Published in final edited form as:

J Aircr. 2019 July ; 55(4): 1388–1400. doi:10.2514/1.C034938.

Statistical Analysis of Solutions from the Sixth AIAA Computational Fluid Dynamics Drag Prediction Workshop

Joseph M. Derlaga^{*}, Joseph H. Morrison[†]

NASA Langley Research Center, Hampton, Virginia, 23681

Abstract

A graphical framework is used for statistical analysis of the results from an extensive N-version test of a collection of Reynolds-averaged Navier-Stokes computational fluid dynamics codes. The solutions were obtained by code developers and users from North America, Europe, Asia, and South America using both common and custom grid sequences as well as multiple turbulence models for the June 2016 6th AIAA CFD Drag Prediction Workshop sponsored by the AIAA Applied Aerodynamics Technical Committee. The aerodynamic configuration for this workshop was the Common Research Model subsonic transport wing-body previously used for both the 4th and 5th Drag Prediction Workshops. This work continues the statistical analysis begun in the earlier workshops and compares the results from the grid convergence study of the most recent workshop with previous workshops. Most notably, the results reinforce a lesson learned from the 5th Drag Prediction Workshop regarding the importance of a common grid sequence in decreasing solution variation and demonstrate that predicted drag increments show less scatter between codes than the predicted absolute values of drag for a given configuration.

I. Introduction

In June 2001, the AIAA Applied Aerodynamics Technical Committee (APATC) conducted a Drag Prediction Workshop (DPW) to evaluate the state of the art of Computational Fluid Dynamics (CFD) in predicting transonic aircraft forces and moments. The objectives for the workshop were (1) to assess the state-of-the-art computational methods as practical aerodynamic tools for aircraft force and moment prediction of industry relevant geometries, (2) to provide an impartial forum for evaluating the effectiveness of existing computer codes and modeling techniques using Reynolds averaged Navier-Stokes solvers, and (3) to identify areas needing additional research and development. The focus of the workshop was on drag prediction accuracy. Furthermore, DPW used a statistical framework to assess the comparisons within the computational results.

The success of DPW led to a total of six workshops. All of the workshops maintained the objectives of the first DPW, and a statistical framework was used for each including comparisons with earlier workshops. All six workshops focused on commercial transport configuration force and moment predictions, with each workshop introducing new cases: in

^{*}Research Scientist, Computational AeroSciences Branch, 15 Langley Blvd. MS 128, AIAA Member

[†]Associate Project Manager, Transformational Tools and Technologies, Aeronautics Research Directorate, 6 East Taylor St., MS 254, AIAA Associate Fellow

2001, the first Drag Prediction Workshop (DPW-I) focused on the DLR-F4, [1, 2] in 2003, the second Drag Prediction Workshop (DPW-II) introduced a grid convergence study and focused on total and incremental drag for the DLR-F6, [3, 4] in 2006, the third Drag Prediction Workshop (DPW-III) provided a blind test case for a side of body fairing on the DLR-F6, [5, 6] in 2009, the fourth Drag Prediction Workshop (DPW-IV) focused on blind drag prediction for the NASA Common Research Model (CRM) transonic wing-body-tail, [7, 8] and in 2012, the fifth Drag Prediction Workshop (DPW-V) focused on a common grid study and a buffet study for the CRM. [9, 10]

In 2016, DPW-VI focused on studying a 2D verification test case, predicting the drag increment for the CRM nacelle-pylon, and predicting the static aeroelastic effect for an angle of attack sweep using the aeroelastic deflection measured in the wind tunnel, as Rivers et al. [11] and Hue [12] demonstrated the impact of static aeroelastic deflection for the CRM after DPW-V. The static aeroelastic effect for an angle of attack sweep required substantially more grid generation work as each angle of attack had a different geometry corresponding to the aeroelastic deflection. Optional cases included a grid adaption case and a coupled aerostructural simulation. Roy et al. [13] provide a summary of the verification test case, and Tinoco et al. [14] provide a summary of the nacelle-pylon drag increment, static aero-elastic effect, grid adaption, and coupled aero-structural simulation test cases in DPW-VI. This paper focuses on a statistical analysis of the grid convergence study for the DPW-VI nacelle-pylon increment test case.

This paper is organized in the following manner. Section II provides a description of the statistical analysis. Section III outlines the test cases for the workshop. Section IV details the statistical results. Section V provides some summary comments.

II. Statistical Approach

Hensch [2] introduced the idea of treating different computations of a test case as a collective and using N-version testing in a statistical framework to investigate the submissions. No individual result is considered the “right” or “best” result. This framework is useful for identifying differences between submissions. The dispersion of the results is treated as noise in the collective computational process.

A running record of individual outcomes is plotted for each of the measures of interest and derived quantities reported by participants. Participants reported several quantities including angle of attack (α), total drag coefficient (C_D), pressure drag coefficient (C_{DPR}), skin friction drag coefficient (C_{DSF}), and pitching moment coefficient (C_m). The value of the measure of interest is plotted on the vertical axis and a unique (to this article) integer index is used for each data submission on the horizontal index. The order of the solutions on the horizontal axis is arbitrary since this is not a temporal axis.

An estimate of the population mean $\hat{\mu}$ of the plotted data submissions is made and is shown on the graph as the centerline. Upper and lower scatter limits are placed on the graph as follows:

$$\text{Upper Limit} = \hat{\mu} + K\hat{\sigma} \quad (1)$$

$$\text{Lower Limit} = \hat{\mu} - K\hat{\sigma} \quad (2)$$

where $\hat{\sigma}$ is an estimate of the population standard deviation and K is an appropriate coverage factor. Significant results are outcomes that lie outside the process limits defined in Eqs. (1) and (2). These results, referred to as outliers, represent submissions that are different from the results that lie within the scatter limits (and should be investigated to understand the difference). The population mean is estimated using the sample median which provides a robust estimate when outliers are present. The sample standard deviation is used to estimate the population standard deviation. The coverage factor is estimated for a uniform distribution as $K = \sqrt{3}$. [15] This value of coverage factor was chosen to look for differences in submissions based on past workshop experience.

For this workshop, an additional method of graphically examining the submitted results was developed. Figure 1(a) shows a combination of a violin plot [16] with a box-and-whisker (with notch) [17] plot overlay. The violin plot represents the sample distribution, indicating the levels of solution clustering and behavior of the tails of the data. The box-and-whisker plot groups data by quartiles and indicates outliers. At the neck of the box plot is the sample median, while the upper and lower limits of the box represent the upper and lower quartiles (medians of the upper and lower 50% of the data). The notch around the median is included to aid in comparisons of the median values between different grid levels; nonoverlapping notches are indicative of a significant difference between the median values. The distance between the upper and lower bounds of the box is referred to as the interquartile range (IQR), and contains 50% of the submitted data. The whiskers represent the maximum and minimum values outside of the IQR; however, if the maximum or minimum values lie outside the value of $1.5 \times \text{IQR}$ added or subtracted to the upper or lower quartiles, respectively, then the whiskers represent the maximum or minimum values that are considered statistically significant. Values that are outside the $1.5 \times \text{IQR}$ limit are shown as solid blue circles. In addition, the sample mean is plotted as a solid blue diamond. The horizontal axis shows the grid level and the number of submissions received at that grid level.

Figure 1(b) is a scatter plot of the same data submissions as in Fig. 1(a), and includes the median as a solid line and upper and lower scatter limits, $\hat{\mu} \pm K\hat{\sigma}$, as dotted lines. Note that apparent outliers in Fig. 1(b) are not always treated as outliers in Fig. 1(a). The dotted line in Fig. 1(b) should be interpreted as a relatively broad line, and in order for an outlier to truly be considered an outlier, it must fall significantly outside the range of the scatter limits.

III. DPW-VI Test Cases

Unique to DPW-VI compared to past drag prediction workshops, a code verification study was required of all participants. A subsonic, 2D, NACA 0012 from the Turbulence Modeling Resource was studied as Case 1. More details can be found in Roy. [13]

Cases 2 through 5 all studied the geometry of the Common Research Model (CRM) as described in Vassberg et al. [18] For these studies, the reference conditions were freestream Mach = 0.85, Re = 5 million, and temperature = 100 °F, in free air, i.e., no wind tunnel walls or model mounting systems were included. Boundary layers were modeled as fully turbulent.

Structured overset and unstructured grids were made available by the organizing committee; custom grid systems were welcome. The unstructured grid system was designed for cell-vertex schemes, and therefore resulted in a much higher number of degrees of freedom when used by cell-centered codes. Regardless of the grid system, six grid levels (L1 through L6) were provided by the committee, starting at approximately 20 million nodes for the wing-body (WB) geometry (approximately 25–30 million for the wing-body-nacelle-pylon (WBNP) geometry) and increasing by approximately 1.5x in total vertex count for each subsequent grid level. The coarsest grid level was designed to ensure a viscous wall spacing (y^+) of approximately 1.

A grid convergence study on the middle four grid levels (L2 through L5) was performed for Case 2A and Case 2B. An α -seek was performed in order to match a desired loading condition of $C_L = 0.5 \pm 0.0001$, using a geometry modified for aeroelastic deflection at $\alpha = 2.75^\circ$. Case 2A consisted of a WB geometry, while Case 2B added an engine nacelle and pylon (WBNP) to the geometry of Case 2A in order to understand the predictive capabilities of the various CFD tools in calculating drag increments between configurations.

Case 3 performed an α -sweep study of a model that had undergone static aeroelastic deflections at various angles of attack. In examining previous workshop results, the incorrect definition of wing twist from the as built model and lack of aeroelastic deflections in the workshop grids was found to seriously hamper comparisons with experimental data. [11, 12] Cases 4 and 5 were optional, involving a grid adaptation study of the Case 2A configuration, and a coupled aerostructural simulation of the Case 2A geometry, respectively.

The main focus of this paper is on the results for Case 2A and 2B. Before a detailed analysis was made, an initial study of the submissions identified several outliers. As an example, the total drag coefficient for Case 2B for all submissions on the L3 grid level are shown in Fig. 2(a).

Submissions 43, 44, and 45 were found to have not performed the α -seek in order to match the $C_L = 0.5$ condition and were therefore removed from the statistical analysis. This left one outlier, as shown in Fig. 2(b). Submission 42 utilized a Spalart-Allmaras (SA) turbulence model *, which has been shown in past workshops to suffer from side of body separation, and which was not fully converged on all submitted grid levels. As a result, it has been removed from the statistical analysis as it greatly skews the statistical results. Upon removal of submission 42, the confidence bands have been greatly reduced for all quantities of interest, as shown in Fig. 2(c). All remaining outliers are discussed in detail below.

*Not all submissions utilized the nomenclature of the Turbulence Modeling Resource (TMR) website [19], and so this work refers to a turbulence model by 'class' when the information was not provided.

IV. Results

The data were submitted to the DPW committee before the workshop. After the workshop, authors were given time to evaluate and resubmit their solutions. Several authors took advantage of this time and submitted corrections to their solutions or provided additional data that they did not have time to complete before the workshop. Within the following sections, summary violin/box-and-whisker plots will be shown. These summary plots will be shown side-by-side for Cases 2A and 2B in order to quickly see the trends in the changes of the quantities of interest between the two configurations. Scatter plots (including scatter limits) of a quantity of interest vs. submission number, such as those shown in Fig. 1(b) above, are included in the appendix. Outliers identified in the box-and-whisker plots will be referred to by their integer submission number from the scatter plots.

A. General Summary for Cases 2A and 2B

The IQR for all of the drag coefficients, shown in Fig. 3 through Fig. 5, is universally on the order of four to five drag counts (1 count = $0.0001 C_D$), with the scatter limits reducing by approximately one to two counts per grid level, indicating that while approximately 50% of the submissions were tightly clustered about the median, most of the improvement in reducing the scatter bounds comes from convergence of outliers toward the median. This is not to say that the value of the median doesn't change with grid refinement; just that close to 50% of the data closely tracks the median with grid refinement. Angle of attack, Fig. 6, and pitching moment, Fig. 7, show a similarly constant IQR with grid refinement.

On the finer grid levels, where there are fewer submissions, the IQR typically reduces for the three drag coefficients, but not for the other two quantities. In addition, there were typically 10% to 20% fewer submissions for Case 2B than Case 2A.

Summaries of the medians and standard deviations for Cases 2A and 2B can be found in Table 1 through Table 4. The presence of the nacelle and pylon increase the standard deviation in the drag by less than a count. Compared to the standard deviations computed for DPW-V, [10] the results from DPW-VI show higher deviation for almost all quantities, but this can be attributed to the fact that a common grid sequence was used by the majority of DPW-V submissions. Overall, the medians of the various drag coefficients change very little with grid refinement, with only the skin friction drag for Case 2B showing a change of more the 2 counts, which is well within the standard deviation for that quantity.

B. Summary of Outliers for Cases 2A and 2B

Beginning with angle of attack, one statistically significant outlier is identified on grid levels L2 through L4 for both Cases 2A and 2B (submission 1 as seen in Appendix I). Total pitching moment indicates no outliers for Case 2B due to a larger IQR, but two on the L2 grid level (submissions 1 and 9) and one on the L4 grid level (submission 8) for Case 2A. In general, submissions that had found a higher angle of attack in order to match C_L also had higher pitching moments. For total drag coefficient, submission 6 was the highest outlier for all cases, while submission 11 was second highest for Case 2A on the L2-L4 levels, with submission 21 third highest for Case 2A on the L3 level. For Case 2B, submission 21 was

the second highest outlier on the L2-L4 levels, while submission 19 was the third highest outlier on the L2 grid level. As with C_D , the pressure drag had outliers from submissions 6 and 21 for both cases; submission 6 was the highest outlier for grid levels L2-L4, while submission 21 became the highest for the L5 level. For skin friction drag, submission 27 has the lowest skin friction across all grid levels for both cases (although it is not statistically significant for the L4 and L5 grid levels for Case 2B) followed by submission 8 on the L2 and L3 levels of Case 2A and the L2 level for Case 2B.

Submission 1: While generally showing the highest angle of attack, this submission typically also had a lower magnitude of the pitching moment, close to that of submission 8. Also, while the skin friction value was typically lower than the median, the predicted pressure drag was higher, resulting in total drag coefficients that closely tracked the median. This submission was unique in that it used the k-kL-MEAH2015 turbulence model and used the workshop provided grid system. Submissions 1 and 2 were contributed by the same group, differing only in the use of a reconstruction limiter for the finite volume scheme for submission 2, whereas submission 1 had no such limiter. Submission 2 fell within the scatter limits, but tended to have higher values of angle of attack and lower magnitudes of pitching moment. The difference in turbulence model could well account for the differences between all other submissions.

Submission 6: This submission was typically highest in both total drag and pressure drag, but was within the scatter bounds for all other quantities. An SST- $k\omega$ (Shear Stress Transport) turbulence model was used on custom grids which were much finer, often by a factor of 3x, than the requirements given in the gridding guidelines for the workshop. The solver uses a cell-centered, pressure-based approach, which is unique from all other submissions, and may account for the statistical differences.

Submissions 8 and 9: These submissions were from the same group; submission 8 was unique in that it was the sole submission that used an EARSM (Explicit Algebraic Reynolds Stress Model), and submission 9 used an SA model, both run on the same workshop provided grids. Both submissions typically had higher angles of attack and lower pitching moment magnitudes, even if not statistically significant using the IQR criteria. The EARSM also tended toward having a lower skin friction coefficient. It should be noted that all boundary conditions (including walls) for these submissions are weakly enforced and could therefore account for the differences compared to other solutions.

Submission 19: While considered a statistically significant outlier for the total drag for Case 2B on the L2 grid level, submission 19 is generally within the bounds of the scatter limits but does tend to have a higher value of pressure drag with grid refinement. This submission used an SST turbulence model on Boeing provided unstructured grids.

Submission 21: This submission also used the Boeing provided unstructured grids, but was contributed by another group than that of submission 19. An SA-noft2-QCR2000 turbulence model was used for this submission, but the same group used the same turbulence model and their own grid for submission 25 and fell well within the scatter limits. In addition, this group also contributed submission 20, which used an SA-noft2 model, and

also fell within the scatter limits. This may imply that the Boeing provided grid may be unsuitable for this solver when using the SA-noft2-QCR2000 model.

Submission 27: A cell-centered flow solver using an SA turbulence model on custom grids that followed the gridding guidelines was used for this submission. This has resulted in this submission having the lowest values of skin friction and total drag for both cases at all grid levels. There were no signs of premature side-of-body separation and all cases appeared to be well converged. There was no identified reason why this submission resulted in low values of skin friction and total drag despite having reasonable pressure drag.

C. Delta of Quantities Between Cases 2A and 2B

This workshop provided the ability to study the drag buildup (ΔC_D) between the CRM WB and CRM WBNP configurations; violin/box-and-whisker summary plots for the three drag coefficient quantities are provided in Fig. 8(a) through Fig. 8(c), where the results of Case 2A are subtracted from the results of Case 2B. While the spread between the whiskers for the drag coefficients for Cases 2A and 2B were on the order of 10 to 15 counts, the spread between the whiskers for the ΔC_D s are on the order of 5 counts or less. The IQR is also approximately reduced by a factor of three for each of the drag coefficient quantities as well. While no outliers exist for C_{DSF} , the submissions appear to follow a bimodal distribution, and the overall range between all the submissions is generally on the order of 2 counts.

Outliers in C_{DPR} typically result in higher values of total drag. Submission 6 is the highest outlier for total drag and pressure drag on all grid levels except on grid level L5, at which point submission 18, 19, 20, and 21 all become higher. Submissions 6 and 21 (and by extension 20) were discussed above. Submission 19 was also discussed above, and was contributed by the same group as submission 18. Even though submission 18 was the median value of C_{DPR} for Case 2A, it appears that the pressure drag computed for Case 2B was just high enough to become significant. On the L3 grid level, submission 33 also appears as an outlier for total drag and pressure drag. Submission 33 utilized a custom-hybrid grid system for a node-based finite-volume solver with an SST turbulence model; other submissions from the same group with the same turbulence model were not statistically outliers. As with submission 18, the median value for pressure drag on Case 2A was well matched and the computed value on Case 2B was slightly higher.

Also appearing as outliers on the L3 grid level for total drag are submissions 3 and 26. Submission 3 matched the median value of total drag for Case 2A, but was higher than the median on Case 2B. Unfortunately, submission 26 did not include a drag breakdown in order to determine the contributions from pressure and skin friction drag. Submission 18 is also a total drag outlier on the L4 grid level for the same reasons as noted above for the L5 grid level.

As with Cases 2A and 2B, summaries of the medians and standard deviations are given in Table 5 and Table 6. While skin friction drag accounts for almost 75% of the C_D , it is subject to far less deviation than the pressure drag, which in turns drives the deviation in the total drag. While C_D and ΔC_{DPR} show a slight decrease with grid refinement, α , C_m ,

and ΔC_{DSF} do not. As shown with the violin/box-and-whisker plots above, the C_v of quantities between configurations are somewhat constant with respect to grid level, even if the predicted magnitudes of those quantities are not.

D. Comparison of DPW-VI results with DPW-II through DPW-V

DPW-II through DPW-V each included a grid convergence study. Gridding guidelines were posted before each workshop and included recommended sizes for each of the grids in the grid convergence studies to provide a guide to the participants. For DPW-V and DPW-VI, grids were provided by the committee. Grid sizes tripled over the course of these three workshops. However, the question remains whether increasing grid sizes have improved the results.

The coefficient of variation, [20] $C_v = \sigma/\mu$, provides a measure to compare the variation of populations with different means. The results of DPW-VI are compared to the results of DPW-II, [4] DPW-III, [6] DPW-IV, [8] and DPW-V [10] for the cases that were run at constant lift. Morrison and Hemsch [6] showed that the variation in lift affected the comparisons of drag quantities and pitching moment and showed that quantities that corrected to a constant lift condition, e.g., L/D and idealized profile drag, were required to compare results at different values of lift.

The coefficient of variation for the pitching moment coefficient would be negative since the mean is negative and the estimate of the standard deviation is positive. Therefore, $|C_v|$ is plotted for the pitching moment coefficient to make comparisons easier.

The coefficients of variation for the total drag, pressure drag, skin friction drag, and pitching moment coefficients are shown in Figs. 9–12 for DPW-VI and earlier workshops; the coarsest grid for DPW-V is excluded from the plots for clarity but is included in Derlaga and Morrison [21] and in Table 8 through Table 12 in Appendix II which summarizes all of the $|C_v|$ results. This variation includes different turbulence models, different codes and numerics, different participants, different grids. Table 7 reports the number of submission for grid levels L1 through L6 for both Cases 2A and 2B of DPW-VI.

Compared to DPW-V, the variation in the DPW-VI results are universally higher for similar grid resolution. The variation for Case 2A is lower than the variation for Case 2B, but the L6 grid level for Case 2B had a median value of the grid size that was lower than the L5 grid level and also lower than the L6 grid size for Case 2A. One possible reason for this is that those who could afford to run a sixth grid level also had to use coarser resolutions at the lower grid levels, as compared to those using the workshop suggested guidelines for number of grid points. Despite this, the total variation in all the drag coefficients is actually lower for Case 2B on the L6 grid level compared to the L5 grid level, whereas the opposite is true for the drag coefficients for Case 2A. In general, the total variation for Case 2A shows increased variation at the L6 grid level, and this is most likely due to the decreased number of submissions. The pitching moment variations are also higher than the DPW-III results, but converge at a much faster rate vs. number of grid points.

Given that only DPW-V required a Common Grid Sequence, these results reinforce the impact of grid differences on increasing the variation between submissions. In all other workshops, $|C_v|$ for each coefficient converge in a similar manner, and demonstrate higher values than DPW-V. Given that all the workshops had a wide number of participants using a variety of discretization schemes, turbulence models, and convergence criteria, the Common Grid Sequence clearly had the greatest impact on reducing code to code variation, regardless of the underlying reliance on structured vs. unstructured grids.

V. Conclusions

A statistical analysis, including a new presentation of the submission statistics and outliers, was applied to the results of the 6th AIAA CFD Drag Prediction Workshop. The code-to-code scatter of the total drag coefficient, pressure drag coefficient, and skin friction drag coefficient were substantially higher than for DPW-V, but in line with the scatter from DPW-II through DPW-IV. The lower level of scatter in DPW-V was from the use of the Common Grid Sequence; DPW-VI did not use a common grid and reverted to scatter levels similar to other workshops that also did not use a common grid. However, the drag coefficients did not exhibit a smooth convergence trend and additional analysis of the individual submissions and grids will be required to identify the cause of this behavior. The statistical analysis of the grid convergence study showed approximately 18% of the submissions were outliers, which is consistent with earlier workshops. In addition, the delta between the wing-body and the wing-body-nacelle-pylon configuration had approximately 3x lower levels of scatter than the absolute values. The drag coefficient scatter for the delta between configurations was on the order of a few counts of drag.

Acknowledgments

The authors wish to thank the participants of the DPW series who have spent countless hours making this series of workshops a continuing success. The authors wish to thank the members of the DPW Organizing Committee for their comments and support. This work was supported by the Revolutionary Vertical Lift Technology (RVLT) Project of the NASA Advanced Air Vehicles Program.

Appendix I

Fig. 13.
Coarse (L2) grid measures of interest for Case 2A.

Fig. 14.
Medium (L3) grid measures of interest for Case 2A.

Fig. 15.
Fine (L4) grid measures of interest for Case 2A.

Fig. 16.

Extra Fine (L5) grid measures of interest for Case 2A.

Fig. 17.
Coarse (L2) grid measures of interest for Case 2B.

Fig. 18.
Medium (L3) grid measures of interest for Case 2B.

Fig. 19.
Fine (L4) grid measures of interest for Case 2B.

Fig. 20.
Extra Fine (L5) grid measures of interest for Case 2B.

Appendix II

This appendix contains all the data from present and past DPW's as shown in Figs. 9–12.

Table 8Normalized median grid size, $NPTS^{-2/3}$

Grid Level	DPW-II DLR-F6	DPW-III DLR-F6	DPW-III FX2B	DPW-IV CRM WBT	DPW-V CRM WB	DPW-VI CRM WB	DPW-VI CRM WBNP
L1	4.98E-05	5.24E-05	5.24E-05	3.62E-05	1.35E-04	1.34E-05	1.11E-05
L2	3.13E-05	2.40E-05	2.37E-05	2.02E-05	5.99E-05	1.13E-05	9.30E-06
L3	2.25E-05	1.11E-05	1.11E-05	9.31E-06	3.37E-05	8.51E-06	6.98E-06
L4					1.50E-05	6.27E-06	5.33E-06
L5					8.42E-06	4.62E-06	3.75E-06
L6					3.74E-06	3.52E-06	4.59E-06

Table 9 $|C_v|$ of C_D

Grid Level	DPW-II DLR-F6	DPW-III DLR-F6	DPW-III FX2B	DPW-IV CRM WBT	DPW-V CRM WB	DPW-VI CRM WB	DPW-VI CRM WBNP
L1	0.0290	0.0577	0.0612	0.0314	0.0367	0.0188	0.0200
L2	0.0174	0.0307	0.0275	0.0221	0.0180	0.0173	0.0183
L3	0.0240	0.0249	0.0206	0.0222	0.0128	0.0172	0.0180
L4					0.0108	0.0145	0.0168
L5					0.0090	0.0144	0.0175
L6					0.0085	0.0149	0.0169

Table 10 $|C_v|$ of C_{DSF}

Grid Level	DPW-II DLR-F6	DPW-III DLR-F6	DPW-III FX2B	DPW-IV CRM WBT	DPW-V CRM WB	DPW-VI CRM WB	DPW-VI CRM WBNP
L1	0.0413	0.0379	0.0345	0.0458	0.0335	0.0327	0.0320
L2	0.0396	0.0439	0.0425	0.0358	0.0226	0.0295	0.0290
L3	0.0434	0.0330	0.0296	0.0313	0.0178	0.0262	0.0258
L4					0.0146	0.0235	0.0242
L5					0.0146	0.0220	0.0215
L6					0.0134	0.0220	0.0159

Table 11

 $|C_v|$ of C_{DPR}

Grid Level	DPW-II DLR-F6	DPW-III DLR-F6	DPW-III FX2B	DPW-IV CRM WBT	DPW-V CRM WB	DPW-VI CRM WB	DPW-VI CRM WBNP
L1	0.0653	0.0854	0.0906	0.0558	0.0548	0.0291	0.0319
L2	0.0388	0.0409	0.0327	0.0321	0.0302	0.0282	0.0329
L3	0.0517	0.0441	0.0269	0.0308	0.0237	0.0289	0.0327
L4					0.0219	0.0245	0.0302
L5					0.0200	0.0234	0.0324
L6					0.0171	0.0249	0.0304

Table 12

 $|C_v|$ of C_m

Grid Level	DPW-II DLR-F6	DPW-III DLR-F6	DPW-III FX2B	DPW-IV CRM WBT	DPW-V CRM WB	DPW-VI CRM WB	DPW-VI CRM WBNP
L1	0.0490	0.0880	0.0850	0.1386	0.0376	0.0660	0.0827
L2	0.0450	0.0609	0.0671	0.1462	0.0353	0.0635	0.0781
L3	0.0522	0.0467	0.0593	0.1708	0.0355	0.0598	0.0721
L4					0.0371	0.0536	0.0676
L5					0.0386	0.0448	0.0563
L6					0.0427	0.0437	0.0637

Nomenclature

AR	aspect ratio
C_D	total drag coefficient
C_{DPR}	total pressure drag coefficient
C_{DSF}	total skin friction drag coefficient
C_L	total lift coefficient
C_m	total pitching moment coefficient
K	coverage factor for individual values
n	number of observations in a sample
$NPTS$	number of solution points in mesh
Re	Reynolds number based on reference chord

x_i	value of an observation
$\bar{x}$	sample mean of a set of observations
$\tilde{x}$	sample median
μ	population mean
$\hat{\mu}$	estimate of the population mean
σ	population standard deviation
$\hat{\sigma}$	estimate of the population standard deviation
C_v	coefficient of variation ($\hat{\sigma}/\hat{\mu}$)

References

- [1]. Levy D, Zickuhr T, Vassberg J, Agrawal S, Wahls RA, Pirzadeh S, and Hemsch MJ, "Data Summary from the First AIAA Computational Fluid Dynamics Drag Prediction Workshop," *Journal of Aircraft*, Vol. 40, No. 5, 2003, pp. 875–882.
- [2]. Hemsch MJ, "Statistical Analysis of Computational Fluid Dynamics Solutions from the Drag Prediction Workshop," *Journal of Aircraft*, Vol. 41, No. 1, 2004, pp. 95–103.
- [3]. Laflin KR, Klausmeyer SM, Zickuhr T, Vassberg JC, Wahls RA, Morrison JH, Brodersen OP, Rakowitz ME, Tinoco EN, and Godard J-L, "Data Summary from the Second AIAA Computational Fluid Dynamics Drag Prediction Workshop," *Journal of Aircraft*, Vol. 42, No. 5, 2005, pp. 1165–1178.
- [4]. Hemsch MJ, and Morrison JH, "Statistical Analysis of CFD Solutions from 2nd Drag Prediction Workshop," AIAA Paper 2004–0556, 1 2004.
- [5]. Vassberg JC, Tinoco EN, Mani M, Brodersen OP, Eisfeld B, Wahls RA, Morrison JH, Zickuhr T, Laflin KR, and Mavriplis DJ, "Summary of the Third AIAA CFD Drag Prediction Workshop," AIAA Paper 2007–0260, 1 2007.
- [6]. Morrison J, and Hemsch M, "Statistical Analysis of CFD Solutions from the Third AIAA Drag Prediction Workshop," AIAA Paper 2007–0254, 1 2007.
- [7]. Vassberg J, Tinoco EN, Mani M, Zickuhr T, Levy DW, Brodersen OP, Eisfeld B, Wahls RA, Morrison JH, Mavriplis DJ, and Murayama M, "Summary of the Fourth AIAA CFD Drag Prediction Workshop," AIAA Paper 2010–4547, 6 2010.
- [8]. Morrison J, "Statistical Analysis of CFD Solutions from the Fourth AIAA Drag Prediction Workshop," AIAA Paper 2010–4673, 6 2010.
- [9]. Levy DW, Laflin KR, Tinoco EN, Vassberg JC, Mani M, Rider B, Rumsey C, Wahls RA, Morrison JH, Brodersen OP, Crippa S, Mavriplis DJ, and Murayama M, "Summary of Data from the Fifth AIAA Computational Fluid Dynamics Drag Prediction Workshop," *Journal of Aircraft*, Vol. 51, No. 4, 2014, pp. 1194–1213.
- [10]. Morrison J, "Statistical Analysis of the Fifth Drag Prediction Workshop Computational Fluid Dynamics Solutions," *Journal of Aircraft*, Vol. 51, No. 4, 2014, pp. 1214–1222.
- [11]. Rivers M, Hunter C, and Campbell R, "Further Investigation of the Support System Effects and Wing Twist on the NASA Common Research Model," AIAA Paper 2012–3209, 6 2012.
- [12]. Hue D, "Fifth Drag Prediction Workshop: ONERA Investigations with Experimental Wing Twist and Laminarity," *Journal of Aircraft*, Vol. 51, No. 4, 2014, pp. 1311–1322.
- [13]. Roy C, "Summary of Data from the Sixth AIAA CFD Drag Prediction Workshop: Case 1 Code Verification," AIAA Paper to be published, 1 2017.
- [14]. Tinoco EN, Brodersen OP, Keye S, Laflin KR, Feltrop E, Vassberg JC, Mani M, Rider B, Wahls RA, Morrison JH, Hue D, Roy CJ, Mavriplis DJ, and Murayama M, "Summary Data from the

- Sixth AIAA Computational Fluid Dynamics Drag Prediction Workshop: CRM Cases,” *Journal of Aircraft*, Vol. 51, No. 1, 2018, pp. 1–28.
- [15]. U.S. Guide to the Expression of Uncertainty in Measurement, ANSI/NCSL Z540.2–1997, 1997.
- [16]. Hintze J, and Nelson R, “Violin Plots: A Box Plot-Density Trace Synergism,” *The American Statistician*, Vol. 52, 1998, pp. 181–184.
- [17]. McGill R, Tukey J, and Larsen W, “Variations of Box Plots,” *The American Statistician*, Vol. 32, No. 1, 1978, pp. 12–16.
- [18]. Vassberg JC, DeHaan MA, Rivers SM, and Wahls RA, “Development of a Common Research Model for Applied CFD Validation Studies,” AIAA Paper 2008–6919, 8 2008.
- [19]. Rumsey CL, “Langley Research Center Turbulence Modeling Resource,” <https://turbmodels.larc.nasa.gov/>, 1 2018.
- [20]. Wheeler D, *Advanced Topics in Statistical Process Control*, SPC Press, Knoxville, TN, 1995.
- [21]. Derlaga J, and Morrison J, “Statistical Analysis of CFD Solutions from the Sixth AIAA Drag Prediction Workshop,” AIAA Paper 2017–1209, 1 2017.

(a) Sample violin/box-and-whisker plot

(b) Sample scatter plot

Fig. 1.

Comparison of new summary plot with corresponding scatter plot.

Fig. 2.
Characteristic outlier removal based on C_D for Case 2B, L3 grid level.

(a) Case 2A

(b) Case 2B

Fig. 3.
Summary of C_D .

(a) Case 2A

(b) Case 2B

Fig. 4.
Summary of C_{DPR} .

(a) Case 2A

(b) Case 2B

Fig. 5.
Summary of C_{DSF} .

(a) Case 2A

(b) Case 2B

Fig. 6.
Summary of α .

(a) Case 2A

(b) Case 2B

Fig. 7.
Summary of C_m .

Fig. 8.
Summary of ΔC_D 's of drag coefficients.

Fig. 9.
Coefficient of variation of the total drag coefficient.

Fig. 10.
Coefficient of variation of the pressure drag coefficient.

Fig. 11.
Coefficient of variation of the skin friction drag coefficient.

Fig. 12.
Coefficient of variation of the magnitude of the pitching moment coefficient.

Table 1

Comparison of Case 2A medians.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	2.485	2.485	2.481	2.480
C_{D^*} counts	258	257	257	257
C_{DPR^*} counts	146	145	144	144
C_{DSP} counts	112	112	113	113
C_m	-0.0916	-0.0919	-0.0922	-0.0920

Table 2

Comparison of Case 2A standard deviations.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	0.067	0.062	0.056	0.046
C_{D^*} counts	4.5	4.4	3.7	3.7
$C_{D_{PR^*}}$ counts	4.1	4.2	3.5	3.4
$C_{D_{SP}}$ counts	3.3	2.9	2.6	2.5
C_m	0.0058	0.0055	0.0049	0.0041

Table 3

Comparison of Case 2B medians.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	2.639	2.640	2.645	2.640
C_{D^*} counts	281	280	279	279
$C_{D_{PR^*}}$ counts	153	151	150	149
$C_{D_{SP}}$ counts	129	129	129	130
C_m	-0.0869	-0.0866	-0.0862	-0.0864

Table 4

Comparison of Case 2B standard deviations.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	0.071	0.064	0.061	0.049
C_{D^*} counts	5.1	5.0	4.7	4.9
C_{DPR} counts	5.0	4.9	4.5	4.8
C_{DSP} counts	3.7	3.3	3.1	2.8
C_m	0.0068	0.0062	0.0058	0.0049

Table 5

Comparison of medians of differences.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	0.161	0.164	0.162	0.161
C_{D_s} counts	22.9	22.9	22.7	22.4
$C_{D_{PR}}$ counts	5.9	5.6	5.4	5.1
$C_{D_{SP}}$ counts	16.8	16.9	16.8	17.1
C_m	0.0056	0.0058	0.0057	0.0056

Table 6

Comparison of standard deviations of differences.

	Coarse (L2)	Medium (L3)	Fine (L4)	Extra Fine (L5)
α , degrees	0.007	0.007	0.005	0.005
C_D counts	1.39	1.29	1.31	1.71
C_{DPR} counts	1.32	1.29	1.36	1.79
C_{DSP} counts	0.47	0.45	0.44	0.35
C_m	0.0009	0.0007	0.0007	0.0007

Table 7

Total submission counts for Case 2A and Case 2B.

	Case 2A	Case 2B
L1	41	34
L2	40	35
L3	41	35
L4	40	30
L5	31	23
L6	22	13