The Use of Statistical Image Classification Techniques for the Assessment of Measured Antenna Pattern Functions

J. McCormick*, S.F. Gregson+, C.G. Parini#

*Selex Galileo
Crewe Toll, Edinburgh, EH5 2XS, UK
john.mccormick2@.selexgalileo.com

+Nearfield Systems Inc.
19730 Magellan Drive, Torrance, CA 90502, USA
sgregson@nearfield.com

#Queen Mary, University of London
School of Electronic Engineering & Computer Science
Mile End Road, London, E1 4NS, UK
c.g.parini@eecs.qmul.ac.uk

Abstract— **Attempts to produce robust, objective, quantitative measures of similarity between antenna pattern data sets using statistical methods have been widely reported in the open literature [1, 2, 3, 4, 5]. This paper presents an introduction to the physical validity and use of generalised statistical analysis along with specialist statistical image classification concepts as applied to the assessment of antenna pattern functions. Before presenting results the paper describes some of the more recent developments in the field relating to, nominal, ordinal, interval and the more usual ratio type levels of measurement data available within near-field measurement data sets. Finally the paper describes how these assessment techniques could be implemented in a formal gage repeatability and reproducibility (R&R) [6] numerical quality analysis of a measurement system.**

I. Introduction

When the observations of physical phenomena are designed to extract quantitative information about the physical phenomena being observed these observations are termed measurements. A range of theoretical interpretations about the nature of the measurement process have been advanced, mainly developed from the original concepts stated by Helmholtz [7] In the last century these theoretical interpretations of the measurement process have inclined more and more heavily towards the representational theory of measurement. Where, measurement is defined as, "the correlation of numbers with entities that are not numbers", [8]. That is as shown below in Figure 1. Here, the elements of some physical state set, (Q), *e.g.* mass length time, are mapped onto a representational set, (N), usually the set of real numbers by the measurement process M.

Classically a physical quantity, the extent or extensive nature of a physical phenomenon, can take any real value and this was why the mapping of the physical phenomena onto a representational set was again usually described as an isomorphism from the physical state set onto the real number set.

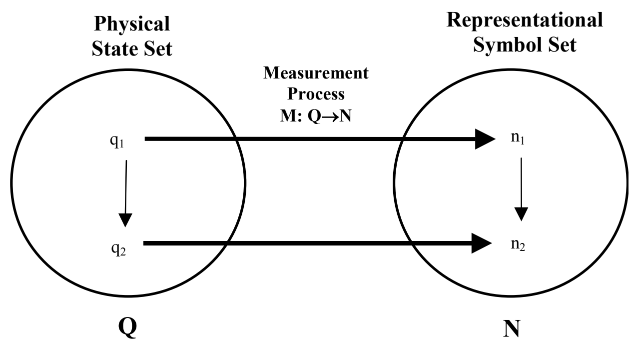

Fig. 1 Pictorial Representation of the Set Theoretic Model of Measurement.

Only in the early twentieth century when the quantised nature of physical phenomena became clear did the concept that physical quantities could not take on any and all real values become valid, [9]. However, even prior to this although it was considered that physical quantities could take on any value any practical measurement system could only map these physical states onto a limited range of numbers, these numbers representing rational multiples of the resolution of the measurement scale. Thus, in practice the information that could be extracted from a measurement process would be restricted to a range of numerical values where a many to one homomorphism mapping would replace the theoretical one to one isomorphism of an imagined perfect measurement process. The validity of the mapping process being confirmed by the equivalence of the real value physical state sets and the limited value representational sets under the same set of automorphisms.

Additionally in practice any measurement system existing in the physical word would be subject to a range of possible systematic and random errors in the measurement process and the mapping from the physical state set to the representational set could not be guaranteed to be reliable and repeatable and ambiguity would be introduced into the measurement result.

If the measurement process is thought of in functional terms where the set of physical states is defined as the domain and the representational set as the image of the domain under the mapping, in fact then in cases were the ambiguity is introduced the result will be that the element in the domain is mapped to a possible set of values in the image set. Formally the mapping transforms the element q into a random variable and N_q is the set of values of this variable. The random variable n, which is an image of the element q in the domain N_q, will have a distribution, $p_q(n)$, where the assignment of a number of elements in N_q is possible but one state in N_q will be preferred.

This illustrates that, for all theoretical and practical measurement systems, measures are in fact probabilities, [10] calculated from mappings of random variables onto a distribution of possible values. Additionally the interpretation of antenna patterns themselves as probability density function for electromagnetic interactions to occur at certain angles relative to an antenna placed at the source of an inertial reference frame, [1] further strengthens the concept that the interpretation and assessment of antenna patterns can best be addressed in the realm of inferential statistics.

However, the choice of statistical approach should be dictated by the nature of the measured data. The mapping between the physical state set and the representational set is only possible if the sets follow the same logical form, *i.e.* relation between the elements in the set Q is the same as the relation in the set N. For extensive measurement of quantities this implies a binary relation between elements based on > or = with gives an order to the sets. For antenna measurements this relation is usually measured relative to an arbitrary reference signal so gains are mapped onto an interval scale and calibrated relative to a known standard. This allows the set of statistical methods relevant to the assessment of data at the interval or metric level. However, the data can be viewed at the ordinal level where no arbitrary interval of > is imposed on the data so ordinal statistical methods are also relevant for assessment. Finally if the measured data of the pattern had specific features that could be extracted objectively it could be treated in terms of categorical statistical methods where the representational set would comprise a number of subsets, defined by intension, that antenna patterns could be categorised to fall within. Thus interval ordinal and categorical statistical methods could all be used to assess antenna measurement data and as is illustrated in the following sections.

II. Assessment of Error Simulations

Many attempts to produce objective quantitative measures of correspondence between data sets, *e.g.* antenna pattern data, that can be used to assess the accuracy, sensitivity and repeatability associated with the production of that data have been have been reported in the open literature [1, 2, 3, 4, 5]. In these, a variety of statistical methods have proved successful in the robust assessment of similarity between respective antenna pattern functions where the comparison has been found to be complicated by three principal factors:

- The large amount of interferometric, *i.e.* complex, data which is used to represent the antenna performance,
- The large dynamic range of that data, typically greater than 70 dB,
- The antireductionist, *i.e.* integral, relationship that exists between the near- and far-fields with each point in one domain affecting every point in the other.

To illustrate the applicability of these statistical data assessment techniques to antenna measurements a partial scan technique [5], which attempts to reduce truncation errors inherent within planar near-field antenna measurements, was simulated. A detailed description of these partial scan techniques is left to the open literature, and is not the specific purpose of this paper. This measurement technique occupies a research area that produces data sets that generally require detailed analysis to assess its applicability and utility as a measurement process. Moving the antenna under test (AUT) between successive partial scans will necessarily involve the disturbance of the reference path of the RF subsystem and introduce imperfections in the alignment between partial scans. It is the subtle impact of these alignment imperfections that will be used here to illustrate the effectiveness of various comparison techniques.

To illustrate the assessment processes a number of measurements were simulated that each included specific alignment errors. A simple physical optics (PO) measurement simulation tool was utilised to produce a series of synthetic measurements containing two kinds of alignment errors that are known to be particularly important to the auxiliary rotation partial scan technique. These were angular and range length, *i.e.* AUT-to-probe distance, errors. In the absence of some overriding definitive standard or infallible model, the only practical methodology for assessing the ability of any test facility to make measurements is by way of repetition of these measurements. This repetition can be accomplished without alteration in the measurement configuration, to simply address repeatability, or with the inclusion of parametric variations to assess sensitivity. As repeatability is inherently a statistical process the validity of any conclusions drawn will greatly depend upon the size of the sample. Thus it is preferable in this case to utilise as large a number of simulations as is practical.

To this end, the assessment of each of these errors entailed the simulation of the ninety-nine tri-scan measurements, *i.e.* two hundred and ninety seven individual partial planes. These measurement sets were transformed to the far-field using the existing transformation computer code assuming that the data sets contained no imperfections in their alignment. The equivalent multipath level (EMPL) was calculated between the ideal pattern and each of the error simulations [5]. This can be thought of as the amplitude necessary to force the different pattern values to be equal. The maximum EMPL, *i.e.* the worst case, value at each angle can be found plotted below in Figure 2, together with the ideal cardinal cut. Similarly, Figure 3 below contains a plot of the optimum cardinal cut together with the maximum EMPL for case of the range length error simulations.

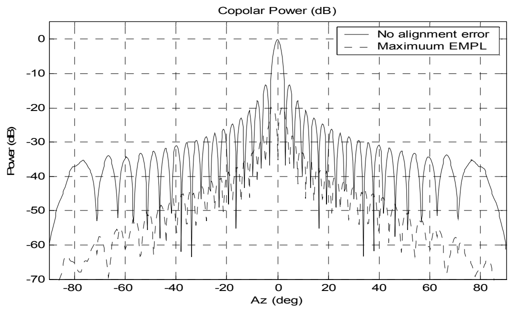

Fig. 2 Far-field azimuth cut of ideal simulation and maximum EMPL (pointing error).

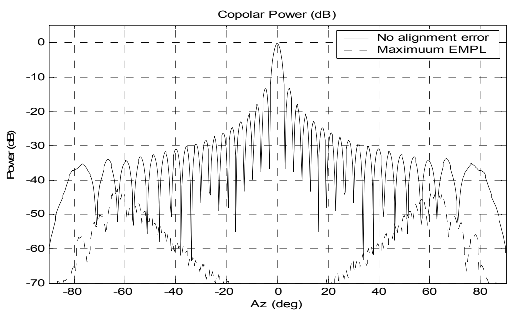

Fig. 3 Far-field azimuth cut of ideal simulation and maximum EMP (range length error).

A comparison between the ideal "error-free" far-field pattern was made with each of the 99 far-field angular error simulations using each of six test correlation coefficients. This comparison process was then repeated for each of the 99 range-length error simulations. The six statistical tests used were:

1) Cross-correlation coefficient is a comparatively computationally expensive, general-purpose technique for obtaining a single quantitative correctly normalised measure of adjacency,
2) Ordinal measure of adjacency constructs a measure of association which is based on the similarity of the permutations of the data when ranked in terms of the amplitude of the data sets.
3) Kendall coefficient of concordance is based upon a measurement of the deviation of the ranked data sets from the mean rank of the data.
4) Interval-ordinal method is a modification to the ordinal measure so that greatest importance is attributed to regions of greatest interest.
5) Categorical-interval method represents a modification to the interval measure such that the data sets are categorised and then the relative frequencies associated with the categorisation are the subject of the cross-correlation.
6) Categorical-ordinal method represents a modification to the ordinal measure such that the data sets are categorised and then the relative frequencies associated with the categorisation are the subject of the ordinal measure of correspondence.

A detailed description of these techniques can be found in [5] with a brief overview of the Kendall coefficient presented in Section III. Crucially, each of these techniques yields a single, normalised, symmetrical (*i.e.* commutative) coefficient. The mean values of these comparison coefficients can be found presented in Table 1 below where the comparisons have been taken over all 99 simulations. Here, all of the correlation coefficients are normalised so that $k = 1$ represents a perfect correlation, and $k = 0$ represents no correlation.

TABLE I
MEAN VALUES OF CORRELATION COEFFICIENTS

Metric	**Mean Value**	
	Angular Error	Range Length Error
Cross correlation	0.9931	0.9982
Ordinal	0.8132	0.8758
Ordinal (Kendall)	0.9612	0.9911
Interval-Ordinal	0.6395	0.8113
Categorical-Interval	0.9991	0.9999
Categorical-Ordinal	0.4792	0.7709

TABLE II
99% CONFIDENCE INTERVALS FOR CORRELATION COEFFICIENTS

Metric	**3 Standard Deviation**	
	Angular Error	Range Length Error
Cross correlation	0.0150	0.0039
Ordinal	0.2638	0.1638
Ordinal (Kendall)	0.0960	0.0218
Categorical-Interval	0.0009	0.0003
Categorical-Ordinal	0.2689	0.1040

Here, the categorisation process involved 101 bins equally spaced spanning the levels –70 dB to 0 dB. Table 2 above contains the 99% confidence interval, *i.e.* the 3 standard deviation of the k values about their respective mean values. Although each of these comparison procedures reveals small but systematic errors introduced into the simulations, the extent with which these variations are reported differ markedly and are discussed in Section IV below.

III. OVERVIEW OF KENDALL COEFFICIENT OF CONCORDANCE

In the above assessment the Kendall coefficient of concordance was used to assess the adjacency of respective antennas patterns. A brief description of this comparison technique follows. Let object i be given the rank $r_{i,j}$ in trial j, where there are in total n objects and m trials. The total rank given to object i is,

$$R_i = \sum_{j=1}^{m} r_{i,j} \tag{1}$$

The mean value of these total ranks is,

$$\overline{R} = \frac{1}{2} m(n+1) \tag{2}$$

The sum of squared deviations of these ranks, S, is defined as,

$$S = \sum_{i=1}^{n} \left(R_i - \overline{R}\right)^2 \tag{3}$$

Whereupon the Kendall W coefficient can be defined as,

$$W = \frac{12S}{m^2\left(n^3 - n\right)} \tag{4}$$

With this definition, if the test statistic W is 1, then all of the trials have been exactly the same, and each trial has assigned the same order to the list of objects. If W is 0, then there is no overall trend of agreement among the trials, and the responses may be regarded as being essentially random. Intermediate values of W indicate a greater or lesser degree of agreement. Thus, W is normalised and occupies the range, $0 \leq W \leq 1$. The principal motivation for using this statistic rather than the previously used ordinal measure of adjacency is that the evaluation of the Kendall coefficient requires significantly less computational effort thus providing a clear incentive for its adoption.

IV GAUGE R & R STUDIES

Additionally the Kendall coefficient is a particularly useful tool if a formal quality Gauge repeatability and reproducibility (R & R) study is to be undertaken on a measurement process that involves the acquisition of antenna patterns. Such a study is a standard quality engineering tool for the quantitative assessment of production test measurement systems. Usually a gauge R & R study attempts to quantify the relative importance of three distinct components of the variation in the recorded data provided by any measurement process, these being:

- Repeatability variation
- Reproducibility variation
- Part to part variation

Repeatability involves the simple repetition of a measurement to assess the extent of any random or systematic variations that occur over the period of the measurement process. Reproducibility often involves the attempt to reproduce the measurement results with a variety of different measurement systems and or procedures. For antenna test ranges it is most usually explained in terms of the definition quoted in [5] where the ability to reproduce the results of a measurement after a period of time that greatly exceeds the period of the measurement with the assumption that in the interim period the range will have been reconfigured and or reused is assessed. Thus, repeatability is mainly concerned with the short term stability of the measurement system and reproducibility is primarily concerned with the ability to reproduce the conditions and procedures in which an original measurement was conducted. Finally the actual variation in the performance of antennas that are in theory produced to have the same characteristics is the part to part variation, the aspect that is of primary importance in the test procedures.

For any production testing it is this part to part variation that is of importance and the other sources of variation only serve to obscure this, and thus must be as small as possible. The detailed R & R gauge study depends on applying the statistical assessment procedure of analysis of variance, (ANOV), [11] to a range of results illustrating the impact of repeatability and reproducibility that can be acquired in a range error budget. ANOV is a linear statistical method that relies on the existence of a single linear measurement variable existing for each measurement trial, e.g length, voltage, time.

Antenna measurements produce thousands of data points in an ordered set so the ability of techniques such as Kendall concordance to produce a single coefficient that can summarise the variation between patterns is vital. Additionally, while the actual value of the coefficient is not linearly related to any specific attribute of the ranking. As the process of acquiring the W statistic is based on variance that the variation in the concordance is not just a function of how many ranking mismatches there are it is also a function of square of the magnitude of the mismatch. Additionally the metric or the overall size of the variation in the concordance is a cubic function of the number of data points in the measurement data set. The existence of these exact functional relationships means that Kendall concordance can produce the kind of single linear variable upon which rigorous ANOV based Gauge R & R studies can be based.

IV. Discussion of Results

The cross-correlation coefficient is essentially "saturated" by the large dynamic range of the data as the coefficient diverged from unity in only the third decimal place. Although this coefficient reported that, on average, the range length error is less critical than the angular error, the discrimination observed is small. In general, such purely interval techniques tend to be highly sensitive with more pronounced differences between patterns occasionally yielding numerical instabilities.

The results of the ordinal measure clearly shows that the small but systematic errors introduced into the simulations can be accurately quantified. However, the ordinal process of ranking the data to produce permutations takes no account of either the absolute amplitude or spatial angles at which the data is found. Thus, every region of the pattern is judged to be equally important in the calculation. This is clearly illustrated by comparison of the mean average values of k determined from the two different error simulations as their values are very similar. This is despite the fact that the range length error principally produces differences only in the wide out, low-level, side-lobe region.

The Kendall concordance coefficient is simple, easily computable, robust, and can form the basis for defining a linear variable that can be used for rigorous R & R production testing studies.

The interval-ordinal technique aims to address the principal deficiency inherent within the ordinal technique (that of ignoring the relative importance, *i.e.* level, of different elements within the data sets) whilst minimising the numerical instabilities that can be encountered when using purely interval techniques. This technique clearly shows that the angular error is of greater importance as it affects large, as well as small field intensities. Thus, this hybrid approach is better able to isolate errors in the data that display amplitude specific traits and thus validates the concept. The extent with which the hybrid interval-ordinal method discriminates between differences in elements corresponding to signal magnitudes can be readily varied on a case by case basis to emphasise or de-emphasise the particular feature under investigation. Unfortunately, it could perhaps have been expected that as less importance is being placed upon differences within small signals that the *k* value for the range length error would remain constant, or perhaps even have increased further towards unity. As the converse is observed it could be concluded that the piecewise polynomial interpolation scheme that is utilised within the re-tabulating process is introducing an additional source of error. If this is true, it is the subject of further work.

The categorical-interval/ordinal schemes remove the requirement for pattern re-tabulation by categorising the elements into a vector of predefined bins. Unfortunately, the categorical-interval method that relies upon the conventional cross-correlation coefficient to determine the similarity between the histograms has yielded correlation coefficients that are so close to unity that they are essentially useless. Despite this, the technique still managed to distinguish between the two forms of error source although this discrimination is observed in the *fourth* decimal place. This is most probably a result of the fact that the histogram contains significantly fewer elements than the data set from which it was abstracted thus the comparison algorithm has fewer elements with which to work.

In contrast, the hybrid categorical-ordinal technique yields results that clearly discriminate differences between small and large signals, does not introduce errors from interpolation, and yields a sensitive correlation coefficient. In general an infinite number of distributions can be chosen. It has been found that by adjusting the distribution of the bins, different emphasis can be placed upon different properties within the data providing additional flexibility in the examination of the way in which the patterns differ. A summary of the broad characteristics of these comparison techniques can be found presented in Table 3 below.

V. Conclusions

It is clear that the new and novel antenna measurement techniques being pioneered at present in various laboratories around the World offer an assessment challenge if the large volumes of data these techniques generate are to be quantitatively, effectively and concisely analysed and summarised.

Almost all data assessment techniques at route depend on reducing the dimensionally of the data sets to make them more easily accessible. Antenna patterns acquired in test ranges may contain tens or hundreds of thousands of individual data points and the quantitative assessment of such large data sets is close to impossible without distilling the data down to manageable levels. Clearly, such data reduction techniques involve the loss of some information from the data sets. However, it should be borne in mind that all inferential statistical methods, be they interval, ordinal or categorical abstract the data to assess specific attributes and/or features in the data sets so in all statistical data assessment information is lost about the specific nature of a the sets being assessed. Thus, the choice of assessment technique or techniques must be guided by an informed understanding of the nature of the antenna parameters that are to be assessed. The research reported in this paper is on-going with further research into antenna pattern comparison intended.

References

[1] J. McCormick, B. M. G. Ghinelli, The Use of Statistical Image Classification Techniques in Antenna Measurements, IEE National Conference on Antennas and Propagation, March 1999, University of York.

[2] S.F. Gregson, J. McCormick, "Image Classification as Applied to the Holographic Analysis of Mis-aligned Antennas", ESA ESTEC Workshop an Antenna Measurements, 1999.

[3] S.F. Gregson, J. McCormick, C.G. Parini, "Advances in the Objective Measure of Comparison Between Antenna Pattern Functions", International Conference on Antennas and Propagation, University of Exeter, 2003

[4] J. McCormick, S.F. Gregson, C.G. Parini, "Quantitative Measure of Comparison between Antenna Pattern Data Sets", IEE Proc.-Microw. Antennas Propag, Vol. 152, No 6, December 2005

[5] S.F. Gregson, J. McCormick, C.G. Parini, "Principles of Planar Near-Field Antenna Measurements", The Institution of Engineering and Technology, UK, 2007.

[6] R.K. Burdick, C.M. Borror, D.C. Montgomery, "Design and analysis of gauge R&R studies: making decisions with confidence intervals in random and mixed ANOVA models", ASA-SIAM Series on Statistics and Applied Probability, 1987.S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[7] J. A., Diez, "A Hundred Years of Numbers. A Historical Introduction to Measurement Theory 1887-1990. Part I: The formation period. Two lines of research: Axiomatics and Real Morphisms, Scales and Invariance", Stud. Hist, Phil. Sci., Vol.28, No.1, pp167-185, Elsevier Science Ltd, 1997.

[8] B. A. W. Russell, "On the Relation of Number and Quantity", Mind, Vol 6, pp 326-241, 1897.

[9] M. S. Longair, "Theoretical Concepts in Physics", Cambridge University Press, Cambridge, 1984, ISBN 0 521 25550 3, pp218-225.

[10] Andrews P. P., Grewal M. S., "Kalman Filters Theory and practice Using Matlab", 2nd Edition, John Wiley and Sons, New York 2001, pp 102-103.

[11] H.T. Hayslett, "Statistics", Heinmann Professional Publishing, 1981, ISBN 0 434 98474 4.

TABLE III
Qualitative Comparison of Pattern Comparison Techniques

Metric	*Interval*	*Ordinal*	*Single Coefficient*	*Domain*	*Holistic*	*Robust*	*Sensitivity to outlying points*	*Absolute ref*
Cross-correlation	Yes	No	Yes	$-1 \le k \le 1$	Yes	No	Yes	Yes
Ordinal	No	Yes	Yes	$-1 \le k \le 1$	Yes	Very Stable	No	No
Ordinal (Kendall)	No	Yes	Yes	$0 \le k \le 1$	Yes	Very Stable	No	No
Interval-Ordinal	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	Stable	No	Yes
Categorical- Interval	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	No	Yes	Yes
Categorical-Ordinal	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	Very stable	No	Yes