Noise TFG: Using Ensemble Averaging to Accelerate Uncertainty Quantification in Automotive Aerodynamics Simulations
ID59

Paul Norman

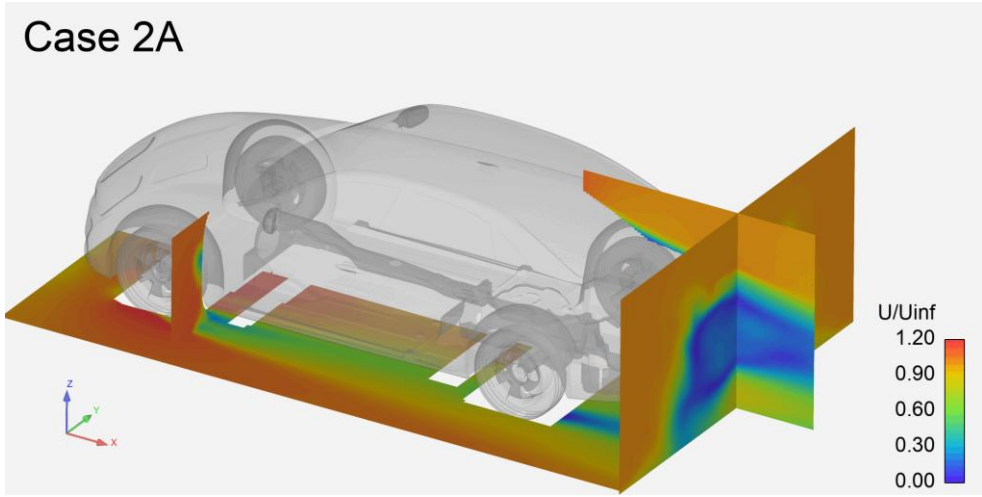
CFD Technical Specialist

Ford Motor Company

Results Overview: ID59

Solver: Icon CFD 5.0.5, all simulations run for 7 s

Config	C_D	C_L	C_{LF}	C_{LR}	Cs
2a	0.2710	0.0257	-0.0702	0.0959	0.0240
2b	0.2633	0.0340	-0.0591	0.0931	0.0273
difference	0.0077	-0.0083	-0.0111	0.0028	-0.0033



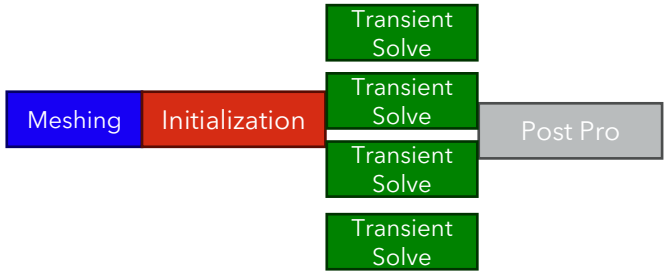
Benefits and Applications for Ensemble of *N* simulations

Potential parallel speed up by a maximum of *N*. Useful when strong scaling limits are reached or there are logistical reasons for running on a small number of CPUs.

Normal CFD Process



Ensemble CFD Process



*Boxes are not to scale
**May want to run multiple meshes/initializations, depending on set up

Benefits and Applications for Ensemble of *N* simulations

Useful for repeat runs, which are often used for high noise vehicles/configs to reduce uncertainty.

Normal CFD Process



Repeat Ensemble CFD Process

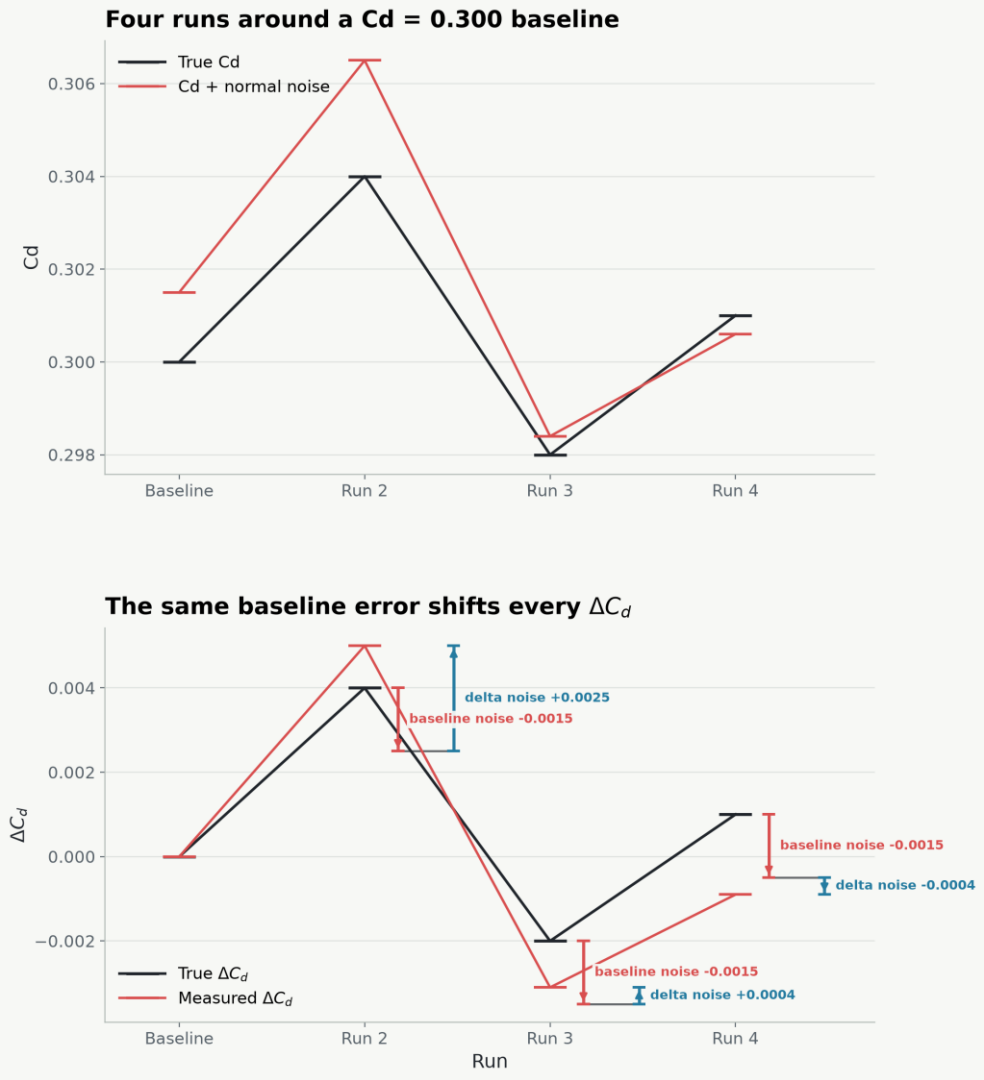


*Boxes are not to scale
**May not want to run multiple meshes/initializations, depending on set up

Benefits and Applications for Ensemble of *N* simulations

Baselines – typically compared to many simulations during development and require fast turn around.

For fixed computation budget and N simulations, minimize error in $\sum|\Delta C_D|$ by running the baseline $\sqrt{N}$ longer than individual aero actions (or run an ensemble of $\sqrt{N}$ simulations)*.



* See backup slides for details

Benefits and Applications for Ensemble of *N* simulations

Potential architectural speedups for FVM CFD codes designed to run ensemble cases[1]

$$\begin{pmatrix} A & & \\ & A & \\ & & \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} \qquad A(\mathbf{u}_1 \;\; \mathbf{u}_2)$$

Solving an ensemble solution turns many sparse matrix-vector products into one matrix-matrix product, which cuts data movement and has the potential to speed up LES.

[1] “Ensemble averaging parallel-in-time approach for industrial LES.” Plana-Rui, J.l Trias, X.; Al (asalti-Baldellou, A.; Colomer, G.; Oliva, A. Journal of Physics: Conference Series 2766 (2024) 012089

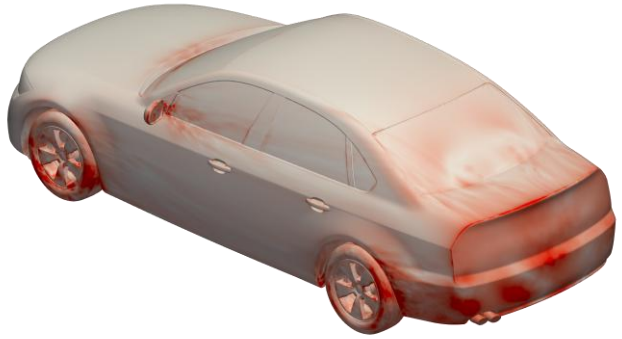
Benefits and Applications for Ensemble of *N* simulations

Individual ensemble make it straight forward to calculate standard deviations of mean values.

- Surface/Volume variables
- Integral values

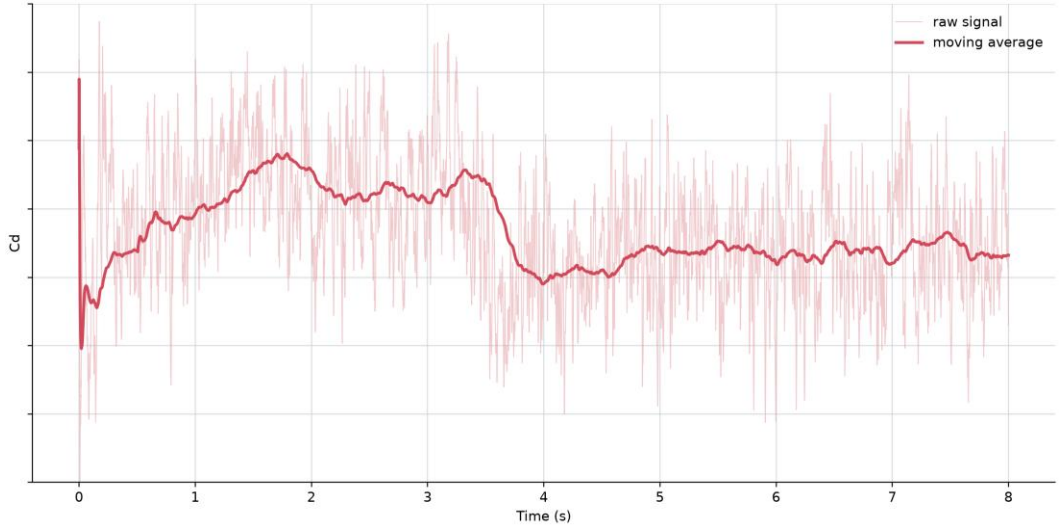


Standard Deviation Surface Cp:



Potential Challenges and Drawbacks

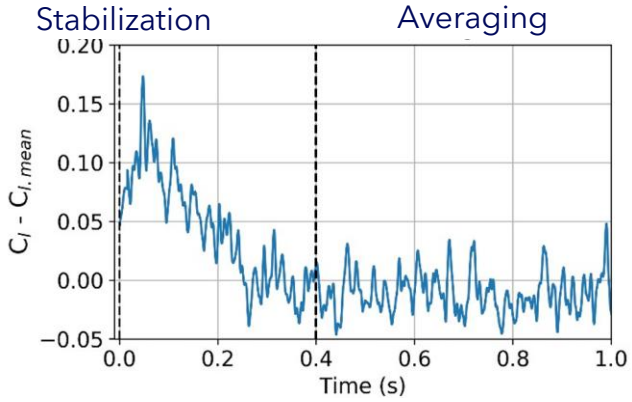
- Many short signals may not represent one long realization, particularly for low-frequency or bimodal behavior. Risks violating the ergodic assumption.



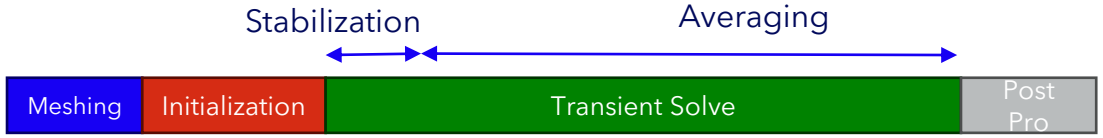
Potential Challenges and Drawbacks

Each simulation must be initialized and stabilized before its time series is usable for averaging. This overhead is repeated (N) times.

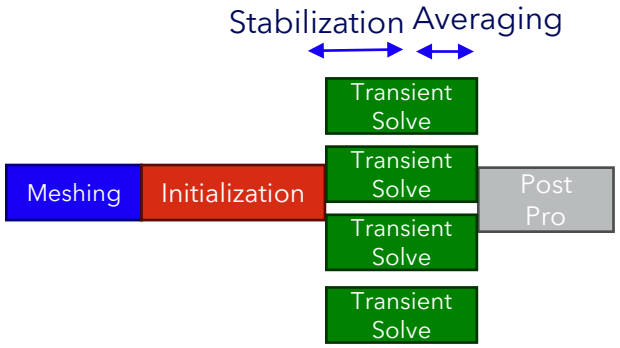
Ensemble members must be uncorrelated before averaging is started.



Normal CFD Process



Ensemble CFD Process



Ensemble Initialization Methods

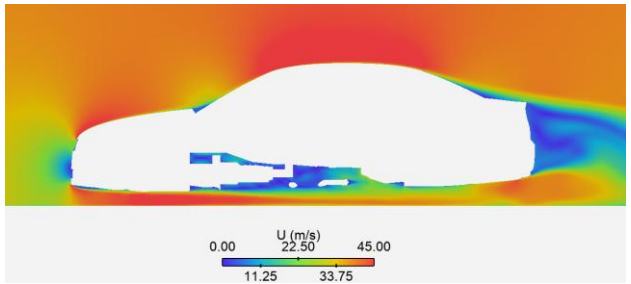
Evaluate ensemble initialization methods with goal to minimize stabilization time and correlation. Create ensemble of 5 cases.

Ensemble Initialization Metho

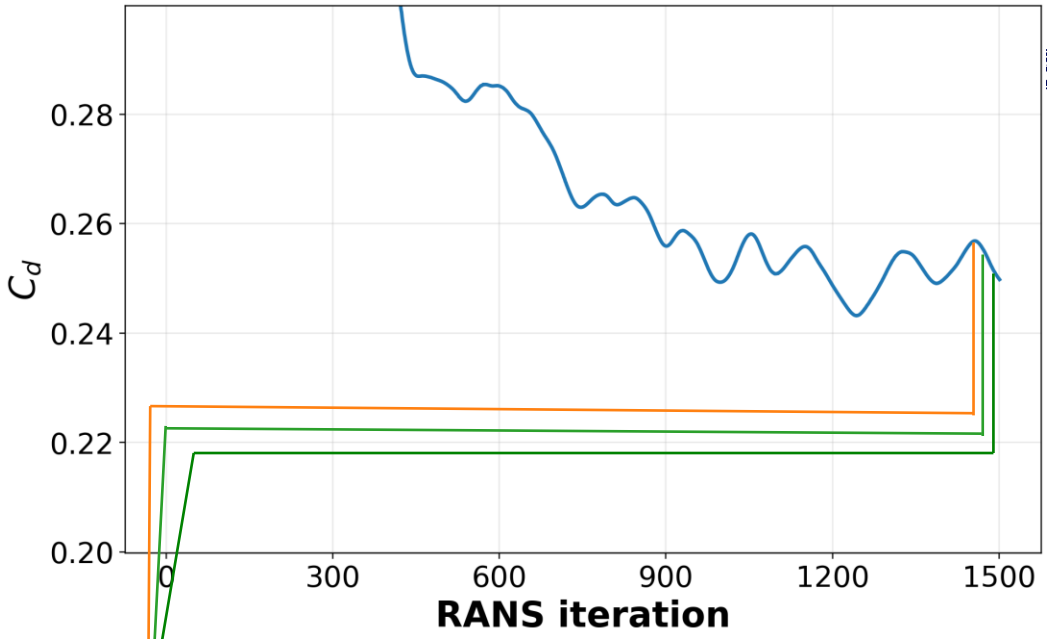
Evaluate ensemble initialization methods with goal to minimize stabilization time and correlation. Create ensemble of 5 cases.

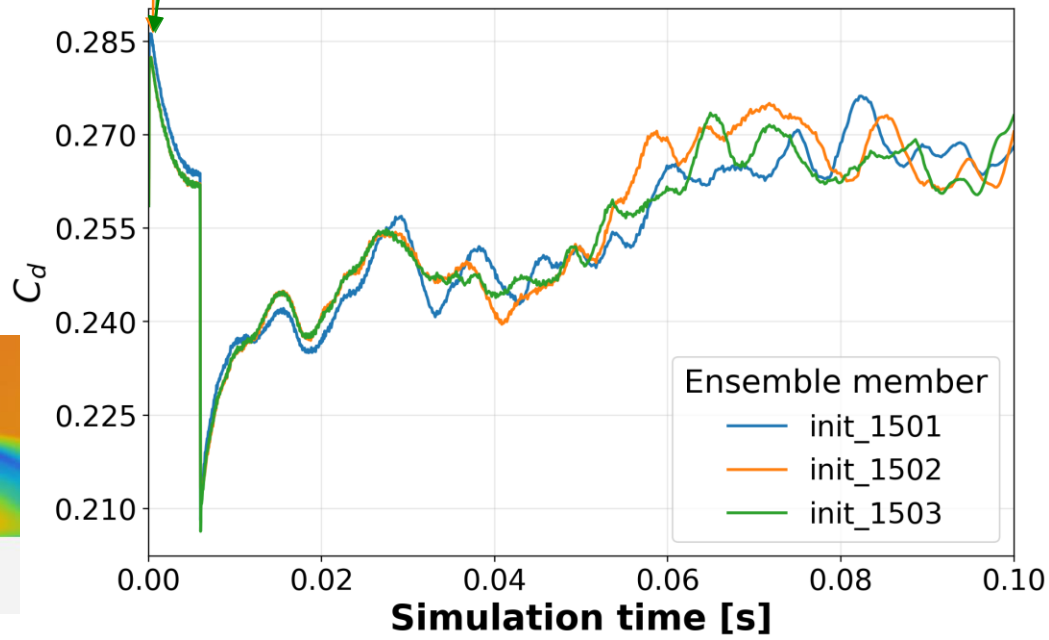
Cold Start

- Initialize from RANS simulations with different number of iterations.



RANS Init



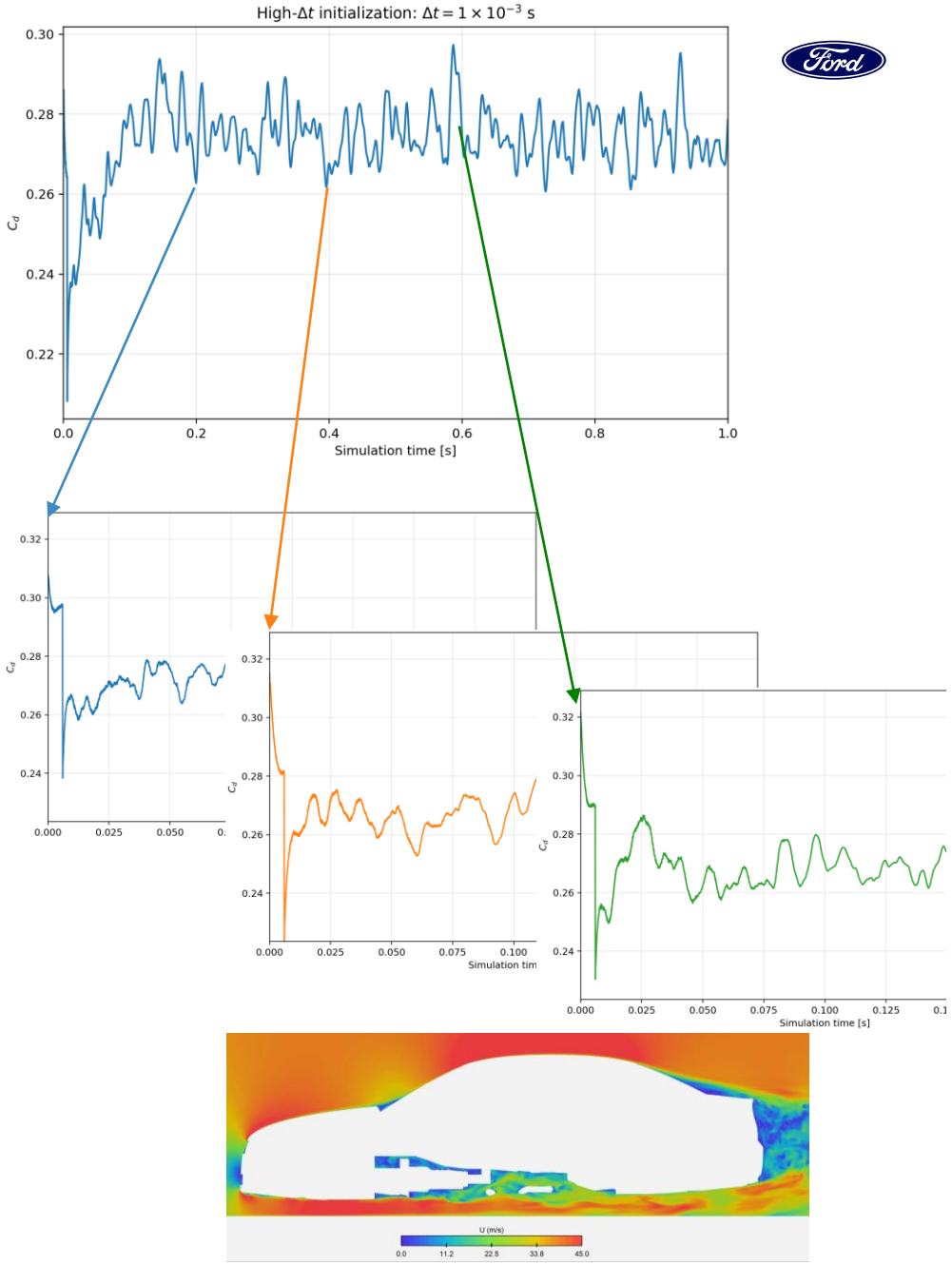


Ensemble Initialization Methods

Evaluate ensemble initialization methods with goal of minimizing stabilization time and correlation. Create an ensemble of 5 cases.

Cold Start

- Initialize from RANS simulations with different number of iterations
- Initialize from DES run at higher time step (2×10^{-3}, 1×10^{-3}, 5×10^{-3} s) for 1000 steps, in intervals of 200 steps.



High TS Init

Ensemble Initialization Methods

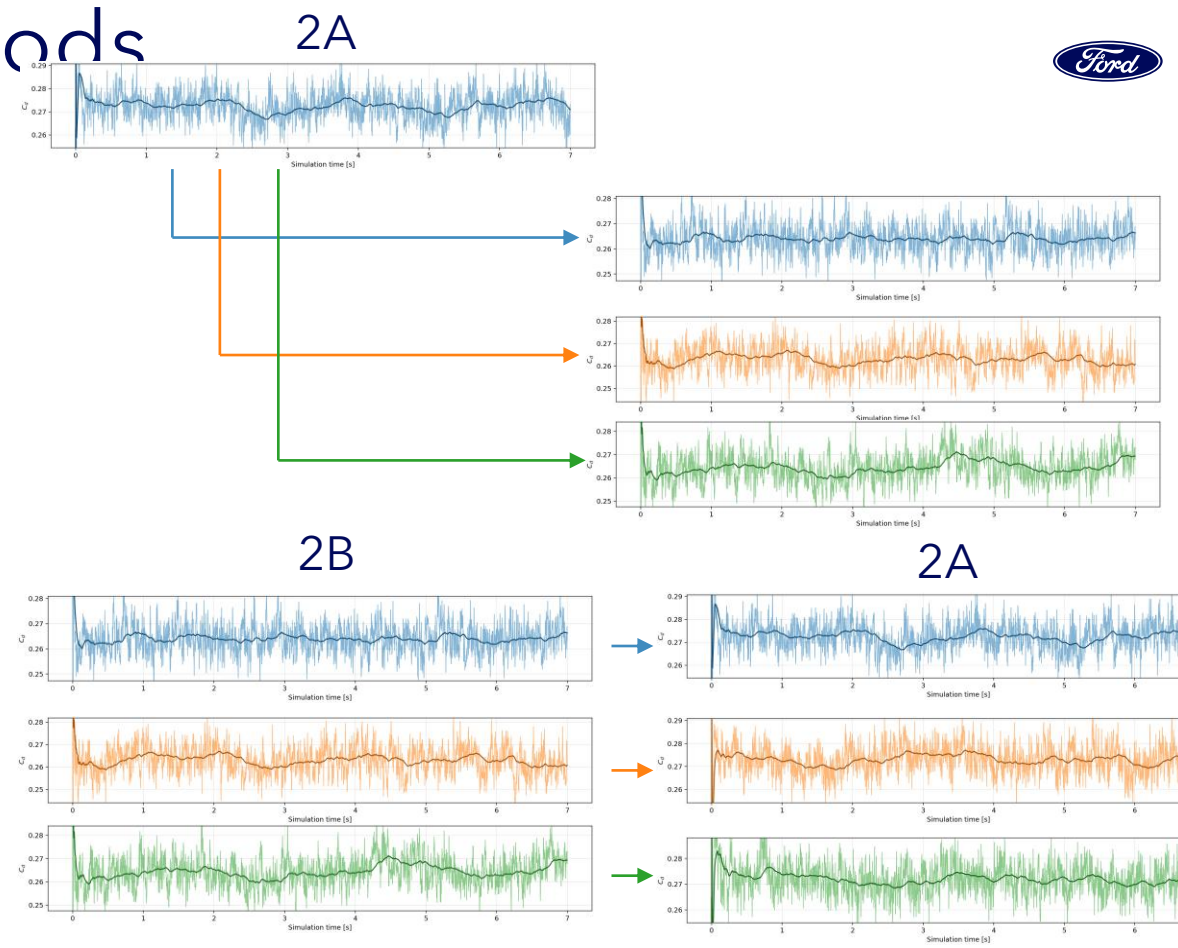
Evaluate ensemble initialization methods with goal of minimizing stabilization time and correlation. Create an ensemble of 5 cases.

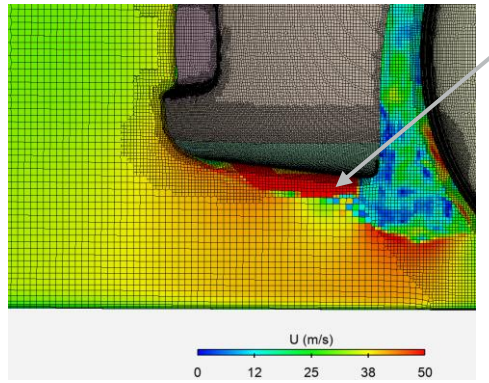
Cold Start

- Initialize from RANS simulations with different number of iterations
- Initialize from DES run at higher time step (2×10^{-3}, 1×10^{-3}, 5×10^{-3} s) for 1000 steps, intervals of 200 steps.

Warm Start

- Initialize from map of transient solution of similar simulation (Case 2A ↔ 2B).
  - Map from 0.5s intervals from Case 2A
  - Map from end of 2B ensemble runs back to 2A

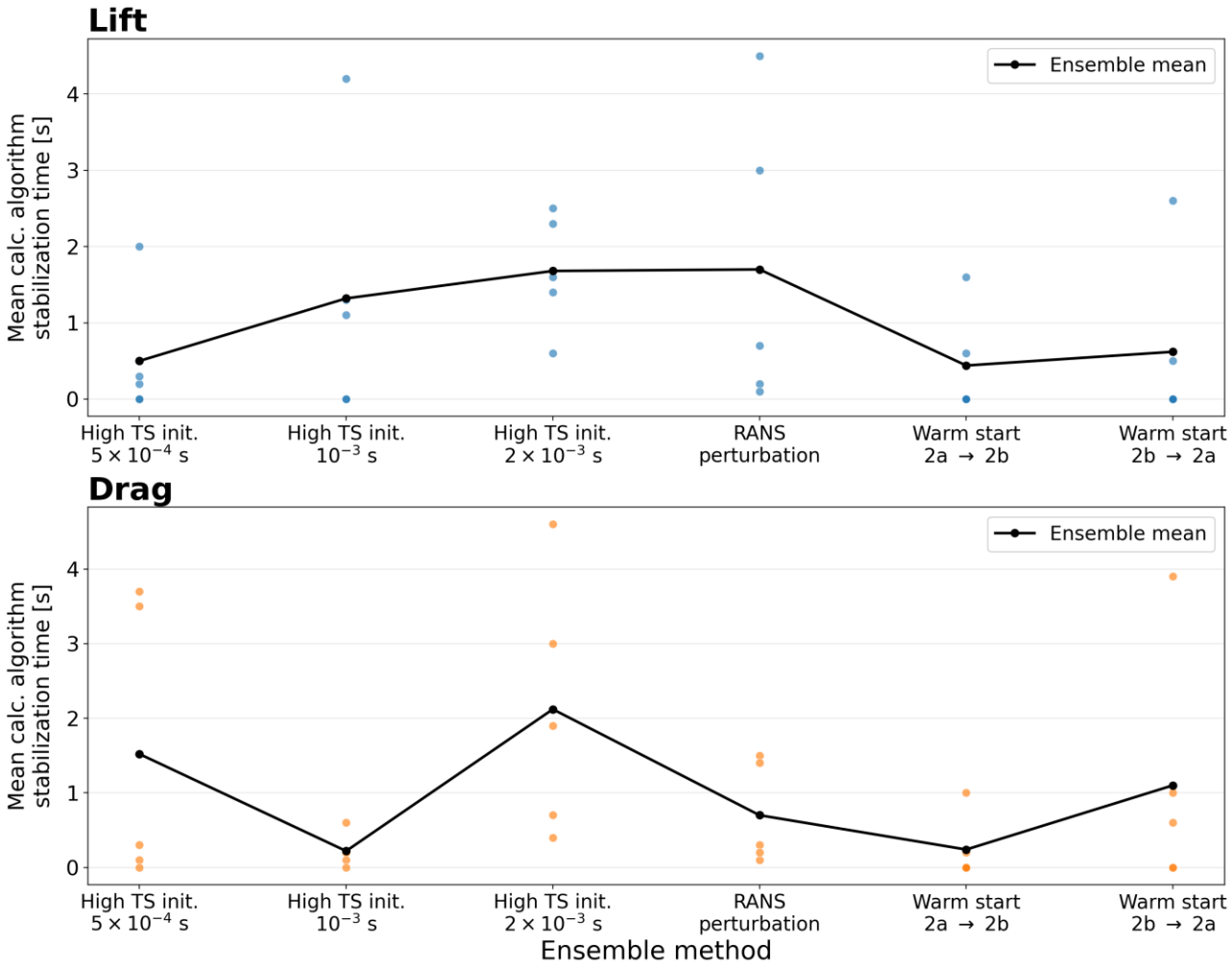


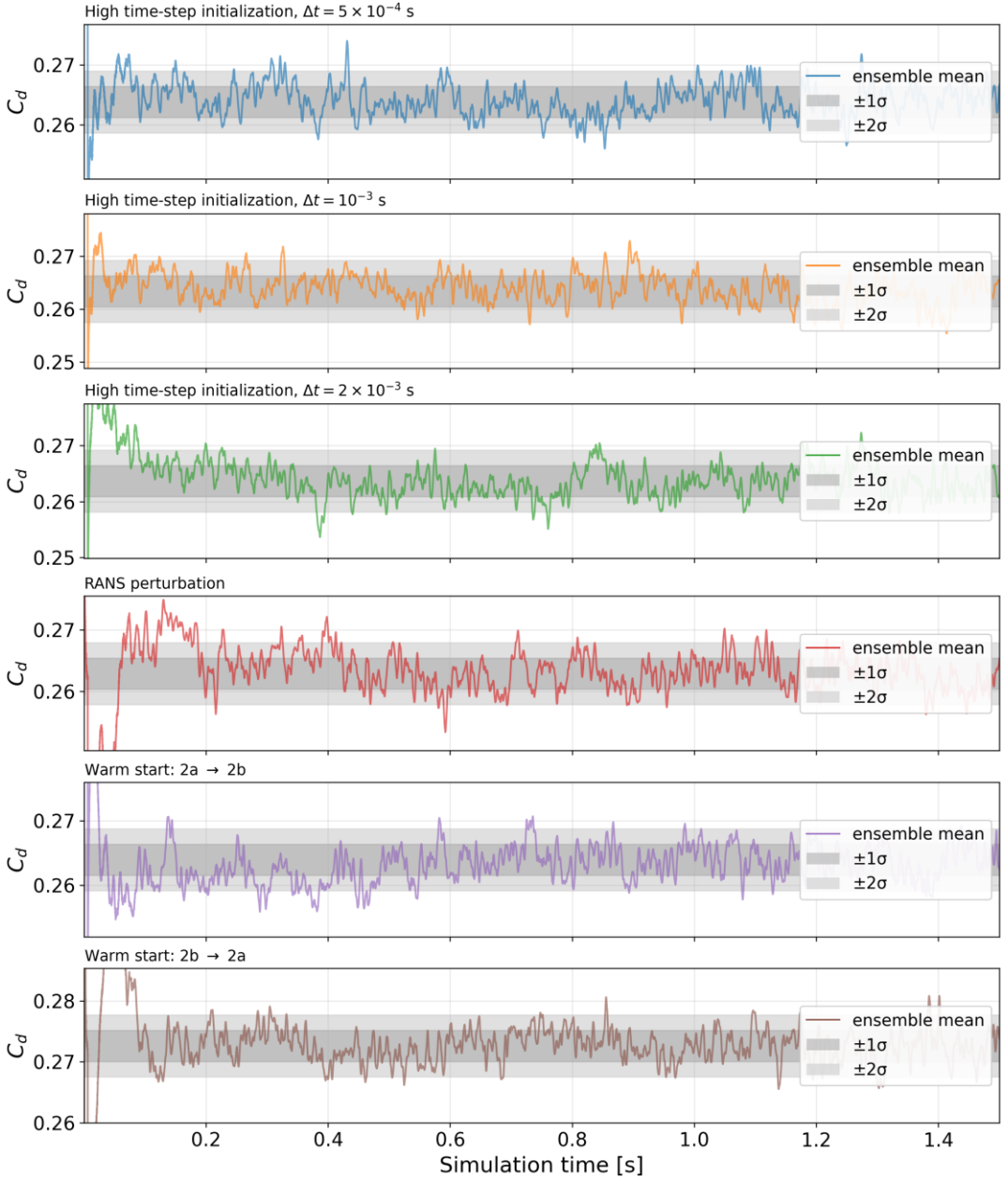


High velocity used for locations with no map

Mapped Init

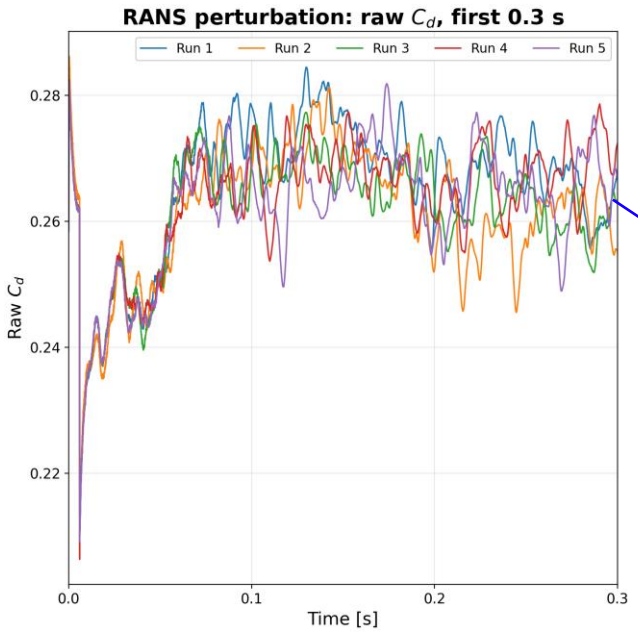
Stabilization Analysis

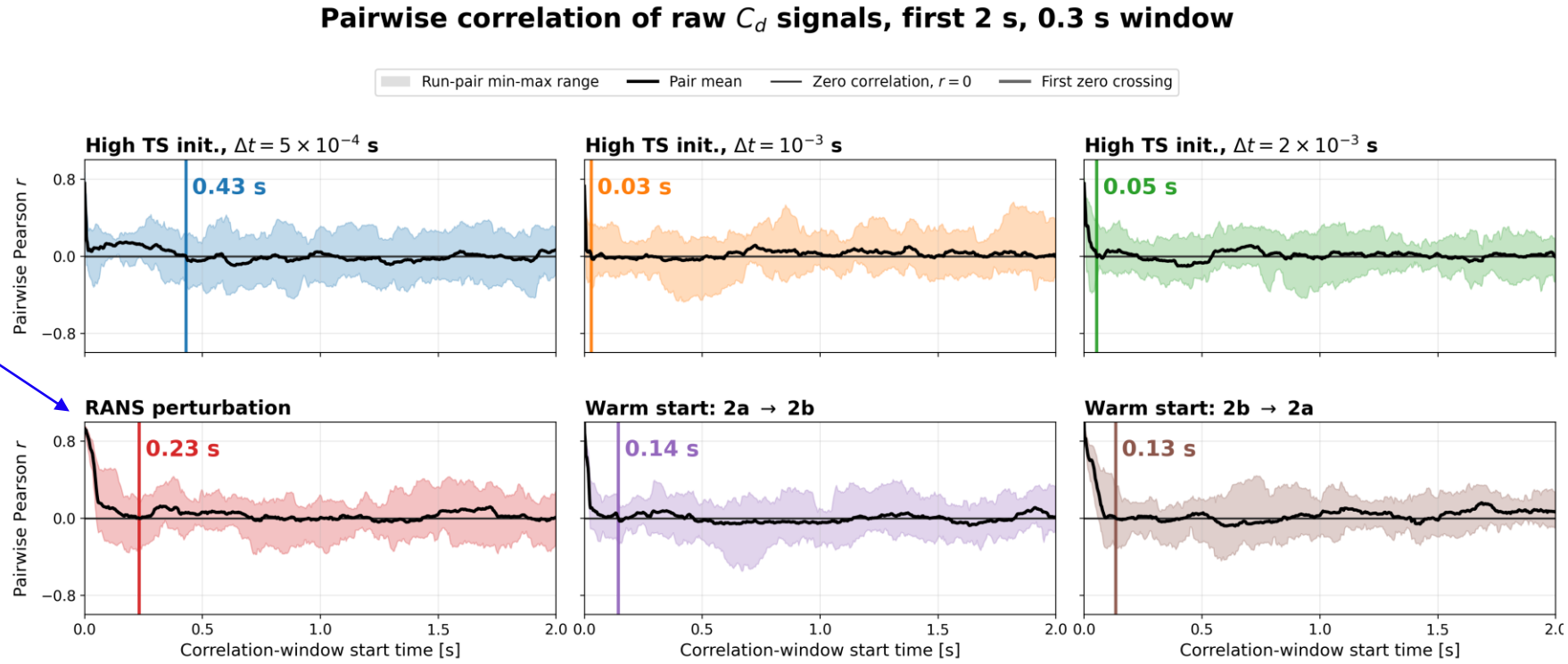




- No method is a clear winner or loser for stabilization time, although warm start is slightly lower than others

Correlation Analysi





$$t_{\text{eff}} = T\left(\frac{S_{\bar{C}_d}}{\sigma_{C_d}}\right)^2 \qquad 10 \times t_{eff} \cong 0.3$$

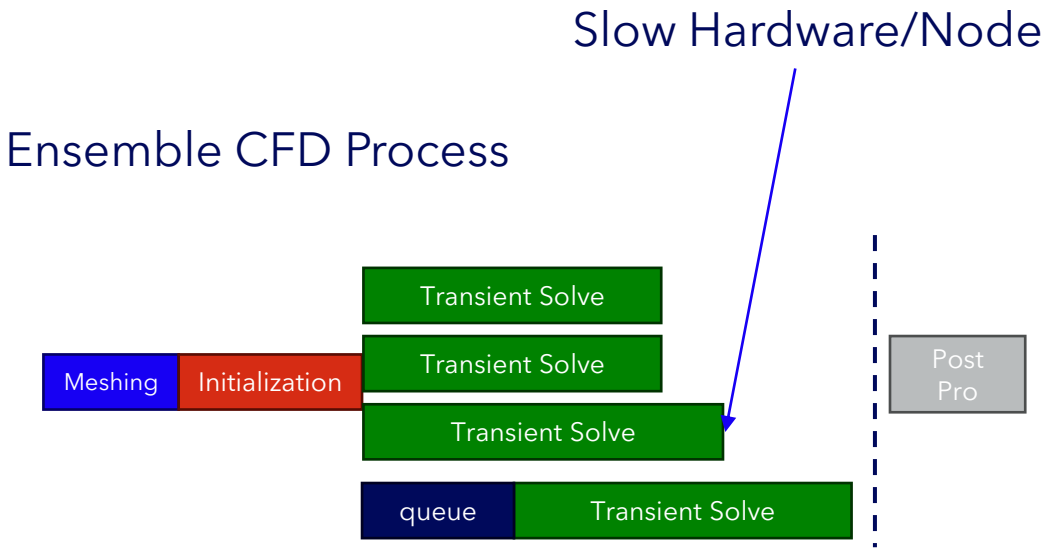
- All method show fast decorrelation, with high time step methods (1×10^{-3}, 2×10^{-3}) being the fastest.
- The decorrelation times are shorter than the MeanCalc algorithm stabilization times for this case.
- Averaged Cd coefficients of ensemble members show deviation equal or greater than time series error, showing that no correlation of average coeffs.
- A better test would be on a noisier case, especially one with strong low frequency modes that could be in phase or out of phase depending on the initialization method

Conclusions

- Running simulations longer lowers the uncertainty in mean ($\propto t^{-\frac{1}{2}}$), and ensemble averaging provides a straightforward method to parallelize ($\propto N^{-\frac{1}{2}}$) .
- Evaluated different methods for initializing ensembles. There was no clear winner in stabilization or correlation time between RANS, High Time Step, and Mapped initializations.
- Decorrelation time is less stabilization time, so small perturbations (RANS initialization or alternative) are acceptable for this case.
- By computational cost cheapest method is mapped initialization.
- Test case with more noise (especially strong low frequency modes) would be a better test.

Potential Challenges and Drawbacks

- Uneven completion complicates workflows. Members can finish at different times, delaying post-processing.



Baselines

A single baseline is used in all *N* comparisons ΔC_D. Its uncertainty appears *N* times and the uncertainty in all aero actions appears once. The variance in a delta between a case *i* and the baseline is:

$$\mathrm{var}(\Delta C_{d,i,b}) = \mathrm{var}(\overline{C_{d,i}}) + \mathrm{var}(\overline{C_{d,b}}) \propto \frac{1}{t_i} + \frac{1}{t_b}$$

The sum of variances for N cases is then:

$$\sum^{N} var(\Delta C_{d,i,b}) \propto \frac{N}{t_i} + \frac{N}{t_b}$$

With a computation total time *T*:

$$T = Nt_i + t_b$$

It can be shown the minimum of $\sum^N Var(\Delta C_{d,i,b})$ occurs at $t_b = \sqrt{N}\, t_i$

So to minimize the sum of errors in the delta to baseline for 9 cases, we would want to run the baseline 3 x longer (or run 3x repeats).

Note that this does not apply to relative error (or rank order) of individual aero actions