upstreamCFD

Scale-Resolving Simulation TFG Context and Summary

C. Mockett[1], M. Fuchs[1], S. Helgesson[1], H. Hetmann[1], A. Walle[2]

1 upstreamCFD

2 PASTEUR LABS

Introduction

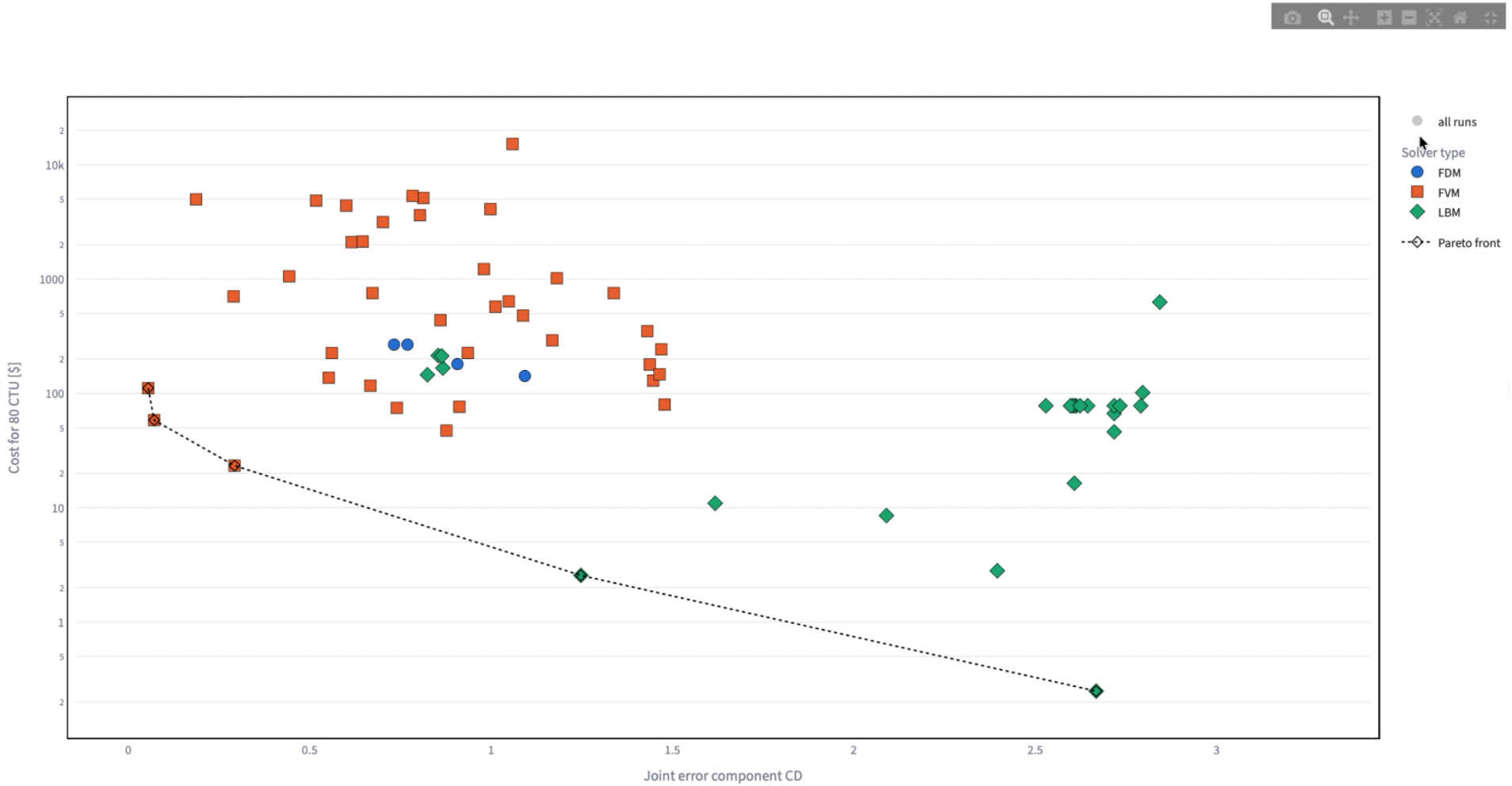
- What's new since AutoCFD4
  - Combined error metric, realising the "HPC cost-vs-error" diagram vision
  - Implemented in the dashboard, together with other upgrades
- Scope of talk
  - Focus on case 2a only
    - Case with the most contributions
    - More than enough for 15-minute presentation
- Overview of case 2 submissions
  - 20 participant IDs: 87 runs for 2a, 58 for 2b
  - Method breakdown (case 2a):
    - 65 hybrid RANS-LES
    - 13 LES
    - 9 (U)RANS
    - Finite-volume, lattice-Boltzmann and finite-difference codes

Animation of Q criterion of AutoCFD case 2:
(Advanced DDES methods, OpenFOAM)



n.b. We are still clarifying issues with the data. Presented results are provisional

Upgrades to the AutoCFD dashboard (screen recording)



provisional

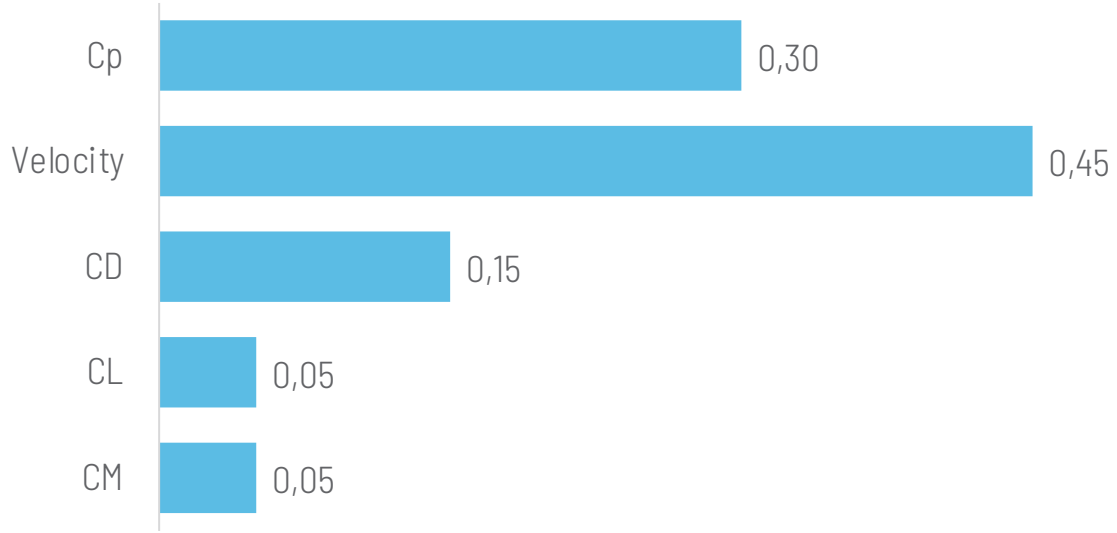
Joint error metric

- Motivation
  - Facilitate analysis: one error value per run
  - Combine multiple quantities consistently, with user-defined weights
- Formulation similar to interlaboratory proficiency tests
  - ISO 13528
  - Deviations to experiment in units of the typical CFD scatter
- Four steps
  1. Deviation d per point: (CFD – experiment)
  2. Typical scatter per point σ_{pt} computed with robust standard deviation ("Algorithm A") and pooled per quantity ("component")
  3. Component error e_k normalised by σ_{pt}
  4. Joint error E = weighted sum of the five components
- Weights adapted from the AI/ML TFG, adjustable in the dashboard
- $e = 1$ means one typical participant spread deviation to the experiment
- Scales can be frozen and applied in future workshops

$$e_k = \sqrt{\frac{\sum_{i \in I_k} d_i^2}{\sum_{i \in I_k} \sigma_{\mathrm{pt},i}^2}}, \qquad E = \sum_k w_k \, e_k$$

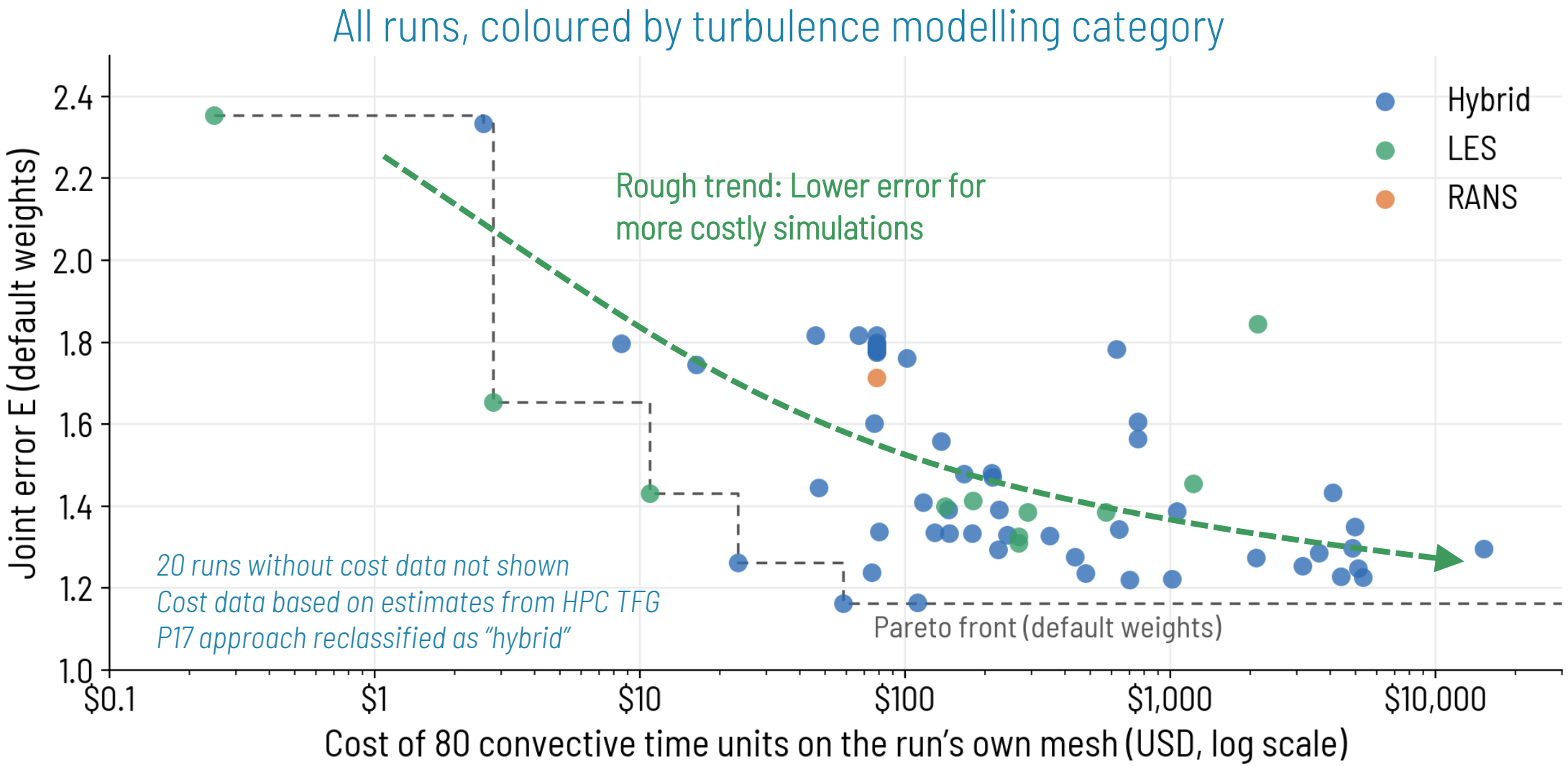
d: CFD – experiment · σ_{pt}: robust spread of all runs · I_k: points of component k · w_k: weight

Default component weights
based on AI/ML TFG's scoring procedure:



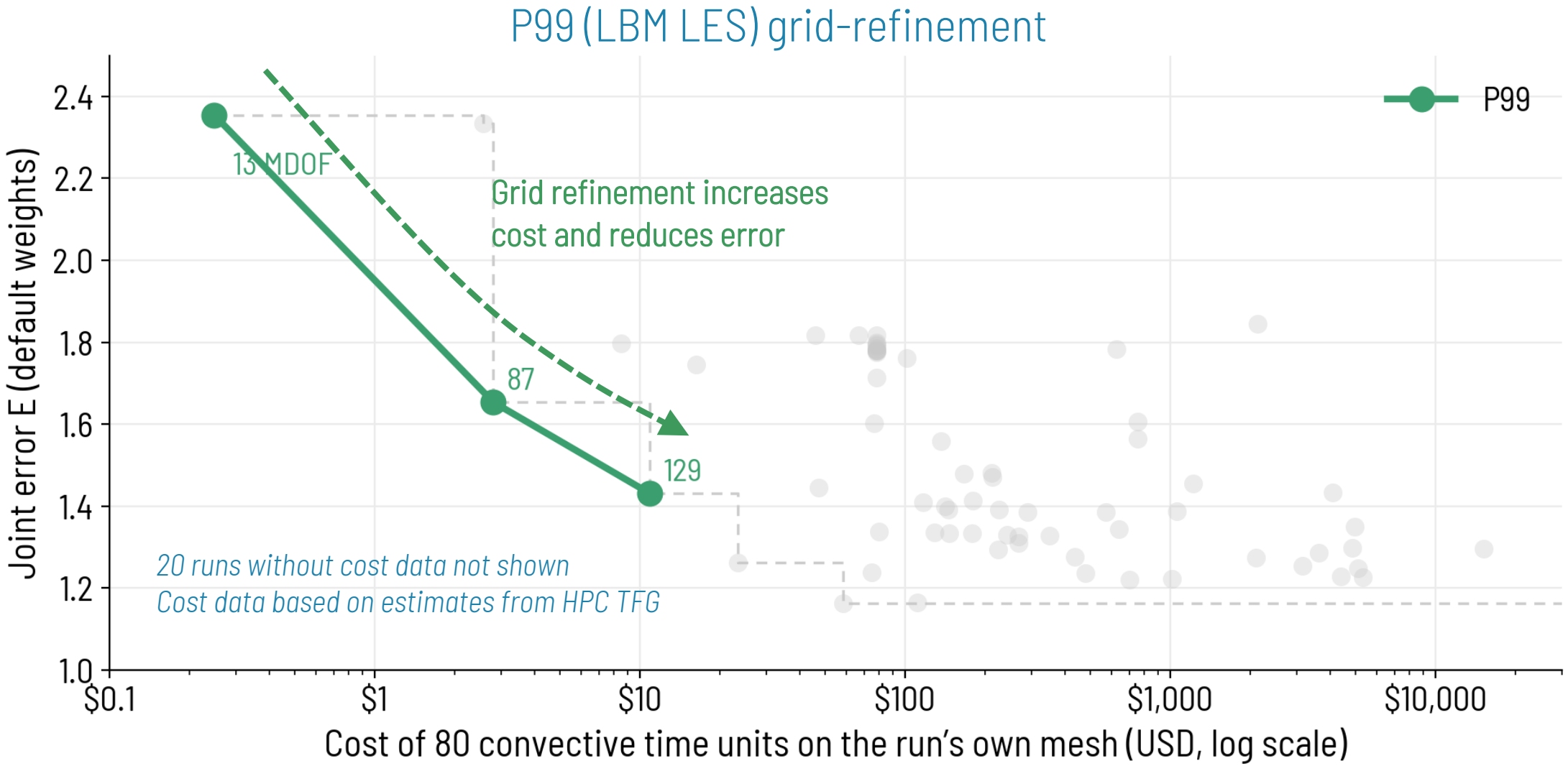
Details: H. Hetmann, C. Mockett (2026)
arXiv link

Cost vs. error diagram, all case 2a submissions with HPC cost data



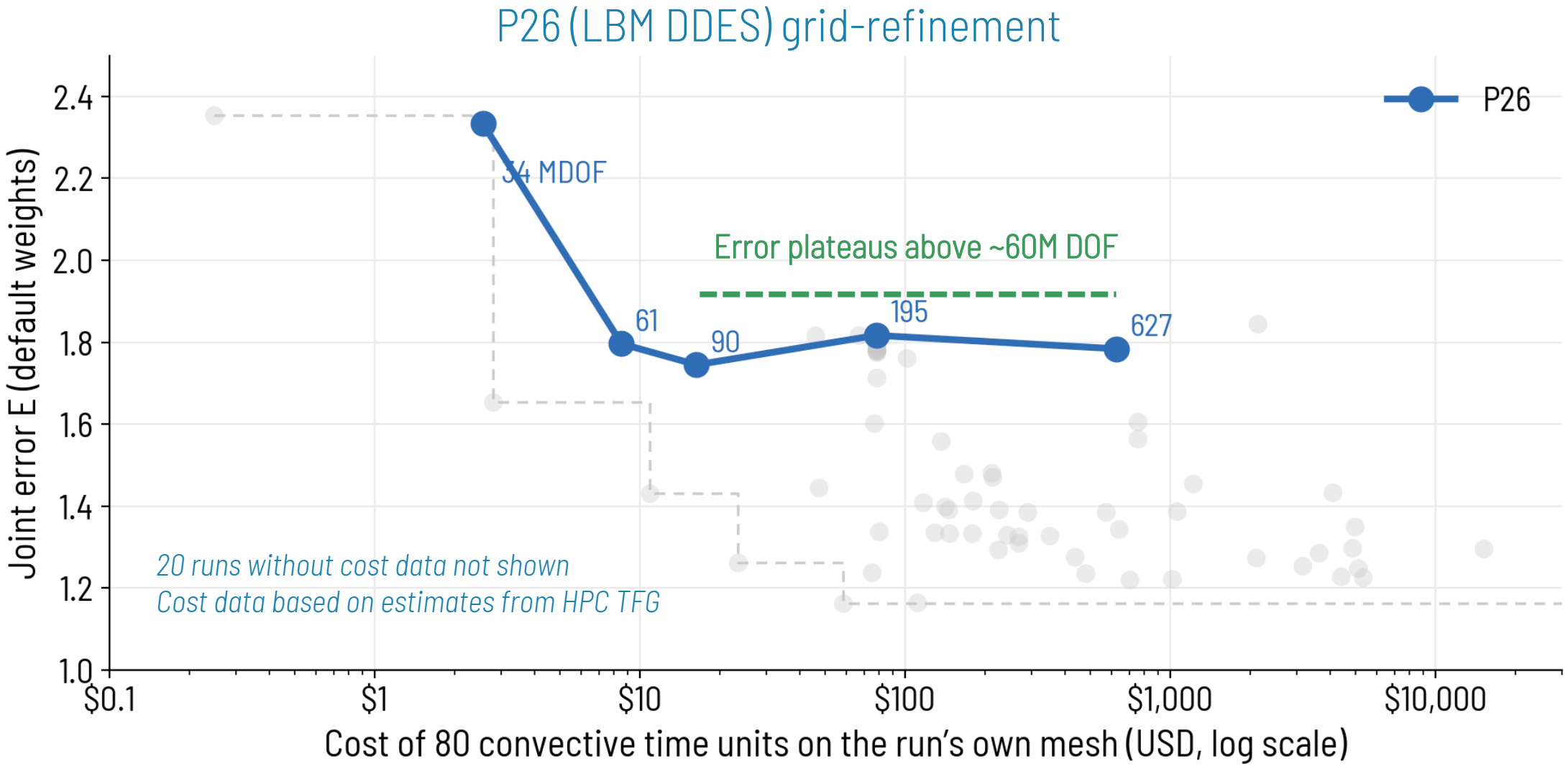
provisional

Cost vs. error diagram, all case 2a submissions with HPC cost data



provisional

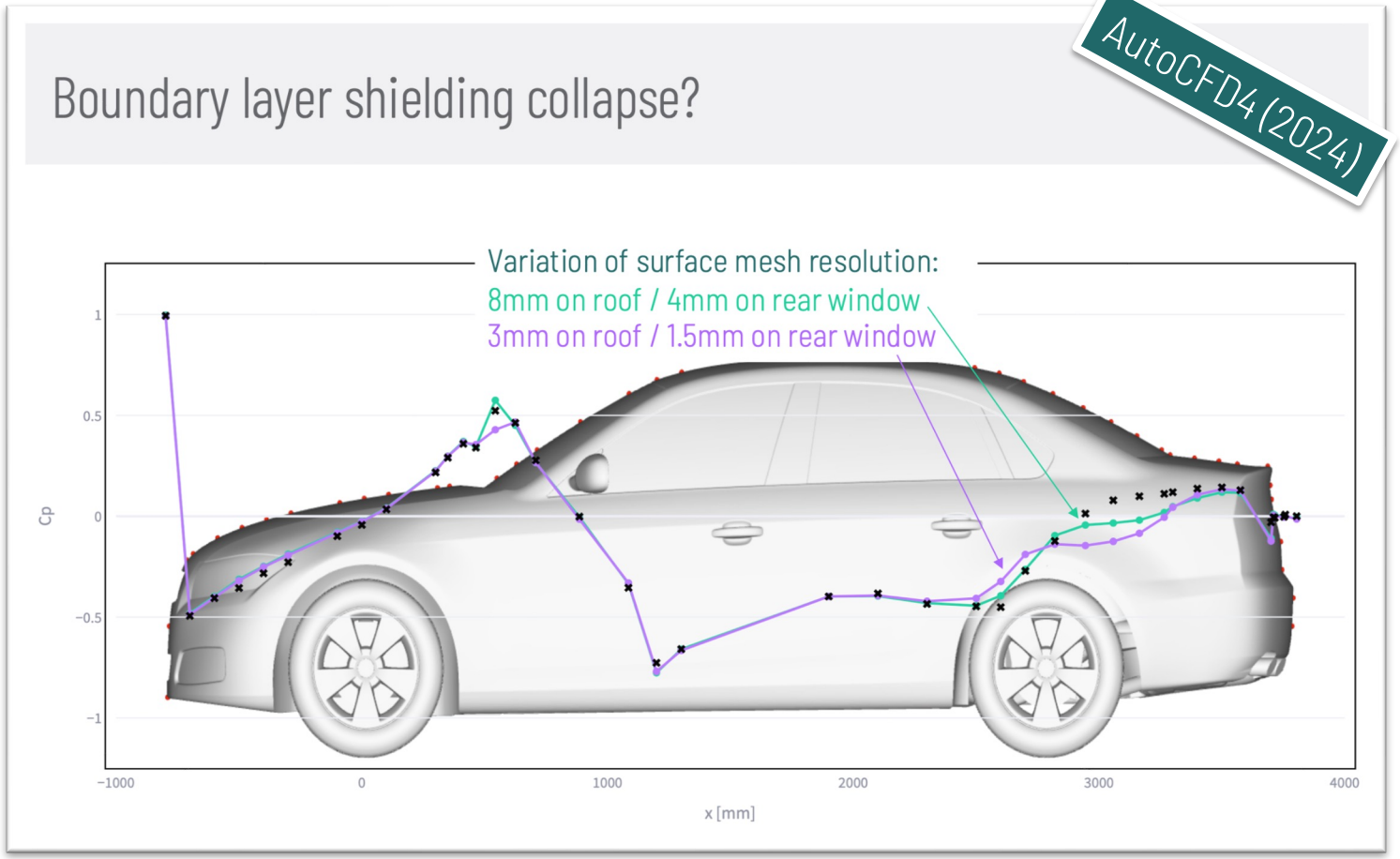
Cost vs. error diagram, all case 2a submissions with HPC cost data



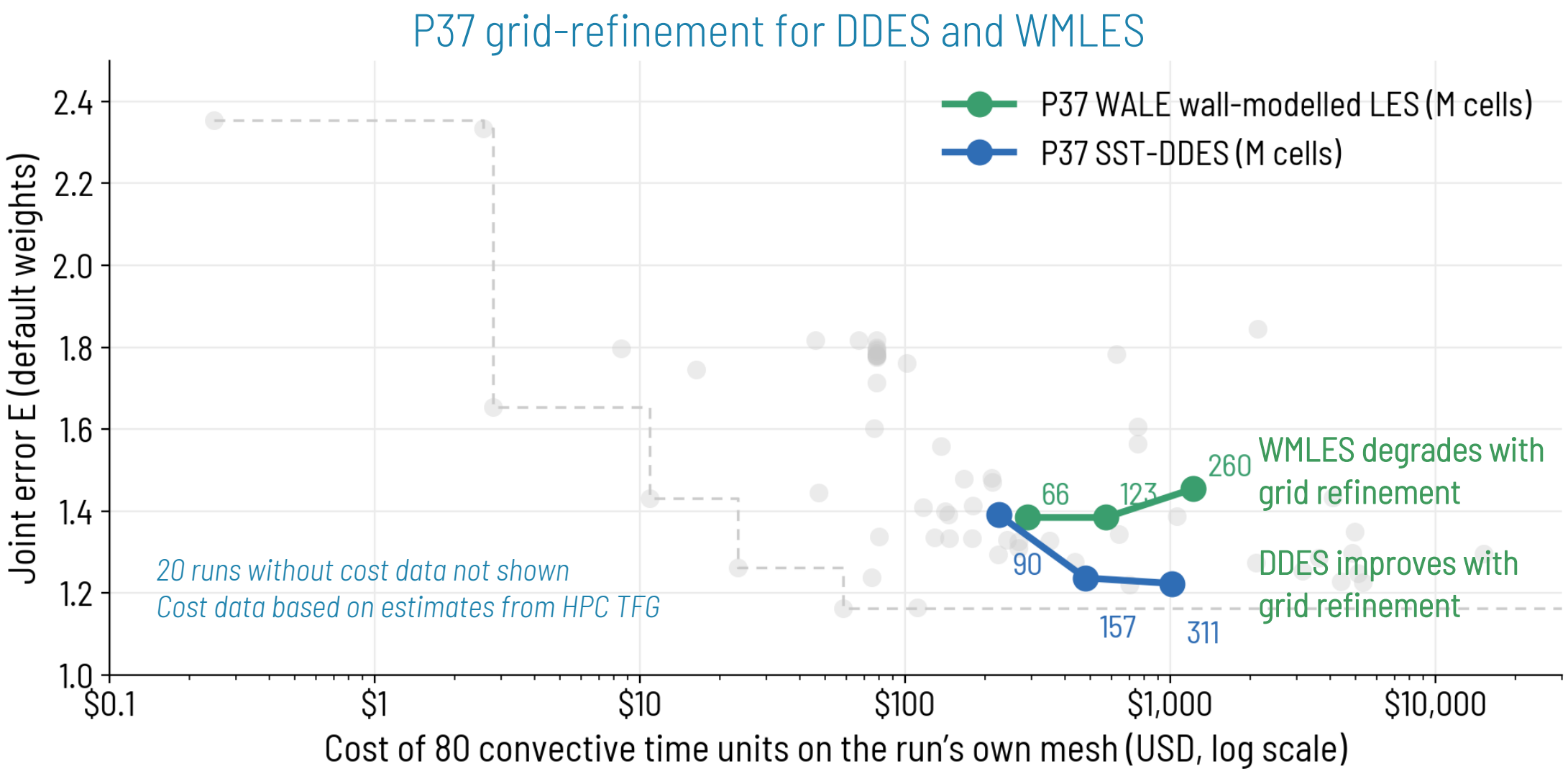
provisional

DES detail – grid-induced separation: No signs found in AutoCFD5!

- Grid-induced separation (GIS):
  - A problem in DES-like methods on fine grids
  - LES mode encroaches inside the boundary layer, causing skin friction deficit
- We found indications of grid-induced separation in DDES results in AutoCFD4
- This time round we included refined committee grids and looked for signs of GIS
  - No indications found
  - Progress! ☺
  - Looking forward to learning how participants solved the problem in their codes

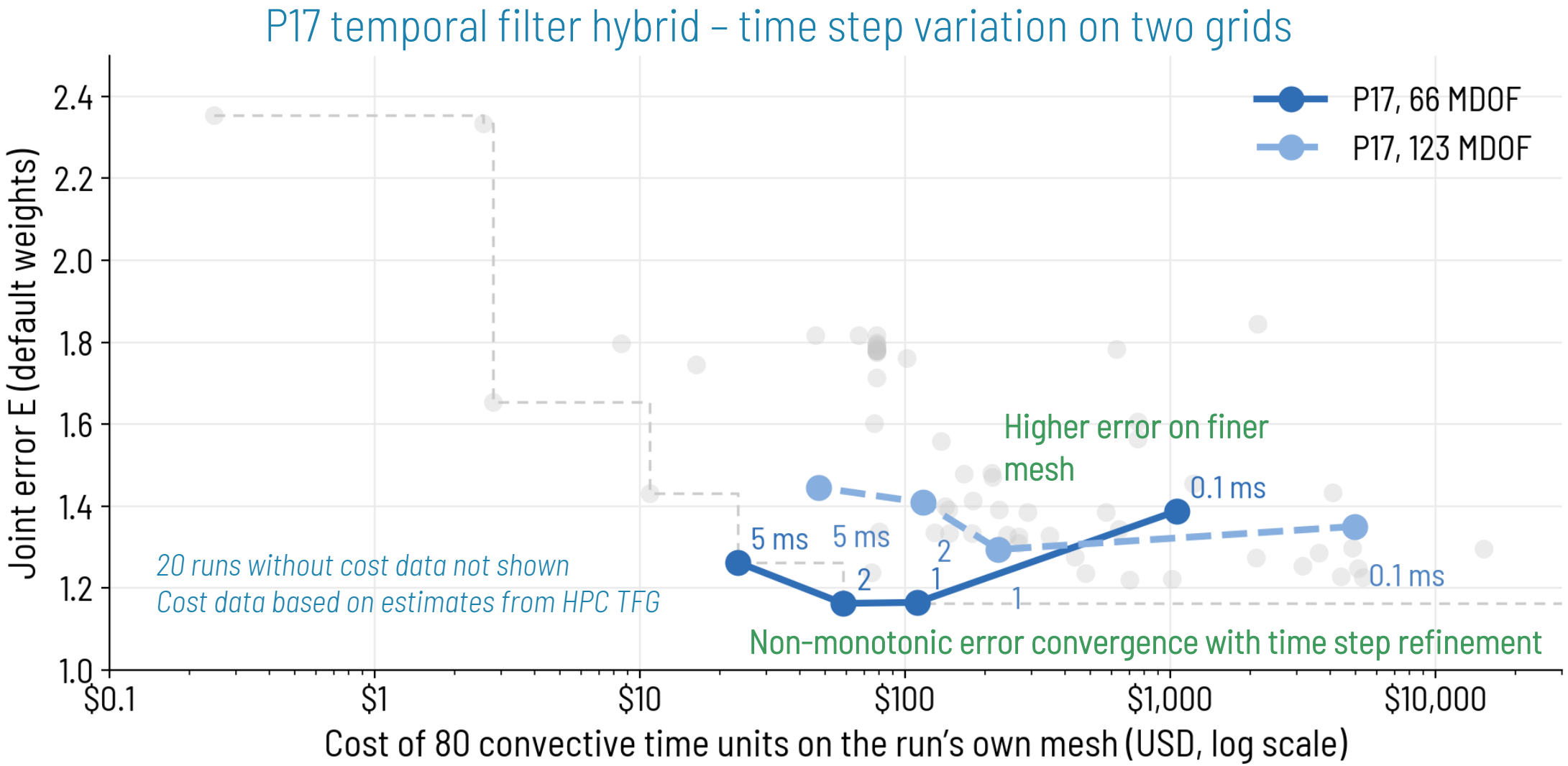


Cost vs. error diagram, all case 2a submissions with HPC cost data



provisional

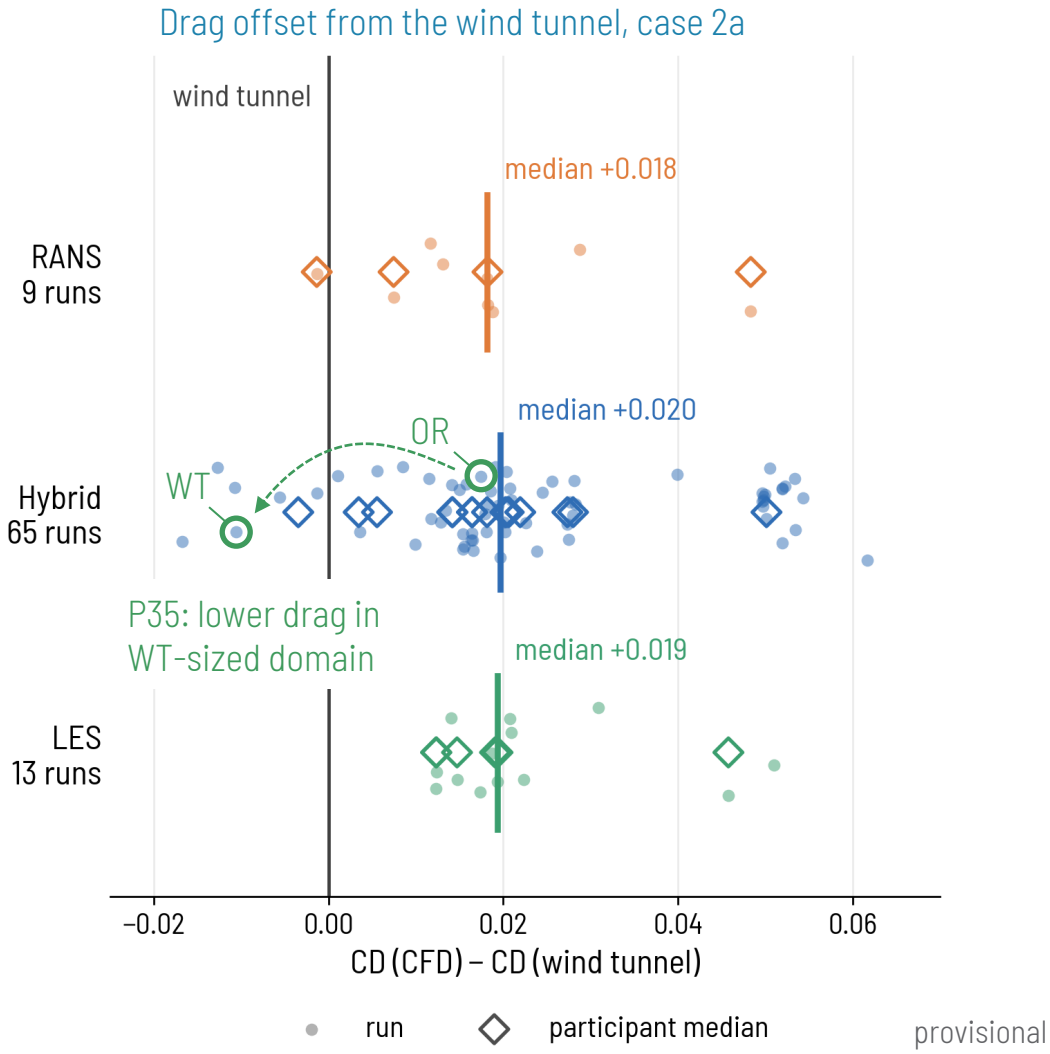
Cost vs. error diagram, all case 2a submissions with HPC cost data



provisional

Drag and wake: CFD vs. wind tunnel

- **Drag:** CFD drag higher than wind tunnel in 80 of 87 runs
  - Median offset +0.019 (+7.5%), about one field spread
  - Valid for every turbulence modelling category
  - One run with tunnel-sized domain: drag falls below the experiment

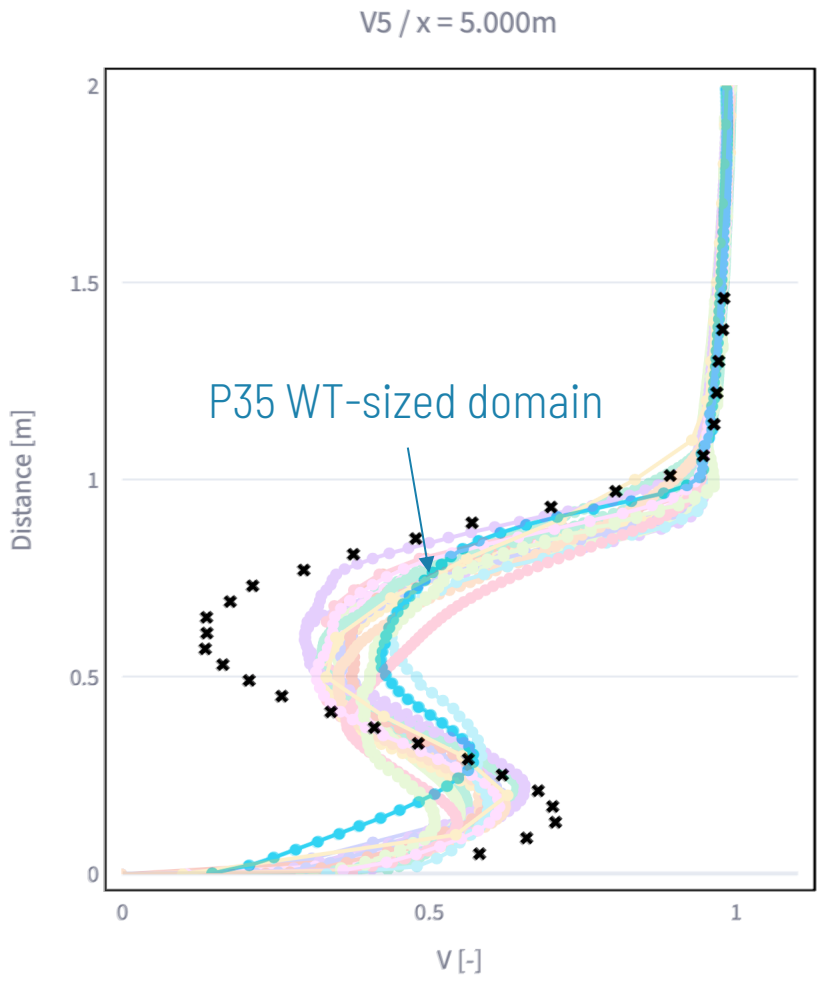


Drag and wake: CFD vs. wind tunnel

- **Drag:** CFD drag higher than wind tunnel in 80 of 87 runs
  - Median offset +0.019 (+7.5%), about one field spread
  - Valid for every turbulence modelling category
  - One run with tunnel-sized domain: drag falls below the experiment

- **Wake at V5 (x = 5 m):** CFD consensus deviates strongly from experiment
  - Also seen at AutoCFD4
  - Wake deviation not strongly affected by WT-sized domain

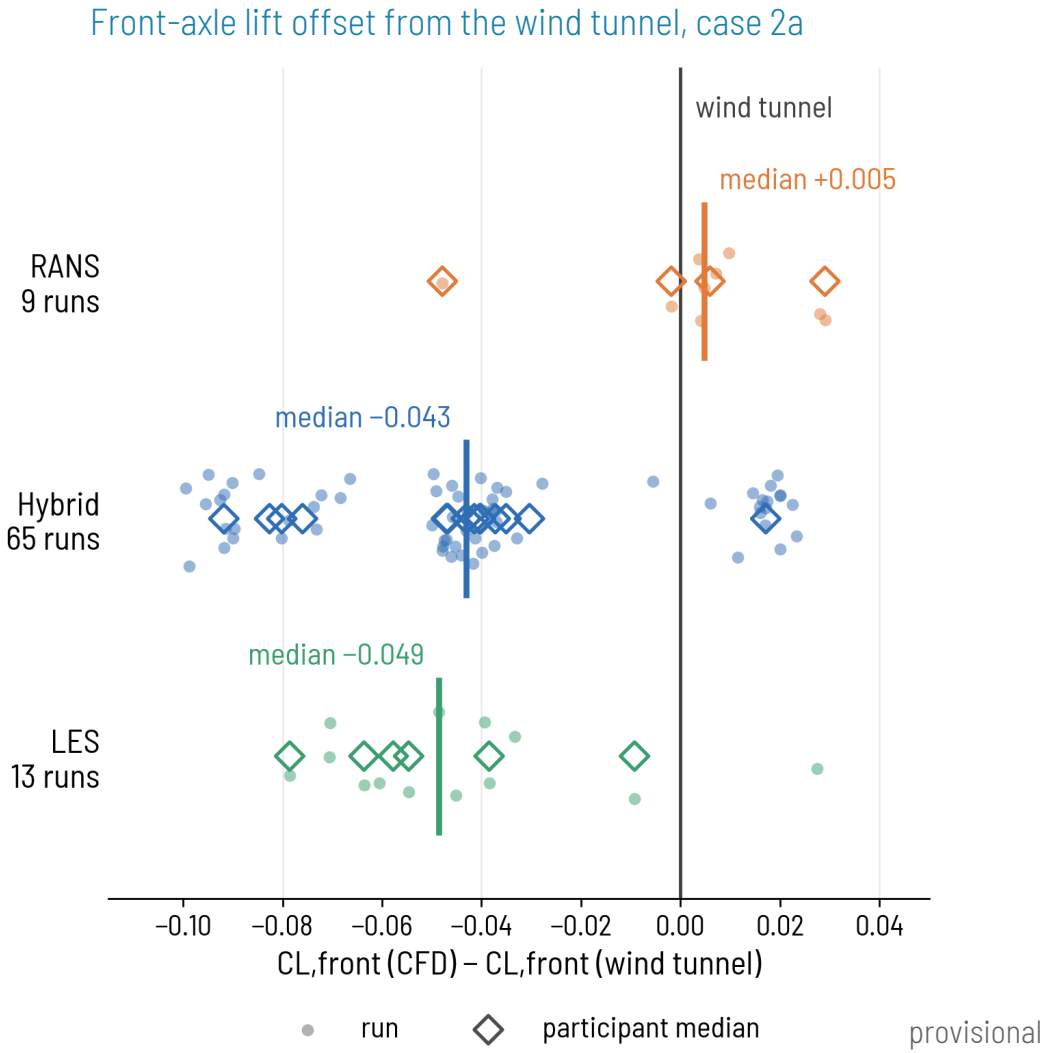
Velocity at V5: experiment vs. all runs



provisional

Front-axle lift: CFD vs. wind tunnel

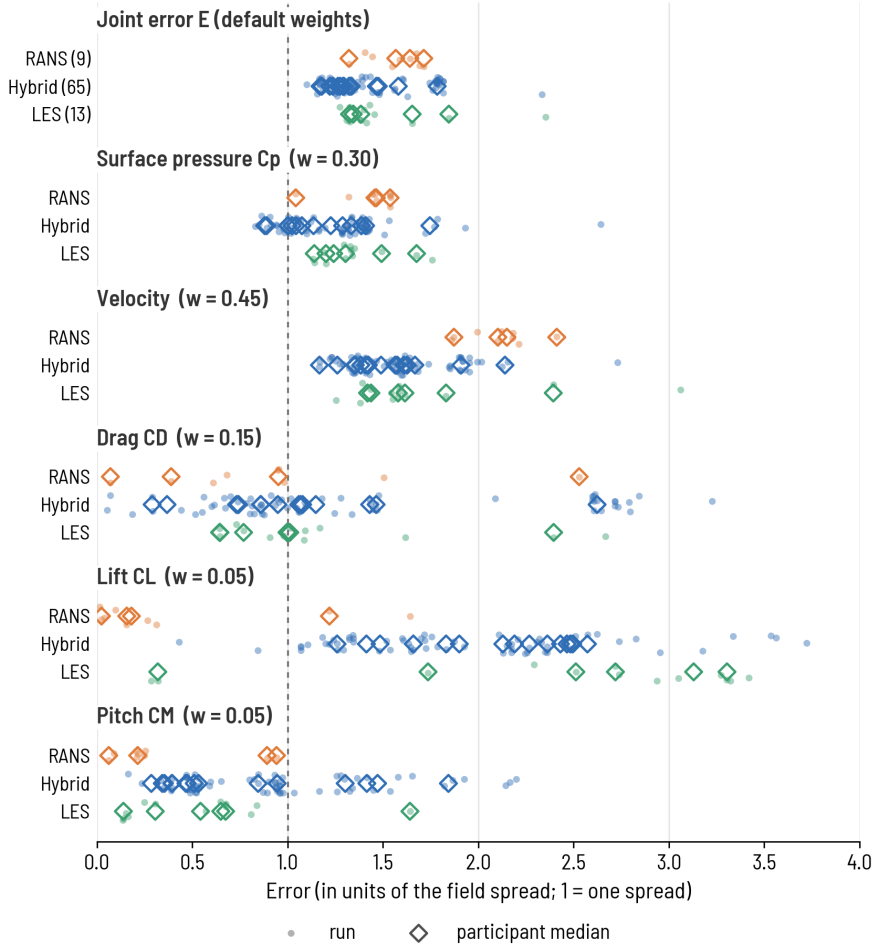
- **Scale-resolving runs below the WT value**, RANS closer
  - 17 of 18 scale-resolving participants below the WT value
  - Rear axle: no offset between categories
  - Low number of RANS data points
- **Refinement does not close the gap**
  - Grid and time-step refinement have little effect, or move away from the experiment



Error breakdown by turbulence modelling category

- Joint error (default weights)
  - As for drag and wake: CFD agrees better with itself than with the experiment
- Fields vs. forces
  - Pressure and velocity errors average over hundreds of probes: no run is lucky everywhere, so none scores near zero
  - A force is one number: pressure errors on different parts of the body can cancel, so many runs land within one spread
  - A force can be right by luck: this is why we gave integral forces lower default weights
- RANS vs. scale-resolving: depends on the quantity
  - RANS best on lift, worst on velocity
- Ranking depends on the weights
  - Default: velocity counts nine times as much as lift, so RANS ranks last
  - Equal weights: RANS first (1.12 vs. hybrid 1.27, LES 1.50)
- Categories also differ in code, grid etc.

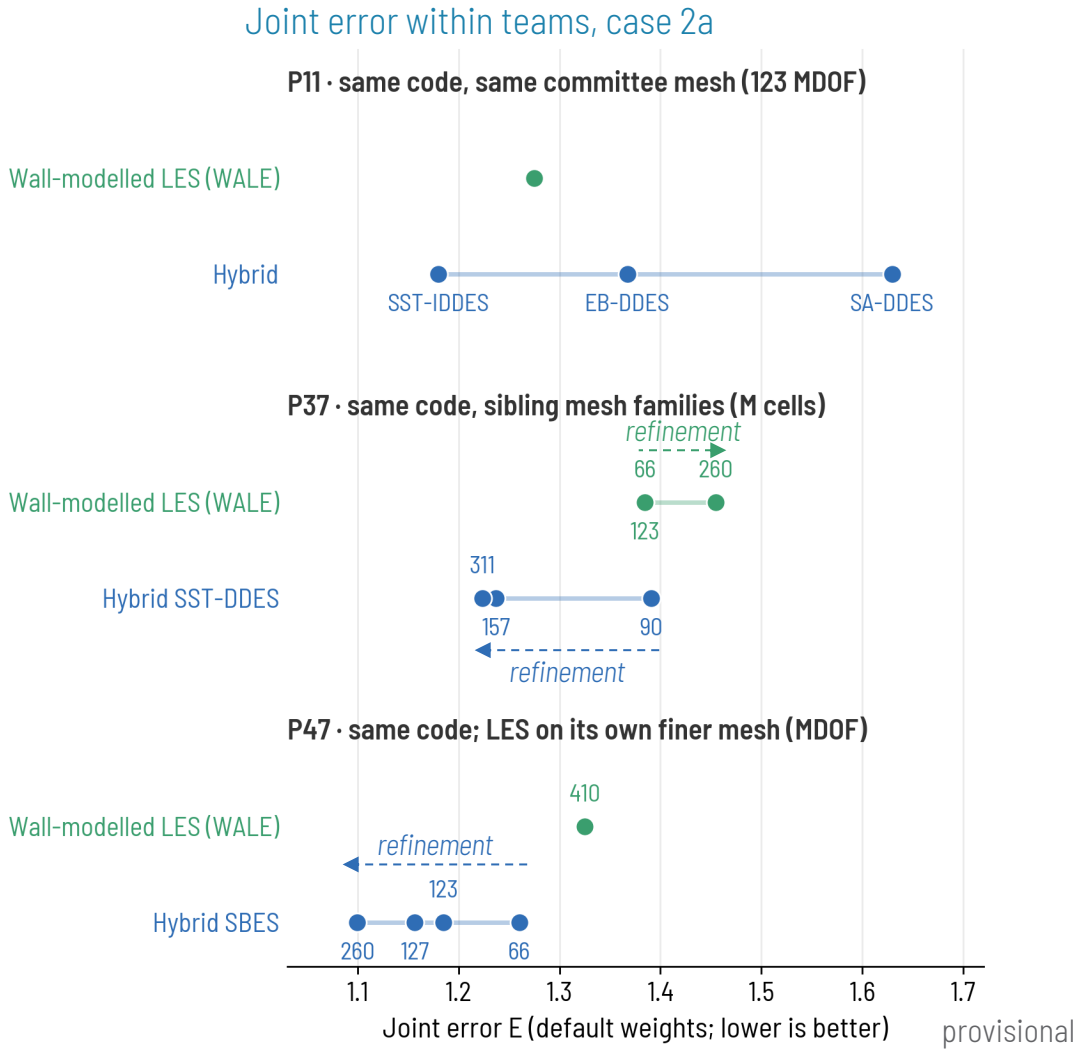
Error by component and category, case 2a



6 of the 9 RANS runs come from P42, whose Cp data are under review. provisional

Wall-modelled LES vs. hybrid RANS-LES

- Three teams compared WMLES and hybrid RANS-LES with the same code
- WMLES did not beat the best hybrid RANS-LES in any of the comparisons
  - P11: LES beats two of three hybrid variants, not the best
  - P37: hybrid improves with refinement, LES does not
  - P47: every hybrid run beats the LES run on a finer grid
- Large scatter between different hybrid RANS-LES models (same code and grid)



Summary and next steps

- Thank you for contributing your simulation data!
- Joint error metric proposed
  - Facilitates analysis of large datasets
  - Proposed default weights emphasise field variables over integral forces, reduces risk of error cancellation
- CFD agrees with itself more than with the tunnel
  - Median drag about 0.02 above experiment
  - No CFD matches the V5 wake
  - Indicates a comparability issue, cause unknown
- RANS vs. hybrid RANS-LES vs. WMLES
  - Within teams, WMLES did not beat the best hybrid method
  - Hybrid RANS-LES model choice very influential
  - RANS predicted lift better than SRS, opposite trend for velocity field
- Next steps
  - Resolve open data issues: results are provisional
  - Fine-tune metric definition to enable comparability between cases 2a and 2b
  - Dashboard publication once issues resolved
    - AutoCFD5 participants only until then
- Explore the data yourself in the dashboard
  - https://autocfd5.cfdsolutions.net

upstreamCFD

Thank you for your attention