CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Workshop - Traitement des données massives en mécanique des fluides

# Quantifying and reducing shape and topological uncertainties in front-tracking problems

Mélanie Rochoux★, Annabelle Collin, Didier Lucor & Philippe Moireau

UNCERTAINTY QUANTIFICATION

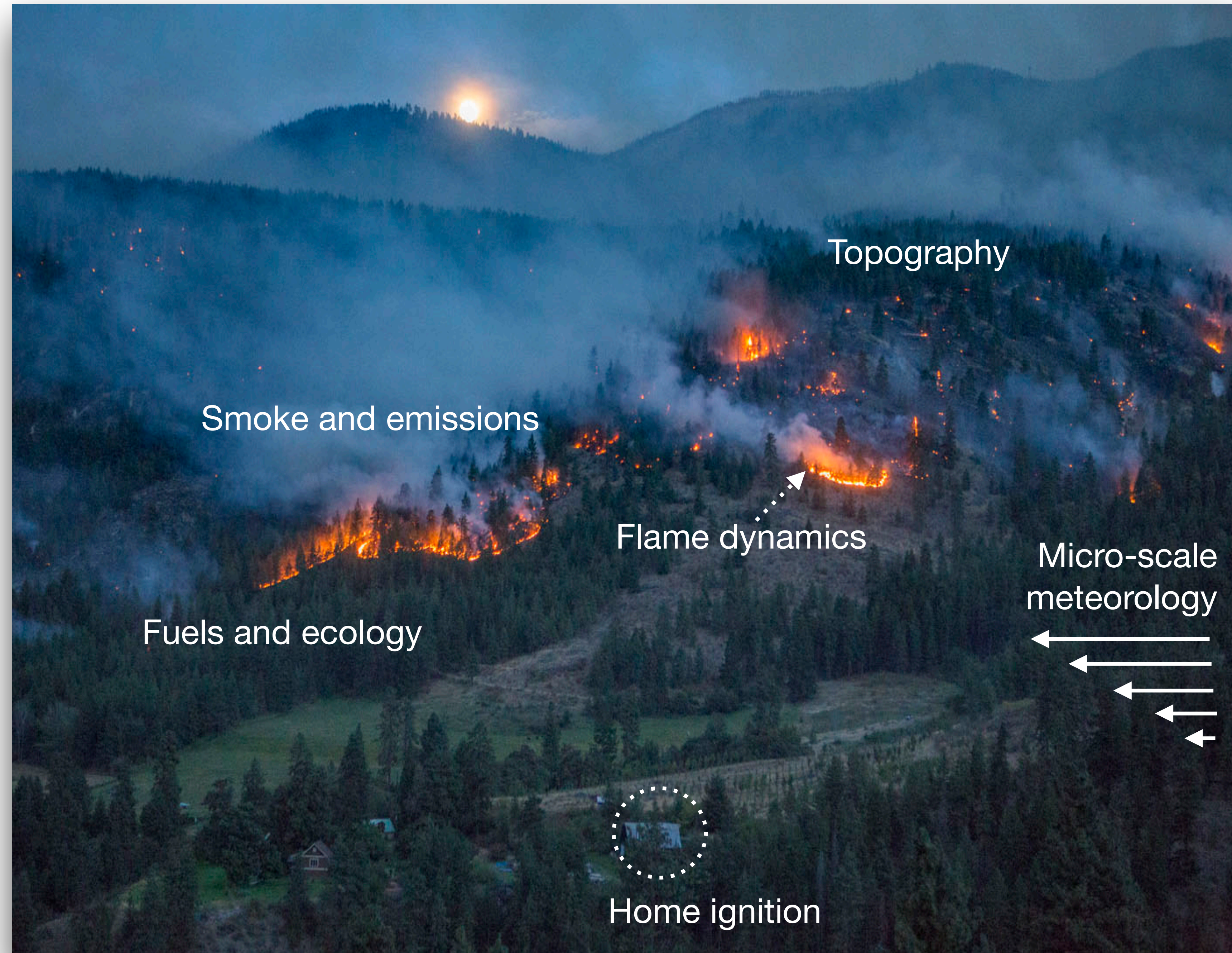
DATA ASSIMILATION

POSITION ERRORS

WILDFIRE

# Introduction

## The wildfire problem: Talk's guideline




### SAFETY ISSUE

Forecast wildfire behavior in real time

- help fire emergency responses
- strengthen firefighting actions

### WILDFIRE COMPLEXITY

Current models are far from being predictive

- wide range of length scales
- a fire creates its own weather (interactions between a fire and the near-surface atmosphere)
- poorly-defined biomass fuels

How can remote sensing help us to design on-demand high-fidelity simulations?

# Introduction

## The wildfire problem: Which modeling viewpoint for safety issue?


**ignition**
- 1 mm - 10 m
- seconds/minutes

spatio-temporal scales

**regional**
- 10-100 km
- hours/days

**BUOYANT FLAME**

✗ Too computationally expensive at large scale

**BURNING AREA**

✓Suitable for regional-scale fire simulations

**PLUME DISPERSION**

✗ No information on the combustion parameters

# Introduction

## Talk's outline

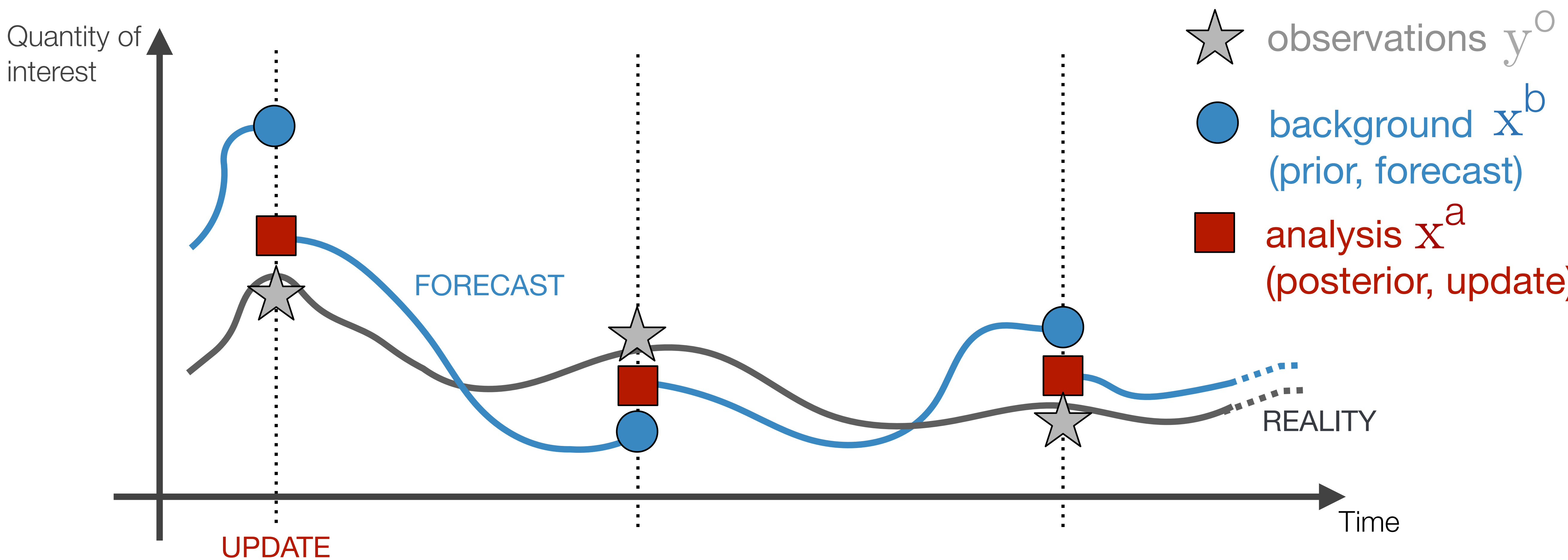
### (1) Data assimilation algorithm
- Ensemble-based Kalman filtering
- Link with uncertainty quantification methods

### (2) Position errors
- Amplitude errors versus Position errors
- New measure to quantity front shape similarity

### (3) Object-oriented data assimilation
- Data assimilation with front shape similarity measure
- Joint state-parameter estimation

The wildfire problem as guideline

# Principles of data assimilation

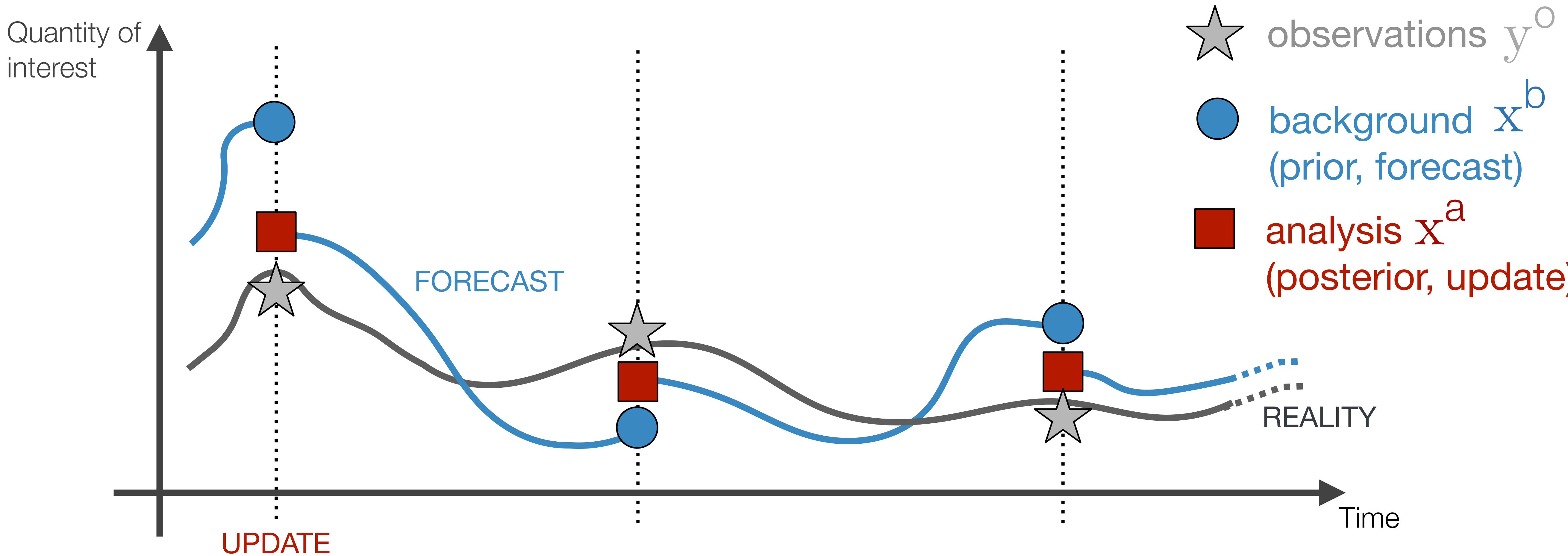



**The analysis minimizes the cost function**

$$\mathcal{J}(x) = ||\mathcal{G}(x) - y^o||^2_{R^{-1}} + ||x - x^b||^2_{B^{-1}} \rightarrow \mathcal{J}(x^a) = \min \mathcal{J}(x)$$

least-squares type data fitting functional with regularization term

# Principles of data assimilation



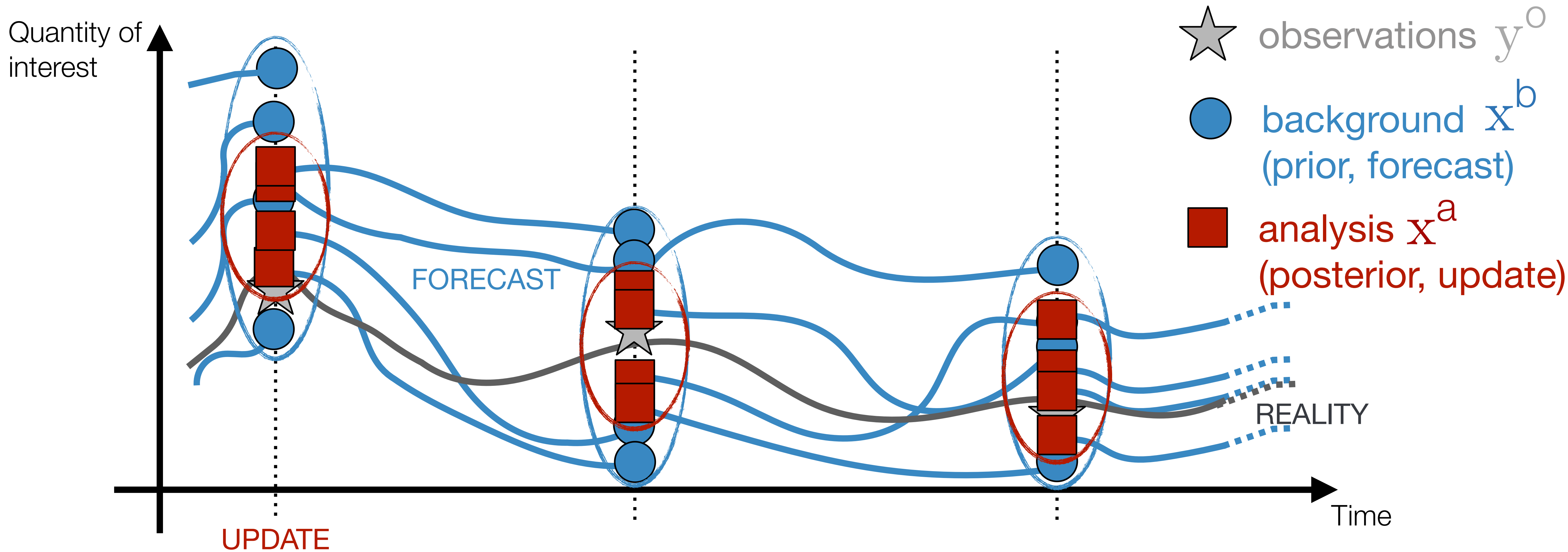

- Each source of information is weighted by its uncertainty
- Assumption of Gaussian error statistics
  - unbiased error
  - error covariance model
  - Euclidean-type norm $||x||^2_{B^{-1}} = x^T B^{-1} x$

**background error covariance matrix *B***

observation error covariance matrix *R*

→ analysis ≡ barycenter

# Ensemble-based Kalman filtering




**Kalman filters formulate the analysis as a correction of the background**

$$x^a = x^b + K\left(y^o - \mathcal{G}(x^b)\right) \rightarrow K = BG^T\left(GBG^T + R\right)^{-1}$$

linear combination of model simulations to find more optimal estimates

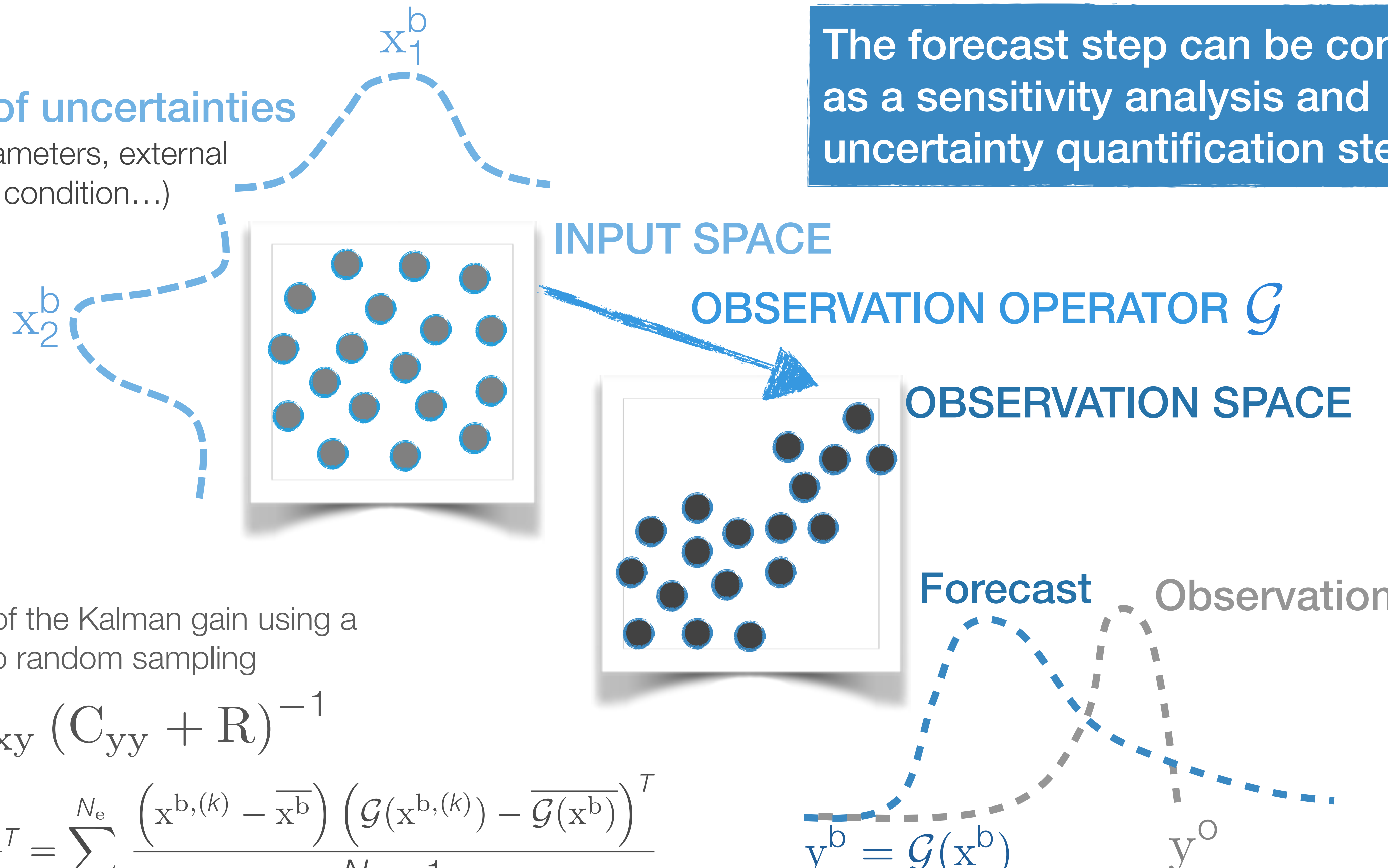
statistical estimation of the gain in the ensemble Kalman filter (EnKF)

# Ensemble-based Kalman filtering

**The forecast step can be considered as a sensitivity analysis and uncertainty quantification step**

## Sources of uncertainties

(physical parameters, external forcing, initial condition…)




Estimation of the Kalman gain using a Monte Carlo random sampling

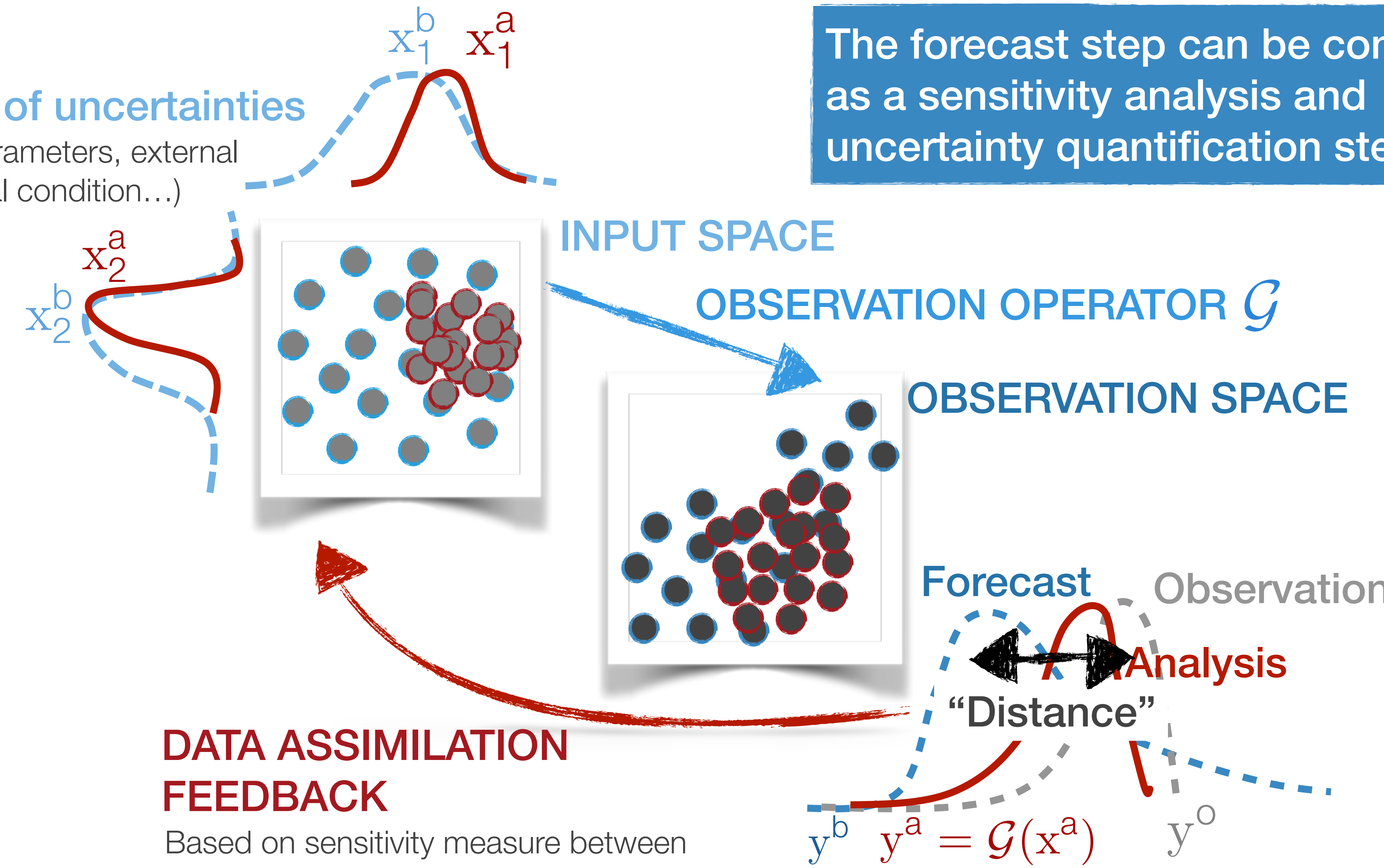
$$K = C_{xy}\,(C_{yy} + R)^{-1}$$

$$\mathbf{C}_{xy} = BG^T = \sum_{k=1}^{N_e} \frac{\left(x^{b,(k)} - \overline{x^b}\right)\left(\mathcal{G}(x^{b,(k)}) - \overline{\mathcal{G}(x^b)}\right)^T}{N_e - 1}$$

$$\mathbf{C}_{yy} = GBG^T = \sum_{k=1}^{N_e} \frac{\left(\mathcal{G}(x^{b,(k)}) - \overline{\mathcal{G}(x^b)}\right)\left(\mathcal{G}(x^{b,(k)}) - \overline{\mathcal{G}(x^b)}\right)^T}{N_e - 1}$$

# Ensemble-based Kalman filtering

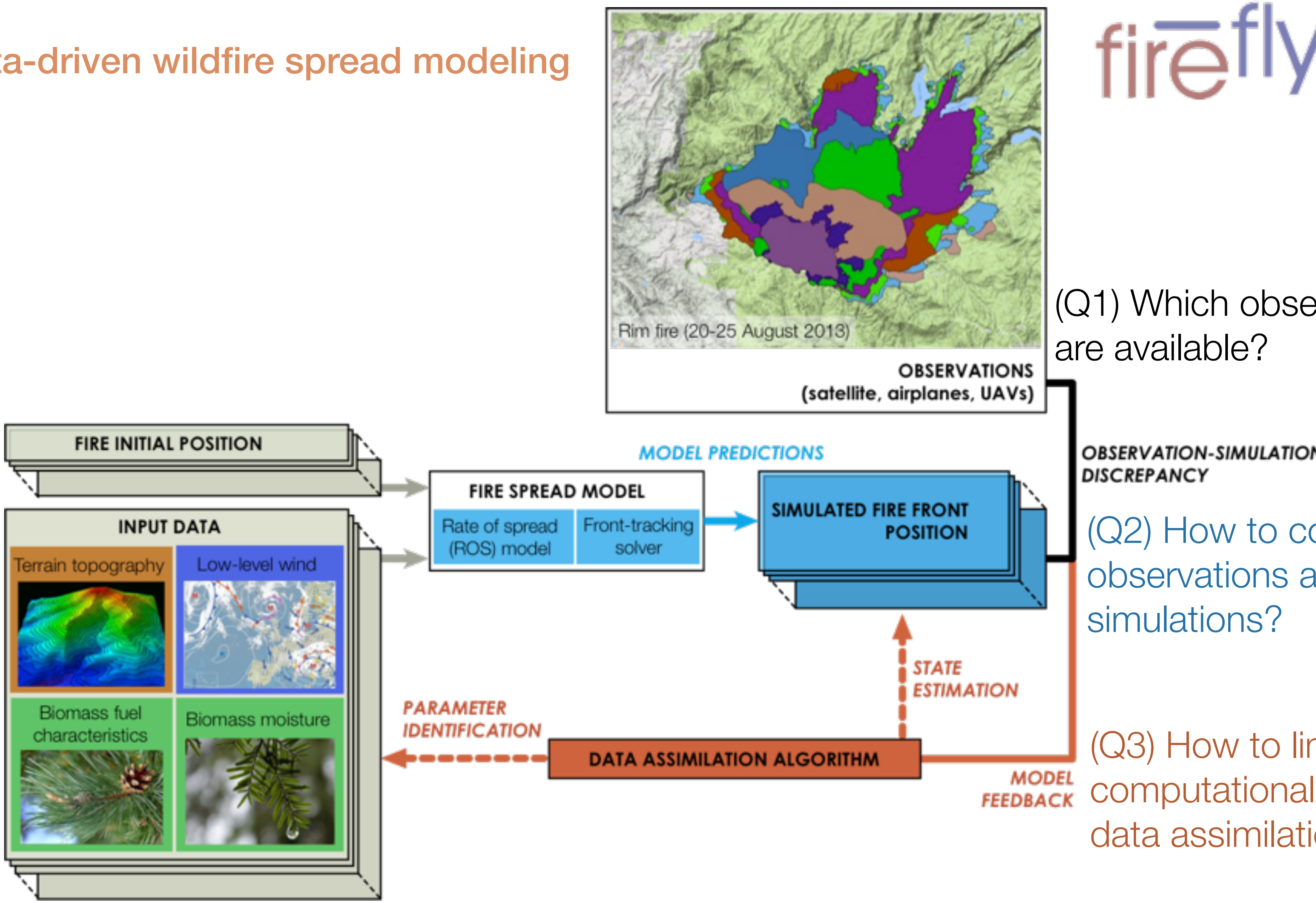
**The forecast step can be considered as a sensitivity analysis and uncertainty quantification step**




Based on sensitivity measure between the control variables and the observed quantities

# Wildfire guideline

## Data-driven wildfire spread modeling



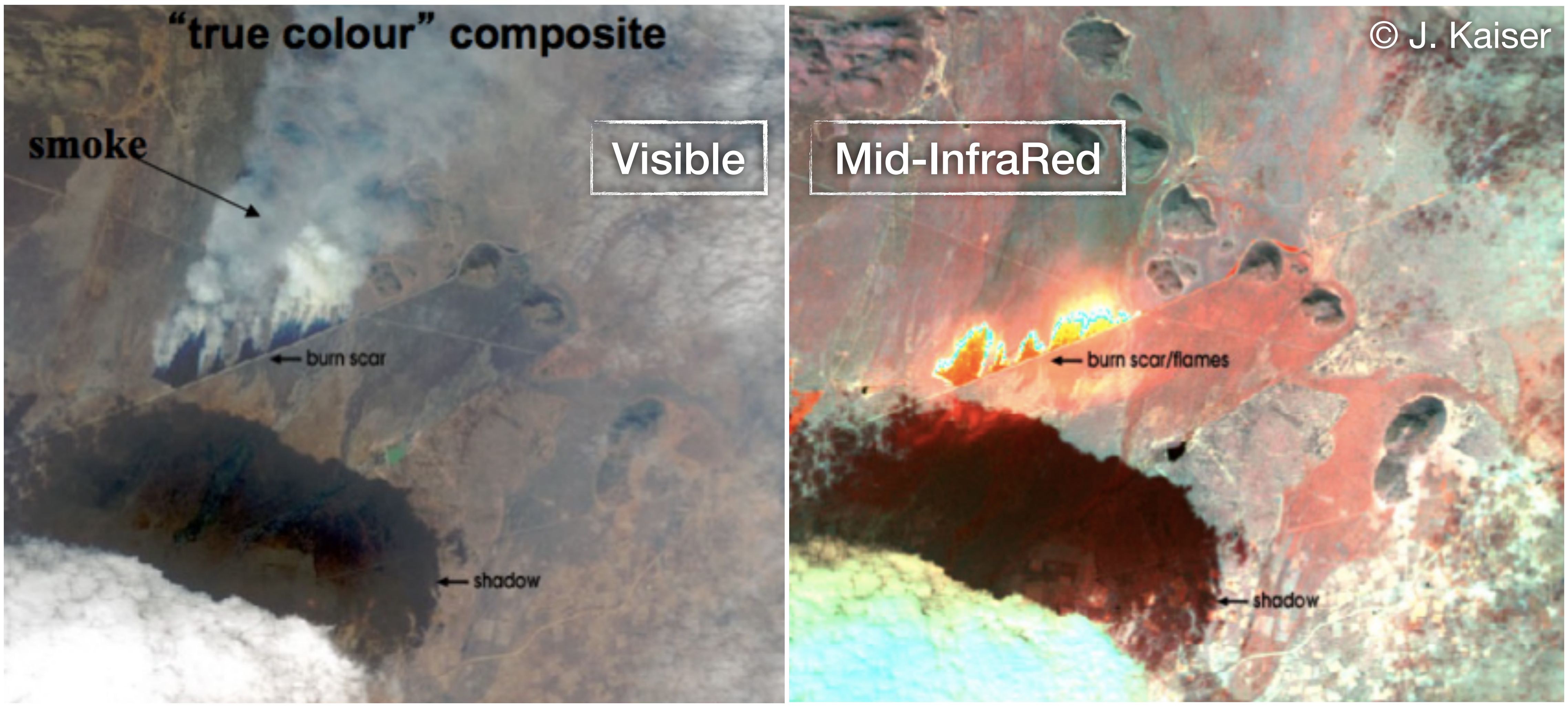

(Q1) Which observations are available?

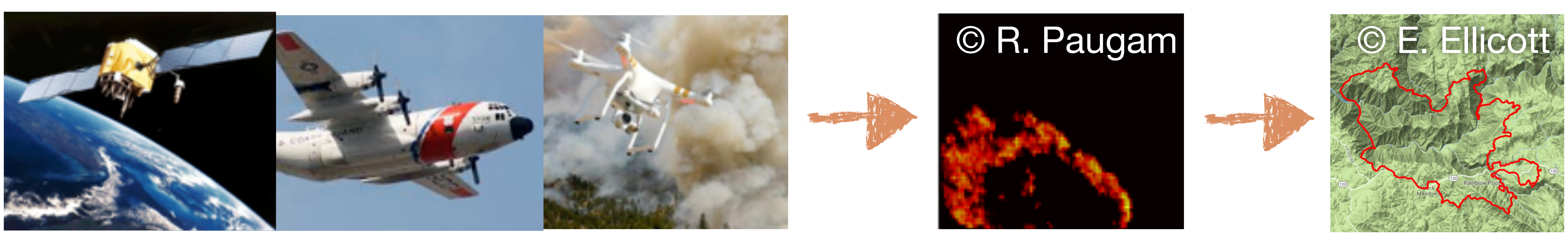
(Q2) How to compare observations and simulations?

(Q3) How to limit the computational cost of data assimilation?

# Wildfire guideline

## Mid-InfraRed (MIR) imagery

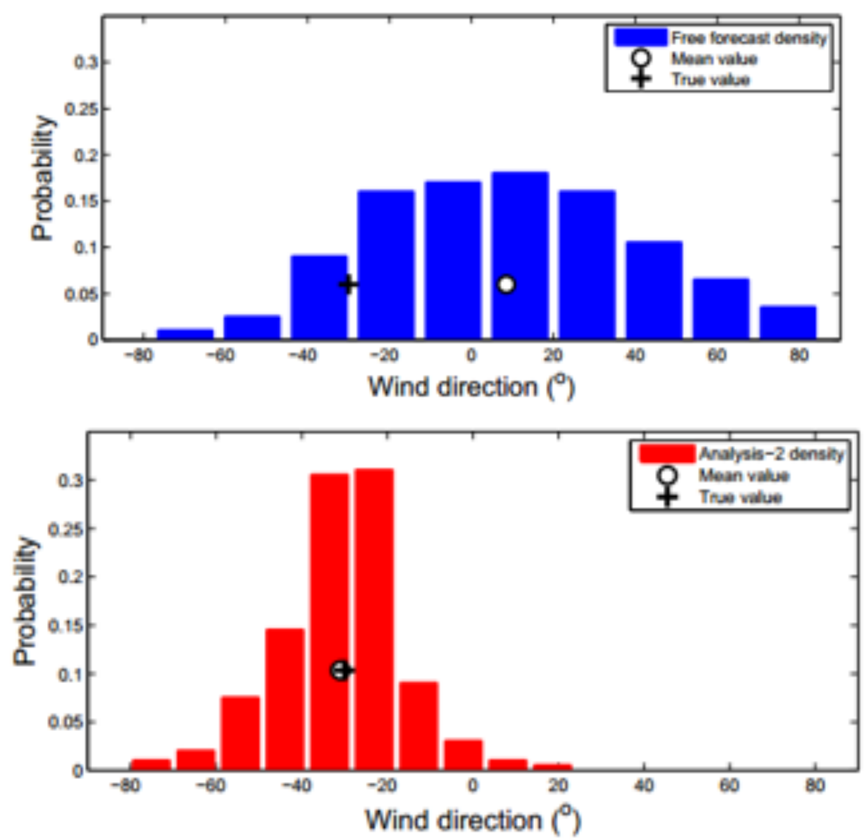


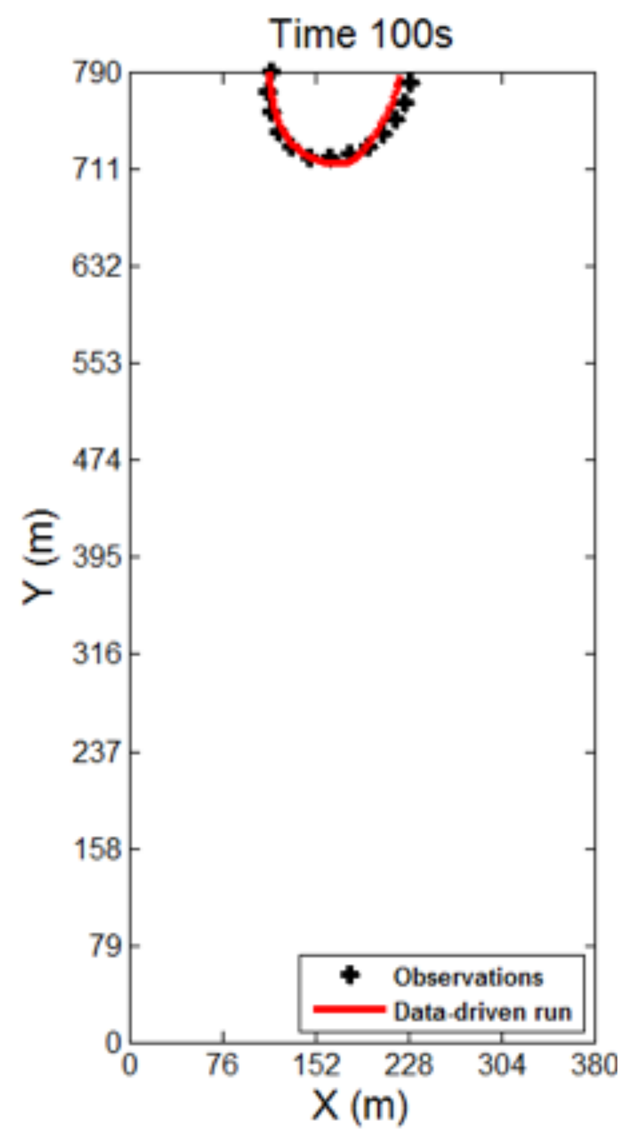



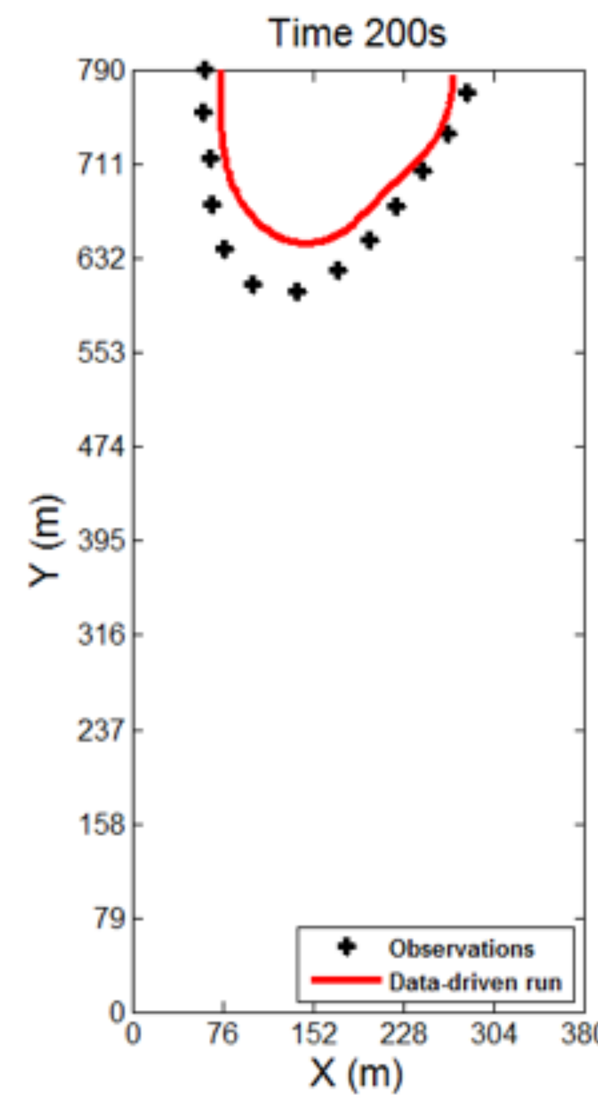
# Wildfire guideline

## Validation test - FireFlux I experiment

- Algorithm: Ensemble Kalman filter (EnKF) for parameter estimation
- Control variables: Spatially-distributed wind magnitude and direction
- Metric: Euclidean distance between observation and simulated front




**Rochoux et al. (2014,2015),** Natural Hazards and Earth System Sciences
**Zhang et al. (2017),** Fire Safety Journal

# Introduction

## Talk's outline

The wildfire problem as guideline

**(1) Data assimilation algorithm**

- Ensemble-based Kalman filtering
- Link with uncertainty quantification methods

**(2) Position errors**

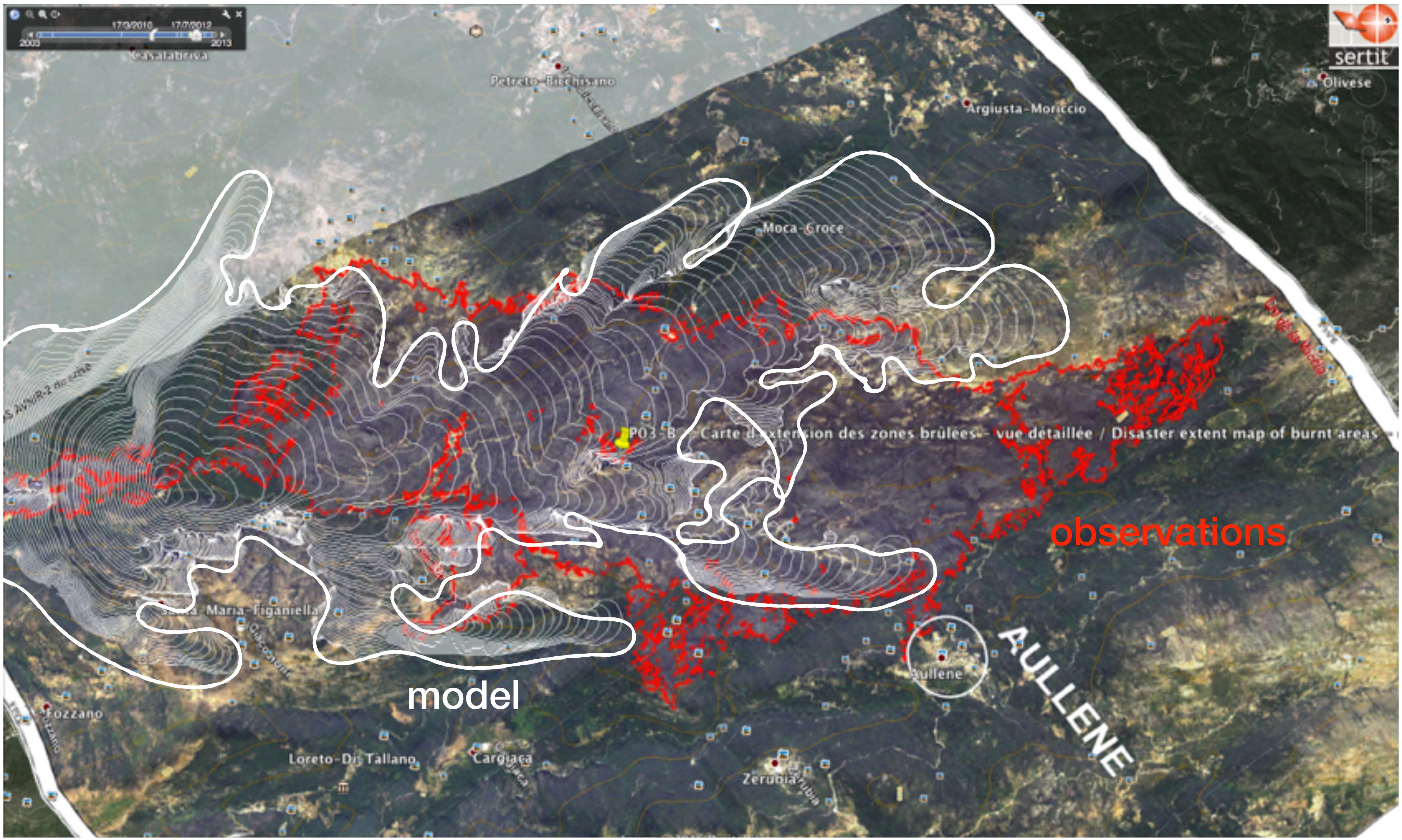
- Amplitude errors versus Position errors
- New measure to quantity front shape similarity

**(3) Object-oriented data assimilation**

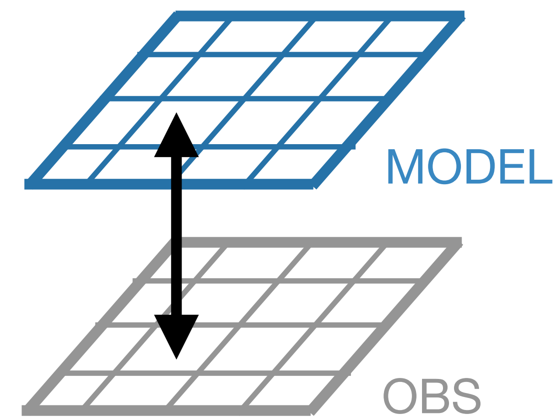
- Data assimilation with front shape similarity measure
- Joint state-parameter estimation

# Data assimilation challenge

How to address position errors for complex front topology?

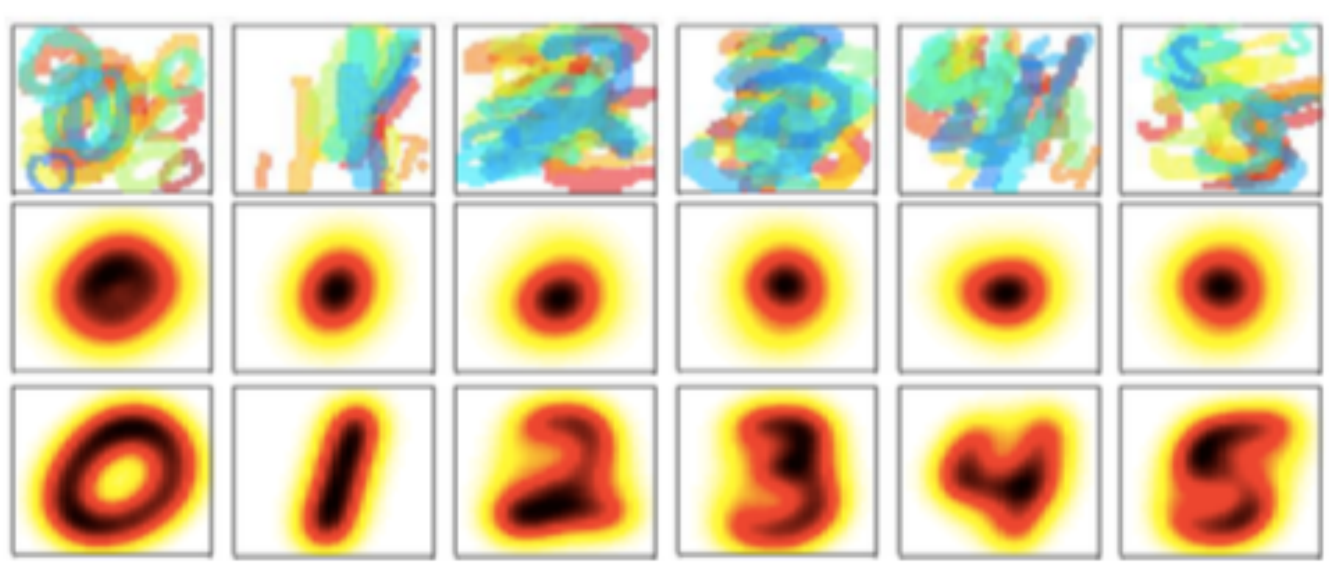


# Data assimilation challenge for tracking structures



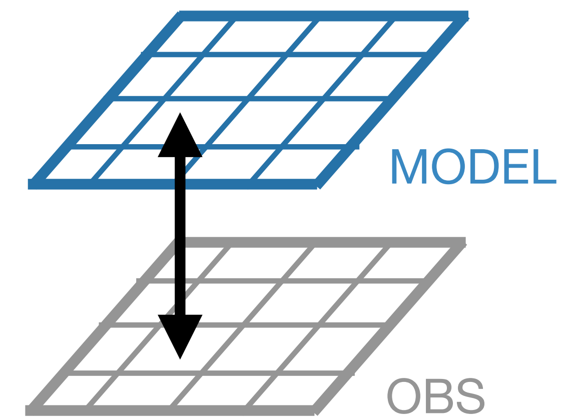

## Limitation of point-wise local metrics

- Several metrics are usually required to satisfyingly compare two fields
- Double penalty effect
  ➡ A misplaced structure is predicted where it should not be and is not predicted where it should be
- Small spatial and temporal shift of the structure position
- Failure of standard data assimilation methods when position errors are large, for instance when observations are infrequent
  ➡ Standard treatment of amplitude errors (Euclidean metrics)
  ➡ Generation of artificial patterns



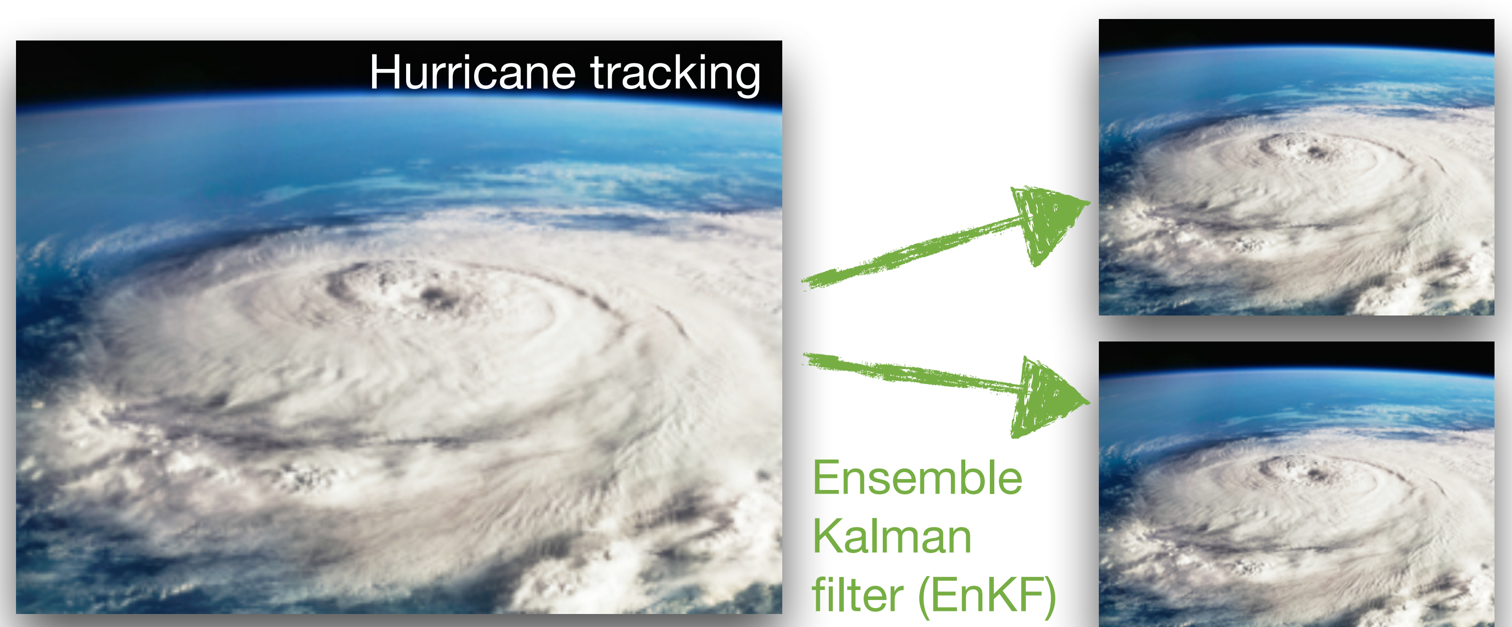

**Doucet and Cuturi (2014),** Proceedings of Machine Learning Research

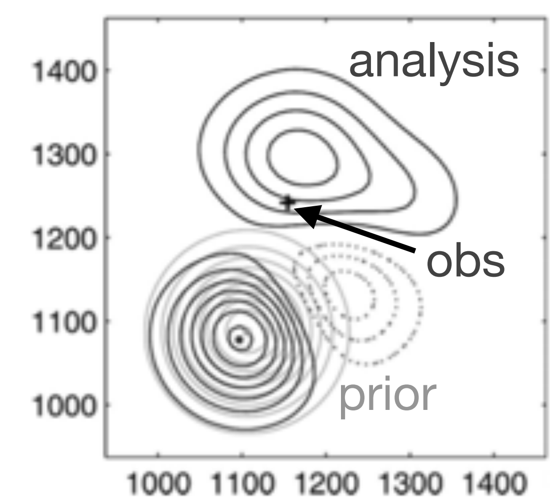
# Data assimilation challenge for tracking structures




## Limitation of point-wise local metrics

- Several metrics are usually required to satisfyingly compare two fields
- Double penalty effect
  ➡ A misplaced structure is predicted where it should not be and is not predicted where it should be
- Small spatial and temporal shift of the structure position
- Failure of standard data assimilation methods when position errors are large, for instance when observations are infrequent
  ➡ Standard treatment of amplitude errors (Euclidean metrics)
  ➡ Generation of artificial patterns

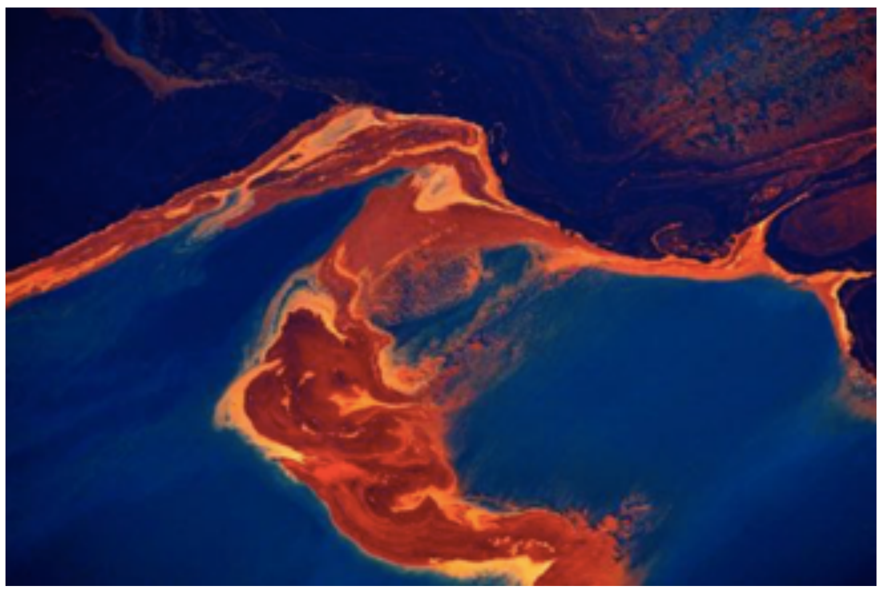



**Chen and Snyder (2007),** Monthly Weather Review
**Beezley and Mandel (2008),** Tellus

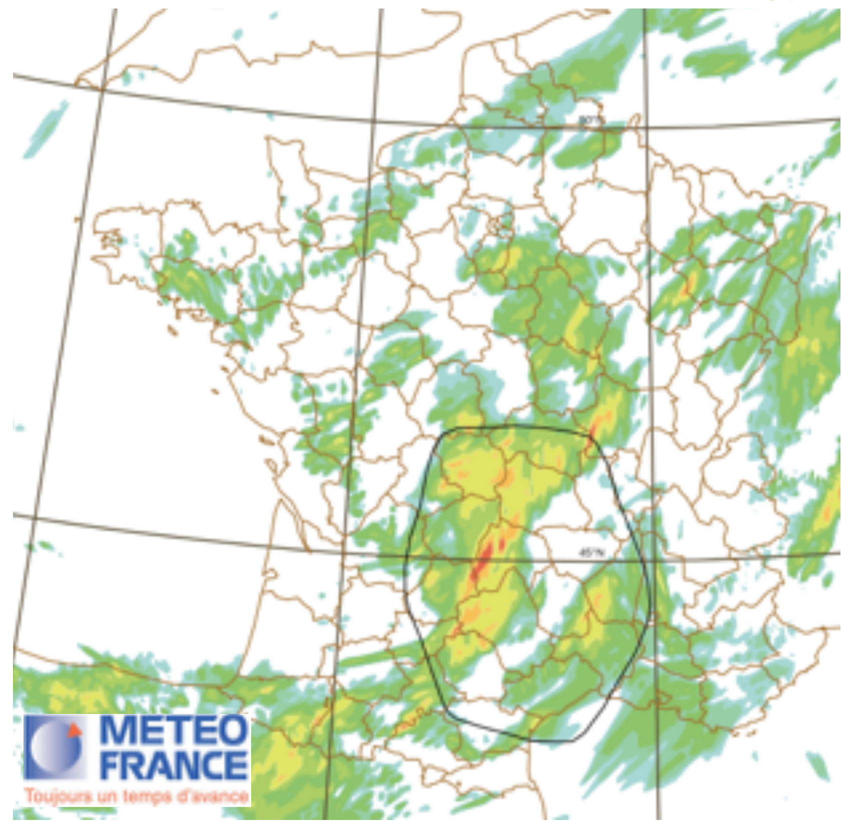
# Data assimilation challenge for tracking structures

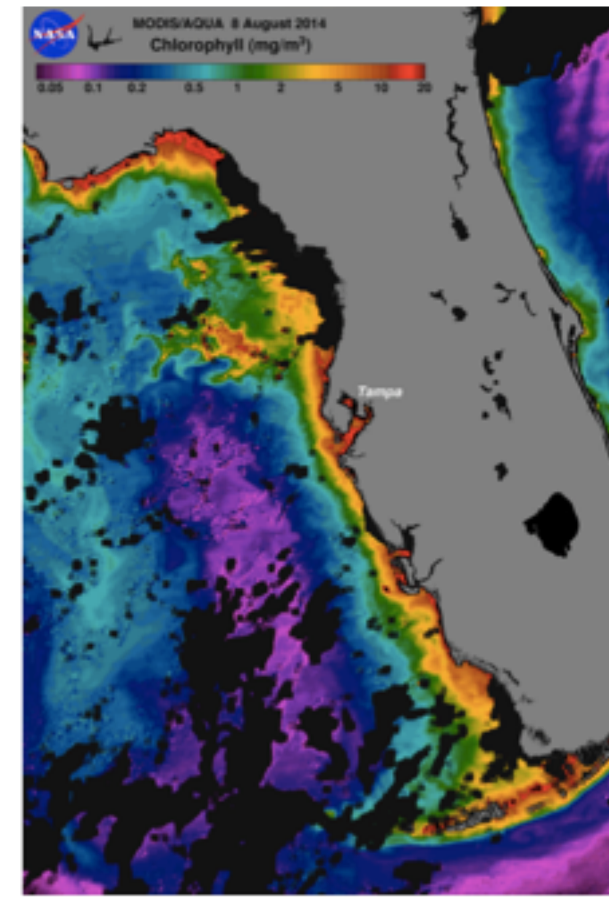
"It is easy to perceive coherent structures by eye, but a full precise mathematical description is still a challenge."



Oil spill in the ocean

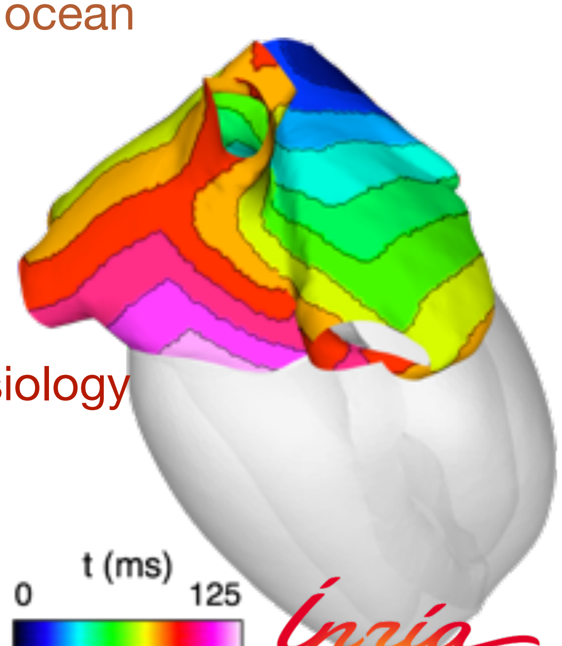
Precipitation pattern in meteorology



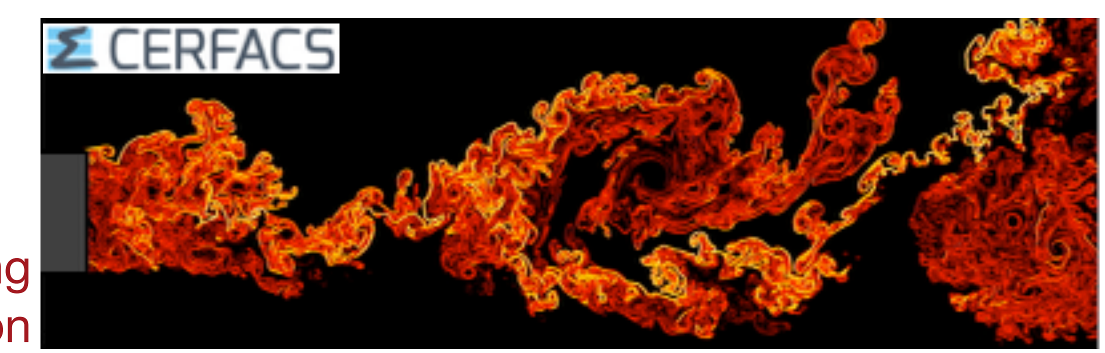




Chlorophylle concentration in the ocean with cloud occlusion

Cardiac electrophysiology




Flaming combustion

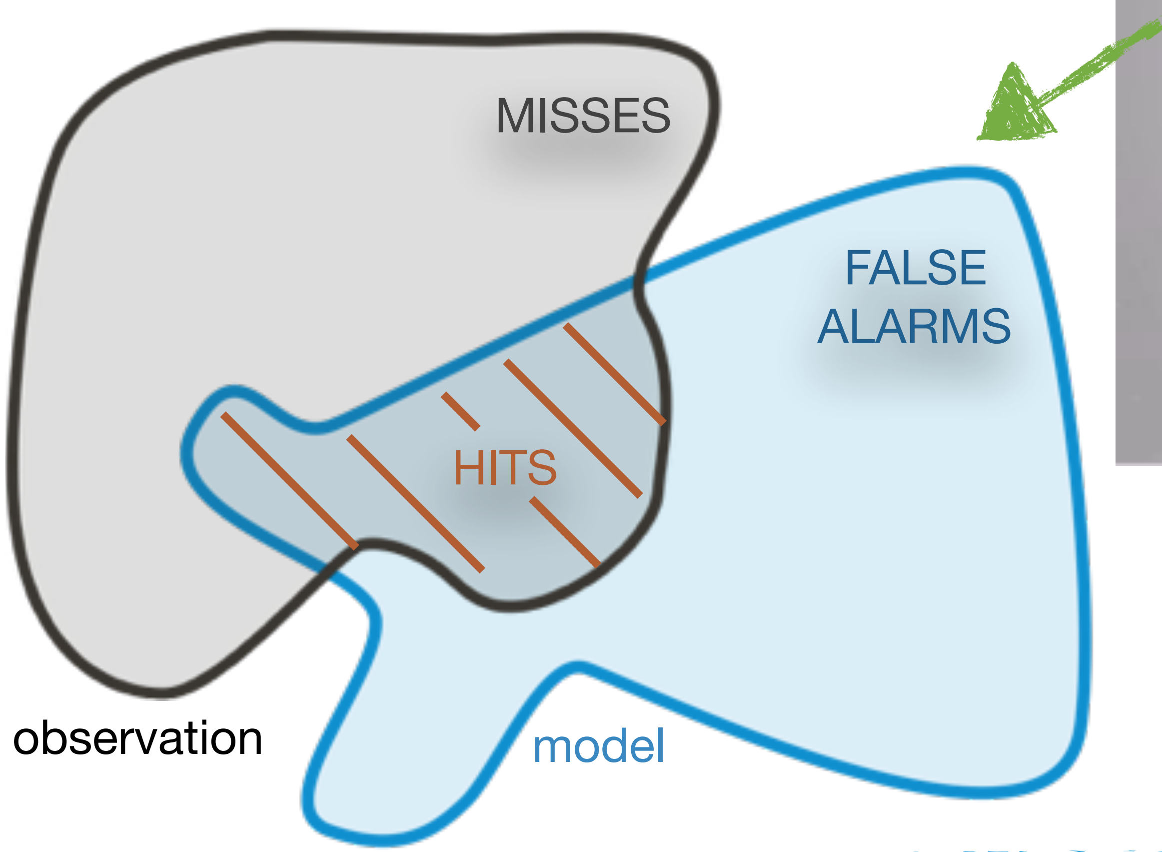


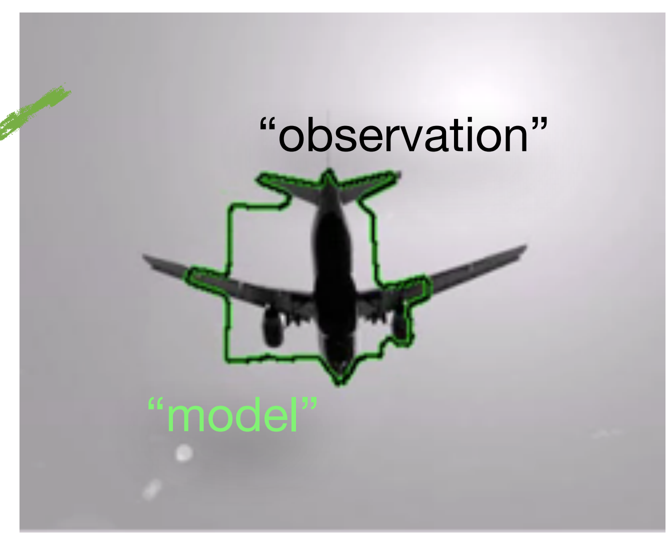
# What can we learn from image segmentation theory?

**How to track a moving object? How to represent uncertainties in the object shape and location?**

- Scale separation (ex: wavelet transform)
- Fuzzy method (ex: prior field smoothing)
- Identification and comparison of main field features
- Field deformation or field displacement (ex: Wasserstein distance, Chan-Vese functional)

FOCUS




**Chan-Vese data fitting functional used in image segmentation**

- minimization of hits
- maximization of misses

**Nelson Feyeux (2016),** Transport optimal pour l'assimilation de données d'images, Thèse de doctorat, Communauté Université Grenoble Alpes

**Collin et al. (2015),** Journal of Computational Physics
**Arbogast et al. (2016),** Quarterly Journal of the Royal Meteorological Society

# Chan-Vese contour fitting functional

## Similarity measure between “target” and simulated fronts

$$\mathcal{J}(\mathrm{x}) = ||\mathcal{G}(\mathrm{x}) - \mathrm{y}^{\mathrm{o}}||^2$$

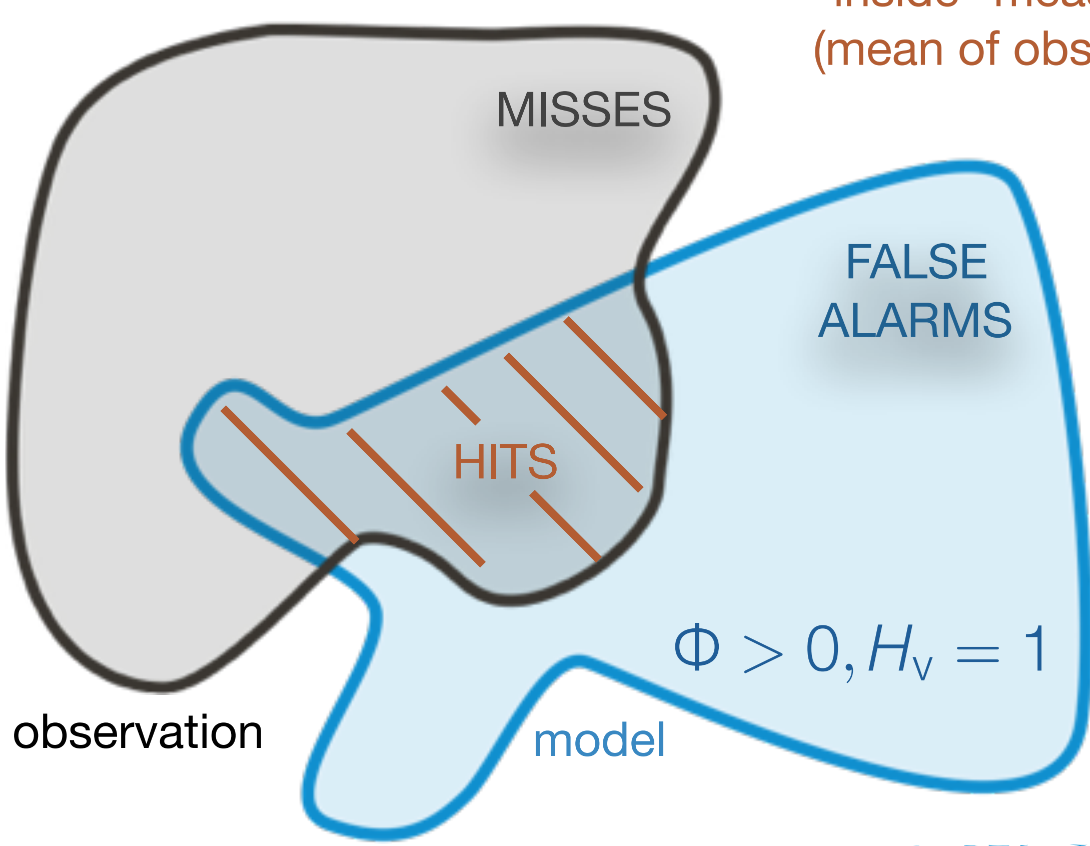
level-set formalism

**Chan and Vese (2001),** IEEE Transactions on Image Processing

$$\mathcal{J}(\phi, \mathrm{y}^{\mathrm{o}}) = \int_{\Omega} H_{\mathrm{v}}(\phi)\left[\mathrm{y}^{\mathrm{o}} - C_1(\mathrm{y}^{\mathrm{o}}, \phi)\right]^2 + (1 - H_{\mathrm{v}}(\phi))\left[\mathrm{y}^{\mathrm{o}} - C_0(\mathrm{y}^{\mathrm{o}}, \phi)\right]^2 \mathrm{dx}$$

“inside” measuring HITS
(mean of obs. in simulated burnt area)

“outside” measuring MISSES
(mean of obs. in simulated unburnt area)




$\Phi < 0, H_{\mathrm{v}} = 0$

**Minimizing the functional acts on the contour of the simulated area to match the shape of the observed front.**

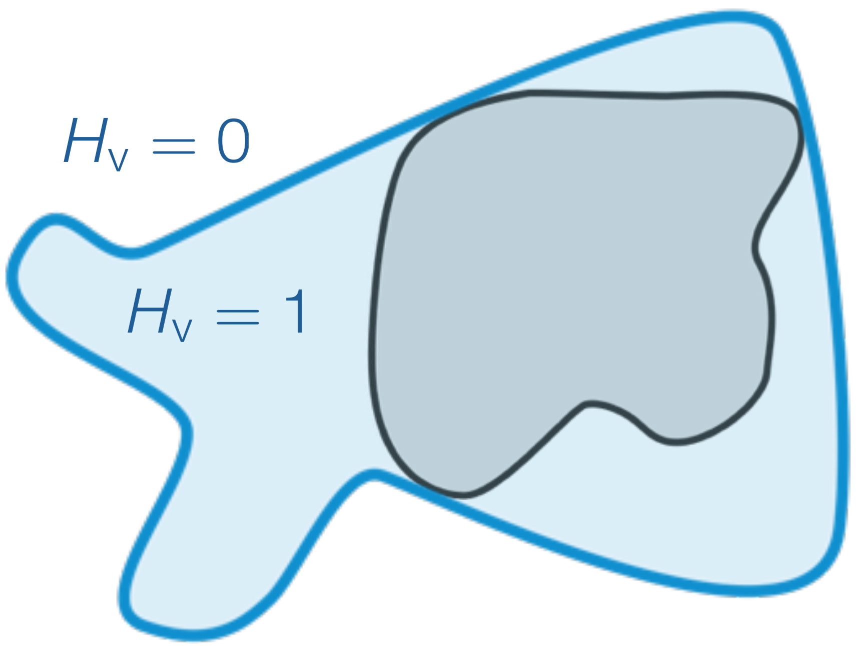
# Chan-Vese contour fitting functional

## Behavior of the front shape similarity measure

$$\mathcal{J}(\phi, \mathrm{y^o}) = \int_\Omega H_v(\phi)\,[\mathrm{y^o} - C_1(\mathrm{y^o}, \phi)]^2 + (1 - H_v(\phi))\,[\mathrm{y^o} - C_0(\mathrm{y^o}, \phi)]^2 \,\mathrm{dx}$$

"inside" measuring HITS
(mean of obs. in simulated burnt area)

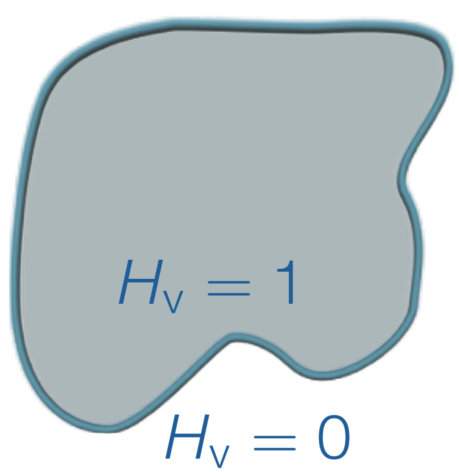
"outside" measuring MISSES
(mean of obs. in simulated unburnt area)




**Case 1:** observations included in simulated burnt area

$H_v = 1, 0 < C_1 < 1$

$H_v = 0, C_0 = 0$




**Case 2:** coincidence of the simulated burnt area with observation

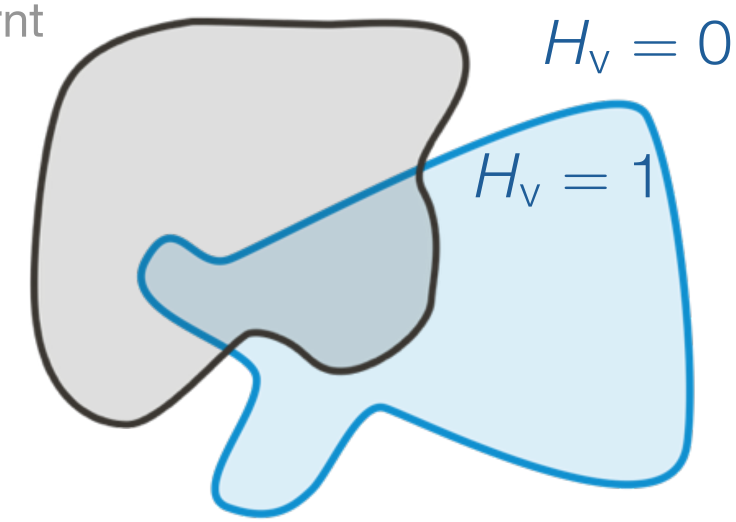
$H_v = 1, C_1 = 1$

$H_v = 0, C_0 = 0$

**Case 3:** partial overlap between simulated burnt area and observation

$H_v = 1, 0 < C_1 < 1$

$H_v = 0, 0 < C_0 < 1$



# Introduction

## Talk's outline

### (1) Data assimilation algorithm
- Ensemble-based Kalman filtering
- Link with uncertainty quantification methods

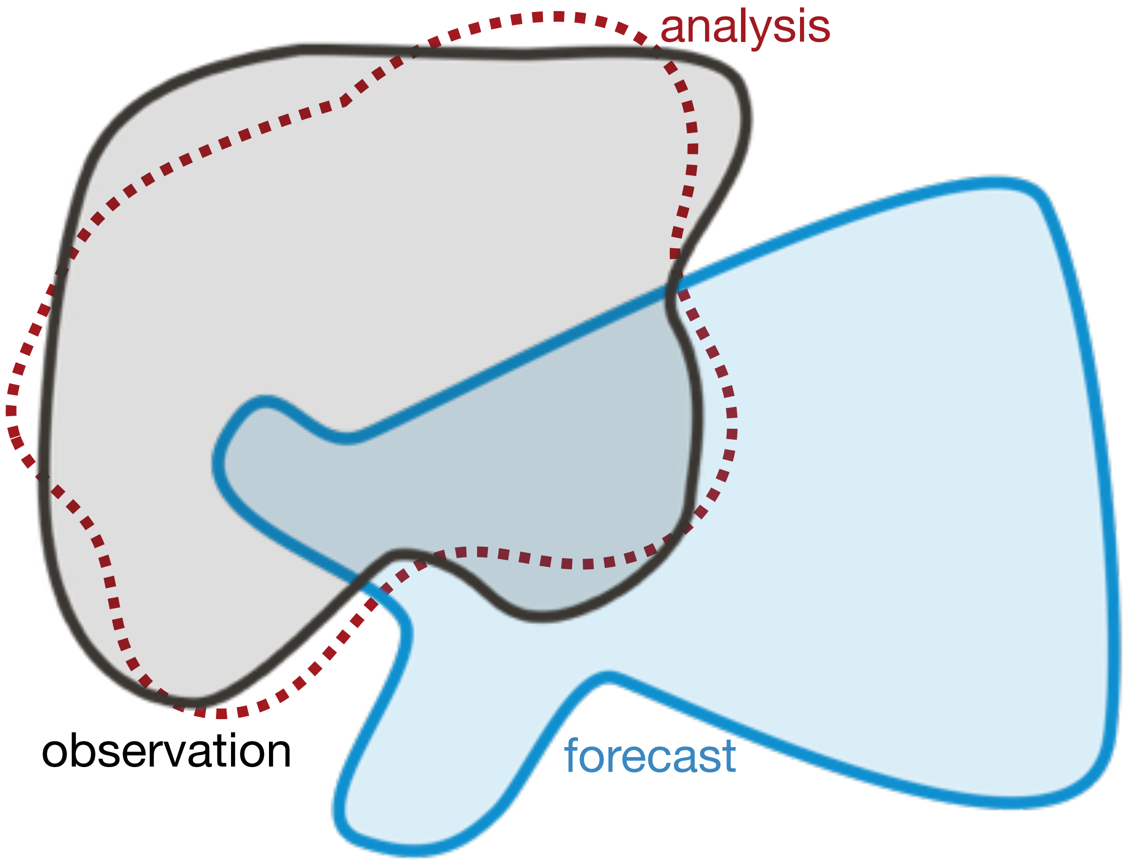
### (2) Position errors
- Amplitude errors versus Position errors
- New measure to quantity front shape similarity

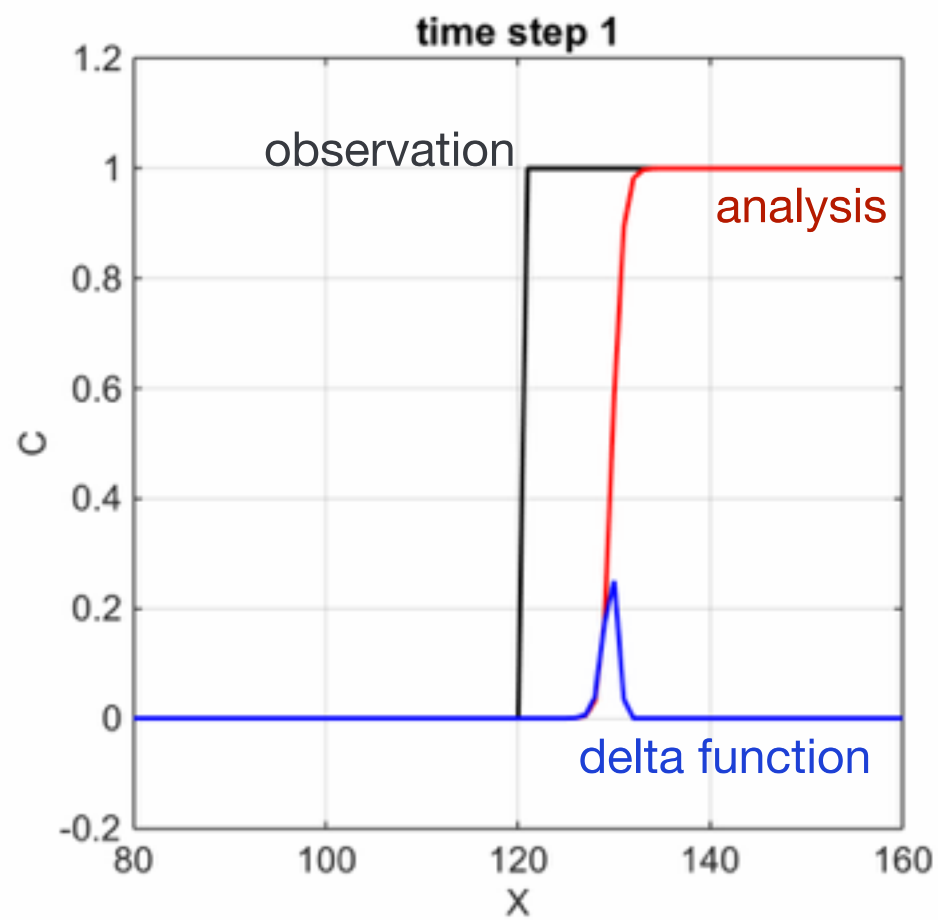
### (3) Object-oriented data assimilation
- Data assimilation with front shape similarity measure
- Joint state-parameter estimation

The wildfire problem as guideline

# State estimation problem

## Formulate the analysis using the front shape similarity measure



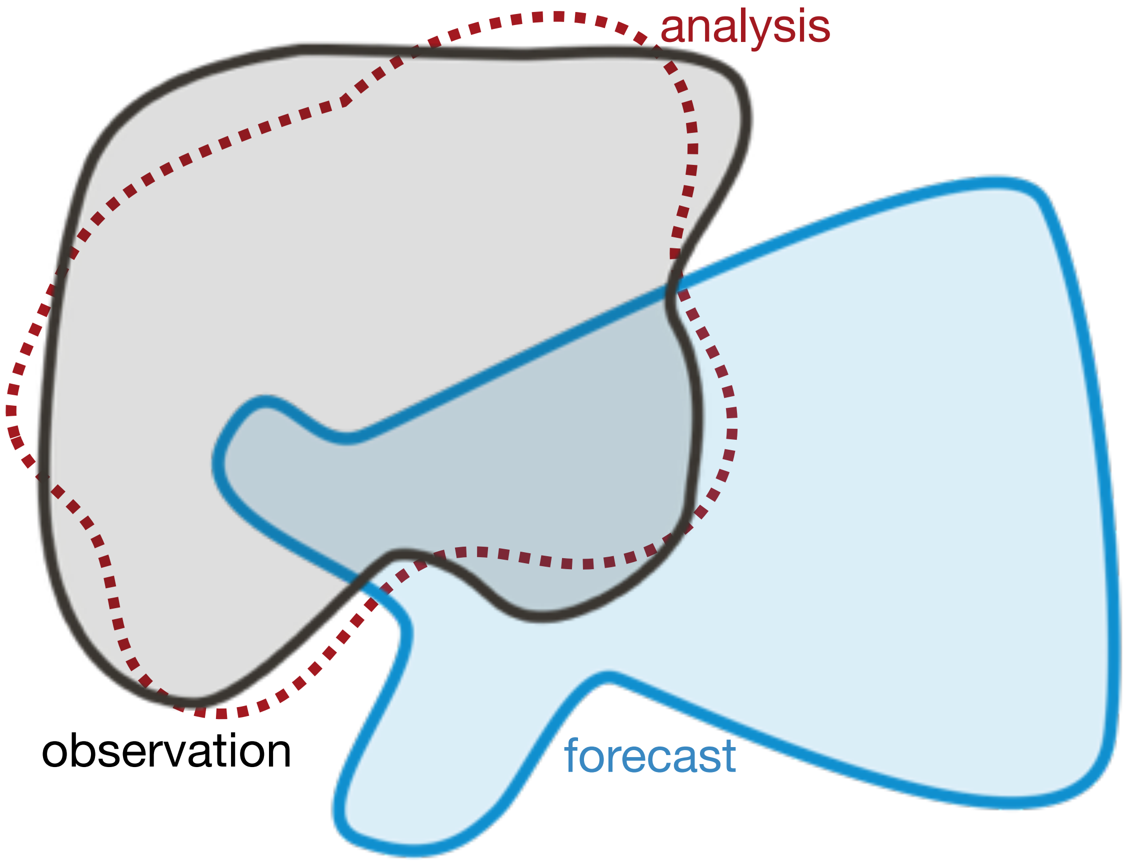

### Estimation of the progress variable *c*

$$\frac{\partial \widehat{c}}{\partial t} + \mathrm{V} \cdot \nabla \widehat{c} = -\lambda\, \delta(\phi) \left\{ [\mathrm{y}^o - C_1(\mathrm{y}^o, \phi)]^2 - [\mathrm{y}^o - C_0(\mathrm{y}^o, \phi)]^2 \right\}$$

Dirac δ function localizing the data assimilation feedback on the simulated front

# Parameter estimation problem

## Adaptation of Kalman filtering to the front shape similarity measure




### FRONT-TRACKING PROBLEM

Progress variable c = c(*x*,*y*,*t*)

- Front marker → contour line $c_{fr}$
- Level set function → $\phi = c - c_{fr}$
- Propagation equation

$$\frac{\partial c}{\partial t} + \mathrm{V} \cdot \nabla c = 0$$

### ENSEMBLE-BASED KALMAN FILTER

Introduction of a discrepancy operator *D*

- Based on the gradient of the Chan-Vese data fitting functional
- Analysis still formulated as a correction of the forecast

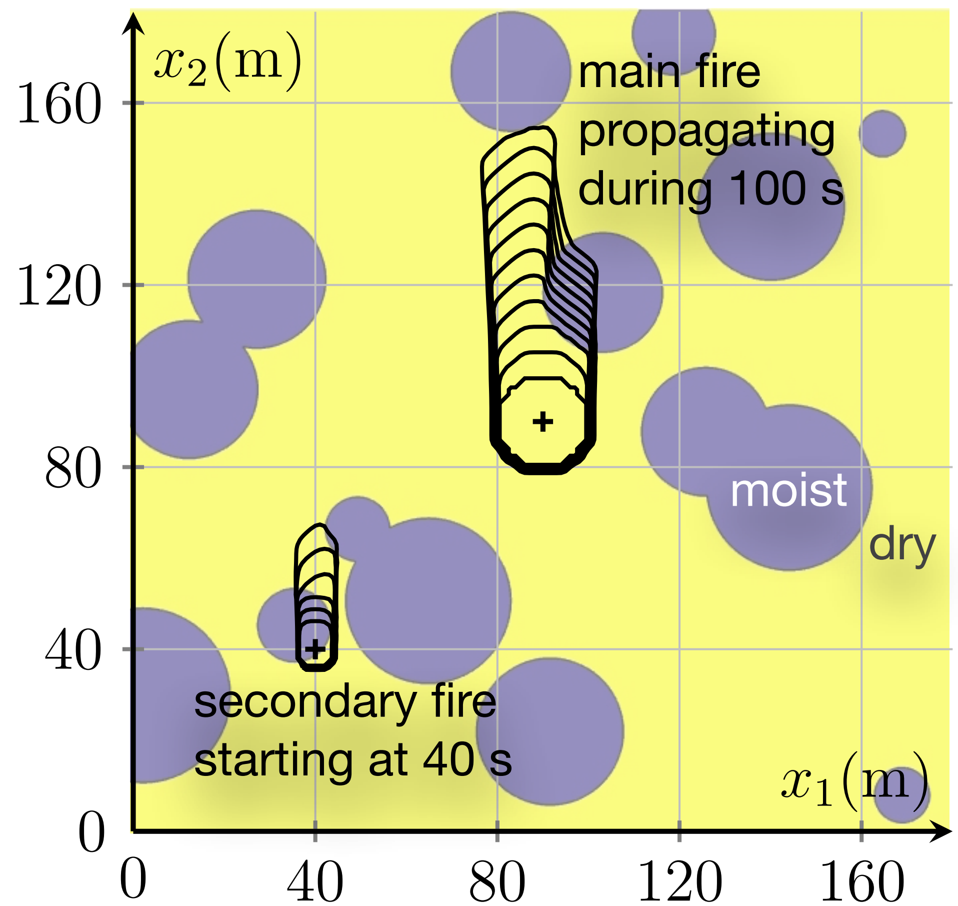
**Estimation of the physical parameters that are inputs to the velocity V**

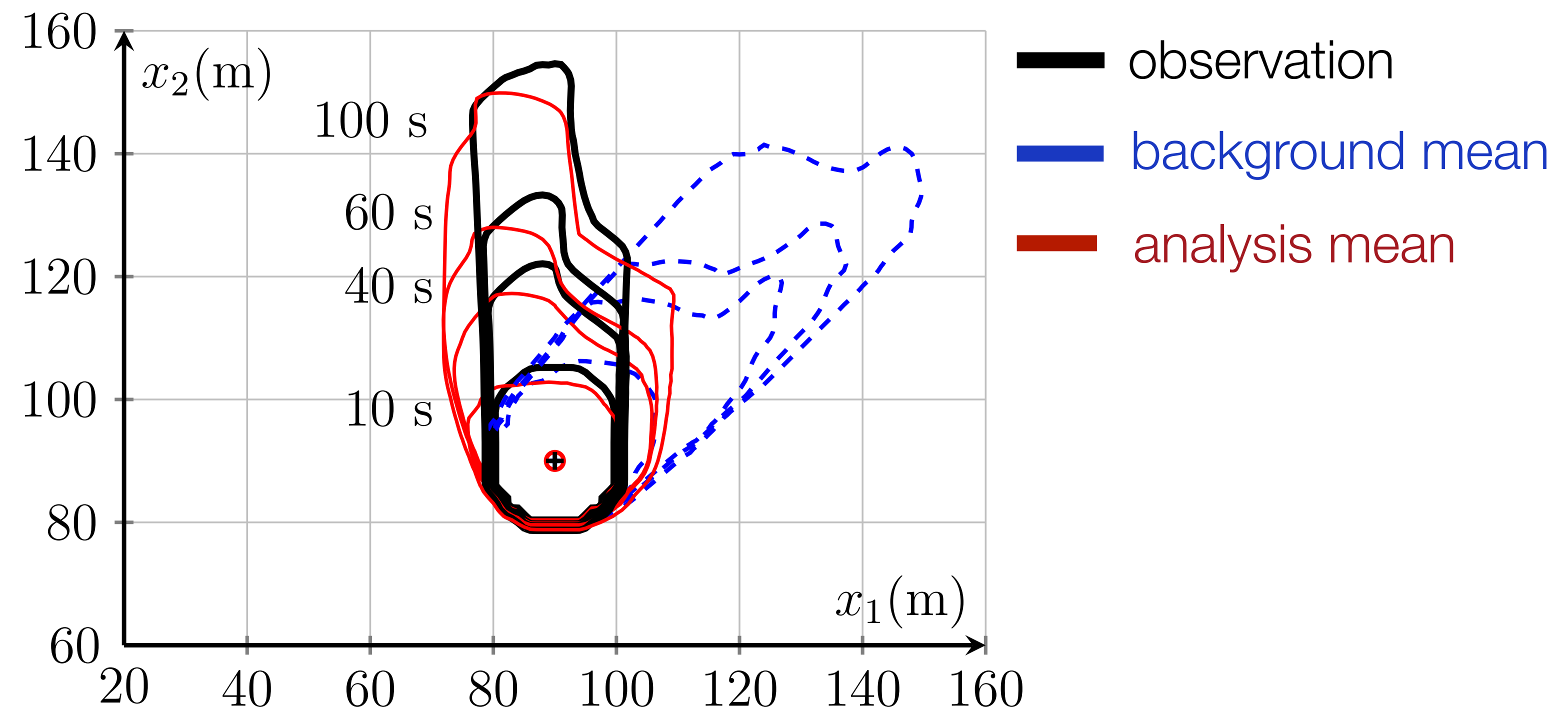
$$\mathbf{x}^{a}_{n+1} = \mathbf{x}^{b}_{n+1} + \mathbf{K}_{n+1}\left[D\left(\mathbf{y}^{o}_{n+1}, G(\mathbf{x}^{b}_{n+1})\right)\right]$$

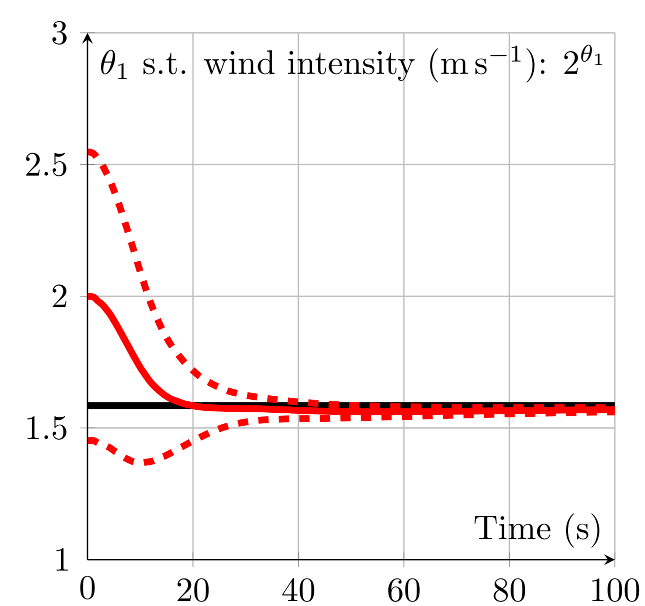
Discrepancy operator that represents front shape discrepancies and that can assimilate image data directly

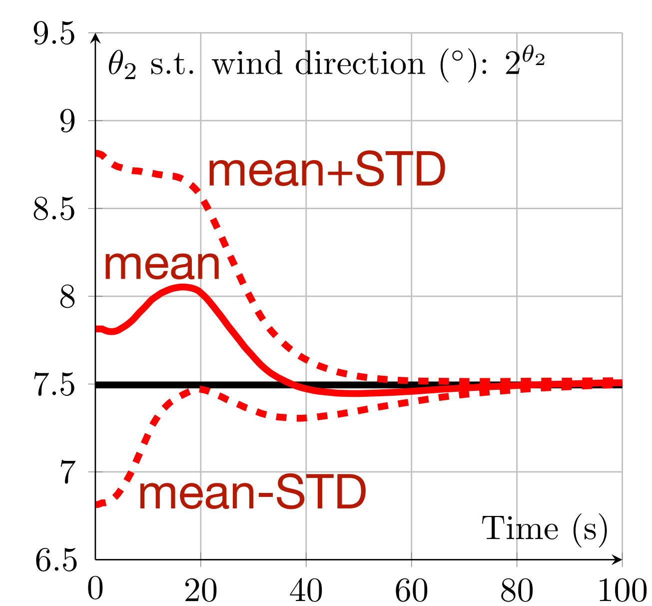
# Verification test

## Parameter estimation with wrong wind (intensity, direction)



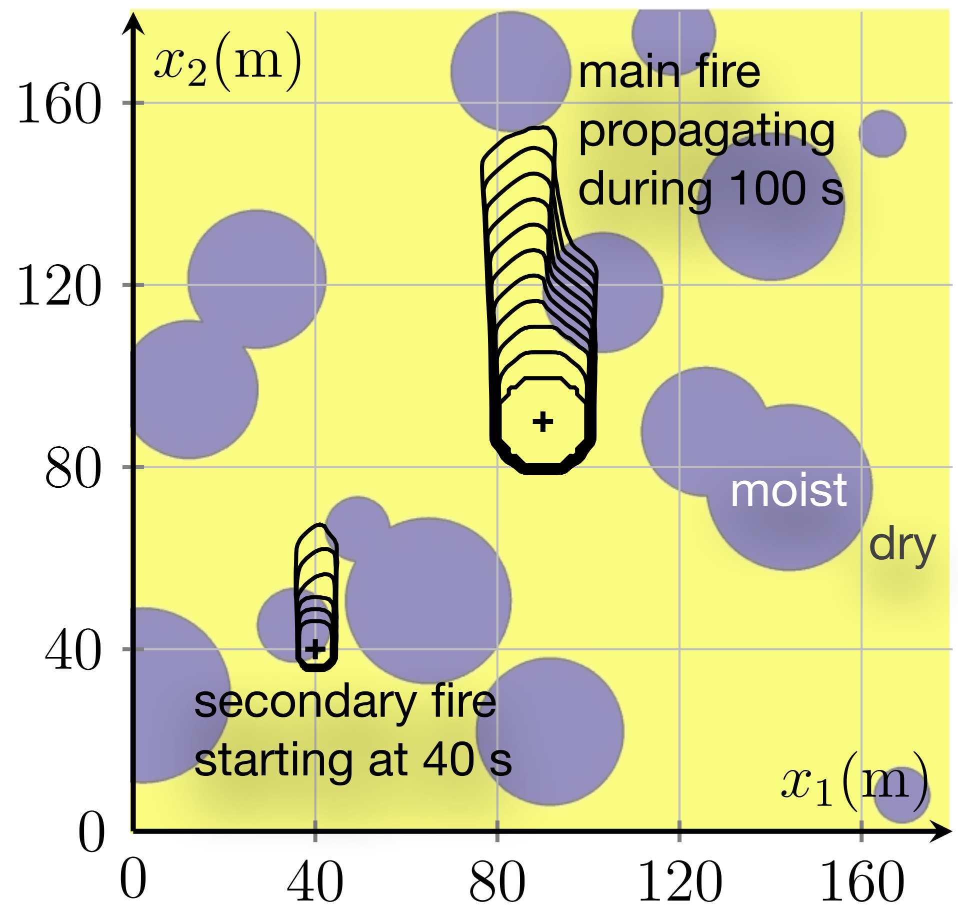



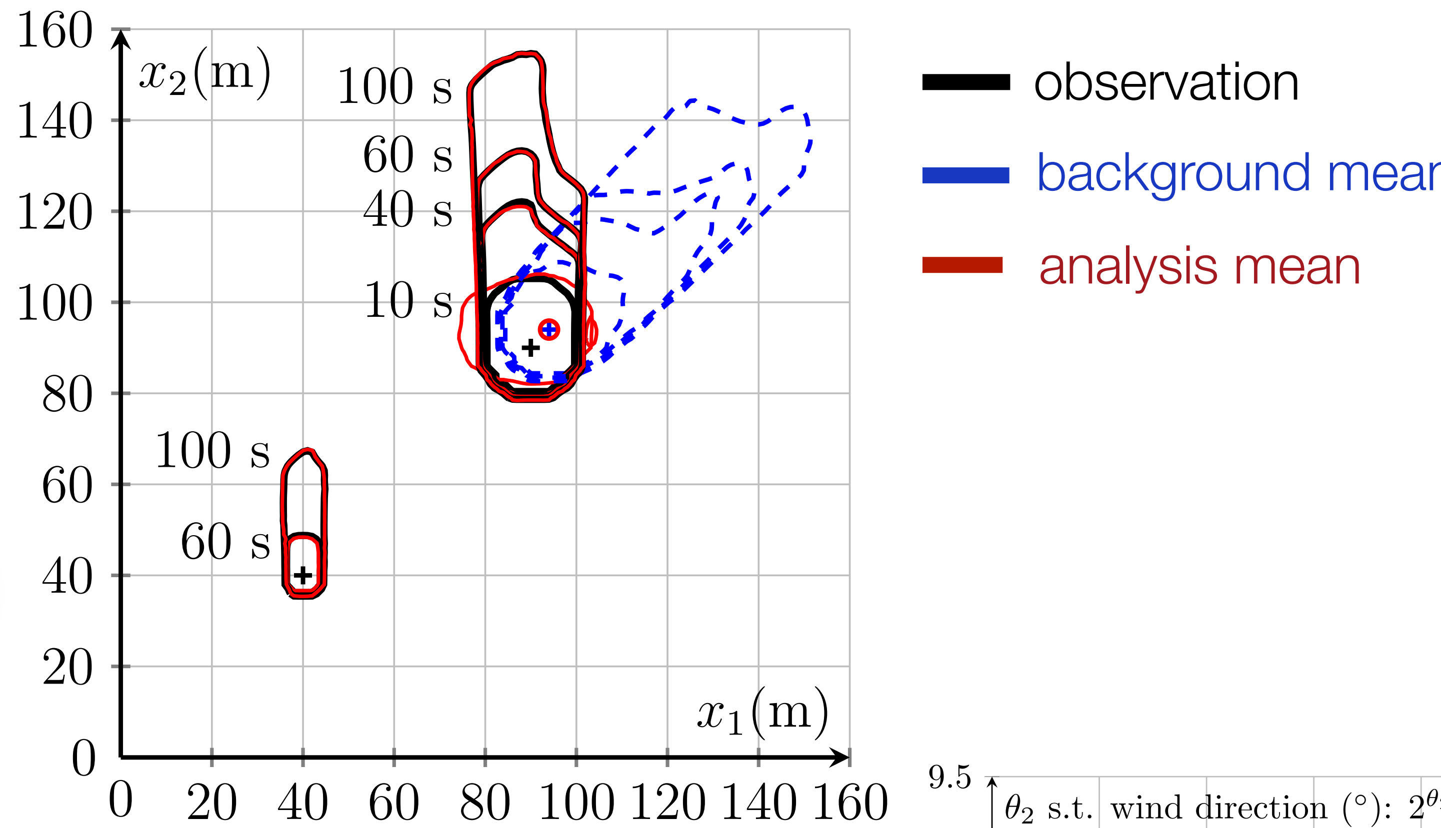

**Rochoux et al. (2017),** ESAIM: Proceedings and Surveys

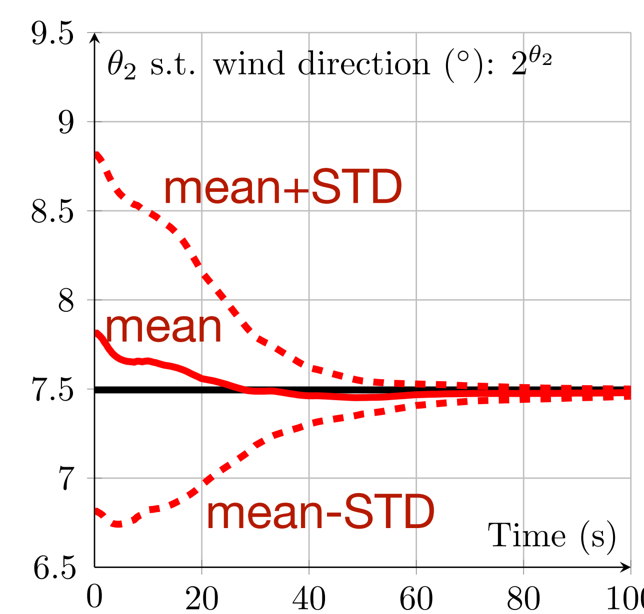
**Collin et al. (2015),** Journal of Computational Physics

# Verification test

## Joint state-parameter estimation with wrong wind and initial condition



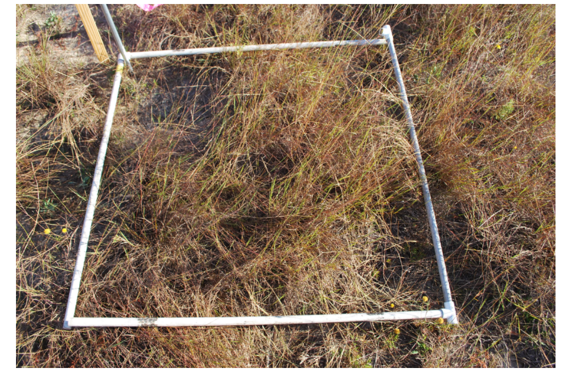

**Rochoux et al. (2017),** ESAIM: Proceedings and Surveys

**Collin et al. (2015),** Journal of Computational Physics

- More physical parameter values
- Able to using an additional topological gradient in the state estimator

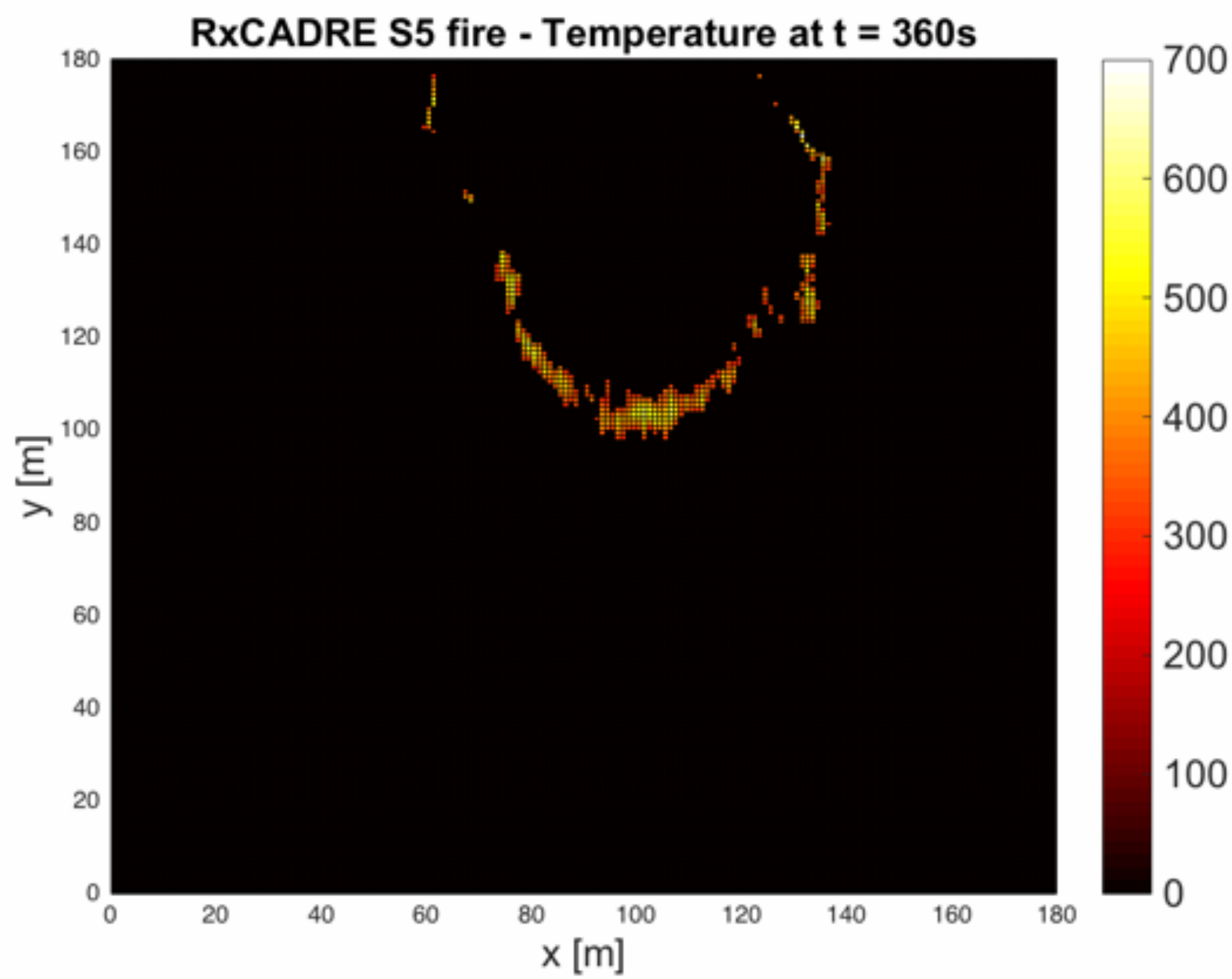
➡ More complex front topology

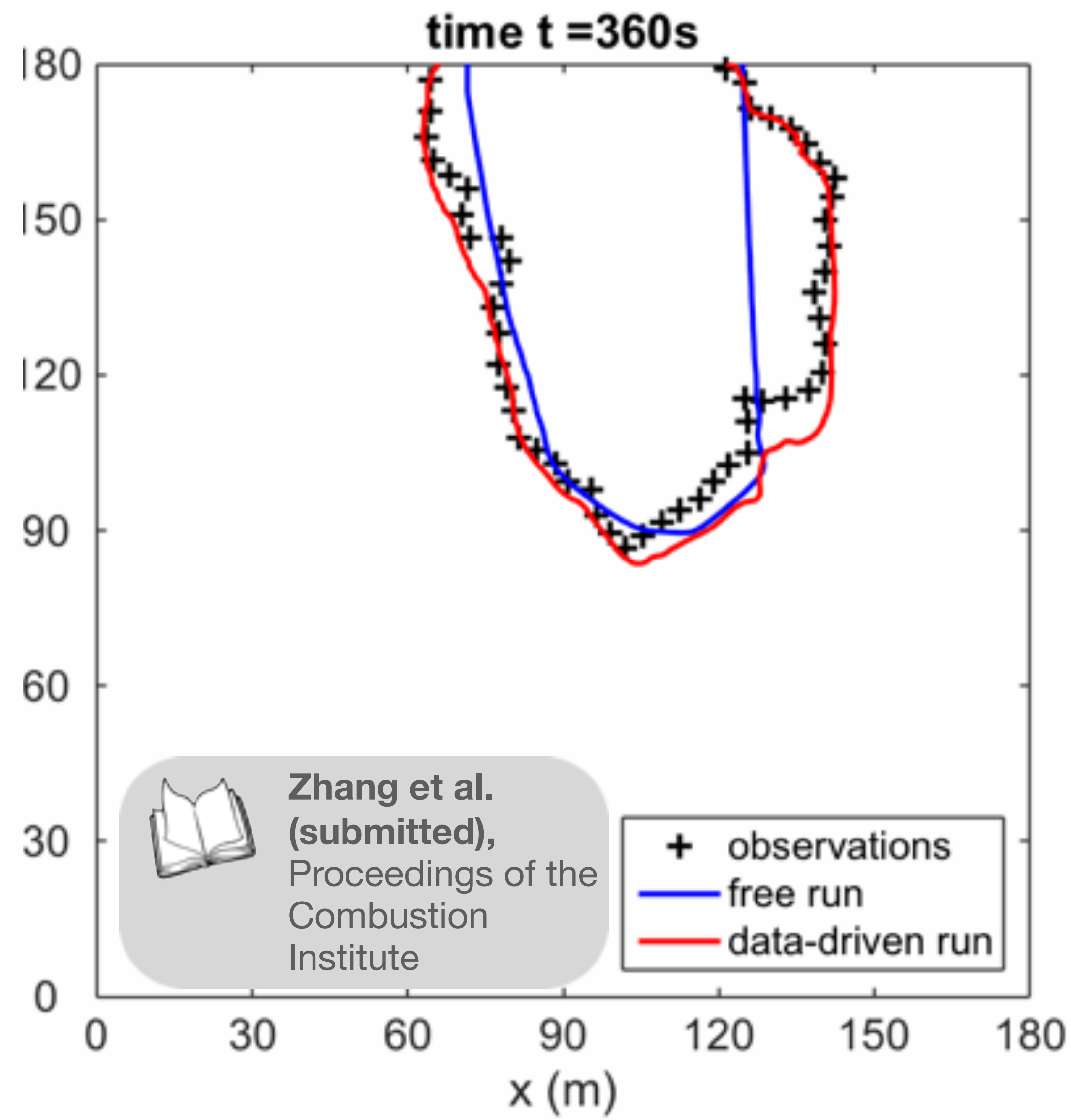
# Validation test




## 3-hectare RxCADRE controlled burn experiment

- Algorithm: Luenberger observer for state estimation
- Control variables: Progress variable *c*
- Metric: Front shape similarity measure
- Fire: 8-min fire propagation over mix of grass and shrub

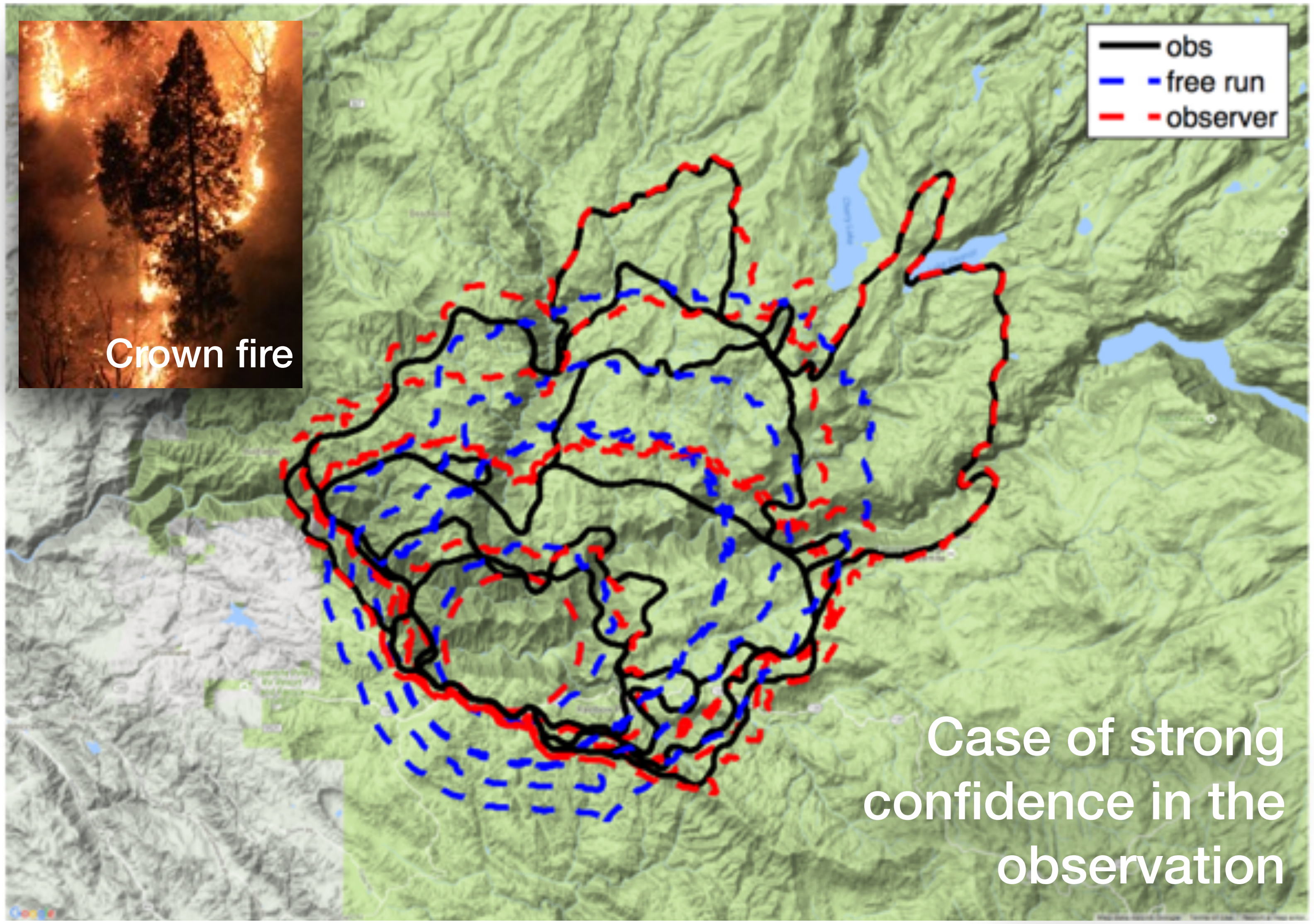


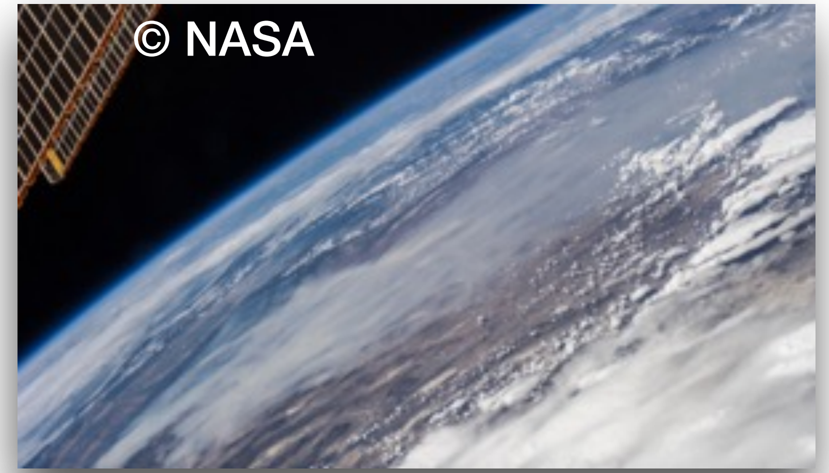



# Validation test

## 70 km x 50 km RIM grass/forest wildfire (California, 2013)

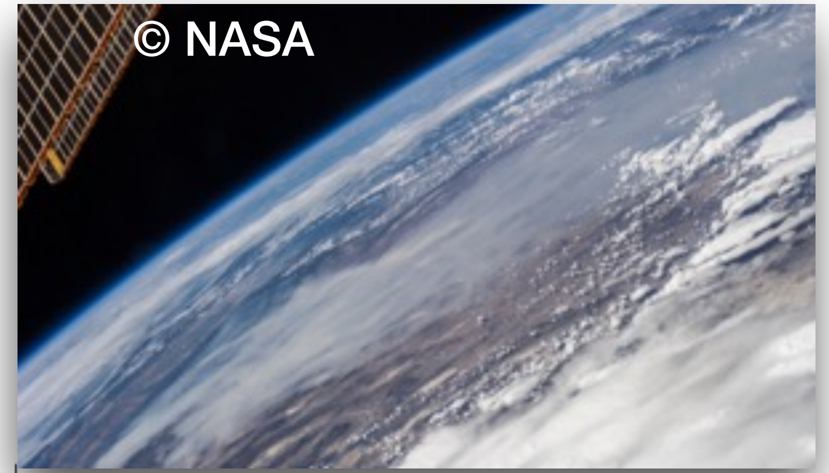
- Algorithm: Luenberger observer for state estimation
- Control variables: Front marker positions (Lagrangian model)
- Metric: Front shape similarity measure
- Fire: 11 observations, August 20-25, started from illegal campfire



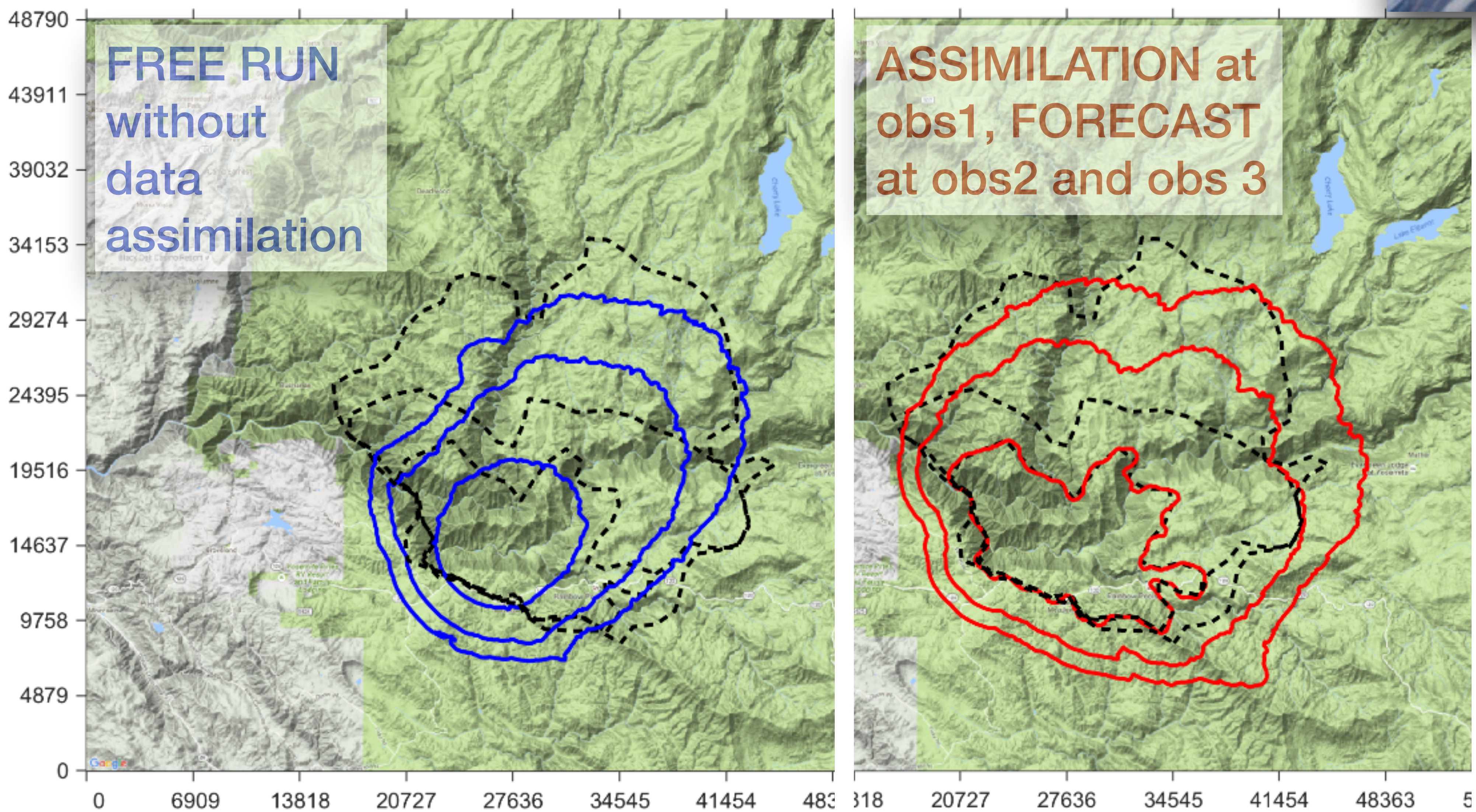


# Validation test




## 70 km x 50 km RIM grass/forest wildfire (California, 2013)

- Algorithm: Luenberger observer for state estimation
- Control variables: Front marker positions (Lagrangian model)
- Metric: Front shape similarity measure
- Fire: 11 observations, August 20-25, started from illegal campfire




**Observations**

- **obs1**: Aug 20, 11:30
- **obs2**: Aug 21, 23:50
- **obs3**: Aug 22, 21:00

**Model computational cost**

- spatial resolution ~70 m
- Time step ~60 s
- Simulation time ~100 hr ↔ CPU time ~20 min

# Conclusions

## New front shape similarity measure for data-driven front-tracking modeling

➡ **Data assimilation:** Find more optimal values of the estimation targets by minimizing the misfit error with respect to the observations
- Ensemble-based Kalman filtering for parameter estimation
- Deterministic Luenberger observer for state estimation

➡ **Design of an adapted misfit error measure for front-tracking problem**
- Position and shape errors, not only amplitude errors
- United framework for Eulerian and Lagrangian models

➡ **Application to wildland fires**
- Observation simulation system experiments
- RxCADRE experiments, RIM wildfire

The wildfire problem as guideline

**Rochoux et al. (2017),** ESAIM: Proceedings and Surveys

**Zhang et al. (submitted),** Proceedings of the Combustion Institute

Thank you for your attention.
Any question?

Contact
Melanie.Rochoux@cerfacs.fr