

TFG-1 Ice Shape Comparison Metric Analysis Applied to TFG-2

Ariane Vieira, Alessandro Donizetti

AeroTex GmbH; Politecnico di Milano

Presented 22 September 2026

CC BY-NC-ND 4.0

This work is licensed by the author(s) under a **Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0)**. To view a copy of this licence, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Suggested citation: Vieira, A., Donizetti, A. (2026). *TFG-1 Ice Shape Comparison Metric Analysis Applied to TFG-2*. Presented at the 3rd AIAA Ice Prediction Workshop, online, 22 September 2026.



TFG-1 Ice Shape Comparison Metric Analysis Applied to TFG-2

Ariane Vieira — AeroTex GmbH

Alessandro Donizetti — Politecnico di Milano

3rd Ice Prediction Workshop

22/09/2026

Outline

Part 1 — The Jaccard index (Politecnico di Milano)

- Possible weight choices and their interpretation, applied to the IPW3 results

Part 2 — ε : possible applications (AeroTex)

- 3 suggested uses for ε , applied to the IPW3 results

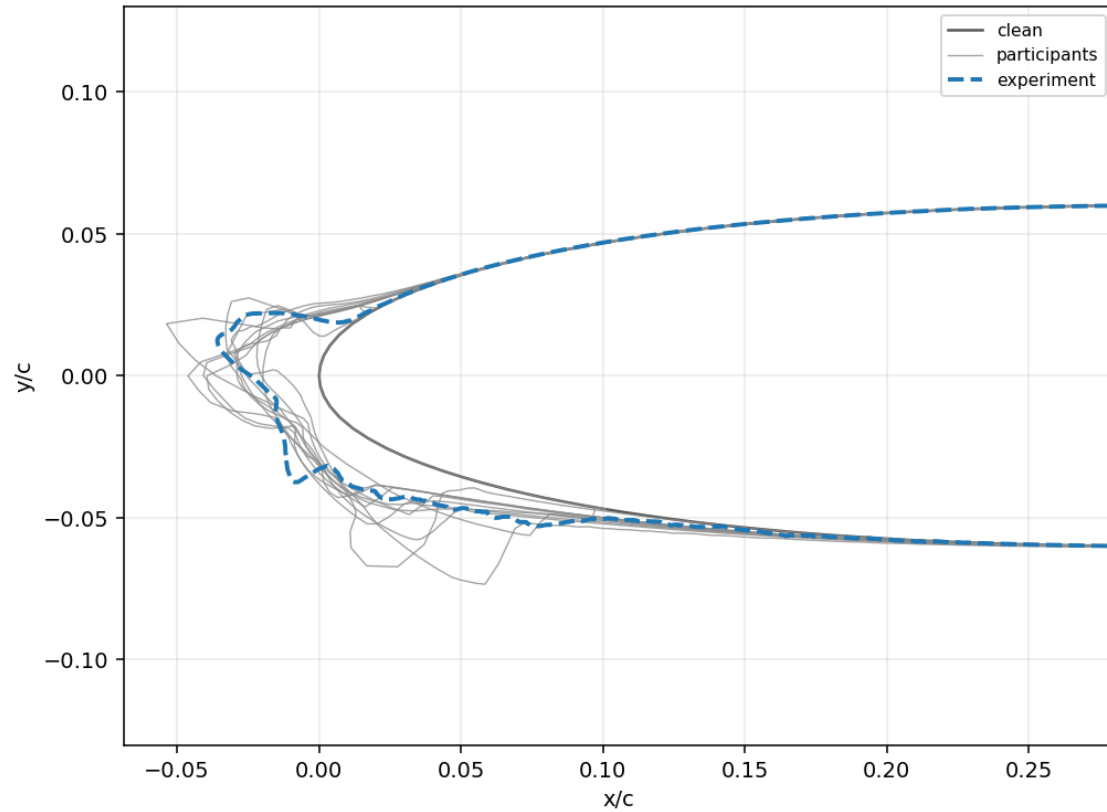
Part 3 — Conclusions

The Jaccard index

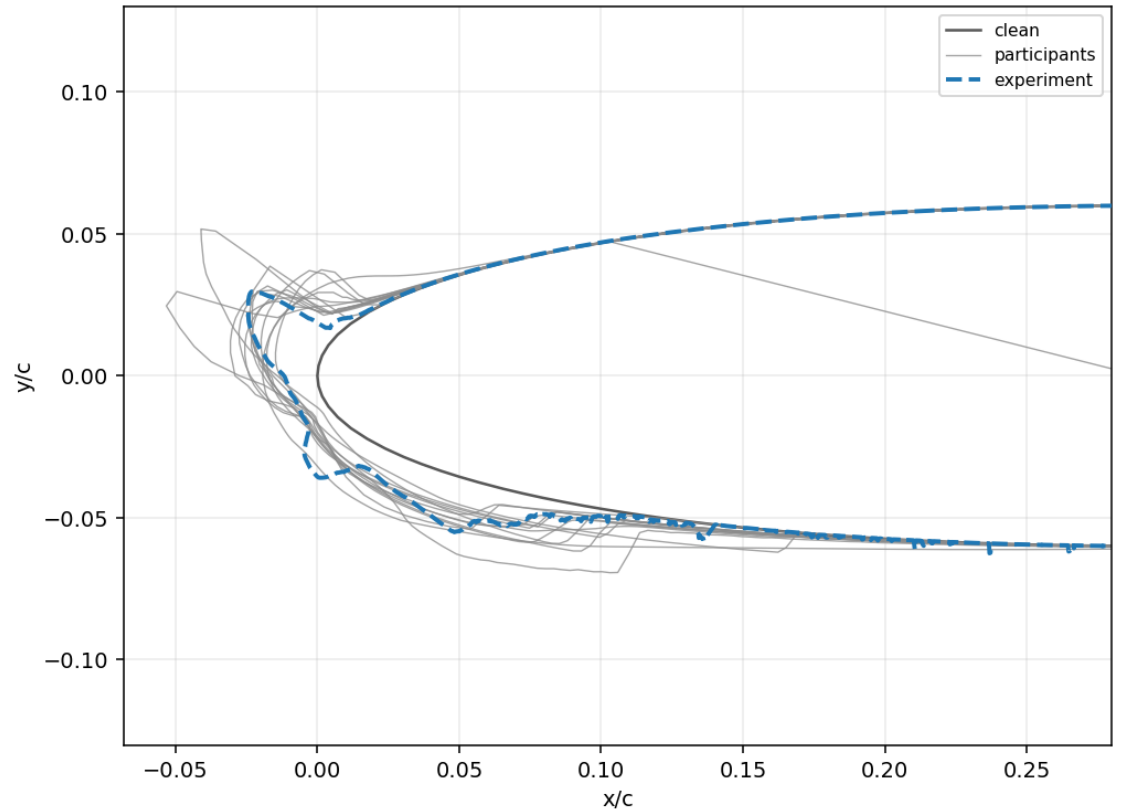
PoliMI – A. Donizetti

AE3932 and AE3933

IPW3-AE3932



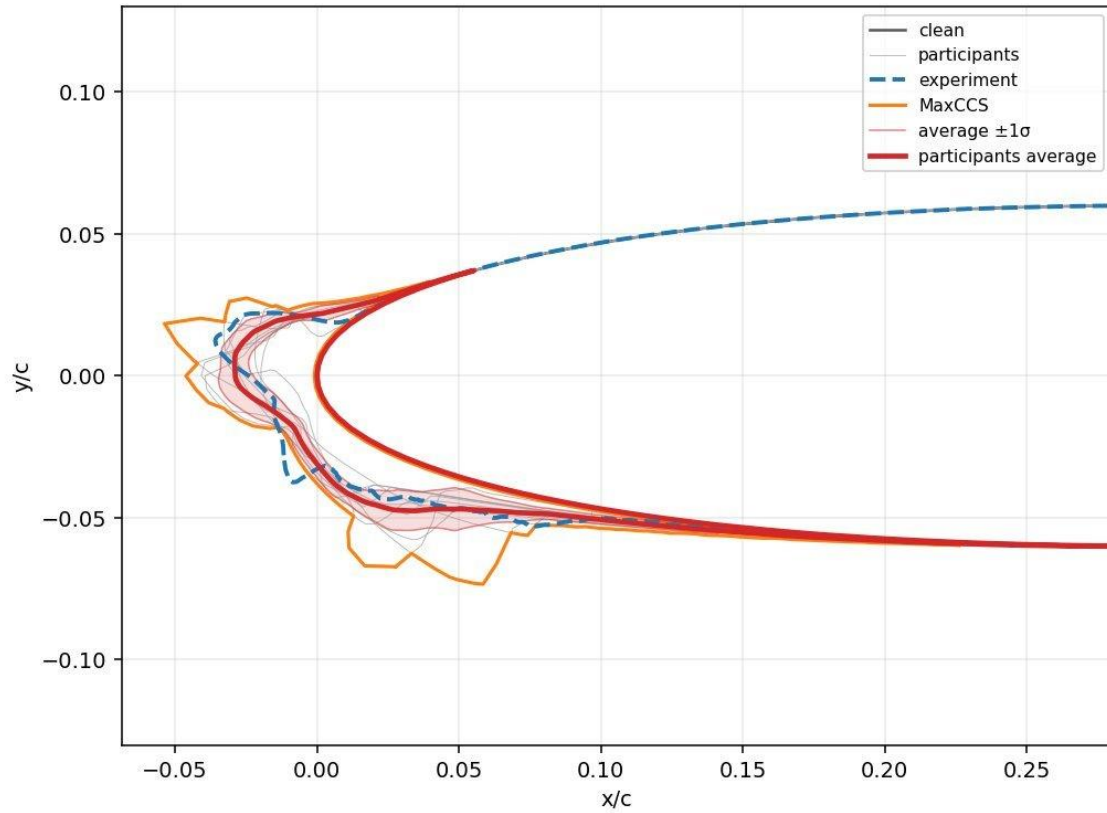
IPW3-AE3933



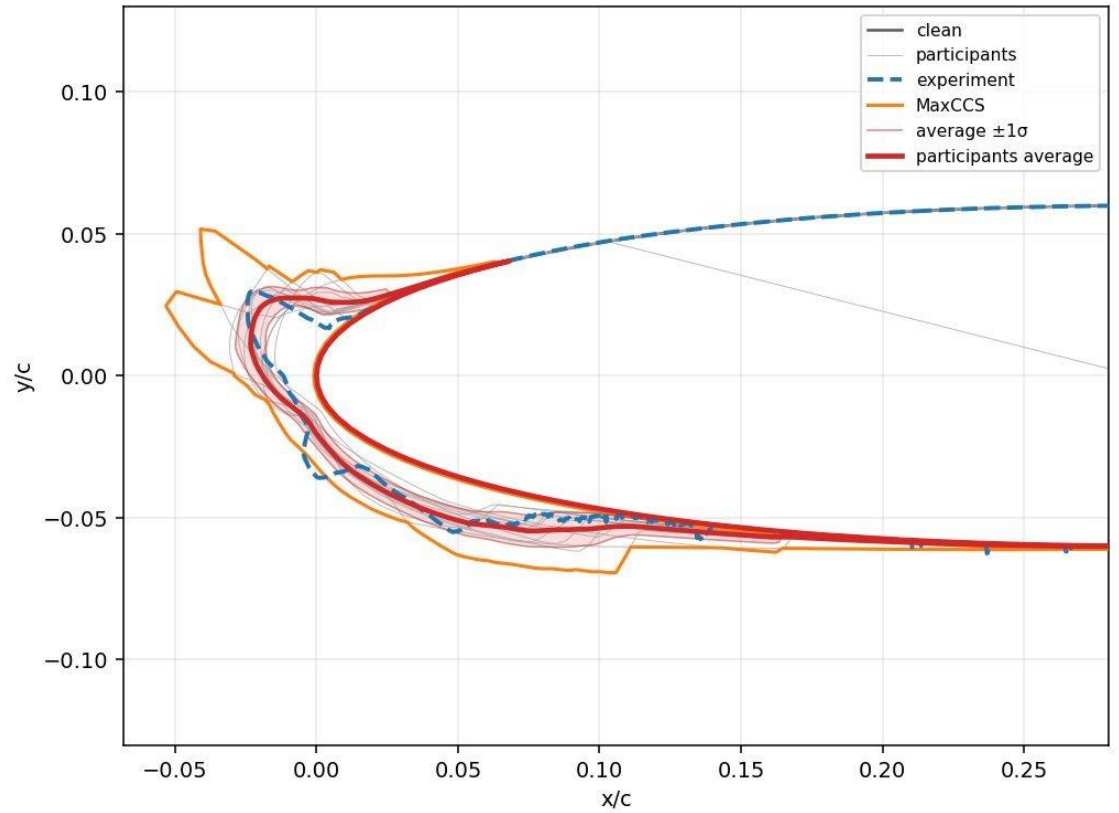
Finest grids, most bins per code

AE3932 and AE3933

IPW3-AE3932



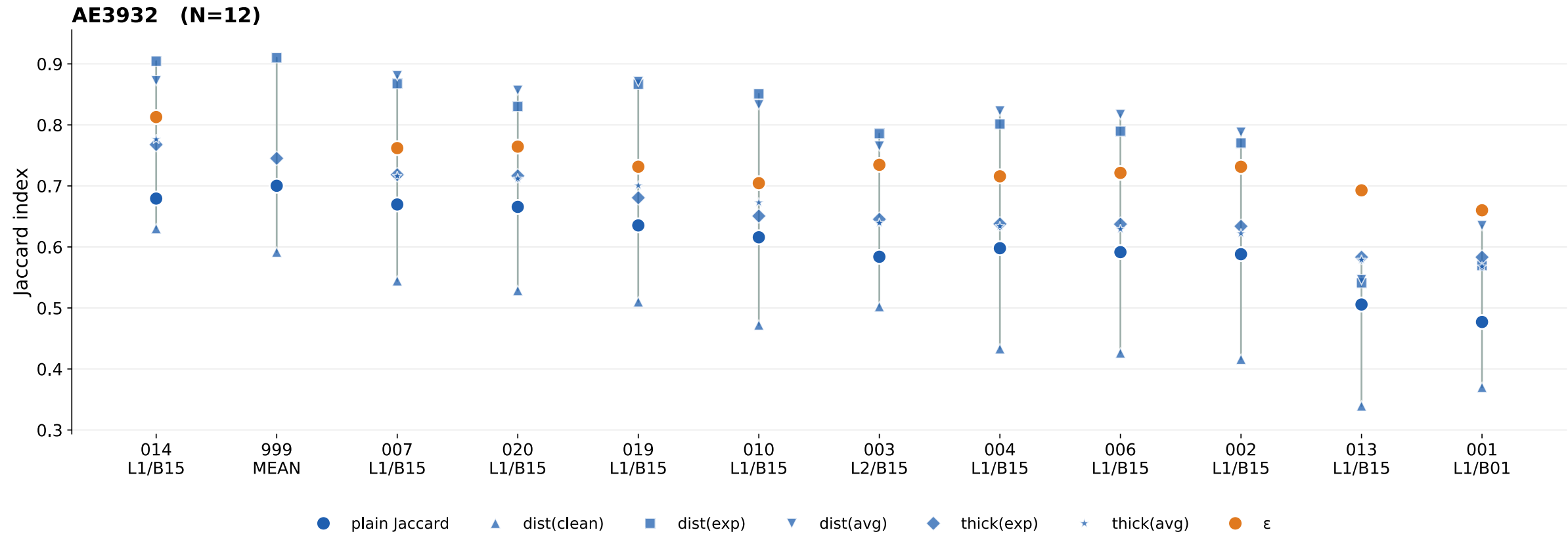
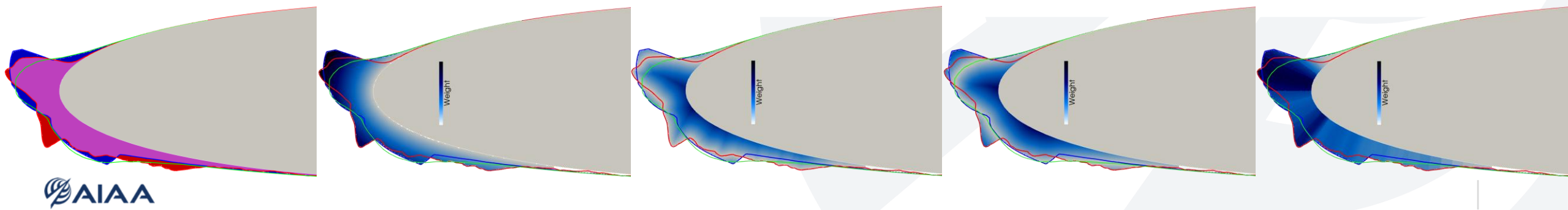
IPW3-AE3933



Finest grids, most bins per code

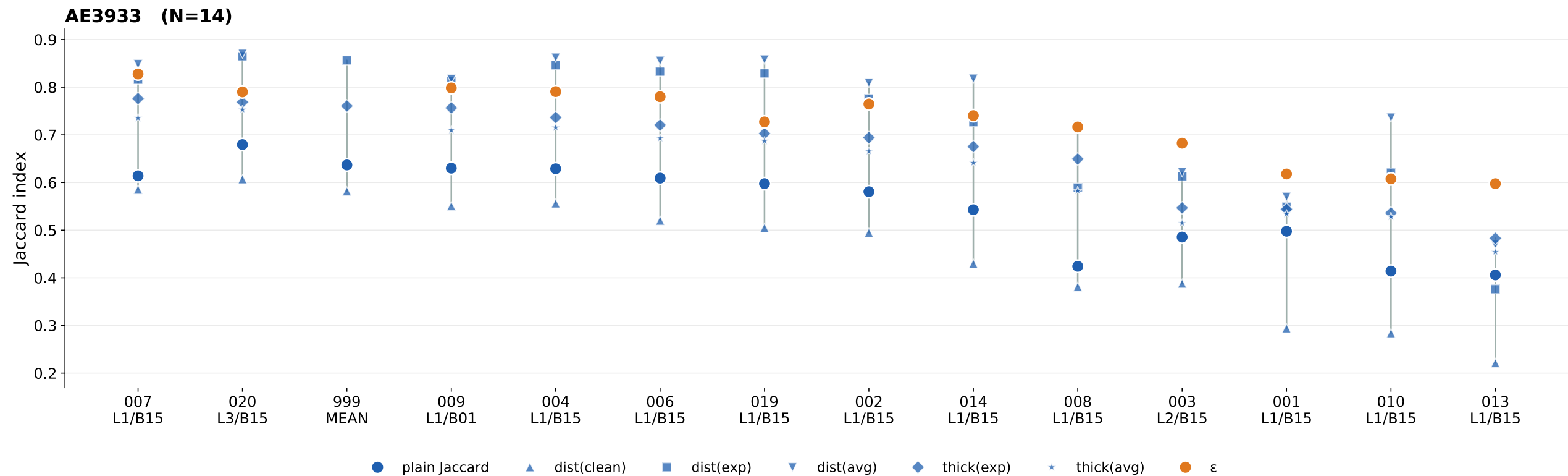
AE3932

NACA0012 AE3932 — score bands by participant

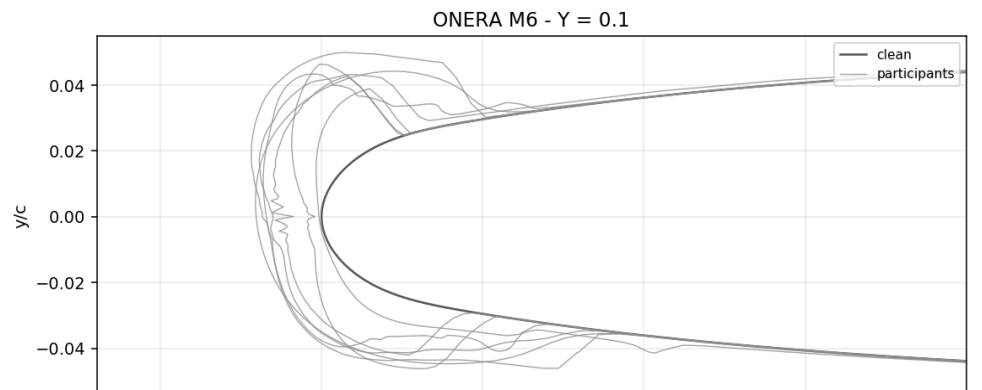
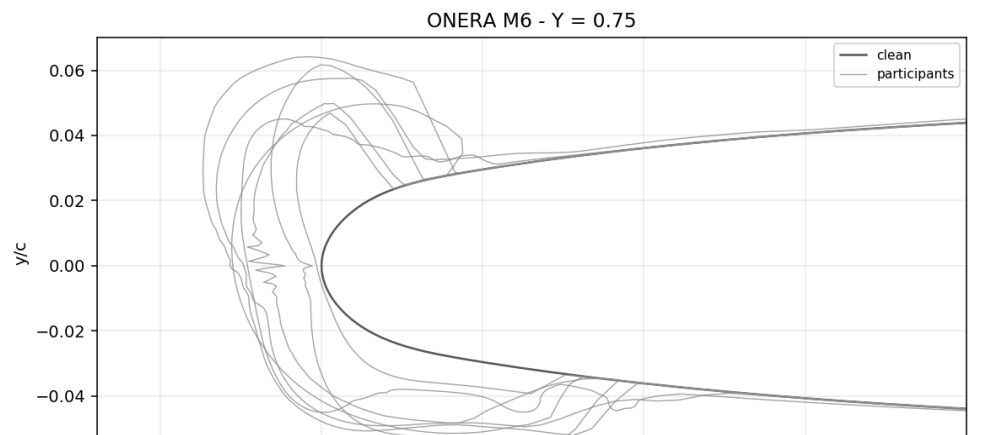
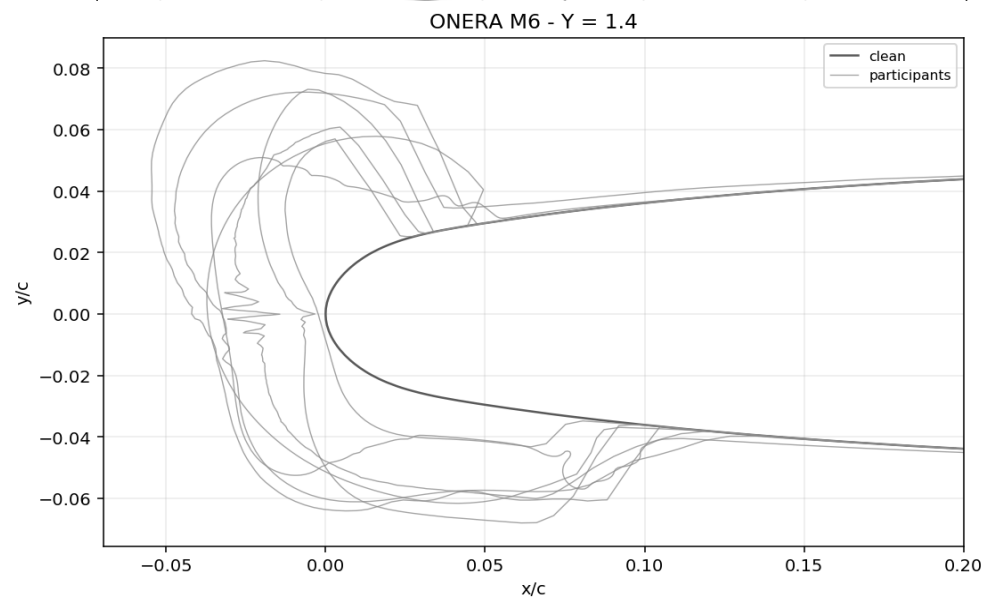
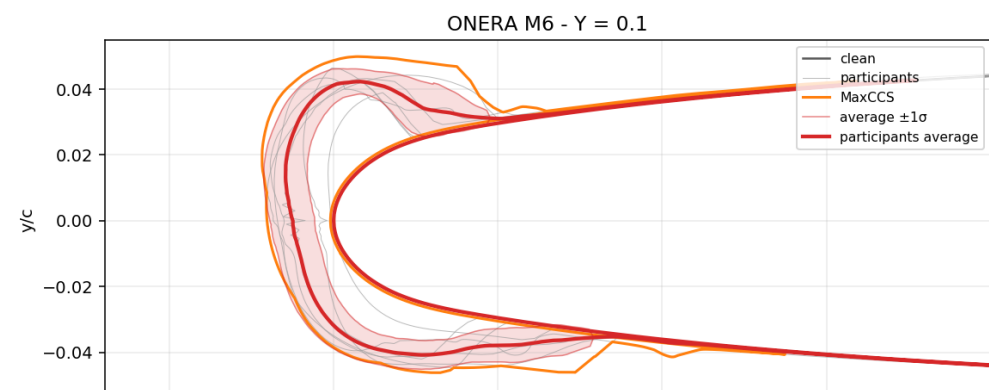
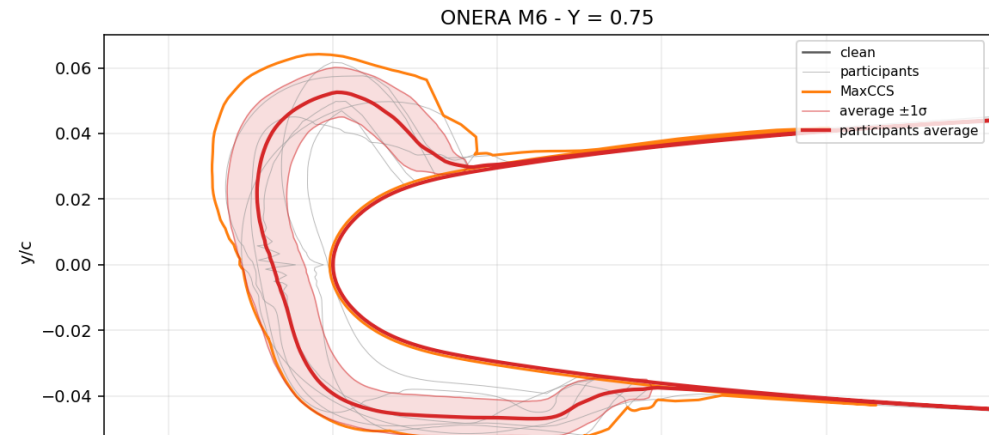
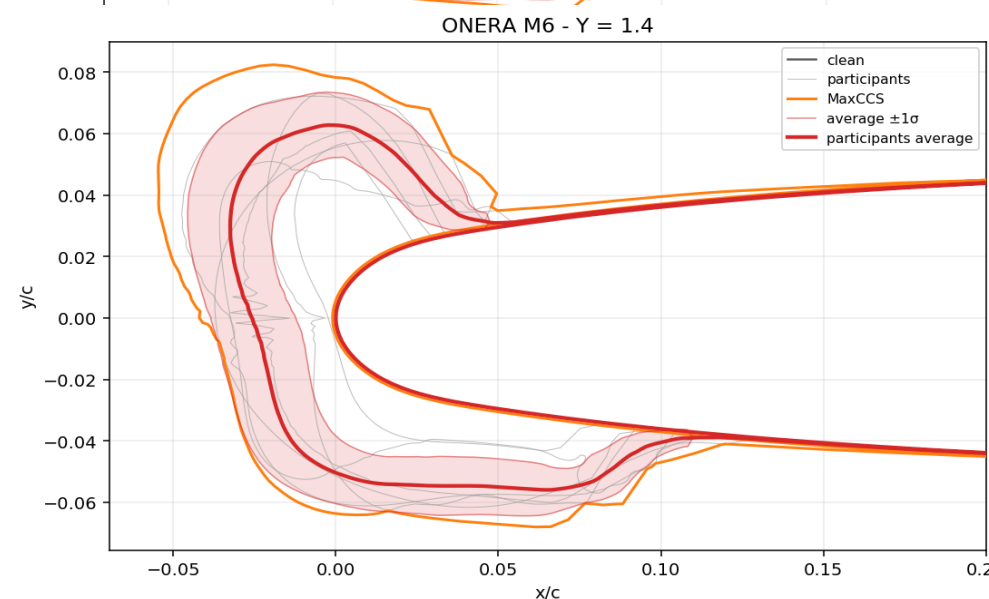


AE3933

NACA0012 AE3933 — score bands by participant

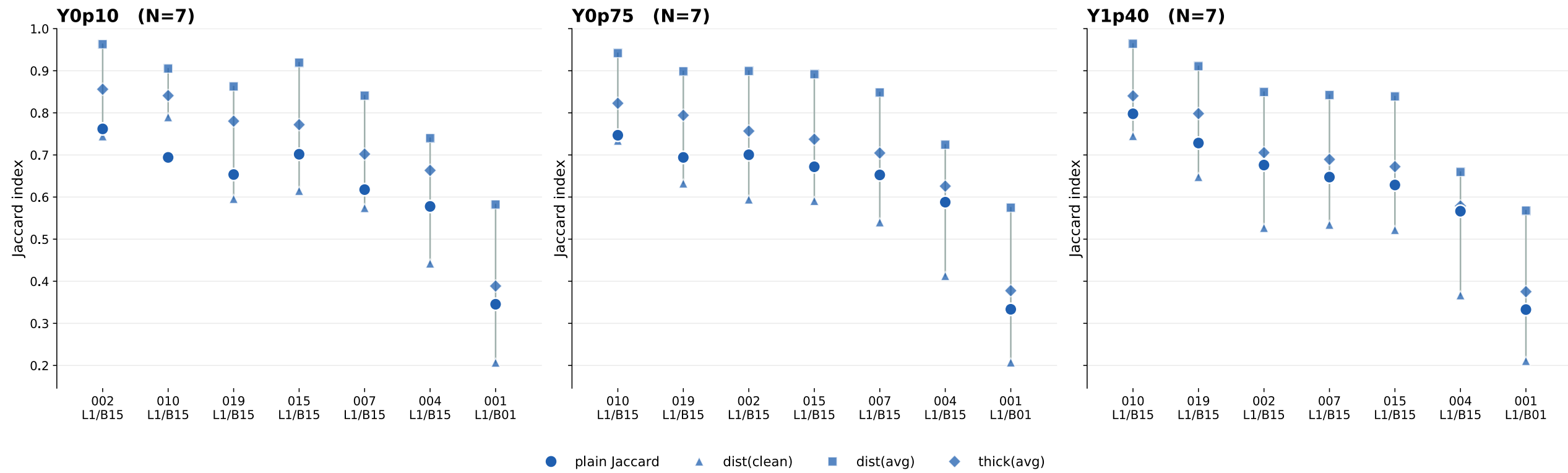


ONERA M6



ONERA M6

ONERA M6 — score bands by participant (reference = consensus mean)



Finest grids, most bins per code

ε: suggested uses

Part 2 · AeroTex — A. Vieira

Suggested uses for the ϵ parameter

Use 1 • Experiment to simulation

- Submitted shapes against a reference (IPW3 NACA0012 versus NASA Min / Mean / Max sections)

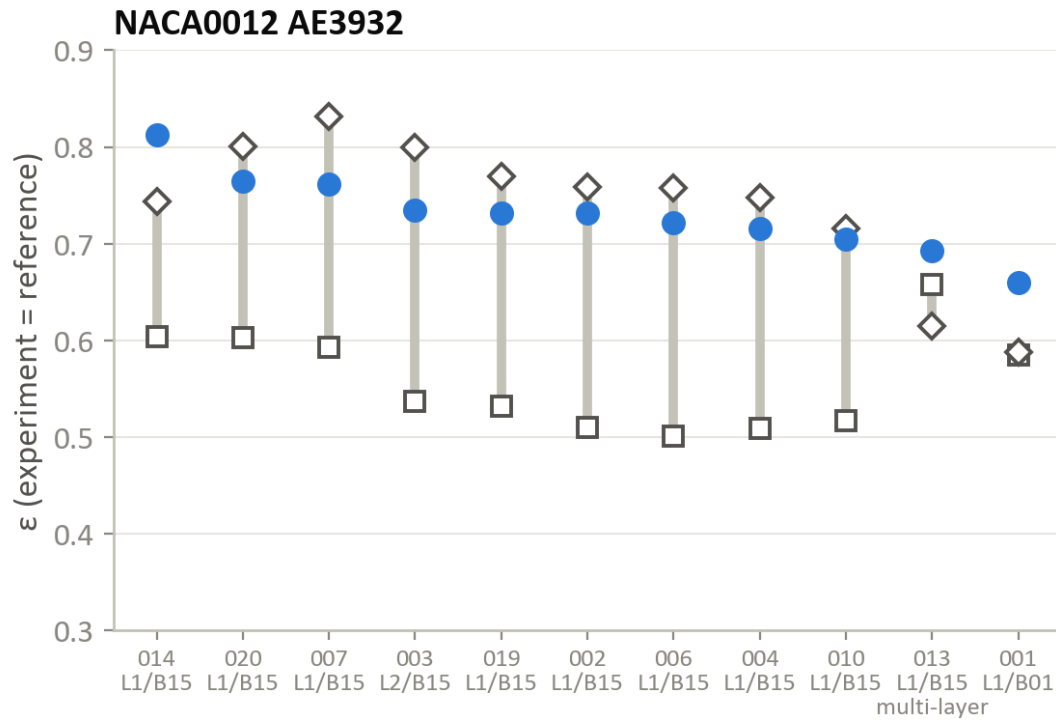
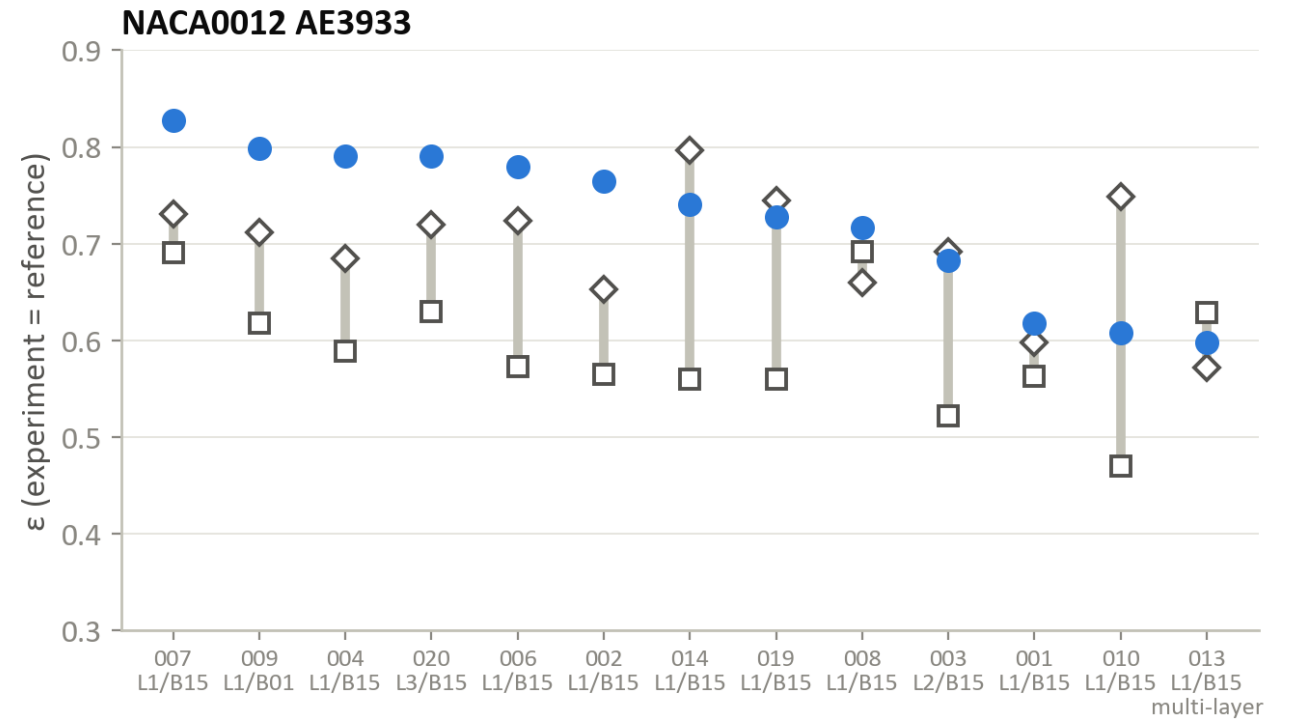
Use 2 • Grid and bin convergence

- Each code scored against itself, from one level to the next

Use 3 • Code-to-code

- Every pair of codes in both directions, a consensus map of the field

Use 1: Experiment to simulation



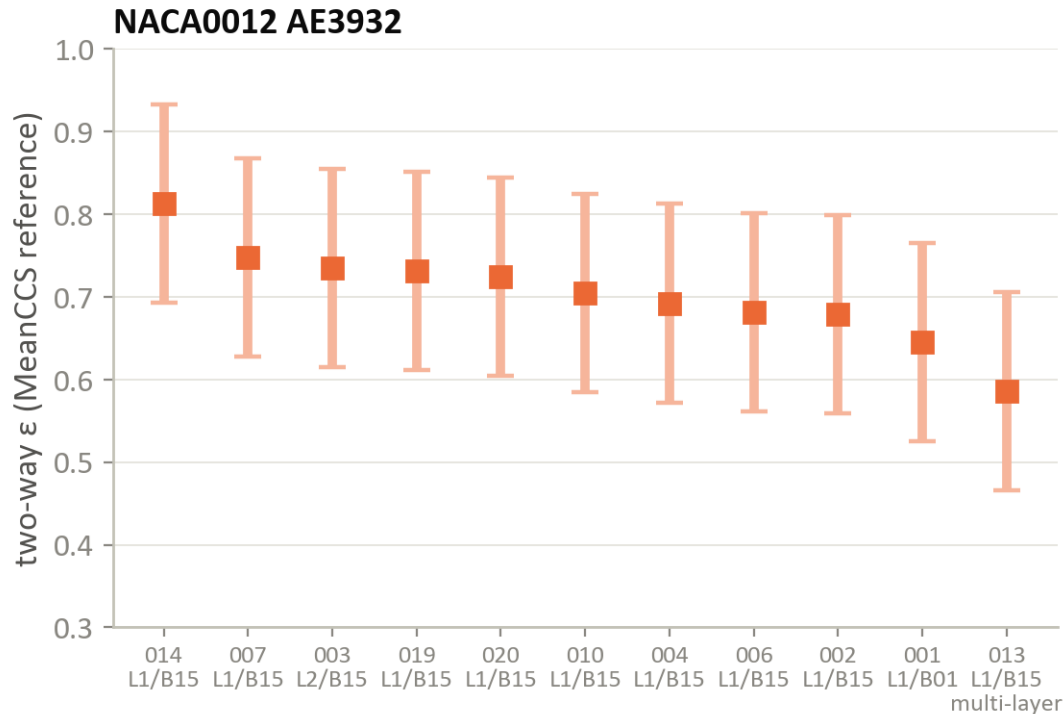
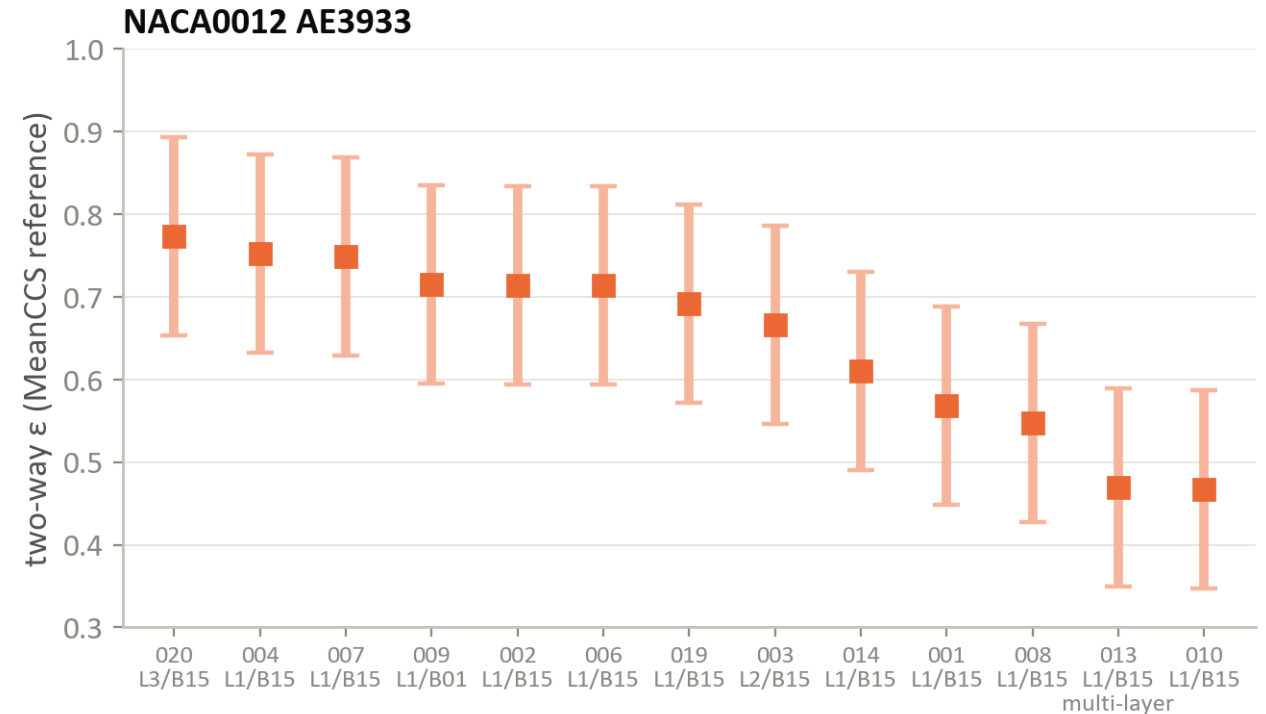
□ ε vs MaxCCS (thickest section)

◇ ε vs MinCCS (thinnest section)

● ε vs MeanCCS

Original ϵ , as published (one-way: experiment = reference) · IPW3 NACA0012 AE3932 / AE3933 · NASA Min / Mean / Max sections · finest grid, most bins per code

Use 1: Two-way ϵ against the Mean section

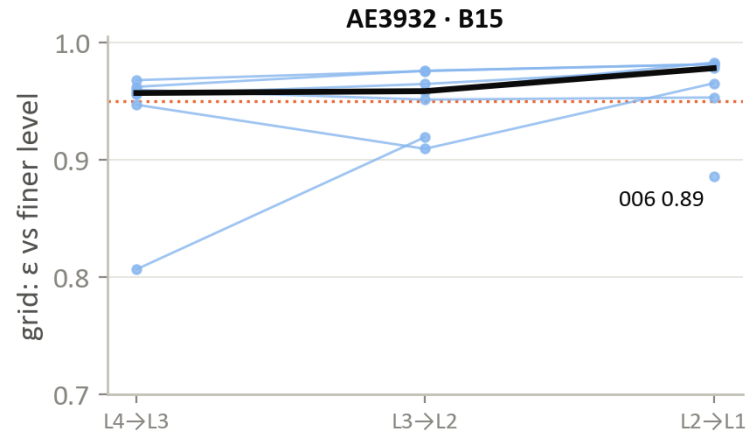
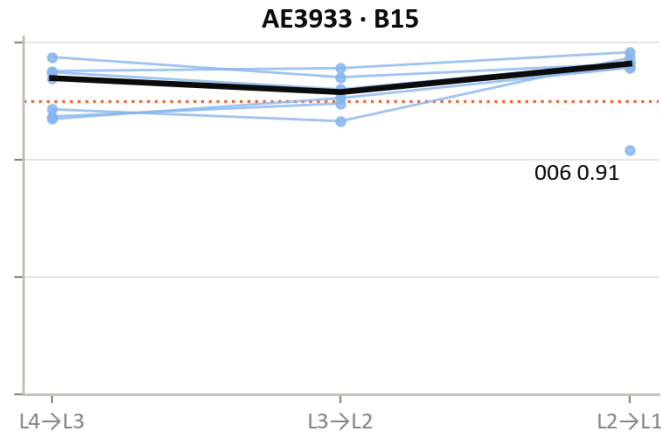


■ two-way ϵ vs MeanCCS (stricter of both directions)

I ± 0.12 band (survey-based reliability of ϵ^*)

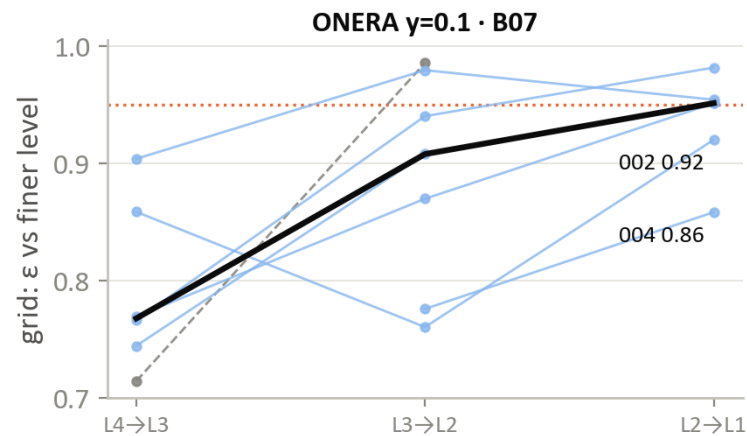
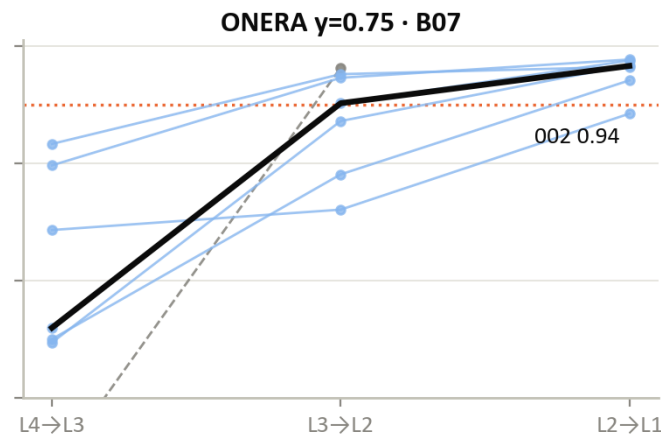
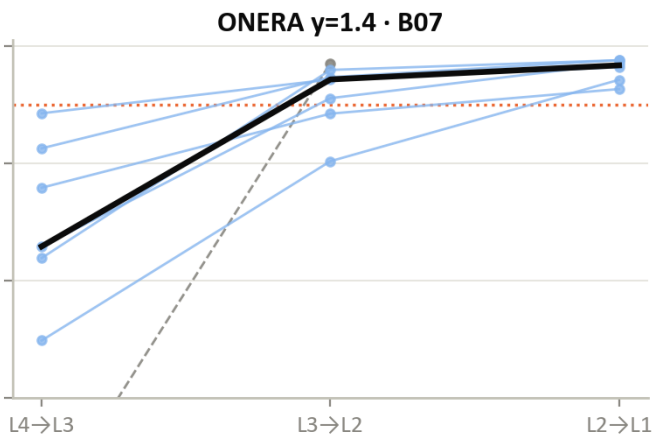
Two-way ϵ = the stricter of both scoring directions · NASA MeanCCS as reference · same cells as the previous slide, re-ranked
 ± 0.12 band = the reliability of ϵ^* proposed in the survey talk: within it, most codes agree

Use 2: Grid convergence



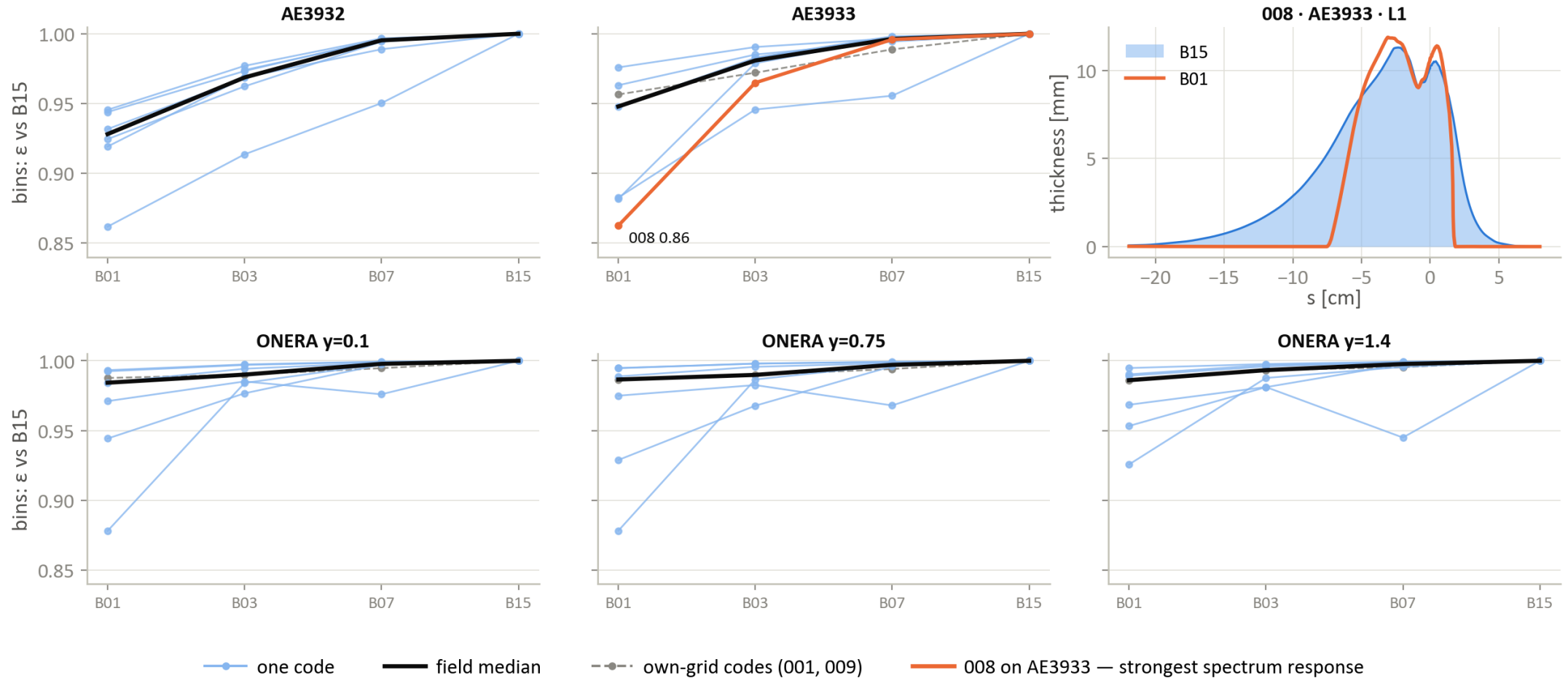
how to read

thin line: one code
 bold line: field median
 dashed: own-grid codes (001, 009)
 dotted: $\varepsilon = 0.95$ adequacy line
 labels: final steps below 0.95



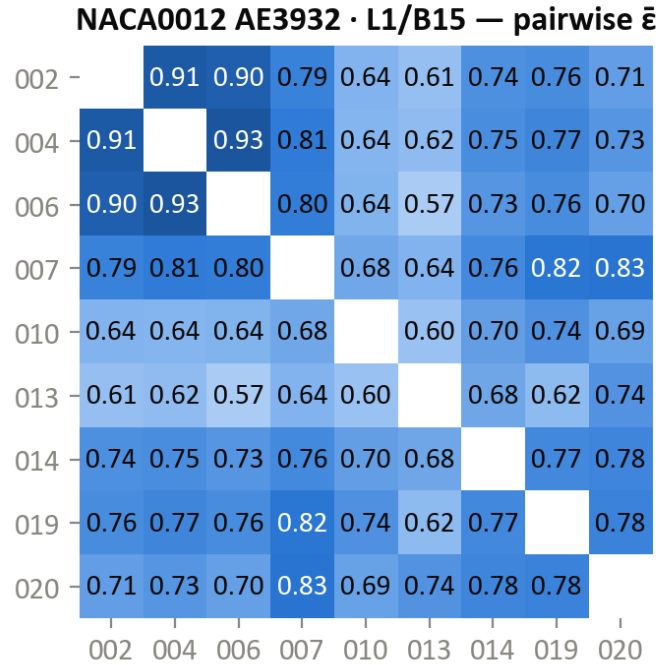
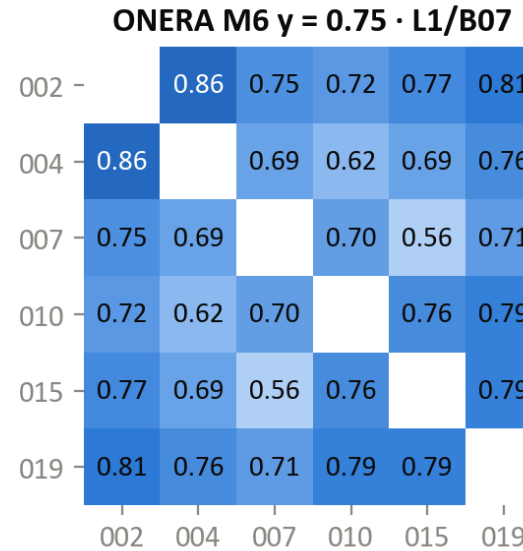
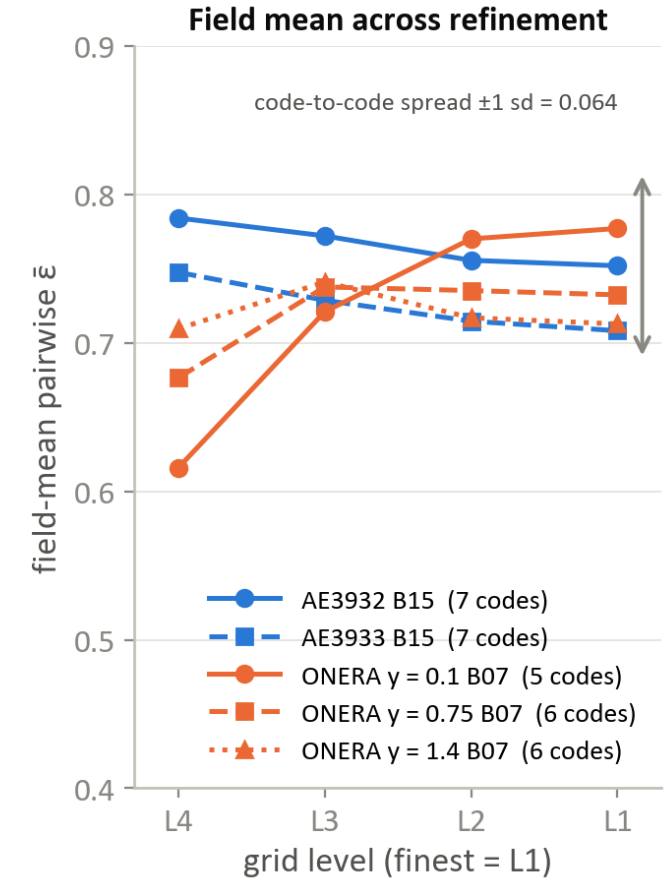
Reading: $\varepsilon \gtrsim 0.95$ against the next finer level = "the shape has stopped changing"

Use 2: Bin convergence



Reading: $\epsilon \gtrsim 0.95$ against the next bin level = "the shape has stopped changing"

Use 3: Code-to-code



Mean of both scoring directions

Refinement converges each code to its own answer: the field mean barely moves, far less than the code-to-code spread

Pros and cons

Use	Pros	Cons	Suitable for
1 • experiment to simulation	one number per shape • ranks like the experts (ϵ^*) • spreadsheet-reproducible	reference choice • one-way: blind outside the footprint	ranking vs a named reference • shapes shown alongside
2 • grid / bin convergence	adequacy screen $\epsilon \gtrsim 0.95$ • localizes refinement needs	saturates near 1 • own grids • mixed datasets / code versions	convergence check before any comparison
3 • code to code	reference-free • symmetric, no blind spot • refinement trends	roughness / grid flags • consensus $\neq$ correctness	clusters • outliers • field trends

Conclusions

Part 3 · Moving forwards

Wrap-up

Both methods provide deterministic ways to evaluate ice shapes. Possible uses:

- **Experiment to simulation:** provides a single number; depends on the reference shape used (Min, Max, Mean); open to evaluation
- **Grid / BIN convergence:** a suggested adequacy screen ($\varepsilon \gtrsim 0.95$)
- **Code-to-code:** a consensus map for the community; open for discussion
- **Other possibilities:** code development; machine learning optimization functions...

Full results, scripts and interactive graphs: to be added to the GitHub page