Measurement of Force Balance Repeatability and Reproducibility in the NTF

Article · February 2004

DOI: 10.2514/6.2004-771 · Source: NTRS

Michael Hemsch
NASA Langley Research Center (Retired)

Measurement of Force Balance Repeatability and Reproducibility in the NTF

M. J. Hemsch* and D. G. Tuttle†
NASA Langley Research Center, Hampton, Virginia

H. P. Houlden‡ and A. B. Graham‡
ViGYAN, Hampton, Virginia

Abstract
A recently published statistical approach for measuring and evaluating wind tunnel force balance repeatability and reproducibility is applied to three check standard tests in the National Transonic Facility at NASA Langley Research Center. Two different airframe models and force balances were used. The short-term repeatability and within-test reproducibility are separately estimated and correlations with tunnel parameters are carried out. Conjectures are presented for the development of scaling laws for predicting the repeatability and reproducibility of other force balance tests in the tunnel.

Nomenclature

AFmax	balance maximum load in axial force
b	model span
CAF	axial-force coefficient in balance coordinates
CNF	normal-force coefficient in balance coordinates
CPM	pitching-moment coefficient in balance coordinates
DACAF	group average of CAF minus the grand average of CAF at a given angle of attack
DACNF	group average of CNF minus the grand average of CNF at a given angle of attack
DACPM	group average of CPM minus the grand average of CPM at a given angle of attack
FS	fuselage station or full-scale
mR	two-point moving range
MAC	model mean aerodynamic chord
M_∞	test-section Mach number
NFmax	balance maximum load in normal force
PMmax	balance maximum load in pitching moment
q_∞	test-section dynamic pressure
R	range (maximum – minimum)
S	standard deviation
Sref	model reference area

*Aerospace Engineer, Associate Fellow, AIAA
†Aerospace Technologist
‡Senior Engineer, Member, AIAA

Introduction
A framework for statistical evaluation, control and improvement of wind tunnel measurement processes was previously developed by the first author and others.[1] The framework is derived from elements of the Measurement Assurance Plan concept[2] developed by the National Bureau of Standards (NBS) and is based on modern notions of measurement statistical process control[3] (SPC) and periodic repeat testing of a standard artifact such as a wind tunnel reference model, herein referred to as a check standard.[1,2,4] The present work is part of an on-going data quality assurance project being conducted by the Wind Tunnel Enterprise (WTE) at Langley Research Center (LaRC). This paper presents results for three recent check standard tests in the National Transonic Facility (NTF) at the NASA Langley Research Center (LaRC) using two different models and balances. The force balance short-term repeatability and within-test reproducibility are separately estimated as a function of tunnel parameters and conjectures are presented for the development of scaling laws for predicting the repeatability and reproducibility of other force balance tests in the tunnel.

The next section describes the LaRC WTE data quality assurance agenda with the following section presenting the facility and models used in the present study. The main section describing the experimental results and their analysis is presented next followed by some final remarks.

Data Quality Assurance Agenda
Ideally, each wind tunnel facility should be able to give customers, before their tests, the estimated repeatability and reproducibility (R&R) for all of the research measurements proposed based on repeat measurements of one or more standard (reference) models. The pre-test predictions could then be used as quality-check limits for repeat sets to be obtained during the customer tests. If those in-test checks are satisfied, then the customers would be assured that the pre-test estimated R&R values were valid for their tests.

One of the key problems with such an agenda, and one that is not generally appreciated in the wind tunnel testing community, is that a large number of replicates are needed to put reasonable bounds on any estimates of the standard deviations for R&R. Figure 1 shows the confidence limits for 95% coverage for the ratio of the actual population standard deviation to the measured sample standard deviation as a function of the number of observations, assuming a stable Normal distribution for the measurements of interest.[5] Figure 1 shows that at least 30-40 observations are required to achieve a bound that is within 20-30% of the actual value.

Obtaining such a large number of repeats for the large test matrices routinely requested in today's testing is clearly out of the question. In order to obtain good estimates of R&R for customer tests without such a burden of repeats, the LaRC WTE has adopted an agenda similar to that previously developed by the National Bureau of Standards for precision calibration laboratories.[2] The WTE agenda is as follows:

1. Periodically measure R&R using standard models, also known as check standards, in special facility tests. Test only in the attached-flow regime so that the results are more likely to be airframe independent and work with the balance outputs in balance coordinates only to avoid mixing results from the six output channels.
2. Track the means and dispersions of the results over time to assure that the measurement system is stable and to build up a reasonable number of observations.
3. Characterize the measured R&R in terms of tunnel, balance and model parameters.
4. Use the above characterization to predict (pre-test) and confirm (in-test) the customer test R&R.

To carry out this agenda, each check standard test consists of 5 groups of repeat angle-of-attack (AOA) runs obtained at 5-10 test section conditions. Each of the groups contains the following back-to-back runs:

1. Inverted
2. Upright
3. Upright
4. Upright
5. Inverted

The first two and last two runs are used to make flow angularity measurements with the two measurements forming a repeatability group. The three upright runs form a repeatability group for estimating force and moment repeatability. Discussion of the flow angularity results and across-test reproducibility is beyond the scope of this paper and will be presented in a later report.

Assessment of the within-group variation gives the repeatability. Similarly, assessment of the across-group variation gives the within-test reproducibility. The across-test variation is given by comparing grand averages after flow angularity changes have been taken into account.

Facility and Models

Facility

The NTF is a closed-circuit, continuous, single-return, transonic, pressure and cryogenic wind tunnel with a slotted test section that can be operated with the slots partially or fully closed. The tunnel and its operation are described in Ref. 6 and 7. For the tests analyzed in this paper, the test section was operated in the 6% open condition and the warm air (non-cryogenic) mode.

Pathfinder I Model

The primary check standard (reference) model for the NTF is known as Pathfinder I (PFI) which is described in detail in Ref. 8. A photograph of the model mounted in the test section is shown in Fig. 2. A sketch, taken from Ref. 8, showing planform and side views of the model, is given in Fig. 3. The characteristics of the model and balance are given in Table 1. Two tests of this model (T121, T137) were conducted within 6 months. Five repeatability groups were obtained in each test. The groups consisted of five back-to-back angle-of-attack (AOA) sweeps at 7-9 values of dynamic pressure* ranging from 15 to 2414 psf (see Table 2 for the nominal test conditions).

The sweeps were obtained at nominal AOA values ranging from -2.0 to 2.5 degrees with zero slideslip and roll. These AOA values are sufficiently small to ensure that the flow over the model is essentially attached. This condition is a requirement for attempting to characterize the check standard results for prediction of the repeatability and reproducibility for other model and balance combinations. The AOA

* Previous unpublished check standard measurements in the LaRC tunnels had suggested that force balance repeatability correlates with test section dynamic pressure.

sweeps cover a range of normal-force coefficient from roughly zero to 0.5.

HSR M5 Model

Another check standard model known as the HSR M5 model was also tested in the NTF in the present study about two months after the completion of the second PFI test. A photograph of the model mounted in the test section is shown in Fig. 3. A planform sketch of the model is given in Fig. 4. In this single test entry, five repeatability groups similar to those obtained for the PFI were obtained. The set of tunnel conditions was different however. The groups were obtained at five values of dynamic pressure ranging from 58 to 710 psf as shown in Table 3. The sweeps were obtained at nominal AOA values ranging from -0.5 to 4.0 degrees with zero slideslip and roll. These AOA values are sufficiently small to ensure that the flow over the model is essentially attached.

Analysis of Experimental Results

Analysis Steps

The U. S. Guide to the Expression of Uncertainty in Measurement[9] (GUM), defines repeatability to be the "closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement". For the purposes of normal wind tunnel force balance testing, "successive" is taken to mean back-to-back sweeps through some parameter such as AOA.

In order to statistically evaluate force balance R&R in a consistent way, the following analysis steps were taken for each test:

1. The normal- and axial-force coefficients, CNF and CAF, and the pitching-moment coefficient, CPM, were reduced in balance coordinates. As will be seen below, the three separate measurements of normal force, axial force and pitching moment exhibit somewhat different repeatability and reproducibility characteristics. Hence, this step is needed for eventual scaling of the individual channel results to other models and balances.

2. For each of the three upright AOA sweeps, the results for each force or moment coefficient of interest were interpolated to the nominal AOA setpoint values. Second-order interpolation was used. The purpose of this step is to remove as much of the variation due to setpoint error as possible. Since interpolation can't be perfect, there will be some residual setpoint error mixed with the actual measurement variation but it should be small.

3. For each nominal value of AOA in each group, the group range and group average of the three coefficient values are obtained, e.g.

$$
\begin{aligned}
R_{CAF} &= Max(CAF_1, CAF_2, CAF_3) \\
&\quad - Min(CAF_1, CAF_2, CAF_3) \qquad (1) \\
\overline{CAF} &= (CAF_1 + CAF_2 + CAF)/3
\end{aligned}
$$

For quick evaluation, the authors use plots like the one shown in Fig. 5 which shows three frames: (1) lower left --- the original data from the three-back-to-back runs (green symbols), the interpolated points (black symbols), and the average (the solid black line), (2) upper left --- the residuals about the interpolated averages, and (3) the ranges given by Eq. 1.

4. For each coefficient of interest and each tunnel condition, the group ranges are plotted as a function of group number (increasing in time as the test progresses) and AOA value.

5. The ranges for all of the groups at a given condition are checked for statistical consistency with respect to the different groups and the different AOA values. This check is carried out using a modified statistical control chart.[1-5] The centerline of the chart is given by the average of all of the ranges for each coefficient for a given test condition. The statistical control chart lower range limit for groups of three is zero. The upper range limit is given by[4]

$$URL = 2.575\,\overline{R} \qquad (2)$$

where $\overline{R}$ is the average range. For consistency, all of the individual group ranges should fall within the limits and be clustered reasonably about the centerline.[4]

6. The averages for all of the groups at a given condition are also checked for statistical consistency with respect to the different groups and the different AOA values. This check is carried out using a modified statistical control chart known as a difference chart. For this plot, the grand average of the group averages at a given angle of attack is subtracted from the group averages. Hence, for this kind of plot the centerline is given by zero. The limits for this chart are computed in two ways: (1) using the average range from the corresponding group range chart and (2) using the average of the moving ranges.[1-4] If there is no significant additional source of variation between the groups, the two sets of limits will be roughly the same. If there is additional between-group variation, the group averages will tend to be outside the limits given by the average

range too often and the limits given by the average moving ranges should be used to determine statistical consistency.

Example Within-Group Control Charts

Example within-group charts are given in Fig. 6 for the three balance output coefficients of normal force, axial force and pitching moment. The condition shown is for PFI, T121 at $M_\infty = 0.45, q_\infty = 904\,\text{psf}$. Each of the colored lines represents the ranges of the five groups obtained at one angle of attack. Then arrayed across the chart are similar results for the other angles of attack. Fig. 6 shows that a single centerline and upper limit for each coefficient contain all of the ranges for those five sets of back-to-back runs, indicating statistical consistency for the test at those conditions. Figure 6 is a typical chart for all of the conditions with about 2% of the ranges being above the upper limit. For the testing being conducted, it would be reasonably expected that roughly that many ranges would be outside the limits by chance.

It should be noted that the ranges appear to be uncorrelated across angle of attack, suggesting that the within-run scatter is dominant for short-term repeats.

Example Between-Group Control Charts

Example between-group charts are given in Fig. 7 for the same results described in the previous subsection. The results show statistical consistency for the limits given by the average moving range but not for the limits given by the average range. This result shows that significant between-group variation was added to the (short-term) within-group variation. Note also that the groups appear to be closely correlated. This simply shows that the averages, if plotted versus angle of attack, would appear to be roughly parallel. Fig. 7 is a typical chart for all of the test conditions.

Repeatability Results

The average ranges obtained from the control charts for each test condition can be used to estimate the standard deviation for the within-group repeatability as follows:[1-4]

$$\hat{\sigma}_{wg} = \bar{R}/1.693 \tag{3}$$

for groups with three observations (per angle of attack). Those values were computed for each test condition and converted to a percentage of the full-scale balance output using the nominal dynamic pressure and the model reference parameters. The results for the three tests are given in Fig. 8. Also shown in Fig. 8 is the typical reported standard uncertainty for balance calibration. Note the good agreement (for scatter) of the two PFI tests.

The results of Fig. 8 follow a pattern also observed in unpublished studies of other LaRC WTE tunnels. At the lowest dynamic pressures, the balance output repeatability is about the same for all three channels and is roughly equal to the resolution of the measurement system. However, after some threshold value of dynamic pressure, the standard deviation begins to increase with increasing dynamic pressure to roughly an order of magnitude larger than the resolution or more. The weaker balance used for the HSR M 5 test shows increasing repeatability at a much lower threshold than the balance used for the PFI.

The observed repeatability behavior casts suspicion on model dynamics as somehow being the cause since it would be expected that the amplitude of any dynamics, absent separated flow, would be proportional to q_∞. For this particular set of balances and models, a simple stretching of the abscissa for the HSR M5 data

$$q_{stretched} = 2.5\,q_\infty \tag{4}$$

is sufficient to satisfactorily collapse the results for all three channels as shown in Fig. 9. This value may simply be coincidence, but it is interesting to note that

$$2.5 = \sqrt{k_{PFI}/k_{HSR}} \tag{5}$$

where k is the balance spring constant in yaw. (NTF model dynamics are known to occur primarily in yaw.)

Reproducibility Results

The average moving ranges, together with the results from the previous subsection, can be used to estimate the standard deviation for the between-group reproducibility. The standard deviation for the group averages is given by[1-4]

$$\hat{\sigma}_{\bar{x}} = \overline{mR}/1.128 \tag{6}$$

But, the between-group variation is defined to be that which is *in addition to* the within-group variation. Hence, $\hat{\sigma}_{bg}$ is defined by (for three observations per group)

$$\sigma_{\bar{x}}^2 = \sigma_{bg}^2 + \sigma_{wg}^2/3 \tag{7}$$

Solving for $\hat{\sigma}_{bg}$ gives

$$\hat{\sigma}_{bg} = \sqrt{\hat{\sigma}_{\bar{x}}^2 - \hat{\sigma}_{wg}^2/3} \tag{8}$$

The results for all three tests are given in Fig. 10. Recall that the effective number of observations for the reproducibility is only five which puts large confidence limits on the estimate for capturing the true population standard deviation. Given that large confidence interval (0.6 S to 2.9 S), the results seem to indicate that the reproducibility for the NTF is independent of the dynamic pressure, at least for the procedures used in the test. It is the authors' tentative conjecture that the between-group reproducibility is due to temperature and temperature gradient changes during the testing which change the balance zeros. This is something which can be controlled and perhaps even modeled to reduce the variation. Even without additional control, the between-group reproducibility appears to be controllable to a level below the calibration bias error.

Closing Remarks

Initial replication data sets from the NTF obtained in three separate tests using two models and two balances indicate that the short-term balance repeatability is statistically consistent within and across tests and is a function primarily of the test section dynamic pressure if considered as a fraction of the full-scale balance output, e.g.

$$\sigma_{wg,NF} = NFmax\ f(q_\infty) \qquad (9)$$

and that a mathematical model for $f(q_\infty)$ might be found by considering the tunnel forcing function that causes the model dynamics.

In addition, the between-group reproducibility appears to be independent of q_∞ and, for the procedures followed in the present tests at least, is no bigger than the calibration bias error.

References

1. Hemsch, M., Grubb, J., Krieger, W., and Cler, D., "Langley Wind Tunnel Data Quality Assurance - Check Standard Results. AIAA Paper 2000-2201, June 2000.
2. Belanger, B., "Measurement Assurance Programs, Part 1: General Introduction". NBS SP 676-I, May 1984.
3. Wheeler, D. J., and Chambers, D. S. [1992]: *Understanding Statistical Process Control*, SPC Press, 1992
4. Eisenhart, C., "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems. *Precision Measurement and Calibration—Statistical Concepts and Procedure*s, NBS SP 300–Vol. 1, February 1969, pp. 21–47.
5. Anon., "Manual on Presentation of Data and Control Chart Analysis: 6th Edition", ASTM Manual Series: MNL 7, 1995.
6. Foster, J. M., Adcock, J. B., "User's Guide for the National Transonic Facility Research Data System", NASA TM 110242, April 1996.
7. Bruce, W. E., Jr., Gloss, B. B., "The U.S. National Transonic Facility", *Special Course on Advances in Cryogenic Wind Tunnel Technology*, Ch. 3 in AGARD-R-774, November 1989.
8. Jacobs, P. F., Gloss, B. B., "Longitudinal Aerodynamic Characteristics of a Subsonic, Energy-Efficient Transport Configuration in the National Transonic Facility", NASA TP 2922, 1989.
9. Anon., *U. S. Guide to the Expression of Uncertainty in Measurement*, ANSI/NCSL Z540-2-1997, October 1997.

	Pathfinder I (T121, T137)	**HSR M5 (T141)**
NFMAX, lbf	6500	1200
AFMAX, lbf	400	120
PMMAX, in-lbf	13000	2400
SREF, sq. ft.	1.988	2.385
MAC, in.	5.74	19.085
b, in.	52.97	26.383
Length, in.	50.0	52.74
Aspect Ratio	9.8	2.0

Table 1. Balance characteristics and model and reference quantities.

q_∞, psf	M_∞	**T121**	**T137**
15	0.09	X	X
34	0.14	X	X
77	0.21	X	X
176	0.19	X	X
398	0.29	X	X
904	0.45	X	X
1425	0.60	X	X
1690	0.80		X
2414	0.50		X

Table 2. Nominal test conditions for Pathfinder I model.

q_∞, psf	M_∞
58	0.2
125	0.3
212	0.4
418	0.6
710	0.9

Table 3. Nominal test conditions for HSR M5 model.

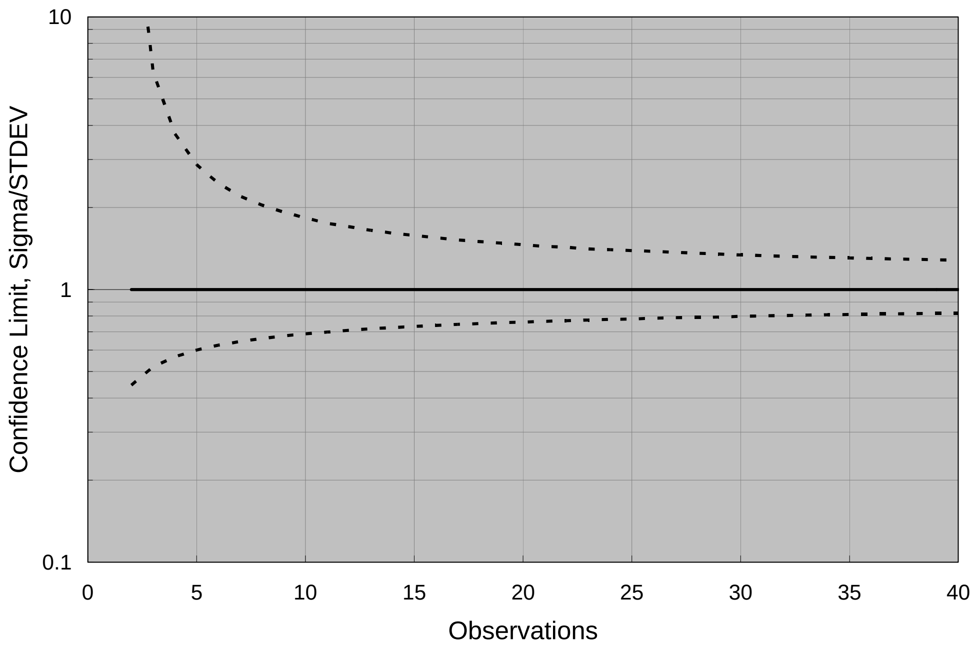
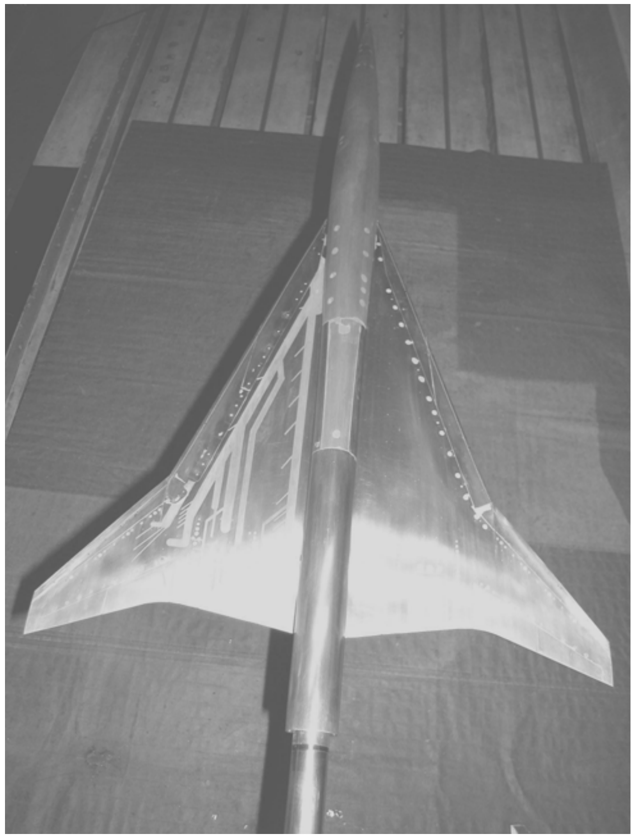
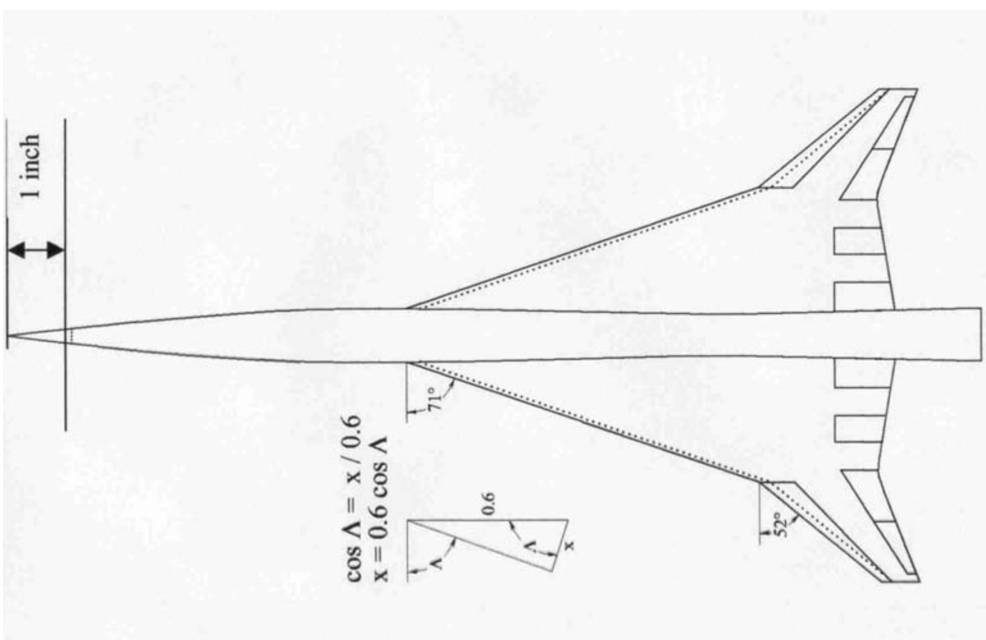


Figure 1. Confidence limits for the standard deviation (Normal distribution).

Figure 2. Photograph of the Pathfinder I model mounted in the NTF.



(a) Photograph of the model in the NTF test section



(b) Planform sketch of the model
Figure 4. HSR M5 model used in T141.

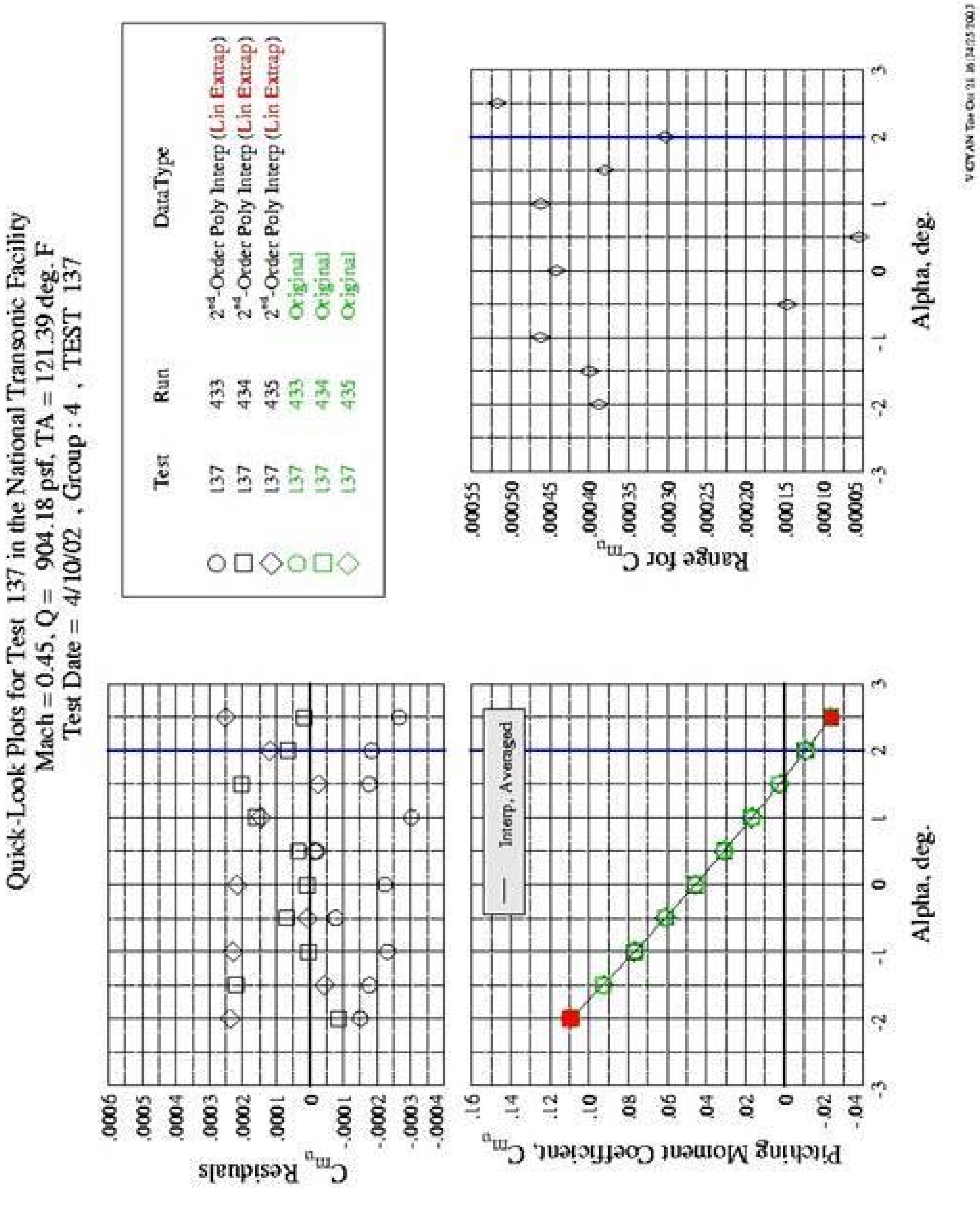
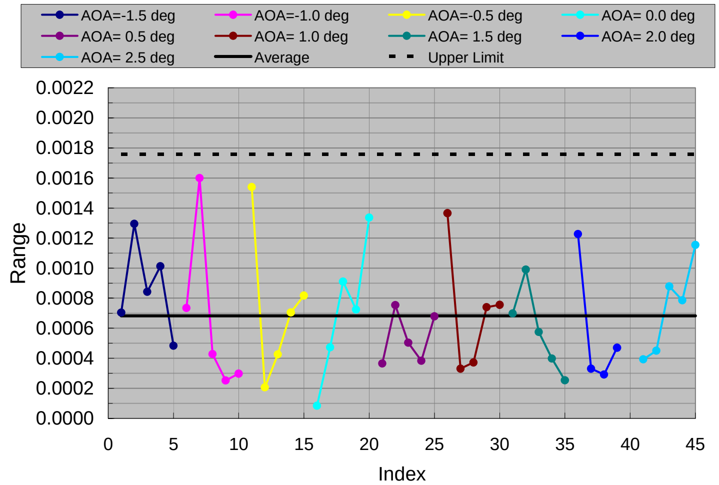
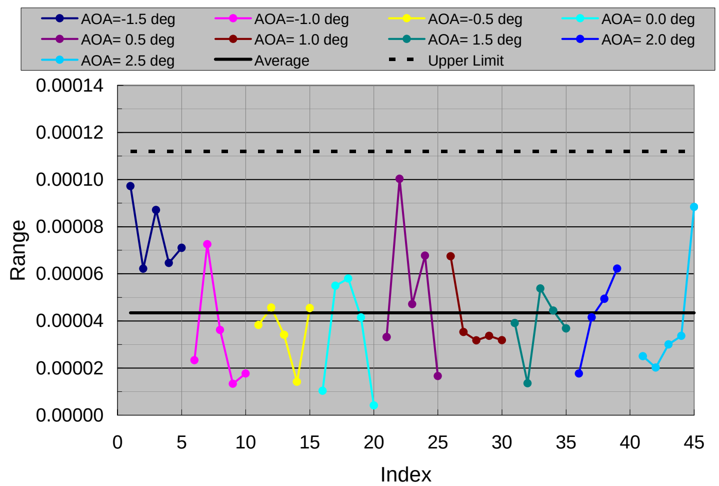


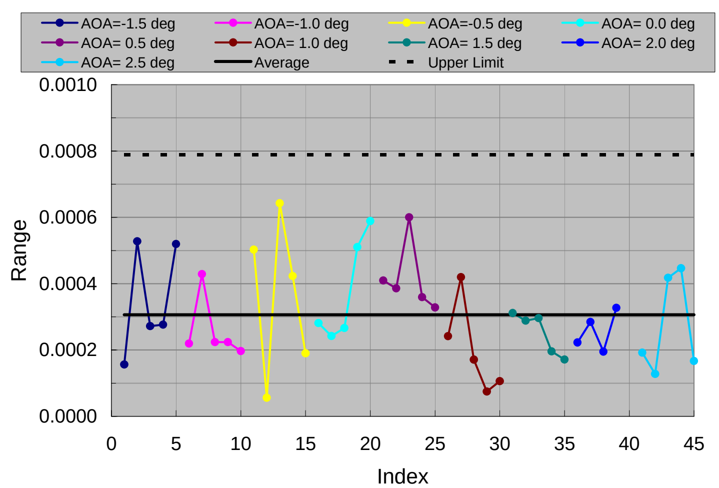
Figure 5. Quick-look plots for NTF T137, Group 4, $M_\infty = 0.45$, $q_\infty = 904$ psf



(a) Control chart for CNF ranges.

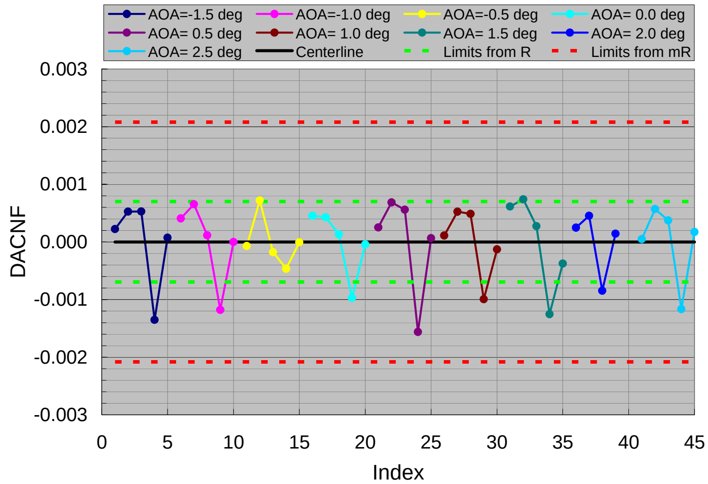


(b) Control chart for CAF ranges

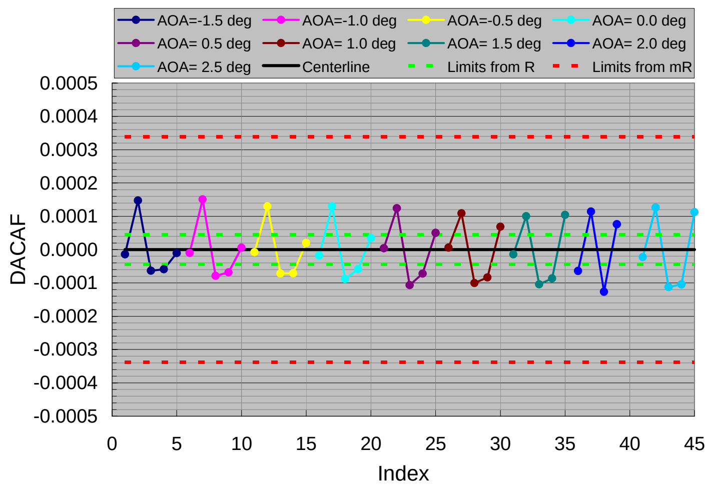


(c) Control chart for CPM ranges

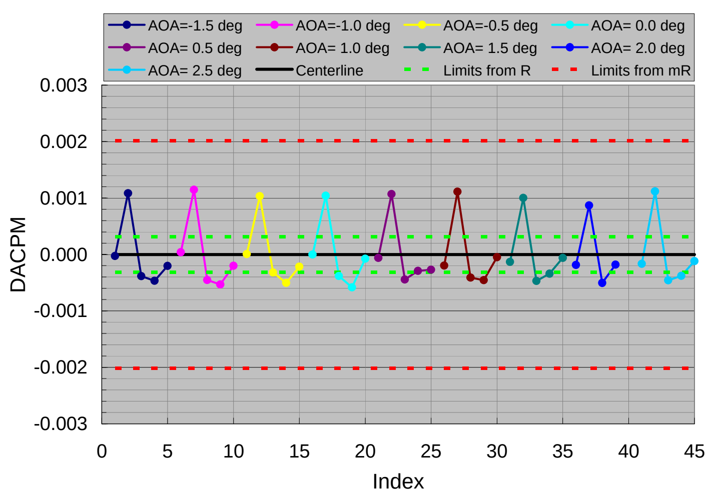
Figure 6. Example statistical control charts for group ranges. PFI Test 121, $M_\infty = 0.45$, $q_\infty = 904\,\text{psf}$.



(a) Control chart for DACNF

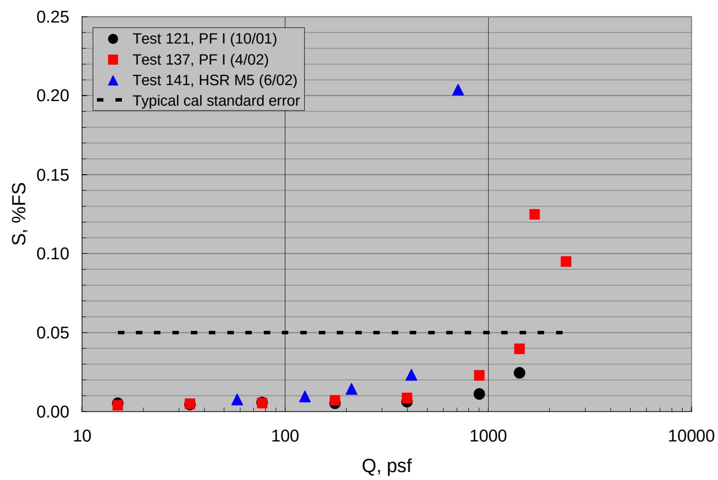


(b) Control chart for DACAF

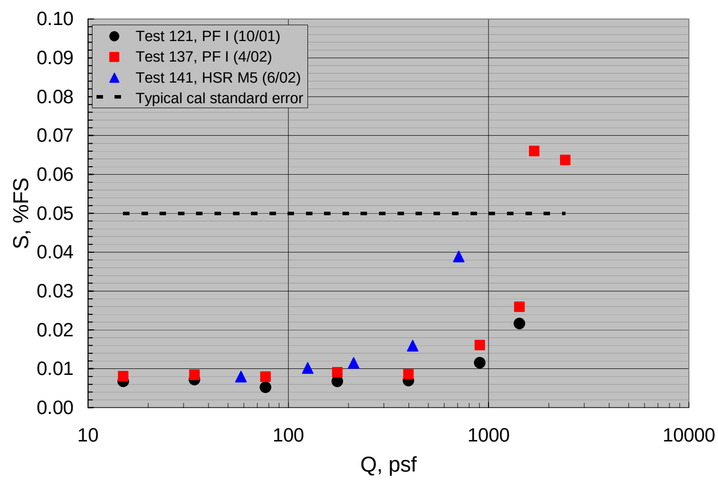


(c) Control chart for DACPM

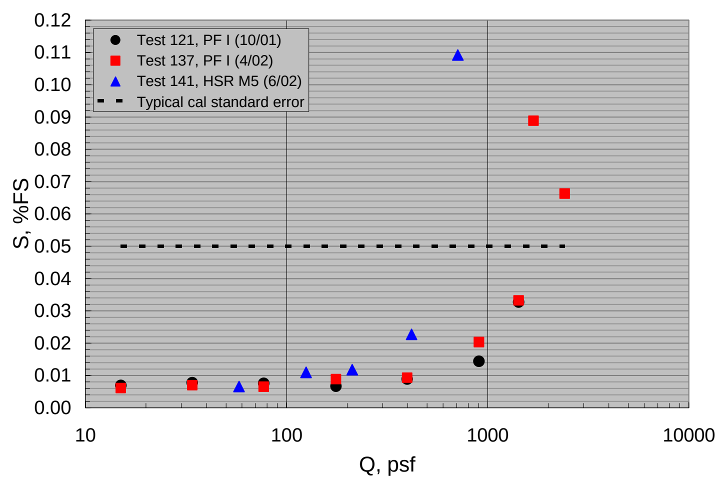
Figure 7. Example statistical control charts for group delta averages. PF I Test 121, $M_\infty = 0.45$, $q_\infty = 904\,\text{psf}$.



(a) Normal force

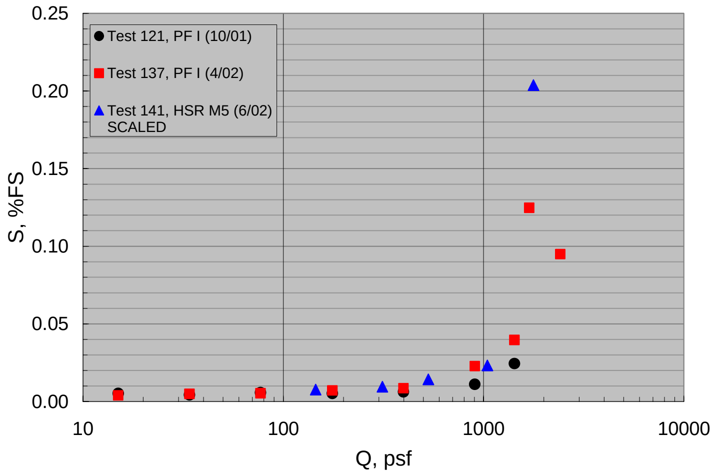


(b) Axial force

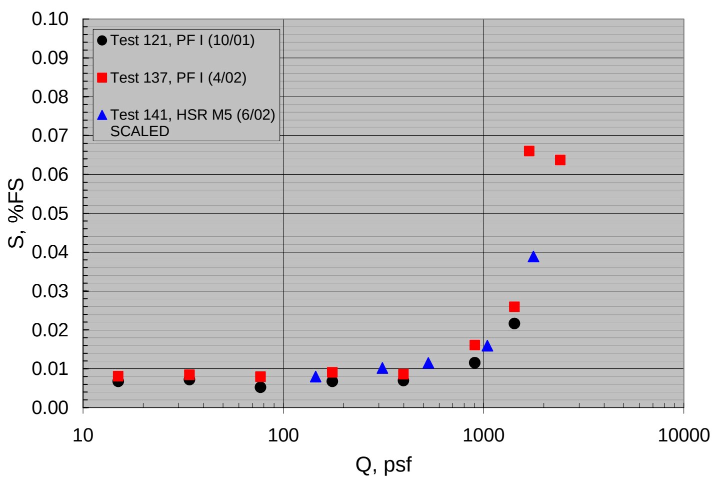


(c) Pitching moment

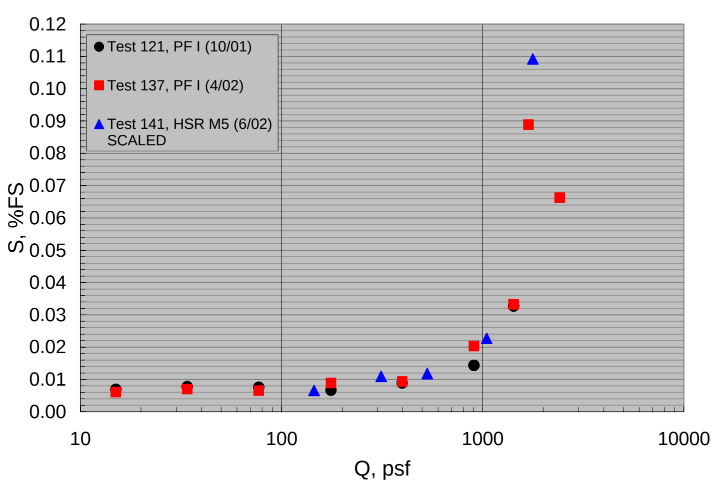
Figure 8. Estimated standard deviations for within-group variation (repeatability), %FS.



(a) Normal force

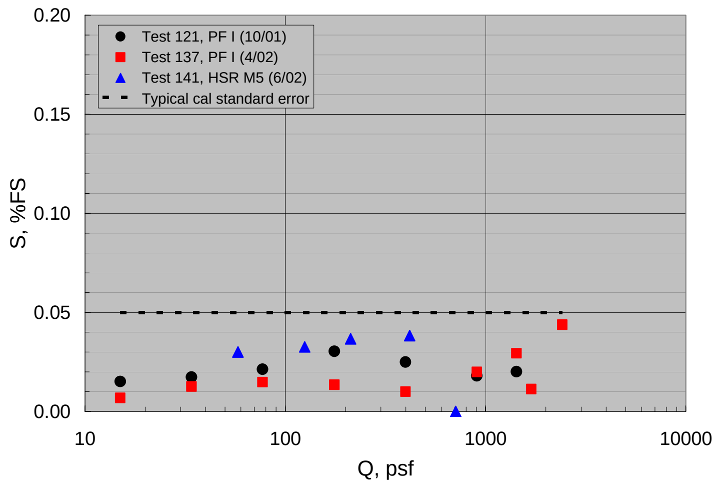


(b) Axial force

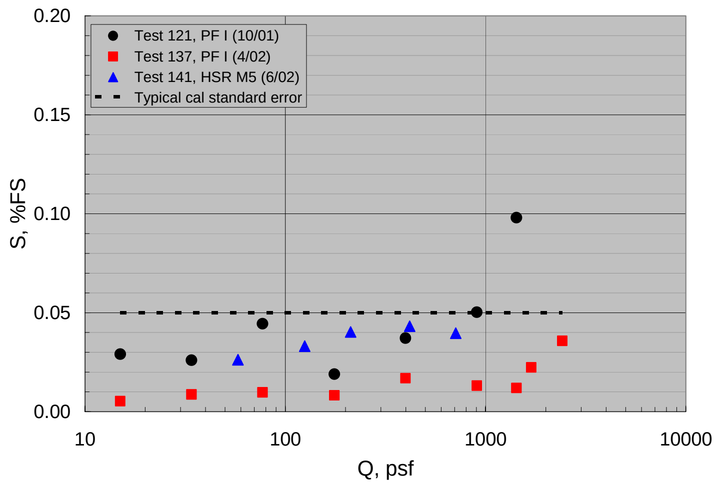


(c) Pitching moment

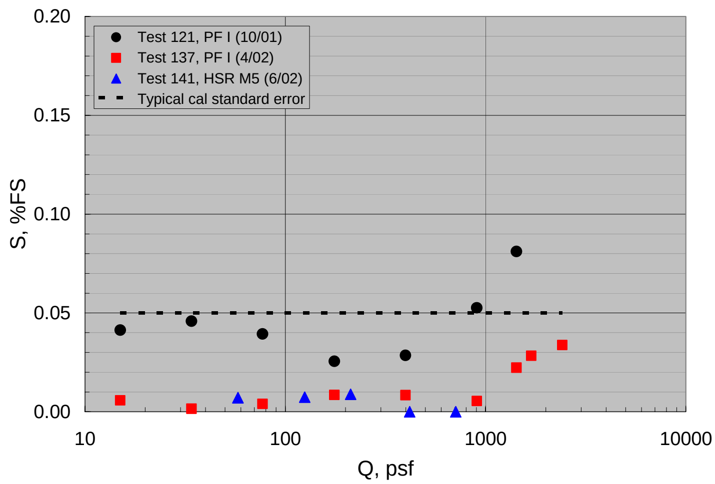
Figure 9. Within-group standard deviations with stretching, %FS.



(a) Normal force



(b) Axial force



(c) Pitching moment

Figure 10. Estimated standard deviations for between-group variation (within-test reproducibility).