DE LA RECHERCHE À L'INDUSTRIE

www.cea.fr

Improvement of atmospheric dispersion simulations in case of an accident or a malevolent action using data assimilation methods in CERES® CBRN-E

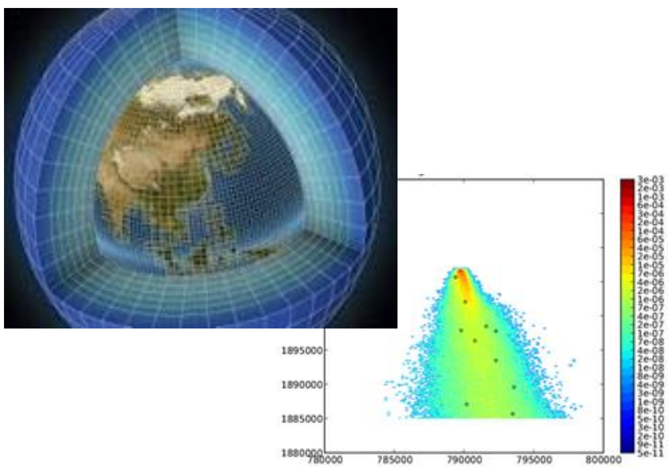
Robin LOCATELLI[3], Vivien MALLET[2], and Patrick ARMAND[1]

[1]Atomic and alternative Energies Commission, France
[2]INRIA, France
[3]STRATHOM, France

Harmo'17 | Budapest (Hungary) | 9-12 May 2016

Context of the study

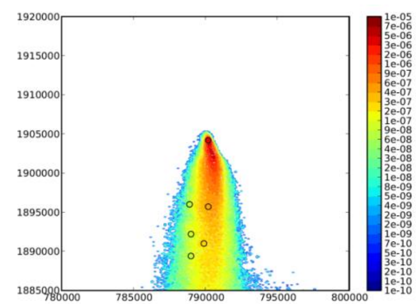
- Modelling of the dispersion at regional or local scale, possibly in built-up environments (e.g. 10 x 10 km domain at a resolution of 100 m or 0.5 x 0.5 km domain at a resolution of 5 m)
- CERES® CBRN-E modelling and decision-support system – A major feature is the choice between different atmospheric dispersion models (MITHRA, SIRANERISK, PMSS…)
- There are several sources of uncertainty in the computations which cannot be ignored: meteorological and source term input data, dispersion and impact assessment models…
- There is a need to reduce these uncertainties for the first responders and decision-makers
  >>>> On idea is to take account of all available information by performing data assimilation!



Flow and dispersion model

Measurements

Data assimilation theory



Optimized state

Choice of a data assimilation method and set up of the study

- Data assimilation - Use of outputs from a dispersion model and atmospheric measurements
- Brief overview of the data assimilation methods
  - Statistical methods: optimal interpolation, Kalman fiter, ensemble Kalman filter…
  - Variational methods: 3D-Var, 4D-Var…
- Differences in the computing time, quality of the optimized state, and system development
- Concentration field optimization without source term modification >>>> Optimal interpolation
- Optimal interpolation is governed by the Best Linear Unbiased Estimator (BEST) equation
- Main goals of this study:
  - Implement the BEST algorithm in conjunction with an atmospheric dispersion model
  - Test the capacity of the BEST algorithm for improving simulations
- « Parallel experience » using observations in a simple case (simple topography and no building):
  - Synthetic observations coming from a reference simulation
  - Perturbations of meteorological conditions and source caracteristics
- Derivation of the reference simulation in the optimized state using the optimal interpolation
- PMSS system (combination of PSWIFT and PSPRAY) but others models could be used…

BLUE matrix equation

Gain matrix

$$X^a = X^b + \underbrace{BH^t(HBH^t + R)^{-1}}_{\text{Gain matrix}}(y_0 - HX^b)$$

Optimized state vector

Prior state vector

Observations vector

Prior state sampled at the observation sites

B : variance / covariance matrix of the errors on the prior

R : variance / covariance matrix of the errors on the observations

H : observation operator (relates the state vector **X** with the observations $\mathbf{y_0}$)

BLUE matrix equation (cont'd)

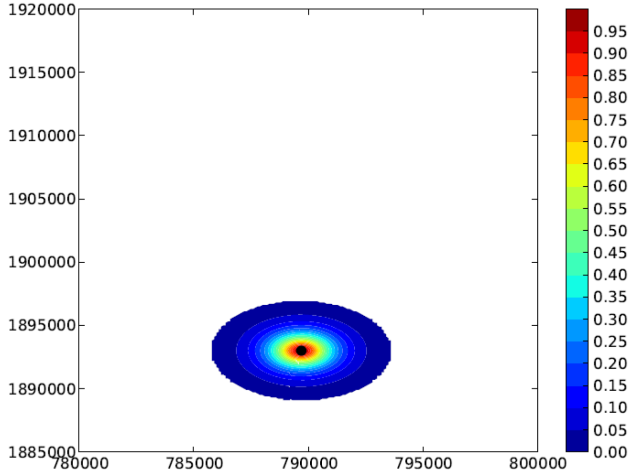
$$\mathrm{X^a = X^b + BH^t(HBH^t + R)^{-1}(y_0 - HX^b)}$$

- How $\mathbf{y_0}$ is built?
  - Localisation of the receptors: randow drawing in the (x,y) domain
  - Sampling of the simulation reference at the receptor localisations
- How **R** is built?
  - **R** is supposed to be diagonal
  - Knowledge on the instrumental errors is needed to determine the variances
- How **B** is built?
  - All sources of uncertainties between $\mathbf{x}^t$ (true state) and $\mathbf{x}^b$ (prior state) must be accounted!
  - Uncertainties come from the source location, wind direction and speed, turbulence model…
  - A satisfactory **B** matrix is essential for the quality of the optimisation!
  - But, how to determine **B**? Analytical formulations? Others methods?…

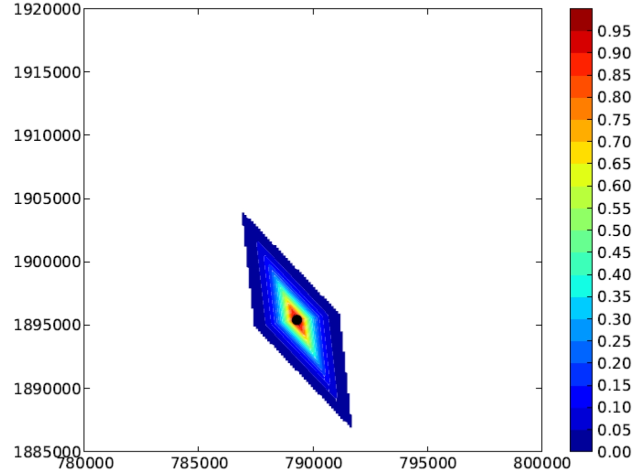
Analytical formulations for the B matrix

Correlation between the grid points in the domain and the reference black point

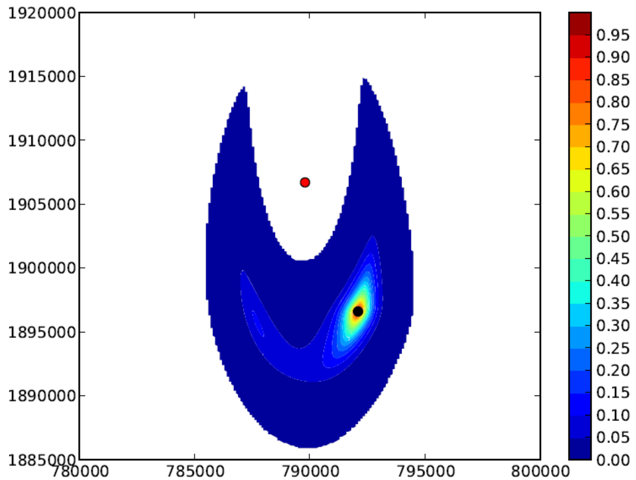
1 – Balgovind formula



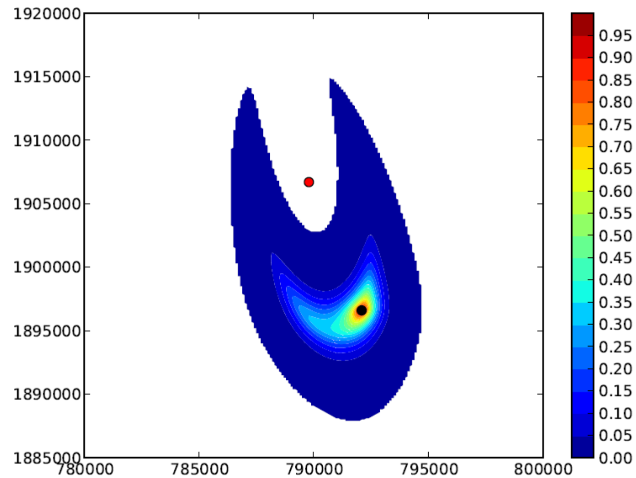
2 – Frydendall formula



3 - « Plume front » formula



4 - « Plume front » + Frydendall



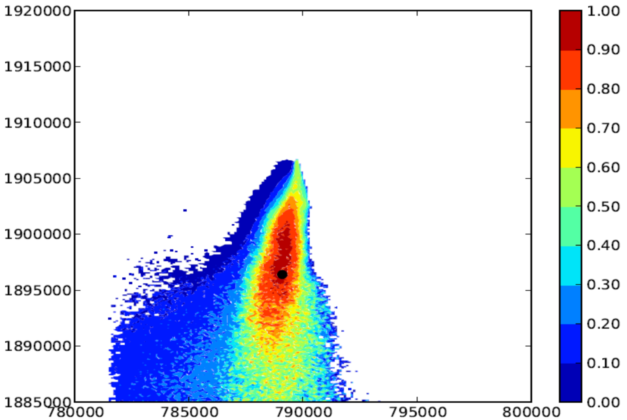
- Analytical formulations are not satisfactory:
  - Need to determine several parameters
  - Dependant to the study case!

Ensemble approach for the B matrix

- B is determined using an ensemble of perturbed simulations
- The pertirbated simulations are created by pertubing the meteorological conditions, the source characteristics, and the numerical choices
- Quantifying discrepancies between perturbed simulations is a way to sample the **B** matrix
- B is computed according to the following equations:

$$B = \frac{1}{N_{members} - 1} \sum_{m=1}^{N_{members}} (x^m - \bar{x})(x^m - \bar{x})^T \qquad \bar{x} = \frac{1}{N_{members}} \sum x^m$$

Correlation between all grid points in the domain and one observation site in black



Computations of the B and R matrices

Several methods exist for specifying variances of error matrixes

1 - Desroziers method:

$$E[d_b^o(d_b^o)^t] = R + HBH^t$$
$$E[d_b^a(d_b^o)^t] = HBH^t$$
$$E[d_a^o(d_b^o)^t] = R$$

E[ ] : mathematical expectation operator

$$d_b^o = y_o - Hx^b$$
$$d_a^o = y_o - Hx^a$$
$$d_b^a = Hx^b - Hx^a$$

$$\sigma = \frac{\text{variances}_{\text{diagnosed}}}{\text{variances}_{\text{specified}}}$$

If $\sigma < 1$, variances need to be decreased

If $\sigma > 1$, variances need to be increased

2 - χ^2 approach:

$$\chi^2 = (y_o - Hx^b)^t \times R + HBH^{t^{-1}} \times (y_o - Hx^b)$$

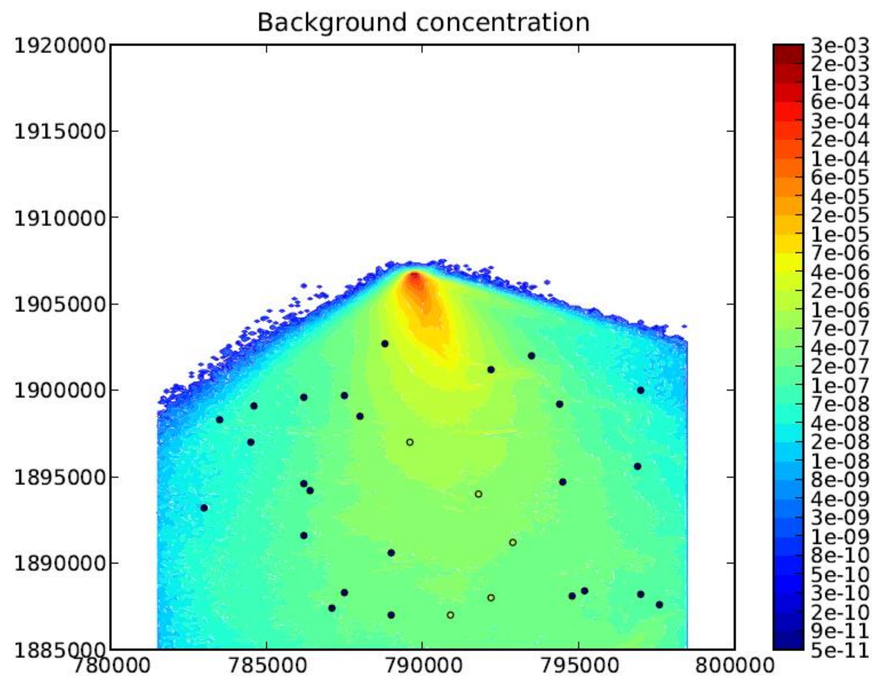
If **R** and **B** are correctly specified, the χ^2 value is in theory equal to the number of observations!

BLUE matrix equation (cont'd)

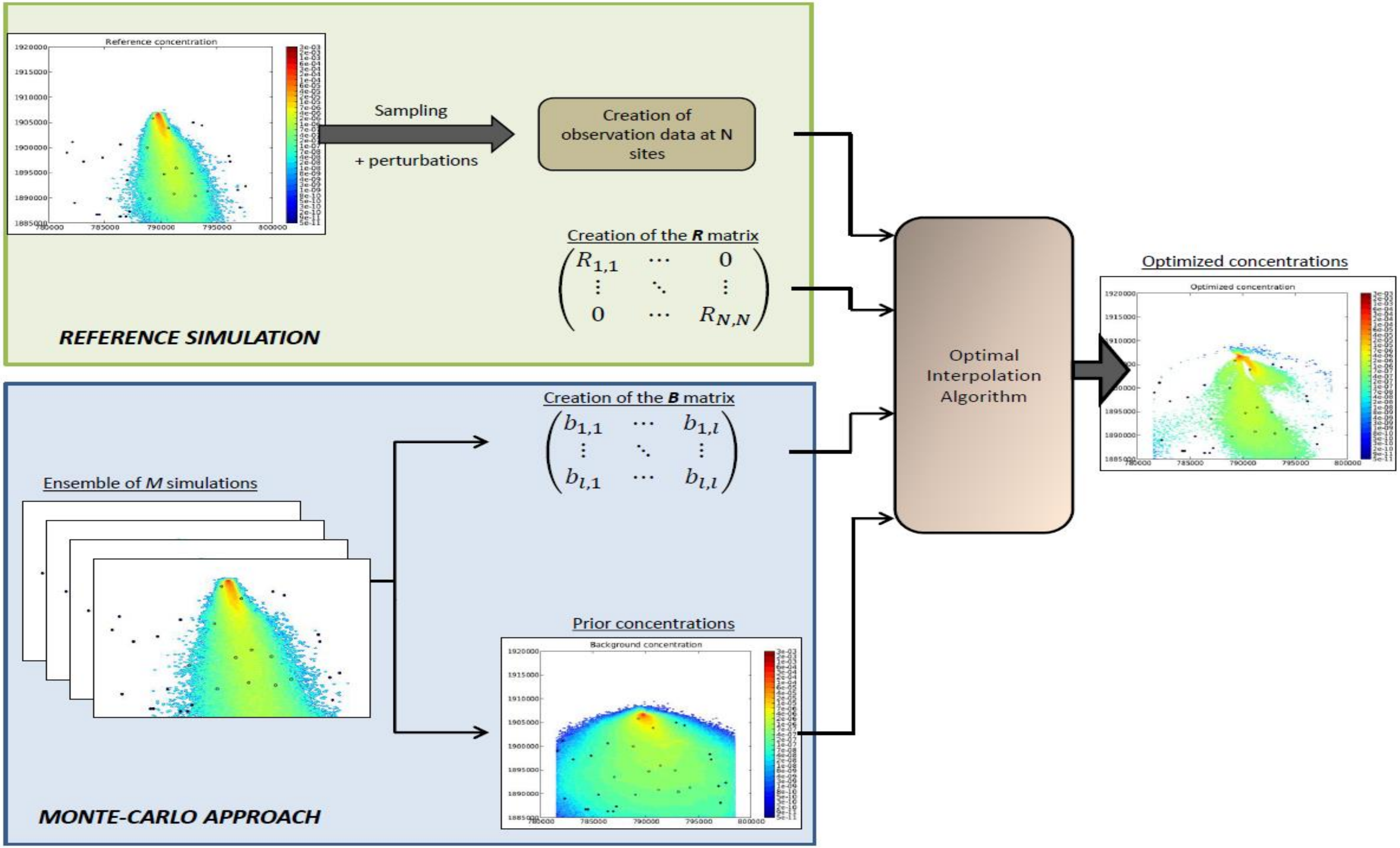
$$X^a = X^b + BH^t(HBH^t + R)^{-1}(y_0 - HX^b)$$

- How **B** is built?
  - **B** is built as the average field of the generated ensemble
- How **H** is built?
  - **H** is the observation operator whose dimension is (N_{obs}, N_{pts})
  - $H_{i,j} = 1$ if the i^{th} observation is located on the j^{th} point of the domain and $H_{i,j} = 0$ otherwise

Prior concentration field is here >>>>



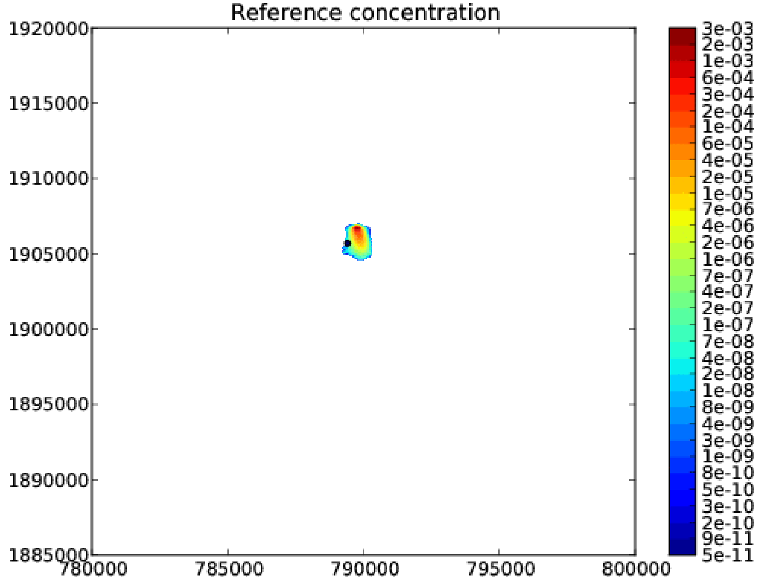
Synoptic view of the method

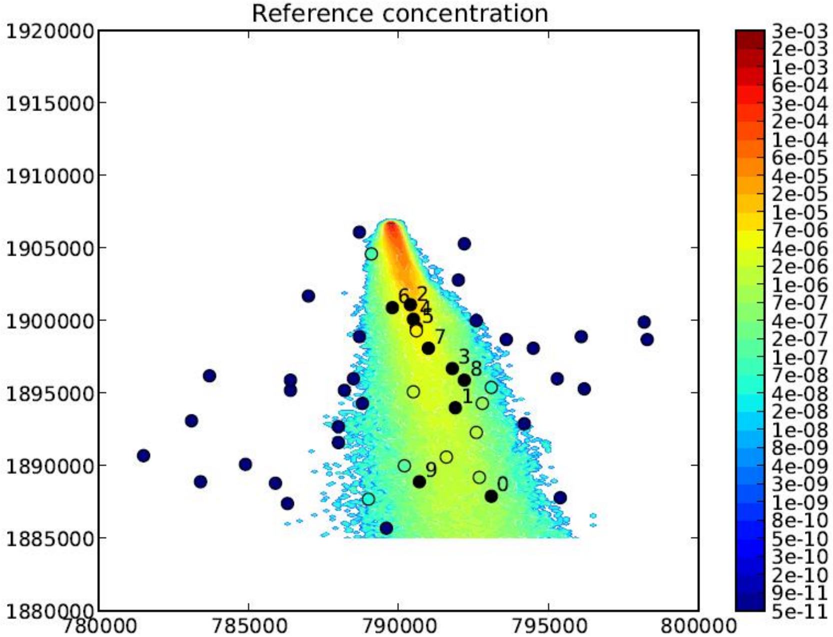


Simulation parameters and definition of the receptors

- Simulation the 12th of September 2011 between 10 am and 11 am
- Output frequency: 300 sec (5 minutes)
- Domain: 17 x 35 km – Horizontal resolution: 100 meters
- Non-reactive tracer emitted by a source whose characteristics are:

  T = 50°C ; w = 12 m/s ; H = 15 m ; Q = 1.2 10^5 Bq emitted



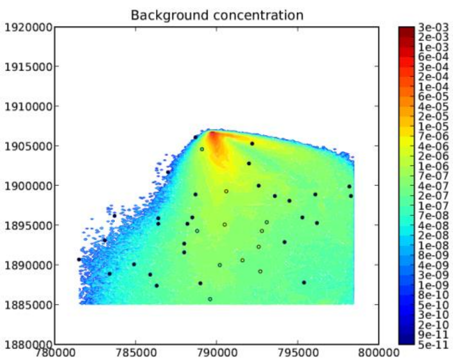


- Assimilation process is tested with 10 to 50 receptors (here, 40)
- 10 « independant » receptors are also used for cross-validation (in black with the numbers)

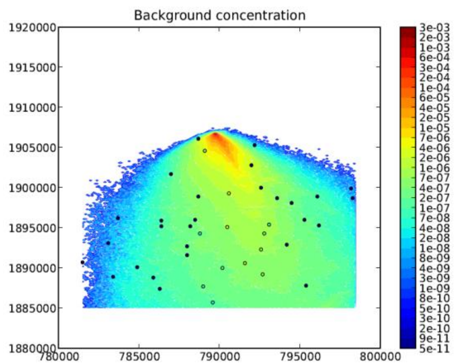
Background concentration field

Background concentrations are computed as the average concentrations of the ensemble

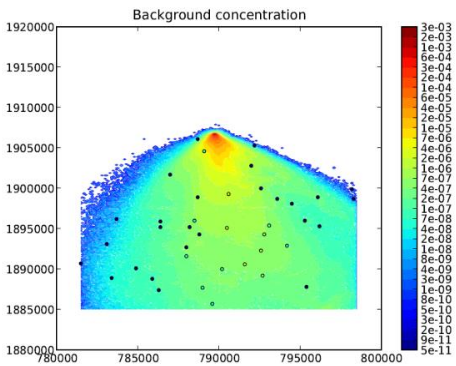
10 members



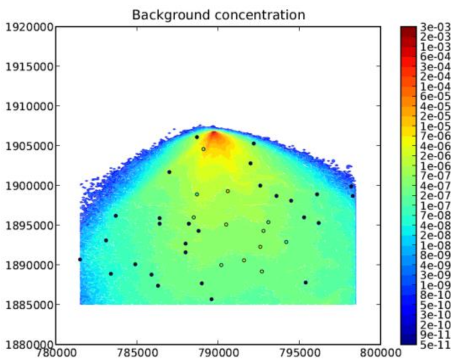
20 members



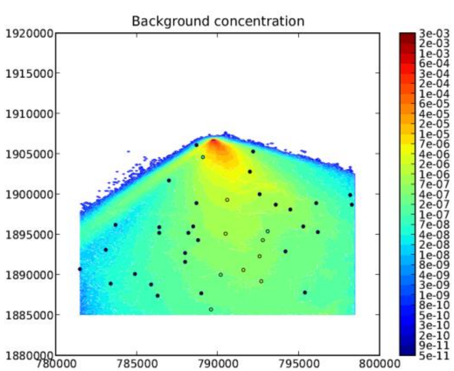
30 members



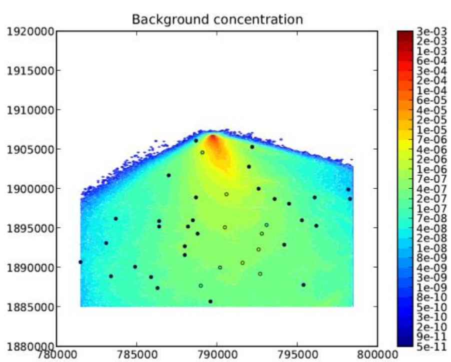
40 members



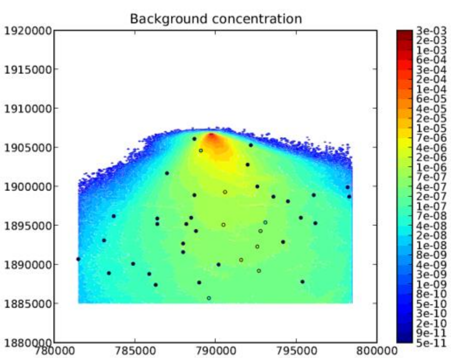
50 members



60 members



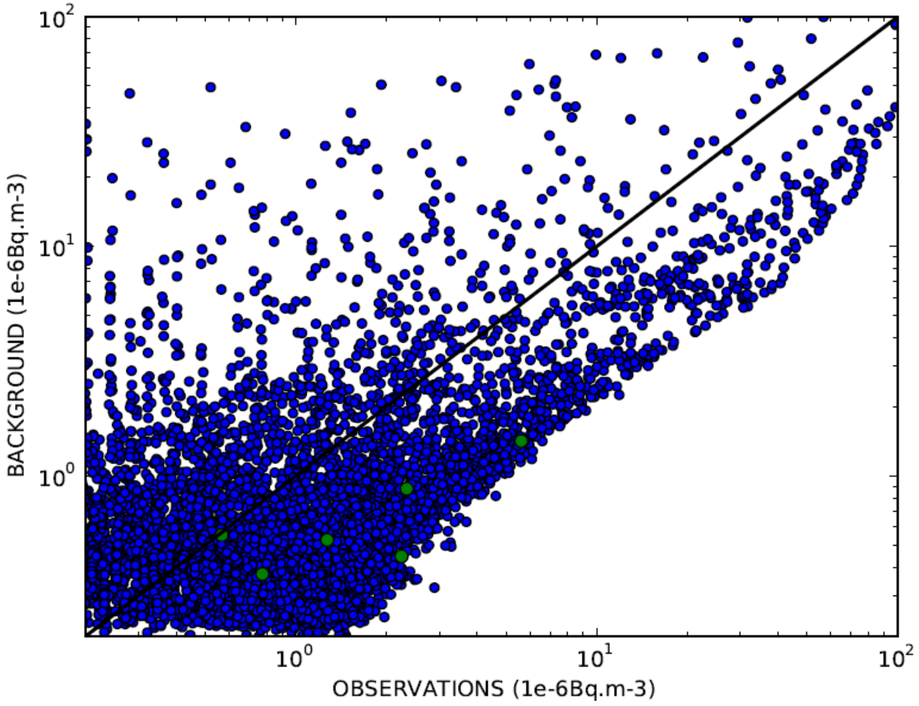
120 members

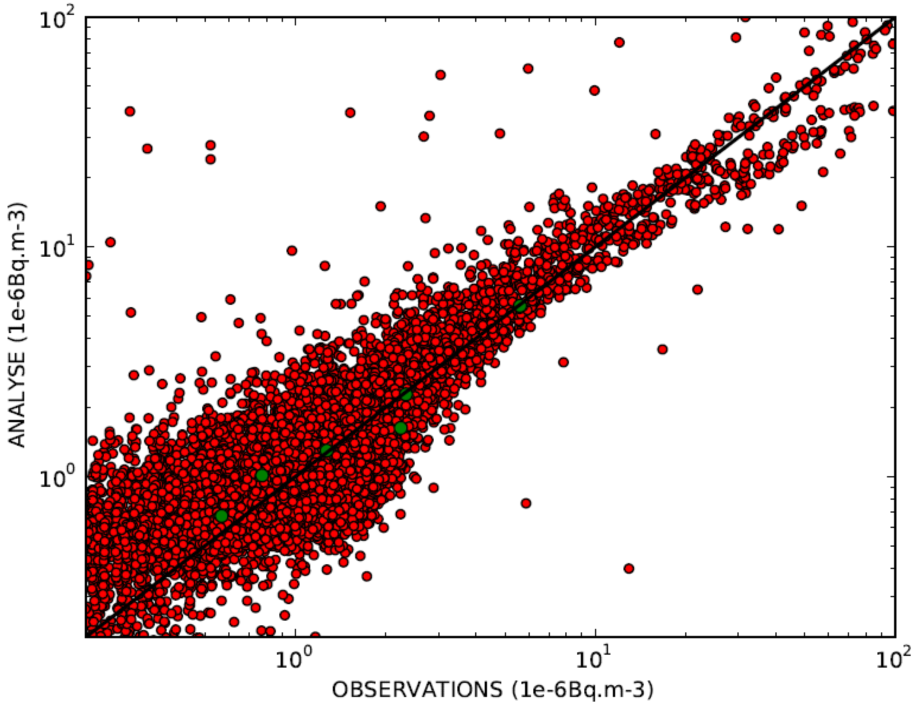


Scatter plots – Prior versus analyse

The distribution of analyzed concentrations fit much better…

The distribution of measured concentrations

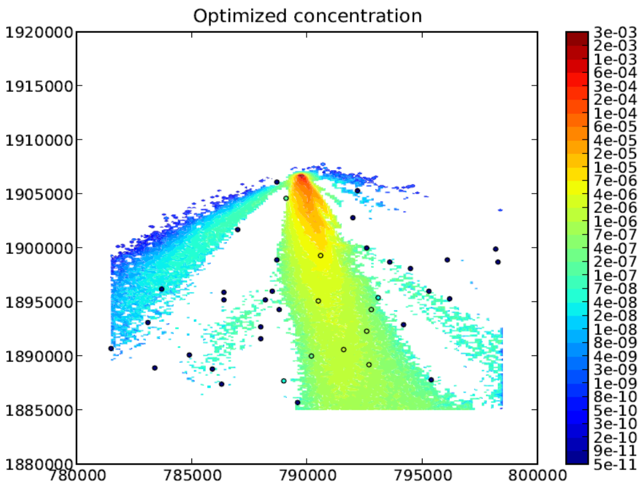


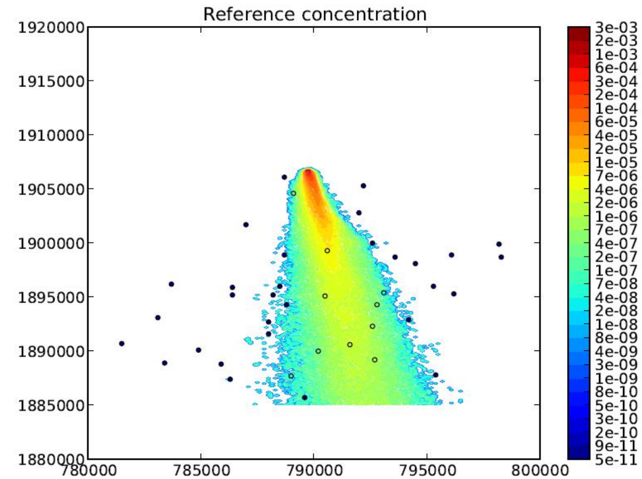


Optimized versus reference concentration

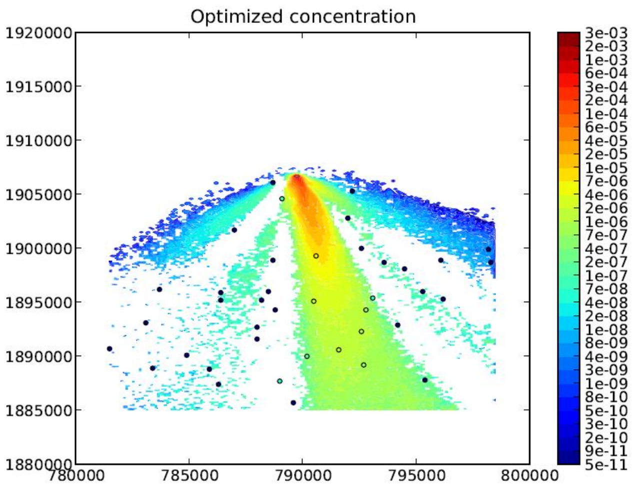
Spatial distribution of the optimized and reference concentrations at two different time steps

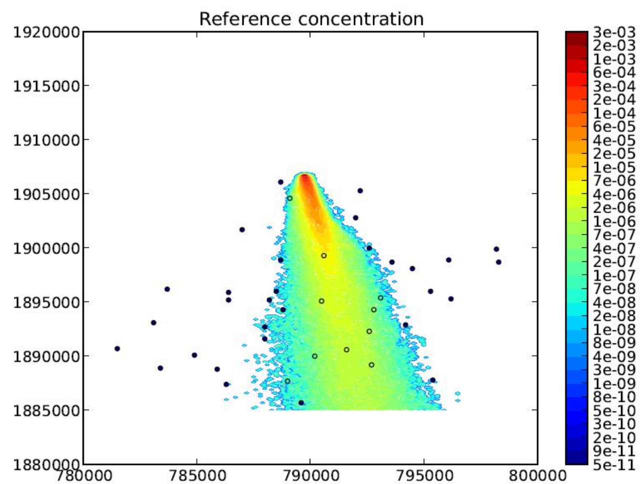
Time step #11





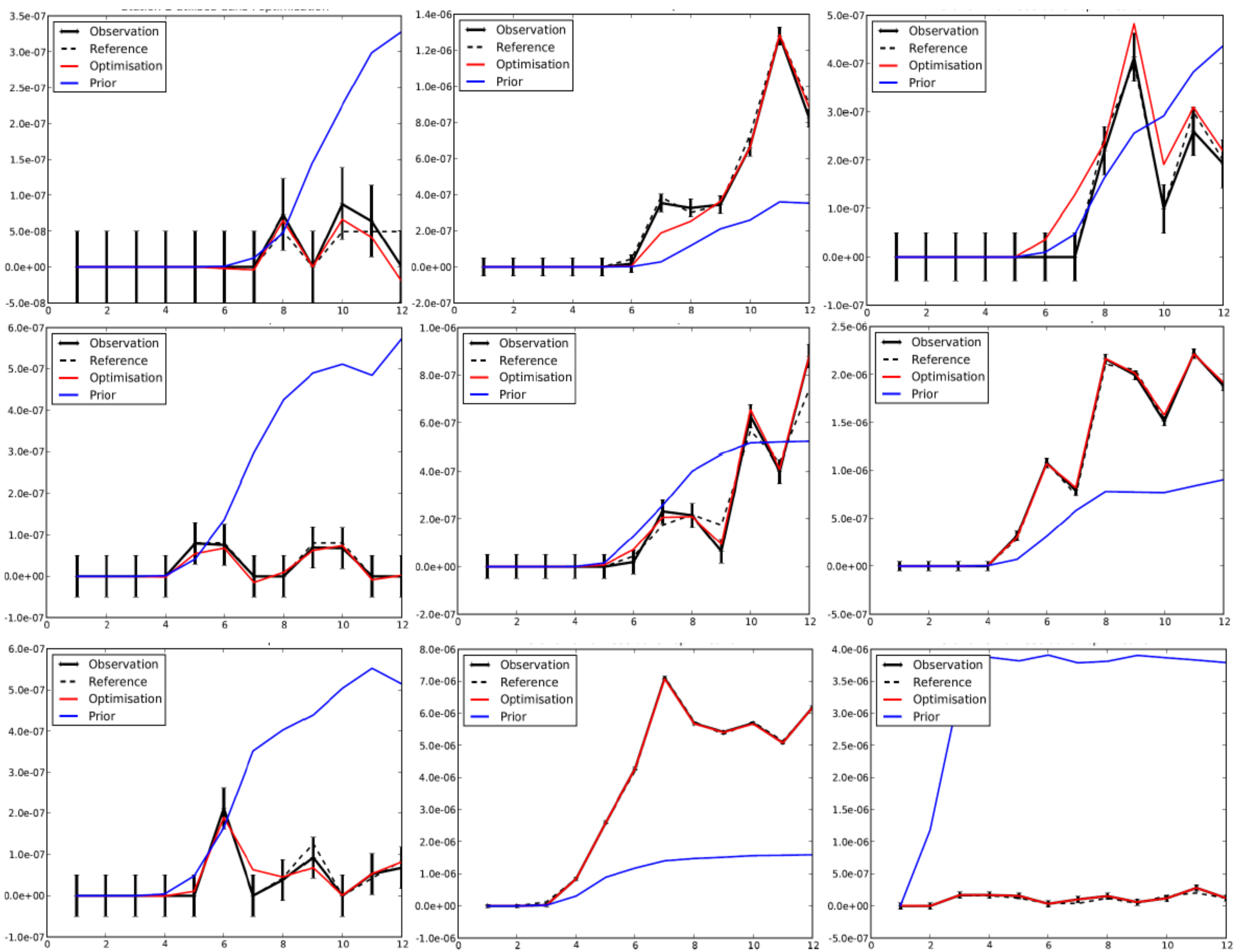
Time step #12





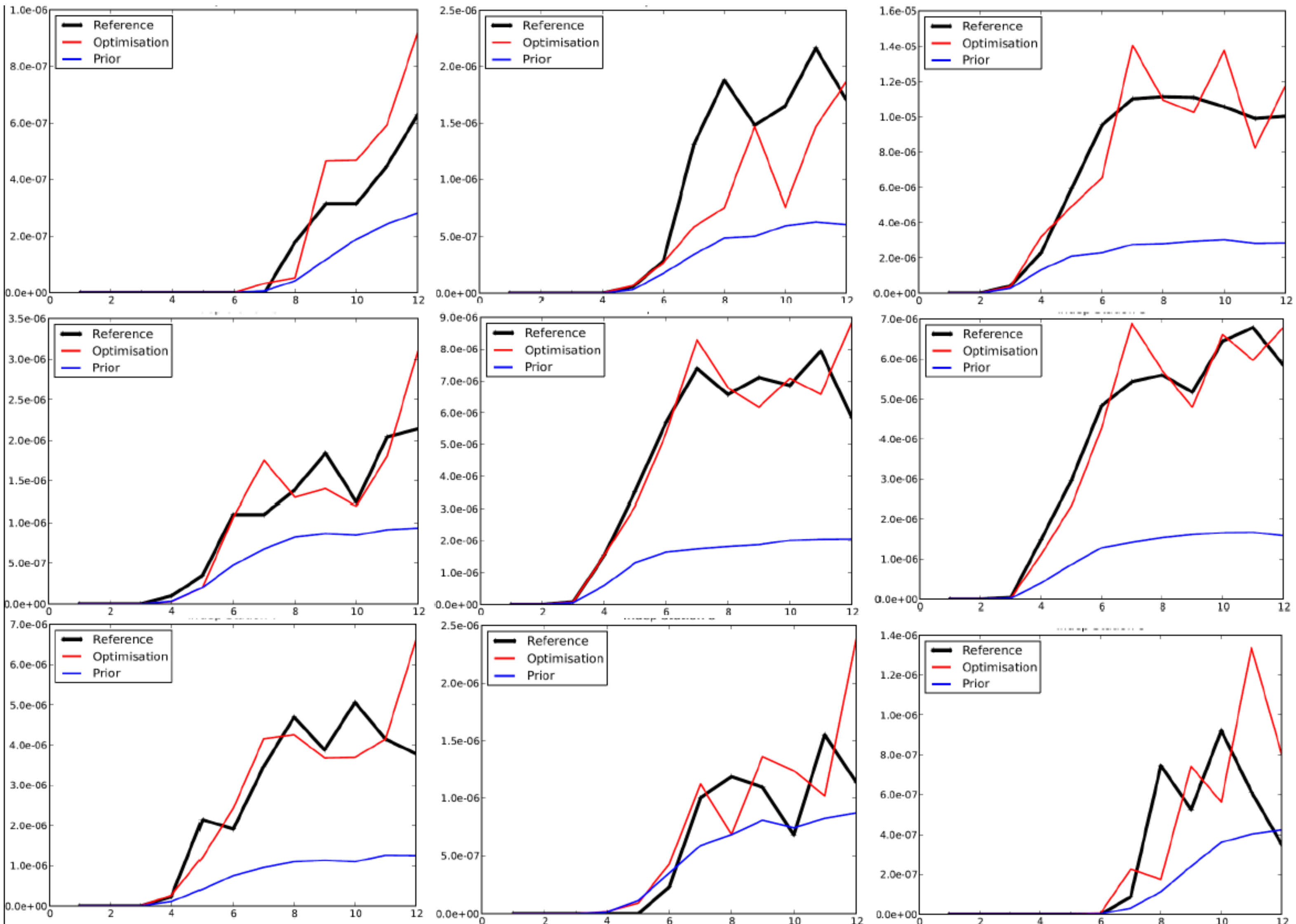
DE LA RECHERCHE À L'INDUSTRIE

Time series at the assimilated stations



- Good skills in the optimized concentrations:
  - Able to derive the correct order of magnitude
  - Able to reproduce the time variations
- However, these concentrations are mathematically built to fit the measurements!
- What about time series at no assimilated station?

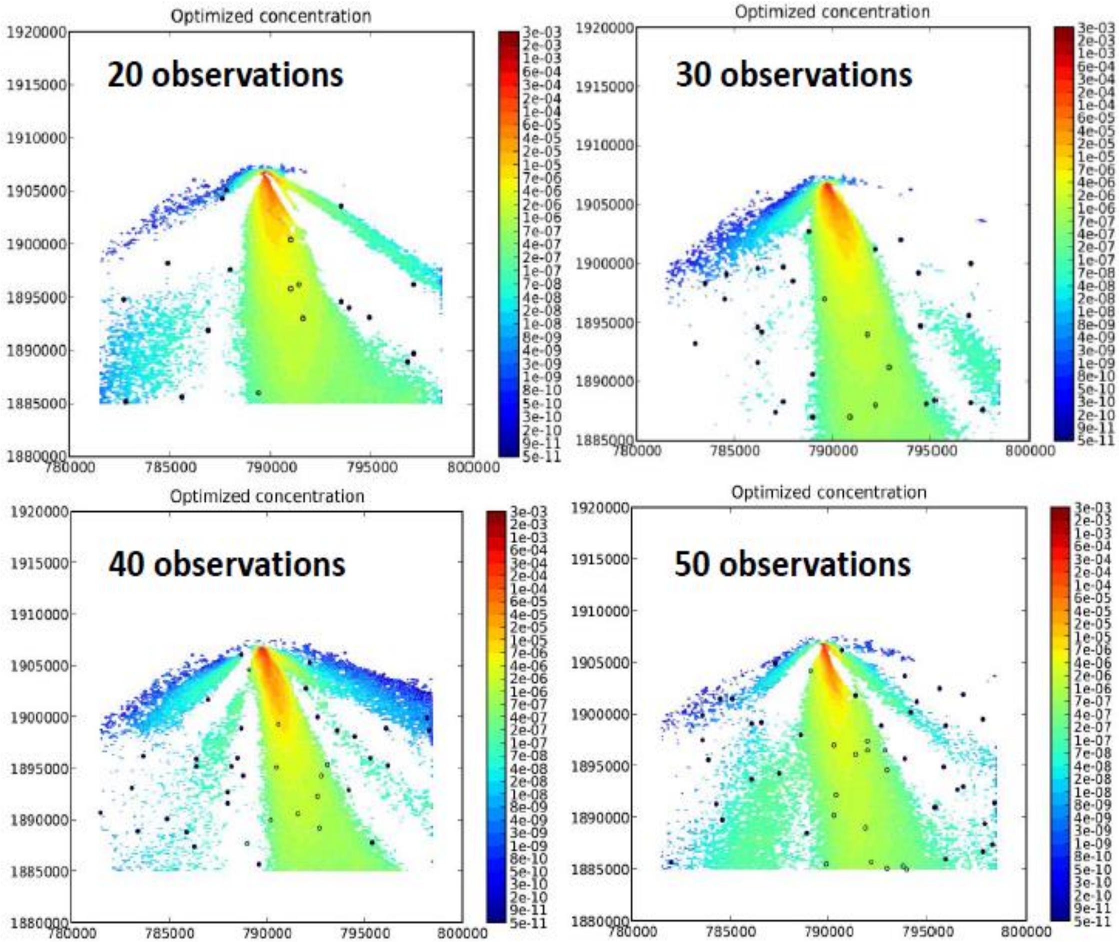
Time series at the not assimilated stations



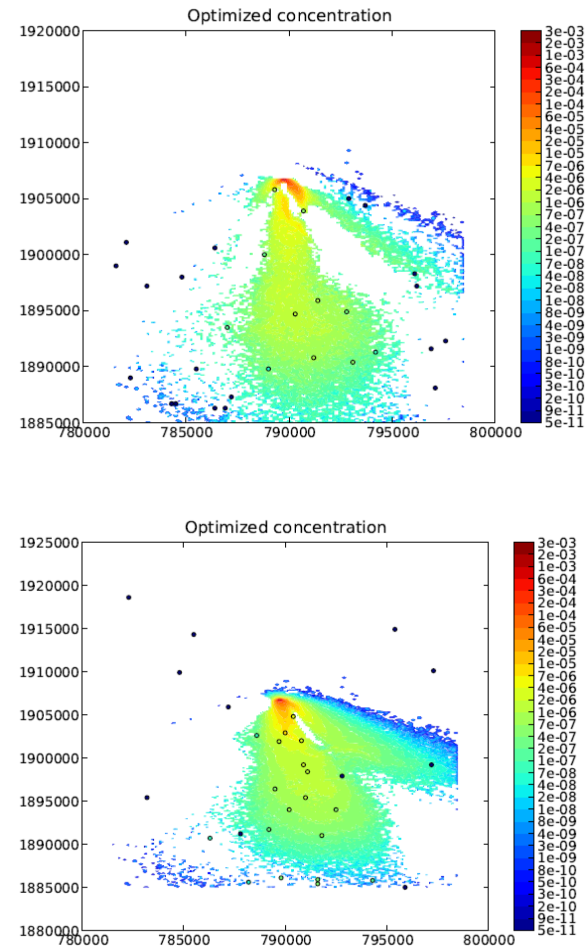
Outside the assimilated stations, the time series of the optimized concentrations fit quite well the reference concentration!

Sensitivity to the observations

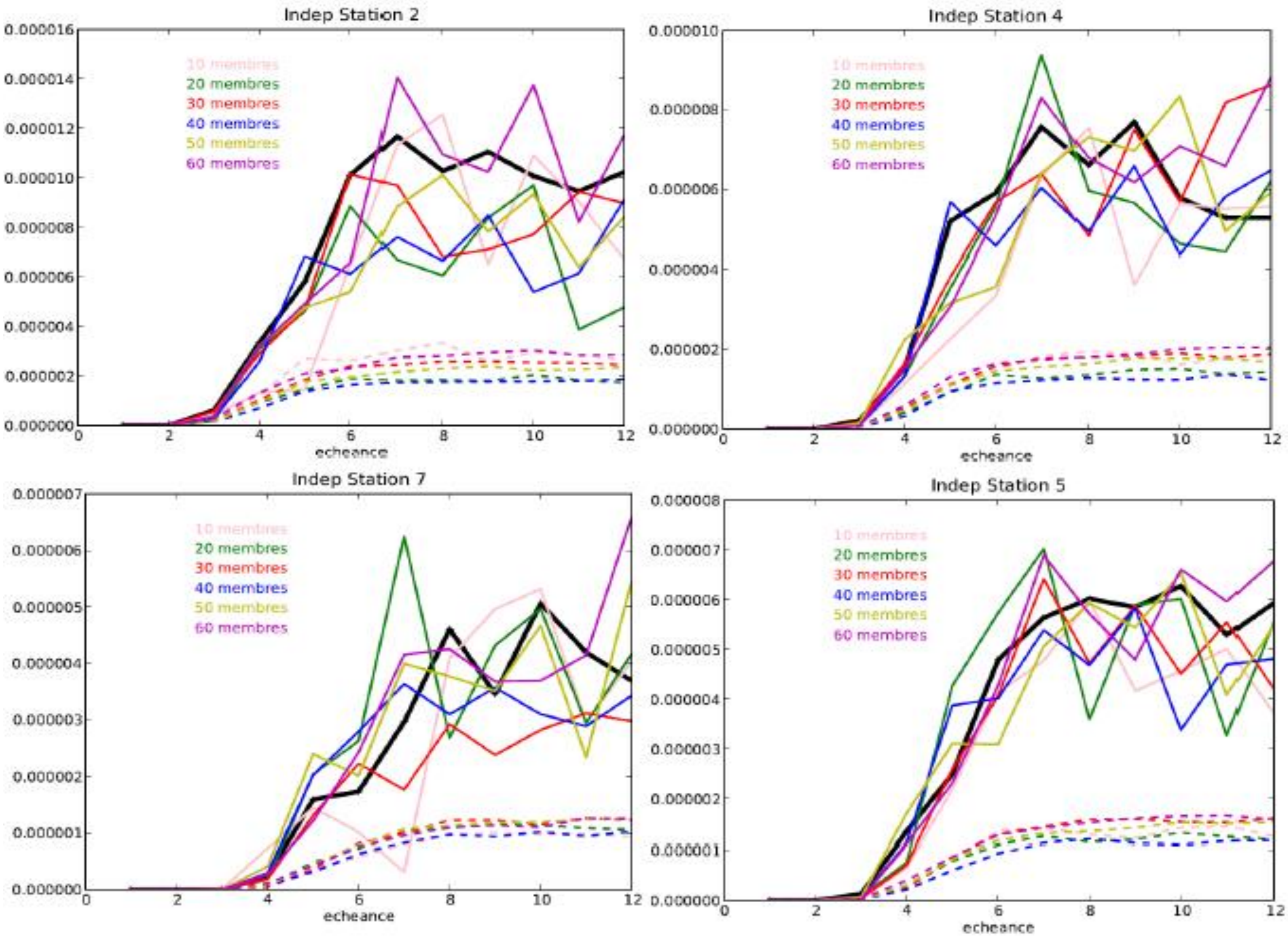
Sensitivity to the observation numbers



Sensitivity to the observation localisations



Sensitivity to the number of members



- All the configurations derive optimized concentration closer to the reference concentrations compared to the prior concentrations
- Large range in the optimized concentrations

Conclusions and outlooks

- Computing time is an essential criterion! (Figures in our simple test-case…)
  - Creation of the ensemble of N simulations using N cores >> 3 min.
  - Reformatting of output files >> 5 min. for 60 members
  - Optimal interpolation combined with the ensemble approach >> 2 min. (Short time!)
- Regarding data assimilation, the optimal interpolation method has several advantages:
  - Good quality of the optimized state
  - Computing time is fast
- One challenging issue is to find the best formulation for the **B** matrix
  - Analytical formulations are not satisfactory enough
  - The ensemble approach is an objective way to compute the elements of **B**
- The system is able to derive optimized concentrations in agreement with the reference ones

  >>>> The optimal interpolation with the ensemble approach for **B** gives encouraging results!
- Much more work is needed!...
  - To test our algorithm in real cases with complex topography, buildings, complex meteorological situations…
  - To find the optimal configuration (number of observations, members in the ensemble, etc.)

Questions?

Corresponding author: Patrick ARMAND
Commissariat à l'énergie atomique et aux énergies alternatives
Centre DAM Île-de-France – Bruyères-le-Châtel | DASE / SRCE
Laboratoire Impact Radiologique et Chimique
91297 Arpajon CEDEX
T. +33 (0)169 26 45 36 | F. +33 (0)1 69 26 70 65
E-mail: patrick.armand@cea.fr