Statistical Comparison of Computational Results

Michael J. Hemsch
Research Facilities Branch
m.j.hemsch@larc.nasa.gov
757-864-2882

AIAA APA TC CFD Drag Prediction Workshop
Hilton Hotel, Anaheim, CA
June 9-10, 2001

Outline of the talk

1. Why do a statistical analysis?

2. Description of several methods for estimating the location and scale.

3. Typical (design) customer requirements for experimental and computational simulations.

4. Experimental results for CL=0.5 and M=0.75.

5. Statistical analysis of the present results.
   A. Drag, AOA, pitching moment at CL=0.5 and M=0.75
   B. Drag rise curves at CL=0.4, 0.5, 0.6
   C. Drag polars at M=0.75

Why do a statistical analysis?

DPW

1. Although repeatability is not an issue, reproducibility is.
2. It gives us a credible quantitative estimate of the scatter and a reasonable sense of the true value.
3. It allows us to make quantitative predictions for each code by using all of them taken together as a collective.
4. It allows us to determine if the scatter is small enough to be useful
    - to designers
    - for valid comparison with experiment.
5. Allows us to see into the scatter to discern otherwise hidden effects.
6. It allows us to use uncertainty measures to compare with experiment.

Some Methods for Estimating Population Parameters from Sample Data

DPW

Location $\hat{\mu}$	Scale $\hat{\sigma}$
Mean $\frac{1}{n}\sum_{i=1}^{n} x_i$	SSD $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$
Median n odd: $x_{(n+1)/2}$ n even: $0.5(x_{n/2}+x_{(n/2)+1})$	AAD $\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}(\lvert x_i-\bar{x}\rvert)$ or $\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}(\lvert x_i-\tilde{x}\rvert)$
Mode Maximum of the pdf, $f(x)$	MAD $1.483\sqrt{\frac{n}{n-1}}\ \text{median}\{\lvert x_i-\tilde{x}\rvert\}$

We will use the median for all estimates of the location and either AAD or MAD for all estimates of the scale, unless stated otherwise, and we will use 100:1 limits (sigma multiplier is 2.576).

Typical (design) customer uncertainty goals ($\pm 2\sigma$) for performance simulations.

Coefficient	Increments	Absolute
C_L	0.005	0.01
C_D	1/2 count	1 count
C_m	0.0005	0.001

Experimental Results for $C_L = 0.5, M_\infty = 0.75$ DPW

Source	Variable	NLR	ONERA	DRA	Mean	Range	Median	Use
Original	Alpha, deg	0.153	0.192	0.181	0.175	0.038	0.181	0.18 ± 0.04
Enhanced	C_D, cts	288	289	281	286	8	288	286 ± 8
Original	C_m	-.130	-.126	-.137	-.131	.011	-.130	-0.13 ± 0.01

--- Notes ---

1. The individual values were obtained by linear interpolation.
2. The estimates given by each of the tunnels are roughly half of the observe scatter values.
3. The above estimates are recommended for risk and validation purposes.

Some stats on the drag point solutions

- 14 codes were used:
  - 7 structured
  - 6 unstructured
  - 1 Cartesian
- 35 solutions for the drag point at CL=0.5, M=0.75
  - 24 structured
  - 10 unstructured
  - 1 Cartesian

  - 17 used Spalart- Allmaras turbulence model
  - 17 used a two-equation turbulence model
  - 1 used Euler-Integral Boundary-Layer method

  - 21 used the provided grid
  - 14 used other grids

More stats on the drag point solutions

Turbulence Model	Structured Grids	Unstructured Grids
Spalart-Allmaras	8	9
Two-Equation (6)	16	1

*Not including the Cartesian Euler/BL solution.

Cautionary notes on the statistical inferences DPW

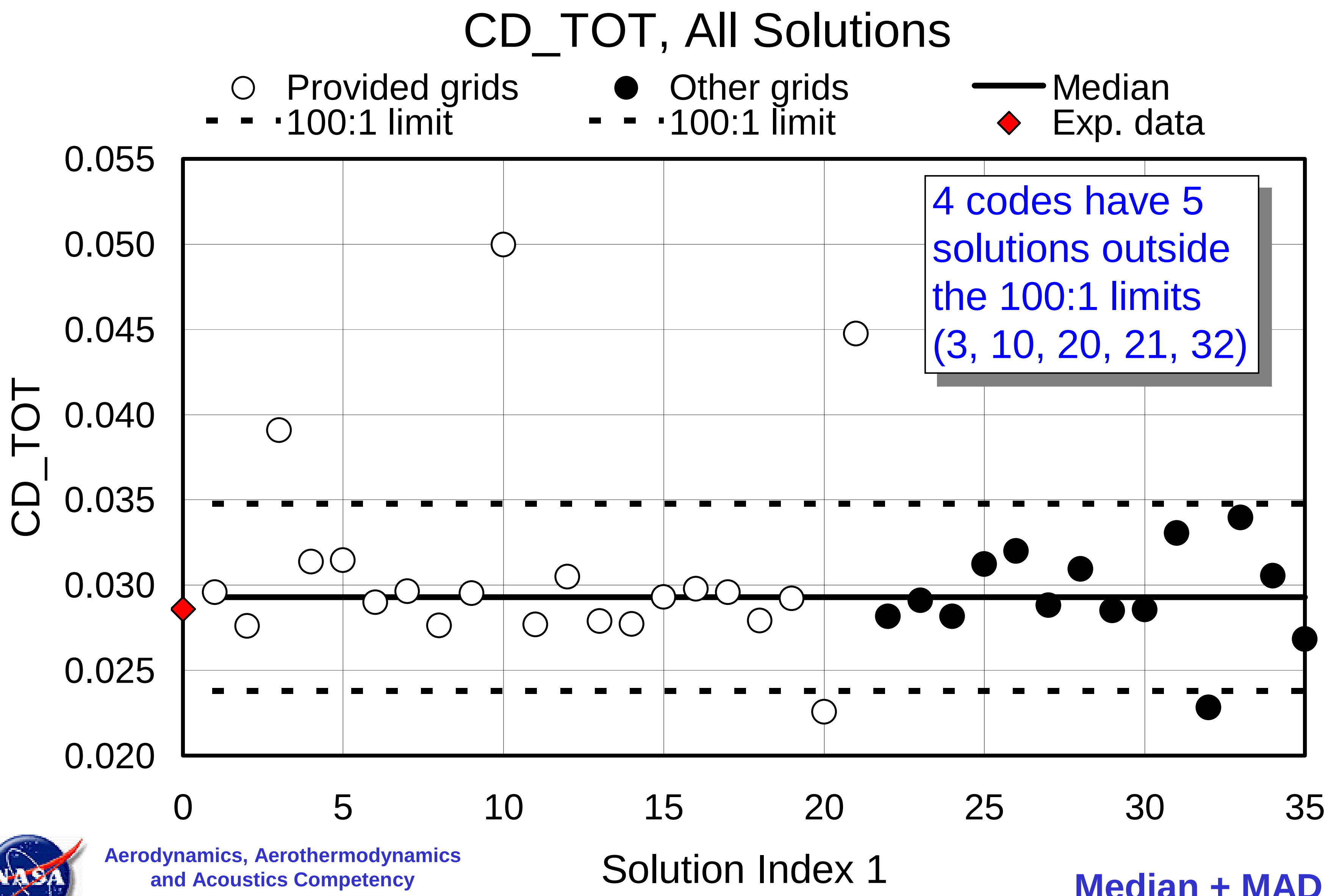
- I will assume that the solutions are random and independent draws from a stable population with a single location and a single scale. Of course, I HAVE NO IDEA IF ANY OF THIS IS TRUE.

- The world of statistics is gray, not black and white. You should treat my inferences as tentative --- to be confirmed later with further work.

- In the language of Hertz, this was NOT EXACTLY a designed experiment in the sense of statistics. But that doesn't mean that we can't use statistics to guide our conclusions.

ALL total drag solutions at CL=0.5, M=0.75

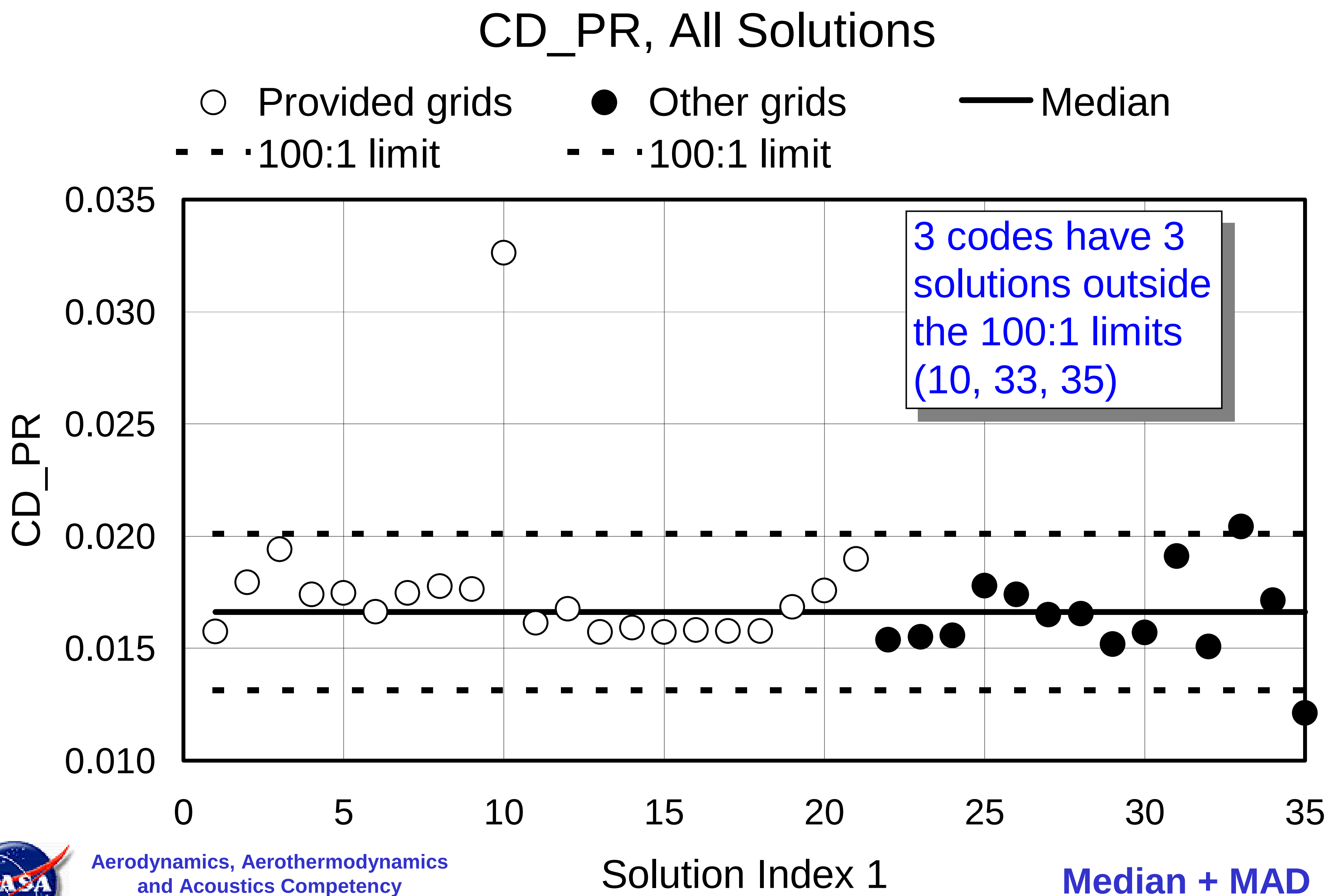
DPW



Median + MAD

ALL pressure drag solutions at CL=0.5, M=0.75

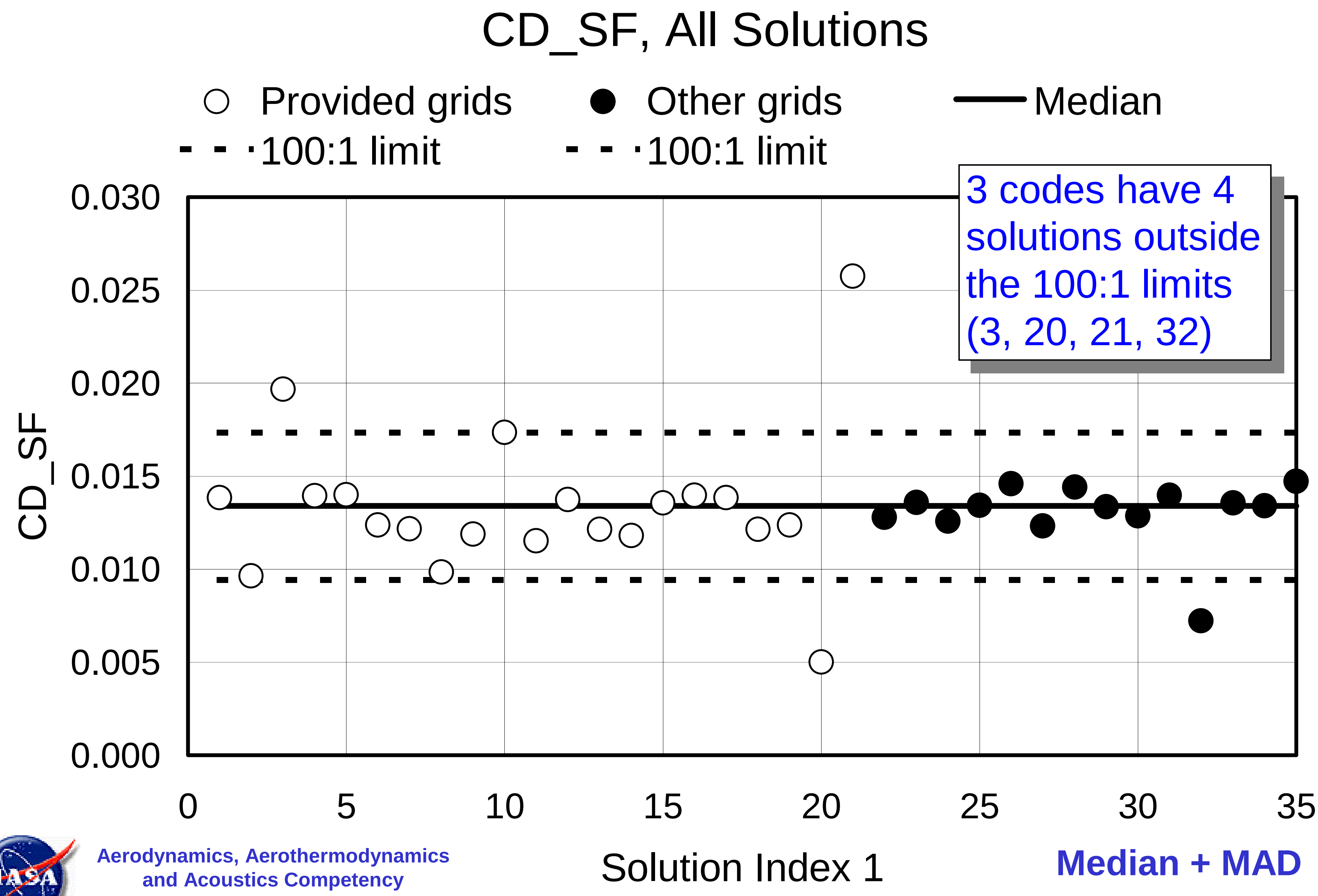
DPW



Median + MAD

ALL skin friction drag solutions at CL=0.5, M=0.75

DPW



Median + MAD

Drag point solutions at CL=0.5, M=0.75

Solutions by Index for Drag Point Table	CD_TOT	CD_PR	CD_SF
3	high		high
10	high	high	
20	low		low
21	high		high
32	low		low
33		high	
35		low	

None of the codes for which the provided grids were designed had any outlier solutions.

Hence, it seems that the provided grids were somehow not perfectly suited for some of the other codes.

Solutions 3, 10, 20, 21, 32 will not be used for the deeper analysis.

Statistics Summary I

- Solutions/codes which had all or part of the drag outside the 100:1 limits:
  - 7 out of 35 solutions (20%)
  - 6 out of 14 codes (43%)
  - 4 out of the 21 solutions on provided grids (19%)
  - 3 out of the 14 solutions on other grids (21%)

- There appears to be no significant difference in either location or scale between the drag solutions carried out on the provided grids and solutions carried out on grids developed by the participants, except possibly the scale for the skin friction.

Drag point solutions at CL=0.5, M=0.75

DPW

Statistics Summary II

	CD_PR	CD_SF	Sum/RSS	CD_TOT	Exp.
Location $\hat{\mu} = \tilde{x}$	166.2	133.9	300.1	292.9	286
Dispersion $\hat{\sigma} = 1.483\sqrt{\frac{n}{n-1}}MAD$	13.5	15.3	20.4	21.4	4
Solutions within 7 (10) counts of median	7 (20%)	15 (43%)		13 (37%)	
Solutions within 14 (20) counts of median	27 (77%)	25 (71%)		24 (69%)	

The inclusion rates for a Gaussian pdf would be roughly 38% and 68% respectively.

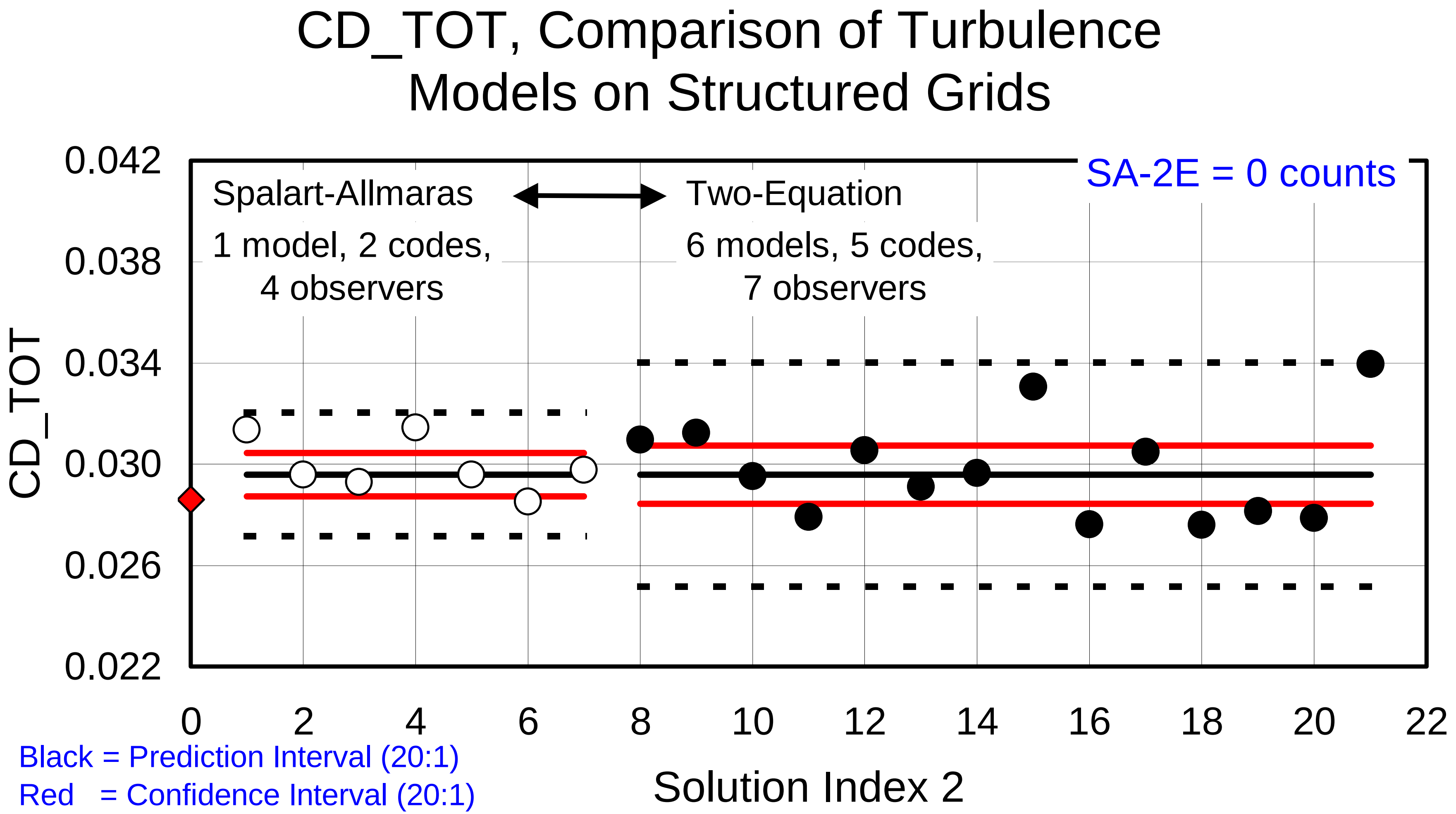
Effect of models and grid types

Turbulence Model	Structured Grids	Unstructured Grids
Spalart-Allmaras	8 (7*)	9 (7*)
Two-Equation (6)	16 (14*)	1

*Not including the outlier solutions found previously.

Effect of models on total drag (CL=.5, M=.75)

DPW

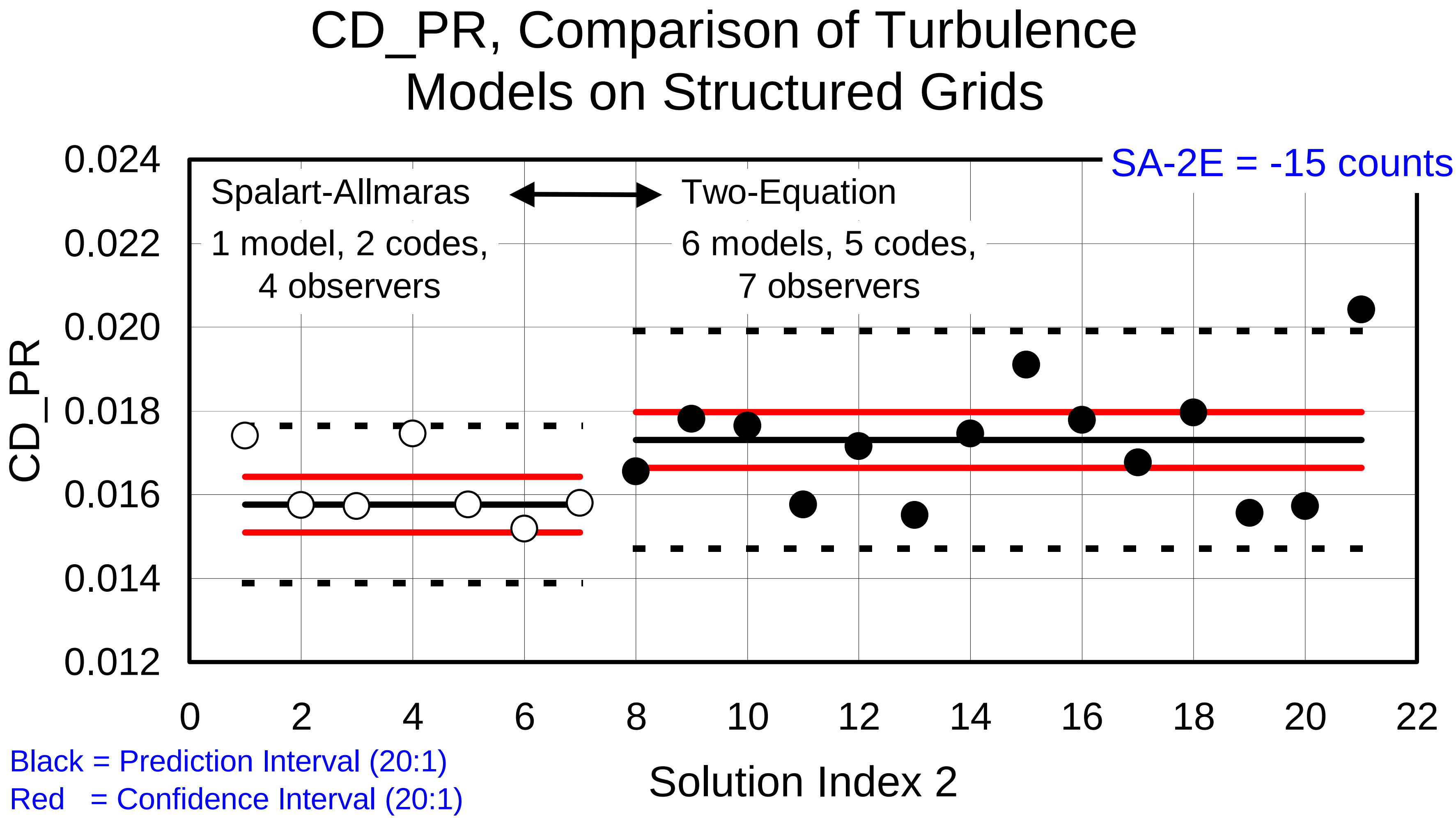


Black = Prediction Interval (20:1)
Red = Confidence Interval (20:1)

Median + AAD

Effect of models on Pressure drag (CL=.5, M=.75)

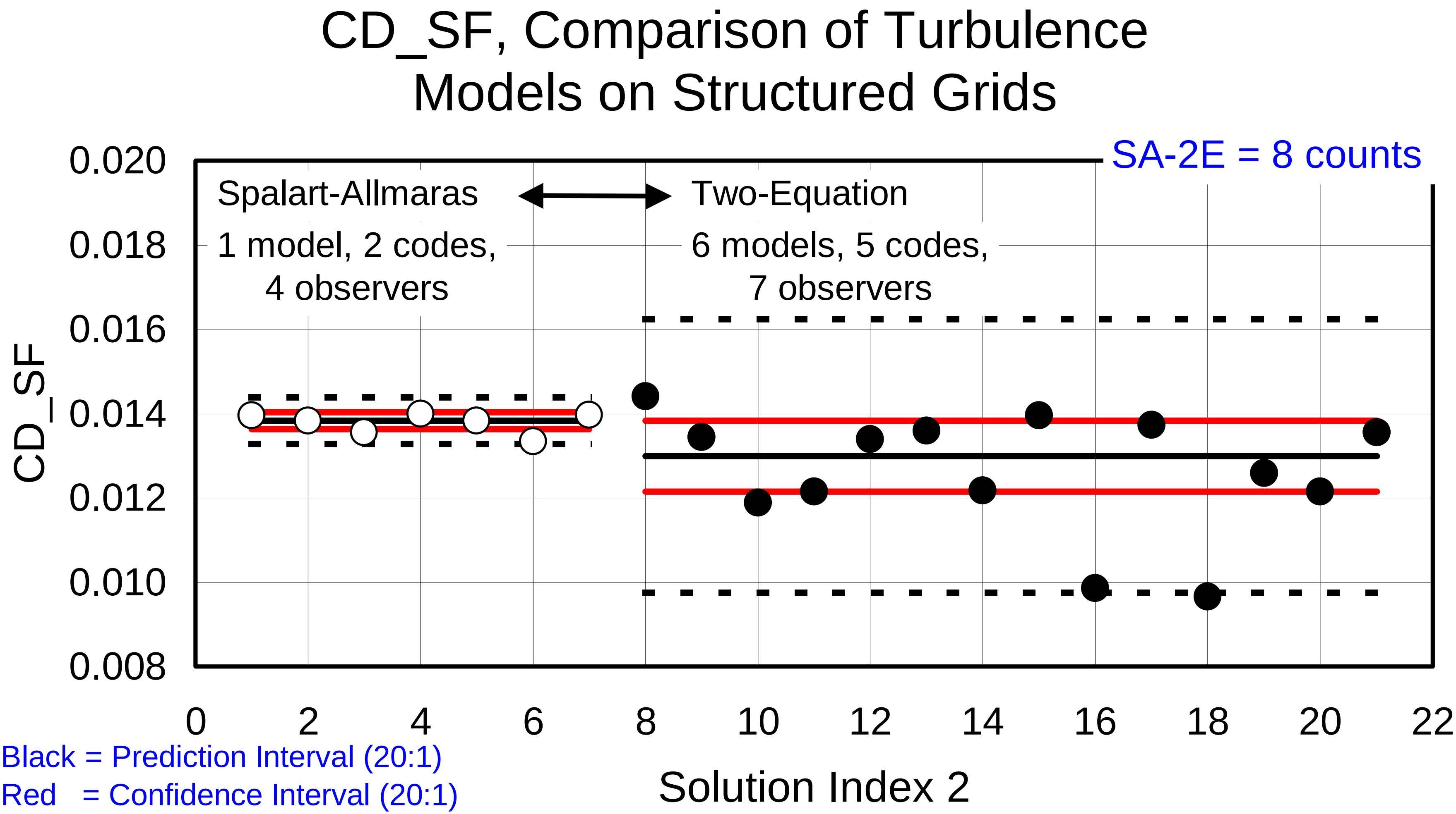
DPW



Median + AAD

Effect of models on skin drag (CL=.5, M=.75)

DPW

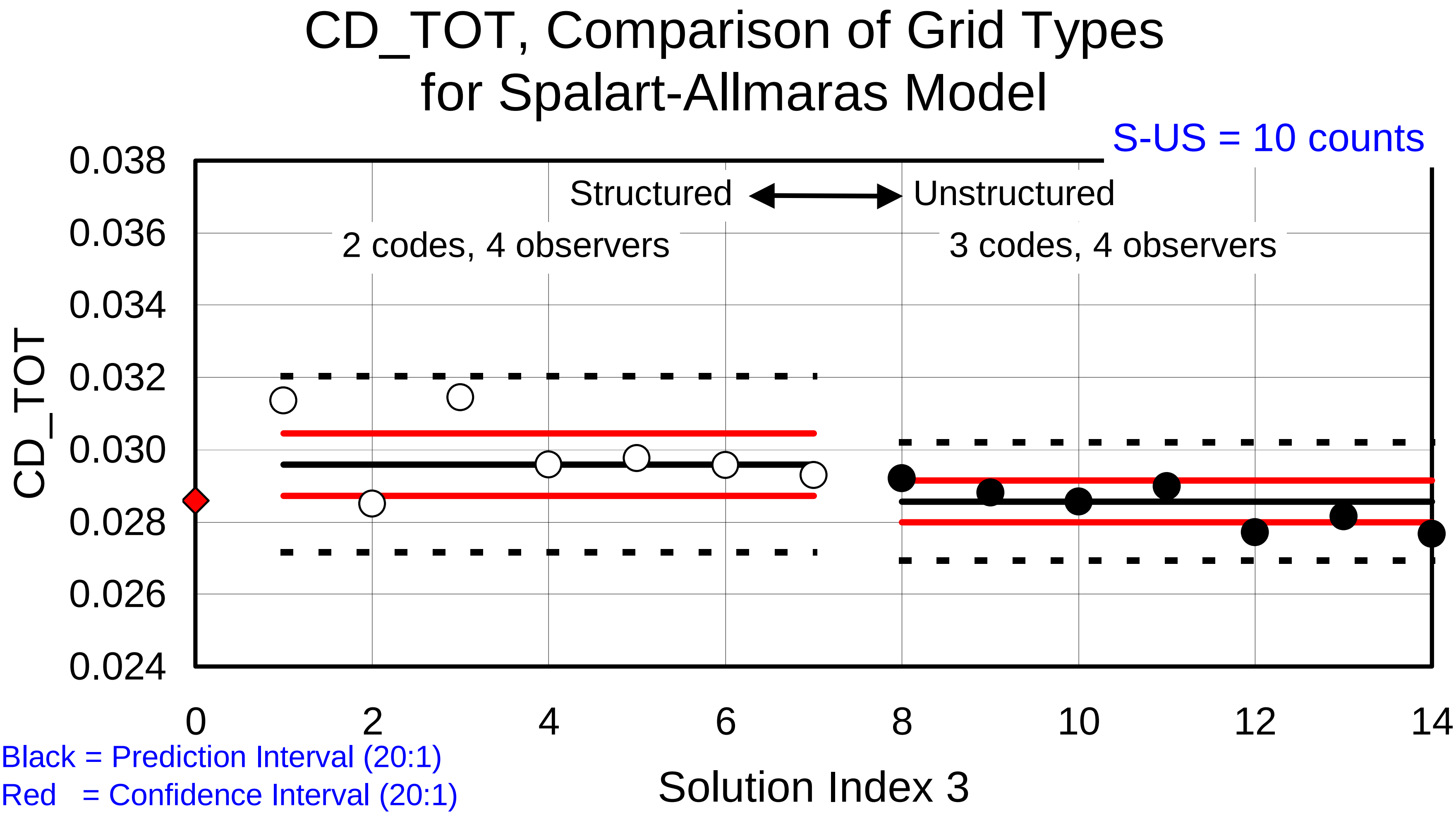


Black = Prediction Interval (20:1)
Red = Confidence Interval (20:1)

Median + AAD

Effect of grid type on total drag (CL=.5, M=.75)

DPW

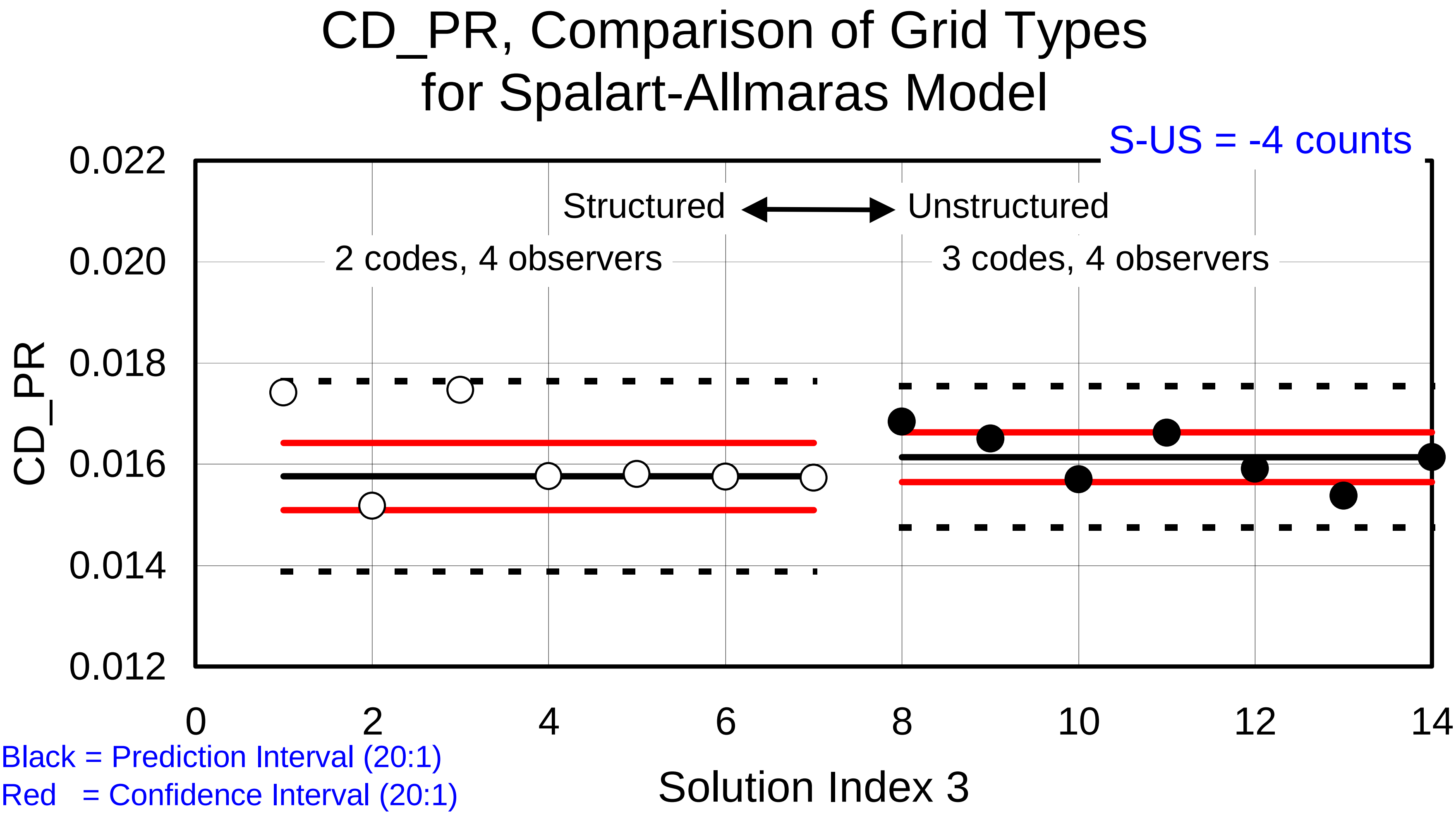


Black = Prediction Interval (20:1)
Red = Confidence Interval (20:1)

Median + AAD

Effect of grid type on pressure drag (CL=.5, M=.75)

DPW

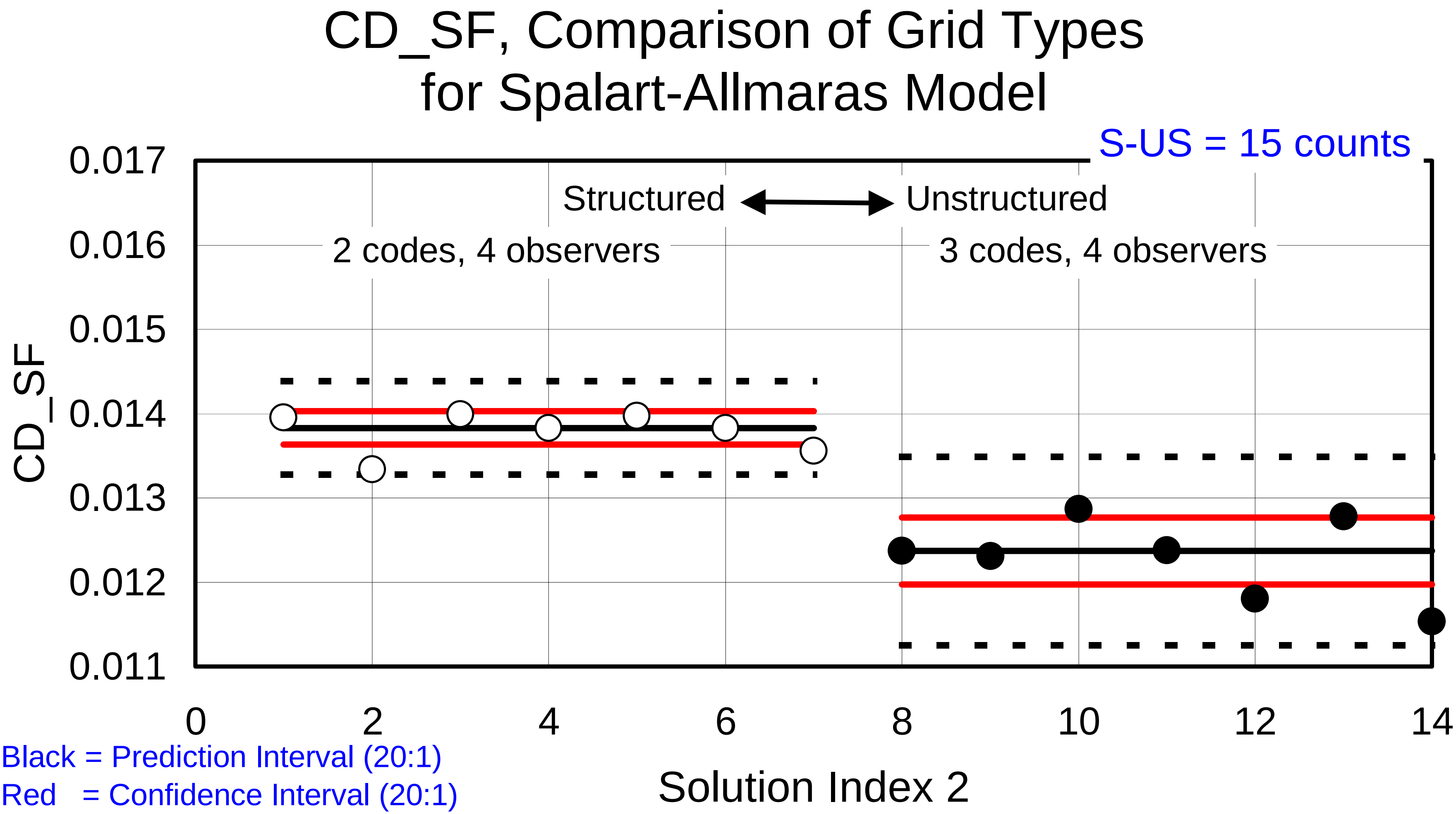


Black = Prediction Interval (20:1)
Red = Confidence Interval (20:1)

Median + AAD

Effect of grid type on skin drag (CL=.5, M=.75)

DPW



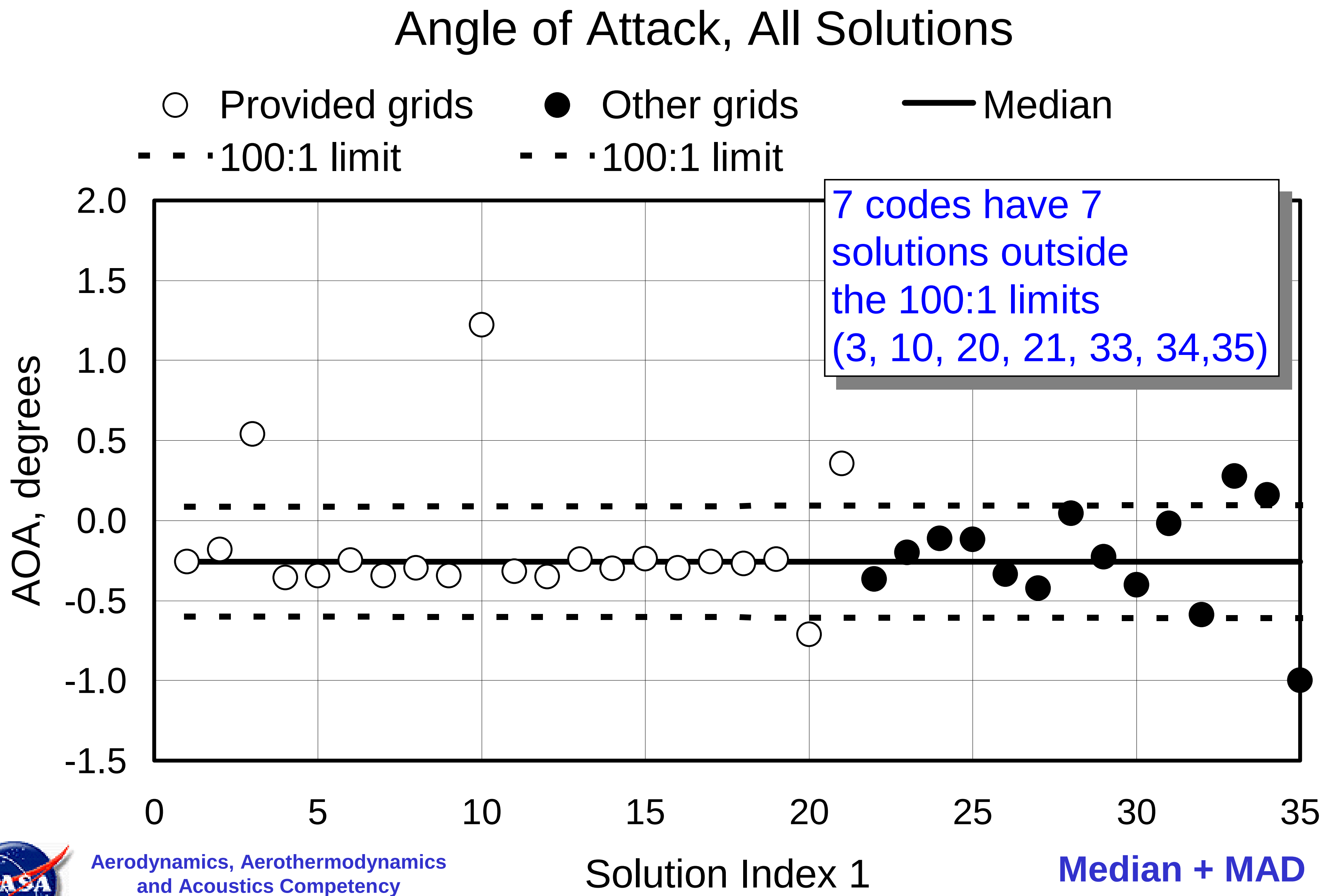
Black = Prediction Interval (20:1)
Red = Confidence Interval (20:1)

Median + AAD

Conclusions from drag point analysis

- All data
  - The skin friction and pressure components seem to be independent.
  - The central group pdf's seem to resemble a Gaussian with a standard deviation of roughly 20 counts.
  - The collection location is roughly 7 counts higher than that of the experimental result. That difference is probably not statistically significant.
- Effect of models on structured grids
  - The Wilcox and Menter models seem to give significantly different locations for the skin friction (approximately 13 counts).
- Effect of grid type for the SA model
  - Grid type seems to have a strong effect on the skin friction (approximately 15 counts).

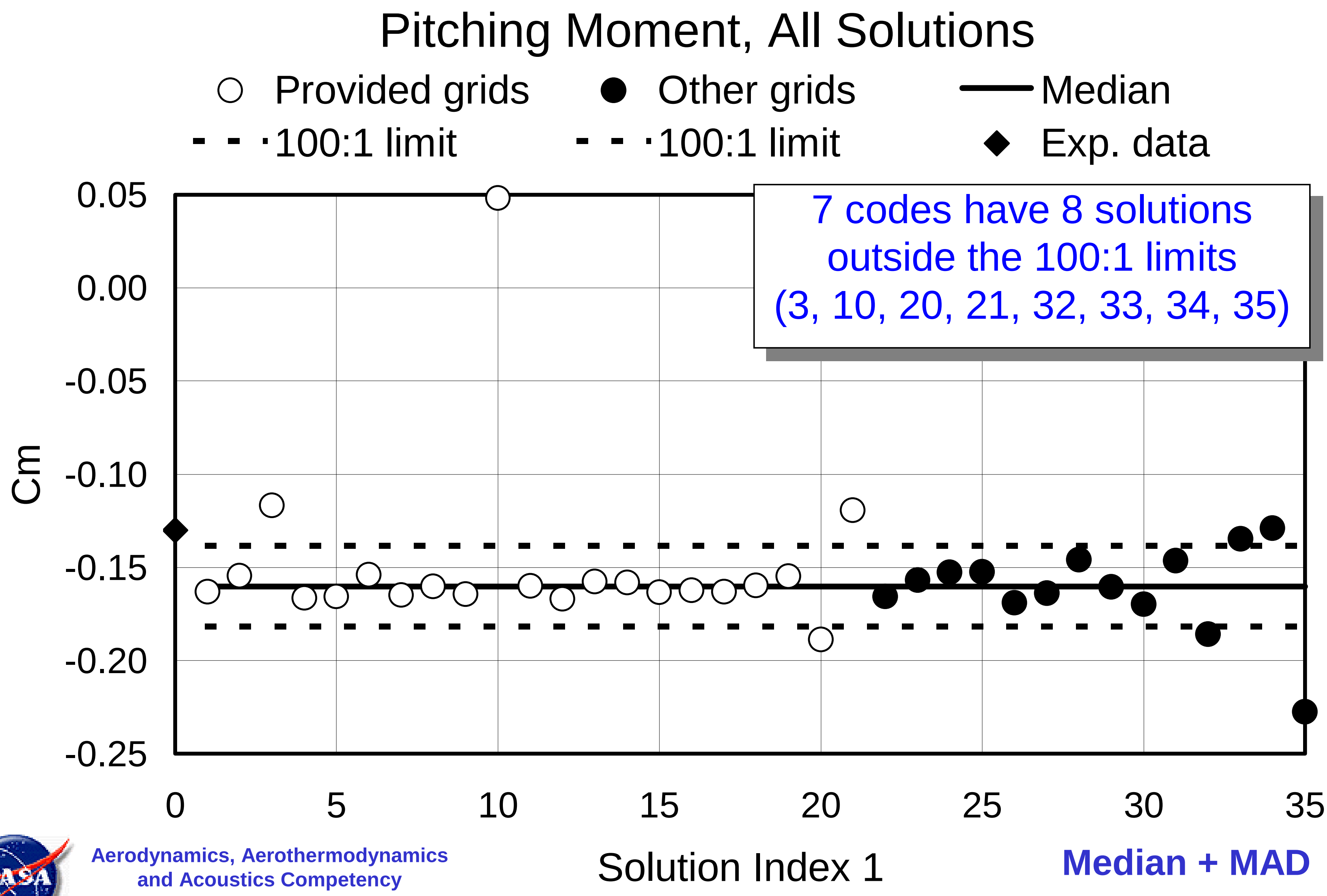
ALL alpha values at CL=0.5, M=0.75



Median + MAD

ALL pitching moment values at CL=0.5, M=0.75

DPW



Median + MAD

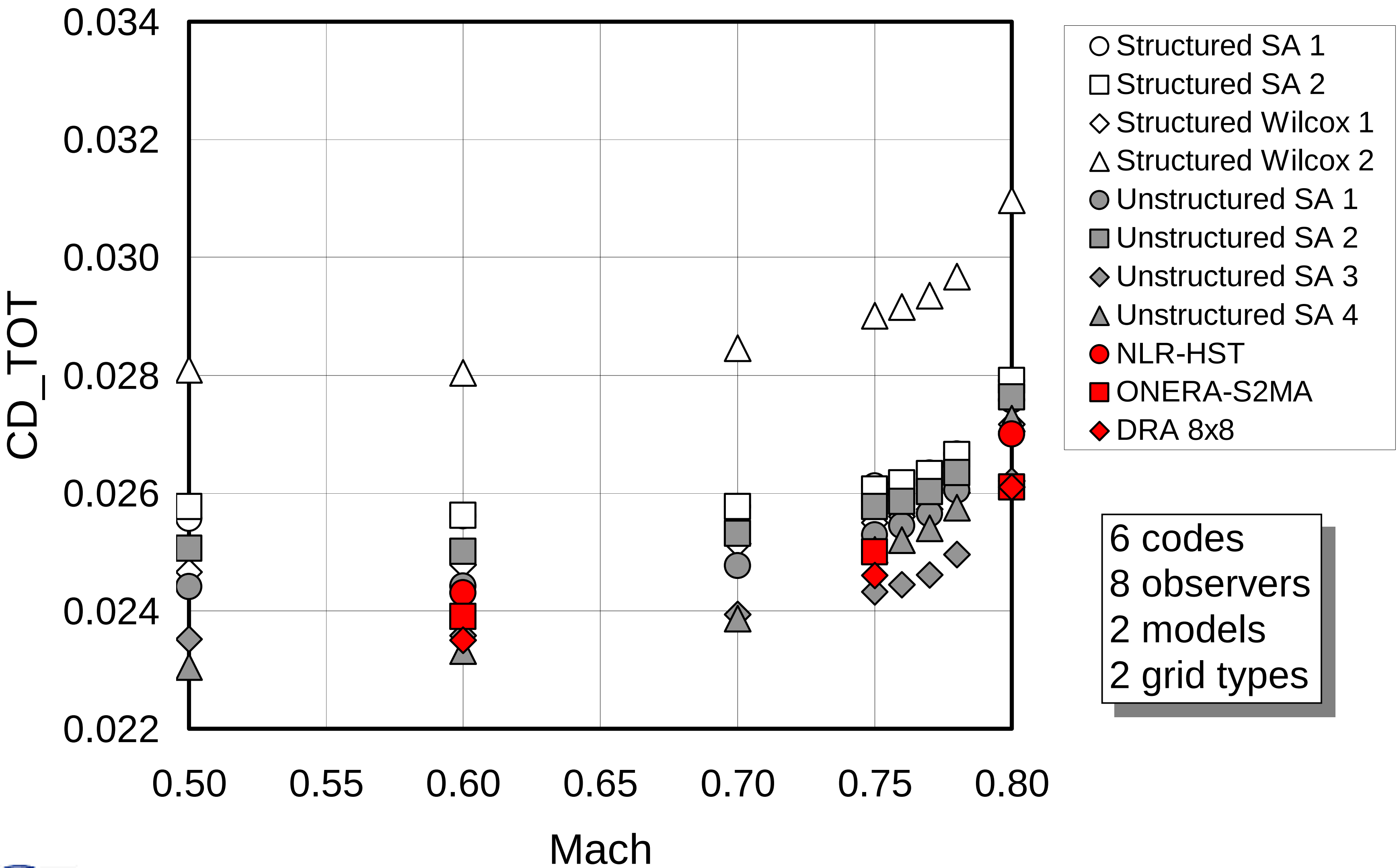
Conclusions from alpha and pitching moment point analysis

- **Alpha**
  - 7 out of 35 solutions were outside the 100:1 limits (20%).
  - 7 out of 14 codes had solutions outside the limits (50%)
  - The offset of the estimated location for alpha relative to the experimental value reflects calculation of the lift coefficient at roughly 10-15% too high.

- **Pitching moment**
  - 8 out of 35 solutions were outside the 100:1 limits (23%)
  - 7 out of 14 codes had solutions outside the limits (50%)
  - The offset of the estimated location for pitching moment relative to the experimental value reflects calculation of the aerodynamic center at roughly 6% of the MAC too far aft.

Drag Rise Curves for CL=0.4

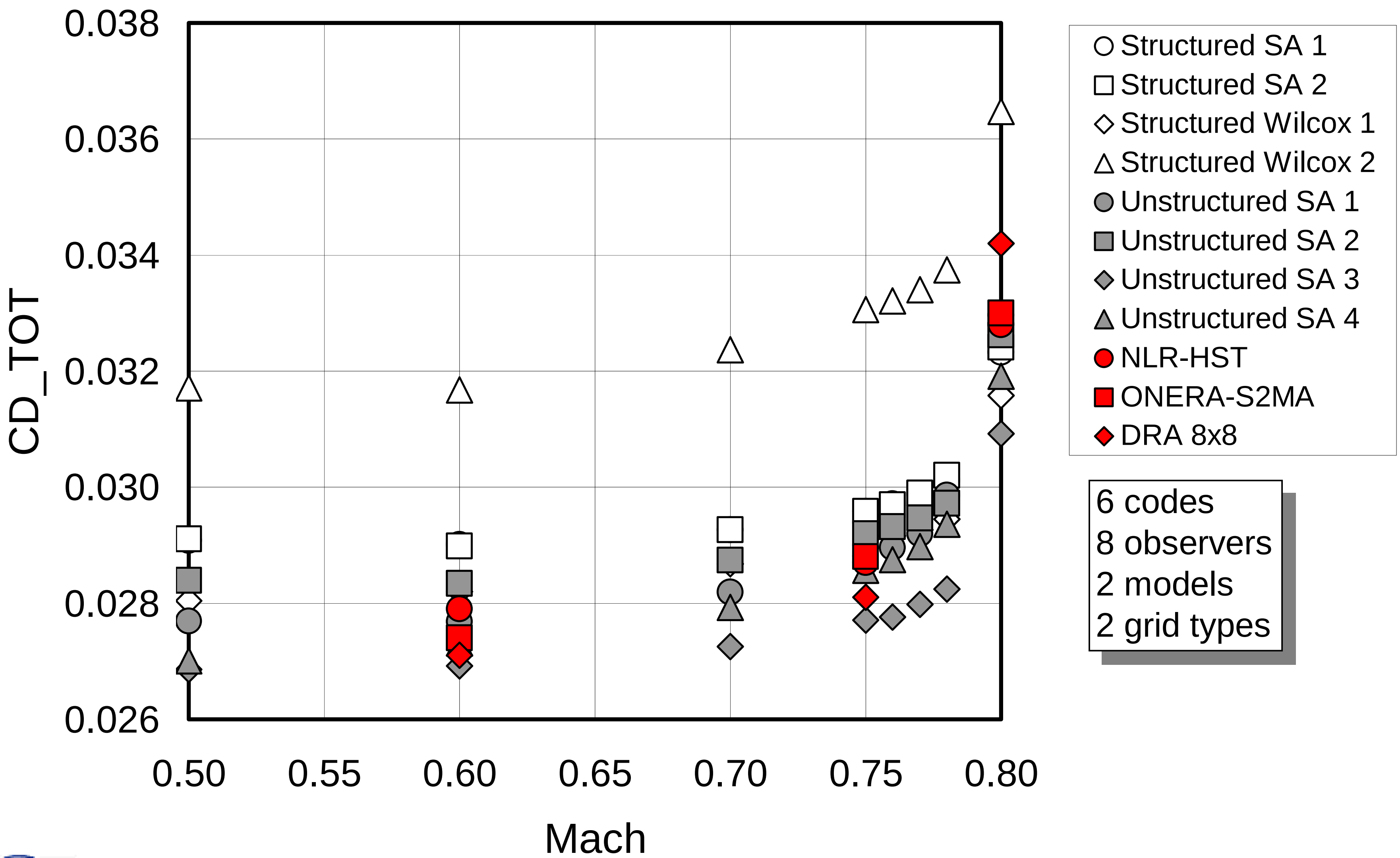
DPW



6 codes
8 observers
2 models
2 grid types

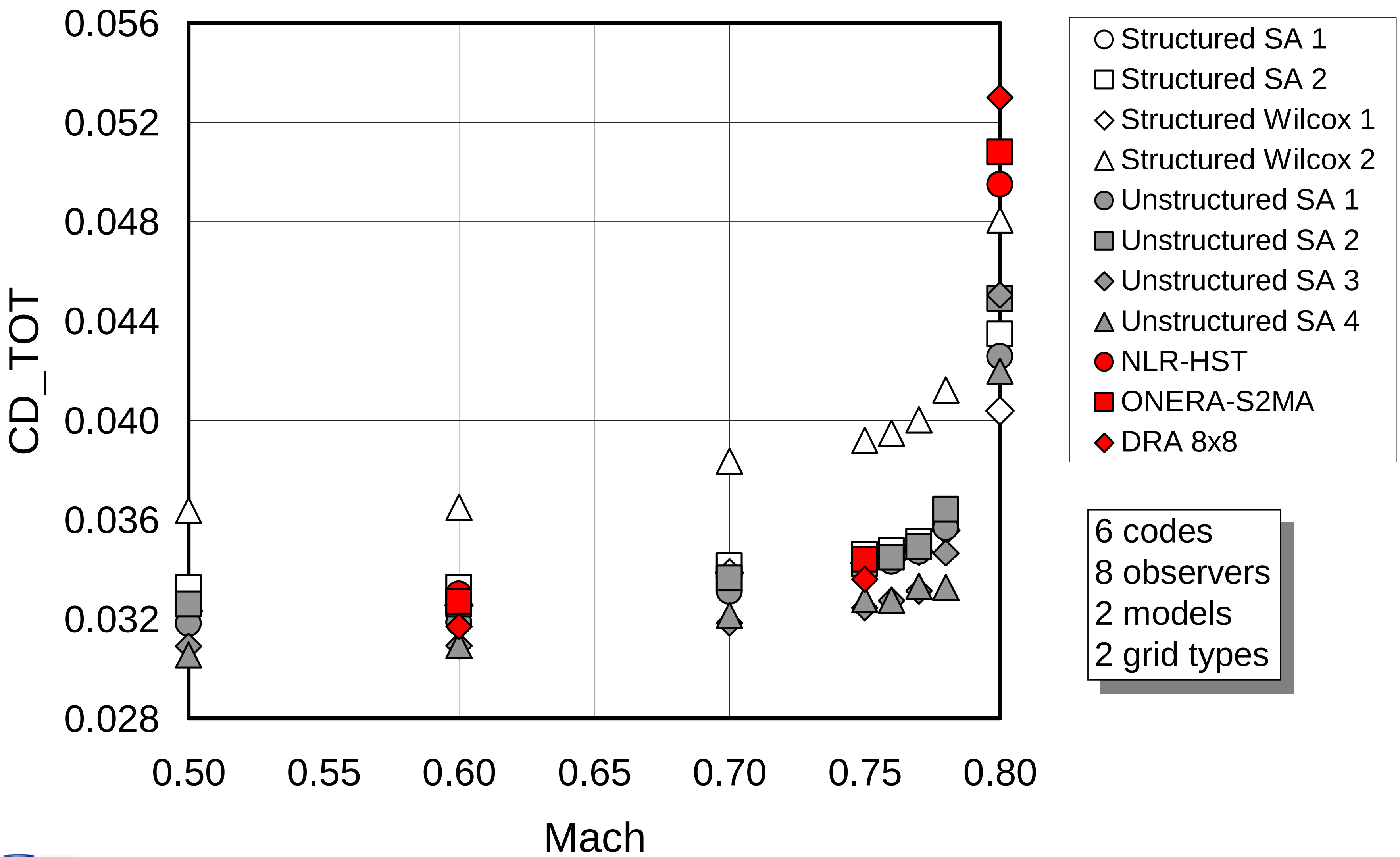
Drag Rise Curves for CL=0.5

DPW



Drag Rise Curves for CL=0.6

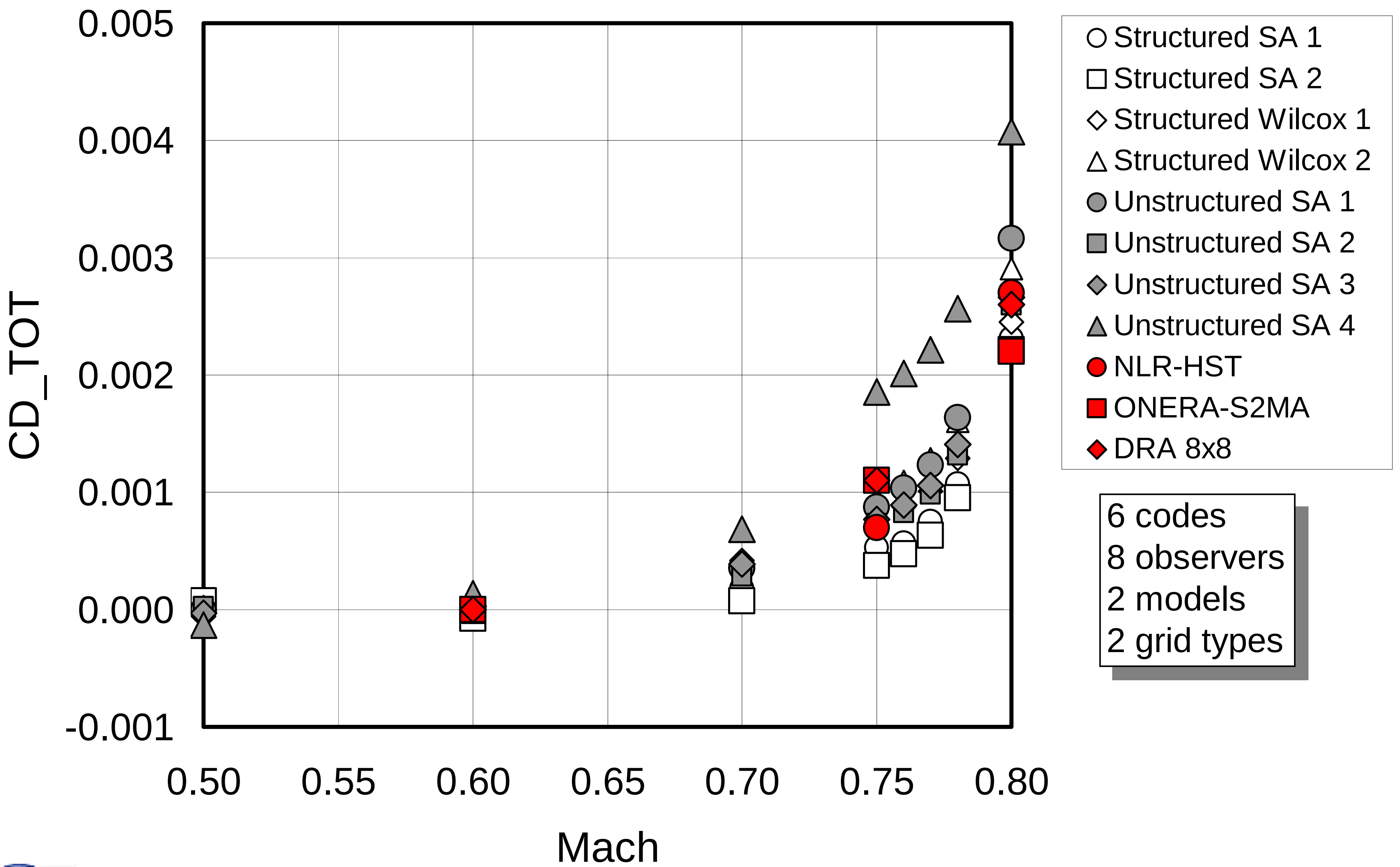
DPW



6 codes
8 observers
2 models
2 grid types

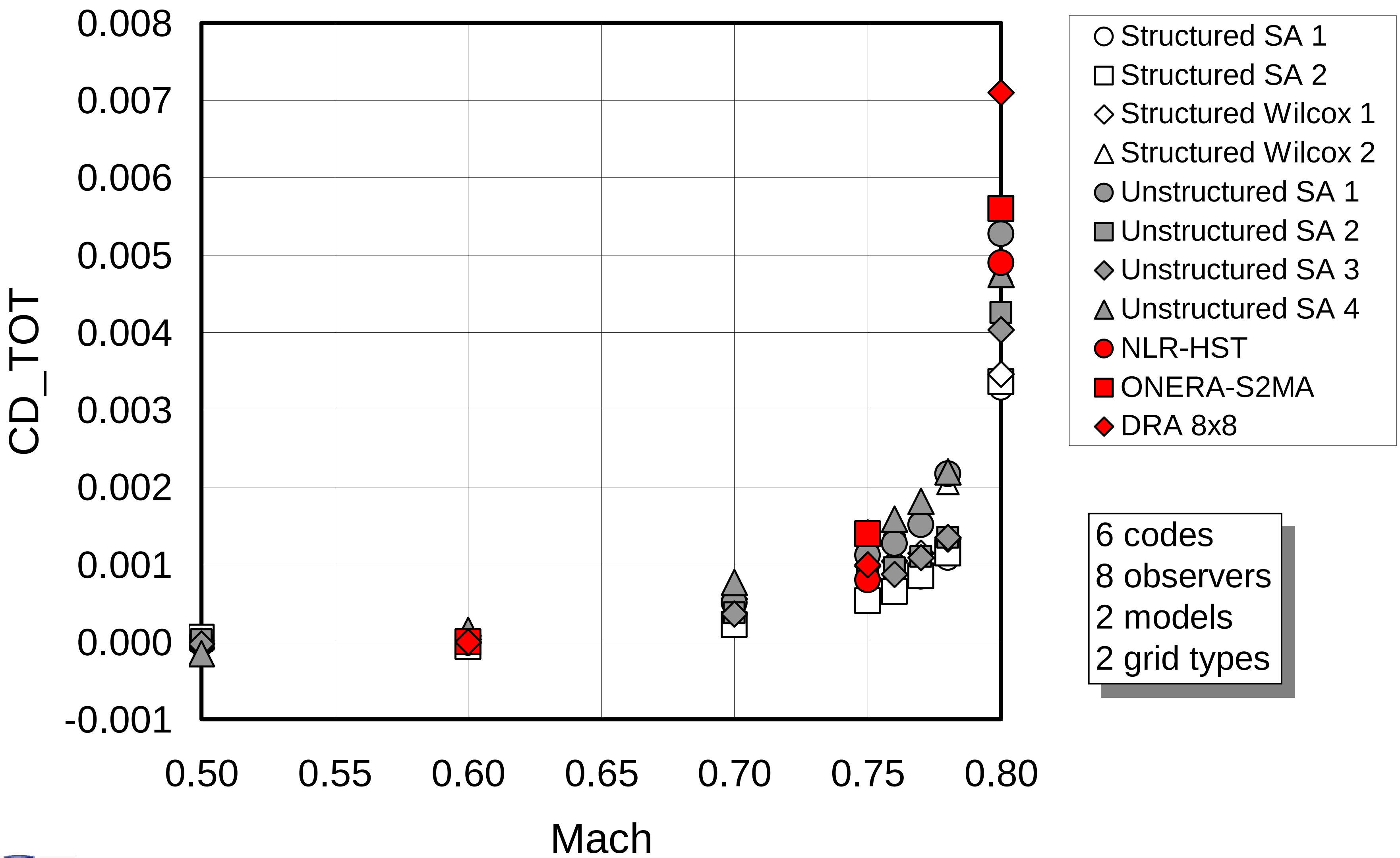
Adjusted Drag Rise Curves for CL=0.4

DPW



Adjusted Drag Rise Curves for CL=0.5

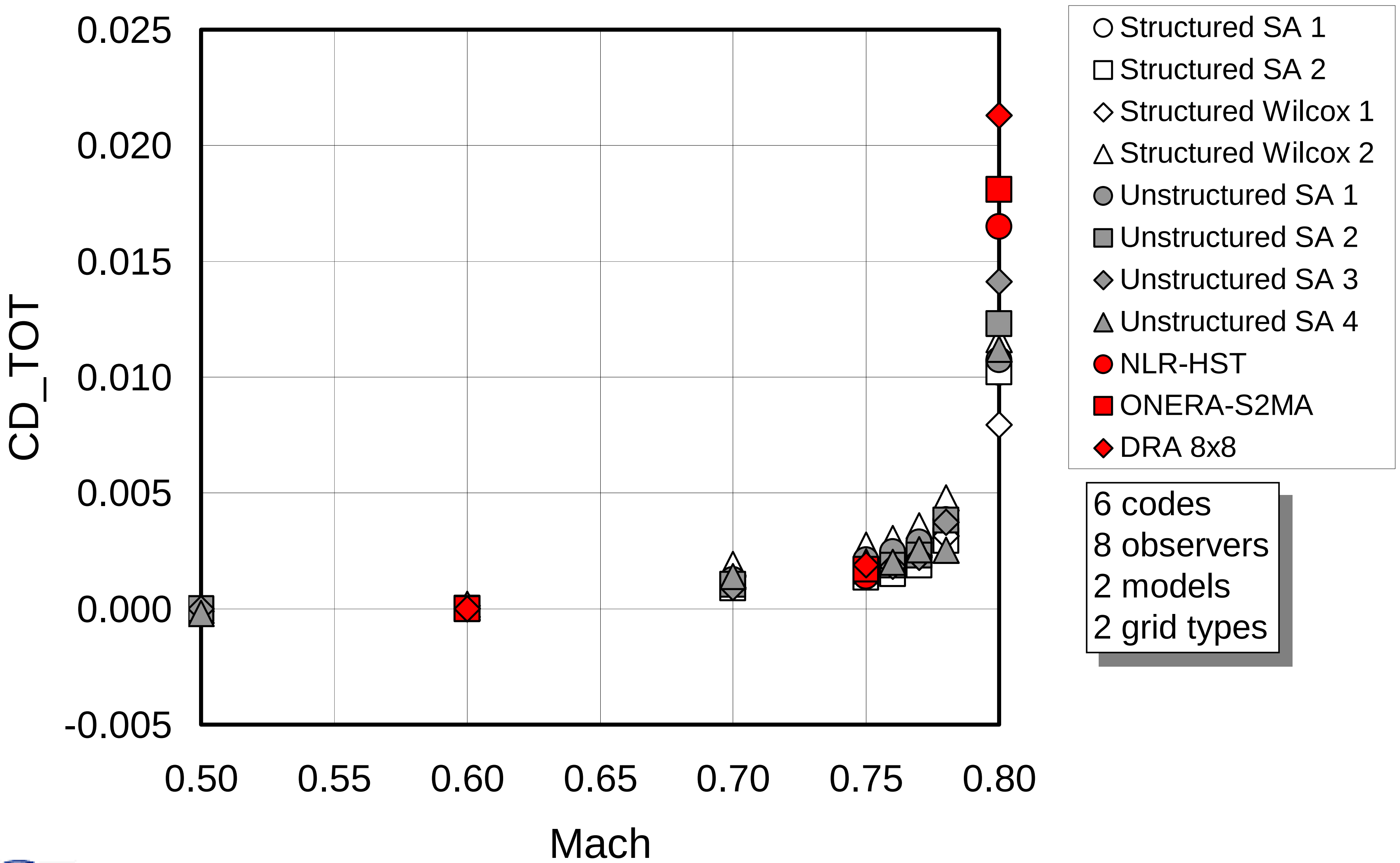
DPW



6 codes
8 observers
2 models
2 grid types

Adjusted Drag Rise Curves for CL=0.6

DPW



6 codes
8 observers
2 models
2 grid types

Conclusions from drag rise analysis

- Removing the scatter at the lower Mach numbers reveals that the solution scatter increases dramatically as the shock Mach number increases.
- For CL=0.4, 0.5, the solutions scatter about the wind-tunnel results.
- For CL=0.6 and M=0.8, the solutions strongly under-predict the wind-tunnel results.

Drag Polar Fits, M=0.75

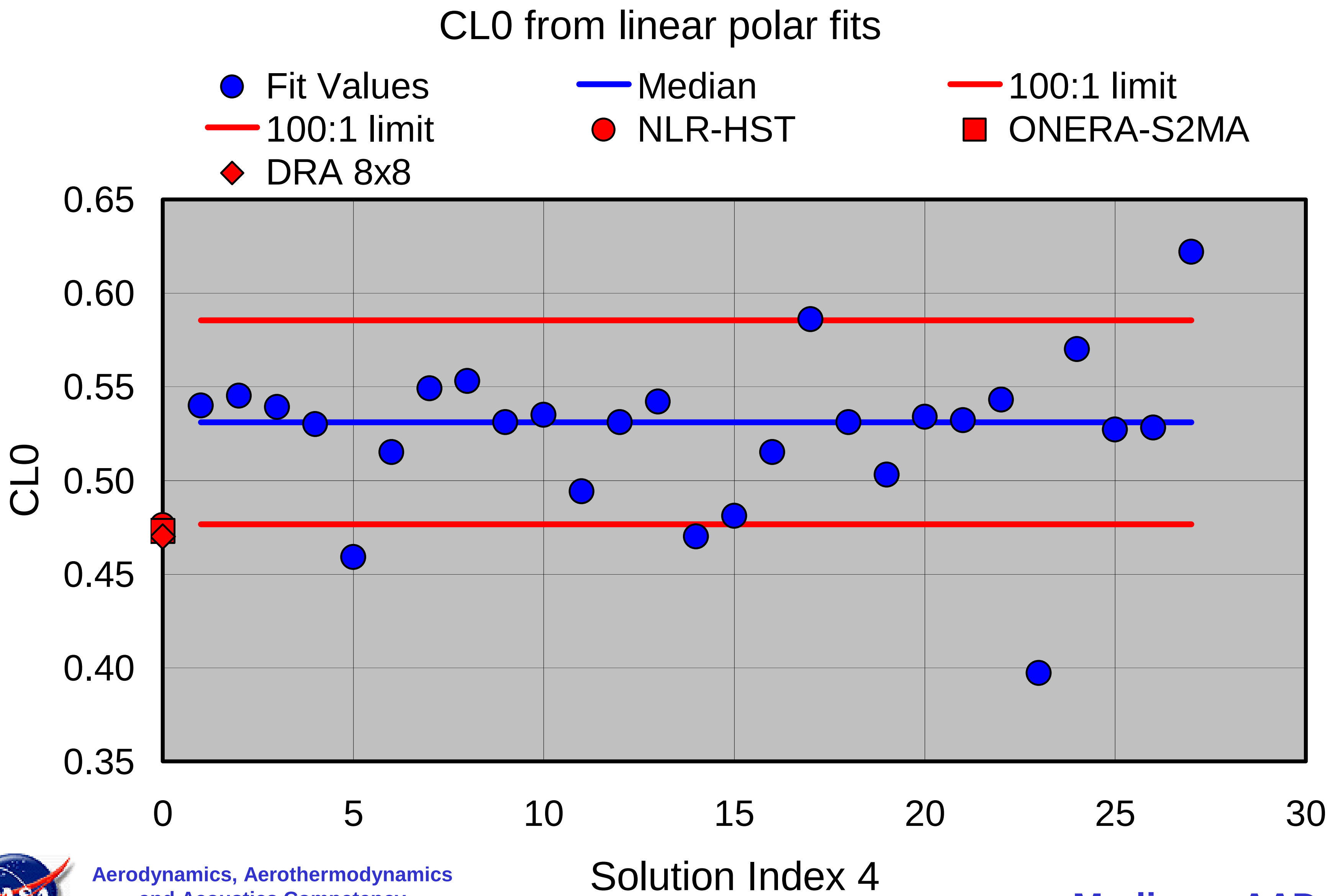
I used the following equations to fit the polars in the linear range (**CL=0.15 – 0.4**):

$$C_L = C_{L_0} + C_{L_\alpha}\,\alpha$$

$$C_D = C_{D_{C_L=0}} + k\,C_L^2$$

Drag Polar Fits, M=0.75

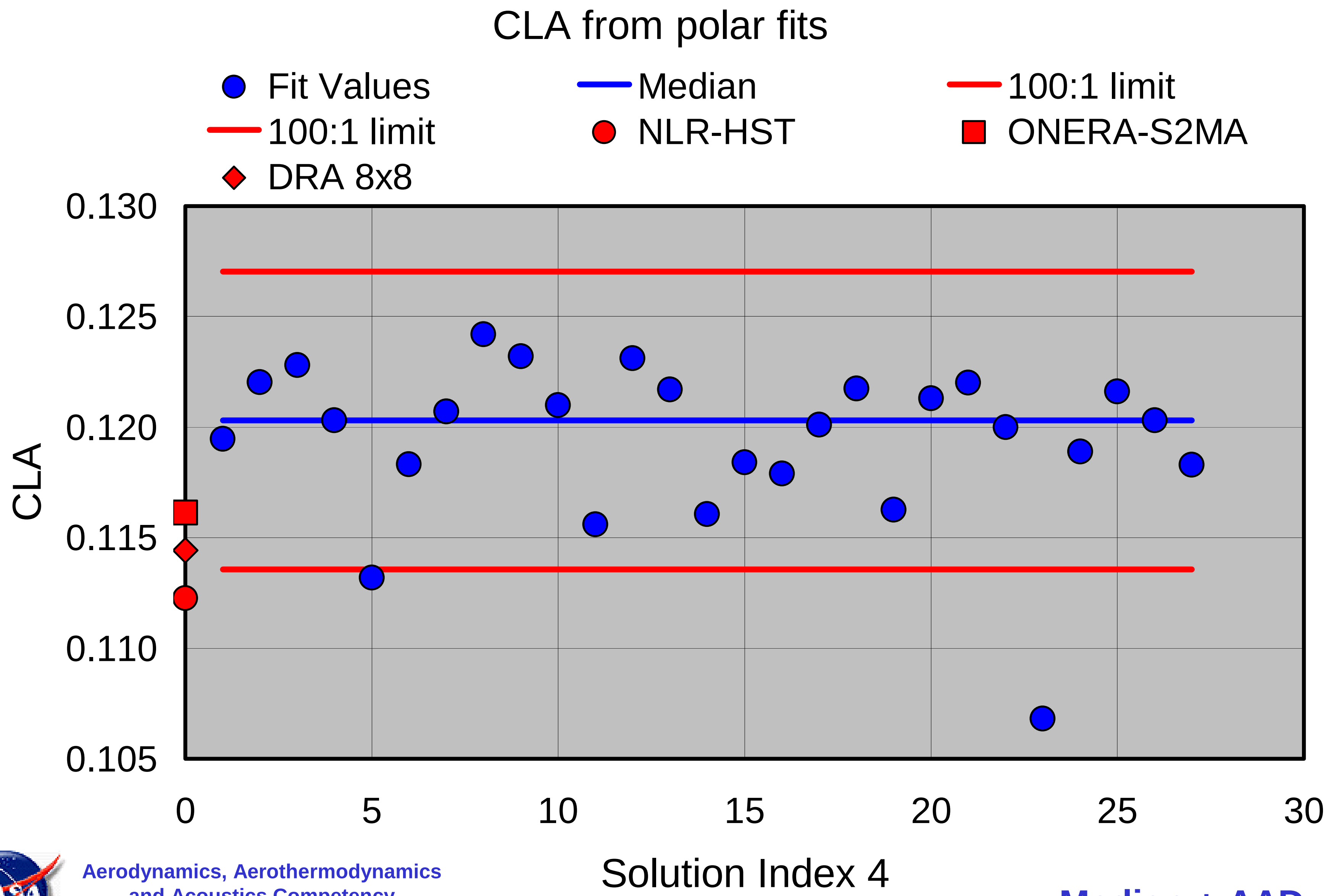
DPW



Median + AAD

Drag Polar Fits, M=0.75

DPW

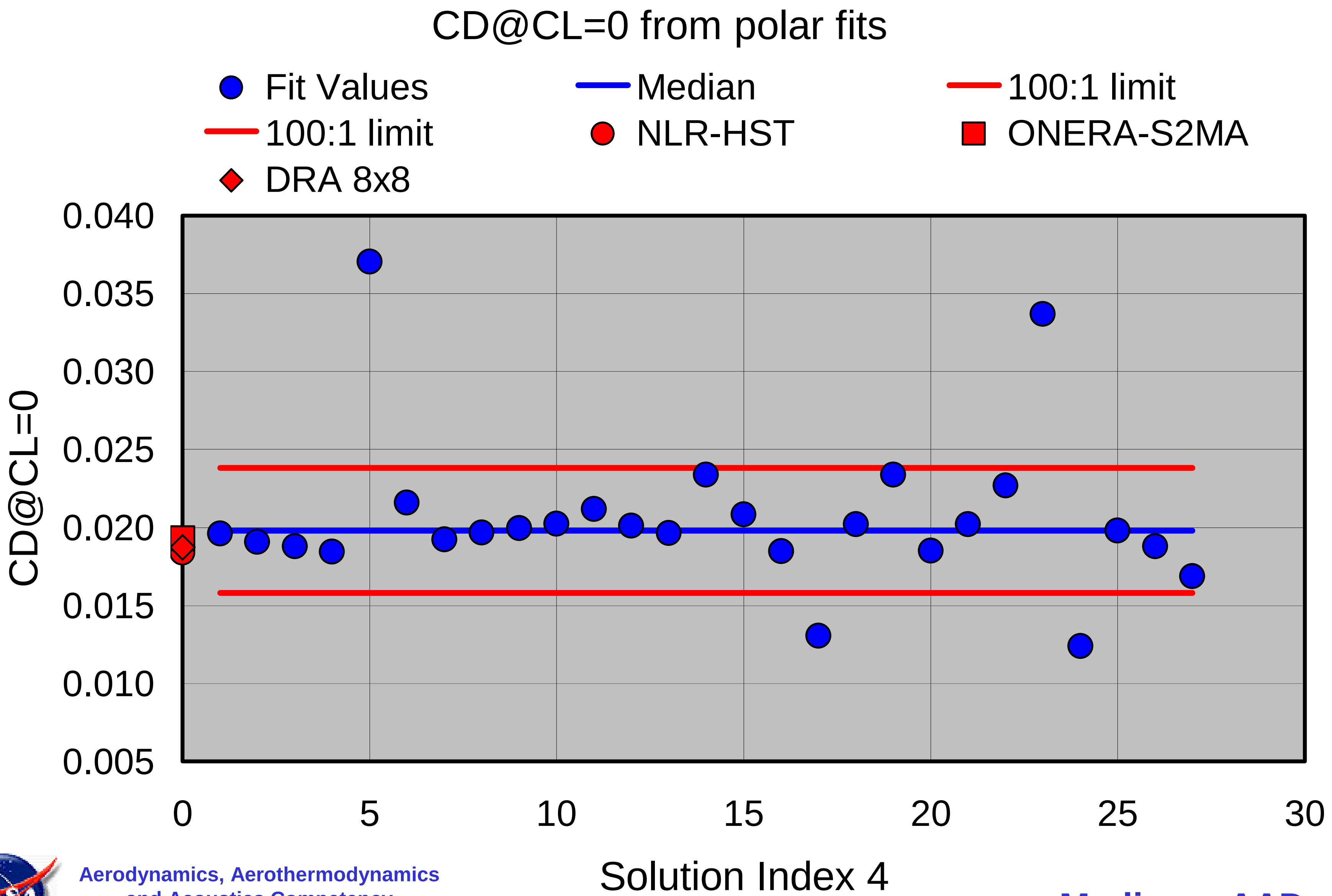


Aerodynamics, Aerothermodynamics
and Acoustics Competency
Langley Research Center

Median + AAD

Drag Polar Fits, M=0.75

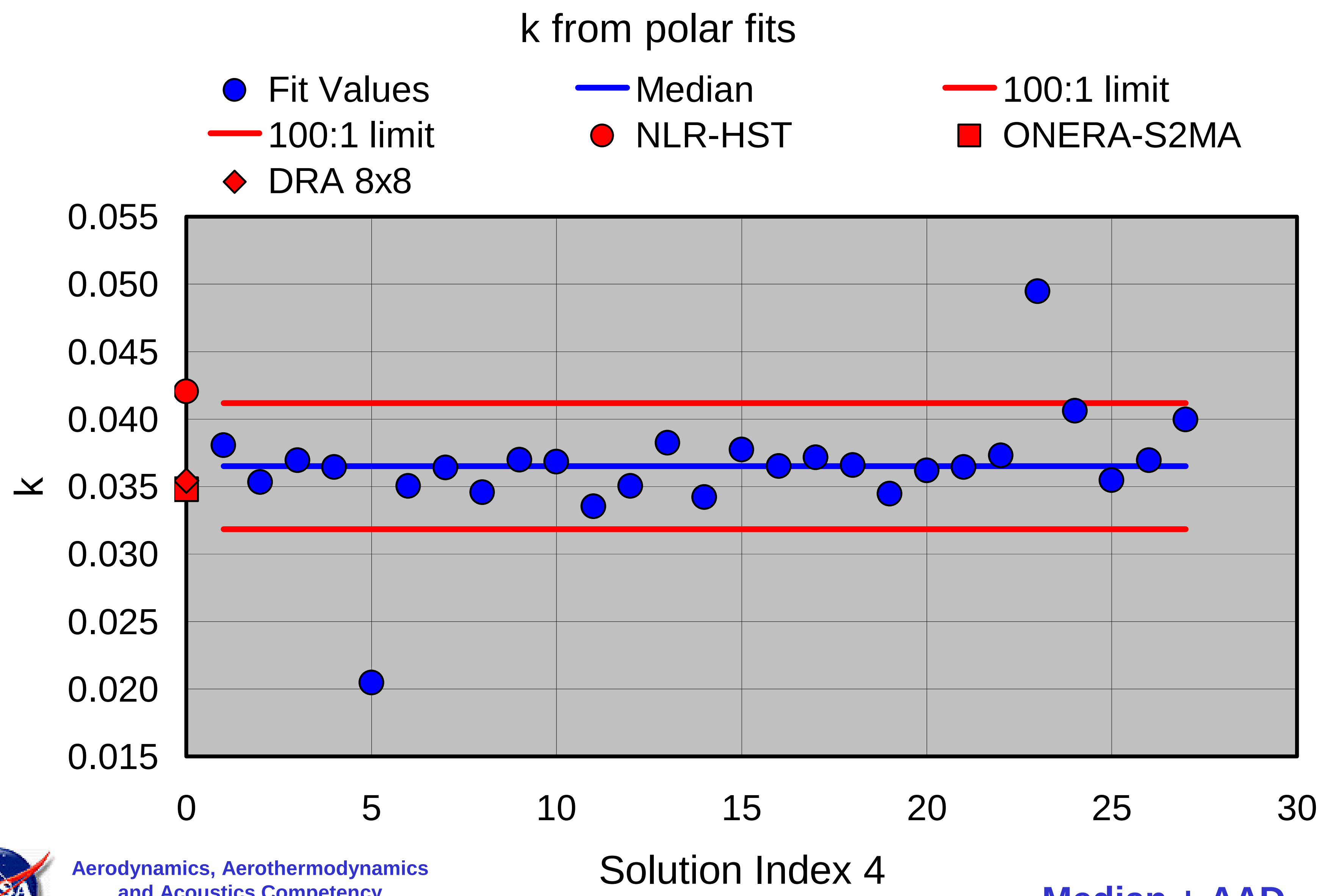
DPW



Median + AAD

Drag Polar Fits, M=0.75

DPW



Median + AAD

Drag Polar Fits, M=0.75

DPW

Summary 1

Solution	C_{L_0}	C_{L_α}	$C_{D_{C_L=0}}$	k
5	Low	Low	High	Low
14	Low			
17			Low	
23	Low	Low	High	High
24			Low	
27	High			

Summary 2

4 of the 27 polars were outliers in lift (15%).

4 of the 27 polars were outliers in drag (15%).

6 of the 27 polars were outliers in either lift or drag (22%).

Drag Polar Fits, M=0.75

DPW

Summary 3

	CFD	EXP	CFD	EXP
	$\hat{\mu}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\sigma}$
C_{L_0}	0.531	0.473	0.021	0.0035
C_{L_α}	0.120	0.114	0.0026	0.0023
$C_{D_{C_L=0}}$	200 cts	188 cts	18 cts	5.5 cts
k	365 cts	374 cts	18 cts	43 cts

Final Remarks

- Using the median to estimate the location and the MAD or AAD to estimate the scale allowed us to discern the outlier solutions without losing the meaning of the comparable core solution values.

- There does seem to be a credible CFD true value and standard deviation. Whether these numbers are durable can only be seen by repeating this exercise.

- It appears that we need some set of best practices and quantitative sanity checks to avoid outliers. The continued existence of such outliers would force us to accept much bigger numbers for the scatter.

Final Remarks

- The scatter for the core solutions is much too large for acceptable validation.
- Comparing CFD solutions to each other, in a collective sense, for diverse codes, grids, turbulence models, and observers, is probably the best way to determine the best practices needed to reduce the scatter to acceptable levels.
- We are probably not going to be able to reduce the drag scatter until we reduce the lift, pitching moment AND pressure distribution scatter.
- And, I must ask, Why does the type of grid make a difference in the skin friction?