Ensemble forecast of analyses with uncertainty estimation

Vivien Mallet, Gilles Stoltz, Sergiy Zhuk, Alexander Nakonechniy

Ensemble Forecast of Analyses With Uncertainty Estimation

Vivien Mallet[1,2], Gilles Stoltz[3,4,1], Sergiy Zhuk[5], Alexander Nakonechniy[6]

[1]INRIA
[2]CEREA, joint laboratory ENPC - EDF R&D, Université Paris-Est
[3]DMA, École normale supérieure, CNRS
[4]HEC Paris
[5]IBM Research, Dublin
[6]Taras Shevchenko National University of Kiev

International Conference on Ensemble Methods in Geophysical Sciences

Toulouse, November 2012

Objective

To produce the best *forecast* of a model state using

- a data assimilation system, which produces analysis state vectors $\mathbf{z}_t$ using one or several models, observations and errors description; and
- a *given* ensemble of forecasts $\mathbf{x}_t^{(m)}$, possibly provided by the data assimilation system.

Ensemble Forecast of Analyses (EFA)

Notation

$\mathbf{x}_t^{(m)}$	State vector forecast by model/member m at time t
$\mathbf{z}_t$	Analysis state vector at time t

Strategy

- *Forecasting* the analysis state vector $\mathbf{z}_t$ computed by the data assimilation system
  - Rationale: The analyses are the best a posteriori knowledge of the state
  - Aggregated forecast:

$$\widehat{z}_{t,i} = \sum_{m=1}^{M} w_{t,i}^{(m)} x_{t,i}^{(m)}$$

  - Success if the ensemble forecasts $\widehat{\mathbf{z}}_t$ beat any sequence of forecasts $\mathbf{x}_t^{(m)}$

Ensemble Forecast of Analyses (EFA)

Notation

$\mathbf{x}_t^{(m)}$	State vector forecast by model/member m at time t
$\mathbf{z}_t$	Analysis state vector at time t

Principle

To produce an aggregated forecast $\widehat{\mathbf{z}}_t$ as efficient as possible, using the linear combination:

$$\widehat{z}_{t,i} = \sum_{m=1}^{M} w_{t,i}^{(m)} x_{t,i}^{(m)}$$

$t-2$	$t-1$	t	$t+1$
$\mathbf{x}_{t-2}^{(m)}$	$\mathbf{x}_{t-1}^{(m)}$	$\mathbf{x}_{t}^{(m)}$	$\mathbf{x}_{t+1}^{(m)}$
$\mathbf{w}_{t-2}^{(m)} \to \widehat{\mathbf{z}}_{t-2}$	$\mathbf{w}_{t-1}^{(m)} \to \widehat{\mathbf{z}}_{t-1}$	$\mathbf{w}_{t}^{(m)} \to \widehat{\mathbf{z}}_{t}$	$\mathbf{w}_{t+1}^{(m)} \to \widehat{\mathbf{z}}_{t+1}$
$\mathbf{y}_{t-2} \to \mathbf{z}_{t-2}$	$\mathbf{y}_{t-1} \to \mathbf{z}_{t-1}$	$\mathbf{y}_{t} \to \mathbf{z}_{t}$	$\mathbf{y}_{t+1} \to \mathbf{z}_{t+1}$

Ensemble Forecast of Analyses (EFA)

Notation

$\mathbf{x}_t^{(m)}$	State vector forecast by model/member m at time t
$\mathbf{z}_t$	Analysis state vector at time t

Principle

To produce an aggregated forecast $\widehat{\mathbf{z}}_t$ as efficient as possible, using the linear combination:

$$\widehat{z}_{t,i} = \sum_{m=1}^{M} w_{t,i}^{(m)} x_{t,i}^{(m)}$$

$t-2$	$t-1$	t	$t+1$
$\mathbf{x}_{t-2}^{(m)}$	$\mathbf{x}_{t-1}^{(m)}$	$\mathbf{x}_{t}^{(m)}$	$\mathbf{x}_{t+1}^{(m)}$
$\mathbf{w}_{t-2}^{(m)} \to \widehat{\mathbf{z}}_{t-2}$	$\mathbf{w}_{t-1}^{(m)} \to \widehat{\mathbf{z}}_{t-1}$	$\mathbf{w}_{t}^{(m)} \to \widehat{\mathbf{z}}_{t}$	$\mathbf{w}_{t+1}^{(m)} \to \widehat{\mathbf{z}}_{t+1}$
$\mathbf{y}_{t-2} \to \mathbf{z}_{t-2}$	$\mathbf{y}_{t-1} \to \mathbf{z}_{t-1}$	$\mathbf{y}_{t} \to \mathbf{z}_{t}$	$\mathbf{y}_{t+1} \to \mathbf{z}_{t+1}$

EFA Using Machine Learning

Computing the Aggregation Weights

- Ridge regression with discount in time ($\lambda > 0$ and $\beta > 0$):

$$\forall i \quad \mathbf{w}_{t,i} = \operatorname*{argmin}_{\mathbf{v} \in \mathbb{R}^M} \left[\lambda \|\mathbf{v}\|_2^2 + \sum_{s=1}^{s<t} \left(\frac{\beta}{(t-s)^2} + 1 \right) \left(z_{s,i} - \sum_{m=1}^{M} v^{(m)} x_{s,i}^{(m)} \right)^2 \right]$$

Theoretical Comparison With the Best Linear Combination With Constant Weights

$$\frac{1}{t} \sum_{s=1}^{s \le t} \left(z_{s,i} - \sum_{m=1}^{M} w_{s,i}^{(m)} x_{s,i}^{(m)} \right)^2$$

$$- \operatorname*{argmin}_{\mathbf{v} \in \mathbb{R}^M} \left[\frac{1}{t} \sum_{s=1}^{s \le t} \left(z_{s,i} - \sum_{m=1}^{M} v^{(m)} x_{s,i}^{(m)} \right)^2 \right] \lesssim \mathcal{O}\left(\frac{\ln t}{t} \right)$$

EFA Using Filtering

Formulation in Terms of Filtering

- State equation:

$$\mathbf{w}_{0,i} = \mathbf{c} + \mathbf{e}_i$$

$$\mathbf{w}_{t+1,i} = \mathbf{A}\mathbf{w}_{t,i} + (\mathbf{I} - \mathbf{A})\mathbf{c} + \mathbf{e}_{t,i}$$

- "Observations" (i.e., analyses in our case):

$$z_{t,i} = \mathbf{E}_{t,i}\mathbf{w}_{t,i} + \eta_{t,i}$$
$$\mathbf{E}_{t,i} = \left(x_{t,i}^{(1)}, \ldots, x_{t,i}^{(m)}\right)$$

EFA Using Filtering

Kalman Filtering

- Assignment of variances to initial weight errors, (weight) model errors and analyses errors
- The filter computes a variance $\mathbf{P}_{t,i}$ for the weight error at time t
- The aggregated forecast has variance $\mathbf{E}_{t,i}\mathbf{P}_t\mathbf{E}_{t,i}^T$

Minimax Filtering

- Bounds on errors, described by an ellipsoid

$$\mathbf{e}_i^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{e}_i + \sum_{t=0}^{T-1}\mathbf{e}_{t,i}^{\mathrm{T}}\mathbf{Q}_t^{-1}\mathbf{e}_{t,i} + \sum_{t=0}^{T} A_t^{-1}\eta_{t,i}^2 \leq 1$$

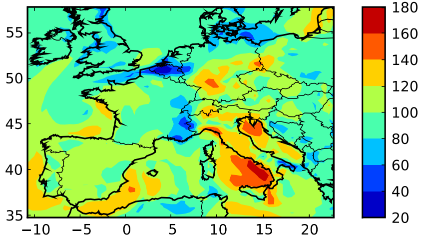
- Admissible weights are compatible with weight model, "observations" and errors bounds
- Weights defined such that, for any direction ℓ,

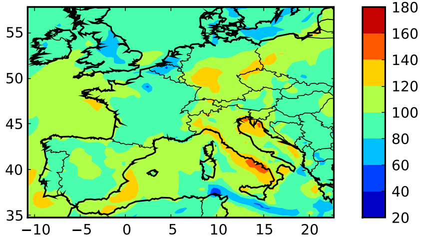
$$\sup_{\mathbf{e},\mathbf{e}_0,\ldots,\mathbf{e}_{t-1},\boldsymbol{\eta}_0,\ldots,\boldsymbol{\eta}_t} \ell^{\mathrm{T}}(\mathbf{w}_t^{\mathrm{true}} - \widehat{\mathbf{w}}_t) \leq \sup_{\mathbf{e},\mathbf{e}_0,\ldots,\mathbf{e}_{t-1},\boldsymbol{\eta}_0,\ldots,\boldsymbol{\eta}_t} \ell^T(\mathbf{w}_t^{\mathrm{true}} - \mathbf{w}_t)$$

Application to Air Quality Forecast

Simulations Description

- Forecasting ground-level ozone at 15h00 UTC (peak) over Europe
- Ensemble with 20 members
- One reference member in the ensemble benefits from data assimilation and actually provides the analyses





Application to Air Quality Forecast

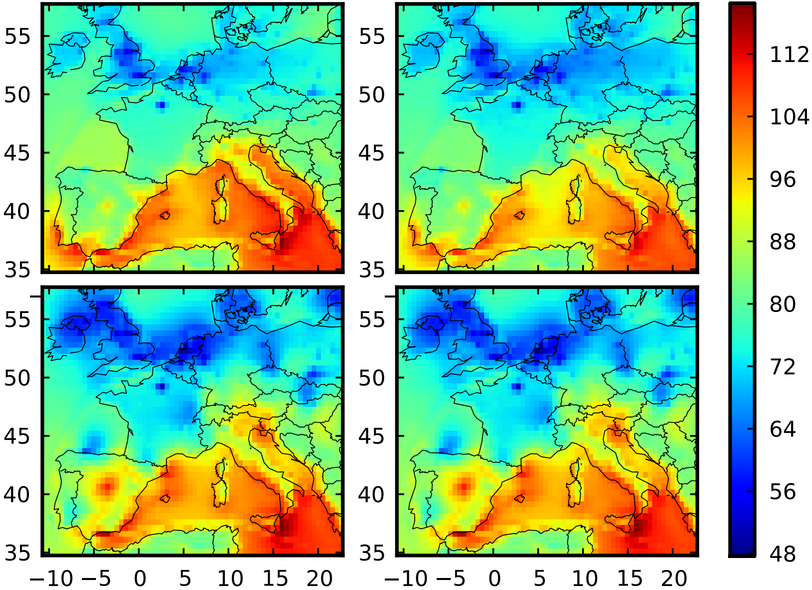
Simulations Description

- Forecasting ground-level ozone at 15h00 UTC (peak) over Europe
- Ensemble with 20 members
- One reference member in the ensemble benefits from data assimilation and actually provides the analyses

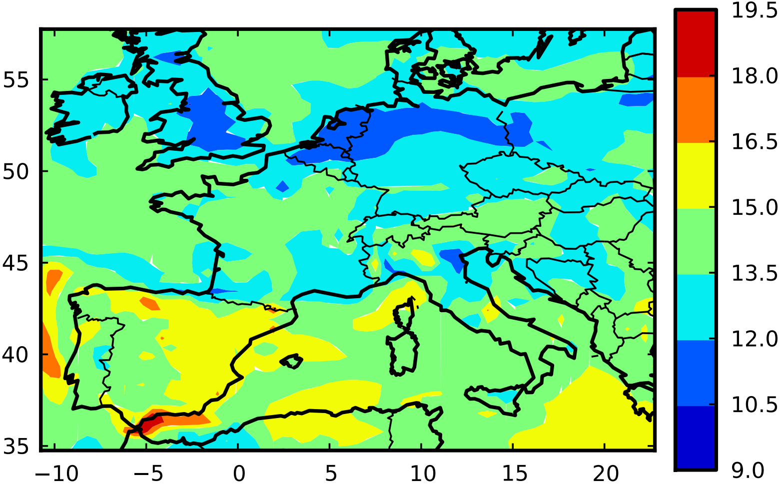
RMSE ($\mu g\ m^{-3}$), With Respect to Analyses and Observations

	Analyses	Observations
Reference model without assimilation	15.8	21.6
Reference model with assimilation	13.5	19.8
EFA with machine learning	11.3	15.6
EFA with filtering	10.9	15.7

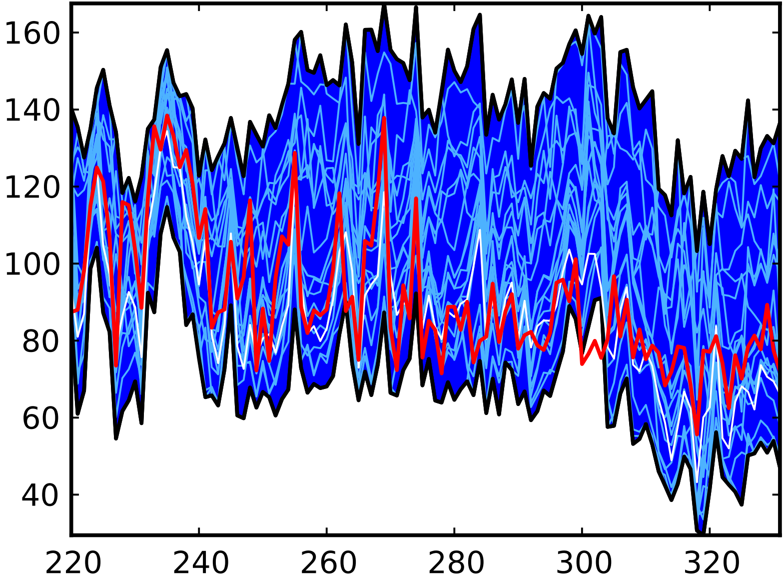
Ozone Maps ($\mu g\ m^{-3}$) Averaged For One Year



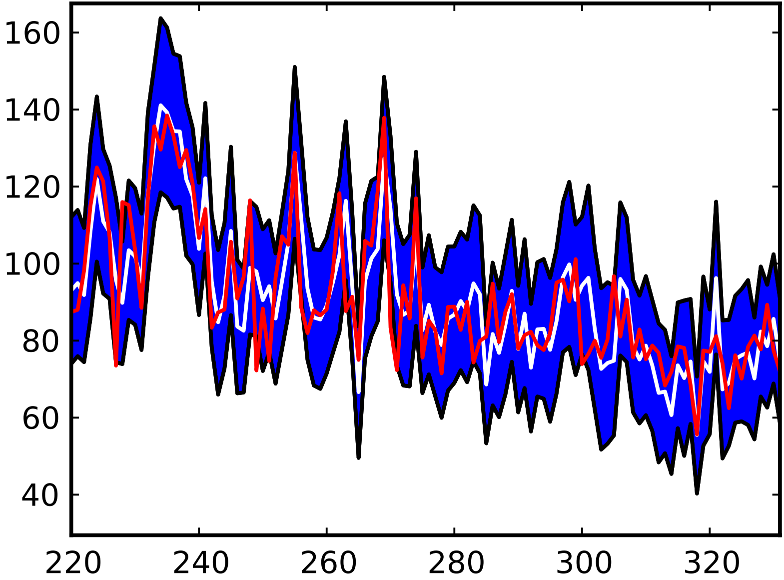
Uncertainty Map (Standard Deviation in µg m^{-3})



Ozone in One Grid Cell ($\mu g\ m^{-3}$)



Ozone in One Grid Cell ($\mu g\ m^{-3}$)



Conclusions

Ensemble Forecast of Analyses

- With machine learning: guaranteed to beat any linear combination with constant weights
- With filtering: access to uncertainty quantification

Some Perspectives

- Machine learning with robust uncertainty quantification
- Some focus on aggregation of fields (with patterns)

- Ensemble forecast of analyses: Coupling data assimilation and sequential aggregation. Mallet, JGR, 2010.
- Ozone ensemble forecast with machine learning algorithms. Mallet, Stoltz & Mauricette, JGR, 2009.

- Air quality simulations with Polyphemus, `http://cerea.enpc.fr/polyphemus/`
- Algorithms from data assimilation library Verdandi, `http://verdandi.gforge.inria.fr/`

Time Evolution of the Weights

Machine Learning and Filtering