



# Prendre en compte les incertitudes dans la modélisation de dispersion atmosphérique de polluants en milieu urbain

Mélanie Rochoux, Eliott Lumet\*, Thomas Jaravel et Simon Lacroix

*Collaboration Cerfacs / CNRS-LAAS, Toulouse*

*\* Thèse de doctorat soutenue en janvier 2024 (<https://theses.fr/s295419>)*



# La dispersion atmosphérique de polluants en milieu urbain

## Évènements de grande ampleur

- ▶ Forts enjeux environnementaux et de santé publique
- ▶ Complexité de l'évènement
  - Méconnaissance de la source (substances chimiques, débit d'émission)
  - Besoin de cartographier les niveaux de toxicité et l'exposition des individus dans une large zone autour de la source



*Incendie d'un entrepôt à Aubervilliers en 2023*



*Incendie de Notre Dame de Paris en avril 2019*



*Incendie de l'usine Lubrizol à Rouen en septembre 2019*

# La dispersion atmosphérique de polluants en milieu urbain

## Les problématiques à micro-échelle

- ▶ Caractériser la source
- ▶ Caractériser les processus en champ proche (1 km autour de la source)
- ▶ Besoin de développer des approches de modélisation détaillées
  - *Micro-météorologie urbaine (résolution décamétrique, < 100 m)*
  - *Emission et transport d'espèces chimiques potentiellement toxiques (gaz, aérosols)*

# La dispersion atmosphérique de polluants en milieu urbain

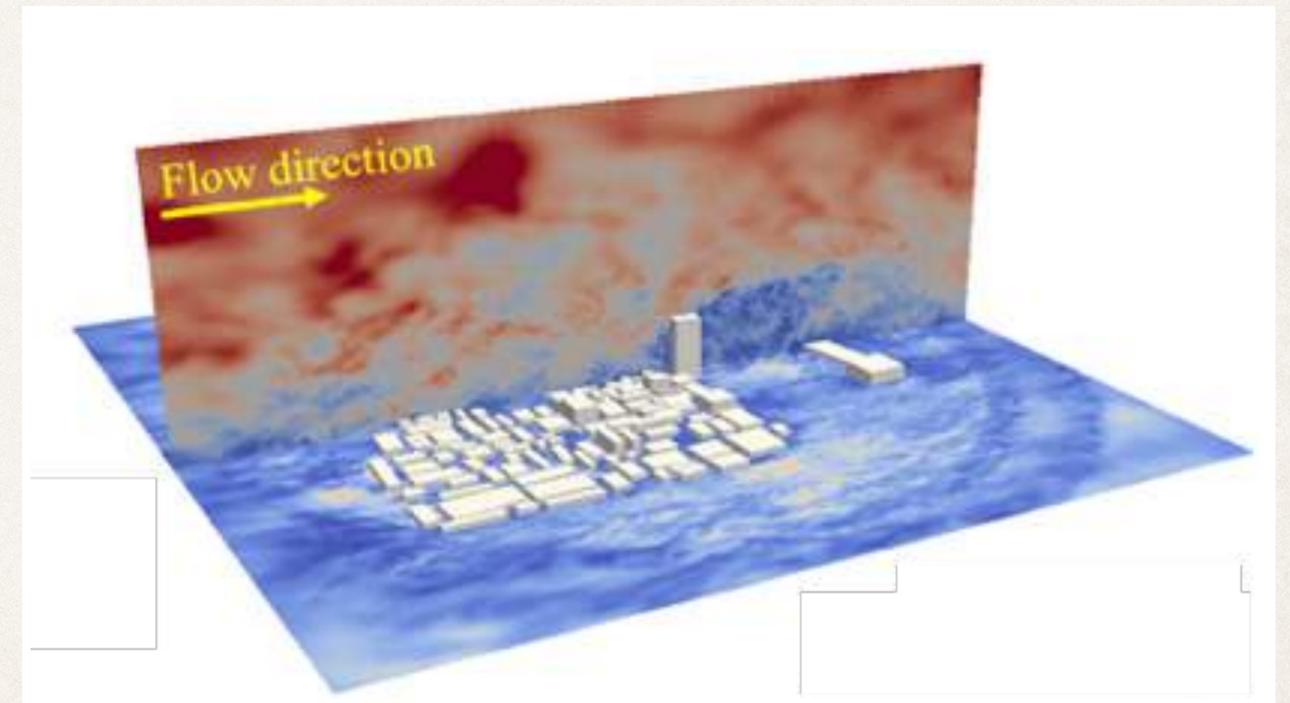
## Les problématiques à micro-échelle

- ▶ Caractériser la source
- ▶ Caractériser les processus en champ proche (1 km autour de la source)
- ▶ Besoin de développer des approches de modélisation détaillées
  - *Micro-météorologie urbaine (résolution décimétrique, < 100 m)*
  - *Emission et transport d'espèces chimiques potentiellement toxiques (gaz, aérosols)*

## Spécificités de la micro-échelle

- ▶ 1. L'effet du milieu bâti
- ▶ 2. Les interactions d'échelles : méso-échelle (10-100 km) et micro-échelle (10-100 m)
- ▶ 3. Le caractère turbulent de l'écoulement

➔ **Nécessité d'intégrer ces phénomènes complexes dans les modèles de dispersion pour capturer le « champ proche »**

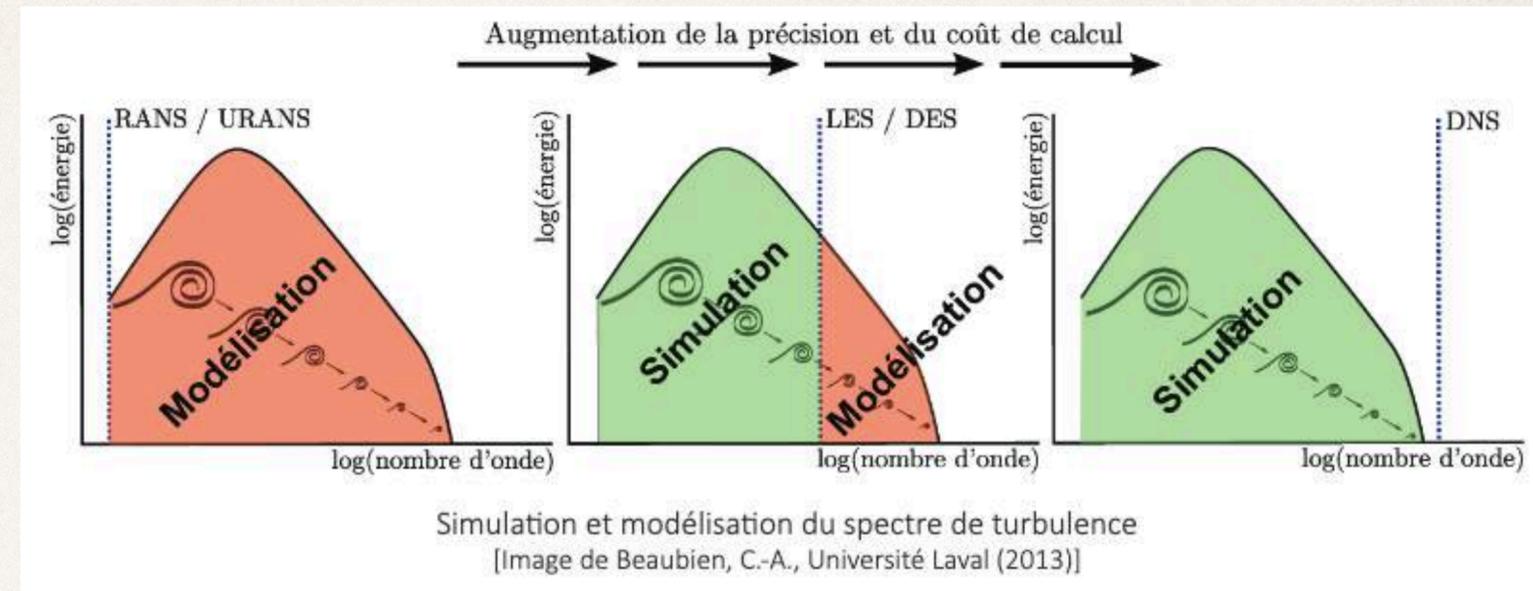


*Simulation de l'écoulement atmosphérique instantanée dans un bidonville de Dacca (Bangladesh)  
Hwang et Gorlé (Flow, 2023)*

# La dispersion atmosphérique de polluants en milieu urbain

## Les simulations aux grandes échelles (LES)

- ▶ Résolution des équations de Navier-Stokes filtrées
- ▶ Avantages des simulations LES
  - Limitent les erreurs de modélisation liées à la turbulence
  - Permettent une représentation spatio-temporelle des processus



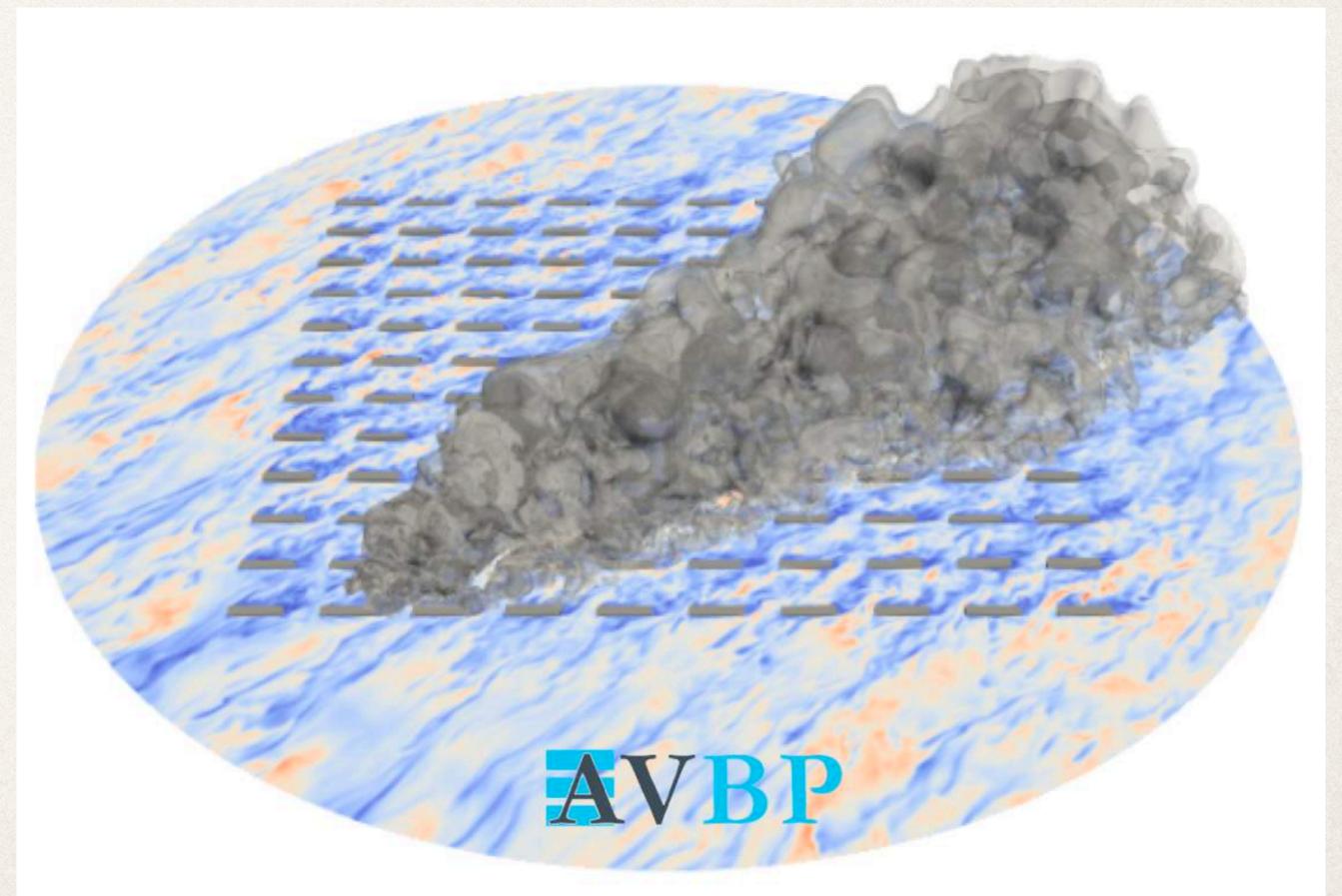
# La dispersion atmosphérique de polluants en milieu urbain

## Les simulations aux grandes échelles (LES)

- ▶ Résolution des équations de Navier-Stokes filtrées
- ▶ Avantages des simulations LES
  - *Limitent les erreurs de modélisation liées à la turbulence*
  - *Permettent une représentation spatio-temporelle des processus*

### Les objectifs du Cerfacs

- ▶ 1. Evaluer la **qualité** (et les limites) des simulations LES
  - ▶ 2. Proposer une **méthodologie** pour fournir aux modèles de dispersion opérationnels, les informations détaillées de la LES tout en tenant compte des incertitudes
- ➔ **Quantification et réduction des incertitudes associées aux simulations LES de dispersion atmosphérique**

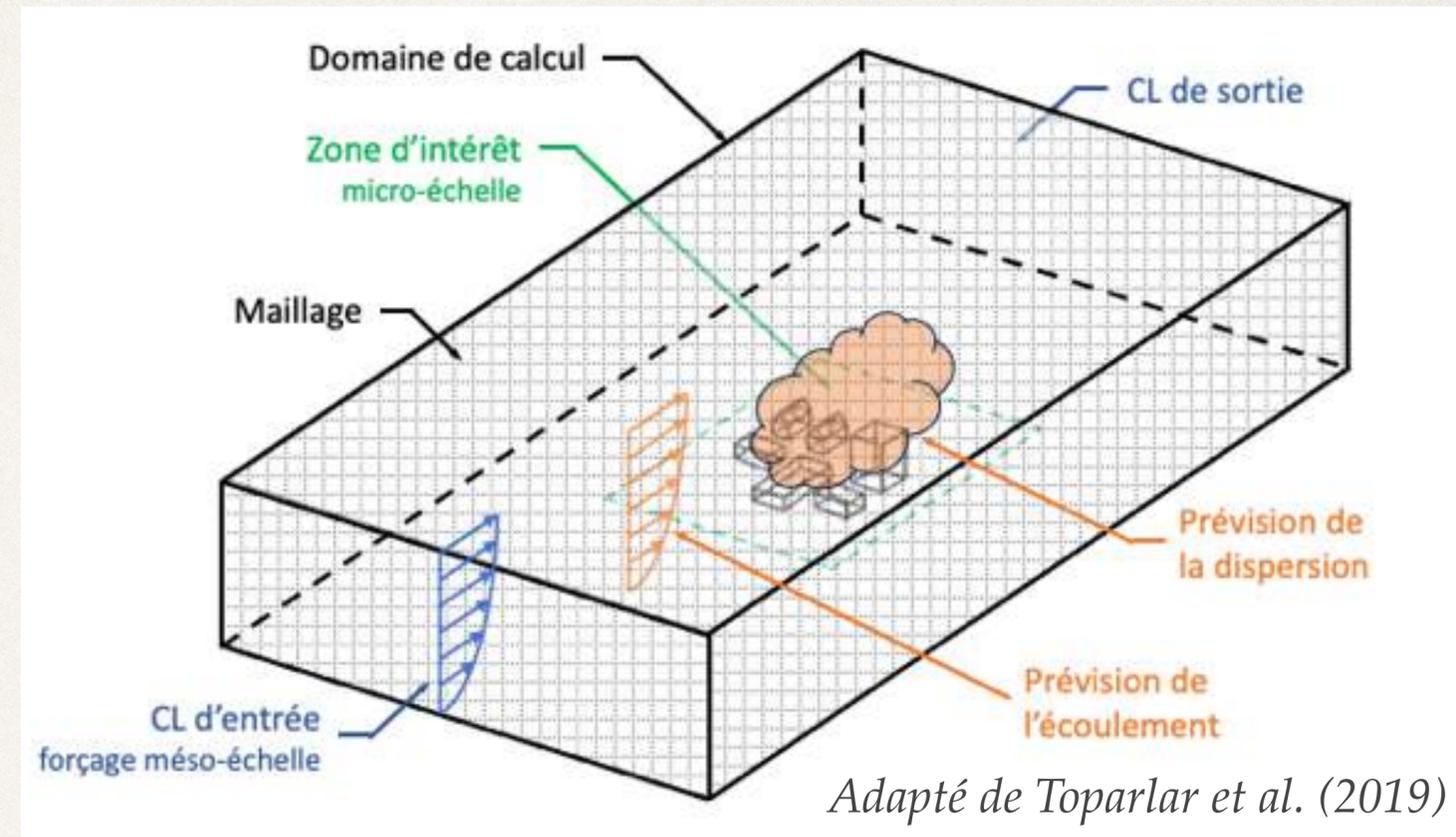


*L'expérience de terrain MUST comme cas de démonstration  
Lumet et al. (BLM, 2024)*

# La dispersion atmosphérique de polluants en milieu urbain

## Les incertitudes liées à la dispersion à micro-échelle

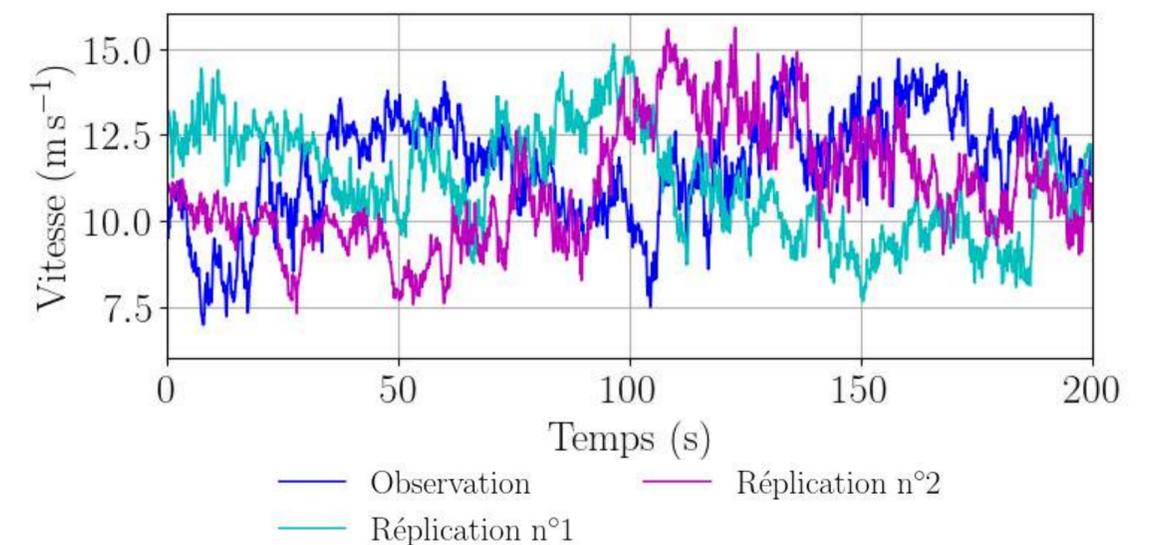
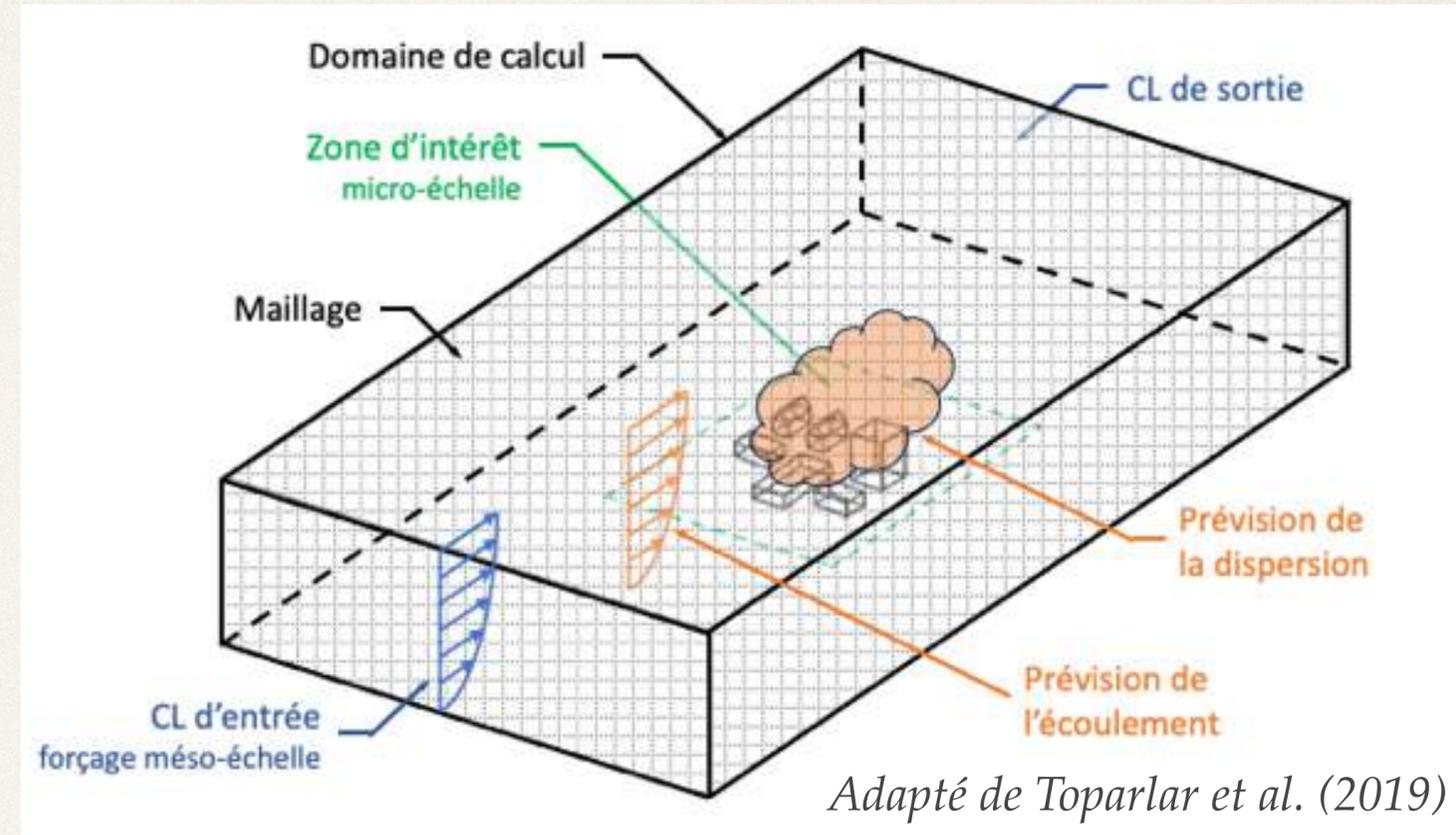
- ▶ Modélisation
  - *Modèle de turbulence de sous-maille, schémas numériques...*
- ▶ Conditions aux limites
  - *forçage météorologique, source, représentation du milieu bâti*
- ▶ Variabilité interne atmosphérique
  - *Dauxois et al. (2021, Physical Review fluids)*



# La dispersion atmosphérique de polluants en milieu urbain

## Les incertitudes liées à la dispersion à micro-échelle

- ▶ Modélisation
  - *Modèle de turbulence de sous-maille, schémas numériques...*
- ▶ Conditions aux limites
  - *forçage météorologique, source, représentation du milieu bâti*
- ▶ Variabilité interne atmosphérique
  - *Dauxois et al. (2021, Physical Review fluids)*



*Illustration de la variabilité interne sur le vent en un point de la canopée urbaine MUST - Lumet et al. (BLM, 2024)*

# La dispersion atmosphérique de polluants en milieu urbain

## Les incertitudes liées à la dispersion à micro-échelle

- ▶ Modélisation
  - *Modèle de turbulence de sous-maille, schémas numériques...*
- ▶ Conditions aux limites
  - *forçage météorologique, source, représentation du milieu bâti*
- ▶ Variabilité interne atmosphérique
  - *Dauxois et al. (2021, Physical Review fluids)*

## Problématiques étudiées sur le cas MUST

➔ Focus sur les incertitudes liées aux conditions atmosphériques

- ▶ 1. Quantifier l'importance de la **variabilité interne** sur les mesures de terrain et les simulations LES
- ▶ 2. Quantifier l'incertitude sur le **forçage atmosphérique** de grande échelle sur les prédictions de micro-échelle

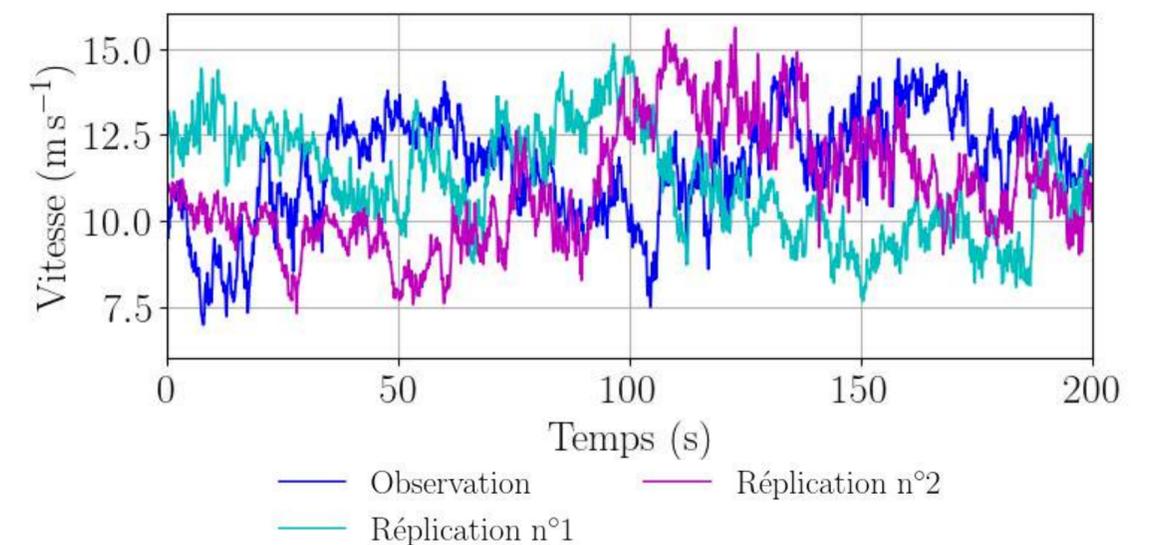
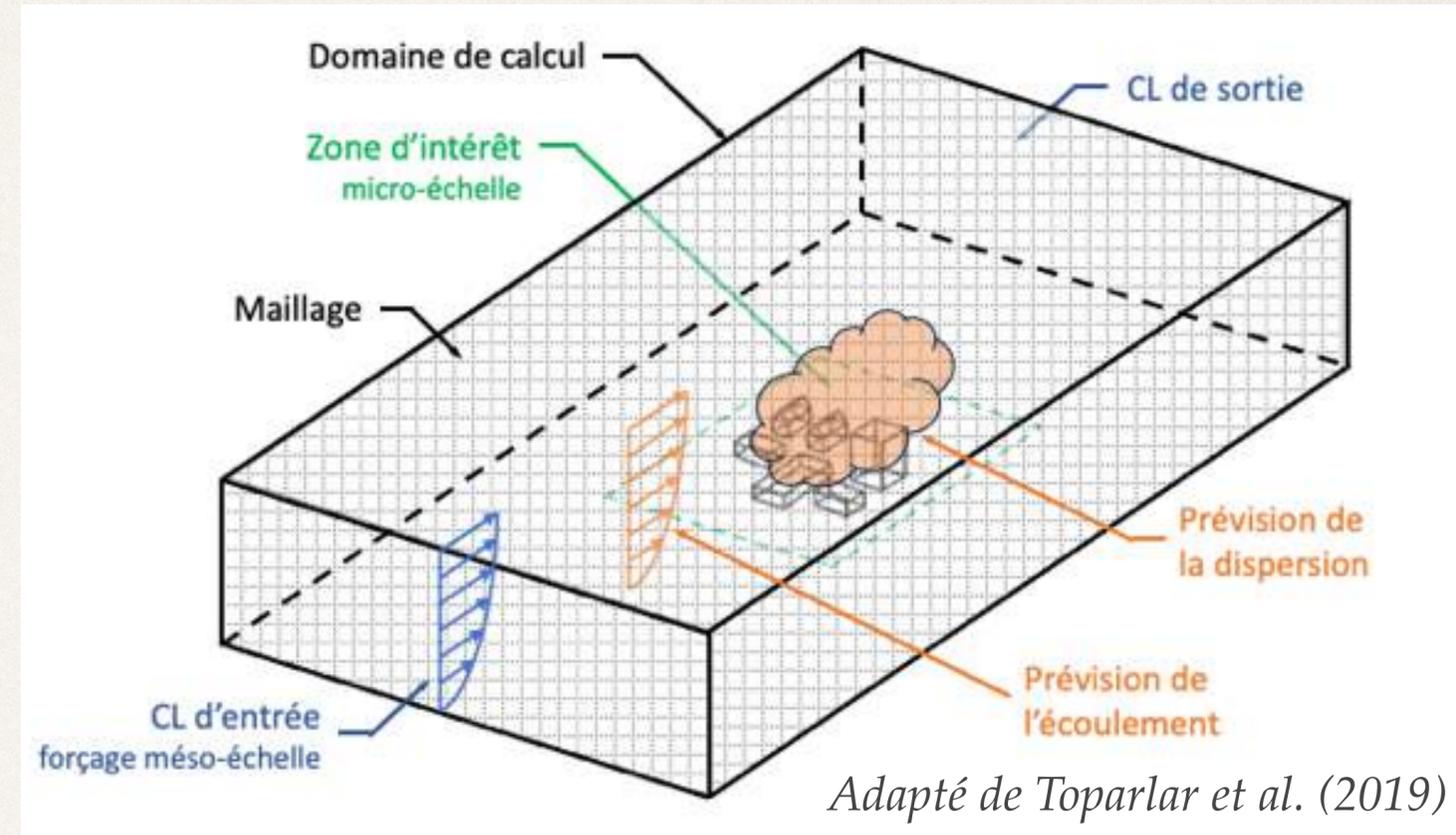


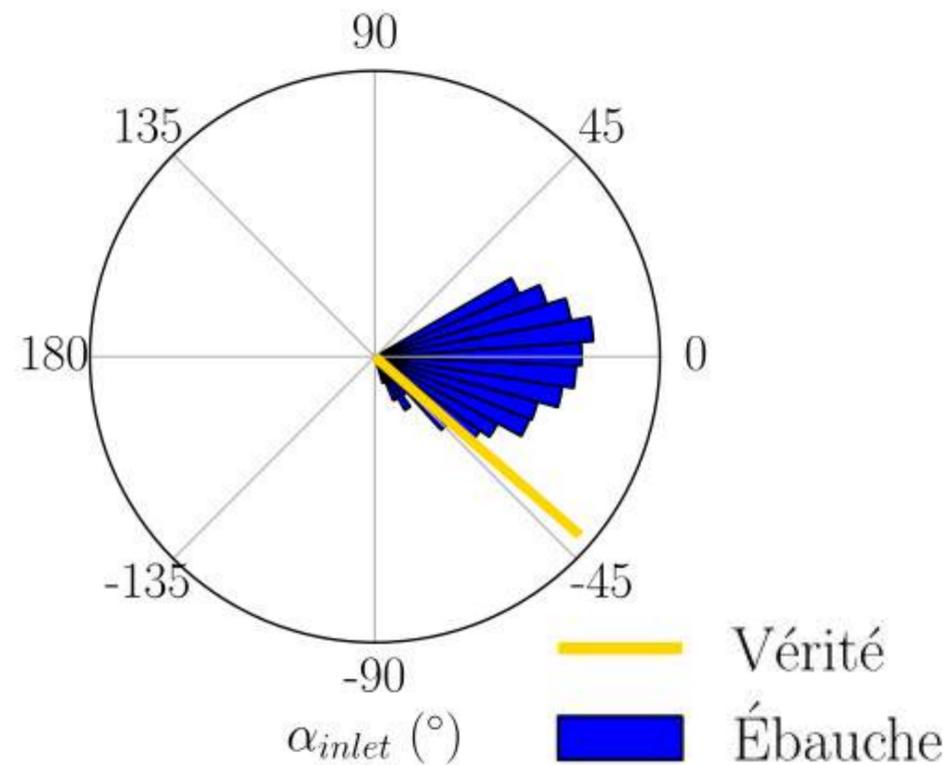
Illustration de la variabilité interne sur le vent en un point de la canopée urbaine MUST - Lumet et al. (BLM, 2024)

# La dispersion atmosphérique de polluants en milieu urbain

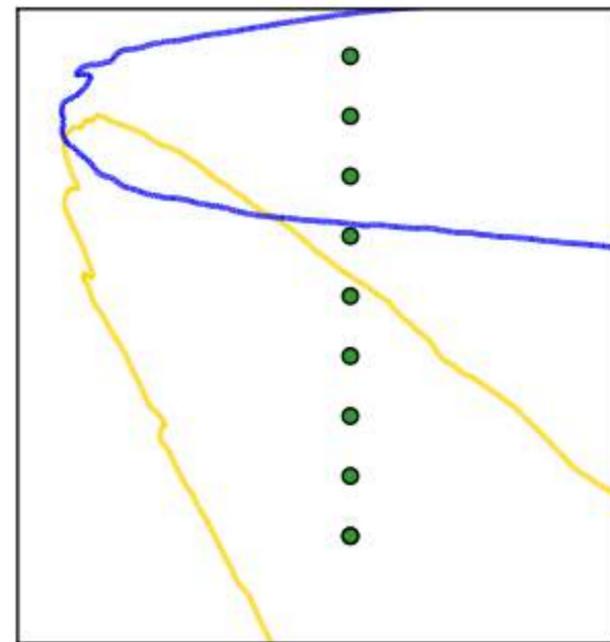
## Illustration de notre problématique pour la direction du vent

### ► Quantification des incertitudes

*Paramètre incertain*



*Iso-contour de concentration*



*Tiré de la thèse d'Eliott Lumet (2024)*

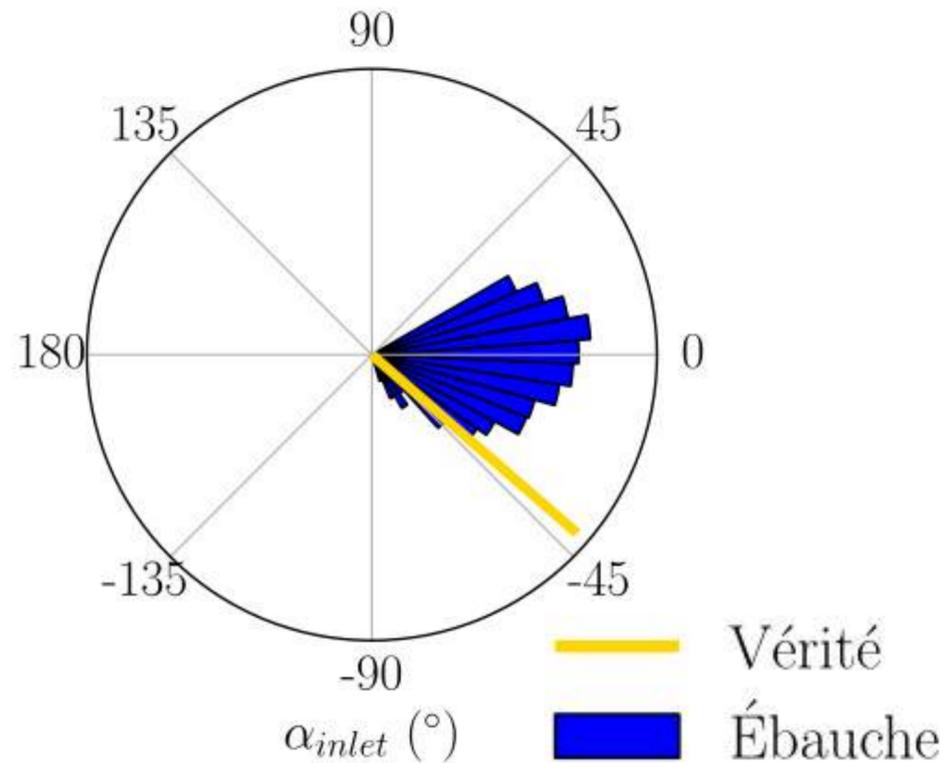
# La dispersion atmosphérique de polluants en milieu urbain

## Illustration de notre problématique pour la direction du vent

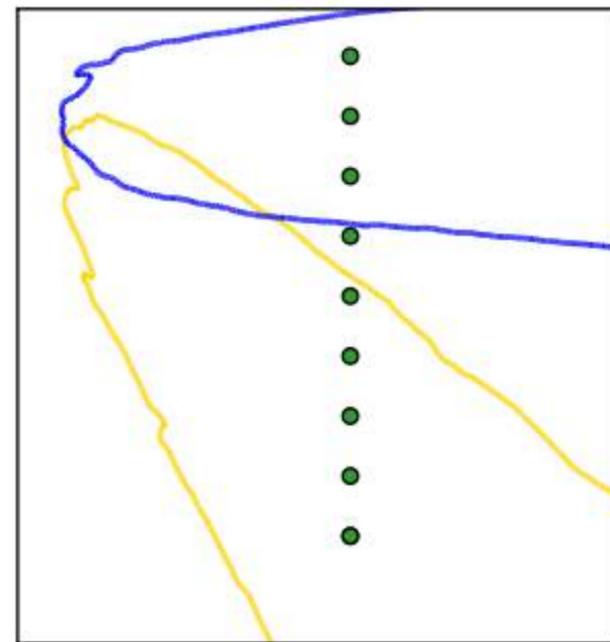
► Quantification des incertitudes

► Réduction des incertitudes (assimilation de données)

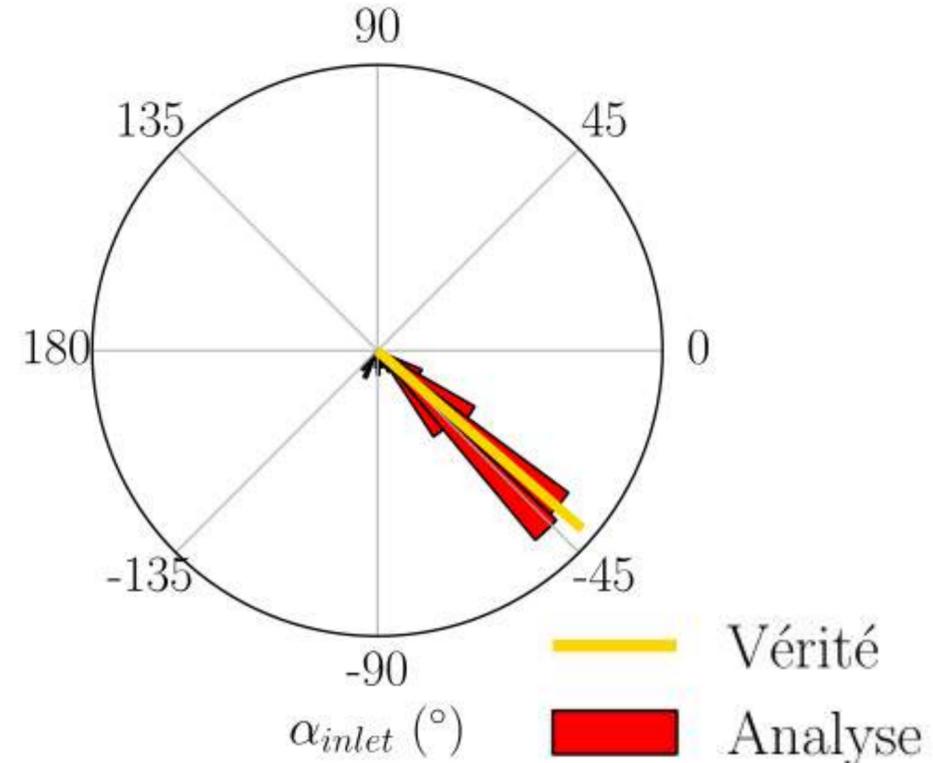
*Paramètre incertain*



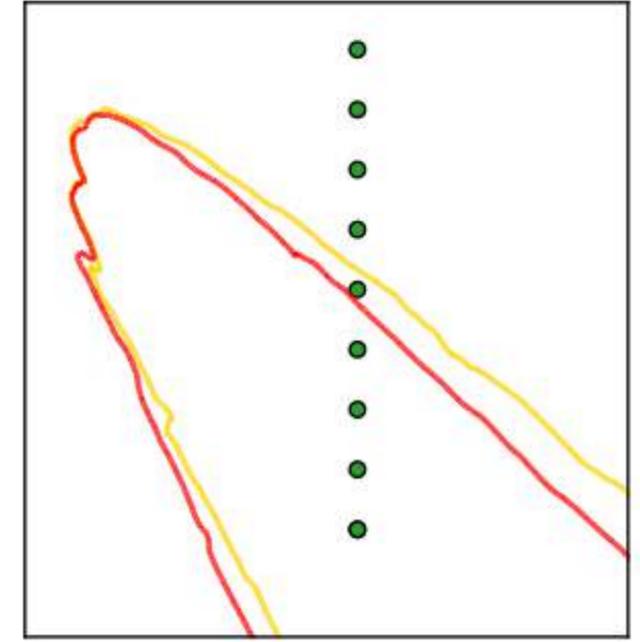
*Iso-contour de concentration*



*Paramètre corrigé*



*Iso-contour de concentration*



Tiré de la thèse d'Eliott Lumet (2024)

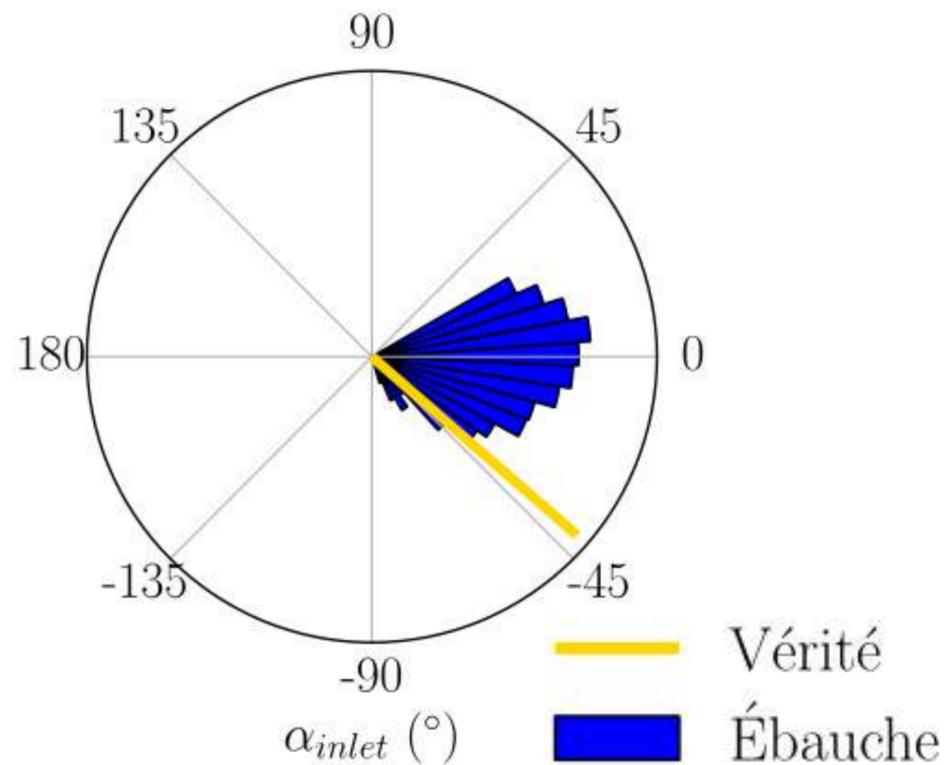
# La dispersion atmosphérique de polluants en milieu urbain

## Illustration de notre problématique pour la direction du vent

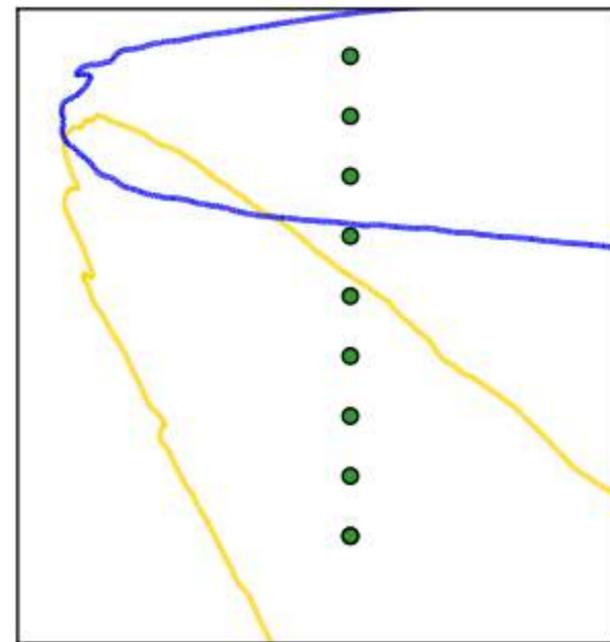
► Quantification des incertitudes

► Réduction des incertitudes (assimilation de données)

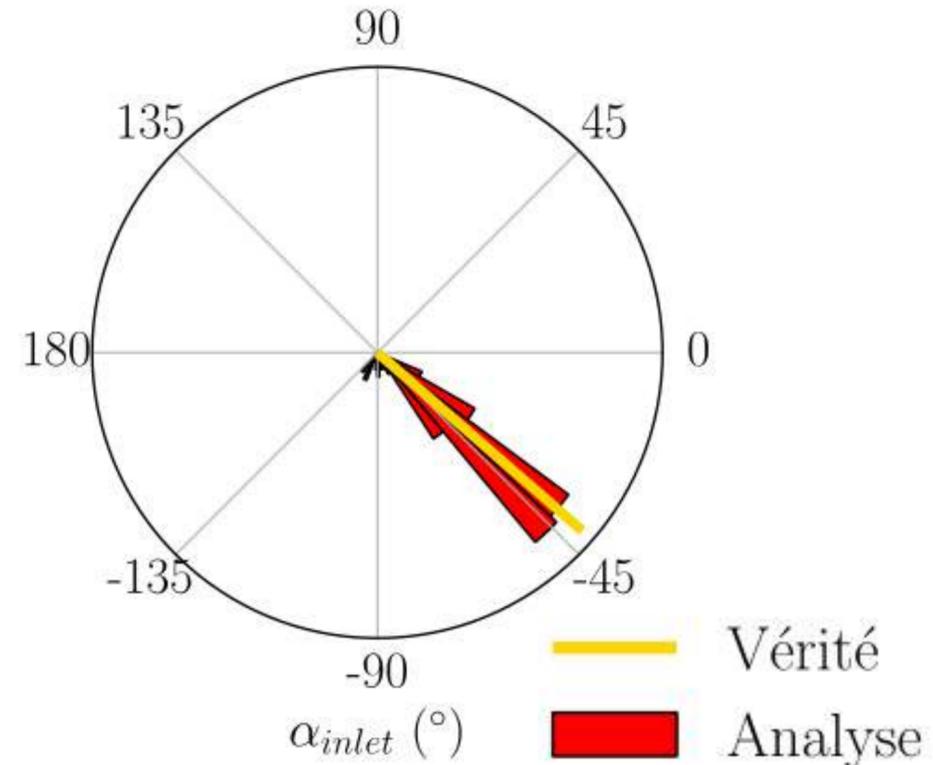
*Paramètre incertain*



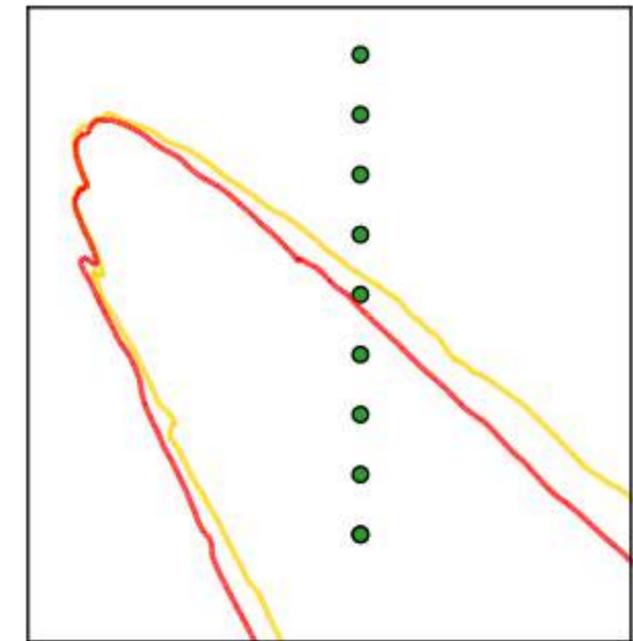
*Iso-contour de concentration*



*Paramètre corrigé*



*Iso-contour de concentration*



Tiré de la thèse d'Eliott Lumet (2024)

► Focus de l'exposé



# Prendre en compte les incertitudes dans les prévisions de dispersion atmosphérique de polluants en milieu urbain

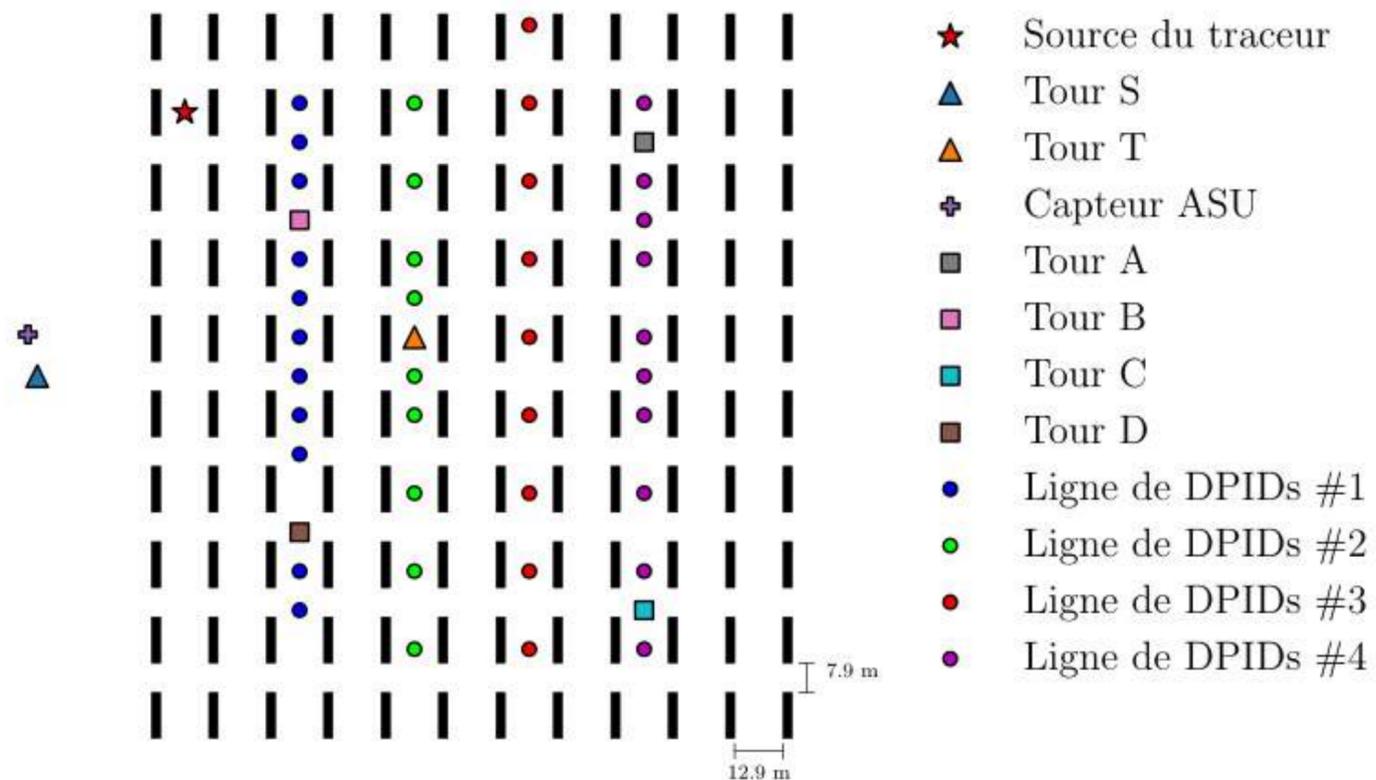
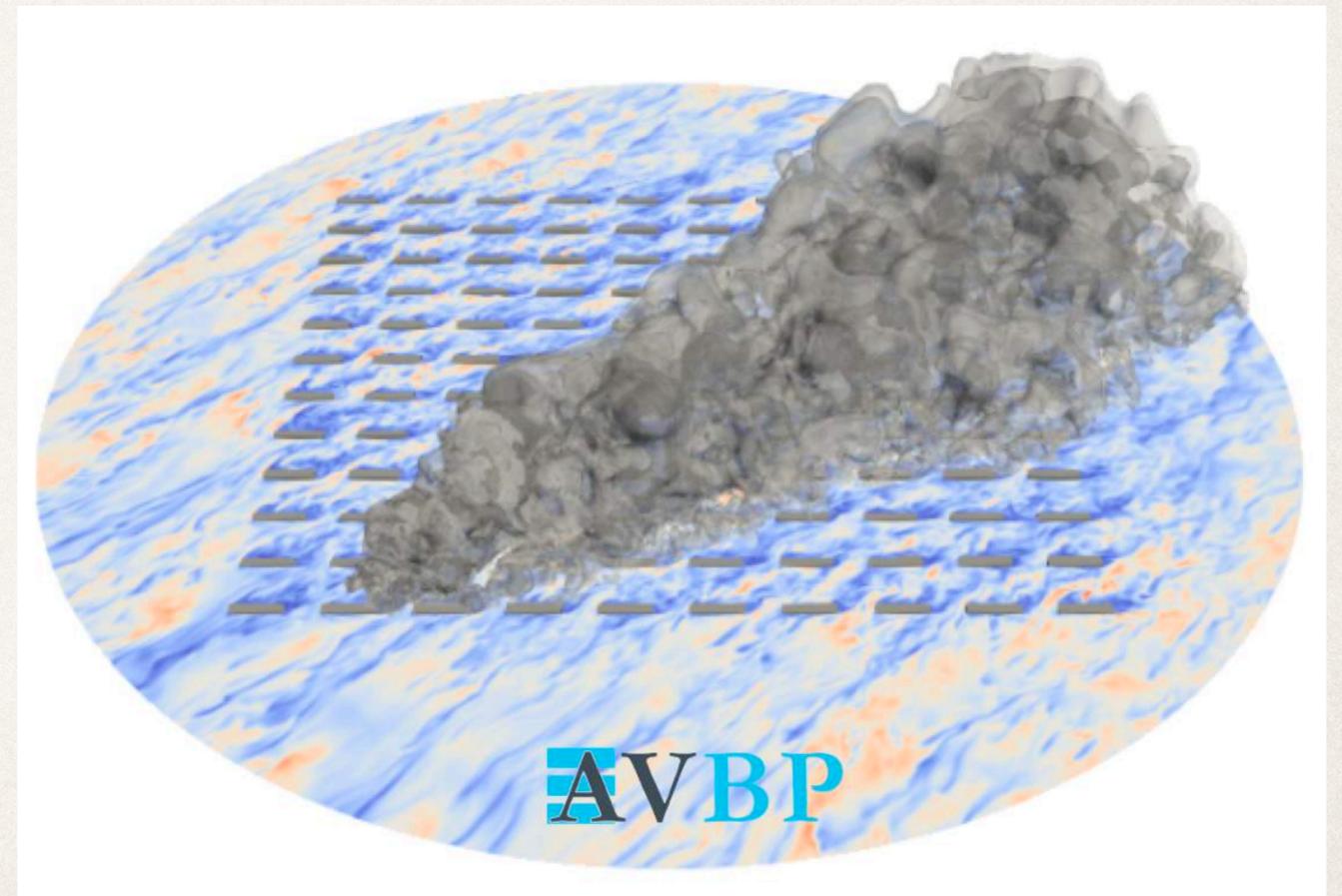
## I- La variabilité interne



# I/ Estimation de la variabilité interne

## Quelques éléments sur le cas d'étude MUST

- ▶ Cas de référence pour l'évaluation de modèles
- ▶ Code LES AVBP développé au Cerfacs et étendu aux applications micro-météorologiques
  - compressibilité artificielle
  - maillage non-structuré (90 millions de tétraèdres)
  - Coût d'une simulation : 20 000 hCPU



Simulation de l'essai 2681829 de l'expérience de terrain MUST dans le désert de l'Utah en 2001 (Lumet et al., BLM, 2024)

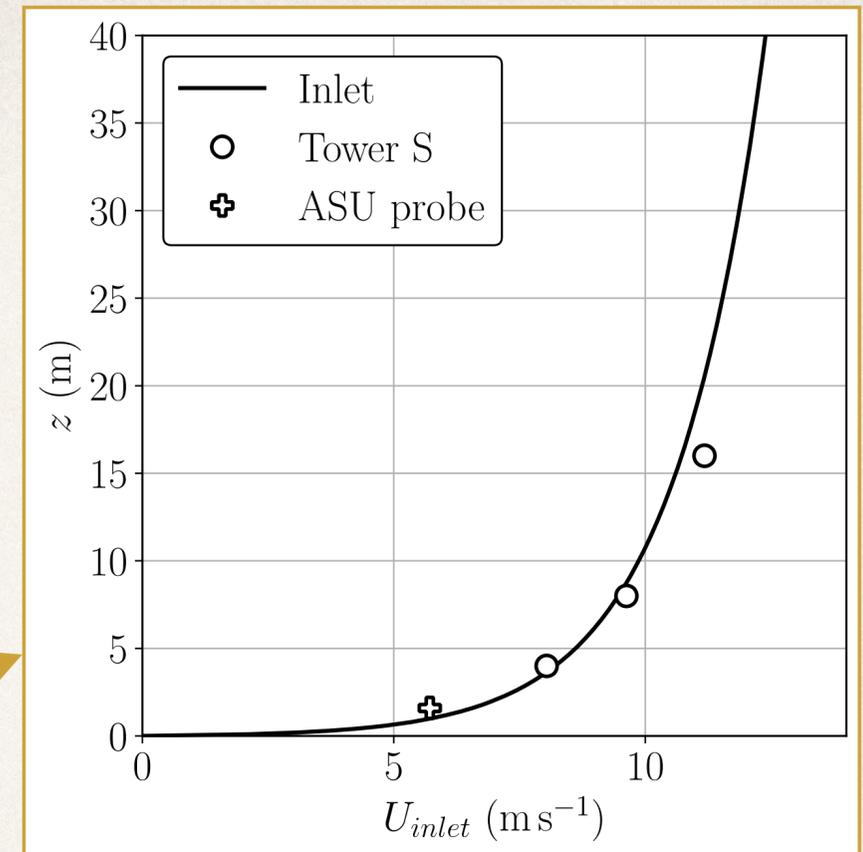
# I/ Estimation de la variabilité interne

## Calibration du forçage atmosphérique

### ► Profil moyen de vitesse

- Direction uniforme et homogène (-41°)
- Vitesse suivant un profil logarithmique typique de conditions neutres

$$U_{\text{inlet}}(z) = \frac{u^*}{\kappa} \ln \left( \frac{z + z_0}{z_0} \right)$$



*Profil vertical de vitesse imposé en entrée du domaine calé sur les mesures à la tour S en amont de la canopée urbaine*

# I/ Estimation de la variabilité interne

## Calibration du forçage atmosphérique

### ► Profil moyen de vitesse

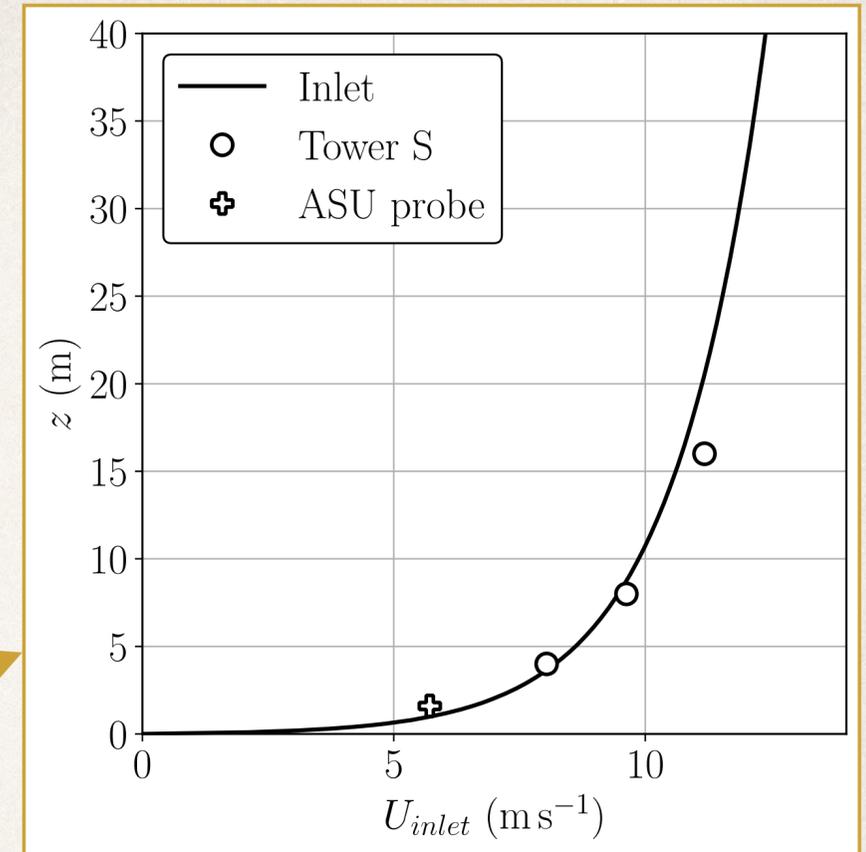
- Direction uniforme et homogène (-41°)
- Vitesse suivant un profil logarithmique typique de conditions neutres

### ► Injection de turbulence synthétique :

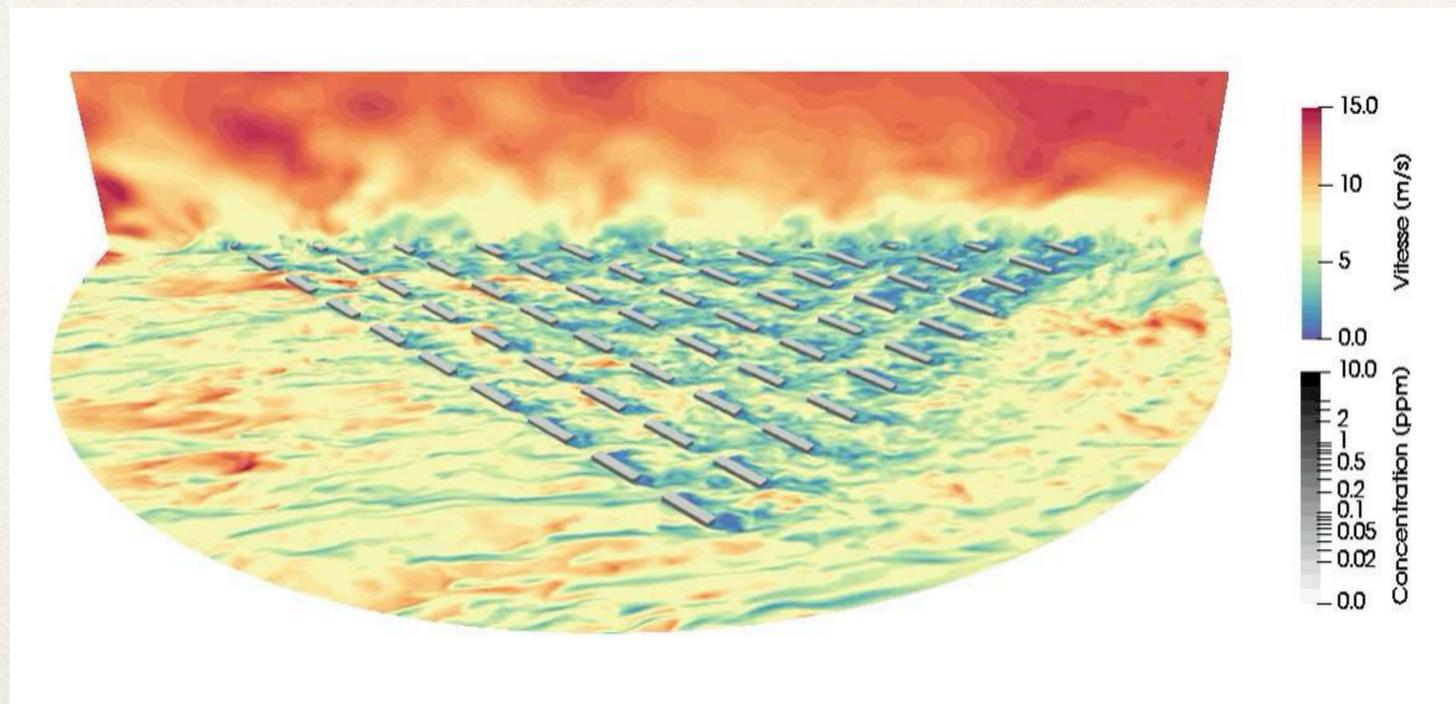
statistiques de turbulence obtenues d'une simulation précurseur en champ libre

*Vasaturo et al. (JWEIA 2018)*

$$U_{\text{inlet}}(z) = \frac{u^*}{\kappa} \ln \left( \frac{z + z_0}{z_0} \right)$$



*Profil vertical de vitesse imposé en entrée du domaine calé sur les mesures à la tour S en amont de la canopée urbaine*



# I/ Estimation de la variabilité interne

## Calibration du forçage atmosphérique

### ► Profil moyen de vitesse

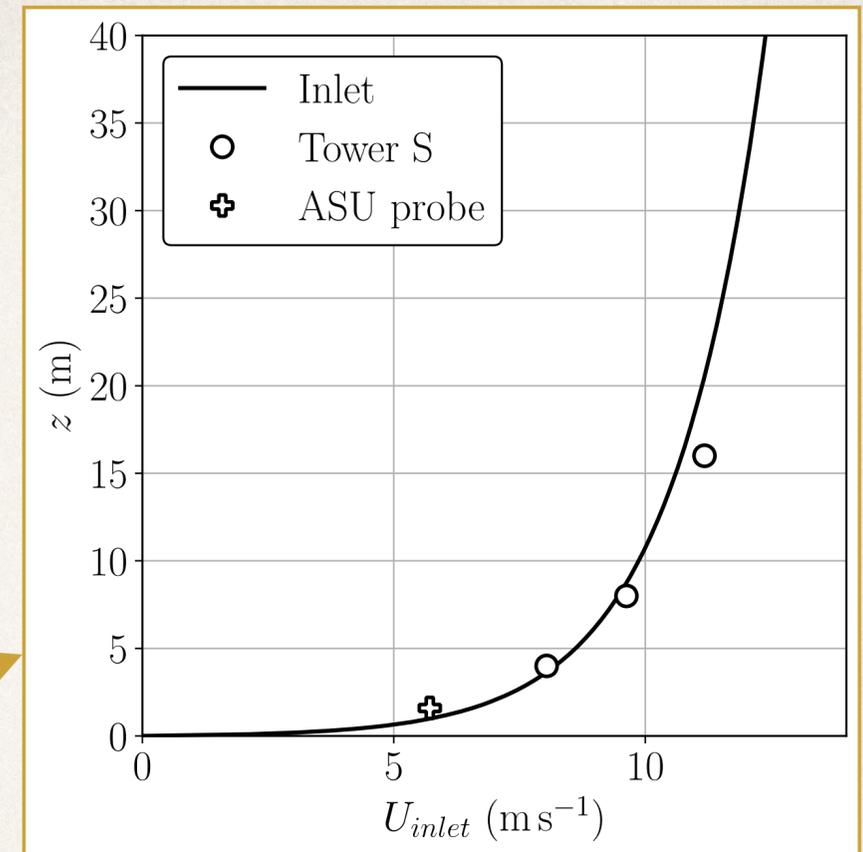
- Direction uniforme et homogène (-41°)
- Vitesse suivant un profil logarithmique typique de conditions neutres

### ► Injection de turbulence synthétique :

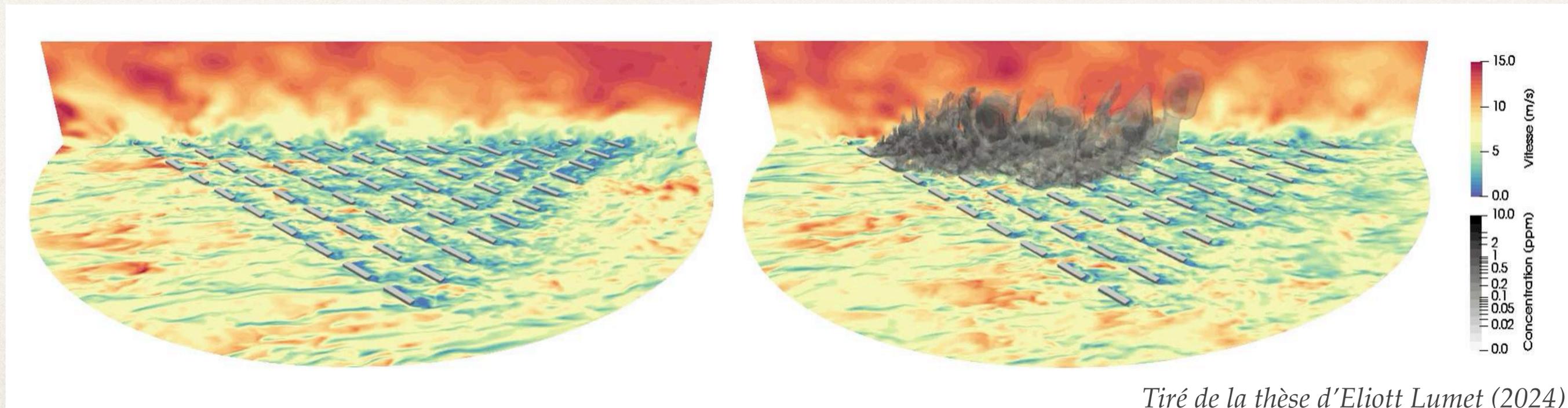
statistiques de turbulence obtenues d'une simulation précurseur en champ libre

*Vasaturo et al. (JWEIA 2018)*

$$U_{\text{inlet}}(z) = \frac{u^*}{\kappa} \ln \left( \frac{z + z_0}{z_0} \right)$$



*Profil vertical de vitesse imposé en entrée du domaine calé sur les mesures à la tour S en amont de la canopée urbaine*



*Tiré de la thèse d'Eliott Lumet (2024)*

# I/ Estimation de la variabilité interne

## Processus de validation du modèle LES

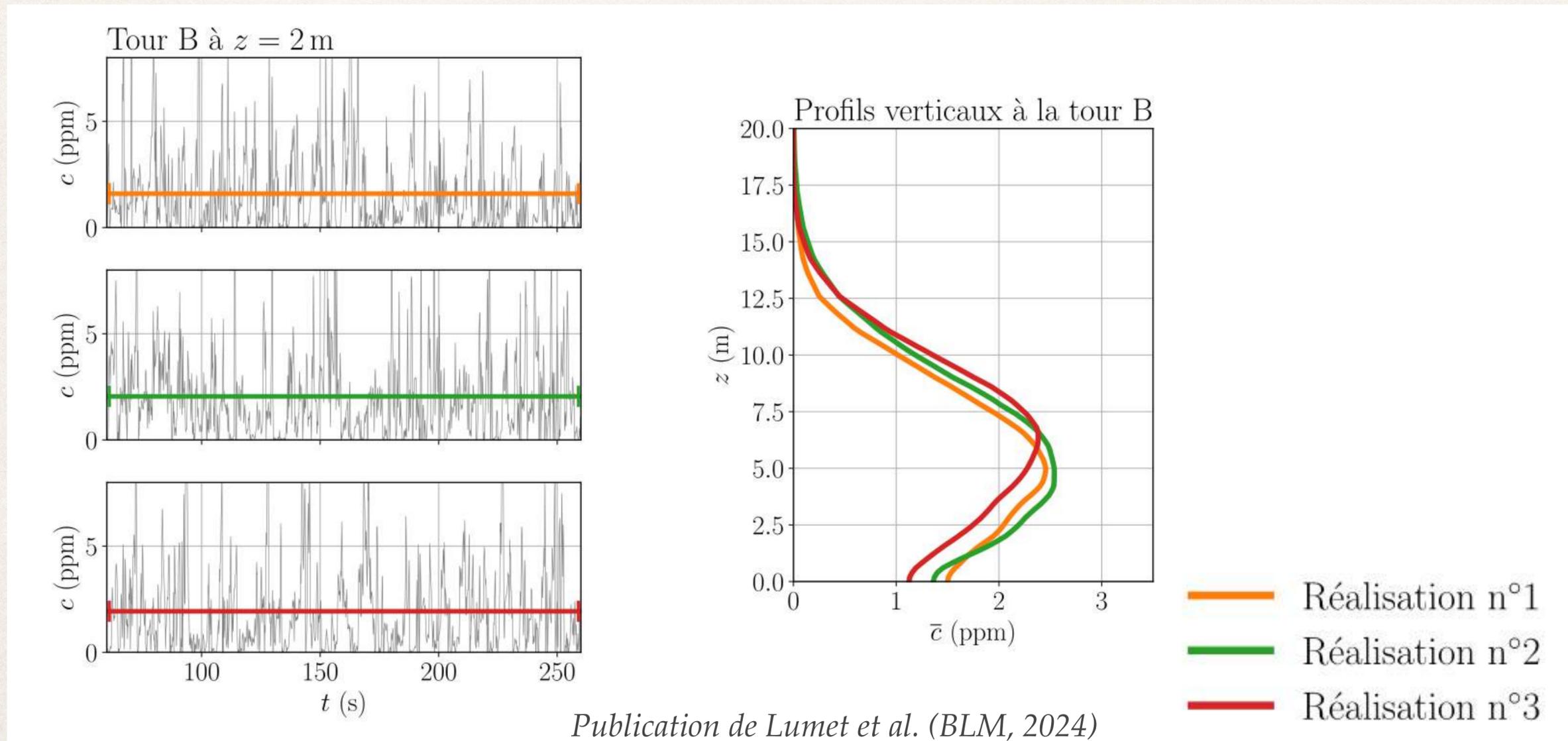
- ▶ **Comment comparer mesures et sorties du modèle LES ?**
  - Accès à des séries temporelles simulées en différents points de la canopée
  - Mesures de référence : moyennes et fluctuations RMS sur une période d'analyse de 200 s

# I/ Estimation de la variabilité interne

## Processus de validation du modèle LES

### ► Comment comparer mesures et sorties du modèle LES ?

- Accès à des séries temporelles simulées en différents points de la canopée
- Mesures de référence : moyennes et fluctuations RMS sur une période d'analyse de 200 s

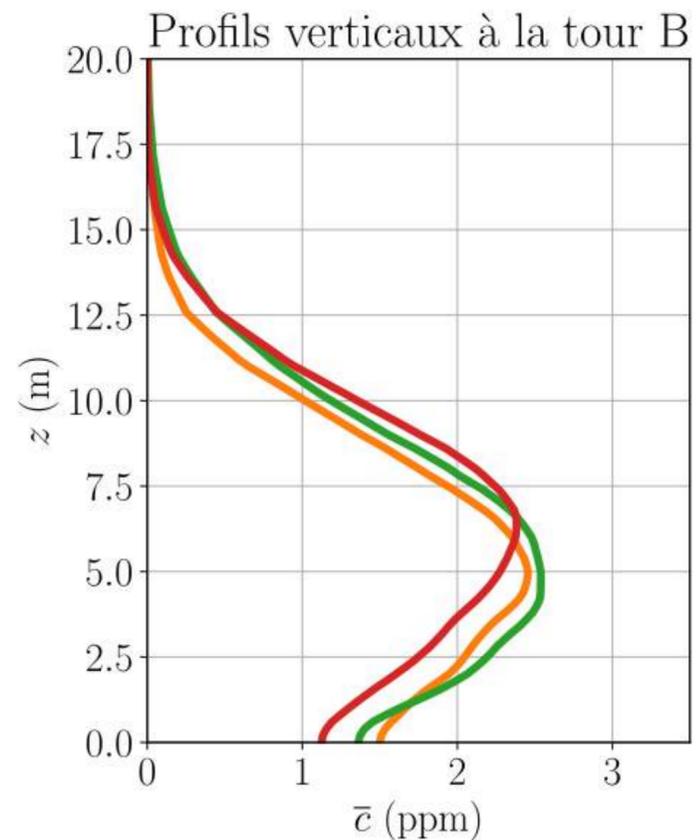
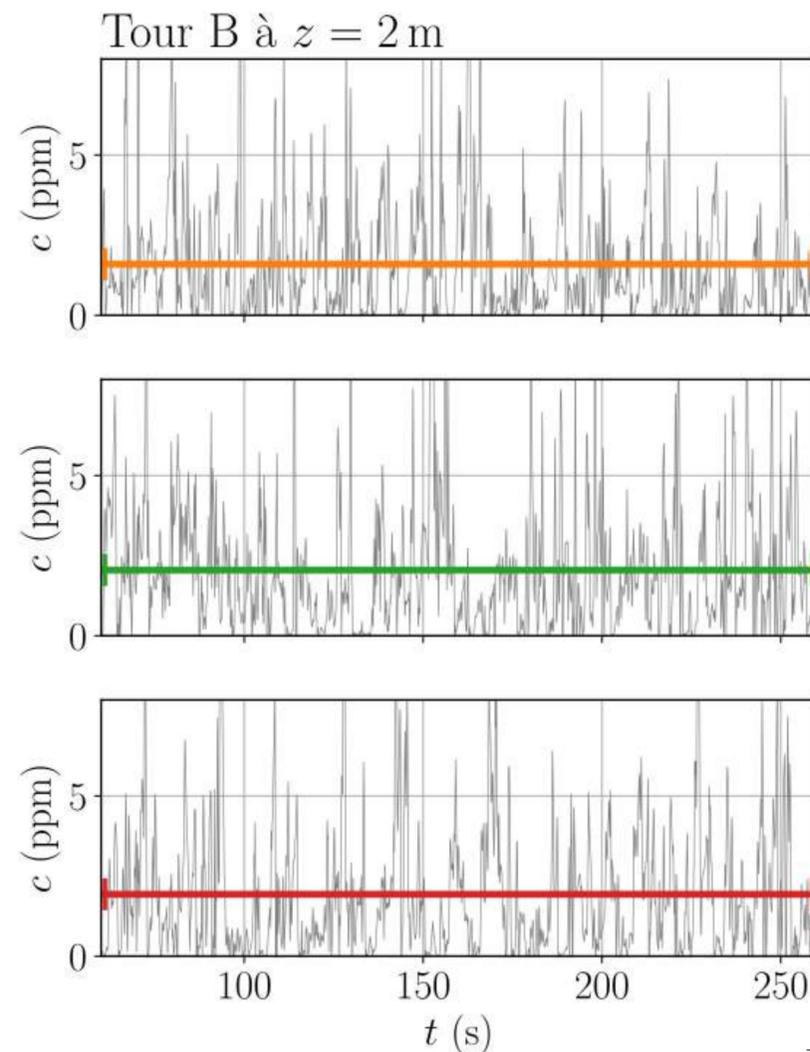


# I/ Estimation de la variabilité interne

## Processus de validation du modèle LES

### ► Comment comparer mesures et sorties du modèle LES ?

- Accès à des séries temporelles simulées en différents points de la canopée
- Mesures de référence : moyennes et fluctuations RMS sur une période d'analyse de 200 s



### Quelle fenêtre d'analyse choisir ?

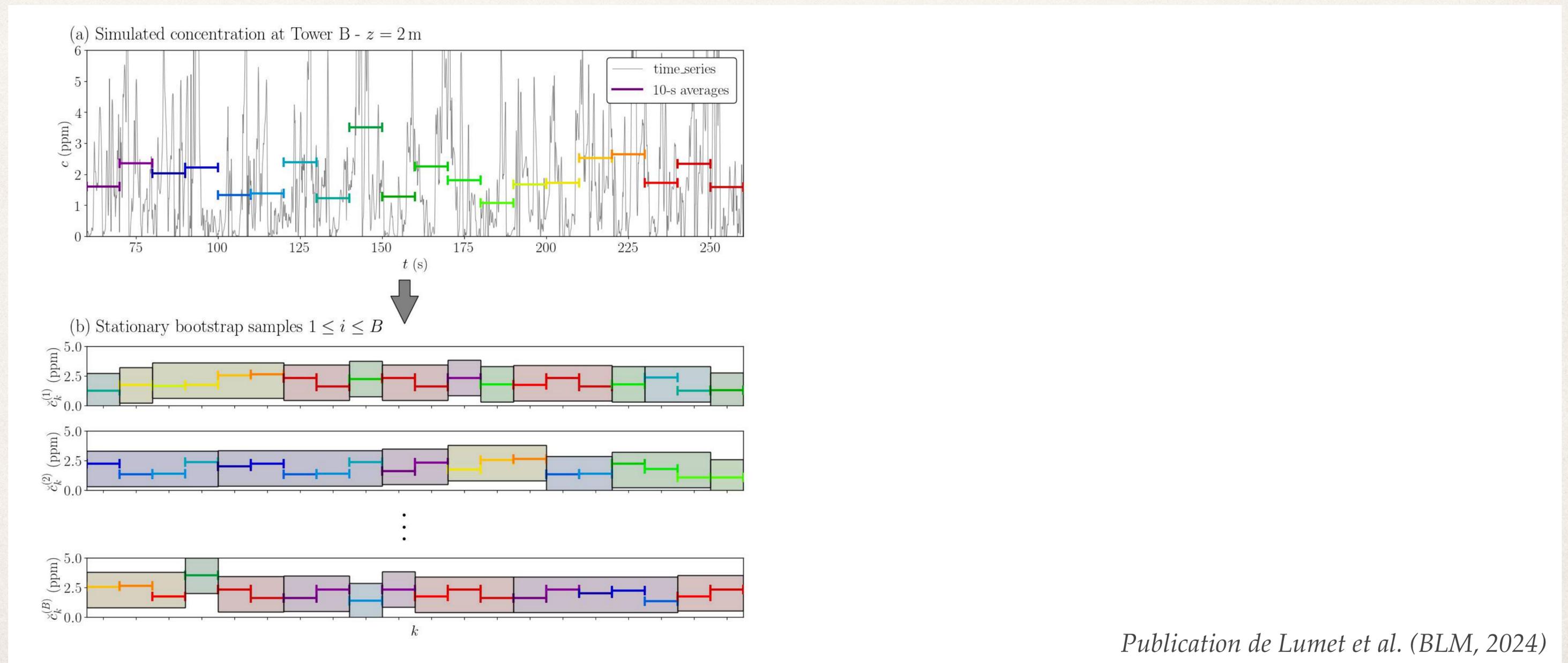
- La variabilité interne affecte significativement la concentration moyenne sur la fenêtre d'analyse
- Résultats LES caractérisés sous l'effet d'enveloppe plutôt qu'une unique simulation ➔ Comment ?

- Réalisation n°1
- Réalisation n°2
- Réalisation n°3

# I/ Estimation de la variabilité interne

## Approche par « stationary bootstrap » (*Politis and Romano, 1994*)

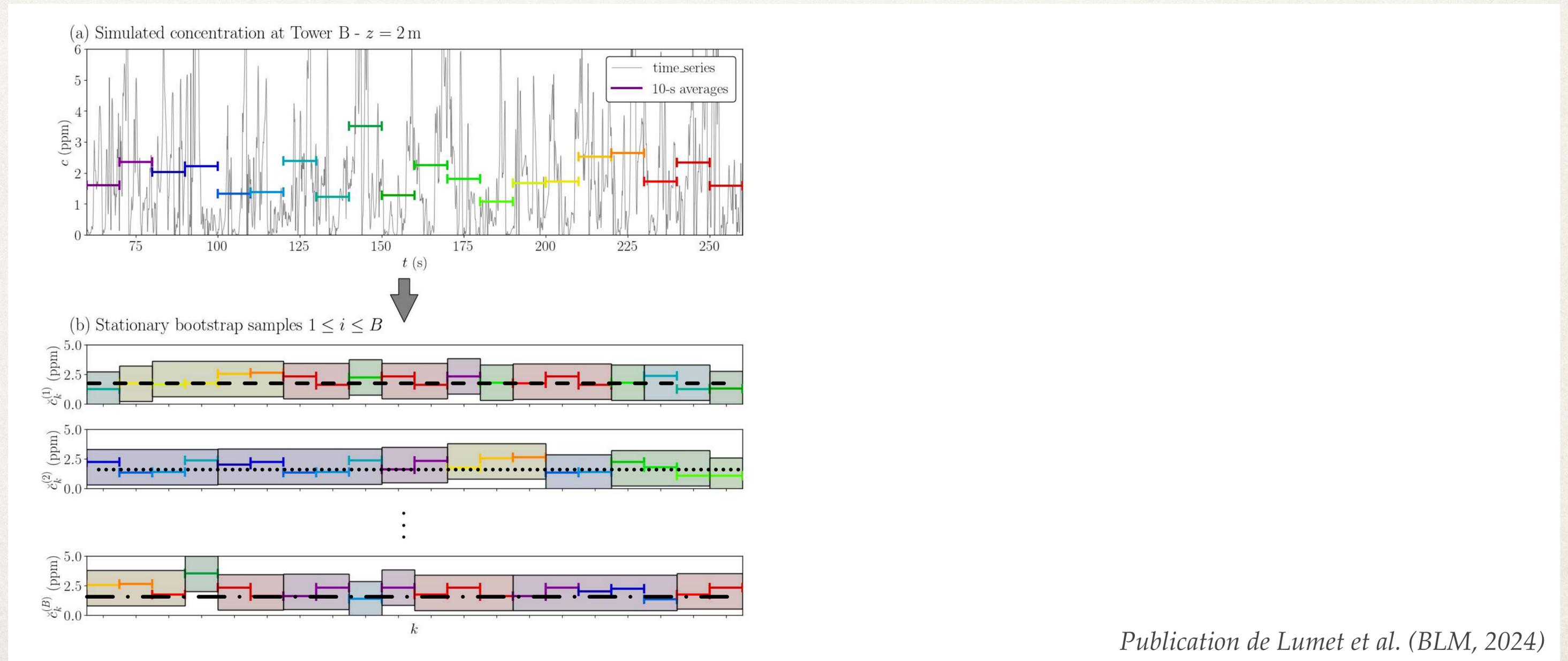
- ▶ Statistiques à partir de répliquions bootstrap sur la fenêtre simulée de 200 s



# I/ Estimation de la variabilité interne

## Approche par « stationary bootstrap » (*Politis and Romano, 1994*)

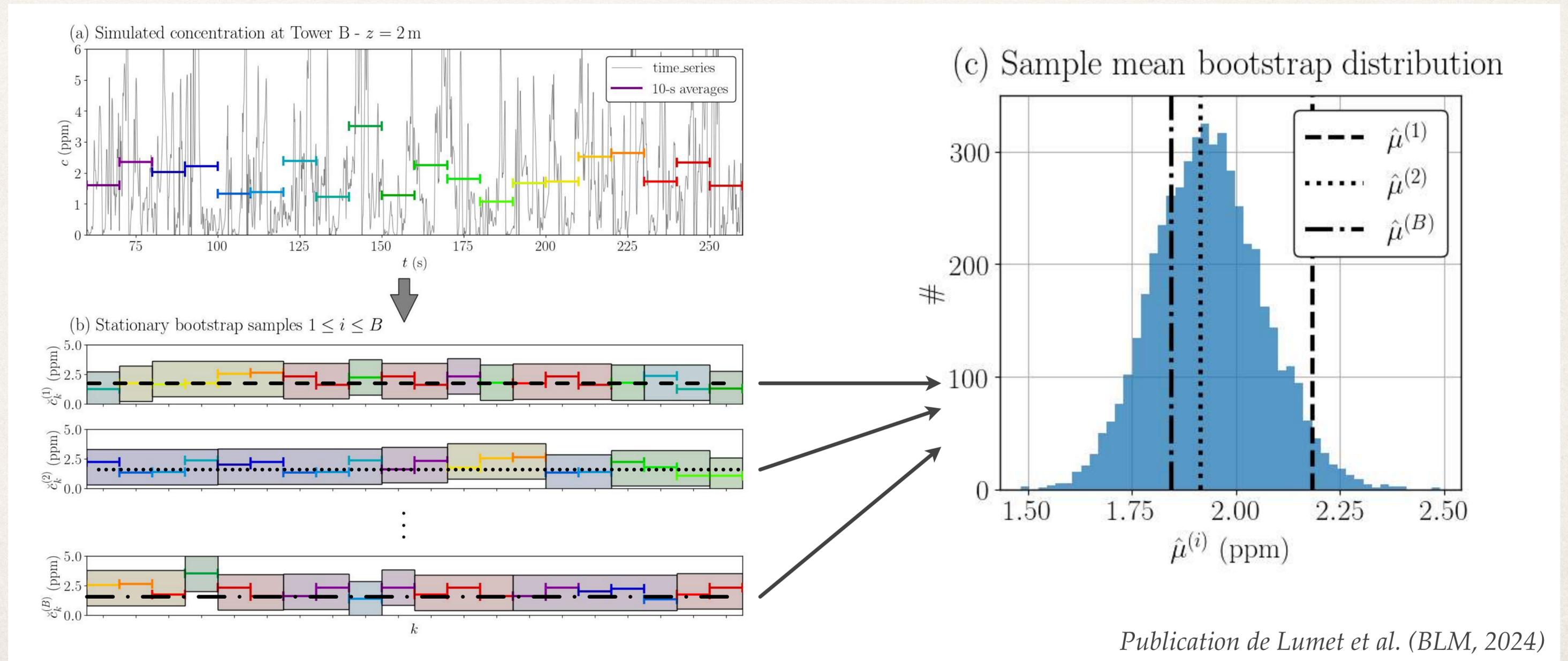
- ▶ Statistiques à partir de répliquions bootstrap sur la fenêtre simulée de 200 s



# I/ Estimation de la variabilité interne

## Approche par « stationary bootstrap » (*Politis and Romano, 1994*)

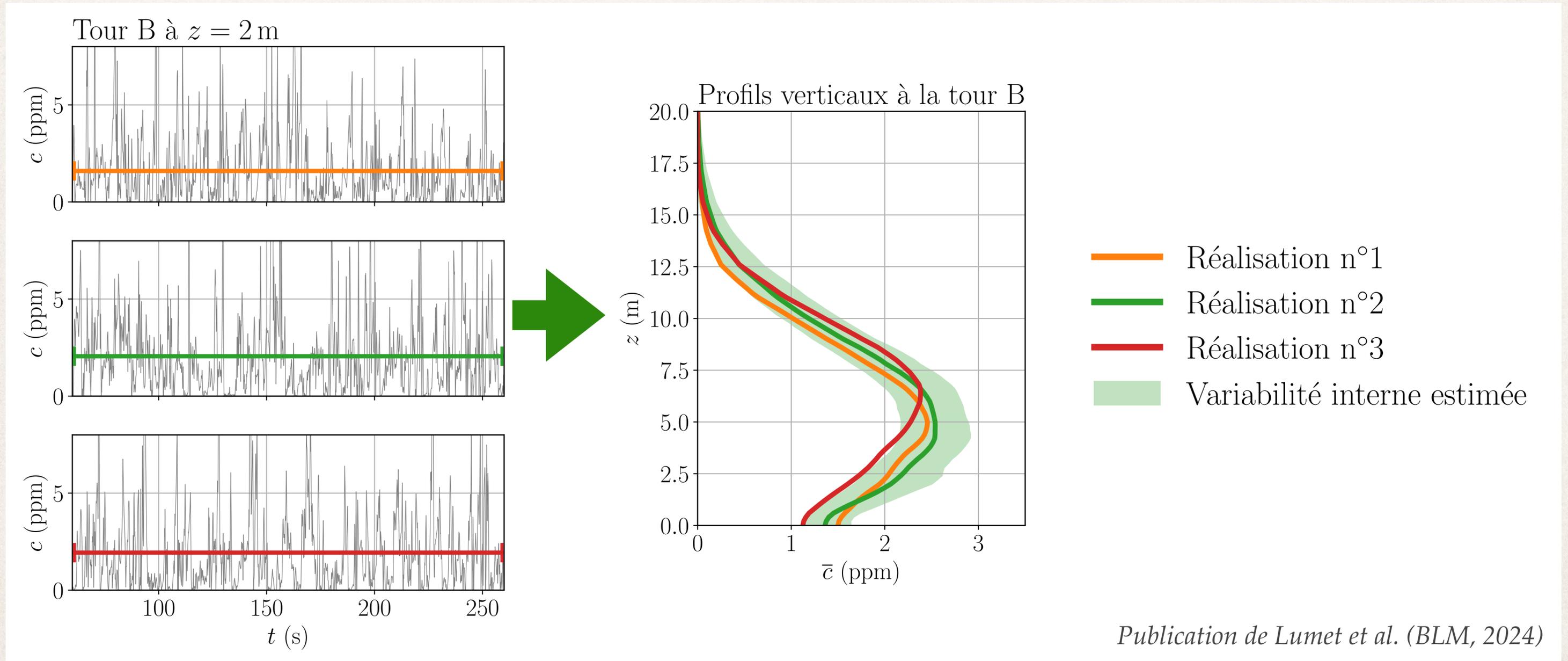
- ▶ Statistiques à partir de répliquions bootstrap sur la fenêtre simulée de 200 s



# I/ Estimation de la variabilité interne

## Approche par « stationary bootstrap » (*Politis and Romano, 1994*)

- ▶ L'approche bootstrap fournit un intervalle de confiance réaliste sans besoin d'allonger les simulations

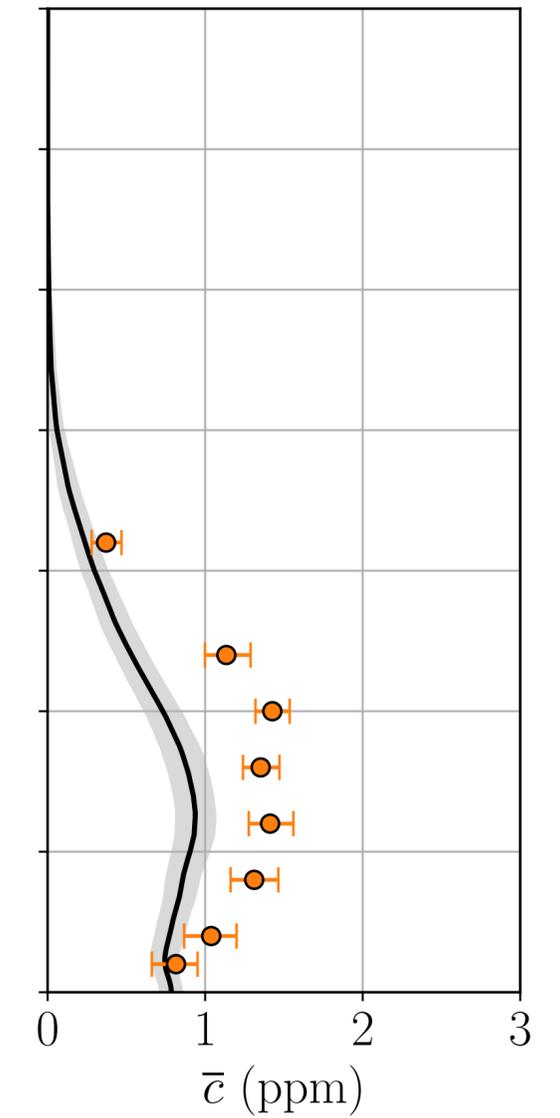
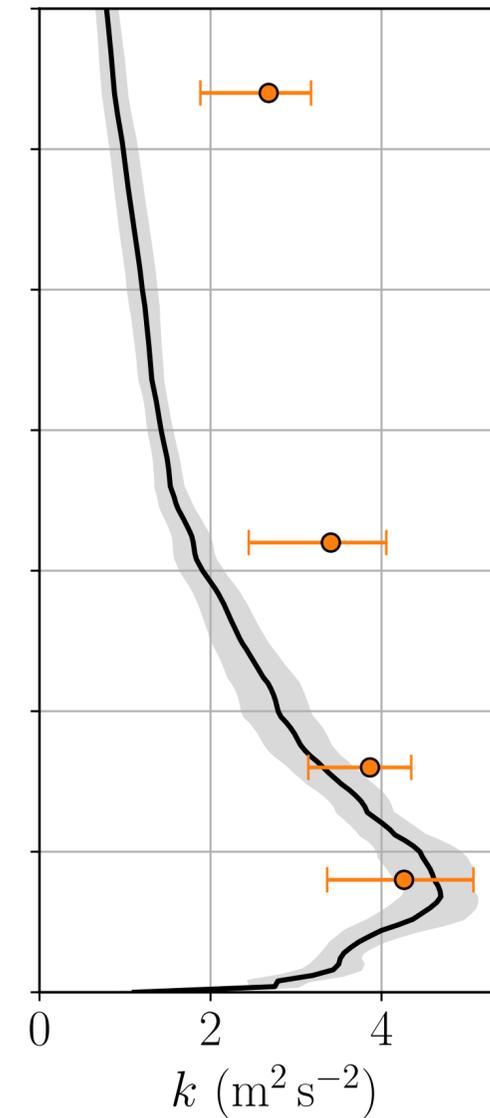
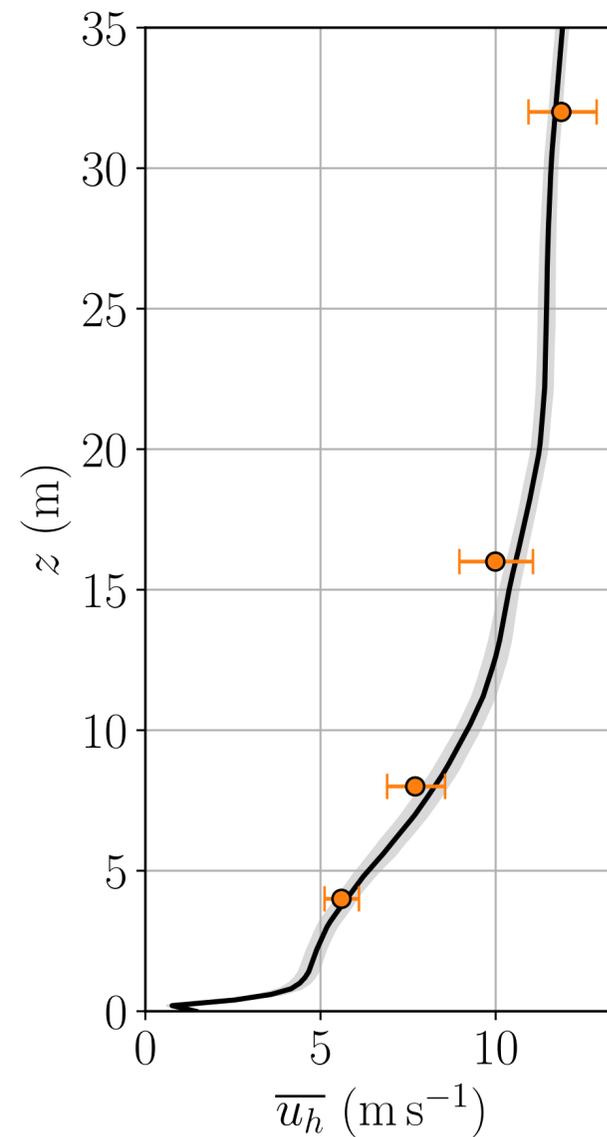


# I/ Estimation de la variabilité interne

## Approche par « stationary bootstrap » (*Politis and Romano, 1994*)

► Application à la validation du modèle LES, ici à la tour T au sein de la canopée urbaine

- Profil moyen estimé
- Variabilité interne simulée
- Donnée expérimentale
- ┃ Variabilité interne observée



# I/ Estimation de la variabilité interne

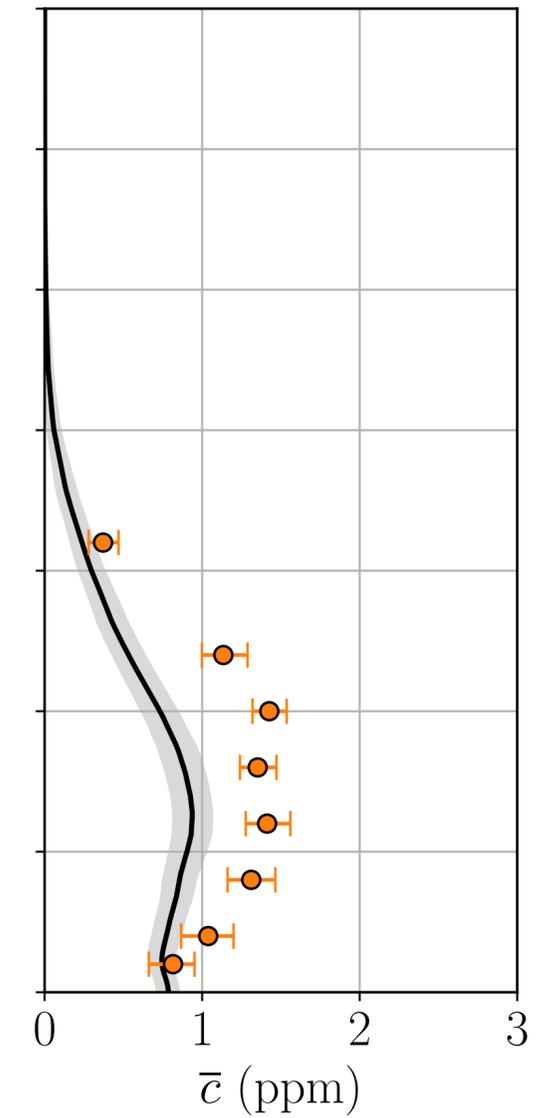
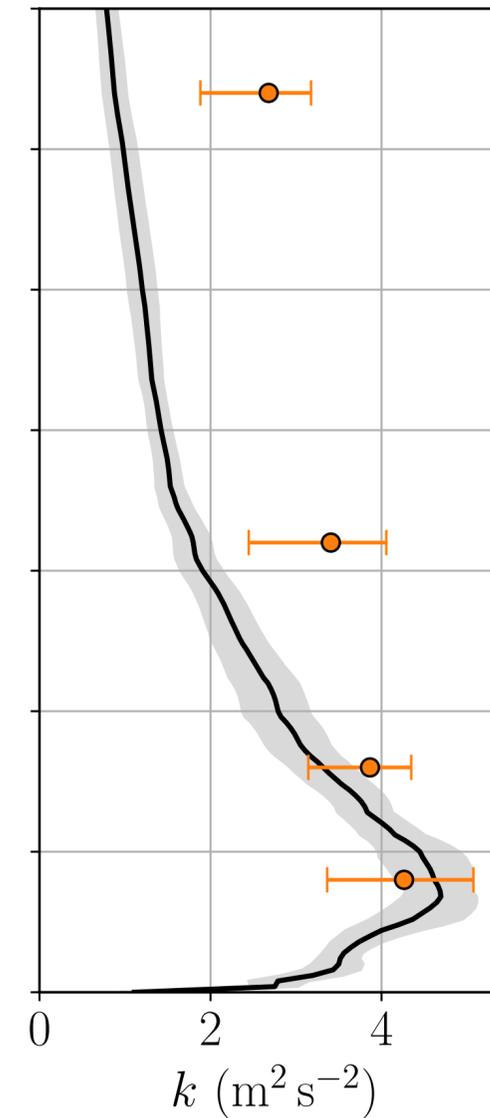
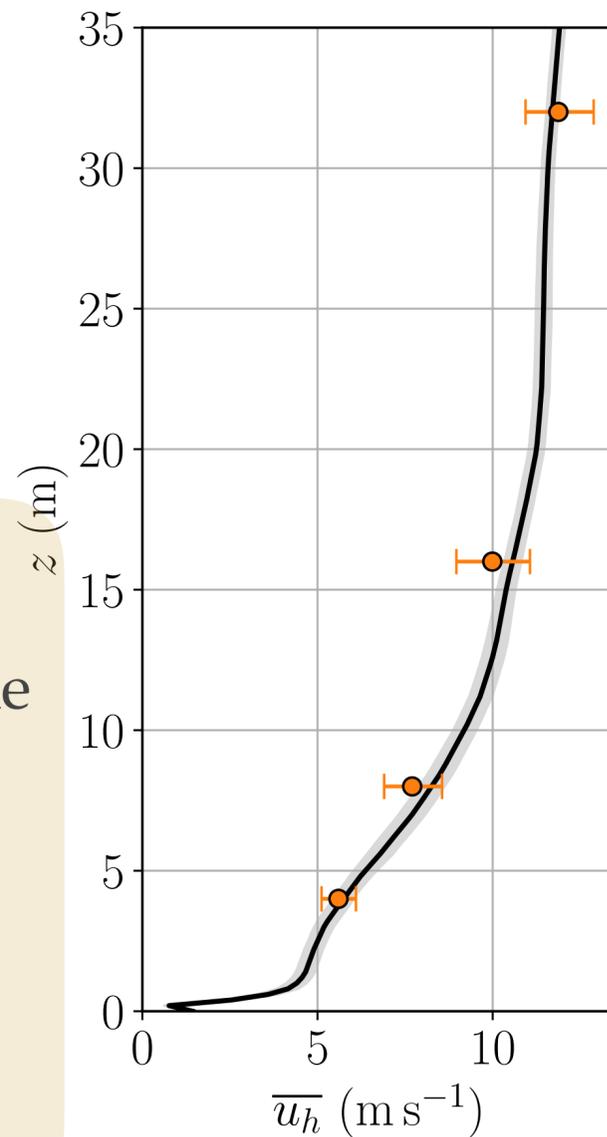
## Approche par « stationary bootstrap » (Politis and Romano, 1994)

► Application à la validation du modèle LES, ici à la tour T au sein de la canopée urbaine

- Profil moyen estimé
- Variabilité interne simulée
- Donnée expérimentale
- ┃ Variabilité interne observée

### Principales conclusions

- La LES permet de reproduire une grande partie de la variabilité interne observée.
- Mais elle n'explique pas tous les écarts entre modèle et observations ➔ Biais





# Prendre en compte les incertitudes dans les prévisions de dispersion atmosphérique de polluants en milieu urbain

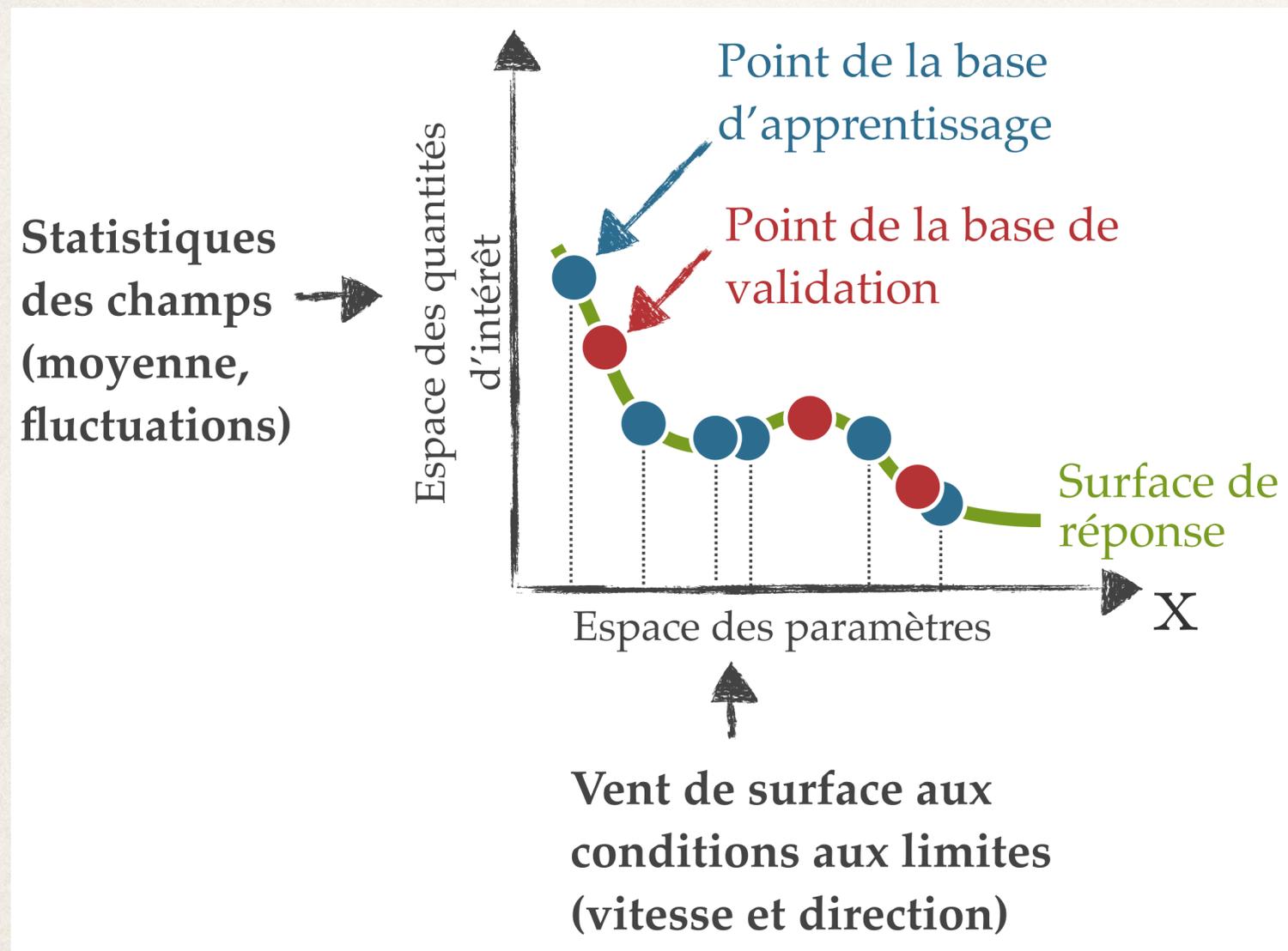
## II- L'incertitude liée au forçage atmosphérique



## II/ Quantification de l'incertitude liée au forçage atmosphérique

### Approche par émulation statistique

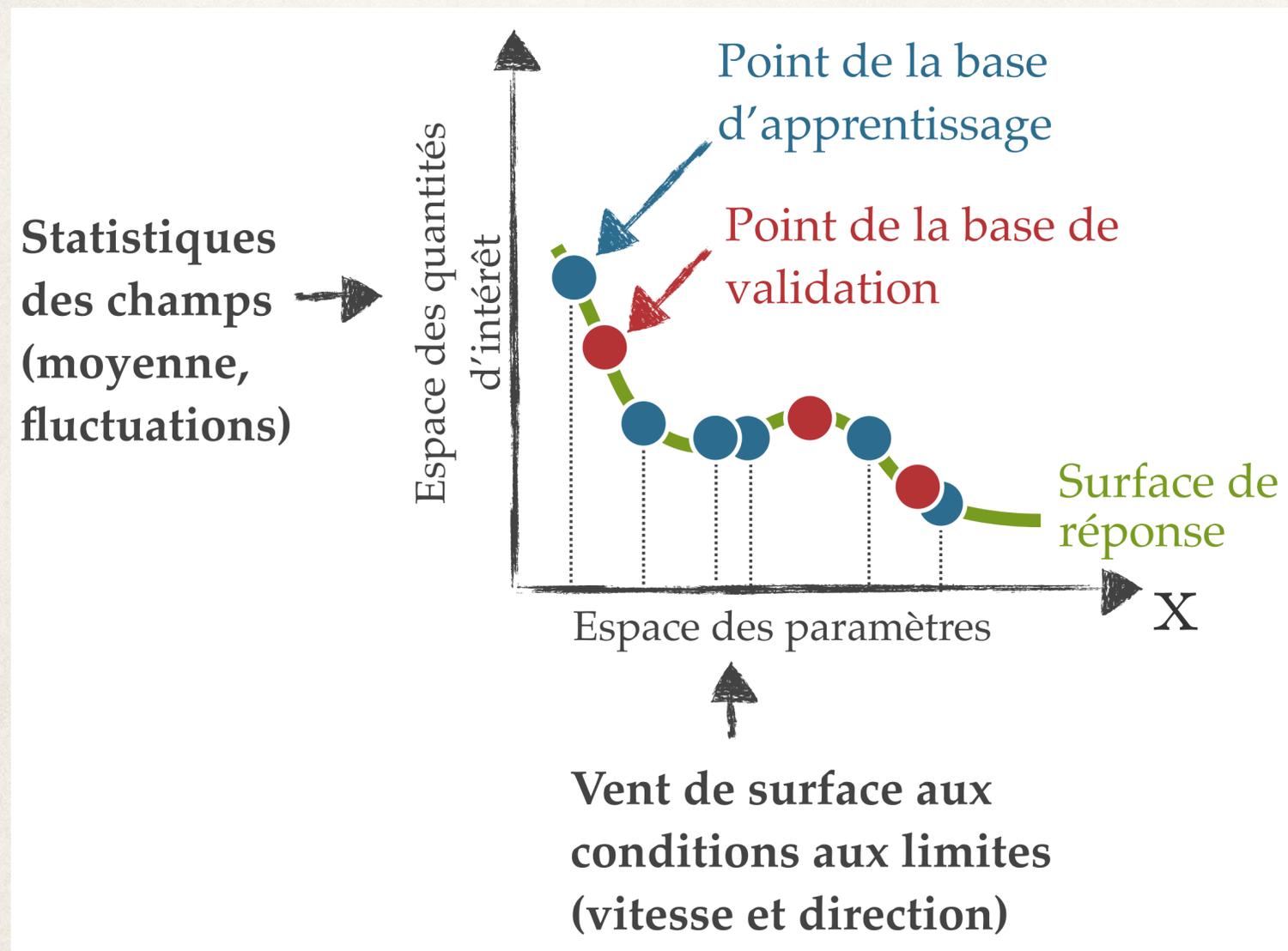
- ▶ Apprendre la surface de réponse du modèle LES par un modèle statistique



# II/ Quantification de l'incertitude liée au forçage atmosphérique

## Approche par émulation statistique

- ▶ Apprendre la surface de réponse du modèle LES par un modèle statistique



## Problématiques

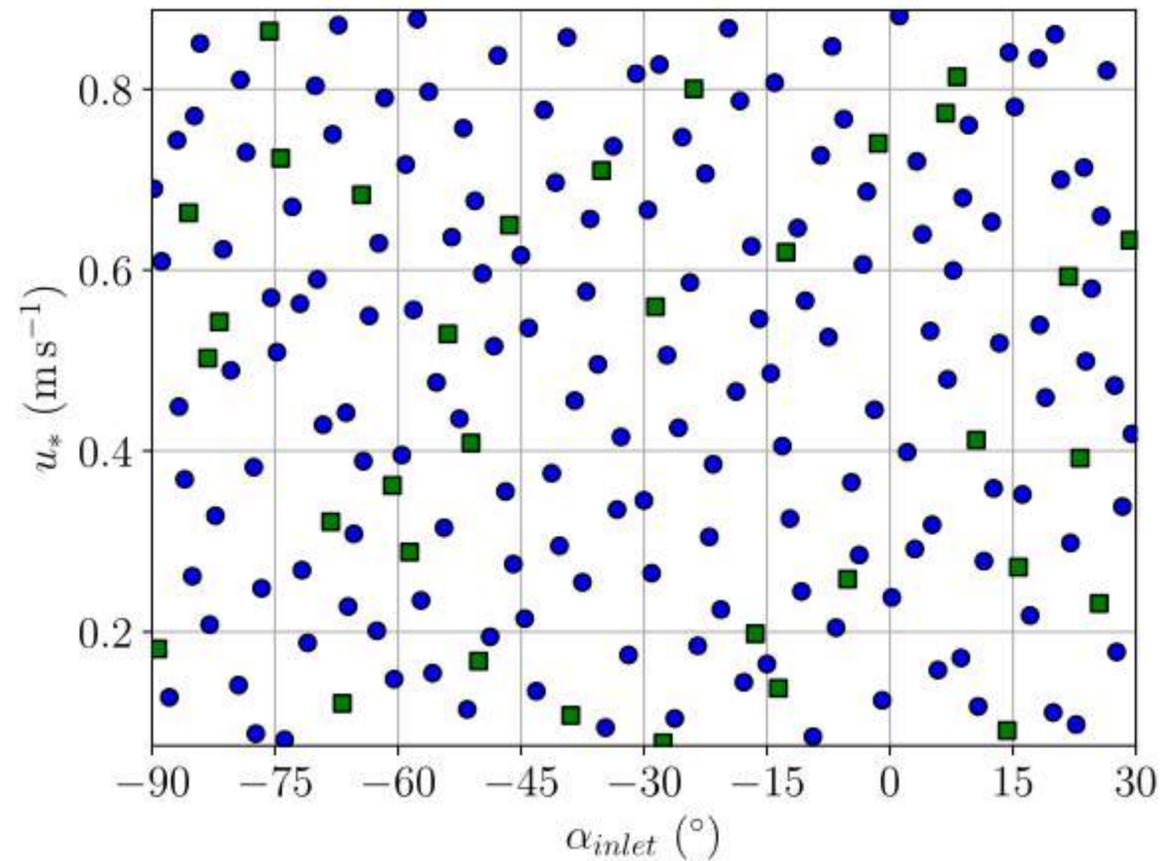
- ▶ Construire une **base de données d'apprentissage** de simulations LES pour représenter la variabilité des forçages atmosphériques
- ▶ Traiter la **grande dimension** des champs moyens de concentration pour apprendre la variabilité spatiale

# II/ Quantification de l'incertitude liée au forçage atmosphérique

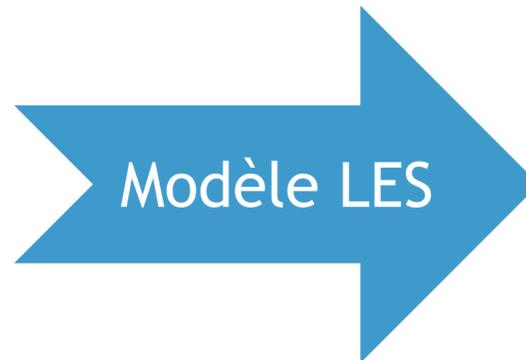
## Base de données LES

- ▶ Budget de 200 simulations LES (6 millions d'hCPU, Cerfacs/GENCI/Météo-France, environ 10 tCO<sub>2</sub>eq)

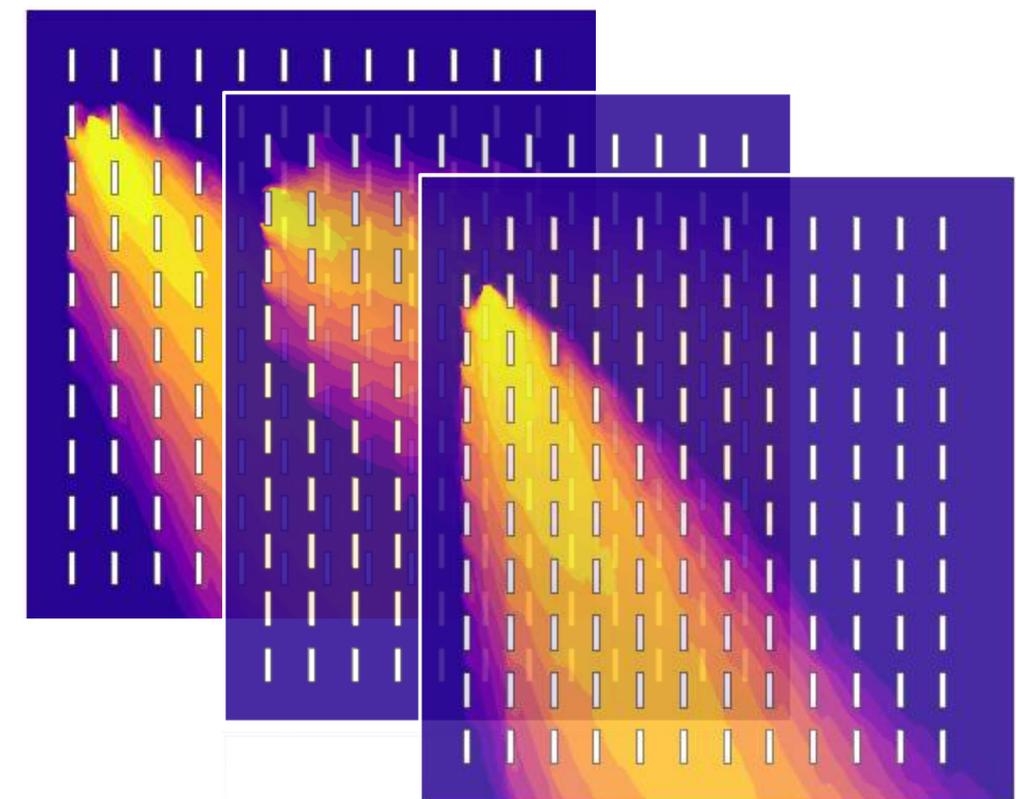
*Méthode de Quasi Monte Carlo (Halton)*



- Échantillon d'apprentissage
- Échantillon de validation



*Champs 3-D de concentration de traceur  
(pas de dimension temporelle)*

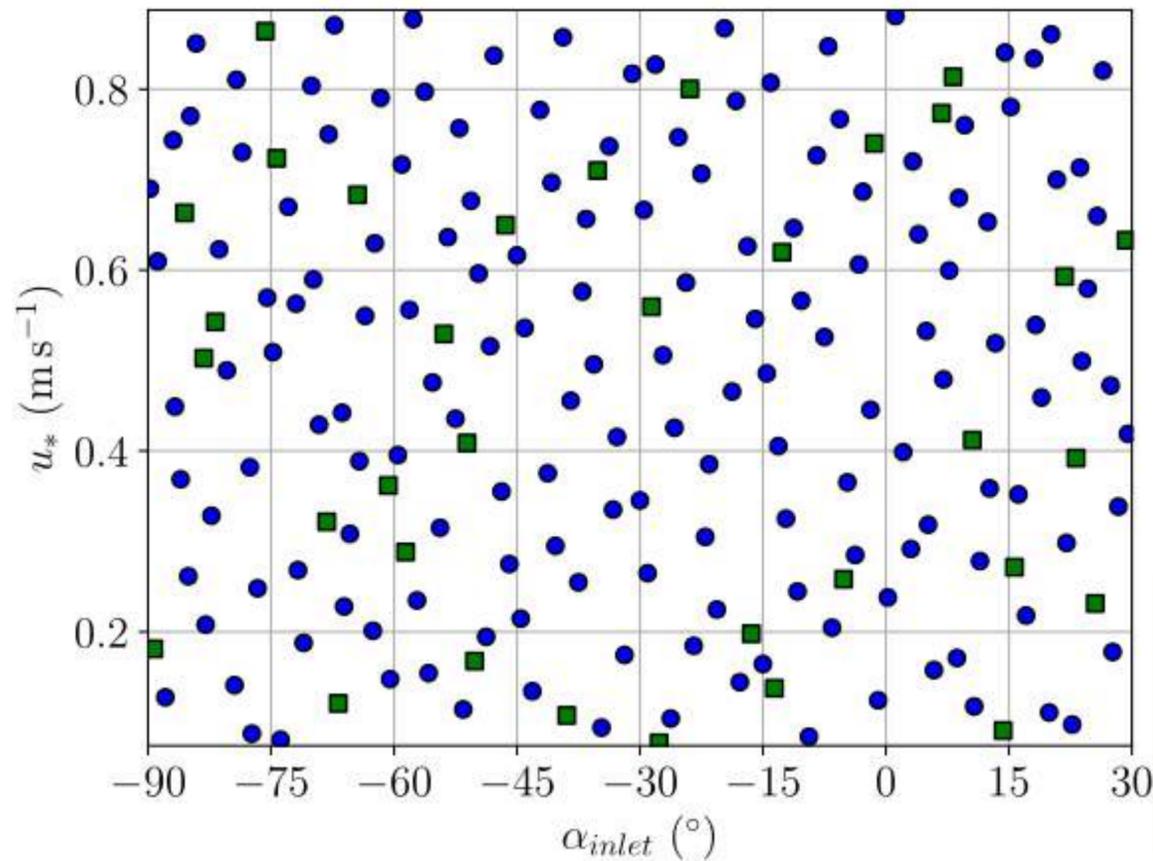


# II/ Quantification de l'incertitude liée au forçage atmosphérique

## Base de données LES

- ▶ Budget de 200 simulations LES (6 millions d'hCPU, Cerfacs/GENCI/Météo-France, environ 10 tCO<sub>2</sub>eq)

*Méthode de Quasi Monte Carlo (Halton)*

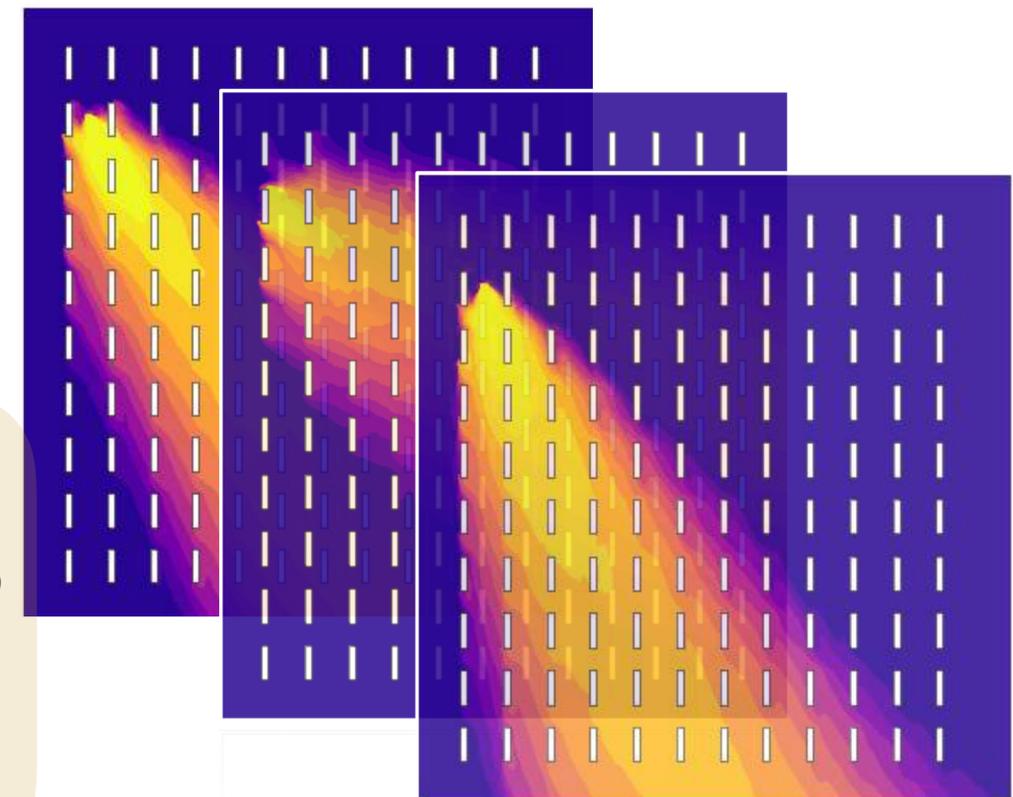


- Échantillon d'apprentissage
- Échantillon de validation

Modèle LES

*Bientôt : données disponibles sur Zenodo et réflexion en cours sur un data paper*

*Champs 3-D de concentration de traceur (pas de dimension temporelle)*



## II/ Quantification de l'incertitude liée au forçage atmosphérique

Processus d'émulation statistique POD-GPR en 2 étapes (Nony et al. SERRA 2023)

$$\bar{c} = \bar{c}(x, y, z, \boldsymbol{\theta}) = \sum_{l=1}^L k_l(\boldsymbol{\theta}) \psi_L(x, y, z)$$

## II/ Quantification de l'incertitude liée au forçage atmosphérique

Processus d'émulation statistique POD-GPR en 2 étapes (Nony et al. SERRA 2023)

$$\bar{c} = \bar{c}(x, y, z, \boldsymbol{\theta}) = \sum_{l=1}^L k_l(\boldsymbol{\theta}) \psi_L(x, y, z)$$

1. Réduction de dimension des champs par analyse en composantes principales (POD)  
➔ construction d'un espace latent de dimension réduite représentant les principaux modes de variabilité de la concentration moyenne

## II/ Quantification de l'incertitude liée au forçage atmosphérique

Processus d'émulation statistique POD-GPR en 2 étapes (Nony et al. SERRA 2023)

$$\bar{c} = \bar{c}(x, y, z, \boldsymbol{\theta}) = \sum_{l=1}^L k_l(\boldsymbol{\theta}) \psi_L(x, y, z)$$

**2. « Interpolation » de la réponse du modèle LES à la variation de la vitesse et de la direction du vent par processus gaussiens dans l'espace latent (GPR)**

➔ apprentissage de la dépendance aux paramètres incertains (Braconnier et al. 2011, Marrel et al. 2015)

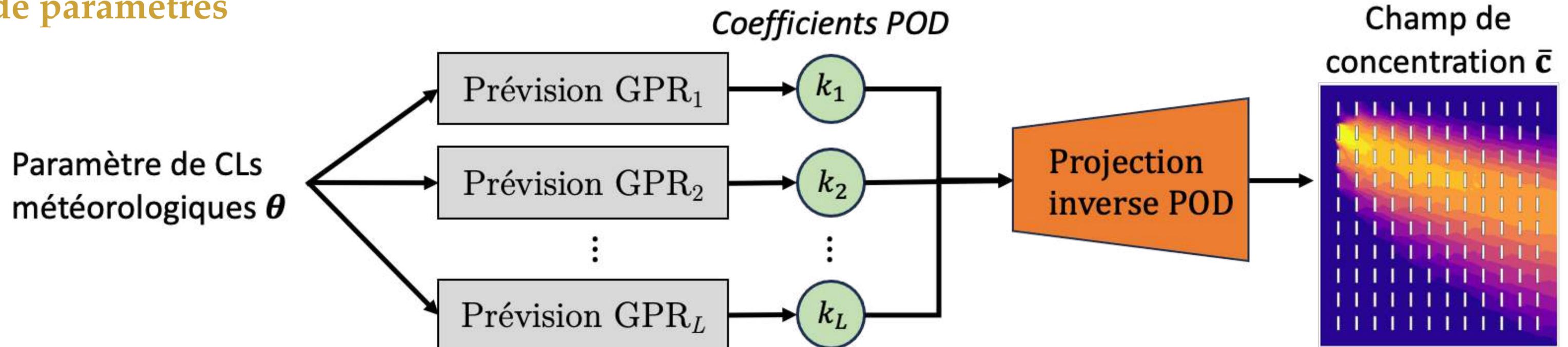
**1. Réduction de dimension des champs par analyse en composantes principales (POD)**  
➔ construction d'un espace latent de dimension réduite représentant les principaux modes de variabilité de la concentration moyenne

## II/ Quantification de l'incertitude liée au forçage atmosphérique

Processus d'émulation statistique POD-GPR en 2 étapes (Nony et al. SERRA 2023)

$$\bar{c} = \bar{c}(x, y, z, \boldsymbol{\theta}) = \sum_{l=1}^L k_l(\boldsymbol{\theta}) \psi_L(x, y, z)$$

Nouvelle prédiction  
pour un nouveau jeu  
de paramètres



# II/ Quantification de l'incertitude liée au forçage atmosphérique

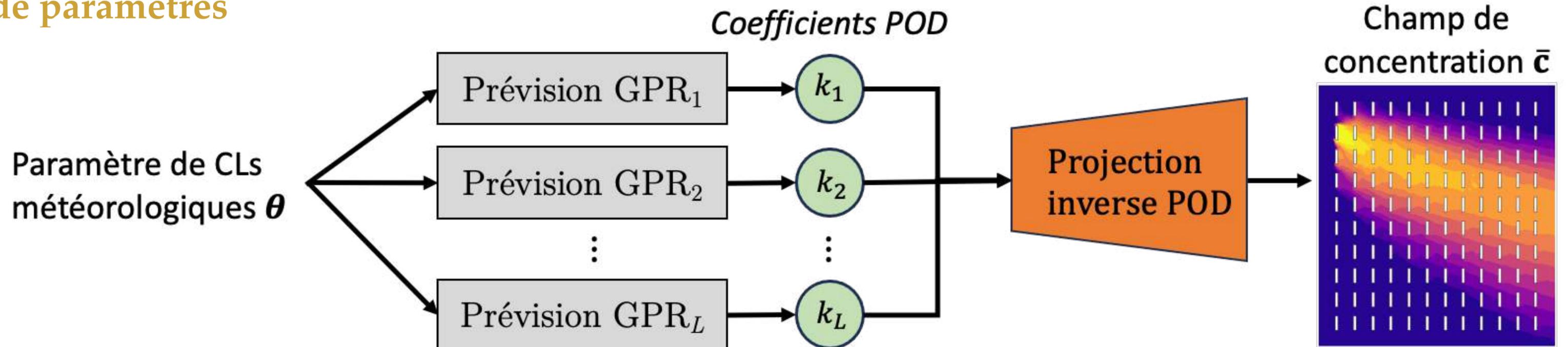
Processus d'émulation statistique POD-GPR en 2 étapes (Nony et al. SERRA 2023)

$$\bar{c} = \bar{c}(x, y, z, \boldsymbol{\theta}) = \sum_{l=1}^L k_l(\boldsymbol{\theta}) \psi_L(x, y, z)$$

Nouvelle prédiction  
pour un nouveau jeu  
de paramètres

## Enjeux

- ▶ Spécificités de la variable concentration
- ▶ Choix de la dimension de l'espace latent
- ▶ Prise en compte de la variabilité interne



## II/ Quantification de l'incertitude liée au forçage atmosphérique

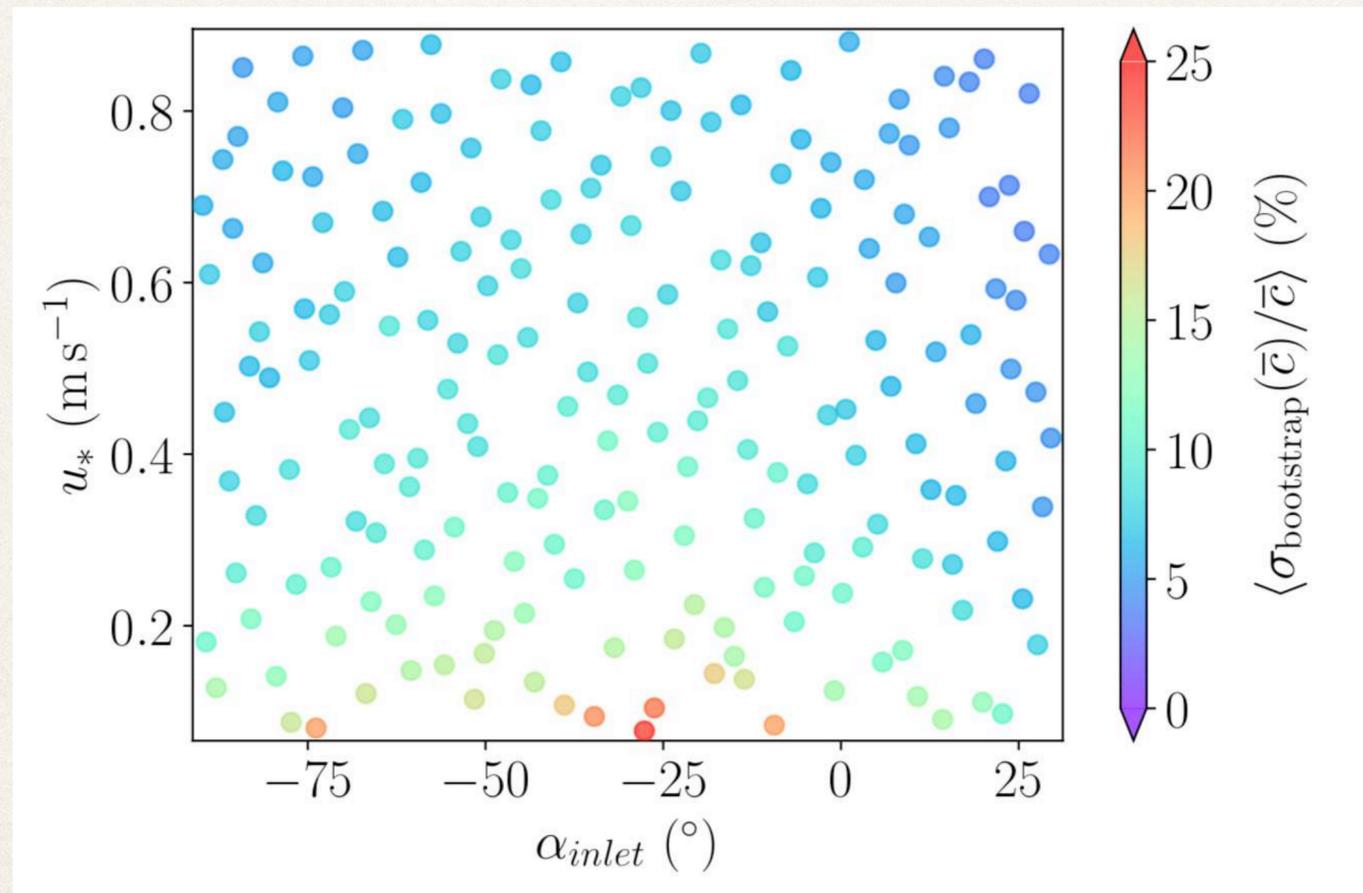
### Dimension de l'espace latent liée à la variabilité interne

- ▶ La base d'apprentissage est bruitée du fait de la variabilité interne : on peut estimer la variabilité interne dans l'espace latent.

# II/ Quantification de l'incertitude liée au forçage atmosphérique

## Dimension de l'espace latent liée à la variabilité interne

- ▶ La base d'apprentissage est bruitée du fait de la variabilité interne : on peut estimer la variabilité interne dans l'espace latent.

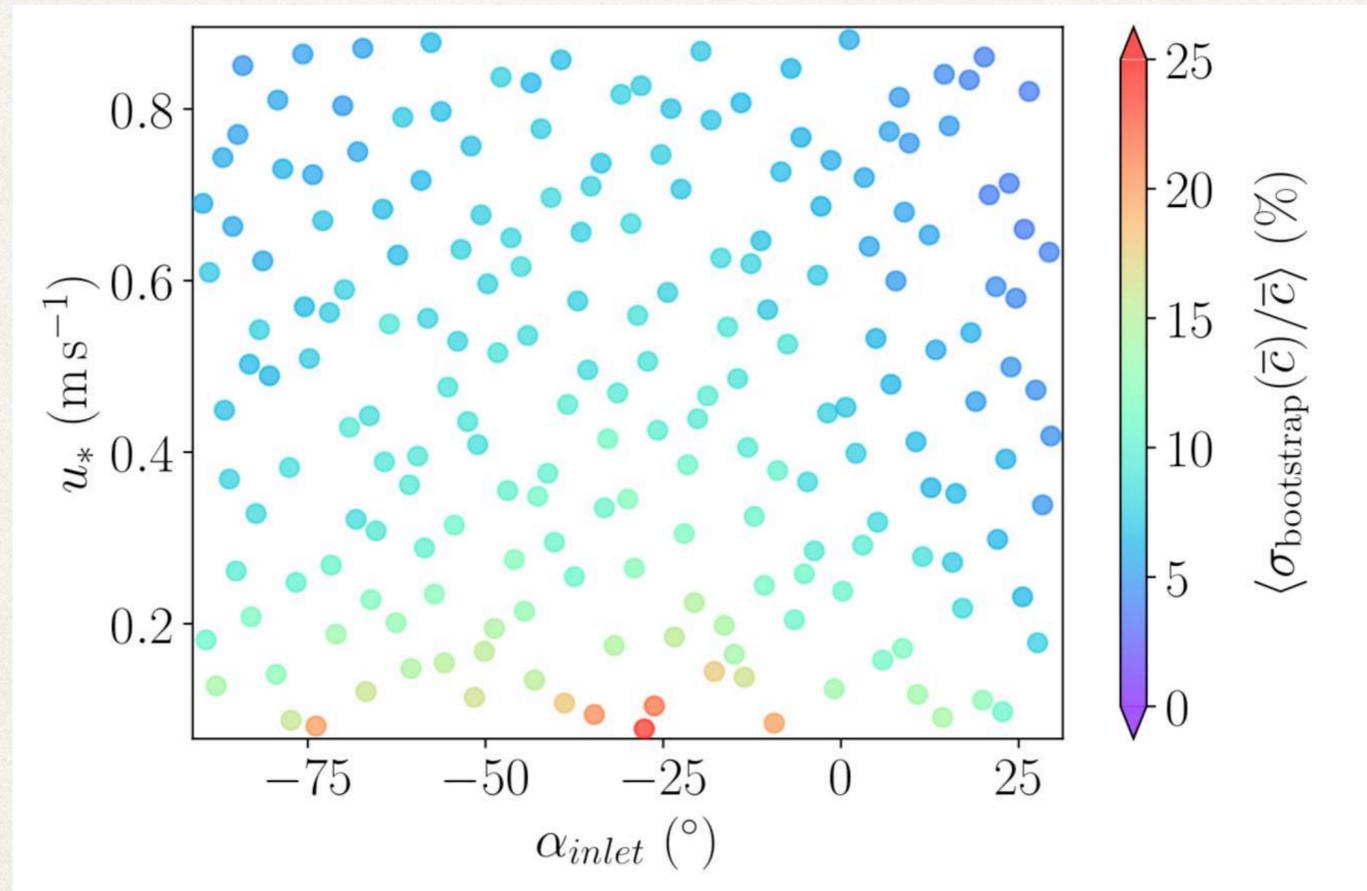


*Dépendance de la variabilité interne aux paramètres incertains : chaque point représente l'incertitude moyenne (en espace) d'un échantillon LES de la base d'apprentissage.*

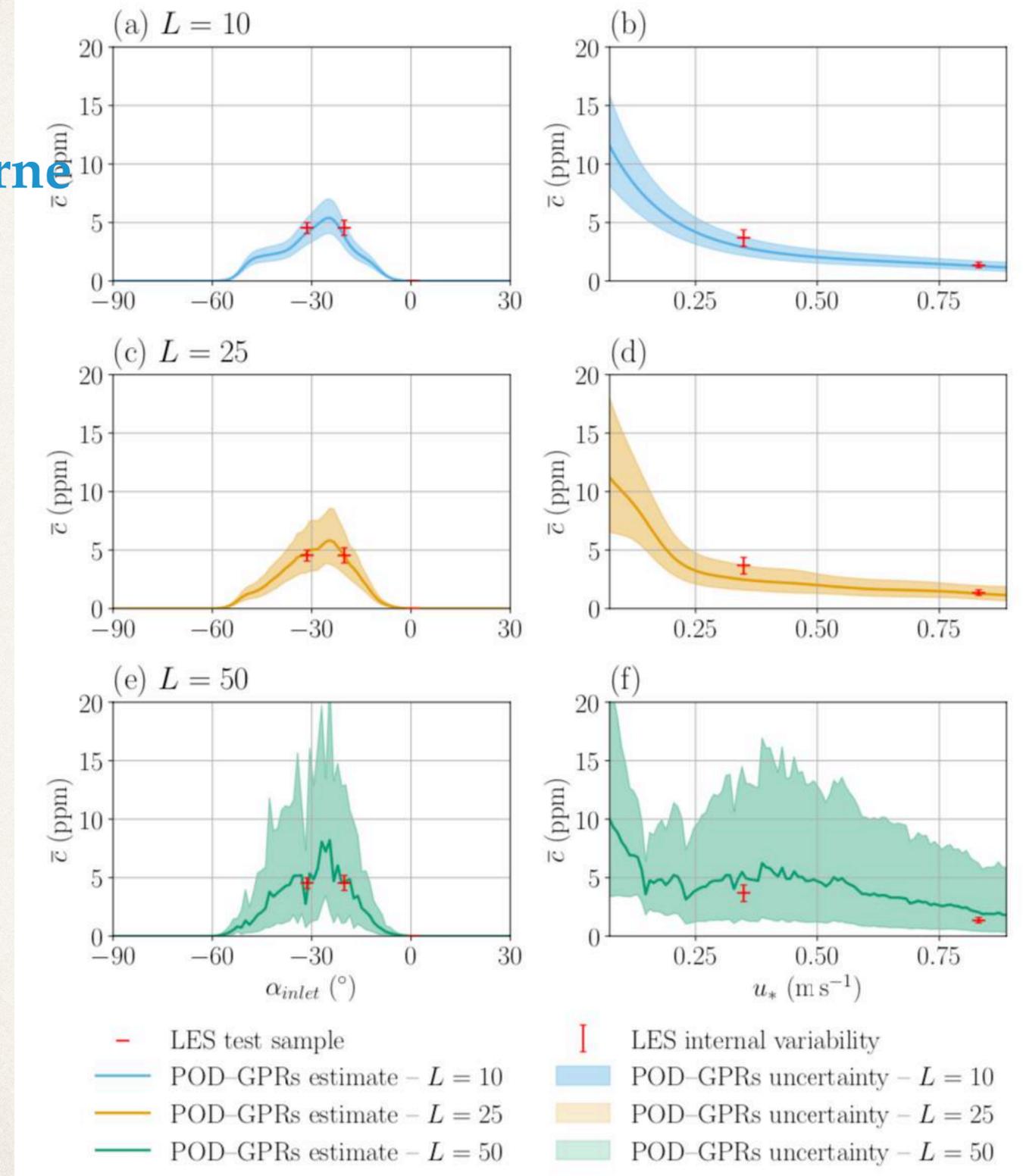
# II/ Quantification de l'incertitude liée au forçage atmosphérique

## Dimension de l'espace latent liée à la variabilité interne

- La base d'apprentissage est bruitée du fait de la variabilité interne : on peut estimer la variabilité interne dans l'espace latent.



Dépendance de la variabilité interne aux paramètres incertains : chaque point représente l'incertitude moyenne (en espace) d'un échantillon LES de la base d'apprentissage.

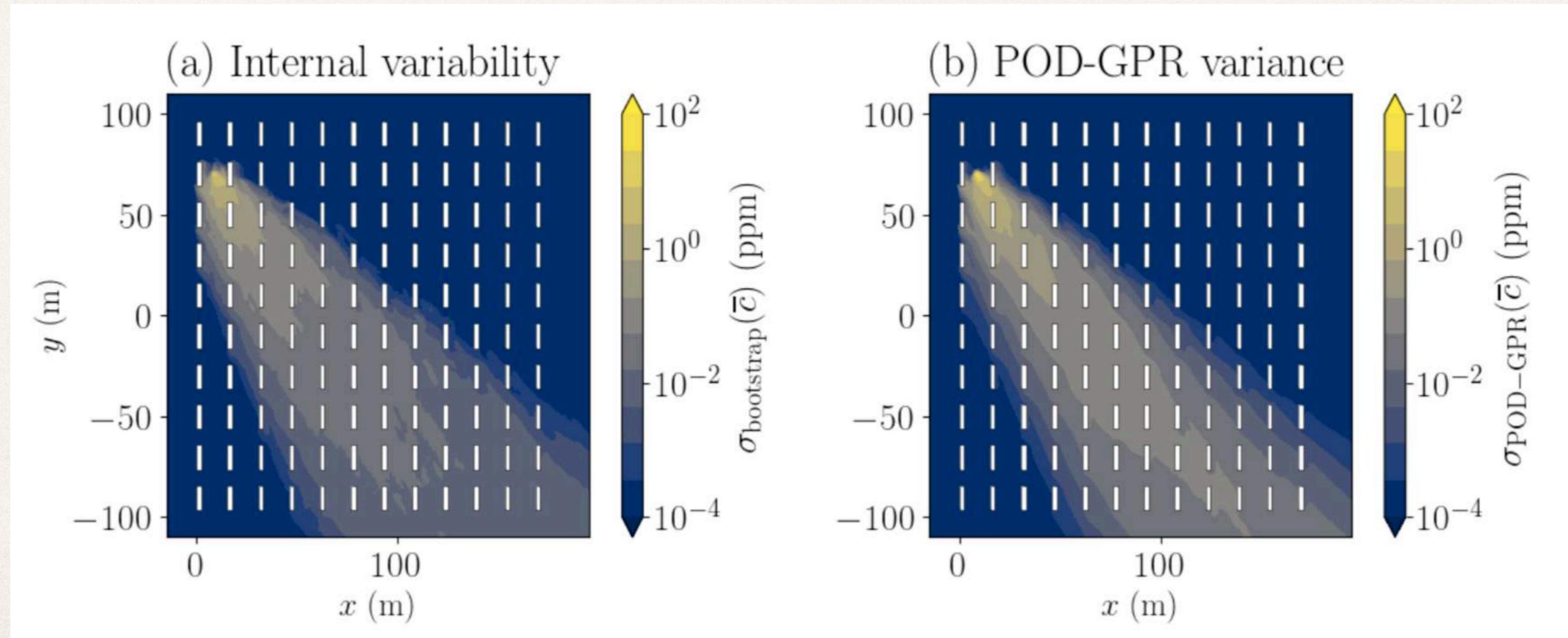


Prédiction de la concentration moyenne à la tour B à  $z = 2$  m 19

## II/ Quantification de l'incertitude liée au forçage atmosphérique

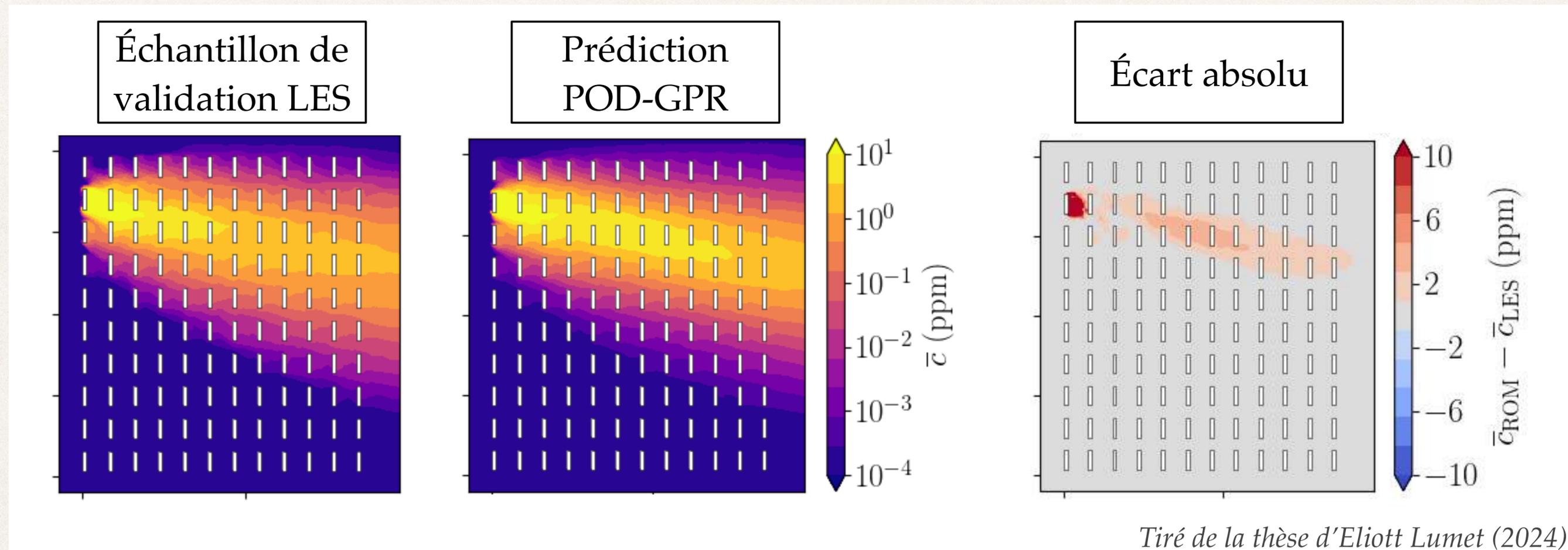
### La variabilité interne vue par le modèle statistique

- ▶ L'approche POD-GPR est capable de bien estimer la variabilité interne (amplitude et structure)



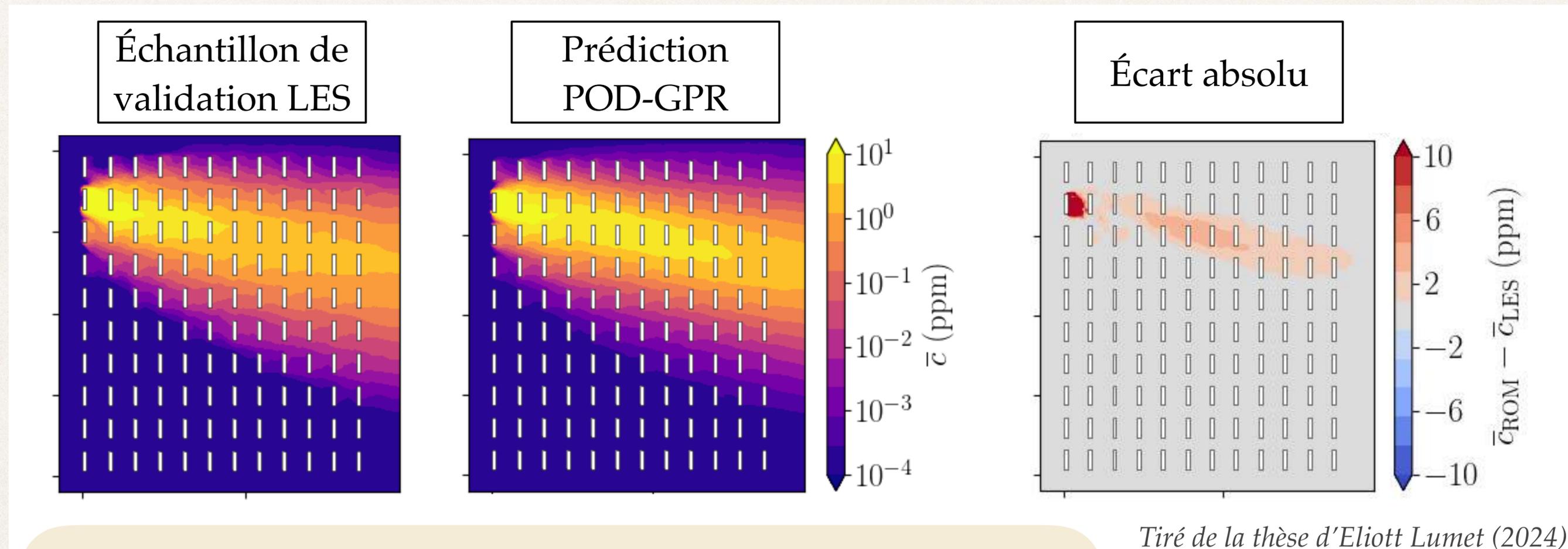
## II/ Quantification de l'incertitude liée au forçage atmosphérique

Exemple d'émulation du champ moyen de concentration (espace latent de dimension  $L = 10$ )



## II/ Quantification de l'incertitude liée au forçage atmosphérique

Exemple d'émulation du champ moyen de concentration (espace latent de dimension  $L = 10$ )

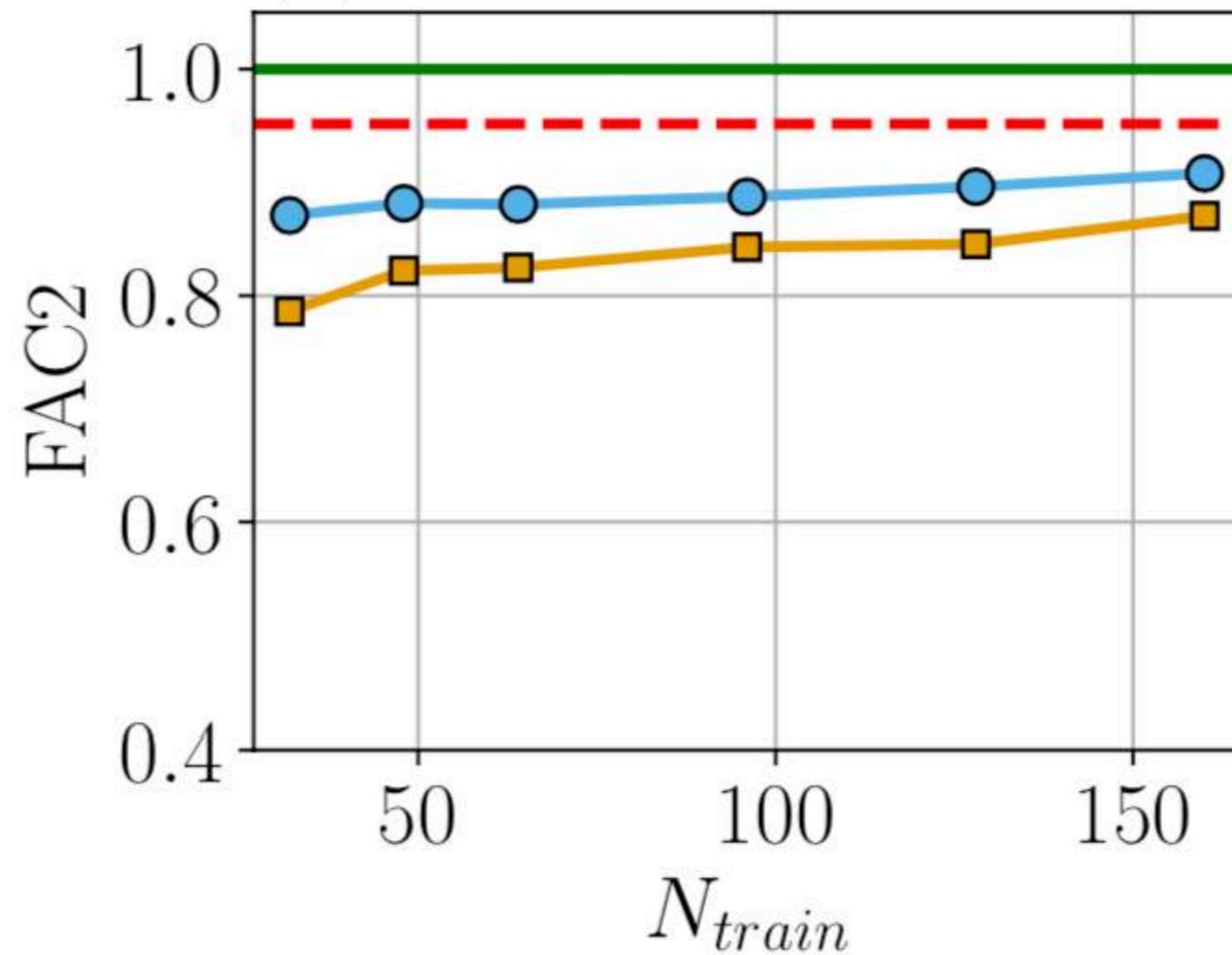


### Principales conclusions

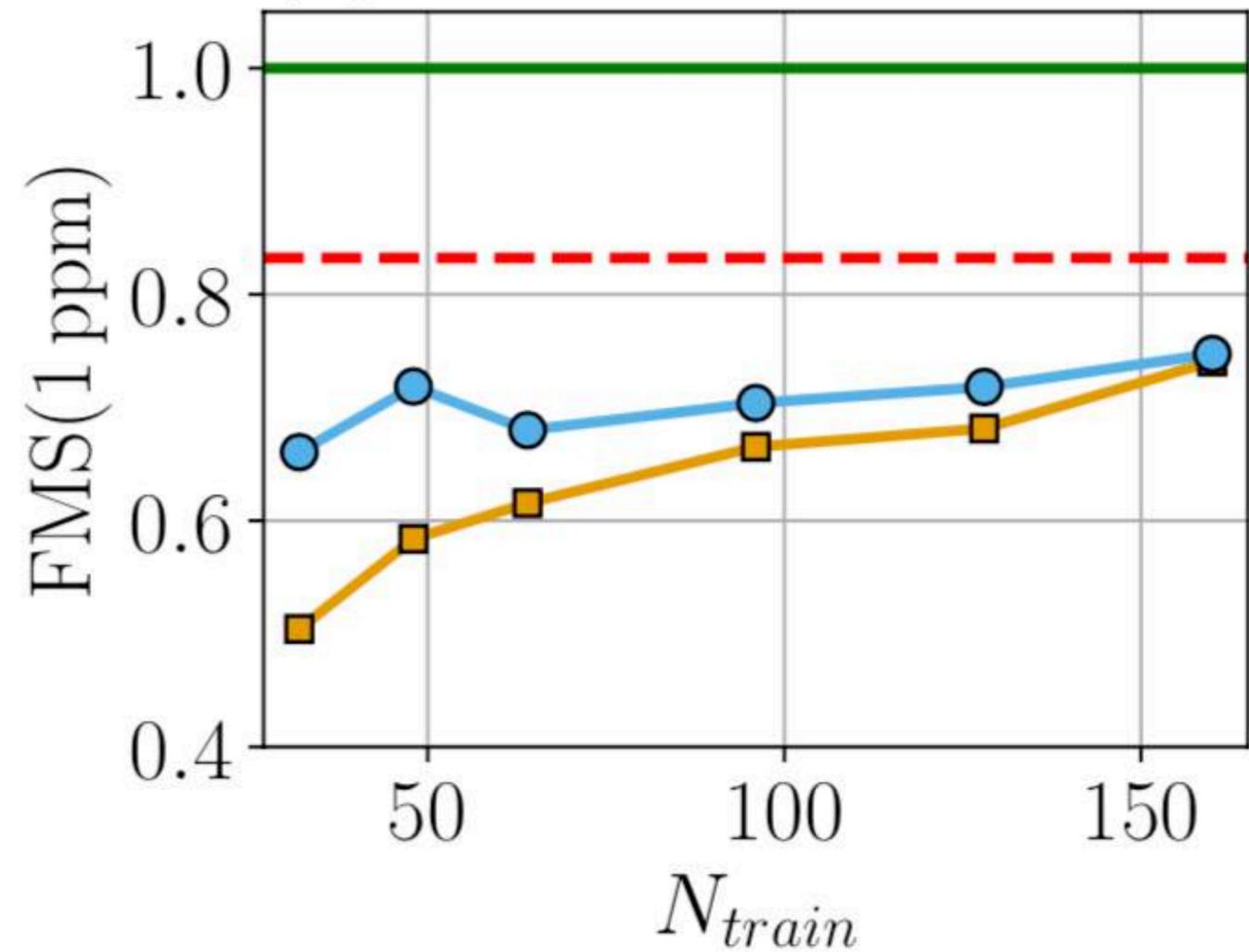
- ▶ Bonne représentation de la structure du panache
- ▶ Difficulté à bien reconstruire la concentration au très proche voisinage de la source (potentiel du mélange d'experts)

## II/ Quantification de l'incertitude liée au forçage atmosphérique

Approche d'émulation robuste à la taille de la base d'apprentissage



— Perfect score  
- - - Mean internal variability error



—●— POD-GPR  
—■— Nearest Neighbor

# Mots de conclusion

- ❖ **Construction et validation d'un modèle détaillé de dispersion en milieu urbain (LES)**
- ❖ **Quantification de la variabilité interne à micro-échelle par approche bootstrap** (*Lumet et al., Boundary-Layer meteorology, 2024*)
- ❖ **Génération d'une grande base de données de simulations pour étudier la dépendance au forçage atmosphérique** (*Dépôt Zenodo et data paper en cours de préparation*)
- ❖ **Approche d'émulation statistique basée sur les processus gaussiens prometteuse** (*Lumet et al., en cours de préparation*)

# Mots de conclusion

- ❖ **Construction et validation d'un modèle détaillé de dispersion en milieu urbain (LES)**
  - ➔ Aller vers des cas d'étude plus complexes (conditions non-neutres) et réalistes (villes réelles)
- ❖ **Quantification de la variabilité interne à micro-échelle par approche bootstrap (*Lumet et al., Boundary-Layer meteorology, 2024*)**
  - ➔ Quid de la variabilité interne de méso-échelle (Nagel et al. BLM 2022) ?
- ❖ **Génération d'une grande base de données de simulations pour étudier la dépendance au forçage atmosphérique (*Dépôt Zenodo et data paper en cours de préparation*)**
  - ➔ Quid de plans d'expériences adaptatif pour la LES ? Quid des approches multi-fidélité ?
- ❖ **Approche d'émulation statistique basée sur les processus gaussiens prometteuse (*Lumet et al., en cours de préparation*)**
  - ➔ Quid des réseaux de neurones ? Et de l'assimilation de données ? (travaux en cours)



**Merci pour votre attention !**

**Contact : [melanie.rochoux@cerfacs.fr](mailto:melanie.rochoux@cerfacs.fr)**

