UNIVERSITY OF CALIFORNIA, SAN DIEGO

Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering

A thesis submitted in partial satisfaction of the
requirements for the degree
Masters of Science

in

Mechanical Engineering

by

Thayjes Srivas

Committee in charge:

Professor Raymond De Callafon, Chair
Professor Thomas Bewley
Professor Phillip Gill

2016

The thesis of Thayjes Srivas is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2016

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Foremost, I would like to express my appreciation to my advisor, Professor de Callafon, for his guidance throughout my masters study. I especially appreciate his efforts of helping me became an independent researcher. Without him, I would not have been able to make as much progress as I have achieved in my research. I would like to thank my other committee members: Thomas Bewley and Philip Gill for making my graduate school experience a more rewarding one. Thanks to all the members of the WiFire team for their support and discussion on my work : Ilkay Altintas, Tomas Artes and Daniel Crawl especially for providing various help and suggestions for my research.
Last but not the least, I would like to thank my family for their understanding and encouragement throughout my study.

Chapters 1, 2, 4, 5 and 8 are, in almost full, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering in Procedia Computer Science Vol 80 (2016). Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

Chapters 6 and 7, in part, have been submitted for publication of the reformatted material as it may appear in American Control Conference 2017, Srivas, Thayjes; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this material.

ABSTRACT OF THE THESIS

Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering

by

Thayjes Srivas

Masters of Science in Mechanical Engineering

University of California, San Diego, 2016

Professor Raymond De Callafon, Chair

This thesis extends FARSITE (a software used for wildfire modeling and simulation) to incorporate data assimilation techniques based on noisy and limited spatial resolution observations of the fire perimeter to improve the accuracy of wildfire spread predictions. To include data assimilation in FARSITE, uncertainty on both the simulated wildfire perimeter and the measured wildfire perimeter is used to formulate optimal updates for the prediction of the spread of the wildfire. For data assimilation, Wildfire perimeter measurements with limited spatial resolution and a known uncertainty are used to formulate an optimal adjustment in the fire perimeter prediction. The adjustment is calculated from the Kalman filter gain in an Ensemble Kalman filter that exploits the uncertainty information on both the simulated wildfire perimeter and the measured wildfire perimeter. The

approach is illustrated on a wildfire simulation representing the 2014 Cocos fire and presents comparison results for hourly data assimilation results. In later chapters we extend FARSITE with the ability to update both fire perimeters and fuel adjustment factors to further improve the accuracy of wildfire spread predictions. To show the effectiveness of fuel adjustment factor updates, a comparison is made using an EnKF with fixed adjustment factor on a wildfire simulation representing the 2014 Cocos fire. The performance of the EnKF technique for tracking time varying fuel adjustment factors based on noisy and limited spatial resolution observations of the fire perimeter is also investigated.

Chapter 1

Introduction

1.1 Impact of Wildfires

As the occurrences of wildfires has increased over the recent past it is important to understand the effect they have on the society as a whole. While the wildfires are sometimes desirable from an ecological standpoint there is no denying the considerable and increasing social and economic costs. Wildfires are often described in terms of lives threatened, structures and homes lost or damaged, overall suppression costs and damage to the natural resources. The most publicized costs associated with wildfire are those to fight, or suppress, large wildfires. Average annual fire suppression expenditures by the U.S. Forest Service alone totaled 580 million from 1991 to 2000, and more than doubled to 1.2 billion annually from 2001 to 2010 (USDA Forest Service 2011c). State expenditures related to wildfire have also increased substantially in recent years. According to a biannual survey conducted by National Association of State Foresters (NASF), more than 1.6 billion annually is spent by State forestry agencies on wildfire protection, prevention, and suppression (including Federal funding expended by State agencies)and that number has more than doubled in the past 10 years (NASF 2010).These figures do not include the cost to local fire departments across the country, which, according to a survey by the National Fire Protection Association (NFPA), responded to an average of 36,700 fires annually in forests, woodlands, or other wildlands from 2004 to 200810 (Ahrens 2010). However, fire suppression expenses represent

only a fraction of the monetary value spent on or lost in damages due to wildfires. Numerous other costs include: the costs of restoring burned areas, lost tax and business revenues, property damage and/or devaluation, and costs to human health and lives. As an example, soil erosion and flash flooding following Colorados 1996 Buffalo Creek fire resulted in more than 2 million in flood damage as well as more than 20 million in damage to the Denver water supply system. Human lives lost or injured in the course of a wildfire are an incalculable societal cost. In the most extreme case to date in North America, the 1871 Pestigo Fire killed more than 1,200 people, destroyed numerous settlements, and burned 2,400 square miles across Wisconsin and Michigan [24]. Among these several costs the wildfire suppression cost has risen the most over the past few decades as the frequency, size and severity of wildfires has increased. In order to efficiently distribute resources and plan for these scenarios its important to predict and simulate these wildfires.

1.2 Models for Prediction and Simulation of Wildfires

Wildfire behavior is described by complex physical and chemical processes whose interactions depend on coupling between the atmosphere, topography, fuels and the fire itself. Fire models can generally be classified into three types: empirical, semi empirical and theoretical based on how they attempt to describe the fire behavior. The most commonly used operational models for wildfire spread use empirically derived relations to predict the spread rate of a fire. Examples are BehavePlus and FARSITE which are commonly used in the United States. These models basically describe the local spread rate as a function of wind speed, terrain slope, humidity and fuel characteristics. The spread rates used in these models are based on semi-empirical relations developed by Rothermel. Empirical models while simple and easy to execute do not work well in situations where strong fire-atmosphere interactions occur. The prediction of local wind and fire-wind interaction is of major importance to simulating fire behavior accurately. Physics or Theoretical based models attempt to address this by including fire/atmosphere

and fuel/fire interactions as well. They seek to approximately solve equations governing these complex processes and use these solutions to predict the wildfires [16]. While these models attempt to account for the many physical processes and appropriate scales it comes at the cost of computational complexity and longer periods for simulations/predictions. For faster predictions the empirical models would be more suitable. While these models do indeed have inaccuracies in their rate of spread functions these can be accommodated for by incorporating data assimilation techniques that make adjustments based on real data.

1.3 Data Assimilation in Wildfires

In the previous section we looked at wildfire modeling which involves the numerical simulation of wildfires in order to understand the properties and predict the fire behavior. The problem with these numerical models is that they are almost always a poor approximation of the actual behavior of the wildfire. Also even if the numerical model was a very good approximation of the actual wildfire, in order to predict or forecast the fire accurately we would need to know the ignition boundary of the wildfire with complete certainty, which is almost never the case. Based on this we can understand the need to include a data assimilation model, in order to use observations of the model to constantly adjust the forecast. Data assimilation is a technique used to incorporate data into a running model using sequential statistical estimation. Data assimilation is made necessary by the facts that no model is perfect, the available data is spread over time and space, and it is burdened with errors. Mandel's work on incorporating data assimilation models into wildfire modeling focuses on this aspect by finding an adequate balance between the sophistication of the model and execution time. The model uses two partial differential equations for prediction of heat and fuel combined with a simple combustion model. Finally they incorporate a version of the Ensemble Kalman Filter to provide data assimilation capabilities to the model [13]. In Trouve's work, an Eulerian front propagation solver FIREFLY, which uses a description of the local rate of fire spread (ROS) as a function of environmental conditions based

on Rothermel's model, is used as a simulator and a data assimilation model based on an Ensemble Kalman Filter for parameter estimation [20].

A software widely used for wildfire forecasting purposes by the U.S Forest Service and other federal and state agencies is FARSITE [7]. While FARSITE's rate of fire spread model is sophisticated but still computationally inexpensive, it should be noted that it does not incorporate any stochastic aspects when simulating the actual wildfire. Simulations in FARSITE, under a given set of inputs and parameters and without the probabilistic generation of embers, is largely a deterministic process. FARSITE also does not possess any features for incorporating noisy and finite spatial resolution measurements of the fire perimeter during the simulation and thus has no data assimilation capabilities.

The use of data assimilation techniques in wildfire spread models is an active field of research [14, 15]. Due to the highly large-scale spatial-temporal simulations required in wildfire simulation, merging simulations and measurements may use Monte Carlo methods for data assimilation techniques [26]. Other approach provided detailed models of the wildfire in interaction with the atmosphere [12] where Tikhonov regularization is used to avoid nonphysical states. Some of the most recent data assimilation techniques for wildfire spread prediction [20, 19] heavily rely on Ensemble Kalman filtering [6, 9, 3, 11] also use in many earth science applications [5]. Although most of these methods can assimilate gridded data, the use of noisy and finite spatial resolution measurements via the explicit incorporation of uncertainty in both the simulate and measured fire perimeter for optimal adjustment of the predicted fire perimeter is often overlooked.In addition to this, these methods use a finite and fixed dimension of the state during the assimilation process, while in this thesis we allow the dimension of the state to increase over time improving resolution of the updated perimeters. The aim of this thesis is to include data assimilation in FARSITE via standard ensemble averaging and optimal adjustments via Kalman filter computations. Uncertainty on both the simulated wildfire perimeter and the measured wildfire perimeter is used to formulate optimal updates for the prediction of the spread of the wildfire. For that purpose, the estimate of the initial fire perimeter is augmented with a

confidence region that is characterized by a covariance matrix. Ensemble sampling based on the mean and covariance information is then used to propagate the uncertainty through FARSITE for a stochastic update of the wildfire spread simulation. Subsequently, wildfire perimeter measurements with limited spatial resolution and a known uncertainty are used to to formulate an optimal adjustment in the fire perimeter prediction. The adjustment is calculated from the Kalman filter gain in an Ensemble Kalman filter that exploit the uncertainty information on both the simulated wildfire perimeter and the measured wildfire perimeter. FARSITE also allows the option of using "adjustment factors" to account for frequent variation in wind speed and incorrect modeling. In later sections we will look into augmenting these factors to the state to improve accuracy and prediction of wildfires.

Chapter 1 is, in some part, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering in Procedia Computer Science Vol 80 (2016).. Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.;Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

Chapter 2

FARSITE

2.1 Forward Model for Wildfire Simulation

FARSITE [7], largely based on Rothermel's model [21], is widely used by the U.S. Forest Service as an effective tool of simulating the growth of natural fires in wilderness areas. It can be seen as a dedicated forward-prediction model taking the form of equation 2.1 which uses spatial and temporal information on the parameters θ_k and driving inputs u_k to predict a wildfire perimeter on a two-dimensional plane, denoted by $\bar{x}_{k+1|k}$.

$$\begin{aligned} \hat{x}_{k+1|k} &= f(\hat{x}_{k|k}, \theta_k, u_k) \\ \hat{y}_{k+1|k} &= C_{k+1}\hat{x}_{k+1|k} \end{aligned} \tag{2.1}$$

For the considered wildfire data assimilation, the (measured) output y_{k+1} refers to a spatially downsampled (coarse) measurement of the actual wildfire perimeter, whereas wildfire related parameters θ_k may include topography and fuel parameters, and driving input u_k can refer to weather and wind conditions.

The FARSITE function $f(\cdot)$ in (2.1) is an implicit and high dimensional forward model that models fire growth via a vector approach and includes fire behavior models for surface fire spread [21], crown fire initiation [25], crown fire spread [22] and dead fuel moisture [4, 17]. Furthermore, the dimension n_k of the two-dimensional perimeter $\hat{x}_{k-1|k-1} \in \mathbb{R}^{2n_{k-1}}$ typically changes to $n_k \geq n_{k-1}$ for

$\hat{x}_{k|k-1}$ over the course of the fire simulation. The implicit knowledge of the forward model $f(\cdot)$ and changing dimension of the state $\hat{x}_{k|k-1}$ makes FARSITE and ideal application candidate for ensemble based state estimation to extend FARSITE with data assimilation capabilities to improve wildfire simulation and any success in this regard will enhance the FARSITE-based data assimilation capabilities for improved wildfire simulation.

2.2 FARSITE data inputs

FARSITE requires an input parameters set that describes the environment where the wildfire is taking place. The parameters can be classified depending on variability of such values (in time and/or space). The static values (which varies mostly spatially) describing the topography, are raster maps which are combined in one single file known as FARSITE landscape file. This file contains digital elevation model information as slope, elevation or aspect as well as the description of the vegetation land cover and the fuel map [1], [2], [23].

The ignition boundary is introduced into FARSITE in a vectorial file format. In such vectorial format, some input parameters as the fire line intensity(FLI) and the rate of spread(ROS) for each point of the boundary are included. Finally, the time varying parameters are mostly weather values. This Linux FARSITE version uses a brief daily description of the weather except for the wind values. Those values are introduced every 10 minutes using a gridded format file. In this manner, an important parameter for wildfires as the wind could be updated frequently.

The weather information used in this work has been obtained from weather stations of the High Performance Wireless Network for Education and Research (HPWREN) [18, 10]. HPWREN allows requesting data in real time from the most near station to the centroid of the initial wildfire boundary. The weather data is obtained in XML format and the request is parsed and written for FARSITE.

Chapter 2 is, in full, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman

Filtering in Procedia Computer Science Vol 80 (2016). Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

Chapter 3

Describing Uncertainty in Wildfire Perimeters

In this section we will describe how we characterize the uncertainty of the wildfire perimeter.

1.1 Representing uncertainty information at a point on a perimeter :

Uncertainty at a point can either be represented as an ellipse or as a 2×2 covariance matrix.
The covariance ellipse is defined by 4 parameters namely :
1. α (the angle between the major axis and the positive x axis in the counter-clockwise direction)
2. σ_x (the standard deviation of the x co-ordinate)
3. σ_y (the standard deviation of the y co-ordinate)
4. s (this parameter determines the certainty/confidence of the ellipse, NOTE : this does not affect the covariance matrix but affects the plotting of the ellipse)
The equation of the ellipse is given by

$$(x/\sigma_x)^2 + (y/\sigma_y)^2 = s \tag{3.1}$$

$$s = \left\{ \begin{array}{ll} 9.210 & \text{, for 99\% certainty} \\ 5.991 & \text{, for 95\% certainty} \end{array} \right\}$$

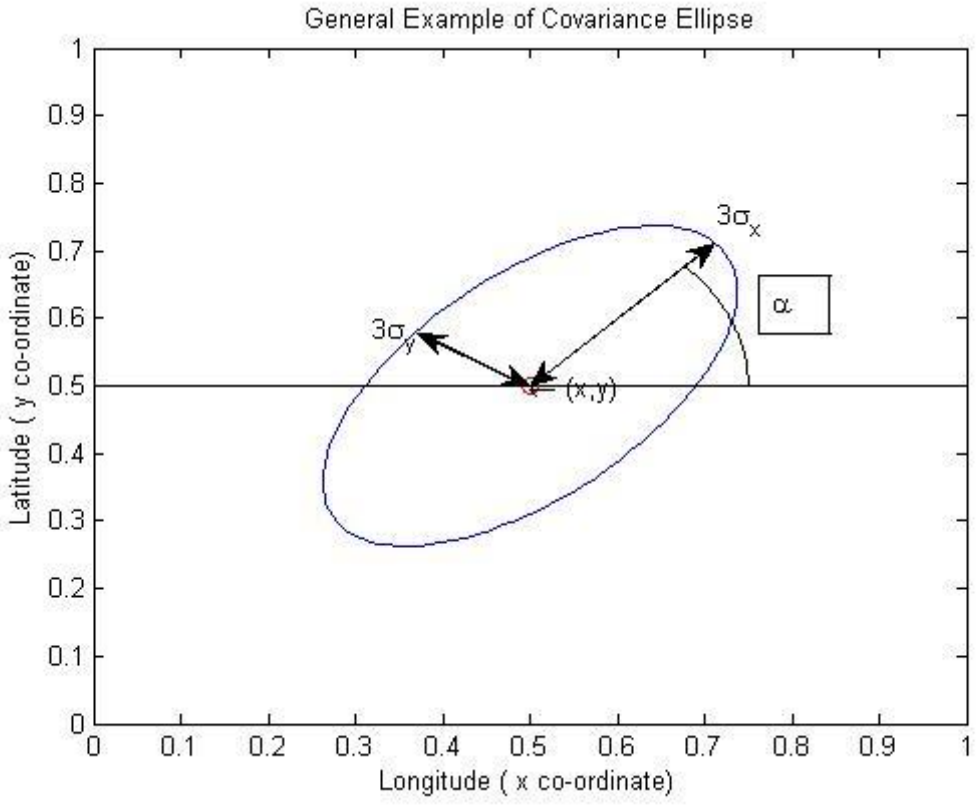


Figure 3.1: Description of Covariance Ellipse

The above figure is a general covariance ellipse of the point (x, y), with length of semi major axis $= 3\sigma_x$, length of semi minor axis $= 3\sigma_y$ and angle of rotation α.

The co-ordinates of the ellipse $(\widetilde{x}, \widetilde{y})$ in the figure are the result of a rotation matrix (with rotation angle α) on the ellipse given by equation (3.1).

$$\begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix}$$

$$P(x,y) = E[\begin{bmatrix} x \\ y \end{bmatrix} \cdot \begin{bmatrix} x^T & y^T \end{bmatrix}]$$

$$P(x,y) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \times P(\widetilde{x},\widetilde{y}) \times \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{3.2}$$

Where

$$\begin{aligned} P(\widetilde{x},\widetilde{y}) = & E[\begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix} \cdot \begin{bmatrix} \widetilde{x}^T & \widetilde{y}^T \end{bmatrix}] \\ = & \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \end{aligned}$$

The above result comes from the fact that

$$\begin{aligned} E[\widetilde{x}\widetilde{x}^T] &= E[\widetilde{x}^T\widetilde{x}] \\ &= var(\widetilde{x}) \\ &= \sigma_x^2 \end{aligned}$$

Expanding (3.2) we get ,

$$P(x,y) = \begin{bmatrix} \sigma_x^2\cos\alpha^2 + \sigma_y^2\sin\alpha^2 & (\sigma_y^2 - \sigma_x^2)\cos\alpha\sin\alpha \\ (\sigma_y^2 - \sigma_x^2)\cos\alpha\sin\alpha & \sigma_y^2\cos\alpha^2 + \sigma_x^2\sin\alpha^2 \end{bmatrix} \tag{3.3}$$

.

Equation (3.3) is a general equation for calculating a covariance matrix for a point (x, y) on the boundary given those parameters at the point.

It is important to note here that given (3.3) we can also obtain the parameters of the ellipse by using a Singular Value Decomposition of $P(x, y)$

$$svd(P(x,y)) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \times \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{3.4}$$

(3.4) is in fact the singular value decomposition of $P(x, y)$. The command $svd(P)$ can be used to get these 3 matrices and hence the α, σ_x *and* σ_y values for the covariance matrix P.

CONCLUSION :
Given the covariance ellipse or the parameters α, σ_x, σ_y at a point, we can calculate the covariance matrix P using (3.3).

$$P(x,y) = \begin{bmatrix} \sigma_x^2 \cos\alpha^2 + \sigma_y^2 \sin\alpha^2 & (\sigma_y^2 - \sigma_x^2)\cos\alpha\sin\alpha \\ (\sigma_y^2 - \sigma_x^2)\cos\alpha\sin\alpha & \sigma_y^2 \cos\alpha^2 + \sigma_x^2 \sin\alpha^2 \end{bmatrix}$$

Also given a covariance matrix P at a point (x, y) we can use the Singular Value Decomposition :

$$svd(P(x,y)) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \times \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

to obtain the parameters and use them to plot the covariance ellipse of the point.

1.2 Example
This covariance ellipse was plotted using the following parameters :

$$\alpha = 60^\circ \quad \sigma_x = 0.1 \quad \sigma_y = 0.05 \quad s = 9$$

Using the results in Section 1.1, we can calculate the covariance matrix at the above point :

$$P(-1,-1) = \begin{pmatrix} 0.0044 & -0.0032 \\ -0.0032 & 0.0081 \end{pmatrix}$$

Note that given this matrix P, we could have use the SVD to get the parameters and plot the above ellipse.

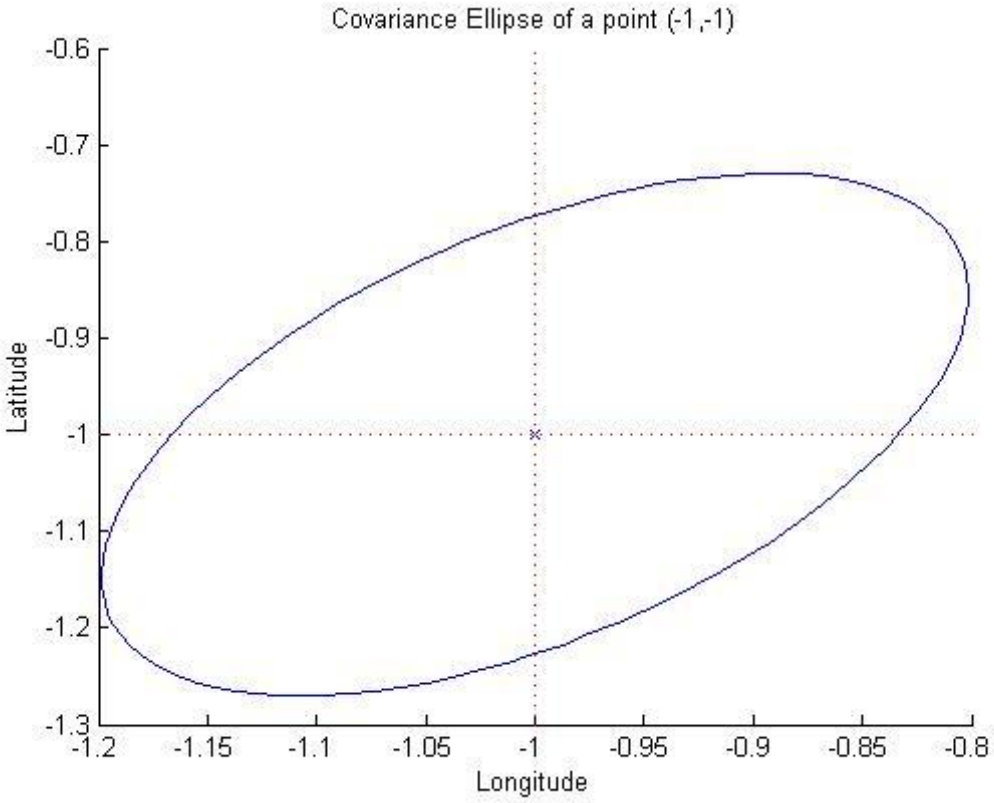


Figure 3.2: Covariance Ellipse Example

Chapter 4

Ensemble Kalman Filter with FARSITE

The Ensemble Kalman Filter (EnKF) [6] is a Monte-Carlo implementation of the Bayesian update problem. Given a probability distribution of the system (the prior) and measurement likelihood, the Bayes theorem is used to obtain the probability distribution with the data taken into account (the posterior). The state estimate is the mean of the posterior distribution. The data likelihood is the conditional probability distribution of the measurement given the current state. The EnKF assumes a Gaussian Distribution for the state variables implying that the state can be entirely characterized by the mean and the covariance of the ensemble.

4.1 Forward simulations

Unlike the Kalman Filter which uses linear equations to propagate the mean and covariance in time, the EnKF propagates the uncertainty by advancing each ensemble member through the forward model. This advantage of the EnKF provides an incentive to use it in high spatial-temporal simulations such as the FARSITE simulation software. For a brief overview of using the EnKF in FARSITE, consider the following computational steps [9]:

1. Initialize the distribution by defining a mean and a covariance of the state.

$$x_{k|k} \sim (\bar{x}_k, P^x_{k|k}), \quad x_{k|k} \in R^{n_k}.$$

2. Generate the ensembles by sampling from this distribution.

$$X_{k|k} = (x^1_{k|k}, x^2_{k|k}, x^3_{k|k}, .., x^N_{k|k}), \quad X_{k|k} \in R^{n_k \times N}$$

We now define the ensemble perturbation/error matrix $E_{x_{k|k}}$

$$E^x_{k|k} = (x^1_{k|k} - \bar{x}_k, x^2_{k|k} - \bar{x}_k, x^3_{k|k} - \bar{x}_k, ..., x^N_{k|k} - \bar{x}_k), \quad E^x_{k|k} \in R^{n_k \times N} \tag{4.1}$$

3. Advance each ensemble member through the forward model (FARSITE)

$$x^i_{k+1|k} = f(\bar{x}_{k|k} + w_k, u_k), \quad i = 1, 2, \ldots, N, \quad x_{k+1|k} \in R^{n_{k+1}} \tag{4.2}$$

$$w_k \sim (0, P^x_{k|k})$$

4. Calculate the mean and sample covariance of the forward ensemble. The mean is calculated simply by taking an average of the members of the forward ensemble. The sample covariance can be calculated using the forward ensemble error matrix, $E^x_{k+1|k}$, which can be calculated similar to (4.1), replacing $x_{k|k}$ by $x_{k+1|k}$:

$$\bar{x}_{k+1|k} = \frac{1}{N} \sum_{i=1}^{i=N} x^i_{k+1|k}$$
$$P^x_{k+1|k} = \frac{1}{N-1} \; E^x_{k+1|k} (E^x_{k+1|k})^T$$

We will now shortly describe how the above variables and steps are defined and carried out in a wildfire data and assimilation context using FARSITE. The mean $\bar{x}_{0|0}$ at $k = 0$ is the state which describes the ignition boundary (initial fire

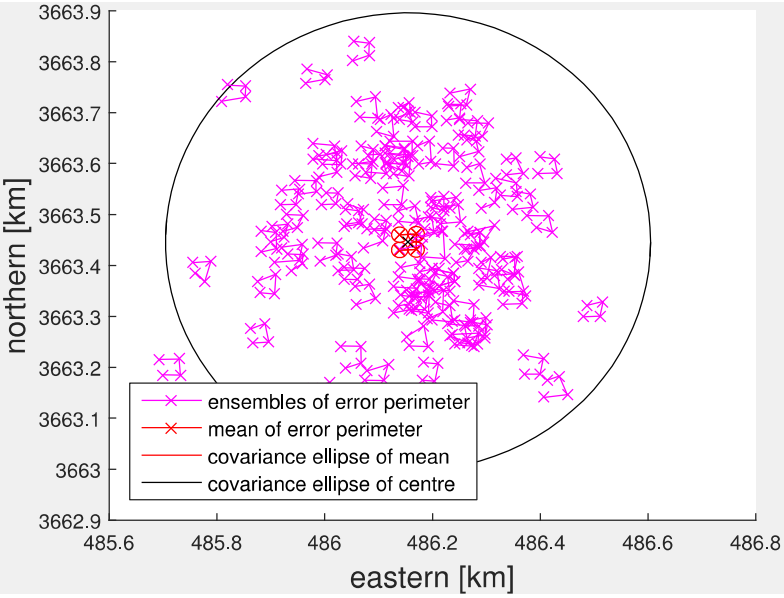


Figure 4.1: Initial Ensemble with Centroid Standard Deviation 150m and Perimeter/State Co-ordinates Standard Deviation 5m.

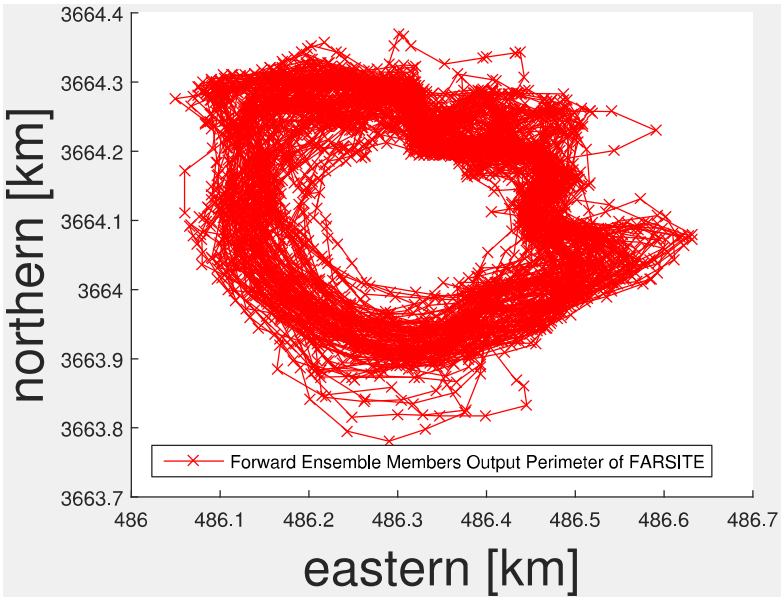


Figure 4.2: An example Forward/Forecasted Ensemble Members at k=3. These are the output perimeters of FARSITE which are the results of the simulation from each ignition boundary.

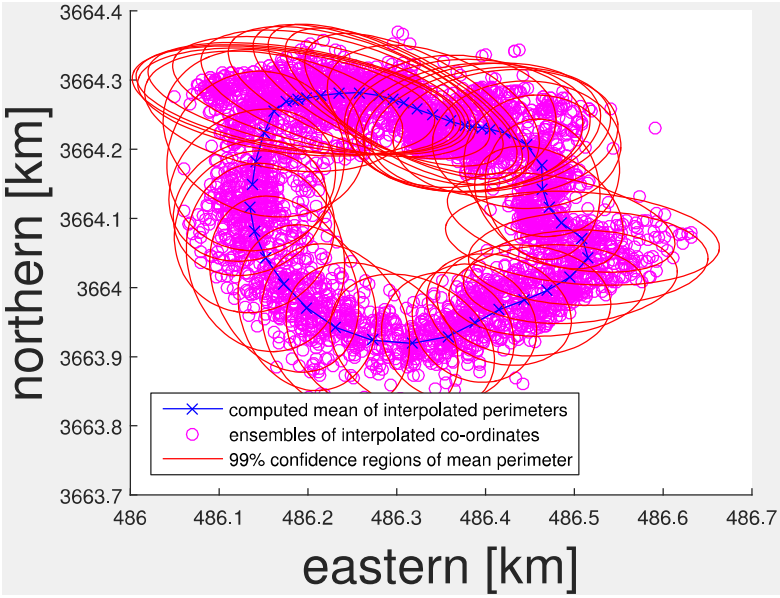


Figure 4.3: Example of Interpolated Ensemble Members at k=3. These are the interpolated perimeters which are used to calculate the sample mean and sample covariance matrix.

perimeter) in eastern e_i and northern n_i coordinates

$$\bar{x}_{0|0} = \begin{bmatrix} e_1 & n_1 & e_2 & n_2 & \cdots & e_m & n_m \end{bmatrix}^T \tag{4.3}$$

where (e_j, n_j) is the jth coordinate of the ignition boundary and $n^0_{0|0} = 2m$ is the dimension of the state variable $\bar{x}_{0|0}$. Let (e_c, n_c) indicate the centroid of the ignition boundary in eastern/northern coordinates, then the covariance of the state $P^x_{0|0}$ at $k = 0$ is the matrix

$$P^x_{0|0} = AP^{x_c}_{0|0}A^T + P^{\bar{x}}_{0|0} \tag{4.4}$$

where $P^{x_c}_{0|0}$ is the 2×2 covariance matrix for the center point $x_c = [e_c \; n_c]^T$ and $P^{\bar{x}}_{0|0}$ is the $n^0_{0|0} \times n^0_{0|0}$ covariance matrix of the state $\bar{x}_{0|0}$. The matrix A is a $n^0_{0|0} \times 2$ transformation matrix that relates $\bar{x}_{0|0}$ to the vector $x_c = [e_c \; n_c]^T$ via $\bar{x}_{0|0} = A[e_c \; n_c]^T$.

Describing the initial state covariance matrix as in (4.4) allows us to maintain the overall shape and size of the ignition boundary while still giving us freedom to impose a large uncertainty on the actual location of the ignition boundary (Fig

4.1) via a covariance matrix $P_{0|0}^{x_c}$ on the center point $x_c = [e_c\ n_c]^T$. Having defined the initial mean and covariance matrix, step 2 is implemented by simply sampling from a normal distribution described by tits mean and covariance. Each of the ensemble members $x_{k|k}^i$ constitutes an independent ignition boundary which is used to compute an update ensemble $x_{k+1|k}^i$ through a forward model (FARSITE) simulation. The simulations in (4.2) can be carried out in parallel n order to increase computational speed. Note that each of the runs use the same parametric and input conditions during the FARSITE simulation. As the dimension of the state increases, there are no clear cut guidelines for selecting the number of ensemble members. There are results indicating that the ensemble size would need to be increased in accordance with the state size to ensure convergence of the trace of covariance matrix and reduction of bias in the estimate [8]. Unfortunately a large ensemble size is not computationally viable and hence certain heuristics, such as covariance localization and inflation may need to be implemented to account for the gross under sampling used to approximate the covariance matrix. It is to be noted, in our case we maintain a sufficient ensemble size and spurious correlations do not appear in our sample covariance matrix for present simulations.

4.2 Fire Perimeter Adjustment Using Observations

The output perimeters $x_{k+1|k}^i$ obtained from the forward model (FARSITE) simulation constitute the members of the forward/forecasted ensemble (Fig4.2). These forward perimeters are generally vectors with a size different and larger than the initial size $n_{0|0}^0$. Furthermore, at each time time step $k+1$, the output perimeters $x_{k+1|k}^i$ may even have different sizes $n_{k+1|k}^1, n_{k+1|k}^2, \ldots, n_{k+1|k}^N$ depending on the ensemble $x_{k|k}^i$ used to compute $x_{k+1|k}^i$. To allow state updates with the EnKF approach for a varying state dimension, each new output perimeters $x_{k+1|k}^i$ is re-interpolated in 2D (eastern/northern) to a new size

$$n_{k+1|k} = \max(n_{k+1|k}^1, n_{k+1|k}^2, \ldots, n_{k+1|k}^N)$$

in order to preserve resolution at each time step $k+1$. Finally the mean $\bar{x}_{k+1|k}$ and sample covariance $P^x_{k+1|k}$ are calculated using the interpolated output perimeters of the forward ensemble (Fig4.3) and allow us to continue the computational steps for data assimilation as follows.

5. We now define the distribution of observations with a mean, $y_{k+1} \in R^m$ and the observation covariance matrix, $V_{k+1} \in R^{m\times m}$ and generate the ensemble of observations by sampling from this distribution:

$$Y_{k+1|k} = (y^1_{k+1|k}, y^2_{k+1|k}, ..., y^N_{k+1|k}), \quad Y_{k+1|k} \in R^{m\times N}$$

and the ensemble of perturbations/error for the observations:

$$E^y_{k+1|k} = (y^1_{k+1|k} - y_{k+1}, y^2_{k+1|k} - y_{k+1}, ..., y^N_{k+1|k} - y_{k+1}), \quad E^y_{k+1|k} \in R^{m\times N}$$

The sample covariance of the observations and sample cross covariance between the state and observations can now be calculated via

$$\begin{aligned} P^y_{k+1|k} &= \frac{1}{N-1} E^y_{k+1|k} (E^y_{k+1|k})^T \\ P^{xy}_{k+1|k} &= \frac{1}{N-1} E^x_{k+1|k} (E^y_{k+1|k})^T \end{aligned}$$

6. The Kalman gain is now calculated using the above computed sample covariances

$$K_{k+1} = P^{xy}_{k+1|k} (P^y_{k+1|k})^{-1} \tag{4.5}$$

7. The next step is the update step given by

$$x^i_{k+1|k+1} = x^i_{k+1|k} + K_{k+1}(y^i_{k+1} - C_{k+1} x^i_{k+1|k})$$

$$\bar{x}_{k+1|k+1} = \frac{1}{N} \sum_{i=1}^{i=N} x^i_{k+1|k+1}$$

where $y^i_{k+1} = y_{k+1} + v^i_{k+1}$ and y_{k+1} is the actual (noisy and downsampled) observation of the fire perimeter and v^i_{k+1} is a zero mean random variable

with $N \sim (0, V_{k+1})$.

8. Finally the updated sample covariance is calculated using the ensemble of the updated state via

$$P^x_{k+1|k+1} = \frac{1}{N-1} E^x_{k+1|k+1} (E^x_{k+1|k+1})^T$$

where, $E^x_{k+1|k+1}$ is again similar to (4.1) with $x_{k+1|k+1}$ replaced by $x_{k+1|k}$ and $\bar{x}_{k+1|k+1}$ replaced by $\bar{x}_{k|k}$. At any time instant, the estimate of the state, $\hat{x}_{k+1|k+1}$, is the mean of the ensemble, $\bar{x}_{k+1|k+1}$.

For the continued iteration along the time index k, one replace $k \to k+1$ and repeat steps 2 to 8. The above steps can be described in context for wildfire data assimilation. It is clear that y_{k+1} in in step 5 is defined as the downsampled version of the "true" fire perimeters given by $y_{k+1} = C_{k+1} x_{k+1}$ where C_{k+1} is the spatial downsampling matrix and x_{k+1} is the true (yet unknown) state data of the fire perimeter.

At this step we also introduce the observation covariance matrix V_{k+1}. In the wildfire context, this characterizes the variance on our measurements of the fire perimeters y_{k+1}. The mean and covariance matrix of the observations allows us to apply uncertainty on our measurements which is very important in calculating the Kalman gain K_{k+1} in step 6. In the next step we update each of our interpolated perimeters, $x^i_{k+1|k}$ which we obtained from the output of the FARSITE, using the Kalman gain and a sample measurement from the ensemble of observations. The ensemble of observations is characterized by its mean y_{k+1} and covariance matrix V_{k+1} described earlier. These updated perimeters $x^i_{k+1|k+1}$ are the fire perimeters which will be used to resume our simulation from the $k+1$th time step. In the remaining steps we do a similar calculation as in the first 4 steps and obtain the updated mean of the fire perimeters $\bar{x}_{k+1|k+1}$ and the updated sample covariance matrix $P^x_{k+1|k+1}$.

Chapter 4 is, in most part, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble

Kalman Filtering in Procedia Computer Science Vol 80 (2016). Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

Chapter 5

Application in WildFire Data Assimilation

5.1 Reference Data for Simulation

In this section we collect topography data and weather conditions from the May 2014 Cocos Fire in San Marcos and use this to generate parametric conditions and input data for a FARSITE wildfire simulation. These parametric conditions and input data are used to produce the "true" fire perimeters x_{k+1} that will be used as a reference for the performance evaluation of the data assimilation tools in this paper. The simulated fire perimeters are depicted in Fig. 5.1 over a 18 hour time period with a one hour time resolution and a 90m spatial resolution along the perimeter, starting from a given square 30m$\times$30m ignition boundary x_0 at $t = k = 0$.

For testing the data assimilation tools, the "true" data with a 90m spatial resolution along the perimeter is down-sampled to generate the output y_{k+1} by

$$y_{k+1} = C_{k+1}x_{k+1} + v_{k+1},$$

where C_{k+1} is a downsampling matrix, and measurements y_{k+1} are produced at a spatial resolution of only 360m along the perimeter. The measurements are perturbed by a white noise v_{k+1} with a standard deviation of 50m. In addition,

it is assumed that the ignition boundary x_0 is not known at the start of the data assimilation procedure. Instead, an initial estimate $\hat{x}_{0|0} \neq x_0$ along with a covariance matrix is used.

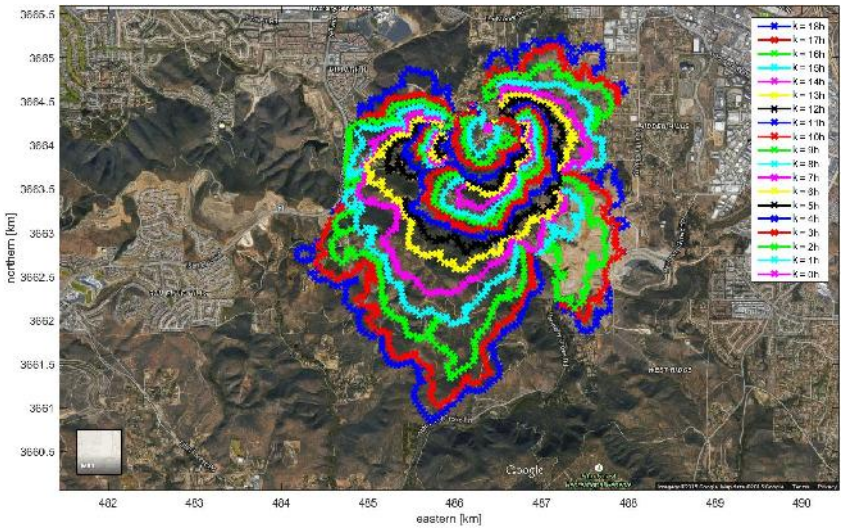


Figure 5.1: Noise free reference data of the "true" hourly fire perimeters x_k starting at x_0 over a 18 hour time period and a spatial resolution of 90m along the fire perimeters.

5.2 Forward Simulation Without Data Assimilation

To illustrate the need for data assimilation, a forward simulation from FARSITE is initialized at $\hat{x}_0 \neq x_0$. The ignition boundary (initial fire perimeter) $\hat{x}_0$ is also characterized by a square 30×30m fire perimeter, but with the center of the perimeter 215m off in eastern direction and 730m off in in northern direction compared to the "true" x_0. A side-by-side comparison can be made between the reference data (the "true" fire perimeters) x_{k+1} in Fig. 5.1 and the fire perimeters $\hat{x}_{k+1}$ obtained by the forward simulation from FARSITE initialized at $\hat{x}_0 \neq x_0$ in Fig. 5.2. It is clear that the relatively small initial error due to $\hat{x}_0 \neq x_0$ between the initial fire perimeters leads to a growing divergence of the fire perimeters over time.

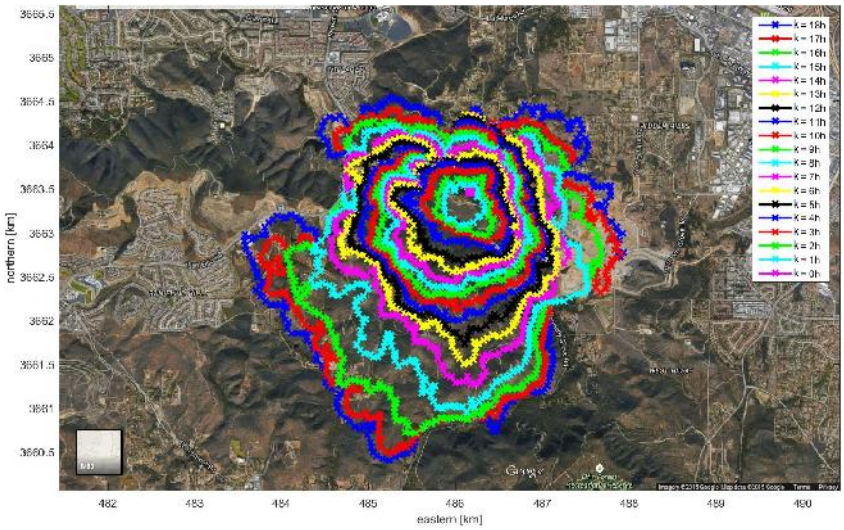


Figure 5.2: Forward (FARSITE) simulation of hourly fire perimeters $\hat{x}_k$ over a 18 hour time period with a spatial resolution of 90m along the fire perimeters started at an off-set initial fire perimeter $\hat{x}_0 \neq x_0$.

The diverging error can be characterized via the Root Mean Square (RMS) error

$$E_{rms_k} = \left(\frac{\sum_{i=1}^{i=n_k} (x_k - \hat{x}_k)^2)}{n_k} \right)^{1/2} \tag{5.1}$$

where n_k is the state size, representing the number of points on the fire perimeter. Evaluating the RMS error for the forward simulation of hourly fire perimeters $\hat{x}_k$ shown earlier in Fig. 5.2, leads to the progress of the RMS error summarized in Fig. 5.3. It is clear from this figure that a simple forward simulation with FARSITE, without the corrections provided by data assimilation, does not improve the RMS error and eventually leads to a diverging RMS error. Data assimilation is needed to stabilize the RMS error and correct for the initial error in the fire perimeter $\hat{x}_0 \neq x_0$.

5.3 Data Assimilation with Hourly Updates

Using noisy and down-sampled measurements y_{k+1} at time index $k+1$, an estimate $\hat{x}_{k+1|k+1}$ is formulated of the "true" fire perimeter x_{k+1} via the ensemble

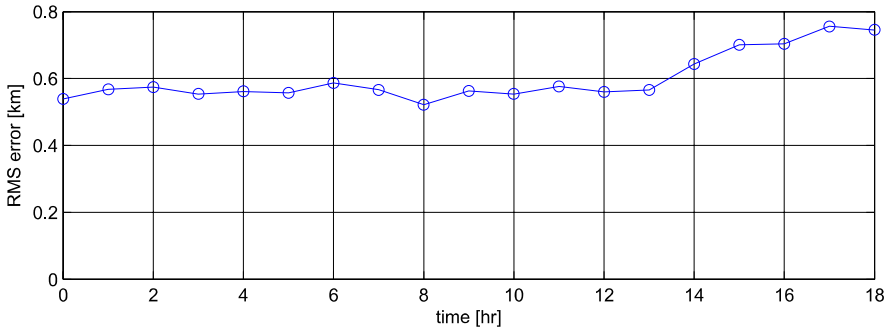


Figure 5.3: RMS error of hourly fire perimeters $\hat{x}_k$ computed via forward (FARSITE) simulation started at an off-set initial fire perimeter $\hat{x}_0 \neq x_0$.

Kalman filter (EnKF) approach outlined earlier in Section 4. To initialize the data assimilation procedure, the same inaccurate value of the initial fire perimeter $\hat{x}_{0|0} \neq x_0$ will be used, where $\hat{x}_{0|0}$ is a square 30×30m fire perimeter, but with the center of the perimeter 215m off in eastern direction and 730m off in in northern direction.

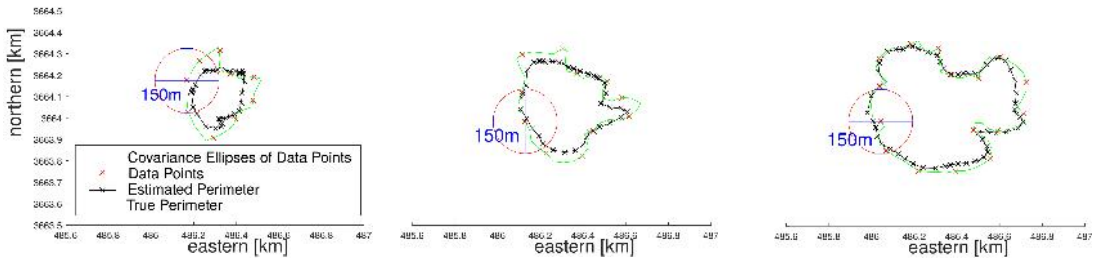
As each point on the fire perimeter consists of a eastern and northern coordinate, the initial square fire perimeter of 4 perimeter points requires an 8 dimensional state estimate $\hat{x}_{0|0}$. The ensembles are created using an 8×8 covariance matrix $P^{\bar{x}}_{0|0}$ for the mean $\bar{x}_{0|0} = \hat{x}_{0|0} \neq x_0$ and a 2×2 covariance matrix $P^{x_c}_{0|0}$ for the center x_c of the square ignition boundary. The covariance matrices are given by

$$P^{\bar{x}}_{0|0} = diag\{P, P, P, P\}, \quad P = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, \quad P^{x_c}_{0|0} = \begin{bmatrix} 150 & 0 \\ 0 & 150 \end{bmatrix}$$

and indicate a relative large uncertainty on the center point and a smaller uncertainty on the individual corner points of the square ignition boundary. The covariance information is combined via (4.4) to get the complete initial covariance matrix $P^{x}_{0|0}$ to create initial ensembles $x^i_{0|0}$, $i = 1, 2, \ldots, N$ that are advanced through the forward model (FARSITE) in (4.2) where $N = 100$.

During the subsequent steps of the EnKF method outlined in Section 4, the comparison of the updated perimeter and the perimeter of the reference fire along with a confidence region on one of the coordinates of the reference fire is depicted in Fig. 5.4. It can be observed that the predicted and reference fire perimeter

converge fairly quickly. This is also confirmed by the plot of the RMS Error convergence rate for this simulation in Fig. 5.5. Clearly the error decreases by a very large amount in the first update and then decreases by only small amounts in subsequent updates, showing the quick convergence and effectiveness of the data assimilation technique to account for errors in the initial fire perimeter.



(a) (b) (c)

Figure 5.4: Comparison of Updated Perimeter $\hat{x}_{k|k}$ and Reference Perimeter x_k for time steps (a) $k = 1$ (b) $k = 2$ and (c) $k = 3$. The circles indicate the 99% confidence interval (3 times standard deviation of 50m or variance of 2500m^2) of the observations.

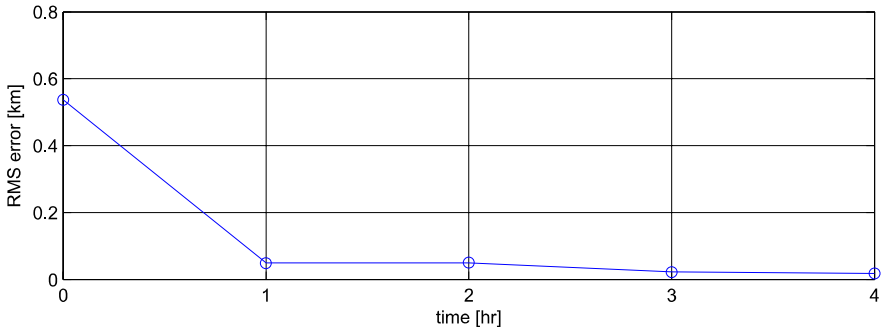


Figure 5.5: RMS Error Between the Updated Perimeters $\hat{x}_{k|k}$ and the "true" Fire Perimeters x_k with hourly updates and a measurement variance of 2500m^2. It is assumed the first measurement and data assimilation update is performed at $k = 1$.

5.4 Effect of Measurement Uncertainty on Convergence Rate

For the hourly updates, the convergence was shown to occur in only a single update step of the EnKF algorithm, despite the error on the initial fire perimeter. The fast convergence is due to the relatively large covariance on the initial fire perimeter and relatively small covariance on the fire perimeter observations. Next we investigate the effect of varying this uncertainty in the measurements of the reference fire perimeter on the convergence rate. The simulation is initialized with a smaller initial uncertainty to see a more pronounced effect of this variation. We apply the same covariance matrix on the coordinates of the mean perimeter and then give a smaller 2×2 covariance matrix for the center of the ignition boundary compared to the previous simulation.

$$P_{0|0}^{\bar{x}} = diag\{P, P, P, P\}, \quad P = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, \quad P_{0|0}^{x_c} = \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}$$

For the first simulation we update the perimeters with a time resolution of 1 hour but with measurements, $y_{k+1} + v_{k+1}^{50}$, where v_{k+1}^{50} is zero mean white noise with 2500m^2 variance. Whereas for the second simulation we use noisier measurements, $y_{k+1} + v_{k+1}^{200}$, where v_{k+1}^{200} is zero mean white noise with a 40000m^2 variance. A comparison of the respective RMS progression curves can be found Fig. 5.6.

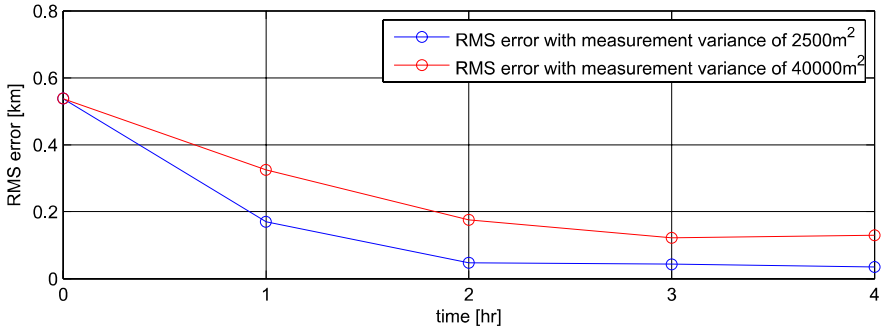


Figure 5.6: Comparison of RMS Errors between data assimilation using measurements of 2500m^2 variance and measurements of 40000m^2 variance

From Fig. 5.6 we can clearly see that the simulation that uses fire perimeter observations with a 40000m^2 variance would take a larger number of steps to converge to the same error when compared to the simulation which uses more accurate fire perimeter observations with a 2500m^2 variance.

5.5 Data Assimilation with Reduced Update Frequency

We repeat the procedure with the same initial fire perimeter and covariance matrices as used in the simulation presented in Section 5.3. For performance evaluation we now plot the (mean) RMS Error *and* the variance of the RMS Error during the course of the simulation. In this case the assimilation steps occur do not occur hourly, but are further apart. Compared to Figure 5.4, data assimilation is only done at the time steps $k = 1$ and $k = 4$ instead of every hourly time step. The results in Figure 5.7 show the effect of both the mean and variance of the RMS error and indicate a significant drop in the mean RMS error and its variance whenever a data assimilation step is performed. Even though the mean RMS error does not increase considerably when no data assimilation step is performed, the uncertainty does increase significantly during time steps without data assimilation. This large uncertainty in the RMS error informs us that even though the (mean) RMS error itself may remain small between the reference and offset perimeters, the uncertainty on the fire perimeter may grow without frequent data assimilation steps.

Chapter 5 is, in full, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering in Procedia Computer Science Vol 80 (2016). Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

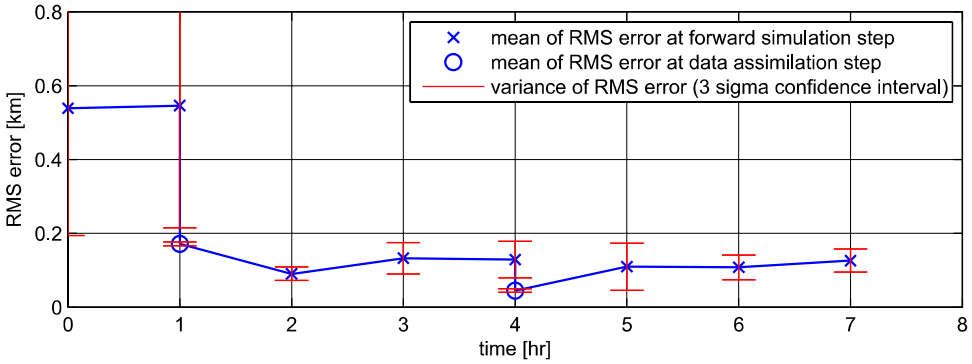


Figure 5.7: Progress in mean RMS Error and variance of RMS Error when data assimilation steps are performed only at the time steps $k = 1$ and $k = 4$ hours.

Chapter 6

Data Assimilation with Adjustment Factors

6.1 Forward Model with Adjustment Factors

Apart from the various parameters and inputs such as wind, weather and topography as mentioned earlier, FARSITE also provides us with Adjustment Factors in order to tune the simulation. These factors are fuel model specific numbers which can alter the rate of spread of the fire. In most fire simulations which occur over long spaces and periods of time FARSITE tends to over predict the fire spread due to the coarse nature of the spatial and temporal data available to us. Thus over these longer periods it tends to use average values of wind speed to predict the average rate of spread. But due to the non-linear relationship between these variables this prediction cannot be expected to work accurately. On the other hand spread rates can also be under predicted because of error in fuel models, inaccurate fuel moistures and improperly represented local winds. Hence these factors provide us with a very efficient way and extra degree of freedom to address these shortcomings and improve the accuracy of the fire spread predictions. The Adjustment Factors are generally assumed to be constant and their values decided based on the judgement/experience of the user. However by augmenting these adjustment factors to our state we would be able to constantly adjust these values

in an optimal manner by assimilating measurements of the fire perimeter.

In this section for the purpose of data assimilation and improvement of fire perimeter prediction, we combine the fire perimeter and adjustment factors in a single state vector

$$\hat{z}_{k|k} = [\hat{x}_{k|k}\ \hat{\alpha}_{k|k}]^T$$

rewriting (2.1) in the format

$$\begin{aligned} \hat{z}_{k+1|k} &= g(\hat{z}_{k|k}, \theta, u_k) \\ \hat{y}_{k+1|k} &= \begin{bmatrix} C_{k+1|k} & 0 \end{bmatrix} \hat{z}_{k+1|k} \end{aligned} \tag{6.1}$$

for forward prediction of the "true" fire perimeter x_k by the forward simulation $\hat{x}_{k+1|k}$. Where $\hat{z}_{k+1|k} = g(\cdot)$ is implemented as,

$$\begin{aligned} \hat{x}_{k+1|k} &= f(\hat{x}_{k|k} + w_k^x, \hat{\alpha}_{k|k} + w_k^\alpha, \theta, u_k) \\ \hat{\alpha}_{k+1|k} &= h(\hat{\alpha}_{k|k} + w_k^\alpha, k) \end{aligned} \tag{6.2}$$

where $f()$ was given in (2.1) and the adjustment factors are kept the same during a forward simulation, e.g. $h(\hat{\alpha}_{k|k}, k) = 1$. It should be noted that both $\hat{x}_{k|k}$ and $\hat{\alpha}_{k|k}$ will be updated via a data assimilation step in between the forward simulation steps. The (measured) output y_{k+1} in (6.1) refers to a spatially downsampled (coarse) measurement of the actual fire perimeter just as before.

6.2 Reference Data for Simulation with Adjustment Factors

As in Section 5.1 we use the same parametric conditions and input data to produce the state z_{k+1} which contains the "true" fire perimeters x_{k+1} and "true" adjustment factors α_{k+1} that will be used as a reference for the performance evaluation of including adjustment factors in data assimilation tools in this section. To generate the reference data for the EnKF-based data assimilation for FARSITE, fuel type number 5 is the most dominant over the topography covering the May

2014 Cocos Fire in San Marcos, CA. For the purpose of illustrating fuel adjustment data assimilation, the entries of the (time dependent) vector of adjustment factors α_k in (6.6) are all set to a 1, except for the 5th entry $\alpha_{k+1,5}$ that is set to $\alpha_{k+1,5} = 0.5$ and assumed to be unknown. As a result, $\alpha_k = [1_{1\times 4}\ 0.5\ 1_{1\times 8}]^T\ \forall k$ and the simulated fire perimeters are depicted in Fig. 5.1 over a 18 hour time period with a one hour time resolution and a 90m spatial resolution along the perimeter, starting from an initial square 30m×30m ignition boundary x_0 at $t = k = 0$.

For the purposes of combined fire perimeter and fuel adjustment factor data assimilation, we will include only the 5th entry $\alpha_{k+1,5}$ of the full size adjustment factor vector α_{k+1} along with the fire perimeters x_{k+1} in the state z_{k+1} and leave the remaining adjustment factors as inputs to the forward model. So our state is now given by

$$z_{k+1} = \begin{bmatrix} x_{k+1} \\ \alpha_{k+1,5} \end{bmatrix}$$

and the data assimilation will attempt to estimate the full state z_k which contains x_k representing the fire perimeter and a scalar adjustment factor α_k representing only the 5th (unknown) entry of the fuel adjustment factors. For testing the data assimilation tools, the "true" data with a 90m spatial resolution along the perimeter is down-sampled to generate the output y_{k+1} given by

$$y_{k+1} = [C_{k+1}\ \ 0]\, z_{k+1} + v_{k+1} \tag{6.3}$$

where C_{k+1} is a downsampling matrix, and measurements y_{k+1} are produced at a spatial resolution of only 360m along the perimeter. The measurements are perturbed by a white noise v_{k+1} with a standard deviation of 50m. As before, it is assumed that the initial state z_0 is not known at the start of the data assimilation procedure. Instead, an estimate $\hat{z}_{0|0} \neq z_0$ along with a covariance matrix $P^z_{0|0}$ is used. In the EnKF context these are defined as follows, The mean $\bar{z}_{0|0} = \hat{z}_{0|0}$ at $k = 0$ is the state which describes the ignition boundary ($\bar{x}_{0|0}$) in eastern e_i and northern n_i coordinates as well as the values of the initial adjustment factors ($\bar{\alpha}_{0|0}$)

for the fuel models present.

$$\bar{z}_{0|0} = \begin{bmatrix} \bar{x}_{0|0} & \bar{\alpha}_{0|0} \end{bmatrix}^T \tag{6.4}$$

$$\bar{x}_{0|0} = \begin{bmatrix} e_1 & n_1 & e_2 & n_2 & \cdots & e_m & n_m \end{bmatrix}^T \tag{6.5}$$

$$\bar{\alpha}_{0|0} = \begin{bmatrix} a_1 & a_2 & \cdots & a_l \end{bmatrix}^T \tag{6.6}$$

where (e_j, n_j) is the jth coordinate of the ignition boundary, a_i is the adjustment factor for the ith fuel model.

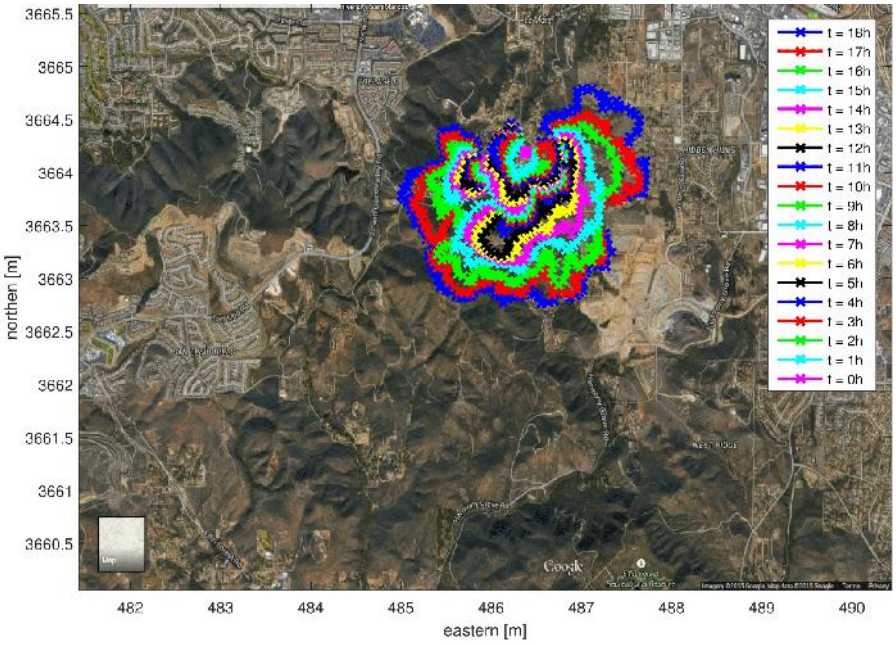


Figure 6.1: Simulation of the use case reference data: hourly noise free "true" fire perimeters x_k from the "true" state $z_k = [x_k \ \alpha_k]^T$ where x_0 is an initial square 30m×30m ignition boundary and $\alpha_k = [1_{1\times4} \ 0.5 \ 1_{1\times8}]^T \ \forall k$.

6.3 Data Assimilation Setup

Using noisy and down-sampled measurements y_{k+1} at time index $k+1$, an estimate $\hat{z}_{k+1|k+1}$ is formulated of the "true" state z_{k+1} via the ensemble Kalman filter (EnKF) approach outlined earlier in Section 4. To initialize the data assimilation procedure, an inaccurate value of the initial state $\hat{z}_{0|0} \neq z_0$ as described

earlier will be used, where $\hat{z}_{0|0}$ contains the initial fire perimeter estimate $\hat{x}_{0|0}$ an 8×1 vector which is a square 30 × 30m fire perimeter, but with the center of the perimeter offset and the 5th entry of the initial adjustment factor $\hat{\alpha}_{(0|0),5}$ off by 0.5.

Having described the mean of the initial state

$$\bar{z}_{0|0} = \hat{z}_{0|0} = \left[\hat{x}_{0|0}\ \hat{\alpha}_{(0|0),5}\right]^T$$

we now need to define the covariance matrix associated with this mean. These matrices that were described in Section 4.1 are given by

$$P_{0|0}^{\bar{x}} = diag\left\{P, P, P, P\right\}, \quad P = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \tag{6.7}$$

and

$$P_{0|0}^{x_c} = \begin{bmatrix} 150 & 0 \\ 0 & 150 \end{bmatrix} \tag{6.8}$$

The covariance matrices in (6.7) and (6.8) are used in (4.4) to get the covariance matrix of the initial fire perimeter $P_{0|0}^{x}$. The uncertainty of the 5th entry of the adjustment factor is described by a variance $\sigma_\alpha^2 = 0.04$. Combining the two we get covariance of the initial state $\hat{z}_{0|0}$ as

$$P_{0|0}^{z} = \begin{bmatrix} P_{0|0}^{x} & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \tag{6.9}$$

6.4 Effect of Adjustment Factors on Data Assimilation with Reduced Update Frequency

The use of an EnKF-based data assimilation for FARSITE where the state z_k contained only the fire perimeter x_k has shown siginificant fire perimeter prediction performance improvement in the previous sections. In this case, the "true" fire perimeter x_k is not only changing over time from the initial (unknown) x_0, but also due to the possibly time varying behavior of the fuel adjustment factors contained in α_k. To analyze the implications of including adjustment factors in

the state, two data assimilation approaches are compared:

- Approach A: partial data assimilation. For data assimilation purposes, the state is defined as $\hat{z}_{0|0} = \left[\hat{x}_{0|0}\right]$ only containing the fire perimeter, for which the covariance is given by $P^z_{0|0} = P^x_{0|0}$.

- Approach B: full data assimilation. In this data assimilation approach the state is defined as $\hat{z}_{0|0} = [\hat{x}_{0|0}\ \hat{\alpha}_{(0|0),5}]^T$ as a combination of both fire perimeter and fuel adjustment factor, for which the mean and covariance were defined in Section 6.3.

For both data assimilation approaches we provide the same measurements y_{k+1} and conduct update steps every 4 hours starting from $k = 1$. For performance comparison we compute the mean RMS error on the estimation of the fire perimeter x_k during the course of the simulations and the results are summarized in Fig. 6.2.

The results in Fig. 6.2 illustrate two important observations with respect to the performance of the data assimilation:

- First of all, it can be seen that the mean RMS error on the estimation of the fire perimeter x_k slowly increases at the hourly time steps when no data assimilation is performed. Only at data assimilation steps performed every 4th hourly time step the mean RMS error of the state estimation drastically drops to a lower value, illustrating the performance of the EnKF-based data assimilation for FARSITE.

- Secondly, it can be observed from Fig. 6.2 that after a few data assimilation steps, the mean RMS error on the estimation of the fire perimeter x_k during assimilation with fuel adjustment factors (Approach B) is always less than the mean RMS error of the state without adjustment factors (Approach A). From Fig. 6.3 we see that the adjustment factor initially increases to compensate for a large initial error in the initial fire perimeter, but decreasing afterwards. This phenomena causes the mean RMS error in data assimilation with adjustment factors to be higher initially.

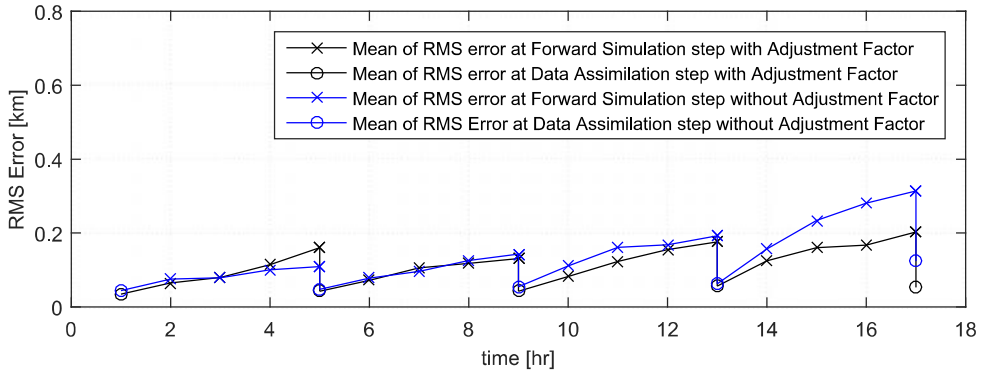


Figure 6.2: Variation in mean RMS Error for assimilation with adjustment factors versus assimilation without adjustment factors. The updates are carried out every 4 hours after $k = 1$.

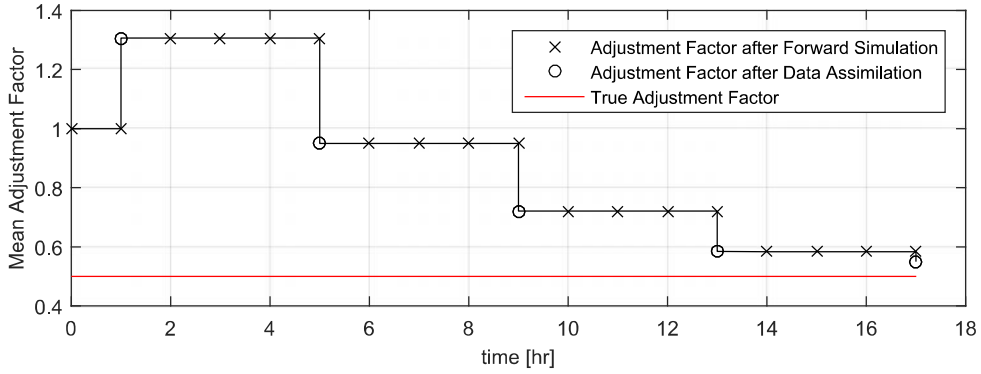


Figure 6.3: Mean Adjustment Factor versus True Adjustment Factor is plotted over the period of the simulation. As mentioned before the adjustment factor is assumed to be a constant over forward simulation time steps.

As a final note it can be observed that the updates are much more effective in decreasing the state RMS error in data assimilation with adjustment factors (Approach B) as we improve our estimate of the adjustment factor over time. This improvement tends to grow over the course of the data assimilation.

Chapter 6 has, in part, been submitted for publication of the reformatted material as it may appear in American Control Conference 2017, Srivas, Thayjes; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this material.

Chapter 7

Adjustment Factor Tracking using Data Assimilation with Hourly Updates

Along with reducing the (mean RMS error) on the estimation of the fire perimeter x_k, it is also important to keep track of any (time varying) fuel adjustment factor α_k. For that pupose, the EnKF-based data assimilation for FARSITE can be used to track the "true" 5th entry $\hat{\alpha}_{0,5} \neq \alpha_{0,5}$ of the adjustment factor vector α_k using hourly down-sampled noisy measurements of the true fire perimeter. For tracking time varying behavior of the fuel adjustment factor $\alpha_{0,5}$ is modeled as a random walk (Fig 7.1 as below,

$$\alpha_{k+1,5} = \alpha_{k,5} + w_k \tag{7.1}$$

where w_k is white noise with zero mean and standard deviation $\sigma_w = 0.1$ and α_0 =0.5.

We model the forward simulation of the estimated adjustment factor in the same manner as the true adjustment factor :

$$\hat{\alpha}_{k+1|k,5} = \hat{\alpha}_{k,5} + w_k \tag{7.2}$$

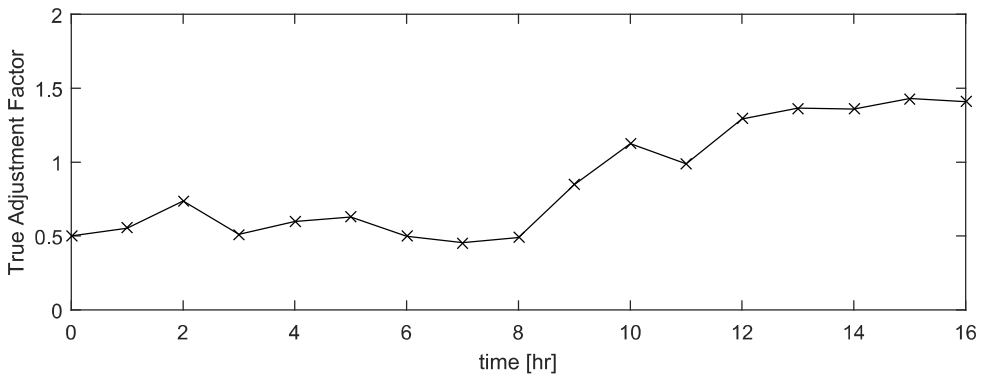


Figure 7.1: The "true" 5th entry of the Adjustment Factor vector $\alpha_{k+1,5}$ modeled as a random walk. We track this using the down-sampled noisy measurements y_{k+1}

7.1 Effect of Initial Error on Tracking

To illustrate the performance of the EnKF-based data assimilation for FARSITE, initial errors in either fire perimeter and/or the adjustment factor is assumed according the the following two scenarios:

1. Scenario 1: full initial state error. The initial error in the entire initial state $\hat{z}_{0|0} = [\hat{x}_{0|0}\ \hat{\alpha}_{(0|0),5}] \neq [x_0\ \alpha_{0,5}]$ as in Section 6.3.

2. Scenario 2: only initial fuel adjustment factor error. Initialization of the state is done via $\hat{z}_{0|0} = [x_0\ \hat{\alpha}_{(0|0),5}]$ with $\hat{\alpha}_{(0|0),5} \neq \alpha_{0,5}$, thus removing any initial error on the initial fire perimeter in the state.

We apply the same uncertainty for both simulations as in Section 6.3 and follow Approach B: full data assimilation. For performance comparison we compare the "true" scalar 5th entry $\alpha_{k,5}$ of the adjustment factor vector α_k with the estimated $\hat{\alpha}_{k,5}$ for both scenarios listed above and the results are summarized in Fig. 7.2.

As intuitively expected and confirmed by the results in Fig. 7.2, the tracking of the (5th entry of the) fuel adjustment factor $\alpha_{k,5}$ is more accurate for Scenario 2 (bottom figure) with only initial fuel adjustment factor error $\hat{z}_{0|0} = [x_0\ \hat{\alpha}_{(0|0),5}]$, $\hat{\alpha}_{(0|0),5} \neq \alpha_{0,5}$. It can also be observed that the estimate of (5th entry of the) fuel adjustment factor $\alpha_{k,5}$ for Scenario 1 (top figure) maintains an error at the first

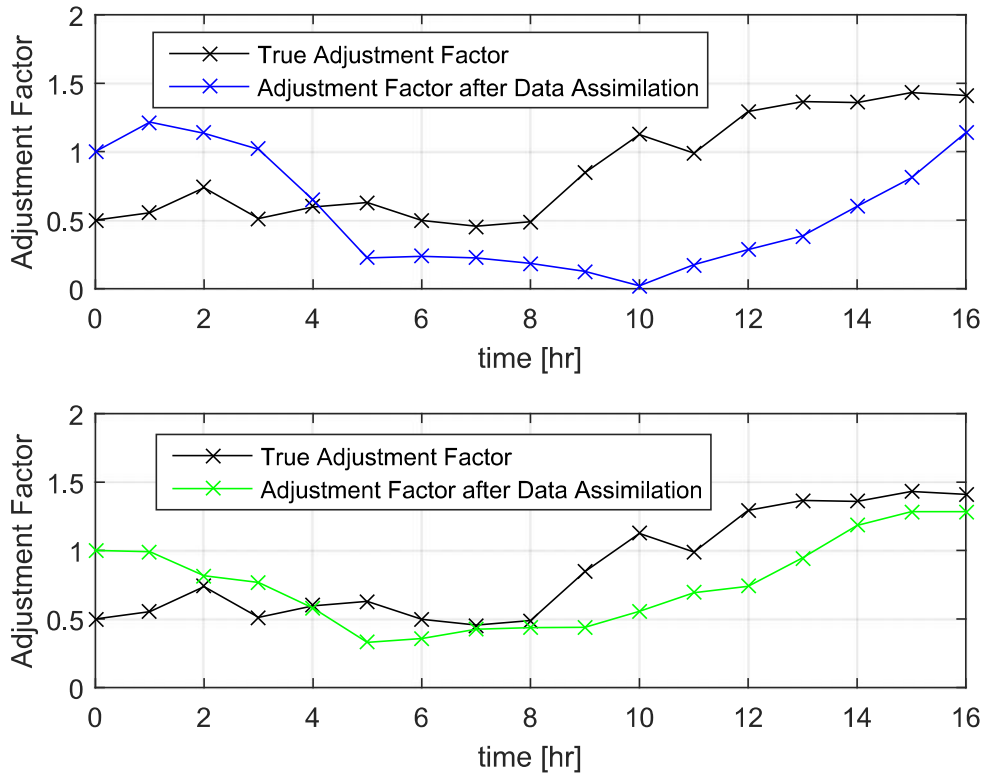


Figure 7.2: Time traces of "true" scalar adjustment factor $\alpha_{k,5}$ and estimated scalar adjustment factor $\hat{\alpha}_{k,5}$ for hourly data assimilation steps. Top figure: full state error, Bottom Figure: only fuel adjustment factor error.

few hourly data assimilation steps. The value of the fuel adjustment factor $\hat{\alpha}_{(k|k),5}$ actually increases (becomes larger than 1) to accommodate the large initial error $\hat{x}_{0|0} \neq x_0$ to "accelerate" the fire growth and reduce the errors on the estimate on the fire perimeter x_k.

7.2 Effect of Measurement Uncertainty on Tracking

The variance of the noise v_{k+1} present on the perimeter measurement y_{k+1} of the "true" state z_{k+1} in (6.3) has a direct effect on the size of the Kalman gain K_{k+1} given in (4.5). A Smaller variance of the noise v_{k+1} tends to give a matrix $P^y_{k+1|k}$ with a smaller (minimal) singular values, creating a Kalman gain K_{K+1} with

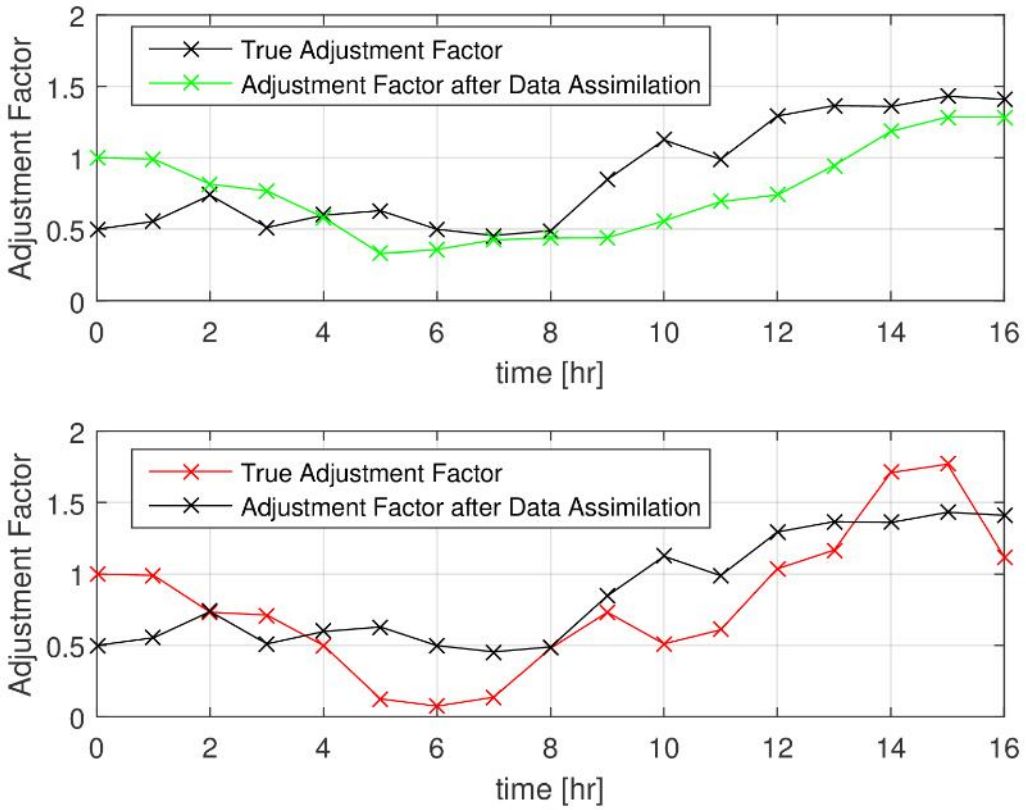


Figure 7.3: Time traces of "true" scalar adjustment factor $\alpha_{k,5}$ and estimated scalar adjustment factor $\hat{\alpha}_{k,5}$ for hourly data assimilation steps using different variance of noise on the measurements.

larger numerical values. A Kalman gain with larger numerical values gives rise to a stronger weighting of the measurements in the update step of $z^i_{k+1|k+1}$. This effect can be easily illustrated with the EnKF-based data assimilation for FARSITE by comparing two data assimilation scenarios with identical initial state error (only in the adjustment factor), but with a diference in the variance the noise v_{k+1} present on the perimeter measurement y_{k+1}.

The results are confirmed in Fig. 7.3 where we observe that for a variance of v_k of 2500m^2 on the measurements, we see gradual updates in the estimated adjustment factor and the tracking of $\alpha_{k,5}$ is much slower. On the other hand, for a measurement variance of 100m^2 we see that the updates of $\hat{\alpha}_{k,5}$ are much faster and follow the trends in $\alpha_{k,5}$ more closely.

Chapter 7, in part, has been submitted for publication of the reformatted material as it may appear in American Control Conference 2017, Srivas, Thayjes; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this material.

Chapter 8

Conclusions

Data assimilation in FARSITE is accomplished by characterizing both simulated and actual measured fire perimeter with a mean and covariance matrix (confidence regions) to formulate optimal updates for the prediction of the spread of the wildfire. Optimal updates are computed via a fire perimeter adjustment, weighted by a Kalman filter gain that is computed via an Ensemble Kalman filter approach. Application of the proposed FARSITE data assimilation to a wildfire simulation representing the 2014 Cocos fire confirmed an inverse relation between the rate of convergence of the fire perimeter and the uncertainty on the fire perimeter measurements. In the presence of incorrect ignition boundary, it is shown that convergence to the actual wild fire perimeter is obtained in only a few data assimilation steps in case of a relatively small (50m) standard deviation on the measured fire perimeters measured at a resolution of 360m. The simulation study also includes results on convergence for larger uncertainty in the measured fire perimeters and when data assimilation steps are not performed regularly. In the final chapters it is shown how the presence of an adjustment factor can improve the RMS error of the fire perimeter estimate, even for reduced hourly updates during the data assimilation steps. The same reference data of the May 2014 Cocos fire is used to show performance of the EnKF-based data assimilation for FARSITE for tracking time varying fuel adjustment factors. It is also shown that the noise variance on the limited spatial resolution observations of the fire perimeter influences the estimate of the fuel adjustment factor in an intuitive way. In summary we observed that

we can successfully track the adjustment factor using only noisy down sampled measurements of the wildfire perimeter and noticed that this tracking improved with reduced initial error and lower measurement uncertainty.

Chapter 8 is, in most part, a reformatted reprint of the material as it appears in Wildfire Spread Prediction and Assimilation in FARSITE using Ensemble Kalman Filtering in Procedia Computer Science Vol 80 (2016). Srivas, Thayjes; Artes, Tomas; de Callafon, Raymond A.; Altintas, Ilkay. The thesis author was the primary investigator and author of this paper.

Bibliography

[1] F.A. Albini. Estimating wildfire behavior and effects. Technical Report GTR-INT-30, Dept. of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT, 1976.

[2] H. E. Anderson. Aids to determining fuel models for estimating fire behavior. Technical Report GTR-INT-122, U.S. Dept. of Agriculture, Forest Service, Intermountain Forest Range and Experiment Station, Ogden, UT, 1982.

[3] J.D. Beezley and J. Mandel. An ensemble Kalman-particle predictor-corrector filter for non-gaussian data assimilation. *Lecture Notes in Computer Science*, 5545:470–478, 2009.

[4] L.S Bradshaw, J.E. Deeming, R.E. Burgan, and J.D. Cohen. The 1978 national fire-danger rating system: technical documentation. Technical Report GTR-INT-169, U.S. Dept. of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT, 1984.

[5] G. Evensen. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.

[6] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag, Berlin, 2009.

[7] M.A. Finney. FARSITE: Fire area simulator-model development and evaluation. Technical Report RMRS-RP-4 Revised, U.S. Dept. of Agriculture, Forest Service, Rocky Mountain Research Station, 2004.

[8] R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227–255, 2007.

[9] S. Gillijns, O.B. Mendoza, J. Chandrasekar, B.L.R. De Moor, D.S. Bernstein, and A. Ridley. What is the ensemble Kalman filter and how well does it work? In *American Control Conference*, pages 4448–4453, Minneapolis, MN, 2006.

[10] T. Hansen, P. Yalamanchili, and H-W Braun. Wireless measurement and analysis on HPWREN. In *Proceedings of Passive and Active Measurement Workshop*, pages 222–229, Fort Collins, CO, 2002.

[11] J. Mandel, J.D. Beezley, L. Cobb, and A. Krishnamurthy. Data driven computing by the morphing fast Fourier transform ensemble Kalman filter in epidemic spread simulations. *Procedia Computer Science*, 1:1221–1229, 2010.

[12] J. Mandel, J.D. Beezley, J.L. Coen, and M. Kim. Data assimilation for wildland fires: Ensemble Kalman filters in coupled atmosphere-surface models. *IEEE Control Systems Magazine*, 29:47–65, 2009.

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

[14] J. Mandel, L.S. Bennethum, M. Chen, J.L. Coen, C.C. Douglas, L.P. Franca, C.J. Johns, M. Kim, A.V. Knyazev, R. Kremens, V. Kulkarni, G. Qin, A. Vodacek, J. Wu, W. Zhao, and A. Zornes. Towards a dynamic data driven application system for wildfire simulation. *Lecture Notes in Computer Science*, 3515:632–639, 2005.

[15] J. Mandel, M. Chen, L.P. Franca, C. Johns, A. Puhalskii, J.L. Coen, C.C. Douglas, R. Kremens, A. Vodacek, and W. Zhao. A note on dynamic data driven wildfire modeling. *Lecture Notes in Computer Science*, 3038:725–731, 2004.

[16] W. Mell, M. A. Jenkins, J. Gould, and P. Cheney. A physics-based approach to modelling grassland fires. *International Journal of Wildland Fire*, 16(1):1–22, 2007.

[17] R.M. Nelson. Prediction of diurnal change in 10-h fuel stick moisture content. *Canadian Journal of Forest Research*, 30:1071–1087, 2000. doi:10.1139/CJFR-30-7-1071.

[18] J. Otero, P. Yalamanchili, and H.-W. Braun. High performance wireless networking and weather. http://hpwren.ucsd.edu/info/images/weather.pdf, 2001.

[19] 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. *Nat. Hazards Earth Syst. Sci.*, 15, 2015.

[20] 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. *Nat. Hazards Earth Syst. Sci.*, 14:2951–2973, 2014.

[21] R.C. Rothermel. A mathematical model for predicting fire spread in wildland fuels. Technical Report RP-INT-115, U.S. Dept. of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT, 1972.

[22] R.C. Rothermel. Predicting behavior and size of crown fires in the northern Rocky Mountains. Technical Report RP-INT-438, U.S. Dept. of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT, 1991.

[23] J.H. Scott and R.E. Burgan. Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model. Technical Report RMRS-GTR-153, U.S. Dept. of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 2005.

[24] S. M. Stein, J. Menakis, M.A. Carr, S. J. Comas, S. I. Stewart, H. Cleveland, L. Bramwell, and V. C. Radeloff. Wildfire, wildlands, and people: understanding and preparing for wildfire in the wildland-urban interface-a forests on the edge report. 2013.

[25] C.E. Van Wagner. Conditions for the start and spread of crownfire. *Canadian Journal of Forest Research*, 7(1):23–34, 1977. 10.1139/x77-004.

[26] H. Xue, F. Gu, and X. Hu. Data assimilation using sequential Monte Carlo methods in wildfire spread simulation. *ACM Transactions on Modeling and Computer Simulation*, 22(23):1–25, 2012.