Lies, Damned Lies, and the Statistics of Data Fitting

Richard Steven, CEESI rsteven@ceesi.com
Eric Harman, CEESI eharman@ceesi.com
Terry Cousins, Taskflo Ltd taskflo.ltd@gmail.com

1. Introduction

A core requirement for achieving low uncertainty flow metering is that each step of the quality control process has integrity and is rigorously adhered to. Calibration is a critical part of many flow meter quality control procedures. The calibration process consists not only of testing the flow meter at a reputable flow laboratory across the appropriate flow conditions, but also of how that recorded data is subsequently implemented into the flow metering system's calculation routine. This second part of the calibration process, often called 'linearization', 'curve fitting', or 'data fitting', is a *critical* part of flow meter quality control. There are multiple ways of linearizing a flow meter, and yet, few meter manufacturers, end users, standards boards, or regulators have stated comprehensive published rules regarding this process.

Some standards text touches on linearization, e.g. AGA 9 Ed 2 says of ultrasonic meter (USM) linearization "… use a polynomial algorithm, piece-wise linear interpolation, *or other industry accepted method*". But what qualifies a linearization technique as an 'industry accepted method'? Such text is open to interpretation. ISO 17089 USM standard also leaves the choice open suggesting a flow-weight mean error, piece-wise linear interpolation, or a polynomial fit. ISO 5167 states for low uncertainty, Venturi and cone meters should be calibrated, but the text does not give any guidance on the subsequent data fit methodology. There are many choices but no published consensus. Although flow meter linearization tends to be thought of as an objective exercise, it is in reality rather subjective.

The integrity of a flow meter's linearization technique is critical to the integrity of the overall flow metering process. Lax attention to the technical details of flow meter linearization undermines the integrity of the flow meter. It represents a potential hole in flow metering quality control. It is arguable that flow meter linearization techniques should ideally be more regulated, and discussed in more detail in flow meter standards and contracts. A standard would set out benchmarks for linearization techniques to be compared to. However, this may be very difficult to practically implement. A pragmatic compromise is a call for far more awareness, disclosure and clarity between manufacturer and end user regarding which linearization technique is being applied, why, and what the pros and cons of that choice are.

2. Calibration Data, Data Fitting and Computers

Flow meters produce an output signal that is relatable to the flow rate via some theoretical relationship. For example, a turbine or ultrasonic meter output signal is theoretically directly proportional to flow rate, and a Differential Pressure (DP) meter output signal has a theoretical parabolic relationship with flow rate. If these theoretical relationships were the whole story in practice there would be no need to calibrate and linearize flow meters. However, in practice this is not the case. In reality there are always various issues, such as meter manufacturing tolerances, secondary physical influences, installation effects etc. that are not accounted for by the theory. For low uncertainty flow metering these influences have to be removed from the meter output, i.e. they are 'calibrated out', sometimes referred to as 'linearizing the flow meter'.

Regardless of standard bodies definitions of flow meter 'calibration', in colloquial terms it tends to mean one of two things. It can mean testing a flow meter against a reputable standard to produce a data set of a correction parameter (e.g. a 'K-factor', 'meter factor' etc.) vs. a correlating parameter (e.g. Reynolds number, volume flow rate etc.). Or it can also mean the follow on acts of then choosing a mathematical expression to represent that calibration data set, i.e. the linearization technique, and then entering that expression into a computer. This computer carries out the flow rate calculation using inputted geometries, fluid property values, the meter's primary signal/s, and a combination of the relevant theoretical calculation and the chosen linearization technique. That is, the flow rate calculation is comprised of two concurrent parts, i.e. the standard theoretical flow rate equation, and the chosen calibration based mathematical expression that corrects for the difference between theory and practice. Hence, *the chosen calibration data based mathematical expression is central to the meter's flow rate prediction output, and therefore that flow rate prediction integrity*. A poor choice of calculation routine means a poorly performing flow meter, with perhaps a relatively high uncertainty or bias.

A modern generic flow meter system comprises of a meter body (the primary component), instrumentation (the secondary component/s), and a computer to gather data and calculate the flow rate. There are generally three alternative locations for such a computer. There is the 'processor' (sometimes called the 'transmitter' or 'meter head') embedded in a flow metering system, and is an integral indelible component of a manufacturer's flow meter package. Then there is the generic stand-alone 'flow computer' products supplied by reputable suppliers that can be installed on many generic flow meters such as Venturi, cone, turbine meters etc. Users can procure the meter body and instrumentation separately from this flow computer. Then there is the third option for the end user to use the mainframe computer that monitors and controls the overall process in which the metering system is a sub-system.

The end user typically programs the curve fit into the mainframe computer or stand-alone flow computer product and hence in those cases they have a fully transparent calculation method. Such flow computers also tend to have an audit function that produces clear reports. Processors that are embedded components of flow meter systems can contain manufacturer inserted undisclosed inaccessible linearization code, and are by nature not very transparent. In such a scenario the end user is not automatically guaranteed a fully transparent flow calculation routine.

Manufacturers of flow computers can be held accountable to standards (e.g. API 21.1 for gas application flow computers). However, manufacturers of embedded processors are outside the scope of such documents and such scrutiny. There are no equivalent regulatory documents for embedded processors. Furthermore, embedded processors tend to produce audit information that is arcane, i.e. unusable by most end users. If the flow rate calculation routine is not fully disclosed this introduces and element of mystery, a potential source of bias, a hole in the uncertainty statement.

In this paper various choices of linearization techniques are discussed, and the pros and cons of different methods are considered. The opportunity offered by lack of rules to subtly manipulate the meter output through the chosen linearization techniques to give positive or negative biases at certain flow rates is described. First though, before curve fitting is discussed, it is necessary to discuss what data is being fitted, i.e. what correlating parameter should be chosen, and why. Only then, can the method of fitting be properly addressed.

3. K factor vs. Flow Rate / Velocity or Reynolds Number

"It is difficult to get a man to understand something when his salary depends upon his not understanding it." - Upton Sinclair

Flow metering is a competitive market. There is considerable commercial pressure to keep the price of the product down. Flow meter calibration can be expensive. However, for custody transfer or fiscal metering applications where so much money is at stake virtually nobody argues against the need for flow meter calibration. What is debated is just how this necessary calibration and linearization is carried out.

Some flow meter manufacturers have in-house calibration facilities. These facilities tend to be liquid flow facilities, usually water facilities, and occasionally oil facilities. They are rarely gas facilities. Meter manufacturers find liquid flow facilities more attractive than gas flow facilities for in-house calibrations for the following reasons:

1) they are generally simpler, quicker, safer, and less expensive to operate than gas facilities, and

2) the liquid reference flow rate is usually very accurate compared to gas flow reference meters.

Hence, some meter manufacturers offer an in-house liquid flow calibration, but not a gas flow calibration, as part of the meter deliverables. As long as the end user is okay with the calibration not being carried out by an independent 3rd party, and the calibration flow condition range is appropriate for the meter's application, there is nothing inherently wrong with such practice. However, a problem arises if the manufacturer's in-house liquid calibration flow condition range is *inappropriate* for the meter's application. Unfortunately this is an all too common scenario.

Meter manufacturers may not want to admit to themselves, or their prospective client, when their in-house liquid calibration will cover an inappropriate flow condition range for some specified meter application. Acceptance of such a problem would mean accepting the extra time and expense of sending the meter to a 3rd party gas flow calibration facility, perhaps making the meter less attractive to the prospective client. Some meter manufacturers seem to be genuinely ignorant of this issue, perhaps a case of "it is difficult to get a man to understand something when his salary depends upon his not understanding it". Others may perhaps operate under 'willful ignorance', or at least have an unofficial 'if they don't ask / don't tell' policy.

The two most common examples of this issue are water calibration data vs. gas or oil flow application performance. At the core of this issue is the significant difference in viscosities between gas, oil, and water. Flow meters are fluid mechanics devices. All fluid mechanics device performances are influenced in some way by fluid viscosity. Internal flow in pipes is influenced by viscosity via the Reynolds number. The Reynolds number describes the relationship (i.e. the relative quantity) of the opposing inertia force (that can be thought of as the flows propensity to keep moving) and the viscous forces (that can be thought of as the flows propensity to not keep moving). That is,

$$\mathrm{Re} = \frac{inertia\ \ force}{viscous\ \ force} = \frac{\rho UD}{\mu} = \frac{4\dot{m}}{\pi\mu\ D} = \left(\frac{4}{\pi}\right)\left(\frac{1}{D}\right)\left(\frac{\dot{m}}{\mu}\right) \tag{1}$$

where *Re* represents Reynolds number

- ρ represents fluid density
- U represents average fluid velocity
- D represents pipe diameter (an arbitrary chosen length)
- μ represents fluid absolute viscosity
- m represents mass flow

Therefore, for any given pipe diameter the Reynolds number is dictated by the ratio of mass flow to viscosity:

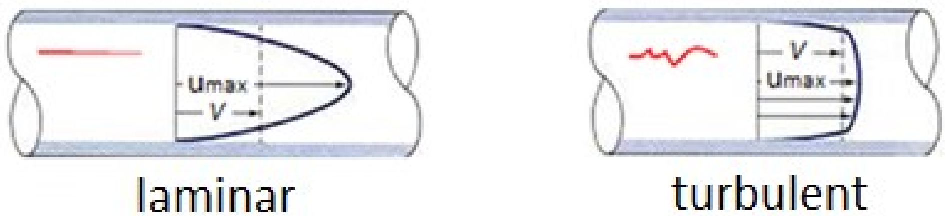

Fig 1. Influence of Reynolds Number on Flow Profile which has Knock-On Effect on Meters.

The Reynolds number dictates the velocity profile (i.e. the velocity distribution) of pipe flow. A low Reynolds number corresponds to a laminar parabolic velocity profile, and a high Reynolds number corresponds to a turbulent flattened ('power law') velocity profile, as sketched in Fig 1. A moderate Reynolds number corresponds to the transitional region between laminar and turbulent velocity profiles. The Reynolds number dictates the velocity profile and the velocity profile influences flow meter performance. To correctly test most flow meter designs performance you have to create the same range of velocity profiles, i.e. test the meter over the same range of Reynold numbers, it will see in the field.

An inconvenient truth that further complicates matters is there is no fixed critical Reynolds number value where flow suddenly switches from laminar to turbulent flow. Transition between laminar and turbulent flow occurs over a Reynolds number range. Furthermore, flow in pipes is rather fickle, with initiation of transition taking place in different pipes at different Reynolds numbers. Transition from laminar to turbulent flow can be delayed by very smooth pipe with no components or protrusions. Transition can be prematurely induced by rough pipe, or components, or protrusions. It is generally assumed that a flow meter will have a reasonable inlet pipe run that is similar to that used during the calibration, and hence the effect on the velocity profile will be similar.

The text books confidentially state that at $Re < 2000$ flow is, at $Re > 4{,}000$ flow is turbulent, and transition takes place at $2{,}000 < Re < 4{,}000$. Unfortunately in the real world it is not as simple as that. There is plenty of evidence (as we will see) that in industrial pipe flows transition from laminar to turbulent may not be complete until Re > 12,000. Transition is a particularly difficult range for flow metering. With no precise predictable Reynolds number range that covers the transitional region, there is no way other than calibration to predict any one flow meter's performance across a given Reynolds number range.

Unfortunately transition can also be unsteady where, within the transitional Reynold number range, the velocity profile can randomly switch between laminar and turbulent, or local pockets of laminar and turbulent flow can periodically coexist. It is generally assumed from convenience, hope, and some limited calibration repeat points that the velocity profile transition is repeatable. Hence, the calibration facilities results over the transition range are assumed representative of future performance in the field. The authors certainly do not have a better alternative to suggest.

As 'seeing is believing' let us now view data sets of various meter calibrations, where it should become evident that the Reynolds number, and not flow rate or velocity is the more appropriate correlation parameter for flow meters.

3a. Water vs Gas Flow Calibrations

Flow meters for gas applications are sometimes calibrated at water flow facilities. Consider again equation 1. For a given pipe diameter and average velocity, while the liquid density is significantly higher than gas density, the liquid viscosity is a couple of orders of magnitude higher than that of gas. This means that even high liquid flow rates produce relatively low Reynolds numbers compared to typical gas flows.

In order to account for the velocity profile influence it is imperative that a flow meter's calibration covers the application's full Reynolds number range. A water flow facility cannot typically cover the full Reynolds number range of gas flow applications. In the rare case where the water flow calibration Reynolds number range does match that of the gas flow application then the water calibration is a valid representation of the meter performance. However, if the water calibration Reynolds number range does not match that of the gas flow application, then it is not a valid representation of the meter's gas flow performance in the field. This is a common problem. The following is a typical example.

A cone DP meter was supplied to a natural gas flow facility (see Fig. 2). The application's gas flow conditions were supplied to the cone meter manufacturer who subsequently supplied a meter with an in-house water flow calibration. In service at the highest gas flow rates it was noticed that the cone meter was significantly under-reading the gas mass flow rate. The cone meter was therefore calibrated by the facility with gas flow.

Fig. 3 shows the manufacturer supplied water calibration data in terms of discharge coefficient (C_d) vs. velocity (or volume flow rate). For a constant cross sectional area the velocity and volume flow rate are interchangeable parameters. This flow range corresponded to a water flow Reynolds number range of approximately 50,000 to 415,000. This water flow range was high enough that the velocity profile would have been always turbulent. A constant C_d fitted the water calibration data to 0.25% uncertainty. Fig. 3 also shows the subsequent gas flow calibration data. This corresponded to a gas flow Reynolds number range of approximately 700,000 to 3,500,000. The C_d vs. velocity plot clearly shows that the different fluids produce distinctly different C_d values at the same flow velocity. This is a clear indication that velocity (or flow rate) is an incorrect correlating parameter.

Fig 4 shows the water and gas data plotted as C_d vs Reynolds number. The relative uncertainty bars for the water and gas reference flow data is also shown. The application (and corresponding gas calibration data) has a maximum Reynolds number about eight times larger than achieved by the manufacturers water calibration facility. It is clear that extrapolating the water calibration data does not give the correct C_d at higher Reynolds numbers. At that maximum Reynolds number the discharge coefficient differs from the water calibration data by nearly 2%. Where the water and gas Reynolds number test ranges overlap the results were similar. That is, there is

Fig 2. Water Flow calibrated Cone DP Meter Installed in a Gas Pipeline.

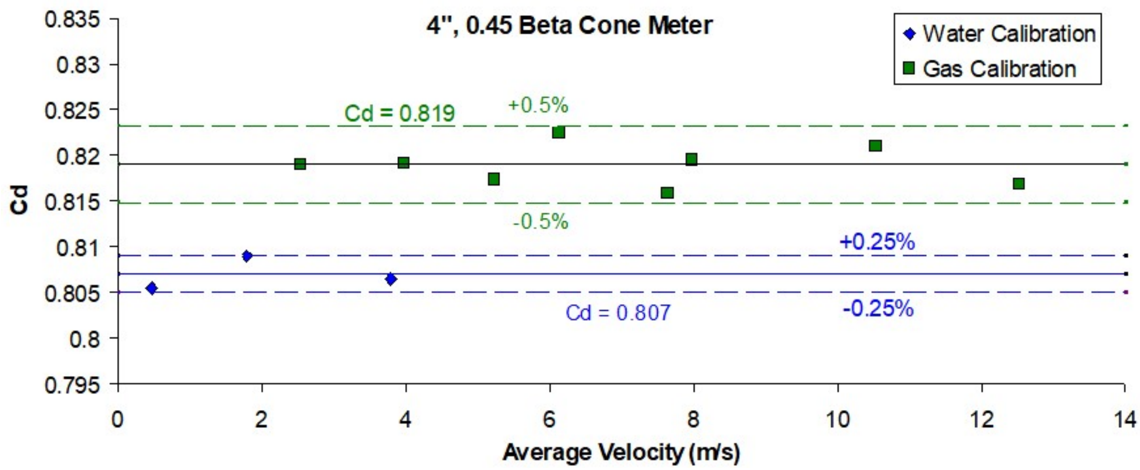

Fig 3. Cone Meter Water and Gas Data Plotted in Terms of Velocity.

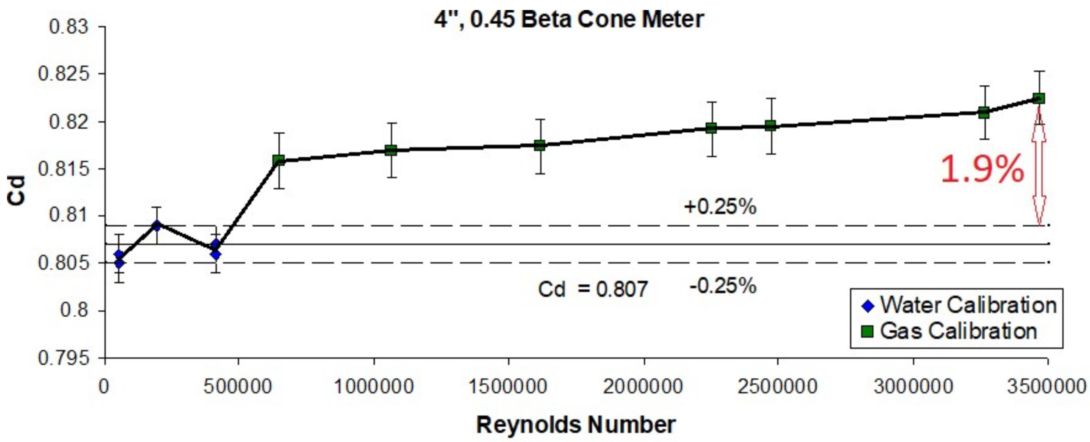

Fig 4. Comparison of Cone Meter Water and Gas Calibration Data, C_d vs. Reynolds Number.

nothing wrong with the water calibration data *within its Reynolds number range*, but it is just not possible to extrapolate the result to much higher Reynolds numbers and expect to maintain no bias. Note that as all the water and gas data are well inside the turbulent flow region the performance shift is due to changes in the turbulent velocity profile (and possibly other influences) and not transitional flow. Taking the cone meter water and gas calibration data

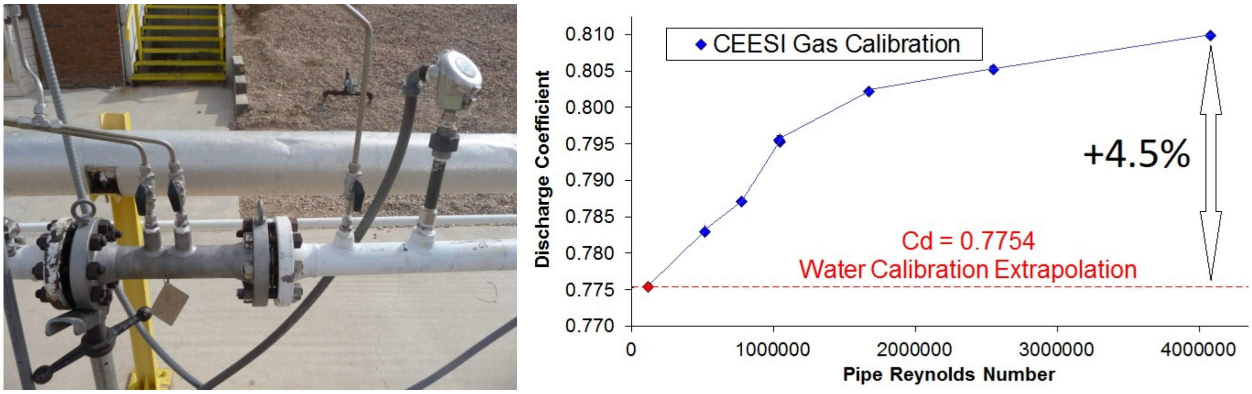

Fig 5. 2”, 0.75β Cone Meter Water and Gas Calibration Data

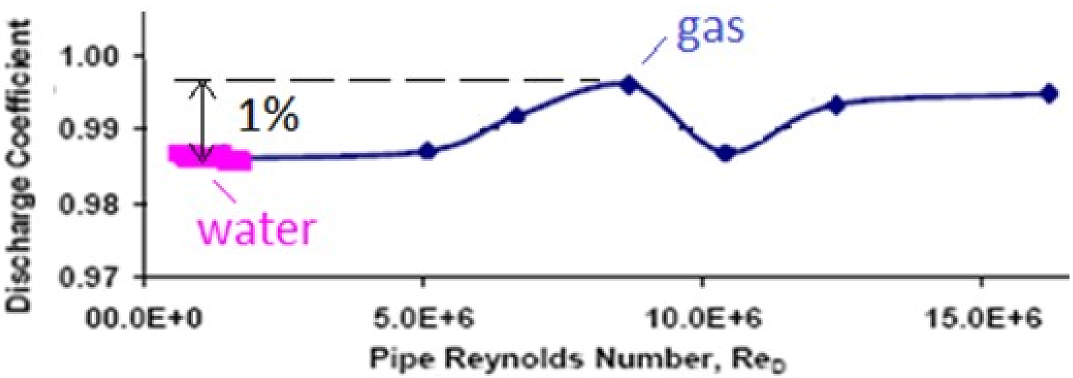

Fig 6. ConocoPhillips 8” Venturi Meter Water and Gas Calibration Data

together Fig 4 shows that a piece wise linear interpolation data fit can easily combine the two calibration data sets with low uncertainty.

The 2% shift shown in Fig 4 is just for that particular example. Other individual meters (of various designs) can have greater or less shifts in case by case basis. For example, Fig 5 shows a 2”, 0.75β cone meter with water and gas calibration data. Extrapolating the manufacturer’s in-house water calibration with a maximum Reynolds number of 100,000 to 4 million Reynolds number gas flow caused a bias > 4%. Fig 6 reproduces ConocoPhillips data (see Geach et al [1]) for an 8” Venturi meter. Here extrapolating the water data to higher gas flow Reynolds numbers caused a maximum bias of approximately 1%.

These water vs. gas calibration examples only show Reynold number ranges that create turbulent flow. This issue becomes accentuated if we consider water vs. oil or different grade of oil calibrations. The significant difference in water and various oil viscosities coupled with the lower flow velocities of liquid flows means such Reynolds number ranges often straddle laminar, transition, and turbulent velocity profiles. This makes the issue of using Reynolds number instead of flow rate or velocity as the correlating parameter all the more acute. Several examples of this will now be shown.

3b. Calibrations Using Different Liquid Viscosities

Figs 7 and 9 show an 8” cone meter and a 4” Venturi meter at the CEESI and TUVNEL oil facility respectively. Both meters were calibrated by using different oils of significantly different viscosity across the same flow rate/ velocity range. Figs 8 & 10 show both meter’s calibration

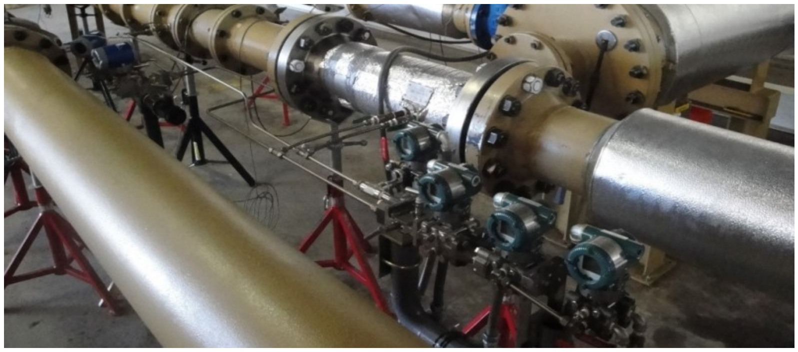

Fig 7. 8”, 0.75β Cone Meter at the CEESI Oil Facility.

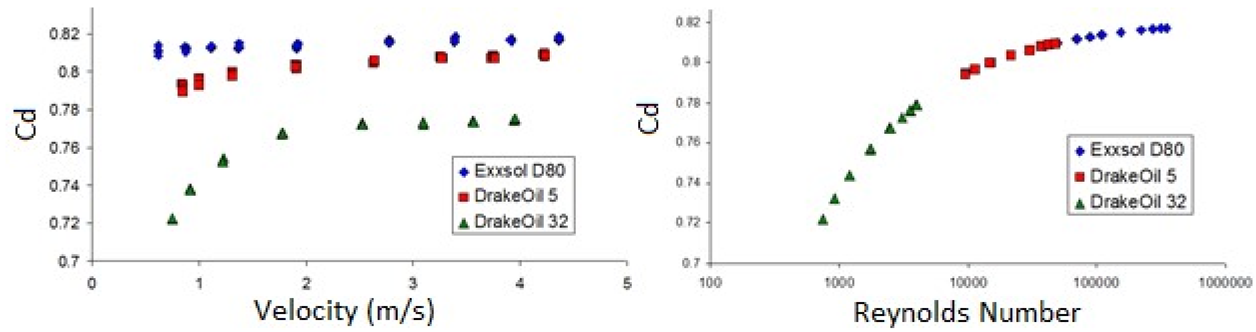

Fig 8. 8”, 0.75β Cone Meter Data with Different Viscosity Oils, C_d vs. Velocity & Reynolds No.

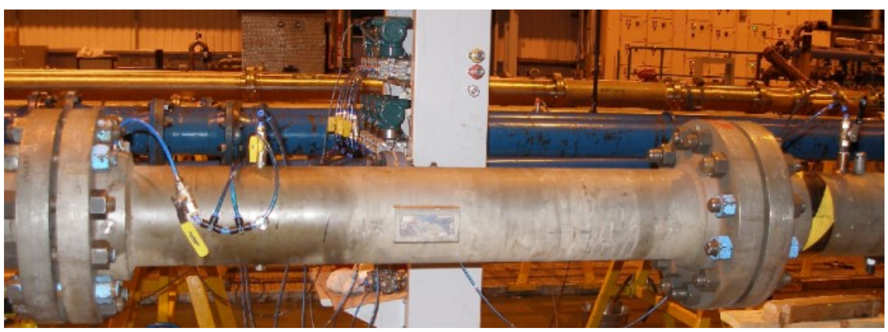

Fig 9. 4” Venturi Meter at the TUVNEL Oil Facility.

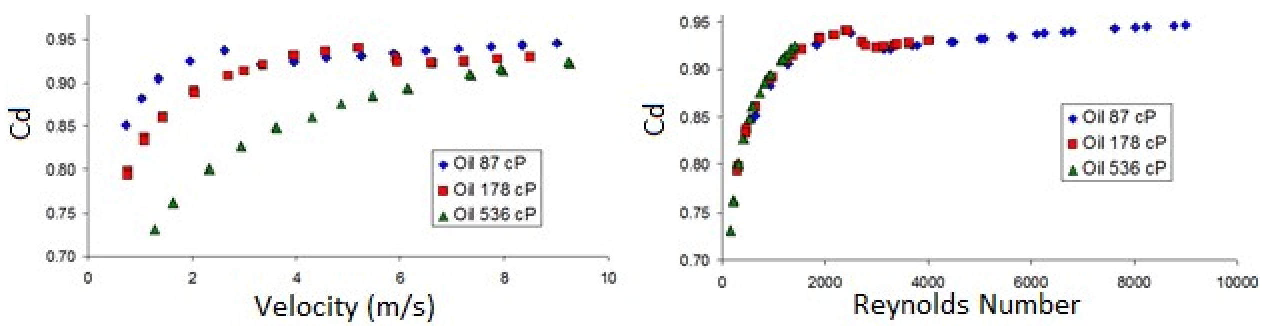

Fig 10. 4”, Venturi Meter Data with Different Viscosity Oils, C_d vs. Velocity & Reynolds No.

results when using C_d vs. velocity (or volume flow rate) and C_d vs. Reynolds number. Clearly, velocity is not an appropriate correlating parameter. For both DP meters a given velocity (i.e. volume flow rate) can produce three distinctly different discharge coefficients. That is, if either of these meters are calibrated with one oil viscosity, and the C_d correlated to velocity, the meter’s flow rate prediction will have a significant bias if the meter is subsequently used with oils of a significantly different viscosity. In contrast, the C_d vs. Reynolds number results are quite remarkable. In each case the separate C_d vs. velocity curves that looked unrelated now look like one curve, fitting together like parts of a jigsaw. This DP meter data clearly highlights the importance of calibrating DP meter C_d vs. Reynolds number.

Fig. 11. 6" Liquid Turbines at the CEESI Oil Facility.

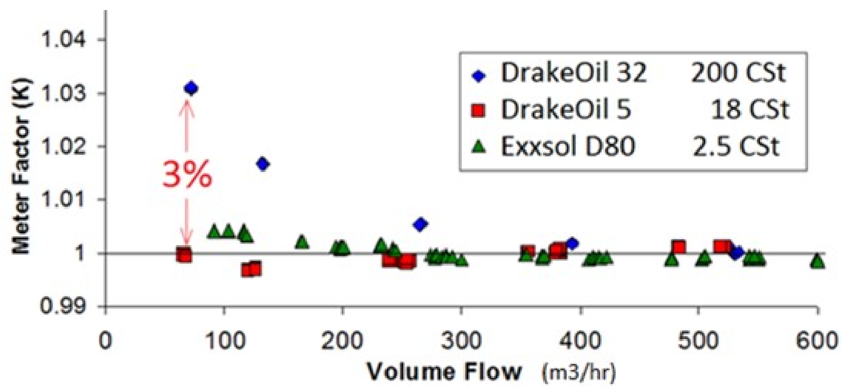

Fig 12. 6" Liquid Turbine Meter Oil Calibration Data, K vs. Velocity.

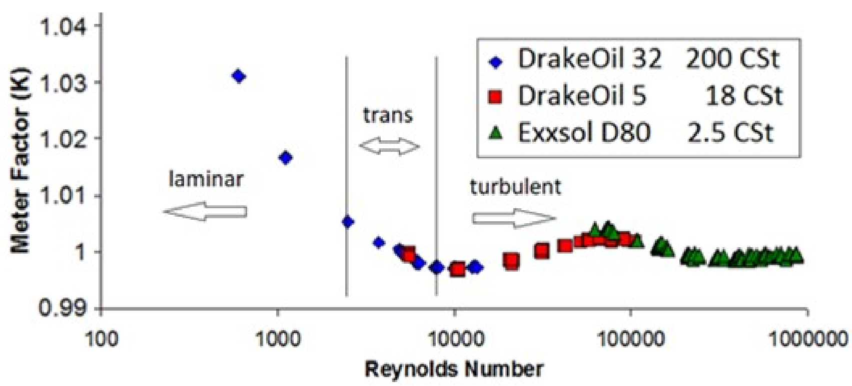

Fig 13. 6" Liquid Turbine Meter Oil Calibration Data, K vs. Reynolds Number

This issue is not restricted to DP meters. Many meter designs performances are influenced by viscosity, i.e. Reynolds number. The following are examples for liquid flow turbine, ultrasonic and Coriolis meters. Fig 11 shows 6" liquid turbine meters at the CEESI oil facility. Fig 12 shows turbine meter factor (K) vs. velocity data for three different viscosity oil. At low velocity different oils give different meter factors. In this case velocity is not a good correlating parameter. Figure 13 shows the same data plotted as K vs. Reynolds number. Again, not only

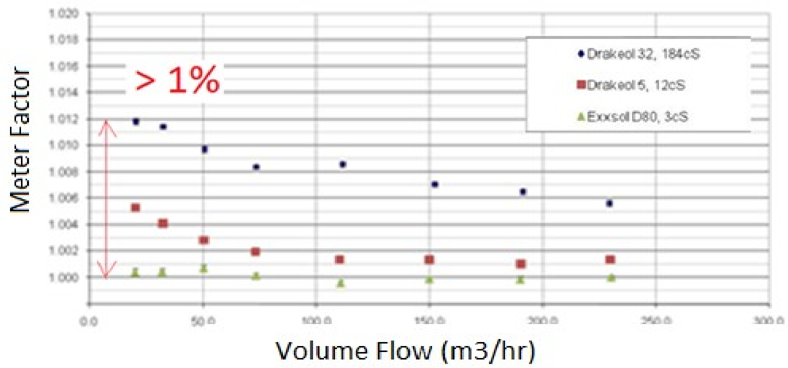

Fig 14. 6” Liquid Coriolis Meter Oil Calibration Data, Velocity vs. Meter Factor

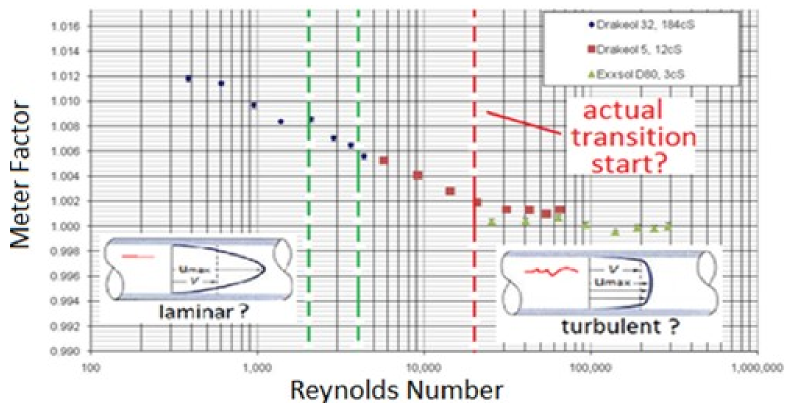

Fig 15. 6” Liquid Coriolis Meter Oil Calibration Data, Reynolds Number vs. Meter Factor

does the data now look like one curve, the three separate K vs. velocity curves which looked unrelated, now fit together like parts of a jigsaw. Furthermore, when considering gas turbine meters AGA 7 (2007) say in its Section 6.3.2 “… the expected operating Reynolds number range and / or density for a meter needs to be taken into account when designing a calibration program”.

Fig 14 shows blinded data from an 8” liquid Coriolis meter calibrated at a 3rd party oil test facility. Again, three different viscosity oil calibration data sets are shown where the correlating parameter was volume flow rate. At lower flows different oils give different meter factors. That

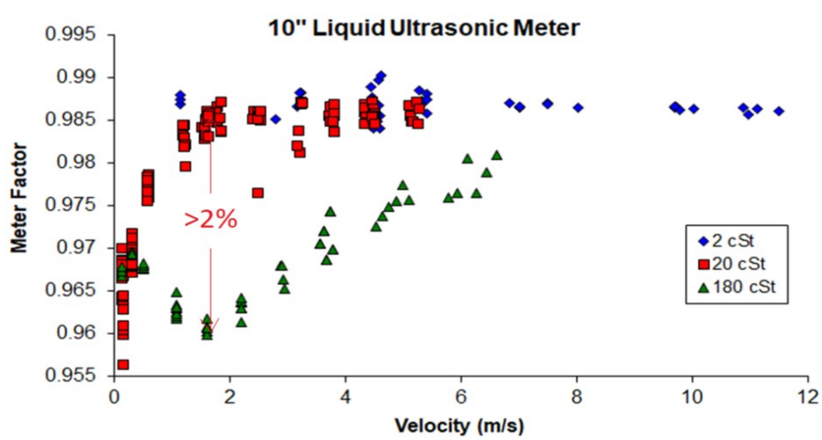

Fig 16. 10” Liquid Ultrasonic Meter Oil Calibration Data, Velocity vs. Meter Factor

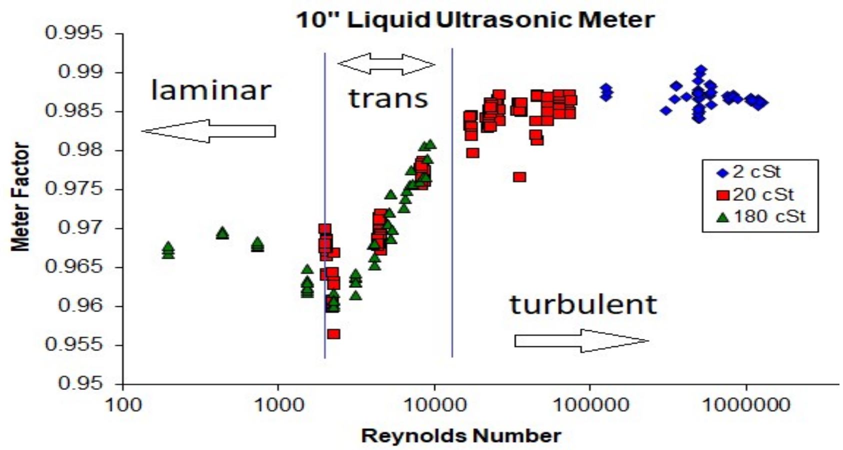

Fig 17. 10" Liquid Ultrasonic Meter Oil Calibration Data, Reynolds Number vs. Meter Factor

is, flow rate is not a good correlating parameter. Fig 15 shows the same data plotted to Reynolds number. Again, the three separate curves created by plotting volume flow rate vs. meter factor which looked unrelated, are now seen to fit together well. Other 3rd parties have reported the same Coriolis meter phenomenon, e.g. NEL (Mills [2]).

Fig 16 shows a blinded 10" liquid ultrasonic meter data set from the CEESI oil facility. Again, three different viscosity oil data sets are shown where the correlating parameter was velocity. The high viscosity oil gives a different meter factor than the two lower viscosity oils. Again, velocity is not a good correlating parameter. Fig 17 shows the same ultrasonic meter data plotted to Reynolds number. Yet again, using Reynolds number makes the data fit one continuous curve.

Summary on Reynolds Number vs. Flow Rate:

"*For a successful technology, reality must take precedence over public relations, for nature cannot be fooled.*" Richard Feynman

The authors are not suggesting that Reynolds number is guaranteed to be the single correlating parameter for all flow meter designs in all conditions. E.g., at low flow rates turbine meters are influenced by bearing friction issues, and it is arguable other correlating parameters should be included. There are meter manufacturers that claim their meter operates better when calibrated to the velocity or flow rate and not Reynolds number. Some give various counter arguments to why they believe velocity is better than Reynolds number for their meter. However, for many meters in many flow conditions there is significant theoretical and calibration evidence that Reynolds number is the most appropriate correlating parameter.

It is not by coincidence that the international standard for the orifice meter (ISO 5167-2) expresses the C_d as a function of Reynolds number. ISO 5167-5, the cone meter standard, states when calibrating a cone meter it should be calibrated over its entire Reynolds number range of operation. The Venturi meter standard ISO 5167-4, is more ambiguous stating a C_d over a given Reynolds number range, not flow rate, but stating, "… for optimum accuracy Venturi tubes for use in gas should be calibrated over the required flowrate range".

Although ISO does not explicitly state that Reynolds number should be the correlating factor for a vortex meter, they do imply this by saying "… K-factor may be presented as a function of either the pipe Reynolds number or flow rate at specific set thermodynamic conditions."

However, for a given homogenous fluid, setting the thermodynamic conditions means setting the density and viscosity and therefore the Reynolds number range. ASME MFC-6-2013 shows turbine meters being calibrated to Reynolds number.

Many oil meters (e.g. turbine meters) are periodically re-calibrated (i.e. 'proved') in the field. Provers offer accurate volume flow reference data. However, a 'prove' is carried out at whatever the liquid viscosity happens to be. The temperature and viscosity of oil can change significantly between proves. Hence, proves over a constant volume flow range at different oil viscosities constitutes proves over different Reynolds number ranges. Although API do not explicitly state that the Reynolds number should be the correlating factor, they do imply this by saying a meter should be reproved (over the same flow rate) whenever there is a change of fluid viscosity.

It is an unfortunate accident of nature that the typical laminar, transition, and turbulent flow profile Reynolds number range falls in the typical flow ranges of many production oil flows. The associated complexity of the linearization techniques can potentially produce increases in uncertainty or metering biases. This complexity in flow metering overlaps the typical oil flow ranges where industry demands extremely low metering uncertainty. Meter manufacturers are under commercial pressure to be seen to meet the tight specification demands. As such, for meters with embedded processors, the temptation to add extra non-transparent linearization techniques within the processor's inaccessible code is no doubt considerable.

A flow meter performance influenced by Reynolds number is in practice very inconvenient. It potentially makes flow meter calibration more expensive and time consuming. It also means that the end user is required to know the fluid viscosity in the field. This is not trivial. With gas composition and temperature known it is relatively straight forward to predict gas viscosity to a low uncertainty, although it is still another burden on the operator. It can be significantly more problematic to predict oil viscosities in the field to low uncertainty. Being required to supply a low uncertainty viscosity value makes operating the flow meter more complex. Nevertheless, physical law cannot be ignored simply for the convenience of manufacturer / end user relations. As Richard Feynman [3] told NASA in the Challenger Space Shuttle disaster report "…for a successful technology, reality must take precedence over public relations, for nature cannot be fooled." Inconvenient as it may be, the reality for many flow meter end users is that their meters are best correlated to Reynolds number, and not velocity or flow rate. In many cases:

Calibration of a flow meter using a velocity or flowrate range that does not coincide with the flow meter applications entire Reynolds number range can potentially result in an inappropriate linearization of the meter and an associated flow rate prediction bias.

We have now discussed ***what*** should be data fitted. It is now time to discuss ***how*** we are fitting data, and ***why*** are we choosing particular methods. Although the authors advocates the use of Reynolds number the following discussion holds for any correlating parameter one chooses.

4. What Constitutes an Appropriate Linearization Technique?

"I apologize for lying to you. I promise I won't deceive you except in matters of this sort."
Spiro Agnew, Vice-President of the United States 1969-73.

The authors do not have any pretensions that they are particularly skilled in the mathematical techniques of 'regression analysis', i.e. 'curve fitting', compared to many highly qualified and experienced engineers. Nor do the authors intend to delve into the finer mathematical details of curve fitting. That is the subject for a dedicated long mathematical text, and is not only not possible in a single paper, but not required here. The core point the authors want to make is that without set rules there is opportunity for the curve fitter to adopt a pick and mix attitude to the curve fit methodology. This gives 'wiggle room' that allows the curve fitter, if he deems it necessary, to manipulate the fit such that the perceived meter performance tends to whatever performance is claimed without breaking any rules or accepted norms. However, in such cases the real performance in service may not be as perceived from the calibration report. Industry should pay more attention to the non-trivial issue of meter linearization if the flow meter uncertainty claims are to have the integrity expected by the hydrocarbon production industry.

Curve fitting (i.e. linearization) can be carried out in an end user mainframe computer, flow computers, or in the embedded processors integral to some flow meter system packages. Mainframe computers and programmable flow computers can be programmed to apply any linearization technique the end user deems suitable. Also, although not always fully transparent, meter system manufacturers can program embedded processors with any linearization technique they deem suitable. Hence, it is valid to discuss general data fitting techniques here, as any data fitting technique, i.e. any correlating equation form, could be applied.

Flow computers tend to only offer piece wise linear interpolation (i.e. a 'look up table') as a pre-programmed linearization methodology. Many end users of flow computers do not utilize the programmable capability, and default to this method. Also, some ultrasonic meter manufacturers use piece wise linear interpolation in their embedded processors. Two of the authors have historically been verbal opponents of such practice, arguing that it produces an artificially good result (i.e. a seemingly perfect coefficient of determination of $R^2 = 1$, and no apparent data fit uncertainty) more to do with show than reality. However, the closer one looks, for all it has flaws, the more benefits piece wise linear interpolation seems to have relative to the alternatives.

4a. Fully-Empirical Fitting, Semi-Empirical Fitting, and Modeling

If the governing physical law of some phenomenon is completely understood then it is possible to create a mathematical expression (or 'model') to describe it. Such a model is theoretical, and if it is correct then the model will describe high quality (i.e. low uncertainty) data accurately. This is the procedure that forms most theoretical flow meter flow rate calculations. However, flow meters are calibrated because in virtually all cases the theoretical understanding is not complete. There are inevitably some secondary influences that are not fully understood. A meter factor is required to correct for these secondary influences, and as they are not fully theoretically predictable, this meter factor cannot be expressed by a theoretical mathematical model. Hence, all meter factor calibration curve fits are somewhat empirical.

There are two forms of empirical data fits, i.e. fully and semi empirical fits. Fully empirical (aka 'blind') data fits express no understanding of the underlying physical laws. They are no more than mathematical expressions that reproduce the data set (and facilitate interpolation between the points). Semi-empirical data fits express some limited (but incomplete) understanding of the underlying physical laws governing these secondary influences. Semi-empirical data fits are a blend of theoretical modeling and empirical data fitting They are therefore fundamentally more robust for interpolation and extrapolation than fully empirical data fits. However, unfortunately most flow meter calibrations are carried out because the theoretical understanding has been exhausted. Hence, a largely unspoken reality is most calibration data fits are fully empirical.

There are a few exceptions. For example, the use of Reynolds number instead of flow rate or velocity to correlate some meter factors is due to the theoretical considerations of the influence of viscosity. Another example is, due to further theoretical considerations beyond that used for the DP meter general equation, Miller [4] shows that based on the work of Murdock, for turbulent flow the form of a DP meter's discharge coefficient calibration fit could be:

$$C_d = C_{d\infty} + \frac{b}{\mathrm{Re}_D^n} \tag{2}$$

where $C_{d\infty}$ is the discharge coefficient at infinite pipe Reynolds number (Re_D), and 'n' and 'b' are data fit values. Murdock modified theoretical work of Blasius to suggest n=1/5 was a suitable value. Nevertheless, most flow meter calibration curve fits are fully empirical. In the following text all curve fit discussions are for fully empirical fits.

4b. Curve Fitting is More Challenging than Many Presuppose

Curve fitting is often thought of as a trivial exercise, but it can be problematic. For example, take the hypothetical gas ultrasonic meter data set example offered by AGA 9 Ed 1 (shown as blue diamond points in Fig 18). This standard has been recently updated and this hypothetical data changes, but this doesn't matter, we are using this data set as an example only.

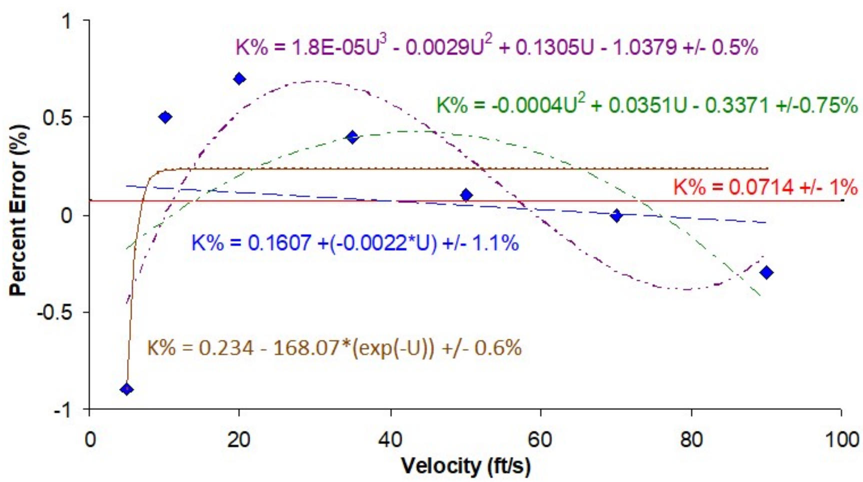

Fig 18. Hypothetical Gas Ultrasonic Meter Data with Typical Curve Fits

Superimposed on the data plot are five sample curve fit equations. There is the averaged constant value, linear line, 2nd & 3rd order polynomials, and the 'best fit' proposal from a commercial data fitting software. Although to the eye the data looks like a simple curve none of the equations fit the data particularly well. There is nothing wrong with this hypothetical data set, it is a realistic response for such a meter. This is a common problem for all flow meter calibrations. The data may look like a smooth curve but it may still be problematic to find a suitable simple mathematical expression to describe it. Furthermore, if every calibration data set required a different equation form then each meter's calibration software would then be bespoke, which would be very time consuming and inconvenient to the various meter manufacturers. This is one reason many in industry lean towards piece wise linear interpolation (aka as a look up table).

This technique is all but guaranteed to fit any and all data sets. This alleviates the problem of finding a suitable curve fit per meter. It is a method of standardizing the calibration data fit procedure. A piece wise linear interpolation fit of this hypothetical data is discussed in Section 5.

4c. Best Coefficient of Determination or Maximum Flow Rate Uncertainty?

"*We want certainty to our uncertainty*" Phil Robbins, Petronas Consultant

There is no single 'best' or 'most appropriate' way to carry out a generic data fit. What constitutes the 'best fit' is subjective. Academics, researchers, and students writing theses tend to default to using the coefficient of determination (i.e. 'R^2') as the best method for assessing the quality of a curve fit. Equation 3 shows the R^2 calculation,

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_i^n \{K_i - f(\mathrm{Re}_i)\}^2}{\sum_i^n \{K_i - K_{av}\}^2} \tag{3}$$

where, for a calibration data set of 'n' calibration points of K vs. Re (or alternative flow rate parameter):

SS_{res} is the sum of squares of residuals
SS_{tot} is the total sum of squares
K_i is the individual meter factor of each calibration point
$f(Re_i)$ is the curve fit prediction of the meter factor of each point
K_{av} is the average meter factor of the calibration data set

The coefficient of determination will generally be bound by $0 \leq R^2 \leq 1$. An R^2 between 0 and 1 indicates the extent to which the meter factor is predictable. An R^2 of zero indicates that the meter factor cannot be predicted by the Reynolds number (or other flow rate parameter) fit, and an R^2 of one indicates that the calibration data meter factors are *perfectly* predicted by the fit.

The commonly used (but limited capability) curve fitting function in Microsoft Excel™ and more sophisticated commercial curve fitting software (such as TableCurve™) use R^2 as the default quality indicator of the suggested curve fit/s. However, the hydrocarbon production industry does not automatically use R^2 as the indicator of the best fit. And indeed they are free to choose any linearization technique because…

As far as the authors are aware there are no rules set by standards, regulators or contracts as to what constitutes an industry accepted linearization / curve fitting method.

The hydrocarbon production industry is generally more interested in finding a flow meter linearization technique that meets the end users practical requirements than following what academia does. This requirement tends to be that across a stated flow range the flow rate prediction uncertainty will be guaranteed to some specified 'x%' uncertainty to 95% confidence. That is what Phil Robbins means by "we want certainty to our uncertainty". Using the coefficient of determination (R^2) does not guarantee this. The coefficient of determination gives a macro view of the overall quality of the curve fit across the whole data set. It considers the data set as a whole, not the individual points. In effect it can 'sacrifice' a point (or for large data sets a few points) in the interest of getting the best overall fit. This can result in the highest ('best') R^2 fit failing to meet the end user requirements while a different fit with a lower R^2 can comply. The following random example highlights the point.

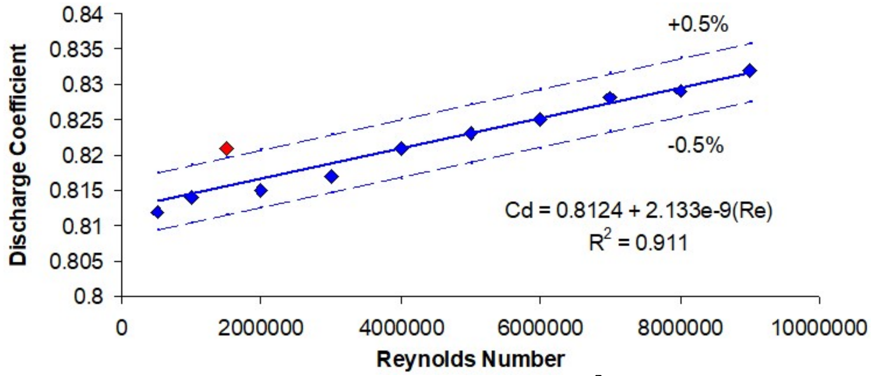

Fig 19. 8" Cone Meter Calibration Best R^2 Linear Data Fit.

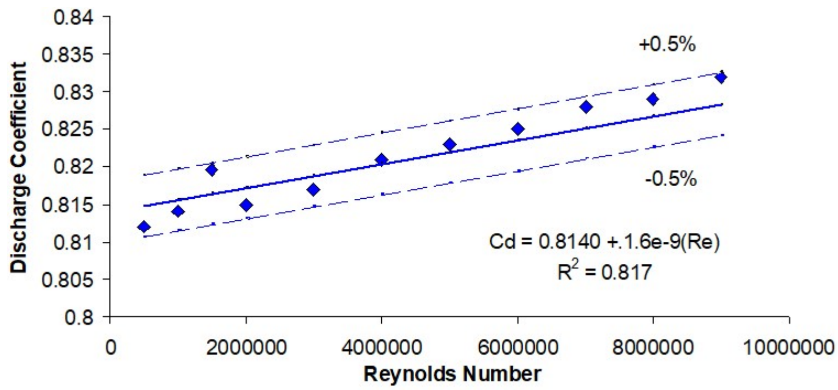

Fig 20. 8" Cone Meter Calibration 0.5% Uncertainty Liner Data Fit.

Fig 19 shows a random 8", 0.55β cone meter eleven point gas calibration data set of C_d vs. *Re*. The type of meter, the particular meter factor, the required uncertainty, and the form of the data fit are irrelevant to the general curve fitting point being made here. The data shows that one point looks like an outlier. However, this was not a 'bad point'. This example is not describing any error by the calibration facility. Repeat tests gave the same result. Such anomalies in data sets are relatively common in various flow meter results. They can be caused by various physical phenomena and they cannot be ignored because they are inconvenient. They will occur in the field as well as the calibration facility and need to be accounted for.

In this example, say the cone meter is required to have a C_d predictable to 0.5% uncertainty at 95% confidence. As way of example the Microsoft Excel linear fit option is selected. The resulting fit offered had a R^2 value of 0.911. Fig 19 shows that over the data set this curve fit is indeed very close to the majority of the points. For ten of the points this fit agrees with the data to < 0.25%. However, there is the single point that disagrees with this fit to > 0.5%. This point has been 'sacrificed' by the R^2 method in order to achieve the best fit of the majority of the data. But, this means that *by choosing the highest (best) R^2 fit the 0.5% uncertainty to 95% confidence requirement has not been met*. Fig 20 shows an alternative fit on the same data. The new fit has a significantly reduced R^2 value of 0.817, but now meets the 0.5% uncertainty stipulation across all calibration data points. This is done at the expense of the average fit quality. By fitting the outlying point within the 0.5% band the other ten points now have a fit that agrees with the data to only < 0.45%. But in many cases industry does not seem to care. It is more important to meet the performance required by the application across the entire Reynolds number range than to have on average the best overall fit with an outlier or two.

There is no right or wrong fit. How anybody fits calibration data depends on what they wish to achieve. It is a subjective exercise. However, there is a cost to choosing the 0.5% uncertainty guaranteed fit over the best R^2 fit. Most meters are calibrated across a range where the maximum calibration range is slightly in excess of the expected maximum in the field. In this case the maximum Reynolds number the meter is to encounter in the field can reasonably be expected to be about 7e6. For a natural gas flow at 50 Bar, 25°C, and 50 MMSCFD, for gas priced at $2.90 per million BTU, the corresponding 7e6 Reynolds number flow has a value per day of approximately $143.5K. If the 0.5% uncertainty guaranteed fit is chosen instead of the best R^2 value then the corresponding predicted discharge coefficient at normal meter flow conditions drops from 0.827 to 0.825, i.e. a -0.26% shift. This means the meter will predict $366 less flow per day, i.e. $133.6K less flow per annum, or for a meter life span of ten year, $1.34 million less flow. This is not due to meter performance per se, but by the choice of linearization technique.

A requirement to linearize the meter such that some percentage uncertainty guarantee is met across the entire applications flow range can inadvertently cause an increase in metering bias at the typical flow conditions in which the meter usually operates.

4d. Loading the Sweet Spot

Statistical tools are blind to physical laws. To achieve x% uncertainty at 95% confidence you only need nineteen out of every twenty points to meet the stipulated x% uncertainty performance. Therefore, if one point out of twenty is 'an outlier', even if it represents a real reproducible physical phenomenon at those flow conditions, the meter's curve fit can still be said to meet the expected performance, even though it really doesn't. This can and does happen by accident, although not many flow meters have a twenty or more point calibration.

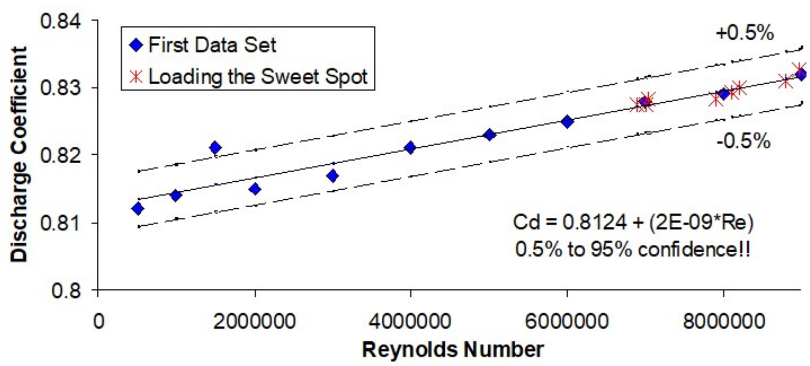

Fig 21. A Random Meter Calibration Result with Extra Sweet Spot Data Points.

End user demands for 'x%' uncertainty at 95% confidence are not always very detailed. This can give scope for interpretation when linearizing a flow meter. With no curve fitting rules the practice of deliberately loading a 'sweet spot' can allow a problem meter to 'pass' a calibration. As a random example let us continue with the example in Section 4c, although again, the same scenario can be carried out on any meter type. The problem in Fig 19 is that there are eleven points and one is outside the stipulated 0.5% uncertainty, i.e. only 91% of the data falls within the required < 0.5%. The outlier is not a bad point, it is caused by some unspecified physical cause. So, rather than fail the meter, without any stated linearization rules the manufacturer could hypothetically decide to test the meter at nine more flow rate points that are in the 'sweet spot', i.e. where the initial data indicates that the meter can be reasonably expected to 'behave itself'. As the new nine points give similar results to the previous calibration points in that same flow rate region the manufacturer could then claim that 19/20 of the calibration data now fits the 0.5% uncertainty to 95% confidence. Fig 21 shows the same data as Fig 19, but with extra sweet spot data added. Now the original best R^2 data fit as shown in Fig 19 can be used and it would meet the 0.5% uncertainty at 95% confidence requirement. This wouldn't necessarily tell the whole story, as at a low Reynolds number the meter will again have the outlier. But without linearization rules forbidding it, this practice is technically allowable.

Whether such a practice is acceptable is a decision for the meter manufacturers and their end user clients. However, it is interesting to note that the aim of such end user stipulated requirements is to attempt to assure good measurement. But Section 4c showed that such a stipulation can sometimes inadvertently create the opposite effect in practice. In this example by loading the data set with sweet spot data, and thereby choosing the highest R^2 linear fit, the likely bias discussed in Section 4c is removed. This is another example of how flow meter linearization techniques are not always objective but actually rather subjective.

4e. Issues with Polynomials (and Other) Curve Fits

Fig 22 shows the AGA 9 Ed 1 hypothetical ultrasonic meter data set with an Excel fitted 3rd order polynomial curve. For all Excel is not considered a commercial curve fitting software package it is undoubtedly one of the most commonly used generic data fitting software. The "Excel Data Fit Function Statement" shown in Fig 22 is exactly as it is produced by Excel. *It is wrong***.** The drawn curve appears correct, but the corresponding stated equation has too few significant figures. The stated equation actually has a R^2 value of only 0.114, not the claimed 0.614. To achieve an R^2 value of 0.614 you need to use more significant figures.

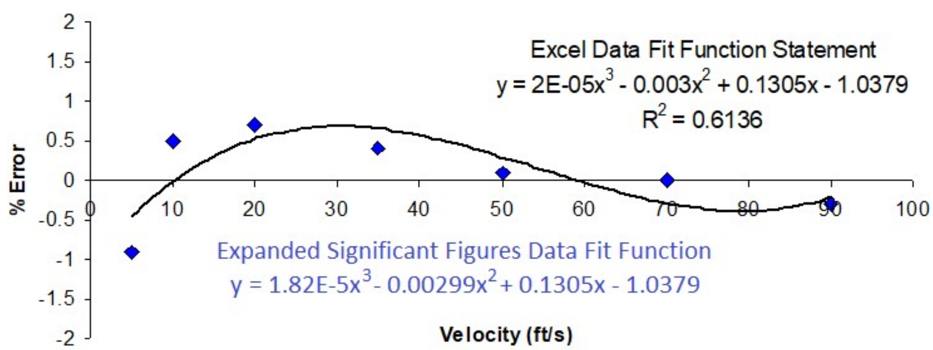

Fig 22. Hypothetical Ultrasonic Meter Calibration Data Set With 3rd Order Polynomial.

A polynomial fit, and various other curve fit equations, are very sensitive to the number of significant figures included in the constants. The Excel software has defaulted to showing one significant figure in the constant associated with the third power term. This stated equation produces a relatively poor result. Table 1 shows the performance of three equations, Equation 1 being the stated equation by Excel, Equation 2 using two significant figures in the third power term, and Equation 3 showing more significant figures in the second and third power terms. The number of significant figures clearly affects the result. As the number of significant figures

Calibration Velocity (ft/s)	Calibration Found % Error	Equation 1 Prediction % Error	Equation 2 Prediction % Error	Equation 3 Prediction % Error
90	-0.3	0.99	0.17	-0.06
70	0.0	0.26	0.43	0.31
50	0.1	0.49	-0.14	-0.19
35	0.4	0.71	-0.23	-0.25
20	0.7	0.53	0.18	0.18
10	0.5	-0.01	0.51	0.51
5	-0.9	-0.46	-0.44	-0.44

Table 1. Relative Effects of Significant Figures in Polynomial Fits

Equation 1: $y = 2E\text{-}5x^3 - 0.003x^2 + 0.1305x - 1.0379$ $R^2 = 0.114$
Equation 2: $y = 1.8E\text{-}5x^3 - 0.003x^2 + 0.1305x - 1.0379$ $R^2 = 0.580$
Equation 3: $y = 1.82E\text{-}5x^3 - 0.00299x^2 + 0.1305x - 1.0379$ $R^2 = 0.614$

increases the curve fits performance improves. It takes Equation 3 to get the performance (i.e. R^2 value) claimed by Excel. Use of the stated Equation 1 would have caused on average higher flow rate prediction biases. There are many cases where such an induced bias is financially significant but subtle and may well go unnoticed. A lesson from this is it is not wise to blindly accept a curve fit equation without first double checking it by an independent means.

4e.1 The Effect of Using Different Number of Significant Figures

As polynomial (and other curve) fits are sensitive to the number of significant figures used then there is potential for flow rate prediction biases to be introduced by the choice of number of significant figures. For example, let's assume in this hypothetical example that Fig 22 represents a calibration of a 16", sch 80 ultrasonic meter that is to be linearized using a 3rd order polynomial. If the meter is being used at say a pressure of 50 bar and 27 m/s (90 ft/s), this corresponds to 453 MMSCFD, i.e. approximately \$1.3 million per day. Table 1 shows that depending on if Equation 2 or Equation 3 is chosen the curve fit will predict +0.17% or -0.06% bias respectively. This is a difference (i.e. bias) of 0.23%, i.e. approximately \$1.1 million per annum, or \$7.7 million per a seven year recalibration cycle.

A little spoken fact about all non- piece wise linear interpolation curve fitting is once a curve fit is selected and implemented into the relevant computer it in effect replaces the actual data set. Although it is in reality an imperfect expression of the actual calibration data it is treated as the data set. Whereas the actual calibration data will be logged in a report, the operation of the meter uses the assigned curve fit and does not reference the original data.

Do you use equation 2 and predict 0.23% more gas flow than equation 3, or equation 3 and predict 0.23% less gas flow than equation 2? This is not an increase in uncertainty. This is a guaranteed curve fit induced bias. If equation 2 is chosen then the meter will predict more flow and the seller wins. If equation 3 is chosen then the meter will predict less flow and the buyer wins. Yet, the authors are unaware of any formal rules regulating this issue.

4e.2 The Effect of Choosing Different Order Polynomials (or Different Generic Fits)

Fig 23 shows real calibration data from a randomly selected and blinded data set from a 12" ultrasonic meter. This meter meets the AGA9 'as found' performance limits. It could be said to

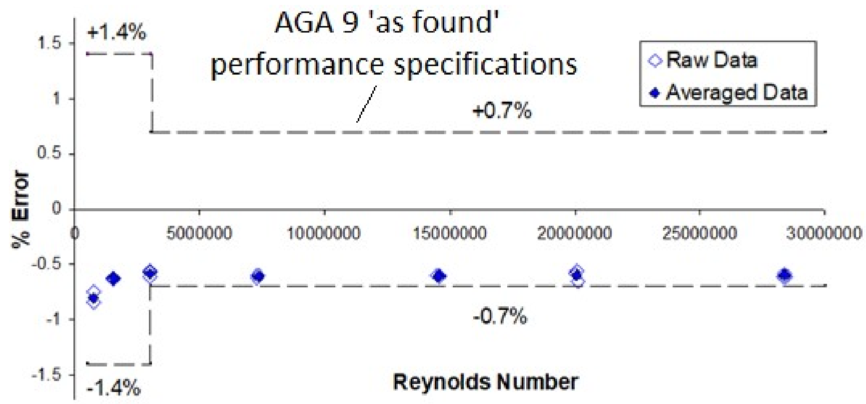

Fig 23. Random Selected Blinded 12” Ultrasonic Meter Calibration Data

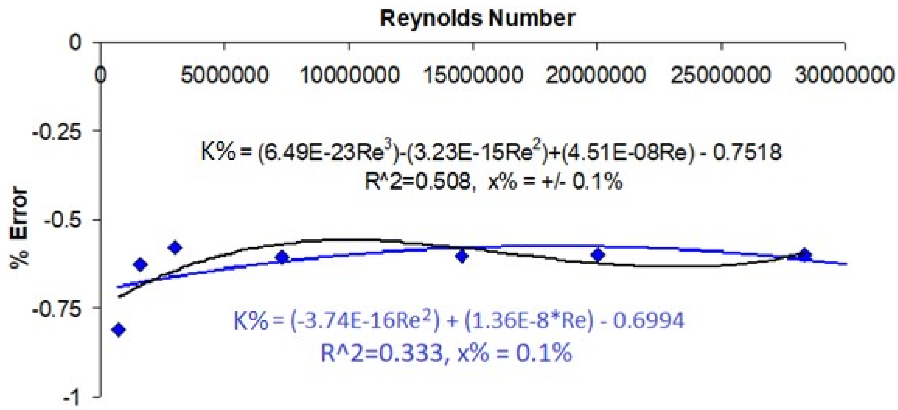

Fig 24. Second & Third Order Polynomial Curve Fits on 12” Ultrasonic Meter Calibration Data.

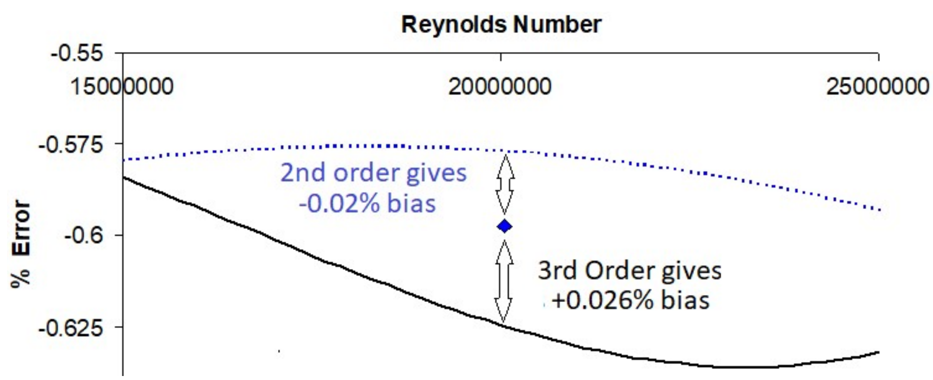

Fig 25. Effect of Second & Third Order Polynomial Curve Fits on 12” Ultrasonic Meter Calibration Data.at Applications Expected Flow Range.

be rather ‘well behaved’ meter, i.e. the percentage error is rather constant across much of the flow range. The data is expressed as meter percentage error (K%) vs. Reynolds number, although the same point could be made if the data was shown as K% vs. flow rate or velocity.

AGA 9 and ISO allow polynomial curve fitting as an option. There is no stipulation on what order of polynomial should be used. ISO does not discuss this. AGA 9 suggests that ‘normally a 2nd order polynomial is used”, but there is no restriction on using higher orders. 3rd order polynomials are relatively common. Fig 24 shows a 2nd & 3rd order polynomial fit. The 2nd order polynomial produces a R^2 of 0.333 and a 0.1% uncertainty. The 3rd order polynomial produces a R^2 of 0.507 and again a 0.1% uncertainty. There isn't much difference, but the 3rd order polynomial can be said to offer a *marginally* better performance (shown by the higher R^2 value).

Does it practically matter which of the two curve fits are chosen? Let us assume for the sake of argument that this meters is to be used in the field at around 70% of the maximum flow tested. That is, the meter will operate around the 2.1e7 Reynolds number calibration point (i.e. 21 m/s). Fig 25 looks at the two polynomial fits in detail around that point. As the fit is not perfect we see that the 3rd order polynomial fit over-predicts the error by 0.026%, while the second order polynomial fit under-predicts the error by -0.02%. However, the choice is to use the 2nd **or** the 3rd polynomial curve fit. Once implemented the fit is no longer referenced to the actual data set. Therefore, the choice is do you choose the 2nd order polynomial and predict 0.046% less gas flow, or choose the 3rd order and predict 0.046% more gas flow. Again, this is not a small increase in uncertainty. This is a guaranteed curve fit induced bias. For a custody transfer meter, if the 3rd order polynomial is chosen then the meter will predict more flow and the seller wins. If the 2nd order polynomial is chosen then the meter will predict less flow and the buyer wins. Regardless of which of the two fits is chosen there is a winner and a loser. If the loser is informed, to accept this is to be complicit in the choice.

Perhaps 0.046% confirmed bias sounds trivial? At 70 Bar this flow rate is approximately 340 MMSCFD which has a market value of approximately $986K dollars per day. Therefore, a 0.046% curve fit induced bias amounts to approximately $165K per annum, or approximately $1 million guaranteed bias over a seven year re-calibration interval. Note that this is a random example using a moderate sized ultrasonic meter (12") with a relatively linear calibration data set, comparing two different polynomials. Larger meters (with more flow), more variable data (leading to bigger differences in curve fit predictions) and different fit comparisons can produce significantly larger biases. Yet, the authors are unaware of any formal rules regulating this issue.

4e.3 Extrapolating and Interpolating Polynomials

Polynomials (and some other) fits are well known to be a poor choice for curve fitting calibration data that may have to be extrapolated. They have the potential to diverge immediately outside the upper and lower limits of the data set. The higher the order of polynomial the more likely and severed this issue is likely to be. An extrapolated diverged curve fit produces nonsense. For all curve fits extrapolate at your own risk, but polynomials are known to be particularly sensitive to the issue. However, polynomials (and some other) fits are generally assumed to be a suitable tool for interpolation. This is an essential requirement of all suitable curve fits. As the calibration shows the meter's performance at the calibration points, the need for curve fitting comes from the requirement to predict the meters performance *between* the calibration points. The fundamental assumption is that the span of unknown performance between the calibration data points is reasonably linear. This is an unspoken reality of flow meter linearization, nobody knows for sure what the meter performance is between the calibration points, and the fit is therefore an educated guess.

Fig 26 shows a hypothetical turbine meter calibration data set. The K factor is correlated to Reynolds number as suggested by AGA 7. The shape is a typical 'turbine curve', showing a relatively linear relationship at high Reynolds number and then rising up and falling off as the viscosity and bearing friction effects begin to be more significant. The graph also shows 2nd, 3rd,

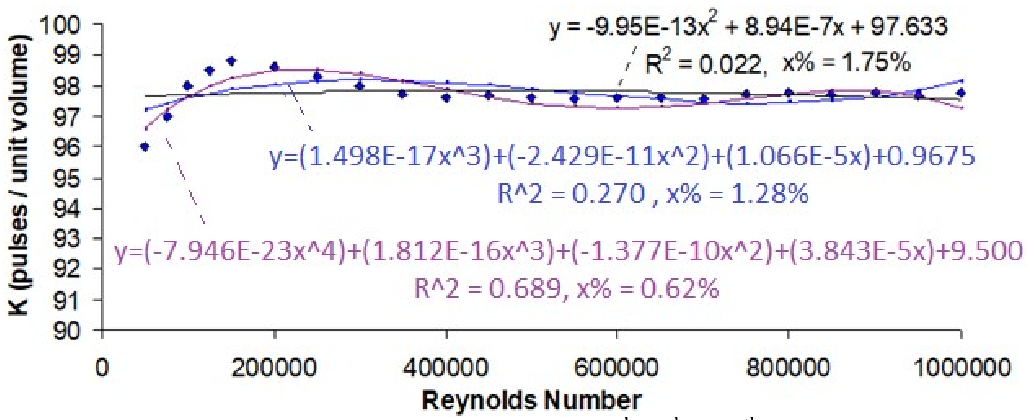

Fig 26. Hypothetical Turbine Meter Curve with 2nd, 3rd, & 4th Order Polynomial Fits.

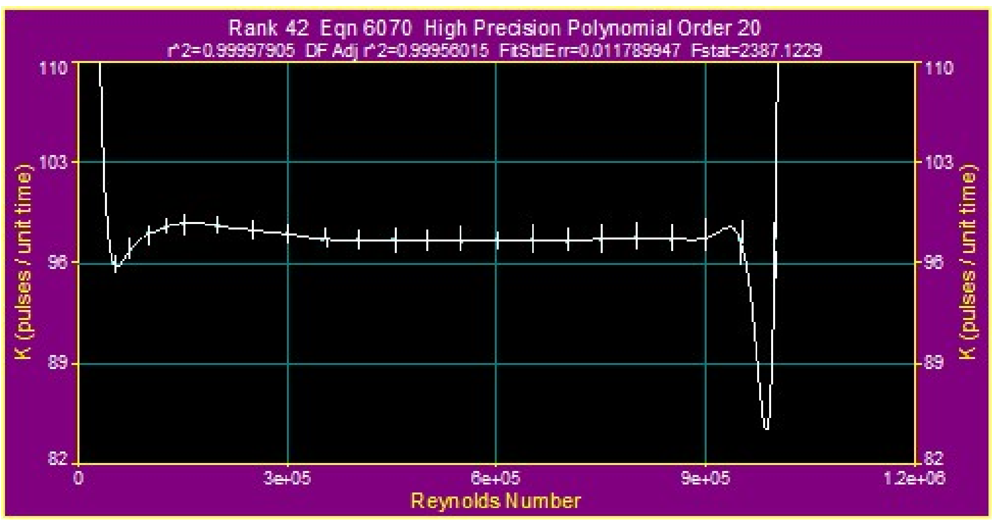

Fig 27. Hypothetical Turbine Meter Curve with 20th Order Polynomial Fit.

and 4th order polynomial fits. As the order of the polynomial increases the R^2 value increases and the uncertainty in the K factor prediction reduces. That is, both common quality of fit checks are indicating that the higher the order of polynomial, i.e. the more complex the fit, the better the fit. So by that rationale the higher the order of the polynomial (or more complex any fit) the better the fit. Except, this is not so…

Fig 27 shows the same data fitted by TableCurve 2D™ to a 20th order polynomial. Markings show where the calibration data is. The R^2 is virtually one (i.e. the coefficient of determination is suggesting the fit is near perfect). The uncertainty of the calibration K-factor data predicted by the 20 order polynomial is virtually zero. However, the fit is obviously terrible. This example is obviously an exaggeration, but such exaggerations are good for making a point. Fig 27 shows the divergence of the polynomial at the limits of the data set. But it also shows that the R^2 and x% uncertainty checks do not necessarily account for interpolation failings. This fit has got a major problem between the two highest Reynolds numbers of the data set. Anybody looking at it would know that despite the very high R^2 and very low x% uncertainty it is not a suitable fit. Common sense tells you so. There is not anything wrong with checking a fit to R^2 or x% uncertainty, *they just don't tell the whole story*. But, how do you convert this common sense observation into a formal mathematical check? In many real world examples the issue will be far more subtle, not as visually obvious, but yet still have a significant effect on measurement. Perhaps different meter manufacturers have their own curve fit check methods to account for such issues, but they are not commonly discussed or agreed upon. This is an example of how the fitting of non-piece wise linear interpolation curve fits tends to be a bit of an unregulated black art.

4f. Fitting A Constant Value Meter Factor

"*Plans based on average assumptions are wrong on average.*" Prof. Sam L. Savage, Stanford University

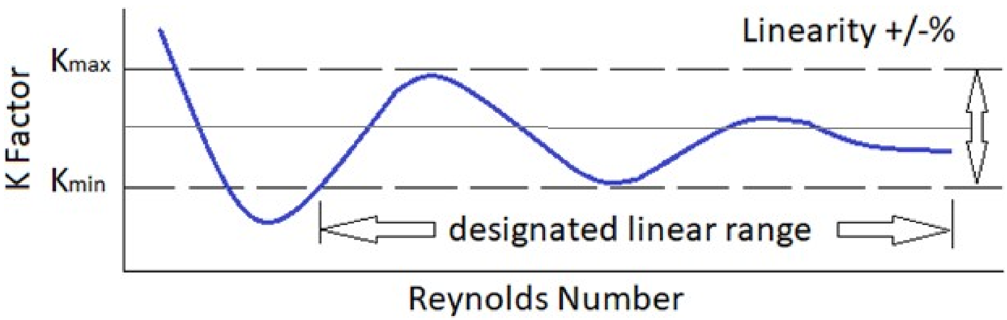

Fig 28. Reproduction of ASME MFC-6-2013 [5] Vortex Meter Standard Graph.

In some applications some flow meters (e.g. vortex meters) are used with a constant meter factor (K_{con}). Fig 28 reproduces a vortex meter graph by ASME. A constant K-factor is to be chosen that will produce a linearity of ±x% across the designated linear range. Once the meter is calibrated, the maximum (K_{max}) and minimum (K_{min}) values across the designated range shall be used to find the constant K-factor (K_{con}) via equation 4 below. ASME calls this factor the mean K factor (K_{mean}) although this can be misleading. It is only the mean of the maximum and minimum points, not the data set as a whole. This procedure is commonly followed for constant meter factors. It is implied that this K_{con} will produce ±x% across the designated range. It does not. In practice it is a good approximation, but *strictly speaking* this is incorrect.

$$K_{con} = \frac{K_{\max} + K_{\min}}{2} \quad \text{-- (4)} \qquad\qquad K_{equal\%} = \frac{2(K_{\max} K_{\min})}{K_{\max} + K_{mid}} \quad \text{-- (5)}$$

Equation 4 actually produces equal magnitude differences. This is different to equal relative (percentage) differences. Equation 5 produces the desired equal percentage difference. To highlight this let us again consider an exaggerated case. Say there was a (very poor performing hypothetical) flow meter that gave a maximum and minimum K-factor of 150 and 50 respectively. Fig 29 shows the effect of using equations 4 and 5. Equation 4 does not produce equal ± x% uncertainty. It produces an equal magnitude difference (i.e. ±50) and different relative differences (i.e. -33% /+100%). Equation 5 produces an unequal magnitude difference (i.e.± -75/+25) and equal relative differences (i.e. -50% / +50%).

There is no correct scientific constant K-factor, it is a choice based on what the meter operator wants to achieve. Again, the choice is subjective. If the operator wants to be exposed to equal ± magnitudes of uncertainty then Equation 4 is required. If the operator wants to be exposed to equal ± percentages of uncertainty then Equation 5 is required.

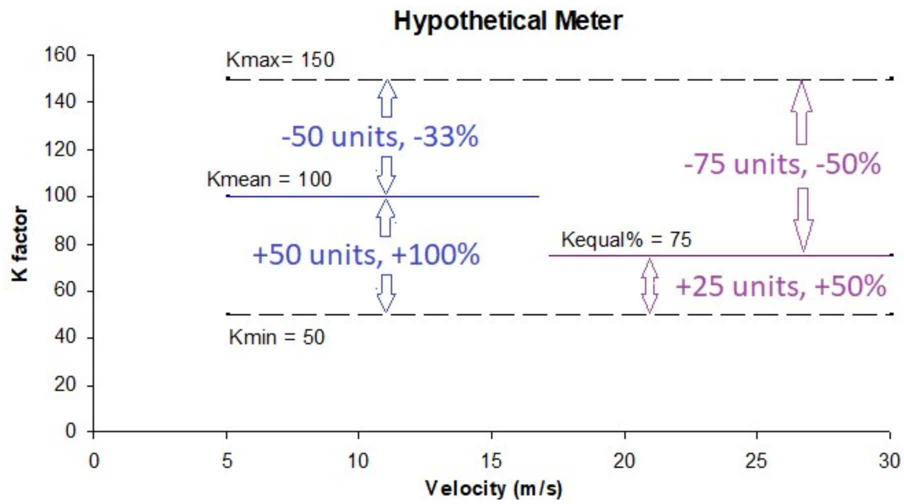

Fig 29. Comparison of K_{con} and $K_{equal\%}$ Effects.

Flow meter end users usually talk in terms of percentage, and hence should use equation 5 to predict the constant K-factor. The authors understand that most use Equation 4. With real flow meters the maximum and minimum K-factors are so close that the difference in K_{con} and $K_{equal\%}$ is very small and the issue goes largely unnoticed. However, technically the common procedure of using Equation 4 to produce a value that will give equal ±x% result is ever so slightly wrong.

4.g An Honorary Mention – A DP Meter Curve Fit is Fixed to a Precise Flow Meters Geometry

A DP flow meter calibration curve fit corresponds to the precise meter geometry used in the calibration calculation that predicts the meter factor (i.e. discharge coefficient). When entering the curve fit and geometry into the computer, if even a slightly different geometry is used then the curve fit is not valid. This may sound obvious and trivial, but this happens more than one may think, and is mentioned here on the request of an operator that had a real problem due to this. The operator is undisclosed, and the following example using a real cone meter calibration data set has nothing to do with the real incident.

Fig 30. Dermaga 20” Cone Meter at CEESI Iowa.

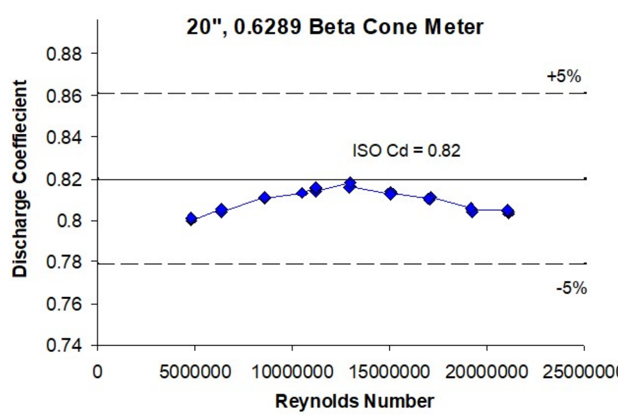

Fig 31. 20” Cone Calibration Data.

Consider a 20”, sch 40, 0.6289β cone meter (seen under calibration at CEESI Iowa in Fig 30). The inlet diameter was 18.793” and the cone diameter was 14.612”. Fig 31 shows the calibration data (that was to be linearized by piece wise linear interpolation). When entering the cone meter data into a flow computer the inlet (D) and cone diameters (d_c) are required. The cone diameter is 14.612”. What would happen if by accident, laziness, or design the cone diameter entered into the flow computer was rounded up to 14.6”? This is only a cone diameter change of -0.08%. It doesn’t sound like that should matter that much? It actually changes the beta (see equation 6) by +0.125%, which in turn causes a step change in the beta dependent geometry terms in the DP meter flow rate calculation (see bracketed term in equation 7) of +0.298%.

$$\beta = \sqrt{1 - \left(\frac{d_c}{D}\right)^2} \quad \text{--- (6)} \qquad\qquad \dot{m} = A\left\{\frac{\beta^2}{\sqrt{1-\beta^4}}\right\} Y C_d \sqrt{2\rho \Delta P} \quad \text{--- (7)}$$

If the cone diameter entered into the calibration data C_d calculation was also rounded up to 14.6” then the resulting bias would be automatically accounted for in the corresponding C_d values. But that is not what usually happens in practice. The calibration facilities virtually always use the precise geometry. If when configuring the flow computer the cone diameter gets rounded off then the calibration curve fit will not correct for this induced geometry bias. In this example, a +0.298% bias on this 20” cone meter when operating at 70 Bar and 329.7 MMSCFD, i.e. a Reynolds number of 17.1e6, is a bias of approximately +$803K per annum.

When configuring a flow computer it imperative to enter the precise geometry values to the same significant figures as was used when calibrating and linearizing the meter. Otherwise the calibration data fit may not be valid and significant flow rate prediction biases can be induced. The authors are unaware of any guidelines explicitly stating this should be checked.

4.h Batch Calibrations and the Inconvenience and Impracticality of Individual Curve Fitting

In 2016 CEESI calibrated a batch of forty DP Diagnostics 2”, 0.45β cone meters. The meters were nominally identical, i.e. manufactured from the same drawing, but in accordance with ISO 5167-5, were to be individually calibrated. ISO predicted that uncalibrated cone meters should have a performance of C_d = 0.82 ±5% to 95% confidence. When calibrated they should individually have a curve fit that has a ≤ 0.5% uncertainty. Other than stating the cone meter should be calibrated to Reynolds number ISO 5167-5 does not state what curve fit should be used. Fig 32 shows one of the meters at CEESI during set up.

Fig 33 shows the massed calibration results. There are three out of forty meters with results obviously outside the ISO prediction limits. There are a few meters around the limits, but when accounting for the uncertainty of the reference flow these are within the ISO prediction. That is, 92.5% of the meters had average discharge coefficients within the ISO prediction. However, ISO also stated “… a simultaneous use of extreme values for D, β, and Re_D shall be avoided as otherwise the uncertainties might increase”. These 2”, 0.45β cone meters have extreme limits of both diameter and beta. Hence, this result generally supports the predictions of ISO 5167-5 for uncalibrated cone meters. In the interest of an independent check on the ISO cone meter statements CEESI chose to curve fit each meter. All forty cone meters had individual curve fits that predicted the discharge coefficient < 0.5% at 95% confidence. This is one of the first independent checks on the new ISO 5167-5 cone meter standard.

Calibration facilities tend to only produce the calibration data. They are not liable for choosing any particular linearization method. That is the responsibility of the client. The client may *instruct* the calibration facility to apply a particular curve fit, e.g. piece wise linear interpolation

Fig 32.One of a Batch of Forty 2”, 0.45β Cone Meters at CEESI

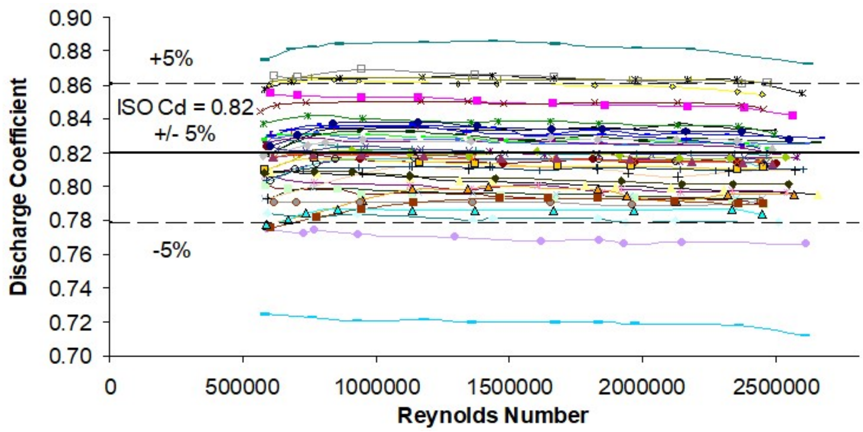

Fig 33. Forty 2”, 0.45β Cone Meter Calibration Results at CEESI.

in the case of many ultrasonic meters, but the responsibility for the choice is with the client. This was the case with these forty cone meters. Whereas CEESI analysis showed that ten data sets could be fitted to a constant discharge coefficient, seventeen to a linear line, four to a second order polynomial, and nine needed a 3rd order polynomial, the client did not wish to use any such fits. (No other curve fits were tested, although other forms would work for individual cases.)

If such individual curve fits were to be chosen, the practical reality is it is time consuming to data fit each meter and then double check the quality of that fit. It is also time consuming monotonous detailed work to program each individual fit into each computer. It is not an attractive option to meter manufacturers making meters on mass. Furthermore, not all computers have software that can take different curve fits. A programmable flow computer would be required. In this case the end user did not want the linearization done in a mainframe computer but in flow computers. The flow computer the client had chosen only allowed piece wise linear interpolation. So in practical this was the only choice.

Two of the authors have historically had reservations regarding the piece wise linear interpolation methodology. It is a rudimentary ‘curve fit’, if indeed it can be called a curve at all. It is a mindless procedure that can be argued to only superficially suggest no associated prediction error, i.e. by definition R^2 is one and the meter factor prediction uncertainty looks like zero percent. There is in reality the hidden uncertainty that occurs on interpolation. However, when considering the pros and cons of other linearization techniques above the authors have come around to accepting that there are some strong arguments *for* piece wise linear interpolation.

5. Piece Wise Linear Interpolation

"When you have them by the balls their hearts and minds will follow"
Theodore Roosevelt, President of the United States 1901-09.

For meters using flow computers it is rather difficult to argue against piece wise linear interpolation (otherwise known as 'using a look up table'), when virtually all flow computer products only offer this methodology. Any other choice requires significant effort, first in curve fitting the data, and then programming the flow computer. Such effort would have to offer clear practical advantages. And it is arguable that it does not.

This issue is an example of the largely unspoken, inherent, innate power held by flow meter computer supply companies. Whether it is flow computer companies, or meter manufacturers with embedded processors, the software they choose to add or not add dictates what linearization techniques (and other calculations) industry can and cannot practically apply. In this respect these computer suppliers and not the end users dictate how meters will be linearized. As such, unless these computer suppliers can be persuaded to add other linearization options, or the end user is willing to spend time entering linearization code into programmable computers, the argument on which linearization method is best is a moot point. Most end users are 'stuck' with piece wise linear interpolation. And with no choice many do not concern themselves greatly with the validity or invalidity of using that method. That is what is on offer, it is what most others do, nobody is going to get fired for following the crowd, it is one less thing to worry about. And indeed there is a lot going for piece wise linear interpolation.

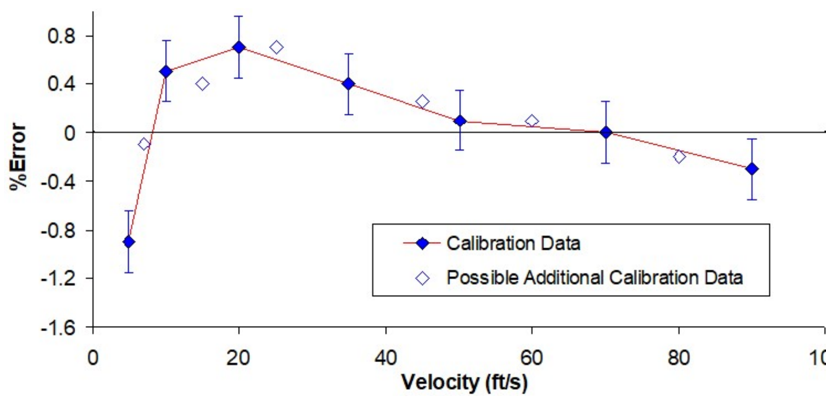

Fig 34. AGA 9 Ed 1 Hypothetical Data With Piece Wise Linear Interpolation Curve Fit.

Consider Fig 18. The data is shown to be difficult to fit with simple curve fit techniques. A lot of work may be needed to satisfactorily curve fit any flow meter data set. Fig 34 shows the same data fitted to piece wise linear interpolation. This method easily fits the data, and should fit any data set, and no great effort is required for it to do so. This means its use alleviates any requirement for individual bespoke flow curve fitting analysis.

Many linearization techniques such as a constant meter factor, linear, or polynomial curve fits have some tendency to focus on a central trend and ignore the details of variations around that trend. Such variations are not always 'bad points', i.e. poor data, but can be indications of some underlying physical phenomenon occurring in that flow region. They are ignored or played down to the detriment of the quality of the flow meter linearization. In the words of the renowned scientist Stephen Jay Gould: "Our culture encodes a strong bias either to neglect or ignore variation. We tend to focus instead on measures of central tendency, and as a result we make some terrible mistakes, often with considerable practical import”. Such can be the case with flow meter calibration curve fits. But not so much when using piece wise linear interpolation.

The authors are not certain of the origins of the trend towards piece wise linear interpolation. It is suspected that it was at least in part due to convenience rather than any detailed technical considerations. Regardless, whether it was by design or chance, this linearization technique has some significant benefits (with a few arguable disadvantages). Advantages of piece wise linear interpolation include:

1) Avoiding the possibility that no curve fit is found that is simple and satisfactory (see Section 4b).

2) Avoiding the engineering time required for individual meter curve fitting analysis and programming a computer with that unique fit. There is a guarantee that this method will work on virtually any data set with low scatter on each repeat data point (see Section 4h).

3) Avoid having to choose between R^2 vs. lowest overall x% uncertainty (see Section 4c). The method produces both at the same time, i.e. the ‘perfect’ R^2=1 and x% = 0%.

4) There is no concern about more data in the sweet spot skewing the result, across a set range this method is immune to this issue (see Section 4d).

5) This method at least guarantees the correct results at the calibration data points. That is, unlike other curve fits there is no residual bias in the prediction of the meter factor at the calibration points (see Section 4e.2). The buyer and seller do not have to deal with a bias, i.e. one does not have to be complicit on the curve fit choice. This method only produces the more palatable uncertainty.

There are a few arguable disadvantages to piece wise linear interpolation. Regarding point three above, the R^2=1 and 0% uncertainty result is an intrinsic inevitable consequence of how the fit relates to the known calibration data points. It gives an *illusion of perfection*. It is of course *only* an illusion. There is the flow lab uncertainty, the repeatability uncertainty, and like all curve fitting the interpolation between points is an educated guess with an associated guesstimated uncertainty. There is a tendency to use this technique and then *falsely* imply there is no associated data fit uncertainty. For example, the ISO 17089 Ed 1 Section 7.7 uncertainty calculation suggests “*calibration curve correction ‘on’: 0%*” curve fit uncertainty. AGA 9 3rd Ed Table A.5 also implies this. However, it is an absurdity to imply any curve fit is infallible. Fig 34 shows hypothetical extra data taken between the original calibration points. Such points have virtually no chance of falling on the linear lines connecting the original data, i.e. there will be some unknown bias, i.e. additional uncertainty. It is for the experienced flow metering pragmatist to estimate just what amount of uncertainty should be expected.

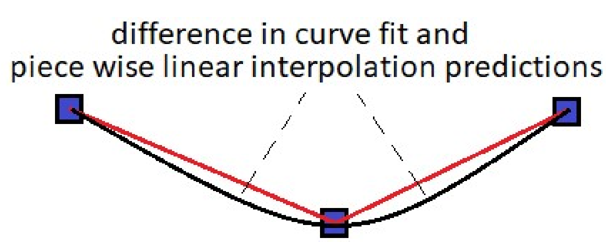

Fig 35. Sketch Showing Differences Between Curve Fit and Linear Interpolation Predictions.

Another potential disadvantage to piece wise linear interpolation is that it could be argued that a linear line between any two calibration points is less likely to be as realistic as a natural curve being traced out between a group of points showing a general trend. Fig 35 shows this. However, the difference is typically small, more so with calibrations with many equally spaced points, although they are both still educated guesses.

The flow meter standards do not tend to dictate which linearization method is used, and the authors do not know of any formal text commonly used that discusses detailed pros and cons of different linearization methods. However, piece wise linear interpolation does seem to be the dominant method used. Most flow computers have this as the only option. Most ultrasonic meter manufacturers tend to use this method. From the authors limited knowledge some contracts imply that this linearization method should be used, although it appears to be due to habit and practical ease of use rather than any scientific preference.

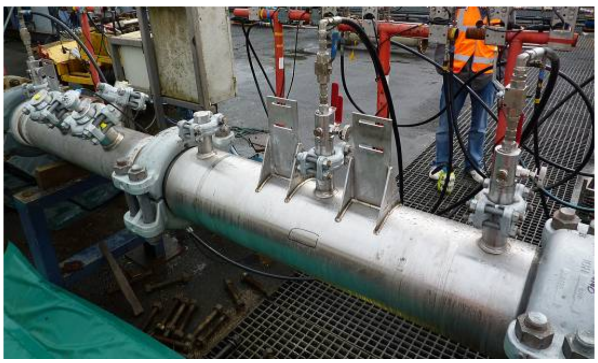

Fig 36. 10” Venturi Meter Under Calibration

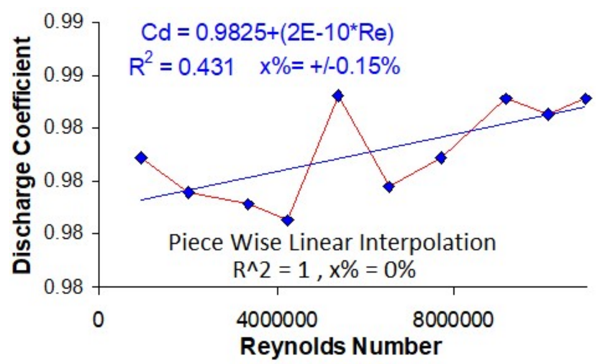

Fig 37. 10” Venturi Meter Calibration Fit

As way of example, consider a 10”, 0.4β Venturi meter calibrated at GLIS (as shown in Fig 36). Fig 37 shows the resulting data. One of the authors witnessed this test and considered a linear fit suitable, giving a ±0.15% C_d prediction over the 12:1 turndown tested. However, the meter manufacturer, skid contractor, and end user all insisted that a ‘look up table’, i.e. piece wise linear interpolation, would be used. The reason was that the contract ‘sort of’ stipulated it. The contract stated that the calibration data must be entered into the flow computer. As the only option for that flow computer was to fill in the look up table with the calibration data the contract was in effect implying (if not explicitly stating) that piece wise linear interpolation was to be used. As far as the authors are aware it wasn’t a statement due to well thought out technical reasons, but rather it was solely due to the practical ease of implementation. But from a technical stance, this does indeed appear to be a defensible linearization choice.

6. Conclusions

“*The only certainty is that nothing is certain*” – Pliny the Elder, Roman Scholar (23-79 CE)

The integrity of a flow meter’s linearization technique is a critical part of any flow meter’s uncertainty calculation. Lax attention to the details of a linearization technique undermines the integrity of a flow meter’s claimed performance. Surprisingly, for all its importance, flow meter linearization is not discussed in much detail, not comprehensively regulated, and not always required to be fully transparent. One important issue is the correlating parameter being chosen. Another important issue is that although the choice of flow meter linearization technique may superficially appear to be objective, in reality it can be rather subjective.

Different linearization techniques can be chosen that subtly modify flow meter output without necessarily violating any contract, standard, or accepted norm. If linearization techniques are not fully disclosed then in reality the end user does not know the true flow rate prediction uncertainty of that meter across its flow range. This is a particular concern with the modern trend for flow metering systems to be supplied as a complete package, meter, sensors, and computer inclusive as embedded parts. Whereas such systems offer many advantages, not least the plug and play ease of use, they also tend to have embedded processors. These can have inaccessible, un-auditable code that could carry out extra undisclosed linearization techniques. Such lack of rules or transparency can give any such metering system manufacturer wiggle room to make a surreptitious ‘correction’ on a problematic meter. In effect the lack of required transparency can potentially be a ‘get out of jail free card’ for the manufacturers. In such a scenario the end user or auditor do not have a full understanding of what linearization techniques are applied and why, and what the ramifications of these techniques may be. Exasperating this issue is unlike stand-alone flow computer products (which have API 21.1) such metering systems have no guidelines dictating what linearization and audit information they must supply. Most modern flow metering system packages have very poor audit reporting capabilities. They do not necessarily disclose all linearization techniques applied. Although such modern metering systems are of great benefit to industry part of the maturing process should surely be some regulation, or at least end user stipulated greater transparency, on the full suite of linearization techniques employed.

Transparency is much less of a concern when the meter linearization techniques are carried out in the end user’s mainframe computer or in a reputable stand-alone flow computer product. However, even here, various linearization techniques programmed into these computers can have various veiled issues. Such issues tend to be out of sight and out of mind in an industry that does not much monitor or strongly regulate linearization techniques. Thankfully by design, or rather fortuitously (if you don’t believe it was by design), most flow computers default to using piece wise linear interpolation as the linearization technique. Although superficially unsophisticated, closer inspection of this technique highlights several technical and commercial advantages over more complex curve fitting techniques. Nevertheless, as with flow metering system packages, it would be to industries advantage if more attention was paid to the pros and cons of linearization techniques available in flow computers.

Finally, the authors suggest that flow meter linearization is important enough that it should have some reputable guideline, technical report, or standard. Failing that, industry should at least be more aware of these issues and require linearization techniques are made more transparent.

7. References

1. Geach D. and Jamieson A. “Wet Gas Measurement in the Southern North Sea”, North Sea Flow Measurement Workshop, St Andrews, UK, Oct 2005.
2. Mills C. et al “Measurement of Flow in Viscous Flows”, International Symposium of Fluid Flow Measurement, Arlington, Virginia, USA, April 2015.
3. Feynman R.: Rogers' Commission Report into the Challenger Crash Appendix F - *Personal Observations on Reliability of Shuttle* (June 1986)
4. Miller, R. “Flow Measurement Engineering Handbook” Ed 3, ISBN-10: 0070423660, Published by McGraw Hill.
5. American Society of Mechanical Engineers, “Measurement of Fluid Flow in Pipes Using Vortex Flowmeters”, MFC-6-2013