A PRELIMINARY ASSESSMENT OF THE EFFECTIVENESS OF 5700A ARTIFACT CALIBRATION

Les Huntley
Les Huntley, Metrologist, Inc.
3715 17th Street
Lewiston, ID 83501

ABSTRACT

As Found and As Left data obtained from approximately 260 5700As returned to Fluke for service were analyzed to assess whether artifact calibration provides specified uncertainties. While the work is not complete because adequate information about measuring system contributions to observed errors is not yet available, this black-box analysis provides preliminary confirmation that artifact calibration operates as it was designed to do and the 5700A performs well within specifications.

INTRODUCTION

About seven years ago, Fluke introduced the 5700A MultiFunction Calibrator. Among its many innovations was the introduction of Artifact Calibration, an interesting internal-metrology approach to maintaining its ambitious accuracy specifications. About the same time, Hewlett Packard introduced its Model 3458 Digital MultiMeter, supported by its own implementation of artifact calibration.

Artifact Calibration as implemented by Fluke incorporates the metrology functions usually performed by a standards or calibration laboratory into the instrument. Thus, the 5700A contains internal reference standards, an internal ratio device and a sensitive null detector. Under the control of its internal software, the 5700A performs the metrology procedures necessary to determine offsets from nominal of its many outputs, and make internal corrections to drive the offsets essentially to zero. Artifact calibration should be capable of providing superior accuracy because the calibration environment and procedures are fixed by the design of the instrument, and thus are not subject to the range of conditions, attitudes and capabilities found in the community of calibration laboratories.

Performing artifact calibration is designed to return the Direct Voltage (DV), Direct Current (DC) and Resistance functions to essentially the accuracy the instrument had when it was manufactured. This is accomplished by adjusting internal references by comparison to external standards, then verifying ratio accuracy, and finally performing the measurements and making the adjustments needed to bring the outputs to nominal. Alternating Current (AC) and Alternating Voltage (AV) outputs are also returned to nominal, referenced to an internal

AC/DC Difference Standard. Artifact calibration of AV and AC differ from the other functions only in the fact that the internal reference is not measured or adjusted. Long term accuracy of these parameters is ultimately determined by long term stability of the internal AC/DC Difference standard.

Artifact Calibration is performed at intervals determined by the specification to be maintained. For example, artifact calibration must be performed each 90 days to maintain 90 day specifications. But no electronic instrument can be relied upon to function exactly as designed forever. Further, the internal AV reference standard is not tested or adjusted by artifact calibration. For these reasons, Fluke requires a complete accuracy verification at two year intervals no matter which specification is maintained.

The 5700A also provides a Cal Check function, which may be used to monitor outputs relative to the internal standards. This feature provides users with a means for implementing a measurement quality program such as that required by paragraph 9.6 of ANSI/NCSL Z540-1, which requires "in-service checks between calibrations and verifications".

Engineers, especially designers, tend to want detailed descriptions and specifications for the internal workings of an instrument. Others who see the instrument as a tool to be used in daily operations are inclined to view the instrument as a "black box", which provides specified outputs in response to specific inputs. The "Black Box" approach to instrument evaluation is concerned only with inputs and outputs and whether the instrument performs to specifications. Details of internal functions are important only when corrective action is required. To practical users, the real issue for artifact-calibrated instruments is whether they perform to specification.

As Found and As Left data for a large number (~260) 5700As obtained by Fluke's Customer Support Services lab from instruments returned for service has been subjected to a black box analysis to determine whether the 5700A performs as specified. This present report concentrates on AV and DV performance, as they provide a good indication of performance to spec. The current functions are so obviously conservatively specified that further analysis is pointless until complete data is available. There is strong but inconclusive evidence that resistance measurements are strongly influenced by offsets and variability in the measuring system. Therefore complete analysis of that function must also await information about measurement system contributions to measured errors.

The conclusions presented here are preliminary because the information needed to separate the effects contributed by the measuring system from effects originating in the instrument is not yet available. Incorporating that information into the analysis will lead to more certain conclusions, which, it is anticipated, will be the subject of another paper.

AS FOUND DATA

Fluke has established a capability for calibration and repair of the 5700A in its Northwest Technical Center, a function of its Customer Support Services (CSS) Group. As customers require service for the 5700A, they are returned to CSS for appropriate processing. As Found data is taken for operational instruments, and data for approximately 260 instruments (a small fraction of total production) was available. The number is approximate because some of the instruments were returned for repair, so not all returned instruments operated properly in all functions.

As Found data was processed in EXCEL spreadsheets, using the statistical and graphing functions. As noted above, not all data is useable and in some cases outlying data points more than four standard deviations from the mean were discarded. As a result, the data set analyzed consists of those points within 4s of the mean.

AS LEFT DATA

Instruments returned to customers are calibrated, using Artifact Calibration, then are verified prior to their return, generating As Left data. As Left data is useful for demonstrating that Artifact Calibration is still operating properly at various times after manufacture. When As Found and As Left data is available for the same instrument, the difference provides an indication of 5700A stability. Analyzing the difference between As Found and As Left data not only can provide useful information about time stability of the 5700A, but clues to the operation of the measuring system as well. Regretfully, this data was provided too late for complete analysis results to be included in this report.

THE PROBLEM OF INTERPRETATION

Interpretation of the data is complicated by the need to separate instrument performance from measuring system performance. Complete analysis of instrument performance requires complete knowledge of the contributions of the measuring system. While the measuring system was analyzed and determined to be capable of testing 5700As to appropriate specs, the detailed analysis of standard deviations and offsets required by the present investigation has not been completed, so the needed information about system performance is not yet available. However, many of the instruments performance parameters can be at least bounded by analyzing the available data. The results of that analysis are the subject of this paper.

Each measurement of each parameter is affected by many individual error contributions. These include errors in the artifact calibration process (at manufacture or later), errors in the instrument's internal metrology, errors in the instrument's internal standards, gain and zero instabilities, noise in all these processes, plus offsets and noise in the measuring system.

For the purposes of this analysis, the measured value of a 5700A parameter can be described by a combination of two linear equations as in the following equation

$$\boldsymbol{V} = \boldsymbol{V}_{nom} + \alpha_{inst} + \beta_{inst}\boldsymbol{T} + \alpha_{meas} + \beta_{meas}\boldsymbol{T} + \varepsilon_{inst} + \varepsilon_{meas}$$

where for both subscripts

α = offset at time T = 0 (line intercept)
β = slope of the line (time drift rate)
ε = random noise contribution
T = time

Thus, the measurement results are considered a linear combination of instrument performance and measuring system effects. In the absence of information about the measuring system, this worst-case analysis of the results will provide upper bounds on errors contributed by the instrument itself.

DATA ANALYSIS

"Black Box" users of the 5700A will want answers to two questions: is the instrument in spec As Found, and does artifact calibration plus verification each two years assure the instrument will continue to meet specifications? Engineers and others who worry about internal details will be interested in stability of the internal reference standard of AC/DC difference, since artifact calibration does not measure or adjust it.

Observed Performance to Specifications

We begin by analyzing As Found measurements to see whether they are within specifications. Figure 1 graphically presents the results of the analysis for AV. In this, and the following four figures, plotted points are average values and error bars represent the range of values expected to contain 99% of individual values assuming normal distributions. Two-sided probabilities are computed for points near nominal. For points more than one-half standard deviation from nominal, one-sided probabilities are computed. This is acceptable since only about 0.1% of instruments would be outside spec limits in the direction opposite to the offset if a two-sided probability were to be computed.

As can be seen, nearly all the voltage/frequency points are well within 90 day specifications. Fluke instruments are designed to provide at least 99% probability that they, if properly maintained, will meet all specifications throughout their useful life. The one voltage/frequency which has greater than 1% probability of OOT is 20 mv, 1 MHz. Here, we see both a large offset, and large variability in As Found measured values, and perhaps 2.5% probability of being OOT As Found.

Keep in mind that these are worst-case results -- the measured values contain offsets and noise contributed both by the instrument and by the measuring system. Until further information is available, there is no way to determine whether more than the expected 1% of instruments were really outside 90 day specifications.

Figures 2 through 5 present similar information for Alternating Current, Direct Current, Resistance and Direct Voltage, respectively. Even this worst-case analysis of the current ranges shows the instrument to be performing extremely well with respect to specifications. Uncritically accepting the As Found results for Direct Voltage would lead one to believe that something like 1.5% of instruments were out of tolerance at lower voltages instead of the expected 1%. But variability in the measurements is large relative to spec, and the variability is as likely to be in the measuring system as in the instrument. Final judgment here, as for the other functions, must await more detailed information about measurement system performance.

The preliminary answer to the first question -- is the instrument in tolerance as found? -- is "very likely". Worst case, at least 97% will be in tolerance on recall, a number which will more than satisfy requirements of most calibration systems. Removing measurement system errors will improve that percentage, perhaps to the 99% targeted in the design of the instrument.

Does Artifact Calibration Work?

That Artifact Calibration does adjust values closer to nominal can be shown by comparing As-Found to As-Left measurements on instruments returned for calibration. As-Left measurements are taken immediately after Artifact Calibration is performed, so offsets should have been adjusted essentially to zero. Figures 7 and 8 present this information for DV in a form which is especially interesting.

While the analysis of Direct Voltage performance is clouded by the lack of information about measuring system error contributions, we do have for this function an analysis not available for the others. Since both positive and negative voltages are measured at each level, we can perform a correlation analysis to assess separately the presence of elements affecting voltage magnitudes and zeros.

The sign of the reported deviations from nominal are such that elements which increase either positive or negative voltage magnitude increase the measured deviation, so these deviations are positively correlated. Such elements include errors in calibrating the internal reference, drift in the internal reference, and ratio errors. Zero errors, for example from thermal emfs, affect positive and negative voltages in opposite directions, so deviations caused by zero errors are negatively correlated.

These effects can be displayed graphically by plotting deviations for negative voltages versus deviations for positive voltages. Figures 7 and 8 present the plots for As Found and As Left voltages for 72 instruments judged not to have been repaired. In all cases, comparing the As Found to As Left plots shows that artifact calibration shifts the average closer to zero and decreases the range of variability of correlated effects, as expected. Somewhat surprisingly, artifact calibration appears not to have affected deviations perpendicular to the trend line, corresponding to negatively correlated effects (zero errors). This suggests the zero errors may originate in the measuring system. The shift in average, about 2 ppm, is likely due to a difference in values assigned to the two or more standards used in artifact calibrating the instruments.

It is clear from these plots that artifact calibration does properly adjust the parameters affecting Direct Voltage magnitudes. If the adjustments (and the measurements) were perfect, the As Left plots would have a nearly circular appearance, such as the plot for 1V. While it appears from the plots that the adjustment is not perfect, all instruments were adjusted to within 24 hour specs (except for a couple of outliers at 100 mv). Again, it must be kept in mind that the residual errors are as likely to be in the measuring system as in the instruments measured.

It is interesting that in no case does variability perpendicular to the trend line appear to be improved by artifact calibration. The 100 mv plots show the effect best. What is the reason for this given that 5700A artifact calibration does adjust zeros? It is likely that the observed scatter perpendicular to trend lines is due to variable zero errors in the measuring system, possibly caused by variable thermal emfs.

Stability Analysis

An important requirement for a measurement standard, one such being the 5700A, is that it be stable over time. Proper calibration can always adjust offsets nearly to zero, and we have seen that artifact calibration is proper calibration, at least for DV. The big question is whether offsets adjusted to zero will stay there. Accordingly, As Found data was analyzed for evidence of time stability.

For Alternating Voltage, and to a lesser extent, for Alternating Current, stability of the internal AC/DC Difference standard is a primary contributor to instrument stability. While the reference cannot be accessed directly, upper bounds on its stability can be established by determining average worst-case instrument drift and variability in the time since the instruments were manufactured. For this analysis we ignore offsets and take all time drift and variability as originating in the instrument. Analyzing stability of the internal reference would more properly done using As Left data, since artifact calibration removes effects not related to the standard. Since that data was not available, As Found data was used, recognizing that it will be noisier than As Left.

Linear regression over the data set, using time since manufacture as the independent variable, provides the needed information. We want to estimate the time after manufacture that 1% of the instruments will go out of spec, assuming perfect initial calibration. This 1% OOS point is reached when either the upper or the lower 99% prediction limit intercepts the spec limit. See Figure 6 for a graphical description of how this point is calculated.

The prediction interval is, from standard linear regression analysis,

$$Y = \alpha + \beta T \pm ts\sqrt{1 + \frac{1}{N} + \frac{\left(T - \bar{T}\right)^2}{(N-1)s_x^2}}$$

where

α = intercept (offset)
β = slope of the line (time drift)
T = Time since manufacture
t = student's t (for 99% confidence)
s = standard deviation of points about the regression line
s_x = standard deviation of the regression times
N = number of points in the regression
and the bar indicates an average

Ignoring offsets ($\alpha = 0$), the 1% OOS point is reached when

$$USL = |\beta| T + ts\sqrt{1 + \frac{1}{N} + \frac{\left(T - \bar{T}\right)^2}{(N-1)s_x^2}}$$

where USL signifies the Upper Spec Limit.

Rearranging and expanding this equation yields a quadratic equation which can be solved for time, T, that fulfills the intersect conditions

$$T = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

where

$$A = \beta^2 - \frac{t^2 s^2}{(N-1)s_x^2}$$

$$B = \frac{2t^2 s^2 \overline{T}}{(N-1)s_x^2} + 2USL|\beta|$$

$$C = (USL)^2 - t^2 s^2\left(1 + \frac{1}{N} + \frac{\overline{T}^2}{(N-1)s_x^2}\right)$$

The rather startling results for Alternating Voltage at five frequencies are presented in Figures 9 through 15.

The metrologist must rely on analysis of past history in predicting future performance, but it would be foolish to place too much faith in predictions of time to 1% OOS measured in decades, or in some cases, in centuries. Since the underlying function may not really be linear, and because the effects of long term (non-white) noise may be underestimated, the prudent metrologist will not extrapolate too far beyond his data base -- as a rule of thumb, never extrapolating for periods exceeding the base period for which data is available. The base period for these regressions is 6.2 years. Therefore, projections beyond about 12 years from manufacture should be viewed with appropriate skepticism.

Even with the above caveat, the 5700A Alternating Voltage functions are seen to be remarkably stable relative to specifications. This is not the whole story, of course, since offsets in the internal reference, and in the internal metrology, must be included in the estimate of time to 1% OOS. Such additions will shorten the time. Removing offsets, drifts and variability corrected by artifact cal, and offsets and noise contributed by the measuring system, will lengthen it. It is clear that one need not be concerned about the stability of the 5700A internal AV standard.

A similar analysis was performed for the other 5700A functions. Figures 16 through 20 present results for Alternating Current, Figure 21 results for Direct Current, Figure 22 for Resistance and Figure 23 for Direct Voltage. In all cases where the computation is possible, this worst-case analysis shows 5700A stability to be more than ample to support the instrument's most demanding specifications. As noted above, determining stability of low values of Direct Voltage will require further analysis.

The complete (preliminary) answer to the second question, then, is "yes!" -- artifact calibration does adjust the instrument nearly to nominal, and stability is such that properly adjusted instruments will maintain those adjustments for very long times. Does this mean we should extend 5700A calibration intervals indefinitely? No, it does not! Even the best instruments may occasionally malfunction in ways that are not obvious to their operators. For this reason, Fluke requires a complete verification of instrument performance each two years to assure that the internal metrology continues to operate as it was designed to do. Fluke has also provided the Cal Check function to provide control chart information as an aid to continuous monitoring of instrument functions.

CONCLUSIONS

A preliminary analysis of As Found and As Left data for over 250 5700As, obtained over a period of about five years provides important information about the instruments' performance to specifications. This preliminary worst-case analysis showed only five of 109 parameters to have greater than the expected 1% of instruments outside specified tolerances. The analysis is preliminary and worst-case because adequate information about the error contributions of the measuring system is not available at the time this report is written. A more definitive analysis will be the subject of a future paper.

The analysis provides evidence that artifact calibration does in fact operate as it was designed to do. Further, performance to specifications for those parameters where validity of data is not clouded by uncertainty about measuring system error shows the design to have been very conservatively specified, in accordance with long standing Fluke tradition.

Plotting deviations from nominal versus time since manufacture graphically demonstrates the instrument's time stability. Performing linear regression versus time provides the average time drift of the instrument and variability in the drift. Using this information, one can project the time from manufacture until 1% of instruments will drift out of tolerance. Except for low values of Direct Voltage, two combinations of alternating voltage/frequency and two values of resistance, time to 1% out of spec for all parameters can appropriately be expressed in decades rather than months or years. Again, in all cases, removing offset and variability contributions of the measuring system is expected to significantly improve these exceptions.

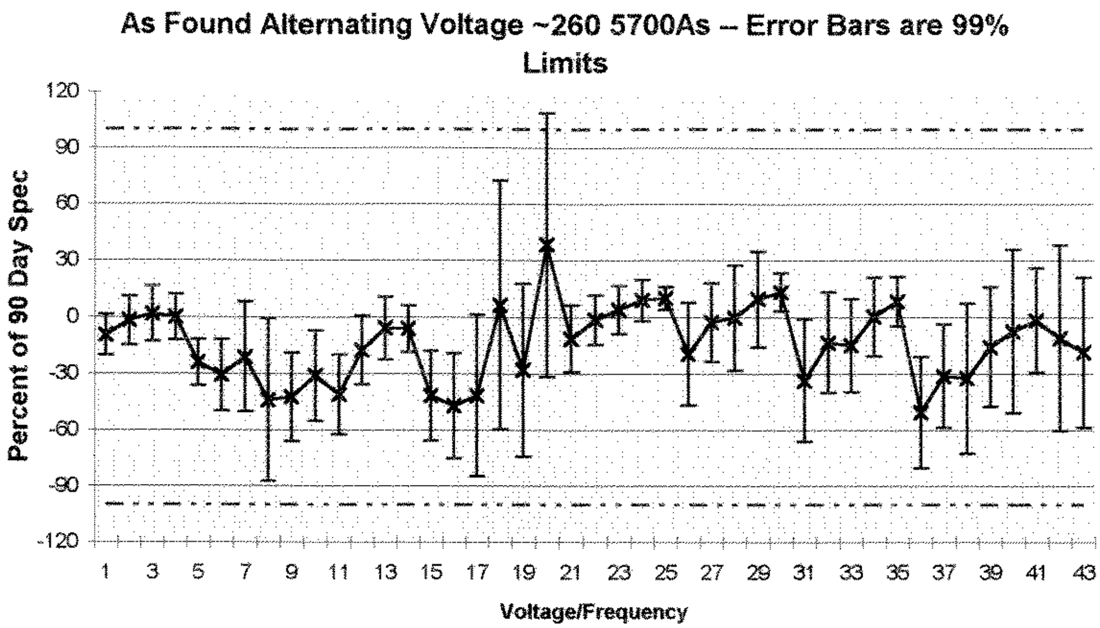

Figure 1. Deviations from nominal and 99% limits, all verified voltages and frequencies.

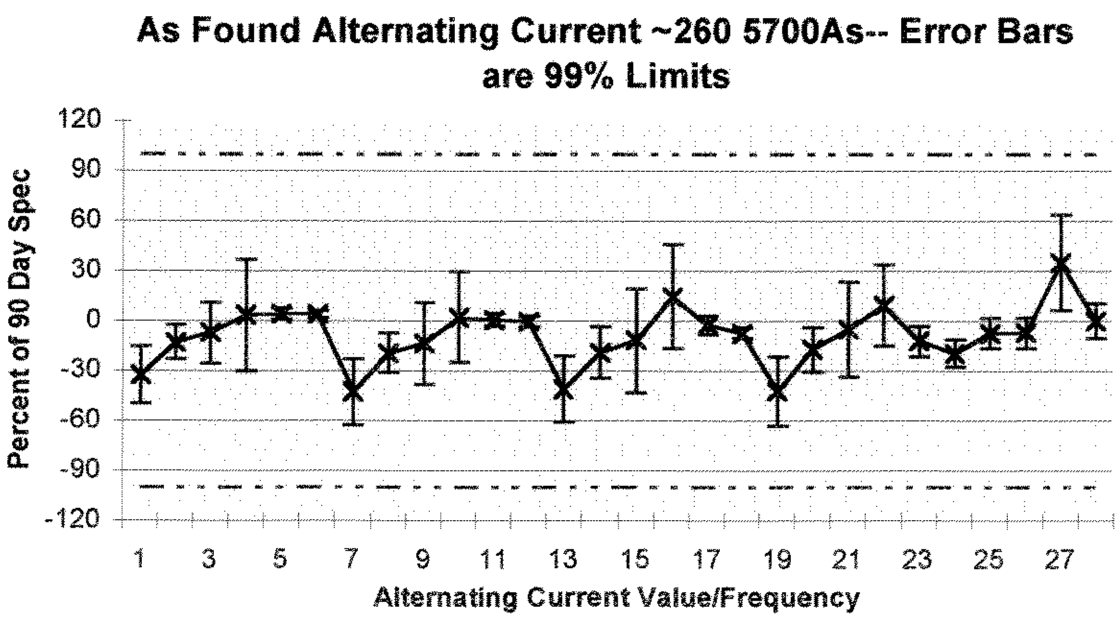

Figure 2. Deviations from nominal and 99% limits, all verified currents and frequencies.

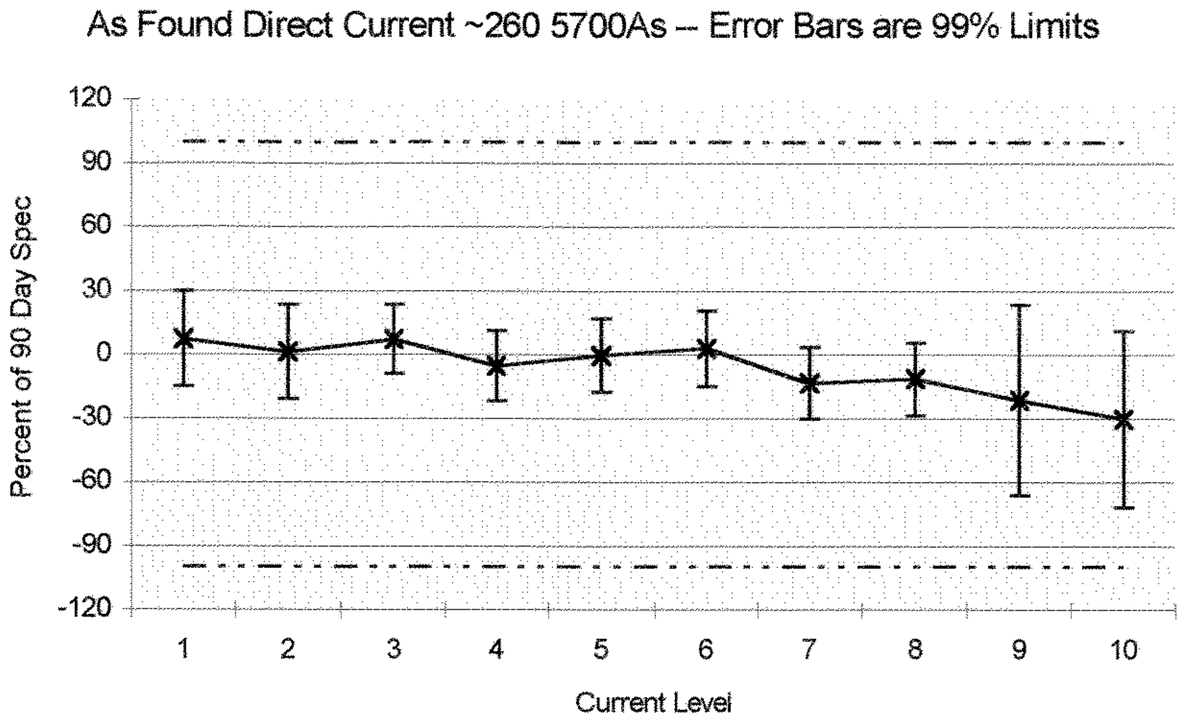

Figure 3. Deviations from nominal and 99% limits, all verified Direct Current levels.

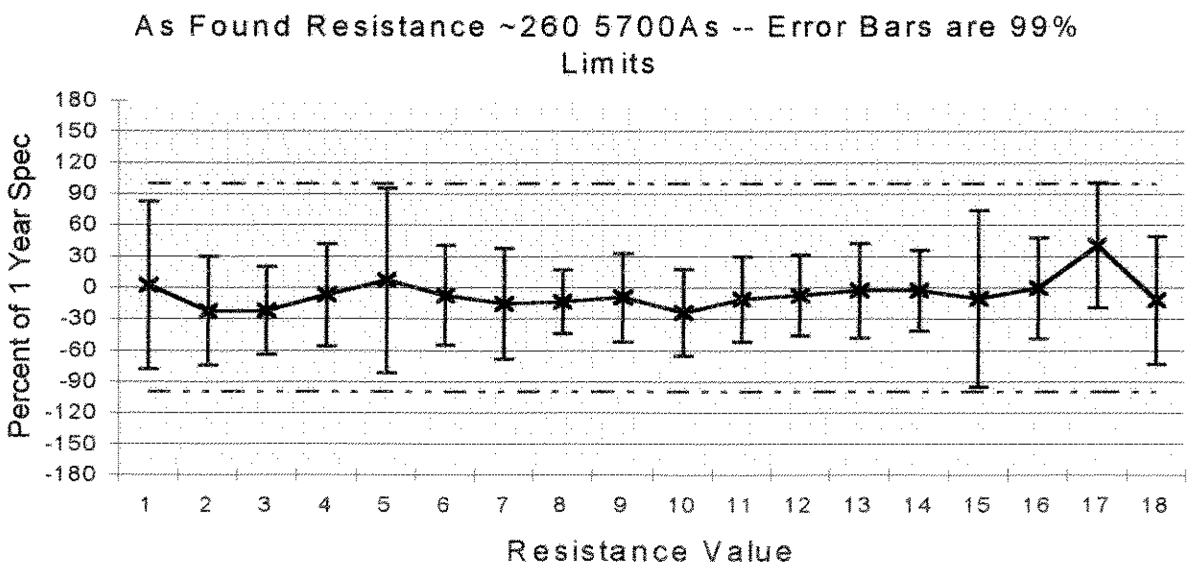

Figure 4. Deviations from nominal and 99% limits, all Resistance values.

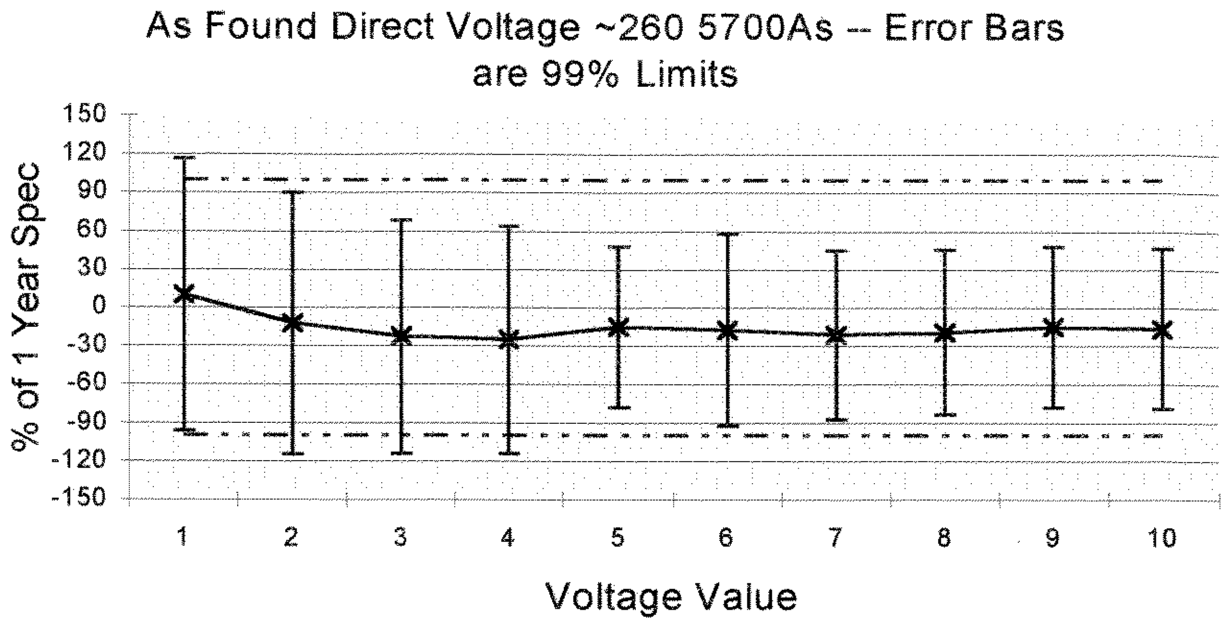

Figure 5. Deviations from nominal and 99% limits, all verified Direct Voltages.

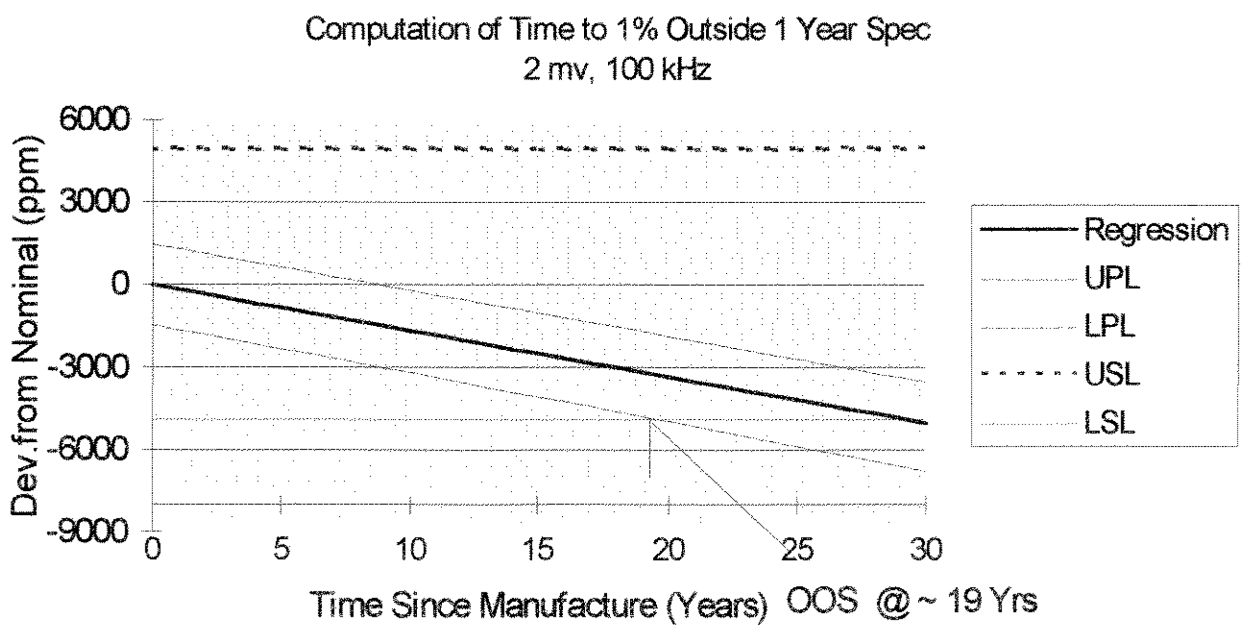

Figure 6. Computation of time to 1% outside spec limits.

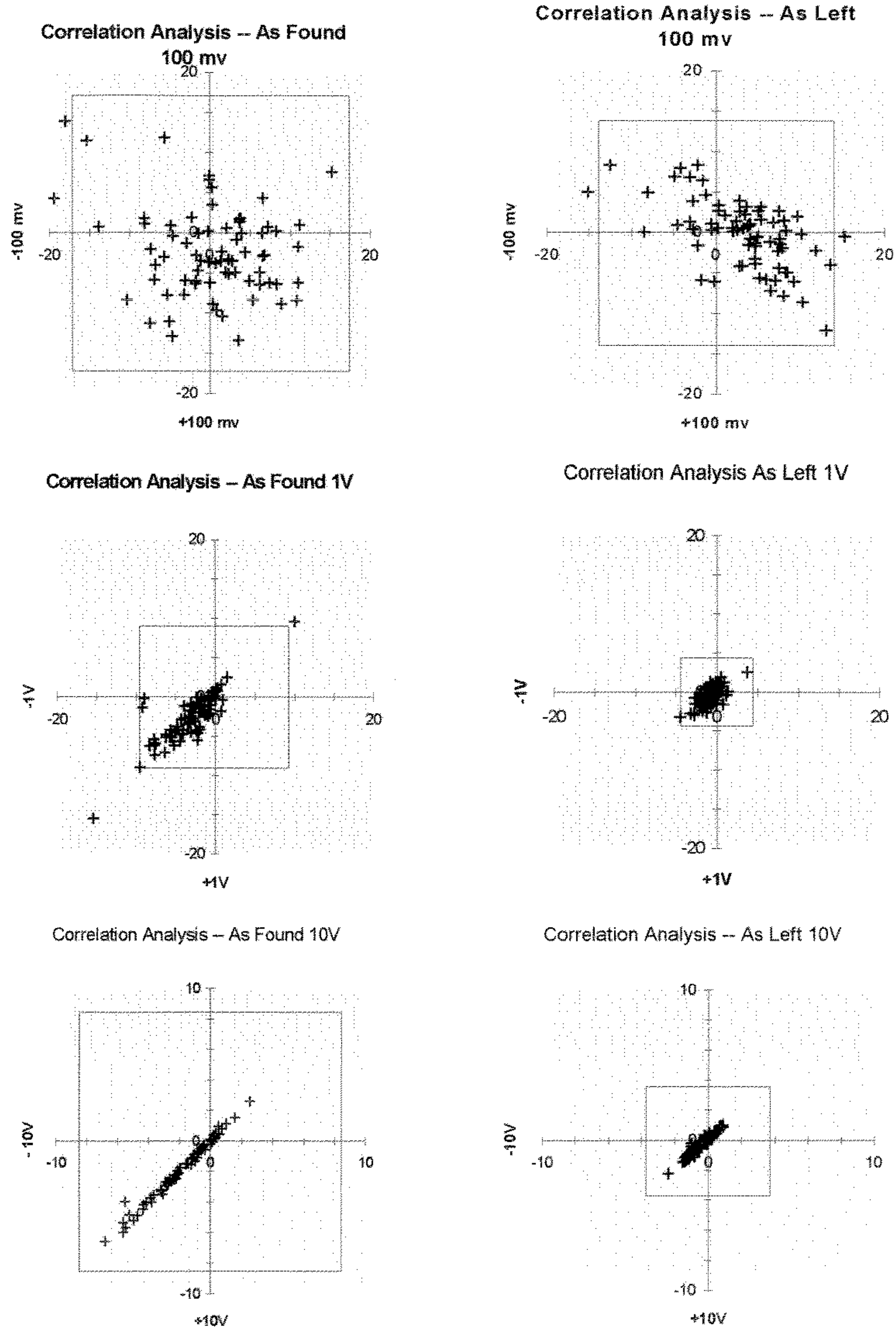

Figure 7. ~72 "Not Repaired" 5700As. Scale values are ppm. Boxes show 1 year spec, As Found and 24 hour spec, As Left.

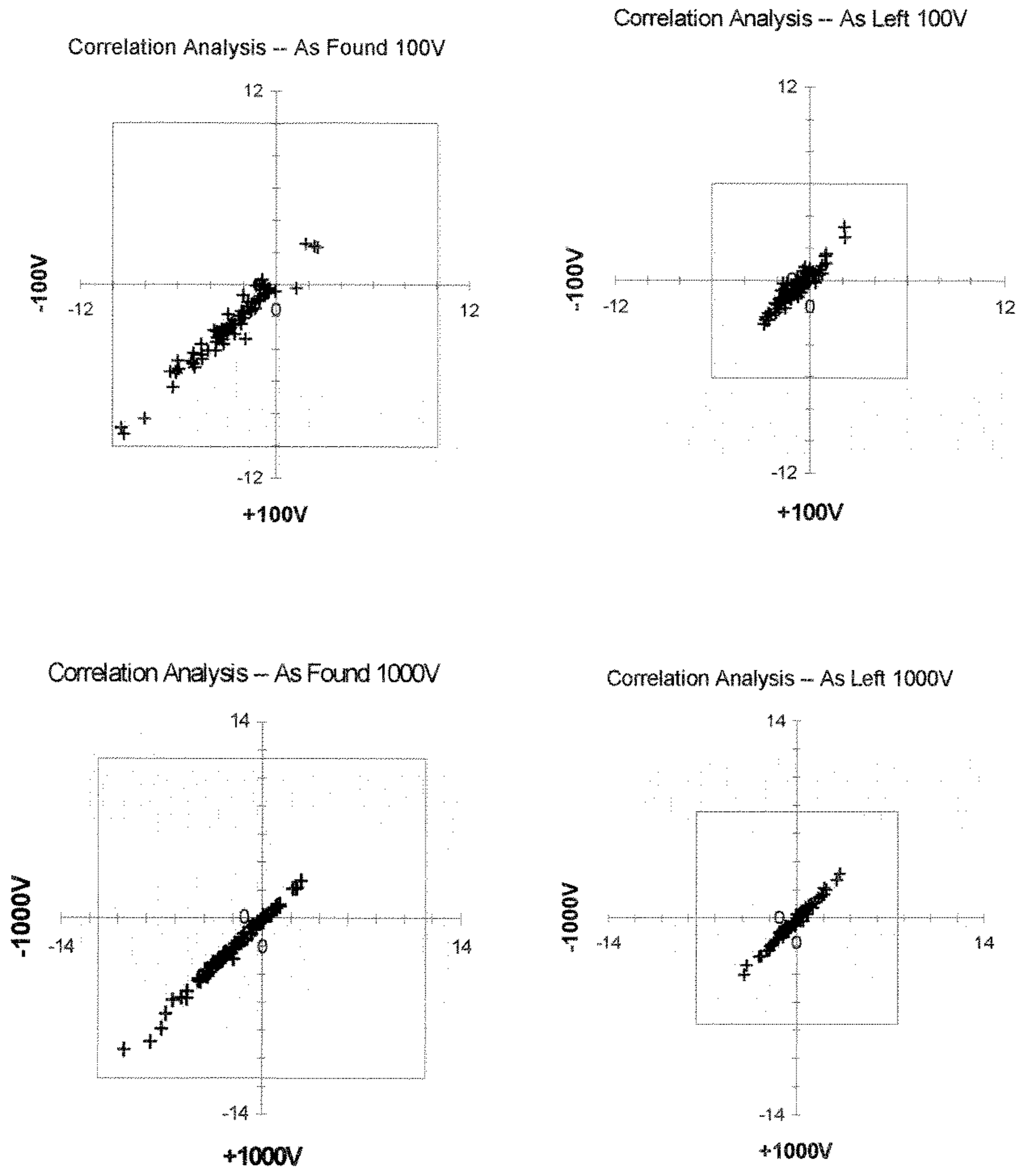

Figure 8. ~72 "Not Repaired" 5700As. Scale values are ppm. Boxes show 1 year spec, As Found and 24 hour spec, As Left.

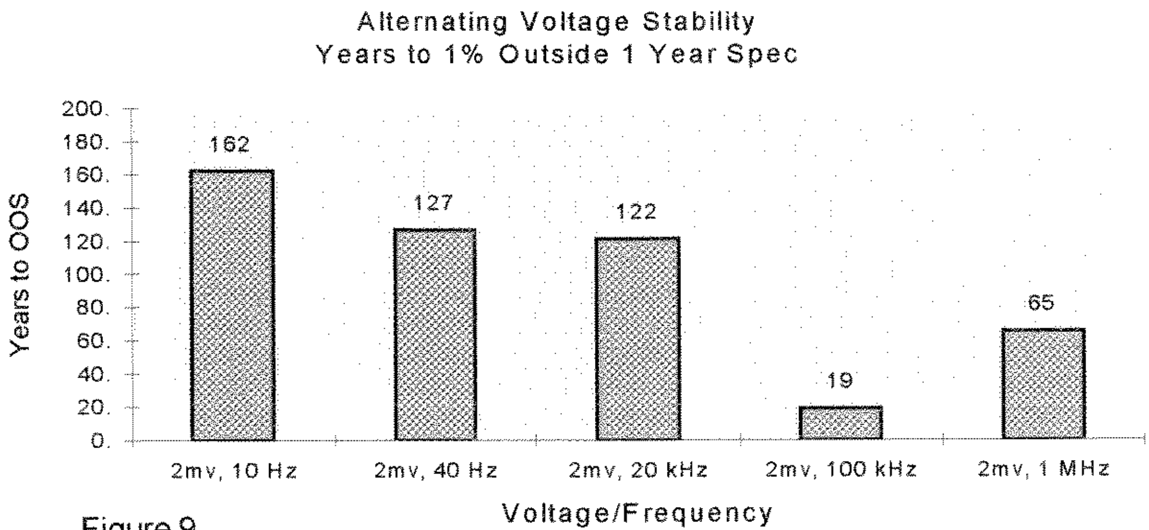

Figure 9.

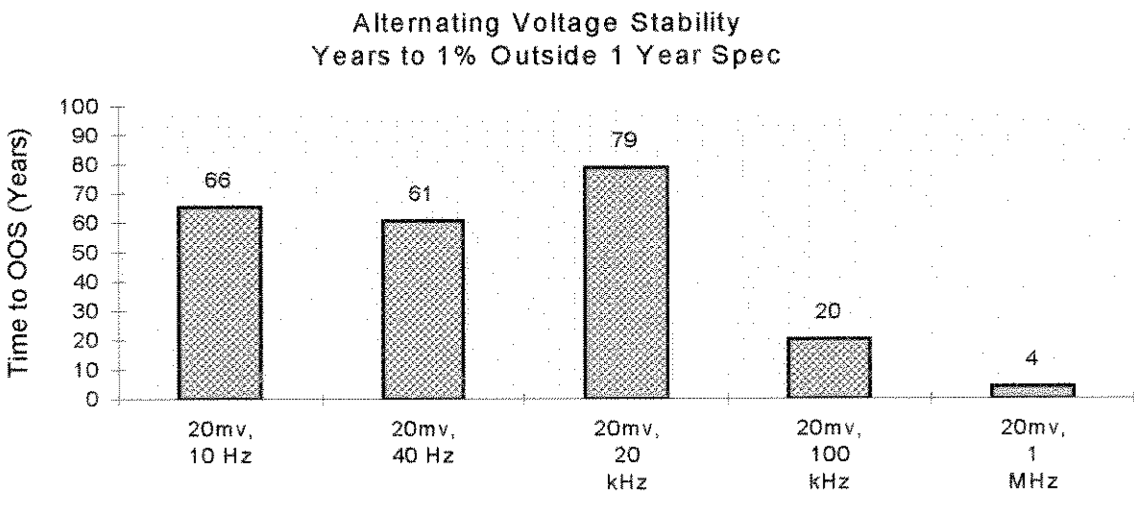

Figure 10.

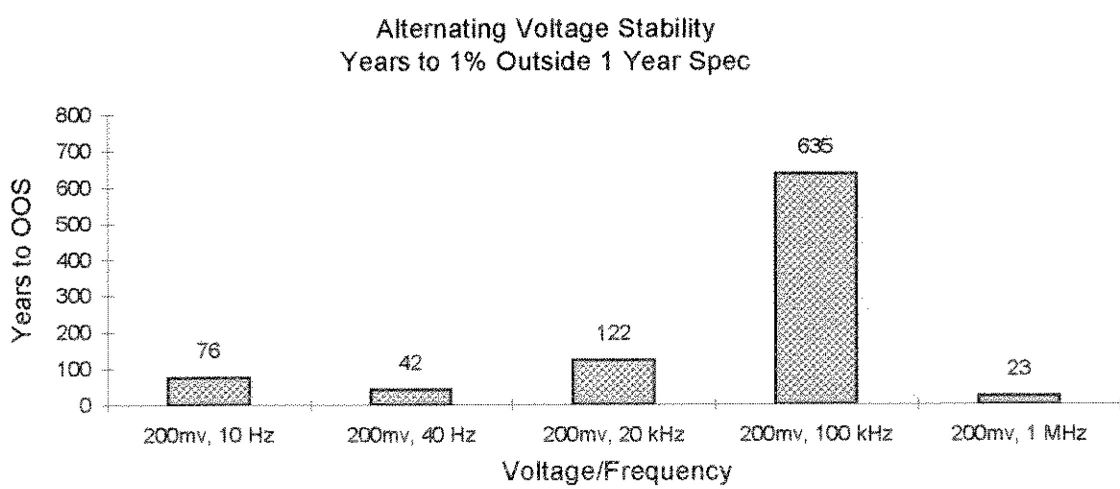

Figure 11.

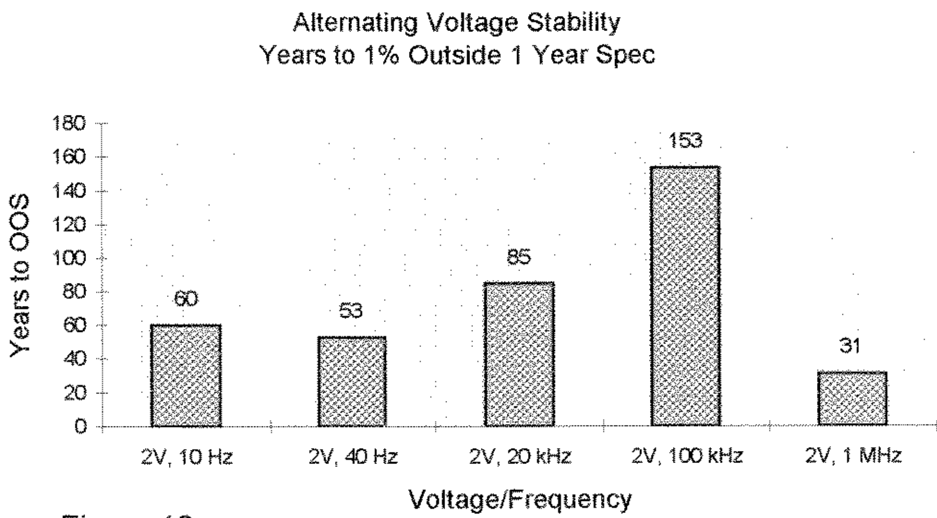

Figure 12.

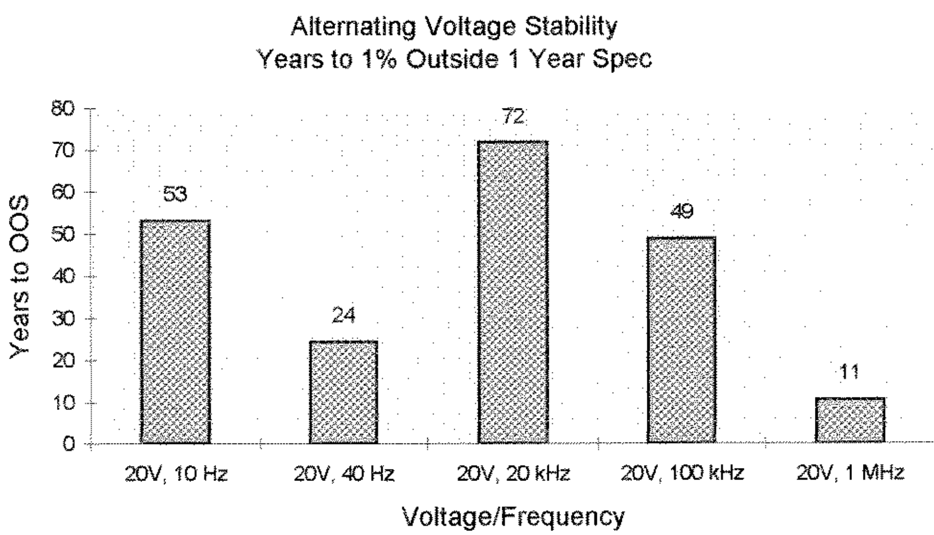

Figure 13.

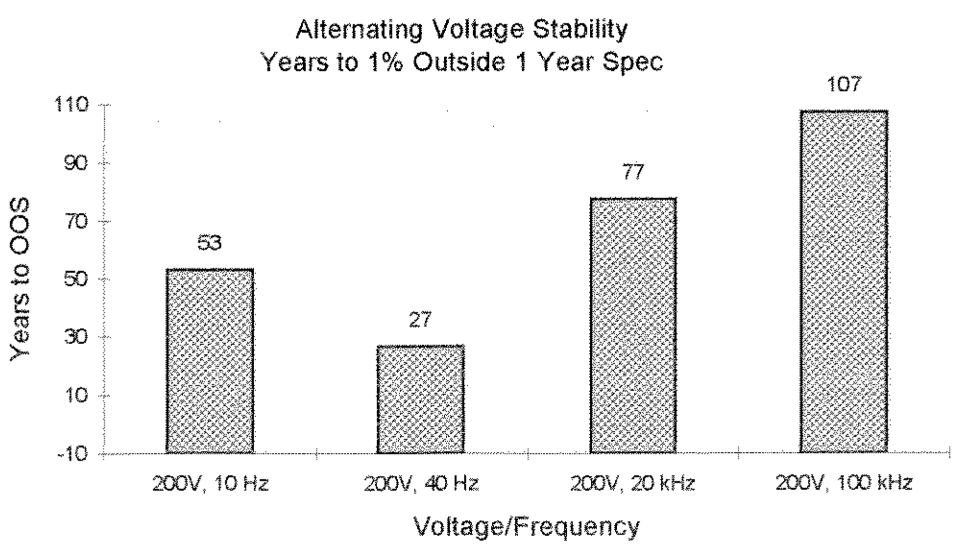

Figure 14.

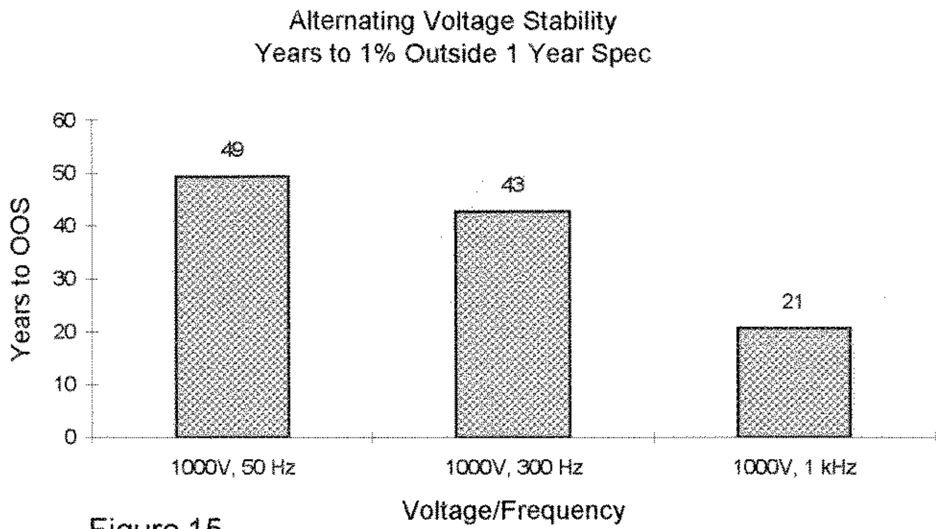

Figure 15.

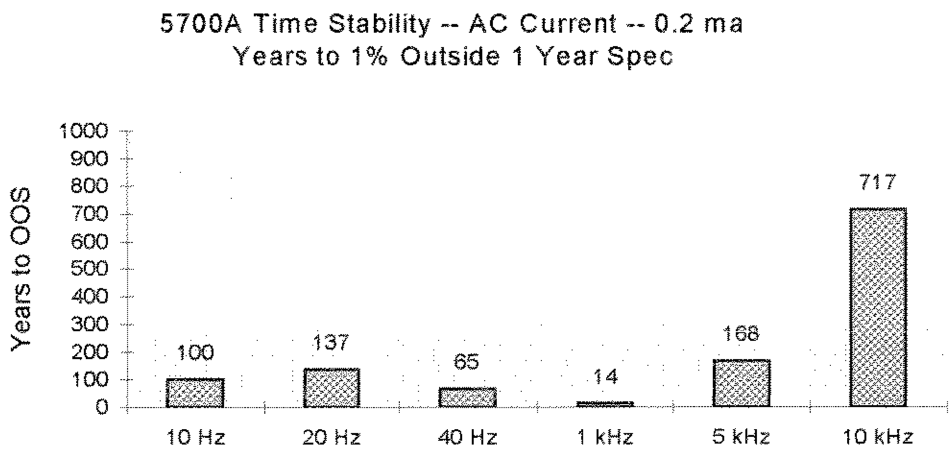

Figure 16.

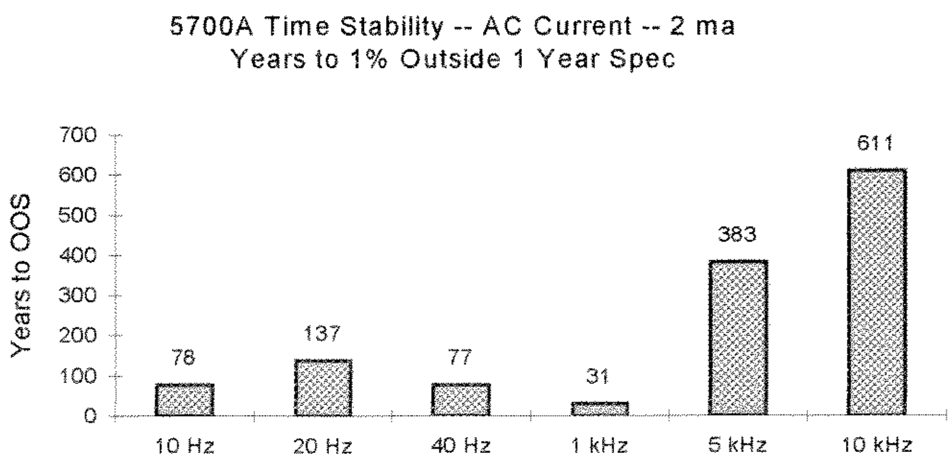

Figure 17.

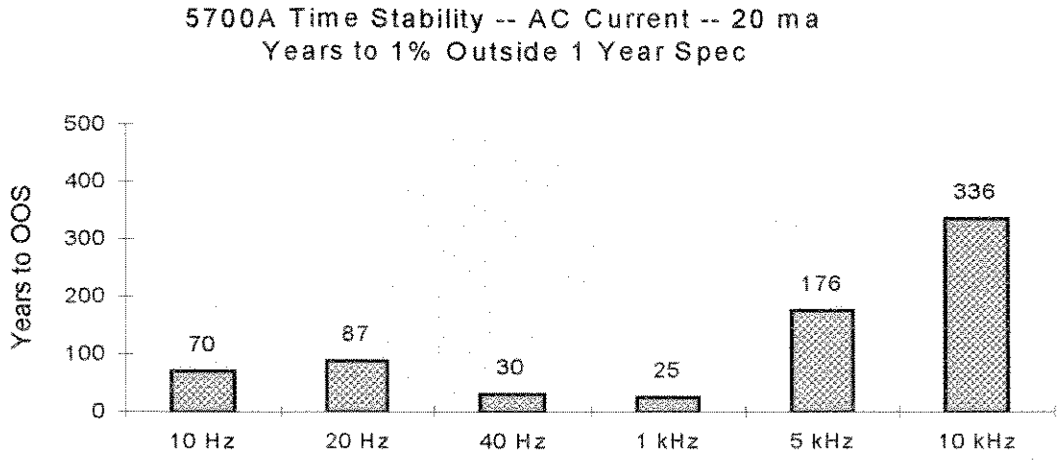

Figure 18.

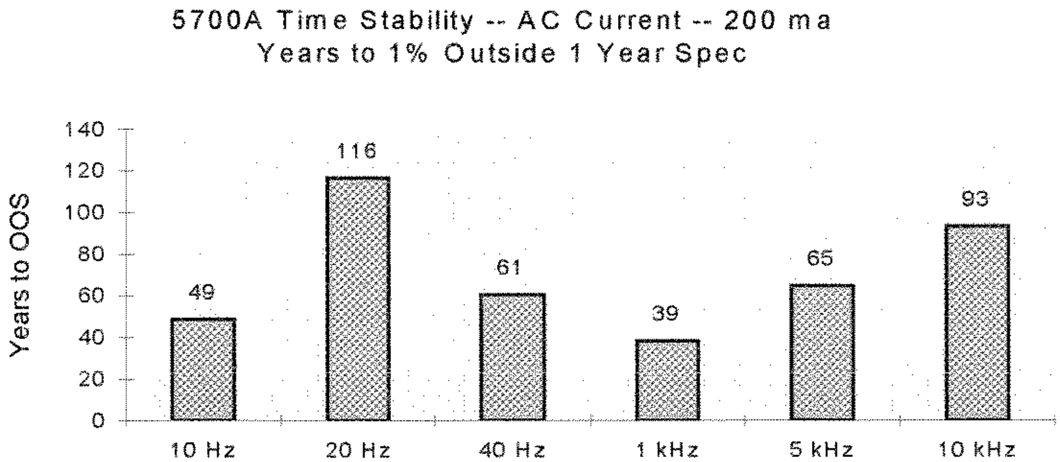

Figure 19.

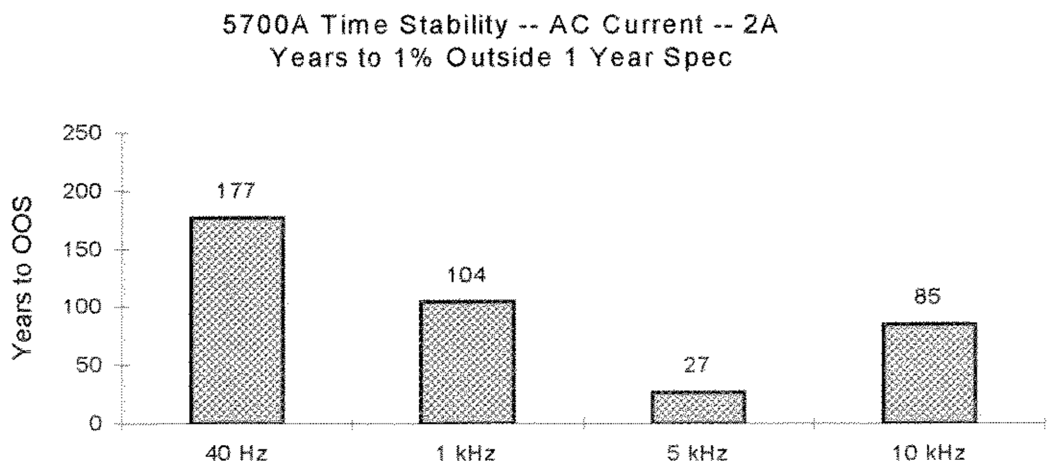

Figure 20.

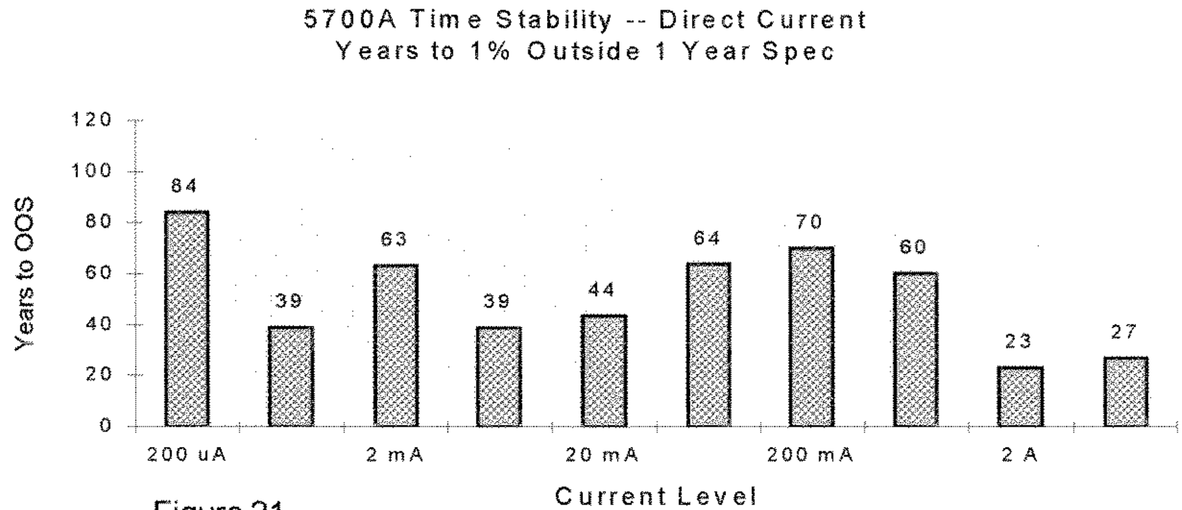

Figure 21.

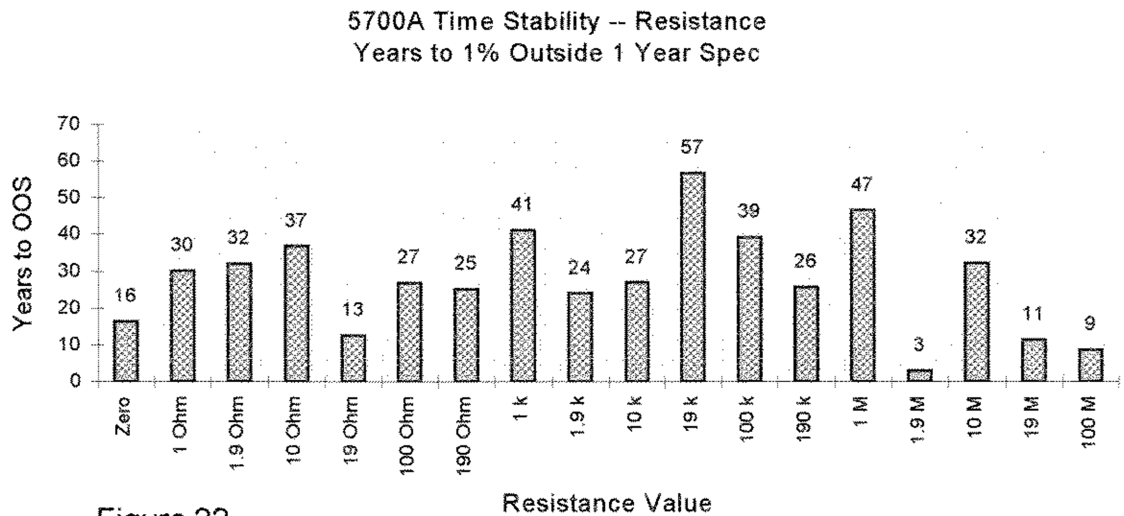

Figure 22.

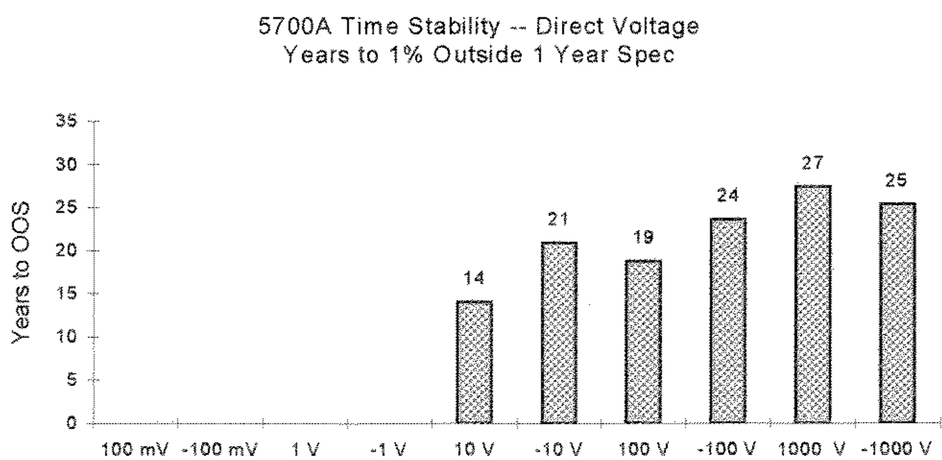

Figure 23. Data for lower voltages has too much variability for a solution.