scales, our ability to predict the behavior of wildfires at large regional scales (i.e., at scales ranging from a few tens of meters up to several kilometers) remains limited. Current wildfire spread simulators mainly rely on front-tracking techniques where the propagation speed of wildfires, also called the Rate Of Spread (ROS), is the main physical quantity of interest. The ROS is treated as a simplified function of vegetation, topographical and meteorological properties, based on Rothermel's model (Rothermel 1972). A first limitation in this approach is that the feedback of the fire on the atmosphere and on the fuel is not accounted for. A second limitation is that the actual input variables that determine the ROS are only known with limited accuracy. For the wildfire spread simulation to be predictive and compatible with operational applications, the uncertainty in the ROS semi-empirical model should be quantified and reduced.

In this study, a Data Assimilation (DA) methodology is considered in order to minimize the uncertainty on the parameters for the ROS model using fire front observations, and thus over-come some of the current limitations of regional-scale wildfire modeling. DA provides indeed an attractive framework for integrating fire sensor observations into computational models, for accounting for the effects of both observation and modeling errors and thereby for providing optimal estimates of poorly known parameters and improved predictions of fire spread dynamics. Previous works (Rochoux *et al.* 2012) describe a cost-effective DA prototype using an Extended Kalman Filter (EKF) algorithm that relies on

† CERFACS, 42 avenue G. Coriolis, Toulouse, France.
‖ Ecole Centrale Paris, Grande voie des vignes, Châtenay-Malabry, France.
‡ CNRS, URA-1875, 42 avenue G. Coriolis, Toulouse, France.
†† Institut d'Alembert, UPMC-CNRS, UMR-7190, 4 place Jussieu, Paris, France.
¶ Dept. of Fire Protection Engineering, University of Maryland, College Park, USA.

the assumption that the relation between a perturbation in the model parameters and the resulting change in the fire front position (the observation operator) is linear. The EKF linearises, sometimes unrealistically, this operator based on local derivatives, which are often difficult to compute reliably. While this EKF-based prototype showed very good results for a controlled grassland fire experiment, for realistic cases of regional-scale wildfire spread where the wind direction and magnitude may vary and the vegetation properties may be strongly heterogenous, the local linearity hypothesis may not be satisfying and the resulting analysis may not be optimal (Ros & Borga 1997). Here, to better account for non-linearities in the observation operator, an Ensemble Kalman Filter (EnKF) was implemented (Evensen 1994). This ensemble-based DA approach was originally developed for dynamic state estimation and has already been used in the wildfire modeling community for temperature state correction (Mandel *et al.* 2008). It was recently extended to sequential parameter estimation in the field of hydraulics (e.g., Durand *et al.* 2008, Moradkhani *et al.* 2005). In the present study, the EnKF is used to stochastically characterize the non-linear relationship between the fire front position and the ROS model parameters, thus allowing for a temporal correction of the vegetation and atmospheric characteristics as the fire propagates.

The obvious source of numerical errors of EnKF stems from sampling. In fact, Li & Xiu (2008) showed that more frequent data assimilation by EnKF does not intuitively lead to a more accurate estimate of the true control variables due to the accumulation of sampling errors. However, the size of the sample used in the EnKF algorithm may prove computationally burdensome within an operational environment. Efforts have therefore been devoted to designing more efficient EnKF schemes by reducing the sampling errors (e.g., Szunyogh *et al.* 2008, Li & Xiu 2008). In the case of a real fire propagation, the convergence of the EnKF is established for a number of members that is not compatible with operational application; additionally, the required size of the sample significantly increases with the complexity of the physics (multi-parameter estimation), the non-linearity in the model (complex physics), since the sampling error increases with the non-linearity in the observation operator, thus emphasizing the need for a reduced-cost EnKF. For this purpose, and following work from Li & Xiu 2009, a numerical strategy based on Polynomial Chaos (PC) approximation is proposed; the polynomial surrogate model is used to generate a large number of members for the EnKF at a low computational cost.

In this paper, we present an ensemble-based DA approach and computationally efficient algorithm to improve fire front propagation modeling, reducing the uncertainty in the model parameters. The stochastic approach is used to estimate the covariances matrices in the EnKF that describe the relation between the vegetation and atmospheric parameters and the fire front position. The paper is organized as follows. The fire propagation model (also called the forward model) is presented in Section 2. Section 3 explains the hybrid PC-EnKF algorithm developed for the wildfire application. The performance of the resulting wildfire simulation capability is demonstrated in Section 4 using synthetically-generated and real observations.

2. The fire spread model (the forward model)

The DA algorithm primarily relies on a numerical tool for the fire spread simulation, denoted by $M_{[t_i,t_i+T]}(\mathbf{x})$, where $\mathbf{x}$ designates the set of n control parameters and $[t_i, t_i+T]$ is the time interval for the model integration (also denoted as cycle c_i). This model simulates the wildfire spread as a thin flame zone that self-propagates normal to itself towards unburnt vegetation, using the following two components:

- **a sub-model for the ROS Γ based on Rothermel's model**: $\Gamma\,[m/s]$ is written as a function of the fuel depth (e.g., the vegetation thickness) $\delta\,[m]$ such as:

$$\Gamma(x,y,t) = P\Big(M_f, \Sigma, \mathbf{u}_w(x,y,t), ...\Big)\,\delta(x,y), \tag{2.1}$$

where the coefficient $P\,[s^{-1}]$ depends on the local vegetation characteristics (especially on the fuel moisture content M_f and the fuel particle surface-to-volume ratio $\Sigma\,[m^{-1}]$) and on the meteorological properties (i.e., the wind velocity at mid-flame height $\mathbf{u}_w\,[m/s]$).

- **a level-set-based solver for the front propagation**: A progress variable c ranging from 0 in the unburnt vegetation to 1 in the burnt vegetation is introduced as a flame marker; the flame front is identified as the iso-contour $c_f = 0.5$. The two-dimensional variable c is calculated as a solution of the following propagation equation using a total variation diminishing scheme (Rehm & McDermott 2009; Rochoux *et al.* 2012):

$$\frac{\partial c}{\partial t} = \Gamma\ |\nabla c|\,, \tag{2.2}$$

with Γ specified, locally, along the normal direction to the front $c_f = 0.5$ using Eq. (2.1).

3. The ensemble-based Data Assimilation algorithm using Polynomial Chaos

3.1. *Stochastic estimate of the observation operator*

Let assume the EnKF algorithm is applied for one assimilation cycle c_i between times t_i and $t_i{+}T$. Starting from an *a priori* ensemble of control parameters $\mathbf{x}_{c_i}^{-,k}$ ($k = 1, \cdots, N_{ens}$ where N_{ens} is the ensemble size), the fire spread model produces an ensemble of predictions of the time-evolving fire front location (predicted measurements) designated as $\mathbf{y}_{c_i}^{-,k} = H(\mathbf{x}_{c_i}^{-,k})$. More precisely, this observation operator H is the composition of the fire spread model $M_{[t_i,t_i+T]}$ (that provides the spatio-temporal variations of the progress variable c over the time window $[t_i, t_i + T]$) with the selection operator $S(c_f)$ (that provides the (x_f, y_f)-coordinates of the p points representing the discretized $c_f = 0.5$ contour at the observation time $t_i + T$).

In order to provide an estimate of the model deviation from the observations denoted by $\mathbf{y}_{c_i}^{o}$, we compute the distance between the prediction $\mathbf{y}_{c_i}^{-,k}$ and the observation vector $\mathbf{y}_{c_i}^{o}$ on top of which a noise ξ^k is added to avoid the ensemble collapse (Burgers *et al.* 1998). This distance called the innovation vector and denoted by $\mathbf{d}_{c_i}^{k}$ reads

$$\mathbf{d}_{c_i}^{k} = \mathbf{y}_{c_i}^{o} + \xi^k - \mathbf{y}_{c_i}^{-,k} \tag{3.1}$$

3.2. *The Ensemble Kalman Filter update*

Based on the Bayesian theory, the EnKF algorithm assumes that the control parameters $\mathbf{x}_{c_i}^{-,k}$ and the observations $\mathbf{y}_{c_i}^{o}$ are random variables defined by Gaussian probability density functions (PDF) with a zero mean value and an error covariance model. The prior error covariance matrix is denoted by $\mathbf{P}_{c_i}^{-}$ with variance σ_x^2, and the observation error covariance matrix is denoted by $\mathbf{C}_{\mathbf{y}^o\mathbf{y}^o}$ with variance σ_o^2. The DA procedure offers to combine these PDF in order to estimate the set of parameters that minimizes the error variances of the control parameters, knowing the innovation vector $\mathbf{d}_{c_i}^{k}$. It provides, for cycle c_i, an ensemble $\mathbf{x}_{c_i}^{+,k}$ of optimal estimates of the true vector $\mathbf{x}_{c_i}^{t}$ satisfying

$$\mathbf{x}_{c_i}^{+,k} = \mathbf{x}_{c_i}^{-,k} + \mathbf{C}_{\mathbf{xy}}(\mathbf{C}_{\mathbf{yy}} + \mathbf{C}_{\mathbf{y}^o\mathbf{y}^o})^{-1}\mathbf{d}_{c_i}^{k}, \tag{3.2}$$

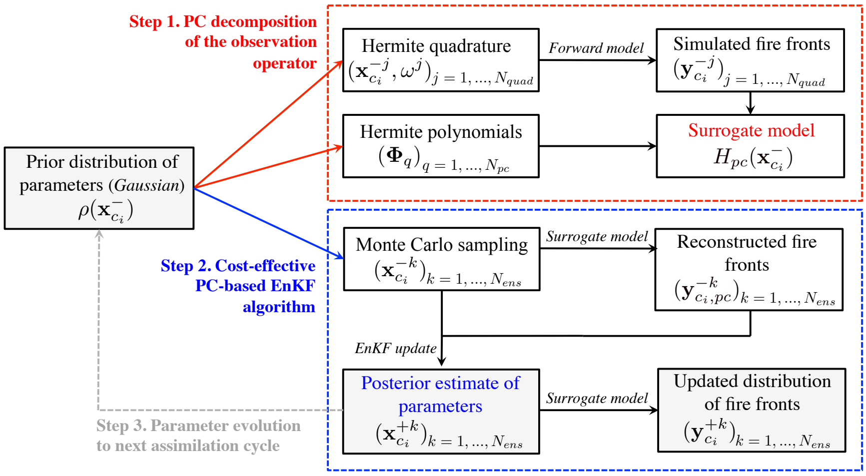


FIGURE 1. Schematic of the PC-based EnKF algorithm at cycle c_i.

where $\mathbf{C_{xy}}$ and $\mathbf{C_{yy}}$ are the covariance matrix of the model parameters with the predicted measurements of fire front positions, and the covariance matrix of the predicted measurements, respectively.

Using the standard notation of the Kalman Filter, the gain matrix $\mathbf{K}_{c_i}$ reads

$$\mathbf{K}_{c_i} = \mathbf{C_{xy}}(\mathbf{C_{yy}} + \mathbf{C_{y^o y^o}})^{-1}, \tag{3.3}$$

where $\mathbf{C_{xy}} = \mathbf{P}_{c_i}^{-}\mathbf{H}^T$ with $\mathbf{H}^T$ the linearised observation operator and $\mathbf{C_{yy}} = \mathbf{H}\mathbf{P}_{c_i}^{-}\mathbf{H}^T$. The formulation (3.3) avoids the estimation of $\mathbf{P}_{c_i}^{-}$ and $\mathbf{H}$, difficult to compute reliably.

Using the ensemble of model input parameters $\mathbf{x}_{c_i}^{-,k}$ and output results $\mathbf{y}_{c_i}^{-,k}$, the relationship between the fire front positions and the model parameters can be stochastically characterized. This method allows to better take into account the non-linearity in the observation operator H than a local estimation $\mathbf{H}$ achieved with a finite difference scheme as in EKF. Such approach provides an ensemble of posterior estimates $\mathbf{x}_{c_i}^{+,k}$ easily used to simulate an ensemble $\mathbf{y}_{c_i}^{+,k}$ of retrospective posterior estimates of the fire front positions over $[t_i, t_i + T]$ as well as an ensemble of forecasts of the fire spread beyond $t_i + T$. In order to allow for a temporal correction of the model parameters between the cycles, the EnKF algorithm is sequentially applied. Along the cycles, the parameter evolution is artificially set up (Moradkhani *et al.* 2005):

$$\mathbf{x}_{c_i+1}^{-,k} = \overline{\mathbf{x}}_{c_i}^{+} + \epsilon_{c_i}^{k} \tag{3.4}$$

where $\overline{\mathbf{x}}_{c_i}^{+}$ is the mean of the posterior estimates $\mathbf{x}_{c_i}^{+,k}$ and $\epsilon_{c_i}^{k}$ is a randomly-generated mean zero normal variable with standard deviation (STD) σ_x.

3.3. *The Polynomial Chaos-based Ensemble Kalman filter update*

In the classical EnKF algorithm, a Monte Carlo sampling is used to generate the prior members $\mathbf{x}_{c_i}^{-,k}$. While this provides an accurate access to the full statistics of the modeling uncertainties provided a large enough sample, it involves a large number of forward model integrations. In order to maintain the EnKF computational cost compatible with the fire application, a numerical strategy based on a PC expansion is outlined in Fig. 1.

The key idea in the PC expansion technique is to build a polynomial approximation of the forward model response (i.e., of the observation operator H) to a prior estimate of the control parameters $\mathbf{x}_{c_i}^-$. Therefore, the simulated positions of the fire front $\mathbf{y}_{c_i}^- = H(\mathbf{x}_{c_i}^-)$ are projected onto a stochastic space spanned by a set of orthogonal polynomials $\Phi_q(\mathbf{x}_{c_i}^-)$ ($q = 1, ..., N_{pc}$) that are suitably selected in accordance with the PDF of $\mathbf{x}_{c_i}^-$. As these n control parameters $\mathbf{x}_{c_i}^-$ are assumed to follow a Gaussian distribution in EnKF, the surrogate model of the observation operator (denoted by H_{pc}) is built upon the Hermite polynomials basis (Ghanem & Spanos 2003). H_{pc} can be formulated as follows:

$$\mathbf{y}_{c_i}^- \cong \mathbf{y}_{c_i,pc}^- = H_{pc}(\mathbf{x}_{c_i}^-) = \sum_{q=1}^{N_{pc}} \hat{\mathbf{y}}_q \Phi_q(\mathbf{x}_{c_i}^-), \tag{3.5}$$

where the unknowns are the N_{pc} vectors $\hat{\mathbf{y}}_q = ((\hat{y}_1)_q, ..., (\hat{y}_p)_q)$, with p the number of points along the fire front $c_f = 0.5$, $N_{pc} = (n + Q_{po})!/(n!\, Q_{po}!)$ the number of Hermite polynomials and Q_{po} the maximum order of the polynomial approximation.

Due to the orthogonality of the PC basis, the q-th PC coefficients are given by:

$$\hat{\mathbf{y}}_q = \frac{\mathrm{E}[H_{pc}(\hat{\mathbf{x}}_{c_i}^-)\Phi_q(\hat{\mathbf{x}}_{c_i}^-)]}{\mathrm{E}[\Phi_q(\hat{\mathbf{x}}_{c_i}^-)^2]}, \tag{3.6}$$

where $\mathrm{E}[\cdot]$ refers to the expectation operator satisfying $\mathrm{E}[\Phi_q(\hat{\mathbf{x}}_{c_i}^-)\Phi_l(\hat{\mathbf{x}}_{c_i}^-)] = 0$ if $q \neq l$, $\mathrm{E}[\Phi_q(\hat{\mathbf{x}}_{c_i}^-)^2]$ is a normalization factor, and $\hat{\mathbf{x}}_{c_i}^-$ is the quadrature roots vector of size N_{quad}. The numerator $\mathrm{E}[H_{pc}(\hat{\mathbf{x}}_{c_i}^-)\Phi_q(\hat{\mathbf{x}}_{c_i}^-)]$ is computed using a Gauss-Hermite quadrature rule; that is,

$$\mathrm{E}[H_{pc}(\hat{\mathbf{x}}_{c_i}^-)\Phi_q(\hat{\mathbf{x}}_{c_i}^-)] = \int_{\mathbf{R}^n} H_{pc}(\hat{\mathbf{x}}_{c_i}^-)\Phi_q(\hat{\mathbf{x}}_{c_i}^-) dP(\hat{\mathbf{x}}_{c_i}) \tag{3.7}$$

$$= \sum_{j=1}^{N_{quad}} H(\hat{\mathbf{x}}_{c_i}^{-,j})\, \Phi_q(\hat{\mathbf{x}}_{c_i}^{-,j})\, \omega^j, \tag{3.8}$$

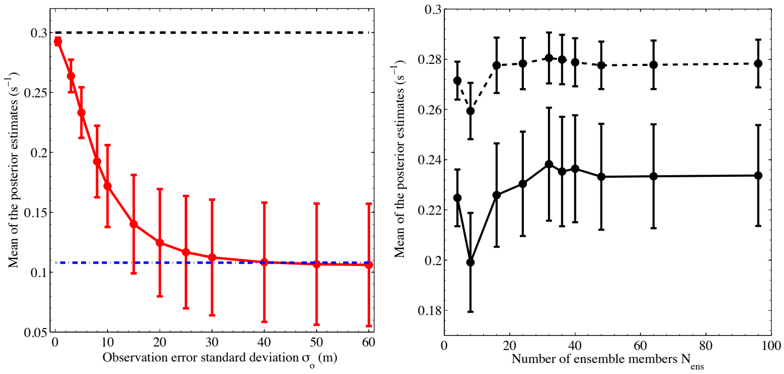
where $\mathbf{y}_{c_i}^{-,j} = H(\mathbf{x}_{c_i}^{-,j})$ is provided by the forward model integration evaluated at the j-th quadrature root $\hat{\mathbf{x}}_{c_i}^{-,j}$ with its associated weight ω^j. It is worth noting that this quadrature rule provides an exact integration of Eq. (3.8) if the order of the polynomial approximation is such that $2Q_{po} < 2(N_{quad} - 1)$.

The construction of the surrogate model H_{pc} requires therefore a limited number N_{quad} of forward model integrations. H_{pc} is then used in the EnKF algorithm, instead of the forward model H, to compute the predictions of the time-evolving fire front locations $\mathbf{y}_{c_i,pc}^-$ for a large number of members N_{ens} so as to properly estimate $\mathbf{C}_{\mathbf{xy}}$ and $\mathbf{C}_{\mathbf{yy}}$. Thus, the EnKF update can be performed with reliable covariances matrices, leading to the posterior estimates of the control parameters $\mathbf{x}_{c_i}^{+,k}$.

4. Results: Evaluation of the data-driven simulation capability

4.1. *Validation of the Ensemble Kalman Filter implementation*

As a validation step, the EnKF algorithm is applied to correct the parameter $P\,[s^{-1}]$ in Eq. (2.1), that accounts for all the uncertainties in the model. In the following experiments, the fire is ignited as a circular front and spreads upon a random fuel distribution. Observations are synthetically-generated each $100s$ with the forward model H and a chosen *true* value $\mathbf{x}_{c_1}^t = 0.3s^{-1}$; an observation error is also introduced, $\mathbf{y}_{c_i}^o = H(\mathbf{x}_{c_i}^t) + \epsilon^o$, where ϵ^o is a random variable with a Gaussian distribution characterized by the STD σ_o.



(a) Mean and STD of the posterior estimates as a function of the observation error STD σ_o.

(b) Convergence of the mean of the posterior estimates, for $\sigma_o = 2m$ (dashed line) and $\sigma_o = 5m$ (solid line).

FIGURE 2. Estimation of the parameter P for cycle 1. The black dashed line corresponds to the true parameter 0.3s^{-1}, the blue dashed-dotted line corresponds to the mean value 0.1s^{-1} of the prior estimates; and the red solid lines correspond to the mean of the posterior estimates and the vertical error bars to the associated STD σ_x.

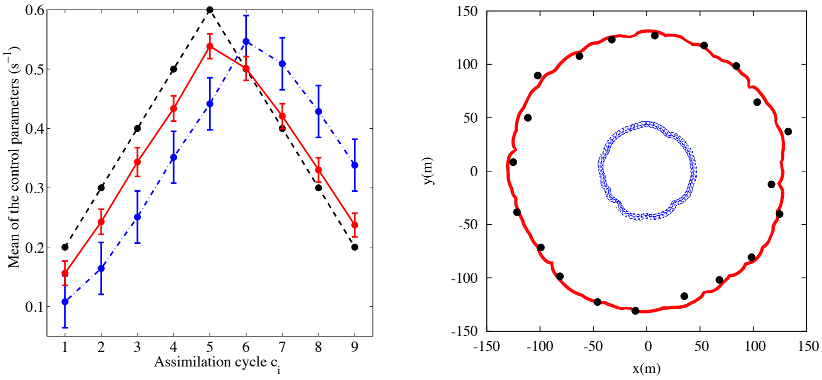
The ensemble of prior values for the first cycle $\mathbf{x}_{c_1}^{-,k}$ is drawn from a Gaussian distribution centered in 0.1s^{-1} with a STD $\sigma_x = 0.05s^{-1}$.

The red solid curve in Fig. 2(a) displays the mean of the posterior estimates over 1 cycle of EnKF when the observation error STD σ_o increases from $0m$ to $60m$. When σ_o is small (x-axis), the mean of the posterior estimates (y-axis) is equal to the true value $\mathbf{x}_{c_1}^t$ (black dashed line) with a small error STD σ_x (vertical error bars), meaning that the estimation of the parameter is reliable. As σ_o increases, the posterior estimates remain closer to the background value (blue dashed-dotted line) and the ensemble spread is larger (up to $0.05s^{-1}$). These results show that the DA algorithm allows to retrieve the true values of the control parameter even when the observation error is large. For this validation case, it was shown that the EnKF converges for a minimum of 48 members and that the convergence is reached more easily when σ_o is small, see Fig. 2(b).

The sequential application of the EnKF allows for a temporal correction of the parameter P for an experiment where the true time-varying value (black dashed curve) was artificially set as illustrated in Fig. 3(a) over 9 cycles for $N_{ens} = 48$ and $\sigma_o = 5m$. For cycle c_1, the mean value of the prior estimates is set to $0.1s^{-1}$, while for cycle c_i ($i \in [2, 9]$), it is set to the mean of the posterior estimates from the previous cycle c_{i-1} (blue dashed-dotted curve). The EnKF solution (red curves) allows to identify an optimal mean value of the parameter (see Fig. 3(a)) that results in an ensemble of fire fronts that is coherent with the observation error statistics, while the model without data assimilation (blue curves) significantly underestimates the rate of spread (see Fig. 3(b)).

4.2. *Application to a controlled grassland fire*

The EnKF is applied to a real-world case study corresponding to a reduced-scale controlled grassland fire ($4m \times 4m$) occurring under moderate wind conditions. The observed fire front locations are extracted from thermal-infrared imaging every $28s$, with a measurement error $\sigma_o = 0.047m$ (based on the spatial resolution of the camera).



(a) Mean of the prior (blue dashed-dotted curve) and posterior estimates (red solid curve) compared to the true value (black dashed line).

(b) Ensemble of fire fronts at cycle 9 with data assimilation (red curves) and without assimilation (blue curves) compared to the observations (black dots).

FIGURE 3. Sequential application of the EnKF over 9 cycles using $N_{ens} = 48$ members and observations generated with a time-varying true parameter and a constant error STD $\sigma_o = 5m$.

The control parameters are the fuel moisture M_f, the fuel particle surface-to-volume ratio Σ and the wind properties (both magnitude $\mathbf{u}_w$ and direction d_w). The prior estimates of these parameters are described in table 1 along with the associated STD. A prior ensemble of 512 members is generated and corrected by assimilating the fire front at time $t = 78s$. Results are presented in Fig. 4(a): the black dots represent the observed front (discretized with 40 points), the blue trajectory represents the spread of the forward model simulations obtained without data assimilation (i.e., using the prior estimates of the 4 control parameters), while the red trajectory derives from the data assimilation update using the posterior estimates in the forward model integrations. It is found that the EnKF algorithm allows to significantly decrease the distance between the observations and the simulated fronts at $t = 78s$. It should also be noted that the uncertainty on the front positions is reduced as the STD of posterior front positions (red error bars) is smaller to that of the prior front positions (blue error bars).

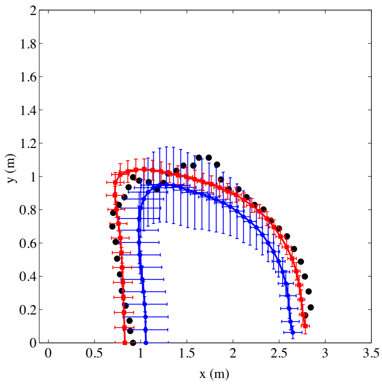
Fig. 4(b) compares the forecast of the front position at $t = 106s$ obtained using the prior estimates (blue curve) and the posterior estimates from DA at $t = 78s$ (red curve). The EnKF algorithm appears to properly represent the forecast trajectory at $t = 106s$ with a small uncertainty. This result illustrates the improved accuracy of the fire spread simulation and forecast using the EnKF. Still, this calibration is made at the expense of a heavy computational cost (512 members for 4 parameters).

4.3. *Towards a cost-effective Ensemble Kalman Filter*

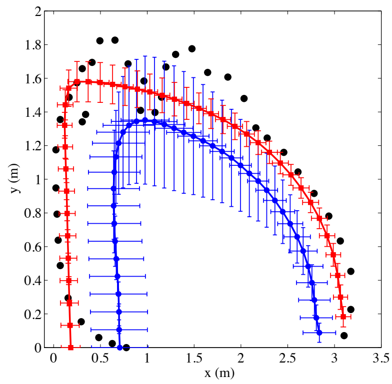
The application of the PC-based EnKF is illustrated for a wind-aided grassland fire spread directed northward ($\mathbf{u}_w = 0.8m/s$, $d_w = 0deg$) during $100s$. The control parameters are the moisture content M_f and the surface-to-volume ratio Σ; the means of the prior estimates are 15% and $13000m^{-1}$ with 3.3% and $3000m^{-1}$ error STD, respectively. The observations are synthetically-generated by adding an artificial noise $\sigma_o = 1m$ to the true trajectory obtained for $M_f = 10\%$ and $\Sigma = 14500m^{-1}$. A PC approximation (with a polynomial order $Q_{po} = 4$ and a quadrature order $N_{quad} = 5$) is used to build the model response surface to the prior control parameters. Figure 5(a) illustrates the observation

TABLE 1. Real-world case study: Prior and posterior ensemble statistics.

Control parameter	Prior mean value	Prior STD	Posterior mean value	Posterior STD
Moisture content M_f	$22.0\,\%$	$10.0\,\%$	$14.7\,\%$	$3.10\,\%$
Surface area/volume Σ	$11500\,m^{-1}$	$3000\,m^{-1}$	$13720\,m^{-1}$	$2364\,m^{-1}$
Wind magnitude $\mathbf{u_w}$	$1.0\,m/s$	$0.20\,m/s$	$0.88\,m/s$	$0.15\,m/s$
Wind direction d_w	$307.0\,deg$	$16.0\,deg$	$304.3\,deg$	$7.4\,deg$



(a) Assimilation time, $t = 78s$.



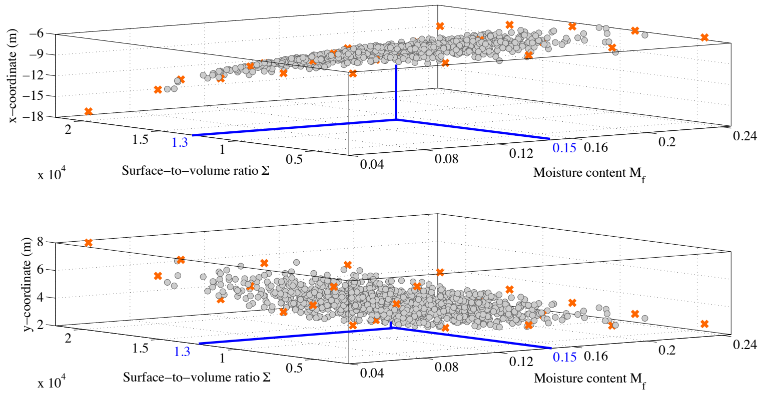
(b) Forecast time, $t = 106s$.

FIGURE 4. EnKF application to the grassland burning. The observations are represented with black dots; the simulations are represented with DA (red squared curves) and without DA (blue circled curves), the error bars (scale-factor of 1/2) are indicative of the ensemble spread.

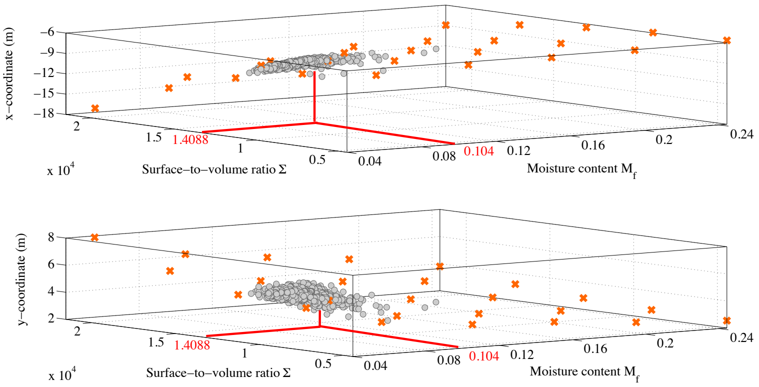
operator mapping onto the space spanned by the control parameters (for x_f- and y_f-coordinates of the fire front positions, top and bottom panels respectively). In the present case, the results are shown for the 5th point on the front position starting from the west side of the front. The orange crosses represent the simulated fire front positions associated to the 5×5 quadrature roots.

The surrogate model is used with $N_{ens} = 1000$ members to evaluate the prior fire front positions that are represented in Fig. 5(a) with grey circles for the 1000 prior estimates. The mean value of the prior estimates is shown in blue. It should be noted that the grey circles are contained within the surface response described by the orange crosses, meaning that the PC decomposition properly approximates the observation operator. Figure 5(b) shows the 1000 posterior estimates of the control parameters obtained using the EnKF. The means of the posterior estimates (shown in red) are 10.4% for M_f and $14088m^{-1}$ for Σ, which is more consistent with the true values. It also shows that the posterior ensemble STD is significantly reduced compared to the prior ensemble STD, meaning that the PC-based EnKF algorithm allows to retrieve reliable statistical information for only 25 forward model integrations. For this example, the computational time is reduced by a factor of at least 5 compared to the classical EnKF.

Figure 6 compares the simulated fire front using the mean value of the posterior ensemble with the most probable fire front derived from the PDF of the posterior estimates. It



(a) Prior estimates $\mathbf{y}_{c_1}^{-,k}$ $(k = 1, ..., 1000)$, mean value in blue.



(b) Posterior estimates $\mathbf{y}_{c_1}^{+,k}$ $(k = 1, ..., 1000)$, mean value in red.

FIGURE 5. Surrogate model of the observation operator H_{pc}. Prior and posterior estimates (grey circles) of the x_f- (top) and y_f-coordinates (bottom) of the fire front positions mapped onto the PC-based model surface response (orange crosses).

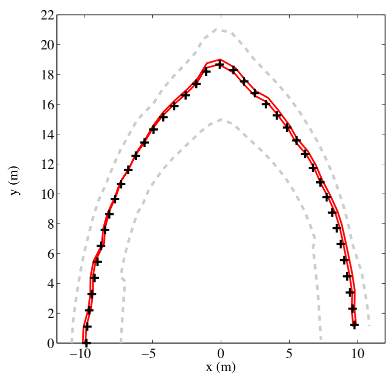


FIGURE 6. Comparison of the mean posterior front (black crosses) with the most probable front (red solid lines) at $t = 100s$. The ensemble spread is delimited by the grey dashed lines.

shows that these fronts feature a similar topology, meaning that the Gaussian assumption on the error statistics inherent in the EnKF remains valid for the present configurations.

5. Conclusions

A data assimilation strategy based on the Ensemble Kalman Filter (EnKF) is demonstrated to account for both experimental and modeling uncertainties in wildfire spread modeling. Based on a sequential multi-parameter estimation, this prototype is efficient at reducing the uncertainty in the numerical predictions of fire spread for synthetical and real measurements. The results show that, in the present configurations, a PDF sampling based on Polynomial Chaos allows to significantly reduce the computational cost of the EnKF and to provide accurate error statistics on model inputs and outputs.

The authors would like to acknowledge R. Paugam (King's College London) for sharing the experimental data, F. Duchaine and T. Morel (CERFACS) for their support on OpenPALM, as well as G. Iaccarino and P. Constantine (CTR) for the helpful discussions.

REFERENCES

Burgers, G., van Leeuwen, P. & Evensen, G. 1998 Analysis scheme in the ensemble kalman filter. *Mon. Weather Rev.* **126**, 1719–1724.

Durand, M., Andreadis, K., Alsdorf, D., Lettenmaier, D. & Moller, D. 2008 Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Ltrs.* **35**, L20401.

Evensen, G. 1994 Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics. *Journal of Geophysics Research* **99**, 10143–10162.

Ghanem, R. & Spanos, P. 2003 *Stochastic Finite Elements: a Spectral Approach (revised edition)*. Springer-Verlag.

Li, J. & Xiu, D. 2008 On numerical properties of the ensemble kalman filter for data assimilation. *Comput. Meth. Appl. Math. Eng* **197**, 3574–3583.

Li, J. & Xiu, D. 2009 A generalized polynomial chaos based ensemble kalman filter with high accuracy. *Journal of Computational Physics* **228**, 5454–5469.

Mandel, J., Bennethum, L., Beezley, J., Coen, J., Douglas, C., Kim, M. & Vodacek, A. 2008 A wildland fire model with data assimilation. *Mathematics and Computers in Simulation* **79**, 584–606.

Moradkhani, H., Sorooshian, S., Gupta, H. & Houser, P. 2005 Dual state-parameter estimation of hydrological models using ensemble kalman filter. *Advances in Water Resources* **28**, 135–147.

Rehm, R. & McDermott, R. 2009 Fire front propagation using the level-set method. *Tech. Rep.* 1611. National Institute of Standards and Technology.

Rochoux, M., Delmotte, B., Cuenot, B., Ricci, S. & Trouvé, A. 2012 Regional-scale simulations of wildland fire spread informed by real-time flame front observations. *Proc. Combust. Inst.* **34**.

Ros, D. D. & Borga, M. 1997 Adaptive use of a conceptual model for real time flood forecasting. *Nordic Hydrology* **28**, 169–188.

Rothermel, R. 1972 A mathematical model for predicting fire spread in wildland fuels. *Tech. Rep.* INT-115. US Department of Agriculture Forest Service, Intermountain Forest and Range Experiment.

Szunyogh, I., Kostelich, E., Gyarmati, G., Kalnay, E., Hunt, B., Ott, E., Satterfield, E. & Yorke, J. 2008 A local ensemble transform kalman filter data assimilation system for the ncep global model. *TellusA* **60**, 113–130.