Applying the Feature Selective Validation (FSV) method to quantifying rf measurement comparisons

H.G. Sasse hgs@dmu.ac.uk
A.P. Duffy apd@dmu.ac.uk
Department of Engineering
De Montfort University
LE1 9BH UK

A. Orlandi Antonio.orlandi@univaq.it

EMC Laboratory
University of L'Aquila
Italy

Abstract: The FSV method was initially developed to provide an objective quantification of the data typical to the validation of computational electromagnetics in EMC applications. FSV has subsequently been embodied in IEEE Standard 1597.1: Standard for Validation of Computational Electromagnetics Computer Modeling and Simulations. The purpose of FSV is to provide a range of numerical and 'natural language' outputs from a global 'goodness of fit' value to a point by point analysis of the quality of the data comparison. From this, improvements can be quantified and areas that need to be addressed identified. The data to be compared often involves measured data. Hence, FSV can be applied to comparing measured data alone. The purposes of such comparisons include repeatability tests, aging analysis, 'portability' between facilities and assessing variations due to configuration changes. This paper presents the background to FSV, demonstrating that it compares favourably with the consolidated opinions of groups of engineers (in validation tests) and detail of its implementation. Through a series of data comparisons obtained from reverberation chamber tests, the interpretation of FSV results will be presented: showing the range of outputs from a single 'natural language' summary to a point-by-point analysis.

1. Introduction

This paper discusses the application of the Feature Selective Validation (FSV) method to measured data and provides a review of the range of outputs available from the technique that can be used for assessing such data. The reasons for this assessment are *inter alia* to quantify:

- Measurement repeatability – same measurement with the same personnel in the same facility.
- Measurement 'portability' - same measurement at a different facility.
- Personnel dependence - the effect of different people on the results.
- Method variation – the effect of small (or gross) changes to a configuration.

Where the results are structurally and visually simple, quantification of any comparison is straightforward. For example an antenna resonance may be *x*% or *y* MHz from the designed parameter or an amplifier may differ by *z* dB from its designed performance. However, where the results contain a number of resonance or resonant-type structures, such single valued assessment becomes more difficult and becomes more dependent on the experience of the engineers performing the analysis to be able to adequately assess such comparisons. Even then, it is common for a group of engineers to offer different assessments of the data based on their personal background and experiences; this can also be influenced by mood and other environmental and physiological conditions. However, the combined opinions and group experience will provide a very valuable assessment tool that emphasises agreed assessment but which reduces the impact of outliers.

The FSV technique was initially developed to provide quantification of comparisons of numerical modelling data against experimental or other modelled data for the purpose of validation. It is a heuristic approach that has now been included as the central comparison technique in IEEE Standard 1597.1 [1]. However, the fact that the nature of the data is the same, whether modelled or measured, indicates that FSV could be as useful where only measured data is considered.

This paper reviews the FSV method, presents some results to be analysed and applies FSV to those results in order to demonstrate the FSV method. First, a visual rating scale is presented that helps to define some of the common categories (natural language descriptors) used in the subsequent analyses.

2. Visual Rating Scale

One major problem with asking for an engineer's subjective, quantified, rating of any results is the lack of any definitions of 'good' or 'bad' (if using qualifiers) or '2' or '7' if using number scales. This leads to situations where individuals can be self-consistent but there are offsets between individuals. There is a general attraction to such terms as 'good' or 'bad' because humans like to categorise and are comfortable using natural language descriptors. This is something that FSV has adopted by categorising results in six 'bins' ranging from 'Excellent' to 'Very Poor'. In order for a group opinion to be captured adequately, some rating scale is required that minimises inter-subject variability. This was done [2] and the rating scale is presented in Figure 1.

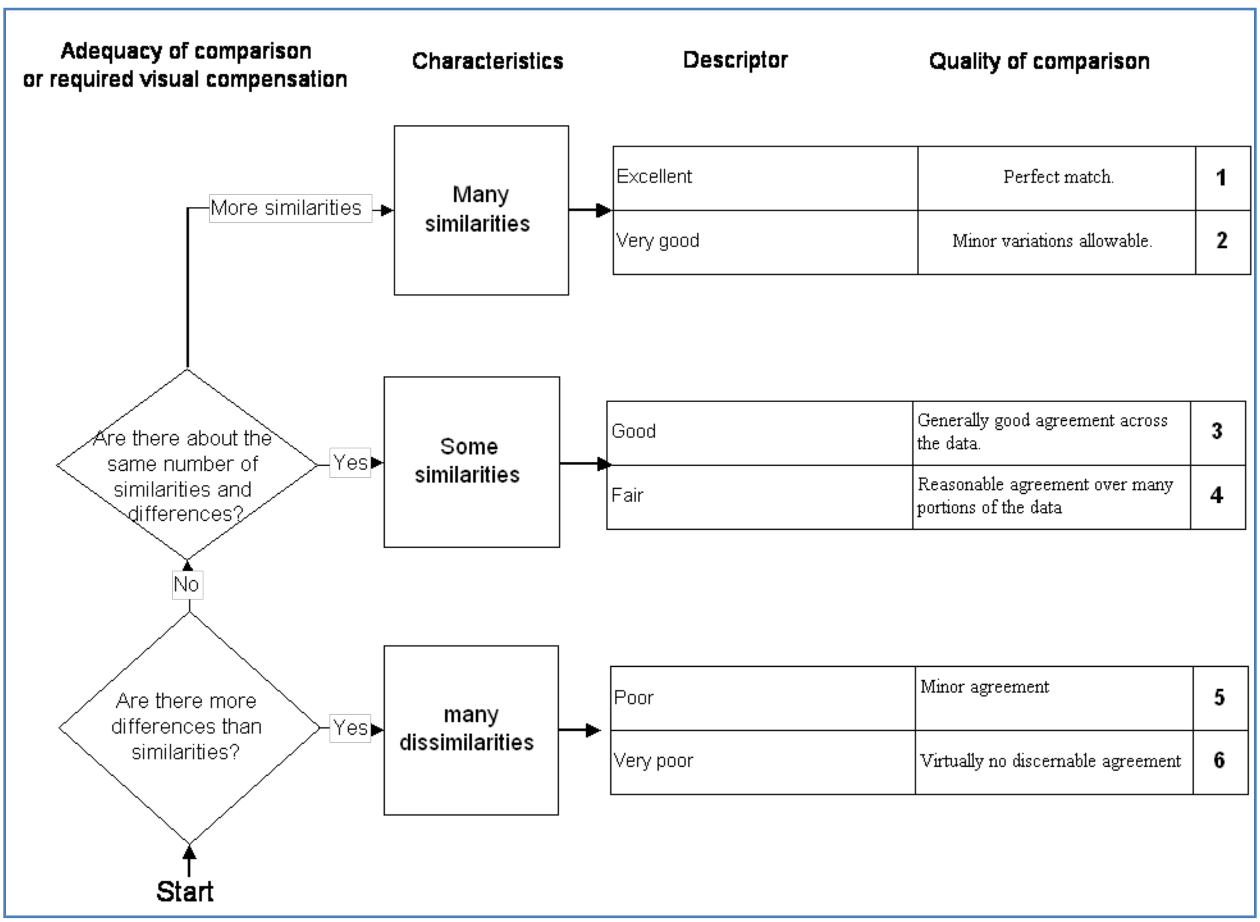

Figure 1. Visual rating scale

This scale relies on a series of binary decisions and has been shown to reduce the offset variations between engineers but not to homogenise the group opinion. This, effectively, provides definitions for terms such as 'good' and 'poor' as well as establishing a traceable approach to canvassing group opinion, which could be used to rate any comparison data or verify the performance of FSV or other comparison approaches.

3. The Feature Selective Validation (FSV) method

FSV is based on the creation of a number of quanta from the original data. The basis is that a majority of comparisons undertaken by humans rely on the assessment of the overall shape / envelope of the data and the comparison of the individual, fast moving, features [3][4]. The FSV measures are the Amplitude Difference Measure (ADM) and the Feature Difference Measure (FDM) which can be combined into a Global Difference Measure (GDM) which is an overall goodness-of-fit measure. Those point-by-point comparisons are constructed as described below. The average value of these provides single value summaries that are helpful to attribute an overall quality to the comparison and use as the basis for rank-ordering several comparisons. There are several other information sets that can be extracted from FSV and these are described below and illustrated in the next section.

Implementation of the FSV method is based on the following procedure:

1. Determine the region of overlap for the two data sets to be compared and ignore everything outside this window. This is because it is inappropriate to extrapolate such data.
2. Resample the data to the lowest point density such that the data points of the two data sets to be compared are coincident.
3. Filter into three ranges. This is done by Fourier Transforming the data and the 'DC' set is the DC term plus the lowest four data points. The 'Lo' region is from the next point to the approximate location where 40% of the total area under the transformed curve occurs. The 'Hi' region is everything else. These regions are then transformed back to the original domain.
4. Create an Amplitude Difference Measure (ADM) from the normalized difference of the low pass data. This can be represented as a point-by-point graph or as a single goodness-of-fit value by taking an average over the domain of interest. The ADM is given in equation (1), Where Lo_1 and Lo_2 are the low pass filter components of the original data and N is the total number of data points being considered.

$$ADM(n) = \left|\frac{\alpha}{\beta}\right| + \left|\frac{\chi}{\delta}\right| \exp\left\{\left|\frac{\chi}{\delta}\right|\right\} \tag{1}$$

where

$$\alpha = \left(\left|Lo_1(n)\right| - \left|Lo_2(n)\right|\right)$$

$$\beta = \frac{1}{N}\sum_{i=1}^{N}\left(\left(\left|Lo_1(i)\right| + \left|Lo_2(i)\right|\right)\right)$$

$$\chi = \left(\left|DC_1(n)\right| - \left|DC_2(n)\right|\right)$$

$$\delta = \frac{1}{N}\sum_{i=1}^{N}\left(\left(\left|DC_1(i)\right| + \left|DC_2(i)\right|\right)\right)$$

5. Create a Feature Difference Measure (FDM) from the normalized difference of a combination of derivatives of the band and high pass data. This can be represented as a point-by-point graph or as a single goodness-of-fit value. This is given in equation 2.

$$FDM(n) = 2(| FDM_1(n) + FDM_2(n) + FDM_3(n)) \quad (2)$$

where

$$FDM_1(n) = \frac{|Lo_1'(n)| - |Lo_2'(n)|}{\frac{2}{N}\sum_{i=1}^{N}\left(\left(|Lo_1'(i)| + |Lo_2'(i)|\right)\right)}$$

$$FDM_2(n) = \frac{|Hi_1'(n)| - |Hi_2'(n)|}{\frac{6}{N}\sum_{i=1}^{N}\left(\left(|Hi_1'(i)| + |Hi_2'(i)|\right)\right)}$$

$$FDM_3(n) = \frac{|Hi_1''(n)| - |Hi_2''(n)|}{\frac{7.2}{N}\sum_{i=1}^{N}\left(\left(|Hi_1''(i)| + |Hi_2''(i)|\right)\right)}$$

6. Combine the ADM and FDM to give a Global Difference Measure (GDM) (both as point-by-point and as an overall goodness-of-fit measure). This is given in equation 3.

$$GDM(n) = \sqrt{(ADM(n))^2 + (FDM(n))^2} \quad (3)$$

7. Rate the ADM, FDM and GDM summary values according to the natural language descriptors in Table I.

Table I Natural language descriptors

FSV value (quantitative)	FSV interpretation (qualitative)
Less than 0.1	Excellent
Between 0.1 and 0.2	Very good
Between 0.2 and 0.4	Good
Between 0.4 and 0.8	Fair
Between 0.8 and 1.6	Poor
Greater than 1.6	Very poor

8. Calculate the proportion of the point-by-point difference measures that fall into the categories listed in Table I. These are the confidence histograms that are used to estimate the collective assessment of a group of engineers.
9. Calculate the Grade and Spread of the confidence histograms [5]. The Grade is the number of adjacent categories that include at least 85% of the cumulative data in the confidence histogram and Spread is the number of categories (starting with Excellent) that need to be summed in order to achieve a cumulative 85%. The basis of these is that as the mean used to obtain the FSV single figure values is calculated and used in the way a statistical mean is, Grade is an FSV analogue of statistical variance and Spread is an FSV analogue of statistical skew. They can also be used to automatically weight the ADM and FDM in the calculation of the GDM. Equation 3 assumes that the ADM and FDM have equal significance. However, it is possible to use the Spread, which gives

a measure of how much confidence can be placed in each of the measures, to establish a better estimate of that significance. The approach is:

```
If
```
$$\{\mathrm{Spread}_{\mathrm{ADM}} < \mathrm{Spread}_{\mathrm{FDM}}\}$$

```
Then
```
$$\mathrm{K}_{\mathrm{ADM}} = 1$$

$$\mathrm{K}_{\mathrm{FDM}} = \mathrm{Spread}_{\mathrm{ADM}} / \mathrm{Spread}_{\mathrm{FDM}}$$

```
Else if
```
$$\{\mathrm{Spread}_{\mathrm{ADM}} > \mathrm{Spread}_{\mathrm{FDM}}\}$$

```
Then
```
$$\mathrm{K}_{\mathrm{ADM}} = \mathrm{Spread}_{\mathrm{FDM}} / \mathrm{Spread}_{\mathrm{ADM}}$$

$$\mathrm{K}_{\mathrm{FDM}} = 1$$

```
Else
```
$$\mathrm{K}_{\mathrm{ADM}} = 1$$

$$\mathrm{K}_{\mathrm{FDM}} = 1$$

```
End if
```
$$GDM = \sqrt{(K_{ADM} \cdot ADM)^2 + (K_{FDM} \cdot FDM)^2}$$

4. Example

In order to demonstrate the FSV performance, the following example takes the real part of the reflection coefficient of a monopole antenna in a reverberation chamber [6] with two different stirrer positions. The results were taken from part of the antenna's efficiency measurement. The purpose of the comparison undertaken here is to ensure that there is difference between the two measurements (the frequency range of the measurements was 120 – 200 MHz and the notional lower working frequency of the chamber is 180 MHz – although it should be recognised that this value is based on cumulative frequency density). The reason for undertaking this comparison is to determine if there is noticeable mode stirring / movement at all frequencies. The monopole resonant frequency was approximately 160 MHz. The stirrer positions used were the reference position and 18° (step 10 out of 200 stirrer positions for a full revolution). The original data is shown in Figure 2.

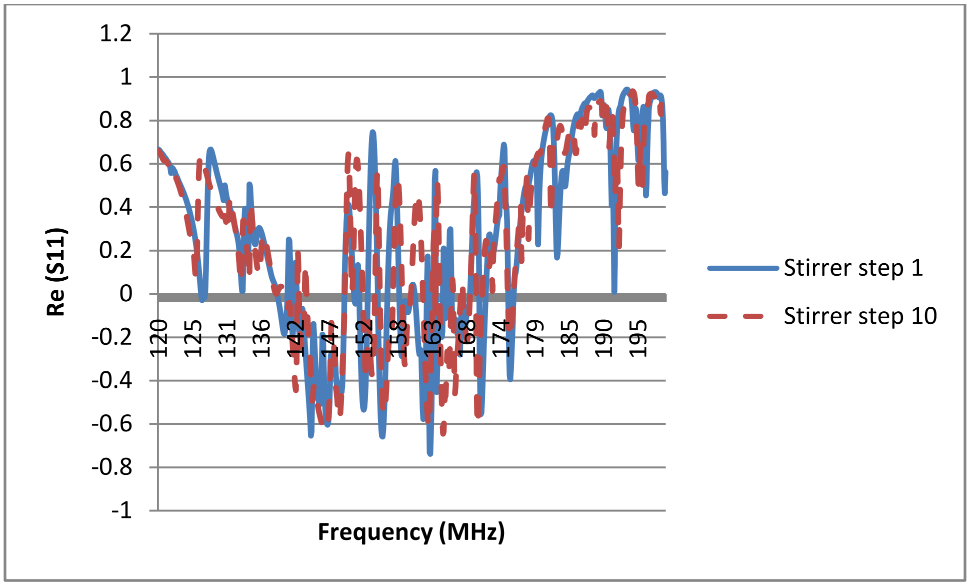

Figure 2 Original data for comparison (real part of reflection coefficient) over two different positions.

While it is evident that there are some differences between the two data sets, it is difficult to quantify this. In order to try to quantify the differences, the point-by-point comparisons were undertaken resulting in the data of figures 3 – 5

In order to help interpret these data, it should be remembered that 'Very Poor' is anything above 1.6, 'Poor' is anything between 0.8 and 1.6 and 'Fair' is between 0.4 and 0.8. These descriptions are only helpful categorisations and not indications of levels to be set for agreement. In this case, we could arbitrarily say that, because we are looking for stirring and therefore a poor level of agreement and we are only taking two stirrer positions (out of 200) that anything Fair or worse is acceptable indication of mode movement: i.e. anything higher than the 0.4 line (as in [8]) is a good enough indication.

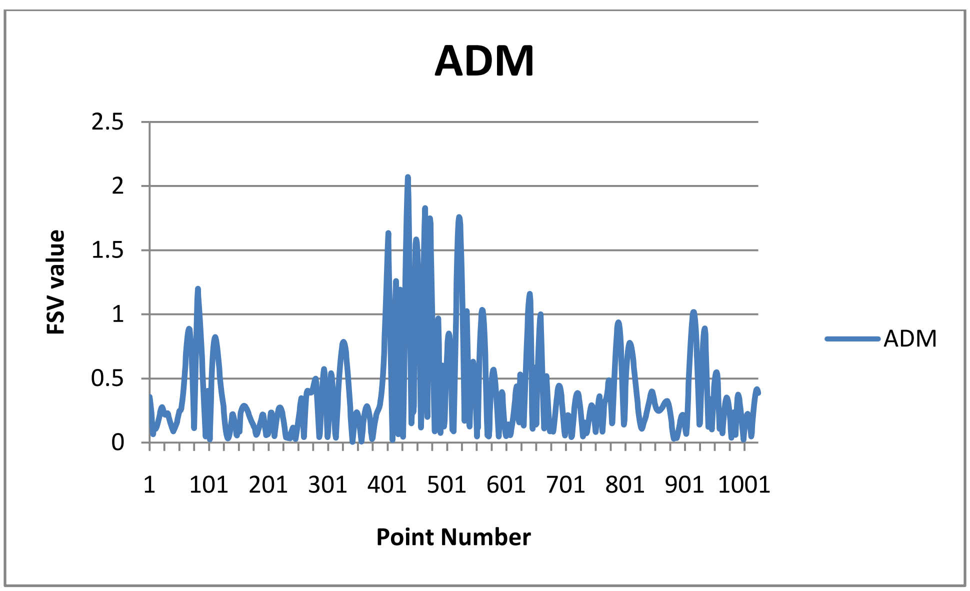

Figure 3 Amplitude Difference Measure of the data in Figure 2

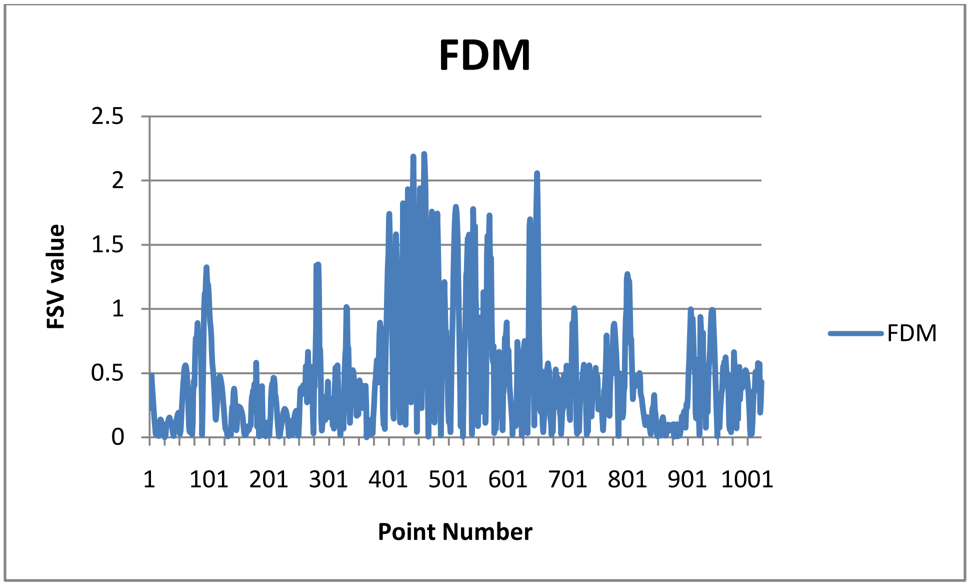

Figure 4 Feature Difference Measure of the data in Figure 2

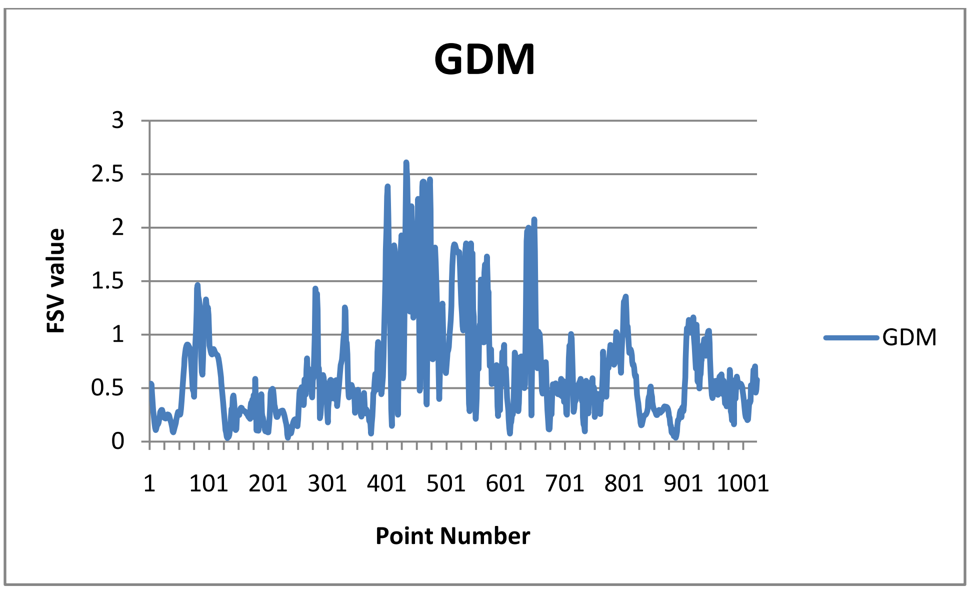

Figure 5 Global Difference Measure of the data in Figure 2

Taking the GDM as a summary, it shows that in the resonant portion of the antenna (approximately the middle of the above graphs) stirring is occurring very well but similarly, even below the nominal lowest working frequency, there is evidence of adequate stirring.

The summary values for FSV are given in Table II.

Table II FSV summary values

ADM	0.393	Good
FDM	0.465	Fair
GDM	0.663	Fair

These show that, overall, the amplitude difference is not significant: while there are differences due to local amplitude variations, the shape is reasonably constant. It also shows that, overall, there is a reasonable range of stirring across the graph.

The confidence histograms are shown in Figures 6 – 8. These represent a measure of the spread of opinion that a group of engineers, assessing the original data, would decide [7].

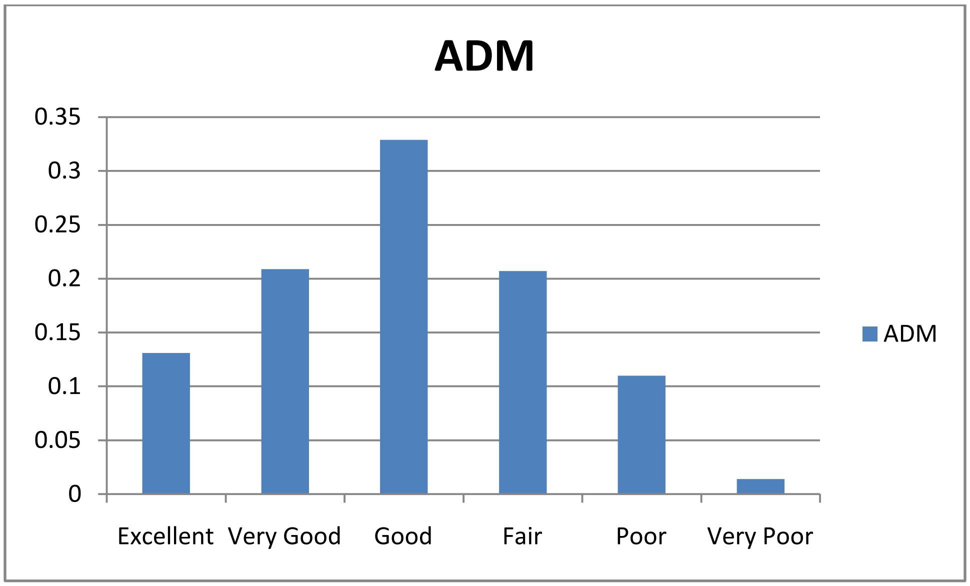

Figure 6 Amplitude Difference Measure confidence histogram

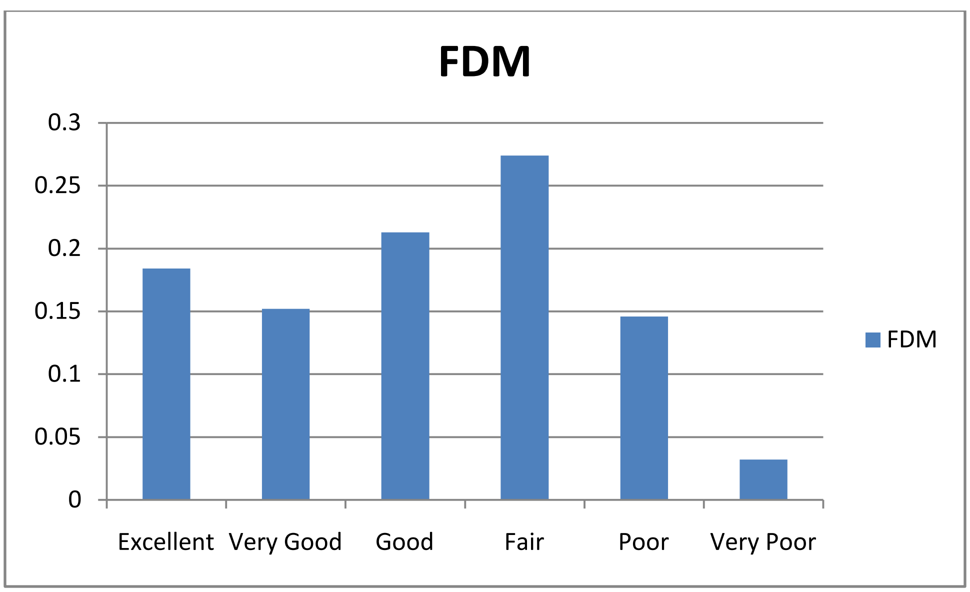

Figure 7 Feature Difference Measure confidence histogram.

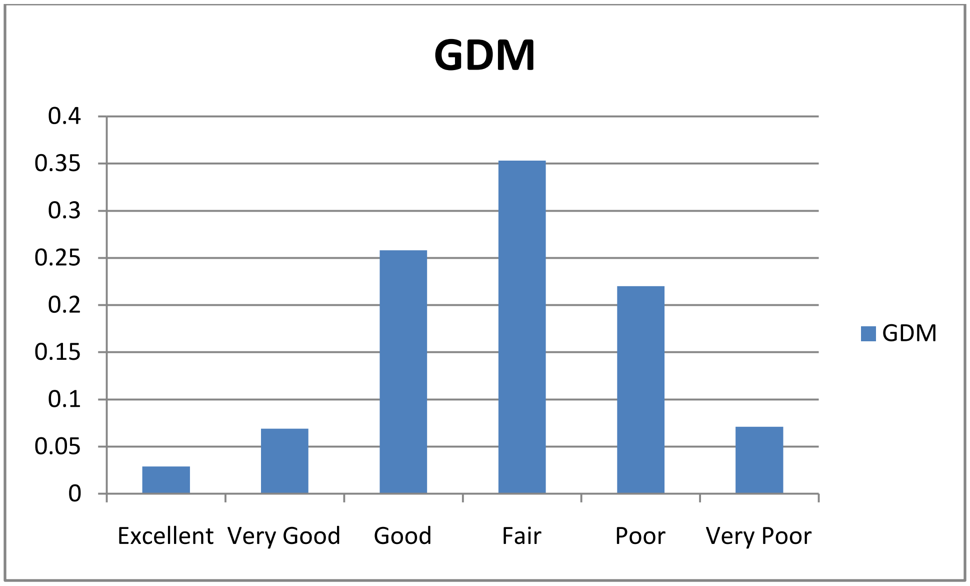

Figure 8 Global Difference Measure confidence histogram.

These confirm the opinions voiced above, but it is interesting to note the shape of the FDM confidence histogram separately. This is much broader than the other two with a much less well defined 'bell' shape.

Finally, the Grade and Spread are given in Table III.

Table III Grade and Spread for the data of Figure 2

	ADM	FDM	GDM
Grade	4	5	5
Spread	4	5	3

From this table, we can infer that there is a lot of difference between the two data sets (Grade) and this is similar for all the measures. However, the Spread, for the GDM, is only three categories, indicating a relatively high level of confidence that can be placed on the assessment that, over the whole range of the graph, the agreement between the two sets of results is Fair and, as was indicated earlier, anything Fair or worse would be acceptable.

5. Conclusions

The Feature Selective Validation (FSV) method has been presented and has been applied to some measured data arising from reverberation chamber measurements. The detail of the FSV method was discussed and the various measures that can arise from the full method applied to this data. In many cases, only a sub-set of this information would be presented or used. However, the whole range can be used to build up a comprehensive picture of how the original data does compare.

FSV can be used directly in decision making about whether a comparison is good enough or not. However, its principal use is as a knowledge management tool: providing a framework for discussion about "what's good and what's not". It has certainly been the authors' experiences that removing the data from its original domain and original form and constructing FSV results can help in discussion about the reasons for various parameters of the comparison. For example, the authors are considering whether the lack of apparent mode movement between about 130 MHz and 150 MHz is an artefact of the small stirrer sampling used (i.e. two) or an actual feature of the chamber. It would be helpful, in this case to look at other stirrer steps and repeat the comparison and, on the basis of that further analysis, accept results in this range or agree that antenna efficiency measurements in this range are not likely to be accurate and, therefore, not recorded.

References

[1] IEEE Standard 1597.1 Standard for Validation of Computational Electromagnetics Computer Modeling and Simulations, 2009.

[2] DE Coleby, AP Duffy "A Visual Interpretation Rating Scale for Validation of Numerical Models", COMPEL: Int J for Computation and Maths in Electrical and Electronic Eng.,, Vol 24, No 4 (April 2005), pp 1078-92.

[3] AP Duffy, AJM Martin, A Orlandi, G Antonini, TM Benson, MS Woolfson, "Feature Selective Validation (FSV) for validation of computational electromagnetics (CEM). Part I – The FSV method", IEEE Trans-EMC, Vol 48 No 3, Aug 2006, pp 449 – 459

[4] A Duffy, A Orlandi and H Sasse, "Offset Difference Measure (ODM) Enhancement for the Feature Selective Validation (FSV) Method", IEEE Transactions on EMC, Vol 50, No 2, May 2008, pp. 413 - 415

[5] A Orlandi, G Antonini, C Ritota, AP Duffy "Enhancing Feature Selective Validation (FSV) interpretation of EMC/SI results with Grade-Spread" IEEE Int Symp on EMC, Portland Oregon, 2006

[6] CS Lee, AP Duffy, C Lee, "Antenna Efficiency Measurements in a Reverberation Chamber without the need for a Reference Antenna", Antenna and Wireless Propagation Letters, in press.

[7] A Orlandi, AP Duffy, B Archambeault, G Antonini, DE Coleby, S Connor "Feature Selective Validation (FSV) for validation of computational electromagnetics (CEM). Part II – Assessment of FSV performance" IEEE Trans-EMC Vol 48 No 3, Aug 2006 pp 460 – 7

[8] A Coates, H Sasse, D Coleby, A Duffy, A Orlandi, "Validation of a three dimensional Transmission Line Matrix (TLM) model implementation of a mode stirred reverberation chamber" IEEE Transactions on EMC, 2007