CERFACS

CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Workshop - Traitement des données massives en mécanique des fluides

Quantifying uncertainties in large eddy simulations of pollutant dispersion using surrogate models

Mélanie Rochoux*, Géraldine Rea, Nicolas Frebourg, Claire Lamotte, Matthias De Lozzo & Olivier Vermorel

UNCERTAINTY QUANTIFICATION

LARGE EDDY SIMULATIONS

SURROGATE MODEL

AIR QUALITY

Introduction

What are the challenges in pollutant dispersion numerical simulations?

SAFETY ISSUE

Map the areas with peak concentration of air pollutants

COMPLEXITY

Multi-physics multi-scale problem

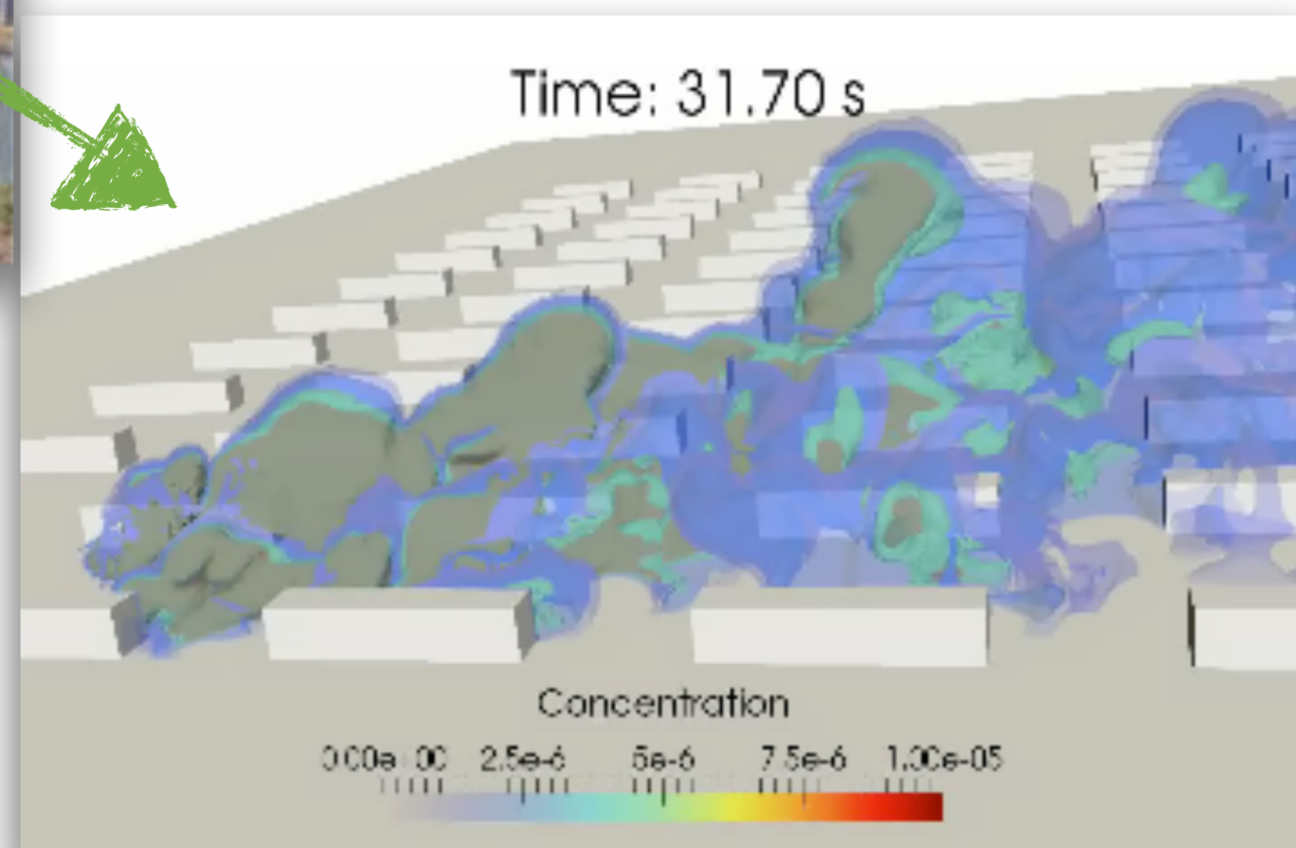
- poor information on the emission sources
- strong interaction between the land surface and the near-surface atmosphere

How to represent and characterize the early stage of the smoke plume near the source?

Introduction

The Mock Urban Setting Test (MUST): The case study

Objective: **High-fidelity** simulations and **Uncertainties** for **micro-scale** meteorology and air quality



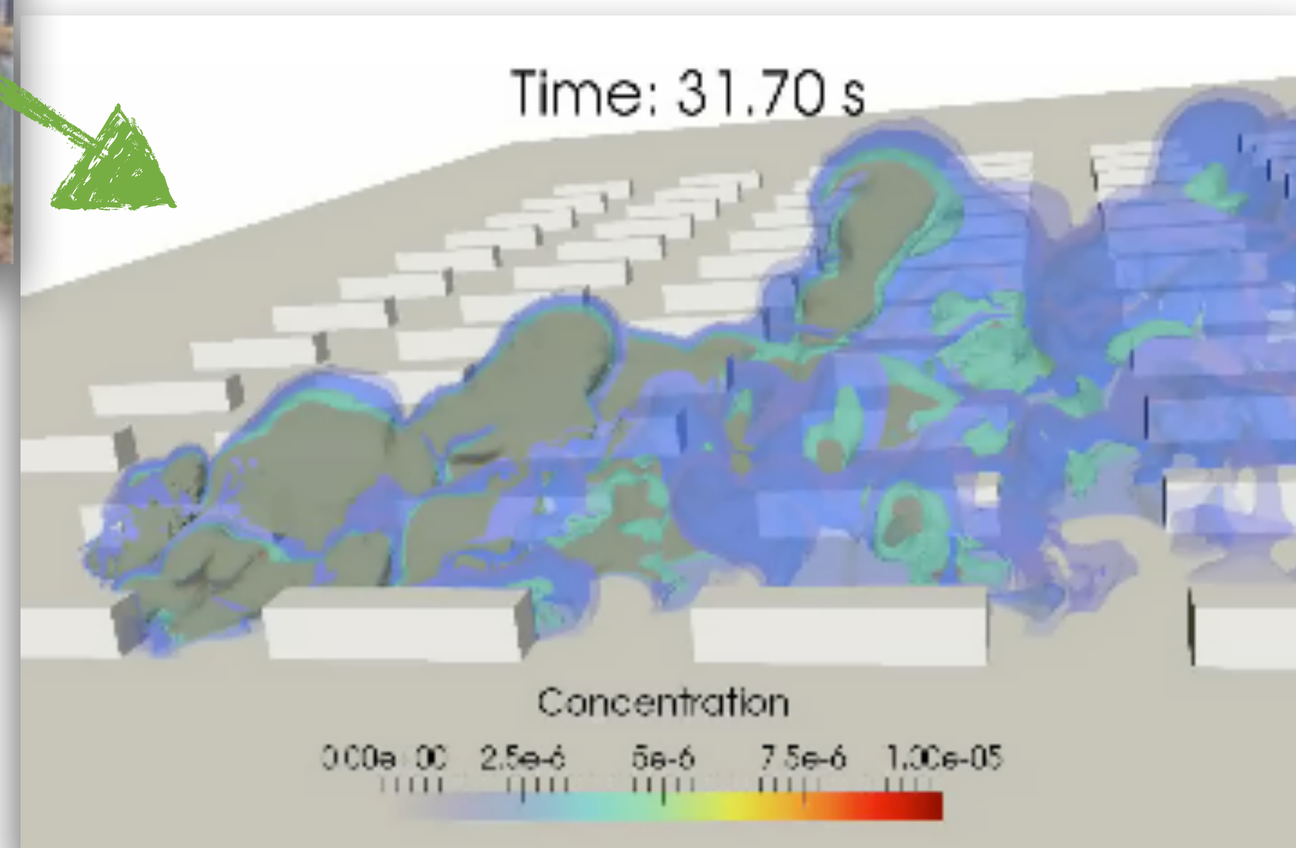
micro-scale

- explicit simulation of the plume made of air pollutants near the emission source
- explicit representation of surface-atmosphere interactions

Introduction

The Mock Urban Setting Test (MUST): The case study

Objective: **High-fidelity** simulations and **Uncertainties** for **micro-scale** meteorology and air quality



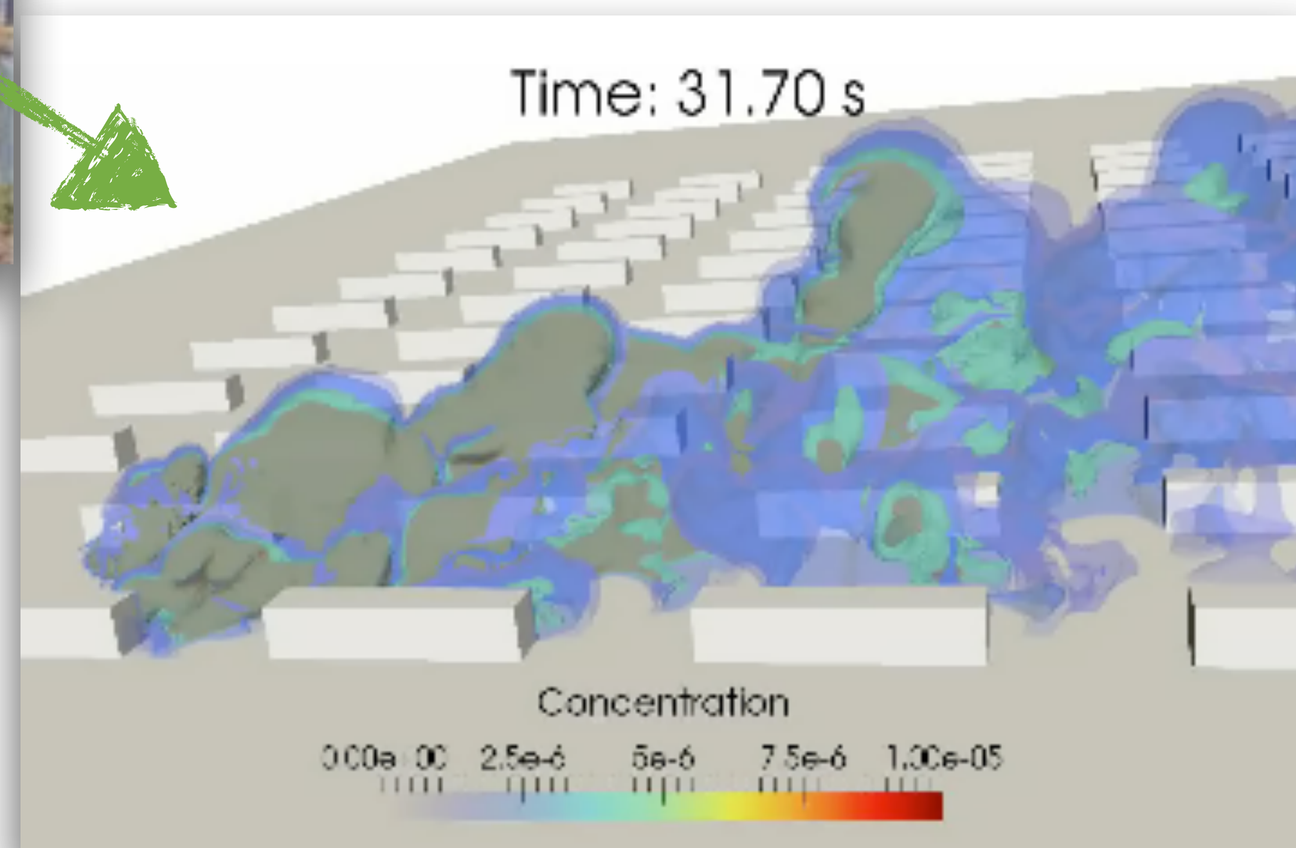
high-fidelity

- added value of large eddy simulations (LES)
- 3-D unsteady turbulence
- spatial resolution ~1 m
- massively parallel simulations

Introduction

The Mock Urban Setting Test (MUST): The case study

Objective: **High-fidelity** simulations and **Uncertainties** for **micro-scale** meteorology and air quality



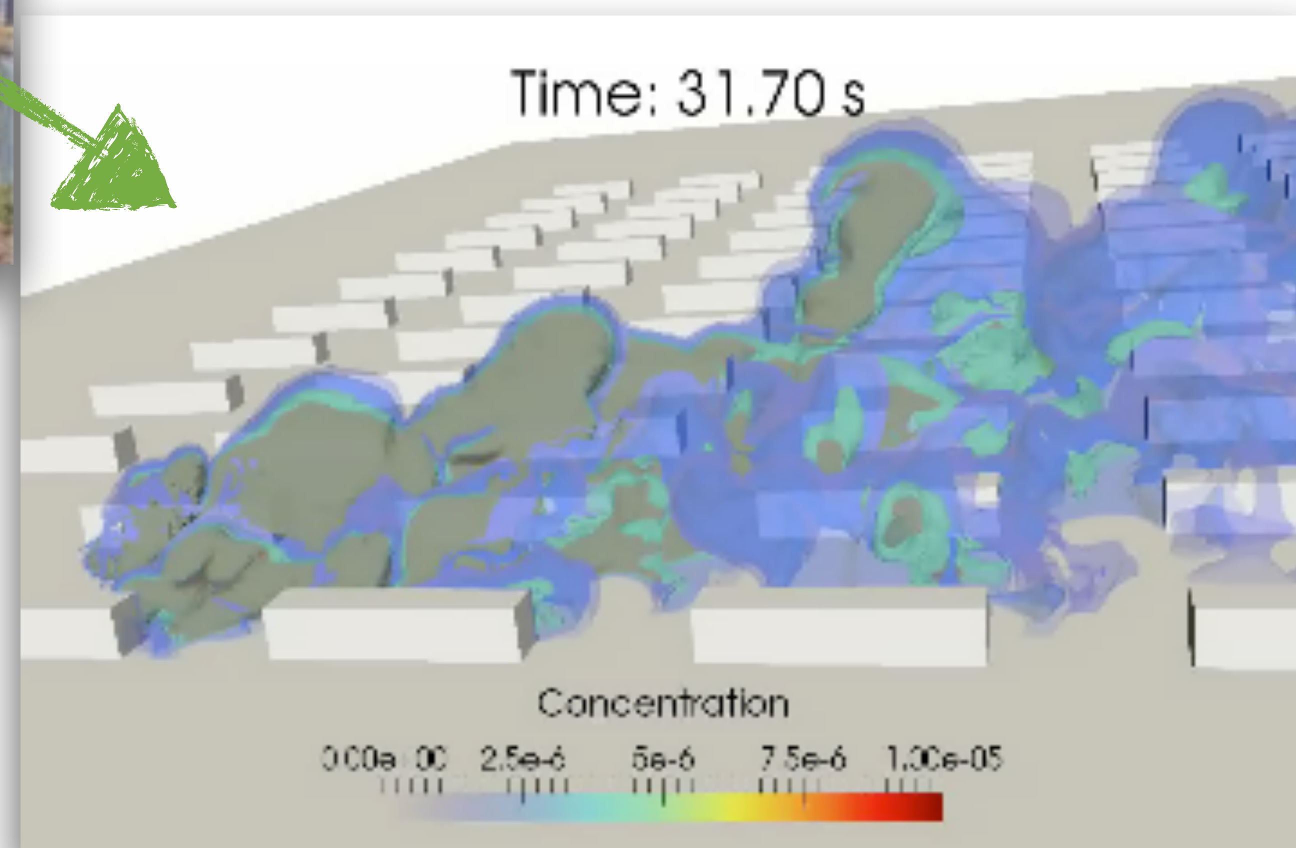
uncertainties

- ensemble of LES-type simulations
- sensitivity to users' choices (physical and numerical parameters)
- impact of meteorological hazards (spatial and temporal intrinsic

Introduction

The Mock Urban Setting Test (MUST): The case study

Objective: **High-fidelity** simulations and **Uncertainties** for **micro-scale** meteorology and air quality



Meso-NH mesoscale non-hydrostatic model

YALES2

AVBP CERFACS

Intercomparison of LES-type simulations and sensitivity to inlet wind

Introduction

Talk's outline

(1) MUST case study
- Experimental settings
- Initial and inlet wind conditions

(2) Best known large eddy simulations
- Solvers: AVBP, YALES2-AE, Meso-NH
- Numerical settings
- Diagnostics

(3) Sensitivity to inlet wind conditions
- Uncertainty quantification in a nutshell
- Inlet wind statistics
- Surrogate models
- Diagnostics

STAE TOULOUSE

Introduction

Talk's outline

(1) MUST case study
- Experimental settings
- Initial and inlet wind conditions

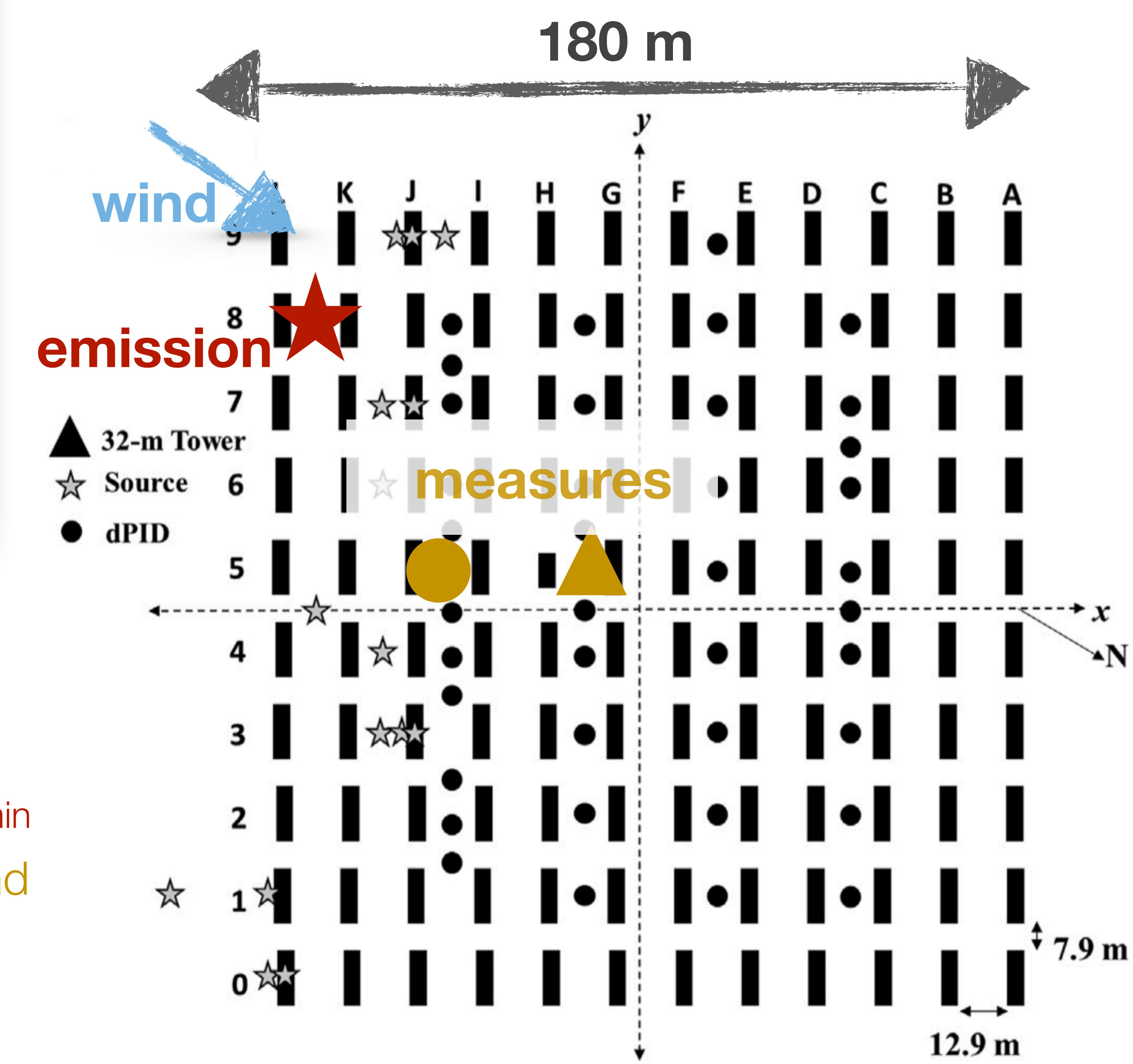
(2) Best known large eddy simulations
- Solvers: AVBP, YALES2-AE, Meso-NH
- Numerical settings
- Diagnostics

(3) Sensitivity to inlet wind conditions
- Uncertainty quantification in a nutshell
- Inlet wind statistics
- Surrogate models
- Diagnostics

STAE TOULOUSE

MUST trial 2681829

Experimental settings - Near-neutral conditions (6:30 PM)



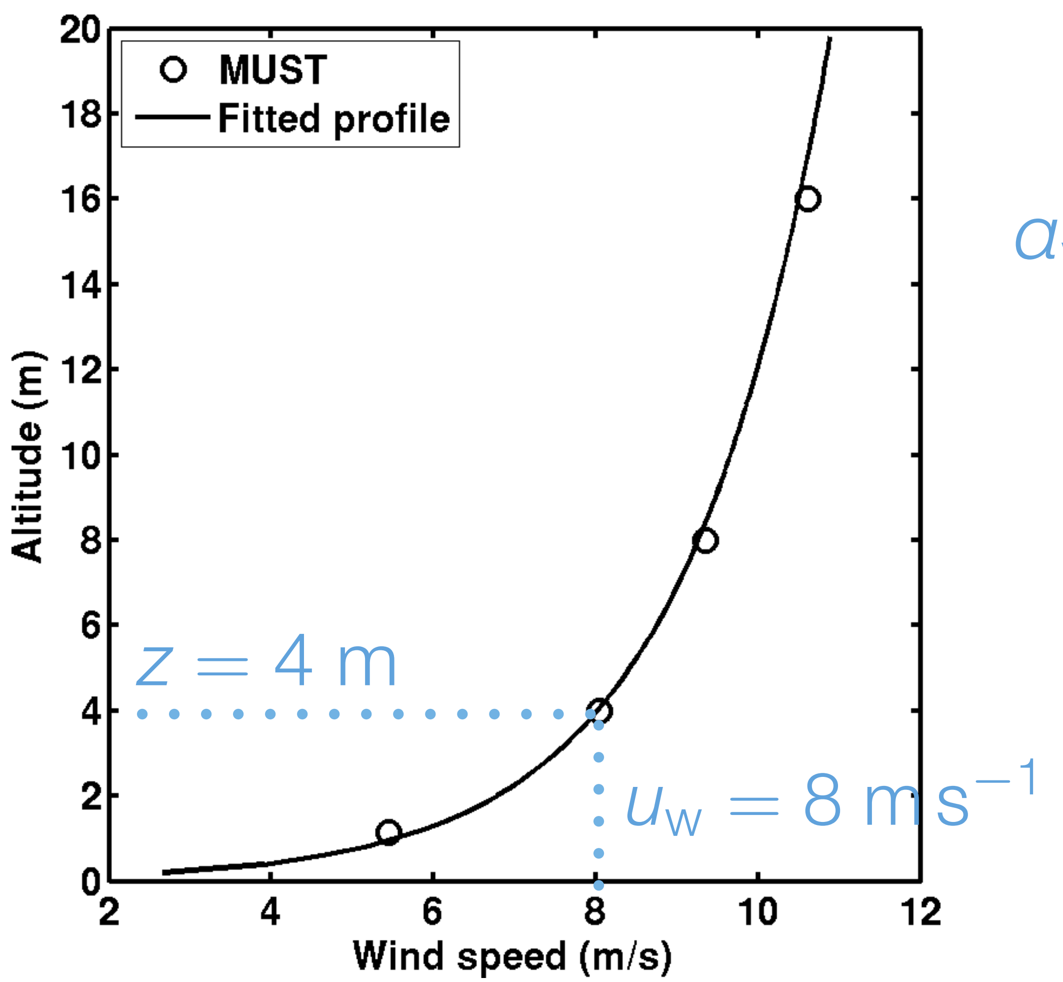
container (12 m x 2,4 m x 2,5 m)
➡ regular array of 120 containers

local emission at z = 1,8 m
➡ 200 s, passive tracer (propylene), 225 L/min

available measurements for wind and tracer concentration across the container array, upstream and downstream

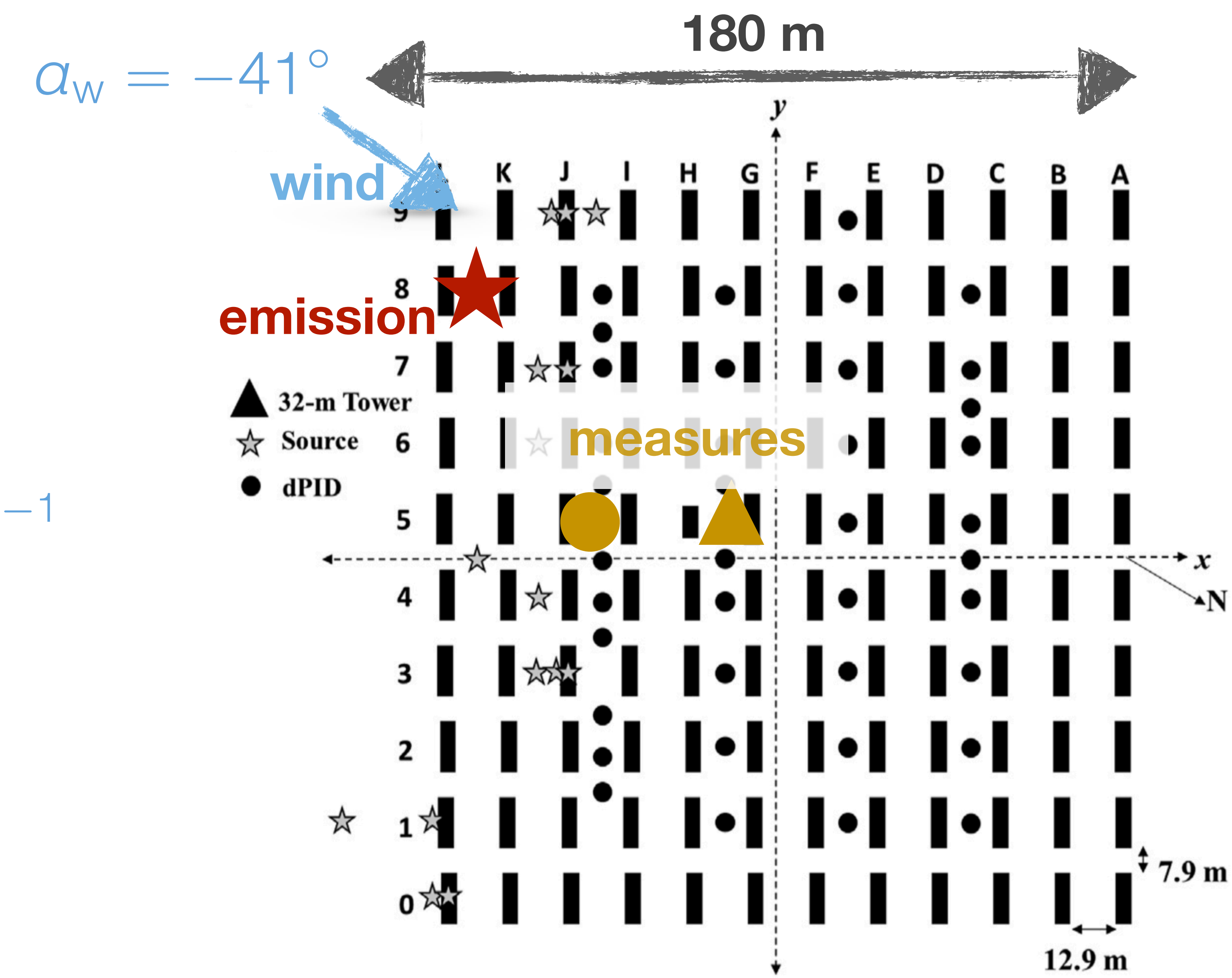
MUST trial 2681829

Initial and inlet wind profile



$$\overline{u}(z) = \frac{u_*}{\kappa} \ln\left(\frac{z + z_0}{z_0}\right)$$

$$u_* = 0.715\ \mathrm{m\,s^{-1}}$$



Introduction

Talk's outline

(1) MUST case study
- Experimental settings
- Initial and inlet wind conditions

(2) Best known large eddy simulations
- Solvers: AVBP, YALES2-AE, Meso-NH
- Numerical settings
- Diagnostics

(3) Sensitivity to inlet wind conditions
- Uncertainty quantification in a nutshell
- Inlet wind statistics
- Surrogate models
- Diagnostics

Large eddy simulation solvers

Resolution of 3-D Navier-Stokes equations with unsteady turbulence

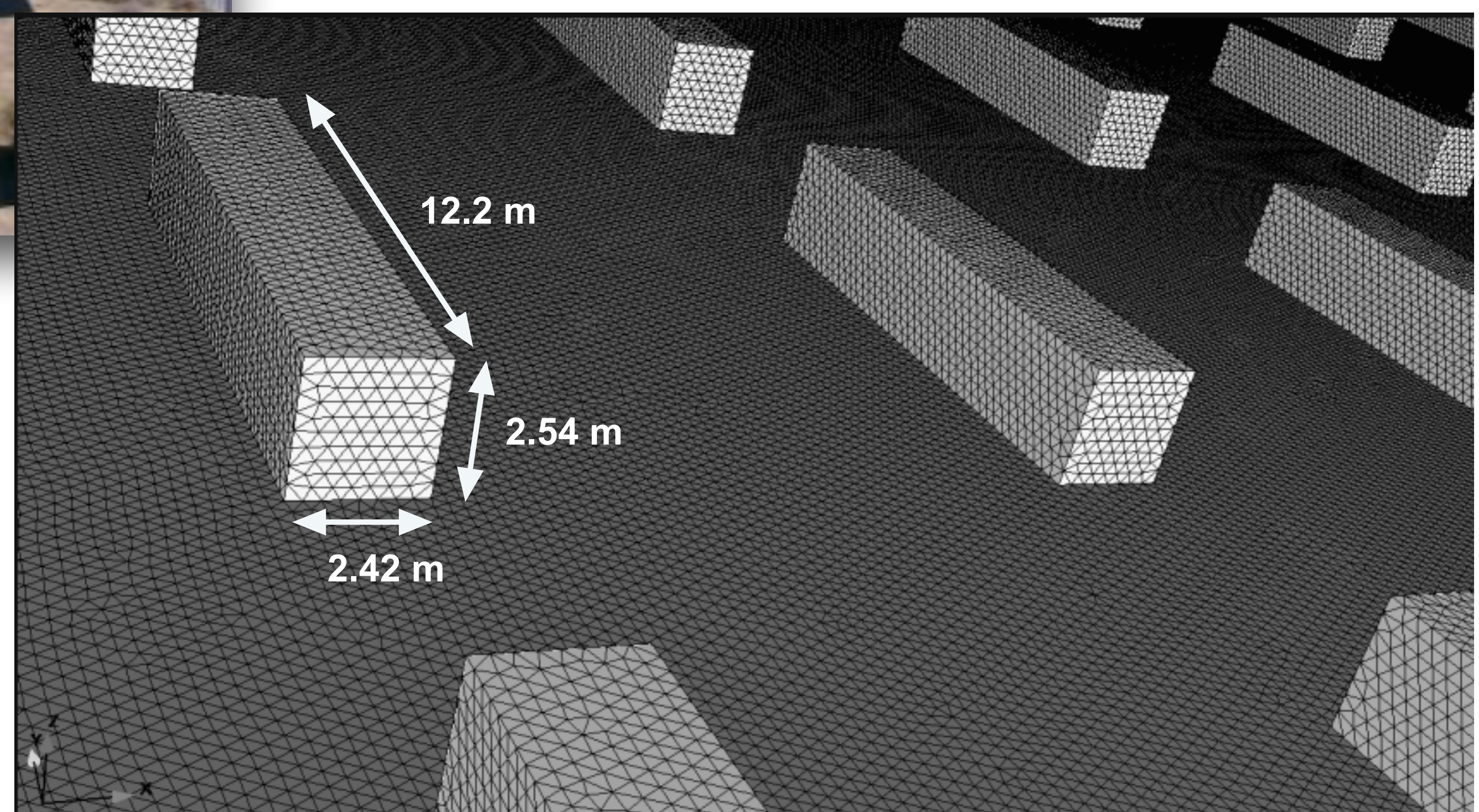
	Meso-NH (mesoscale non-hydrostatic model)	YALES2	AVBP
	Météo-France, Laboratoire d'Aérologie	*CORIA - CERFACS*	*CERFACS - IFPEN*
Equations	Structured grid Incompressible (anelastic approximation)	Unstructured grid Low Mach approx.	Unstructured grid Compressible
Container boundary condition	Immersed Boundary Method (Auguste et al.)	Boundary fitted	Boundary fitted
Numerical schemes	• Space: WENO 5 (Lunet et al. 2017) • Time: Runge-Kutta 4	• Space: 4th order centered scheme • Time: Runge-Kutta 4 (TFV4A)	3rd order in space and time, explicit, two-step Taylor-Garlerkin (TTGC)
Subgrid-scale turbulence model	TKE 1.5	WALE (Wall Adaptative Local Eddy Viscosity, Nicoud and Ducros 1999)	WALE (Wall Adaptative Local Eddy Viscosity, Nicoud and Ducros 1999)

Numerical settings

Computational domain and grid

Unstructured mesh

- 350 m x 350 m x 50 m
- min. resolution = 30 cm
- 71 million of grid cells
- YALES2 = 27 000 h CPU (~16 h, low Mach)
- AVBP = 165 000 h CPU (~68 h, compressible)

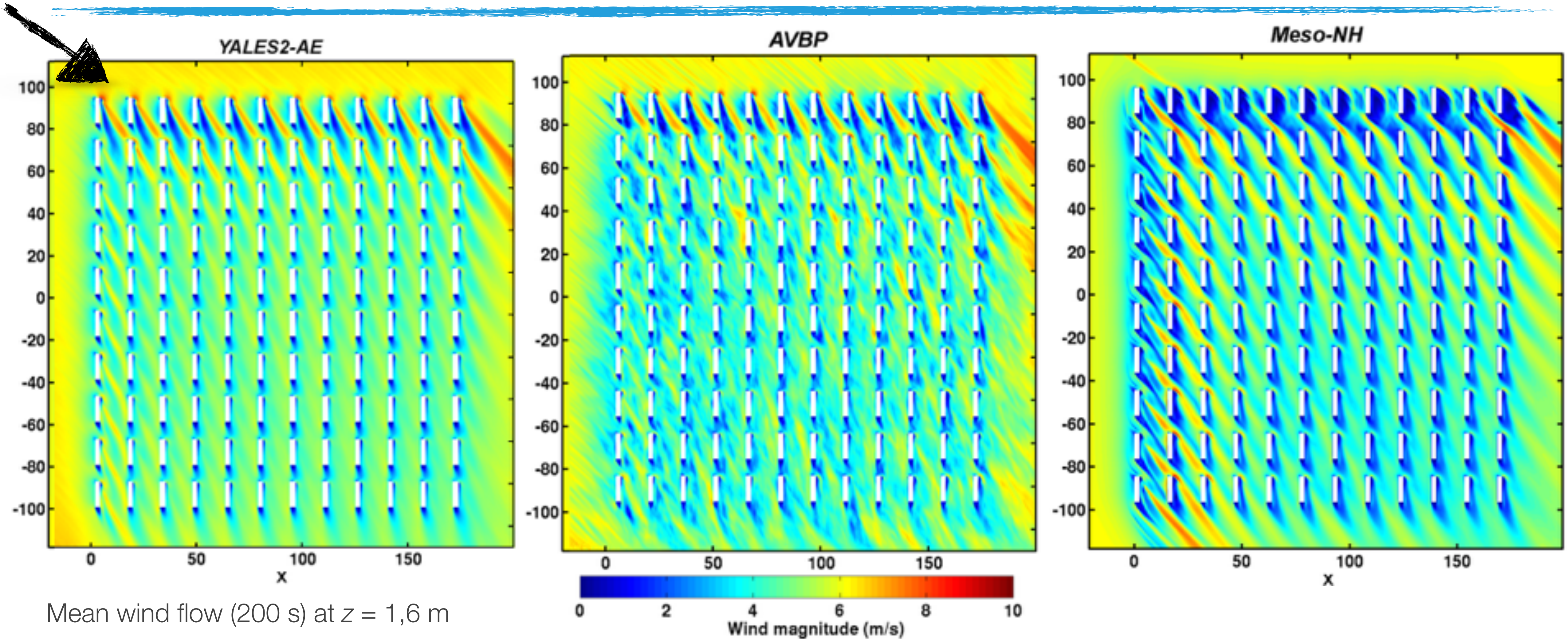


Meso-NH
mesoscale non-hydrostatic model

Structured mesh

- 300 m x 300 m x 40 m
- min. resolution = 20 cm
- 150 million of grid cells
- MesoNH = 40 000 h CPU (~20 h) (incompressible anelastic)

Wind flow dynamics

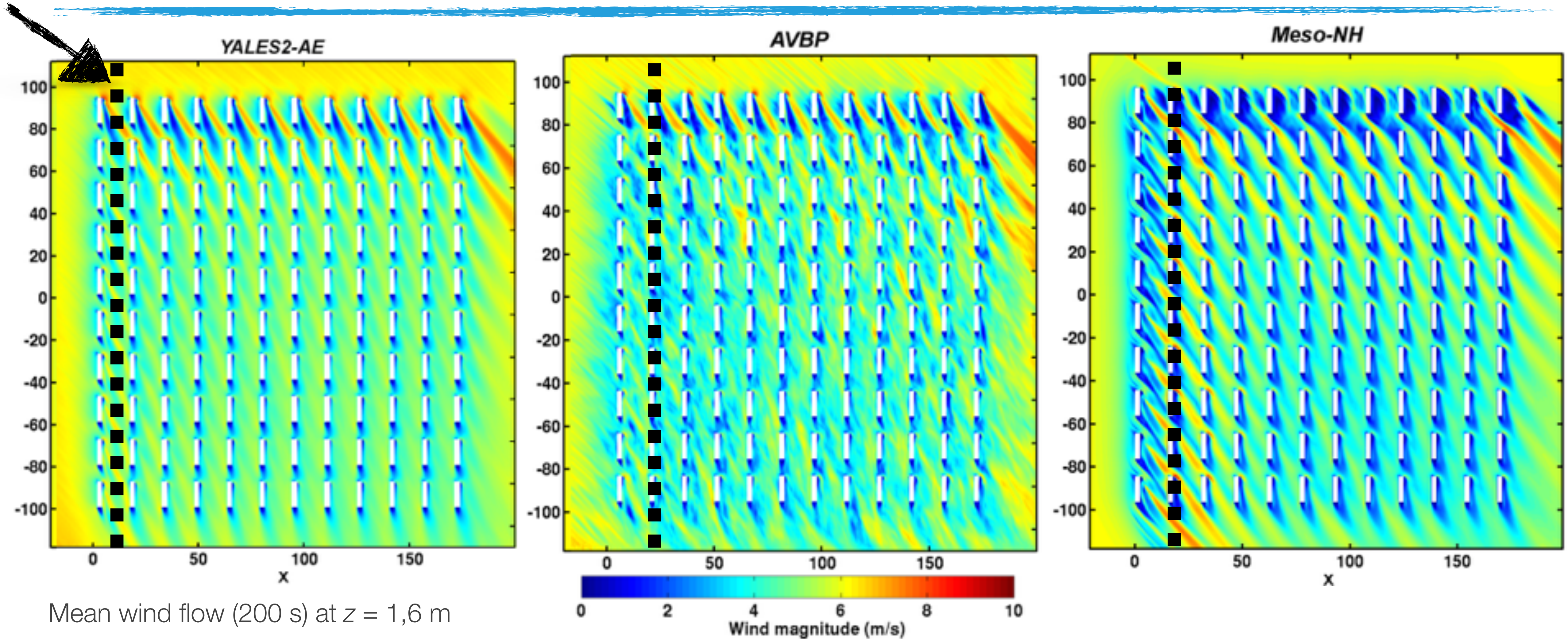


Mean wind flow (200 s) at z = 1,6 m

MAIN RESULTS

- Acceleration (channel effects) and slowing down induced by the containers, which were measured and captured by LES
- Persistent influence until z = 10 m
- Deviation of the flow direction induced by the containers, which is more important for YALES2-AE and AVBP than for MesoNH

Wind flow dynamics

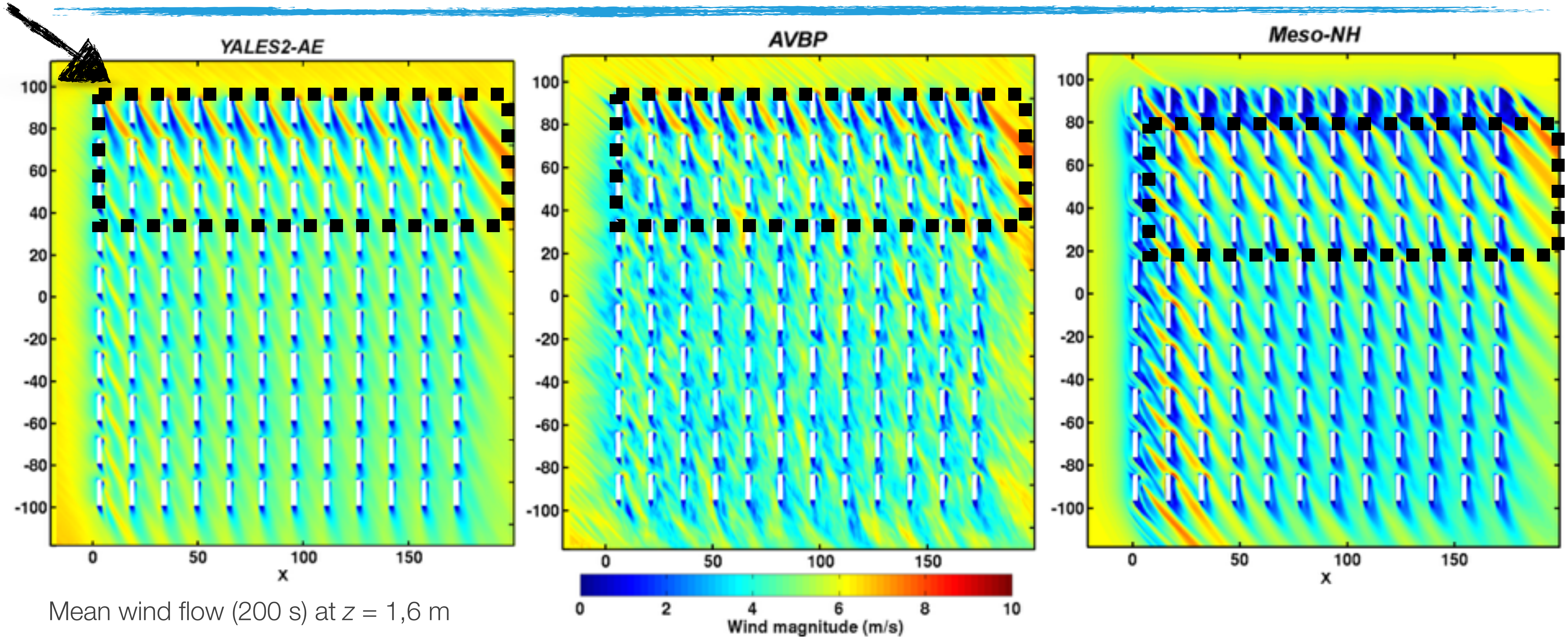


Mean wind flow (200 s) at z = 1,6 m

MAIN RESULTS

- Acceleration (channel effects) and slowing down induced by the containers, which were measured and captured by LES
- Persistent influence until z = 10 m
- Deviation of the flow direction induced by the containers, which is more important for YALES2-AE and AVBP than for MesoNH

Wind flow dynamics

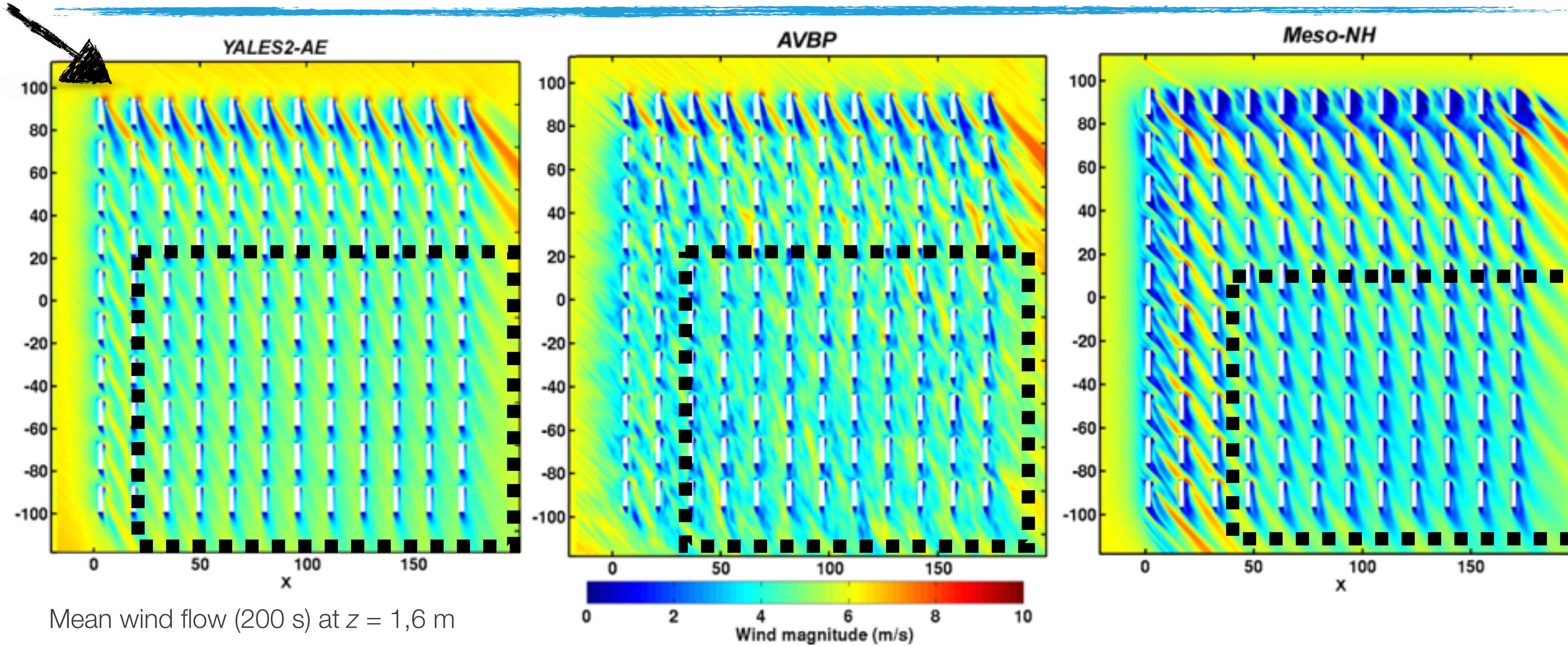


Mean wind flow (200 s) at z = 1,6 m

MAIN RESULTS

- Acceleration (channel effects) and slowing down induced by the containers, which were measured and captured by LES
- Persistent influence until z = 10 m
- Deviation of the flow direction induced by the containers, which is more important for YALES2-AE and AVBP than for MesoNH

Wind flow dynamics

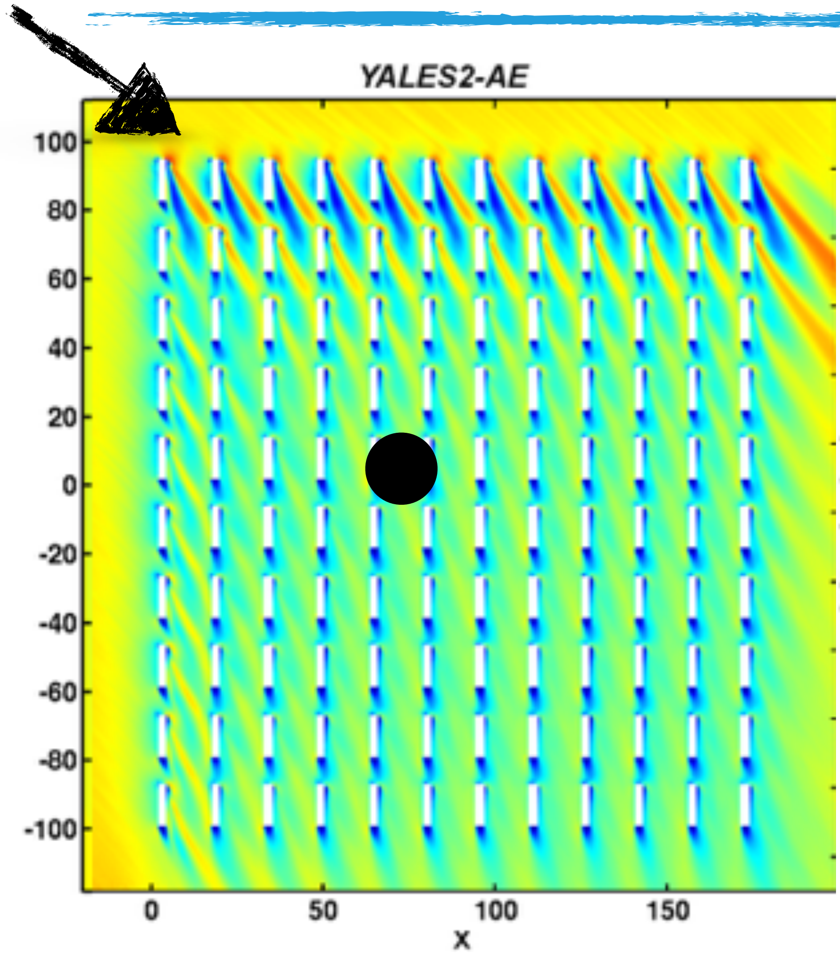


Mean wind flow (200 s) at z = 1,6 m

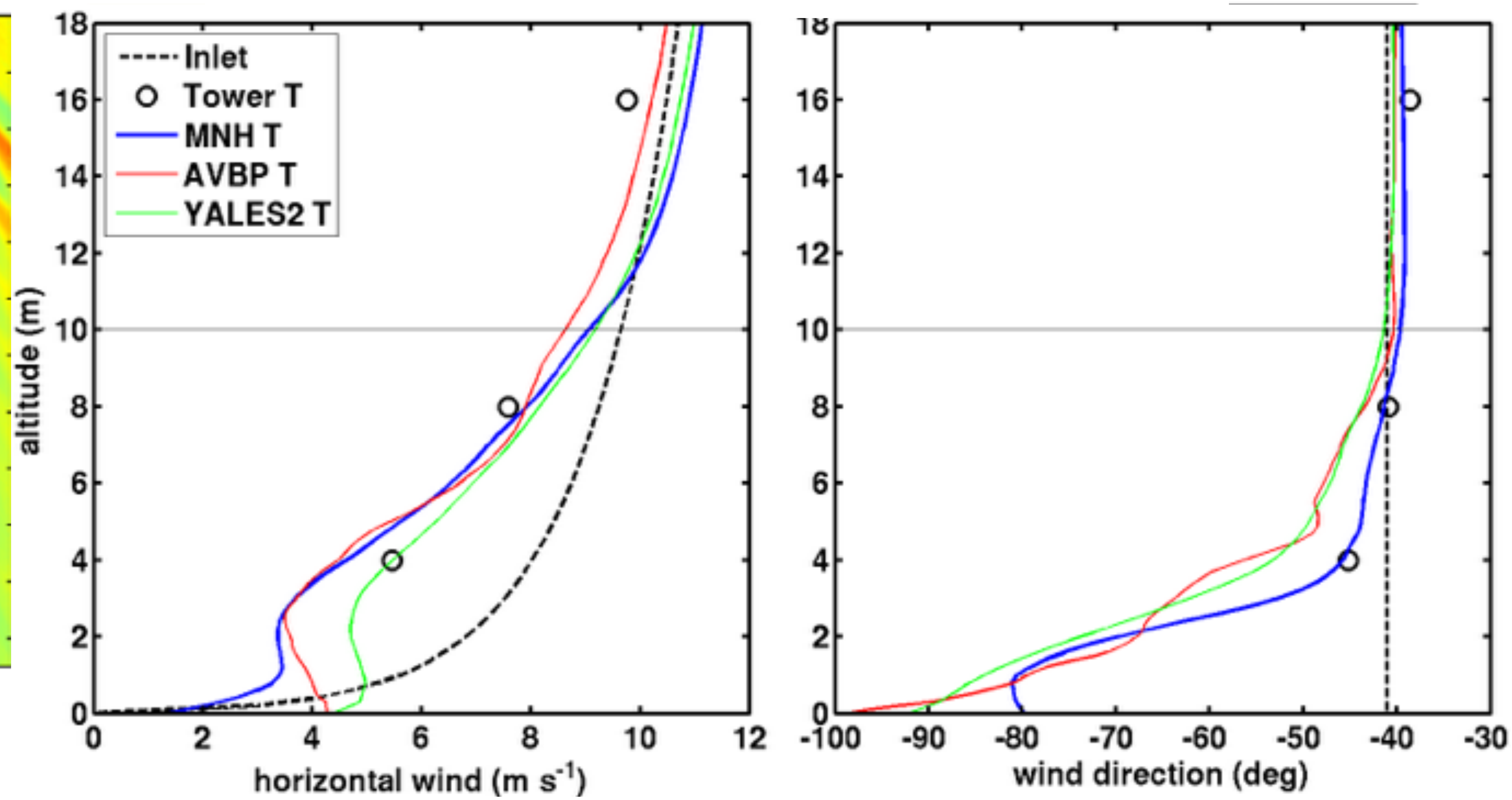
MAIN RESULTS

- Acceleration (channel effects) and slowing down induced by the containers, which were measured and captured by LES
- Persistent influence until z = 10 m
- Deviation of the flow direction induced by the containers, which is more important for YALES2-AE and AVBP than for MesoNH

Wind flow dynamics



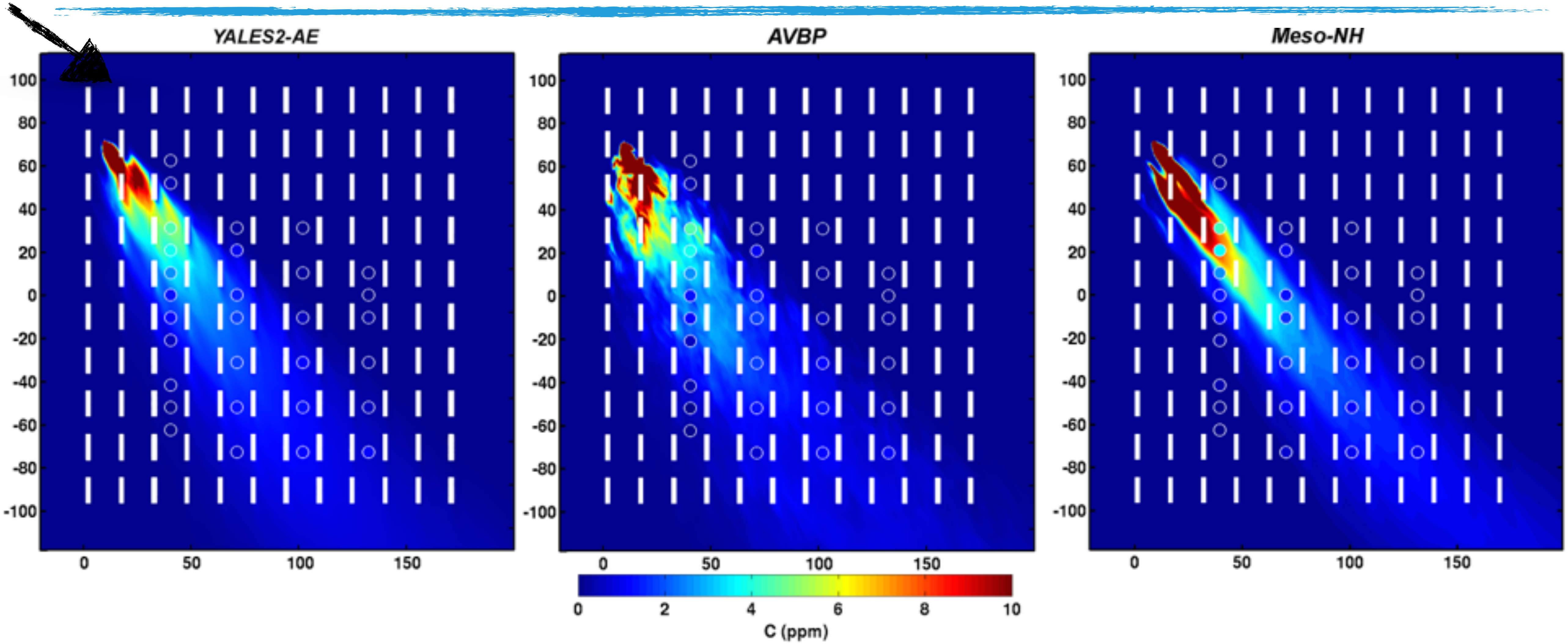
Mean wind flow (200 s) at z = 1,6 m



MAIN RESULTS

- Acceleration (channel effects) and slowing down induced by the containers, which were measured and captured by LES
- Persistent influence until z = 10 m
- Deviation of the flow direction induced by the containers, which is more important for YALES2-AE and AVBP than for MesoNH

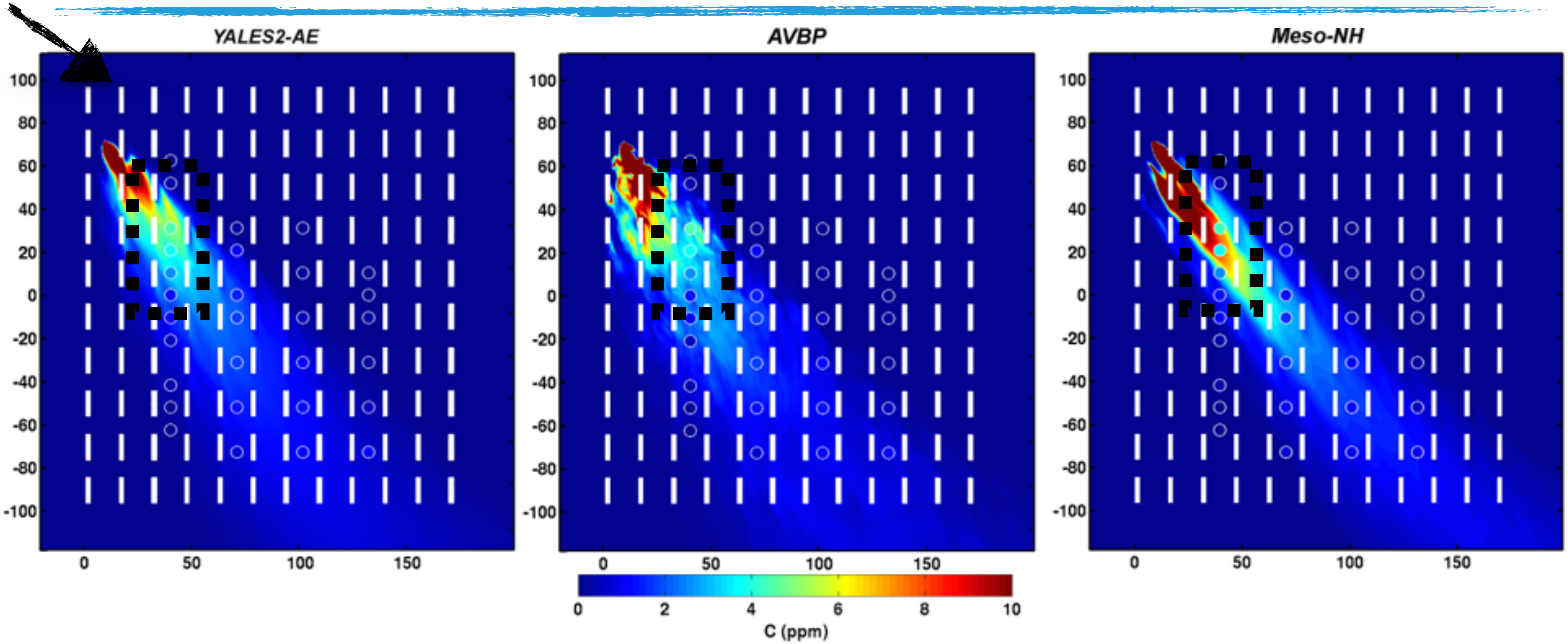
Tracer concentration



MAIN RESULTS

- Deviation of the plume main axis with respect to the inlet wind direction
- Impact on the plume shape and on the location of min./max. tracer concentration
- Good concentration statistics (Chang and Hanna, 2004)
- Good match in terms of time series when high tracer concentration (> 1 ppm)

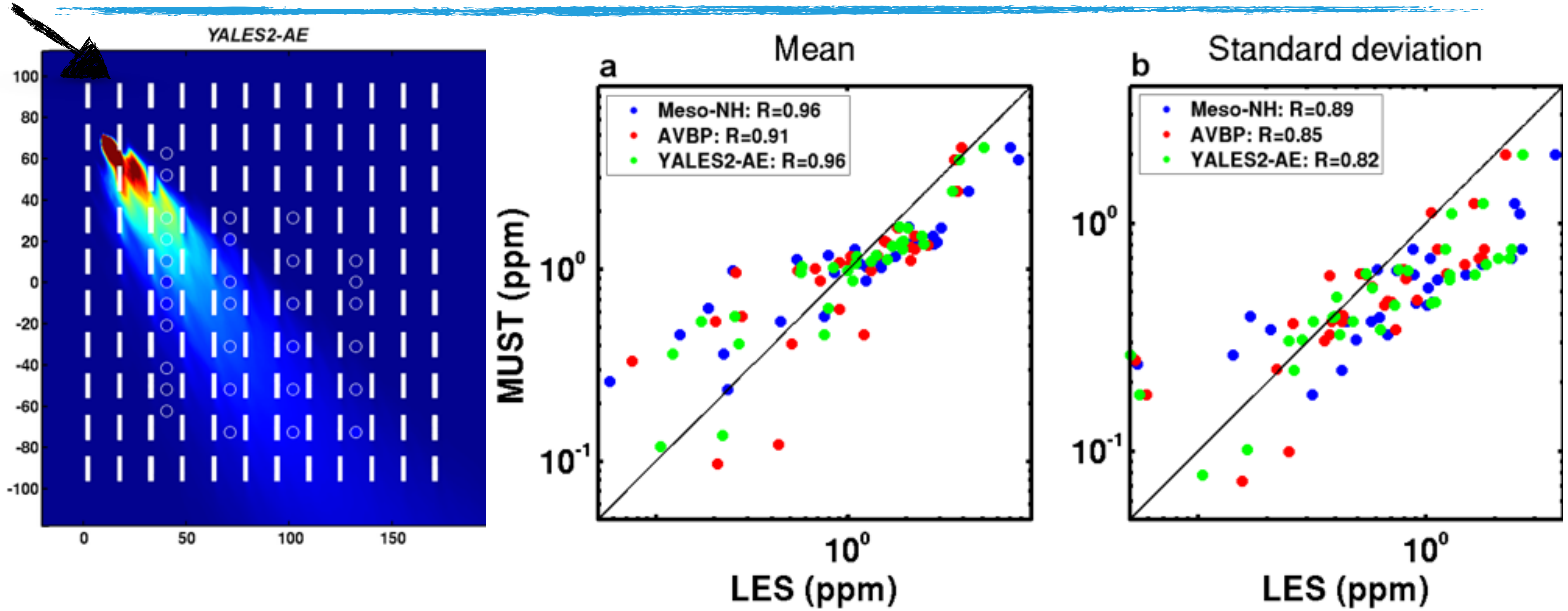
Tracer concentration



MAIN RESULTS

- Deviation of the plume main axis with respect to the inlet wind direction
- Impact on the plume shape and on the location of min./max. tracer concentration
- Good concentration statistics (Chang and Hanna, 2004)
- Good match in terms of time series when high tracer concentration (> 1 ppm)

Tracer concentration



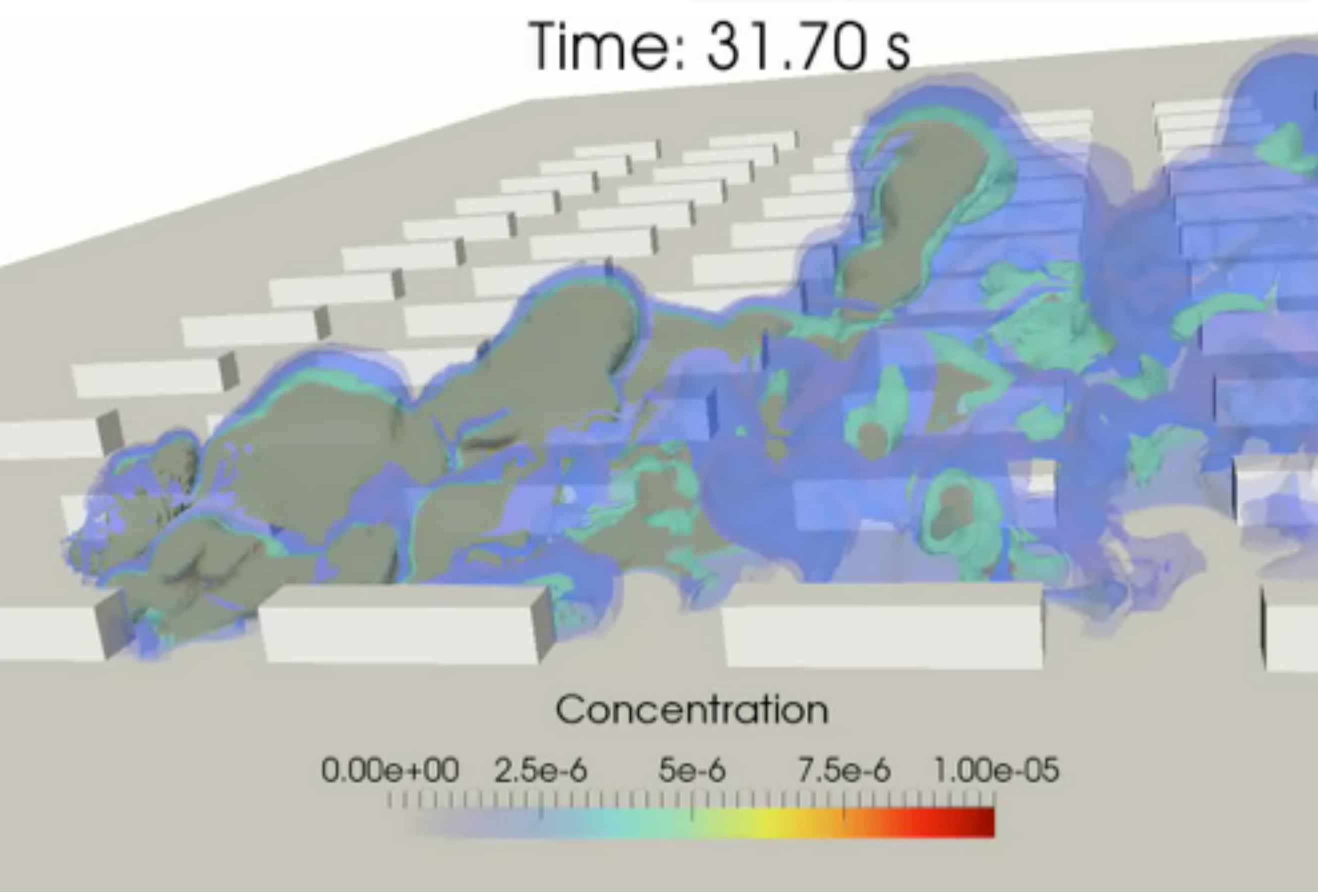
MAIN RESULTS

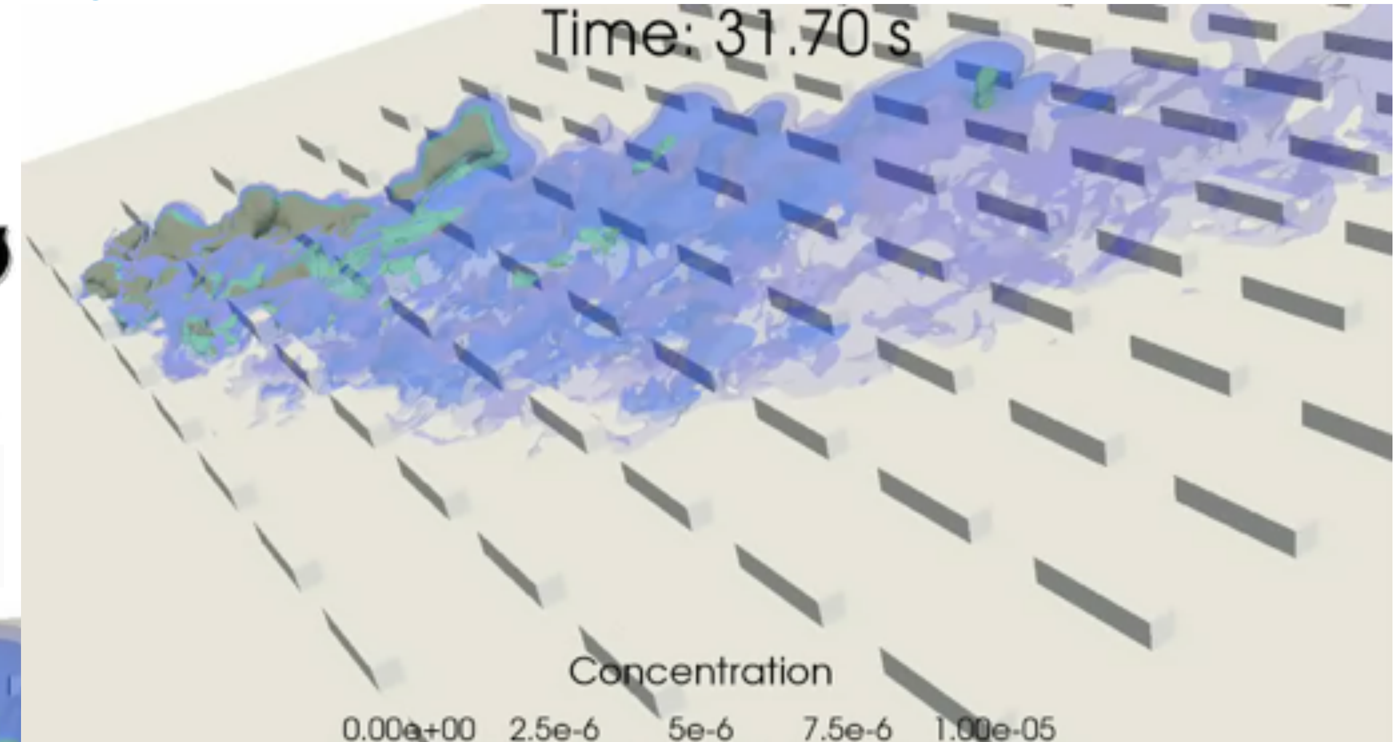
- Deviation of the plume main axis with respect to the inlet wind direction
- Impact on the plume shape and on the location of min./max. tracer concentration
- Good concentration statistics (Chang and Hanna, 2004)
- Good match in terms of time series when high tracer concentration (> 1 ppm)

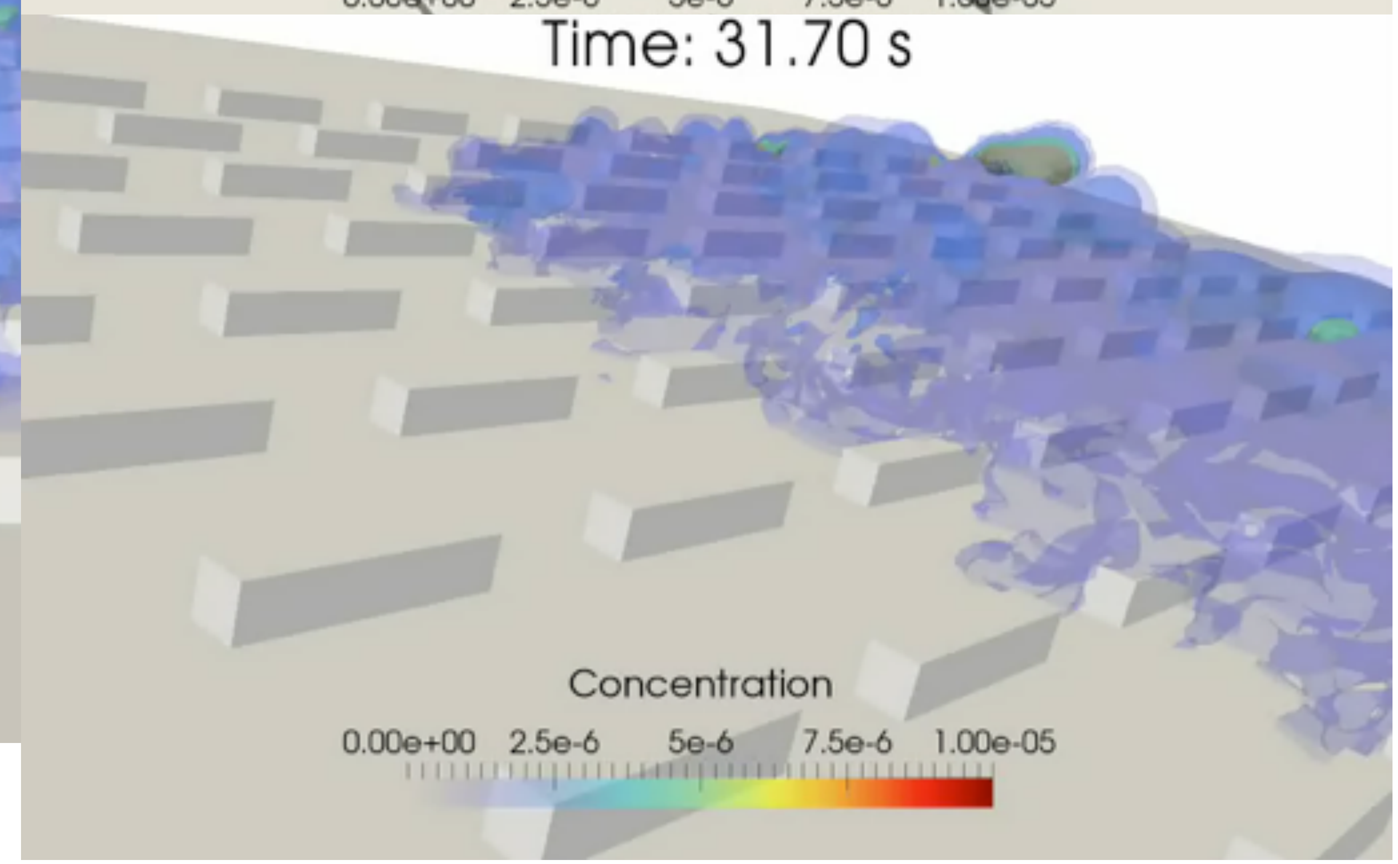
A computing challenge!

AVBP
CERFACS

GENCI







Introduction

Talk's outline

(1) MUST case study

- Experimental settings
- Initial and inlet wind conditions

(2) Best known large eddy simulations

- Solvers: AVBP, YALES2-AE, Meso-NH
- Numerical settings
- Diagnostics

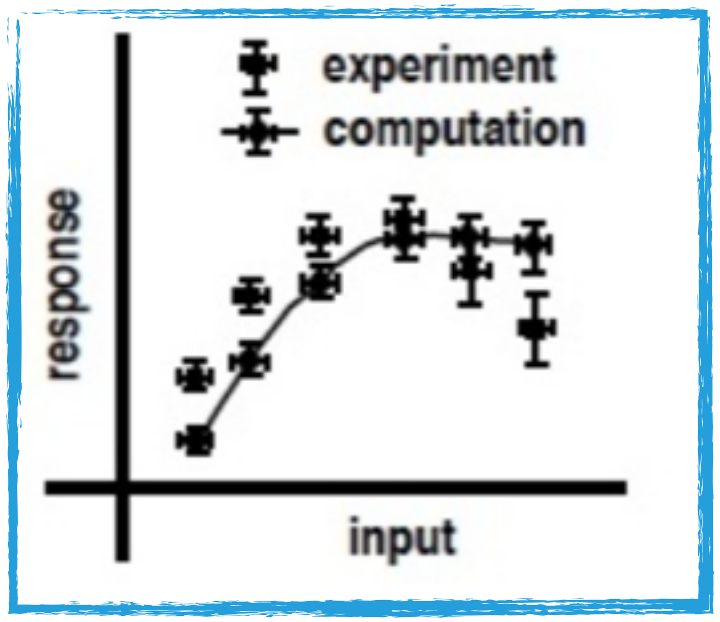
(3) Sensitivity to inlet wind conditions

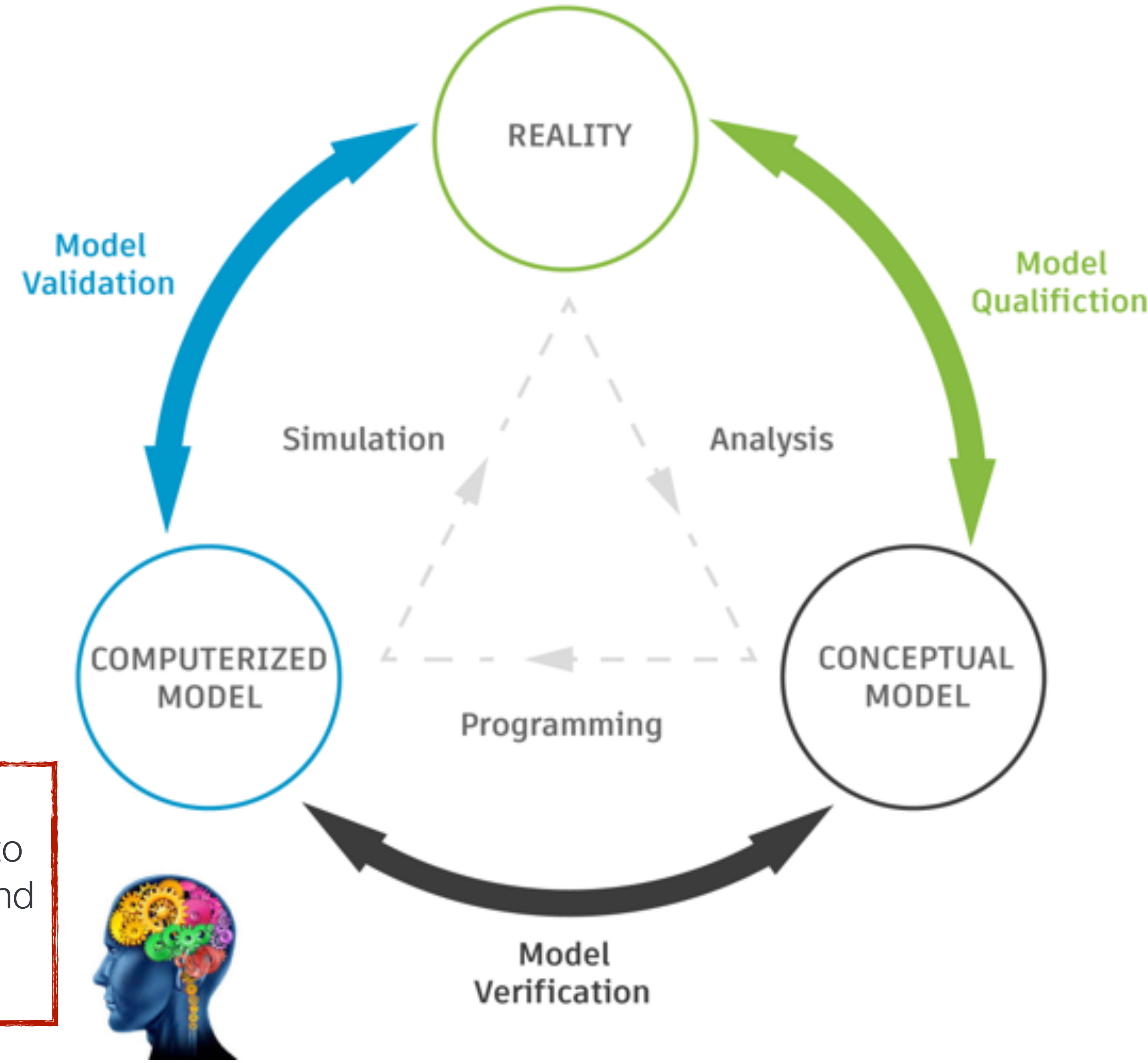
- Uncertainty quantification in a nutshell
- Inlet wind statistics
- Surrogate models
- Diagnostics

STAE TOULOUSE

Uncertainty quantification in a nutshell

How to take into account uncertainties in the process of model validation? What are the levels of confidence in the model outcomes?





Aleatory vs. epistemic uncertainty

Aleatory: Intrinsic variability present in the system that is irreducible but needs to be characterized

Epistemic: Potential deficiency that is due to a lack of knowledge and that is reducible via more investigations

Uncertainty quantification in a nutshell

Main steps

Inputs → Model → Output

$\mathcal{M}$

Step 1
Characterize uncertainties in the inputs

$\mathbf{x} \in \mathbb{R}^d$

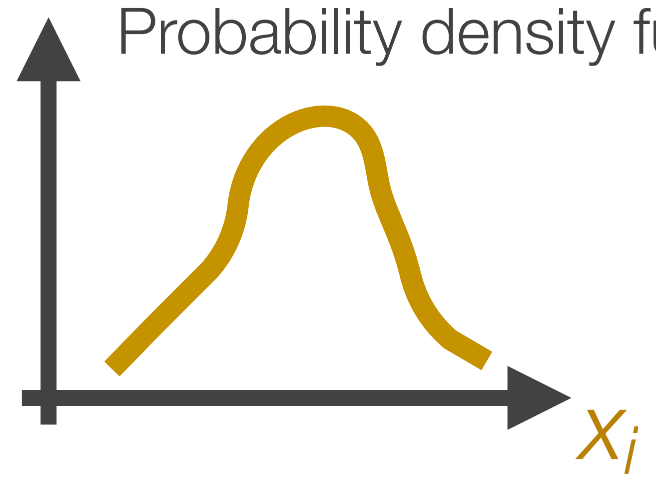
$d >> 0$

$\mathbf{y} = \mathcal{M}(\mathbf{x}) \in \mathbb{R}^M$

- Based on available information: direct or inverse methods (inference, calibration)
- Identification of the *d* explicit and hidden parameters of the computational model
- Characterization of the associated level of knowledge

Interval

Probability density function (PDF)



Uncertainty quantification in a nutshell

Main steps

$\mathcal{M}$

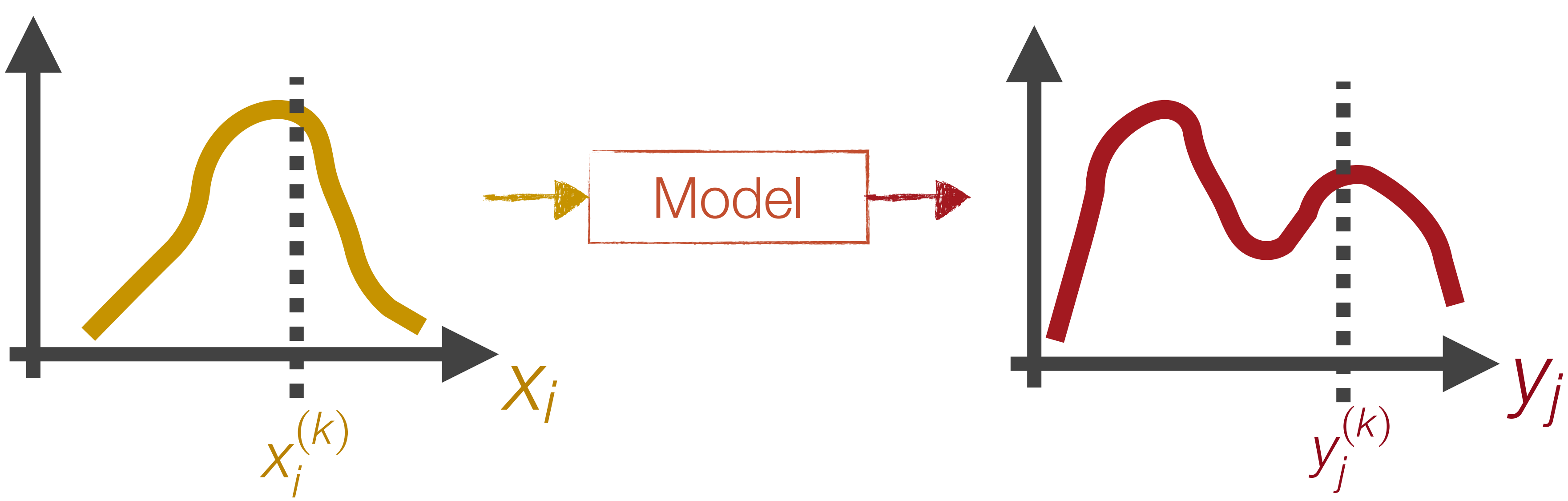
$\mathbf{x} \in \mathbb{R}^d$

$d >> 0$

$\mathbf{y} = \mathcal{M}(\mathbf{x}) \in \mathbb{R}^M$

Non-intrusive methods
- The model is used as a black box
- Need to define a training set to approximate the model response

Step 2: Perform simulations while accounting for the identified uncertainties



Uncertainty quantification in a nutshell

Main steps

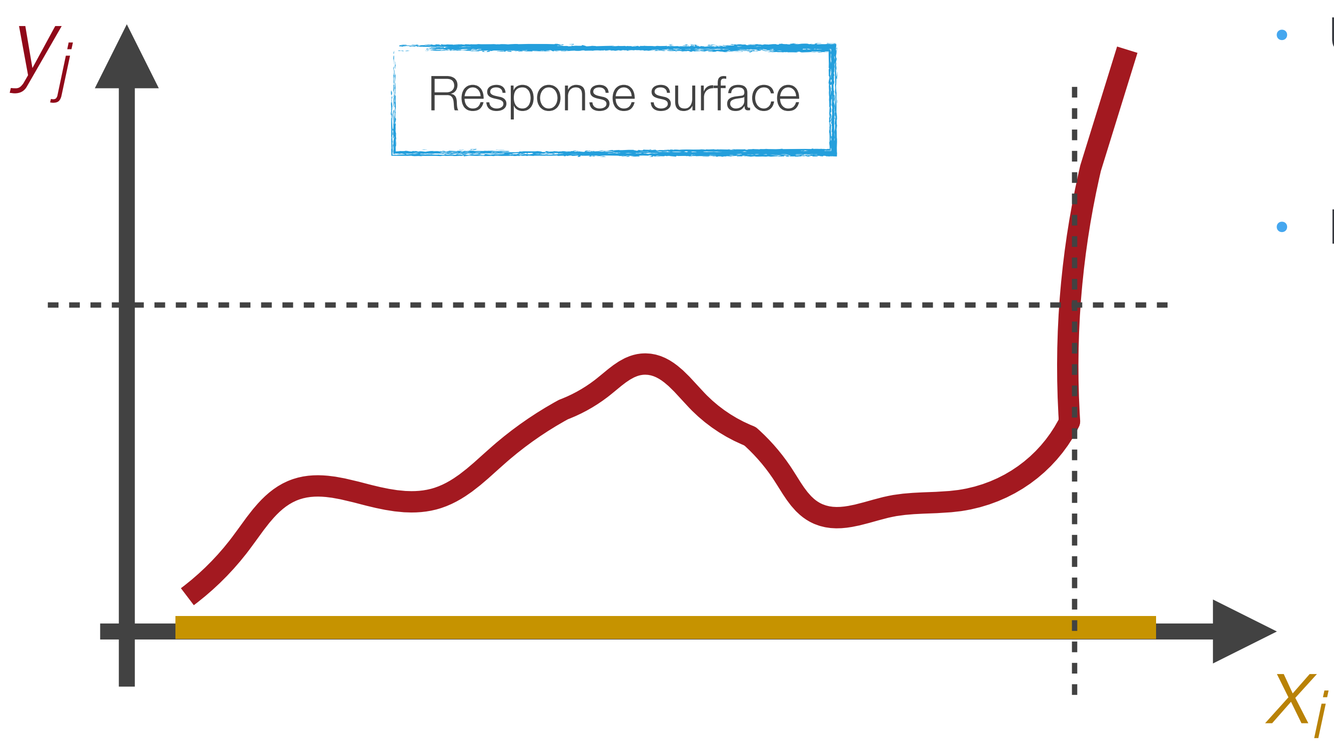
Inputs → Model → Output

$\mathcal{M}$

$\mathbf{x} \in \mathbb{R}^d \quad \longleftrightarrow \quad \mathbf{y} = \mathcal{M}(\mathbf{x}) \in \mathbb{R}^M$

$d >> 0$

Step 3: Investigate the connection between the inputs and outputs of the model



- Understand model structure
  - ➡ How the variability of the model response is affected by the variability of each input variable?
- Rank the input parameters by order of importance

Sobol' indices (0 < Si < 1)
= contribution of the *i*th input variable to the output variance

Uncertainty quantification in a nutshell

Main steps

Surrogate model

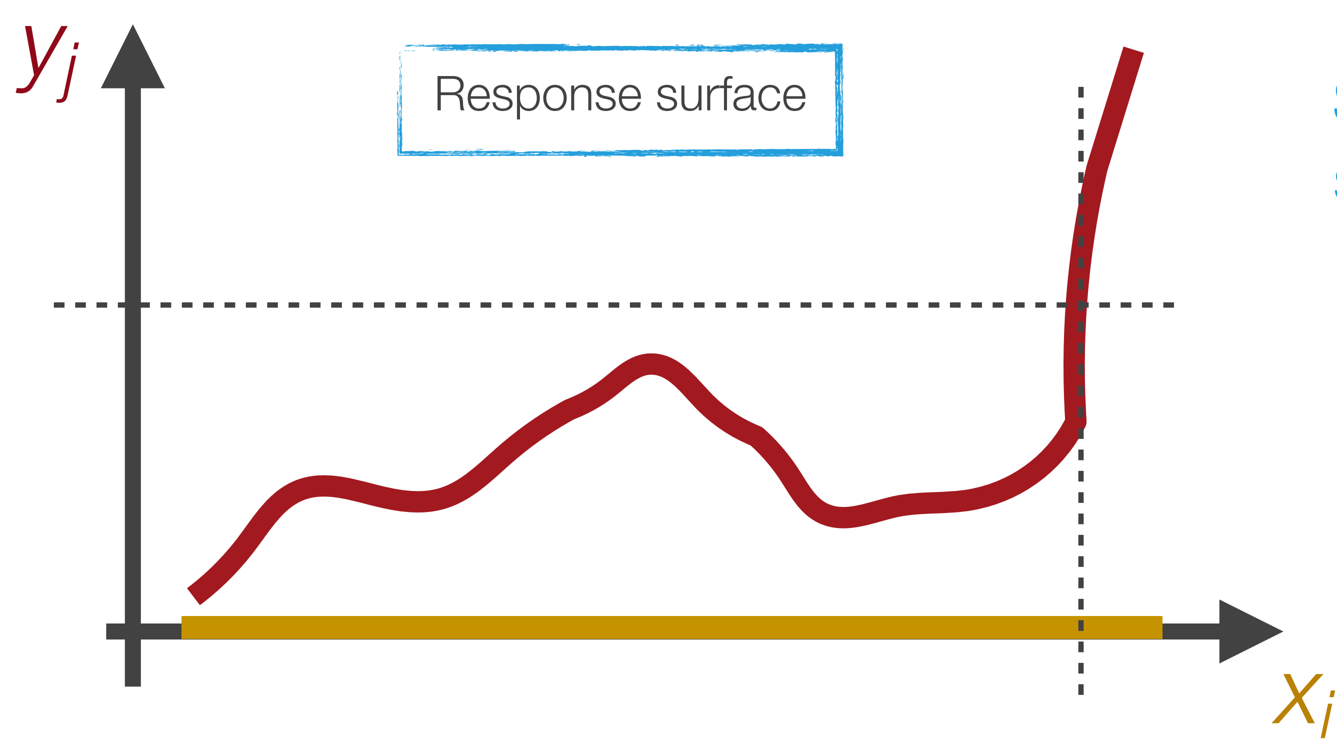
Inputs → ~~Model~~ → Output

$\mathcal{M}$

$\mathbf{x} \in \mathbb{R}^d$

$d >> 0$

$\mathbf{y} = \mathcal{M}(\mathbf{x}) \in \mathbb{R}^M$



Strategy: Build a surrogate using a training set to obtain statistics at reduced cost

- Polynomial Chaos Expansion
- Gaussian Process Surrogate

Uncertainty quantification in a nutshell

Polynomial Chaos Expansion

Multi-variate polynomials

$$\mathrm{y} = \mathcal{M}\left(\mathrm{x}(\zeta)\right) = \sum_{\alpha=0}^{N_p} \boxed{\gamma_\alpha} \boxed{\Psi_\alpha} \left(\mathrm{x}(\zeta)\right)$$

Coefficients to be determined

POLYNOMIAL BASIS

- Choice of the input distribution (uniform ↔ Legendre polynomials)
- Choice of the total polynomial order
- Truncation strategy (full or sparse basis)

$$\boxed{< \Psi_\alpha(\zeta), \Psi_\beta(\zeta) >= \int_\Gamma \Psi_\alpha(\zeta)\, \Psi_\beta(\zeta)\, \rho_{\mathrm{X}}(\zeta)\, \mathrm{d}\zeta}$$

$$< \Psi_\alpha(\zeta), \Psi_\beta(\zeta) >= \delta_{\alpha\beta} \parallel \Psi_\alpha \parallel^2$$

orthogonality

COEFFICIENTS

Galerkin projection

- least-squares problem (linear system)
- spectral projection (Gaussian quadrature)

$$\mathrm{y}^{(k)} = \mathcal{M}(\mathrm{x}^{(k)}), \quad k = 1, \cdots, N_e$$

training set

Uncertainty quantification in a nutshell

Gaussian Process (Kriging)

Gaussian Random process

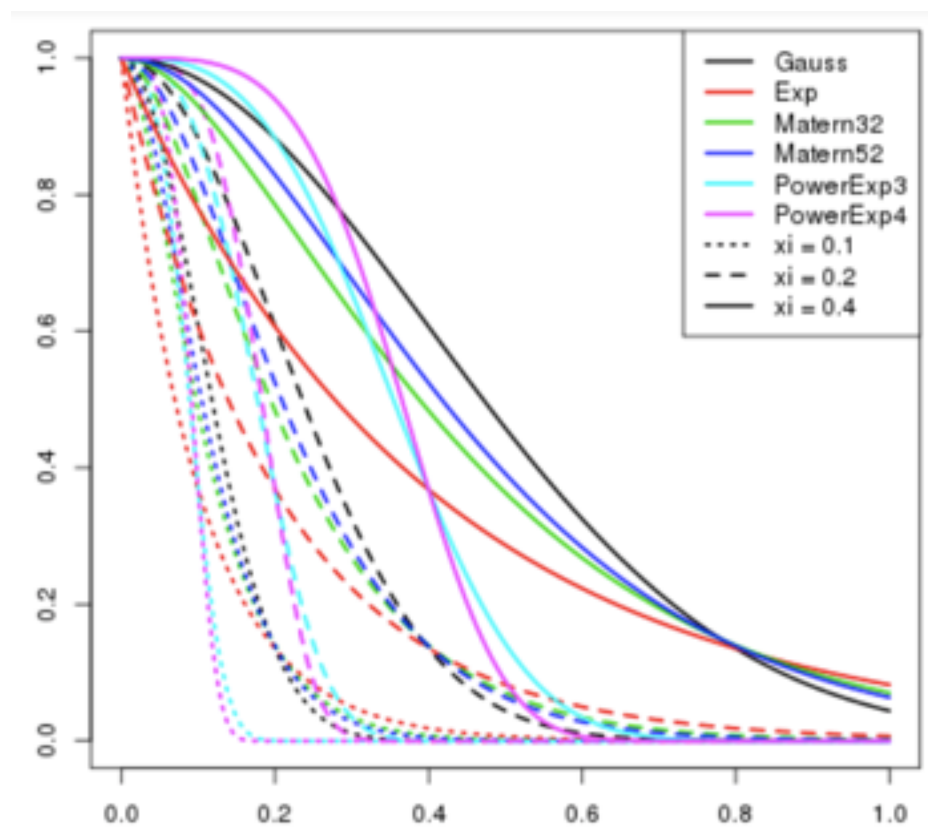
$$y = \mathcal{M}(x) = \sum_{\alpha=1}^{N_p} \gamma_\alpha \boxed{\Psi_\alpha}(x)$$

GAUSSIAN RANDOM PROCESS

Fully characterized by zero mean and correlation structure

- Choice of the correlation structure
- Optimization of the hyper parameters (length scale, variance, …) using maximum likelihood

$$\boxed{\Psi_\alpha(x) \sim \mathrm{GP}\left(0, \sigma_\alpha^2 \pi_\alpha(x, x')\right)}$$



TRAINING SET

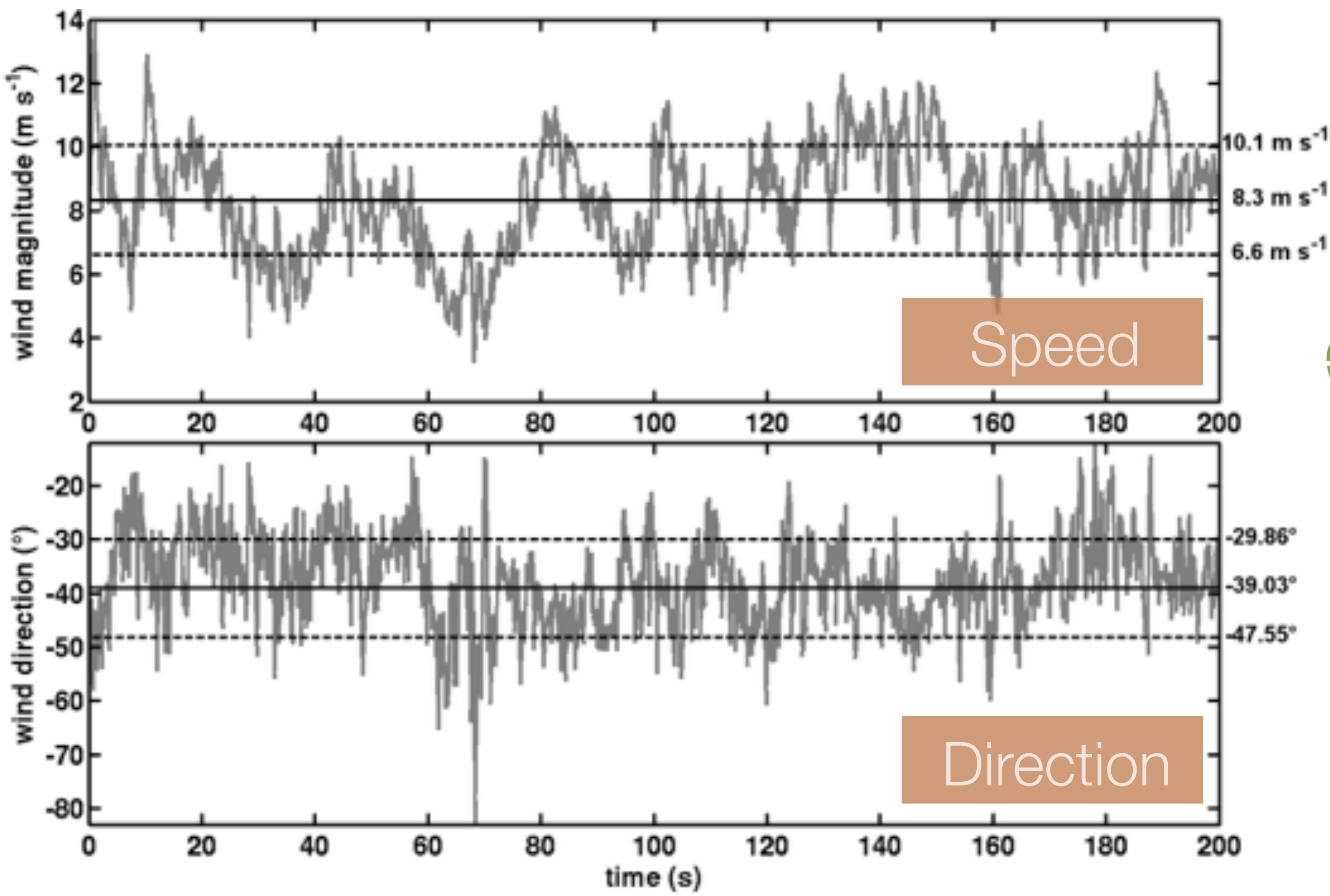
Any finite collection of process values has a joint Gaussian distribution

$$\boxed{y^{(k)} = \mathcal{M}(x^{(k)}), \quad k = 1, \cdots, N_e}$$

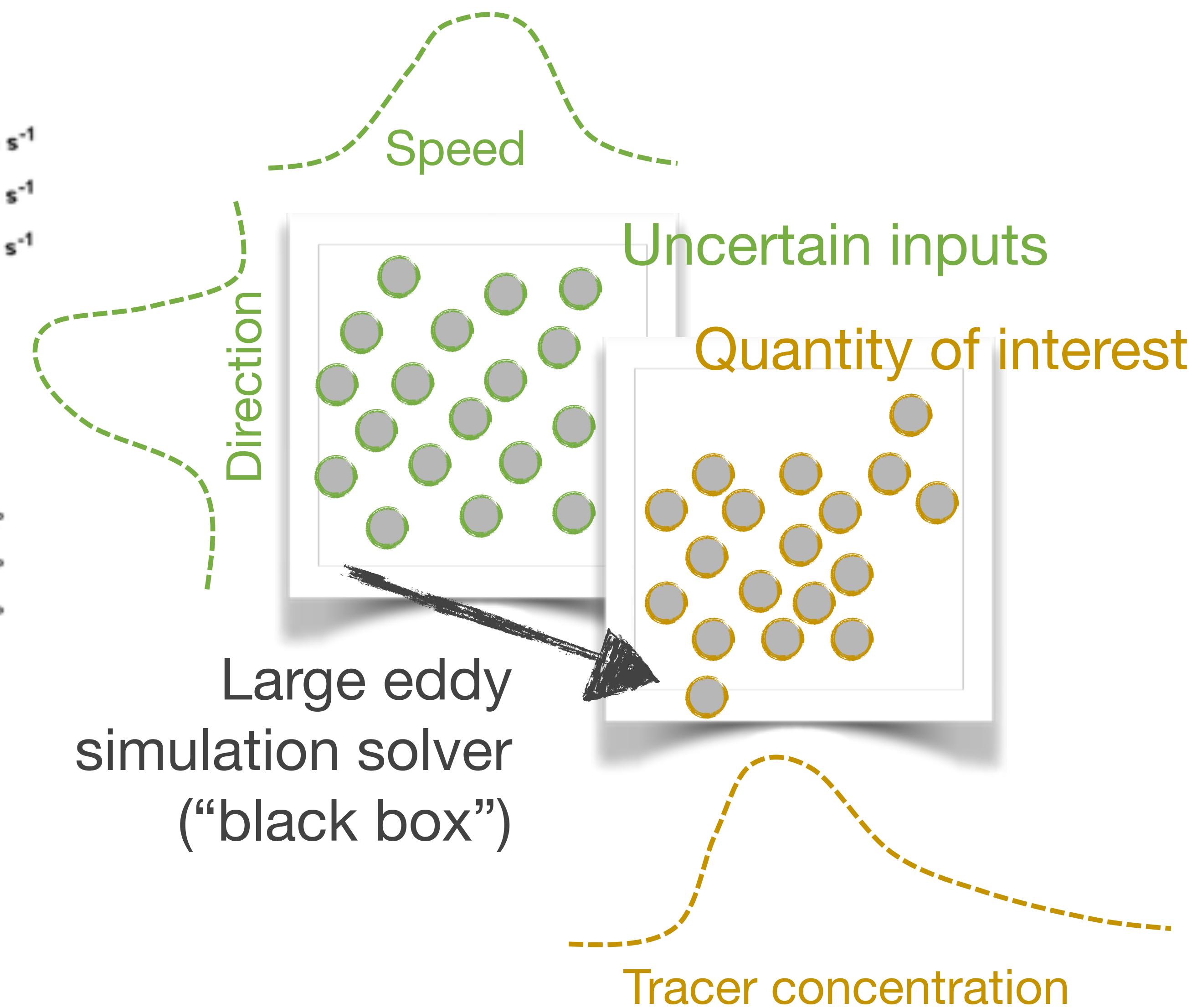
Inlet wind statistics

Ensemble of large eddy simulations to characterize the sensitivity of the numerical predictions to the variability in the inlet wind

INLET WIND VARIABILITY



STATISTICAL APPROACH



Inlet wind statistics

Ensemble of large eddy simulations to characterize the sensitivity of the numerical predictions to the variability in the inlet wind

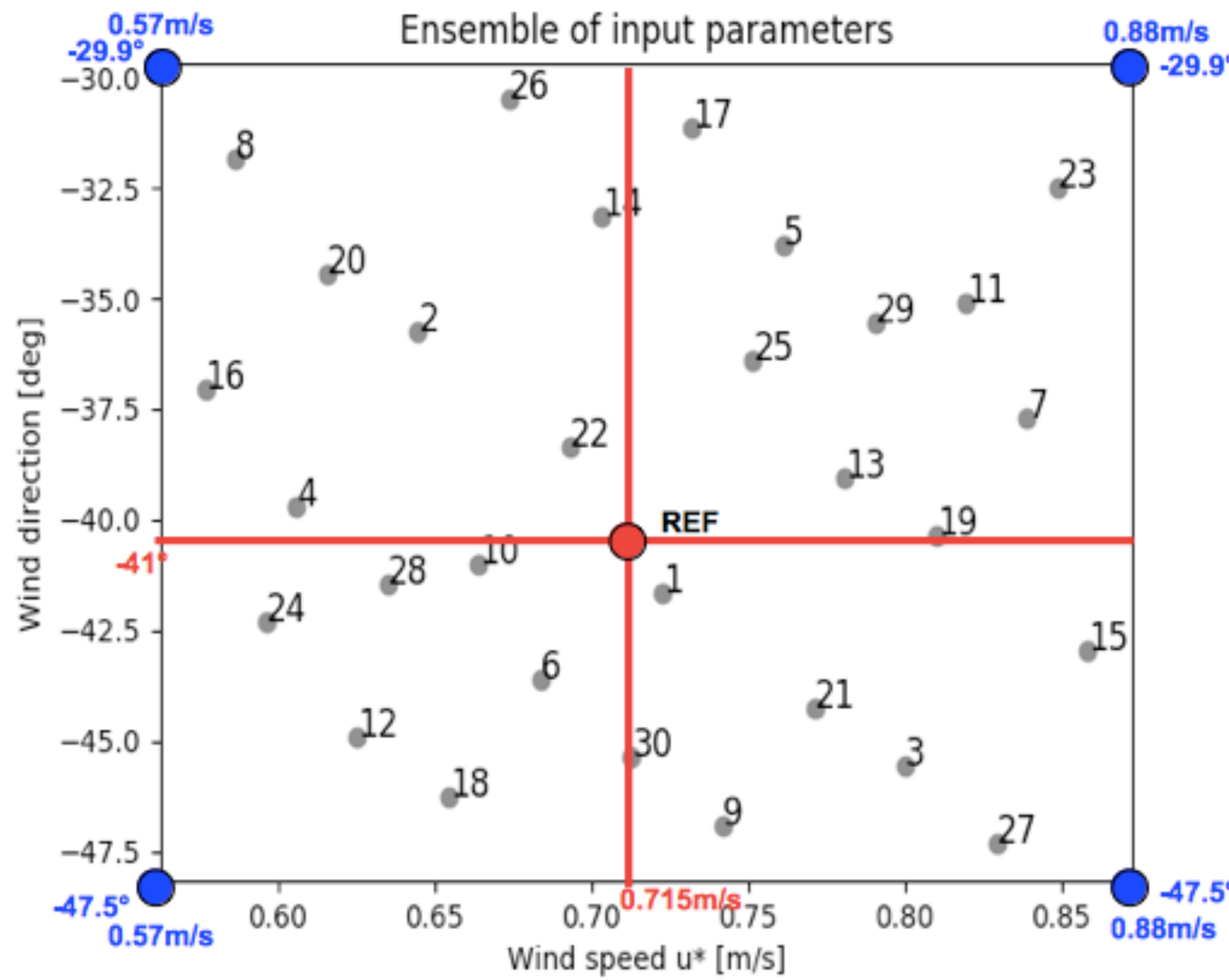
BUDGET FOR TRAINING SET

- 30 YALES2-AE simulations
- Each LES corresponds to a different wind inlet condition.

DESIGN OF EXPERIMENT

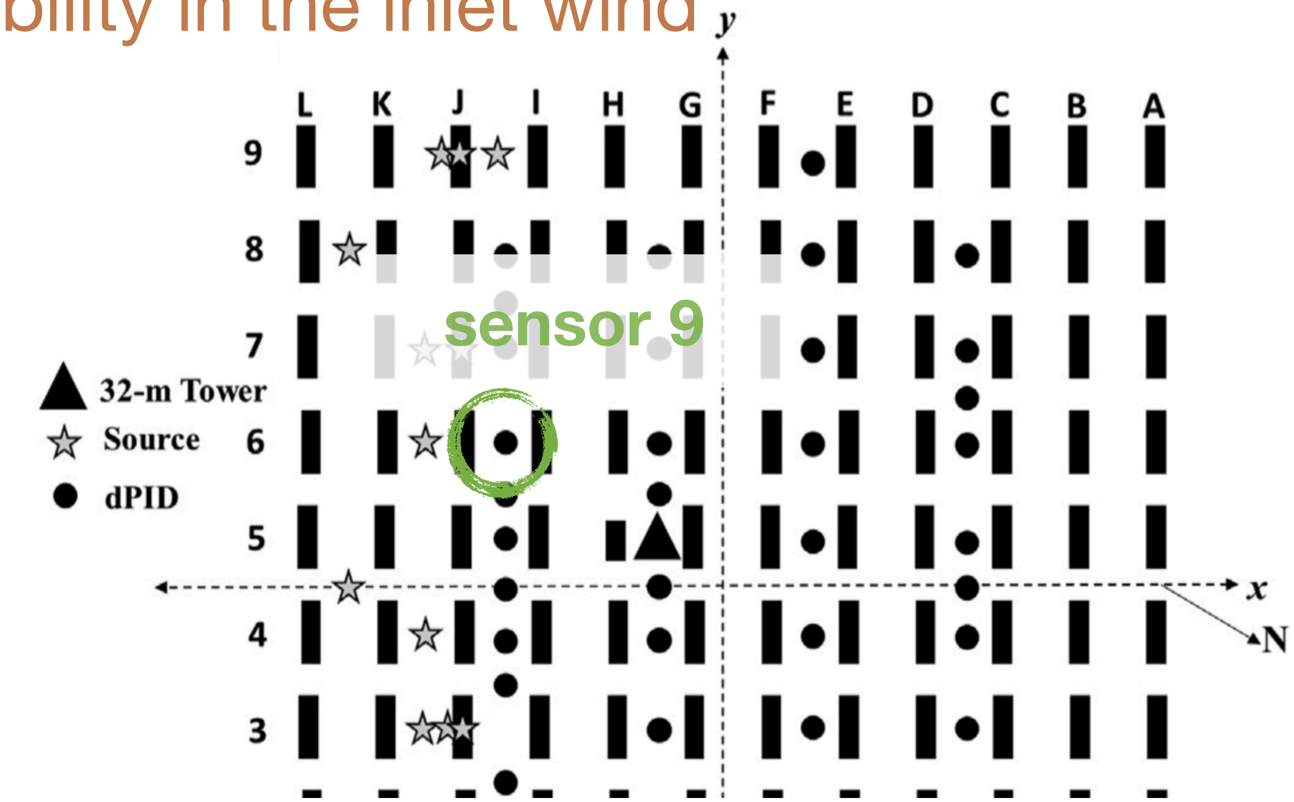
➡ How to select the points of the training set?

- Halton's low discrepancy sequence
- Uniform distributions for the inlet wind speed and direction



Inlet wind statistics

Ensemble of large eddy simulations to characterize the sensitivity of the numerical predictions to the variability in the inlet wind



BUDGET FOR TRAINING SET

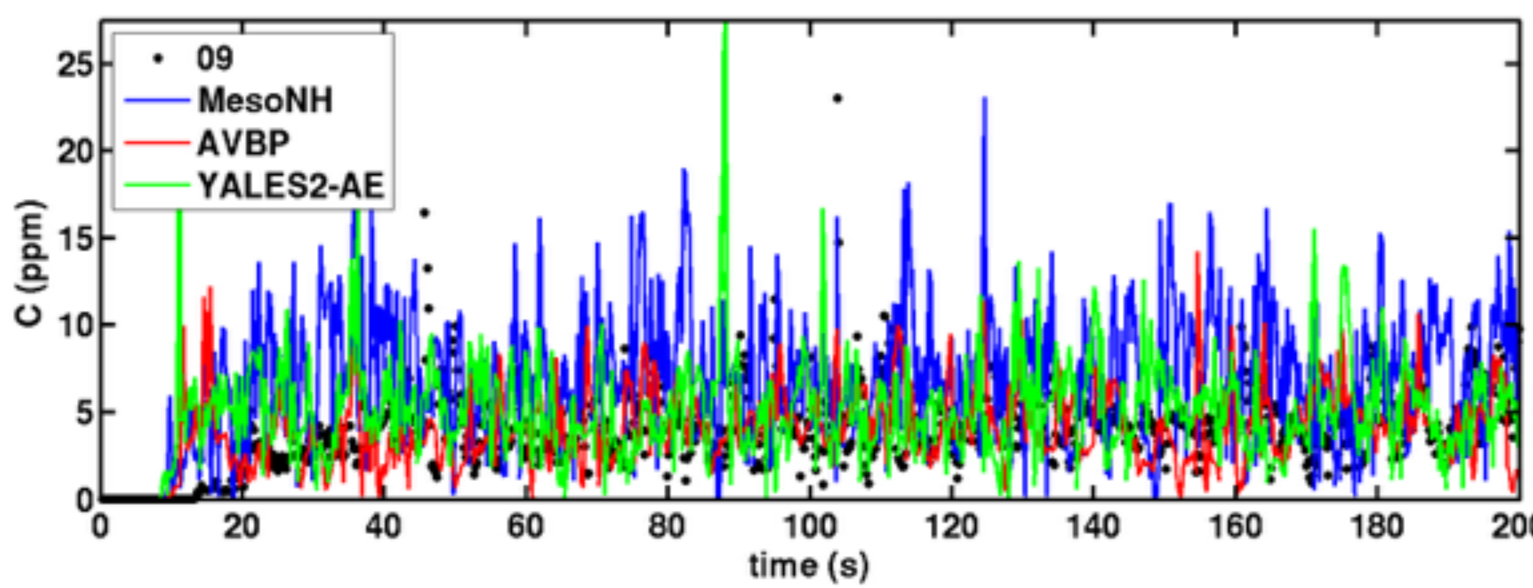
- 30 YALES2-AE simulations
- Each LES corresponds to a different wind inlet condition.

YALES2 GENCI

QUANTITY OF INTEREST

➡ For which variable we need to assess uncertainty?

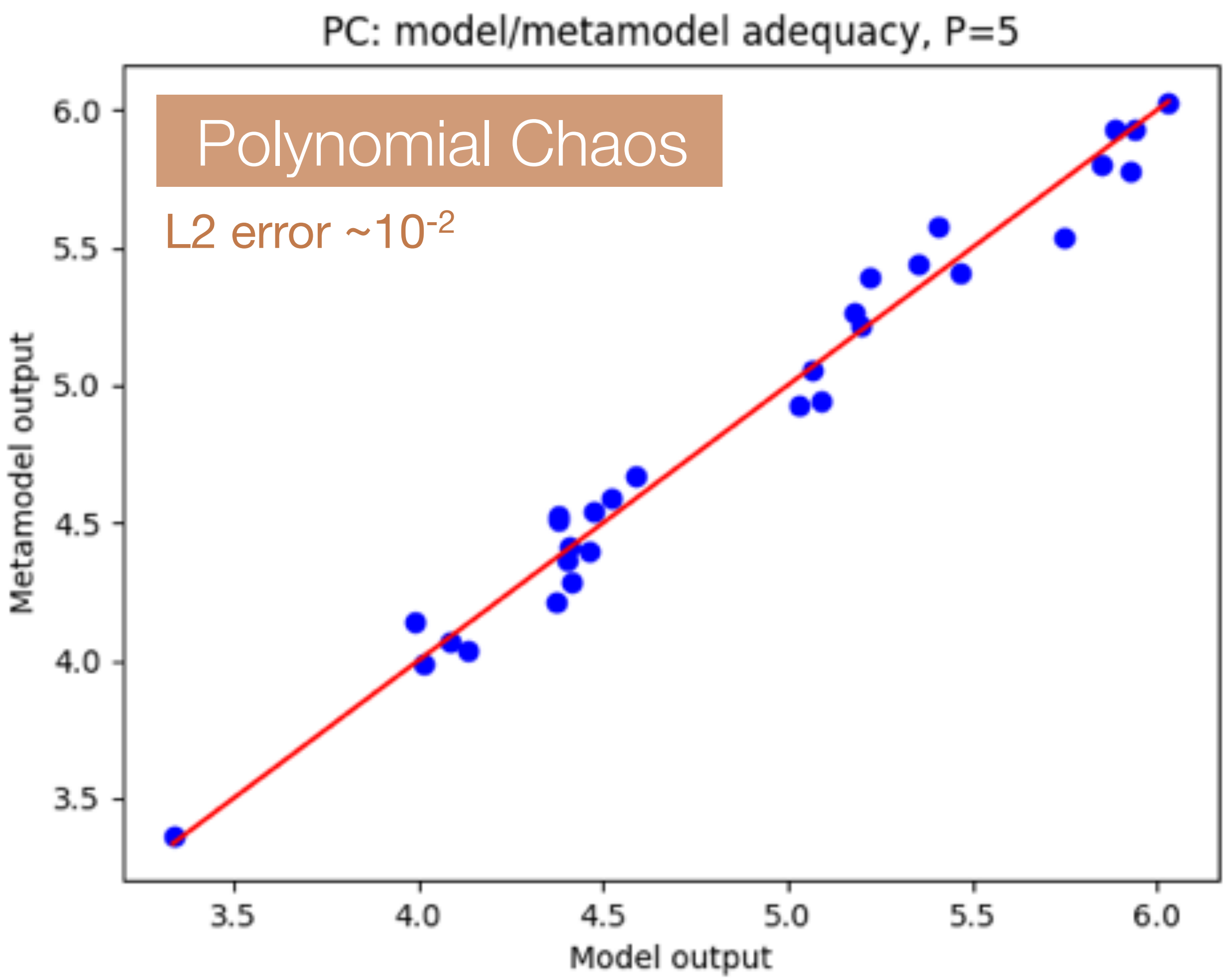
- Mean tracer concentration
- Focus on a given sensor (no. 9)

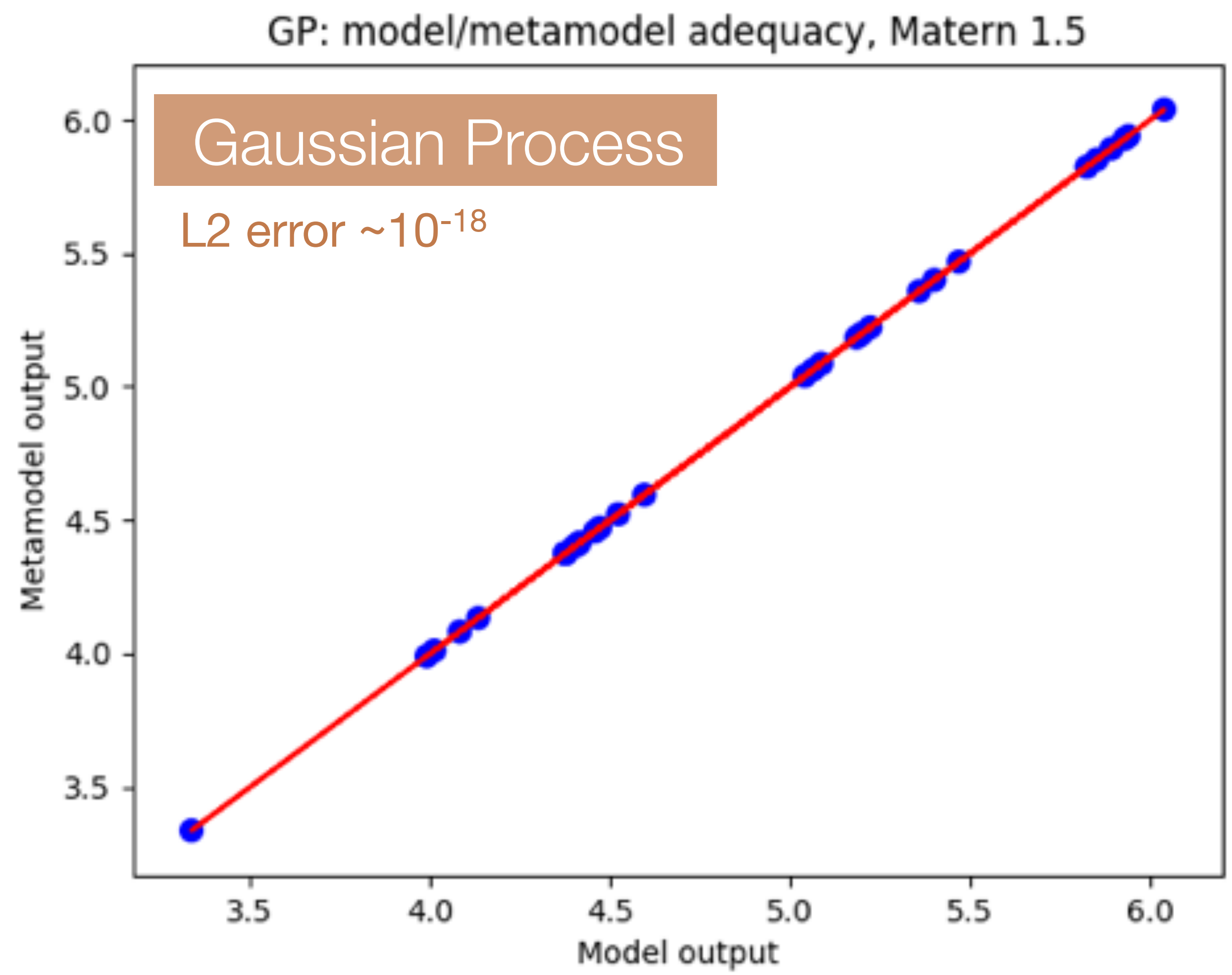


Diagnostics

Model-Surrogate adequacy → Training error

▶ Accuracy with which the surrogate reproduces the experimental design model evaluations





● Training error of the surrogate (evaluated against the 30 YALES2AE simulations)

— Target

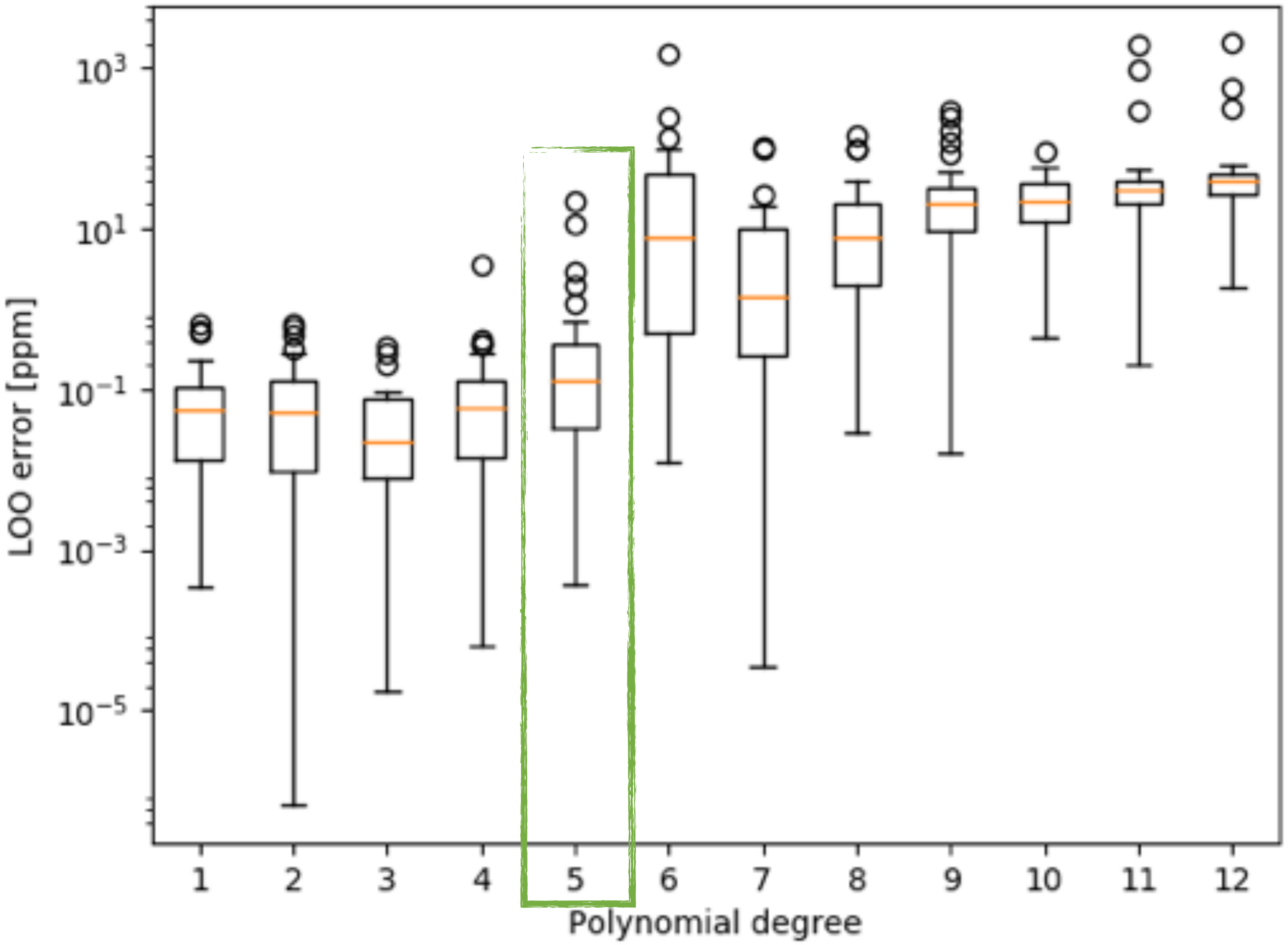
Diagnostics

Model-Surrogate adequacy → Generalization error

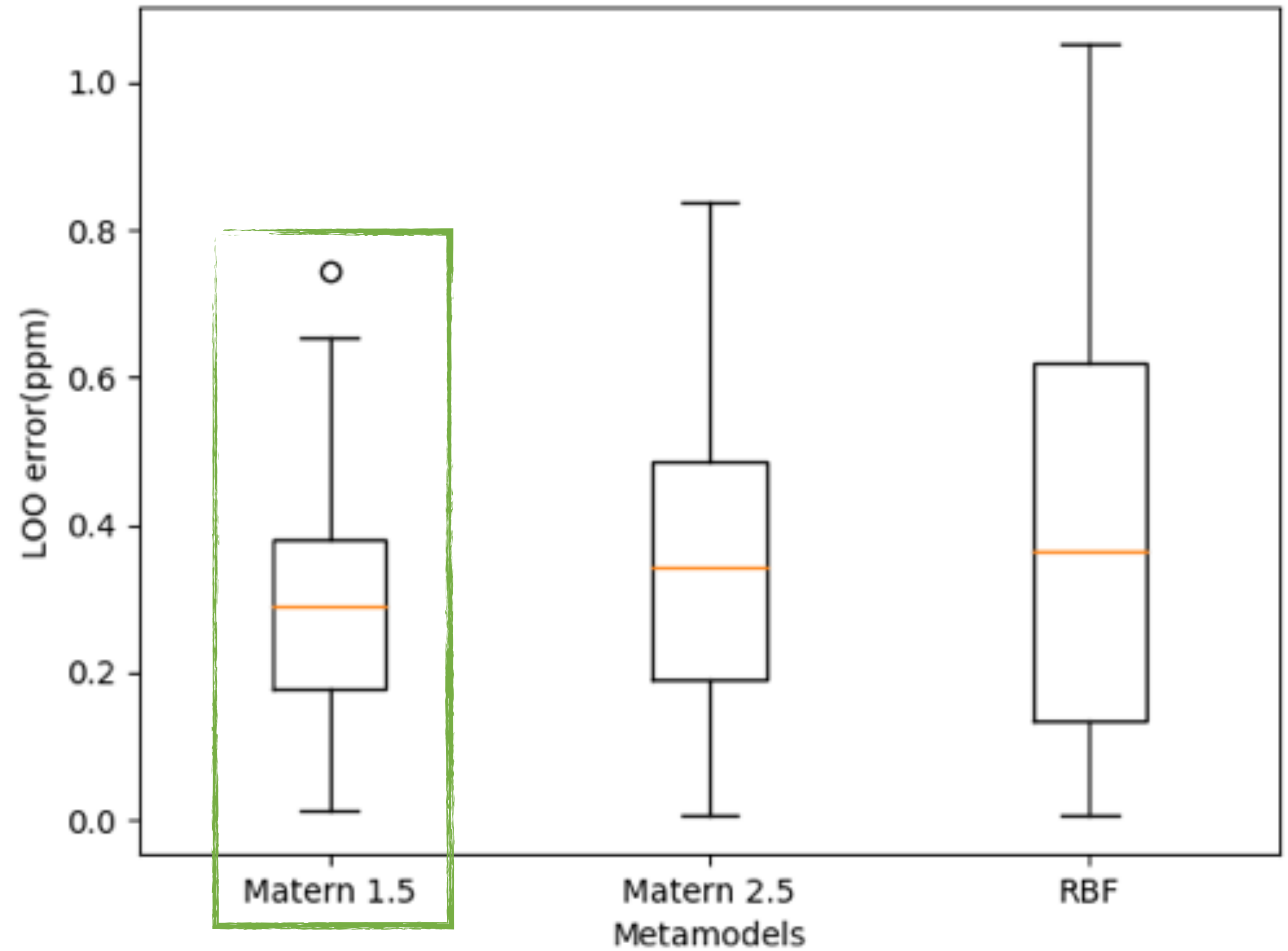
- Cross-validation (Leave-One Out, LOO)

Construction of 30 metamodels, each metamodel using 29 elements of the training set and the remaining element being used for validation

Polynomial Chaos



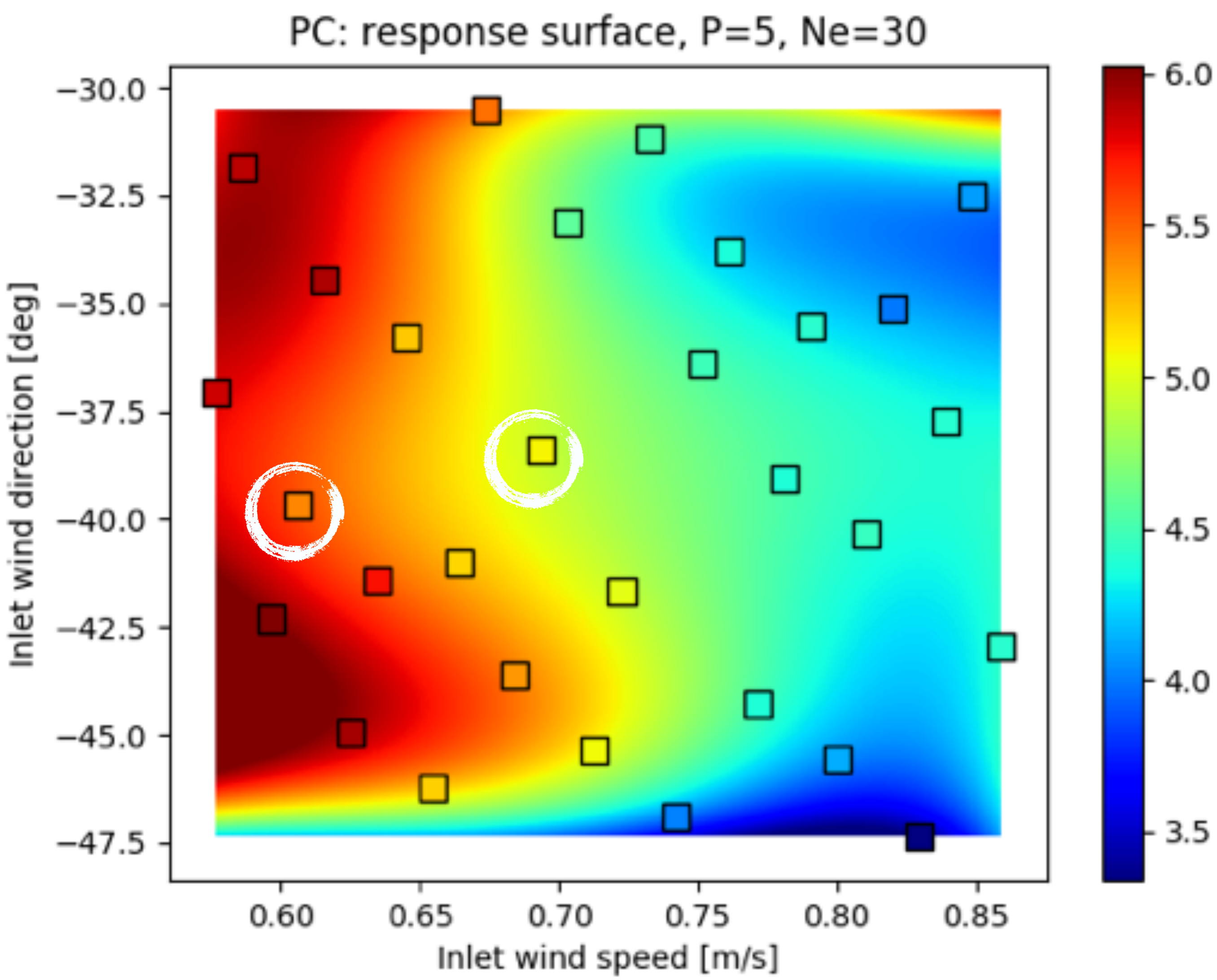
Gaussian Process



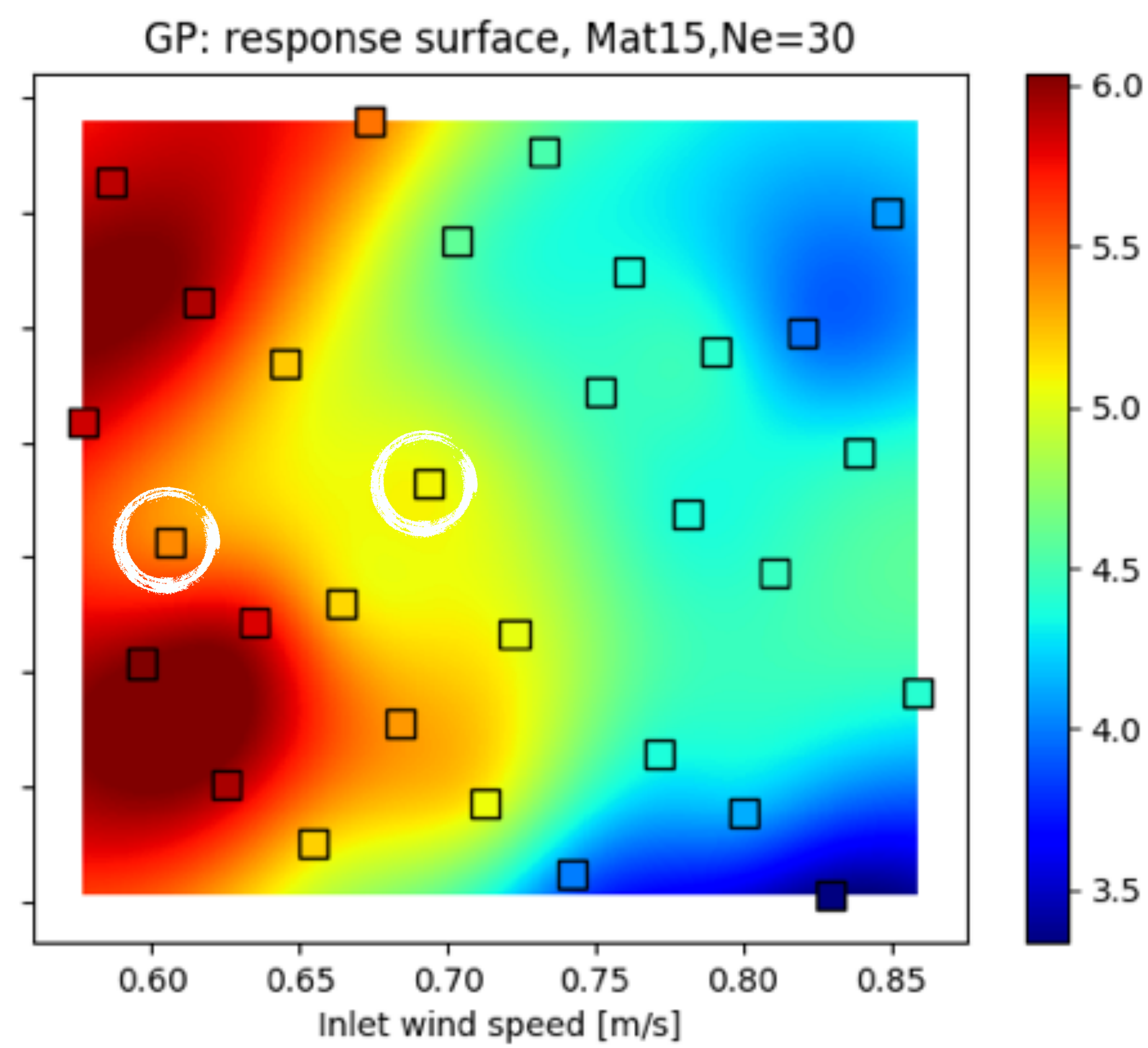
Diagnostics

Response surface

Polynomial Chaos



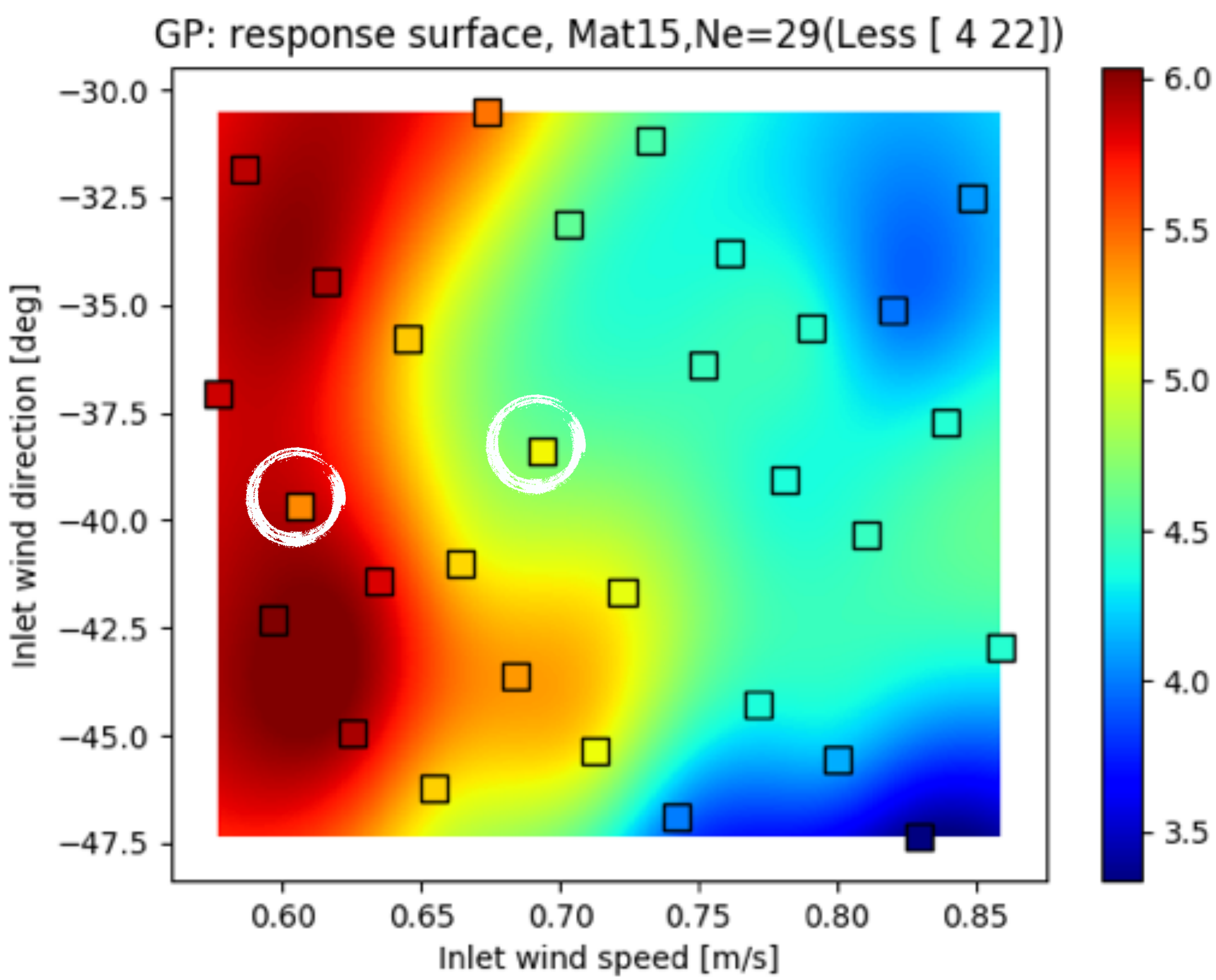
Gaussian Process



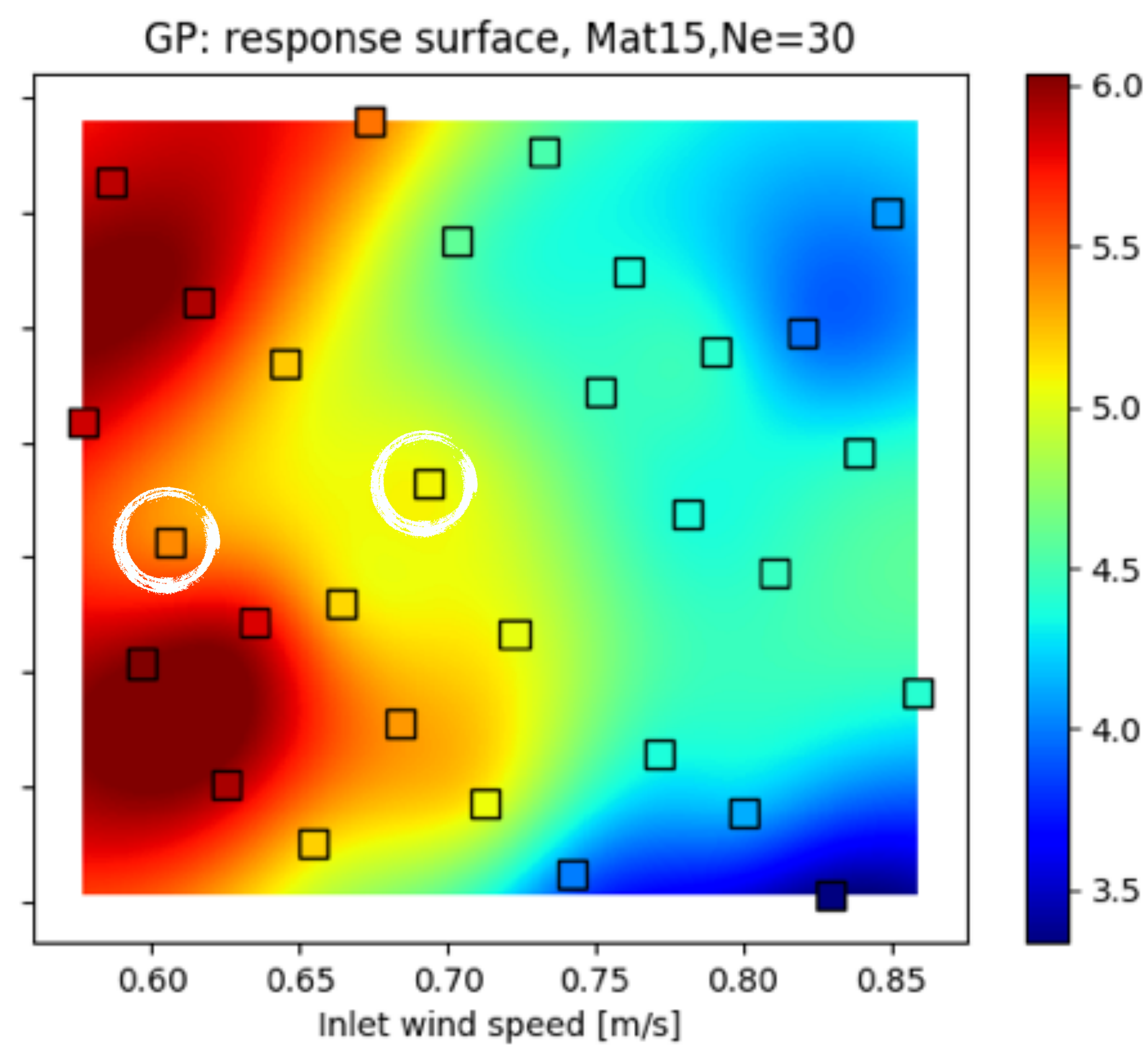
Diagnostics

Response surface

Gaussian Process



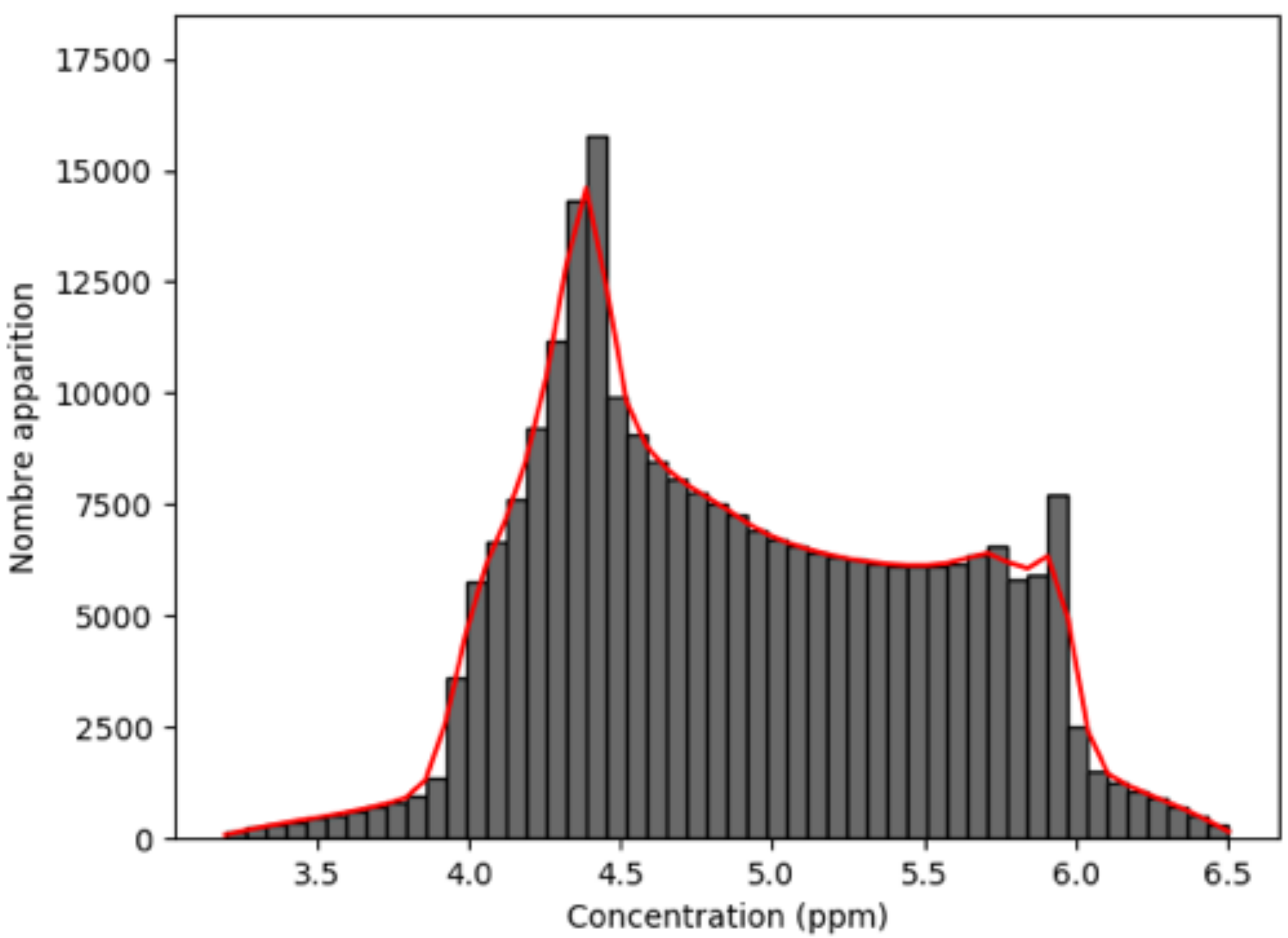
Gaussian Process



Diagnostics

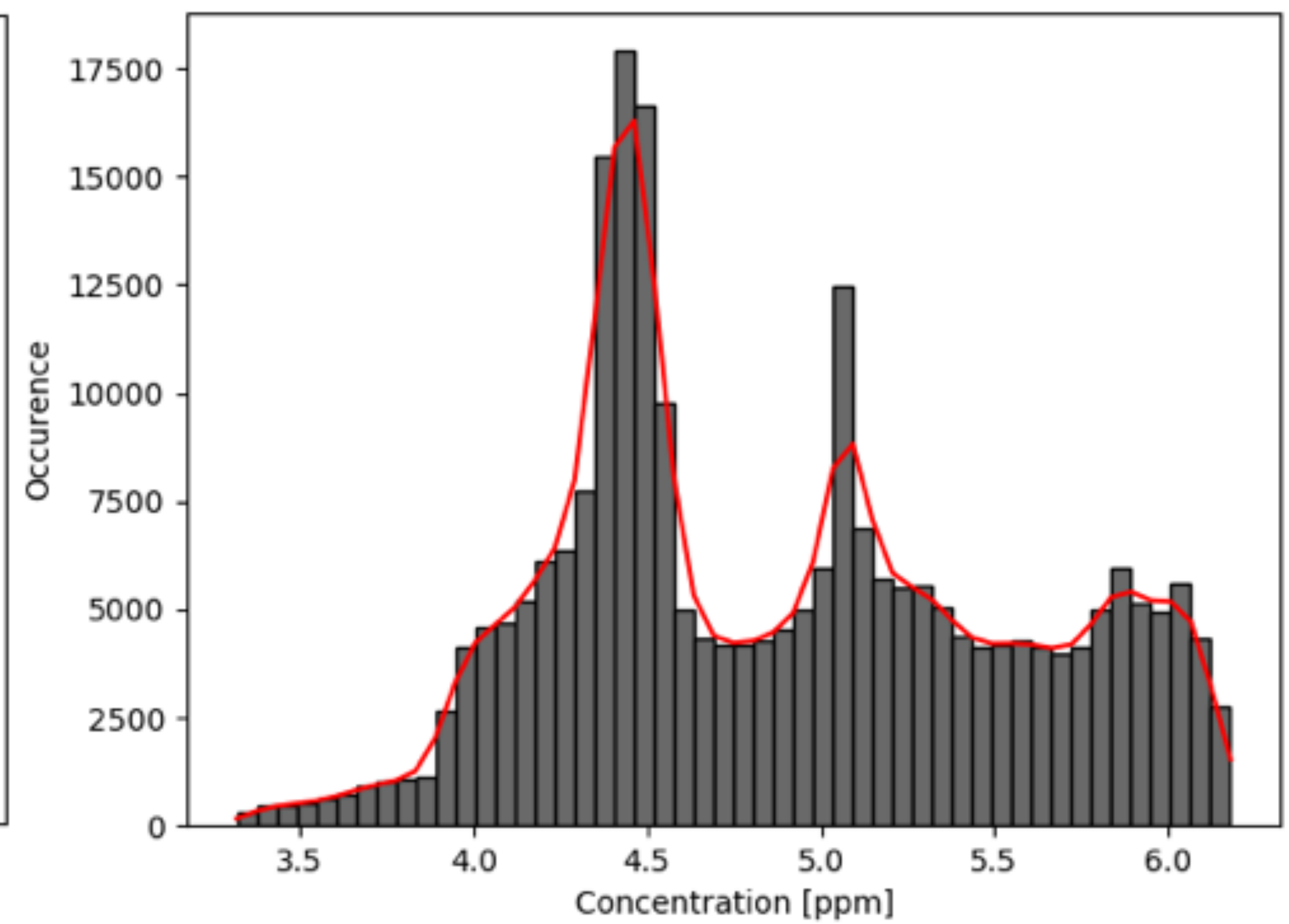
PDF of the mean tracer concentration

Polynomial Chaos



Mean	Standard deviation	Sobol’ index (wind speed)	Sobol’ index (wind direction)
4.9	0.6	83 %	11 %

Gaussian Process



Mean	Standard deviation	Sobol’ index (wind speed)	Sobol’ index (wind direction)
4.9	0.6	88 %	6 %

Conclusions

Added value of large eddy simulations for micro-scale meteorology and air quality

➡ **Metric-scale large eddy simulations:** Evaluation of epistemic uncertainties

- Intercomparison of AVBP, MesoNH and YALES2-AE
- Sensitivity to physical and numerical parameters (computational grid, subgrid-scale model, numerical schemes…)

➡ **Design of suitable surrogate models:** Evaluation of aleatory uncertainties induced by inlet wind conditions

- Intercomparison of Polynomial Chaos and Gaussian Process surrogates
- Sensitivity to inlet wind speed and direction

Rea et al. (in preparation), Atmospheric and Environment - Part 1

Rochoux et al. (in preparation), Atmospheric and Environment - Part 2

STAE TOULOUSE

Perspectives

Added value of large eddy simulations for micro-scale meteorology and air quality

➡ **Metric-scale large eddy simulations:** What is the importance of epistemic uncertainties with respect to aleatory uncertainties?

- Mapping of the epistemic uncertainties estimated through multi-model simulations and of the aleatory uncertainties
- Quality of inlet wind conditions (meso/micro-scale)

➡ **Design of suitable surrogate models:** Improve the quality and reduce the cost of building surrogates

- Identification of critical points in the design of experiment
- Accounting for epistemic uncertainties in the construction of the surrogates

Rea et al. (in preparation), Atmospheric and Environment - Part 1

Rochoux et al. (in preparation), Atmospheric and Environment - Part 2

STAE TOULOUSE


Thank you for your attention.
Any question?
Contact
Melanie.Rochoux@cerfacs.fr
© P. Crombette