Technical Review of Measuring and Calculating Balancing Authority Frequency Response in Interconnections with More than One Balancing Authority – April 27, 2011

Comments by Howard Illian

May 12, 2011

Introduction:

Mike Potishnak and other parties who support his efforts should be encouraged to continue with the publication of documents that provide the technical basis for the decision-making supporting the NERC Standards, Policies and Recommendations. Documents of this type are a necessary part of any logical standards development process.

This presentation of some of the work Mike has been performing in support of the Frequency Response Standard Drafting Team and the Resources Subcommittee is well written and understandable in general. It creates a good basis for continuing the discussions of how to implement the measurements necessary to create a new Frequency Response and Frequency Bias Standard.

The following comments are intended to address some issues the technical paper Mike offered addresses. At the same time it is important to insure that the technical decisions made using this work are the correct decisions for the desired result: a Frequency Response and Frequency Bias Standard that will properly support reliability of the interconnections, and the Balancing Authorities that make up those interconnections. Additionally, these decisions will insure the effectively coordinated operations of the interconnections through the use of the appropriate values for Frequency Bias in the Tie-line Bias control equations and systems.

The Big Picture:

The Frequency Response Standard Drafting Team has the charge of developing a Frequency Response and Bias Standard will achieve three important goals. These three goals are significantly different from the single goal envisioned for this standard a few years ago when this effort was initiated and the standard drafting team was formed.

1. The standard must insure that sufficient Frequency Response is delivered to the interconnections to insure an appropriate level of interconnection reliability associated with Frequency Response.
2. The standard must insure that each Balancing Authority delivers its minimum share of Frequency Response to insure that the interconnections will remain reliable, and that each region of the interconnection will also remain reliable.

3. The standard must insure that each Balancing Authority implements a Frequency Bias Setting in their Tie-line Bias control equations that will insure the properly coordinated control among the Balancing Authorities on each interconnection.

The principal short coming with the technical work presented by Mike isit primarily addresses the third goal and fails to address the first two goals of this latest goal set. This is demonstrated in that the discussion of the measurement of Frequency Response only addresses the problem in the context of the measurement of Frequency Response for a single Balancing Authority. One only need look at the draft standard offered to the industry for comment to recognize the weakness in the measurement approach investigated in the technical discussion.

In the draft standard, a minimum Frequency Response is determined for the interconnection and then allocated among the BAs using an as yet to be determined allocation method. The compliance for each BA is then determined based upon whether or not that BA supplied their share of that allocated minimum Frequency Response.

In my judgment, the most important characteristic of measuring Frequency Response to support this standard is the characteristic of the parts summing to the whole. This type of system is called a linear system. The standard must be constructed on a linear system because the starting point is the minimum Frequency Response maintains the desired level of reliability at the interconnection level. This minimum amount of Frequency Response is then allocated among the BAs using some method of apportionment. If the system is not linear, then it is impossible to determine how much Frequency Response each individual BA must contribute to achieve the minimum amount for the interconnection. Under these conditions, a standard cannot assure reliability because it cannot assure that the sum of the individual BA contributions to Frequency Response will provide the minimum interconnection Frequency Response.

Unfortunately, the technical discussion was silent on this consideration. When one examines the three measurement methods investigated: median, mean and linear regression, one finds that the median measurement cannot be used to implement a linear system. Therefore, one cannot be assured that the sum of the individual BA measurements of Frequency Response will equal the interconnection Frequency Response. On the other hand, both mean and regression measures can be used to implement linear systems and those measurement methods can both provide assurance that the sum of the individual BA measurements is equal to the interconnection measurement.

This characteristic of sum ability of individual BA responses to the interconnection response is a necessary condition for the standard as currently drafted. If the individual BA responses do not sum to the interconnection response, it is possible for the interconnection to have insufficient Frequency Response while all BAs on that interconnection have sufficient Frequency Response. It is also possible for all of the BAs on the interconnection to appear to have insufficient Frequency Response while the interconnection has sufficient Frequency Response. Either of these two conditions

will result in a standard that is ineffective. Using a linear system as the basis for measurement is the only way to insure that this condition will not occur.

It is often suggested since the data contains errors, it is unimportant that the sum of the parts equal the whole because those errors will cause differences between the sum of the parts and the whole. This was the logic that industry used before Dr. W. Edwards Deming began developing quality control in Japan. The object of quality control is to reduce or eliminate these errors and improve the quality of the products produced. However, the reduction and elimination of error will only improve the end product in a linear system. If we chose to design the system so that the sum of the parts does not equal the whole, the reduction or elimination of error will have an unpredictable effect on the result. Sometimes it will improve the system and sometimes it will degrade the system. When the reduction or elimination of error has unpredictable results, all incentive to reduce or eliminate error is removed. This makes the measurement system unreliable and of little value.

Specific Analysis:

The following sections provide comments on the analysis provided to support the specific conclusions.

"It is highly recommended that BA's not be allowed to provide corrections for changes in scheduled net interchange."

This conclusion is well supported. In addition, it is a necessary condition for maintaining a linear system for measurement, evaluation and compliance.

"It is conclusive that all BAs need to use a common methodology to avoid 'cherry picking'."

This conclusion is well supported. In addition, it is a necessary condition for maintaining a linear system for measurement, evaluation and compliance.

"It is conclusive that some AGC action will appear as frequency response with the use of the 20 to 52 second metric."

This conclusion is well supported. However, evaluation of a single interconnection will not provide sufficient data to determine the appropriate common sampling interval for all four interconnections. The choice of the common sampling interval is based upon more than just the inclusion of some AGC response in the data. Other considerations were included in the choice of the 20 to 52 second interval.

The choice of a common sampling interval for all four interconnections should be based on analysis of all four interconnections. This is a choice that if appropriate, should be evaluated and modified based on information developed during the Field Trial.

"While more analysis is needed for many more BAs, this may be an indicator that the 20 to 52 second metric will flag squelched response."

This is one of the considerations that may influence the choice of the sampling interval and this should be investigated during the Field Trial.

"While more data is needed for more BAs, there does not seem to be a need to create special metrics based on the peak period, nor should sampling be restricted to on peak periods as was done in prior practices."

Although the data evaluated seems to support this conclusion, the data is limited and only for a single interconnection. This conclusion would also apply to the event selection. This should be investigated in more depth with all four interconnections during the Field Trial.

"Summarizing, the results are sensitive to the method used to choose time zero – largest is easier to administer, but more research is needed in this area."

This conclusion is well supported. More research on the determination of event time zero should be performed during the Field Trial.

"In summary, it seems that events with point B above 59.95 Hz will result in lower scores, but a larger set of BAs are needed to make any firmer quantitative statements. It may make sense to have the majority of frequency events having a point B below 59.95 Hz."

This conclusion is only a guess and not well supported by the analysis. More evaluation should be performed to determine the effect. For example, the effect could be due to larger versus small events. This issue needs to be investigated further during the Field Trial.

"In summary, more analysis is needed to evaluate the impact of time skew related to the point A sampling interval."

This is not a conclusion, it is only a suggestion. I believe that the SDT should investigate this in more detail during the Field Trial.

"In summary, many BAs need to provide scan rate data from 60 seconds before to 90 seconds after the frequency event to support an analysis to determine the best (or maybe just the least evil) sampling interval."

This is simply a recommendation to acquire the necessary data for further analysis during the Field Trial. The SDT needs to be careful that it does not adjust the measurement method for correctable problems within the individual BA EMSs. When the standard measurement method is modified in this way, incentives are removed that would encourage the BA to improve its own systems.

"In summary, it is impractical to rely on a single sample to be above some threshold value, and frequency response values will be meaningful ONLY when used with some type of averaging technique, and the choice of

averaging technique will impact the accuracy of the score, given the data quality problems associated with using actual net interchange values."

Although I agree with this statement, the data provided does not support this conclusion. There is no sound technical basis for removing values below zero from the data. As a consequence, no conclusions can be drawn from the analysis. In a signal with noise problem, if the noise has a larger variation than the signal size, it is logical that some values will show a negative signal. Selective elimination of these individual samples simply biases the result and does not improve the result.

"In summary, 25 samples should be a sufficient sample set size, provided that the right choices are made for the sampling interval and the averaging technique."

This conclusion is supported by the analysis and is consistent with recommendations derived from statistics (30 samples).

"In summary, analysis of the symmetry of contaminated data performed thus far does not at all support an assumption that contamination will balance out for a 25 event sample size."

This conclusion is unsupported. The exclusion criteria are not repeatable. They are based on the judgment of four subject matter experts without further definition. The exclusion criteria are based upon the final calculated value of the Frequency Response instead of the source data used to calculate the Frequency Response. Symmetry has a far different meaning in the context of mean as compared to linear regression. The correct way to investigate symmetry is to perform a density analysis on the data and investigate the shape of the density distribution. The definition of representative value for Frequency Response is not well defined. Is it a value that is most likely to occur or is it a value that is most representative of the risk that the interconnection will experience? One way to investigate the issue of contamination would be to compare the sum of the individual BAs response to the interconnection response for those same events. Differences between the sum and the interconnection value would indicate contamination, other variation would simple be noise and not contamination. More work should be performed to evaluate this issue during the Field Trial.

"In summary, the median is projected to be more resilient than either the mean or linear regression in the inevitable presence of contaminated actual net interchange data, and should be the averaging technique of choice unless a larger data set is analyzed in the future and a different metric is determined empirically to be more resilient to data contamination."

This conclusion is unsupported. It is unsupported for the same reasons given for the previous conclusion.

"In summary, the mean, median, and regression will yield very similar results once contaminated data is removed for all sampling intervals and event groupings."

This is unsupported as a conclusion. It is simply a corollary of the previous conclusion.

"In summary, based on the empirical data presently available, the median is the best averaging technique to use, but this should be checked again as more data becomes available for more BAs."

This is unsupported as a conclusion. It is simply a corollary of the previous conclusion.

"In summary, the sample set size is too small to be conclusive, but further research is warranted at a moderate priority."

Further research should be conducted at a high priority.

Conclusion:

In general, this work is beneficial for the SDT. Mike should be thanked for taking the time to perform the evaluations presented and writing this report. Without work of this nature being attempted by those on the front line in the industry, many of the issues we face will not be adequately addressed.

I agree with most Mike's conclusions with the exception of those that recommend the use of median. I will provide Mike with specific methods that can be used to evaluate whether or not the skew in the Frequency Response data is great enough to justify the use of the median rather than the mean or linear regression. The issues associated with linear regression for this specific data should be addressed by the SDT as a whole.