Quantification and reduction of epistemic uncertainties in flow simulations: tackling the turbulence modeling dilemma

Paola CINNELLA

Laboratoire DynFluid, Arts et Métiers ParisTech

in collaboration with:

R. Dwight, TU Delft, NL,

W.N. Edeling, CWI, NL,

G. Iaccarino, CTR, Stanford, USA

M. Schmeltzer, TU Delft, NL

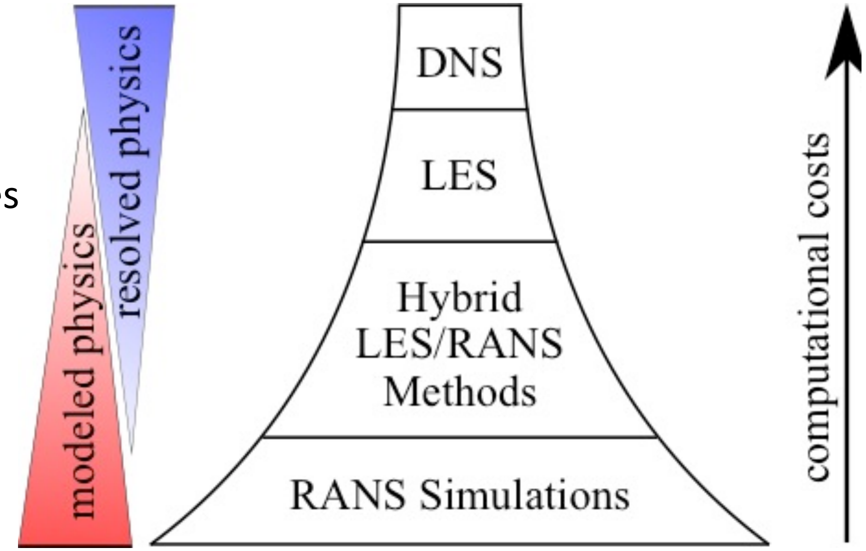
Séminaire ARISTOTE, Mécanique déterministe ou Incertitudes,
Ecole Polytechnique, 21 Février 2019

Talk overview

- Turbulence modeling uncertainties
- RANS models: what is wrong?
    - Uncertainties in RANS models
- Tackling parametric uncertainties
- Tackling structural uncertainties
- Numerical results
- Conclusions and future trends

Hierarchy of turbulence models

- Navier-Stokes equations contain all the necessary information:
  - brute force option: all scales resolved
    → **Direct numerical simulation**
    computationally <u>intractable</u> for most practical cases
  - alternatives:
    several levels of approximation,
    according to the amount of resolved/modelled scales
    LES→ RANS/LES→RANS
- More resolved scales
  → high cost, high sensitivity to
  numerical errors & inflow conditions
- More modeled scales
  → lower cost, more flow-dependent,
  uncertain models



RANS will remain the **workhorse for fluid simulations** of engineering interest for many years to come → from now on, we focus on RANS models

RANS models, what is wrong?

- Reynolds averaged Navier-Stokes (RANS) equations:

$$\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}'$$

$$\nabla \cdot \mathbf{u} = 0$$

$$\frac{\partial \bar{\mathbf{u}}}{\partial t} + \bar{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}} = -\frac{1}{\rho}\nabla \bar{p} + \frac{1}{\rho}\nabla \left(\mu \nabla \bar{\mathbf{u}} \underbrace{- \rho \overline{\mathbf{u}'\mathbf{u}'}}_{\text{Reynolds stresses}}\right)$$

Modeling problems

1. Need to define an average (generally in time)
   ☹ ill posed for some flow configurations
   ☺ *should* be detectable though expert judgement
2. Reynolds stresses $\boldsymbol{\tau}$ need <u>a constitutive law</u>: a **turbulence model**

Sources of RANS uncertainties

- Uncertainties due to the constitutive law
  - Model structure → expert judgment
  - Model constants → imperfectly known/adjustable

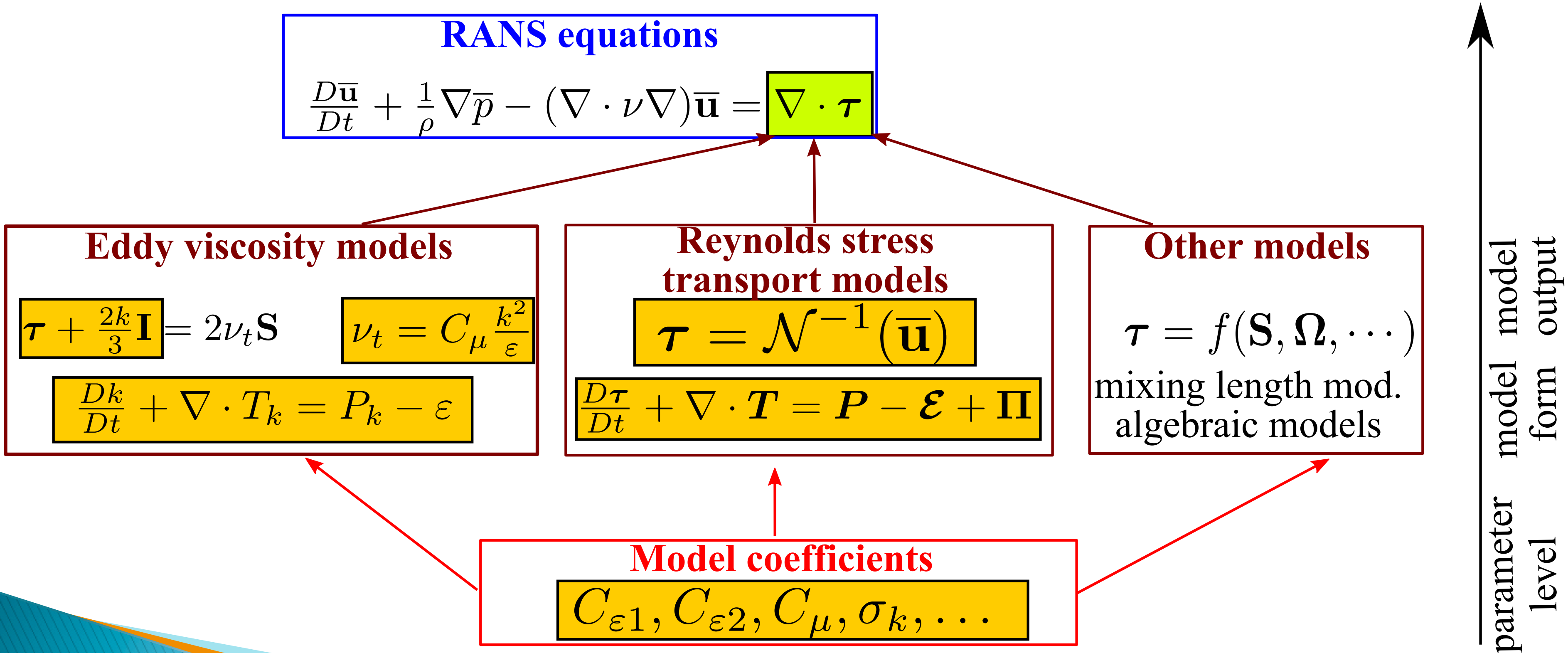
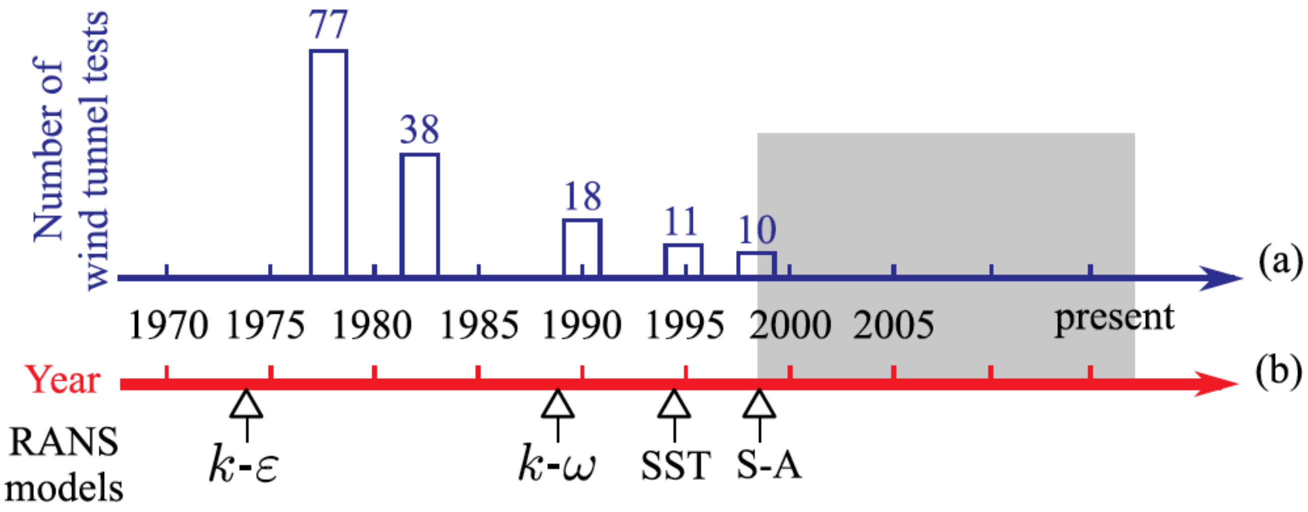


Fig. from Xiao & Cinnella, 2019

RANS closure models

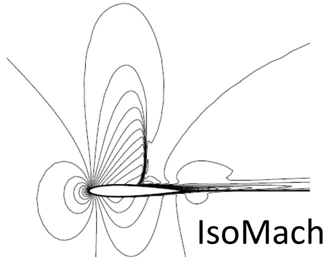
- Rich zoology of models proposed through the years

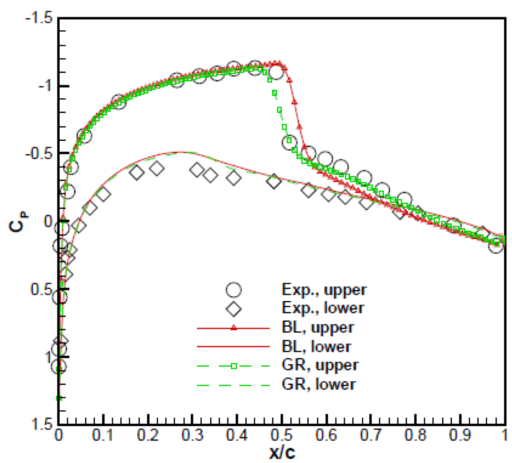


- No universally accepted and valid model yet →pacing item for engineering design
- Model development has been stagnating since the end of the 90's
- Accuracy depends on quantity of interest (QoI)
- Models calibrated for simple flows
  (homogeneous isotropic turbulence, thin shear flows, flat plate boundary layers)

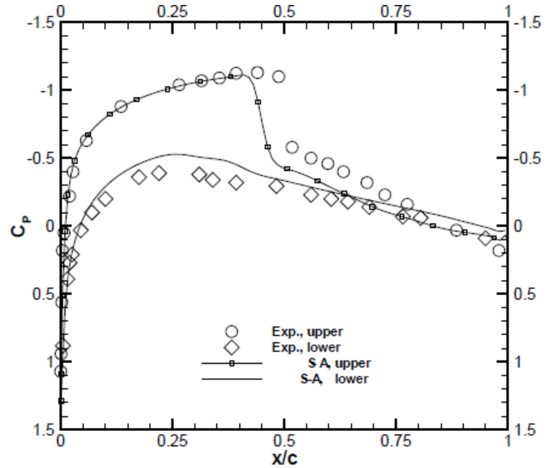
Motivating example

Turbulent flow past an NACA0012 airfoil,
M=0.8, AoA=2°, Re=9e6

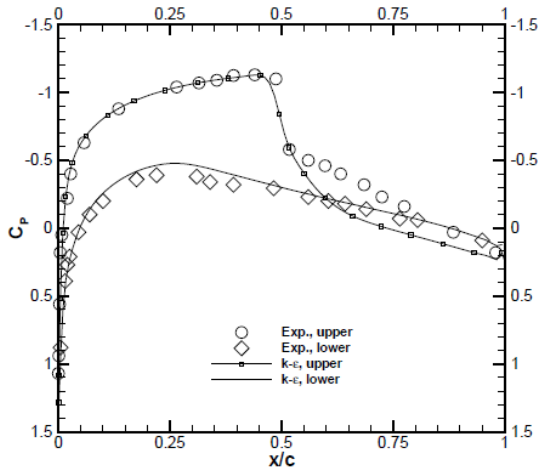




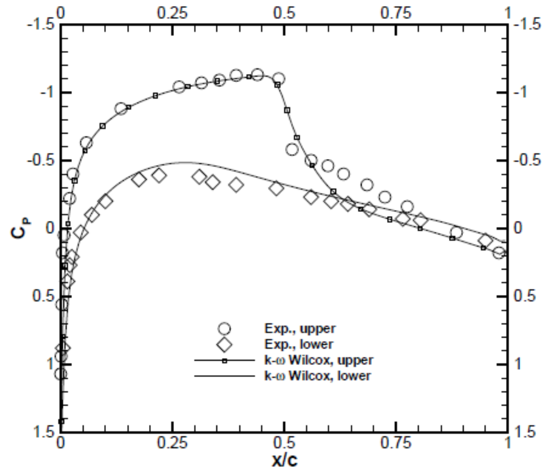
(a)



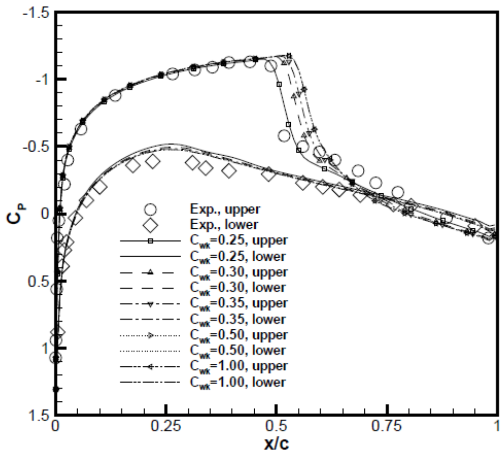
(b)



(c)



(d)



Wall pressure coefficient, BL model, sensitivity to a single closure coefficient (out of 7)

← Wall pressure coefficient, various models

Objectives

- "Improve" RANS models
  - More accurate prediction (not always possible)
  - **Reliability information** (how much wrong am I?)
    - Identify flows/flow regions with high RANS uncertainty
    - Quantify the uncertainty

- Two main approaches

1. Model structure is not modified → parametric approach
   - Update model parameters
   - Bayesian mixture of models
2. Model structure is modified → non parametric approach
   - « Data-driven » modelling, machine learning approaches

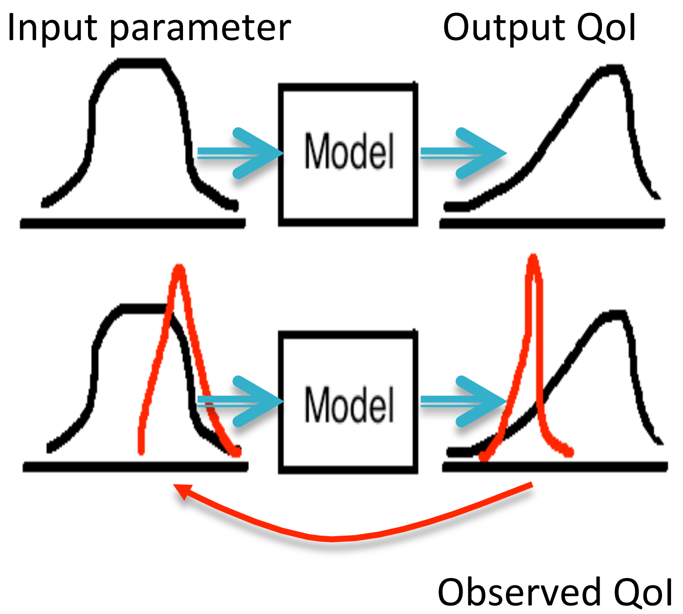
Talk overview

- Turbulence modelling uncertainties
- RANS models: what is wrong?
  - Uncertainties in RANS models
- Tackling parametric uncertainties
- Tackling structural uncertainties
- Numerical results
- Conclusions and future trends

Tackling parametric uncertainty

- Goal: quantifying and reducing uncertainties related with the choice **of closure coefficients**

Direct ou Backward UQ



- Direct UQ
  - Sensitivity analysis of a QoI to model parameters
    [Platteuw et al. 2008, Roy & Oberkampf 2010, Edeling et al., 2013, Margheri et al. 2014]
  - Preliminary information on coefficient pdf required
- Backward UQ
  - Infer posterior pdf of model coefficients from available observations
    [Cheung et al., 2011; Kato & Obayashi 2013; Edeling et al. 2013; Margheri et al. 2014…]

→Solve an inverse probabilistic problem

Backward step: Bayesian inference

- Represents model uncertainty as a probability distribution
- Well suited for scarce/noisy data
- Provides a probability distribution (*pdf*) of the closure coefficients → estimate + measure of confidence in estimate

→ All uncertainties treated in terms of probability, including epistemic

Model calibration relies on Bayes theorem for conditional probability:

$$p(\theta|\mathbf{z}) = \frac{p(\mathbf{z}|\theta)p(\theta)}{p(\mathbf{z})}$$

where $p(\vartheta)$ is the prior belief about parameters ϑ;

$p(\mathbf{z}|\vartheta)$ is the likelihood of observing the data $\mathbf{z}$ given the parameters;

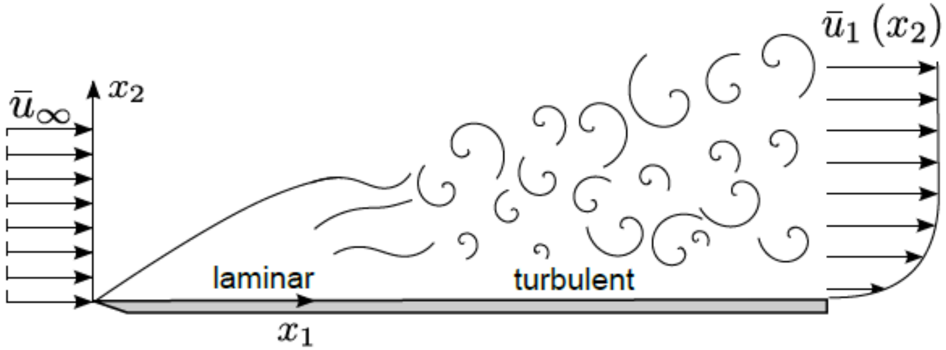
$p(\mathbf{z})$ is the evidence of the data can generally be treated as a normalization constant

→ **OUTCOMES**:

- Probability distributions for the parameters (posterior pdf)
- Probability distribution for unobserved quantities (predictive posteriors)

Exemple: turbulent flow over a flat plate

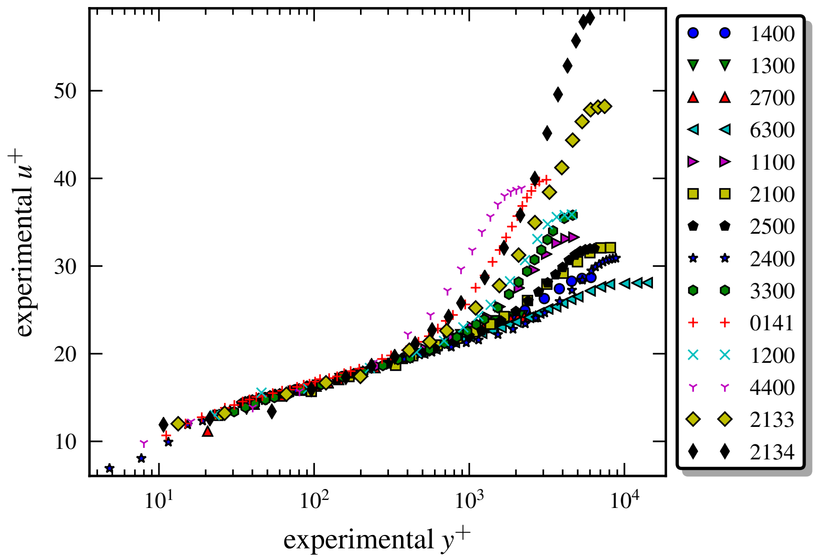
Objective: predict velocity profiles developing in the turbulent boundary layer close to the wall



- **Governing equations**: RANS + turbulence model
  - Wilcox' boundary layer code (fast function evaluations)
  - Launder-Jones's (1972) k-ε model

- **Data** : 13 measured velocity profiles

  [Coles & Hirst, 1968]



Calibration setup

- **Likelihood model**: data **z** related to model outcomes $\mathbf{y}(\theta_p)$, function of the parameters via the multiplicative model error η and the observational error *e*, at each measurement point *i*

$$z_i = \eta_i y_i(\theta) + e_i \quad \text{with, e.g., the assuptions}$$

  - $\eta \sim N(\mathbf{1}, K_{mc})$, squared exponential correlation structure
  - $e_i \sim N(0, \lambda^2)$

- Numerical solutions: quick boundary-layer code

- Use Markov-Chain Monte-Carlo method to draw samples from the posterior

$$p(\theta|\mathbf{z})$$

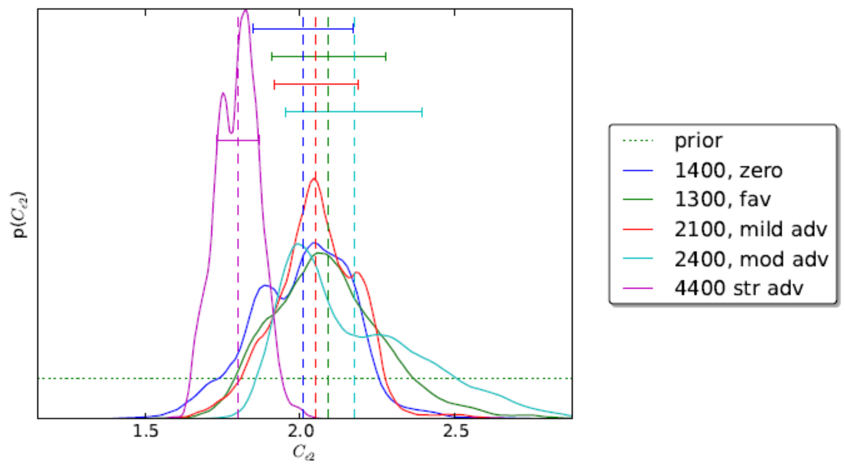
  - Python package pymc https://pymcmc.readthedocs.org/en/latest/#

- Other options: Kalman filters (MAP estimate), particle filters…

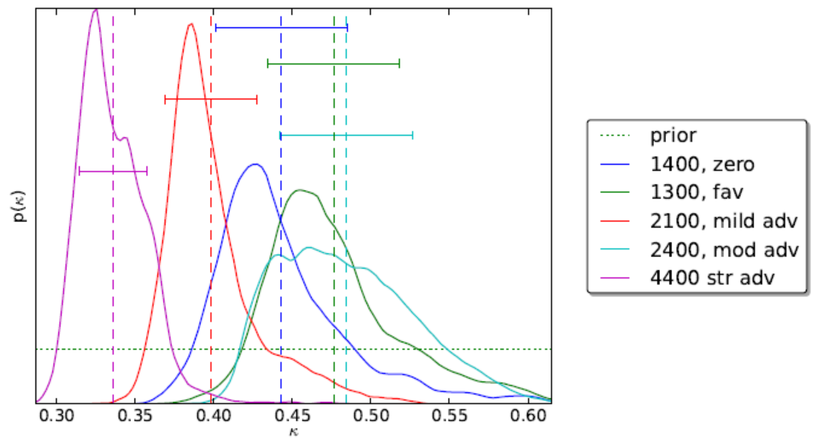
Some results

Sample results for the k-ε Jones-Launder model

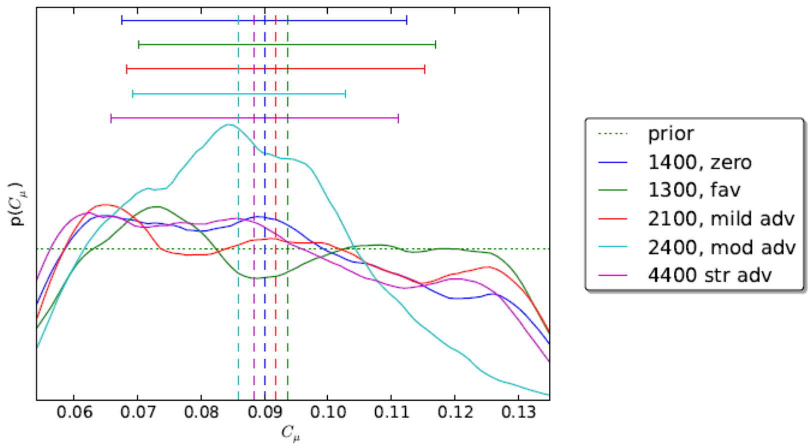
Posterior distributions for $C_{\varepsilon 2}$ for a favorable, zero, mild, moderate and strongly adverse $d\bar{p}/dx$.



Posterior distributions for κ for a favorable, zero, mild, moderate and strongly adverse $d\bar{p}/dx$.

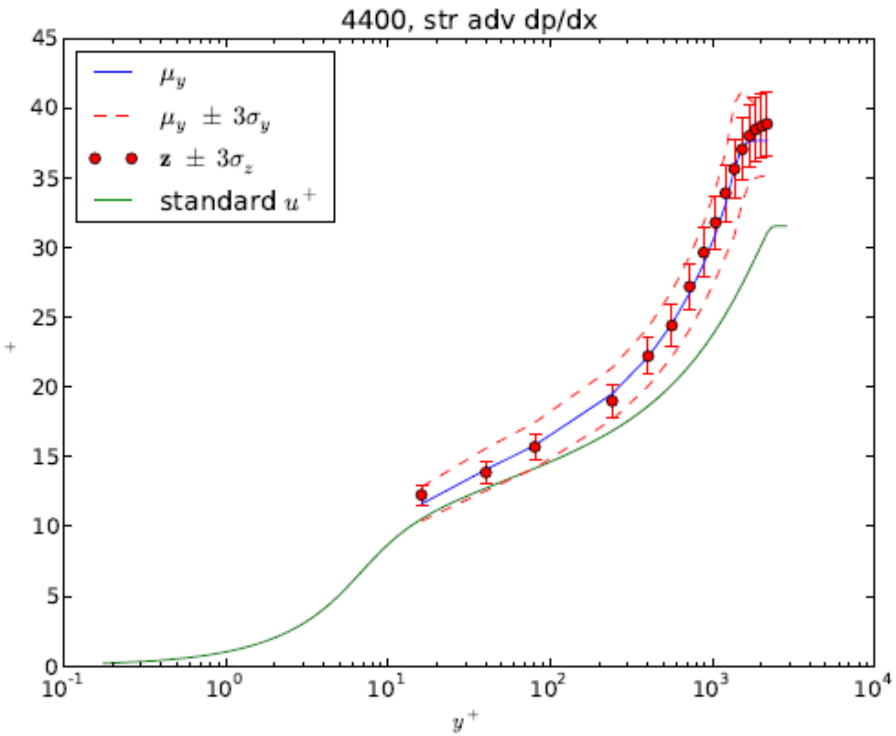


Posterior distributions for C_μ for a favorable, zero, mild, moderate and strongly adverse $d\bar{p}/dx$.

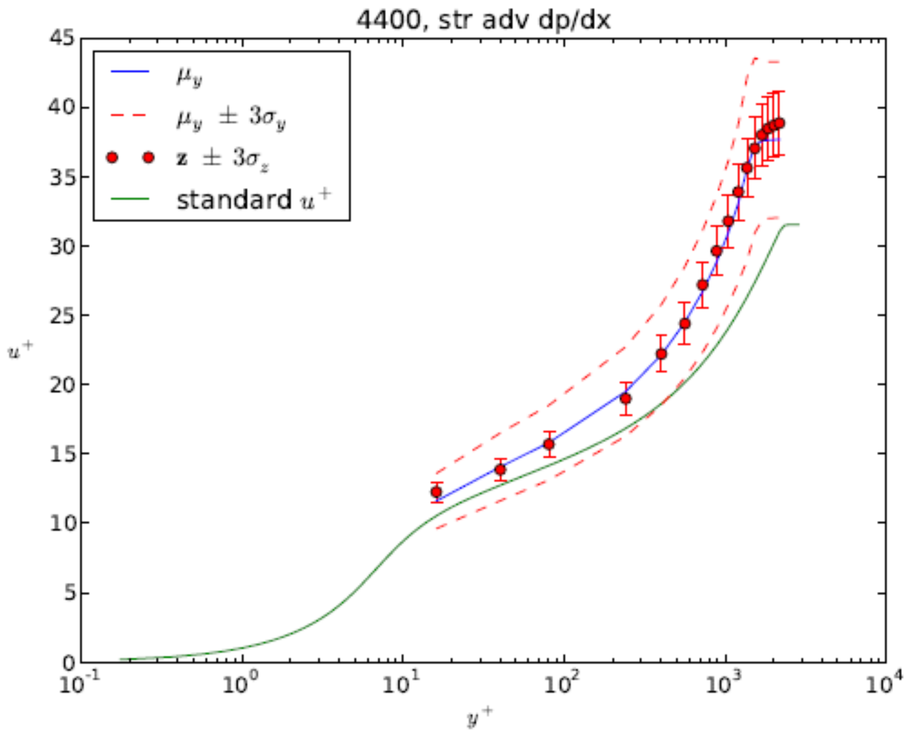


Some results

- Posteriors are *propagated* through the RANS code to get the posterior estimate of the velocity profile
- Samples can also be drawn out of the model indequacy term



Posterior distribution of *y*



Posterior distribution of *ηy*

Lessons learned

1. Coefficients are highly case-dependent
   - This reflects the structural inadequacy of the calibrated model

2. Including a model-inadequacy term partly alleviates overfitting, but it cannot be easily extended to predict new cases or new QoI

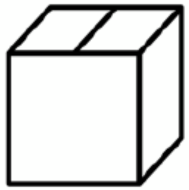
3. How to summarize the effect of both parametric and model-form uncertainty to make predictions of new cases?
   - Parameter variability can be used to estimate structural uncertainty (e.g. via p-boxes, Edeling et al. 2014)
   - Use of multi-model ensembles, Bayesian model averages

Talk overview

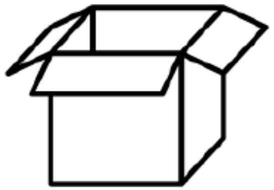
- Turbulence modelling uncertainties
- RANS models: what is wrong?
  - Uncertainties in RANS models
- Tackling parametric uncertainties
- Tackling structural uncertainties
- Numerical results
- Conclusions and future trends

Tackling structural uncertainty

- Much more challenging than parametric uncertainty
- Approaches applied to turbulence modelling
  1. Include a model-inadequacy term in the stochastic model [Kennedy and O'Hagan, 2001, He & Xiu, 2016]
  2. Multi-model ensembles, Bayesian Model Averaging (BMA) [Edeling et al., 2014; 2018]
  3. Non-parametric approaches: find a deterministic or stochastic correction of modeled terms (eddy viscosity, source terms in model transport equations, Reynolds stresses) by some form of machine learning
     - **Black-box** machine learning [Dow & Wang, 2011; Singh & Duraisamy, 2016; Xiao et al., 2016]
       → Difficult to get insight into models



     - **Open-box** machine learning [Weatheritt & Sandberg, 2016; Edeling et al., 2018; Schmeltzer et al., 2018]
       → Tangible mathematical expressions



Bayesian Model Averaging (BMA)

[Edeling, Cinnella, Dwight, JCP 2014, AIAA J 2018]

- Call $\mathcal{M}$ the space of all possible models, M a precise model, y a QoI, z the data:

  → composed by the structure S and the model parameters ϑ (Draper, 1995)

$$\boxed{p\left(y \mid \mathcal{M}, z\right)=\int_{\mathcal{M}} p\left(y \mid M, z\right) p(M \mid z) \, dM=\int\int p\left(y \mid \theta, S, z\right) p(\theta, S \mid z) \, d\theta \, dS}$$

Infinite alternative model structures

- In practice: infinite space approximated by a finite discrete set of models

  → Use several concurrent models for predicting a new configuration

- Predictions from alternative models are weighted by posterior model plausibilities
- May be extended to include scenario uncertainty (**BMSA**)
- Mean of the ensemble expected to outperform individual ensemble members
- However, members of a multi-model ensemble
  - May share common systematic errors (e.g. Boussinesq)
  - Do not span the full range of possible model configurations (kind of bias)

Non parametric approaches

- Example: linear eddy viscosity models based on Boussinesq approximation

$$\tau_{ij} = -\nu_t 2S_{ij} + \frac{2}{3}k\delta_{ij} = a_{ij} + \frac{2}{3}k\delta_{ij}$$

  - Not suitable for flows with separation, streamline curvature, strong gradients, etc.
  - No method to assess reliability of RANS predictions w/o experimental data

- **Internal additive error term**→ correction for model-form inadequacy

$$a_{ij} = -2\nu_t S_{ij} + 2kb_{ij}^{\Delta}$$

- Use high-fidelity data (DNS, LES,…) to inform the correction
- Correction may be subject to physical constraints

Deterministic approach: direct model learning

[Schmeltzer, Dwight, Cinnella, ETMM 2018]

- Use high-fidelity data (DNS, LES,…) to inform the correction
  1. Find b^{Δ}_{ij} : passively solve model equations using S_{ij}, τ_{ij}, k from a high-fidelity database
  2. Assume that τ is a function of the velocity gradient only
  3. Write correction as a function of base tensors and invariants (Pope, 1975):

$$b_{ij}(S_{ij},\Omega_{ij}) = \sum_{n=1}^{N} T_{ij}^{(n)} \alpha_n(I_1, \dots, I_5)$$

  - 10 base tensors and 5 invariants: for simplicity, we choose to use at most quadratic terms, and the first two invariants:

$$T_{ij}^1 = S_{ij}, \qquad T_{ij}^3 = S_{ik}S_{kj} - \frac{1}{3}\delta_{ij}S_{mn}S_{nm}, \qquad I_1 = S_{mn}S_{nm},$$

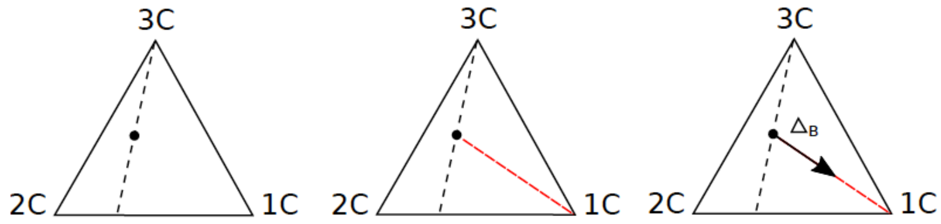
$$T_{ij}^2 = S_{ik}\Omega_{kj} - \Omega_{ik}S_{kj}, \qquad T_{ij}^4 = \Omega_{ik}\Omega_{kj} - \frac{1}{3}\delta_{ij}\Omega_{mn}\Omega_{nm} \qquad I_2 = \Omega_{mn}\Omega_{nm},$$

Relationship between coefficients α and the invariants I established by **symbolic regression**

Stochastic approach: return-to-eddy-viscosity model

[Edeling, Iaccarino, Cinnella, FTaC 2017]

- Direct estimation of the correction requires solving a highly dimensional inference problem + deterministic regression does not provide uncertainty estimates
- Assume model error is transported
- Decompose $\langle u_i u_j \rangle = 2k(V \Lambda V^T + \frac{1}{3} I) \rightarrow$ baseline (bl) model
  Magnitude (k), shape (Λ) & orientation (V) of $\langle u_i u_j \rangle$
- Perturb eigenvalues projected onto Banerjee's (2007) baricentric triangle
  Linear anisotropy relationship: $b_{ij} = C_{1c} b_{1c} + C_{2c} b_{2c} + C_{3c} b_{3c}$



- Perturb shape $\langle u_i u_j \rangle \Leftrightarrow \Delta\lambda_\alpha \Leftrightarrow$ perturb location in barycentric map by defining Δ_B and $C_{\alpha c}$ at every location in space!

- Write transport equations for the perturbations (inspired by the LAG model of Olsen & Coackley)

$$\boxed{\frac{\mathrm{D} C_{\alpha c}}{\mathrm{D} t} = \underbrace{a_{\alpha c} \frac{\epsilon}{k} \left(C_{\alpha c}^{(bl)} - C_{\alpha c} \right)}_{LAG\ \ terms}}, \qquad \alpha = 1 \text{ or } 2$$

Rate of return to eddy viscosity

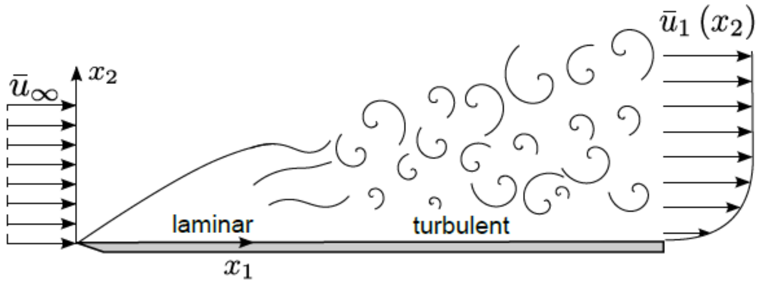
  - BL RANS model recovered where the model does not need to be perturbed (e.g. thin shear flows)
  - Use a priori (data-free) or a posteriori (data-driven) intervals for the coefficient

Talk overview

- Turbulence modelling uncertainties
- RANS models: what is wrong?
    - Uncertainties in RANS models
- Tackling parametric uncertainties
- Tackling structural uncertainties
- Numerical results
- Conclusions and future trends

BMSA: Turbulent flow over a flat plate

Objective: predict velocity profiles for flat plate boundary layer subject to various gradients



Governing equations*:* RANS + turbulence model

- Algebraic Baldwin-Lomax' (1972) model
- Launder-Jones's (1972) k-ε model
- Menter's (1992) k-ω SST model
- Spalart-Allmaras (1992) model
- Wilcox' stress-ω model (2006)

Data: 13 velocity profile measurements from [Coles & Hirst, 1968]

Bayesian model-scenario averaged prediction of the profiles : requires 5x13 UQ runs

- Average prediction:

$$E\left[\Delta \mid Z\right]=\sum_{i=1}^{m}\sum_{j=1}^{s}E\left[\Delta \mid z_j, M_i\right]P\left(M_i \mid z_j\right)P\left(z_j\right)$$

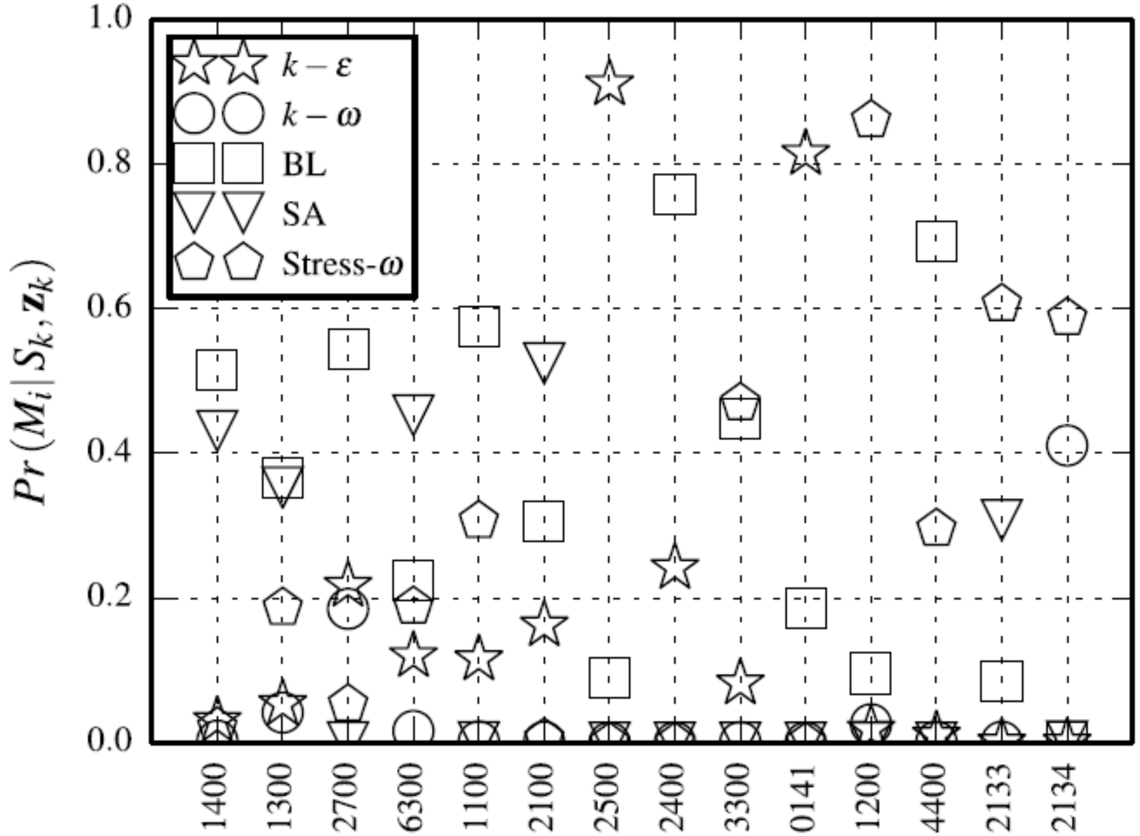
- Analogous formula for the variance

BMSA : model probabilities

- Posterior model plausibilities computed for all models in M for each z_k by sampling

$$p(\theta_k \mid \mathbf{z}_k)$$

- Can be considered as a measure of consistency of calibrated model M_i with data $\mathbf{z}_k$



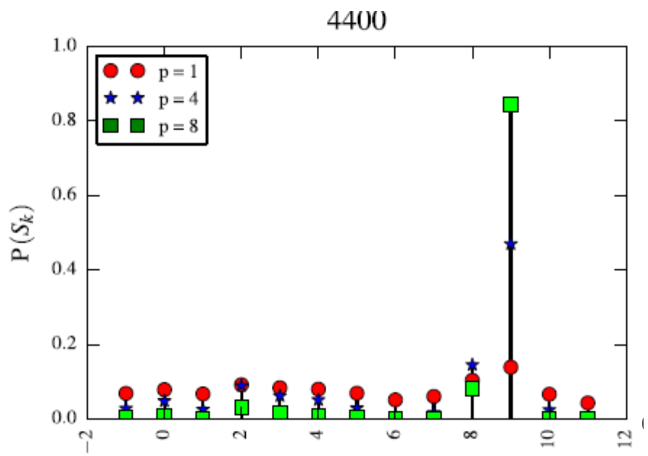
Large spread in model plausibilities, according to the pressure gradient scenario

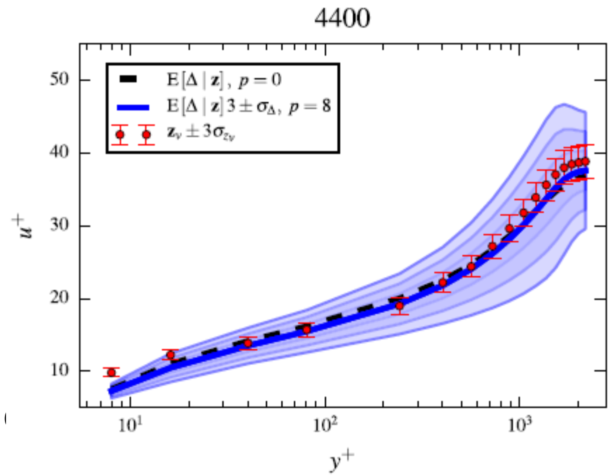
BMSA prediction

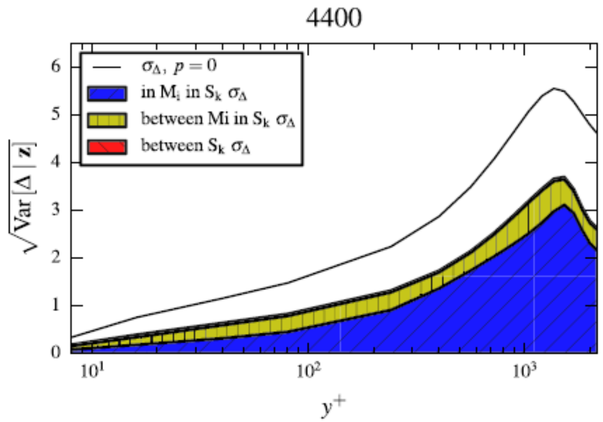
- Scenario pmf uniform (overconservative) or weighted according to an error measure
  → penalizes scenarios with a large between-model, in-scenario variance

$$\varepsilon_j = \sum_{i=1}^{m} \left\| E\left[\Delta \mid z_j, M_i\right] - E\left[\Delta \mid z_j, \boldsymbol{M}\right] \right\|_{L_2} \qquad \forall z_j \in Z$$

$$P(z_j) = \varepsilon_j^{-p} \Big/ \sum_{k=1}^{s} \varepsilon_k^{-p}$$



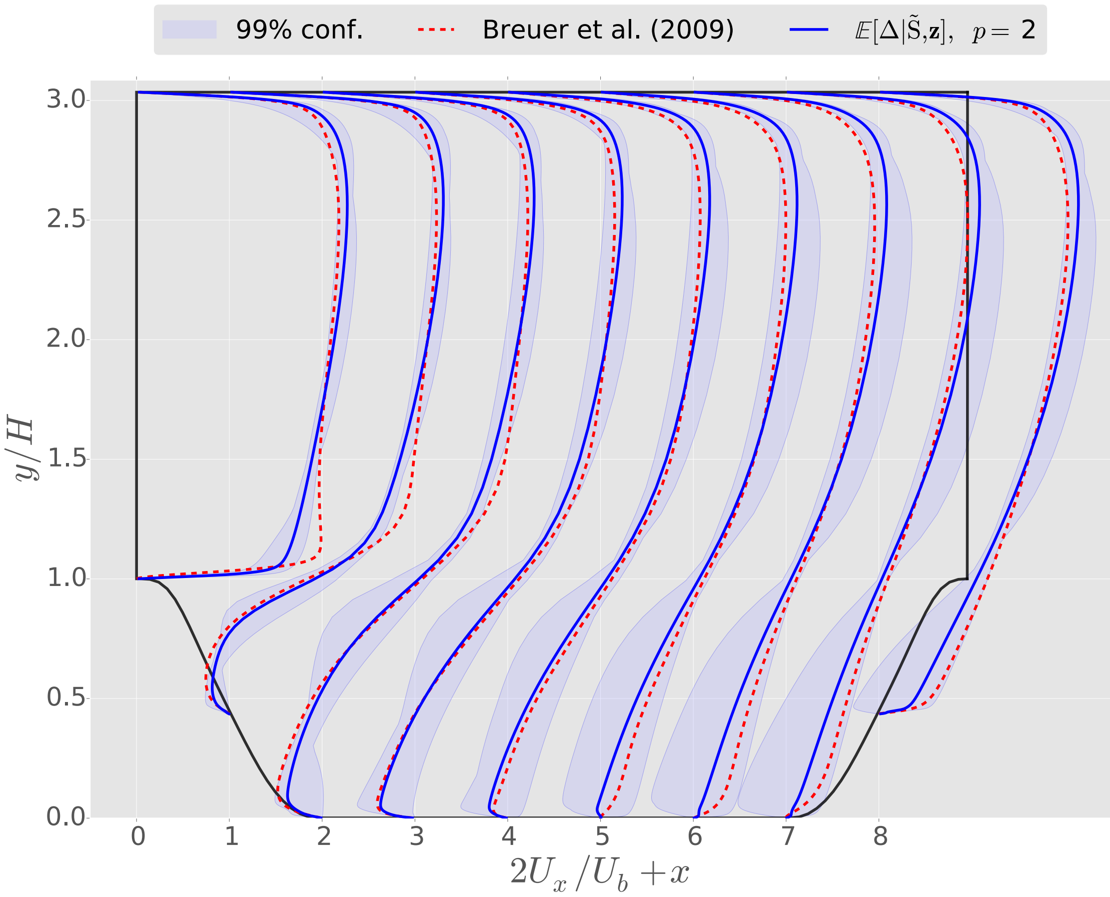




BMSA: Flow over a periodic 2D hill at Re=5600

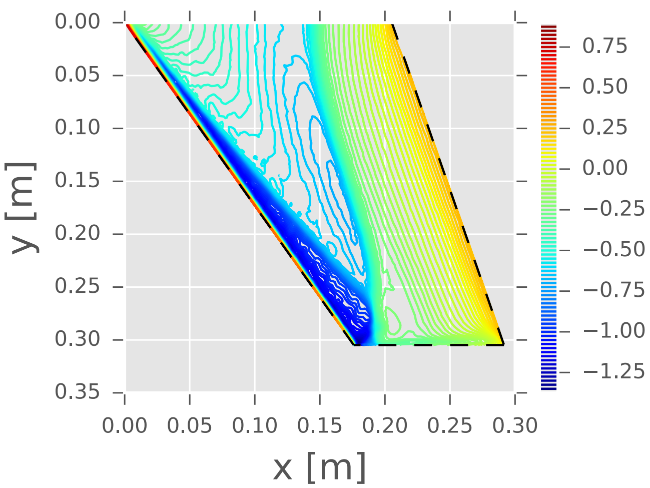
- Prediction of a flow configuration far from RANS safety margins
- Models: Spalart-Allmaras, Jones-Launder, Wilcox
- Propagation of the 13 boundary layer MAP estimates of the parameters through SIMPLEFOAM
- Comparison with DNS data of Breuer et al.

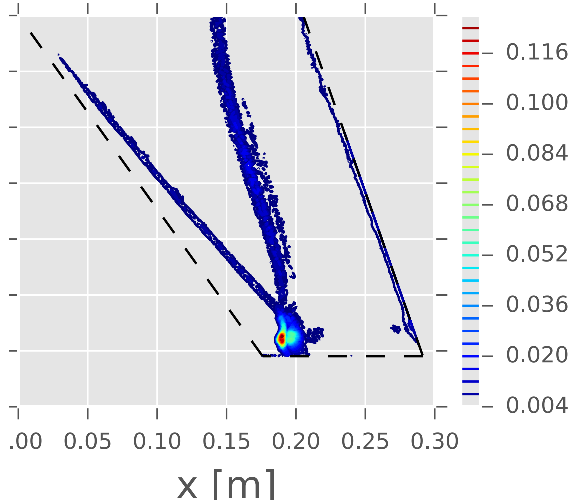
Velocity profiles at various downstream positions.



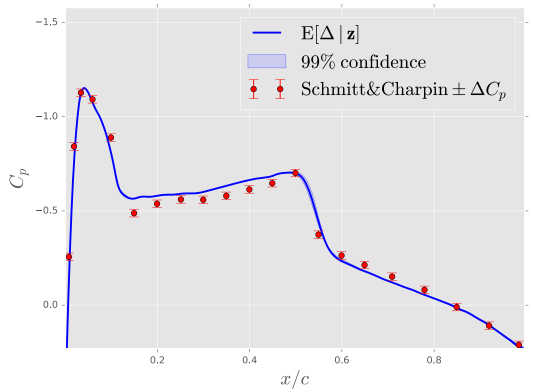
BMSA: Transonic flow past a wing

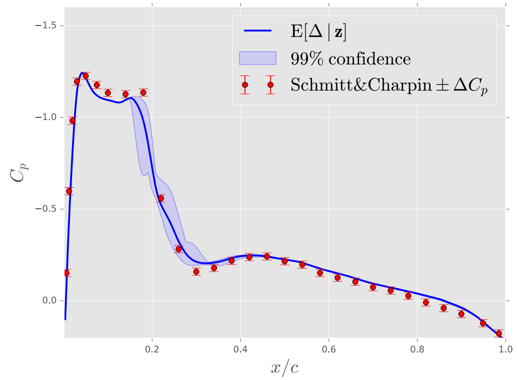
- Coefficient and plausibility database applied to the prediction of the pressure coefficient for transonic flow past the ONERA M6 wing:
  - M=0.8395, AoA=3.06°
- Results based on two models (Jones-Launder & Spalart-Allmaras)
- Propagation of the MAP estimates of the parameters through FLUENT
- Scenario weights computed locally in each section.





Predicted pressure coefficient C_p





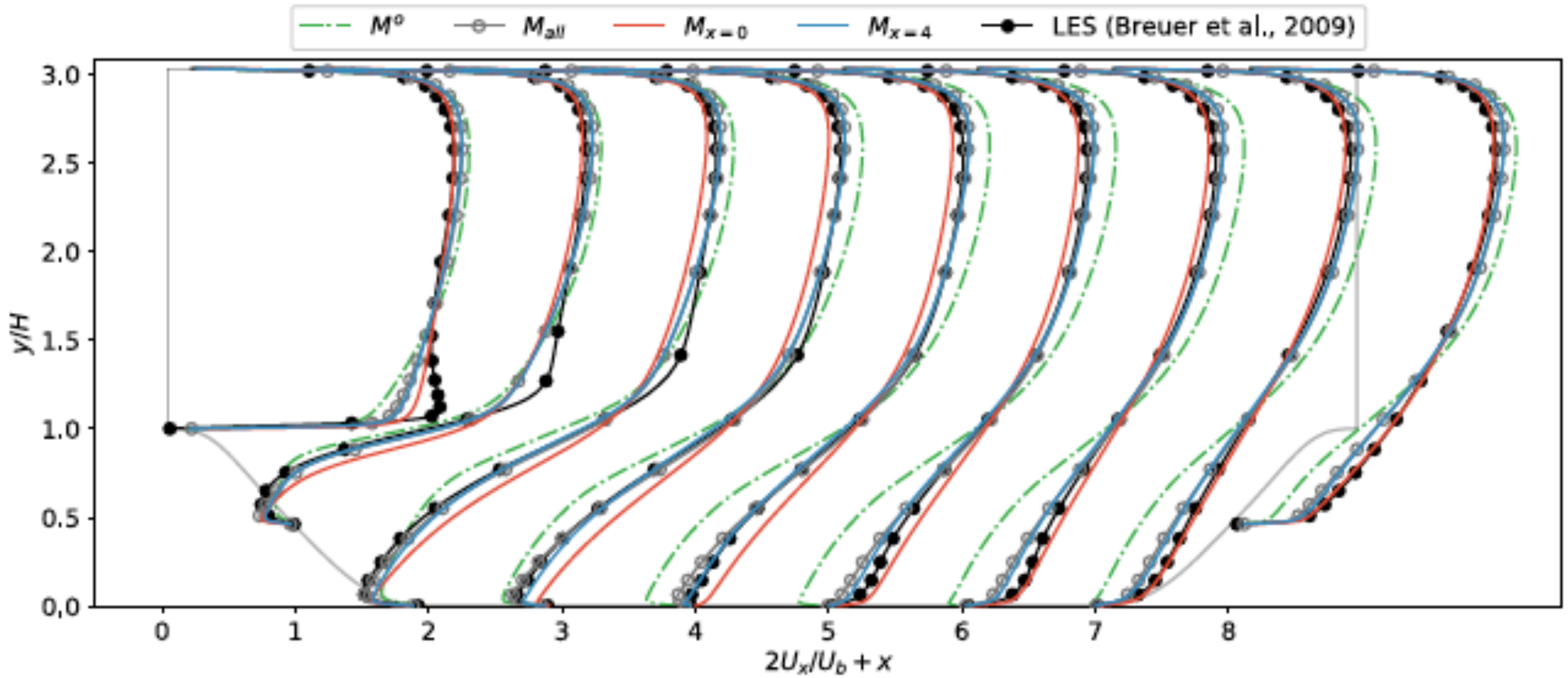
- Low model uncertainty at the root
- Uncertainty in shock locations
- High uncertainty in the tip region (shock/tip vortex interaction)

Direct model learning

- Model corrections to k-ω SST derived using LES data (Breuer 2009) for the 2D-hill flow at Re=10595
  - Learn model on Reynolds-stress profile
  - Propagate through the main system (OPENFOAM solver)
- Several … a i… …at x=0 or x=4, …)

$$M_{all} = (0.2862 + 0.8799\ I_1)\ T_{ij}^{(1)} + 4.1341\ T_{ij}^{(2)} - 2.0649\ T_{ij}^{(4)},$$

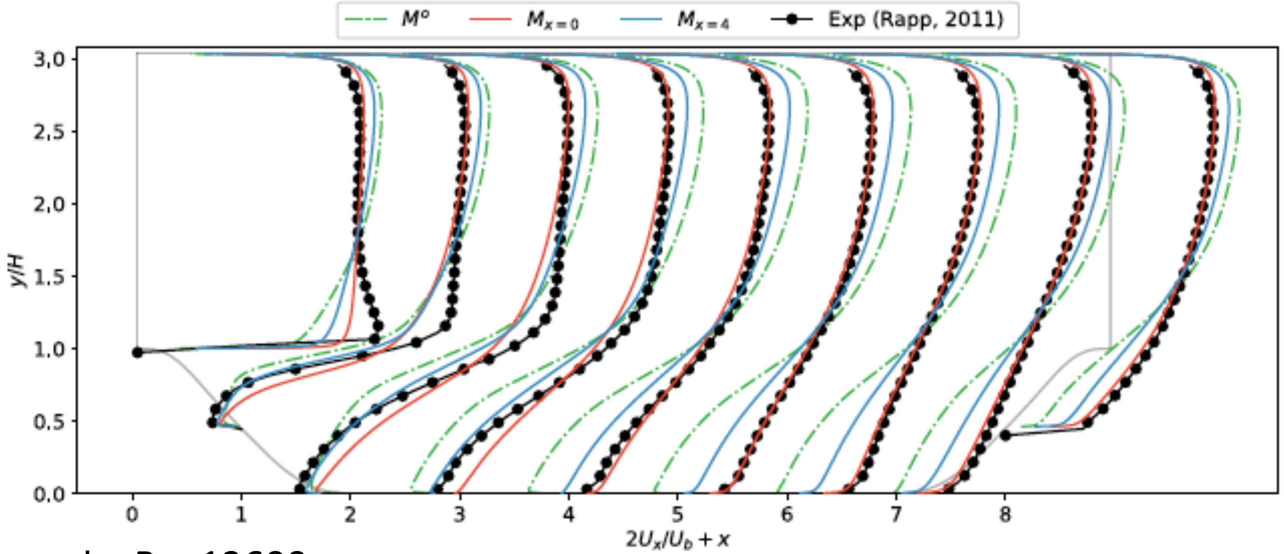
Model	$\epsilon_{\xi=1.0}$	$\epsilon_{\xi=0.1}$
M_{all}	0.0682	0.1035
$M_{x=0}$	0.0057	0.011
$M_{x=4}$	0.0051	0.009



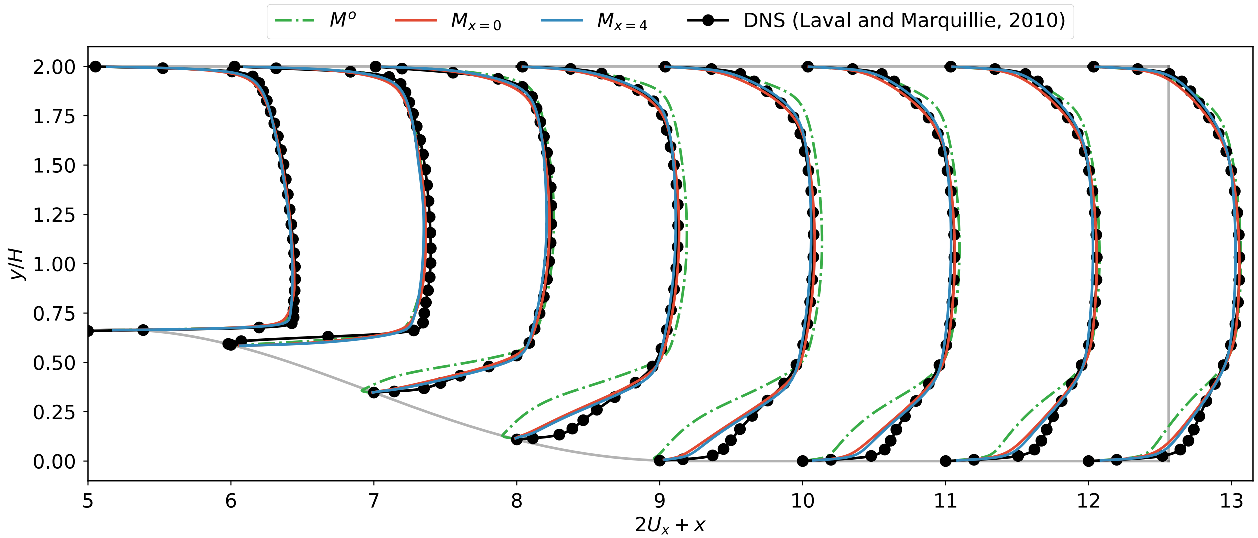
Velocity profile predictions for the training case

Direct model learning

- Velocity profile predictions for different cases
  (same geometry, different Re; different geometry, similar Re)
  - 2D-hill at Re=37000



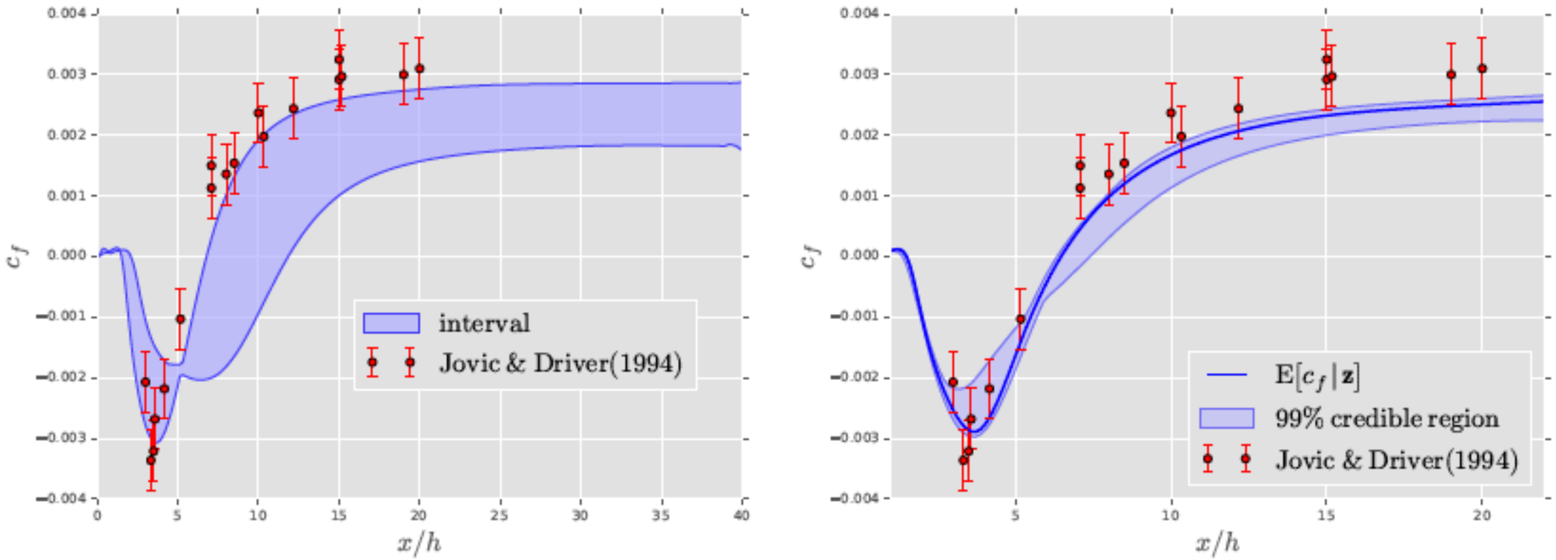
  - Converging-diverging nozzle, Re=12600



Return-to-eddy-viscosity model

- Application: backward facing step at Re=5100
- Data-free prediction: 4 RANS calculation (nominal + 3 perturbed to the corners of baricentric triangle)
- Data-driven prediction:
  - Bayesian calibration of the LAG coefficients using 3 experimental Reynolds-stress profiles [Jovic & Driver, 1994]; use Gaussian likelihood + uniform priors corresponding to theoretical bounds
  - Use stochastic collocation polynomial surrogate (high accuracy with 12x12 tensor grid)

Skin friction coefficient distribution



(a) Data-free Prediction Intervals (b) Data-driven Confidence Intervals

Talk overview

- Turbulence modelling uncertainties
- RANS models: what is wrong?
    - Uncertainties in RANS models
- Tackling parametric uncertainties
- Tackling structural uncertainties
- Numerical results
- Conclusions and future trends

Conclusions

- Bayesian inference provides a systematic framework for updating coefficients associated to turbulence models, and selecting or averaging concurrent models
  - Posterior distributions of the coefficients can be propagated back through the code to make predictions with quantified uncertainty
  - The posteriors maybe strongly dependent on the calibration scenario
- In most cases, no single « best » model can be identified
  - BMSA can be used to account for uncertainties due to multiple model structures/calibration scenarios
  - BMSA is non-intrusive → **readily available for industrial codes**
  - Computational cost may be strongly reduced by using MAP approximation
- But:
  - BMSA acts through model coefficients → global
  - Estimated structural uncertainties biased by the models used in the mixture.
- Is it possible to calibrate an error field?

Emerging trends: «data-driven» frameworks

Characterized by:

- (Spatially) **local** information on model quality
- Use of **lots** of data (DNS/LES)
- High-dimensional parameter-spaces.
- Usually machine-learning to identify mean-flow => error relationship.

Many possibilities are being explored:

- Calibrate correction terms for a baseline turbulence model (e.g. by direct identification)
  - Requires high-quality data
  - Provides an open-box model
  - No uncertainty quantification
- Develop transport equations for Reynolds stress correction
  - Data-free or data-driven application, rough data sufficient
  - Physical realizability constraints
  - Uncertainty bounds provided
  - Computational cost (two additional transport equations + UQ)

Reliability of predictions far away of the training set?

Questions?

[1] Edeling, W.N., Cinnella, P., Dwight, R., Bijl, H., 2014. Bayesian estimates of parameter variability in the k ε turbulence model. J Comp Phys 258 : 73-94.

[2] Edeling, W.N., Cinnella, P., Dwight, R., 2014. Predictive RANS simulations via Bayesian Model-Scenario Averaging. J Comp Phys 275 : 65-91.

[3] Edeling, W.N., Schmeltzer, M., Dwight, R., Cinnella, P., 2018. Bayesian predictions of RANS uncertainties using MAP estimates. AIAA J. 56(5):2018-2029

[4] Edeling, W.N., Iaccarino, G., Cinnella, P., 2017. Data-Free and Data-Driven RANS Predictions with Quantified Uncertainty. Turb Flow Comb 100(3):593-616

[5] Schmeltzer, M., Dwight, R., Cinnella, P., Data-driven discovery of algebraic stress models using elastic net regularization, 12th ETMM, Montpellier, September 26-28 2018

[6] Xiao, H., Cinnella, P., Quantification of model uncertainty in RANS simulations: a review. Progress in Aerospace Sciences, 2019. In press. https://arxiv.org/abs/1806.10434

Bayesian model-scenario averaging (BMSA)

- Let M_i be a model structure in set M of m models
- Be z_j a calibration dataset taken in a training set Z of s calibration scenarios
- BMSA prediction of the expectancy of QoI Δ for a new scenario :

$$E[\Delta \mid Z] = \sum_{i=1}^{m} \sum_{j=1}^{s} E[\Delta \mid z_j, M_i] P(M_i \mid z_j) P(z_j)$$

The scenario of Δ is **NOT** in the calibration set Z

$$E[\Delta \mid z_j, M_i]$$

is the expectancy of Δ for the new scenario, under model M_i calibrated on dataset z_j

Bayesian model-scenario averaging (BMSA)

- Similarly, the variance of Δ may be written as:

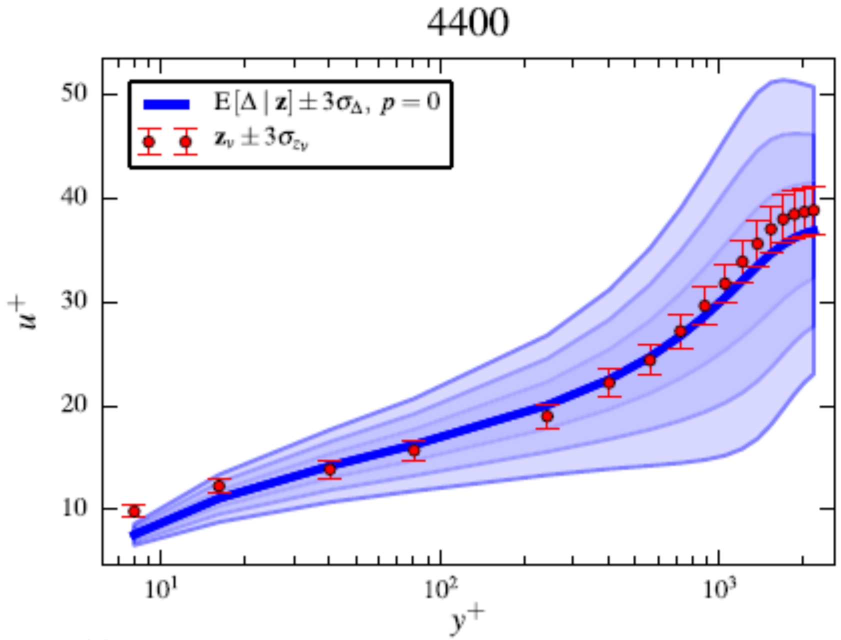
$$\underbrace{\operatorname{var}\left[\Delta \mid Z\right] = \sum_{i=1}^{m}\sum_{j=1}^{s} \operatorname{var}\left[\Delta \mid z_j, M_i\right] P\left(M_i \mid z_j\right) P\left(z_j\right) +}_{\text{In-model, in-scenario variance}}$$

$$\underbrace{\sum_{i=1}^{m}\sum_{j=1}^{s}\left(E\left[\Delta \mid z_j, M_i\right] - E\left[\Delta \mid z_j\right]\right)^2 P\left(M_i \mid z_j\right) P\left(z_j\right) +}_{\text{Between-model, in-scenario variance (model error)}}$$

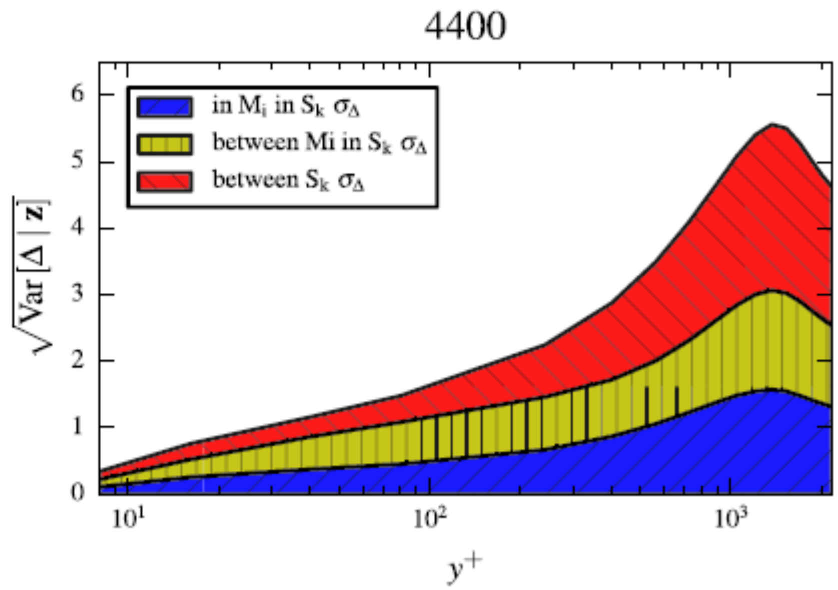
$$\underbrace{\sum_{j=1}^{s}\left(E\left[\Delta \mid z_j\right] - E\left[\Delta \mid Z\right]\right)^2 P\left(z_j\right)}_{\text{Between-scenario variance (spread)}}$$

BMSA prediction

- BMA prediction for a validation case (not in the calibration set)
  - Strong adverse pressure gradient
- Uniform pmf over the calibration scenarios



(c) The BMSA prediction for validation case 4400.

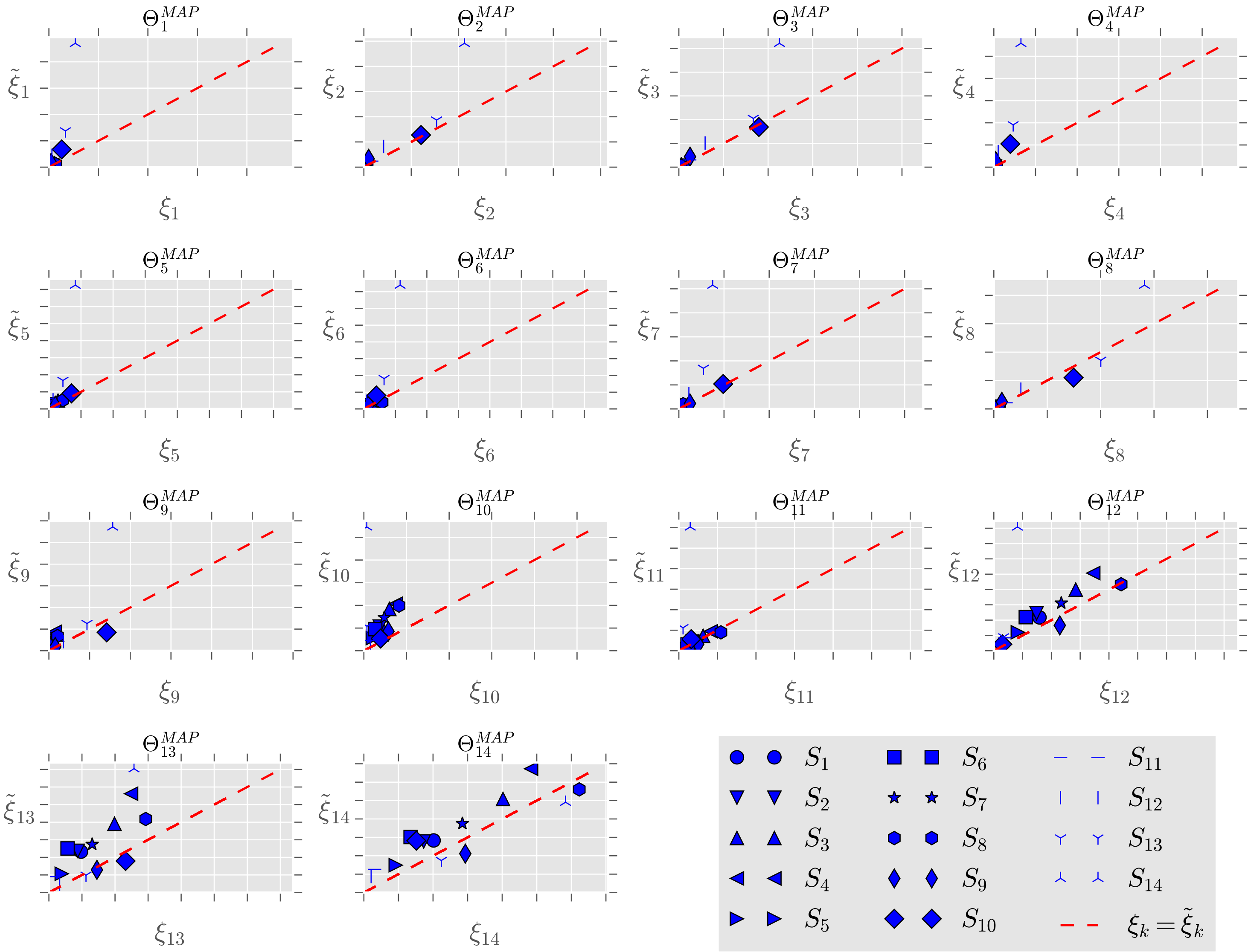


(d) The BMSA standard deviation for validation case 4400.

- Good prediction, but variance strongly over-estimated
- Significant contribution of the between-scenario variance

How good is our scenario weighting?

- For calibration cases we DO HAVE data! Leave one out validation
  - Modelled error versus real error → good agreement in most cases



Deterministic approach: direct model learning

[Schmeltzer, Dwight, Cinnella, ETMM 2018]

- Symbolic regression procedure:
  1. Build a library of weakly nonlinear functions of the invariants

$$\mathcal{B} = \Big\{ c, I_1, I_2, I_1^2, I_2^2, I_1^2 I_2^3, I_1^4 I_2^2, I_1 I_2^2, I_1 I_2^3, \\ I_1 I_2^4, I_1^3 I_2, I_1^2 I_2^4, I_1^2 I_2, I_1 I_2, I_1^3 I_2^2, I_1^2 I_2^2 \Big\}$$

  2. Multiply by each base tensor, store in the library matrix for each mesh point k=1,…,K:

$$\mathbf{C} = \begin{bmatrix} cT^{(1)}_{xx}|_{k=0} & cT^{(2)}_{xx}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{xx}|_{k=0} \\ cT^{(1)}_{xy}|_{k=0} & cT^{(2)}_{xy}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{xy}|_{k=0} \\ cT^{(1)}_{xz}|_{k=0} & cT^{(2)}_{xz}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{xz}|_{k=0} \\ cT^{(1)}_{yy}|_{k=0} & cT^{(2)}_{yy}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{yy}|_{k=0} \\ cT^{(1)}_{xz}|_{k=0} & cT^{(2)}_{yz}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{yz}|_{k=0} \\ cT^{(1)}_{zz}|_{k=0} & cT^{(2)}_{zz}|_{k=0} & \cdots & I_1^2 I_2^2 T^{(4)}_{zz}|_{k=0} \\ \vdots & \vdots & & \vdots \\ cT^{(1)}_{zz}|_{k=K} & cT^{(2)}_{zz}|_{k=K} & \cdots & I_1^2 I_2^2 T^{(4)}_{zz}|_{k=K} \end{bmatrix}$$

  3. Carry out regularize regression (L1-rnom LASSO, L2-norm ridge, mixed L1-L2 elastic net and MLE have been tested) to find the spatially-varying coefficients combining elements of *C*. E.g.:

$$\Theta = \arg\min_{\hat{\Theta}} \left\| \mathbf{C}\hat{\Theta} - \mathbf{b}^{\Delta} \right\|_2^2 + \lambda\rho \left\| \hat{\Theta} \right\|_1 \\ + 0.5\lambda(1-\rho) \left\| \hat{\Theta} \right\|_2^2,$$

Stochastic approach: return-to-eddy-viscosity model

[Edeling, Iaccarino, Cinnella,FTaC 2017]

- Perturb eigenvalues projected onto Banerjee's (2007) baricentric triangle

  Linear anisotropy relationship: $b_{ij} = C_{1c}b_{1c} + C_{2c}b_{2c} + C_{3c}b_{3c}$

  - Only two independent coefficients (convex combination)
  - All realizable states lie within the triangle and can be expressed in terms of the C coefficients (bijection beteen the C's and the l's)

Perturbation tensor $\check{E}_{ij} := \mathrm{diag}\left(\Delta\lambda_1, \Delta\lambda_2, \Delta\lambda_3\right)$

$$\check{E}_{ij} = \mathrm{diag}\left(\frac{2}{3}\left[C_{1c} - \frac{3}{2}C^{(bl)}_{1c}\right], -\frac{1}{3}C_{1c} + \frac{1}{6}C_{2c}, -\frac{1}{3}\left[C_{1c} - 3C^{(bl)}_{1c}\right] - \frac{1}{3}C_{2c}\right)$$

- Data-free prediction: perturb the coefficients toward the triangle corners.

  4 RANS simulations are required to build the **prediction interval**:

$[a_{1c}, a_{2c}]$	state
$[0,0]$	isotropic
$[K,K]$	baseline
$[0,K]$	max axisymmetric contraction
$[K,0]$	max axisymmetric expansion

K initial value (plain strain)

- Data-driven prediction: infer on the LAG coefficients (characteristic time scale for return to eddy viscosity model), propagate posteriors to obtain confidence intervals

Rate of return to eddy viscosity

$$\boxed{\frac{DC_{\alpha c}}{Dt} = \underbrace{a_{\alpha c}\frac{\epsilon}{k}\left(C^{(bl)}_{\alpha c} - C_{\alpha c}\right)}_{LAG\ \ terms}}, \qquad \alpha = 1 \text{ or } 2$$