A Local Ensemble Transform Kalman Particle Filter

Sylvain Robert and Hans R. Künsch
ETH Zürich, Seminar for Statistics

Outline

- Introduction
- Algorithm
- Numerical experiments
- Conclusions

Motivation

- **Target application:** Convective Scale Data Assimilation
- **Challenges:**
  - high resolution (≈1km)
  - non-linear forecasting step $\boldsymbol{x}_t^a \rightarrow \boldsymbol{x}_{t+1}^b$.
  - non-Gaussian background distribution $\pi_t^b(x)$.
  - non-Gaussian analysis distribution $\pi_t^a(x)$ (even if likelihood is linear and Gaussian).

Introduction

- We assume that $y|x \sim \mathcal{N}(Hx, R)$ and skip time index t.
- $\boldsymbol{x}^b$ and $\boldsymbol{x}^a$: background and analysis ensembles.
- $\pi^b(x)$ and $\pi^a(x)$: background and analysis distributions.

Analysis:

- $\boldsymbol{x}^b$ "+" $y \rightarrow \boldsymbol{x}^a$
- **Bayes' formula:** $\pi^a(x) \propto \pi^b(x) \cdot \ell(y|x)$

Introduction

different assumptions = different solutions

- $\pi^b(x)$ Gaussian + unlimited computation $\rightarrow$ Kalman Filter (KF)
- $\pi^b(x)$ non-Gaussian + unlimited computation $\rightarrow$ Particle Filter (PF)
- $\pi^b(x)$ Gaussian + limited computation $\rightarrow$ EnKF
- $\pi^b(x)$ non-Gaussian + limited computation $\rightarrow$???

The EnKPF in a nutshell

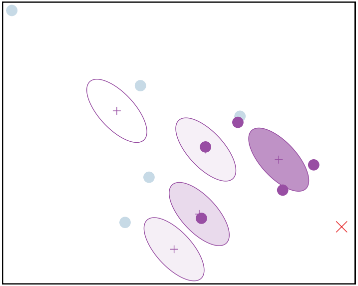
Frei and Künsch (Biometrika, 2013).

- $\pi^a(x) \propto \pi^b(x) \cdot \ell(y|x) = \underbrace{\pi^b(x) \cdot \ell(y|x)^{\gamma}}_{\propto \pi^{\gamma}(x)} \cdot \ell(y|x)^{1-\gamma}$
- Two steps:

$$\pi^b(x) \xrightarrow[\gamma]{\text{EnKF}} \pi^{\gamma}(x) \xrightarrow[1-\gamma]{\text{PF}} \pi^a(x)$$

- $\gamma = 1 \rightarrow$ EnKF
- $\gamma = 0 \rightarrow$ PF

The EnKPF in a nutshell



Analysis distribution $\pi^a_{EnKPF}(x)$:

$$x^{a,j} \sim \sum_{i=1}^{k} \alpha^{\gamma,i} \mathcal{N}(\mu^{\gamma,i}, P^{a,\gamma})$$

Outline

- Introduction
- **Algorithm**
  - **Ensemble space**
  - **Localization**
- Numerical experiments
- Conclusions

Ensemble space

- Split ensembles into mean and deviations:

$$\boldsymbol{x}^b = \bar{x}^b \mathbf{1}' + X^b \qquad \text{and} \qquad \boldsymbol{x}^a = \bar{x}^a \mathbf{1}' + X^a$$

- Use the empirical covariance $P^b = \frac{1}{k-1} X^b (X^b)'$
- **Analysis mean** :

$$\bar{x}^a = \bar{x}^b + X^b m, \qquad m = k \times 1$$

- **Analysis deviations**:

$$X^a = X^b W, \qquad W = k \times k$$

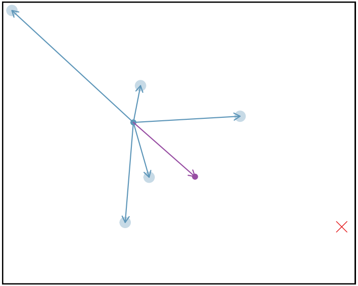
EnKPF in ensemble space

$$m = m^{\mu} + \frac{1}{k} W^{\mu} W^{\alpha} \mathbf{1}$$

$$W = W^{\mu} W^{\alpha} + W^{\epsilon} - \frac{1}{k} W^{\mu} W^{\alpha} \mathbf{1}\mathbf{1}'$$

- m^{μ} and W^{μ}: mixture components $\mu^{\gamma,i}$
- W^{α}: particle resampling
- $\frac{1}{k} W^{\mu} W^{\alpha} \mathbf{1}$: correction of the mean due to resampling.
- W^{ϵ}: individual perturbations to ensure correct covariance:
  - stochastic
  - deterministic: *transform* filter

Step 1: analysis mean

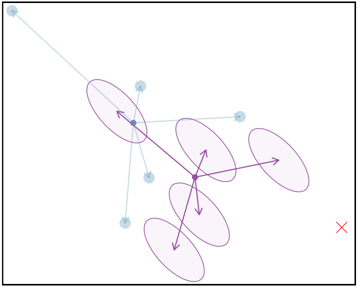


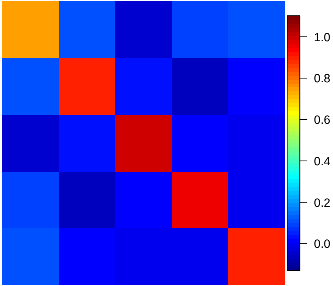
First we move the ensemble mean towards the observation:

$$\bar{x}^a = \bar{x}^b + X^b m^\mu$$

Step 2: mixture components

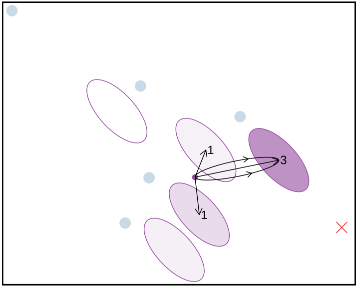
$$(\bar{x}^b + X^b m^\mu)\mathbf{1}' + X^b W^\mu$$

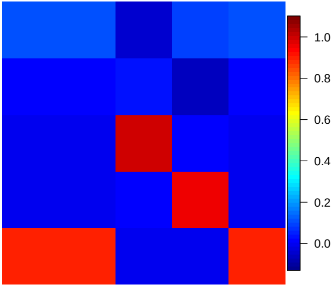




Step 3: weights and resampling

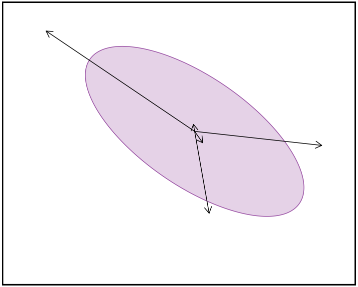
$$(\bar{x}^b + X^b m^\mu)\mathbf{1}' + X^b W^\mu W^\alpha$$

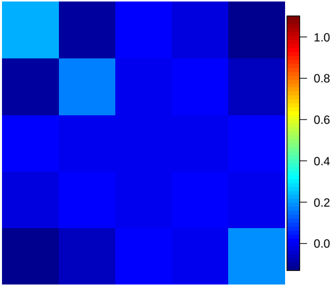




Step 4: individual perturbations

$X^b W^\epsilon \sim \mathcal{N}(0, P^{a,\gamma})$, *Transform* version

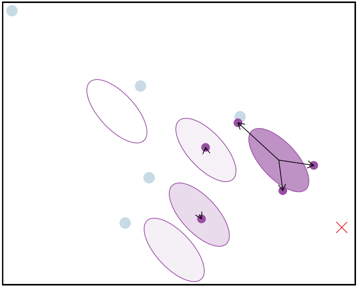


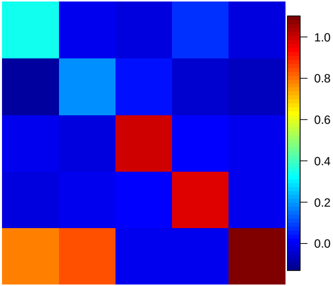


Remark: W^ϵ is not just a square-root of $P^{a,\gamma}$ $\rightarrow$ solve for $Cov(\boldsymbol{x}^a)$.

All together

$$(\bar{x}^b + X^b m^\mu)\mathbf{1}' + X^b W^\mu W^\alpha + W^\epsilon$$





Localization

The curse of dimensionality:

- **PF:** necessary number of particles increases exponentially.
- **EnKPF:** a bit better but not immune to the problem.

Possible remedies:

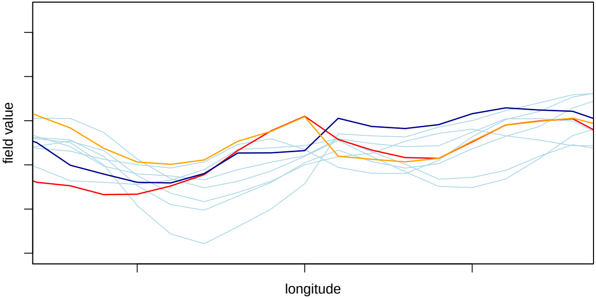
- Carefully chosen proposal distribution.
- **Localization**.

Problem:

- Not easy to apply to PF methods.

Localization: step 1

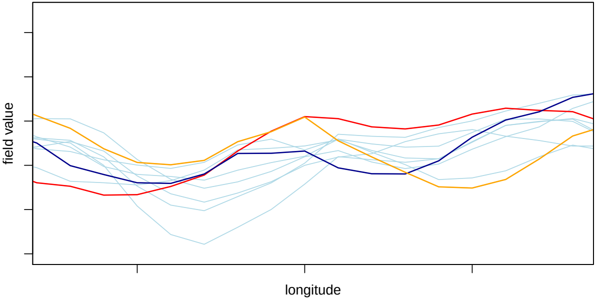
Local weights, resample locally and glue together → discontinuities



Each line is an analysis particle with three cases highlighted in color.

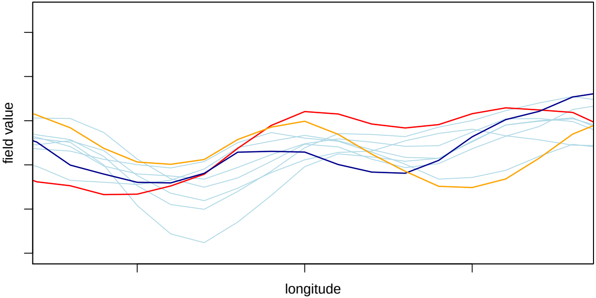
Localization: step 2

Permute indices locally → remove some discontinuities.



Localization: step 3

Interpolate W on finer grid $\rightarrow$ smooth out remaining discontinuities.



Summary of new algorithm

Local Ensemble Transform Kalman Particle Filter: **LETKPF**

- Hybrid between **EnKF** and **PF**: handles non-Gaussian distributions while maintaining sample diversity.
- **Transform:** guarantees exact covariance with deterministic scheme.
- **Local:** uses local weights and local resampling while avoiding problems with discontinuities.

Outline

- Introduction
- Algorithm
- **Numerical experiments: COSMO-KENDA**
  - **Case study (7th of June 2015 12 UTC)**
  - **Cycled experiment (June 04-16)**
  - **Forecast experiment (12 hours)**
- Conclusions

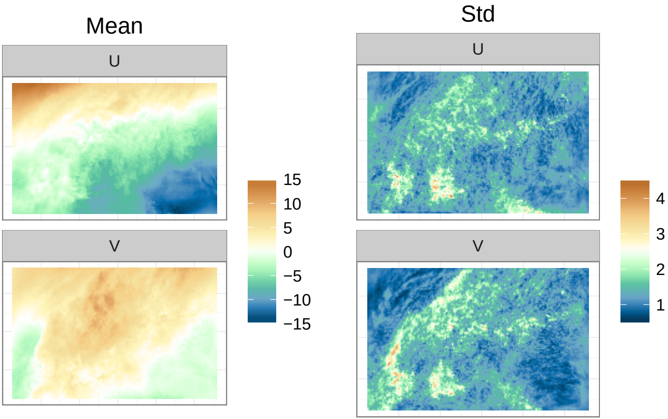
COSMO-KENDA experiment

In collaboration with **Daniel Leuenberger** from Meteoswiss and with the helpful support of **DWD**.

- Area surrounding Switzerland, high resolution ($\approx$ 2.2km).
- 04-16th of June 2015 (period of intense convective activity).
- 40 ensemble members.
- Assimilation of conventional observations (no radar).
- Algorithms: LETKPF and LETKF.

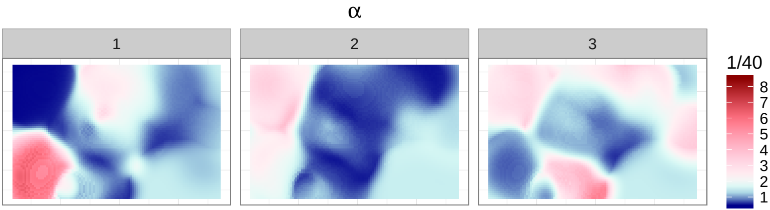
7th of June 12 UTC

Zonal and meridional wind components: ensemble mean and std.



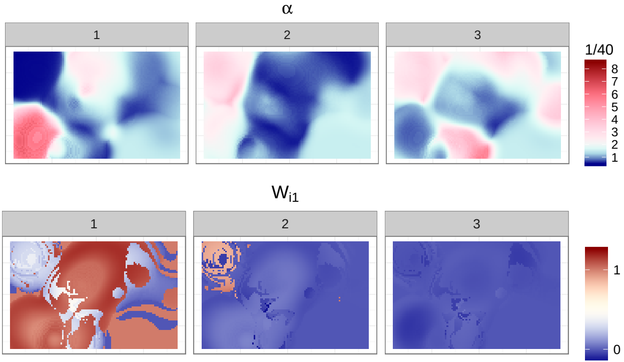
Local particle weights: $\alpha^{\gamma,i}$

- Different particles fit the observations better in different places.
- $1/40 < \alpha < 2/40 \rightarrow$ particle resampled once or twice.
- $2/40 < \alpha < 3/40 \rightarrow$ particle resampled twice or thrice.
- ...



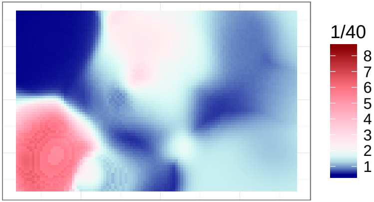
Combination of particles

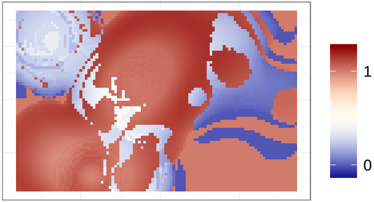
$$(x^{a,1} - \bar{x}^a) = (x^{b,1} - \bar{x}^b)W_{11} + (x^{b,2} - \bar{x}^b)W_{21} + (x^{b,3} - \bar{x}^b)W_{31} + \ldots$$



Discontinuities

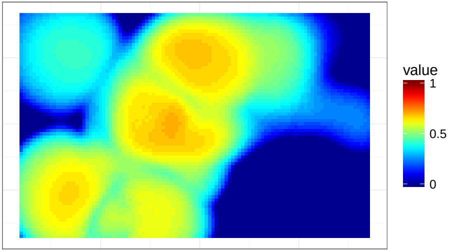
- Contribution of particle 1 to its own analysis: W_{11}.
- Large continuous patches: good.
- Some discontinuous patterns could be fixed, but it would require some global communication (work in progress).





Adaptive choice of γ

- Chosen locally such that ESS=50% ($\approx$ half of the mixture components $\mu^{\gamma,i}$ are used).
- Small γ means more PF, big γ more EnKF.
- Joint property of the background distribution and the observations.



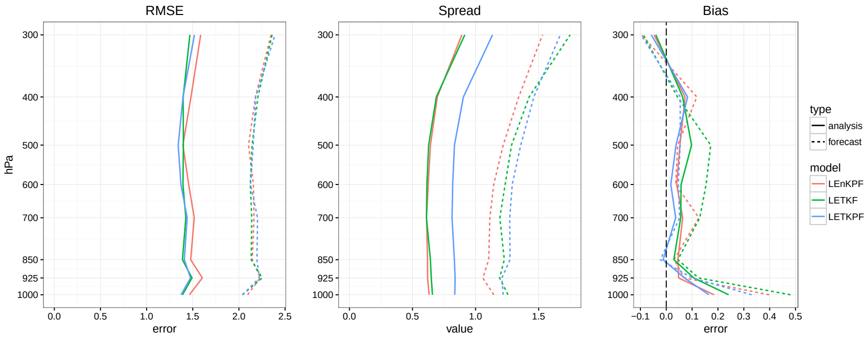
Cycled experiment

- Hourly assimilation of conventional observations.
- LETKPF vs LETKF.
- Vertical profiles of RMSE and spread for T, RH and WIND.
- Averaged over whole period 04-16.06.2015.

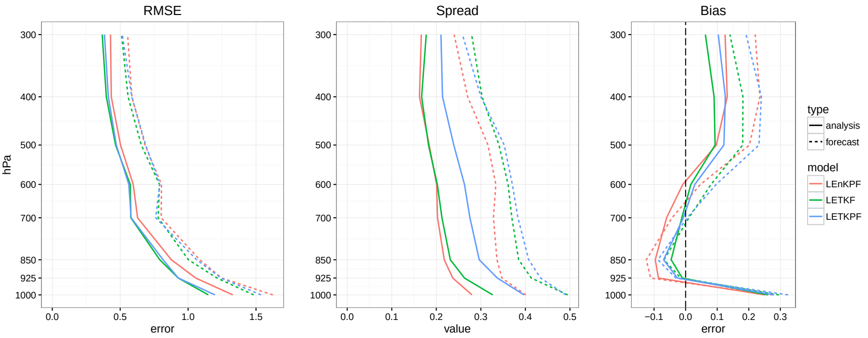
Parameters:

- Parameter γ chosen adaptively such that ESS=50%.
- Localization radius: chosen adaptively.
- Multiplicative covariance inflation.

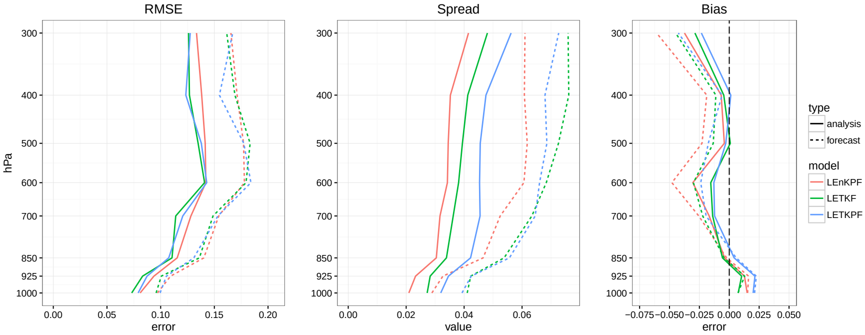
Wind


RMSE
Spread
Bias
hPa
300
400
500
600
700
850
925
1000
error
value
error
0.0 0.5 1.0 1.5 2.0 2.5
0.0 0.5 1.0 1.5
−0.1 0.0 0.1 0.2 0.3 0.4 0.5
type
analysis
forecast
model
LEnKPF
LETKF
LETKPF

Temperature



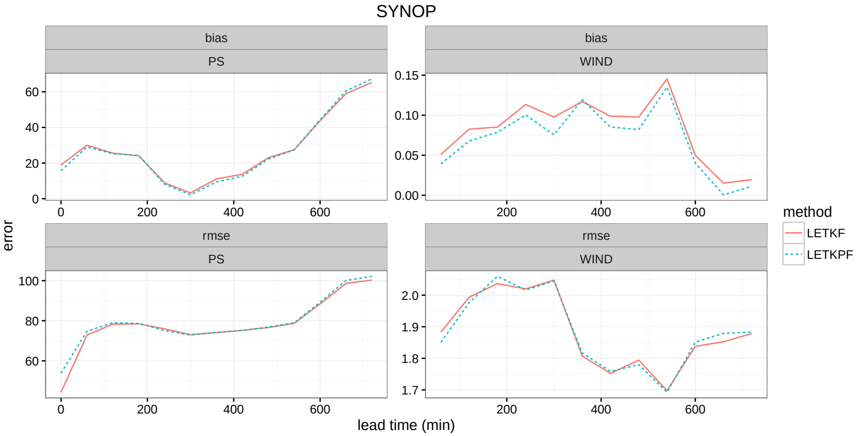
Relative humidity



Outline

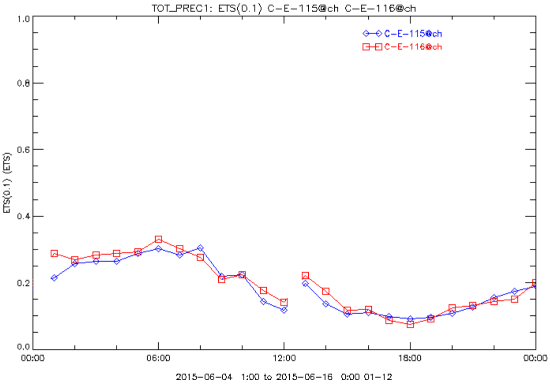
- Introduction
- Algorithm
- **Numerical experiments**
  - Case study (7th of June 2015 12 UTC)
  - Cycled experiment (June 04-16)
  - **Forecast experiment (12 hours)**
- Conclusions

Forecast: RMSE and bias



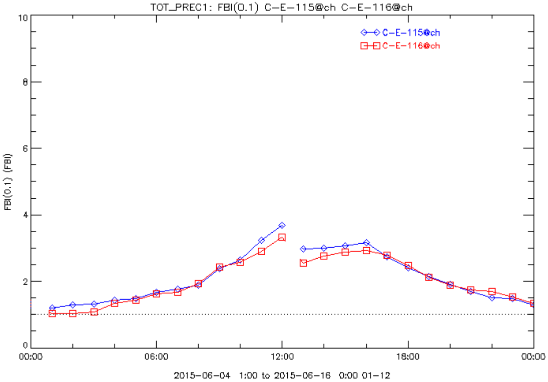
Forecast: ETS of TOT_PREC

How well did the forecast predict rain > 0.1 mm/h (accounting for chance)?



Forecast: FBI

Frequency Bias Index: 1: perfect, >1: over-, <1: under-forecasting.



Conclusions

- New algorithm: **LETKPF**
  - Hybrid EnKF and PF in ensemble space.
  - transform and local.
- Positive results with **COSMO-KENDA**:
  - Case study of 07.06.2015: reasonable behavior.
  - Transform filter better than stochastic version.
  - Results equivalent to LETKF on cycled experiment.
  - Successful 12 hour forecasts.

Thank you!

Questions?