AIAA SciTech
13-17 January 2014, National Harbor, Maryland
16th AIAA Non-Deterministic Approaches Conference

A Pre-Validation Study on Supersonic Wind Tunnel Data Collected from Legacy Aerothermal Experiments

Michael Scott Balch*and Benjamin P. Smarslok†

Before wind-tunnel data can be used to validate a model, it is necessary to confirm that the auxiliary models being used to interpret the data are also accurate. This paper considers the use of existing hypersonic wind tunnel data to validate piston theory. In NASA Technical Paper 2631 by Glass and Hunt in 1986, there is an apparent discrepancy between the static pressure ratio predicted by oblique shock theory and the pressure ratio actually measured. The only explanation conclusively eliminated in the current study is that the apparent bias is due to random error. Furthermore, positive evidence for the existence of a non-negligible edge vortex is established. Development of an efficient edge vortex model is required for further investigation into these data.

I. Introduction

One of the greatest challenges to the development of scramjet-based hypersonic vehicles is accurately modeling the aerothermoelastic response. The interaction of hypersonic aerodynamics, structural deformation, and heat transfer pose multiple analysis challenges.[1] The computational requirements are too large for a coupled, direct computational fluid dynamics and finite element simulation approach, and one cannot simply factor-of-safety out of the problem; the mass constraints to meet mission requirements are too tight. Reduced-order models are needed, and the errors in those models need to be quantified using validation data.

Recent studies by AFRL and its affiliates have focused on a reduced-order model approach to panel flutter (one of the structural challenges of sustained hypersonic flight), using piston theory as the basis for pressure predictions. This group has even pursued validation of piston theory using published wind tunnel data originally obtained for other purposes.[1–3] Their work made use of a 1986 study conducted by Glass and Hunt in NASA's 8' High-Temperature Tunnel (HTT). While the study by Smarslok and Mahadevan provides a framework for validation work, it lacks an essential step in the use of historical data for validation purposes.

No reduced-order model operates on its own. The boundary conditions and inputs fed into said model are derived from predictions produced by other reduced-order models. For example, in validating piston theory using the Glass and Hunt 1986 data, one must establish the "undisturbed" post-shock flow on which the instrumented dome impinges. In the study by Smarslok and Mahadevan,[1] these values are produced using 2D oblique shock theory. If 2D oblique shock theory is not appropriate, or if there are biases in the reported data, then inferences about the validity and accuracy of piston theory drawn from the Glass and Hunt 1986 data will be wrong. It is to avoid this type of error that a pre-validation study is conducted and reported here.

II. Legacy Aerothermal Data from the NASA High-Temperature Tunnel

In 1986, Glass and Hunt studied the flow over a shallow dome protuberance on a flat plate in nominally Mach 6.5 flow to investigate the thermal and structural loads on body panels in extreme environments.[4] The experiments in this study were conducted in the NASA Langley 8' HTT and provide rare data relevant to the validation of piston theory. Originally, these data were intended to shed light on the flow over shuttle tiles during reentry.

*Member of the Technical Staff at Arctan, Inc. Alexandria, VA 22314
†Research Aerospace Engineer, Air Force Research Laboratory, Wright-Patterson AFB, OH 45433

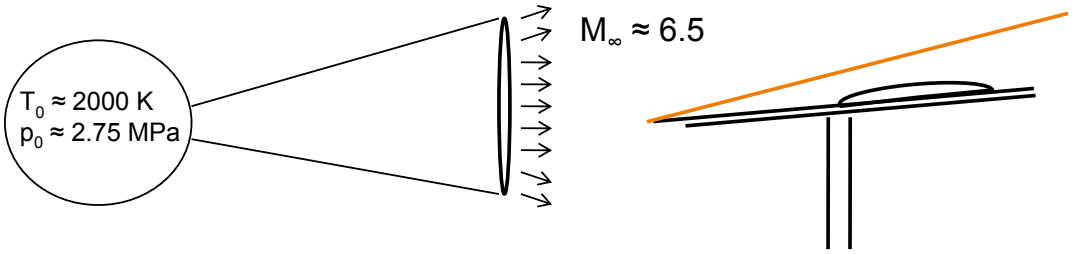

Figure 1. Sketch of Apparatus from Glass and Hunt 1986 HTT Experiments

Domes instrumented with pressure taps and thermocouples were embedded in spherical dome specimens on a flat plate held at an inward five degree incline with respect to the flow. So, between the flow and the dome plate, there was an oblique shock. The inviscid flow region was still supersonic ($M \approx 5.7$) and uniform post-shock until encountering the spherical dome protuberances. Figure 1 illustrates the structure of the flow studied. There were pressure data for 30 experiments at 57 locations on the dome.

Given the simplicity of the flow and the ample data taken, the Glass and Hunt 1986 study would seem an ideal candidate for use in the validation of piston theory. In theory, all one should need to do is compute the post-shock flow conditions and use these as the "undisturbed" flow conditions input into piston theory. The deflections induced by the dome are simple and easy to compute. Therefore, determining the discrepancy between piston theory prediction and real pressures should be a relatively simple task.

However, an initial review of the data reveals that the ratio of the measured post-shock static pressure to the measured freestream pressure does not match the static pressure ratio obtained using oblique shock theory. The Mach numbers reported range between 6.53 to 6.62. The reported deflection angle is five degrees. These conditions should produce a weak oblique shocks with pressure ratios ranging from 2.11 to 2.13. However, the ratios between the measured flat plat pressures and the reported freestream pressures range from 1.85 to 2.00. Something is wrong: a bias, either in the data or in its theoretical interpretation.

The pressure discrepancy is important in that it throws the data, the interpretation of the data, or both into question. If the post-shock pressure predicted by oblique shock theory is wrong, then so is the post-shock Mach number. That Mach number is a key input into the predictions of piston theory. If the pressure measurements are biased, then it throws the pre-shock and post-shock Mach numbers into question. Finally, if the discrepancy is due to 3D flow effects, those same effects will impact the dome pressure measurements. Whether the measurements are in error or the theoretical treatment of the data is in error, there is an error, and that error will affect any comparison between the measurements obtained by Glass and Hunt and the predictions of piston theory. The identification of that error is the focus of this paper and constitutes the "pre-validation study" alluded to in the title.

There are many potential physical explanations for the pressure discrepancy. Two simple explanations are explored here. The first explanation is a bias in the reported freestream pressures. A cursory investigation showed that a 10% bias in freestream pressure could readily account for the discrepancy in pressure ratios. A 1973 study of the facility (NASA 8' HTT) showed that the pressure profile across the test flow does, in fact, vary on the order of 10%. This may be due to the conical nature of the wind tunnel flow. So, if the freestream pressure reported was drawn from a different part of the flow than that encountering the test bed, that would certainly explain the bias. Alternatively, if the deflection (or effective deflection) of the flow were off by half a degree from the nominal five degrees, that too could account for the observed pressure ratio discrepancy.[a]

Another important potential explanation for the pressure ratio discrepancy is pressure relief via a strong edge vortex. (Actually, it would be a pair of edge vortices, one on each side.) This possibility is difficult

[a] A third conceivable explanation is that the ratio specific heats differs from 1.4. However, a cursory exploration demonstrated that the pressure ratio was somewhat insensitive to γ and, at a Mach number of 6.55, $\gamma \leq 1.2$ would be necessary to explain the observed pressure ratios across the shock. Even if the flow were thermodynamically frozen from the stagnation chamber to the test section, the stagnation temperature was less than 2000 Kelvin in almost all of the runs, accounting for a ratio of specific heats no lower than 1.27, still too high. Although not absolutely impossible, this explanation for the pressure ratio discrepancy was dismissed as physically implausible.

to explore quantitatively because the authors are unaware of a computationally efficient edge vortex model. In fact, even CFD-based approaches to supersonic vortex modeling are known to suffer from substantial numerical dissipation. Without a concrete model structure for the edge vortex, it is not possible to statistically test this potential "bias" source.

Finally, pressure measurements on the dome itself are highly suggestive of an edge vortex. In the absence of an edge vortex, pressure taps that reflect each other along the dome diameter aligned with the freestream flow *should* have equal pressure measurements. To be sure, there will be some discrepancy attributable to random error, but at first glance, the scale and consistency of that bias would seem to indicate an edge vortex. The evidence for such an edge vortex is explored on that basis in this study.

III. Mathematical Methods

A. Posterior p-Value

Assessing the hypothetical sources of pressure ratio bias presents a difficult problem. Given the provisional nature of these hypotheses, a p-value approach seems appropriate. However, these are not simple hypotheses corresponding to a physical model with a set of parameters. The goal here is to test the assertion that there is a bias in the freestream pressure measurements or in the deflection angle, without committing to the magnitude of that bias. So, it is a test of model form, rather than a test of parameter values.

Readers unfamiliar with traditional p-values may benefit from a quick review. In short, the "p" in "p-value" stands for plausibility. The p-value accorded to a hypothesis reflects its plausibility in light of the data. Given a test statistic, the p-value for a hypothesis is the probability of having gotten a less favorable test score given that the hypothesis is true.

That is, if T is the test-statistic, and small values of $T(x, H)$ are favorable to hypothesis H, then the p-value for H is as follows:[5]

$$\mathrm{pls}(H|\mathbf{x}) = \mathrm{Pro}_{X'|H}\left(T(\mathbf{x}, H) \leq T(\mathbf{X}', H)\right).$$

If large values of $T(x)$ are favorable to hypothesis H, then the p-value of H is as follows:

$$\mathrm{pls}(H|\mathbf{x}) = \mathrm{Pro}_{X'|H}\left(T(\mathbf{x}, H) \geq T(\mathbf{X}', H)\right).$$

Given a test-statistic, $T(x)$, the user determines what kinds of values are considered favorable or unfavorable. That might seem arbitrary, but a directional preference usually suggests itself. For example, suppose μ were the hypothetical mean of X, and

$$T(\mathbf{x}, \mu) = \sum_{k=1}^{n} (x_k - \mu)^2.$$

Here, T represents a discrepancy, and small values are favorable to the hypothesis in question. Alternatively, suppose instead that T were the following likelihood ratio:

$$T(x, \theta) = \frac{f(\mathbf{x}|\theta)}{max_{\theta'} f(\mathbf{x}|\theta')}$$

where $f(\mathbf{x}|\theta)$ is the probability of $\mathbf{x}$ given θ. In this second example, larger values of T are more favorable to the hypothesized value of θ. While not all test-statistics are as simple as measuring goodness or badness of fit, their directional interpretations tend to remain intuitive, simply because test statistics are *designed* to support a simple directional interpretation.

When dealing with a precise hypothesis, computing a p-value is relatively straightforward. If nothing else, one can generate Monte-Carlo replicates under the assumption that said hypothesis is true and record the resulting Monte-Carlo sample of the test statistic. The p-value is the number of test statistic replicates of the test statistic that are worse than the one obtained with the real data, divided by the total number of Monte-Carlo replicates.

To deal with the compound problem of simultaneously tuning the model parameters *and* assessing the plausibility of the model form, only a slight variant is needed. First, tune the parameters to yield either a Bayesian posterior or confidence structure, assuming that the hypothesized model form is correct. Compute the test statistic for the observation and tuned parameter values. Next, draw a Monte-Carlo sample of the

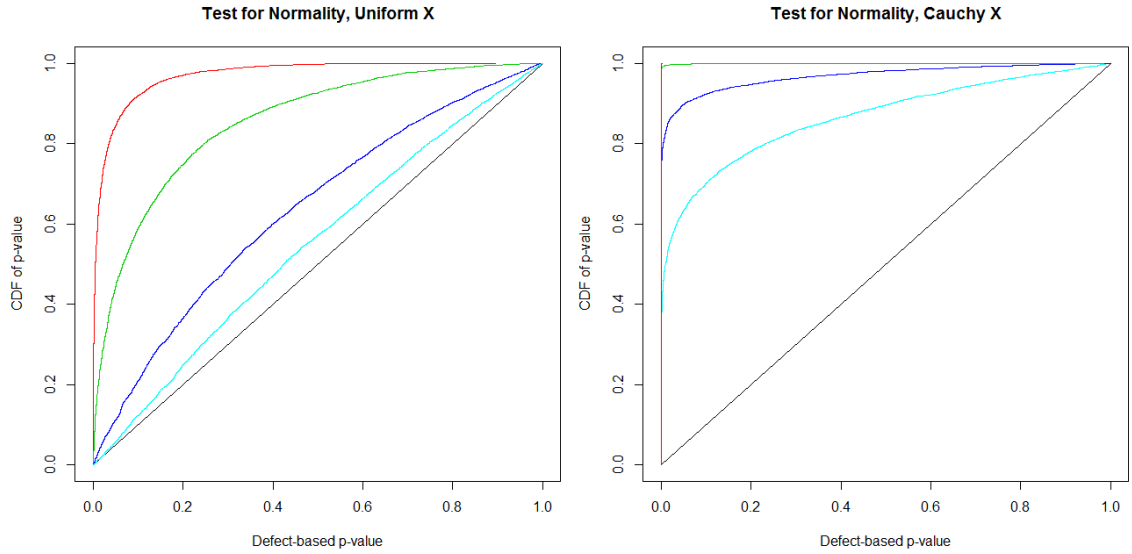

Figure 2. Falsification Power of Posterior p-Value Approach for Various Sample Sizes (Light Blue = 10, Dark Blue = 20, Green = 50, Red = 100)

parameters from the posterior. For each, generate a random value set of the data, x'; re-infer the parameters as though x' were the original observation; and recompute the test statistic value. From here, the calculation of the p-value is the same as for a precise hypothesis. However, instead of testing the hypothesis

$$x \sim f\left(x|\theta\right)$$

for some specified θ, the posterior p-value will test the hypothesis that

$$\exists\theta \;\text{ s.t. }\; x \sim f\left(x|\theta\right).$$

Figure 2 demonstrates the power of this approach to falsify a normality hypothesis (i.e. that a sample has a normal or Gaussian distribution) when the data are in fact uniform or Cauchy. The test statistic used is explained in Section C.

P-values and posterior p-values provide a simple plausibilitic interpretation for statistical results. A high p-value reflects that the hypothesis in question is still plausible (i.e. not yet falsified) in light of the data, and a low p-value reflects the hypothesis in question is now implausible (i.e. falsified) in light of the data. At the risk of redundancy, it should be stressed that a high p-value alone does not prove a hypothesis true. In fact, given a set of data, multiple competing hypotheses may be plausible. It is only when one has eliminated all reasonable alternatives that a given hypothesis may be taken as confirmed by the data.[b]

B. Metropolis-Hastings Algorithm

Bayesian inference was accomplished in this study using a slight variant on the Metropolis-Hastings algorithm. Metropolis-Hastings does not perform Bayesian inference exactly; it yields a Monte-Carlo sample distributed according to a specified likelihood function: in this case, the posterior. The variant used here involves working with iteration on a large sample, rather than selection from a recorded Markov Chain. What this variant lacks in computational efficiency, it makes up for in simplicity and robustness.

The process is simple. Start with a seed sample, $\mathbf{x}_0$, of desired size. A size of 10,000 is sufficient for most Monte-Carlo analyses. Next, iterate as follows:

[b]Philosophically speaking, the world supply of alternative hypotheses is inexhaustible. In epistemology, this is known as the problem of underdetermination. Practically speaking, however, engineers and scientists can sometimes be trusted to specify a finite set of alternative hypotheses, and treat those as provisionally exhaustive.[6] Moreover, p-values and significance tests are often used, not as justification for taking a hypothesis as true, but rather as a sanity check on the assumptions to which the analyst was already inclined.[7]

1. Add a random perturbation, $\mathbf{e}_k$, to current sample, $\mathbf{x}_k$. Record the result, $\mathbf{x}' = \mathbf{x}_k + \mathbf{e}_k$. The perturbation, $\mathbf{e}_k$, must be an iid sample of a distribution that is symmetric about zero.

2. Take the ratio of likelihoods, $\mathbf{R} = \frac{\lambda(\mathbf{x}')}{\lambda(\mathbf{x}_k)}$

3. Take a random selection seed, $\mathbf{u}$, uniformly distributed on $[0, 1]$.

4. Assign members of the current sample and the perturbed sample to the new sample according to whether the ratio of a member of the sample is greater than the selection seed. That is, $x_{k+1,i} = [[R_i > u_i]]\, x'_i + [[R_i \leq u_i]]\, x_{k,i}$, where $[[]]$ returns a Boolean truth value of zero or one. This formula results in a probability of R_i or 1 (whichever is less) of replacing $x_{k,i}$ with x'_i.

Continue this process until the empirical CDF of the Monte-Carlo sample has stopped making significant changes.

When using the algorithm outlined above, the user should be mindful of the following details:

- The perturbation distribution should be symmetric or near-symmetric. A normal distribution with zero mean is a safe option, so long as non-physical samples (if any occur) are rejected.
- It helps to have the perturbation distribution be on the scale of the target distribution. The user may or may not be able to guess what that scale is.
- Since the likelihood ratio drives the selection algorithm, the user does not need to know the normalization coefficient of the distribution you are sampling. This is one of the most useful features of Metropolis-Hastings.
- Another major advantage of Metropolis-Hastings over other methods is its applicability to multidimensional problems. There is nothing, in principle, to distinguish a likelihood over one variable from a likelihood over several.
- *However*, the larger the inference problem (in terms of dimensionality), the more iterations it takes to converge Metropolis-Hastings.[c]

A short Matlab example may be instructive. Suppose the analyst wants to sample from a gamma distribution with scale parameter = 1 and shape parameter = $\frac{1}{2}$. The probability density function for this variable is as follows:

$$f(x) = x^{-\frac{1}{2}} e^{-x} \frac{1}{\sqrt{\pi}}$$

The analyst could map from a uniform random sample to this distribution by mapping with the inverse of the incomplete gamma function. Or, alternatively, the analyst could generate a sample more easily using the Metropolis-Hastings algorithm.

Let N be the desired size of the sample. The following Matlab script will generate an iid sample of size N from the gamma distribution with unity scale and shape of one half:

```
x0 = - sqrt(2) * log( 1 - rand(N,1) );
% initial guess from exponential distr
x = x0;
for k = 1:100
      x_pr = x + sqrt(2) * randn( N , 1 ); % normal perturbation
      R = exp( x - x_pr ) .* sqrt( x ./ x_pr ) .* ( x_pr >= 0 );

      u_slct = rand( N , 1 ); % uniform sample for selection
      x = x_pr .* ( R >= u_slct ) + x .* ( R < u_slct );
            % final selection of new xs
end
```

[c] It is not surprising that Monte-Carlo *inference* would be sensitive to dimensionality, but it is a disappointment, given the insensitivity of Monte-Carlo uncertainty *propagation* to dimensionality.

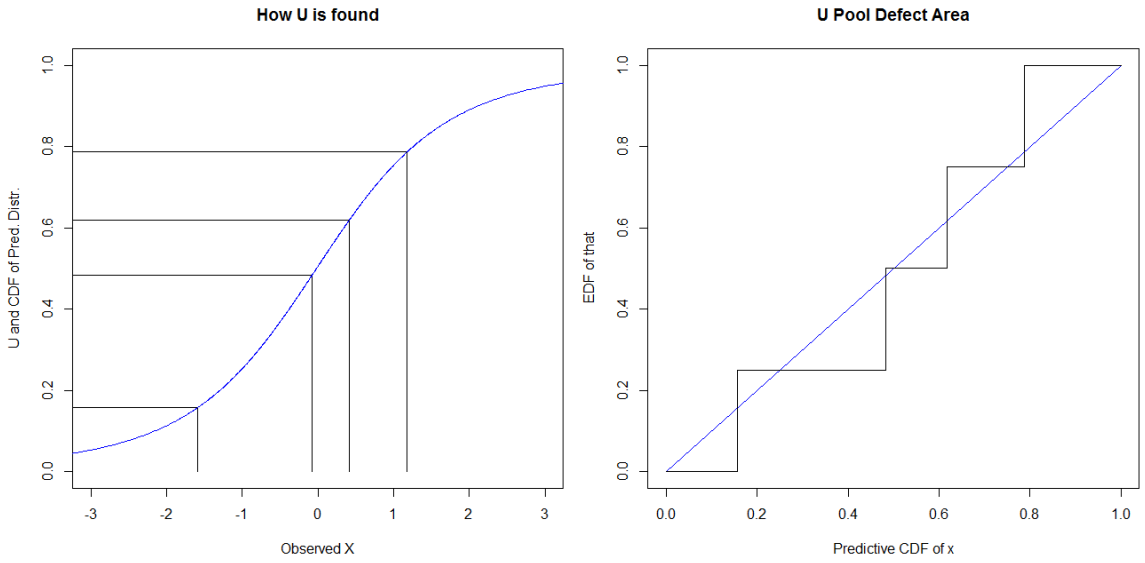

Figure 3. Calculation of the U-Pool Defect Statistic

C. U-pool Defect

The U-pool defect is a simple and widely applicable test statistic. It was originally suggested by Ferson and Oberkampf in their response to the Sandia Challenge Problem and makes use of the uniformity of CDF values.[8] Suppose one has a sample, x, and it is assumed that there is some value of θ such that $x \sim f(x|\theta)$. Then, for some future sample, x', the CDF of the predictive distribution based on x should be uniformly distributed. That is $F(x'|x)$ should be an iid uniform sample. The U-pool defect is the area of the absolute difference between the EDF of $F(x'|x)$ and the line $x = y$ over $[0, 1]$.

In this study, a slight variant of the U-pool defect is used; this variant could be called the Auto-U-pool defect. Instead of taking a second sample, x', x is simply fed back in as though it were a new sample. So, strictly speaking, $F(x'|x)$ with $x' = x$, won't be iid uniform, but as long as the sample size is sufficiently larger than the number of parameters, it will be close enough. Moreover, for small sample sizes, Auto-U-pool defect values tend to be smaller than they would be for an actual iid uniform sample; so, it will tend to give high p-values, i.e. conservative inference.

D. Confidence Structure on the Non-Parametric Difference

Confidence structures are a simple and robust alternative to Bayesian inference.[9] There are three differences between a confidence structure and a Bayesian posterior. First, a confidence structure does not require a prior distribution. Second, while a Bayesian posterior's belief values usually have a subjective interpretation, a confidence structure's belief values have a statistical confidence (i.e. coverage probability) interpretation. Finally, while Bayesian belief values satisfy the Kolmogorov axioms, confidence structures only satisfy the more general Shafer axioms. This generality enables the application of confidence structures to problems of non-parametric statistics, which is useful in this study.

The Mann-Whitney U-statistic is the foundation of a well-known test for the equality of two distributions given no assumptions about the form of those distributions. Moreover, under the assumption that the two distributions considered are identical except for an offset, Mann-Whitney U is a pivot. That is, $U(x - \delta, y)$, where δ is the offset between the distributions of x and y, has a known distribution. The pivot is the key! It has previously been demonstrated that a pivot can be used to readily construct a confidence structure.[9] The confidence structure developed from the Mann-Whitney U-statistic is referred to in this work as the "Non-Parametric Difference."

Figure 4 illustrates how a pair of samples translates to a non-parametric difference. First, one takes all $n_1 \times n_2$ possible values of $x_1 - x_2$ and sorts them in ascending order. This set of points partition of the real

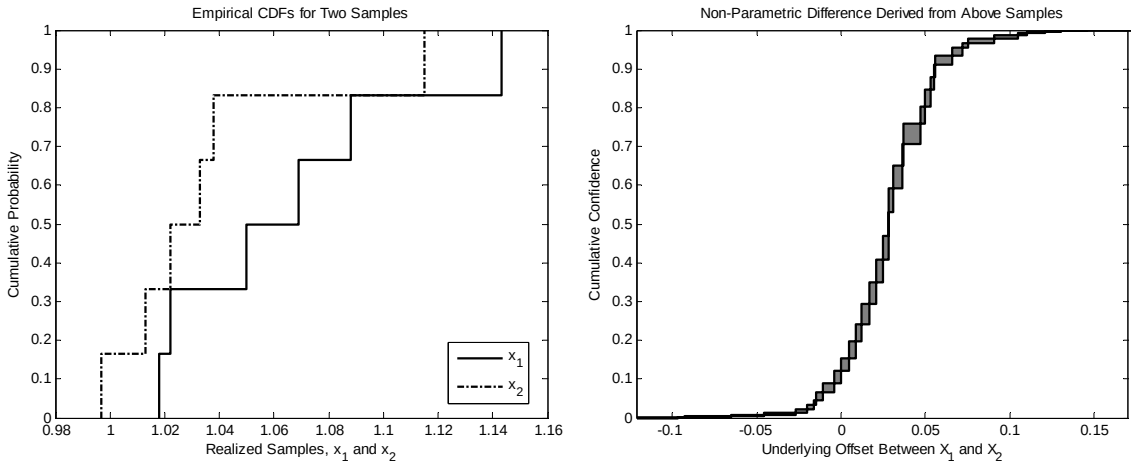

Figure 4. Example of Non-Parametric Difference Confidence Structure

line. This partition is the set of focal elements for the non-parametric difference. Numbering these focal elements, A_k, from 0 to $n_1 \times n_2$, the weight assigned to A_k is equal to the probability that $U(x - \delta, y) = k$, where δ is the true unknown offset.

IV. Data and Models Used

Three of the four potential explanations for the freestream bias are explored using the posterior p-value method described in Section A. Each of these hypotheses leaves several free parameters. First, it is assumed that, even in the absence of bias, there is still a random error in the measurements of the pre-shock and post-shock pressures. The random error is assumed to be normal (i.e. Gaussian). However, the standard deviation of these normal errors are not known. So, in addition to whatever bias term is inferred, the true pre-shock pressure in each flow is inferred, as are the standard deviations of the errors in the pre-shock and post-shock pressure measurements. The following subsection describe the exact mathematical models used to support these inferences.

A. Pre-shock conditions

For the analyses in Sections B-C, it is assumed that there is some error (bias, random, or both) in the reported pressures. As mentioned above, the true unknown pre-shock pressure is inferred via Bayesian inference. To the authors' knowledge, there is no such thing as a pure Mach number measurement device. Moreover, there is no record of direct velocity measurement (e.g. via laser doppler anemometry) in the Glass and Hunt 1986 report. So, it is reasonable to assume that the Mach numbers reported in Table Two of Glass and Hunt[4] were based on the pre-shock pressure measurements, as described in the 1973 wind tunnel calibration report.[10] However, if the reported pressure has an error, so does the reported Mach number.

Inferred values of the true pre-shock pressure imply true pre-shock Mach numbers that are (slightly) different from the reported Mach number. It is assumed that the reported pre-shock Mach number, M_∞, corresponds correctly to the reported pre-shock pressure value, p_∞. The corrected Mach number is adjusted to the inferred pressure on the assumption that the entropy is near equal between reported and corrected conditions. This leads to the following transformation:

$$M_{\infty_T} = \sqrt{\left(\frac{2}{\gamma - 1}\right)\left(\left(1 + \frac{\gamma - 1}{2} M_\infty^2\right)\left(\frac{p_\infty}{p_{\infty_T}}\right)^{\frac{\gamma-1}{\gamma}} - 1\right)}$$

where γ is the ratio of specific heats (approximately equal to 1.4 at test section conditions), p_{∞_T} is the inferred true pre-shock pressure value, and M_{∞_T} is the resulting corrected pre-shock Mach number. Using

this information, the true value of the post-shock (i.e. flat plate) pressure can be calculated as well, according the bias models described in the next three subsections.

In the bias models explored in Sections B-D, it is assumed that the 2D oblique shock relations apply and that either random error, bias in the freestream measurements, or bias in the deflection were responsible. Under these assumptions, the relationship between the true pre-shock (p_{∞_T}) and true post-shock (p_{fp_T}) pressures is given as follows:

$$\frac{p_{fp_T}}{p_{\infty_T}} = 1 + 2\frac{\gamma}{\gamma+1}\left(\left(M_{\infty_T}\sin\beta\right)^2 - 1\right) \tag{1}$$

where β is the shock angle, calculated as follows:

$$\tan\beta = \frac{M_{\infty_T}^2 - 1 + 2\lambda\cos\left(\frac{4\pi+\arccos\chi}{3}\right)}{3\tan\theta\left(1+\frac{\gamma-1}{2}M_{\infty_T}^2\right)}$$

where θ is the deflection angle (i.e. the angle by which the flow is turned inward) and λ and χ are parameters defined as follows:

$$\lambda = \sqrt{\left(M_{\infty_T}^2-1\right)^2 - 3\left(1+\frac{\gamma-1}{2}M_{\infty_T}^2\right)\left(1+\frac{\gamma+1}{2}M_{\infty_T}^2\right)\tan^2\theta}$$

and

$$\chi = \frac{\left(M_{\infty_T}^2-1\right)^3 - 9\left(1+\frac{\gamma-1}{2}M_{\infty_T}^2\right)\left(1+\frac{\gamma-1}{2}M_{\infty_T}^2+\frac{\gamma+1}{4}M_{\infty_T}^4\right)\tan^2\theta}{\lambda^3}.$$

This explicit formulation for the oblique shock relations is available in a commonly used text on compressible aerodynamics[11][p.143].

B. The no-bias hypothesis

Under the "no bias" hypothesis, it is assumed that there is no bias in the measurement of either the pre-shock or post-shock pressure, i.e. that any random is error. Under this hypothesis, the observed bias between the predicted and observed pressure ratios across the shock is explained as a random fluke.

Given sample values for the the true values of pre-shock and post-shock pressures, it is possible to calculate the posterior likelihood of said sample. This is a necessary step in executing Bayesian inference to tune the parameters in the different bias models. In each, a flat prior for the true pre-shock pressures is used. Moreover, a $\frac{1}{\sigma}$ prior is used for the two unknown standard deviations. Given the normality assumptions described above, the likelihood can be computed as follows:

$$f\left(\{p_{fp_T}\},\{p_{\infty_T}\}\,|\,\{p_{fp}\},\{p_\infty\}\right) = \frac{\exp\left[-\frac{1}{2}\sum_{k=1}^{n}\left(\frac{p_{fp_k}-p_{fp_{T_k}}}{\sigma_{fp}}\right)^2 - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{p_{\infty_k}-p_{\infty_{T_k}}}{\sigma_\infty}\right)^2\right]}{(2\pi)^{\frac{n}{2}}\left(\sigma_{fp}\sigma_\infty\right)^{n+1}} \tag{2}$$

where k indexes the runs, σ_{fp} is the standard deviation of the random error in the post-shock (flat-plate) pressure measurements, and σ_∞ is the standard deviation of the random error in the pre-shock pressure measurements. Both of these standard deviations are unknown; so, $n+2$ parameters are being inferred. There are $2n$ pressure observations (one on each side of the shock for each run); so, as long as more than two runs are used, the problem is closed (i.e. determinate). It is this closure that allows the use of non-informative priors.

C. Freestream bias hypothesis

Under the "freestream bias" hypothesis, it is assumed that there is a proportional bias in the measurements of freestream pressure, in addition to random error. The resulting likelihood model is as follows:

$$f\left(\{p_{fp_T}\},\{p_{\infty_T}\}\,|\,\{p_{fp}\},\{p_\infty\}\right) = \frac{\exp\left[-\frac{1}{2}\sum_{k=1}^{n}\left(\frac{p_{fp_k}-p_{fp_{T_k}}}{\sigma_{fp}}\right)^2 - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{p_{\infty_k}-Kp_{\infty_{T_k}}}{\sigma_\infty}\right)^2\right]}{(2\pi)^{\frac{n}{2}}\left(\sigma_{fp}\sigma_\infty\right)^{n+1}},$$

where K is the proportional bias, which is an additional inferred parameter. Interestingly, this bias is equivalent to having a proportional bias in the measured total (i.e. chamber) pressure, meaning that this model covers two (indistinguishable, given the data) possible bias explanations.

D. Deflection bias hypothesis

Under the "deflection bias" hypothesis, it is assumed that the deflection is other than the reported nominal 5 degrees. The pressure values and likelihoods are related as expressed as in Equations 1-2, the deflection angle is merely different. This difference, $\delta\theta$, is the tunable bias parameter.

E. Dome pressure asymmetry

If the pressure bias is due to relief from an edge vortex, this edge vortex will also cause pressure to decrease from the center towards edges. Three cases were examined: 7in, 14in, 28in. The relative position of the pressure dome in these three cases is illustrated in Figure 5. Fortunately, in the 7in and 14in runs, the pressure dome is located off-center. Moreover, on the front half of the pressure dome, each tap has another tap reflected across the centerline oriented with the freestream flow direction. A span-wise decrease in pressure caused by an edge vortex would show up as an asymmetry in the pressures between paired pressure taps. Conversely, in the absence of an edge vortex, the paired pressure taps should produce measurements that are, on average, equal.

The asymmetry of paired dome pressures was explored using the non-parametric difference described in Section D. It is assumed that each pair of taps has the same random error distribution. It is further assumed that the underlying offset between a given pair of pressure taps has a constant value throughout all runs, obscured only by random experimental error. The non-parametric difference may be taken as a direct measure of the asymmetry in each pair.

P-values derived from the Mann-Whitney U-test also provide a holistic assessment of asymmetry across pressure dome. These p-values were split according to dome diameter (7in, 14in, or 28in). If the underlying dome pressures were symmetric, then the Mann-Whitney p-values should be uniformly distributed. The U-pool defect of Section C was then used to assess whether or not they were, and whether their deviation from uniformity was statistically significant.

V. Results

A. Pressure ratio bias results

The methods applied in this study were chosen for their robustness, not their computational efficiency. As such, it was only possible to generate 100 posterior bootstrap samples. Moreover, the limited samples still provided estimates of the p-values for the different hypotheses. Those results are as follow:

- No bias hypothesis: p-value = 0 out of 100 (true pls $< 3\%$ with 95% confidence)
- Freestream bias hypothesis: p-value = 24 out of 100 (true pls btwn 16% and 34% with 95% conf)
- Deflection bias hypothesis: p-value = 15 out of 100 (true pls btwn 9% and 24% with 95% conf)

None of the three hypotheses examined using the posterior p-value approach merited a high plausibility. However, only the "no bias" hypothesis was conclusively falsified. It bears noting that the mismatch statistic (i.e. the U-pool defect) value obtained for the "bias" case is so much larger than its bootstrap counter-parts that using any reasonable curve-fit to those bootstrap samples yields a machine-zero p-value for the "no bias" hypothesis. That is to say, the p-value of the no-bias hypothesis, when approximated using a curve fit to the Monte Carlo statistic sample values, is smaller than the computer is able to represent.

B. Pressure tap asymmetry

Using the aggregated Mann-Whitney p-values described at the end of Section E, it was found that:

- In the 7in case, symmetry is 43.5% plausible.
- In the 14in case, symmetry is 0.67% plausible
- In the 28in case, symmetry is 30.7% plausible.

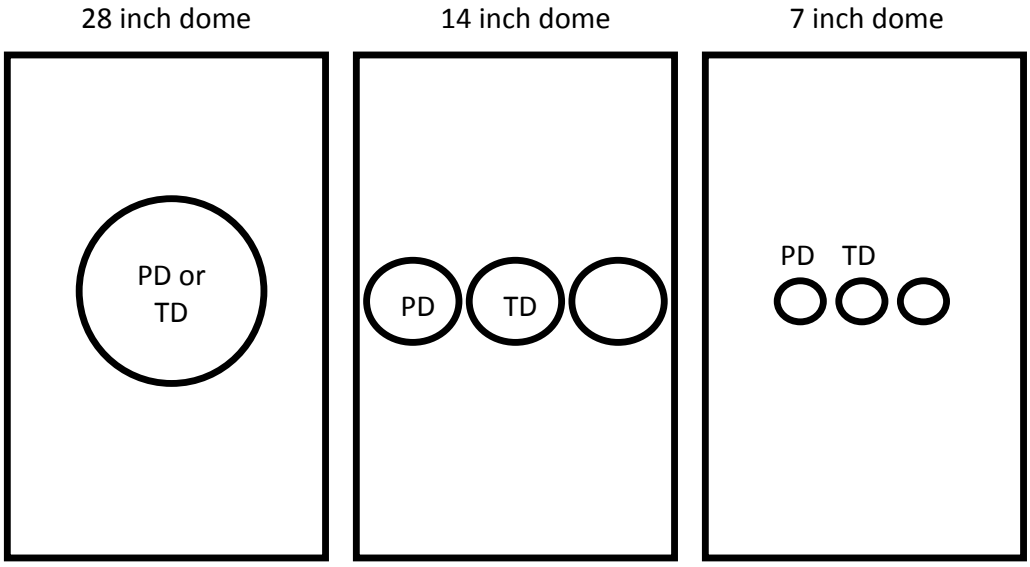

Figure 5. Variations of Dome Size and Position

Symmetry is statistically implausible for the 14 inch case, where the pressure dome is next to the edge, statistically plausible for the 28 inch dome case, in which the dome is centered, and plausible for the 7 inch dome case in which the dome is not centered, but also not close to the edge. These results are consistent with the existence of a significant edge vortex.

Figure 6 illustrates the level of asymmetry at three point-pairs in the 14-inch case. The asymmetry is stronger for points more widely separated. This too is consistent with the existence of a significant edge vortex. The front-most point-pair has a backwards (negative) offset relative to the other two pairs. It is unknown whether the effect is due to a real underlying physical phenomenon (e.g. shock-vortex interaction) or to random measurement error.

VI. Conclusions and Future Work

This paper presented a pre-validation study of legacy aerodynamic data obtained in 1986 by Glass and Hunt in the NASA 8' HTT. An apparent bias in the pressure data confounds the direct application of these data to the validation of piston theory. Advanced statistical methods were applied to explore this bias.

It is unclear how the observed pressure bias in the Glass and Hunt data originates, but it is clear that the bias is the result of unaccounted physics or apparatus bias, and not an artifact of random coincidence. Furthermore, there is positive evidence for an edge vortex in the Glass and Hunt data. Whether that edge vortex is solely responsible for the discrepancy between expected and observed ratios of pre-shock and post-shock pressures remains unknown. Moreover, there may be other complex phenomena that would affect the comparison of reduced-order model predictions to the Glass and Hunt data.

While the conclusions presented here are limited, the methods used in this study should be new to many readers. The posterior p-value enables the assessment of hypotheses about model form, even when model parameters are left free. The Metropolis-Hastings algorithm provides a robust way to accomplish the model tuning involved in both traditional Bayesian inference and the more computationally intense posterior p-value method. In a different vein, the non-parametric difference allows a specialized analysis free from assumptions about distribution form. These are the kind of detailed statistical tools that engineers will need if they are to undertake verification, validation, and uncertainty quantification successfully.

Pre-validation work with the Glass and Hunt data is not complete. It is possible that more severe statistical tests (e.g. Kolmogorov-Smirnov) may provide a basis on which to falsify hypotheses explored in this study. Perhaps more important is the problem of the edge vortex. It is impossible to account for the

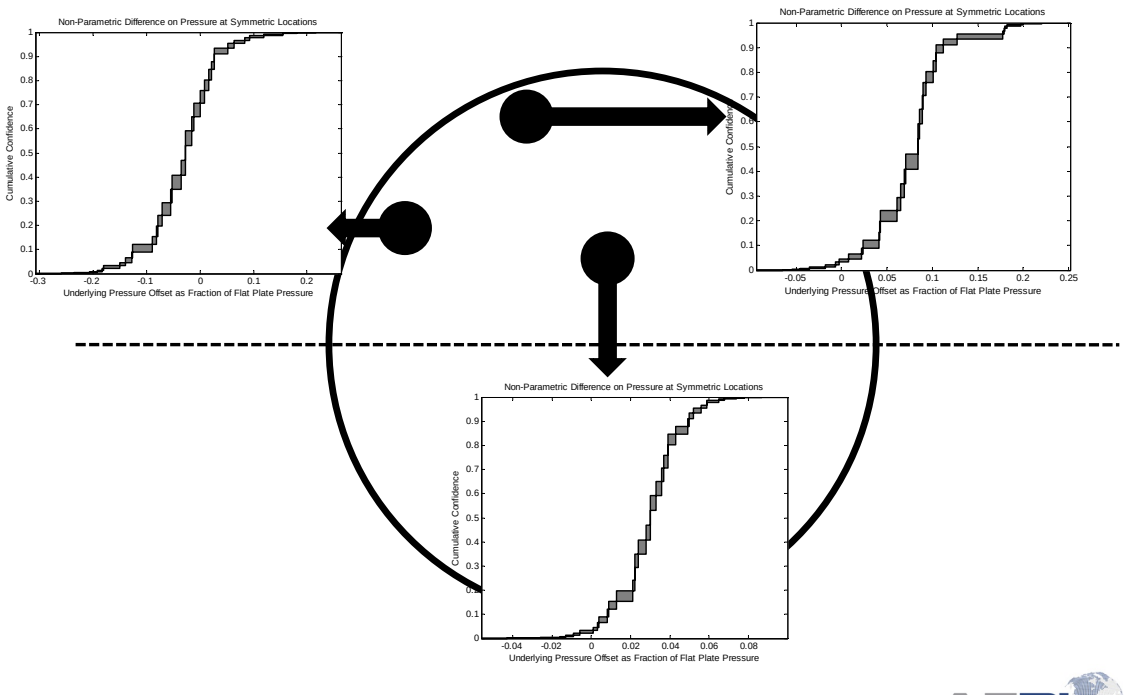

Figure 6. Non-Parametric Difference Across Three Pressure Tap Pairs

effects of an edge vortex without a coherent hypersonic edge vortex model. To the authors' knowledge, this remains one of the outstanding problems in aerodynamics.

While there will certainly be some point at which the analyst must say "enough is enough," right now, the unexplained discrepancies in the Glass and Hunt data are too large to ignore. Moreover, it is anticipated that legacy hypersonic wind tunnel data are replete with such discrepancies. The wind tunnel engineers of previous decades simply did not design their experiments with quantitative model validation in mind. However, judicious pre-validation studies, like the one commenced here, may ultimately provide a framework for using those legacy data in today's validation efforts.

VII. Acknowledgments

This research was sponsored by the Air Force Office of Scientific Research (AFOSR) and the authors gratefully acknowledge the support of AFOSR program managers Drs. Fariba Fahroo and David Stargel.

References

[1]Smarslok, B. P., Culler, A. J., and Mahadevan, S., "Error Quantification and Confidence Assessment of Aerothermal Model Predictions for Hypersonic Aircraft, AIAA-Paper 2012-1817," 53^{rd} *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Honolulu, HI, April 2012, CD-ROM.

[2]Culler, A. J. and McNamara, J. J., "Studies on Fluid-Thermal-Structural Coupling for Aerothermoelasticity in Hypersonic Flow," *AIAA Journal*, Vol. 48, No. 8, August 2010, pp. 1721–1738.

[3]Ostoich, C. M., Bodony, D. J., and Geubelle, P. H., "Fluid-Thermal Response of Spherical Dome Under a Mach 6.59 Laminar Boundary Layer," *AIAA Journal*, Vol. 50, No. 12, December 2012, pp. 2791–2808.

[4]Glass, C. E. and Hunt, L. R., "Aerothermal Tests of Spherical Dome Protuberances on a Flat Plate at a Mach Number of 6.5," Tech. Rep. NASA TP-2631, NASA, NASA Langley Research Center, Hampton, VA, 1986.

[5]Casella, G. and Berger, R. L., *Statistical Inference*, Thomson Learning, Pacific Grove, California, 2nd ed., 2002.

[6]Mayo, D. G., *Error and the growth of experimental knowledge*, University of Chicago Press, Chicago, Illinois, 1996.

[7]Fisher, R., "Statistical Methods and Scientific Induction," *Journal of the Royal Statistical Society, Series B*, Vol. 17, No. 1, 1955, pp. 69–78.

[8]Ferson, S., Oberkampf, W. L., and Ginzburg, L., "Model validation and predictive capability for the thermal challenge problem," *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, No. 29, 2008, pp. 2408–2430.

[9]Balch, M. S., "Mathematical Foundations for a Theory of Confidence Structures," *International Journal of Approximate Reasoning*, Vol. 53, 2012, pp. 1003–1019.

[10]Deveikis, W. D. and Hunt, L. R., "Loading and Heating of a Large Flat Plate at Mach 7 in the Langley 8-Foot High-Temperature Structures Tunnel," Tech. Rep. NASA TN D-7275, NASA, NASA Langley Research Center, Hampton, VA, 1973.

[11]Anderson, J. D., *Modern Compressible Flow: with Historical Perspective*, McGraw-Hill, New York, NY, 3rd ed., 2003.