Urban air quality simulation

Vivien Mallet[1,2]
With contributions by
Anne Tilloy[1,2] Raphaël Périllat[1,2] David Poulet[3]
Fabien Brocheton[3] Frédéric Mahé[4] Pierre Pernot[4] Fabrice Joly[4]

[1]INRIA

[2]CEREA, joint laboratory École des Ponts ParisTech - EDF R&D, Université Paris-Est

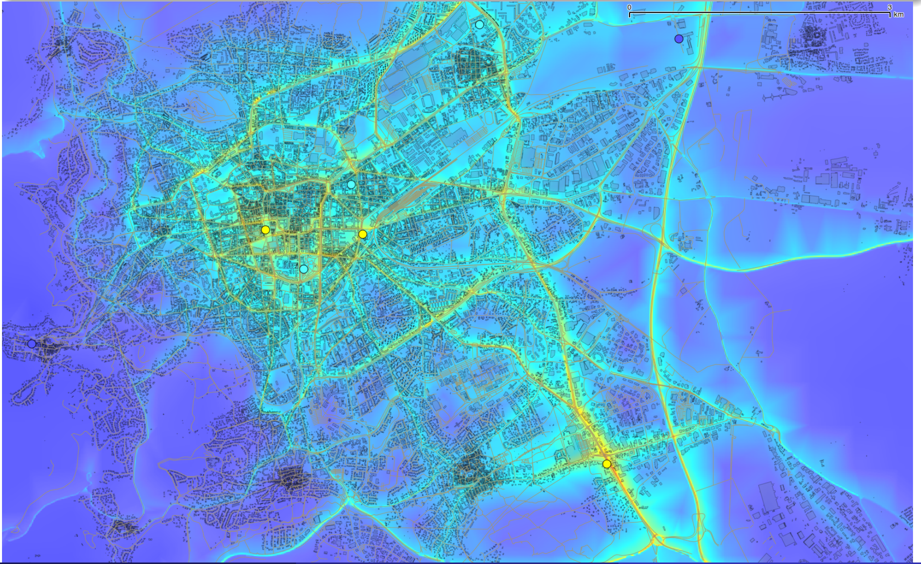
[3]Numtech

[4]Airparif

Berkeley–Inria–Stanford Workshop, Stanford University, May 2013

Quick introduction to urban air quality simulation

Simulation of pollutant concentrations over a city with street resolution.



Urban air quality simulation: what for?

Objectives

1. Evaluating the air concentrations of NO_2, PM_{10}, O_3, ...
   - Analyzing: exposure of population for one or several past years
   - Forecasting: for the next few days
2. Supporting decision making
   - Characterizing: emission sources, local versus regional pollution
   - Testing: scenarios of emissions reduction, new roads or industrial facilities

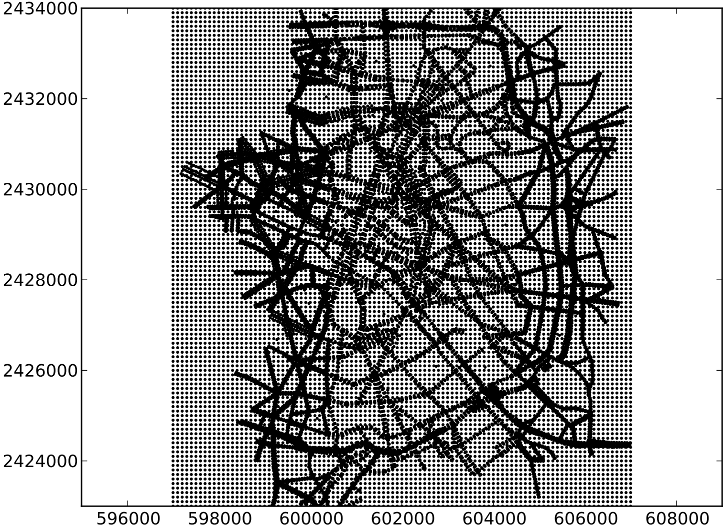
Simulation tools: numerical models with street resolution

Classical model: ADMS Urban

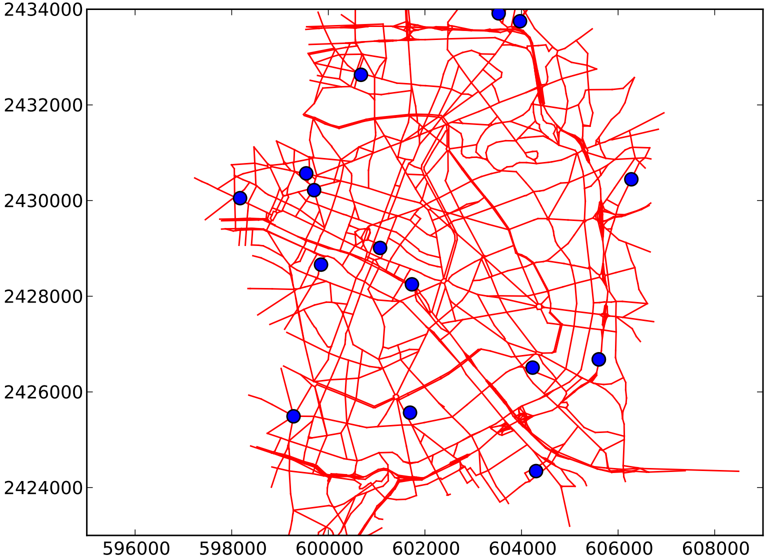
1. Computing the stationary solution of the reactive transport equation
   - Every point source creates a plume, with Gaussian shape crosswind
   - Parameterization for the standard deviations depending on meteorological variables
   - Special treatment within the streets
2. Inputs
   - Time-dependent: spatial distribution of emissions, background pollutant concentrations, meteorological variables (one value for the whole domain)
   - Time-independent: street network
3. High computational costs
   - $\sim 10\,\text{min}$ of computations for a single date, i.e., $\sim 4\,\text{h}$ for a full day

Simulation tools: numerical models with street resolution

Output points of ADMS Urban for Paris (east part)



An important source of information: observations



Merging model outputs and field observations

Data assimilation classical assumptions

- The error on the simulated concentration vector $\boldsymbol{x}^b$ has mean $\mathbf{0}$ and variance $\mathbf{B}$
- The observation vector $\boldsymbol{y}$ can be compared with $\mathbf{H}\boldsymbol{x}^b$ where $\mathbf{H}$ is called the observation operator
- The error on the observation vector $\boldsymbol{y}$ has mean $\mathbf{0}$ and variance $\mathbf{R}$
- No correlation between simulation and observational errors

BLUE: best linear unbiased estimator

- BLUE is the linear estimator $\boldsymbol{x}^a = \mathbf{L}\boldsymbol{x}^b + \mathbf{K}\boldsymbol{y}$ whose error has mean $\mathbf{0}$ and variance $\mathbf{A}$, so that $\mathbf{A}$ has minimal trace
- BLUE reads

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \mathbf{K}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x}^b), \text{with}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^\top(\mathbf{H}\mathbf{B}\mathbf{H}^\top + \mathbf{R})^{-1}$$

Parameterization for the error variances

Observational error

- Observational error variance: $\mathbf{R} = r\mathbf{I}$

Simulation error

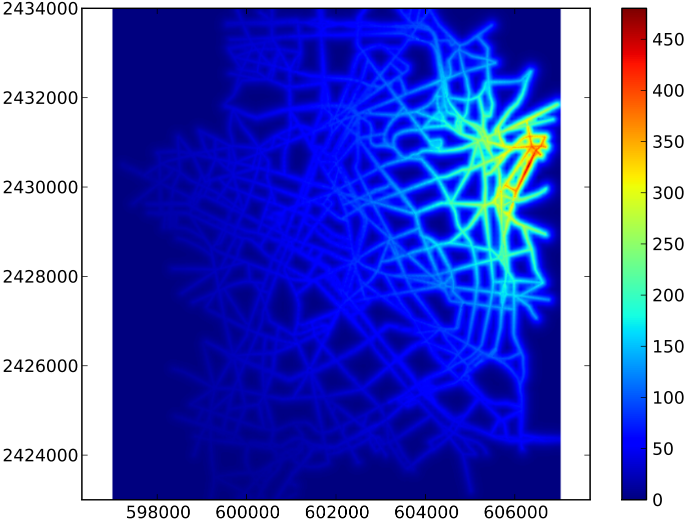
- Simulation error covariance: $B_{ij} = b \exp\left(-\frac{d_{ij}}{L_d}\right) \exp\left(-\frac{|P_i - P_j|}{L_p(i,j)}\right)$
- d_{ij}: distance, along the network, between the projections on the network of the output points i and j
- P_i: distance to the road network
- L_d and $L_p(i,j) = L_p + \alpha \min(P_i, P_j)$: decorrelation lengths

Determination of the parameters

- Statistical study of $\boldsymbol{y} - \mathbf{H}\boldsymbol{x}^b$, whose variance should be $\mathbf{HBH}^\top + \mathbf{R}$
- Leave-one-out cross-validation

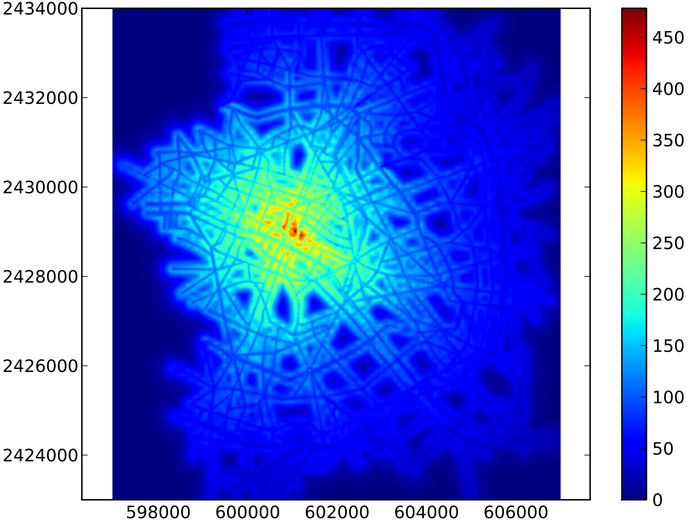
Simulation error covariances

With respect to a traffic station

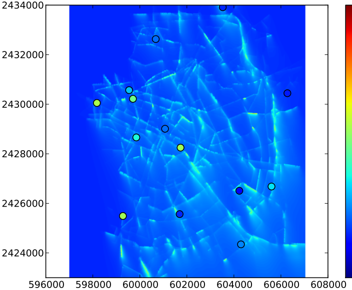


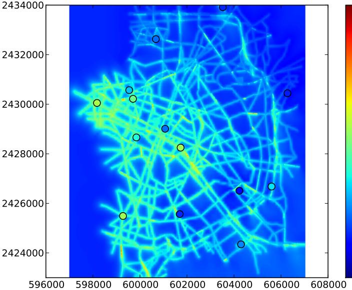
Simulation error covariances

With respect to a background station

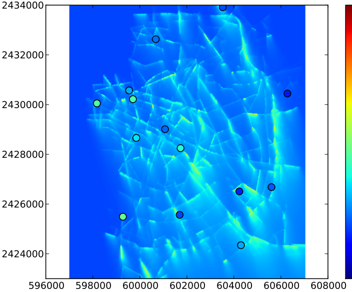


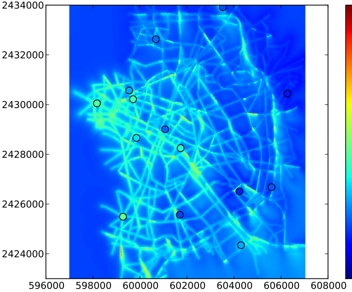
Before and after assimilation (preliminary result)

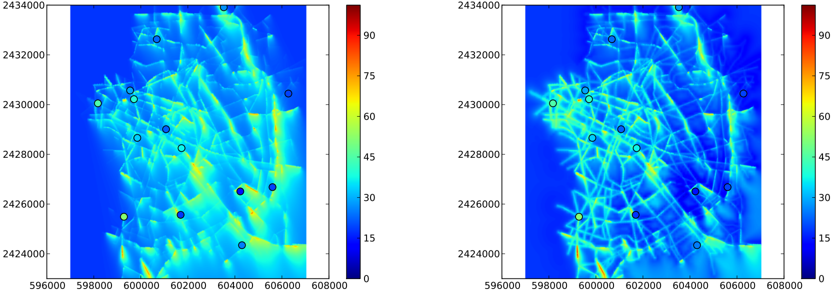




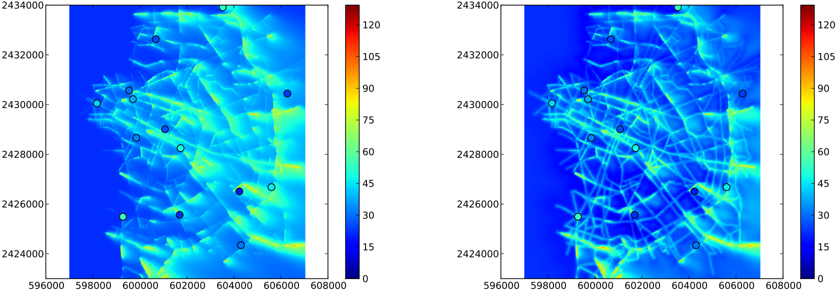
Before and after assimilation (preliminary result)



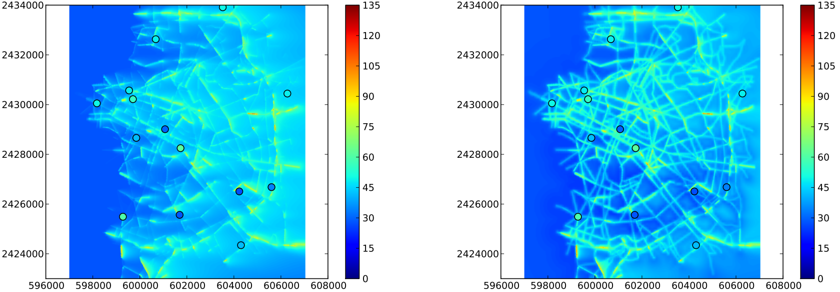



2434000
2432000
2430000
2428000
2426000
2424000
596000
598000
600000
602000
604000
606000
608000
0
15
30
45
60
75
90
2434000
2432000
2430000
2428000
2426000
2424000
596000
598000
600000
602000
604000
606000
608000
0
15
30
45
60
75
90

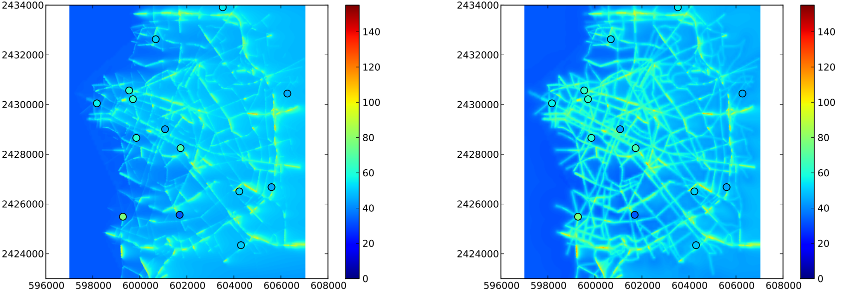
Before and after assimilation (preliminary result)



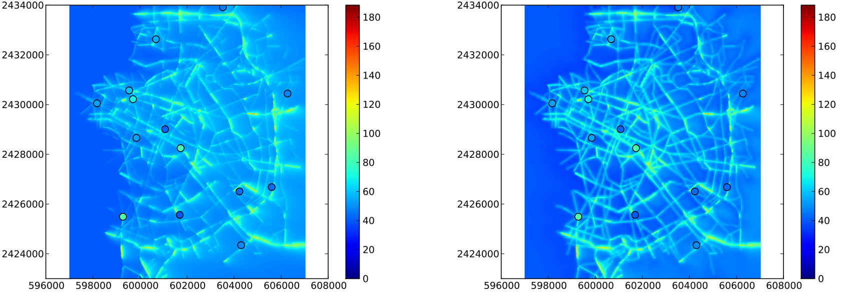
Before and after assimilation (preliminary result)



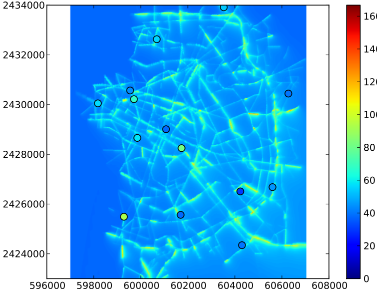
Before and after assimilation (preliminary result)

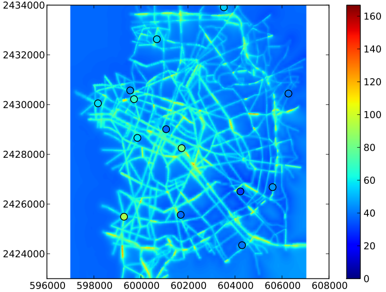


Before and after assimilation (preliminary result)

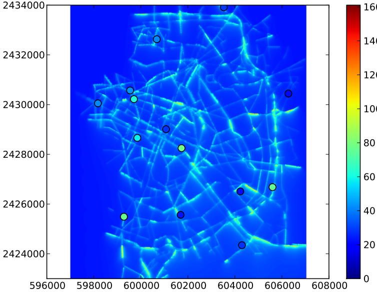


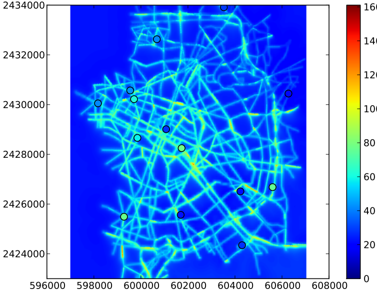
Before and after assimilation (preliminary result)



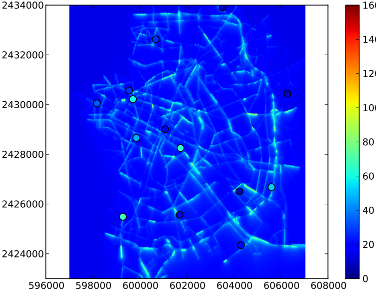


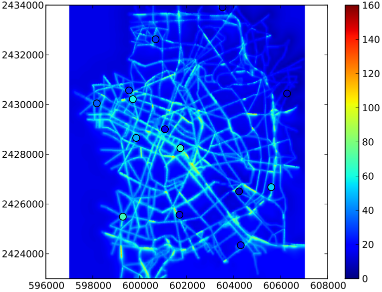
Before and after assimilation (preliminary result)



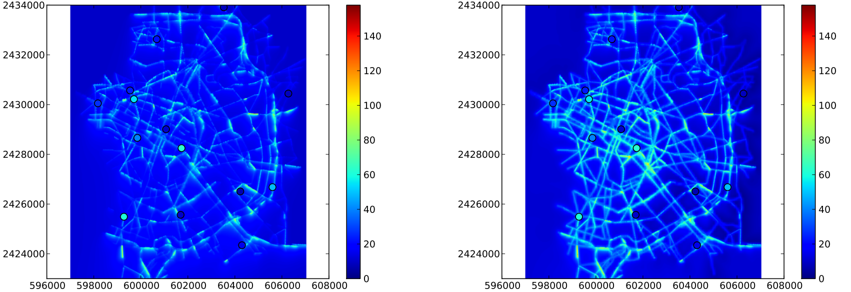


Before and after assimilation (preliminary result)

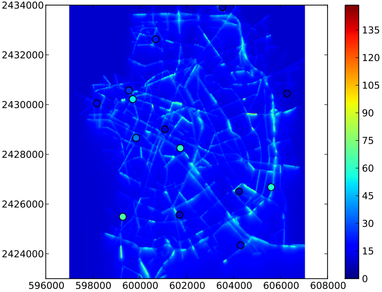


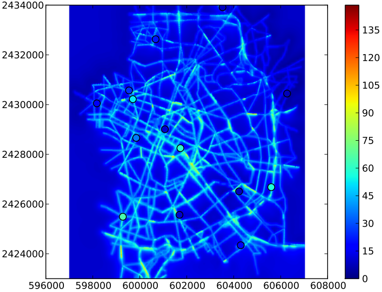


Before and after assimilation (preliminary result)

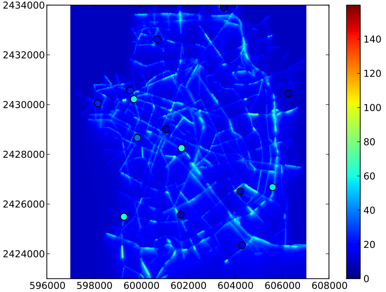


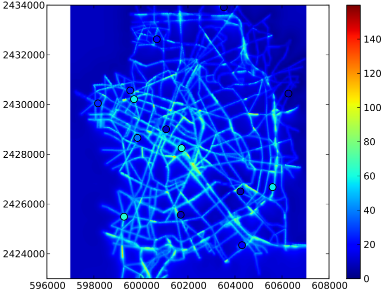
Before and after assimilation (preliminary result)



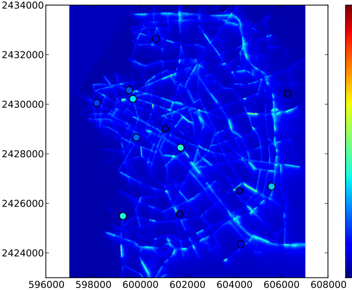


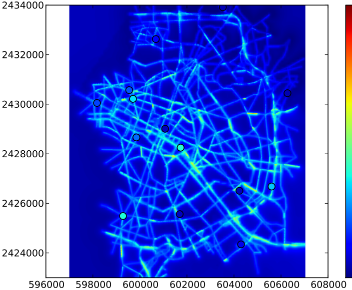
Before and after assimilation (preliminary result)



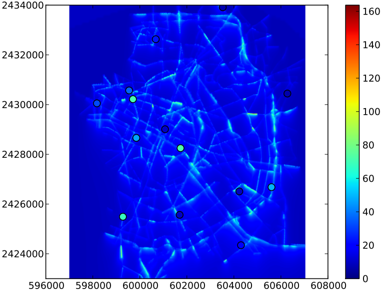


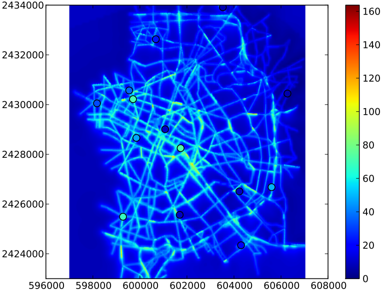
Before and after assimilation (preliminary result)



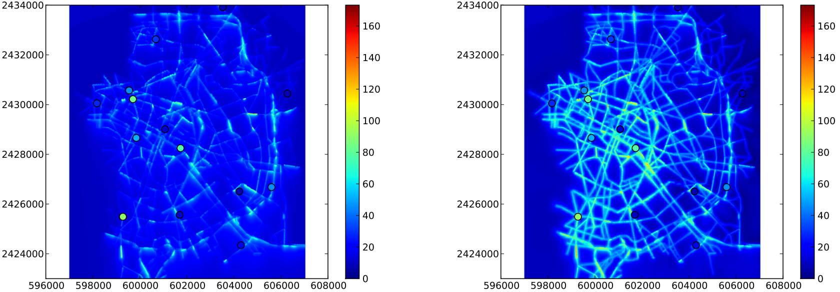


Before and after assimilation (preliminary result)

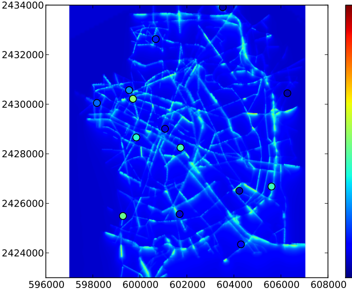


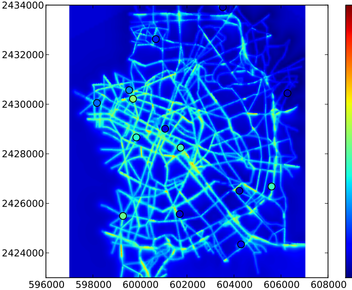


Before and after assimilation (preliminary result)

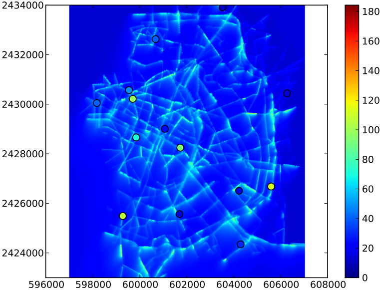


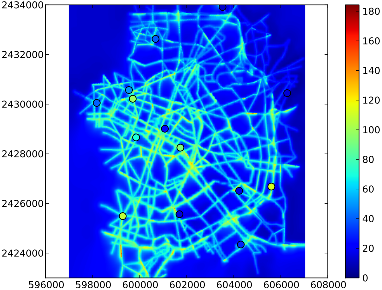
Before and after assimilation (preliminary result)



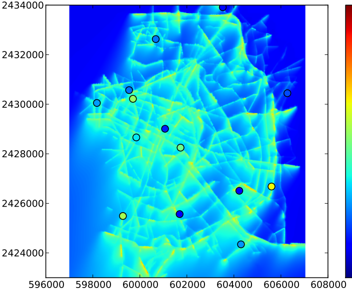


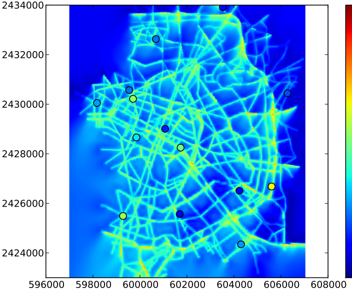
Before and after assimilation (preliminary result)



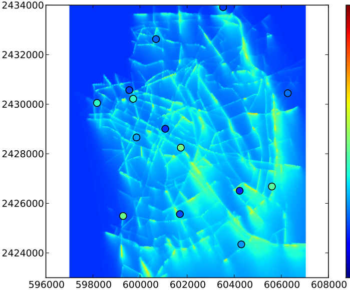


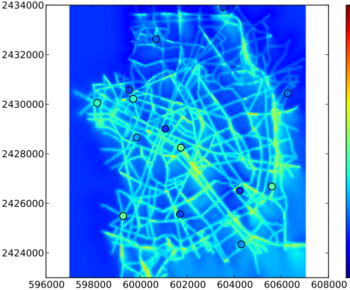
Before and after assimilation (preliminary result)



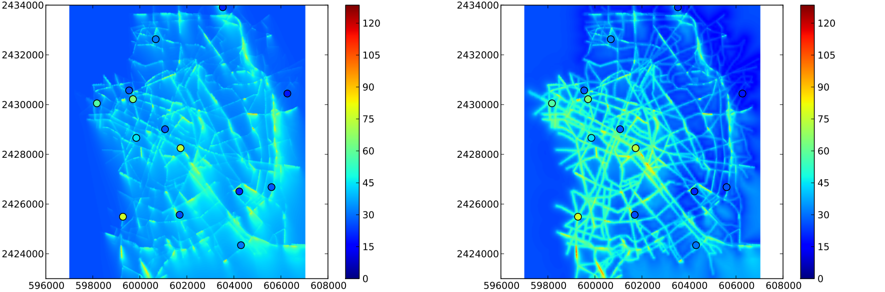


Before and after assimilation (preliminary result)

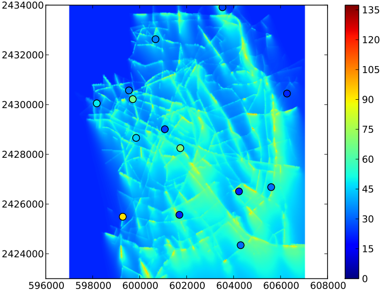


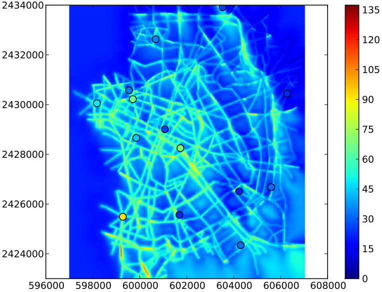


Before and after assimilation (preliminary result)

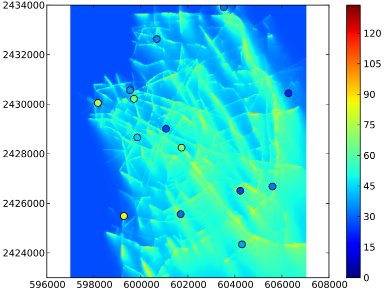


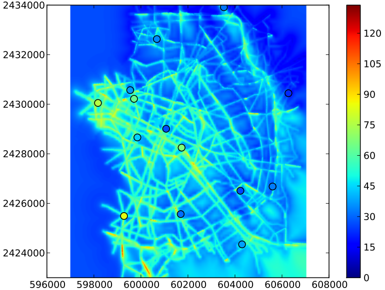
Before and after assimilation (preliminary result)



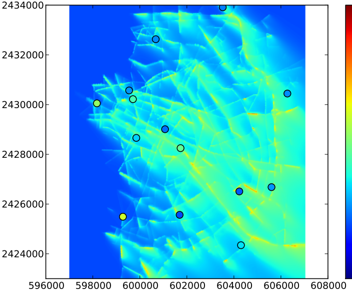


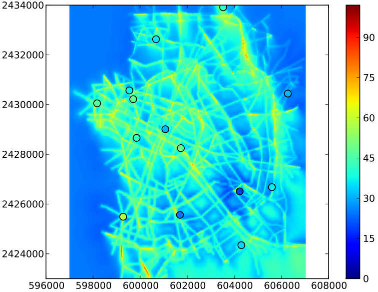
Before and after assimilation (preliminary result)



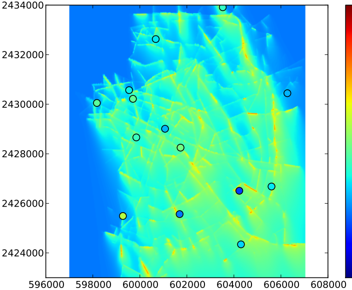


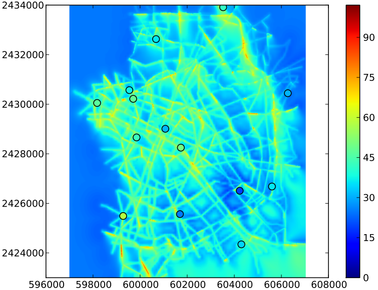
Before and after assimilation (preliminary result)



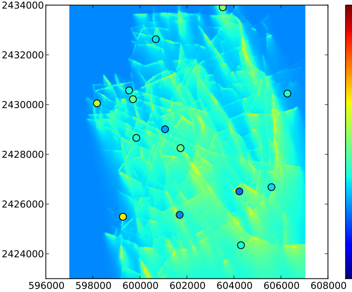


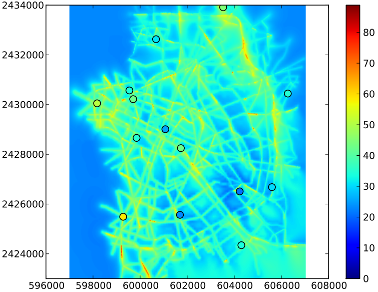
Before and after assimilation (preliminary result)



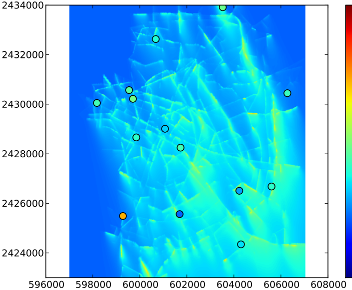


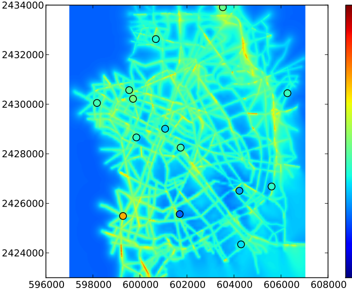
Before and after assimilation (preliminary result)



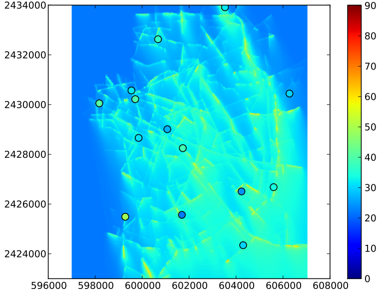


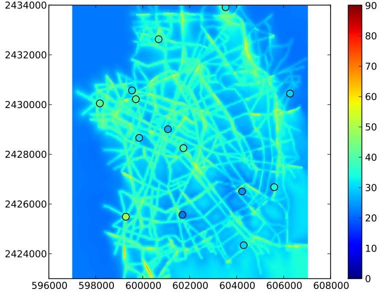
Before and after assimilation (preliminary result)



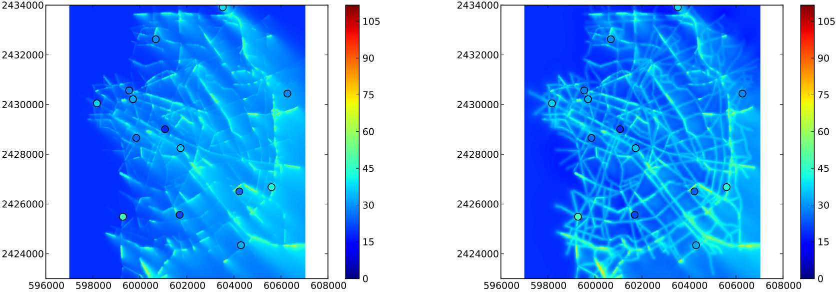


Before and after assimilation (preliminary result)

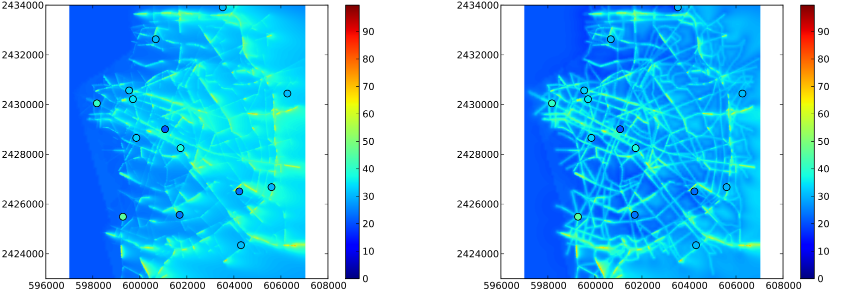




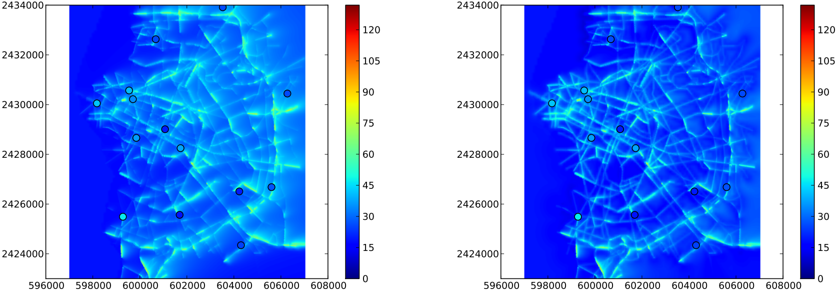
Before and after assimilation (preliminary result)



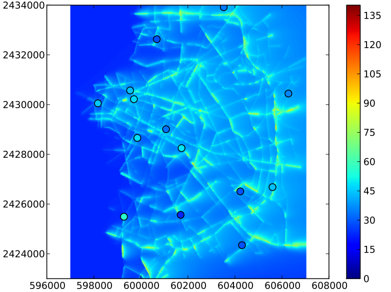
Before and after assimilation (preliminary result)

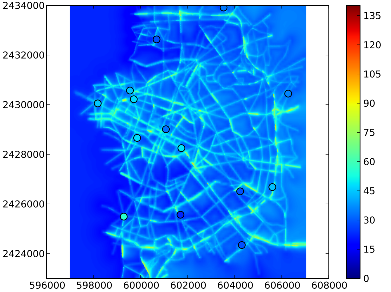


Before and after assimilation (preliminary result)

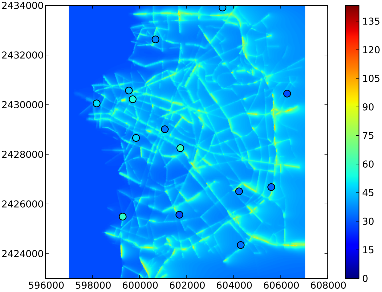


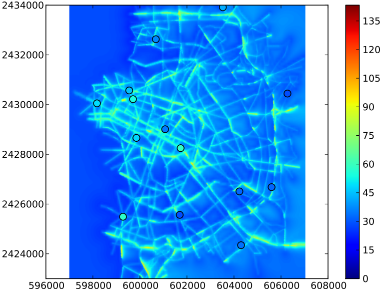
Before and after assimilation (preliminary result)



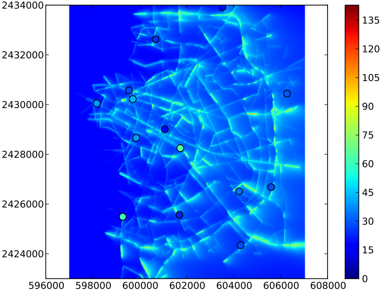


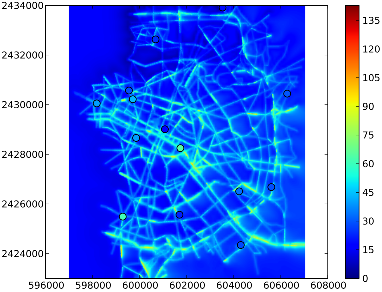
Before and after assimilation (preliminary result)



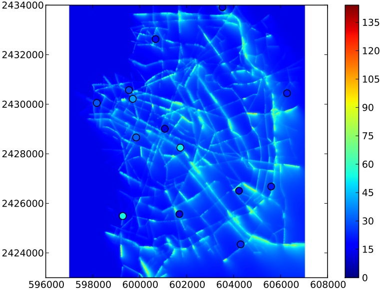


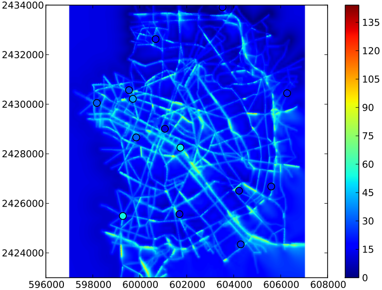
Before and after assimilation (preliminary result)



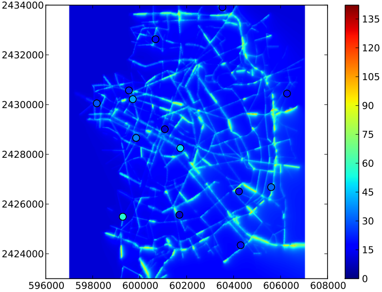


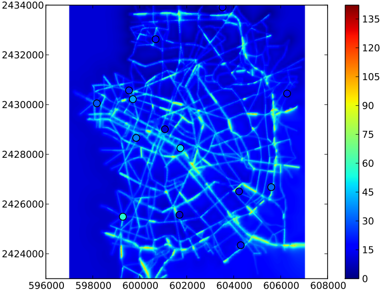
Before and after assimilation (preliminary result)



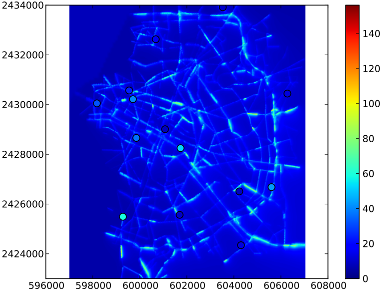


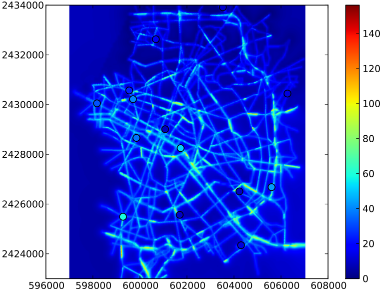
Before and after assimilation (preliminary result)



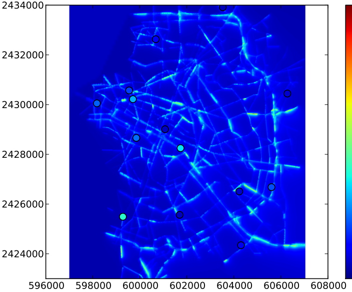


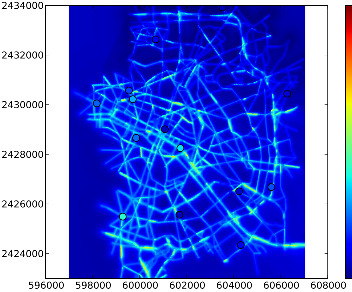
Before and after assimilation (preliminary result)



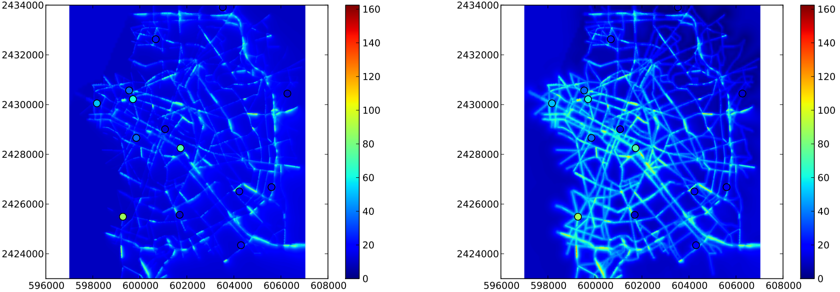


Before and after assimilation (preliminary result)

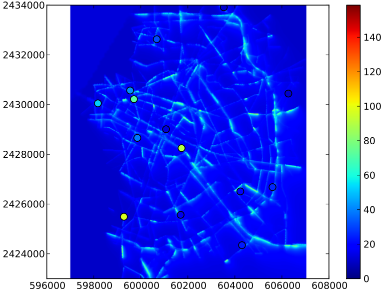


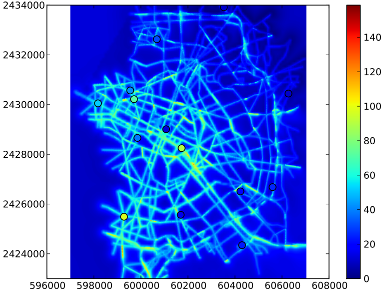


Before and after assimilation (preliminary result)

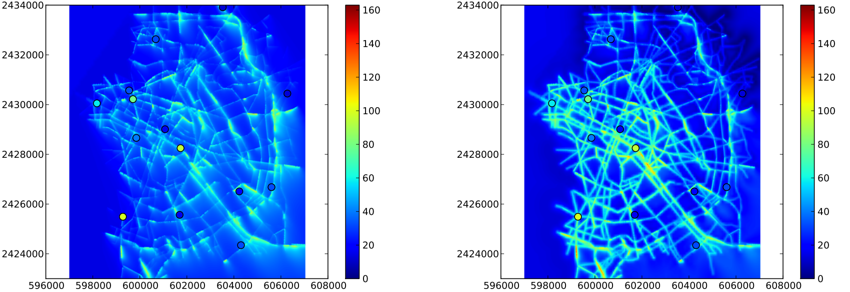


Before and after assimilation (preliminary result)

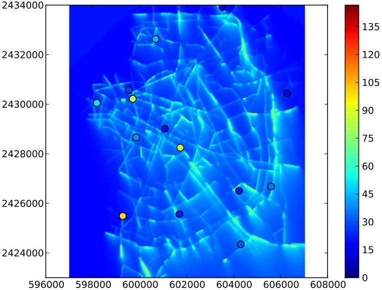


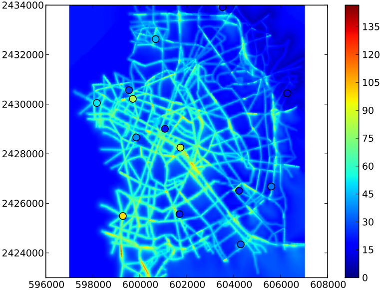


Before and after assimilation (preliminary result)

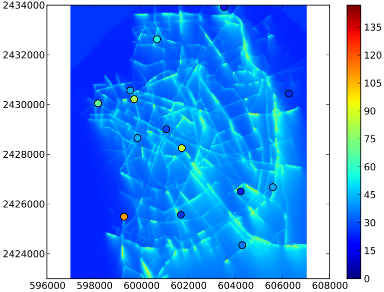


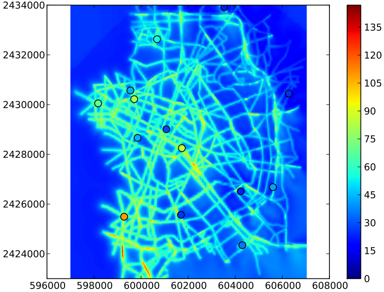
Before and after assimilation (preliminary result)



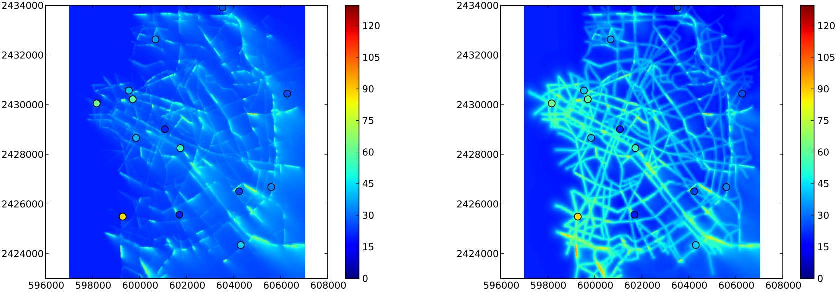


Before and after assimilation (preliminary result)

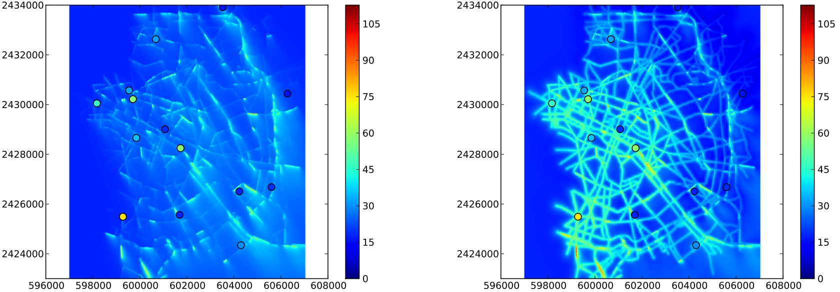




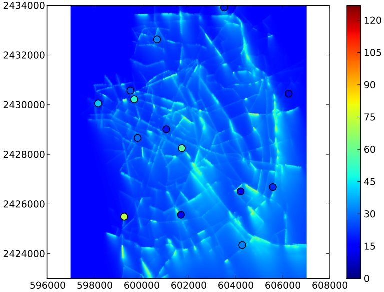
Before and after assimilation (preliminary result)

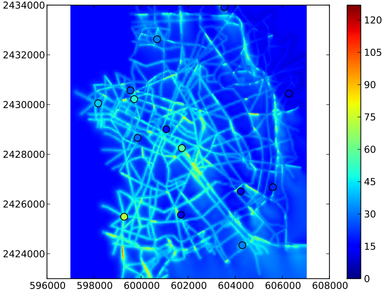


Before and after assimilation (preliminary result)

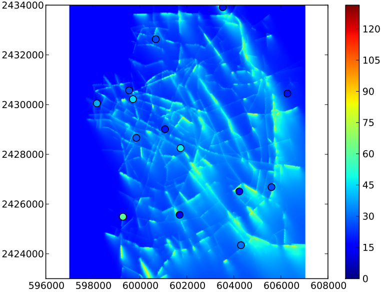


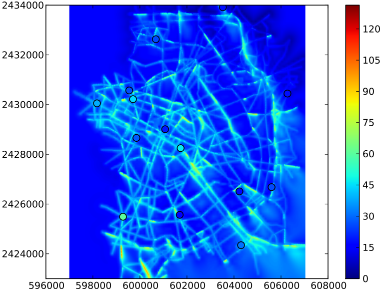
Before and after assimilation (preliminary result)





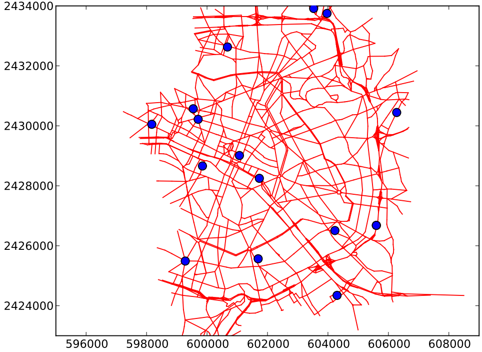
Before and after assimilation (preliminary result)



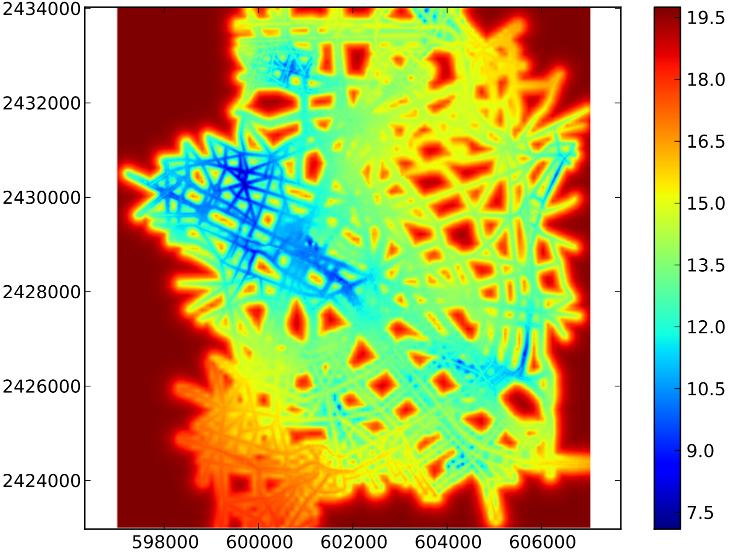


Leave-one-out cross-validation (preliminary result)

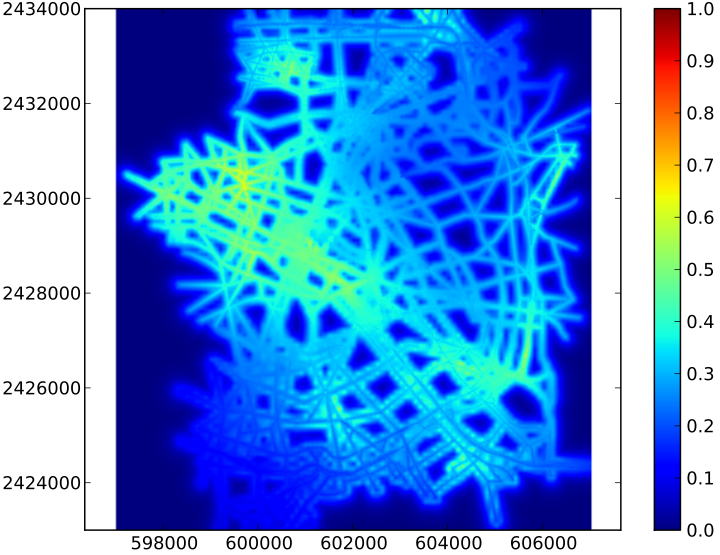
Station	Error change (%)
AUB	-8
BAGN	$+15$
BASC	-44
BONA	-45
CELE	-33
ELYS	-44
ETU6	-44
HAUS	-56
IVRY	-44
PA12	-50
PA13	-7
PA18	-42
PA4C	-49
PERA	-51



Standard deviation of the analysis (preliminary result)



Reduction of the standard deviation (preliminary result)



Data assimilation at urban scale

Running operationally since June 2011

- See `http://votreair.airparif.fr/`
- Real-time traffic $\rightarrow$ emission model $\rightarrow$ ADMS Urban $\rightarrow$ real-time observations $\rightarrow$ data assimilation
- Still a prototype, but will be extended to Paris or Paris region by Airparif

Part of Numtech products

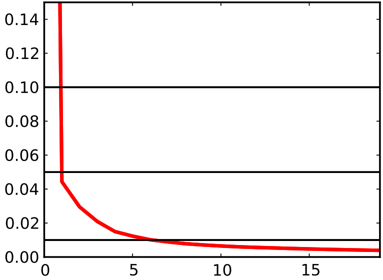
- For air quality agencies and cities
- Might need to assimilate new type of observations

Need for a better uncertainty estimation.

Reduction strategy

Dimension reduction

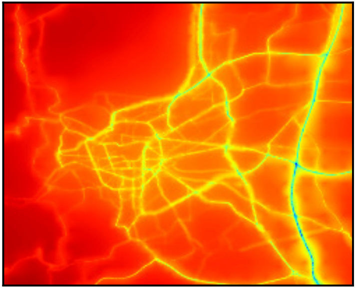
- Projection of inputs $\boldsymbol{p}$ (when necessary) and outputs $\boldsymbol{x}$ into a reduced subspace
- E.g., for outputs, application of principal component analysis on outputs of a one-year simulation
  - 99% of total variance explained with just 8 modes
  - $\boldsymbol{x} \simeq \sum_{j=1}^{8} \alpha_j \boldsymbol{\Psi}_j = \boldsymbol{\Psi}\boldsymbol{\alpha}$

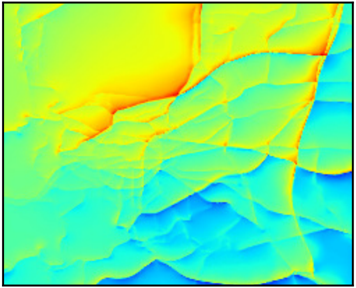


Relative part of unexplained variance against number of modes

Projection modes

Example for Clermont-Ferrand





Reduction strategy

Reduced model

- Complete model: $\boldsymbol{x} = \mathcal{M}(\boldsymbol{p})$
- Reduced model: $\boldsymbol{\alpha} = \boldsymbol{\Psi}^\top \mathcal{M}(\boldsymbol{p})$; note that $\boldsymbol{x} \simeq \boldsymbol{\Psi}\boldsymbol{\Psi}^\top \mathcal{M}(\boldsymbol{p})$
- $\boldsymbol{p} \in \mathbb{R}^{10}$ and $\boldsymbol{\alpha} \in \mathbb{R}^8$ are low-dimensional vectors

Emulation

- Components of $\boldsymbol{\alpha}$ show a smooth dependence on the components of $\boldsymbol{p}$
- Emulation consists in finding a surrogate function m for $\boldsymbol{\Psi}^\top \mathcal{M}$
- We can always reconstruct the full output: $\boldsymbol{x} \simeq \boldsymbol{\Psi} m(p)$

Building the emulator

Training values

- Let us consider the jth component of $\boldsymbol{\Psi}^\top \mathcal{M}$, and its emulator m_j
- We draw M samples $\boldsymbol{p}^{(i)}$, possibly by latin hypercube sampling
- We apply the reduced model to constitute the learning set: $\boldsymbol{\Psi}_j^\top \mathcal{M}(\boldsymbol{p}^{(i)})$

Emulator formulation

- The emulator is made of two parts:

$$m_j(\boldsymbol{p}) = \underbrace{\sum_{k=1}^{10} \beta_{j,k} \boldsymbol{p}_k}_{\text{Regression}} + \underbrace{\sum_{i=1}^{M} w_{i,j}(\boldsymbol{p}, \boldsymbol{p}^{(1)}, \ldots, \boldsymbol{p}^{(M)}) \left(\boldsymbol{\Psi}_j^\top \mathcal{M}(\boldsymbol{p}^{(i)}) - \sum_{k=1}^{10} \beta_{j,k} \boldsymbol{p}_k^{(i)} \right)}_{\text{Interpolation of the residuals}}$$

- Different options for the interpolation of the residuals:
  - Kriging (particular case of Gaussian processes), which also provides an uncertainty estimation;
  - Interpolation in high dimension with radial basis functions
  - Even the closest neighbor(s)

Model reduction applied to urban simulations

Computational costs: dimension reduction, emulator training and prediction

- About 6 months of simulation to determine the reduced subspace spanned by the columns of $\boldsymbol{\Psi}$
- $M = 2000$ samples for the emulator training, i.e, less than 3 months of simulation
- Full ADMS Urban cost: $\sim 10\,\mathrm{min}$ on 12 cores for one date (i.e, one hour)
- Emulator prediction cost: $50\,\mathrm{ms}$

Summary and perspectives

Data assimilation

- Merging model outputs and observations
  1. Strongly improves the evaluation of air quality across the city
  2. Provides insights on the best locations for the monitoring stations
- Requires better uncertainty estimation

Model reduction

- Dimension reduction is very efficient on outputs
- Emulation is possible and is so fast that it dramatically changes the perspectives

A few perspectives

- Propagation of inputs PDFs through the emulator
- Inverse modeling: computing the a posteriori PDFs on the inputs