Statistical Analysis of CFD Solutions from 2nd Drag Prediction Workshop

M. J. Hemsch* and J. H. Morrison#
NASA Langley Research Center, Hampton, Virginia

Abstract

In June 2001, the first AIAA Drag Prediction Workshop was held to evaluate results obtained from extensive N-Version testing of a series of RANS CFD codes. The geometry used for the computations was the DLR-F4 wing-body combination which resembles a medium-range subsonic transport. The cases reported include the design cruise point, drag polars at eight Mach numbers, and drag rise at three values of lift. Although comparisons of the code-to-code medians with available experimental data were similar to those obtained in previous studies, the code-to-code scatter was more than an order-of-magnitude larger than expected and far larger than desired for design and for experimental validation. The second Drag Prediction Workshop was held in June 2003 with emphasis on the determination of installed pylon-nacelle drag increments and on grid refinement studies. The geometry used was the DLR-F6 wing-body-pylon-nacelle combination for which the design cruise point and the cases run were similar to the first workshop except for additional runs on coarse and fine grids to complement the runs on medium grids. The code-to-code scatter was significantly reduced for the wing-body configuration compared to the first workshop, although still much larger than desired. However, the grid refinement studies showed no significant improvement in code-to-code scatter with increasing grid refinement.

Nomenclature

AMR	average moving range
AOA	angle of attack
b	model wing span
CDPR	pressure drag coefficient
CDSF	skin-friction drag coefficient
CDTOT	total drag coefficient
CL	centerline
CM	pitching-moment coefficient
GRIDGEN	GRIDGEN-generated grids
HYBRID	unstructured grids not generated by the workshop organizers
ICEM	ICEM-generated grids
K	confidence interval coverage factor
KE	$k-\varepsilon$ turbulence model
KW	Wilcox $k-\omega$ turbulence model
KWASM	algebraic stress model coupled with *k-w* turbulence model
LARC	grids generated by NASA Langley Research Center
LARCA	modified version of grids generated by NASA Langley Research Center
MAD	estimate of the population standard deviation based on the median of the absolute deviations from the median
MB	Multiblock one-to-one grids
M_∞	Mach number
OV	overset-grid solver
PUFGG	grids generated by PUFGG system
R	range (maximum value minus minimum value)
RANS	Reynolds-averaged Navier Stokes equations
Re	Reynolds number based on the mean aerodynamic chord
SA	Spalart-Allmaras turbulence model
SA1A	1A version of Spalart-Allmaras turbulence model
SAE	Spalart-Allmaras turbulence model with Edwards modification
SST	Menter SST turbulence model
UN	unstructured-grid solver
y	spanwise distance measured from fuselage centerline
$\hat{\mu}$	estimate of the population mean
$\hat{\sigma}$	estimate of the population standard deviation

* Aerospace Engineer, Associate Fellow, AIAA
# Research Scientist, Senior Member, AIAA

Introduction

In June 2003, the AIAA Applied Aerodynamics Technical Committee (APATC) sponsored a second Drag Prediction Workshop (DPW II) for transonic cruise drag predictions of subsonic transports. The workshop was a follow-on to the first Drag Prediction Workshop (DPW I) held in June 2001.[1-6] The objectives for both workshops were (1) to assess state-of-the-art computational methods as practical aerodynamic tools for aircraft force and moment prediction, (2) to impartially evaluate the effectiveness of existing computer codes and modeling techniques, and (3) to identify areas needing additional research and development.

The challenge for DPW I was to compute the lift, drag and pitching moment for the DLR-F4 wing-body configuration[7-9] for three sets of conditions (all at Re=3.0 x 10^6):

1. Cruise at $M_\infty = 0.75, C_L = 0.5$ (required)
2. Drag polar at $M_\infty = 0.75$ (required)
3. Drag rise at $C_L = 0.4, 0.5, 0.6$ (optional)

The DLR-F4 wing-body configuration is shown in Figure 1. For the participants, the primary purposes of the challenge were to determine how accurately each of their individual codes compared to experiment and how well they compared to each other (code-to-code scatter). Although most of the codes performed relatively well, there were two major surprises from the code-to-code statistical analysis: (1) roughly 20% of the solutions were statistical outliers compared to the rest and (2) the code-to-code scatter was more than an order-of-magnitude larger than desired.[6] Thus, a major outcome of the workshop was the realization that some set of best practices and quantitative sanity checks is needed to avoid outliers and to control the code-to-code scatter in industrial practice.

For DPW II the emphasis was on the determination of installed pylon-nacelle drag increments and on grid refinement studies with the hope of seeing reduced code-to-code scatter. The geometries used were the DLR-F6 wing body (WB) and wing-body-pylon-nacelle (WBPN). The configuration is described in Ref. 10 and shown in Fig. 2 mounted in the ONERA S2MA tunnel. The cases run were similar to the first workshop except for additional runs on coarse and fine grids to complement the runs on medium grids. An overview of the results is presented in Ref. 11. Individual papers by the workshop participants are presented in Refs. 12-22.

This paper presents a statistical analysis of the code-to-code scatter of the participants' results from the second workshop and is divided into five sections. The first summarizes the statistical point of view that has been adopted for the analysis presented in Ref. 6 and in this paper. The second section briefly discusses the flow patterns on the challenge model that the workshop participants believed are the primary reason that the code-to-code scatter is so large. In the third section, the code-to-code scatter results for the two workshops are compared for the wing-body configurations on medium grids. In the fourth section, the code-to-code scatter is examined for coarse, medium and fine grids to determine if grid convergence for the collective is manifested. The fifth section presents the surface pressure results of seven different cruise-point solutions on fine grids and the sixth section presents some final remarks.

Statistical Approach for N-Version Testing[a]

N-Version Testing

For this paper, the point of view is taken that the scatter (dispersion) of the results computed by the participants for any given output coefficient at a given set of conditions represents the *reproducibility* of the computational process just as if the individual computed realizations were obtained from a replicated measurement process. In measurement, such a process is called N-th Order Replication.[23] In computation, a similar process is called N-Version Testing.[24] Reproducibility is defined[25] for measurement as *closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement*. The changed conditions of "measurement" for the DPWs are, of course, the different codes, solution methods, turbulence models, grids, computing platforms, observers (people who carried out the computational process) and so on.

For this type of analysis, no individual outcome (computational realization) is considered the "right" answer or "best" result. To be specific, the *collective* computational process is considered to consist of all of the individual processes used and the dispersion of the results to be noise in that collective computational process. This viewpoint has been suggested by Youden[26] of the former National Bureau of Standards for precision measurements of physical constants at different laboratories. It is also consistent with the frequentist[27] interpretation of probability and with the new international standard for reporting measurement uncertainty[25]. There are three requirements for the existence of a collective and, hence, meaningful statistics: (1) a group of "individuals" that are more alike than not (e.g. all RANS codes), (2) the individuals in the collective differ in ways that are not accounted for in the sampling (e.g. grid type, coding approaches, turbulence models, etc.), and (3) the key properties of the collective do not change with time.

The desire to analyze the results of N-Version testing in such a way as to minimize the possibility of drawing incorrect conclusions from the aggregated data for the collective leads to the need for robust

[a] Most of the remarks of this section are taken from Ref. 6.

statistical methods which can answer the following questions:

1. Which solutions constitute a reasonable "core" set and which solutions lie outside that core (i.e which are outliers)?
2. What is a reasonable estimate of the population mean of the core solutions?
3. What is a reasonable estimate of the standard deviation of the core solutions?

Analysis Methods

In order to answer the three questions posed above, it is necessary to use a method for which outliers do not have a large effect on the estimates of the population mean and standard deviation. For the purposes of this paper, statistical control charts are used for comparisons of the code-to-code scatter in the two workshops.[28,29]

Construction of statistical control charts starts with running records of the results to be compared. For experimental measurements, it is useful to display the data values on the vertical axis and the time of the data point acquisition on the horizontal axis. However, since time is irrelevant in the DPW challenges, the individual solution values are plotted herein as a function of an integer index. The order of the plotting of the solution values along the abscissa was randomly assigned.

To enhance the possibility of discerning significant results in the running record, an estimate of the population mean ($\hat{\mu}$) of the plotted data points is made and shown on the graph as a *centerline*. In addition, *scatter limits* are placed about the centerline as follows:

$$\begin{aligned} \text{Lower Limit} &= \hat{\mu} - K\hat{\sigma} \\ \text{Upper Limit} &= \hat{\mu} + K\hat{\sigma} \end{aligned} \tag{1}$$

where $\hat{\sigma}$ is an estimate of the population standard deviation and K is an appropriate coverage factor.[28,29] The area between the scatter limits is considered to be a "noise zone" within which it is not possible to discern significant differences for the individual values - *at least not on the basis of statistics.*

The conventional estimate of the population mean of the collective, the arithmetic average is:

$$\hat{\mu} = \bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

However, conventional estimates of the population standard deviation can be grossly inflated if the sample is contaminated with values which do not belong to the population of interest (outliers). For the purposes herein, the simplest way to achieve robustness is to use the average moving range (AMR) to estimate the population standard deviation

$$\begin{aligned} \hat{\sigma} &= \text{AMR}/1.128 \\ &= \frac{1}{1.128(n-1)}\sum_{i=1}^{n-1}\left|x_{n+1} - x_n\right| \end{aligned} \tag{3}$$

For eight decades, it has been found that using K=3 for statistical control charts is a good compromise between mistaking noise for a signal and vice versa.[28,29]

For the analysis of the grid refinement study results, a different approach has been used. Because there is significant correlation between subsets of the results, particularly with respect to grids and turbulence models, it is difficult to assign proper limits to the code-to-code scatter for a given confidence level. Hence, the authors chose to perform a relatively simple analysis and display all of the data for comparison rather than aggregates. In order to make the analysis robust, the median and the median absolute difference (MAD) were used to estimate the population means and standard deviations respectively. The formulas for those estimates are (for sorted data)

$$\begin{aligned} \hat{\mu} &= \tilde{x} \\ \tilde{x} &\equiv x_{(n+1)/2} \quad (n\ odd) \\ &\equiv 0.5(x_{n/2} + x_{(n/2)+1}) \quad (n\ even) \end{aligned} \tag{4}$$

$$\hat{\sigma} = 1.483\sqrt{\frac{n}{n-1}}\,\text{median}\left(\left|x_i - \tilde{x}\right|\right) \tag{5}$$

The DLR-F6 Model and Some Observed Flow Features

The challenge models for DPW II were the DLR-F6 wing-body and wing-body-pylon-nacelle combinations. The full model is shown in Fig. 2 mounted in the ONERA S2MA wind tunnel.[30] Although it was the intent of the workshop organizers to aim the challenge at the calculation of drag for a transport at cruise conditions and, hence, with essentially no separation regions, it was found in the wind tunnel and in the calculations that there were several significant regions of flow separation on airframe components. The key separation region for the wing-body combination was at the trailing edge of the wing root. This separation region was also likely present for the challenge problem of the first workshop and was detected in many of the calculations. A planview of the wind tunnel model after an oilflow run is shown in Fig. 3.[31] With the

addition of the pylon and nacelle, significant flow regions were observed on the lower surface of the wing and on the pylon as shown in Fig. 4.[31] These flow regions were generally believed by the participants of both workshops to be the main source of variation of the code results. Since neither challenge was designed to distinguish separation effects on code-to-code scatter from attached flow effects, additional study would be needed to determine if that belief is accurate.

Comparison of Results for Wing-Body Combination on Medium Grids

Code-to-Code Comparisons

The solution statistics for the two workshops for the wing-body combinations at the *cruise design point* on medium grids are given in Table 1. The first workshop had almost twice as many solutions for the cruise design point challenge condition as for the second workshop. However, many of those solutions were repeats on the same grids by the same codes for different users and turbulence models. Comparison of the statistics for numbers of individual authors (code users), institutions participating, and codes used shows that the two workshops were similar in diversity. Note that some of the solutions from both workshops analyzed in the paper were resubmitted after the workshops were completed

The control charts for total, pressure and skin-friction drag for both workshops are shown in Figs. 5-7. The scales of the control charts for both workshops are identical so that comparing the figures makes it easy to see that the code-to-code scatter is considerably smaller for DPW II than for DPW I, overall and for the "core" solutions where dispersion is proportional to the interval between the limits (red lines). Note that the pressure drag for solutions 6 and 18 is considerably greater than the other solutions. But the skin friction is considerably less. In the interest of process improvement for the collective as a whole, it would be very useful to determine the cause of the differences. The control charts for AOA and pitching-moment coefficient at the design lift coefficient are given in Figs. 8 and 9 respectively. The reduction in both the ranges and the AMRs is obvious.

Overall, it is clear that, for DPW II, the number of solutions that might be considered outliers is much smaller that for DPW I. Comparisons of the ranges and AMR's for the two workshops are given in Tables 2 and 3 respectively. Direct graphical comparisons are given in Figs. 10 and 11. The reason for the significant improvement in code-to-code scatter between workshops is not known.

Comparisons with Experiment

Table 4 compares the computed medians for the two workshops with the experimental results obtained in the ONEAR S2MA tunnel for the two challenge models. At first glance, it does not appear that there has been much improvement except for the pitching-moment coefficient for which the difference between the computational median and the experimental value has been halved. However, as pointed out in Ref. 6, the most appropriate way to compare the results of the workshops and the experiments is to use the "validation uncertainty" approach of Coleman and Stern[32]. The first step is to compute the difference between the computed and experimental result. For this paper, the median of the computed results will be used.

$$\Delta = \tilde{x}_{computed} - x_{experiment} \tag{6}$$

The next step is to estimate the standard uncertainty[25] for Δ by combining the standard uncertainties for $\tilde{x}_{computed}$ and $x_{experiment}$, i.e.

$$\hat{\sigma}_{\Delta} = \sqrt{\hat{\sigma}_{comp}^2 + \hat{\sigma}_{\exp}^2} \tag{7}$$

The value of $\hat{\sigma}_{comp}$ for the median is estimated using the standard deviation estimate of Eq. 3 divided by the square root of the number of solutions, i.e.

$$\hat{\sigma}_{comp} = \mathrm{AMR} / \left(1.128\sqrt{n}\right) \tag{8}$$

The uncertainty of the experimental data is unknown. Hence, it will be necessary to consider

$$\sigma_{\Delta} = \sigma_{comp} \tag{9}$$

to be a lower bound. The usual coverage for comparisons of this type is K=2.[b] Hence, the question of interest is this: Does the interval $\Delta \pm 2\hat{\sigma}_{\Delta}$ include zero? The results for both workshops are given in Table 5.

Examination of the results of Table 5 shows that the collective medians for AOA and CM are significantly offset outside the code-to-code scatter from the experimental results. It is unlikely that inclusion of the experimental uncertainty will increase those intervals sufficiently to include zero. Hence, as a collective, the workshop solutions are biased for AOA and CM.

[b] For independent random sampling from a static Normal distribution, K=2 would represent a confidence interval coverage of roughly 95%. However, since there is considerable correlation among the solutions, the sampling is neither truly independent nor truly random. Thus, the actual confidence interval coverage is unknown.

The story is different for the total drag. For both workshops, zero is included in the code-to-code scatter interval. The addition of experimental uncertainty can only increase the interval. Hence, according to the notion of validation uncertainty described by Coleman and Stern[32], the drag calculation for the collective as a whole is validated but only to the levels of uncertainty quoted in Table 5. Of course, validating CDTOT while not being able to validate the angle of attack and moment raises the important question – why?

Analysis of Grid Refinement Results for the Cruise-Point Condition

A major emphasis of the second workshop was an investigation of the effects of grid refinement for the cruise point condition. Coarse, medium and fine grids were developed for each of the three major grid types (structured block, structured overset and unstructured).[11] Numerical estimates of grid/solution quality were not made for the workshop. Hence, it is not possible to directly compare the refinement levels of the various grid types. However, an attempt was made to determine the effects of grid refinement for the collective as a whole by simply binning the solutions into "coarse", "medium" and "fine". (The gridding guidelines for the workshop (http://aaac.larc.nasa.gov/tsab/cfdlarc/aiaa-dpw/Workshop2/gridding.guidelines) attempted to establish similar grid resolutions between different grid types by specifying the approximate number of unknowns on each grid level, spacings on the surface and normal to the surface, and grid stretching.)

Ordinarily, grid convergence is considered only for an individual solution[33,34] so a new question must be posed for this analysis: What should grid convergence look like for the collective as a whole? The scatter in solutions is due to numerical error, modeling error (both physics models and computational models), user errors, and code errors. As the grid spacing is reduced, the numerical error should be reduced. Therefore, if grid convergence is demonstrated, the scatter due to numerical error can be reduced to any desired level removing it from consideration as a source of scatter.) Therefore, it seems reasonable to expect that the range (spread) of the solutions would decrease substantially.

Code-to-Code Comparisons

Although 21 solutions were computed by the participants for the cruise point condition on the medium grids, only 16 of those solutions were complemented by corresponding coarse and fine grid solutions for both the WB and WBNP configurations. Those solutions will be referred to in this paper as "nested". See Table 6 for the author/institution/code statistics. The scatter plots for the WB and WBNP solutions and the installed nacelle-pylon increments are given in Figs. 12-16. The three basic types of solvers (1) structured multiblock one-to-one, (2) structured overset, and (3) unstructured are delineated in Figs. 12-16 by red, blue and yellow symbols and lines respectively. (The legend identifies the grid type (MB,UN or OV), grids (ICEM, GRIDGEN, etc) and turbulence model (KW, SA, etc.) used for each contribution. See Nomenclature for description of each of these.) Also shown on the plots of Figs. 12, 15 and 16 are the experimental results from the ONERA S2MA wind tunnel[10,31] shown in green. Note that there is no discernable difference in code-to-code scatter with increasing grid refinement.

Medians. Also shown in Figs. 12-16 as a solid black line and in Table 7 are robust estimates of the population means obtained using the sample medians, for each bin, given by Eq. 4. (Since the sample median is not a linear operator, the increment medians can not be obtained by simple subtraction of the WBNP and WB results but rather must be obtained from the sets of increments themselves.) The largest effects of grid refinement are found in CDTOT, CDPR and AOA. Interestingly, CDSF shows almost no effect of grid convergence.

Standard Deviations. Robust estimates of the population standard deviations for each bin were obtained using the MAD given by Eq. 5. There was no significant difference between the estimates for coarse, medium and fine solutions for each parameter of interest so the estimates were averaged for the three grid densities. The results are given in Table 8. Note that the code-to-code scatter tends to be higher for the more complicated configuration and lower for the increments (except for CDTOT) as shown also in Figure 17. The standard deviation results were multiplied by two[c] and added to and subtracted from the medians to obtain the upper and lower limits respectively that are shown in Figs. 12-16 as dashed black lines. These limits serve as a rough estimate of the code-to-code scatter and help to suggest outliers.

Outliers. Considering first the total drag (CDTOT), it is seen that only one of the solutions is outside the limits for the WB and WBNP results. The same solution is outside the increment limits only for the

[c] See previous footnote.

fine grid case. However, another solution which is inside the limits for the WB and WBNP results lies outside the increment limits for all of the grid densities. These results suggest that the usual generalization that codes can do a better job computing drag increments compared to computing absolute values may be incorrect or a least overly simplified.

For pressure drag (CDPR), one code is above the limits for the WB and WBNP results but within the scatter limits for the increments. For skin friction drag (CDSF), two codes are below the limits but one of them falls within the limits for the increments while another code that is within the limits on WB and WBNP falls outside the limits on increments.

For angle-of-attack (AOA) and moment (CM), the solutions tend to underpredict the experimental data for the WB and WBNP, but are largely clustered about the data for the increments. As seen in the validation study, the limits don't contain the experiment for AOA and CM for the WB configuration, but do for the WBNP and for the increments.

Comparison of the Collective Medians with Experimental Data

The experimental data from the ONERA S2MA wind-tunnel test[10,31] are given in Table 9. The validation methodology described above for comparison of WB results on the medium grids has also been applied to the WBNP and increment results as shown in Table 10. Recall that it is assumed for the purposes of this paper that the experimental data have no uncertainty. Similar results are obtained for the WBNP, i.e., CDTOT is validated to the experimental value while AOA and CM are not.

It is particularly interesting to note that the CDTOT results for the collective as a whole compare favorably for the absolute WB and WBNP configurations, but not for the increment. It is usually believed that drag increments can be computed with more accuracy than the absolute values. These results suggest the opposite. It should be noted, however, that the comparisons apply only to the medians (and their scatter) of the collective as a whole. The performance of individual codes cannot be obtained with the present kind of analysis. It is also interesting to note that, unlike CDTOT, while the absolute median values for AOA and CM do not agree with experiment as expected, they do agree for the median increments.

Surface Pressures

Figs. 19-26 show the experimental and calculated pressures computed for the surface of the wing at eight spanwise locations as shown in Fig. 18[15], 2y/B = 0.15, 0.239, 0.331, 0.377, 0.411, 0.514, 0.638, and 0.847, for both the WB and the WBNP at the Case 1 fixed lift cruise condition (CL = 0.5, M = 0.75). Seven participants contributed fine grid results for both configurations; two solutions using the structured multiblock one-to-one grids, three solutions using the nodal unstructured grids, one solution using the cell-centered unstructured grids with wall functions, and one solution using the structured overset grids. Four of the solutions used the SA turbulence model, two used the SA1A model, and one solution used the $k\varpi$ model. The indexing of the solutions for the pressures has been changed from that of the forces and moments.

The variation of the surface pressures will be shown for three regions: the wing-body juncture, the pylon/nacelle, and the wing.

Wing-Body Juncture

Fig. 19 shows the surface pressures on the 2y/B = 0.15 wing section. This section is close to the fuselage. The seven calculations have small variation on the lower surface of the wing up to about 70% of the local chord and are in close agreement with the experimental pressures. The WBNP configuration has larger variation in the suction peak and upper surface pressures than the WB. The calculations agree reasonably well with the experiment over approximately the first 50% of the upper surface. Both the WB and the WBNP have much larger variation on the aft portion of the wing on both upper and lower surfaces. This region of the wing-body juncture has demonstrated separated flow in both the experiment and calculations.

Pylon/Nacelle

Fig. 20 shows the surface pressures at 2y/B = 0.239. This spanwise location is between the pylon-nacelle and the fuselage for the WBNP configuration. Fig. 20(a) shows very small variation in the calculated pressures over almost the entire lower surface and good agreement with the experimental pressures. Variation in pressures on the upper surface is somewhat larger and the agreement with experiment is somewhat worse, especially between about 20% and 30% of local chord where the experiment appears to show a double shock structure that is smeared out in the calculation. Fig. 20(b) shows substantially larger variation in the pressures on both the lower and upper

surfaces for the WBNP configuration than was present in the WB. The increase in variation is quite possibly an effect of the pylon/nacelle.

Fig. 21 shows the surface pressure at 2y/B = 0.331 which lies just inboard of the pylon/nacelle for the WBNP. Fig. 21(a) shows very small variation in the calculated pressures for the lower surface of the WB and shows good agreement with the experimental data. The upper surface shows much larger variation than the lower surface; the shock location varies by approximately 2-3% of local chord. The calculations under-predict the rooftop pressure, predict the shock too far forward, and miss the pressure recovery downstream of the shock. Fig. 21(b) shows slightly larger variation on the upper surface for the WBNP configuration than the WB. The rooftop pressure is again under-predicted, the shock is too far forward, and the pressure recovery is miss-predicted. The lower surface has extremely large variation. All seven solutions agree with each other for about the first 10% of the local chord. Only two of the solutions agree well with the experimental pressures on the lower surface. A separation region was observed in the experiment on the inboard side of the wing/pylon. It appears that differences in predicting this separation region are contributing to huge variations in the predicted pressures on the lower surface inboard of the pylon.

Figs. 22 and 23 show larger variation for the WBNP configuration than the WB configuration on both the upper and lower surfaces. This difference in variation is due to the presence of the pylon/nacelle. The effect of the pylon/nacelle is evident from 2y/B = 0.239 out to 2y/B = 0.411; it is possible that there may be pylon/nacelle effects on the variation at 2y/B = 0.15 also.

Wing

Figs. 21(a), 22(a), 23(a), 24(a), and 25(a) show the surface pressures on the WB configuration at 2y/B = 0.377, 0.411, 0.514, 0.638, and 0.847 respectively. Figs. 24(b), 25(b), and 26(b) show the surface pressures on the WBNP configuration at 2y/B = 0.514, 0.638, and 0.847 respectively. The computed pressures on the WB for these spanwise locations are very similar to the 2y/B = 0.331 location shown in Fig. 20(a). The variation in the lower surface pressures for all of these locations is very similar. The variation in the upper surface pressure is larger than the lower surface. The upper surface shock location varies from about 2-3% of the local chord at 2y/B = 0.331 to about 6-7% of local chord at 2y/B = 0.847. The mean shock location is about 5% of the local chord too far upstream at 2y/B = 0.331 and grows to about 10% of the local chord near the wing tip.

The shock location and variation in shock location is due in part to the constant lift condition of the workshop Case 1. The calculation with the most forward position of the shock is UN-LARC-SA 3 with the calculated angle of attack of -0.128 degrees. This is the lowest angle of attack of the solutions. The shock moves aft with increasing angle of attack. The order from front to back is OV-SA1A at alpha = -0.0263 degrees, MB-ICEM-KW at alpha = 0.075 degrees, MB-ICEM-SA1A at alpha = 0.108 degrees, UN-LARCWF-SA at alpha = 0.248 degrees, UN-LARC-SA 2 at alpha = 0.263 degrees and UN-LARC-SA 1 at alpha = 0.275 degrees.

The agreement with experiment on the lower surface is generally quite good with the exception of the region around 2-10% of the local chord. All of the calculations miss the pressure in this region and show larger variation. The variation at the mid-chord on the lower surface is very small. This area appears to be insensitive to angle of attack, code, turbulence model, and grid. The leading edge and trailing edge on the lower surface show larger variations.

The agreement with experimental data on the upper surface is not as good. The suction peak is missed, the shock location is consistently predicted too far forward and therefore the shock strength is missed and the pressure is poorly predicted downstream of the shock. The pressure on the entire upper surface shows substantially more variation than on the lower surface.

Final Remarks

The code-to-code scatter was significantly reduced for the wing-body configuration on medium grids compared to the first workshop, although it is still an order-of-magnitude larger than desired by airframe designers. However, the grid refinement studies showed no significant change in code-to-code scatter with increasing grid density due apparently to code-to-code bias that was not reduced with grid refinement.

In addition to the twin workshop surprises for DPW I and II of (1) larger-than-expected code-to-code scatter and (2) apparent lack of reduction in code-to-code scatter with increasing grid refinement, the present authors believe that there is another, perhaps even more important, outcome --- that the two workshops are a rich source of information with which to tackle improving all of the codes. But it is detailed analysis of across-code discrepancies that will point to possible improvements. Furthermore, it

seems likely that uncertainty quantification for real industrial computation problems could fruitfully begin with the running of more than one code for a given problem.

Acknowledgements

The authors are grateful to the workshop organizers for suggesting this effort and to the participants and Dr. James M. Luckring for many helpful suggestions.

References

1. Rakowitz, M., Sutcliffe, M., Eisfeld, B., Schwamborn, D., Bleeke, H., Fassbender, J., "Structured and Unstructured Computations on the DLR-F4 Wing-Body Configuration", AIAA-2002-0837, January 2002.
2. Mavriplis, D., Levy, D., "Transonic Drag Prediction Using an Unstructured Multigrid Solver", AIAA-2002-0838, January 2002.
3. Pirzadeh, S., Frink, N., "Assessment of the Unstructured Grid Software TetrUSS for Drag Prediction of the DLR-F4 Configuration", AIAA-2002-0839, January 2002.
4. Vassberg, J., Buning, P., Rumsey, C., "Drag Prediction for the DLR-F4 Wing/Body using OVERFLOW and CFL3D on an Overset Mesh", AIAA-2002-0840, January 2002.
5. Levy, D. W., Zickur, T., Vassberg, J., Agrawal, S., Wahls, R. A., Pirzadeh, S., Hemsch, M. J., "Data Summary from First AIAA Computational Fluid dynamics Drag Prediction Workshop", J. Aircraft, Vol. 40, No. 5, Sep-Oct 2003, p. 875-882. (Also AIAA-2002-0841.)
6. Hemsch, M. J., "Statistical Analysis of CFD Solutions from the Drag Prediction Workshop", AIAA-2002-0842, January 2002 (Accepted for publication in the Journal of Aircraft).
7. Redeker, G., "DLR-F4 Wing Body Configuration", in Chapter B of *A Selection of Experimental Test Cases for the Validation of CFD Codes*, AGARD-AR-303 Vol. II, August 1994.
8. Redeker, G., Muller, R., Ashill, P. R., Elsenaar, A., and Schmitt, V., "Experiments on the DLR-F4 Wing Body Configuration in Several European Wind Tunnels", in Chapter 2 of *Aerodynamic Data Accuracy and Quality: Requirements and Capabilities in Wind Tunnel Testing*, AGARD-CP-429, July 1988.
9. Elsholz, E., "The DLR-F4 Wing/Body Configuration", in *ECARP - European Computational Aerodynamics Research Project: Validation of turbulence Models*, Notes on Numerical Fluid Mechanics, Vol. 58, 1997, p.429-450.
10. Broderson, O., and Sturmer, A., "Drag Prediction of Engine-Airframe Interference Effects Using Unstructured Navier-Stokes Calculations", AIAA-2001-2414, June 2001.
11. Laflin, K., et al, "Summary of Data from the Second AIAA CFD Drag Prediction Workshop", AIAA-2004-0555, January 2004.
12. Broderson, O., Rakowitz, M., Amant, S., Destarac, D., Larrieu, P., Sutcliffe, M., "Drag Prediction for the DLR-F6 Configuration: Airbus, ONERA, and DLR Results from the 2nd AIAA Drag Prediction Workshop", AIAA-2004-0391, January 2004.
13. Langtry, R., Kuntz, M., Menter, F., "Drag Prediction of Engine-Airframe Interference Effects with CFX 5", AIAA-2004-0392, January 2004.
14. Sclafani, A., DeHaan, M., Vassberg, J., "OVERFLO Drag Prediction for the DLR-F6 Transport Configuration: A DPW II Case Study", AIAA-2004-0393, January 2004.
15. Rumsey, C., Rivers, S., Morrison, J., "Study of CFD Variation on Transport Configurations from the Second Drag Prediction Workshop", AIAA-2004-0394, January 2004.
16. Wurtzler, K., Morton, S., "Aircraft Drag Prediction Using Cobalt", AIAA-2004-0395, January 2004.
17. May, G., van der Weide, E., Jameson, A., Shankaran, S., Martinelli, L., "Drag Prediction of the DLR-F6 Configuration", AIAA-2004-0396, January 2004.
18. Kim, Y., Park, S., Kwon, J., "Drag Prediction of DLR-F6 Using the Turbulent Navier-Stokes Calculations with Multigrid", AIAA-2004-0397, January 2004.
19. Yamamoto, K., Ochi, A., Shima, E., Takaki, R., "CFD Sensitivity to Drag Prediction on DLR-F6 Configuration by Structured Method and Unstructured Method", AIAA-2004-0398, January 2004.
20. Tinoco, E., "Drag Prediction with the Zeus/CFL3D System", AIAA-2004-0552, January 2004.
21. Klausmeyer, S., "Drag, Lift, and Moment Estimates for Transonic Aircraft Using the Navier-Stokes Equations", AIAA-2004-0553, January 2004.
22. Lee-Rausch, E., Mavriplis, D., Frink, N., "Transonic Drag Prediction Using Unstructured Grid Solvers", AIAA-2004-0554, January 2004.
23. Moffat, R. J., "Contributions to the Theory of Single-Sample Uncertainty Analysis", Journal of Fluids Engineering, Vol. 104, June 1982, p. 250-260.
24. Hatton, Les, "The T Experiments: Errors in Scientific Software", IEEE Computational Science and Engineering, Vol. 4, No. 2, April-June 1997, p. 27-38.
25. Anon., *U.S. Guide to the Expression of Uncertainty in Measurement*, ANSI/NCSL Z540-2-1997, October 1997.
26. Youden, W. J., "Enduring Values", Technometrics, Vol. 14, No.1, February 1972, p. 1-11.
27. von Mises, Richard, *Probability, Statistics and Truth*, Dover 1981.
28. Montgomery, D. C., *Introduction to Statistical Quality Control, 3rd Edition*, Wiley 1996.
29. Wheeler, D., *Understanding Statistical Process Control, 2nd Ed.*, SPC Press, 1992.
30. LaVerre, J., Charpin, F., "The S2MA Wind-Tunnel at the Modane-Avrieux Aerodynamic Test Center", ESATT 862 1983 (ONERA NT 1983-5).
31. Godard, J. L., "F6 Model Tests in the ONERA S2MA Wind Tunnel", presented at the 2nd AIAA CFD Drag Prediction Workshop, Orlando, FL, June 2003, http://aaac.larc.nasa.gov/tsab/cfdlarc/aiaa-

dpw/Workshop2/pdf/14Godard__F6_expe_red.pdf, Accessed November 23, 2003.

32. Coleman, H. W., and Stern, F., "Uncertainties and CFD Code Validation", J. Fluids Engineering, Vol. 119, December 1997, p. 795-803.
33. Roache, P., *Verification and Validation in Computational Science and Engineering*, Hermosa, 1998.
34. Ferziger, J., Peric, M., *Computational Methods for Fluid Dynamics*, 2nd Ed., Springer, 1999.

	DPW I	DPW II
No. of Solutions	38	21
No. of Authors	18	20
No. of Institutions	14	16
No. of Codes	13	18

Table 1. Comparison of workshop statistics for wing-body combinations on medium grids.

	DPW I	DPW II
CDTOT	274 counts	48 counts
CDPR	205 counts	70 counts
CDSF	208 counts	52 counts
AOA, degrees	2.22	0.56
CM	0.28	0.05

Table 2. Comparison of ranges for wing-body combinations on medium grids.

	DPW I	DPW II
CDTOT	39 counts	8 counts
CDPR	19 counts	15 counts
CDSF	26 counts	15 counts
AOA, degrees	0.29	0.15
CM	0.024	0.013

Table 3. Comparison of average moving ranges for wing-body combinations on medium grids.

	DPW I		DPW II	
	Computation	**Experiment**	**Computation**	**Experiment**
CDTOT, counts	294	289	292	295
AOA, degrees	-0.26	0.19	0.18	0.52
CM	-0.161	-0.126	-0.138	-0.121

Table 4. Comparison of computed medians for wing-body combinations on medium grids with experimental results from ONERA S2MA tunnel.

$\Delta \pm 2\hat{\sigma}$	DPW I	DPW II
CDTOT, counts	5 ± 11	-3 ± 4
AOA, degrees	-0.45 ± 0.08	-0.34 ± 0.07
CM	-0.035 ± 0.007	-0.017 ± 0.004

Table 5. Comparison of computation-experiment deltas for wing-body combinations on medium grids with confidence intervals for the collective medians.

No. of authors	16
No. of Institutions	14
No. of Codes	15
No. of Solver Types	3
No. of Turbulence Models and Variants	7
No. of Grids and Variants	7

Table 6. Statistics for nested solutions for DPW II.

	WB			WBNP			INCREMENT		
	Coarse	**Med.**	**Fine**	**Coarse**	**Med.**	**Fine**	**Coarse**	**Med.**	**Fine**
CDTOT, counts	299	292	290	354	343	338	53	54	50
CDPR, counts	164	161	159	194	186	181	27	27	22
CDSF, counts	131	130	130	157	158	155	25	25	24
AOA, degrees	0.12	0.18	0.16	0.64	0.66	0.72	0.56	0.52	0.51
CM	-0.138	-0.138	-0.140	-0.135	-0.133	-0.134	0.003	0.001	0.005

Table 7. Sample medians for nested cruise-point solutions.

	Wing-Body	**Wing-Body-Nacelle-Pylon**	**Increment**
CDTOT, counts	7.3	11.4	8.0
CDPR, counts	8.6	13.4	7.1
CDSF, counts	6.3	7.7	2.9
AOA, degrees	0.14	0.19	0.10
CM	0.008	0.014	0.008

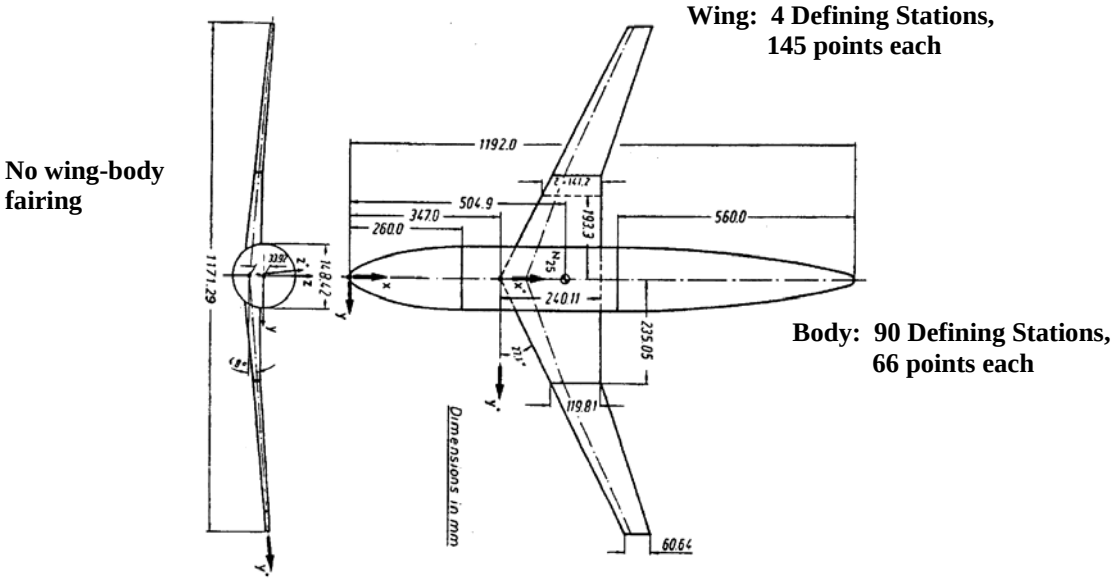
Table 8. Estimated code-to-code population standard deviations for nested cruise-point solutions.

	WB	**WBNP**	**INCREMENT**
CDTOT, counts	295	338	43
AOA, degrees	0.52	1.00	0.48
CM	-0.121	-0.120	0.001

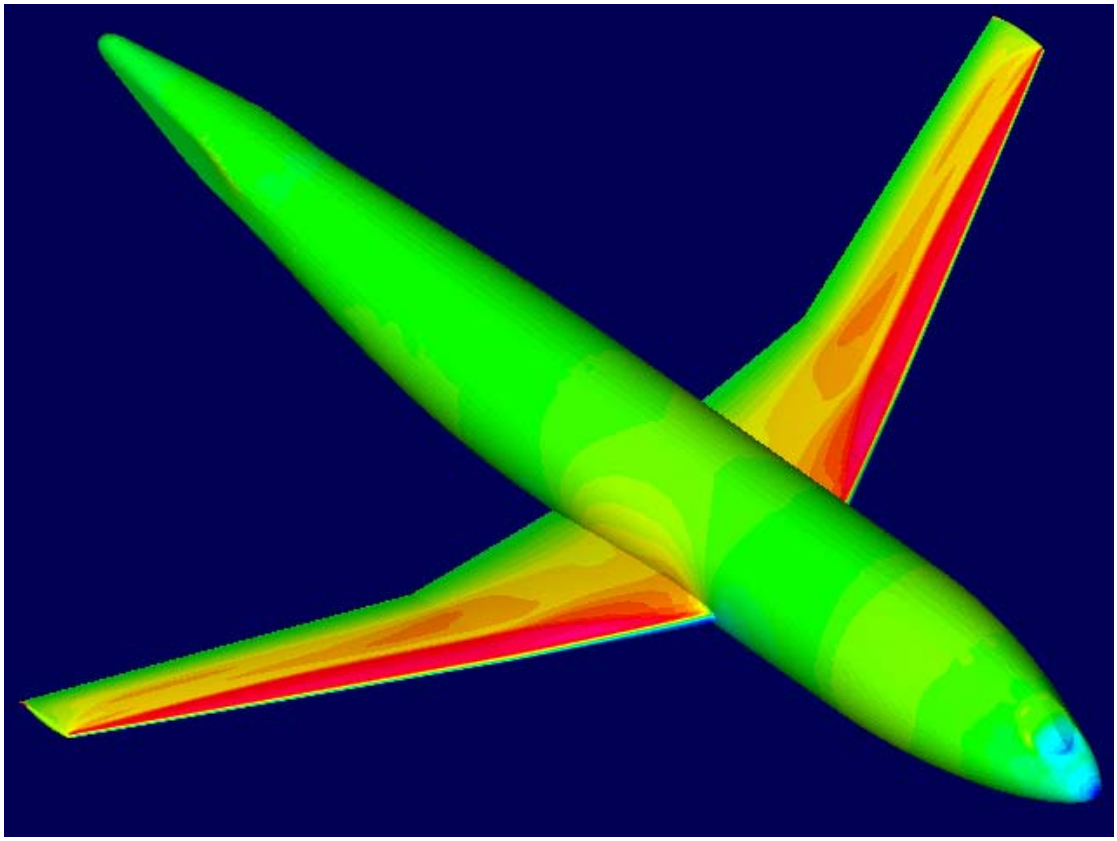
Table 9. Experimental results for cruise-point condition.

$\Delta \pm 2\hat{\sigma}_{\bar{x}}$	**WB**	**WBNP**	**INCREMENT**
CDTOT, counts	-3 ± 4	5 ± 6	11 ± 4
AOA, degrees	-0.34 ± 0.07	-0.34 ± 0.10	0.04 ± 0.05
CM	-0.017 ± 0.004	-0.013 ± 0.007	0 ± 0.004

Table 10. Comparison of computation-experiment deltas for nested solutions on medium grids with confidence intervals for the collective medians.



(a) Geometry sketch



(b) Perspective with example flow solution

Figure 1. DLR-F4 challenge model for DPW I.

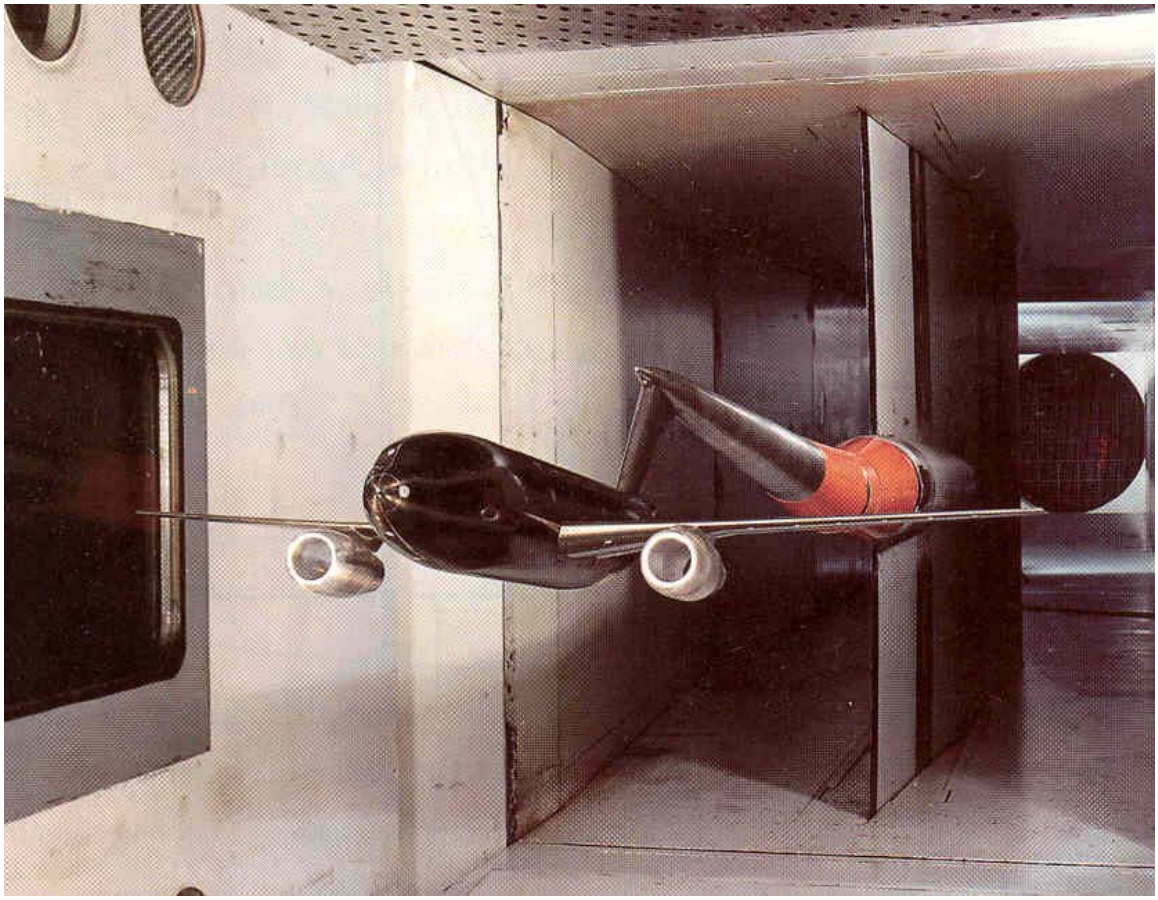


Figure 2. DLR-F6 model in ONERA S2MA tunnel.

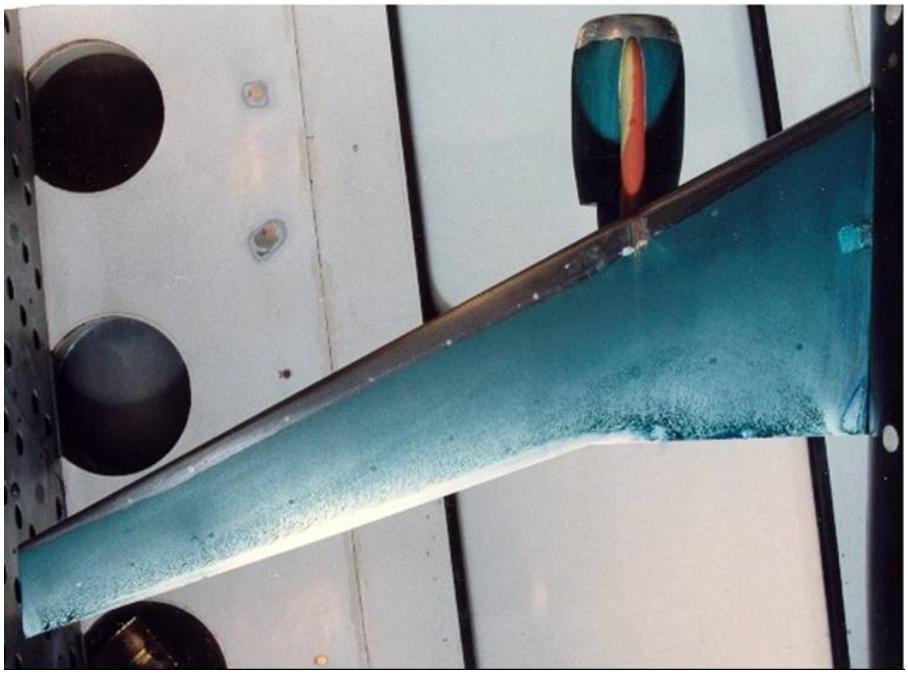


Figure 3. Upper planview of DLR-F6 model with oil flow patterns showing wing-root trailing-edge flow separation.

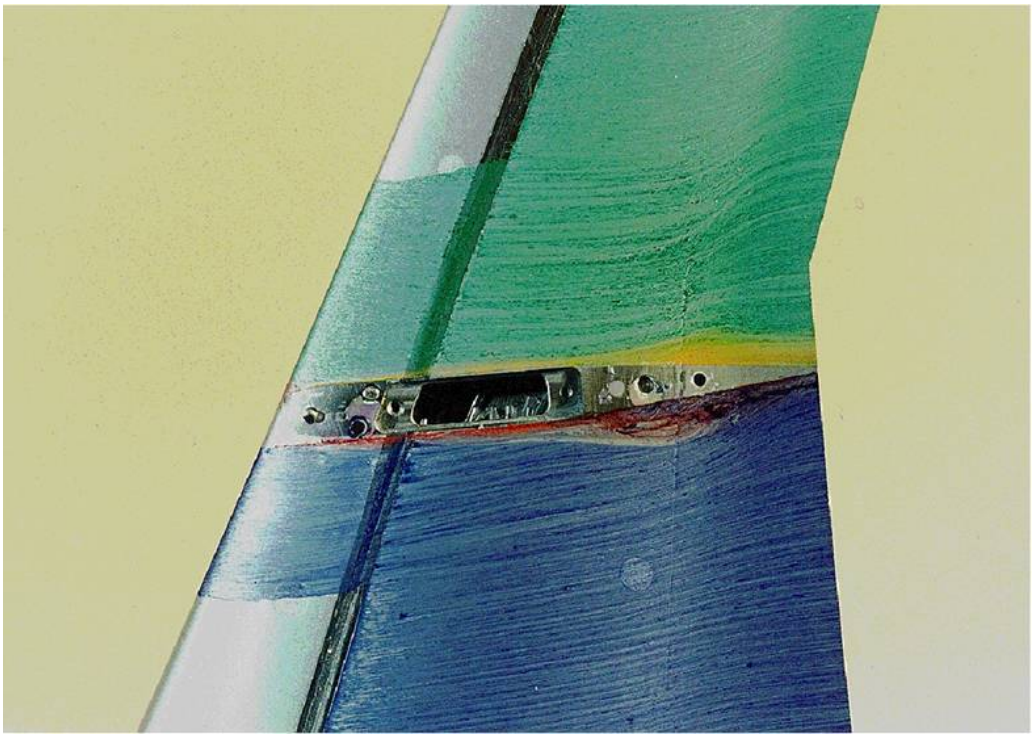
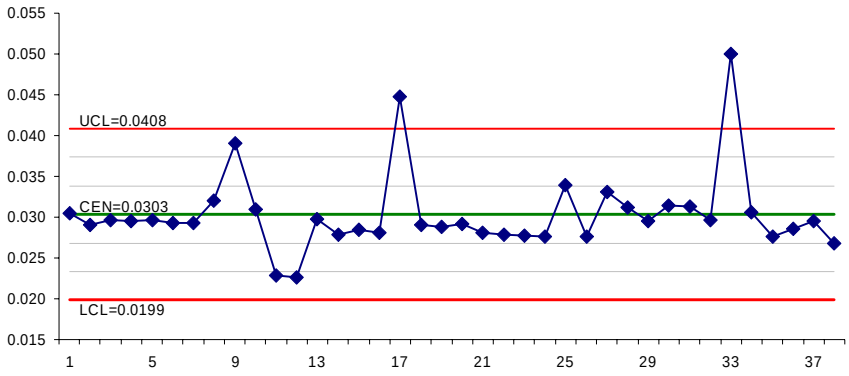
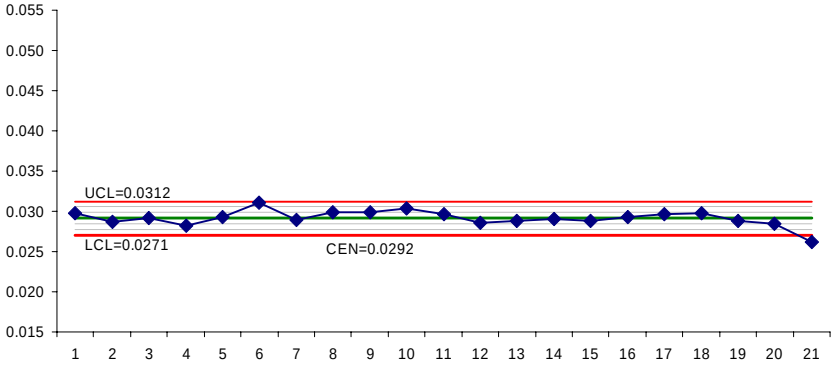


Figure 4. Lower planview of DLR-F6 model with oil flow patterns showing wing lower-surface flow separation near pylon.

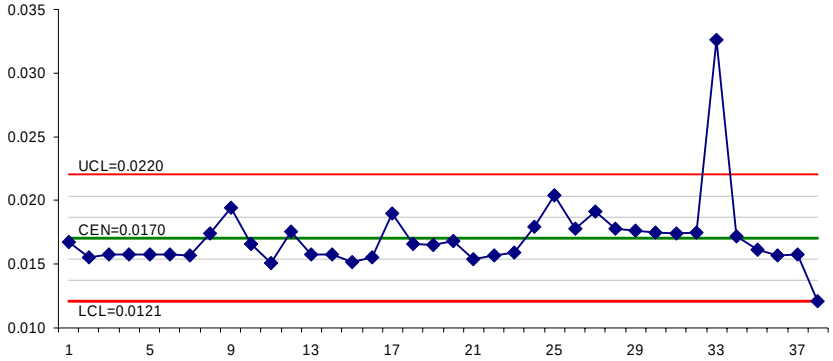


(a) DPW I

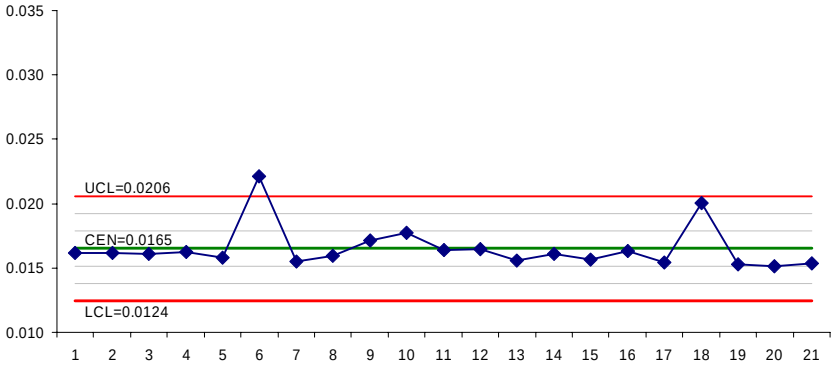


(b) DPW II

Figure 5. Comparison of individuals control charts for total drag coefficient at the cruise point design condition. The solutions are randomly ordered on the abscissa.

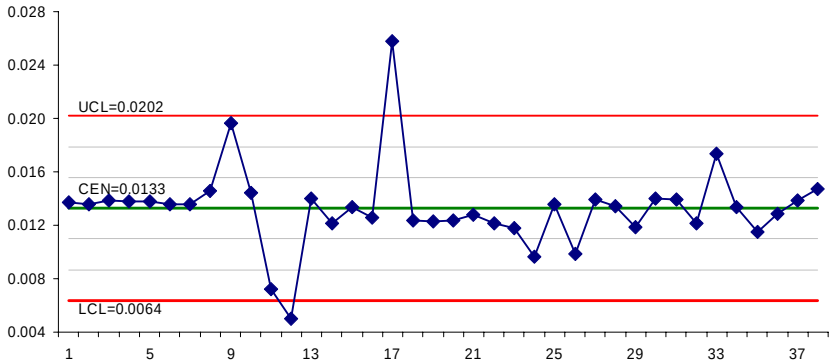


(a) DPW I

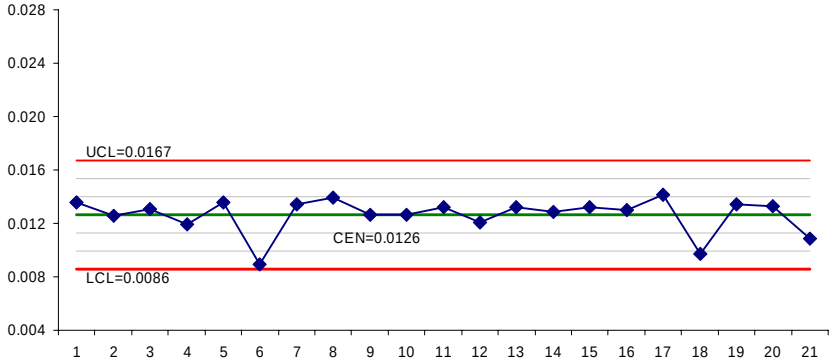


(b) DPW II

Figure 6. Comparison of individuals control charts for pressure drag at the cruise point design condition.

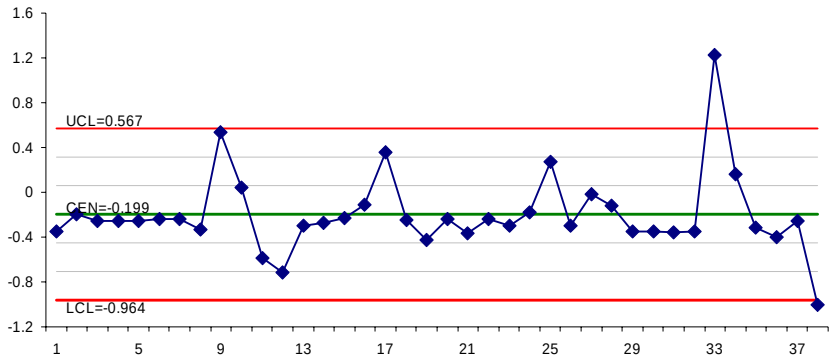


(a) DPW I

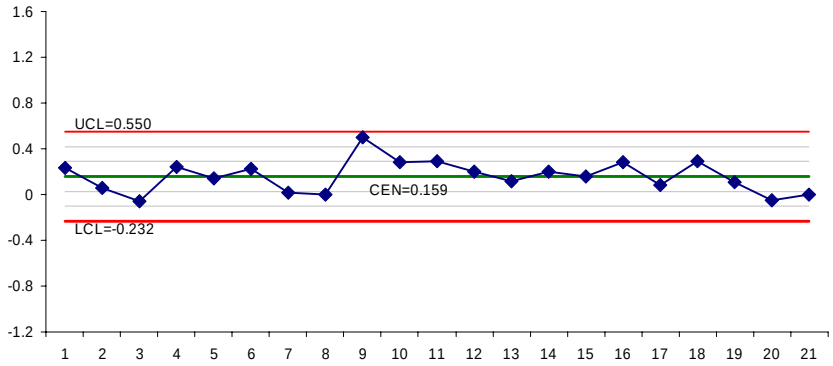


(b) DPW II

Figure 7. Comparison of individuals control charts for skin-friction drag at the cruise point design condition.

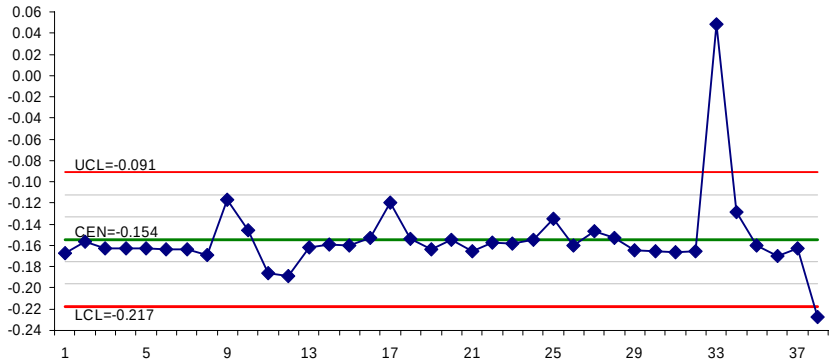


(a) DPW I

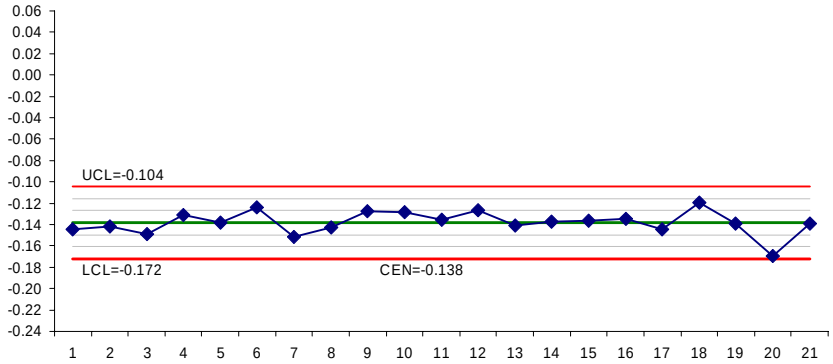


(b) DPW II

Figure 8. Comparison of individuals control charts for angle of attack at the cruise point design condition.



(a) DPW I



(b) DPW II

Figure 9. Comparison of individuals control charts for CM at the cruise point design condition.

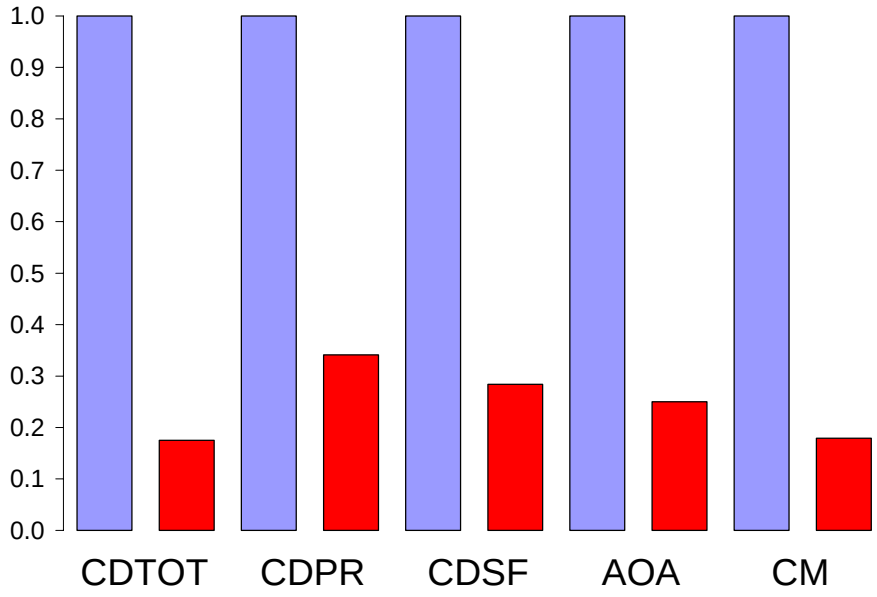


Figure 10. Comparison of ranges for wing-body combinations on medium grids. Blue is DPW I normalized to one, red is DPW II normalized to DPW I.

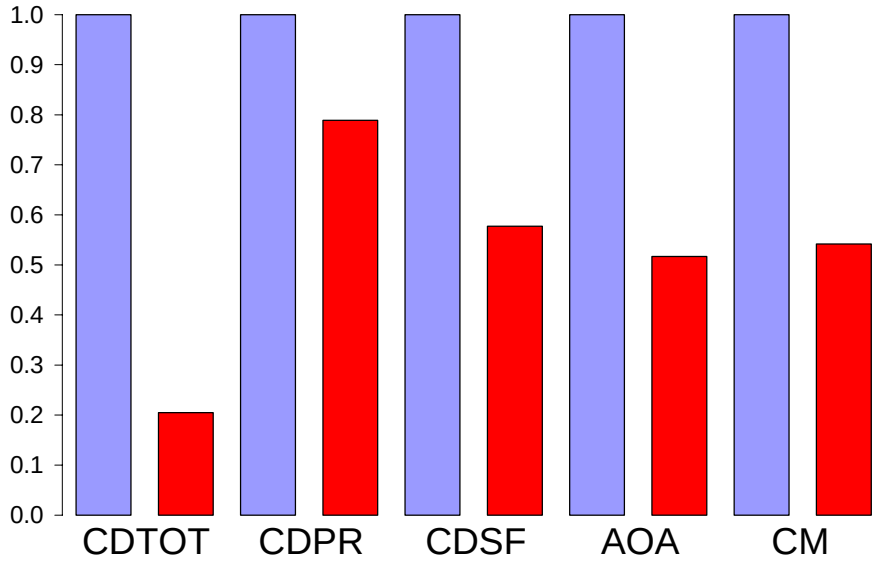
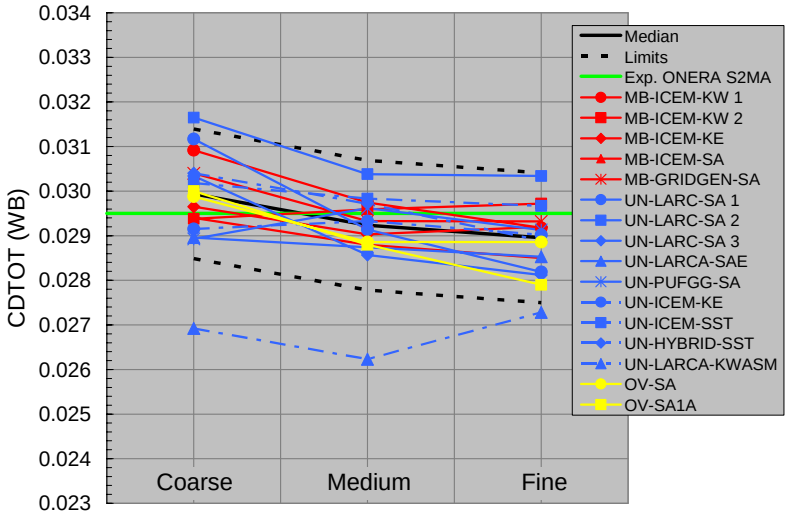
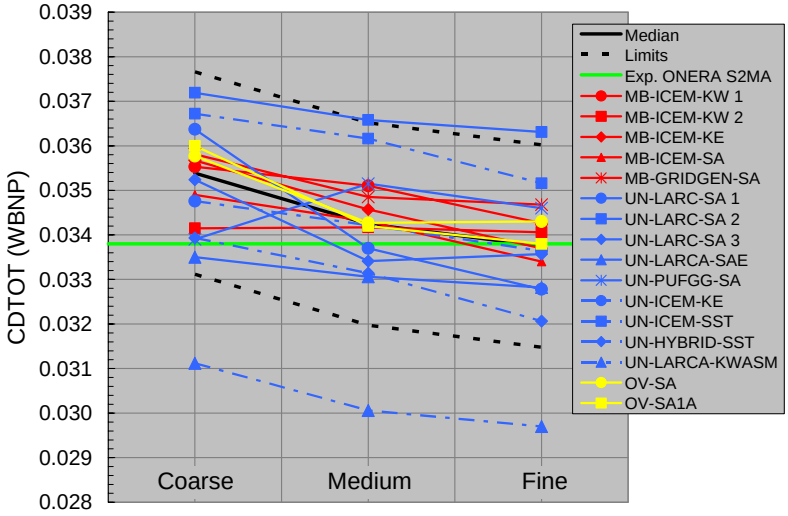


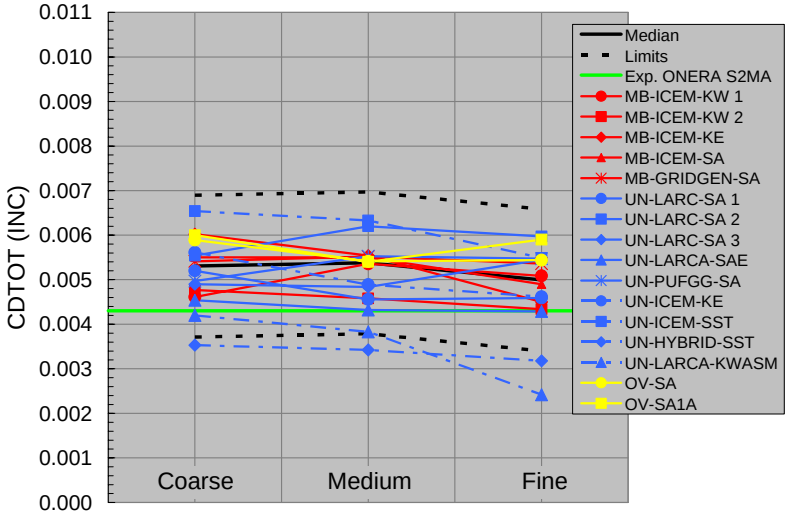
Figure 11. Comparison of average moving ranges for wing-body combinations on medium grids. Blue is DPW I normalized to one, red is DPW II normalized to DPW I.



(a) DLR-F6 wing-body configuration

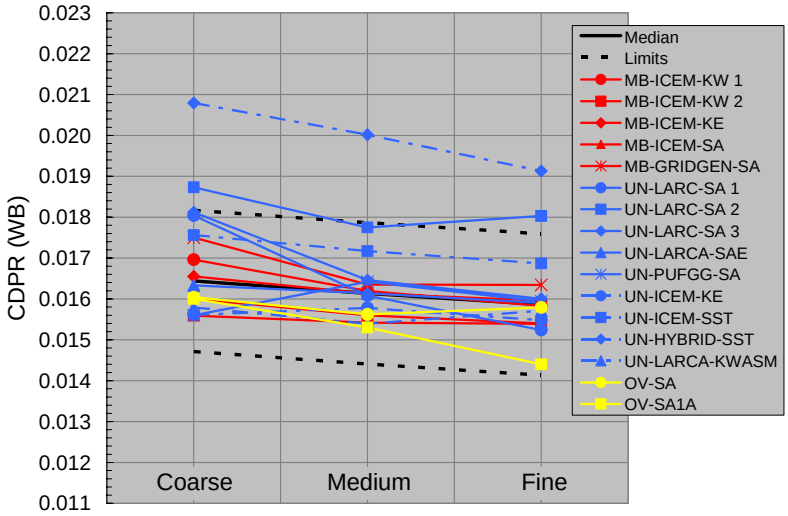


(b) DLR-F6 wing-body-nacelle-pylon configuration

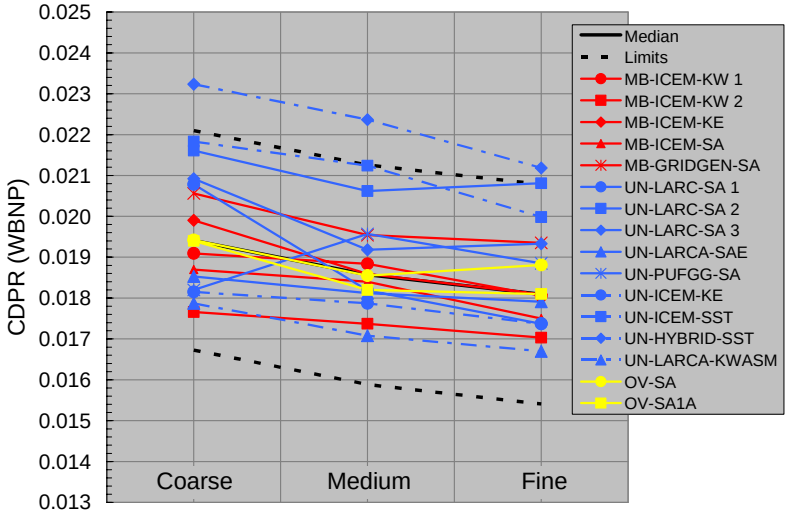


(c) DLR-F6 installed-nacelle-pylon increment

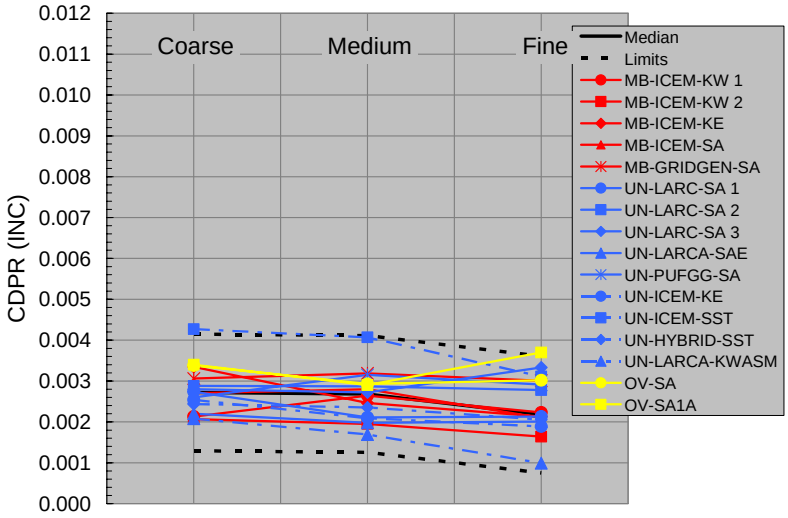
Figure 12. Code-to-code scatter for CDTOT at cruise-point lift condition.



(a) DLR-F6 wing-body configuration

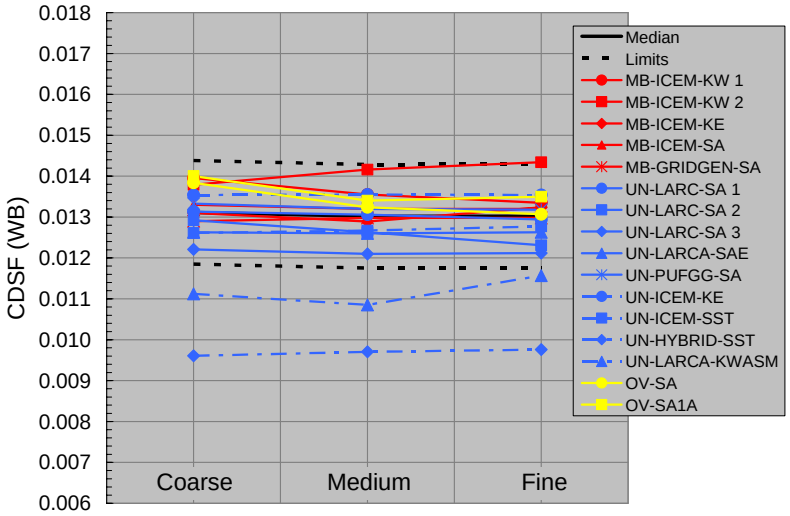


(b) DLR-F6 wing-body-nacelle-pylon configuration

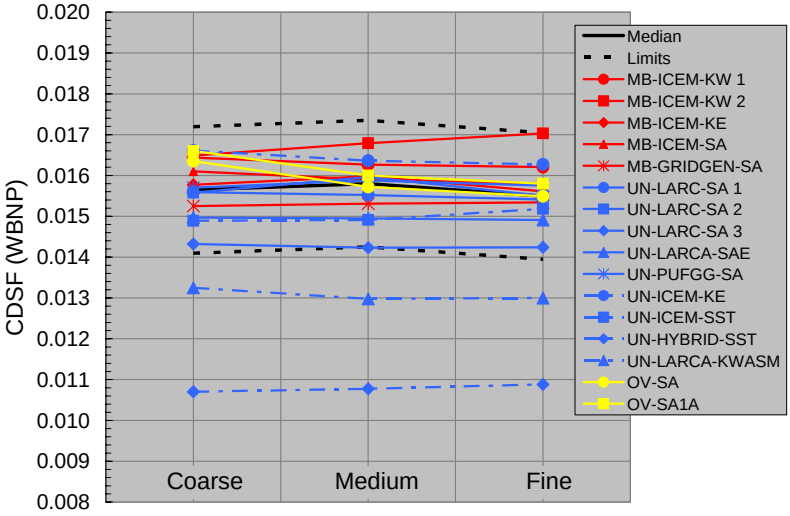


(c) DLR-F6 installed-nacelle-pylon increment

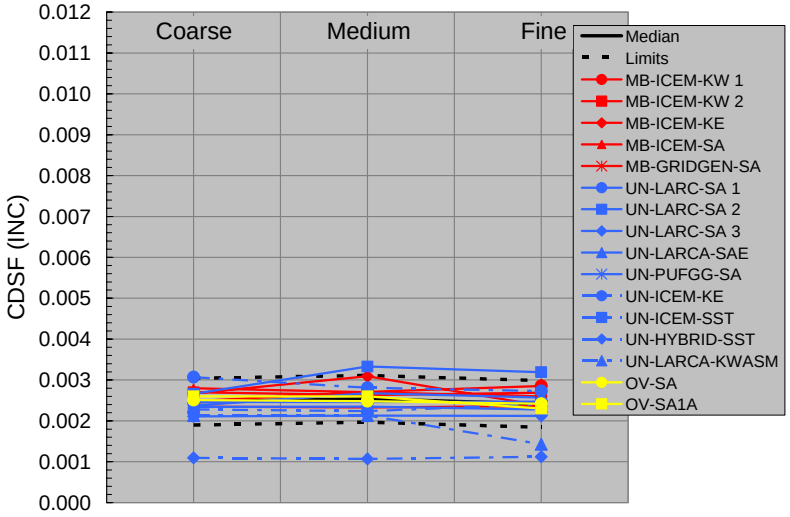
Figure 13. Code-to-code scatter for CDPR at cruise-point lift condition.



(a) DLR-F6 wing-body configuration

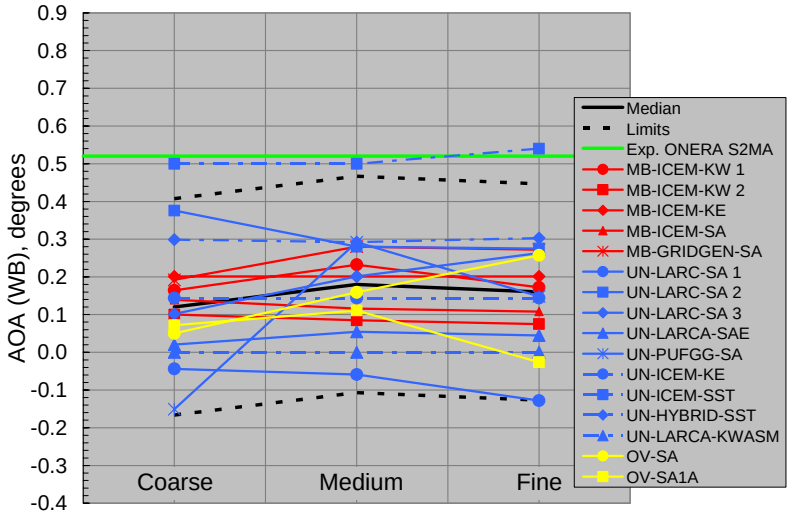


(b) DLR-F6 wing-body-nacelle-pylon configuration

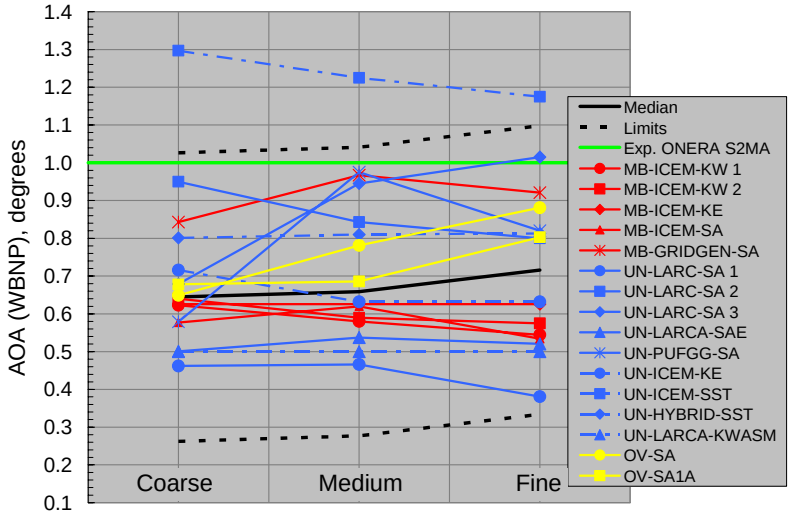


(c) DLR-F6 installed-nacelle-pylon increment

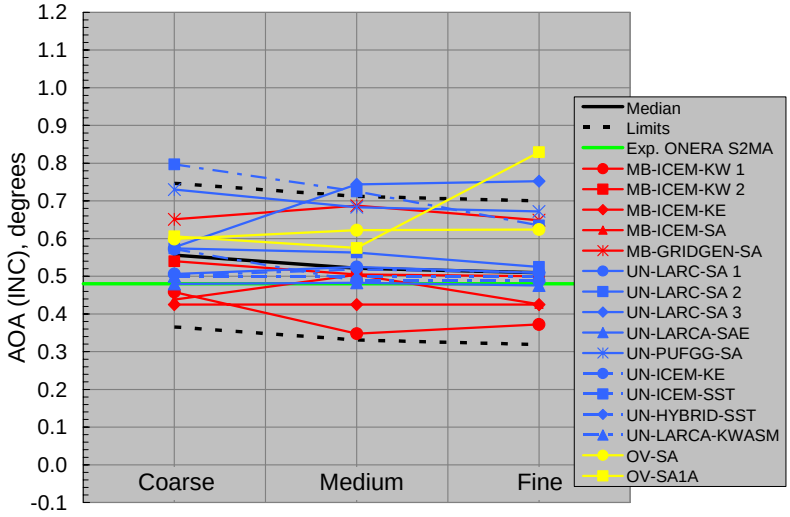
Figure 14. Code-to-code scatter for CDSF at cruise-point lift condition.



(a) DLR-F6 wing-body configuration

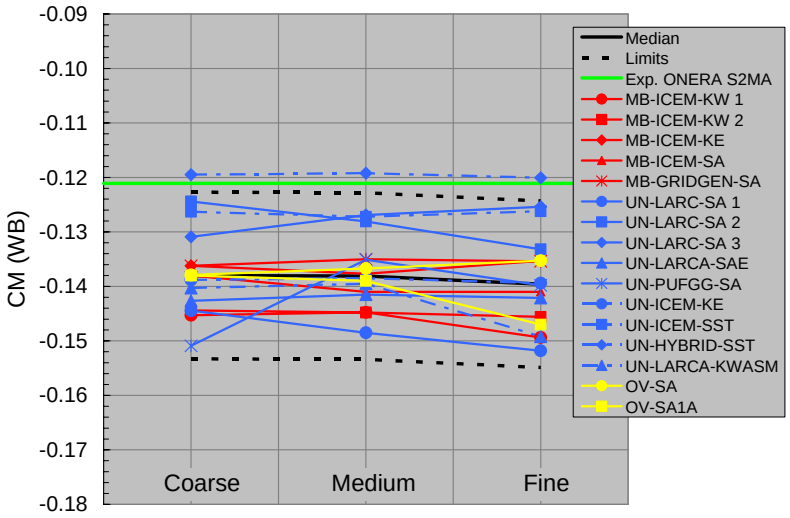


(b) DLR-F6 wing-body-nacelle-pylon configuration

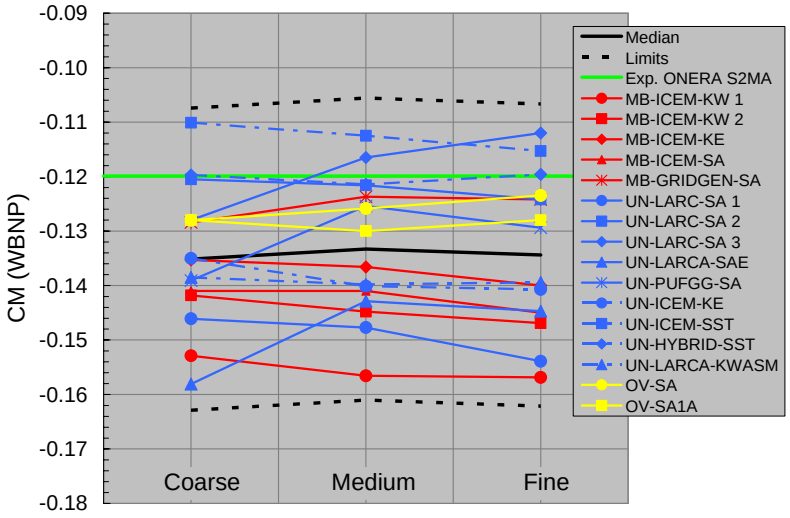


(c) DLR-F6 installed-nacelle-pylon increment

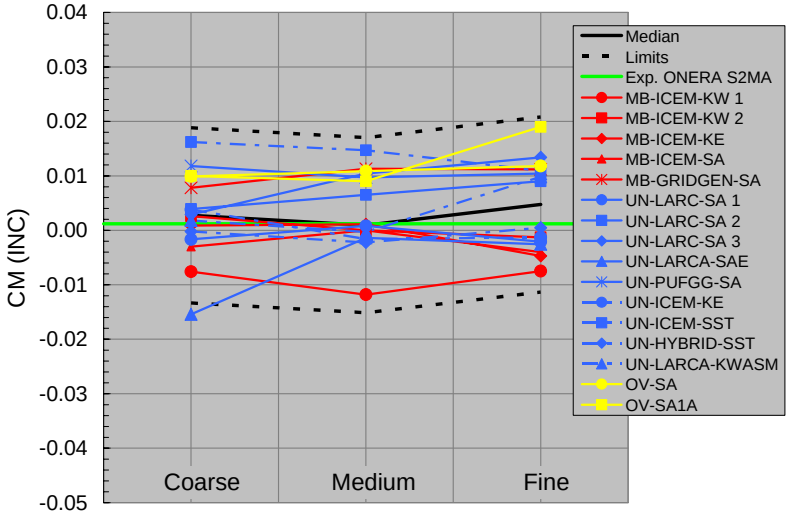
Figure 15. Code-to-code scatter for AOA at cruise-point lift condition.



(a) DLR-F6 wing-body configuration



(b) DLR-F6 wing-body-nacelle-pylon configuration



(c) DLR-F6 installed-nacelle-pylon increment

Figure 16. Code-to-code scatter for CM at cruise-point lift condition.

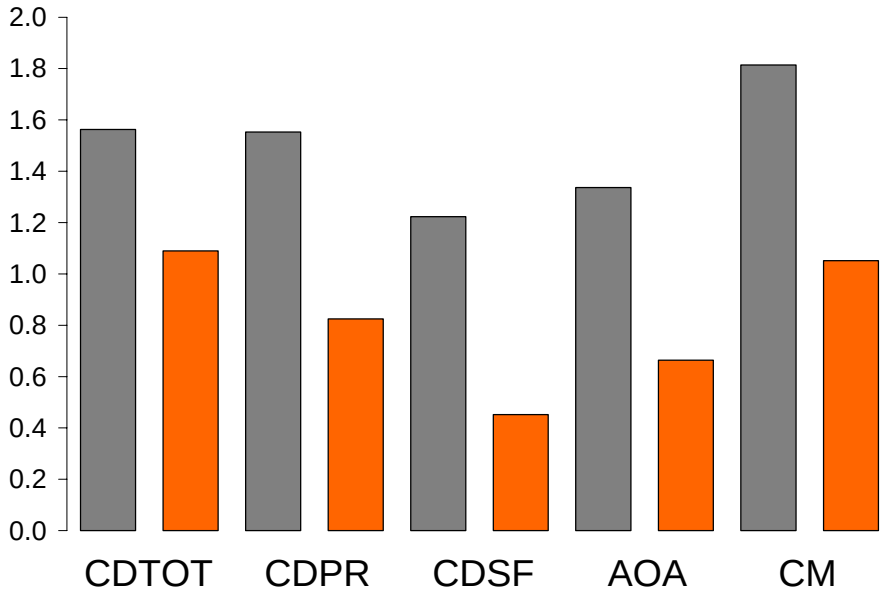


Figure 17. Comparison of code-to-code scatter standard deviations relative to the wing-body combination. Grey and orange are the ratios of the WBNP and Increment values to the WB value respectively.

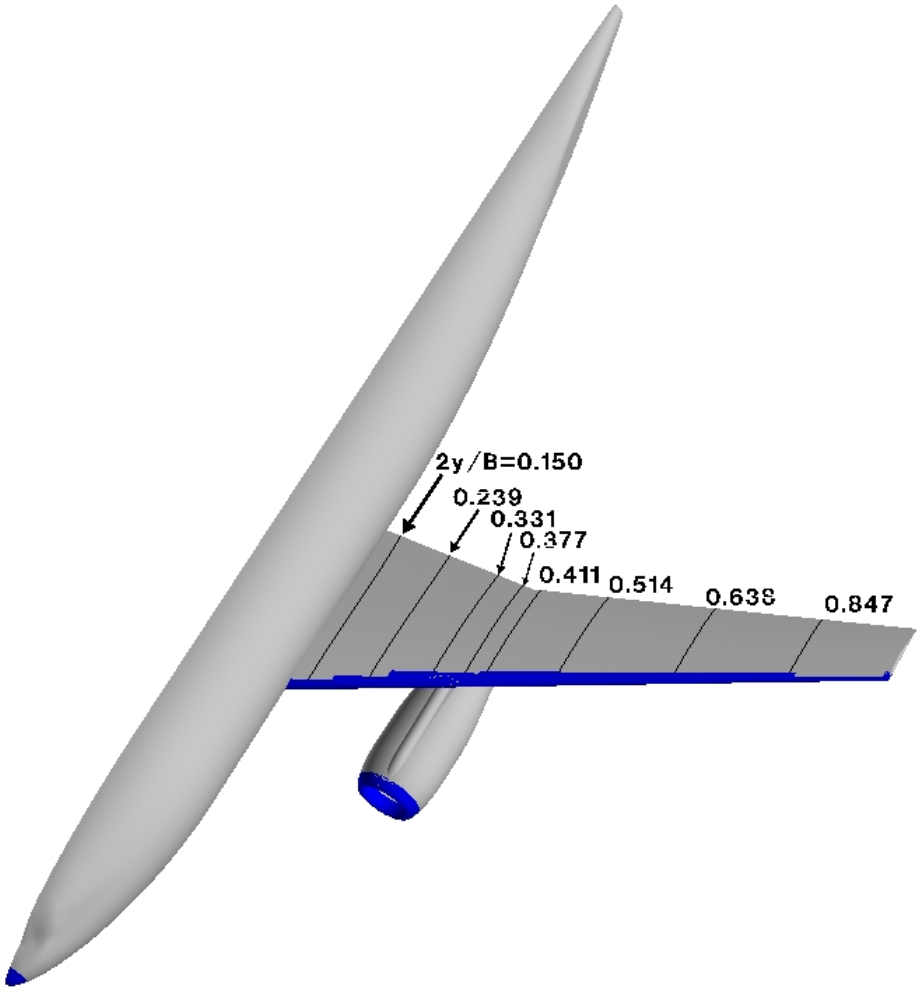
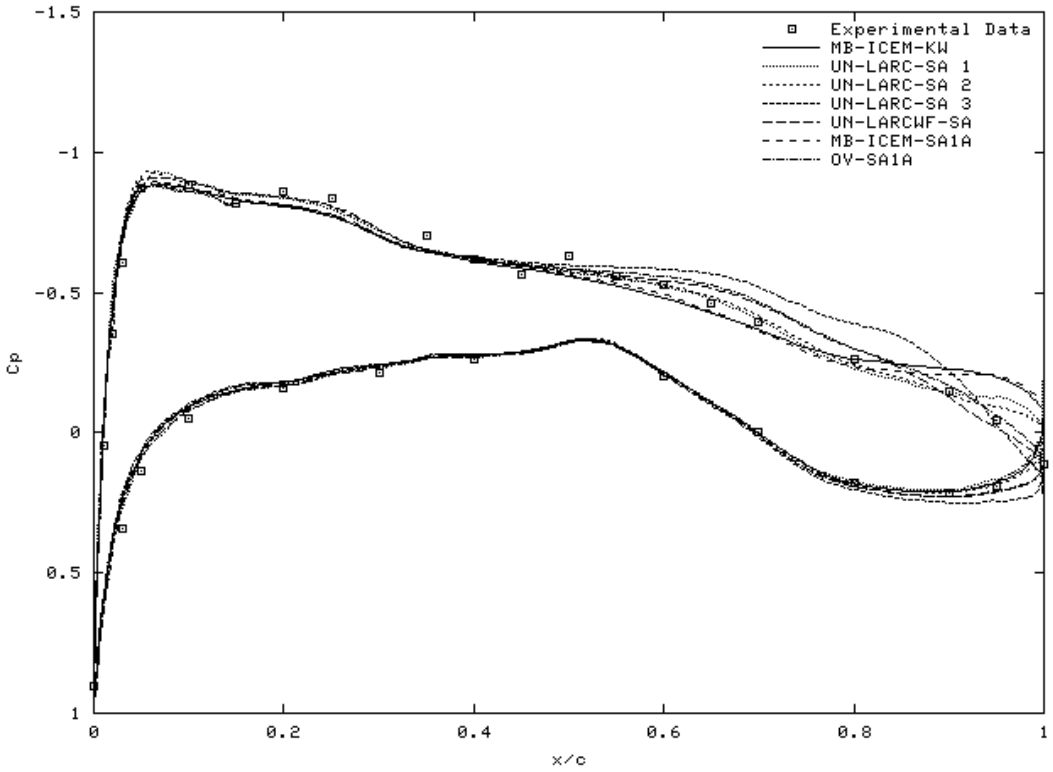
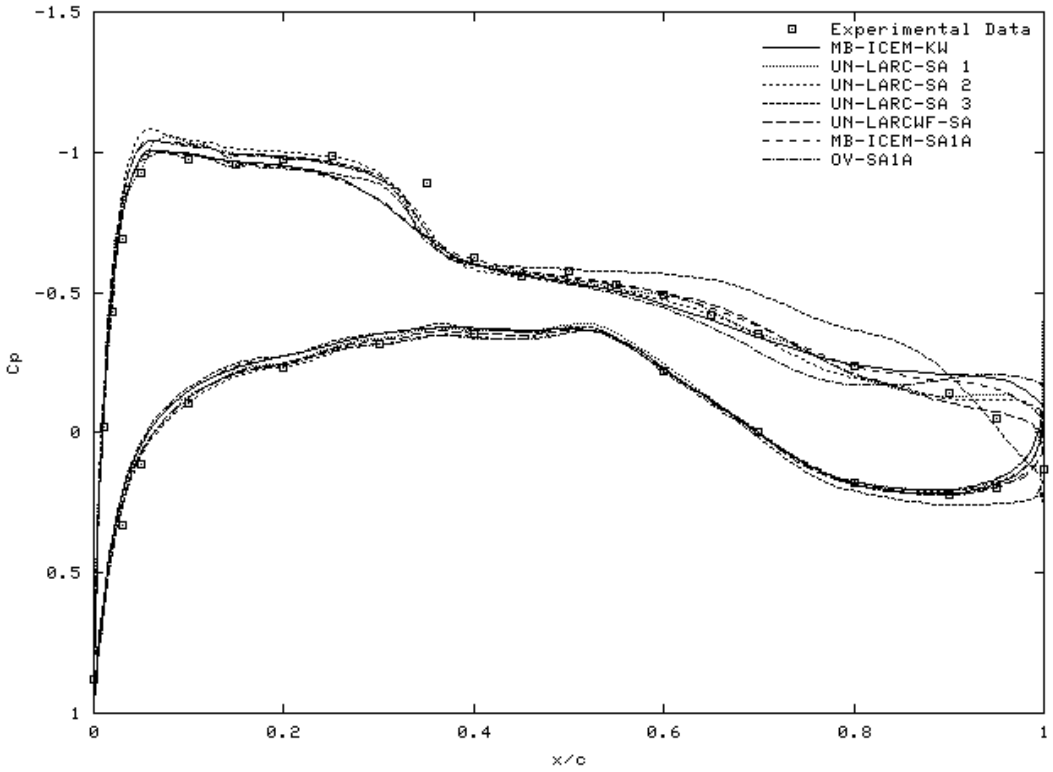


Figure 18. Location of wingspan stations for pressure comparisons.

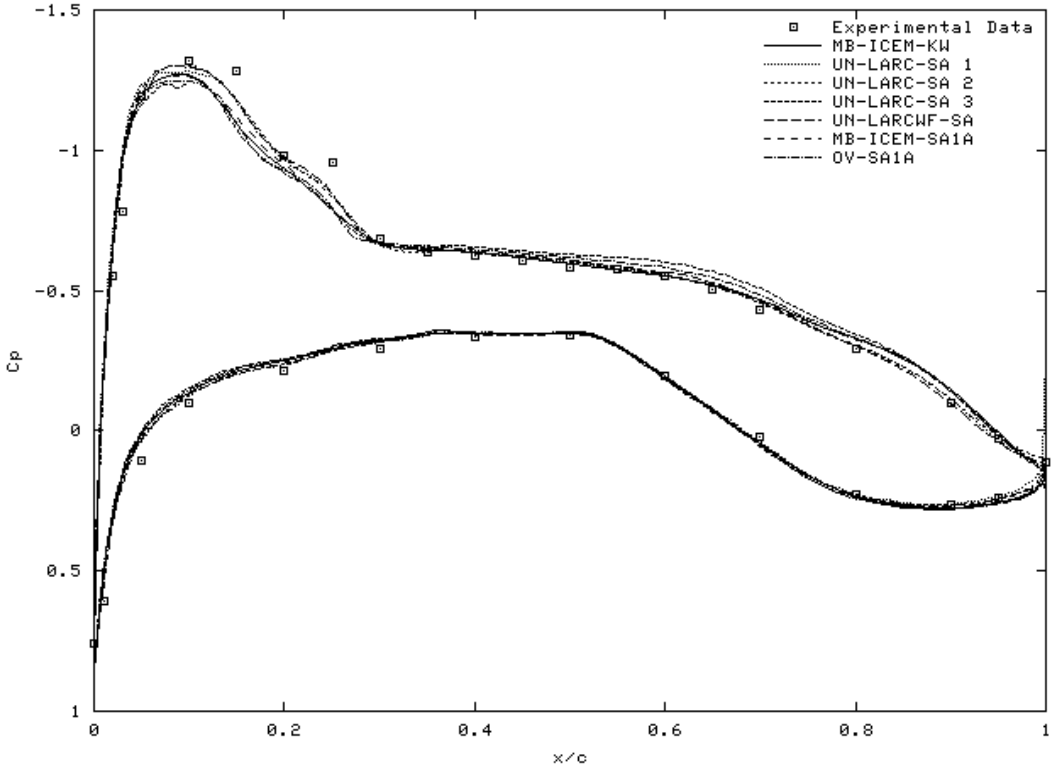


(a) Wing-body combination

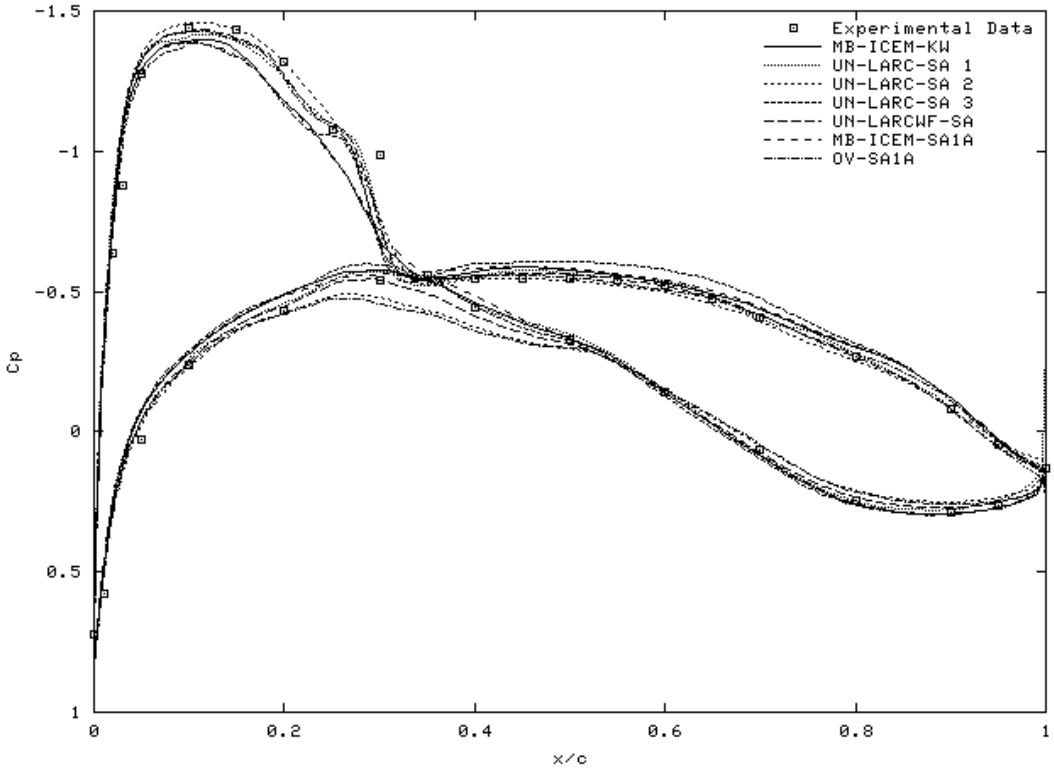


(b) Wing-body-nacelle-pylon combination

Figure 19. Wing surface pressures at 2y/b=0.15 for fine-grid solutions.

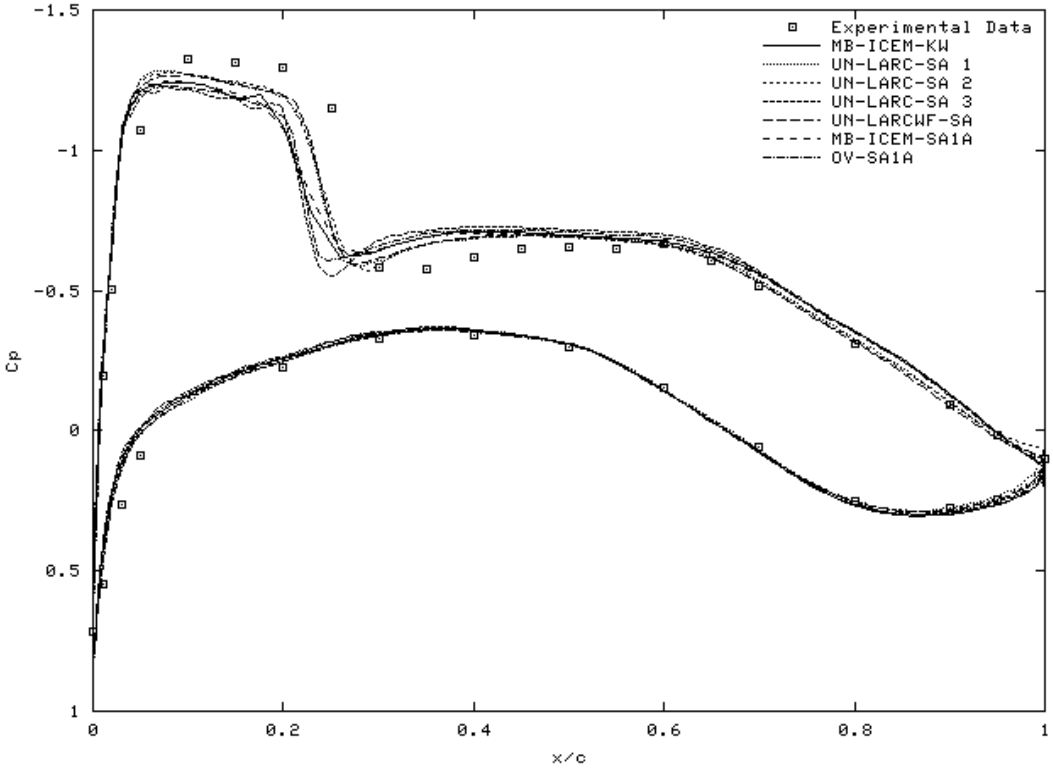


(a) Wing-body combination

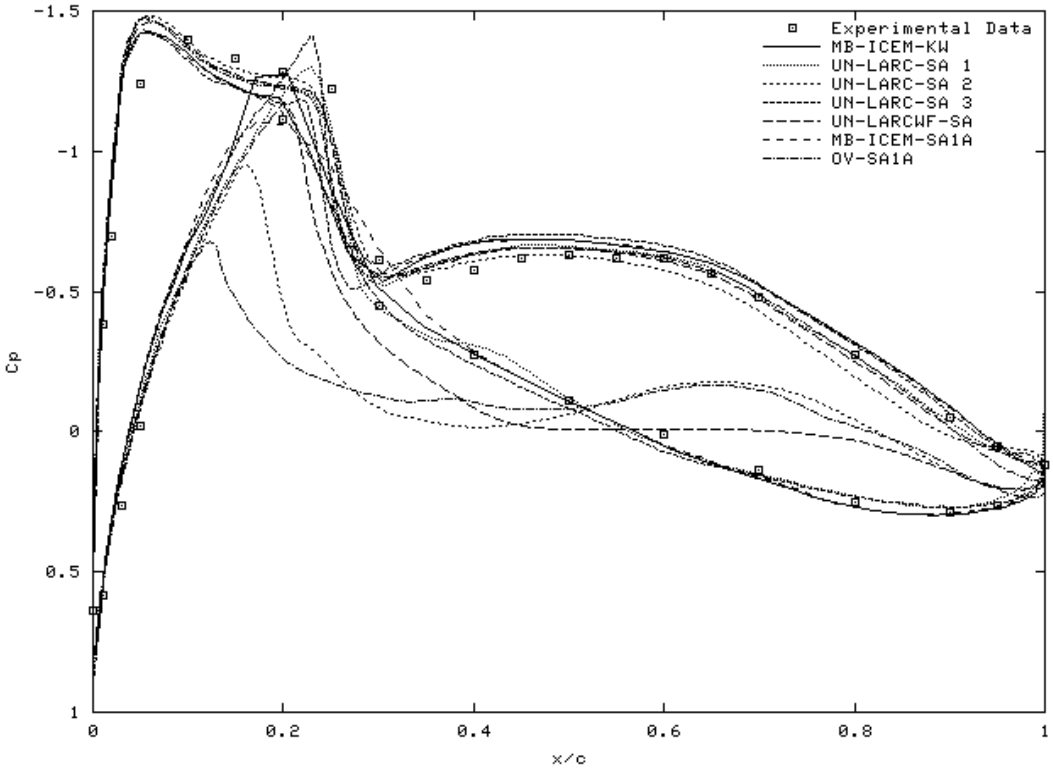


(b) Wing-body-nacelle-pylon combination

Figure 20. Wing surface pressures at 2y/b=0.239 for fine-grid solutions.

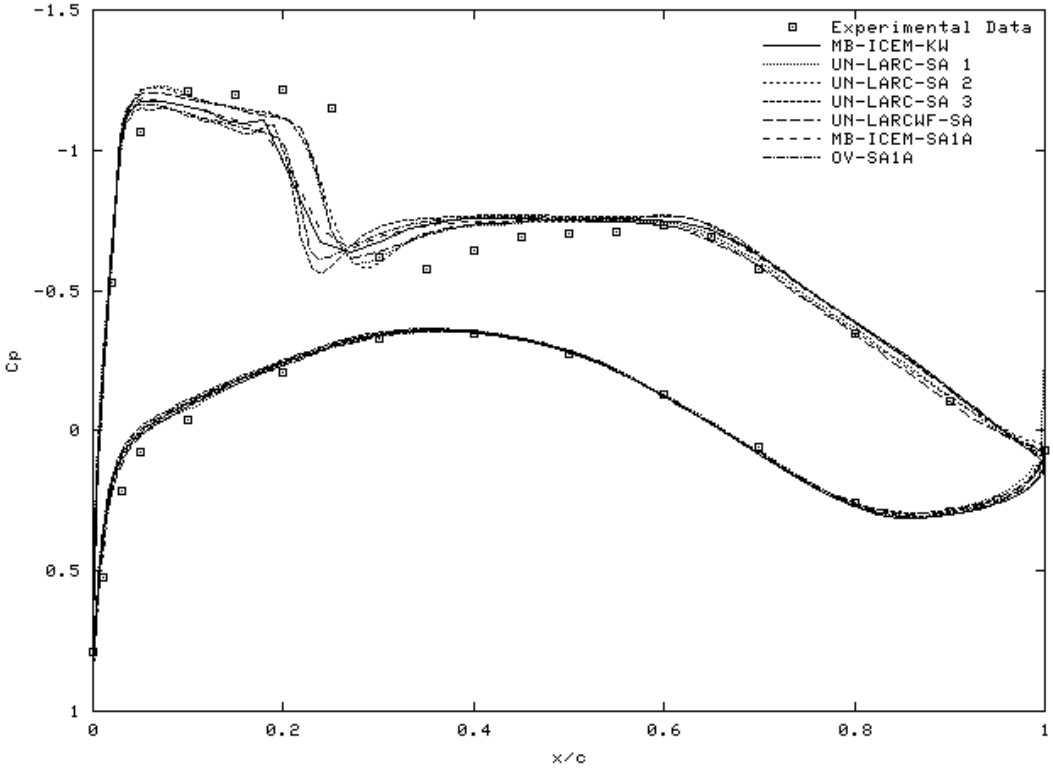


(a) Wing-body combination

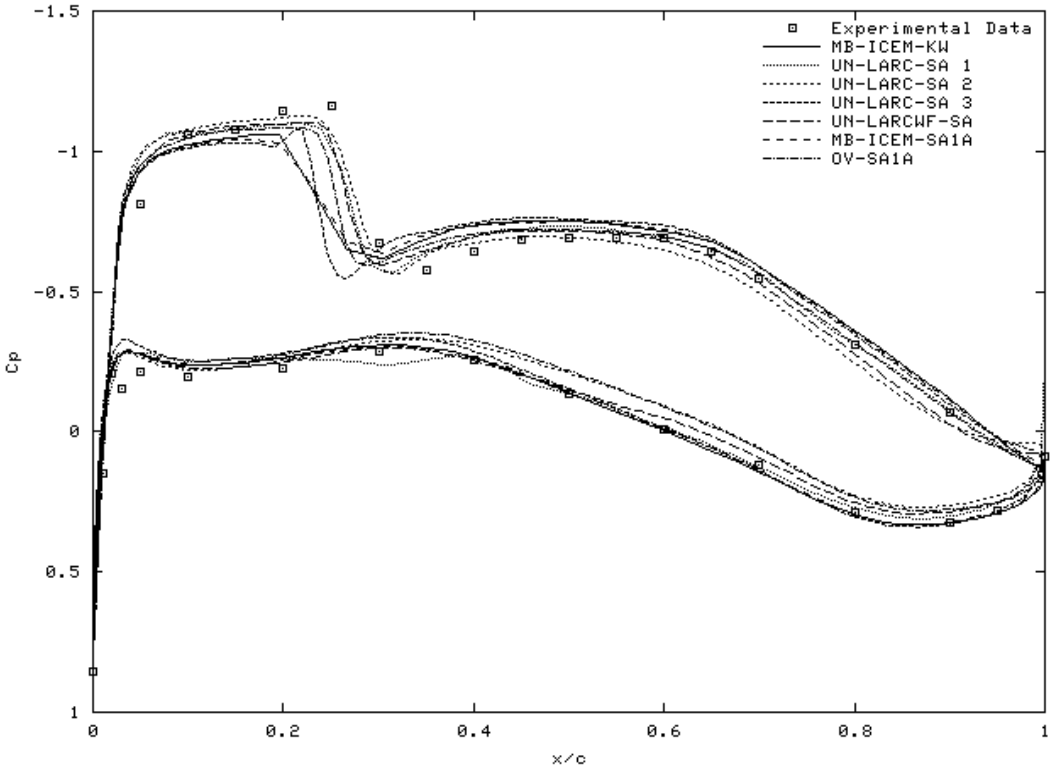


(b) Wing-body-nacelle-pylon combination

Figure 21. Wing surface pressures at 2y/b=0.331 for fine-grid solutions.

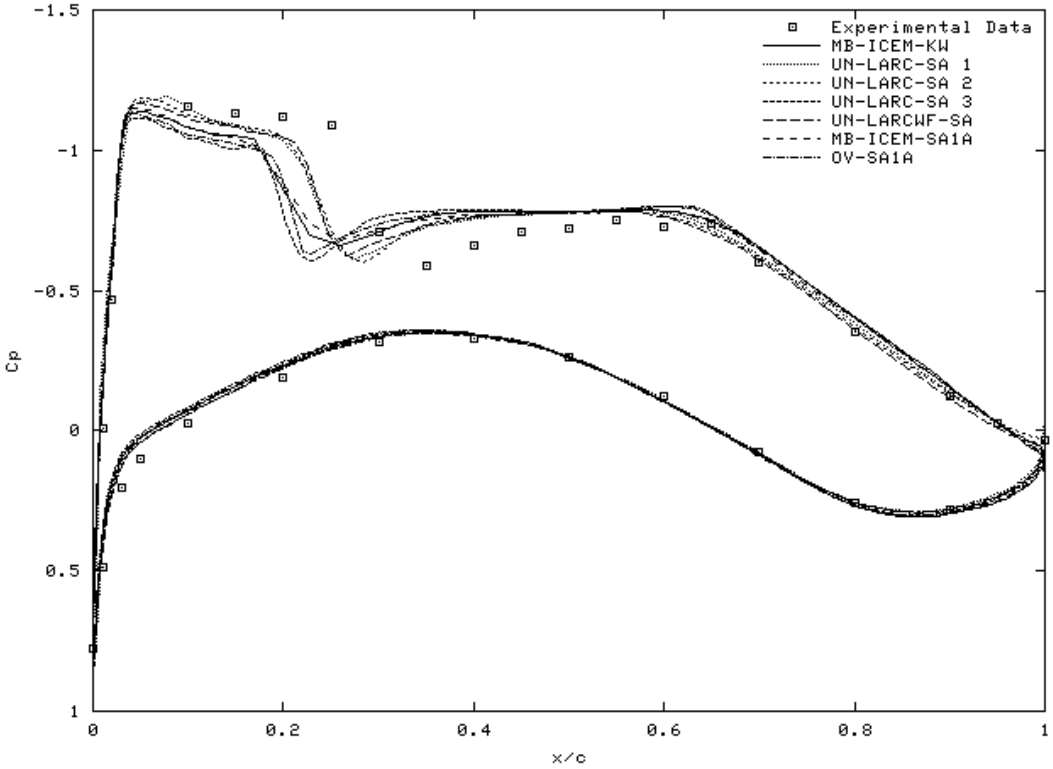


(a) Wing-body combination

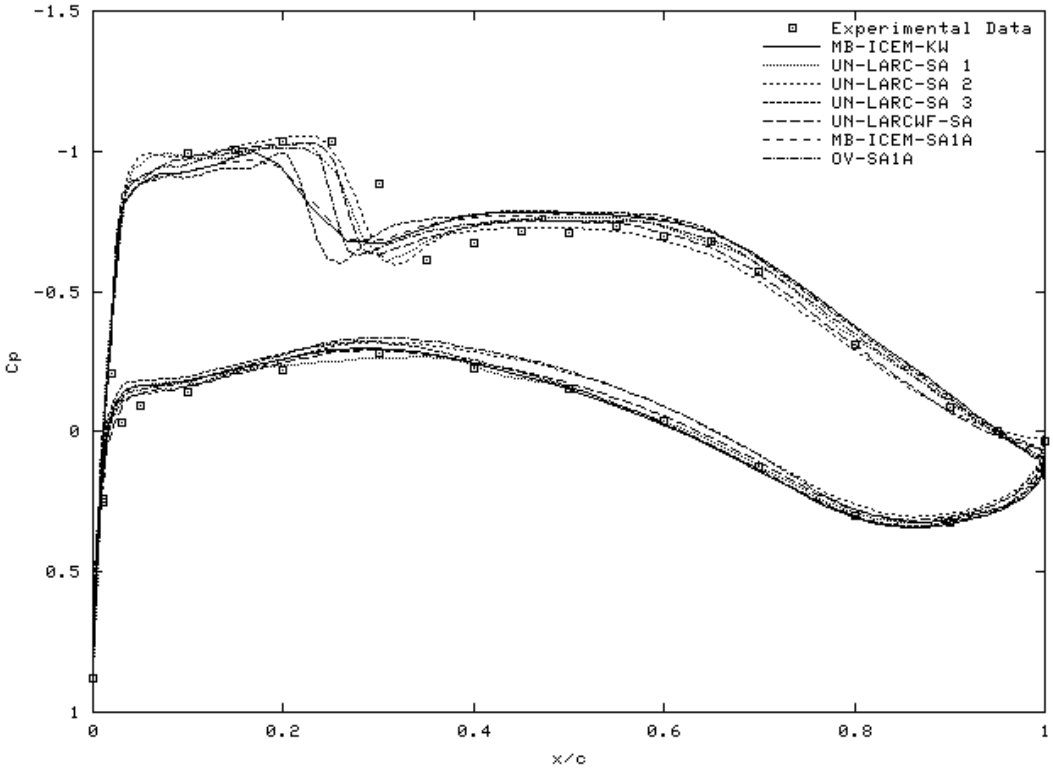


(b) Wing-body-nacelle-pylon combination

Figure 22. Wing surface pressures at 2y/b=0.337 for fine-grid solutions.

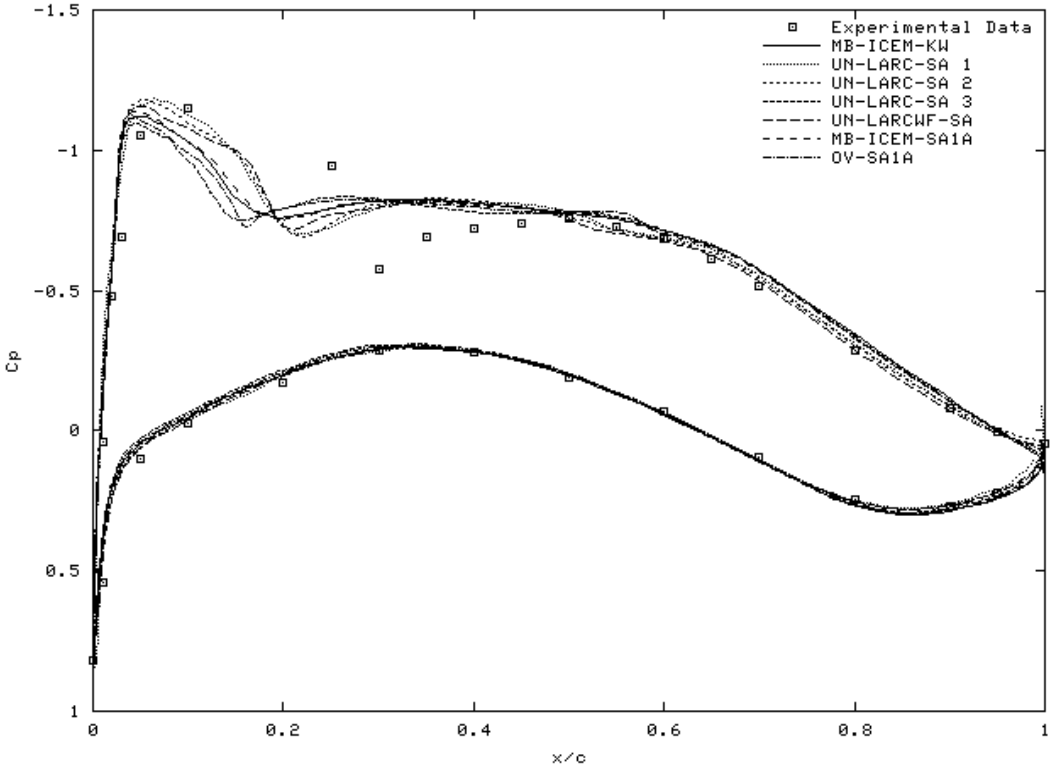


(a) Wing-body combination

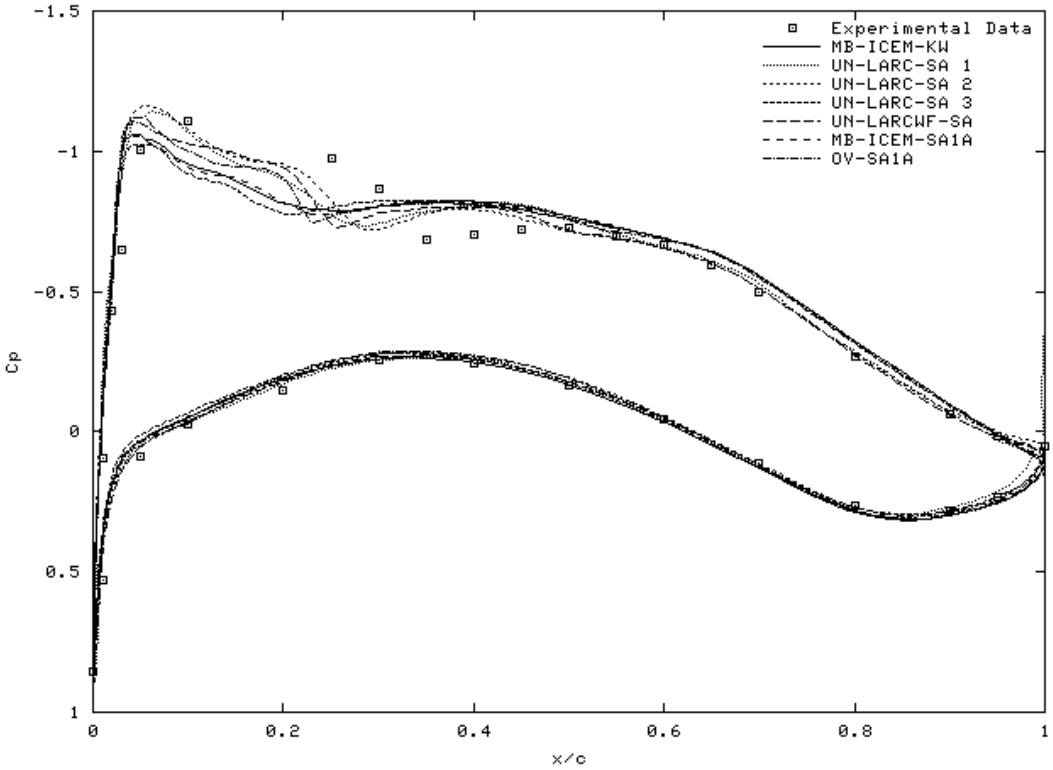


(b) Wing-body-nacelle-pylon combination

Figure 23. Wing surface pressures at 2y/b=0.441 for fine-grid solutions.

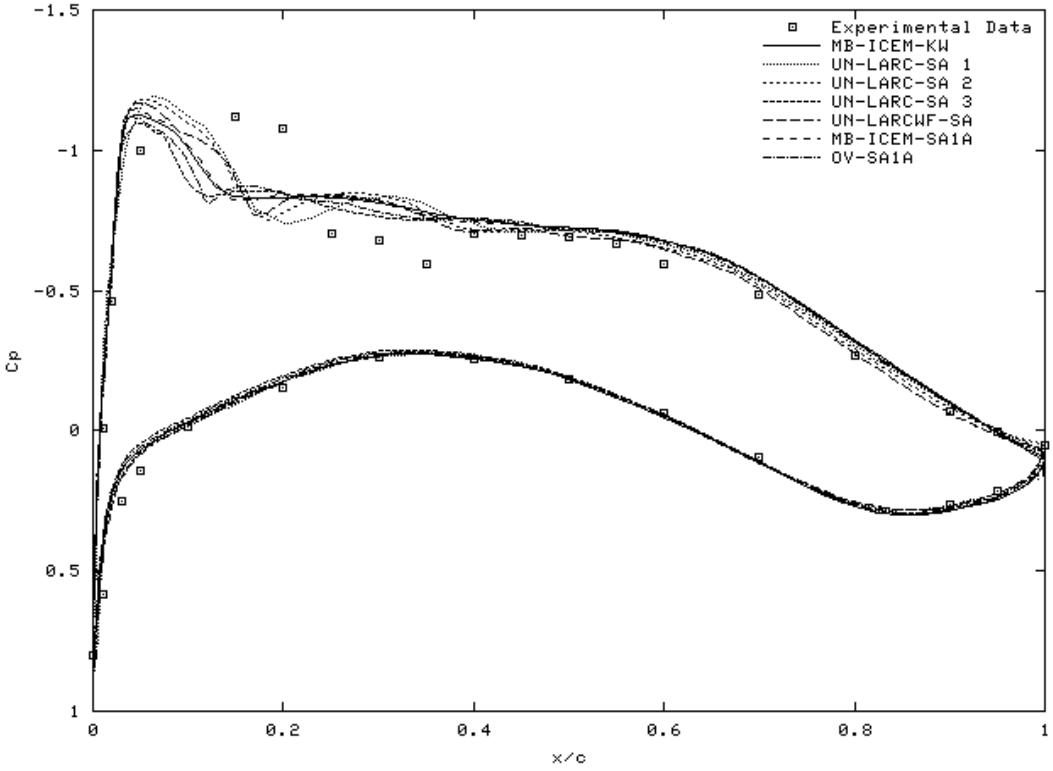


(a) Wing-body combination

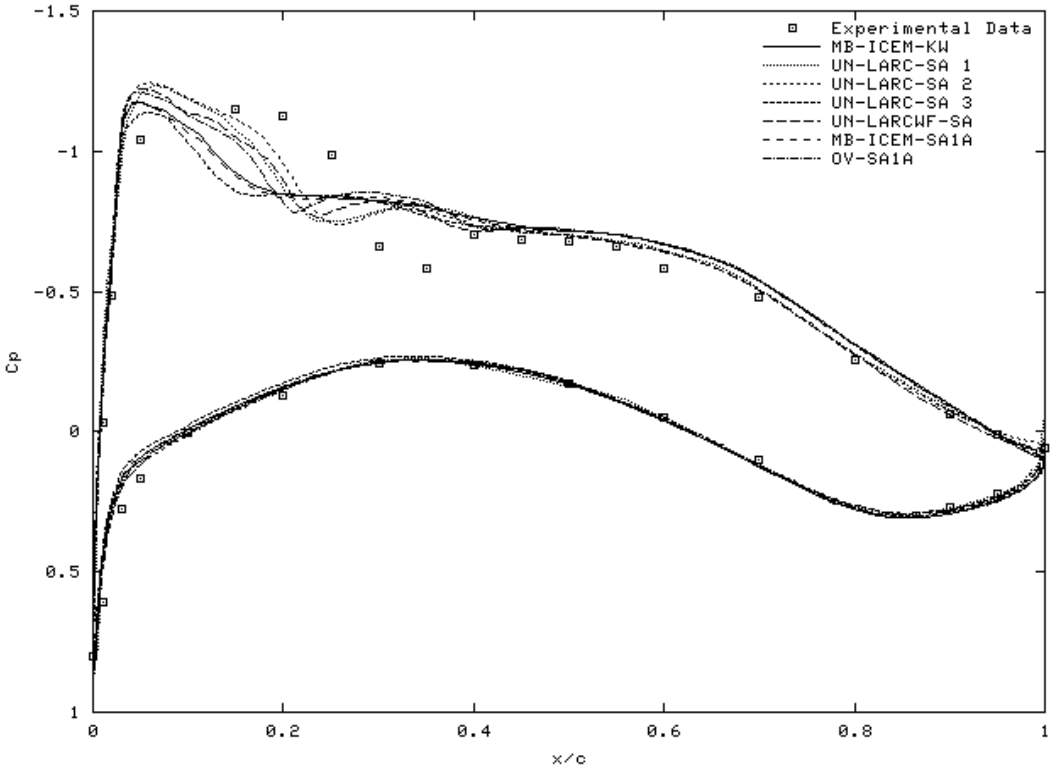


(b) Wing-body-nacelle-pylon combination

Figure 24. Wing surface pressures at 2y/b=0.514 for fine-grid solutions.

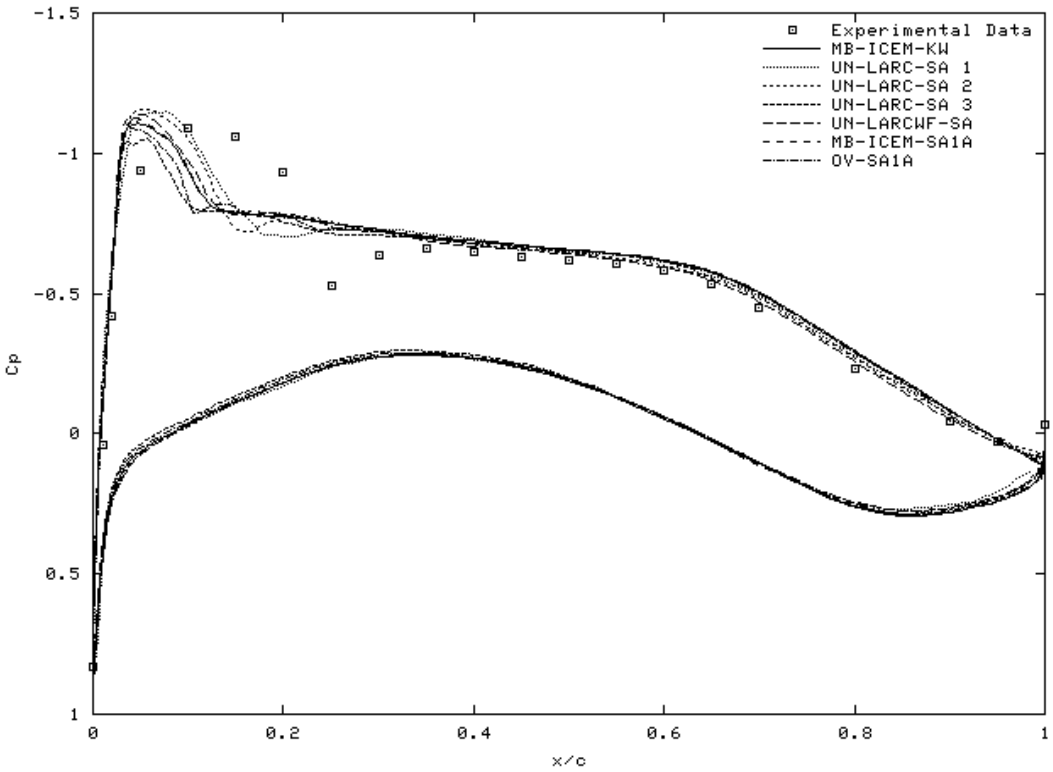


(a) Wing-body combination

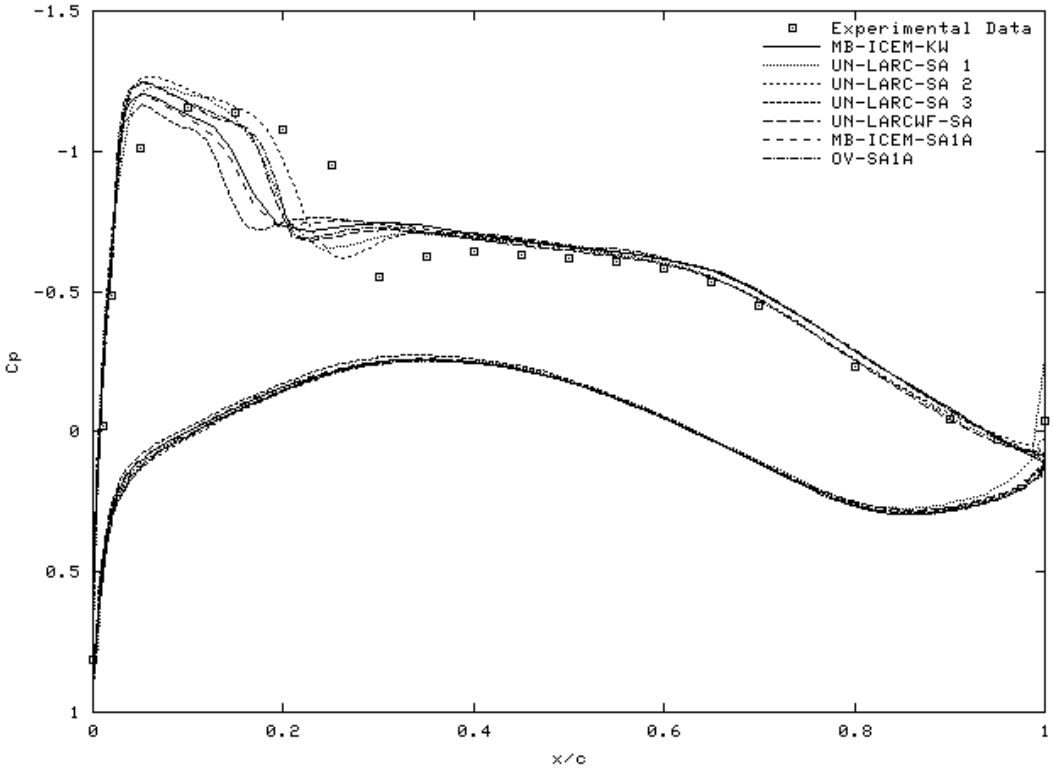


(b) Wing-body-nacelle-pylon combination

Figure 25. Wing surface pressures at 2y/b=0.638 for fine-grid solutions.



(a) Wing-body combination



(b) Wing-body-nacelle-pylon combination

Figure 26. Wing surface pressures at 2y/b=0.847 for fine-grid solutions.