

# Réunion ADOQA

## *Quelques éléments sur la modélisation de la qualité de l'air et l'assimilation de données*

Marc Bocquet, Vivien Mallet et Bruno Sportisse

CEREA, 17 Juin 2005

# Quelques éléments de modélisation

## Équation de dispersion atmosphérique

- Évolution pour la concentration  $c_i$  :

$$\frac{\partial c_i}{\partial t} + \underbrace{\operatorname{div}(V c_i)}_{\text{advection}} = \underbrace{\operatorname{div}(K \nabla c_i)}_{\text{turbulent diffusion}} + \underbrace{\chi_i(c, t)}_{\text{gas-phase chemistry}} - \underbrace{\Lambda_i(x, t) c_i}_{\text{scavenging}} + \underbrace{S_i(x, t)}_{\text{volume sources}}$$

- Conditions aux limites au sol :

$$-K \nabla c_i \cdot n = \underbrace{E_i(x, t)}_{\text{surface emissions}} - \underbrace{v_i^{dep}(x, t) c_i}_{\text{dry deposition}}$$

- Conditions aux limites latérales si modèle à aire limitée.
- Échelle typique : modélisation continentale.

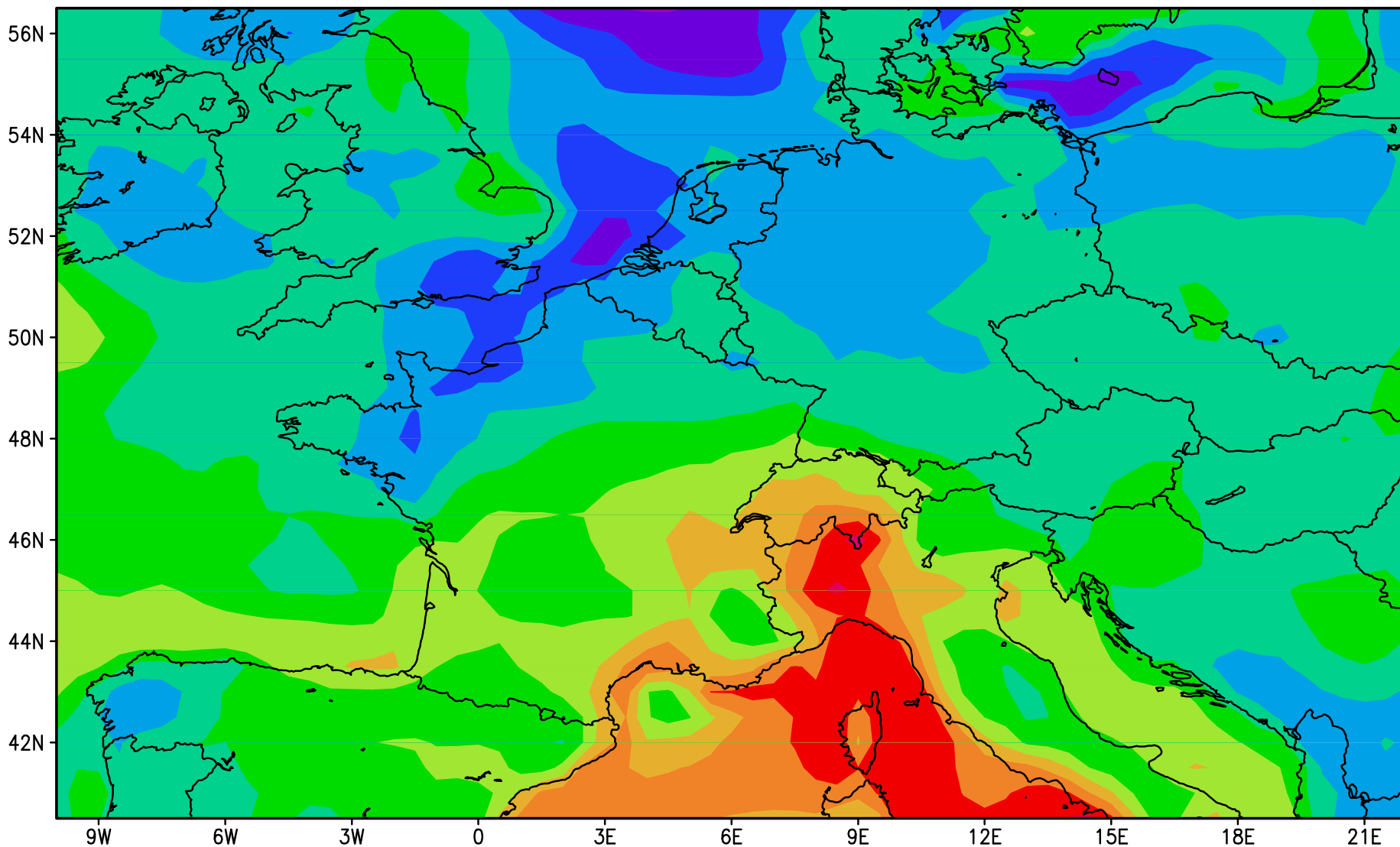
# Quelques problèmes...

- Dimensionnalité : vecteur d'état  $\simeq$  100-300 degrés de liberté (évolution vers les aérosols) ;
- CPU : une journée Europe = 5 minutes CPU (modèle gazeux).
- Aucun test convaincant sur un système "jouet" (il faut restituer la complexité/la dimensionnalité/les incertitudes sur les données).

# “État de l’art”

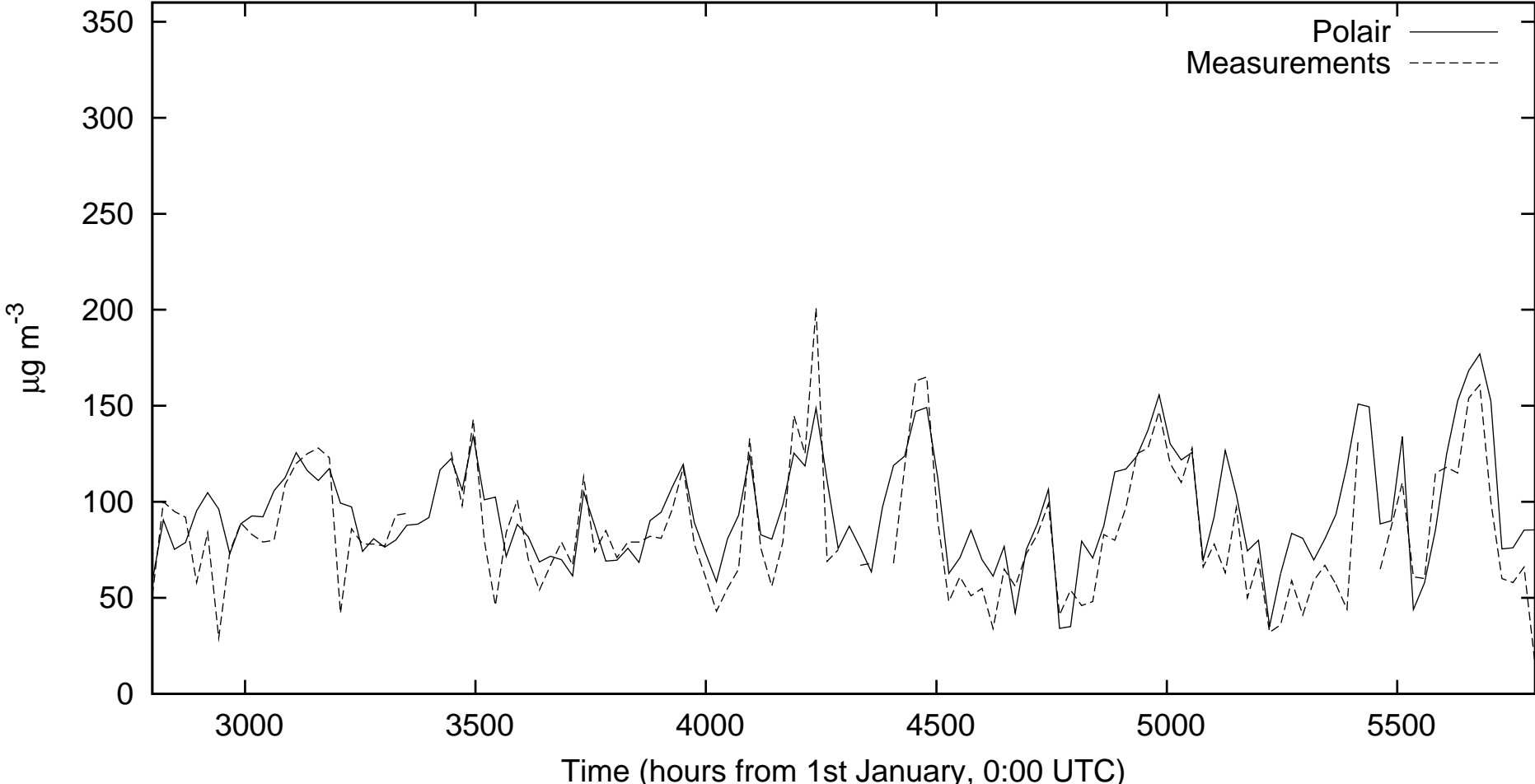
- Modèles relativement “bien validés” notamment pour les données terrestres et quelques polluants (ozone) -> utilisation en prévision opérationnelle (Prév’air) et en modélisation inverse des flux d’émission.
- Très fortes incertitudes : données d’entrée, forçage météo, paramétrisation -> prévision d’ensemble (Mallet).
- Assimilation de données 4D-var (Elbern, Quélo/Mallet pour de la modélisation inverse)
- Assimilation de données séquentielle (filtre de Kalman d’ensemble, éventuellement réduit : Segers, Torres).
- Pas de système “qui tourne”, pas de diagnostic sur 4D-var vs séquentiel.

# Quelques exemples



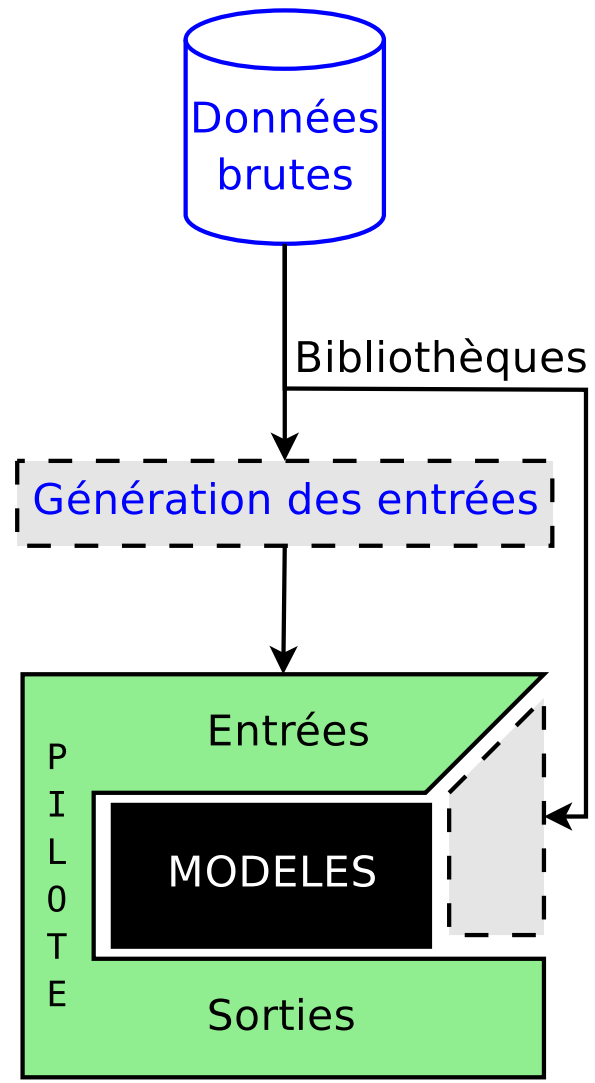
Carte d'ozone

# Quelques exemples



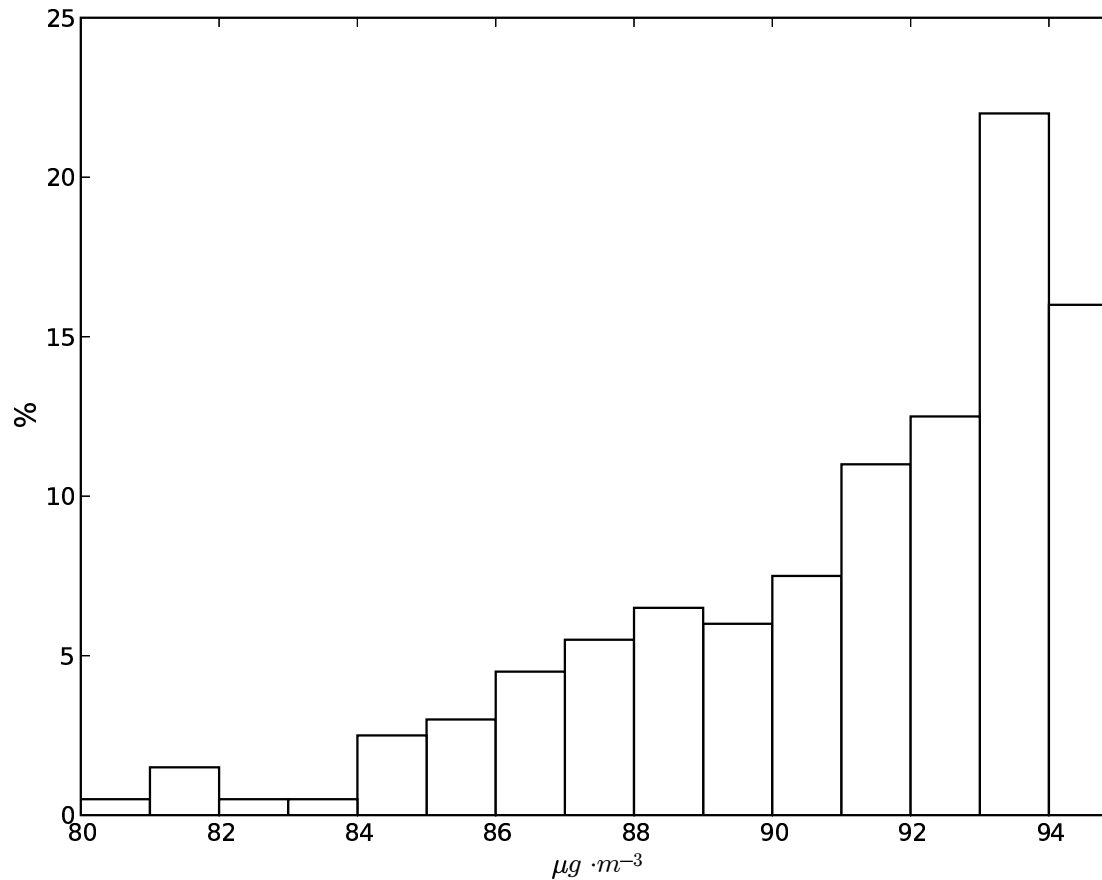
Comparaison typique aux observations (pics)

# Quelques exemples



Structure idéale du système

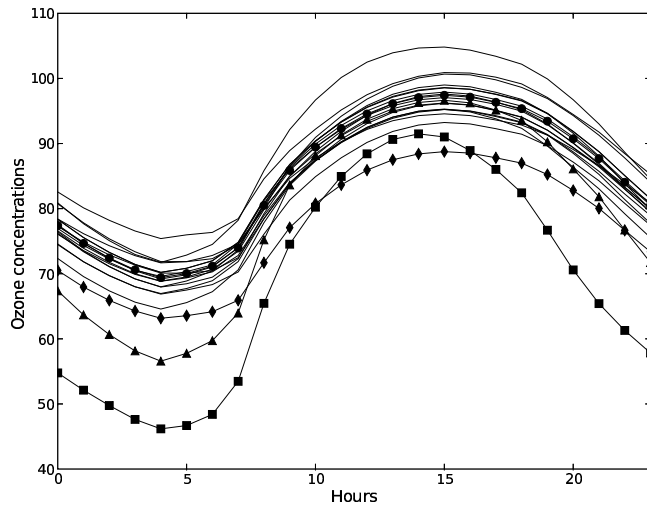
# Quelques exemples



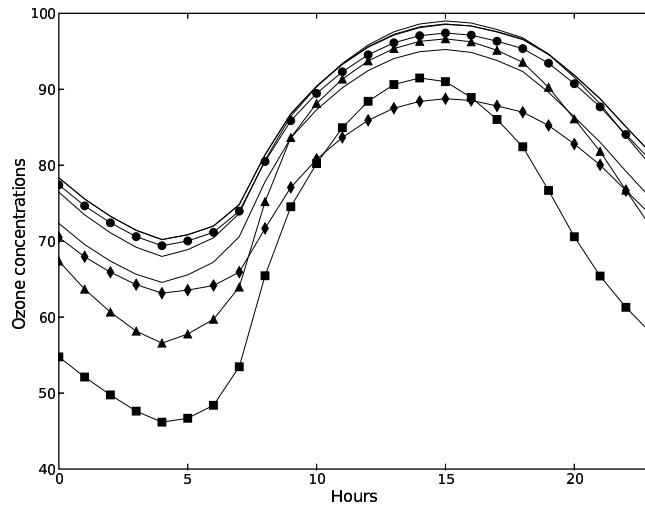
Réponse de l'ozone à une perturbation  
log-normale des émissions de NO



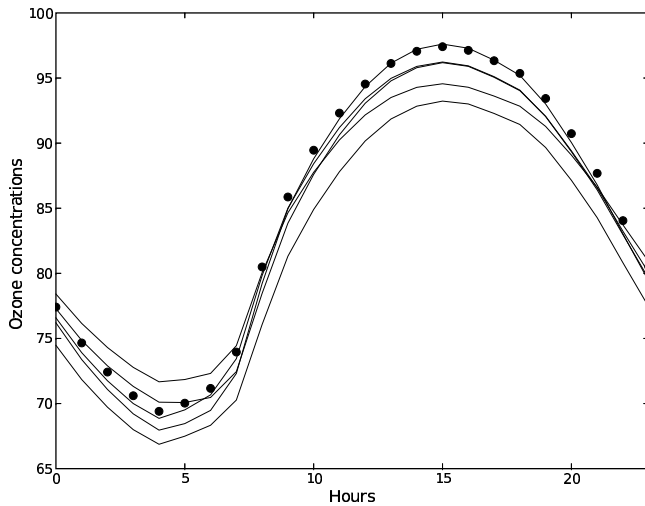
# Quelques exemples



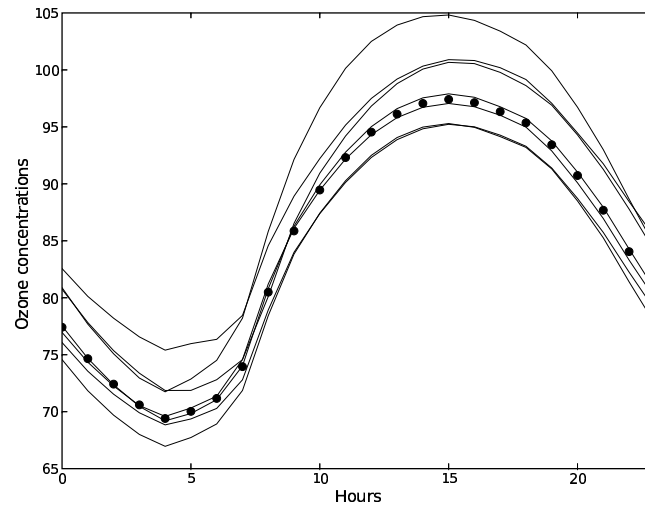
(all)



(param.)



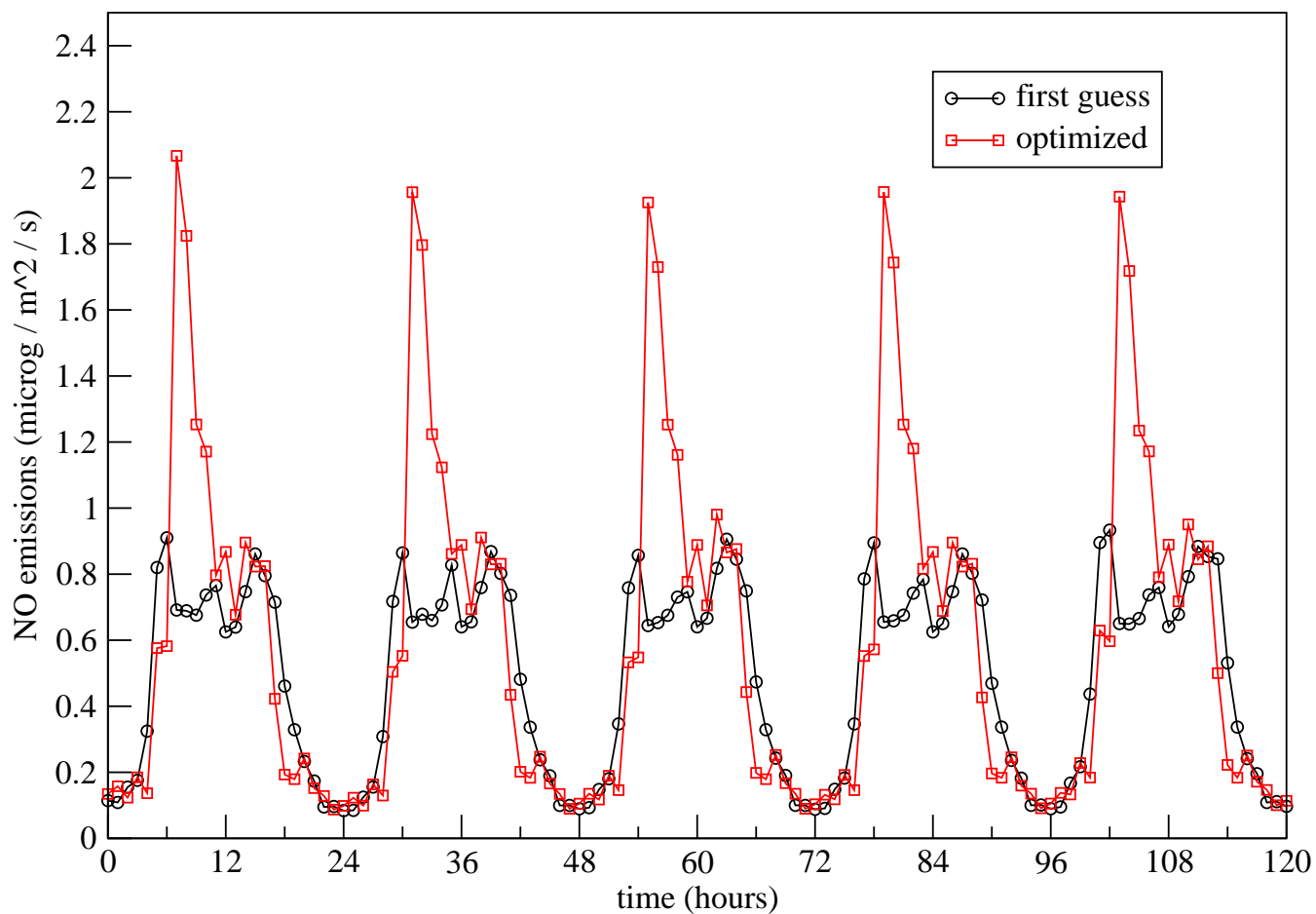
(data)



(num.)

## Multi-modèle

# Quelques exemples



Modélisation inverse d'émissions (4D-var, régional)

# Deux types d'approches

- Approche variationnelle (4D-Var pour un système dynamique). Maintenant utilisé en prévision météo (ECMWF, Météo-France).
  1. + : Traite les non-linéarités sans approximation (en théorie)
  2. + : N'exige pas d'explicitement les matrices de covariance d'analyse
  3. - : Techniquement lourd à mettre en oeuvre (adjoint du modèle, optimisation)
  4. - : Interprétation difficile
- Approche séquentielle (dérivés des filtres de Kalman). Surtout utilisé dans la communauté océanographique, et dans la communauté météo nord-américaine.
  1. + : Facile à mettre en oeuvre (algèbre linéaire + modèle)
  2. + : Plus facile d'interprétation (matrices de covariance calculées)
  3. - : Matrice de covariances explicite de taille colossale.
  4. - : Les non-linéarités ne sont qu'approchées.

# Assimilation séquentielle : les non-linéarités

- Un modèle de chimie-transport est non-linéaire. Comment traiter les non-linéarités locales (chimie dans une maille) ?
  1. Utiliser le modèle direct pour propager les matrices de covariances d'erreur : échantillonnage de la matrice de covariance, puis propagation des (état moyen + modes).
  2. De nombreuses variantes : déguisé dans le filtre de Kalman d'ensemble (EnKF), clairement annoncé dans certains filtres racines ( $\equiv$  qui ne propagent pas les matrices mais les modes) comme le filtre UKF (unscented), filtre SEIK.
- Les non-linéarités engendrent des erreurs qui ne suivent pas toujours des distributions gaussiennes (changement de régime chimique). Comment en rendre compte ?
  1. La propagation des modes ne suffit pas : le filtre EnKF propage correctement les modes, mais ne considère que les moments d'ordre deux des erreurs. Il n'a donc qu'une description gaussienne des statistiques d'erreur.
  2. Actuellement, en géosciences, tous les filtres, mêmes bien justifiés, font cette approximation. Différent des filtres particuliers.
  3. L'impossibilité de faire mieux tient à la dimensionalité des systèmes
- Au bout du compte, est-ce important pour un modèle de chimie-transport ?

# Assimilation séquentielle : la dimension

La dimension a priori d'un système de chimie transport est grande : (# mailles) x (# espèces chimiques). Les matrices de covariances d'erreur sont donc incalculables et instockables. La solution : réduire la dimension du système.

- Représenter les matrices de covariance d'erreur par leurs modes :  $P = S S^T$ .
- Réduire l'espace des états à  $m$  : on ne sélectionne qu'un nombre limité de modes dominants associés aux matrices de covariance (diagonalisation).
- La conjonction de ces deux stratégies fournit les filtres racines : filtres RRSQRT.
- Problème : comment ne pas appauvrir le système dynamique ? Représenter la variabilité du système par de l'erreur modèle (matrice  $Q$ ). Pour un filtre racine, cela est réalisé en augmentant le nombre de modes ( $m + q$ ). A l'étape de réduction, les modes seront ramenés à  $m$ .
- Facteur limitant numérique : étape des  $m + 1$  propagations.
- Nouveau problème : que mettre dans  $Q$ , ou plus exactement ses modes ? La question est ouverte...
- Autre stratégie de réduction : la localisation de l'étape d'analyse (filtre de Kalman d'ensemble local). Pourrait permettre à terme d'utiliser localement des filtres particuliers (Van Leeuwen).

# Le programme de travail tel qu'on le voit

- Objectif : diagnostic des “bonnes méthodes” + développement et test de méthodes “avancées” + application sur une plate-forme “opérationnelle” (Polyphemus/Prév'air).
- Finalisation d'un “terrain de jeu” réaliste (4D-var -sera fait-/EnKF -à finaliser-)  $\simeq$  6 mois.
- Diagnostic et tests comportement 4D-var vs EnKF (6 mois).
- Nouvelles méthodes (localisation ? couplage à de la prévision d'ensemble ? 4D-var avec erreur modèle ?) -> le problème de la réduction (6 mois).