

# Fusion des données simulées et observées en milieu urbain

Vivien Mallet<sup>1</sup>

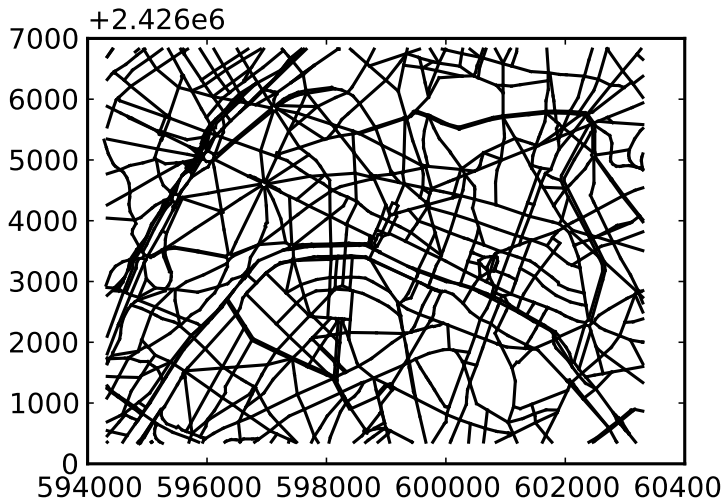
<sup>1</sup>INRIA

École RESCOM, Lyon, juin 2015

- 1 Introduction du problème et notations
- 2 Meilleur estimateur linéaire sans biais : BLUE
  - Opérateur d'observation
  - Hypothèses sur les erreurs
  - Calcul de BLUE
- 3 Détermination des matrices de covariance d'erreur et application
- 4 Erreur a posteriori
- 5 Vers l'application à la pollution sonore

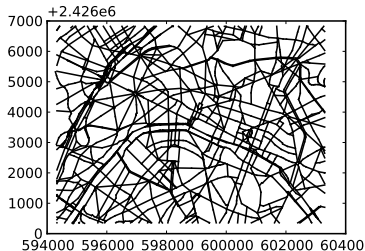
# Réseau routier pour la modélisation du trafic

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



# Entrées du modèle numérique de dispersion atmosphérique

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

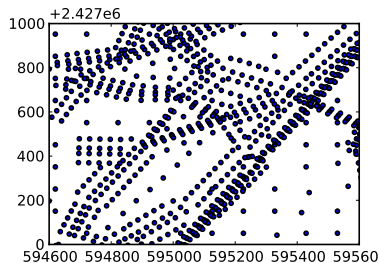
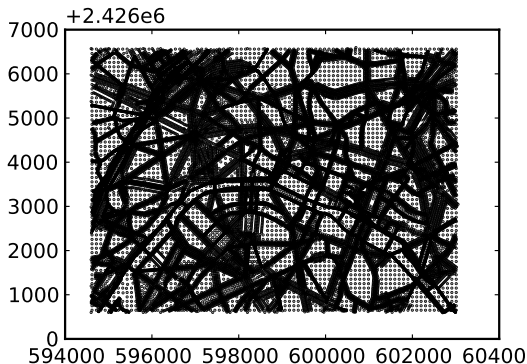


- Émissions du trafic

- Émissions industrielles et résidentielles
- Météorologie : vitesse et direction du vent, température, intensité de pluie, couverture nuageuse
- Pollution de fond :  $[\text{NO}_2]$ ,  $[\text{NO}_x]$  et  $[\text{O}_3]$
- Jour de l'année, type de jour (semaine, samedi, dimanche), heure du jour

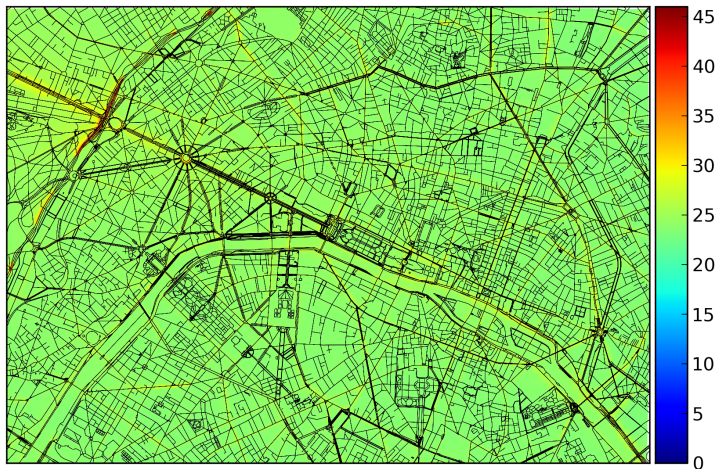
# Sorties du modèle numérique de dispersion atmosphérique

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



# Résultats de la simulation à une heure donnée

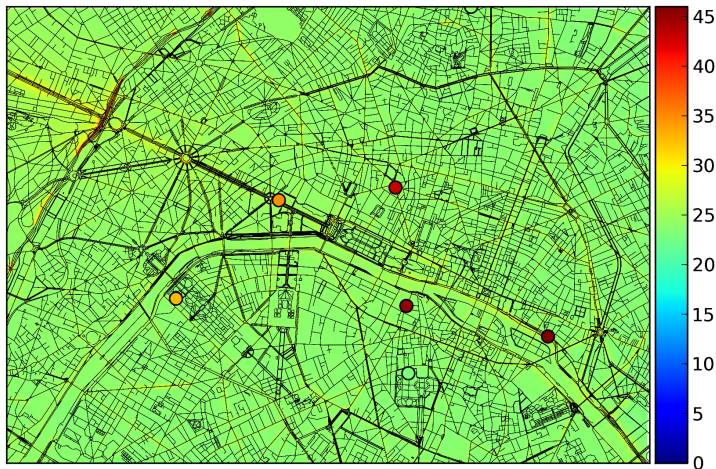
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



Concentrations de NO<sub>2</sub> ( $\mu\text{g m}^{-3}$ ) simulées pour le 26 juin 2011 à 7h00

# Simulation et observations à une heure donnée

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



Concentrations de NO<sub>2</sub> ( $\mu\text{g m}^{-3}$ ) simulées pour le 26 juin 2011 à 7h00

- Comment fusionner la carte simulée et les données observées ?
- La carte est la représentation du vecteur  $\mathbf{x}^b \in \mathbb{R}^N$ , appelé *ébauche de état* du système, ou prévision, à corriger
- Les *observations* ponctuelles sont rassemblées dans le vecteur  $\mathbf{y} \in \mathbb{R}^m$ , généralement de petite taille ( $m \ll N$ )



- 1 Introduction du problème et notations
- 2 Meilleur estimateur linéaire sans biais : BLUE
  - Opérateur d'observation
  - Hypothèses sur les erreurs
  - Calcul de BLUE
- 3 Détermination des matrices de covariance d'erreur et application
- 4 Erreur a posteriori
- 5 Vers l'application à la pollution sonore

# Opérateur d'observation

Comment comparer simulation et observation ?

## Définition de l'opérateur d'observation

- Permet de comparer  $\mathbf{x}^b$  et  $\mathbf{y}$
- Fait passer de l'espace de la simulation à l'espace des observations :  $\mathcal{H}(\mathbf{x}^b) \in \mathbb{R}^m$  se compare à  $\mathbf{y}$
- S'écrit comme la matrice  $\mathbf{H} \in \mathbb{R}^{m \times N}$  dans le cas linéaire

# Opérateur d'observation

Comment comparer simulation et observation ?

## Exemple à deux points, dont un observé

- $\mathbf{x}^b = \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix}$  ne contient que deux valeurs
- $\mathbf{y} = [y_1]$  ne contient qu'une valeur
- Exercices :
  - 1 Écrire  $\mathbf{H}$  si  $y_1$  correspond à  $x_1^b$
  - 2 Écrire  $\mathbf{H}$  si  $y_1$  correspond à  $x_2^b$

# Opérateur d'observation

Comment comparer simulation et observation ?

## Exemple à deux points, dont un observé

- $\mathbf{x}^b = \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix}$  ne contient que deux valeurs
- $\mathbf{y} = [y_1]$  ne contient qu'une valeur
- Exercices :
  - 1 Écrire  $\mathbf{H}$  si  $y_1$  correspond à  $x_1^b$
  - 2 Écrire  $\mathbf{H}$  si  $y_1$  correspond à  $x_2^b$

## Réponse

$$\mathbf{H} \in \mathbb{R}^{m \times N} = \mathbb{R}^{1 \times 2}$$

- 1  $\mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}$
- 2  $\mathbf{H} = \begin{bmatrix} 0 & 1 \end{bmatrix}$

# Opérateur d'observation

Comment comparer simulation et observation ?

## Exemple à trois points, avec une observation moyenne

- $\mathbf{x}^b = \begin{bmatrix} x_1^b \\ x_2^b \\ x_3^b \end{bmatrix}$

- $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

- Écrire  $\mathbf{H}$  si  $y_1$  observe  $x_1^b$  et  $y_2$  observe la moyenne de  $x_2^b$  et  $x_3^b$

# Opérateur d'observation

Comment comparer simulation et observation ?

## Exemple à trois points, avec une observation moyenne

- $\mathbf{x}^b = \begin{bmatrix} x_1^b \\ x_2^b \\ x_3^b \end{bmatrix}$
- $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$
- Écrire  $\mathbf{H}$  si  $y_1$  observe  $x_1^b$  et  $y_2$  observe la moyenne de  $x_2^b$  et  $x_3^b$

## Réponse

$$\mathbf{H} \in \mathbb{R}^{m \times N} = \mathbb{R}^{2 \times 3}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

## Définition des erreurs

- État réel :  $\mathbf{x}^t \in \mathbb{R}^N$ , inconnu et recherché
- Erreur d'ébauche :  $\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t$
- Erreur d'observation :  $\mathbf{e}^o = \mathbf{y} - \mathbf{H}\mathbf{x}^t$

## Erreur d'ébauche $\mathbf{e}^b$

- Vecteur aléatoire de moyenne nulle :  $\mathbb{E}(\mathbf{e}^b) = \mathbf{0}$
- Vecteur aléatoire de variance  
 $\mathbf{B} = \mathbb{E}((\mathbf{e}^b - \mathbb{E}(\mathbf{e}^b))(\mathbf{e}^b - \mathbb{E}(\mathbf{e}^b))^\top) = \mathbb{E}(\mathbf{e}^b \mathbf{e}^{b\top})$ 
  - Élément  $(i, i)$  :  $B_{ii} = \mathbb{E}(e_i^{b2}) = \text{var}(e_i^b)$
  - Élément  $(i, j)$  :  $B_{ij} = \mathbb{E}(e_i^b e_j^b) = \text{covar}(e_i^b, e_j^b)$

## Définition des erreurs

- État réel :  $\mathbf{x}^t \in \mathbb{R}^N$ , inconnu et recherché
- Erreur d'ébauche :  $\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t$
- Erreur d'observation :  $\mathbf{e}^o = \mathbf{y} - \mathbf{H}\mathbf{x}^b$

## Erreur d'observation $\mathbf{e}^o$

- Vecteur aléatoire de moyenne nulle :  $\mathbb{E}(\mathbf{e}^o) = \mathbf{0}$
- Vecteur aléatoire de variance  
 $\mathbf{R} = \mathbb{E}((\mathbf{e}^o - \mathbb{E}(\mathbf{e}^o))(\mathbf{e}^o - \mathbb{E}(\mathbf{e}^o))^\top) = \mathbb{E}(\mathbf{e}^o \mathbf{e}^{o\top})$ 
  - Élément  $(i, i)$  :  $R_{ii} = \mathbb{E}(e_i^{o2}) = \text{var}(e_i^o)$
- Décorrélée de l'erreur d'ébauche :  $\mathbb{E}(\mathbf{e}^o \mathbf{e}^{b\top}) = \mathbf{0}$
- $\mathbf{e}^o = \mathbf{y} - \mathbf{y}^t + \mathbf{y}^t - \mathbf{H}\mathbf{x}^t$ 
  - Erreur instrumentale  $\mathbf{y} - \mathbf{y}^t$
  - Erreur de représentativité  $\mathbf{y}^t - \mathbf{H}\mathbf{x}^t$



# Meilleur estimateur linéaire sans biais, $x^a$ : BLUE

Best Linear Unbiased Estimator

BLUE signifie « Best Linear Unbiased Estimator » et constitue l'*analyse* qui est le résultat de la fusion entre l'ébauche/prévision et l'observation.

## Définition

- **Linéaire** :  $x^a = \mathbf{K}y + \mathbf{L}x^b$

L'estimateur est une combinaison linéaire de l'ébauche et des observations

- **Sans biais** :  $E(e^a) = \mathbf{0}$ , avec  $e^a = x^a - x^t$

On préserve l'absence de biais sur l'ébauche

- **Meilleur** : trace  $\text{Tr}(\mathbf{A})$  minimale, avec  $\mathbf{A} = E(e^a e^{a\top})$

L'estimateur minimise la somme des variances des composantes de l'erreur  $e^a$ , soit  $\sum_{i=1}^N E(e_i^{a2})$

À partir de la forme linéaire,

$$\mathbf{x}^a = \mathbf{K}\mathbf{y} + \mathbf{L}\mathbf{x}^b, \quad (1)$$

on introduit l'erreur

$$\mathbf{x}^a - \mathbf{x}^t = \mathbf{K}\mathbf{y} - \mathbf{K}\mathbf{H}\mathbf{x}^t + \mathbf{K}\mathbf{H}\mathbf{x}^t + \mathbf{L}\mathbf{x}^b - \mathbf{L}\mathbf{x}^t + \mathbf{L}\mathbf{x}^t - \mathbf{x}^t, \quad (2)$$

d'où

$$\mathbf{e}^a = \mathbf{K}\mathbf{e}^o + \mathbf{L}\mathbf{e}^b + (\mathbf{K}\mathbf{H} + \mathbf{L} - \mathbf{I})\mathbf{x}^t. \quad (3)$$

Puisque  $E(\mathbf{e}^a)$  et  $E(\mathbf{e}^o)$  sont nulles, la moyenne de  $\mathbf{e}^a$  satisfait

$$E(\mathbf{e}^a) = E((\mathbf{K}\mathbf{H} + \mathbf{L} - \mathbf{I})\mathbf{x}^t), \quad (4)$$

ce qui est toujours nul si

$$\mathbf{L} = \mathbf{I} - \mathbf{K}\mathbf{H}. \quad (5)$$

On obtient donc

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b), \quad (6)$$

avec  $\mathbf{K}$  à déterminer.

L'erreur de BLUE est

$$e^a = \mathbf{K}e^o + (\mathbf{I} - \mathbf{KH})e^b, \quad (7)$$

donc sa variance est

$$\mathbf{E}(e^a e^{a\top}) = \mathbf{E} \left( (\mathbf{K}e^o + (\mathbf{I} - \mathbf{KH})e^b)(\mathbf{K}e^o + (\mathbf{I} - \mathbf{KH})e^b)^\top \right). \quad (8)$$

En développant l'expression, les termes croisés disparaissent puisque  $\mathbf{E}(e^a e^o^\top) = \mathbf{0}$  (erreurs décorrélées) et on obtient

$$\mathbf{E}(e^a e^{a\top}) = \mathbf{K} \mathbf{E}(e^o e^{o\top}) \mathbf{K}^\top + (\mathbf{I} - \mathbf{KH}) \mathbf{E}(e^b e^{b\top}) (\mathbf{I} - \mathbf{KH})^\top, \quad (9)$$

donc

$$\mathbf{E}(e^a e^{a\top}) = \mathbf{K} \mathbf{R} \mathbf{K}^\top + (\mathbf{I} - \mathbf{KH}) \mathbf{B} (\mathbf{I} - \mathbf{KH})^\top. \quad (10)$$

Autour de la valeur optimale pour  $\mathbf{K}$ , une perturbation  $\delta\mathbf{K}$  augmente la trace de  $\mathbf{E}(e^a e^{a^\top})$ . Considérons

$$\text{Tr} \left( (\mathbf{K} + \delta\mathbf{K})\mathbf{R}(\mathbf{K} + \delta\mathbf{K})^\top + (\mathbf{I} - \mathbf{K}\mathbf{H} - \delta\mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H} - \delta\mathbf{K}\mathbf{H})^\top \right). \quad (11)$$

Après développement and après avoir retiré les termes constants (c'est-à-dire sans  $\delta\mathbf{K}$ ) et de second ordre (c'est-à-dire avec  $\delta\mathbf{K}$  et  $\delta\mathbf{K}^\top$ ), on obtient

$$\begin{aligned} \text{Tr} \left( \mathbf{K}\mathbf{R}\delta\mathbf{K}^\top + \delta\mathbf{K}\mathbf{R}\mathbf{K}^\top + \mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^\top\delta\mathbf{K}^\top \right. \\ \left. + \delta\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^\top\mathbf{K}^\top - \mathbf{B}\mathbf{H}^\top\delta\mathbf{K}^\top - \delta\mathbf{K}\mathbf{H}\mathbf{B} \right). \end{aligned} \quad (12)$$

La trace est linéaire et invariante par transposition, et sachant que les matrices de covariance  $\mathbf{B}$  et  $\mathbf{R}$  sont symétriques, on obtient l'expression

$$2 \text{Tr} \left( \mathbf{K} \left[ \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^\top \right] \delta\mathbf{K}^\top - \mathbf{B}\mathbf{H}^\top\delta\mathbf{K}^\top \right). \quad (13)$$

Cette expression est nulle lorsque  $\mathbf{K} = \mathbf{B}\mathbf{H}^\top(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^\top)^{-1}$ .

# Meilleur estimateur linéaire sans biais, BLUE

## Best Linear Unbiased Estimator

- Le modèle calcule le vecteur d'état  $\mathbf{x}^b$  dont l'erreur est supposée
  - de moyenne nulle,
  - de variance  $\mathbf{B}$ .
- On remonte le vecteur des observations  $\mathbf{y}$  dont l'erreur est supposée
  - de moyenne nulle,
  - de variance  $\mathbf{R}$ ,
  - décorrélée de l'erreur sur  $\mathbf{x}^b$ ,
  - prenant en compte l'erreur instrumentale et l'erreur de représentativité ( $\mathbf{H}\mathbf{x}^t$  n'est pas la réalité aux points d'observation puisque  $\mathbf{x}^t$  est une représentation incomplète de la réalité).
- L'opérateur d'observation  $\mathbf{H}$  est introduit de façon à ce que  $\mathbf{y}$  soit comparable à  $\mathbf{H}\mathbf{x}^b$ .
- BLUE, aussi appelé analyse, est le « meilleur » estimateur linéaire sans biais :

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

# Exemple simple

$$\mathbf{x}^b = \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} R_{11} \end{bmatrix}$$

$$\mathbf{x}^a = \mathbf{x}^b + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left( \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} R_{11} \end{bmatrix} \right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^b)$$

$$\begin{bmatrix} x_1^a \\ x_2^a \end{bmatrix} = \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix} + \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} \left( \begin{bmatrix} B_{11} \end{bmatrix} + \begin{bmatrix} R_{11} \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} y_1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix} \right)$$

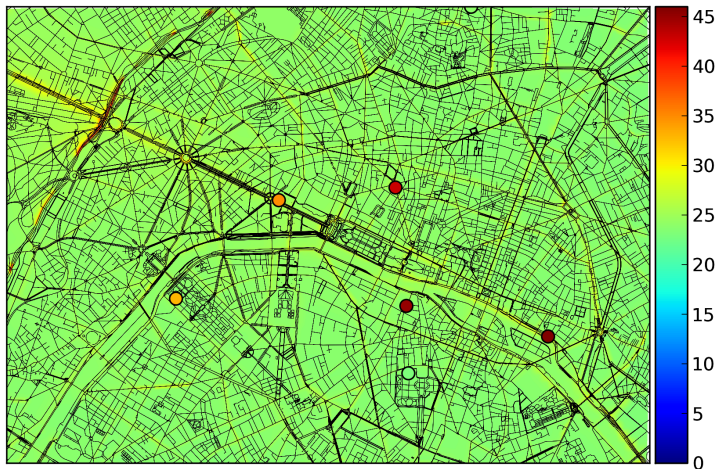
$$\begin{bmatrix} x_1^a \\ x_2^a \end{bmatrix} = \begin{bmatrix} x_1^b \\ x_2^b \end{bmatrix} + \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} \left[ \frac{1}{B_{11} + R_{11}} \right] \begin{bmatrix} y_1 - x_1^b \end{bmatrix}$$

$$\begin{cases} x_1^a = x_1^b + \frac{B_{11}}{B_{11} + R_{11}} (y_1 - x_1^b) \\ x_2^a = x_2^b + \frac{B_{21}}{B_{11} + R_{11}} (y_1 - x_1^b) \end{cases}$$

- 1 Introduction du problème et notations
- 2 Meilleur estimateur linéaire sans biais : BLUE
  - Opérateur d'observation
  - Hypothèses sur les erreurs
  - Calcul de BLUE
- 3 Détermination des matrices de covariance d'erreur et application
- 4 Erreur a posteriori
- 5 Vers l'application à la pollution sonore

# Simulation et observations à une heure donnée

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



Concentrations de NO<sub>2</sub> ( $\mu\text{g m}^{-3}$ ) simulées pour le 26 juin 2011 à 7h00



# Matrices de covariance d'erreur

Les erreurs d'observation sont souvent supposées décorrélées, même si ce n'est pas toujours exact :

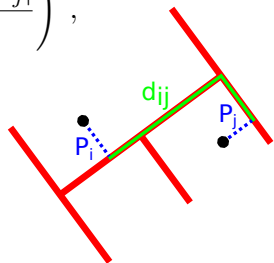
$$\mathbf{R} = r\mathbf{I} .$$

L'erreur de simulation proche du réseau routier est supposée dépendre fortement des émissions du trafic. De plus, les erreurs entre deux points sont supposées d'autant plus corrélées que les deux points sont à une distance similaire du réseau routier (pollution de proximité versus pollution de fond). Finalement, on choisit

$$B_{ij} = b \exp\left(-\frac{d_{ij}}{L_d}\right) \exp\left(-\frac{|P_i - P_j|}{L_p}\right) ,$$

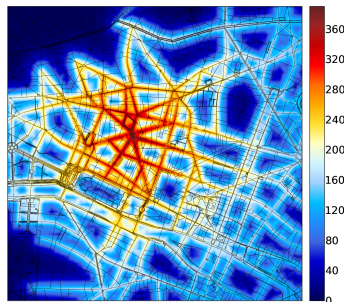
où  $d_{ij}$  est la longueur du plus court chemin entre les projections des points sur le réseau,

$P_i$  et  $P_j$  sont les distances au réseau routier, et  $L_d$  et  $L_p$  sont des distances de décorrélation le long du réseau routier et transversalement au réseau routier.

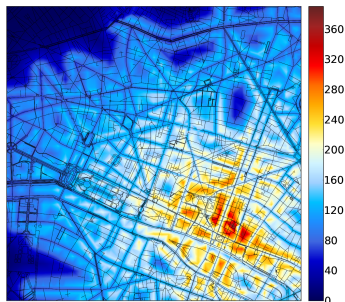


# Illustration de la covariance de l'erreur d'ébauche

Cartes de la covariance avec l'erreur à une station proche d'une route (gauche), et avec une erreur à une station de pollution de fond (droite).



Station OPERA (trafic)

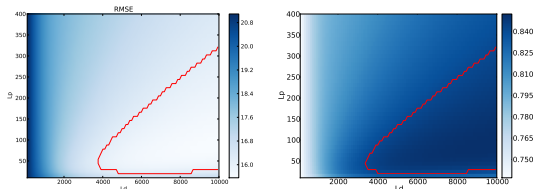


Station PA4C (fond)

Ce sont des lignes de la matrice  $\mathbf{B}$ , en  $\mu\text{g}^2 \text{m}^{-6}$ .

# Calage des paramètres dans les covariances

- Longueurs de décorrélation  $L_d$  et  $L_p$  : scores (RMSE, corrélation, ...) en validation croisée
  - Validation croisée en retirant une station d'observation à la fois
  - BLUE à la station retirée est comparé à l'observation
  - Valeurs typiques :  $L_d = 2000$  m à  $6000$  m et  $L_p = 100$  m

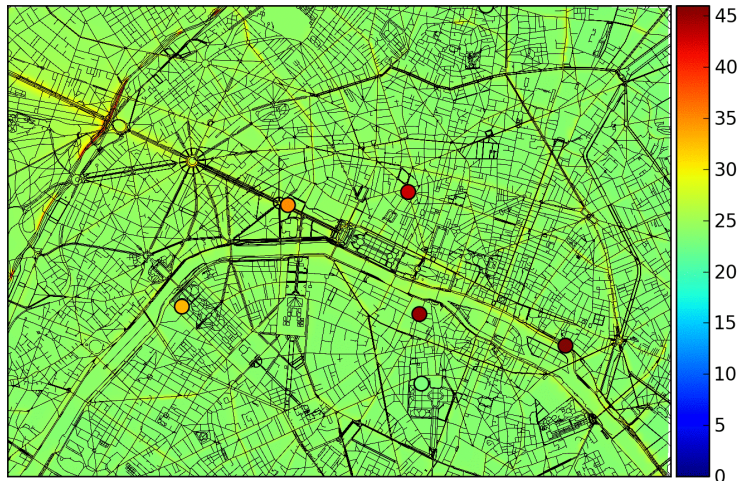


- Variance  $b$  : diagnostic du  $\chi^2$ , avec  $L_d$ ,  $L_p$ ,  $\alpha$  et  $v_o$  déjà fixés
  - Vérifier la cohérence entre  $\mathbf{y} - \mathbf{H}\mathbf{x}^b$  et sa variance  $\mathbf{S} = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^\top$
  - Le scalaire  $\chi^2 = (\mathbf{y} - \mathbf{H}\mathbf{x}^b)^\top \mathbf{S}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^b)$  doit être égal, en moyenne, au nombre d'observations (dimension de  $\mathbf{y}$ ).
  - Exemple :  $b = 290 \mu\text{g}^2 \text{m}^{-6}$ , if  $\alpha = 1$  and  $r = 100 \mu\text{g}^2 \text{m}^{-6}$

# Résultats de la simulation à une heure donnée

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-26 07:00:00

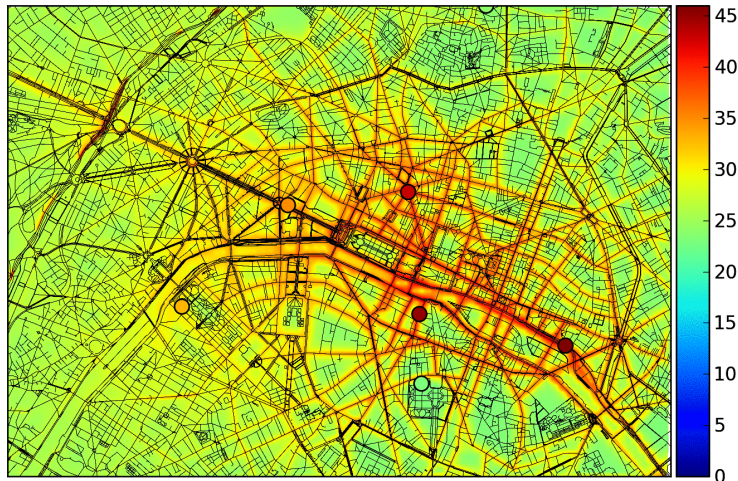


Concentrations de NO<sub>2</sub> ( $\mu\text{g m}^{-3}$ ) simulées pour le 26 juin 2011 à 7h00

# Résultats de l'assimilation de données

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-26 07:00:00

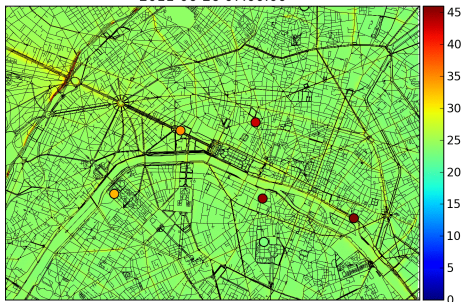


Concentrations de NO<sub>2</sub> ( $\mu\text{g m}^{-3}$ ) analysées pour le 26 juin 2011 à 7h00

# Résultats de l'assimilation de données

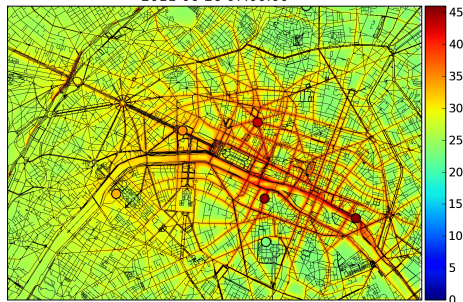
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-26 07:00:00



Carte de [NO<sub>2</sub>] (μg m<sup>-3</sup>), simulée

2011-06-26 07:00:00



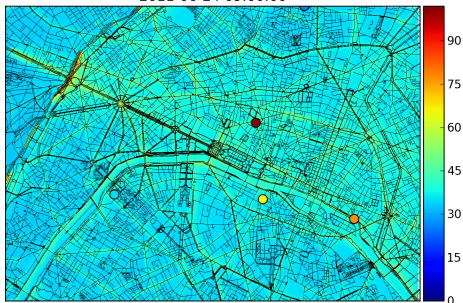
Carte de [NO<sub>2</sub>] (μg m<sup>-3</sup>), analysée

26 juin 2011 à 7h00

# Résultats de l'assimilation de données

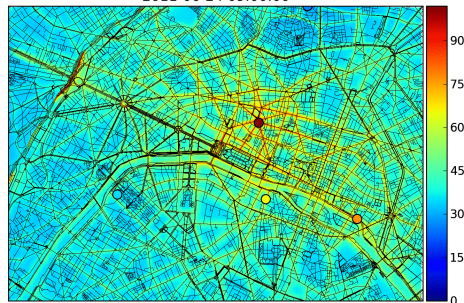
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-24 09:00:00



Carte de [NO<sub>2</sub>] (μg m<sup>-3</sup>), simulée

2011-06-24 09:00:00



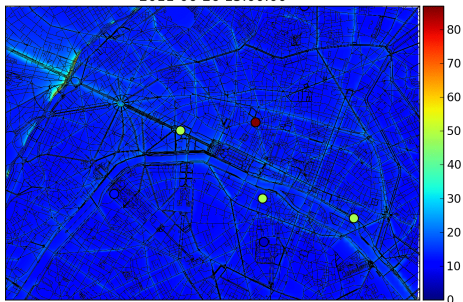
Carte de [NO<sub>2</sub>] (μg m<sup>-3</sup>), analysée

24 juin 2011 à 9h00

# Résultats de l'assimilation de données

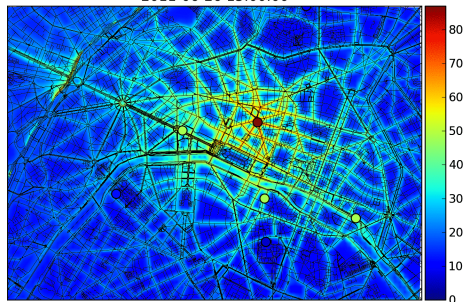
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-26 15:00:00



Carte de  $[\text{NO}_2]$  ( $\mu\text{g m}^{-3}$ ), simulée

2011-06-26 15:00:00



Carte de  $[\text{NO}_2]$  ( $\mu\text{g m}^{-3}$ ), analysée

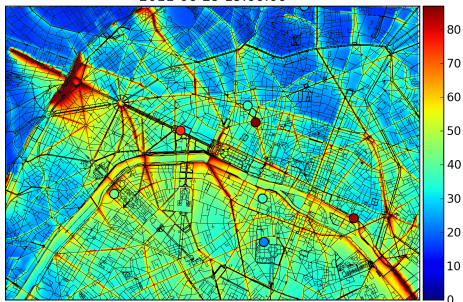
26 juin 2011 à 15h00



# Résultats de l'assimilation de données

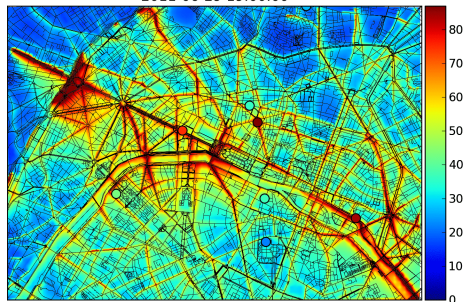
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-29 19:00:00



Carte de [NO<sub>2</sub>] (µg m<sup>-3</sup>), simulée

2011-06-29 19:00:00



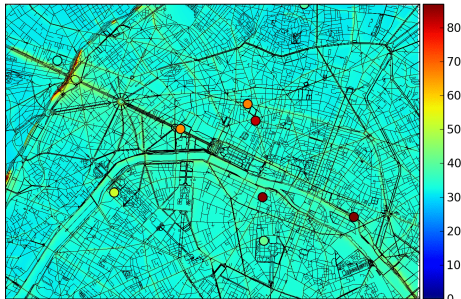
Carte de [NO<sub>2</sub>] (µg m<sup>-3</sup>), analysée

29 juin 2011 à 19h00

# Résultats de l'assimilation de données

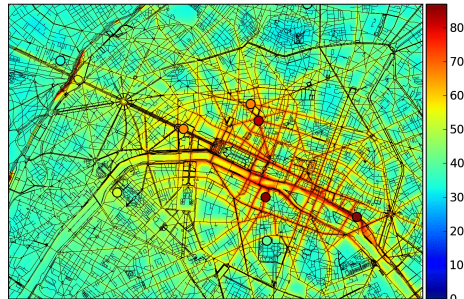
Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)

2011-06-30 08:00:00



Carte de  $[\text{NO}_2]$  ( $\mu\text{g m}^{-3}$ ), simulée

2011-06-30 08:00:00

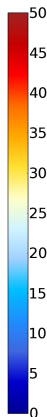
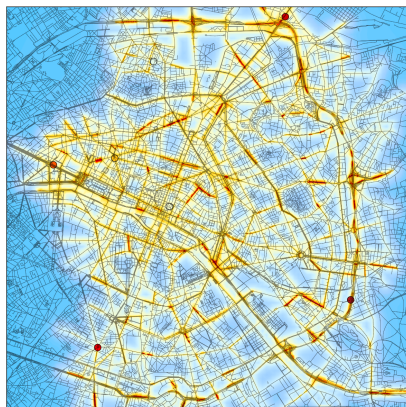
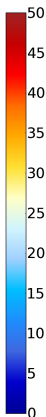
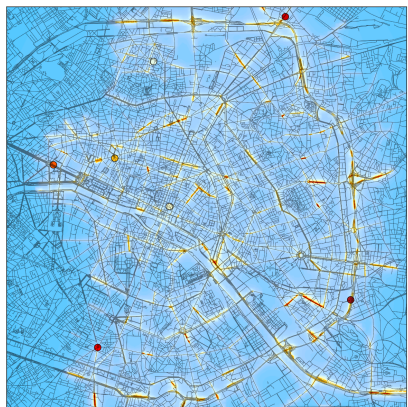


Carte de  $[\text{NO}_2]$  ( $\mu\text{g m}^{-3}$ ), analysée

30 juin 2011 à 8h00

En cours de déploiement sur 12 villes en France par la PME Numtech.

# Partie est de Paris, pour les particules fines



Carte de  $[PM_{10}]$  ( $\mu\text{g m}^{-3}$ ), simulée

Carte de  $[PM_{10}]$  ( $\mu\text{g m}^{-3}$ ), analysée

1er septembre 2012 à 18h00

Approche appliquée à  $NO_2$ ,  $PM_{2.5}$ ,  $PM_{10}$  et aux suies. Aussi à l'échelle de l'Île-de-France.

# Diminution de la RMSE pour NO<sub>2</sub>

Calculée dans le cadre de la validation croisée

Domaine	Trafic	Urbain	Fond	Ensemble
Paris, "Votre Air"	15 %–45 %	0 %–20 %	—	25 %
Paris, partie est	—	—	—	25 %
Clermont-Ferrand	10 %–30 %	30 %–50 %	5 %–20 %	25 %

# Application mobile : exposition à la pollution

Qualité de l'air en temps réel avec le prototype « Votre Air » (Airparif, Numtech, Inria)



- 1 Introduction du problème et notations
- 2 Meilleur estimateur linéaire sans biais : BLUE
  - Opérateur d'observation
  - Hypothèses sur les erreurs
  - Calcul de BLUE
- 3 Détermination des matrices de covariance d'erreur et application
- 4 Erreur a posteriori
- 5 Vers l'application à la pollution sonore

Analyse (BLUE, état a posteriori) :

$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) \\ \mathbf{K} &= \mathbf{B}\mathbf{H}^\top (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^\top)^{-1} \end{aligned}$$

Variance de l'erreur a posteriori  $e^a$  :

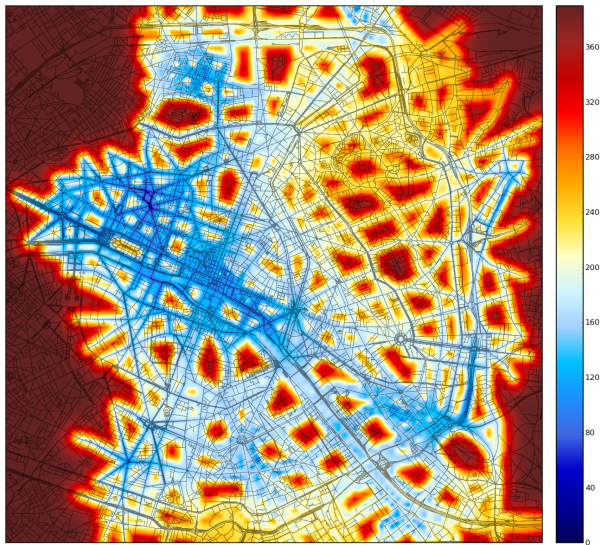
- Telle que calculée précédemment, cf. l'expression (10) :

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

- Après réécriture :

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$$

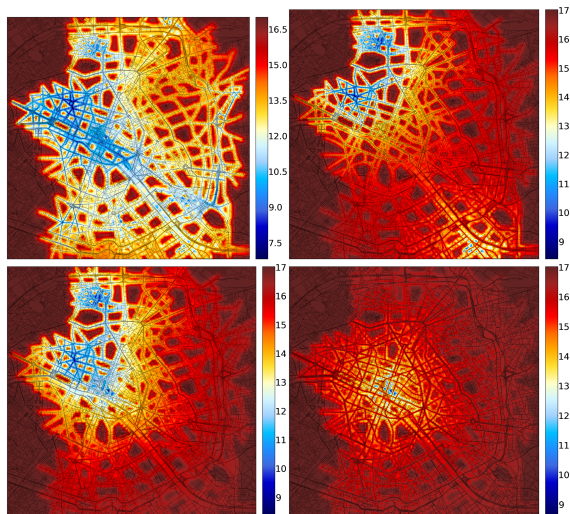
# Variance de l'erreur a posteriori



Variance ( $\mu\text{g}^2\text{m}^{-6}$ ) de l'erreur a posteriori sur  $[\text{NO}_2]$

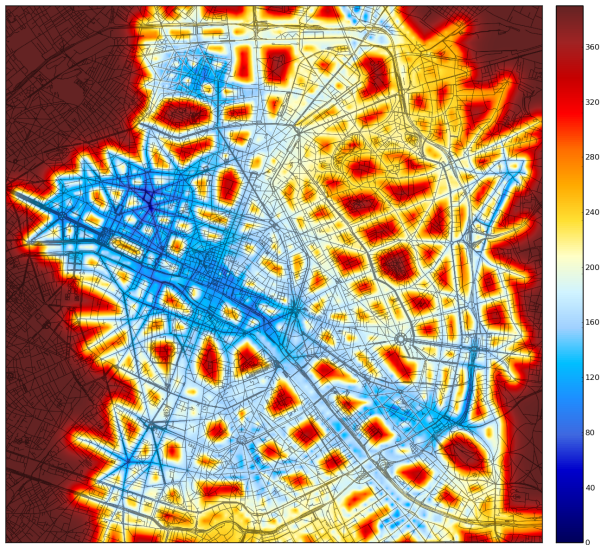


# Variance de l'erreur a posteriori pour quatre polluants



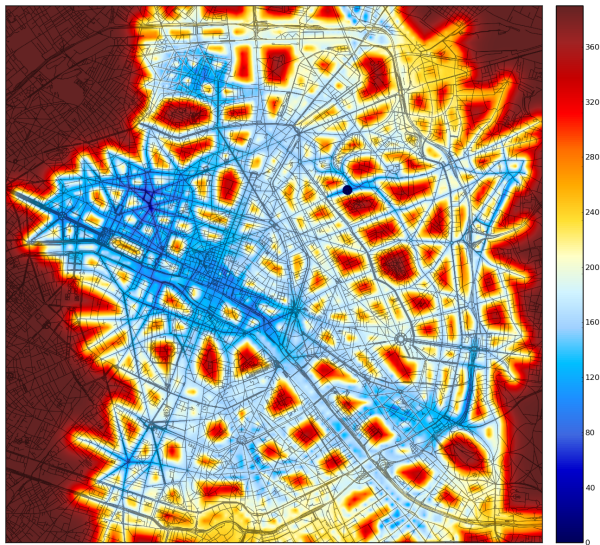
Écarts types ( $\mu\text{g m}^{-3}$ ) de l'erreur a posteriori pour [NO<sub>2</sub>] et [suies] (haut), et [PM<sub>10</sub>] et [PM<sub>2.5</sub>] (bas).

# Vers l'optimisation de réseau



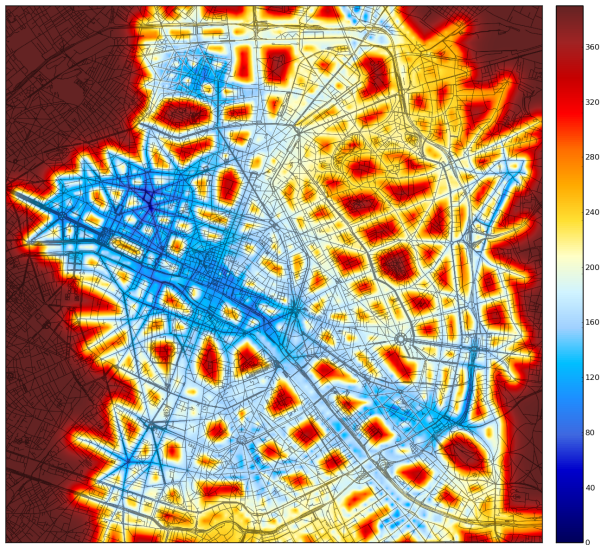
Variance ( $\mu\text{g}^2\text{m}^{-6}$ ) de l'erreur a posteriori sur  $[\text{NO}_2]$

# Vers l'optimisation de réseau



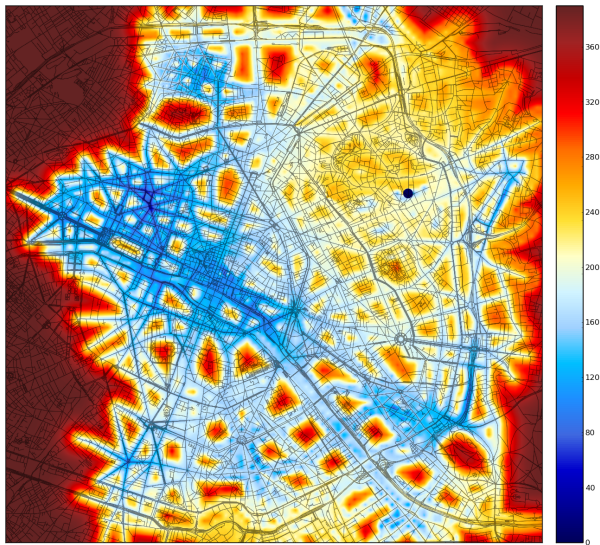
Variance ( $\mu\text{g}^2\text{m}^{-6}$ ) de l'erreur a posteriori sur  $[\text{NO}_2]$

# Vers l'optimisation de réseau



Variance ( $\mu\text{g}^2\text{m}^{-6}$ ) de l'erreur a posteriori sur  $[\text{NO}_2]$

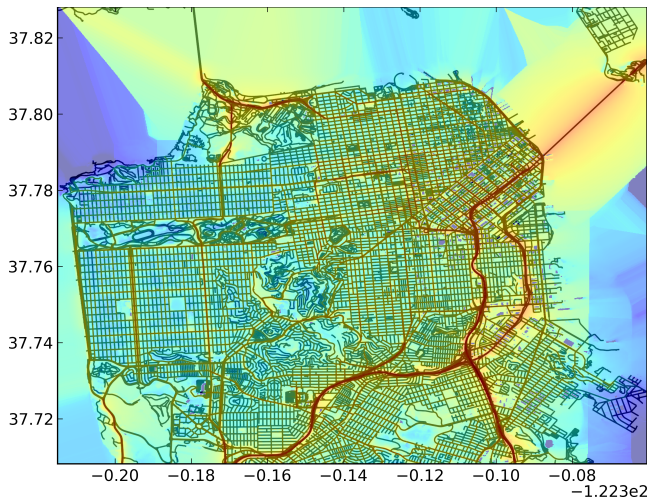
# Vers l'optimisation de réseau



Variance ( $\mu\text{g}^2\text{m}^{-6}$ ) de l'erreur a posteriori sur  $[\text{NO}_2]$

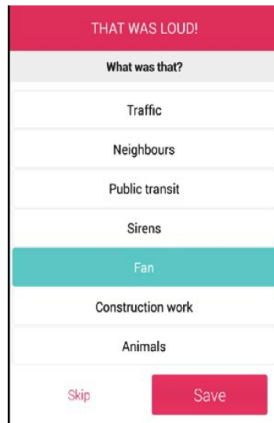
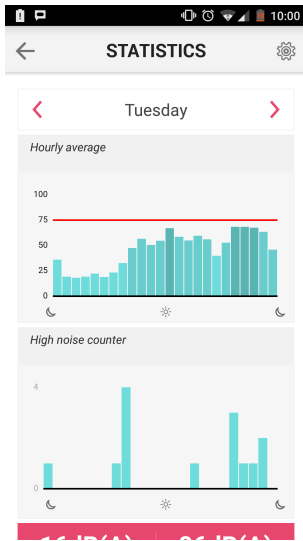
- 1 Introduction du problème et notations
- 2 Meilleur estimateur linéaire sans biais : BLUE
  - Opérateur d'observation
  - Hypothèses sur les erreurs
  - Calcul de BLUE
- 3 Détermination des matrices de covariance d'erreur et application
- 4 Erreur a posteriori
- 5 Vers l'application à la pollution sonore

# Vers l'application à la pollution sonore



Carte de bruit moyen sur la journée, calculée par la ville de San Francisco

# Mesure participative du bruit

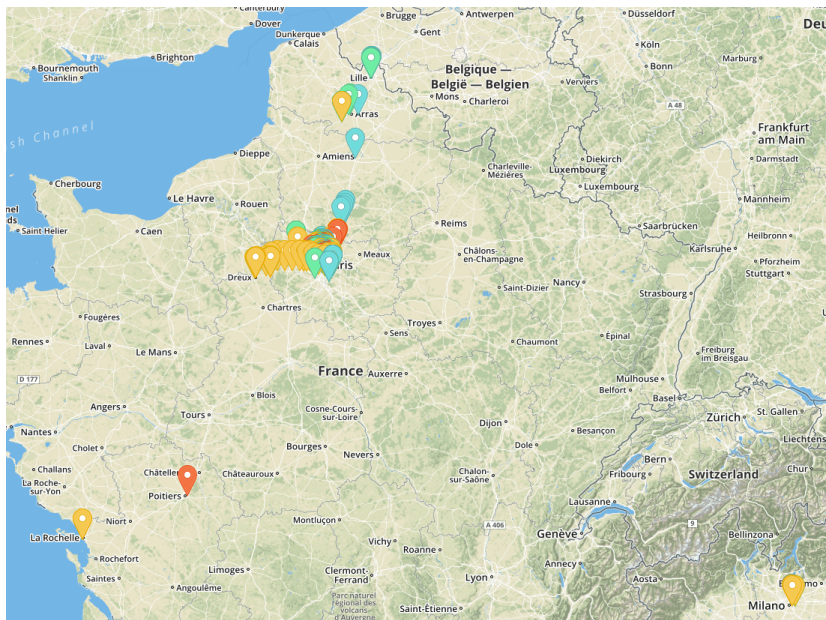




# Démonstration à Futur en Seine (13 and 14 juin)

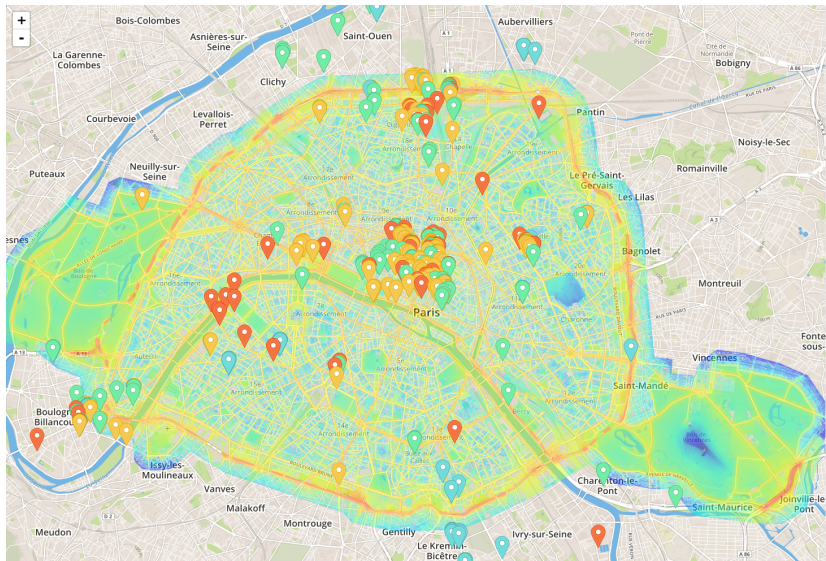


# Mesure de la pollution sonore





# Mesure de la pollution sonore





Introduction à l'assimilation de données : F. BOUTTIER et P. COURTIER (1999). *Data assimilation concepts and methods*. Meteorological training course lecture series. ECMWF



Notes de cours : "Introduction aux principes et méthodes de l'assimilation de données en géophysique", M. Bocquet — <http://cerea.enpc.fr/HomePages/bocquet/teaching.html>



Assimilation de données à l'échelle urbaine : A. TILLOY, V. MALLET, D. POULET et al. (2013). « BLUE-based NO<sub>2</sub> data assimilation at urban scale ». Dans : *Journal of Geophysical Research* 118.4, p. 2,031–2,040. DOI : 10.1002/jgrd.50233



Bibliothèque d'assimilation de données Verdandi : <http://verdandi.gforge.inria.fr/>



Équipes INRIA et ville intelligente : <http://citylab.inria.fr/>



Ville intelligente, implication citoyenne : <http://urbancivics.com/>