



HAL
open science

Surrogate-based data assimilation for microscale atmospheric pollutant dispersion

Eliott Lumet, Mélanie C. Rochoux, Thomas Jaravel, Simon Lacroix

► **To cite this version:**

Eliott Lumet, Mélanie C. Rochoux, Thomas Jaravel, Simon Lacroix. Surrogate-based data assimilation for microscale atmospheric pollutant dispersion. EGU General Meeting 2024, Apr 2024, Vienne, Austria. 2024. hal-04748723

HAL Id: hal-04748723

<https://cnrs.hal.science/hal-04748723v1>

Submitted on 22 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Surrogate-based data assimilation for microscale atmospheric pollutant dispersion

Elliott Lumet^{1,2}, Mélanie Rochoux¹, Thomas Jaravel¹ et Simon Lacroix²

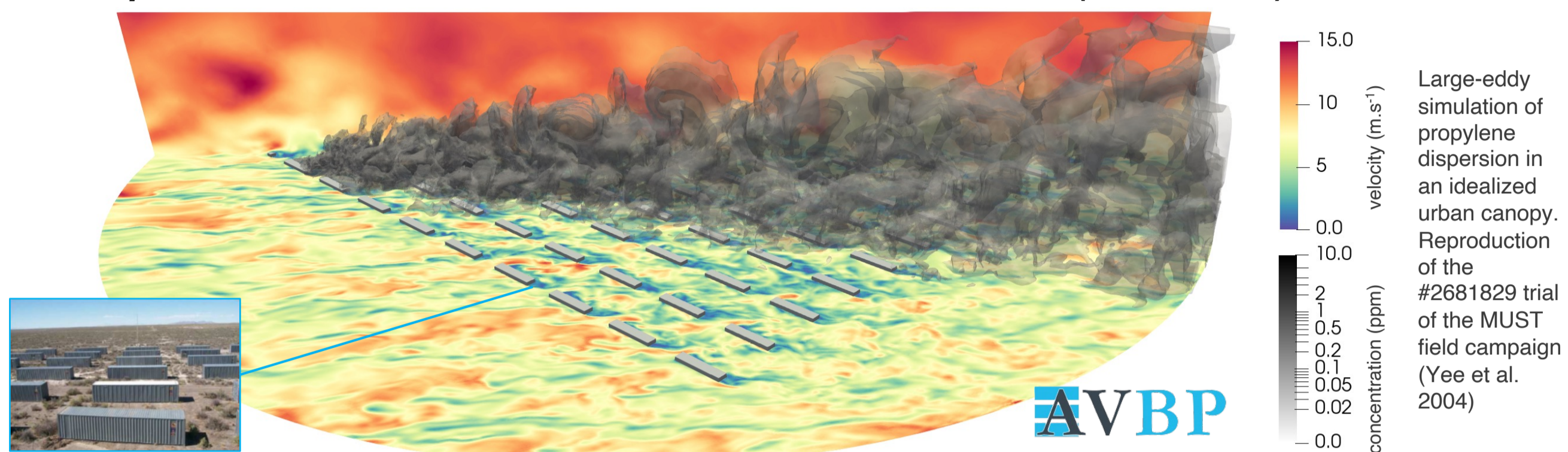
[1] CECI, CNRS, CERFACS, Toulouse, France

[2] LAAS, CNRS, Toulouse, France

Contact: elliott.lumet@cerfacs.fr

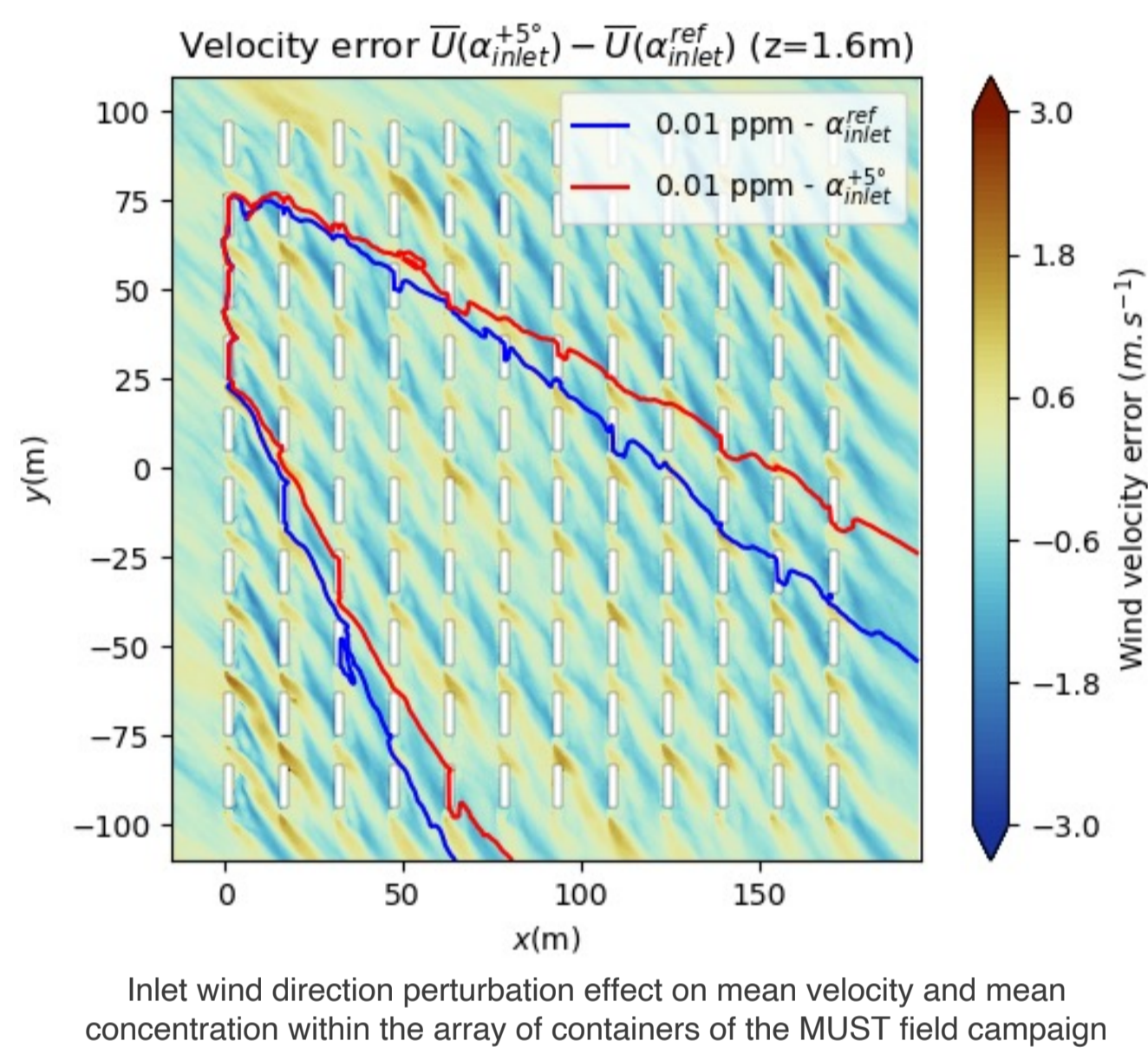
1) LES for atmospheric dispersion

Design, evaluate and improve high fidelity models for pollutant dispersion in urban areas at microscale ($\approx 100m$)



A. Why do we use Large-Eddy Simulation (LES)?

- i. Explicitly takes into account the effect of the urban canopy on atmospheric flow
- ii. Allows to track temporal evolution of the quantities of interest
- iii. Resolves the largest turbulence scales
 - ⊕ Reduced modeling uncertainty
 - ⊕ Better representation the effect of atmospheric variability on the plume



B. Which model limitations?

- ⊖ **Cost:** One 200-s simulation \Leftrightarrow 20 000h_{CPU}
- ⊖ **High uncertainties** (Dauvois et al. 2021):
 - Epistemic uncertainties:** Boundary conditions (Large-scale meteorological forcing, Pollutant source, Urban canopy representation), Structural (Subgrid scale modeling, Numerical scheme)
 - Aleatory uncertainties:** Microscale atmospheric internal variability

LES remains very uncertain despite its substantial computational cost

2) Surrogate modeling

Emulate the response surface of the LES model at reduced computing cost

A. Learning database of LES simulations

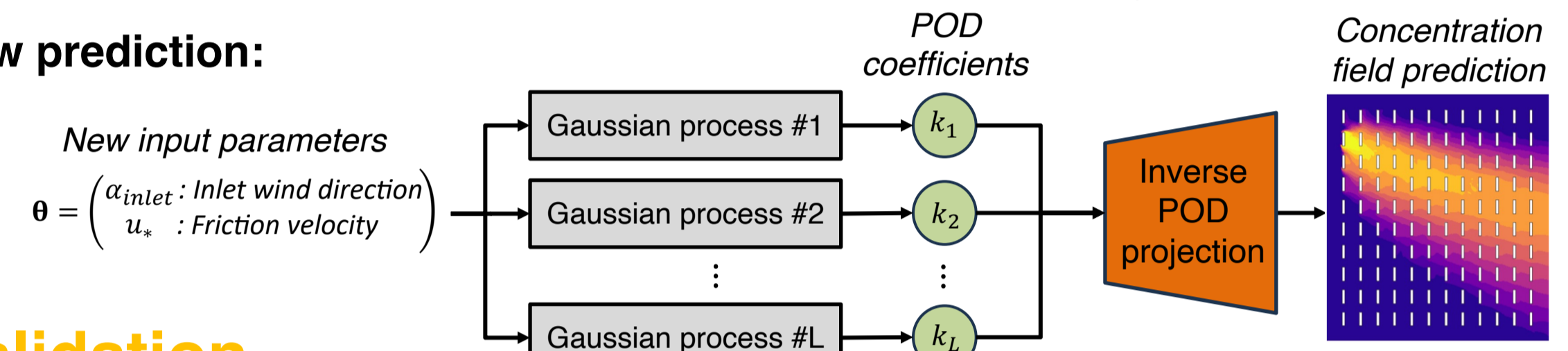
- **Sensitivity analysis:** Wind velocity and direction \gg rugosity, turbulence intensity, SGS model, ...
- **Microclimatology** to define realistic parameter ranges
- **Input parameter space sampling** using the low-discrepancy Halton sequence

Generation of an ensemble of 200 LES for varying wind boundary conditions (≈ 6 Mh_{CPU})

B. The POD-GPR surrogate model

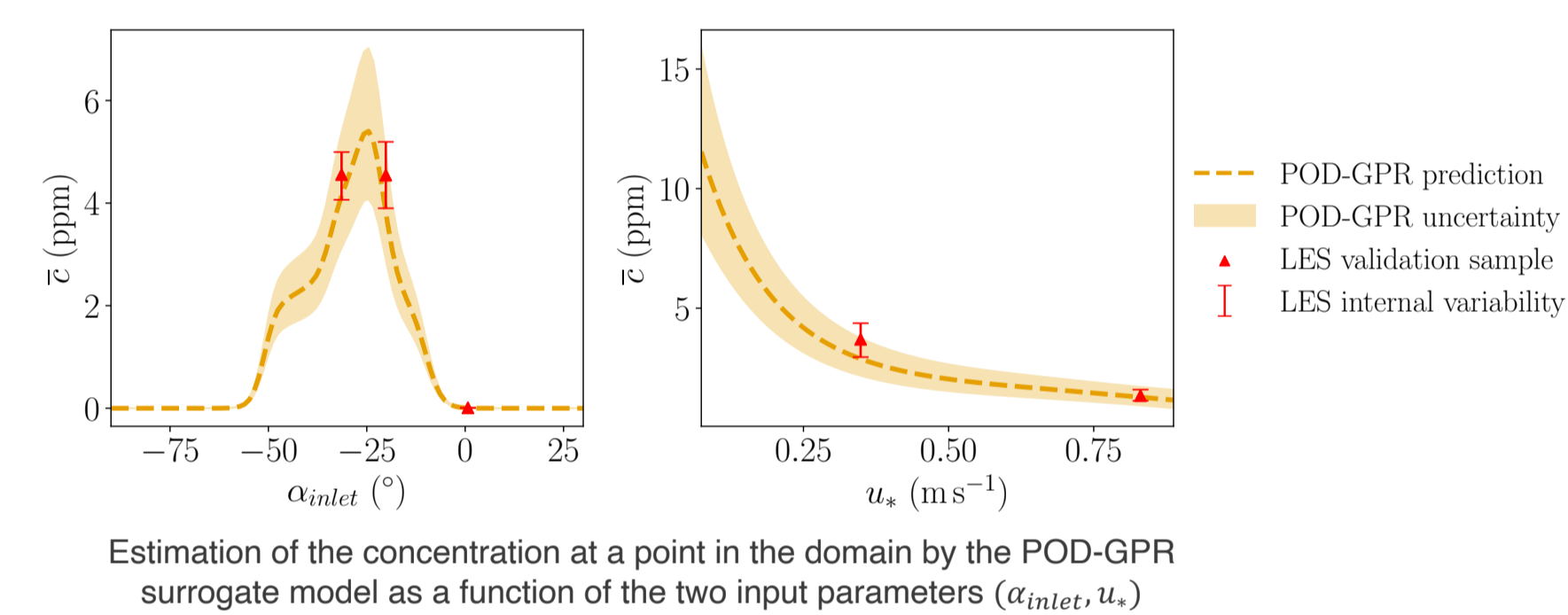
- **Two-step approach:** i. Dimension reduction using Proper Orthogonal Decomposition (POD) (Nony et al. 2023) ii. Estimation of POD coefficients using Gaussian Process Regressors

New prediction:



C. Validation

- i. **Validation** on a test set of 40 LES shows near-maximum accuracy given the noise in the database (except for very high concentrations near the source)
- ii. **Computational costs:** Prediction $\approx 0.03s$ versus Training $\approx 30s$
- iii. **Surrogate model uncertainty** = Model reduction uncertainty + Internal variability



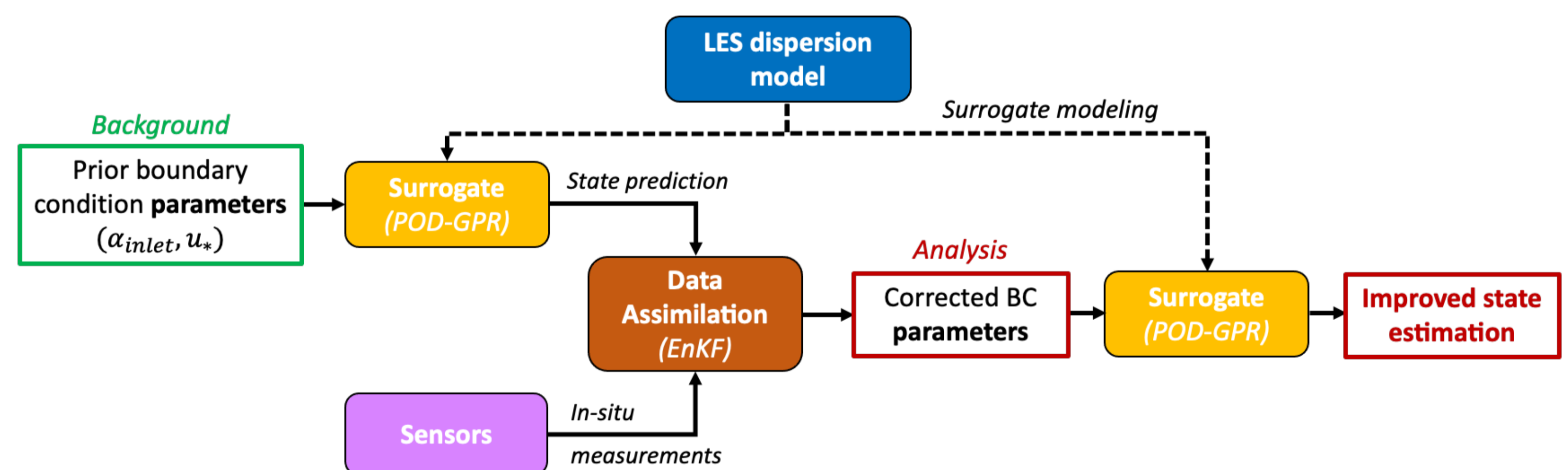
The POD-GPR surrogate model can accurately and efficiently replace the LES model while representing part of the uncertainties involved

3) Data assimilation for wind boundary condition parameters estimation

Assimilate in situ measurements to estimate large-scale wind boundary conditions and improve LES concentration field prediction

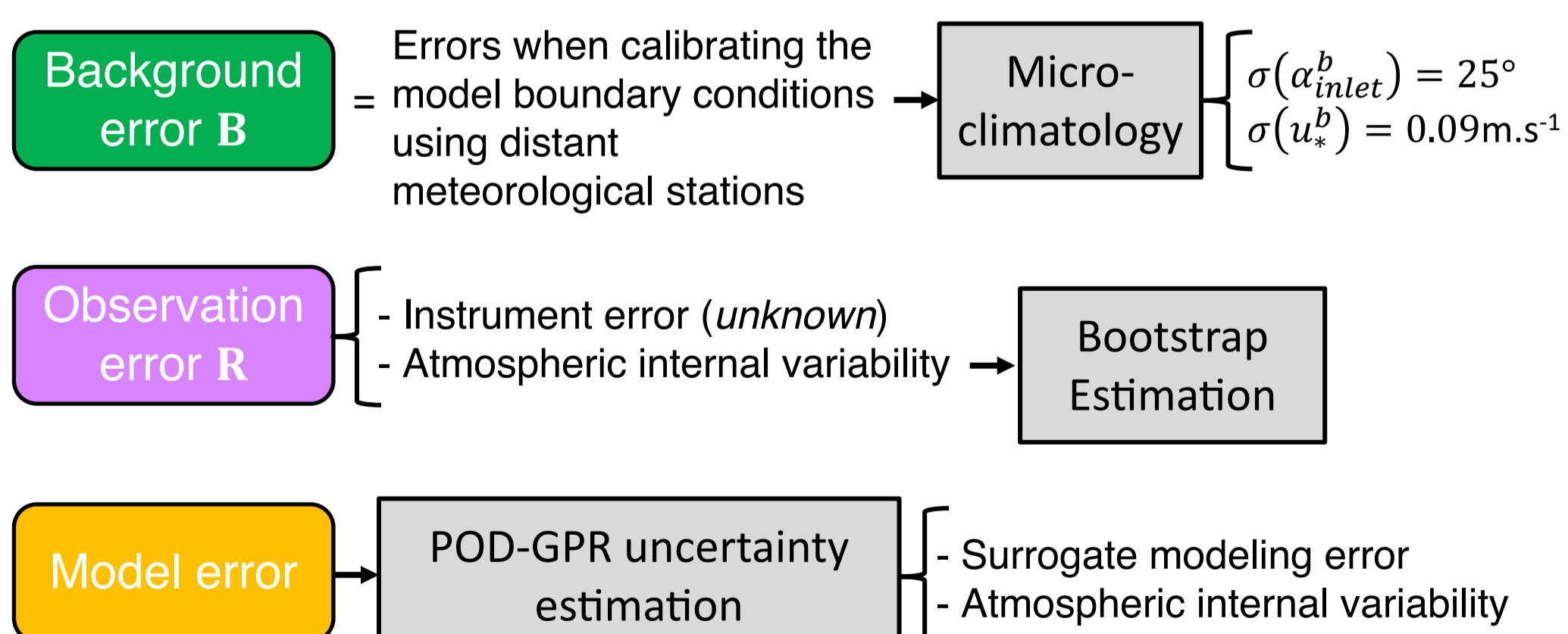
A. Data assimilation framework

- **System state:** Mean concentration field
- **Control vector:** Boundary condition parameters (α_{inlet}, u_*) as initial conditions quickly vanish at microscale (Deforge 2019)
- **Observations:** 13 concentration measurements at different locations
- **Anamorphosis** (Deforge et al. 2021)
- **Data assimilation method:** Ensemble Kalman Filter (EnKF)
- **Ensemble size:** 500 members (not a problem using the surrogate model)

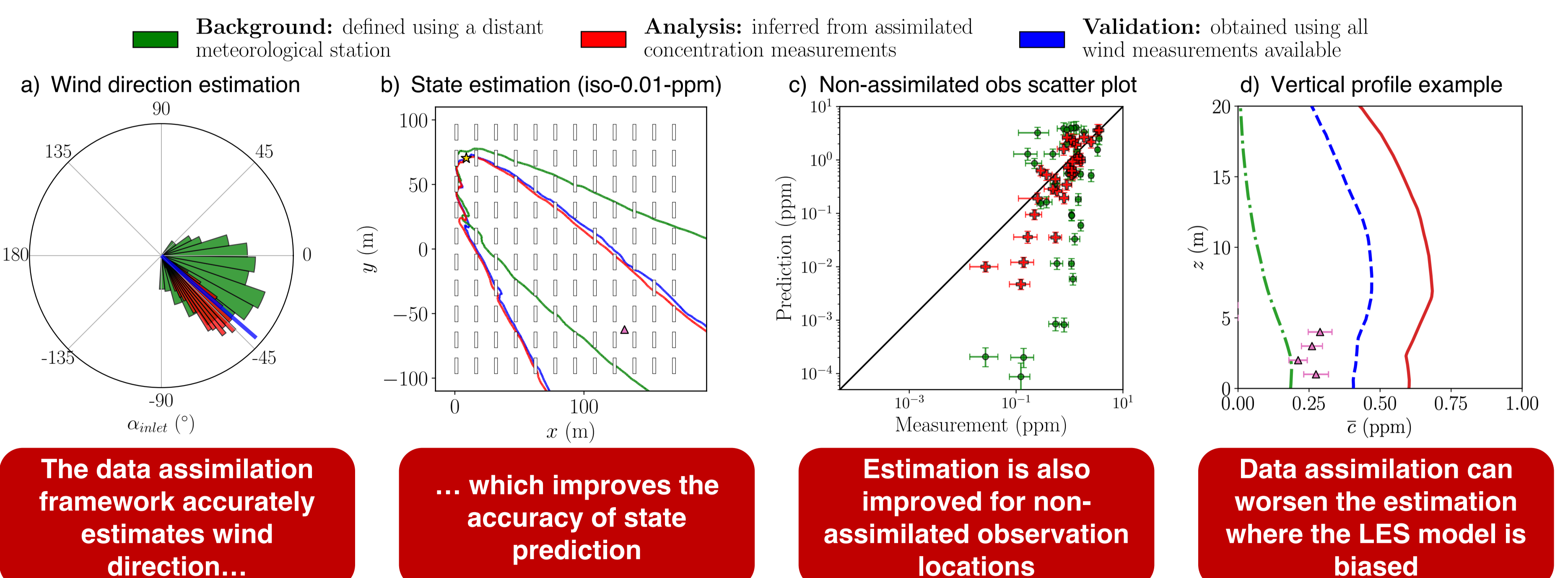


B. Errors modeling

Robust and realistic errors models are necessary to ensure pertinent analysis



C. Assimilation of field measurements



The data assimilation framework accurately estimates wind direction...

... which improves the accuracy of state prediction

Estimation is also improved for non-assimilated observation locations

Data assimilation can worsen the estimation where the LES model is biased

4) Take-home messages

- The use of a surrogate model enables real-time data assimilation and large ensemble size while also providing an estimation of the uncertainties involved
- The proposed data assimilation framework efficiently corrects wind boundary conditions and improve pollutant dispersion predictions
- **Perspectives:** i) state-parameter estimation, ii) optimal sensor placement, and iii) assimilation of plume images

Poster & materials

