JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 36, No. 1, January–February 1999

Engineering Notes

ENGINEERING NOTES are short manuscripts describing new developments or important results of a preliminary nature. These Notes cannot exceed 6 manuscript pages and 3 figures; a page of text may be substituted for a figure and vice versa. After informal review by the editors, they may be published within a few months of the date of receipt. Style requirements are the same as for regular contributions (see inside back cover).

Statistical Comparison of Computed Surface Pressure Predictions

Malcolm S. Taylor* and Walter B. Sturek†
U.S. Army Research Laboratory,
Aberdeen Proving Ground, Maryland 21005

Introduction

THIS Note discusses a statistical approach to the evaluation of results from an international study that applied Navier–Stokes techniques to a complex flowfield.[1] The study included 8 computer codes, 9 turbulent viscous models, up to 15 grid resolutions, and 6 test cases, resulting in more than 90 test-case results for evaluation with respect to each other and to the effects of grid resolution, choice of turbulence model, and computational technique.

Experimental measurements of surface pressure were obtained at 41 circumferential positions ϕ, at each of 9 axial stations x/d, for an ogive-cylinder model mounted in a wind tunnel. In Fig. 1, experimental measurements of surface pressure are plotted, along with computed values from Navier–Stokes solutions using four turbulence models, at a single axial location. The computed results show close agreement with the experiment on the windward side ($0 \le \phi \le 90$ deg) and disagreement on the leeward side ($90 < \phi \le 180$ deg). The leeward-side separated flow is a critical region for evaluating the predictive capability of the computational procedures. Traditionally, an engineer evaluates flowfield computations by visual inspection and qualitative comparison, an often satisfactory process for a small number of data sets, but in this study, the number of data sets was so large as to make the evaluation and comparison of the results a formidable task, and the contribution that contemporary data analysis techniques might provide to the evaluation process is explored.

Statistical Approach

The wide availability of powerful and affordable computing resources has impacted the way in which data analysis is conducted at a fundamental level, and expressions such as exploratory data analysis and data visualization have entered the lexicon.[2,3] The basic idea is that data, properly collected and effectively portrayed, may speak directly to the subject-matter expert without resort to extensive statistical methodology.

A set of experimental data serves as a baseline for comparison of the flowfield predictors. The technique involves the comparison of two data sets: the experimental data and a set of computed results. To facilitate discussion, the experimental data are denoted as

$$Y_{ij}, \qquad i = 3.5(1)11.5, \qquad j = 0(4.5)180 \tag{1}$$

where Y_{ij} is the measured surface pressure at axial station i and circumferential position j; the corresponding computed values are denoted as

$$X_{ij}^k, \qquad k = 1(1)9 \tag{2}$$

where subscripts i and j are the same as for the experimental data and the superscript k indexes the data sets chosen for comparison.

If we define the difference between a computed value of surface pressure and the corresponding experimental value as error, then

$$E_{ij}^k = X_{ij}^k - Y_{ij} \tag{3}$$

is the error at location (i, j) for data set k. A contour plot, such as shown in Fig. 2 for data set $k = 1$, provides an informative display of error over the entire axial-circumferential grid. The approximately planar portion of the surface in Fig. 2 corresponds to the windward region, where experimental and computed values are in reasonably good agreement. Substantial irregularities are again seen to occur on the leeward side ($90 < \phi \le 180$ deg).

The highly irregular surface argues for further data reduction before a simultaneous comparison of flowfield calculations can be effectively undertaken. A useful reduction device, box and whisker plots, is shown in Fig. 3. To construct a box and whisker plot for a data set k, the errors E_{ij}^k over the entire grid are first ordered from smallest to largest. The bottom and top of the box mark the first and third quartiles, respectively, of the ranked errors; the whiskers are vertical lines that terminate at the extreme values of the data set; the median, or middle value, is represented by the dot · inside the box. These five-value data summaries provide a compact description of how the errors are distributed over the entire grid and, as can be seen, facilitate simultaneous comparison across data sets.

If the computed and experimental values coincide at every location (i, j), the box and whisker plot will collapse to zero. Failing that, a thin box with short whiskers indicates good overall agreement. Visual inspection of Fig. 3 suggests that data sets $k = 3$ and 4

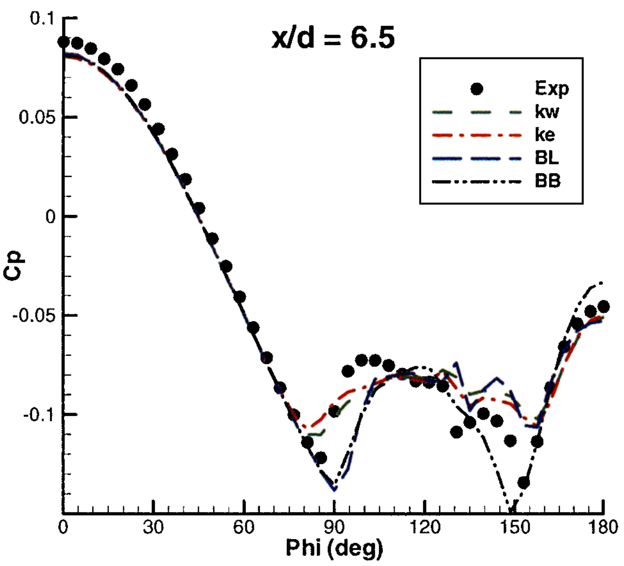

Fig. 1 Surface pressure C_p vs circumferential position ϕ.

Received April 21, 1997; revision received Oct. 14, 1998; accepted for publication Oct. 14, 1998.

*Chief Statistician, Information Science and Technology Directorate; currently Research Associate, Center for Automation Research, University of Maryland, College Park, MD 20742.

†Senior Scientist, High Performance Computing Division, Corporate Information and Computing Center. Associate Fellow AIAA.

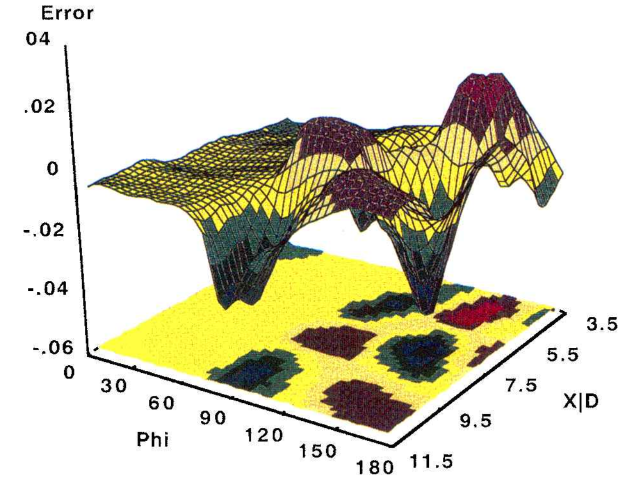

Fig. 2 Contour surface of error for data set k = 1.

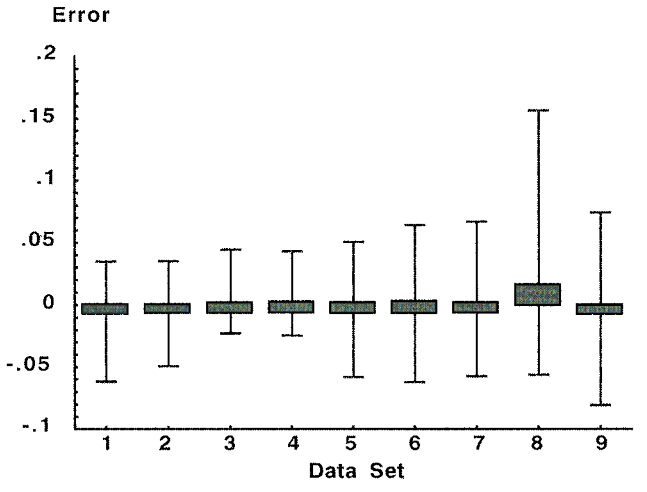

Fig. 3 Box and whisker plots of error for data sets k = 1, . . . , 9.

are closest to this ideal. A spurious observation could distort the whisker length and mislead the viewer, and so the data underlying Fig. 3 were preprocessed for outliers.

The statistical graphics in Figs. 2 and 3, although enlightening, suggestive, and highly appropriate for an initial screening, still do not provide a rigorous assessment of which error sets represent closest agreement between experiment and model. Clearly, we would like the errors to be tightly clustered about zero. Such a distribution of errors would be reflected in a location parameter (a mean, or median) and a dispersion parameter (a standard deviation, or interquartile range) both assuming values close to zero. An attempt to formally rank the effectiveness of the computation procedures as revealed through the error sets should involve a statistic that includes measures of both location and dispersion.

We chose as a statistic the distance between the point determined by the sample mean and standard deviation, denoted as $(\bar{x}_k, s_k)$, and the origin. Recall that the origin (0, 0) corresponds to perfect agreement between calculation and experiment, so that the closer $(\bar{x}_k, s_k)$ is to the origin, the better. The distances $\sqrt{(\bar{x}_k^2 + s_k{}^2)}$, $k = 1, 2, \ldots, 9$, were determined, and the ranking induced under this procedure in terms of the index k is as follows:

$$4 \prec 3 \prec 2 \prec 7 \prec 5 \prec 6 \prec 1 \prec 9 \prec 8 \tag{4}$$

Table 1 Distance measurements for $(\bar{x}_k, s_k)$

Model:	B–L	k–ω	k–ϵ	B–B
Distance	0.01125	0.01131	0.01149	0.01257

Relation (4) asserts that the set of computed values with index $k = 4$ is in closest agreement with the experimental measurements, $k = 3$ is second closest, etc. More complicated ranking procedures are possible, but to engage in them without a compelling reason serves no practical purpose and was not undertaken.

The results of the ranking procedure are entirely consistent with the preliminary graphics. For the statistician, failure to account for error in the measurements Y_{ij} is troublesome, but engineering experience leaves the fluid dynamicist convinced that concern over this point is unwarranted. In this study, only a single set of experimental data is available, and all of the flowfield predictors are deterministic, and so the measurement error, whatever its magnitude, remains confounded with the recorded observations.

Application to Turbulence Models

The statistical approach was applied to data output from four turbulence models: the Baldwin–Lomax (B–L), Baldwin–Barth (B–B), k-omega (k–ω), and k-epsilon (k–ϵ) (Ref. 4). We have already seen in Fig. 1 that visual comparison between experiment and computation over the four models, even when limited to a single axial location, is exceedingly difficult. Box and whisker plots suggested that the B–L model agreed most closely with the experimental data. The mean errors were close to zero for all four data sets, with little apparent advantage of one turbulence model over another. The computed values of the distance statistic are shown in Table 1, where the formal ranking is B–L $\prec$ k–ω $\prec$ k–ϵ $\prec$ B–B.

The seemingly insignificant numerical differences in Table 1 hold substantial practical significance for the engineer. The four turbulence models vary widely in complexity and computer resource requirements. The most computer-efficient model is the B–L model, followed by the B–B model. Models k–ω and k–ϵ are more complex two-equation models that require significantly greater computer resources than either the B–L or B–B model.

Conclusion

A statistical approach, with emphasis on data visualization, was used to assist the engineer in the comparison of Navier–Stokes predictors. The method was applied to a set of computational results obtained from four turbulence models and showed that the use of more computer-intensive models did not provide correspondingly superior results. The statistical procedure makes few assumptions and should be of particular value when a large number of data sets are to be compared and when an impartial quantitative assessment is sought.

References

[1]Sturek, W. B., Birch, T., Lauzon, M., Housh, C., Manter, J., Josyula, E., and Soni, B., "The Application of CFD to the Prediction of Missile Body Vortices," AIAA Paper 97-0637, Jan. 1997.

[2]Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A., *Graphical Methods for Data Analysis*, Wadsworth, Pacific Grove, CA, 1983, pp. 21–24.

[3]Tufte, E. R., *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT, 1983, pp. 123–137.

[4]Taylor, M. S., and Sturek, W. B., "Statistical Analysis of Surface Pressure Measurements vs. Computational Predictions," U.S. Army Research Lab., Rept. ARL-TR-1318, Aberdeen Proving Ground, MD, March 1997.

R. M. Cummings
Associate Editor

Color reproductions courtesy of U.S. Army Research Laboratory