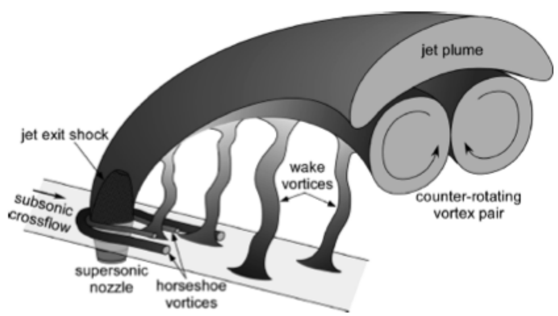
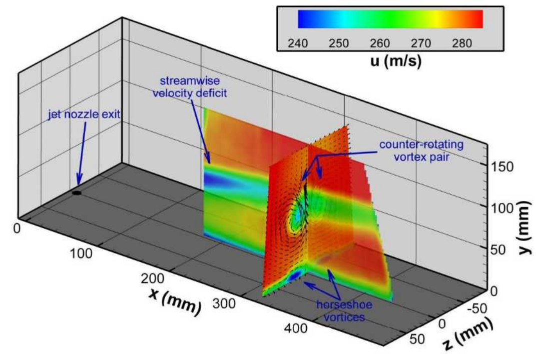
SAND2015-0333PE

Exceptional service in the national interest

Sandia National Laboratories





Bayesian calibration of k-ε parameters for predictive jet-in-crossflow simulations

J. Ray, S. Lefantzi, S. Arunajatesan and L. Dechant

Contact: jairay@sandia.gov

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND2014-2429C

Introduction

- Aim: Develop a predictive RANS model for transonic jet-in-crossflow simulations
  - A strongly vortical flow, often with weak shocks
- Drawback: RANS simulations are simply not predictive
  - They have "model-form" error i.e., missing physics
  - The numerical constants/parameters in the k-ε model are usually derived from canonical flows
- Motivation
  - RANS simulations are still the workhorse for most design activities
  - Jet-in-crossflow is a canonical flow for many maneuvers involving spin rockets and finned aerodynamics bodies
- Hypotheses
  - H1: Source of errors are the unsuitable values used for $\{C\mu, C_2, C_1\}$
  - H2: Model form error in RANS mostly due to the eddy-viscosity model

Technical approach

- H1 – Obtain better values of **C** = {$C\mu$, C_2, C_1} by calibrating to an incompressible flow over square cylinder
  - Strongly vortical, but has little else in common with transonic jet-in-crossflow.
- H2 – Quantify model-form errors by calibrating RANS to transonic jet-in-crossflow measurements
  - Also check robustness of calibration (compare predictive skill at off-calibration points
- Estimate k-ε parameters by posing it as a Bayesian inverse problem
  - Estimate {$C\mu$, C_2, C_1} as a 3-dimension joint PDF (JPDF) by solving the inverse using Markov chain Monte Carlo (MCMC)
    - Capture uncertainty due to (1) limited measurements (2) model limitations
  - Post-calibration, draw 100 samples from the JPDF and probabilistically predict the flow field (develop an ensemble of 100 predictions)
- MCMC will require $O(10^4)$ invocations of the flow solver (to link proposed {$C\mu$, C_2, C_1} with calibration / experimentally observed variable)
  - Develop a surrogate/proxy/statistical response function representation of the RANS simulator

Sections of the talk

- Section 1
  - Show that **C** = {Cμ, C_2, C_1} obtained by calibrating to flow-over-square-cylinder observations are better than the nominal values $\mathbf{C}_{nom}$ = {0.09, 1.42, 1.92}
- Section 2
  - Quantify the improvement in predictions if {Cμ, C_2, C_1} are calibrated to a transonic jet-in-crossflow experiment (M = 0.8, J = 10.2)
    - Are they still more predictive (versus $\mathbf{C}_{nom}$) at other M and J?
  - How big is the disagreement cause by the model-form error on RANS?
    - Where can we isolate the model-form error and how big is it (in that particular variable of interest)?

Section 1

CALIBRATING TO FLOW OVER SQUARE CYLINDER EXPERIMENT

Introduction

- Aim: Develop a predictive RANS model for transonic jet-in-crossflow (JinC) simulations
  - A strongly vortical flow, often with weak shocks
- Approach:
  - Estimate **C** = $\{C\mu, C_2, C_1\}$ from experimental measurements of Reynolds stresses from a flow-over-square-cylinder experiment
  - Check predictive skill versus $\mathbf{C}_{nom}$ in a JinC interaction
- Numerical considerations
  - Describe how one makes a surrogate model for Reynolds stresses generated by a 2D RANS simulator
  - Set up the Bayesian inverse problem, and describe how it's solved using MCMC; describe the estimation error
  - Check predictive skill by developing an ensemble of 100 JinC predictions, post-calibration

The problem

- The model
  - Devising a method to calibrate 3 k-ε parameters **C** = {C_μ, C_2, C_1} from expt. data

$$\frac{\partial \rho k}{\partial t}+\frac{\partial}{\partial x_i}\left[\rho u_i k-\left(\mu+\frac{\mu_T}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right]=P_k-\rho\varepsilon+S_k$$

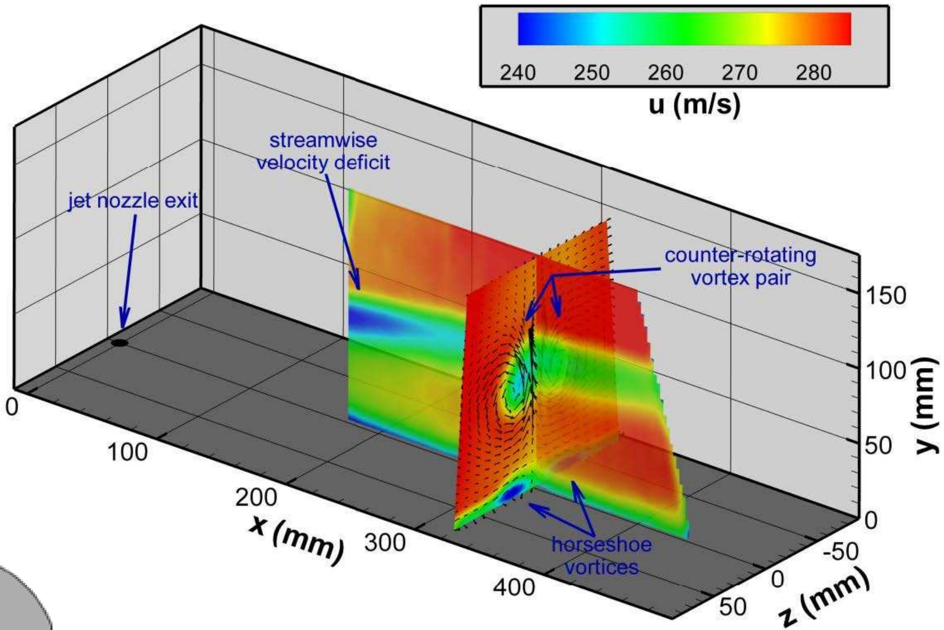
$$\frac{\partial \rho \varepsilon}{\partial t}+\frac{\partial}{\partial x_i}\left[\rho u_i \varepsilon-\left(\mu+\frac{\mu_T}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_i}\right]=\frac{\varepsilon}{k}\left(C_1 f_1 P_k-C_2 f_2 \rho\varepsilon\right)+S_\varepsilon$$

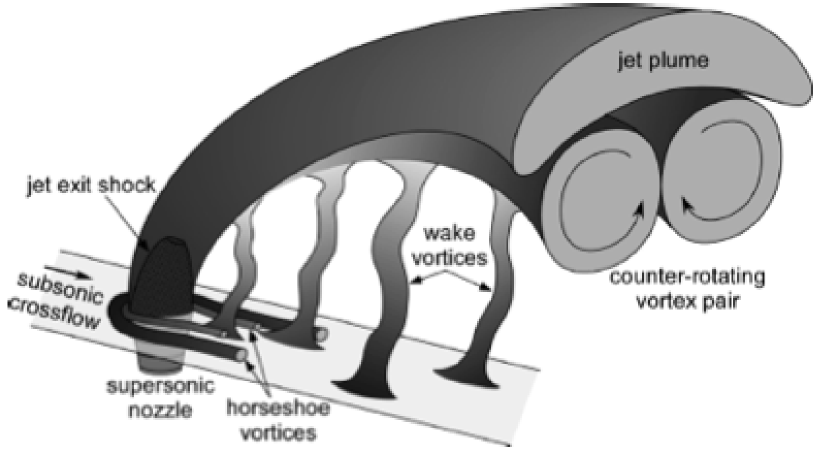
$$\mu_T=C_\mu f_\mu \rho\frac{k^2}{\varepsilon}$$

- Calibration parameters
  - C_μ: affects turbulent viscosity; C_1 & C_2: affect dissipation of TKE
- Calibration method
  - Pose a statistical inverse problem using experimental data for flow-over-a-square-cylinder
  - Estimate parameters using Markov chain Monte Carlo
  - Construct a polynomial surrogate for square-cylinder RANS simulations

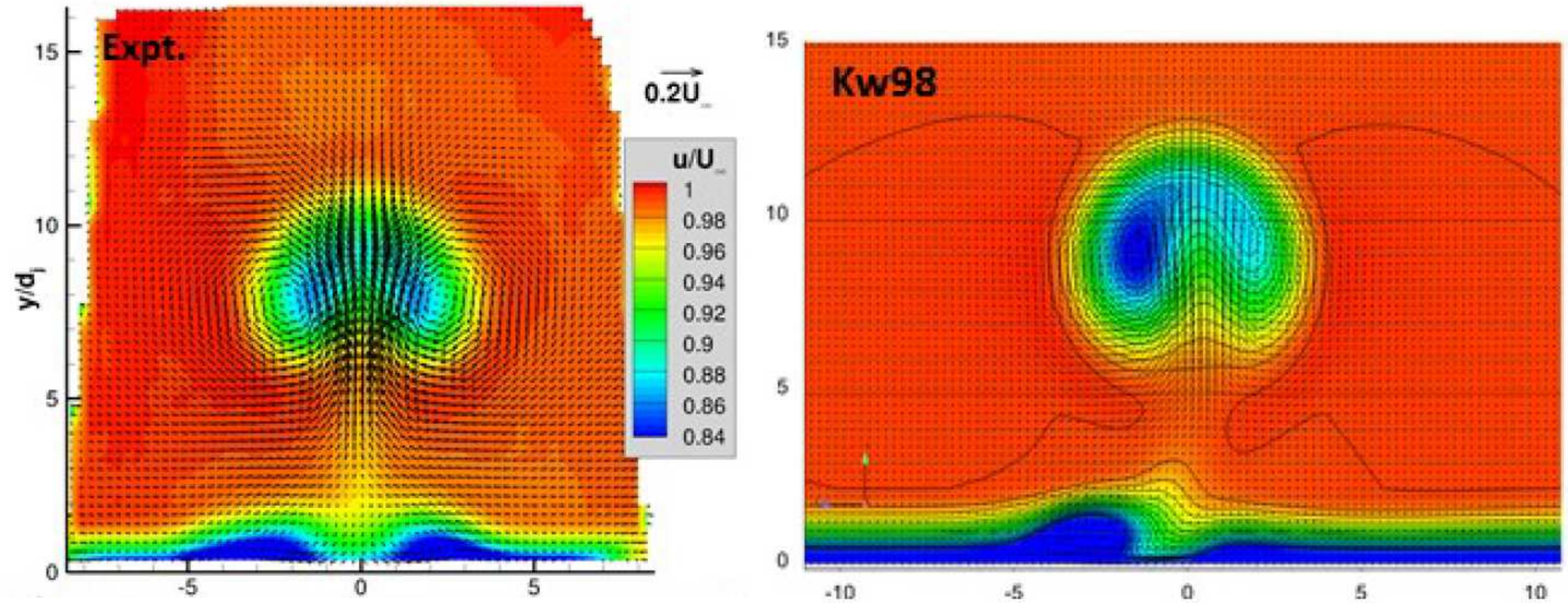
Target problem - jet-in-crossflow

- A canonical problem for spin-rocket maneuvering, fuel-air mixing etc.
- We have experimental data (PIV measurements) and corresponding RANS simulations
- The RANS simulations have stability problems



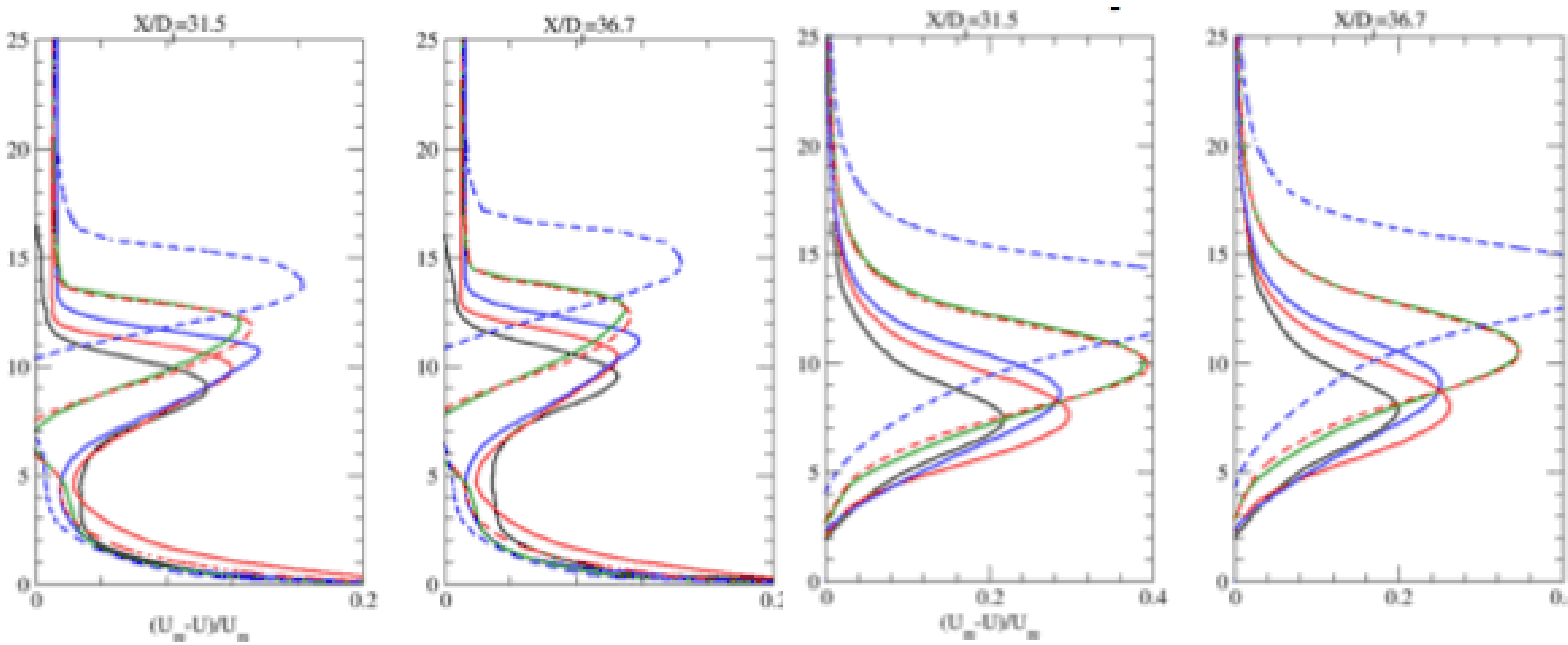


RANS (k-ω) simulations - crossplane results



- Crossplane results for stream
- Computational results (SST) are too round; Kw98 doesn't have the mushroom shape; non-symmetric!
- Less intense regions; boundary layer too weak

RANS (k-ω) simulations – midplane results



U-defect

V - velocity

- Experimental results in black
- All models are pretty inaccurate (blue and red lines are the non-symmetric results)

The desired outcome

- Summary
  - The velocity distribution from RANS at the crossplane is sub-optimal
  - At the mid-plane, the jet sits too high; the vertical velocity is too high indicating a very strong vortex
- Aims of the calibration
  - Get the crossplane vorticity distribution right
    - Correct circulation, position and size of the CVP
  - Match the midplane velocity profiles
- Procedure
  - Use experimental data from a flow-over-square-cylinder experiment
    - Observations of Reynolds stress in the wake behind the cylinder
  - Construct a computationally inexpensive surrogate for the RANS model / predictions of Reynolds stress
  - Use the surrogate for Bayesian calibration of the 3 parameters

Flow over a square cylinder

- Experimental data
  - Water tunnel, 39 cm X 56 cm cross-section
    - Square-cylinder 4 cm per side
  - 96 probes in the wake where $\eta = u'v'$ are measured
- Calibration: Make a map of η to (C_μ, C_2, C_1)
  - Use a statistical (surrogate) model
  - Make a RANS training set using 2744 samples from the (C_μ, C_2, C_1) space
  - Save $\eta = u'v'$ at the 96 probes for each run

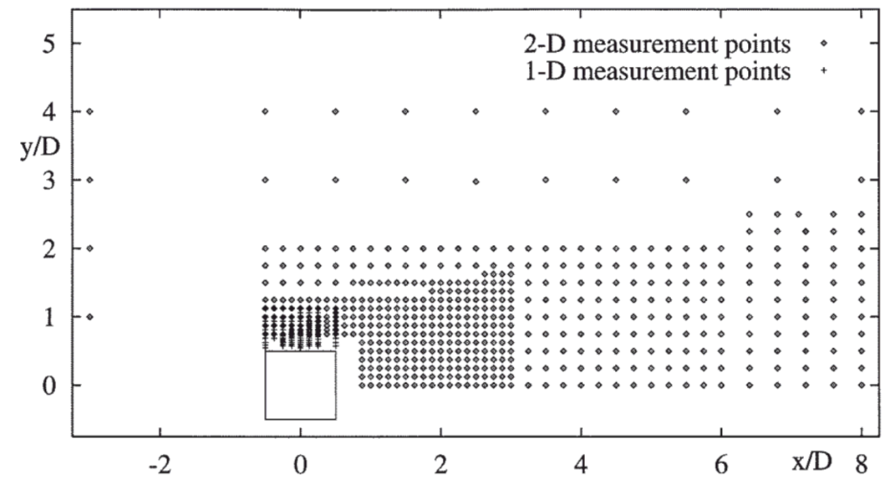


Figure 1: Coordinate system and location of measurement points.

Experimental data and setup from Lyn & Rodi, JFM, 1994

Surrogate models

- Model η as a function of **C** i.e. $\eta = \eta(\mathbf{C})$
  - Approximate this dependence with a polynomial

$$\eta \cong \eta_{trend} = a_0 + a_1 C_\mu + a_2 C_2 + a_3 C_1 + a_4 C_\mu C_2 + a_5 C_\mu C_1 + a_6 C_2 C_1 + \ldots..$$

  - Given η_{exp} at a bunch of probe locations, it should be possible to estimate $\{C_\mu, C_2, C_1\}$ by fitting the polynomial model to data
- But how to get $(a_0, a_1, \ldots.)$ for each of the probe locations to complete the surrogate model for each probe?
  - Divide training data in a Learning Set and Testing Set
  - Fit a full quadratic model for η to the Learning Set via least-squares regression; sparsify using AIC
  - Estimate prediction RMSE for Learning & Testing sets; should be equal
- Final model tested using 100-fold cross-validation; a 10% error threshold was used to select models for the probes

Calibration – in earnest

- Basic idea:
  - Choose 55/96 probes at x/D = 2 ... 8
  - Measured u'v', $(u')^2$ and $(v')^2$
  - minimize $||\eta_{ex} - \eta_{trend}||_2$ by finding 'good' values of (C_m, C_2, C_1)
  - Bayesian calibration: Find $P(C_\mu, C_2, C_1 \mid \eta_{expt})$

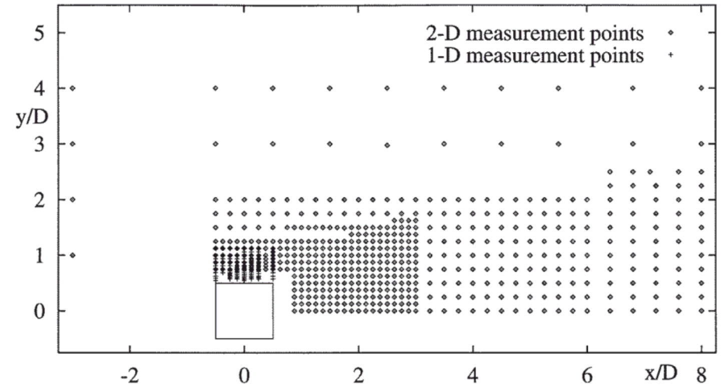


Figure 1: Coordinate system and location of measurement points.

- RANS does not even provide a very good prediction for the wake
  - $(\eta_{ex} - \eta_{trend})$ can be large for many probes
- Choose a set of 'calibration' probes
  - $0.25 < \eta_{ex} / \eta_{trend}(\mathbf{C}_{nominal}) < 4$
- We end up with 28 / 96 probes which we can use for calibration
  - We call this set of 28 probes $\mathcal{P}$

The Bayesian calibration problem

- Model experimental values at probe p as $\eta^{(p)}_{ex} = \eta^{(p)}_{trend}(\mathbf{C}) + \varepsilon^{(p)}$, $\varepsilon^{(p)} \sim N(0, \sigma^2)$

$$\Lambda\left(\eta^{(p)}_{\mathrm{ex}} \mid C\right) \propto \prod_{p \in \mathcal{P}} \exp\left(-\frac{\left(\eta^{(p)}_{ex} - \eta^{(p)}_{trend}(C)\right)^2}{2\sigma^2}\right)$$

- Given prior beliefs π on **C**, the posterior density ('the PDF') is

$$P(C,\sigma \mid \eta^{(p)}_{ex}) \propto \Lambda(\eta^{(p)}_{ex} \mid C,\sigma)\, \pi_{\mu}(C_{\mu})\, \pi_{2}(C_{2})\, \pi_{1}(C_{1}) \pi_{\sigma}(\sigma)$$

- $P(\mathbf{C}|\eta_{ex})$ is a complicated distribution that has to be described/visualized by drawing samples from it
- This is done by MCMC

What is MCMC?

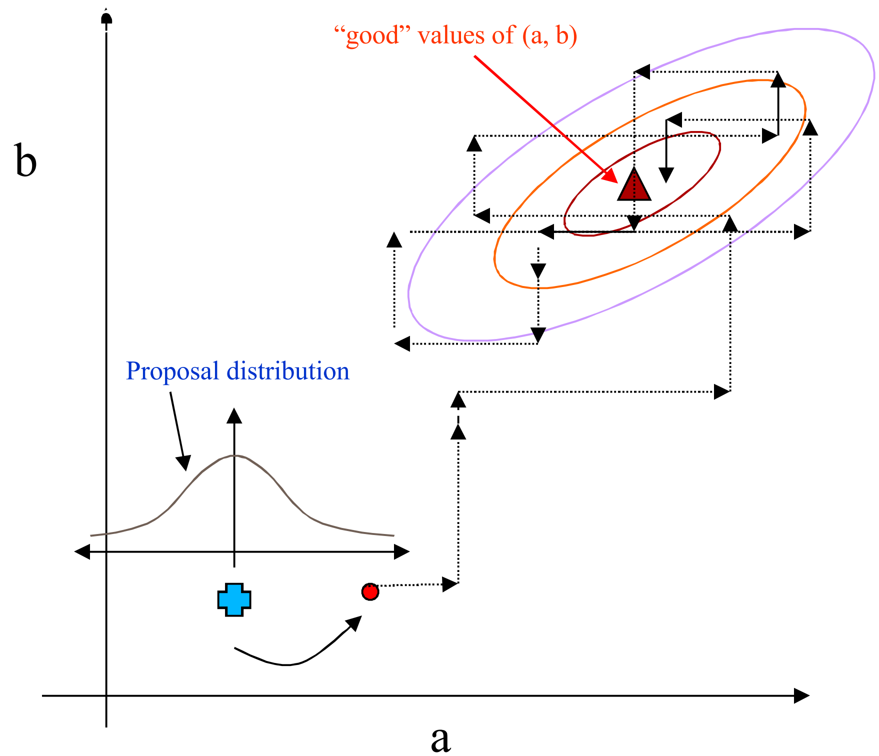
- A way of sampling from an arbitrary distribution
  - The samples, if histogrammed, recover the distribution
- Efficient and adaptive
  - Given a starting point (1 sample), the MCMC chain will sequentially find the peaks and valleys in the distribution and sample proportionally
- Ergodic
  - Guaranteed that samples will be taken from the entire range of the distribution
- Drawback
  - Generating each sample requires one to evaluate the expression for the density π
  - Not a good idea if π involves evaluating a computationally expensive model

An example, using MCMC

- Given: (Y^{obs}, X), a bunch of n observations
- Believed: y = ax + b
- Model: $y_i^{obs} = ax_i + b_i + \varepsilon_i$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$
- We also know a range where a, b and σ might lie
  - i.e. we will use uniform distributions as prior beliefs for a, b, σ
- For a given value of (a, b, σ), compute "error" $\varepsilon_i = y_i^{obs} - (ax_i + b_i)$
  - Probability of the set (a, b, σ) = $\Pi \exp(- \varepsilon_i^2/\sigma^2)$
- Solution: $\pi (a, b, \sigma \mid Y^{obs}, X) = \Pi \exp(- \varepsilon_i^2/\sigma^2)$ * (bunch of uniform priors)
- Solution method:
  - Sample from $\pi (a, b, \sigma \mid Y^{obs}, X)$ using MCMC; save them
  - Generate a "3D histogram" from the samples to determine which region in the (a, b, σ) space gives best fit
  - Histogram values of a, b and σ, to get individual PDFs for them
  - Estimation of model parameters, with confidence intervals!

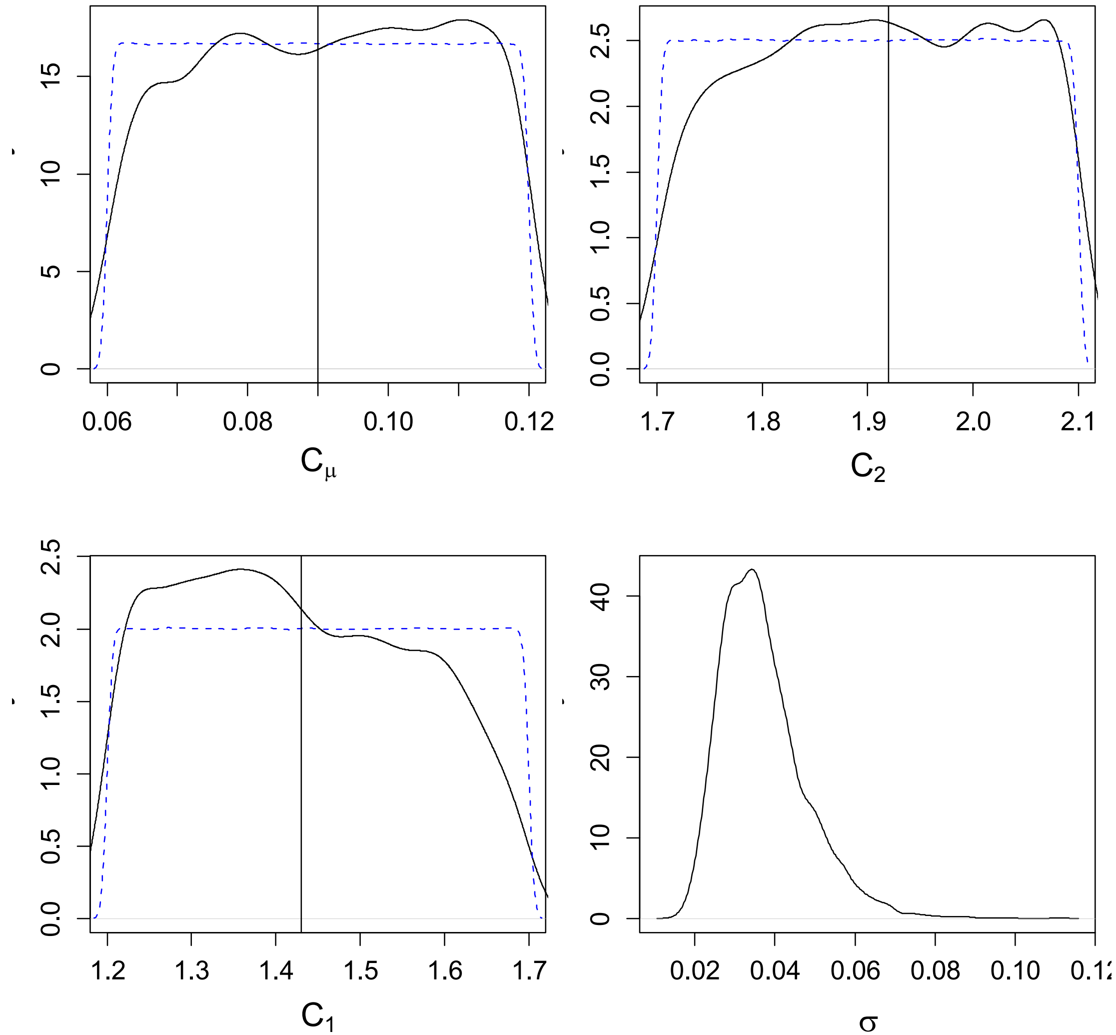
MCMC, pictorially

- Choose a starting point, $P^n = (a_{curr}, b_{curr})$
- Propose a new a, $a_{prop} \sim \mathcal{N}(a_{curr}, \sigma_a)$
- Evaluate $\pi (a_{prop}, b_{curr} \mid \ldots) / \pi (a_{curr}, b_{curr} \mid \ldots) = m$
- Accept a_{prop} (i.e. a_{curr} <- a_{prop}) with probability min(1, m)
- Repeat with b
- Loop over till you have enough samples



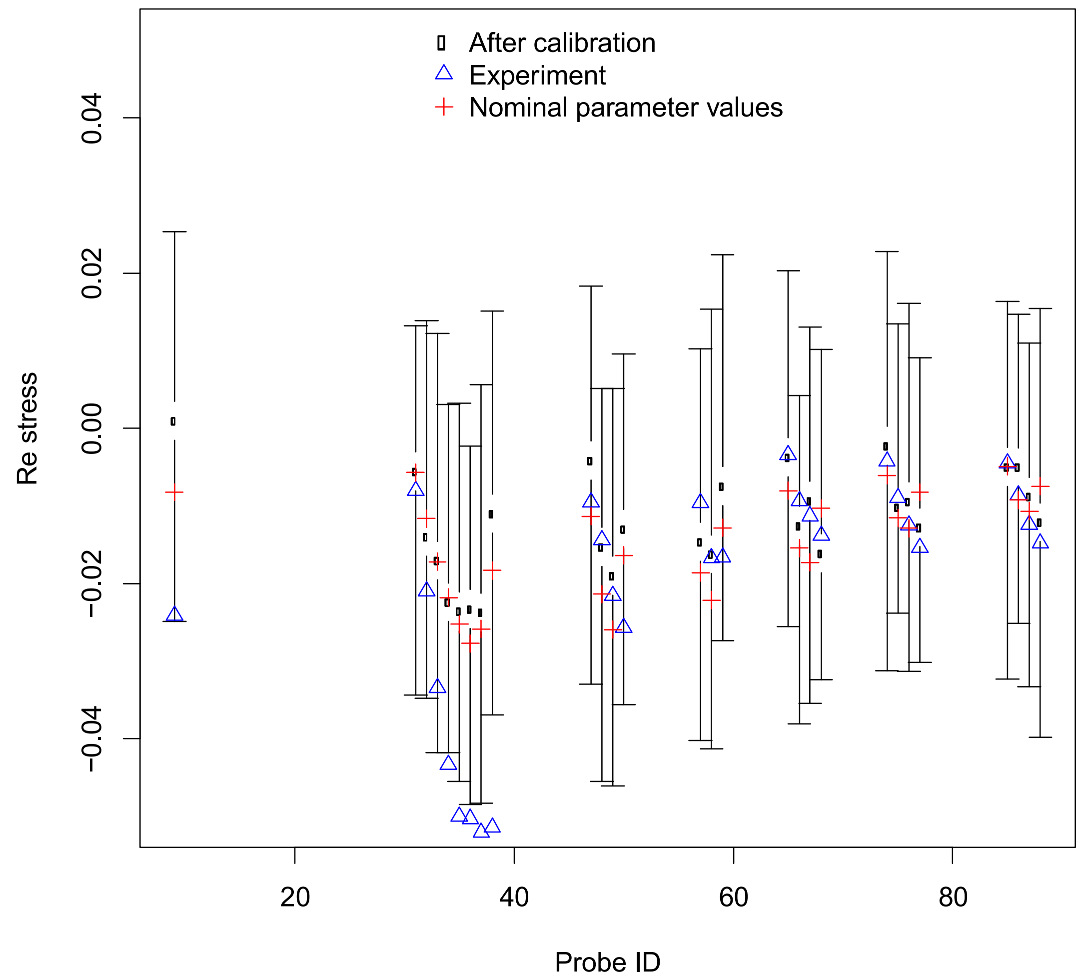
MCMC solution for (C_μ, C_2, C_1)

- Computed using an adaptive MCMC method (DRAM)
- These are marginals – the distribution is 4D
- Nominal values are vertical lines
- Blue dashed lines are prior beliefs
- The model error σ is large



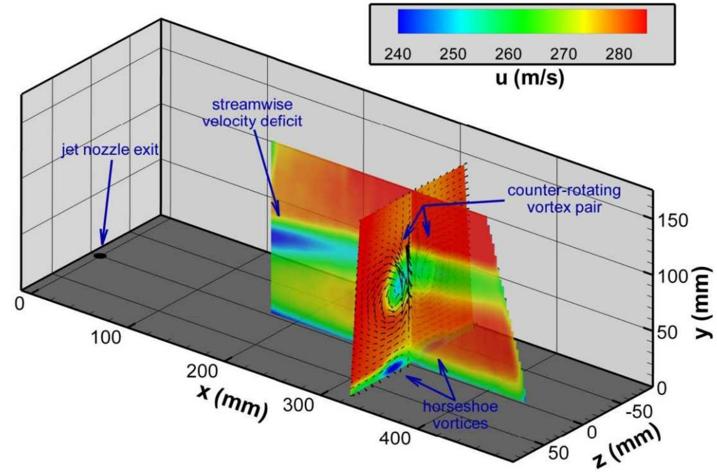
Recreating experimental observations

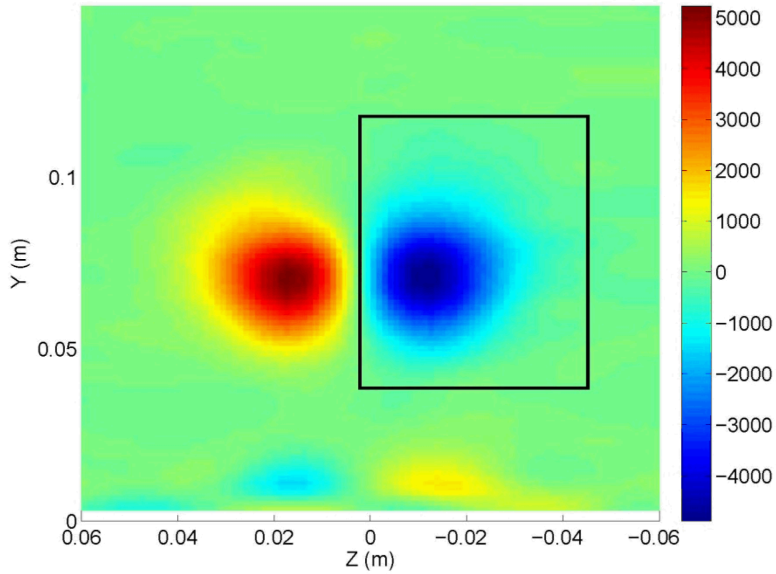
- Post-calibration, we choose 100 **C** samples from the PDF
  - Run the ensemble of 100 RANS runs and plot results at $\mathcal{P}$
- Median predictions close to experimental values
- Error bars capture all measurements



Is the PDF predictive for jet-in-crossflow?

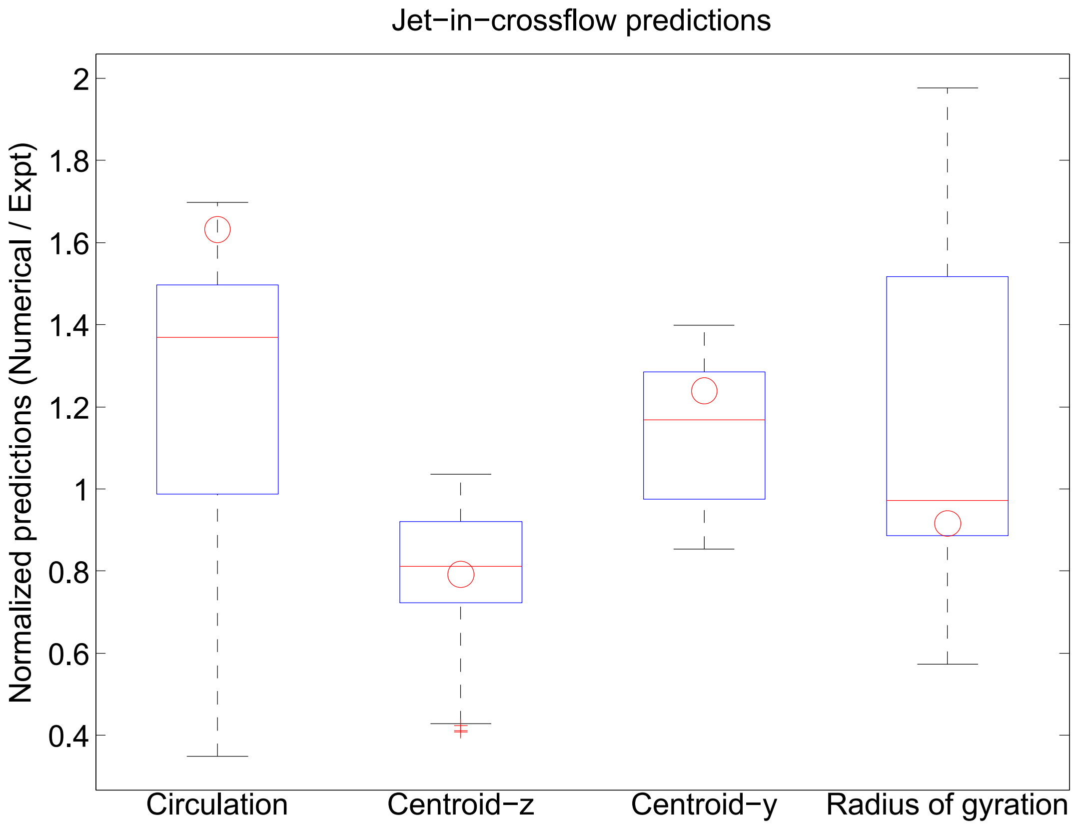
- Pick 100 C samples from the PDF
- Simulate jet-in-crossflow
- In the crossplane, quantify
  - Circulation
  - Centroid of vorticity
  - Radius of gyration
- From the ensemble, calculate median, quartiles etc
- Compare with experimental values



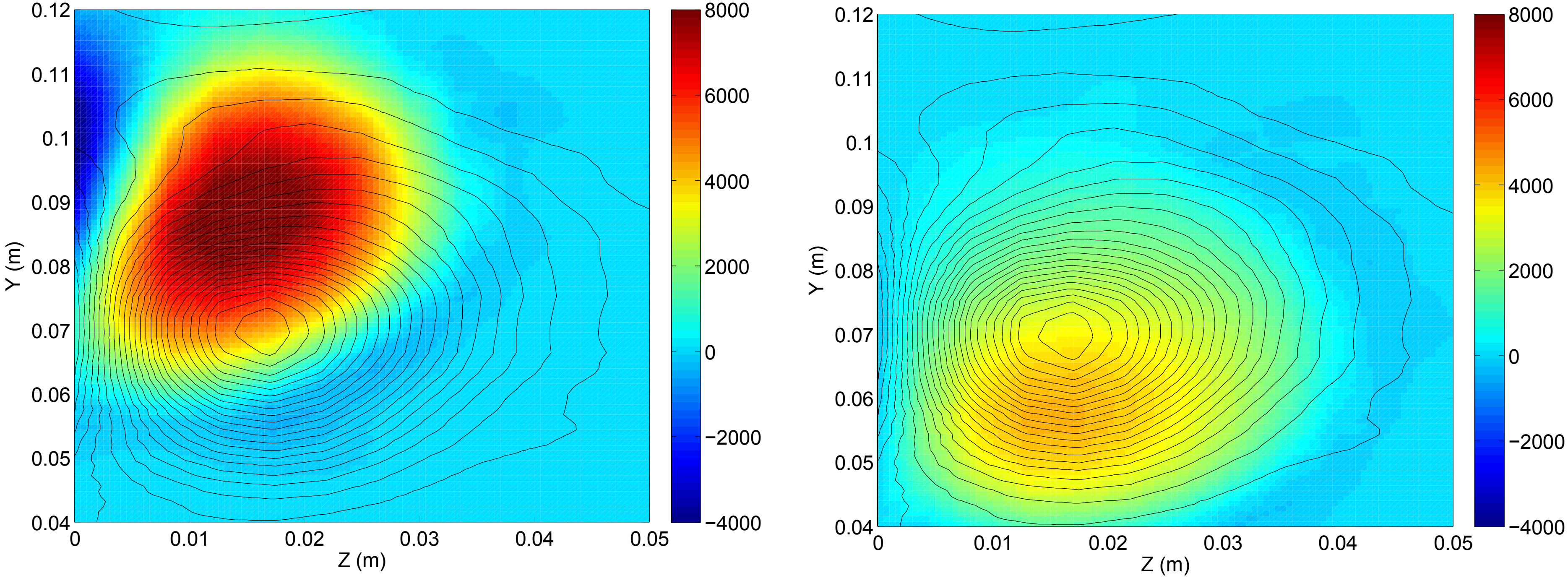


Comparison of predictions and experiments

- Plotting Predictions / Experimental values
- We overpredict circulation
- Location is somewhat off
- Size is somewhat larger
- Big improvements over nominal value
- Also search the 100 ensemble members for best prediction
  - "Optimal" ensemble member



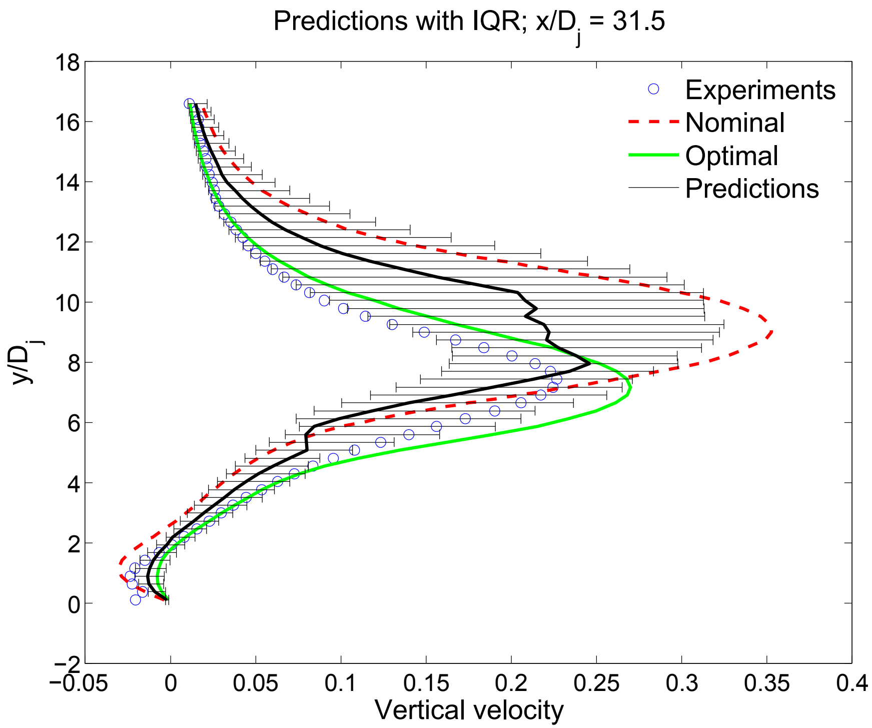
Optimal ensemble member – vorticity



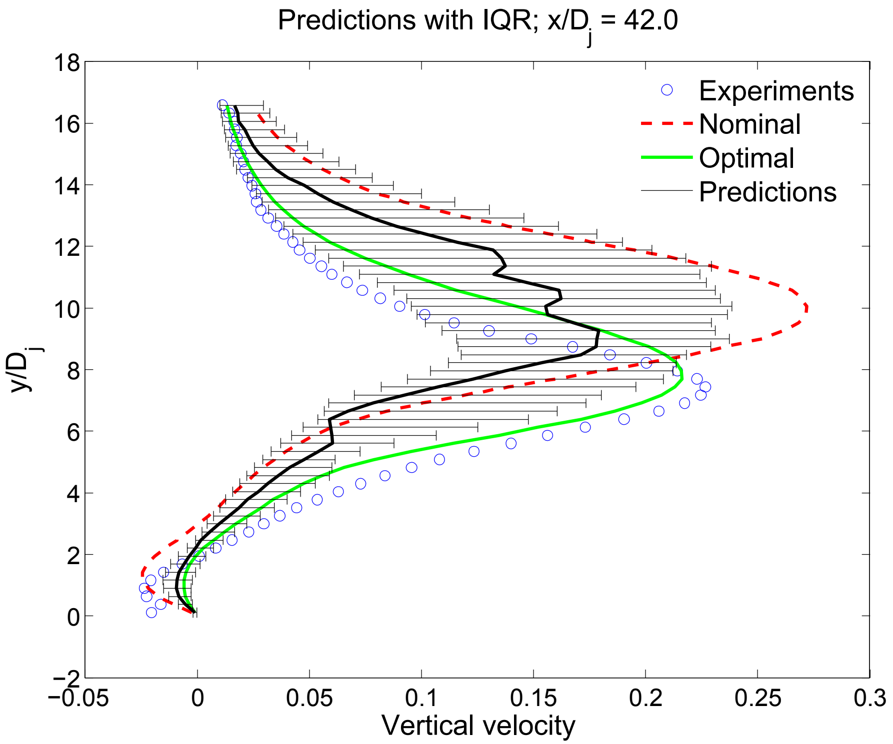
With nominal **C** | With best **C**

- Experimental vorticity as contours
- Calibration positions the vortex better; also gets its strength right
- The circulation, position and size are +/- 15% from experiments

Optimal ensemble member: v velocity



x/D = 31.6



x/D = 42.0

- Improvement over $\mathbf{C}_{\mathbf{nominal}}$
- Nearly nailed the experiment

Conclusions – Section 1

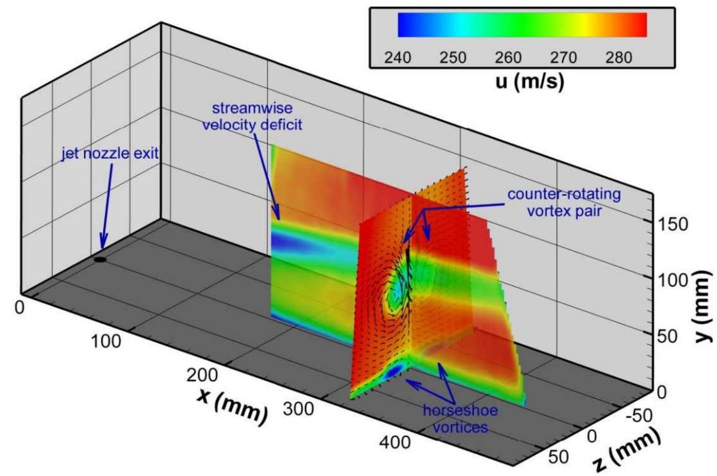
- Our hypothesis of calibrating to a simple vortical flow for predictive jet-in-crossflow proved correct
- Even simple, polynomial surrogates were sufficiently accurate to allow us to calibrate RANS models
  - More elaborate models, with the deficit would probably do somewhat better
  - With surrogates come Bayesian calibration and PDFs of calibrated parameters
- Being able to get a PDF for (C_μ, C_2, C_1) proved to be very convenient
  - Ensemble predictions provide error bars on predictions
  - They allow us to test various (C_μ, C_2, C_1) combinations for predictive power
- *Details: S. Lefantzi, J. Ray, S. Arunajatesan and L. Dechant, "Tuning a RANS k-ε model for jet-in-crossflow simulations", Sandia Technical Report, SAND2013-8158*

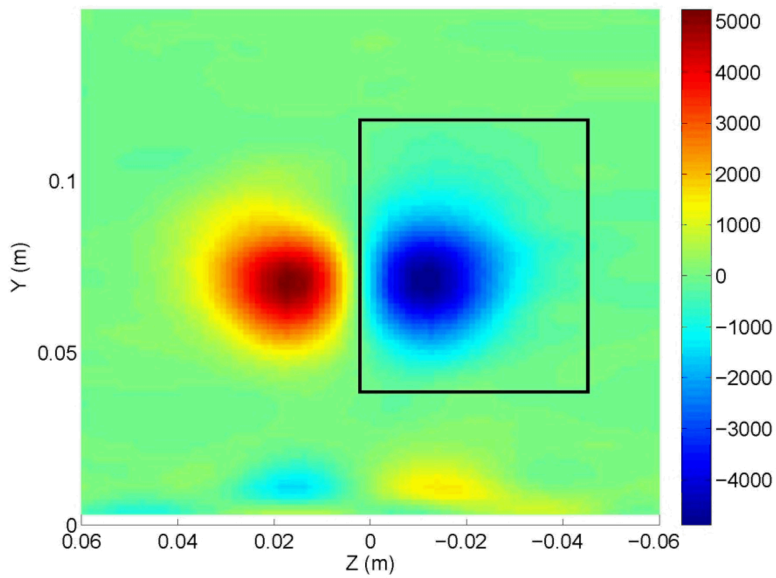
Section 2

UNCOVERING MODEL-FORM ERROR

Introduction

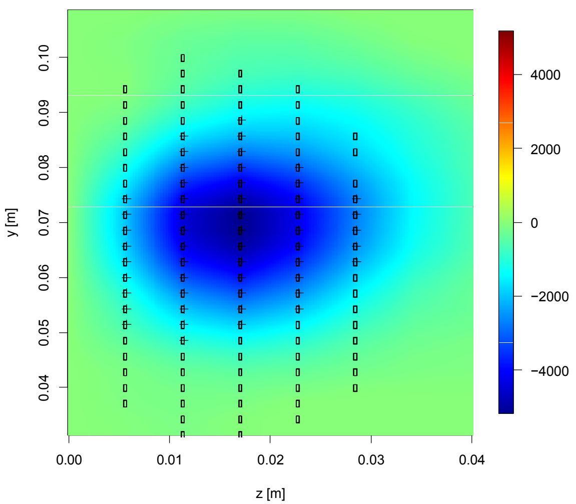
- Aim: Estimate model-form error in 3D RANS simulations of transonic jet-in-crossflow interaction
- Approach
  - Estimate {$C\mu$, C_2, C_1} using Bayesian inference and surrogate models of a 3D RANS simulator
  - Experimental data: Beresh et al, AIAAJ 2005; vorticity on the crossplane
    - M = 0.8, J = 10.2
  - Predict the flowfield (and see improvement over C_{nom} predictions)
  - See predictive skill at off-calibration points (other M and J)
  - Uncover mismatch between predicted and experimentally measured turbulent stresses

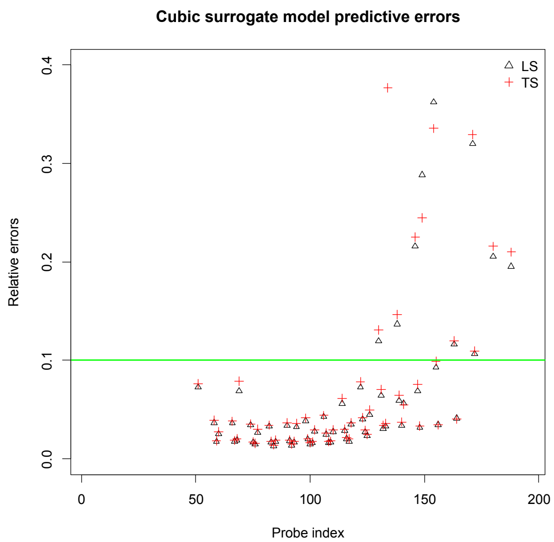




Building surrogate models

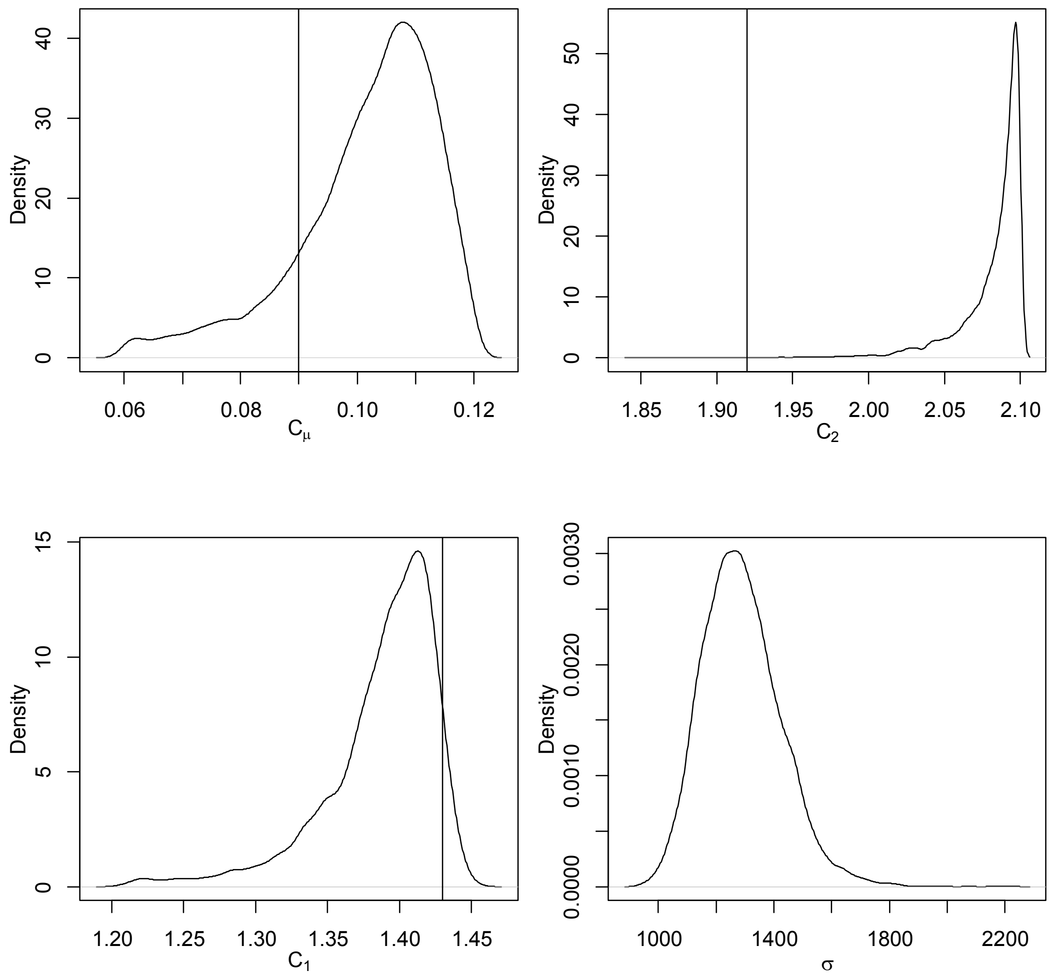
- Sample {Cμ, C_2, C_1} space with 2744 points
  - Run 3D RANS at each and obtain 2744 predicted vorticity fields on crossplane
  - Choose locations with high vorticity (less affected by numerical noise) - 108 "probe" locations
  - Construct a quadratic surrogate F(Cμ, C_2, C_1) for stream-wise vorticity
    - $\omega_x^{(RANS)} \sim F(C\mu, C_2, C_{1;}\ \mathbf{p}) + \eta$
  - Retain only those surrogates that have η < 10%
    - Only 52 / 108 "probes" survive
- Compute vorticity using experimentally observed velocity on crossplane
  - "experimental" vorticity
  - Use them in MCMC calibration





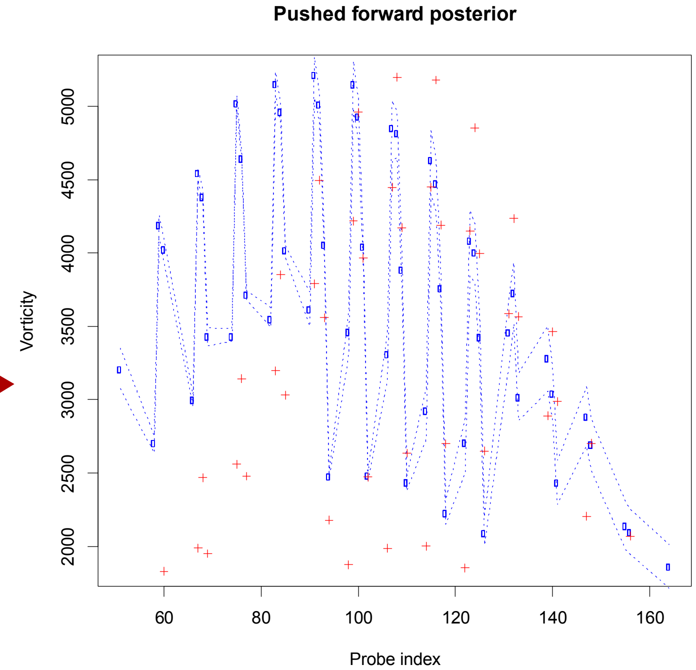
Bayesian calibration

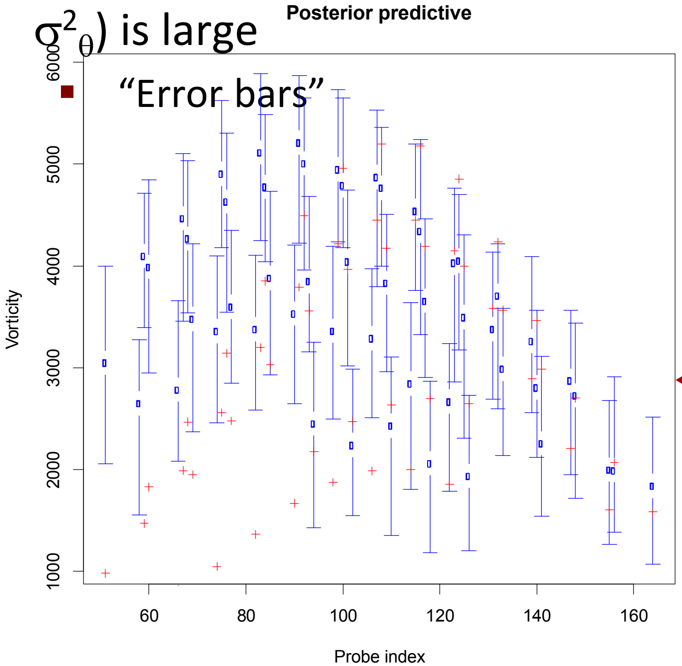
- Use "experimental observations" of vorticity to perform MCMC calibration
- Vertical lines are the nominal values of parameters
- Only C_1 estimates are close to nominal one
- Also compute an estimate of model – data mismatch $\omega_x^{(exp)} - \omega_x^{(RANS)} = \theta \sim N(0, \sigma^2_\theta)$
- Redid calculation using GA; $\mathbf{C}_{GA}$ = {0.105,2.099,1.42}



Can we predict vorticity?

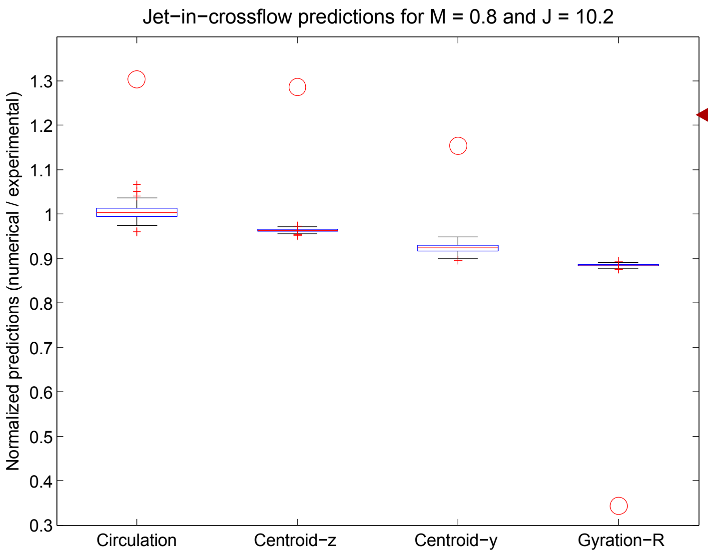
- Take 100 samples from JPDF and run 3D RANS with them
- Compute median prediction & inter-quartile range
- Uncertainty in {Cμ, C_2, C_1} does not lead to a big variation in $\omega_x^{(RANS)}$
- Model-form error $\omega_x^{(exp)} - \omega_x^{(RANS)} = \theta \sim N(0, \sigma^2_\theta)$ is large
  - "Error bars"



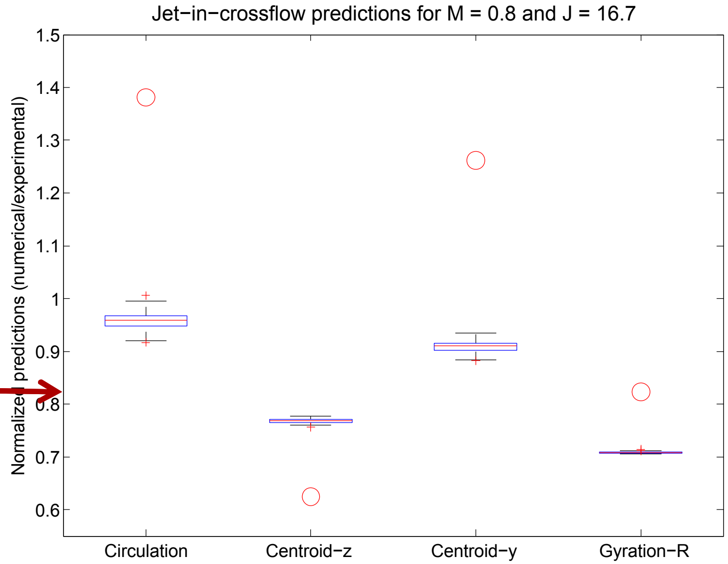


- Now add in the statistical summary of model-form error
- Model-form error (as estimated) is responsible for coming close to the measurements

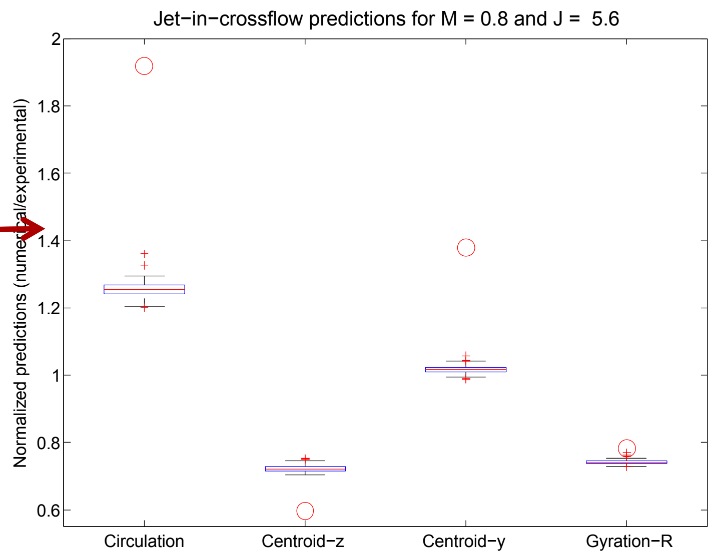
Pre- and post-calibration comparison



M = 0.8,
J = 10.2



M = 0.8,
J = 16.7

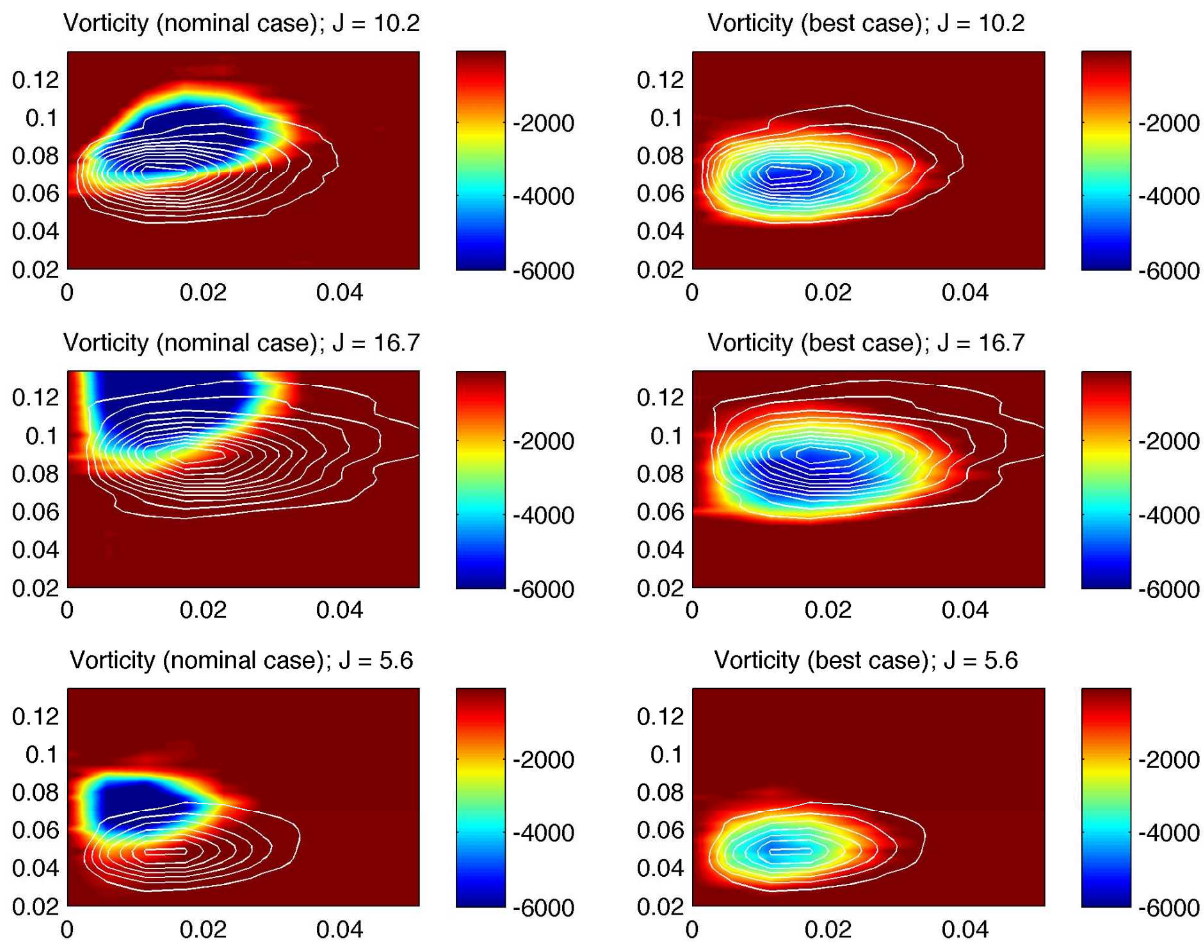


M = 0.8,
J = 5.6

- Summarize vorticity on the crossplane as a point-vortex
- Normalize by experimental values
- Plot predictions using $\mathbf{C}_{nom}$ for comparison
- Compare, pre & post-calibration
- Do for M = 0.8, J = 10.2 (calibration case), J = 16.7 & J = 5.6

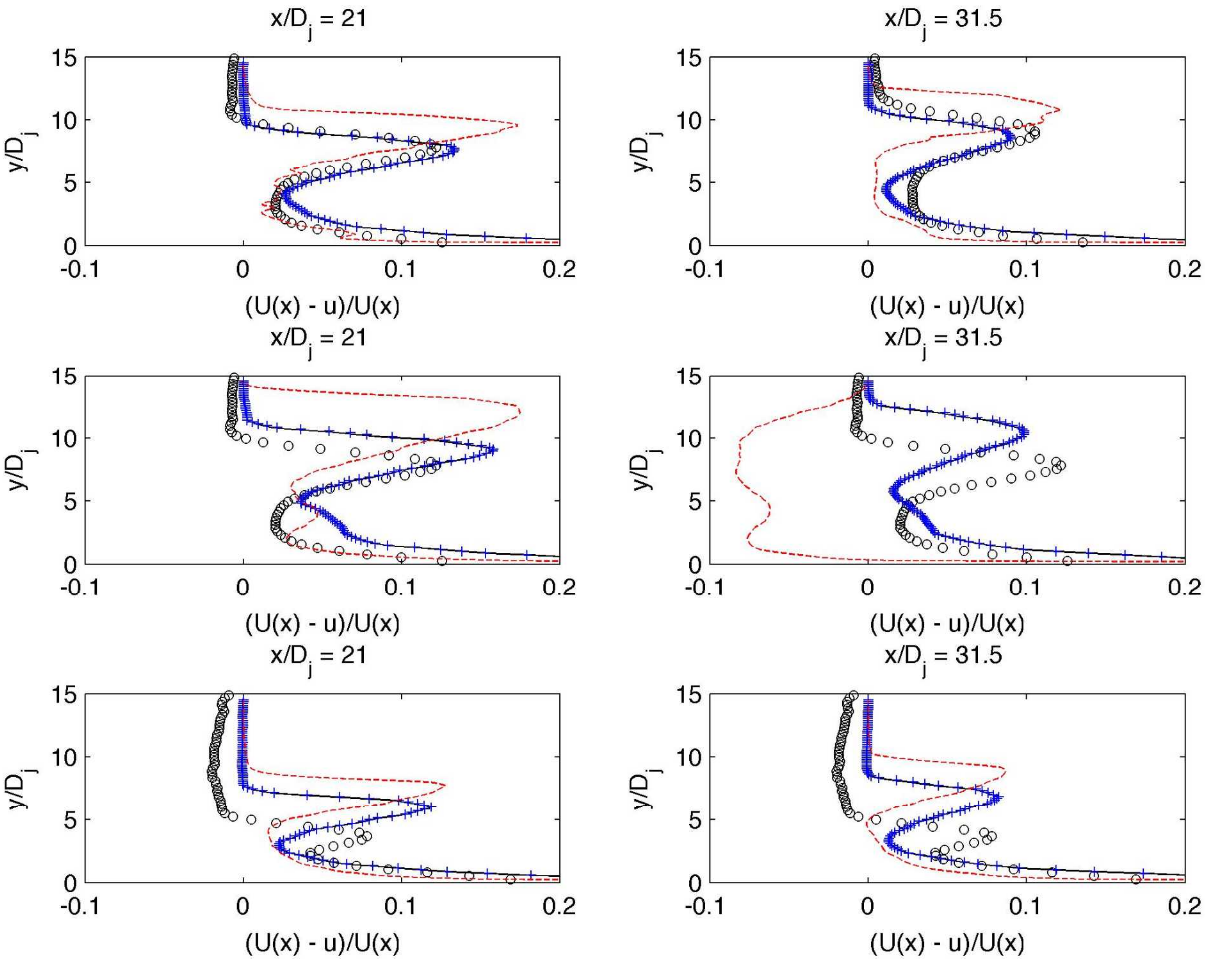
Vorticity distribution

- Keep M constant and vary J
- Use "point-vortex" metrics to compute an optimal $\mathbf{C}_{opt}$
- $\mathbf{C}_{opt}$= {0.1025, 2.099, 1.416}
- Plot predictions with $\mathbf{C}_{opt}$ for comparison



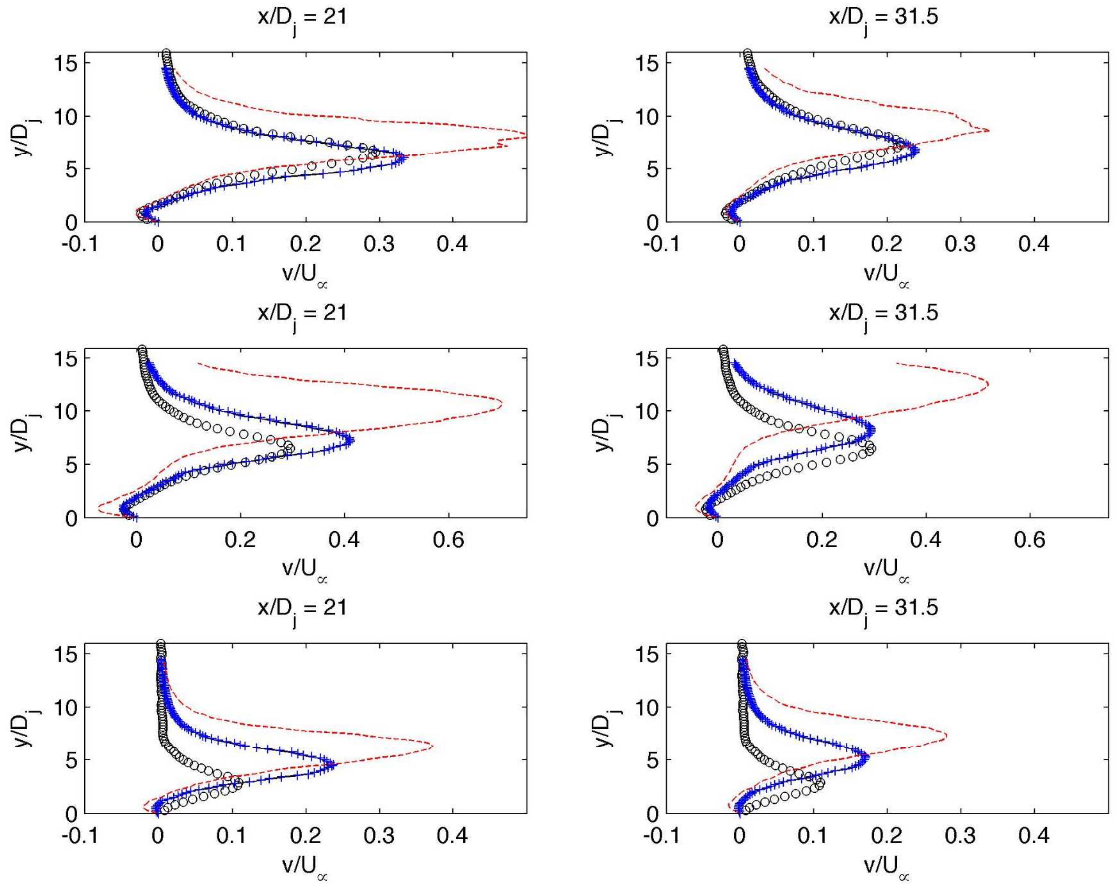
Streamwise velocity deficit

- Keep M constant and vary J
- Computed and compared on the midplane
- NOT used in the calibration
- Compared at 2 locations
  - Experiment, ensemble mean & nominal
- Improvement in predictions persists at off-calibration points



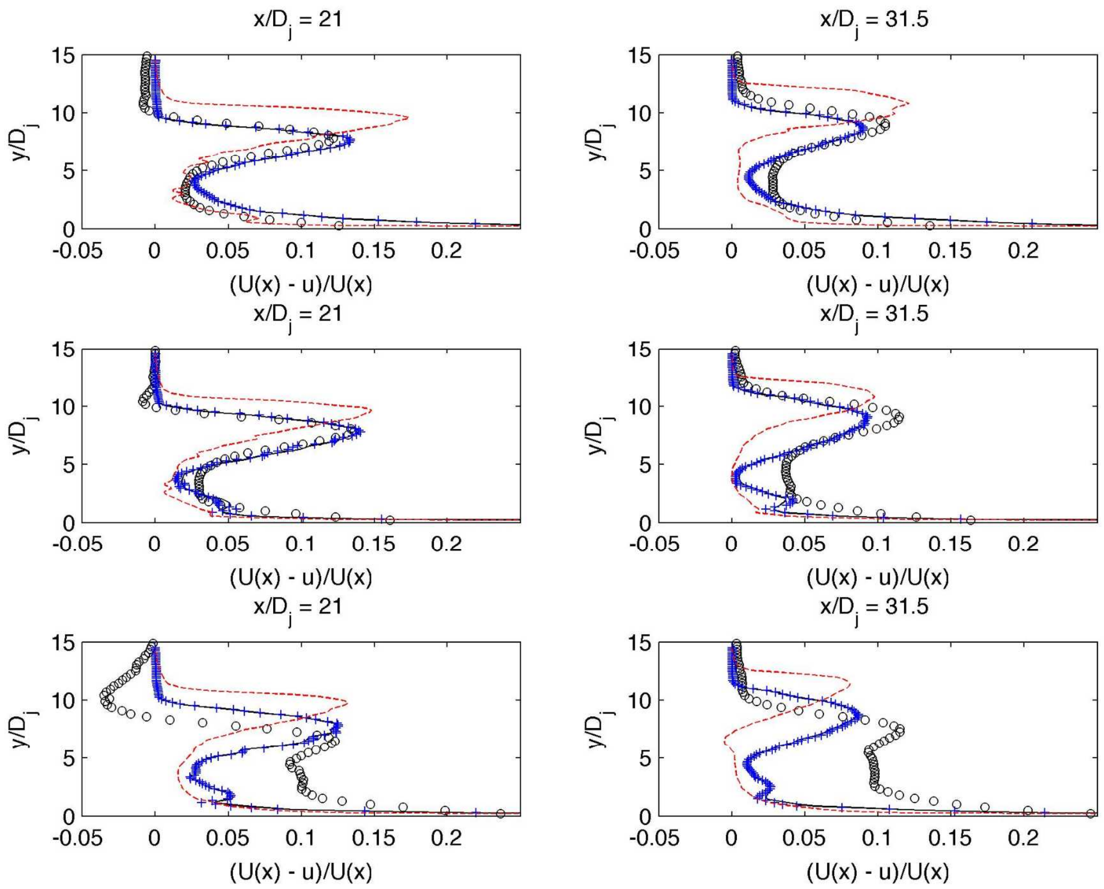
Vertical velocity distribution

- Keep M constant and vary J
- Compared at 2 locations
  - Experiment, ensemble mean & nominal
- Extremely good agreement
- Governed mostly by streamwise vorticity



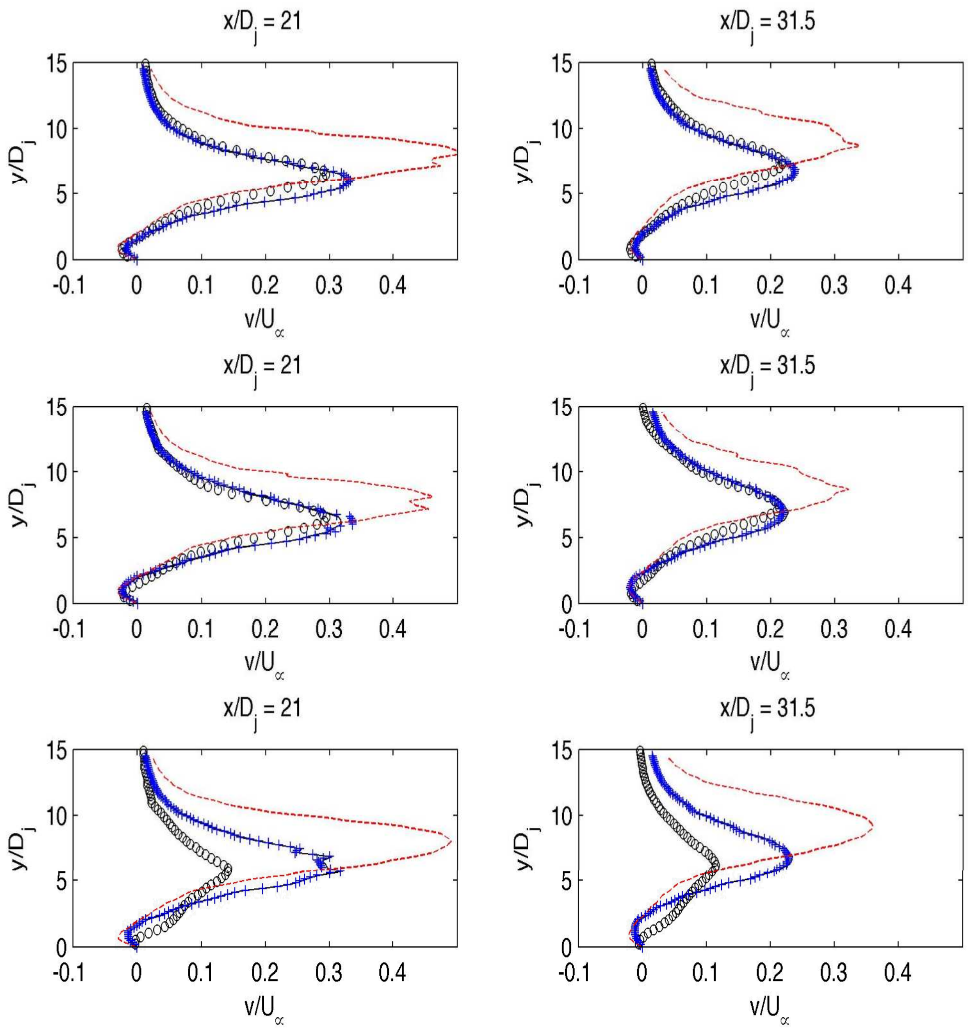
Streamwise velocity deficit

- Keep J constant and vary M (0.8, 0.7, 0.6)
- Computed and compared on the midplane
- NOT used in the calibration
- Compared at 2 locations
  - Experiment, ensemble mean & nominal
- Improvement in predictions persists at off-calibration points

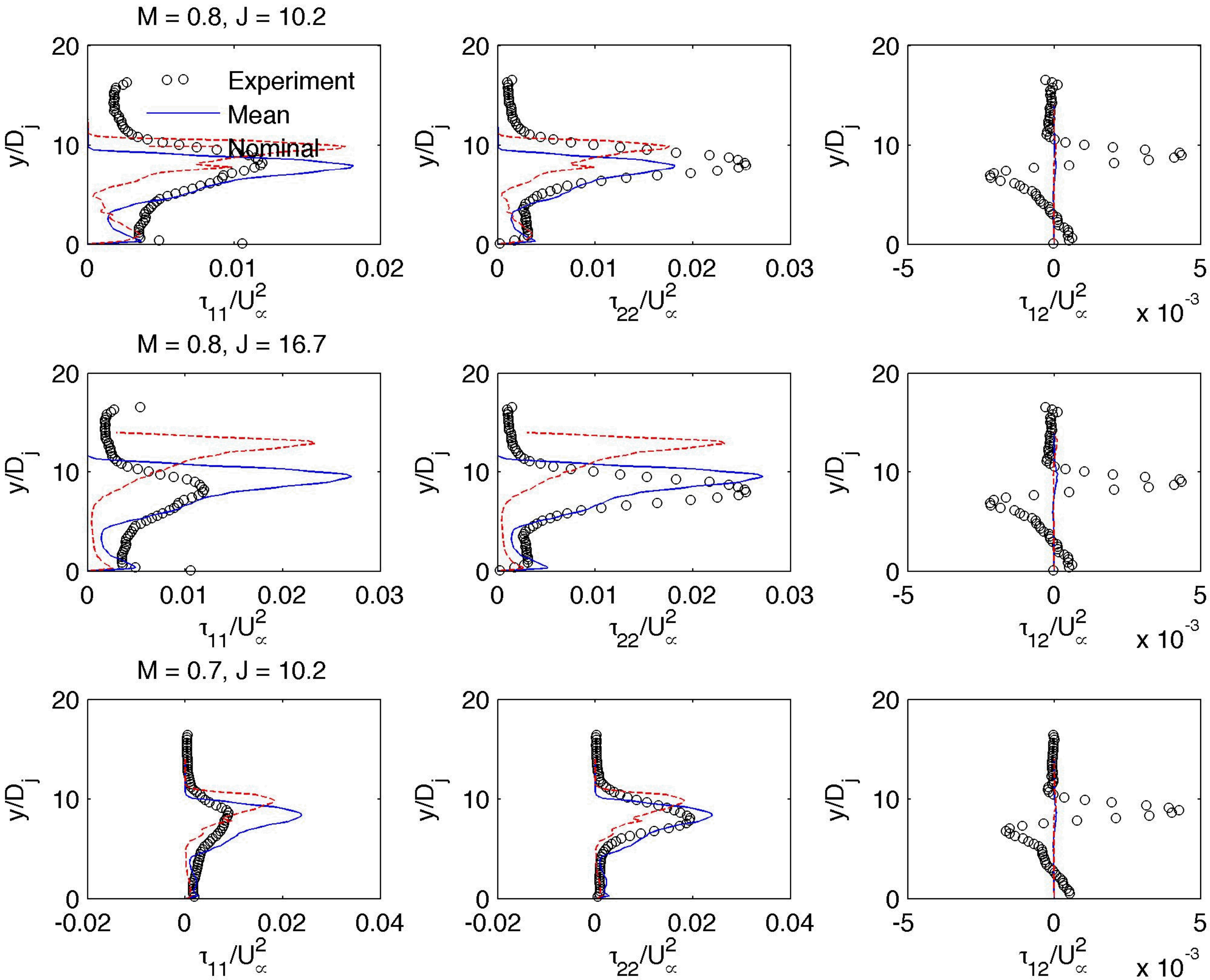


Vertical velocity distribution

- Keep J constant and vary M (0.8, 0.7, 0.6)
- Compared at 2 locations
  - Experiment, ensemble mean & nominal
- Extremely good agreement
- Governed mostly by streamwise vorticity



Turbulent stresses



- Compare normal (τ_{11}, τ_{22}) and shear stresses(τ_{12})
- Strain-rates have very little effect on the stresses – it's mostly 2/3k

Conclusions

- We have developed a way of calibrating RANS models for predictive jet-in-crossflow simulations
  - Based on surrogate models and Bayesian inference
  - Predictions are probabilistic – we predict using an ensemble
- The primary cause of inaccurate RANS JinC predictions was an unsuitable C_{nom}
  - Calibration to flow-over-square cylinder largely fixed it
  - Calibration to JinC data revealed model error – and it's not much, comparatively
  - Calibrated joint PDF predictive even at off-calibration flow interaction
- Cause of model-form error – the linear eddy viscosity model we use