U.S. DEPARTMENT OF ENERGY

PNNL- 24071

Prepared for the U.S. Department of Energy
Under Contract DE-AC05-76RL01830

A Conceptual Framework for Measurement
(with emphasis on phasor measurement)

H Kirkham

February 2015

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401, fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847, fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.
(8/00)

PNNL-24071

A Conceptual Framework for Measurement
(with emphasis on phasor measurement)

H Kirkham

February 2015

Prepared for
U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Executive Summary

This Report explores the processes involved in metrology, the science of measurement. It concentrates on measurements that are made of continuous data streams. Measurement is seen as a process of data compression, performed sometimes to reduce the bandwidth needed in communication and sometimes to reduce storage space for the results of measurements. Importantly, it is shown that the result of a measurement should be viewed as having meaning.

The fact that a measurement result has meaning separates the topic of metrology from the world of communication explored by Nyquist and Shannon. Measurement results are semantic elements, and in metrology the language of semantics is mathematics. The goal of a measurement is seen as that of finding the values of the parameters of a mathematical model: the coefficients of a set of equations.

The device making the measurement is designed to estimate those values, so the equations must be known to the measuring instrument. However, two semantic problems can arise. The realized quantity, the physical thing that is used to represent the conceptual model that is the measurand, may be "colored" by what could be thought of as semantic noise, changing the meaning of the signal. That semantic coloration is arguably a Type B measurement problem in the terms defined by The Guide to the Expression of Uncertainty in Measurements (GUM)[1], whereas random "engineering" noise is Type A.

In addition to coloration, there can be semantic imprecision between the model in the measuring instrument and the measurand equations. To eliminate this imprecision, it is not sufficient to describe the measurand in words: it must be expressed mathematically.

The phasor measurement unit (PMU) is examined in the light of the framework that these conclusion create. It is shown that the calibration of PMUs is done by a method that suffers semantic imprecision, and that can account for some of the problems with creating PMUs that meet the requirements of the standard. In addition, but as a separate aspect of the measurement, semantic coloration is added by (for example) harmonics on the power system.

A new method of implementing the PMU, developed on this DOE project and reported a year ago, is seen to fit within this framework, and because its equations match those used in calibration, it is theoretically capable of achieving good calibration results. The semantic imprecision built into the existing PMU standard means that the PMUs so far made are not likely to be so capable of good results.

Rather than change the calibration method used, the Standard should be re-written. The equations used to generate the calibration signal are far more representative of the power system than the equations it forces the PMU to use. The standard should match them.

[1] GUM: BIPM *Guide to the Expression of Uncertainty in Measurement.* International Organization for Standardization, Geneva. ISBN 92-67-10188-9, First Edition 1993, corrected and reprinted 1995.

Acknowledgements

This Report integrates ideas that I have had over many years, beginning at about the start of the 21st century. As I became more deeply involved in the work of the IEEE Standards working groups for phasor measurements, around 2010, it became clear that those more general ideas were applicable to the synchrophasor problem. It is fair to say that some of the problems faced by that community arose because we did not have the benefit of the framework that is developed in this report. This present work is generally applicable, but focused on phasor measurements. It has benefitted from not only the funding of Phil Overholt of US DOE, but also his constant understanding and encouragement. What follows owes much to Phil's involvement.

Over the last few years, I have had discussions with my metrologist colleagues on many topics relevant to this work. Alex McEachern of Power Standards Lab helped me understand the correctness of the various reactive power definitions. Definitions are important: a definition of "frequency" that would work when frequency is changing is something that many colleagues in the PMU community have asked for. All have been patient as we struggled to understand what that definition might be, and there have been many stimulating and helpful discussions. In particular, I would like to express my appreciation of the support and encouragement of the IEEE Power System Relaying Committee C37 Working Group on the PMU standard. I thank Ken Martin of TPG, who leads that group, for his efforts to keep the standard as open and permissive as possible, Veselin Skendzic of SEL for his careful and patient explanations of the technology, Jay Murphy of Macrodyne for many discussions on the matter, and Jerry Stenbakken of NIST for placing the thought in my mind that digital technology has created a revolution in measurements.

Vive la revolution!

Harold Kirkham, February 2015

Table of Contents

List of Figures

A Conceptual Framework for Measurement

1.0 Context

Metrology exists at the point in some multidimensional space where science, engineering and epistemology (the science of knowledge) come together. The national metrology institute (NMI) deals with the science and the engineering. Most instrumentation users do not get deeply into either, but if there is a bias it is probably toward the engineering. Most of what little teaching exists is biased toward the engineering aspects of the topic. The goal of this paper is to clarify the subject by examining the central area. Its method is to examine the engineering in terms of the science and the epistemology. A framework for understanding measurements is developed.

1.1 Introduction to Measurements

We begin by reviewing some of the basics of measurement. We concentrate on measurements in which there is a stream of data being digitally sampled and used to find a result. This sort of measurement breaks the incoming signal into sections called sampling windows, and calculates results for each window.

Typically, a good measurement involves taking many readings and averaging them. We tell ourselves we are understanding the effect of the measurement uncertainty. If we take just a few readings with our instrument, we do not get a very satisfactory result, especially if the readings seem to differ. Therefore, we take a fairly large number of readings, and use that average.[1] We could plot the readings we obtained as a distribution, by putting the results into bins of suitable size. We might assume that the result showed that the readings had a Gaussian distribution, and that therefore the mean was a good estimator of the location of the distribution. For example, if we took a lot of readings of the voltage on an electric outlet, we might obtain a result like Figure 1. (The word *frequency* to label the ordinate means that the columns show the number of times the result of the measurement fell into each bin.)

The readings are distributed around some value in the middle. They have a *variance*, and if we had enough data we could estimate that variance. The mean is a measure of the *location* of the distribution of our results, the variance is a way to express the *dispersion* of those results.

Noting that two standard deviations covers 95% of the results, we might choose to say that the result of the measurement was that the value measured was x units, with an uncertainty of Δx (the value at 2σ), with a confidence level of 95%. Or we could say that we are 95% sure that the value of the quantity being measured is between $(x - \Delta x)$ and $(x + \Delta x)$, and that its most likely value is x. These expressions are the sort of way that a metrologist presents the result of the measurement. The question of uncertainty is addressed in considerable depth in the *Guide to the Expression of Uncertainty in Measurement* (GUM,

[1] Increasing the number reduces what might be thought of as the uncertainty with which the *location* of the distribution is known. (Usually it is called the standard deviation of the mean.) Just how many readings are necessary is itself a question worthy of study, and the topic is mentioned in several papers. It seems that for many purposes about ten will often be adequate. See Stephanie Bell, "A Beginner's Guide to Uncertainty of Measurement," NPL Measurement Good Practice Guide No. 11 (Issue 2), 1999. Available at https://www.wmo.int/pages/prog/gcos/documents/gruanmanuals/UK_NPL/mgpg11.pdf

1995) but that kind of statement is what it comes down to. Until GUM, the fundamentally statistical nature of measurement had not been recognized by all measurement people. It is now central to metrology.

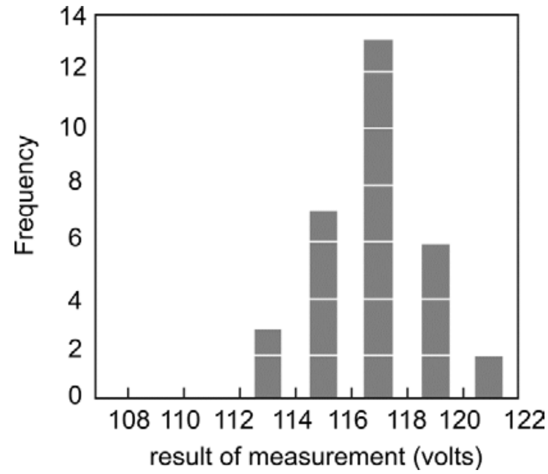

Figure 1 Histogram of voltage measurements

If we were to change the width of the sampling window in the measurement system, we would find the variance of the readings would change in a particular way, as shown in Figure 2. The figure may be familiar as the Allen variance, something much used in horology.

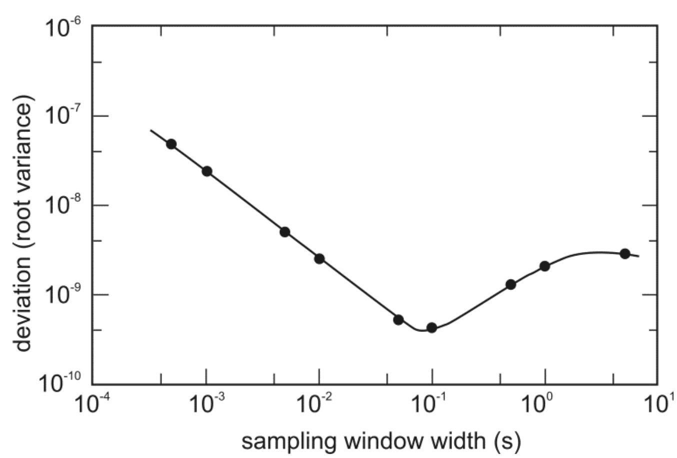

Figure 2 Deviation as a function of window width

If the measurement duration is made too short, the effect of noise on the signal predominates: it is clear that we cannot make a measurement in an arbitrarily short time. If the measurement takes too long, other effects (such as vibrations or temperature) start to become evident in the result. It seems that there is a limit to what can be done this way to reduce the effect of Gaussian noise.

For the made-up example here, the optimum seems to be about a tenth of a second. Realistically, the shape of the curve and the values will depend on the nature of the noise affecting the system. In fact, the graph is something that characterizes the noise in the system, and suggests an optimum window size for the measurement system. But that optimum window size is really important only when the user is trying to measure a small signal down near the level of the noise. If the signal to noise ratio is high, the noise can be expected to have little effect on the accuracy of the measurement.

The window width is related to the detector time-constant in an analog measurement system. In an analog system, the signal is typically not windowed, it is always connected to the instrument. The reading on the display is filtered by the time-constant of the detector and the response-speed of the indicating instrument. It has long been known that if the noise was random, the average value was zero, so designs for analog instruments quite often made the detector time constant as long as possible, consistent with not boring the user with a too-long response time.[2]

It seems, then, that the act of measurement takes a finite amount of time, and it involves estimating values of a distribution. It may be inferred from these features of the process that the signal being measured must not change during the measurement. It is assumed constant for the duration of a window, so the declared value will apply. If it is not constant for longer, the notion of finding a statistical distribution for repeated measurements does not work. The notion of changing the window duration also does not work, as the signal parameters will be different if the signal changes.

But even if the quantity is unchanging, the answer given by the statement of the result is incomplete. The three statistics that describe the output of the measurement device do not constitute a complete statement of what is represented. Those three parameters, the value, the uncertainty and the confidence level, are not always furnished as part of the result – yet even if they were, there is something missing. They do not give a complete picture.

1.2 Still Something Missing

What follows is an anecdote about an engineer designing a dc/dc converter with an output of 400 V. It is a true story.

The designer made a prototype, and demonstrated that the output was very exactly 400 volts. While the output voltage requirements on the design were not very stringent (a value within five or ten volts would have been acceptable), this particular designer was in a competitive mood. He built a prototype that was demonstrated to have 400.00 Vdc on the output terminals, measured with a high-quality dc voltmeter. The voltmeter’s display clearly showed a very stable four digits with an occasional swap between a zero and a one in the last place.

However, when viewed on an oscilloscope, the output was shown to have an 80 Hz oscillation with ripple of about 100 Vp-p on top of the 400 Vdc level. The rather significant discrepancy between the indications on the two instruments originated in the long integration time required for the high precision of the voltmeter. The oscillations were essentially averaged out of existence during data processing, by what

[2] Users of these analog measurement devices sometimes became adept at interpreting the movements of the needle. Vibration might indicate ripple on a power supply, for example. With the advent of digital measurements, things changed. The features that corresponded in the digital instrument were the sampling window, and the refresh rate on the display. But it is scarcely possible to discern the reading on a digital display that is changing rapidly.

was essentially a low-pass filter. The converter never provided the stable 400-V output indicated by the voltmeter.

The conventional statement of the result of the measurement (the voltage, the uncertainty, and the confidence level) does not disclose the fact that this was a measurement that very largely ignored the character of the signal being observed.

The full signal, oscillations and all, was available to the measuring instrument. The output reading – the declared value of the measurement – was indeed a representation of the rms value of the input signal, and there was nothing in the result that suggested a problem. There *could* have been some indication. While an analog meter would likely never have given such an "accurate" reading, the user would likely have noticed the needle vibrating.

It is not implied that a measurement using digital means cannot be as good as an analog one. It does seem fair to say, however, that the usual instruments are not always as informative. To see how to get around the limitation, let us look at what is involved in making a measurement.

2.0 Inside the Instrument

2.1 The measurement Process

George Carey Foster made some wise observations (anticipating Kelvin's remark that "to measure is to know") more than a century ago[3]:

> Before methods of measurement can be devised, it is evident that clear conceptions must be formed of the things to be measured. Such conceptions usually grow up by degrees in many minds from indistinct beginnings, until, in some one mind, they take definite shape and receive the precise expression which makes it possible for them to become the subject of mathematical reasoning.

Carey Foster perceived measurement as a process that starts with a concept in a mind. We argue here that the concept is triggered by an *application*. We also suggest that the modern way of work would reverse the sequence somewhat: what starts as a notion in one mind is nowadays given precise expression by a committee. Once that has happened, the design and manufacture of instruments can be aligned with agreed-to terms. In what follows, we will concentrate on the development of this instrumentality, and in particular on an instrument called a phasor measurement unit, or PMU.

The concept behind the measurement may be, as Carey Foster said, indistinct. Before a measuring instrument can be made, something more definitive is needed. This is the precise expression that Carey Foster had in mind, and that metrologists know as the *measurand*.

The term measurand is sometimes described as the "thing that is to be measured." That is a loose definition, however. The word is better thought of as a *description* of what is to be measured. Realistically, it has to be a careful *definition* of the quantity to be measured. This definition is derived from the concept.

However, a measurand is still a conceptual thing, and cannot itself be measured. It is a *definition*, of the thing to be measured, and so it is used to find something real to present to the measuring instrument. This is called the *realized quantity* [4]. The realized quantity is presented as a stimulus to the measuring instrument.

Inside the measuring instrument, the definition is used to design what is in essence a data-compression algorithm that gives the result of the measurement in some agreed-to conventional units. There are physical and conceptual aspects to the measurement process.

The measuring instrument samples the realized quantity and from the samples (the observations) and it performs the calculation that gives what metrologists call the *result* of the measurement. The method used in the processing of the realized quantity will depend on the description in the measurand, and the kind of signal that is anticipated.

---

[3] George Carey Foster, Inaugural Address of the President for 1881, Society Telegraph Engineers and Electricians, Vol. X – 1881, page 4-20. Accessed at http://books.google.com/books/about/Journal_of_the_Society_of_Telegraph_Engi.html?id=iScFAAAAQAAJ

[4] It has to be said, however, that most metrologists would accept the use of the word *measurand* to mean *realized quantity,* provided the meaning was clear from the context.

We saw earlier that the signal had to be unchanging. “Stationary” is a term normally applied to the statistical parameters of a stochastic process, and means that the parameters do not change between one measurement and the next. For most measurements, and for the purposes of this report, we do not deal with stochastic processes, yet we need to define a similar property of a more deterministic signal. We could say that the coefficients of the variables that describe the signal do not change. That does not necessarily mean that we can only measure signal that is unchanging in all its aspects. Consider the alternating current in a power line. We would (in words) say that the current was not changing if the coefficients of an equation that described it did not change

In the kind of measurement system we shall consider, the instrument compares the samples with some reference quantity in order to express the result in the proper units. We may as well imagine the samples to be converted to volts by some kind of transducer, and the reference quantity to be some kind of band-gap device whose own uncertainty is extremely low. We are not particularly concerned with those details here. The instrument then executes some kind of data compression algorithm to produce the result of the measurement as its *declared value*. This result is then presented to the user, or more generally, the application.

In terms of implementing the measurement, we note that the algorithm (the data compression method) must also be designed with reference to the measurand. The algorithm is designed to match the definition in the measurand. Thus, the measurand plays a central role in the instrumentality. It takes into account the needs of the application for the result, and it guides the selection of the realized quantity as well as the selection (or design) of the algorithm. The application also influences the choice of the algorithm, and even the definition in the measurand, creating a need for consistency among the various conceptual elements of the process.

Thus, there is a sequence from concept to measurand to realized quantity and to measuring instrument. From there the result of the measurement goes to the application. In the background are other connections at the conceptual level. The elements of the process are quite interconnected, with the interconnections forming a framework for viewing the overall process. The various interconnections are shown in Figure 3.

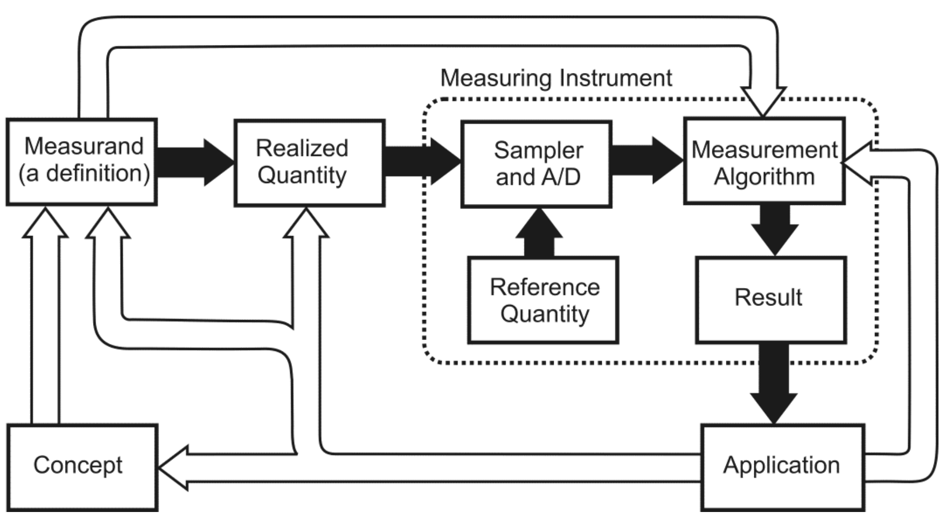

Figure 3. The interconnected elements of the process of measurement

The dark arrows in Figure 3 are the connections that a metrologist might think about when performing a measurement. Knowing the quantity he wants to measure, he selects a realized version of it and connects it to the measuring instrument. The result emerges, ready for whatever the application uses it for. The Figure shows the other connections, usually in the background, with open arrows. The application should probably be thought of as the starting place for the whole process.

The goal of the measurement algorithm shown in Figure 3 is to extract from the observed data (the samples of the realized quantity signal) some simpler representation of the signal being observed.

The representation is the measurand. Therefore, we must think of the measurand not as just a definition; it should be a *mathematical model*. The process of making a measurement is the process of finding (via the algorithm) the values of the parameters of that mathematical model. It is those parameters that must not change if we are to extract value from repeating the measurements.

The idea of finding the value of a parameter of an equation is very robust. The equation is an unambiguous way of expressing the measurand. We have all along had this in the back of our minds: the parameter value is the "true value" that we may never find exactly. At least, we will never *know* if we find it exactly!

In many measurements, the mathematical model can recreate a signal that contains all the information in the original signal. Consider the situation in which we are observing an alternating current signal. We might choose to find just the rms voltage, or we might model the signal with the usual sine or cosine function:

$$v(\mathrm{t}) = V\cos(\omega\mathrm{t} + \varphi) \qquad (1)$$

To make a complete measurement of the signal we would give the values for the parameters V, ω and φ, along with their uncertainties and the confidence level. If the input signal is indeed described by a function such as this, there is no information lost between signal and its description by the result of the measurement. However, if the signal is not so "clean," the measurement cannot accurately describe it if the model is that of Equation 1.

It is thus fair to ask whether the values we give as the declared value are the result of a "good" measurement. The values are estimates: how we can know if we are making a "good" estimate of something from the sampled data? It is certainly not enough simply to say that repeated measurements give the same result. Can we even know if it is possible to make a "best" estimate from a set of input data? The signal includes noise: how well does our estimator do at extracting the desired result from that noise? We need to look at the part of the measurement process that we have so far called the *algorithm*.

2.2 Relation to Shannon

Having said that the equation, with the appropriate parameter values, can recreate the information in the signal, it is instructive to recall the work of Shannon (1948). Shannon was concerned with the capacity of a channel to convey information in the presence of noise. To explain the nature of the problem he was tackling, Shannon used a diagram similar to the one in Figure 4.

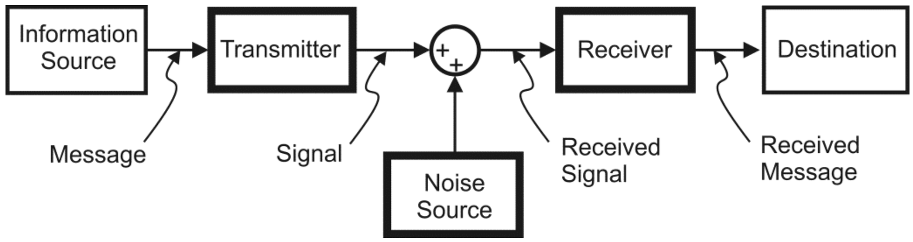

Figure 4 Noisy communications (after Shannon)

In his theoretical development, Shannon was concerned with the elements of the diagram shown here in heavier-weight boxes. The messages outside of these elements are dealt with by the logic of *semantics*, that is, the things to do with the study of *meaning*. Shannon dismissed the semantics as irrelevant to the engineering problem.

However Weaver, in his Introductory Note to the book edition of Shannon's papers (Shannon and Weaver, 1963), imagines the noise source in Figure 4 to add "engineering noise" to the signal. He mentions the possibility of an additional source of noise (which he calls "*semantic noise*" but which he does not explain[5]) adding its contribution to the message before it is encoded by the transmitter.

For our purposes as metrologists, this is a useful viewpoint, because we can adapt this straightforward system representation to the problem of measurement. *We are concerned with what Shannon set aside*. Shannon dismissed the semantics with these words: "Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem." (The italics are in the original.) That connection to "physical entities" becomes our connection to measurement.

Consider Figure 5. Following Weaver's suggestion, we have added a semantic noise source.

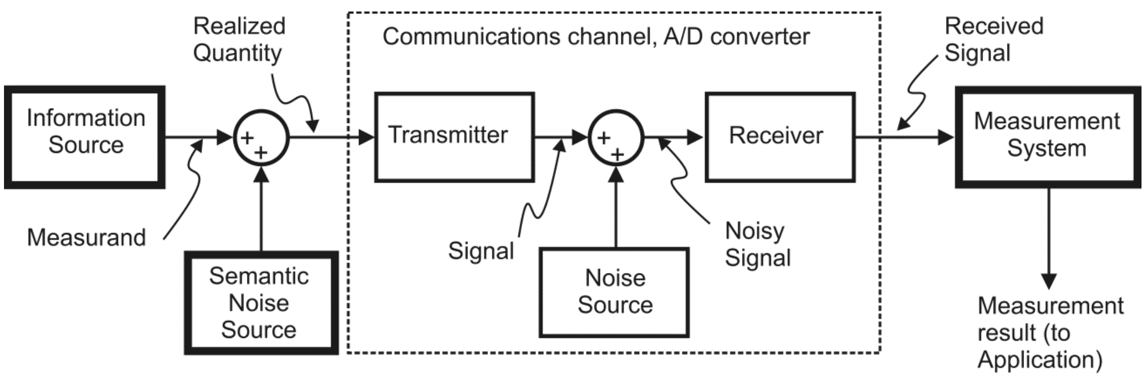

Figure 5 The addition of semantic noise to a measurement system

The semantics that were irrelevant to Shannon's engineering problem, are central to ours. We can accept that Shannon's work has solved the particular problems he had in mind, and we can dismiss the elements

[5] In an earlier work, Weaver (1949) had explored the term, but his interest was mainly human language.

inside the dotted box. We can concentrate on what is left. (That means we will not treat Weaver's "engineering noise" here. We should not forget that it is always present, however.)

In Figure 5, the "destination" of Shannon has been replaced by our measurement system, which passes the result of the measurement on to the application. The input to the measurement system is essentially the realized quantity, unless the communication channel adds some distortion of its own. The semantic noise is therefore something embedded in the realized quantity, which is something we should think of as a semantic entity, something with meaning.

There are three semantic entities shown as boxes in Figure 5: the measurand, the semantic noise source, and the measuring instrument. These three things are concerned with *meaning.*

We can collapse the representation of the communication channel in Figure 5, and expand the diagram as a representation of the process of measurement. See Figure 6 for the complete framework.

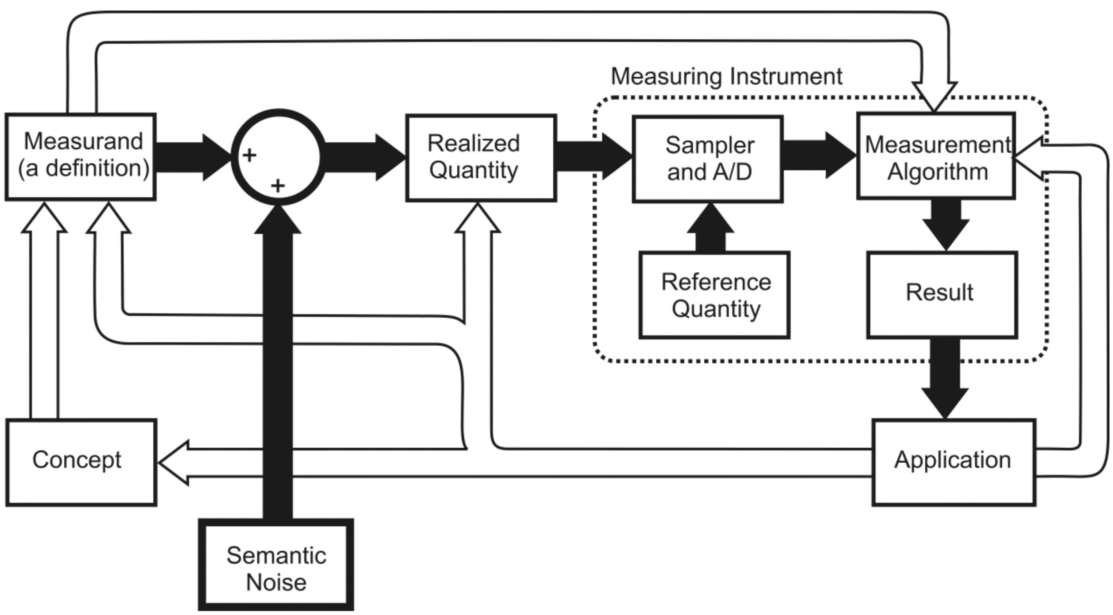

Figure 6 Measurement process including semantic noise

What we see in Figure 6 is a measurement algorithm that is challenged with a realized quantity contaminated with semantic noise. The measurand, the equations whose values are to be estimated, is polluted with this semantic noise.

The measurement algorithm is designed with the goal of finding the parameter values in the measurand, and it has to look "behind" the added semantic noise if it is to be successful. It is the only place in the measurement chain where anything can be done to reduce the effect of this semantic noise. If we do not use the measurement algorithm to correct for the semantic noise, we are not likely to obtain the values we expect for the parameters of the measurand.

The term "semantic noise" is probably not the best way to describe the problem, because in many ways it is not "noise-like" in the usual sense. It would be recognized as "signal" in many senses of the word. It may be a low-frequency thing. Or it may be very regular. It is therefore proposed that the word "coloration" be used to describe these semantic alterations to the signal. That coloration could be caused by vibrations (which might add to the output of a transducer), temperature (which might change the dimensions of parts or the value of electronic components) and, in electric power measurements, nonlinear loads (which add harmonics to signals).

Since these things are (by definition) not part of the equations of the measurand, the effect of these colorations changes the result of the measurement if the measurement process is not designed to take them into account. It is interesting to speculate whether the opposite is also true; that a cleverly chosen definition of the measurand would enable a measurement algorithm to be defined that would reduce the effect of semantic coloration. It may be so.

The matter of the exactness of the definition that is the measurand is of vital importance in measurement. (For example, in the PMU community, there has been a debate going on for a long time about the meaning of "frequency" when it is not constant.) Unable always to fix these things firmly, we have sometimes allowed what might be called *semantic imprecision* to creep into the measurement process. The effect is that the result of the measurement can be repeatable, but various implementations of the measurement equipment do not give the same values.

A report given by Andrew Berrisford[6] of a situation in Canada is an example of an effect of semantic imprecision. The measurand was the power factor, and an assumption made that "everyone knew" what that meant. The digital electric meter of a large consumer was replaced by a different digital meter. Before the replacement, the customer had a power factor of 0.95. After the replacement, the power factor was 0.88, as a consequence of which a penalty was added to the bill. Since the customer load had not changed, the customer was somewhat perplexed (and perhaps annoyed). The power company investigated.

They found that the two meters had used different ways of calculating the power factor. Both ways were technically correct on the assumption of sinusoidal signals, and both had been certified correct by Measurement Canada. However, the harmonic content of the load resulted in the two calculations not giving the same result.

Berrisford's paper indicates that one meter measured Q by evaluating $Q = \Sigma V_h I_h \sin\theta_{h,}$ that is, evaluating the harmonic terms separately and summing them, and hence $S = \sqrt{(P^2 + Q^2)}$, whereas the other meter calculated $S = VI$ (that is, the rms quantities were directly multiplied). In both cases the power factor was then found using $\mathrm{PF} = P/S$. Because the two meters had different definitions, and responded differently to the semantic coloration of the harmonics, they found two different values for the power factor, each based on accepted calculation methods.

This possibility of different results exists because we have not typically thought of the measurand as a mathematical model whose parameters we are estimating. We are dealing with semantics (the study of meaning) *in the language of mathematics*. The equations of the measurand must be "built in" to the measuring device in some form: the device has to know what coefficients it is to furnish as the result of the measurement. In this example, though the meaning of P was clear enough, the meaning of Q was not.

[6] Berrisford, A.J., "Smart Meters should be smarter", paper PESGM2012-001951 presented at the IEEE PES General Meeting, San Diego, CA July 2012.

If the measurand is merely a verbal description, then the semantic entity “declared value” is not well defined. That is why it should be the parameters of a mathematical model.

It is not that there is always a “right” and a “wrong” solution. In the situation described by Berrisford, the goal of the measurement was to incentivize the customer to arrange for his load to minimize the losses associated with delivering the real power to the customer. These losses are proportional to the square of the current, and there is no way to recover their cost because the electric bill is normally based only on the real power delivered. In a sinusoidal situation, there is only one way to minimize the losses: make the current in phase with the voltage. But with harmonics in the circuit, the situation changes.

Since the current waveform would depend on the phase of the various harmonics, it is not possible to say without knowing the details, just how the I^2R losses might be minimized. As Berrisford suggests, it may even be that the current observed by the electric meter could represent harmonic energy being removed from the power system. It would be possible, using digital measurement technology to create a bill based on the 60-Hz component of the load, and bill separately for the harmonics. Further, if the customer was actually removing distortion from the line, the utility might consider paying for the service, rather than billing for it!

The semantic imprecision in this example comes from English-language labeling of the measurand. The quantity measured in each meter was well-enough defined, but the label “power factor” should not be applied to both measurands since they are not the same.

Something similar is going on in the situation of the phasor measurement unit: it assumes the measurand is of the form

$$v(\mathrm{t}) = V\cos(\omega\mathrm{t} + \varphi) \qquad (1)$$

and it attempts to find the parameters V, ω and φ, as well as a parameter $\mathrm{d}\varphi/\mathrm{d}t$. There are two problems. First, the realized quantity is almost certainly rich in harmonics, so the equation does not describe the realized quantity. Second, if the equation is true, then by definition $\mathrm{d}\varphi/\mathrm{d}t$ is zero. *Yet something is measured.*

The impact is significant. The effort of finding the values of these parameters brings to mind a remark of Samuel Johnson’s. It “is like a dog's walking on his hind legs. It is not done well; but you are surprised to find it done at all.” The accuracy of the PMU is indeed a tribute to the designer. We will return to the subject of phasor measurements later.

3.0 Towards Better Measurements

3.1 Relation to GUM

It is tempting to think of the noise source in Figure 5 as the origin of Type A uncertainty in the result of the measurement. In that case, should the semantic coloration (noise) source be considered as giving rise to Type B uncertainty?

The matter is not obvious. First, let us agree to broaden the use of the terms Type A and Type B a little. In the narrowest sense, the terms are applied to the method of evaluating uncertainties (see GUM 3.3.4), but GUM (3.3.3) does allow that the categorization gives rise to two uncertainty groups. Repeated measurements establish a probability distribution that can be treated by the methods of statistics. That is Type A. All else is Type B (see GUM 0.7). However, GUM (3.3.5) states that the estimated variance for a Type B uncertainty is u^2 based on an assumed probability distribution – a statement that at least implies using the methods of statistics! (That is, GUM seems a little inconsistent.)

The difference between Type A and Type B is surely that longer-term averaging or repeated measurements (which reduces the standard deviation of the mean) can reduce the uncertainty resulting from Type A. That would seem to make semantic coloration a source of Type B uncertainty. Treating the matter by longer and longer integration times (or sampling windows) or taking more readings are ways to reduce the effect of “engineering noise” only, and not semantic coloration.

As we saw in Figure 2, at some point the variance of the reading starts to increase. We can say that what is being revealed is indeed semantic coloration: the meaning of the result is changing. See Figure 7.[7]

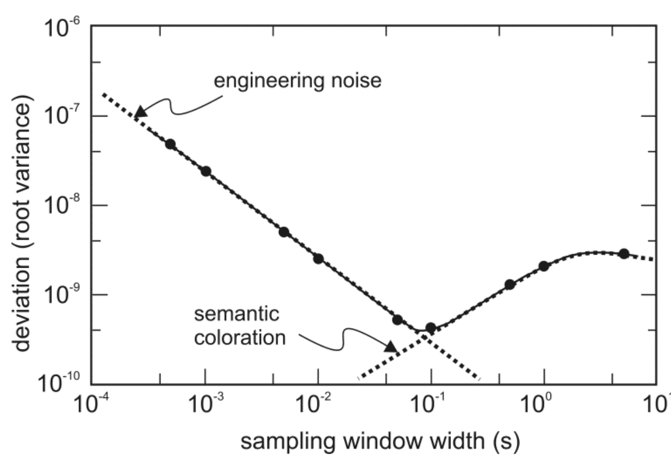

Figure 7 Variance changes identified

It is curious that engineering noise decreases as the sampling window gets wider, but the effect of semantic coloration gets larger. That the semantic coloration occurs on a long time-scale is not likely a

[7] It should not be thought because of this diagram that semantic coloration can be neglected for window widths that makes its effects smaller than that of engineering noise. The diagram shows the effect on the variance of the result, not on the value of the result. Type B errors (or semantic coloration) can have an effect on the value without affecting the variance: a dc offset, for example.

general truth. It just happens that *some* semantic changes do occur slowly, and these are revealed once the effect of engineering noise is sufficiently reduced. (As the developer of an optical measurement system once observed, “that optical current transducer is a very good thermometer!”) Other colorations may exist at higher frequencies and remain hidden in the Allan variance plot.

The readings with sampling windows longer than a tenth of a second or so are changing because the measurement is being influenced by parameters other than the ones intended. That is not noise in the same sense as engineering noise.

It might be well to redefine Type A uncertainties as those caused by the random noise effects, and Type B uncertainties are those that, with a variety of causes, change the meaning of the measurement result. That is a matter that will require discussion among the community of metrologists.

Just as Shannon recognized that the proper encoding of signal in a noisy channel could result in a received message with arbitrarily low error rate, so it may be that the proper treatment in the measurement algorithm could reduce or eliminate errors caused by what we are calling semantic coloration.

With something akin to a matched-filter approach, the algorithm and the measurand might be connected in a way that reduces the coloration effects. Whether the problem always requires anticipation of the kind of coloration in the signal may be a topic worth exploring. If the signal is expected to be cyclic, could the matched filter method produce a useful template for the signal?

While that is an idea that may be worth pursuing, the most obvious source of semantic coloration in any measurement is the fact that in the real world, the quantities being measured are rarely constant. In very general terms, that is a problem that deserves attention.

3.2 Unchanging signals

The process of measurement requires (though it is rarely stated explicitly) that the signal being measured does not change. Recall: the graph of Figure 1 would be meaningless were that not the case, because repeated measurements with different sampling windows would be measurements of different things. The sine-wave of Equation 1 has three parameters; only if they do not change between measurements can the amplitude histogram of Figure 1 be drawn meaningfully.

In the general, ordinary meaning of a measurement, and in particular in a calibration done in an NMI, an unchanging signal is a requirement. In the more ordinary “real” world, we do have to allow for changes. Let us see what it means if we allow for change, beginning with a slow change. (After all, if nothing changed, there would scarce be a point to making a measurement.)

What if the signal really is changing only slowly? That would be the same as a small amount of semantic change in a given period, but it would not be noise in the usual sense. How much should the change be allowed to be? The question is a way of asking how good an approximation is the realized quantity to an unchanging signal. Remember, the language of our semantics is mathematics. What difference does it make if the change in parameter values is (say) 0.1% or 0.5% between the beginning and the end of the analyzed sample stream? Probably the answer depends on the character of the measurement algorithm and the accuracy required of the measurement.

If we imagine the measured parameter changing linearly with the time, it seems likely that the declared value, applicable at the center of the window, would be quite close to correct. On the other hand, it might be good to simulate that condition to be sure, because there might be an error as a result of the way the calculation is done. There may be value in exploring ways to ascertain this condition by further analysis of the signal, but let us leave that aside for the moment.

What if the rate of change is greater? If we assume that the estimator (algorithm) is aligned with a fixed-parameter model (measurand), then the mismatch between the assumed unchanging coefficients of the model and the actual changing signal is a mismatch between the realized quantity and the algorithm, and no statistical treatment of multiple measurements will fix that. That makes it semantic coloration.

It is interesting to examine the phasor measurement unit (PMU) in this light. In the PMU the *voltage* is changing (for example), but it is doing so in the cyclic manner captured by Equation 1. If the rate of cycling (the frequency) is constant, we have met the condition for unchanging coefficients in that the parameters of the model are constant. But suppose the frequency is not constant.

With a signal generator, it is perfectly possible to create a signal that has the appearance shown in Figure 8. It is the sort of waveform that might result from a generator that was coming rapidly to a halt.

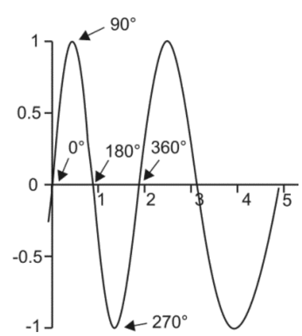

Figure 8 Voltage output from a generator slowing rapidly

Here we have labeled the points on the curve that are "obviously identifiable" by similarity to a sine-wave. If we were to examine the voltage coming out of a power generator that was slowing rapidly we might see something like this, and the angles shown would be related to the physical angle of the rotor. But for most purposes that is a cheat, because the term "phase" as we use it[8] applies to text-book signals that are the same for all time. The textbooks will tell you it has no meaning for signals like this.

The PMU produces its estimate of the parameters and assigns them to the time specified in the appropriate standard. (Usually the window is centered on this pre-determined time.) The estimation

[8] For a discussion of other ways to use the word "phase," the reader is encouraged to read "The Fundamental Principles of Frequency Modulation" by B. van der Pol, Proc IEE (Part III: Radio and Communication Engineering), Vol 93, No 23, pp 153-158, London, 1946. In this paper, van der Pol argues that "*the whole argument of the cosine function, namely* ($\omega t + \psi$) *is the phase*. This definition has, among others, the advantage of enabling one to speak of a phase difference of two oscillations of different frequencies." (The italics are in the original.) In the implementation of the PMU, the phase calculated by any on PMU is always between two signals of different frequencies, since one is a constant (and imaginary) 60 Hz. Figure 8 notwithstanding, phase is *not* defined for such conditions.

problem for the PMU is therefore one of estimating the *local* frequency, that is, the frequency in the region around the time of interest.

In related work, Kirkham and Dagle (2014) modified Equation (1) and created a form in which each parameter was allowed to have a (constant) rate of change. That modification allowed a curve fit to be done of signals like the one in Figure 8. Thus,

$$v(\mathrm{t}) = V\cos(\omega\mathrm{t} + \varphi) \tag{2}$$

becomes

$$v(t) = (V + C_V t)\cos\{(\omega + C_\omega t)t + (\varphi + C_\varphi t)\} \tag{3}$$

which reduces to

$$v(t) = (V + C_V t)\cos\{(\omega + C_\varphi + C_\omega t)t + \varphi\} \tag{4}$$

The work was done to create a new algorithm for solving the phasor measurement problem. Equation (4) allowed the amplitude, the phase, and the frequency to change, and the values could be plugged into a spreadsheet to generate test "waveforms" in the form of tables of data.

Consider the case of the PMU in a situation where just the "frequency" is changing. Using the "signal generator" defined by Kirkham and Dagle, the waveform would have the form

$$v(t) = V\cos\{(\omega + C_\omega t)t + \varphi\} \tag{5}$$

where C_ω is the rate of change of frequency, often called ROCOF.

The PMU, however, is obliged to find the parameter called ROCOF as defined by the following series of equations. First, there is the input signal:

$$x(t) = X_{\mathrm{m}}\cos[\varphi(t)] \tag{6}$$

where $x(t)$ is the instantaneous quantity, X_{m} the peak amplitude and φ the phase. From this, a rate of change of phase is defined, and called frequency:

$$f(t) = \frac{1}{2\pi}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} \tag{7}$$

and from this an equation for ROCOF is derived:

$$\mathrm{ROCOF} = \frac{\mathrm{d}f(t)}{\mathrm{d}t} \tag{8}$$

Equations (6) through (8) are equations (7) through (9) in the relevant IEEE standard.[9] From these, the standard illustrates a finite-difference method of performing the calculation. Phase (assumed constant for the purpose of each step of the calculation) is measured according to Equation (6) in two successive sampling windows. The difference is calculated, and since the time difference is known, Equation (7) is solved in finite difference form. ROCOF can be calculated after one further sampling window.

The algorithm in the PMU is designed to find ROCOF via Equations (6) through (8). If the signal is better represented by (5) there may not be a good match. In the Introduction to the standard is this disclaimer:

> The user shall be aware that in the presence of the previously mentioned undesirable components in the input signal, higher measurement errors could result. These errors may be substantial, particularly where higher order derivatives (such as ROCOF) are used. Signal processing alternatives may be employed to reduce or eliminate these errors. They are difficult to implement in a real-time environment and could adversely affect the measurement latency or the synchrophasor measurement response time.

In other words, the ROCOF results are not to be expected to be very good. The working group attributed all the problems to what we are calling semantic noise. Surely, semantic imprecision is also contributing to the problems. We will return to this matter below, when we discuss calibration.

So difficult is the measurement that few, if any, PMUs were able to meet the standard. Soon after the Standard was issued, the Standard Working Group issued an Amendment[10] that increased the allowed error on ROCOF by a factor of four for some conditions, by a factor of ten for others, and completely suspended the limit for yet others. It is an odd situation that a Standard has to be relaxed so that it can be met!

So far, the situations we have considered have involved relatively well-behaved signals. Sometimes, however, a real signal does things that are not allowed for in the equations of the measurand or the measuring instrument. For example, in a PMU, the signal may represent a step change in phase because of a switching operation in the power system, as seen in Figure 9. In this figure, a step change of phase occurs near the center of the period (time = 0). The frequency is unchanged, and so is the amplitude.

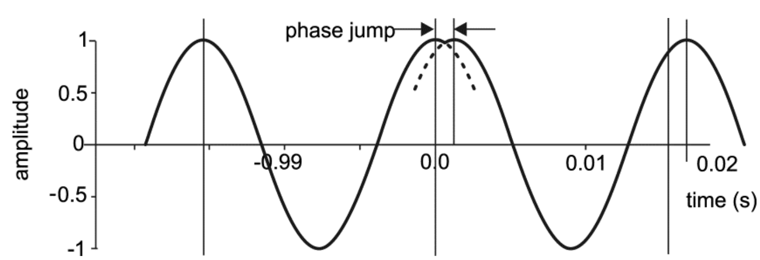

Figure 9 Voltage signal with small change in frequency

[9] IEEE Standard for Synchrophasor Measurements for Power Systems, IEEE Std. C37.118.1 - 2011, Dec. 2011.

[10] C37.118.1a Draft for IEEE Standard for Synchrophasor Measuremnts for Power Systems, Amendment 1: Modification of Selected Performance Requirements, 2014

The change is abrupt: in fact, it is instantaneous. The solution to the equation before the step in phase is not the same as the solution after the step. It is fair to ask what we would *want* the PMU to give as its declared value.

There is surely more than one way to answer that question! The measurand (built into the PMU as the equation whose coefficients it is estimating values of) consists of versions of the phasor equations, and contains no terms for step changes. The question becomes in general What do we want the measurement device to say when the realized quantity does not match the form of the measurand?

Unless we answer the question, we should not be surprised at any PMU response.

3.3 Predictions

When the realized quantity does not "match" the measurand, the available information can be used to estimate the quality of the measurement, and report that the measurement quality dropped. Consider the information available to the measurement device. It knows the form (the equation) of the measurand. It knows the values of the coefficients of that equation. And it knows the sampled values of the realized quantity, because these were the starting point for its calculations.

It follows that the measurement device could calculate the expected values of the equation that is the measurand at the times that correspond to the sampled values. It could calculate the differences, and form the sum of the squares. That number would be indicative of how well the result of the measurement explained the realized quantity. Just how to use that information remains to be seen.

An interesting situation can be explored by considering the world as it is perceived by the PMU. Figure 10 is an example of a signal that is proceeding along looking like a sine-wave, when the next value in the sequence from the A/D converter is not the expected value. The dots in the figure are A/D sample values. The line is a sine-wave fit to them. What are we to conclude based on the sudden non-fit?

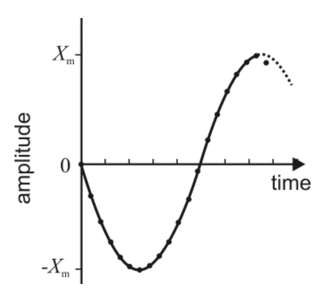

Figure 10 Sampled values changing unexpectedly

Figure 11 indicates that there are four possible causes for this sudden change. It could be that the unexpected value of the sample is caused by (engineering) noise, and the signal will revert to the original track with the next sample. That is shown in (a). Or it could be that the amplitude, the frequency, or the phase of the signal has changed, as shown in Figure 11 (b), (c) and (d).

Except for (a), the change is not associated with *noise*, the change is a semantic one: the meaning of the signal has changed.

Each of these outcomes is associated with a certain probability.[11] As far as the declared value on an instrument is concerned, the result should change from indicating a sine-wave with certain parameters (and a small variance with a high probability) to (suddenly) a set of fifteen alternatives, each with its own statistics. As the signal evolves in time, these probabilities can be updated, and at some point one outcome selected. *A priori* knowledge of the system being observed can be used to assign initial probabilities. These can be updated as more information becomes available.

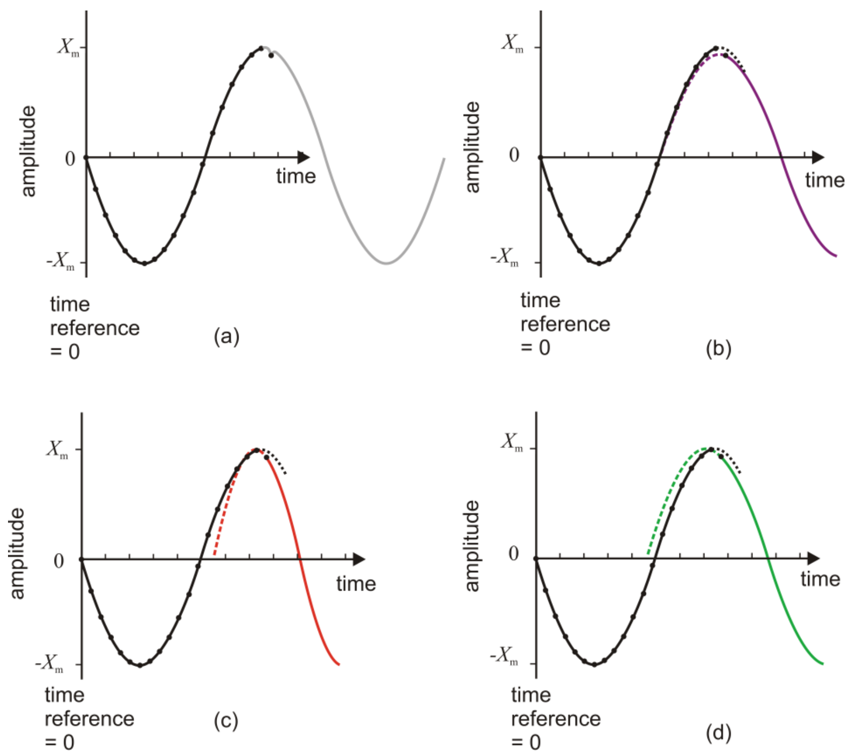

Figure 11 Possible causes for changed sample value

If this sounds familiar, it may be because it is the sort of problem addressed by the Viterbi algorithm. The Viterbi algorithm is used in applications that, like our problem, require the most likely (hidden) cause of a system state to be found. It is a recursive algorithm that reduces the calculation burden of evaluating all the possibilities. It is usually applied to find the (hidden) causes of a set of observations, in association with statements about hidden Markov processes.

Viterbi may not be the best algorithm for all such problems in metrology, but it seems that we do sometimes have (hidden) causes for our observations. Investigation of Viterbi and other methods of assigning probabilities could prove to be valuable adjunct to the familiar least-squares estimators used in measurements.

---

[11] The diagrams were drawn on the assumption that only one parameter changed at a time (*n*, *a*, *f*, *p*) where *n*, *a*, *f*, and *p* stand for noise, amplitude, frequency, phase.. There are thus four parts to Figure 11. There are six ways that the parameters could change two at a time (*na*, *nf*, *np*, *af*, *ap*, *fp*) and four ways they could change three at a time (*naf*, *nap*, *nfp*, *afp*), and one way they could change four at a time. The total number of ways the sample could have come about is thus fifteen.

The single sample in Figure 10 that does not fit expectations increases the number of possible explanations to fifteen. But there is no reason to suppose that the same possibilities did not exist prior to this observation. They were just better hidden. In fact, what we have been thinking of as the declared values for amplitude, frequency and phase in the presence of noise should *all along* be thought of as just a single set of probabilities chosen as the most likely set from a larger set.

This is further evidence that the usual statement of the result of a measurement is *in general* incomplete. A more complete statement would include a statement of *several alternative sets of values* for the declared value, the uncertainty and the confidence level, at least if the probabilities were not vanishingly small. And the lack of such additional information should be taken as indicating that the probabilities were vanishingly small, with the threshold for vanishing defined, and possibly user-selected.

The incompleteness of the usual statement of declared value, uncertainty and confidence is evident in the earlier anecdotal example of the oscillating power supply. The three conventional numbers in the result statement would likely have indicated a declared value of 400.00 with an uncertainty of 0.01 and confidence of 99%. However, they did not convey the nature of the realized quantity. As we saw, since the readings were very consistent, the user would have no clue that the measurement was of low credibility.

There is a problem in that the number of possibilities (of alternative explanations for the observations) is large, and a general-purpose way of expressing the (hidden) causes for the observations may be impracticable. The particular problem of excessive ripple might have been anticipated if the possibility of at least some ripple had been allowed for in the measurement, and if the user had looked for it. But it was not allowed for in the general-purpose instrument chosen to display the voltage.

The PMU example is of a specific instrument for a very particular application, and it may be that in any instrument whose application is known with this level of specificity, it is reasonable to use *a priori* knowledge to anticipate the possibilities given by a changing realized quantity. But in the general case it is not reasonable, nor is it necessary, to find all the possibilities as we did in Figure 11.

A statement of the coefficient of determination (r^2) would be a way to expose problems. Others measures surely exist. The user (whether a person or an application) would be informed by this fourth number that the statistics of the reading no longer explained the variations in the samples as well as before.

Applied to the case of the PMU, this would mean that as soon as the single unexpected sample is received, the r^2 value is decreased. The PMU user may be able to assign new probabilities to the various options, and update them as the time increases. Applied to the general-purpose instrument and the oscillating power supply, the user would be informed that the straight-line dc value was not a very good explanation of the samples. If the user merely wanted the rms value of an alternating signal, that would be understood to be acceptable.

Note that such information is not equivalent to the statement of uncertainty or confidence that would accompany the curve of Figure 1. Those numbers have to do with the output: the distribution of the results of the measurement. Our concern now is with the input: the degree of fit of the samples. What the r^2 value discloses is, in essence, a measure of the match between the defined measurand and the realized quantity. Any observed mismatch could be the result of engineering noise or semantic coloration. That makes the notion worthwhile.

3.4 Calibration

In metrology terms, calibration compares the declared value of an attribute of a calibrating artifact, such as a reference standard, against the declared value of an attribute of a unit under test. There are several ways to do this comparison that are different in principle. The choice depends on the nature of the unit under test.

In the calibration of a measuring device, the calibrating artifact provides a signal to be measured, often termed the stimulus. The calibrating artifact's declared value is its nominal or indicated value. The unit under test is the measuring device, which responds to the stimulus and gives its declared value as output for comparison, as in Figure 12.[12]

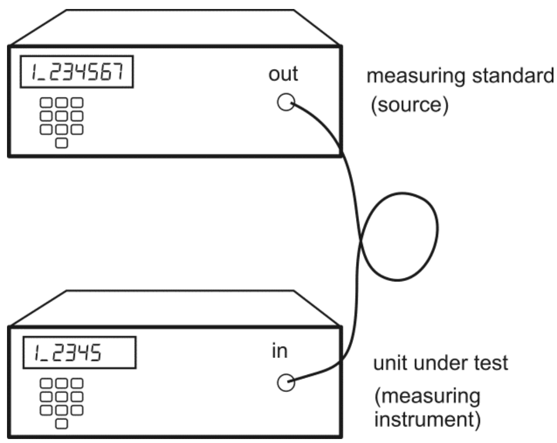

Figure 12 Calibrating a measuring device

In essence, the measurand in Figure 6 is replaced by the nominal value of the source in the calibrating artifact. The stimulus is the realized quantity as far as the instrument is concerned. The goal of the designer of the source is to produce a stimulus that is as "pure" as possible. Semantic coloration can then be (optionally) added in a controlled fashion. Figure 13 shows the framework for calibration.

The calibration process expects the declared values of the source and the measuring device to be similar. Differences are attributed to errors in the measuring instrument. The calibrating artifact is generating its signals from exactly the sort of equations that the device under test is supposed to measure, so that the "correct" response is known: it is the nominal value of the source.

That means that the equation that defines the stimulus must be the same as the equation whose parameters are being estimated by the measuring system. Without agreement between equations, it cannot be asserted that differences between the declared value of the stimulus and the declared value of the measurement device are not the result of semantic imprecision – in this case the mismatch between equations. Semantic imprecision has to be removed so that differences can be said to be due to measurement error.

[12] In another configuration, the stimulus is supplied by a source external to both the calibrating artifact and the unit under test. Each artifact responds to the stimulus and gives a declared value.

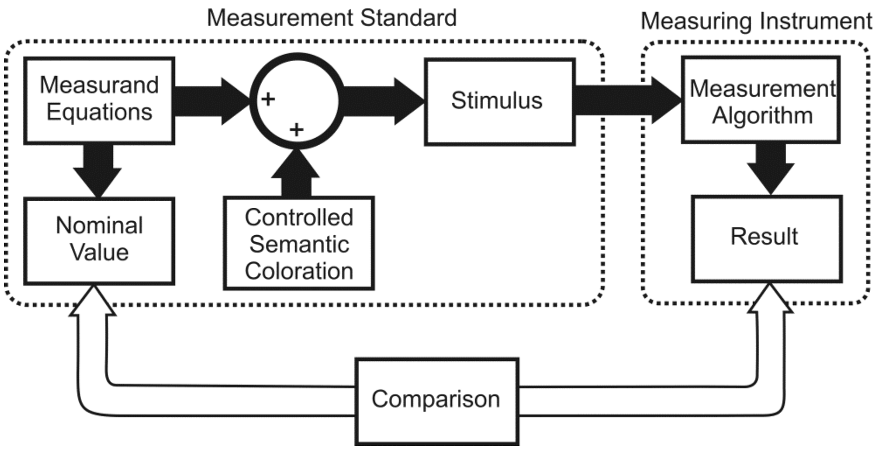

Figure 13 Calibration, shown in the "framework" view

Let us return to consider the PMU. If the signal source used in calibration uses the form given by Kirkham and Dagle, there is semantic imprecision, because those equations are not the ones built into the PMU. At least part of the reason for the debate over the meaning of “frequency” lies in this semantic disconnect. And because of that, the problems with ROCOF measurement are certain to be partly the result of semantic imprecision.

We must expect at least small differences between the source values and the coefficients found by the PMU. But the PMU equations have a completely different form: for example, they allow for no rates of change in the solution inside each measurement window.

In the PMU, the rates of change are calculated using finite differences between windows. If the signal is changing across a sampling window, the measurement in each window is at best approximate, and the result after the finite difference calculation is therefore also only approximate. How good or bad the approximation is emerges from the calibration tests.

To reduce the semantic imprecision, it would be possible to redefine the measurand (the signal source equations) to match the PMU method. But that would certainly reduce the similarity between the calibration signals and the kind of signals seen in the real world. Such a change would not improve the users’ confidence in the devices in the field.

We are left with two options: we could change the way the PMU works so it matches the signals being presented to it, or we can improve our understanding of the effects of the various approximations that are made because of the semantic imprecision. At the time of writing (2015) confidence in the ROCOF values is very low indeed.

4.0 Summary and Conclusions

The act of making a measurement is essentially one of data compression, done so that a human mind can grasp something complex, no matter that the complexity arises from quantum physics or power engineering. Often, the measuring is done so that some further operations can be performed, based on a model of reality, that give us humans something we interpret as "information." Human minds have created many models of reality into which measured information can be inserted, and from which further information can be drawn.

What that means is that we must carefully define what we expect from the process, so that we can trust the result of a measurement when we match it to the models we have created. We have illustrated here a number of interconnections between the elements of the process. We have shown that the measurand is more than a verbal description of the thing to be measured, it is a set of equations whose coefficients are being estimated by the process of measurement.

The debate in PMU circles about the definition of frequency can be easily settled with a very permissive definition: frequency is the name given to the declared value of a device that includes a measurement of something it calls frequency. In a PMU, the measurement is based on a finite-difference calculation, a process that occupies the time of two sampling windows. That is the semantic definition of frequency for a PMU, because that is what its measurement result means.[13]

We have introduced the terms *semantic coloration* and *semantic imprecision* and seen that these are useful ways to think about measurement. Examples of each are given to explain unexpected results. Consideration of the semantics (the meaning) of the result of the measurements is in fact an essential part of the design of an instrument. The language of this semantics is mathematics, not a verbal human language. Such consideration adds to the understanding of what GUM expresses, and may show ways to improve the measurement process. Figure 14 shows the overall process of measurement, and in particular the two semantic challenges, coloration and imprecision, that it faces.

For some while it has been good practice in metrology to quote the result of a measurement as a certain value (the declared value) and the associated uncertainty. More recently, it has been recognized that the uncertainty is associated with a confidence level. The specification of the result of a measurement is therefore a set of three numbers: the declared value, the uncertainty and the confidence level.

We have shown that this generally accepted specification of the result is incomplete. It fails to assess the quality of the measurement process. The use of a parameter such as the coefficient of determination (r^2) to indicate something about how closely the realized quantity represents the measurand, and would give confidence in the choice of the model.

[13] The "new PMU" method of Kirkham and Dagle does not calculate the same way, so it does not measure that frequency. Its definition would not be compliant with the standard for PMUs. However, it would likely give results that were much closer to the calibration signals used to evaluate PMU performance, because those signals almost certainly have very similar equations, and are not based on finite differences.

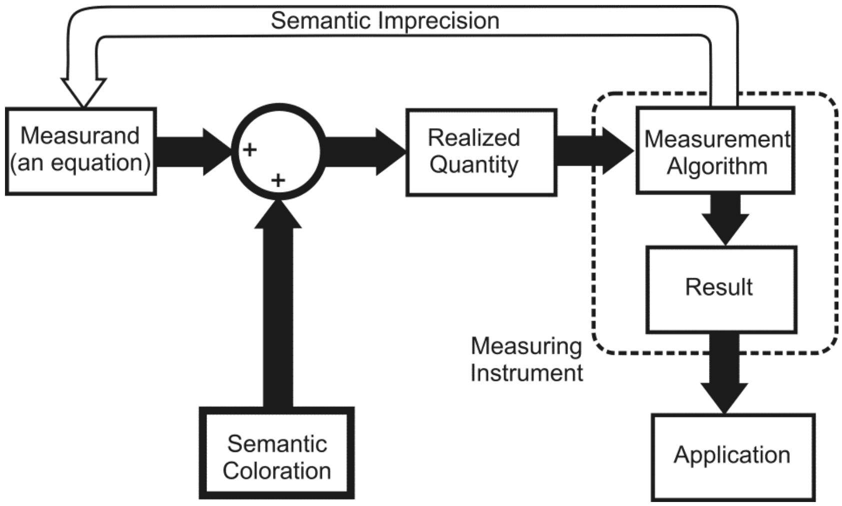

Figure 14 The process of measurement

That assessment of the quality of the measurement is a way to gauge the impact of semantic problems. A solution with a good match between the algorithm and the realized quantity has demonstrated that the levels of semantic imprecision and semantic noise are both low. If the equations in the semantic entity called the measurand are properly reflected in the measurement algorithm, the measuring device has all the information required to calculate the effect of coloration in the realized quantity, whether it is semantic or engineering noise.

We have seen that the conventional result of a measurement (value, uncertainty and confidence) should be considered as an expression of just one of a set of probabilities. If one observes a voltage with an instrument that is suited to measure a constant voltage, that is what one will observe. If one used an oscilloscope instead, one might see a quantity of very different character.[14] In a measurement that does not have a high "quality of fit," the simplest fit measure does not reveal where the problem might lie. In general, there could be many possible explanations.

For some well-understood signals, the quality of fit between the result of the measurement and the observed signal is affected by a small number of known effects, and a change in the quality of fit may reflect a change in any one of the known effects. Since a change in the signal is a common cause of semantic imprecision, the use of an algorithm such as Viterbi may be a way to learn very quickly how the meaning of the signal was changing.

Semantic considerations levy requirements on calibration as a way of assessing and confirming the quality of a device. As applied to measurement, the equipment generating the stimulus must use the same equations (representing the measurand) as the device being calibrated. Without that correspondence, there

[14] It is interesting to speculate that the same apparently-dual nature of light may reflect no more than the way we model and observe the signal. Newton imagined it was a wave-like thing. He was able to devise experimental evidence – but the experimental evidence was in support of his mental model, and showed nothing about what light "was." Einstein had a model of light that was based on "quanta of energy." Much work has been done that shows that model was also a good one. At the time, there was debate over whether light in a vacuum "was" waves, but in interacting with materials "was" quanta. That debate cannot be resolved, for instrumentality is necessarily material. But the models are both conceptual, and both work.

is semantic imprecision. Any observed differences between the declared value of the stimulus and the results from the measurement device may reflect this semantic imprecision in the calibration process, whose value is therefore diminished.

Appendix A. Postscript on PMU technology

If the equations of the calibration source are not reflected in the measuring device, as in the PMU situation, the logical question to ask is "Why is there a difference and how should it be resolved?" Consider: the *measurand* is the entity that Carey Foster said was "the precise expression which makes it possible for [concepts] to become the subject of mathematical reasoning." As described in the IEEE standard the measurand is supposedly the series of equations that give rise to the finite difference definitions. This definition arose as a result of much labor by engineers whose mental model of frequency was the rate of change of phase, because that was what the textbooks said.

The equations in the standard do not, in fact, make a definition of a thing to be measured. They are, rather, an algorithm, a *method of measurement*.

The signal to be measured is surely of the form shown in Equation (4)

$$v(t) = (V + C_V t)\cos\{(\omega + C_\varphi + C_\omega t)t + \varphi\} \tag{4}$$

and so is the signal from the calibrator. In order to generate a signal that matched the equations of the PMU, the signal generator would have to generate a signal of the form known as simple harmonic motion:

$$x(t) = X_{\mathrm{m}}\cos[\varphi(t)] \tag{6}$$

To use this as a calibrating signal, X_{m} and φ would have to be constant for a period *corresponding to the measurement window* in the PMU. At the end of that time, there would be a signal with a different φ, constant for the next measurement window, and so on. It could be done, but it would not be realistic as a representation of the power system. The power system would not heed the sampling windows of the PMUs!

At this point in the technology of the PMU, it is evident that the signals from the calibration devices are well known and understood, and reasonably reflect the sort of signal that would come from a power system. If the semantic imprecision is to be reduced, it is the PMU that has to change.

And that is precisely what the "new PMU method" of Kirkham and Dagle does.

If a manufacturer implements and markets a PMU based on that technology, it will achieve calibration results that are far better than the competition. And yet it would be defined as not complying with the standard.

There is an odd situation indeed! To make compliance possible, the standard should be rewritten so that it describes the model of the power system whose parameters are to be found. That seems unlikely as far as the IEEE standard is concerned. An IEC standard is now being developed, and the present author is one of the US representatives to the IEC working group. It will not be easy to persuade people of the validity of the work in this report, but it is worth the effort.

Bibliography

GUM: BIPM *Guide to the Expression of Uncertainty in Measurement.* International Organization for Standardization, Geneva. ISBN 92-67-10188-9, First Edition 1993, corrected and reprinted 1995.

Kirkham, H. and Dagle, J. (2014) Synchronous phasor-like measurements, Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES, DOI: 10.1109/ISGT.2014.6816510

Shannon C.E., (1948) A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October, 1948. Available at http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf accessed February 2015

Weaver, W. (1949) Recent Contributions to The Mathematical Theory of Communication, available at http://waste.informatik.hu-berlin.de/Lehre/ss11/SE_Kybernetik/reader/weaver.pdf accessed February 2015.

Shannon C.E. and Weaver, W., (1963) The Mathematical Theory of Communication, Illini Books.

Distribution

Joe Eto, Lawrence Berkeley National Laboratory, JHEto@lbl.gov

Jerry FitzPatrick, NIST, fitzpa@nist.gov

Jay Giri, Alstom, jay.giri@alstom.com

Dmitri Kosterev, BPA dnkosterev@bpa.gov

Alex McEachern, PSL Alex@PowerStandards.com

Phil Overholt, US Department of Energy, Philip.Overholt@hq.doe.gov

Tony Weekes, Manitoba Hydro, maweekes@hydro.mb.ca