Summary of Pilot study meter calibration in Colorado

Comments prepared by Dave Thompson: Measurement Technology International Ltd.

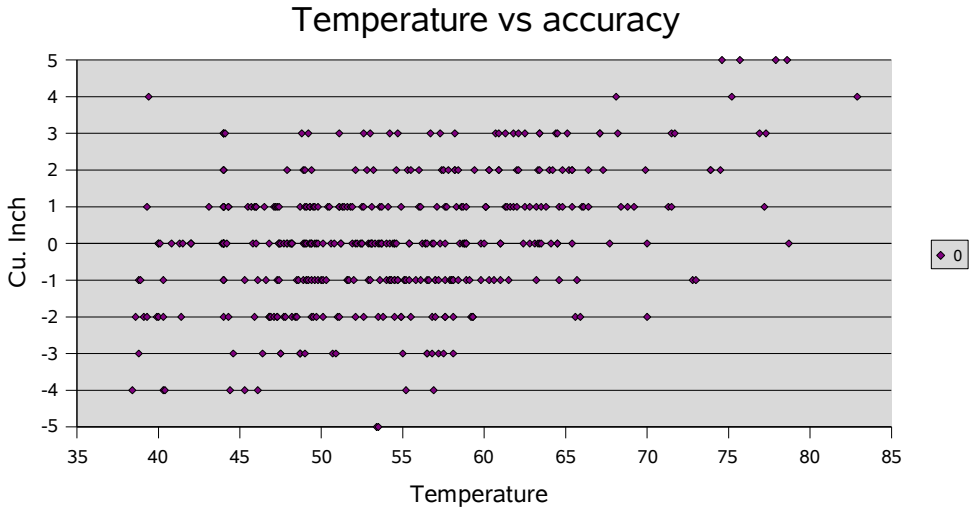
The results of the multi-station study show the expected results of test can comparisons to EPT type accuracy testing. All data gathered confirmed the results we were expecting. The results from testing using the EPT system were very consistent to the dispenser results. The testing using the test can were not so consistent.

This type of testing is very hard to do in the field and get lab quality results because of having no reference to base the tests on. In the field test the only commonality between the two different methods is the dispenser. This then automatically adds in a window of error equal to the repeatability of the dispenser meter. The repeatability of the trade dispenser meters is usually within 0.25 %. In 5 gallon testing this equates to a window of error of 3.125 cubic inches. The comparison between the two methods then has to use this window of error as a baseline.

The following charts show some of what I am describing. The first shows the apparent randomness of the results from the test can method. We charted results based on temperature and results based on fuel type. These results showed the same thing. There is no apparent pattern, although all results fall in the acceptable range. By removing the window of error from the dispenser these results would have been much better.

All of the the following charts show the results that were obtained from data provided by CGRS.

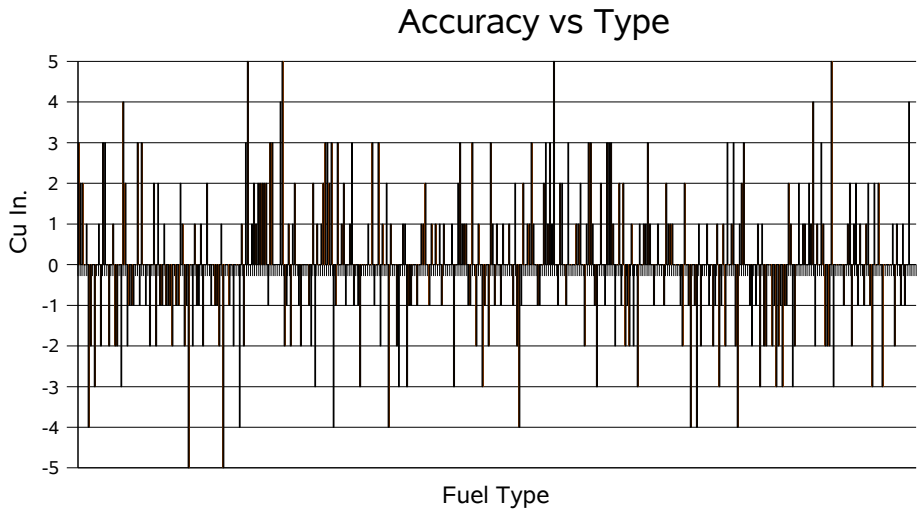
Test can – Temperature vs. Accuracy



Can Calibration

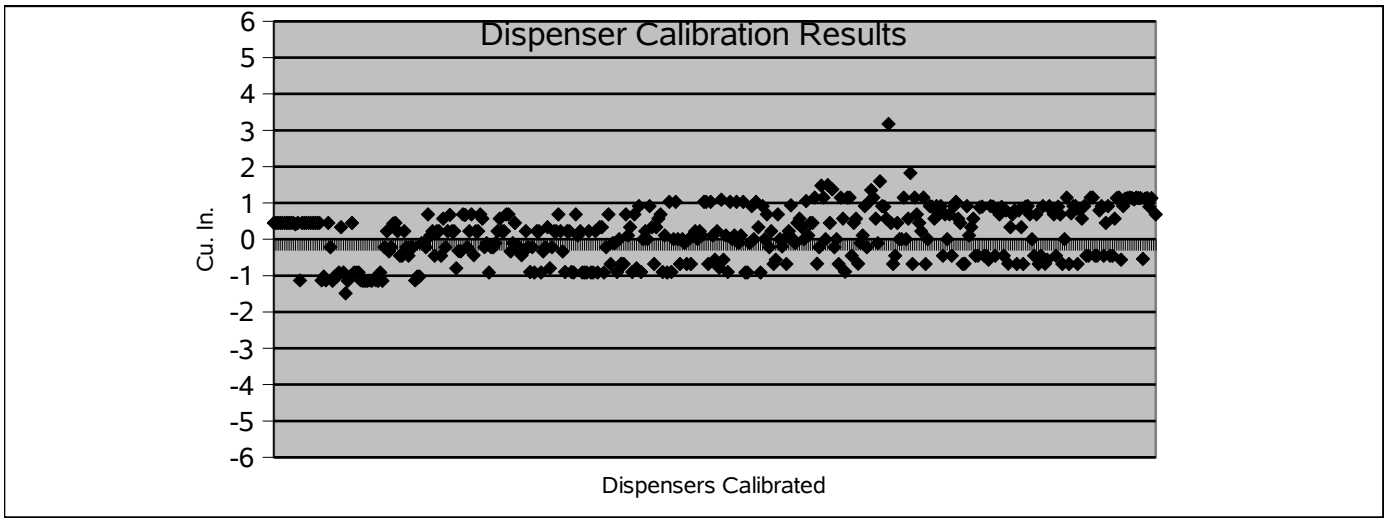
The chart above contains the results of the can procedure based on the temperature of the fuel. The repeatability is all over the map. +-5 cu in. which shows the process of can calibration in the field at being 0.8 percent repeatable. Take out the 0.25 %* being the dispenser repeatability and the final error in the process of using a can for calibration is 0.55 percent. The reason for this is incompatible test

procedures, and too many variables to account for during the test as shown by the charts in this analysis.



There is no consistency of calibration when comparing fuel types The repeatability does not change with the use of different fuel types. The chart below shows the results on one type of fuel. The results are consistent with any of the fuel types. Essentially no pattern is discernible.

EPT System



This at first glance shows the repeatability of the EPT system, yet what it really shows is the repeatability of the dispenser under test. The data set shows a close tracking of the EPT test method to the way a dispenser works. The repeatability of the EPT is much better than the unit under test but the results are masked by the much larger variations in the repeatability of the dispensers being tested. The +- results show the dispensers at approx 1.5 Cu in. which would be approx. 0.25 percent repeatable. This is what we expect to see.

This method of calibration uses a closed loop continuous flow process to minimize external variables which keeps the same conditions in the same meter under test as compared to the EPT system. We then know each system is affected the same by any external influences. The EPT system tracks the results of the dispenser. Even if the repeatability of the dispenser is questionable we can track these results. The variance is the difference in the repeatability of the two systems. The EPT system has a repeatability factor of less than 0.03%. Because it is impossible to calibrate any dispenser to a zero calibration most Oil companies specify a calibration window of 0.2 percent. This window takes into consideration that the dispensers have a repeatability of 0.2 percent and will remain calibrated in this window. Since at any specific time of calibration the dispenser will be in this window of error it is important to make multiple runs at one flow rate to determine the center point of the repeatability window and calibrate this to zero. The dispenser will now be within this 0.2 percent window in trade use. More precise calibration than this is not possible due to the limitations of the trade device.

Observations

The problem with this type of testing or comparison is the fact that the reference is the dispenser under test. This leaves a probable window of error equal to the repeatability of the dispenser plus or minus the repeatability of each individual test method (test can or EPT system).

The repeatability of the unit under test (the dispenser meter) is wider than that of the testing systems. The other problem with testing comparisons is the differences in the testing methods themselves. There are no commonalities in the variables seen by either of the test methods. One system is a closed loop method where the variables are seen by both the unit under test and the testing system (EPT). The other method is open neck can testing. This system has no similarities between the unit under test and the test method. In both cases in order to accurately determine accuracy of the system being tested, any variables that can adversely affect the outcome must be minimized.

As some devices are more accurate than the specification the oil companies set, a required tolerance in approximately the 0.2 percent range is the target specified by the oil Companies. It would be unrealistic to expect a trade type device to hold a 0.0 percent repeatability tolerance when they are designed to meet a 0.20 percent tolerance. Different oil companies use different specifications; all around the 0.20 percent number.

Temperature has to be accounted for.
Environmental conditions have to accounted for.
There are no controlled conditions present for these tests.
The tests were not carried out on the same day for accurate comparison purposes.
The procedure during the tests were not the same. For example:

- No stabilization of the fuel temperature was done with the test can method.
- No repeatability testing was done with the test can method to determine unit under test flow characteristics.
- Temperature measurement was completed using an IR temperature device which would only record surface temperature and then only at the end of the test.
- How many times was the nozzle clicked off at the end of the test to fill the test can to the appropriate marks for evaluation? (note: this puts the meter under test out of its flow range for valid tests)
- What is the minimum resolution read on the test can? (1 cubic inch?)
- Was the drip time when emptying the can consistent?
- Was the can level?
- Was the product recirculated to stabilize temperature prior to testing? (if so how how much

product was recirculated per meter?)
- Was the can wet down prior to testing?
- When was the meniscus read after the test?

Notice the results from the EPT system are within the 0.2 percent window. When using the test can method, there can be no way to determine repeatability from the test can results, as there was no repeatability data available (since the tests were not completed). Without doing these tests the calibration personnel cannot center the calibration around the center of the repeatability factor. For example: If a dispenser repeats 0.2 %, the zero calibration mark should be at the 0.1 percent center of repeatability.

The results from the testing, based on hundreds of hoses, shows that the EPT system corresponded to the units under test in that the repeatability remained consistent throughout all the tests. By looking at the fast flow repeatability numbers over numerous test volumes the numbers matched within the repeatability requirements that NIST specifies for retail trade dispensers in this flow range.

After reviewing Jeff's observations on the calibration procedure used by the Weights and Measures inspectors, it appears that the results of their calibration are affected to a great extent by the human factor. There are no hard and fast rules applied in the calibration of dispensers. Use of laser IR temperature probes is not an effective method of liquid temperature measurement as only the surface temperature is measured.

These results still cannot be taken as absolute due to the fact that the reference in a these comparisons is the dispenser that was under test. The reference in these tests has a moving zero point due in part to the repeatability of the dispensers tested and the difference in technologies used.

A more accurate evaluation cannot be completed, by metrology authorities in the field, to laboratory standards, as there is no way to have maintain controlled or referenced conditions.

Many different data comparisons could be presented.. Unfortunately the bias from the dispenser being used as a reference makes the results of such comparisons difficult to interpret. A fair comparison dictates that the bias needs to be removed.

Trade devices have to meet 40% of absolute tolerance (0.20 percent in the case of retail fuel dispensers) repeatability specification according to NIST Handbook 44 section, section T.3. and N.6.1.1. or Measurement Canada regulation 263 (b) 2/5 of the limit of error for non controlled conditions (0.20 percent for retail fuel dispensers).