CALIBRATION, VERIFICATION & BINNING RESOLUTION CORRECTION TECHNIQUES OF WIM DATA

CHRIS SCHILDHAUER AND ROB SIK

Mikros Systems (Pty) Ltd, P O Box 75034, Lynnwood Ridge, Pretoria, 0040
Tel: 086 111 5393; Email: chris@mikros.co.za and rob@mikros.co.za

ABSTRACT

With the evolvement of modern Weigh-In-Motion equipment both in the field of sensor and logger technology the way in which calibration and verification is undertaken has also changed. This paper discusses some traditional calibration and verification methods and suggests how to implement more reliable in-field and statistical calibration and verification methods. In addition the paper discusses and presents a technique of correcting bias resulting from "binning" recorded axle weights.

Keywords: Weigh-in-Motion, WIM, Calibration, Verification, Front Axle Mass, FAM, Truck Tractor (TT) Method, Binning, Resolution, Bias.

1 INTRODUCTION

With the evolution of High-Speed-Weigh-In-Motion (HSWIM) equipment both in the field of sensor technologies and logger technologies, the way in which calibration and verification of WIM data is undertaken has changed. The authors discuss some aspects of these changes.

Also discussed is the effect of digital axle mass resolution of a logger on the discrete axle mass distribution produced by binning the individual axle masses of vehicles.

Please note that the terms weight and mass are used interchangeable and that by weight is meant the gravitational force exerted downward multiplied by the gravitational constant resulting in an equivalent mass. All weights expressed in this paper Ton refers to the SI metric ton.

2 IN-FIELD CALIBRATION AND VERIFICATION

Most of today's WIM systems use digital methods rather than a physical 'turn-the-knob' process to calibrate a WIM system. When a wheel moves over a given WIM sensor the signal from the sensor is digitized giving a set of signals (s_1, s_2, s_3… s_n) for the response generated by the wheel. These are then processed by the WIM logger using an appropriate signal processing relationship (f) to a single raw response (r) that is related to the weight of the wheel. The signal processing relationship (f) depends on the signals (s), on the type of sensors and may also depend on other variables, such as the speed, road surface deflection etc. (variables x). The actual weight of the wheel (w) or axle is then related to this signal response through a calibration constant (C). One can thus write

$$w = C * r \quad \text{where} \quad r = f(s, x)$$

On old (pre-digital) WIM systems this calibration 'value' (C) was set on the logger by physically turning a knob on a gain amplifier or by adjusting some potentiometer/resistor values. These adjustments were not digitized and often involved setting factors that were not linear. On many modern loggers, all adjustments are digitized and saved as part of the raw vehicle data or as part of the logger setup. The logger thus uses the digital calibration value C only to report on the final weight. It is thus possible to change the calibration value C in any given data set at any time by substituting the value used with a new value.

For modern digital WIM loggers this implies two things. Firstly one can use the same runs needed to calibrate the WIM system to also verify the WIM system. And secondly, the urgency of calibrating a WIM system after installation falls away as one can now post-calibrate all 'non-calibrated' WIM data after the WIM has been calibrated.

Since WIM systems measure the instantaneous in-flight (i.e. dynamic) weight, the weight measured by the WIM (w) differs from the static weight (W) from run to run. Under normal conditions this difference follows a normal distribution and the aim of the **calibration process** is to reduce the mean weight error or mean weight difference to zero.

The aim of the **verification process** is to determine the extent to which the WIM weights differ from the static weight. This is normally expressed as the standard deviation of WIM errors. This deviation gives information on how well the WIM determine the corrections required to relate a WIM weight distribution to the actual static weight distribution.

On older WIM systems that required physical adjustments to the electronics a number of calibration runs had to be done, and the calibration 'knobs' turned, until the system was deemed calibrated; only then would one proceed with the verification runs. To get this calibration 'run' information is time consuming and also expensive. One is thus limited to the number of runs that one has available for either process and very often too few runs are actually done to determine the mean error (i.e. the calibration) with any degree of certainty. On a digital system one can use all runs to both calibrate and verify a system and thus increases the certainty without sacrificing any runs.

Using classical statistics one can show how the number of runs affects the certainty with which one can determine the calibration and why it is advantageous to use both the 'calibration' and 'verification' run information as one set. The relative error (e_i) between the WIM weigh result (w_i) and that of the statically weighed vehicle result (W) is defined as (expressed as a percentage)

$$e_i = 100 * (w_i - W)/W \qquad \text{Equ. 1}$$

If there are (n) samples then the mean error ($\bar{x}$) and standard deviation of error (s) are given by

$$\bar{x} = (\sum e_i)/n \qquad \text{and} \qquad s = ((n \sum e_i^2) - (\sum e_i)^2) \,/\, n(n-1))^{1/2}$$

A WIM system is typically calibrated by adjusting the calibration factor C such that the mean error ($\bar{x}$) is ZERO. If one uses the individual axles for calibration, for example, then the calibration factor C is

$$C = (\sum W_a) \,/\, (\sum r_{ai})$$

If the number of samples (n) is large then the calculated mean error ($\bar{x}$) approximates the true mean error (μ) well, but if the number of samples is small then the difference between $\bar{x}$ and μ can be quite significant. The same applies to the accuracy with which the calculated standard deviation (s) approximates the true standard deviation σ. Classical statistics tells us that the confidence interval $A(1-\alpha)$ for the mean with n samples (and n small) is given by

$$\bar{x} - t_{\alpha/2}s/\sqrt{n} < \mu < \bar{x} + t_{\alpha/2}s/\sqrt{n} \qquad \text{Equ. 2}$$

where $t_{\alpha/2}$ is the value of the Students t distribution with n-1 degrees of freedom and leaving an area of $\alpha/2$ to the right of the distribution

Similarly the certainty with which we can determine the true standard deviation (σ) is also dependent on the sample size (n) and the confidence interval is given by

$$(n-1)s^2 / \chi^2_{\alpha/2} < \sigma^2 < (n-1)s^2 / \chi^2_{1-\alpha/2} \qquad \text{Equ. 3}$$

Where $\chi^2_{\alpha/2}$ is the chi square distribution with n-1 degrees of freedom leaving and area of $\alpha/2$ to the right of the distribution.

For example, if one is dealing with an ASTM Type I WIM system (Reference 2), then one expects that 95% of the individual axles weights fall between ±20% of the static weight i.e. one would expect a standard deviation of errors not to exceed 10.2%. That is, one expects the probability $P(-20\% < e < +20\%)$ to be 0.95. For large samples the Student t distribution reduces to the normal distribution. For a zero mean ($\mu = 0$) one has that $P(-1.96 < z < +1.96) = 0.95$ where $= e/\sigma$. So $\sigma = e/1.96$ i.e. 20%/1.96 or 10.2%.

If one plots the expected uncertainty of the calibration at 95% confidence for the number of available samples and a standard deviation of 10.2% then one gets the plot in Figure 1 where the results of equation 2 is plotted assuming a mean error of zero (which is the target of a calibration). In Figure 2, for the same conditions, the extreme ranges of equation 3 are plotted.

Typically, a 5 axle vehicle with an axle grouping of (1 2 2) is used to calibrate and verify a WIM system. Usually 2 runs are done to 'calibrate' the system and 10 runs to verify the system. If one uses the individual axles to calibrate the system, then each 'run' adds 5 samples to the sample set. After the first run one has a potential error on the calibration of 13%, after the 2nd run the potential error is 7% while after 10 runs (a sample size of 50) the potential error has reduced to 3% (Figure 1).

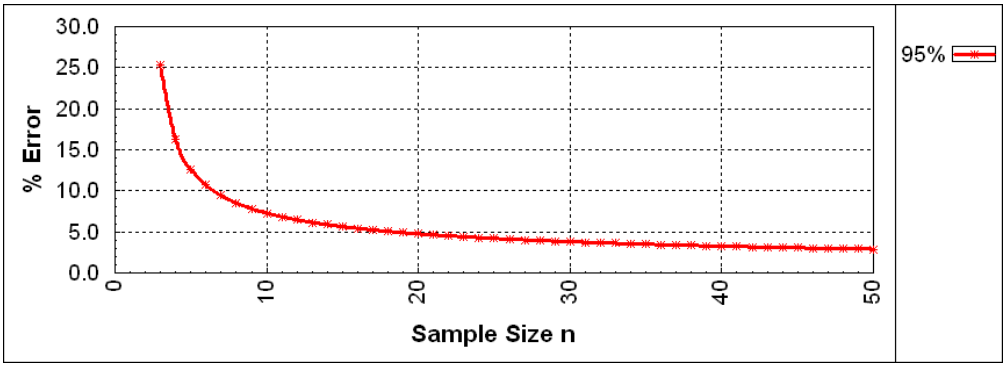

Figure 1 – Potential % Error on the calibration of an ASTM Type I WIM as relates to sample size.

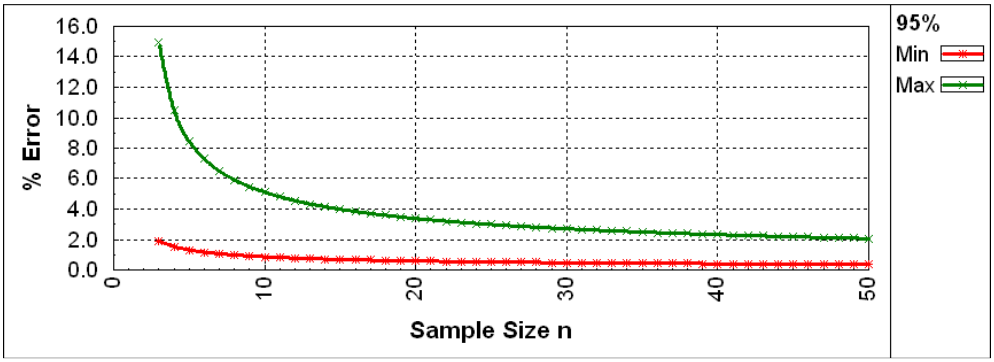

Figure 2 – Potential % Error when estimating the σ for axle weights on an ASTM Type I WIM

To determine whether a system performs within specification the system is verified using 10 runs, thus 50 axles. On an ASTM Type I WIM (± 20%) one could then end up overestimating the WIM performance (s) by 0.4% or underestimating its performance by 2% (Figure 2).

So, if one were to use only two runs to calibrate then there is a large risk that one does not get the WIM calibration right. Clearly more effort should be placed on the calibration of the system, rather than verifying the system. If the system is fully digital, then all sample sets can be used for both the calibration and verification of the WIM system.

3 STATISTICAL VERIFICATION AND STATISTICAL CALIBRATION

The signal from the WIM sensors of a WIM system can drift over time and the performance of the system can also degrade over time. One needs to correct for this drift and catch degradation of the WIM signals in time.

3.1 Front Axle Mass Verification/Calibration Method

Traditionally most WIM users were using the average Front Axle Mass (FAM), also referred to as steer axle weight, from a given class of trucks as a reference to check the calibration of the WIM data and the standard deviation of the FAM to check for failure or degradation (not discussed here) (Reference 5). This method is dependent on the loading of the selected trucks (full or empty) and also dependent on the location of the WIM (i.e. whether it is located on an incline or decline and what the cross-fall at the WIM site is). Never-the-less, with site specific knowledge on the behaviour of the FAM at a given site,

this reference is still a good check to evaluate the potential drift and performance of a WIM system.

As an example the authors chose a WIM site in Australia that has 4 WIM sensors; sensors 1 & 3 in the left and 2 & 4 in right wheel track. The site was calibrated at the end of November 2010 using a 6 axle (123) uniformly loaded truck. In Figure 3 the daily average FAM of all 4 sensors for all 6 axle articulated trucks is plotted. The FAM is fairly stable over short periods (a month) except perhaps on Sundays and Mondays when it abruptly climbs. Over long term the FAM seems to slowly rise.

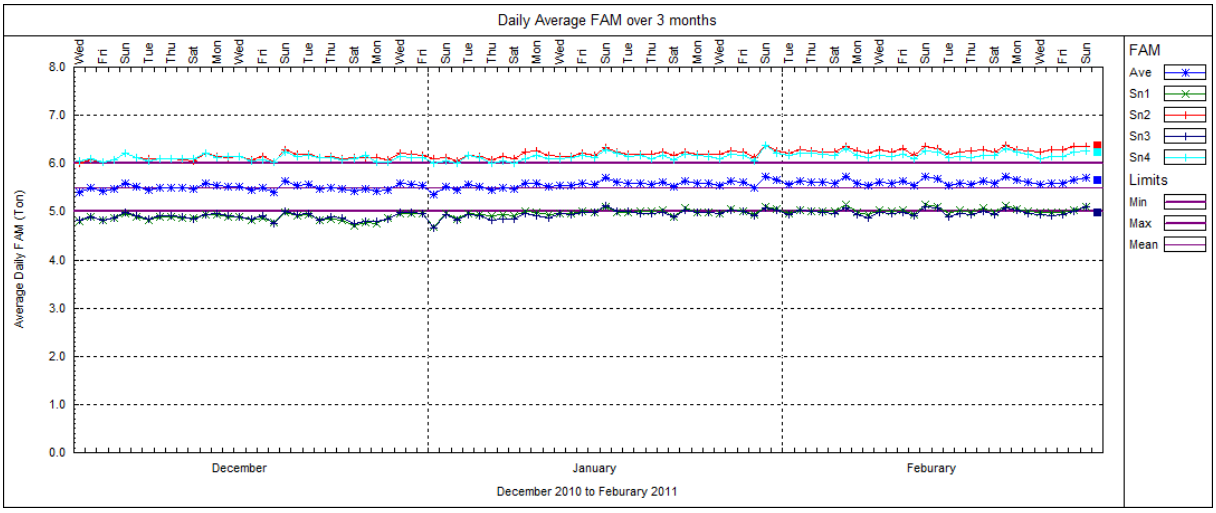

Figure 3 – Average daily FAM for 6 axle articulated vehicles

In Figure 4 the average FAM for December 2010 is shown as a frequency distribution. The mean FAM for these trucks is at around 5.5 Ton and has a standard deviation of 0.7 Ton. (The fact that the mean FAM is 5.5 Ton shortly after a calibration is going to be used later on to determine a value for the TT-Truck).

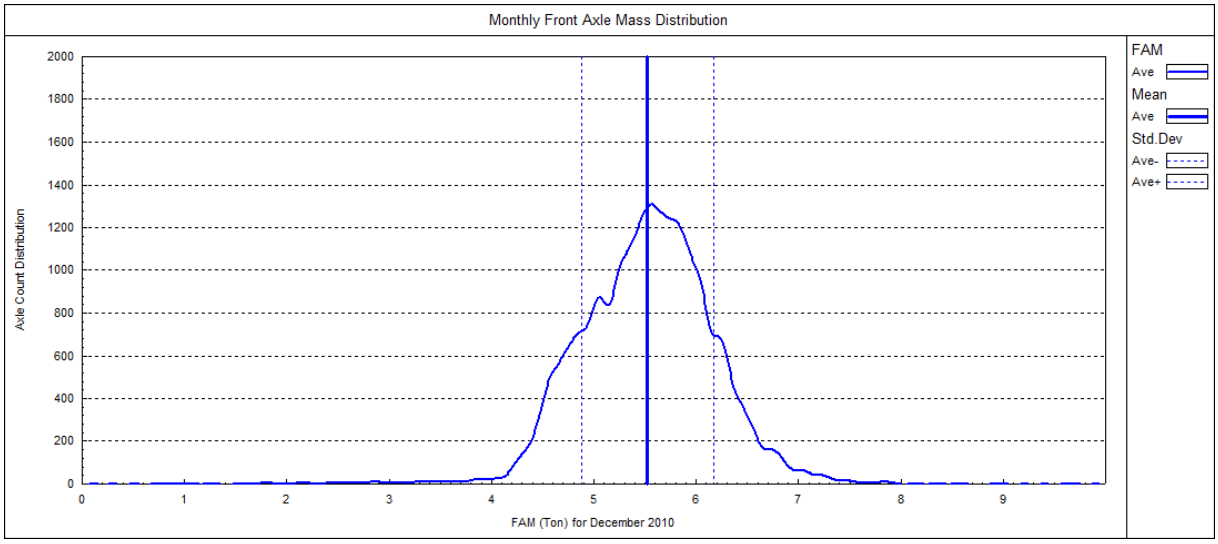

Figure 4 – Average FAM Frequency Distribution of 6 axle articulated vehicles

If one splits up the trucks according to their GVM loading as Full (>80% of GVM legal limit), Half (>50% but <80%) and Empty (<50%) and plots these in the same fashion then on sees that the FAM is load dependent (Figure 5). The mean FAM for Empty, Half and Full trucks are 5.3, 5.6 and 5.8 Ton respectively. Clearly the mean FAM will depend on the loading of the selected trucks on a particular day.

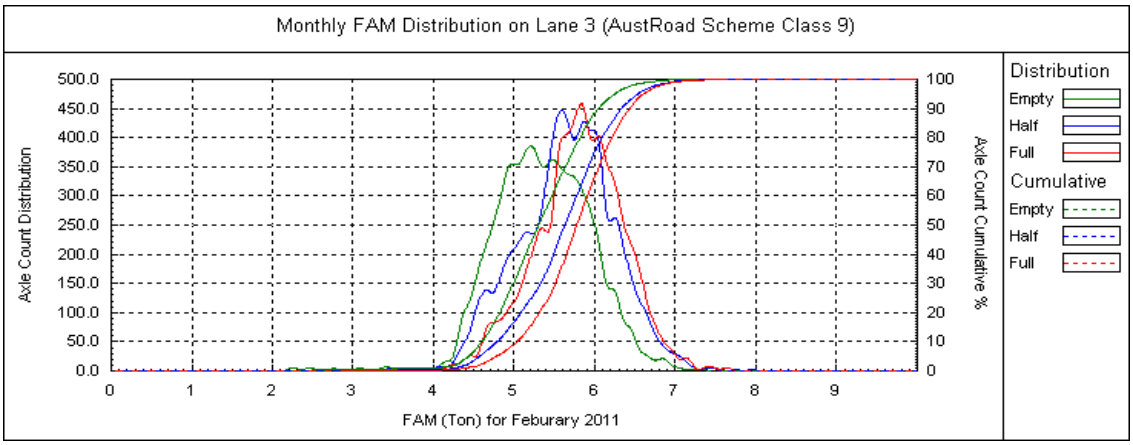

Figure 5 – FAM Frequency Distribution of Empty, Half and Fully loaded vehicles

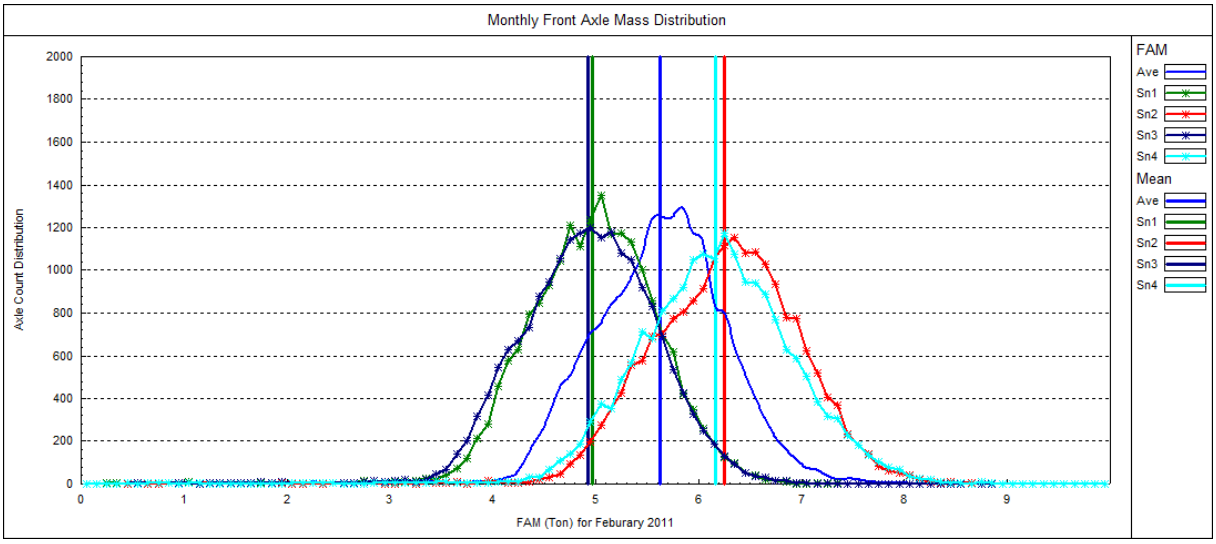

Figure 6 – FAM Frequency Distribution of individual WIM sensors (February 2010)

If one looks at the FAM from each individual sensor then one observes that the sensors in the left wheel track 'under weigh' while those in the right wheel track 'over weigh' (see Figure 6). This is partially due to the fact that the calibration was done using static weights measured on a flat surface while the WIM site is at a camber. The mean FAM for the left sensors is approximately 4.95, while that for the right is approximately 6.25 and the average is at 5.6 Ton. This is a variation of ±10% between what the left and right sensors weigh for the front axle. The authors have observed these differences at a number of other sites in South Africa, Australia and the US too. In the US where trucks drive on the 'other' side of the road (compared to South Africa or Australia) the effect is the same except that left and right swaps. Great care and an understanding of the FAM limits must thus be taken when selecting the FAM as a parameter to check long term 'calibration' drift of the WIM system.

Notice that the means of sensor 2 and 4 do not coincide by approximately 100kg. When the system was calibrated at the end of November 2010 these did coincide (Figure 3). In Figure 3 one can observe a slow long term drift. Also, in December 2010 the average FAM was 5.5 Ton. By February this had drifted to 5.6 Ton. The exact mechanism is not known but the drift can probably be attributed to settling of the WIM installation into the road base as the WIM was only installed in November 2010.

3.2 Truck-Tractor Verification/Calibration Method

With the advent of faster computers, better methods, such as the Truck-Tractor (TT) method, have been developed (by De Wet and others, Reference 1) that allows one to verify WIM systems more reliably and to correct for long term drift. The principle of this method is that the whole tractor weight of a common loaded truck is used as a tracking and correction method. Their research on trucks in South Africa had shown that the mean Truck Tractor weight of 'loaded' trucks for a large sample was 21.8 Ton. The correction method is based on a multi-pass process whereby the 'calibration' of the WIM data is adjusted by a factor k until the mean TT weight of the sample is 21.8 Ton. A summary of the method is as follows:

An *TT-Truck* is defined as a heavy vehicle with 6 or 7 axles with an axle spacing of 2.9 – 3.9 m between the 1st and 2nd axle, 1.2 – 1.6 m between the 2nd and 3rd axle and 4.5 – 9.0 m between the 3rd and 4th axle.

A *Selected TT-Truck* is defined as an *TT-Truck* with average axle mass between 6.5 t and 8.5 Ton.
The average truck-tractor mass of Selected TT-Trucks is used for calibration purposes. The target truck-tractor mass is 21.8 Ton.

The method described by De Wet is for 6 and 7 axle trucks (123 and 1222) that are common in South Africa, but one finds, not so common in Australia and the US. One has to adapt these methods. The authors adapted the method to use a 6 axle articulated trucks that are common in Australia. Assuming that the target TT weight should be the same for left and right sensors the authors determined a TT-weight for the loaded trucks such that after the correction the mean Front Axle Mass (FAM) of all 6 axle articulated trucks for the month of December 2010 (shortly after an on-site calibration) was still 5.5 Ton (see previous section). The target TT weight was then found to be approximately 22.5 Ton. A summary of the method used by the authors is as follows:

An *TT-Truck* is defined as a heavy 6 axle articulated vehicle with an axle spacing of 3.2 – 5.3 m between the 1st and 2nd axle, 1.2 – 1.6 m between the 2nd and 3rd axle and 4.5 – 9.0 m between the 3rd and 4th axle.
A *Selected TT-Truck* is defined as an *TT-Truck* with average axle mass between 6.0 Ton and 8.5 Ton.
The average truck-tractor mass of Selected TT-Trucks is used for calibration purposes. The target truck-tractor mass is 22.5 Ton.

In Figure 7 the frequency distribution of the Truck Tractor portion of the trucks (TT Truck-All) is plotted as well as the selected loaded trucks (TT Truck-SEL) for the raw data in February 2011. In Figure 8 the same information is plotted after post calibration was applied to the February 2011 data.

The same method was applied to the data from December 2010 and January 2011. The result of the post-calibration is that the slow drift has been removed (see Figure 9 and compare to Figure 3), but the average FAM of the left sensors is still below that of the right sensors. Further investigations are required to explain this.

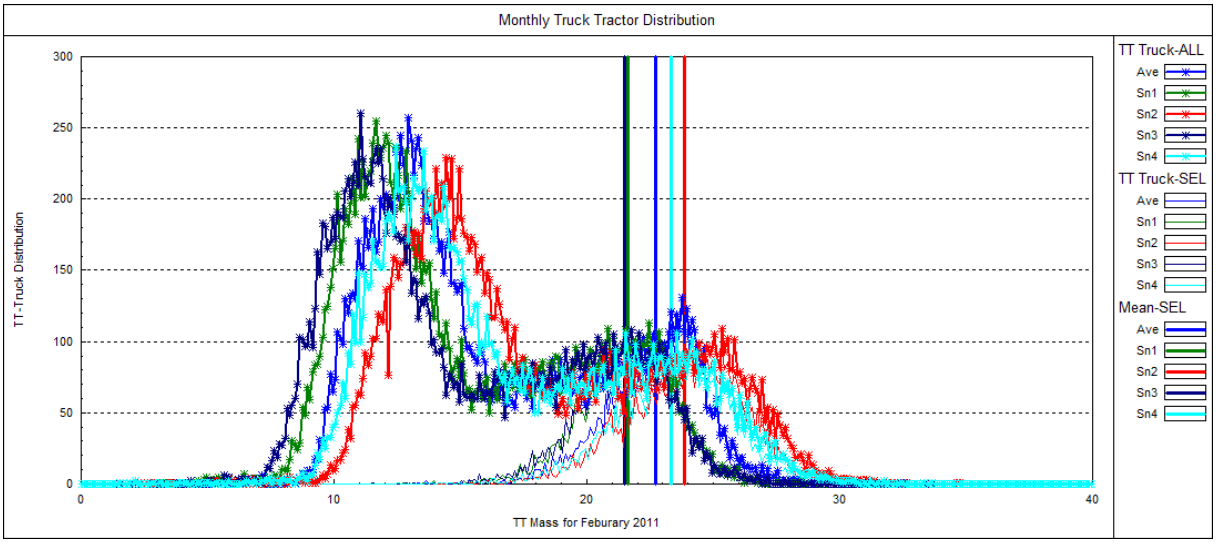

Figure 7 – TT Truck Frequency Distribution of the WIM sensors prior post calibration

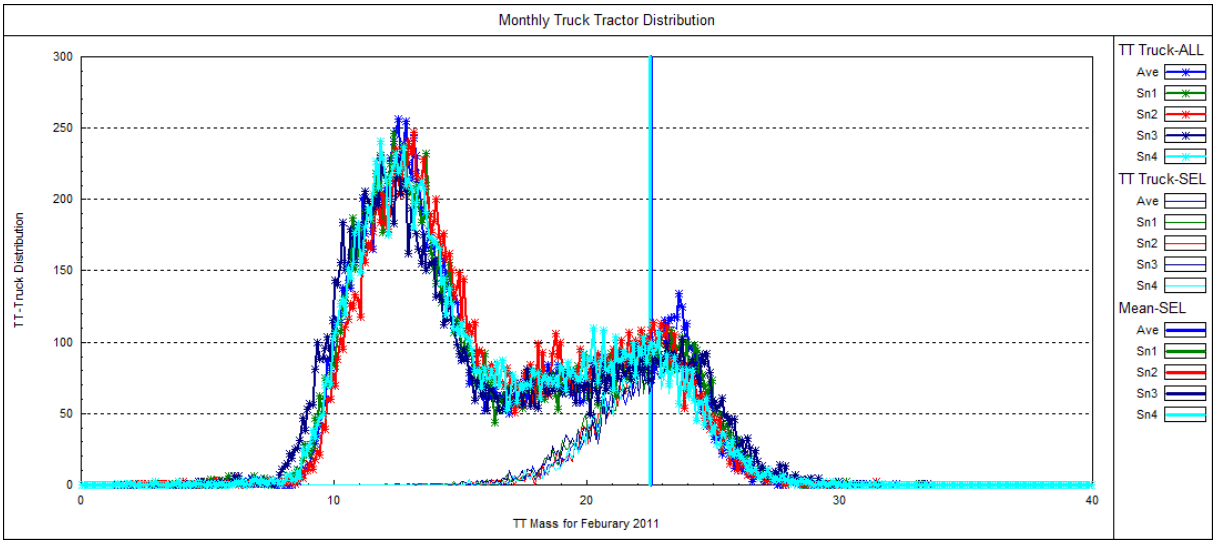

Figure 8 – TT Truck Frequency Distribution of the WIM sensors after post calibration

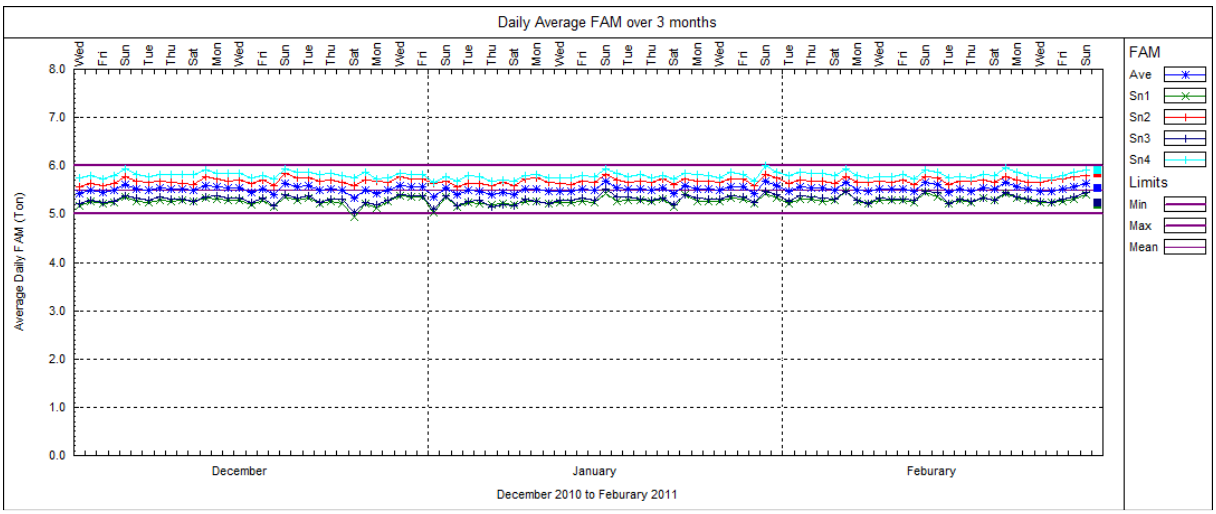

Figure 9 – Average daily FAM after post calibration using the TT Method

Statistical post calibration using the TT-Method can remove slow drifts in the calibration of a WIM system but further research is required on the TT target mass and on how to deal with sensor in the left and right wheel tracks for WIM sites that have a camber, as most sites do.

Statistical post calibration procedures hold great promise in ensuring the stability of WIM systems over a prolonged period of time.

4 EFFECT OF LOGGER RESOLUTION ON AXLE DISTRIBUTIONS

When binning axle weight data to generate an axle distribution from data produced by WIM equipment one sometimes finds that certain weight bins are more likely to occur that others. In Figure 10 an example is given of axle weights from individual vehicles binned into 100 kg bins i.e. counts of how many axles fall into a given 100 kg section. The binned count data is then normalized with the total count and expressed as a percentage.

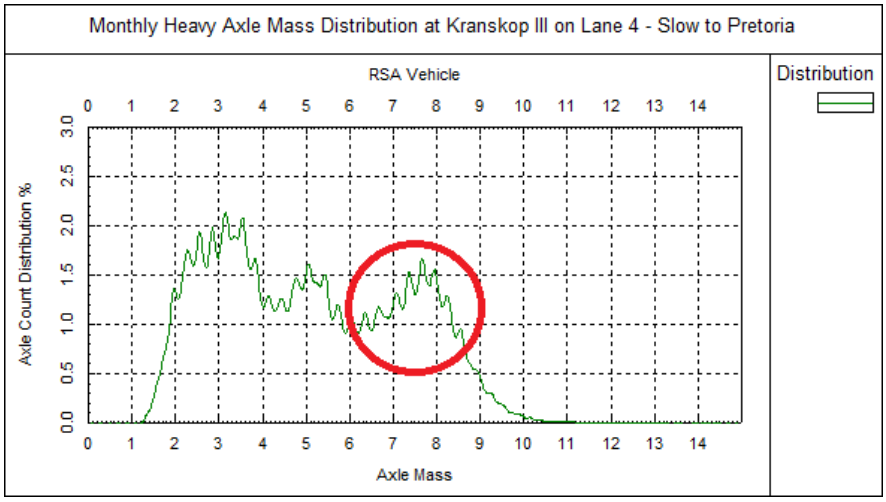

Figure 10 - A typical axle mass distribution produced by a logger with digital resolution of ≈ 15 kg

If one zooms in on the section in the red circle then one gets the result shown in Figure 11, which clearly shows that certain weights seem to occur more than others. There seems to be a bias.

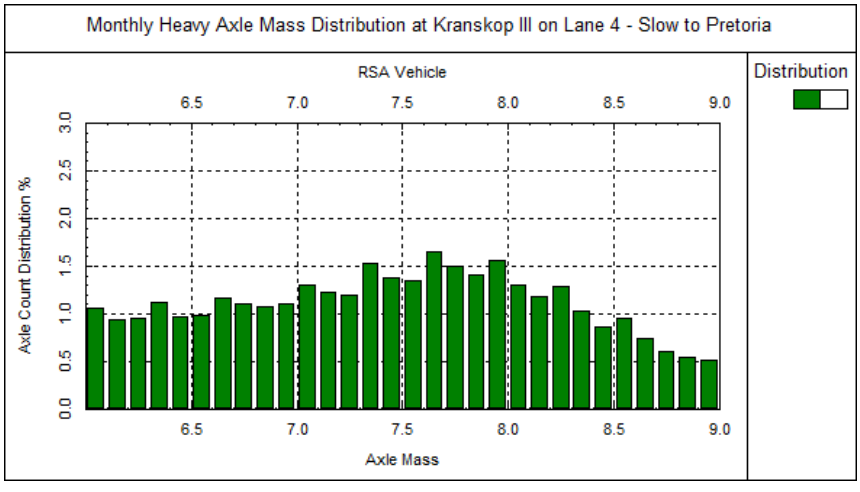

Figure 11 - A portion of the axle distribution showing bias

4.1 DIGITAL LOGGER RESOLUTION LIMIT

Most HSWIM loggers have digital resolutions limits ranging from 10 kg to 50 kg. When data is expected to be accurate to within 100 kg then such a resolution limit is quite acceptable. For a given logger type the actual resolution limit also depends on the individual sensor sensitivity and on the actual installation itself i.e. whether the final resolution is 20 kg or 16 kg for example.

Take an example of a distribution where one has 1 axle per kg in the range 4 to 10 ton. One plots this normalizing each 1 kg bin to 1 Ton i.e. 1000 axles per ton (Figure 13). One then applies this to a logger with 16 kg digital resolution (Figure 13). Finally one takes the raw logger data and bins this into 100 kg bins (Figure 14).

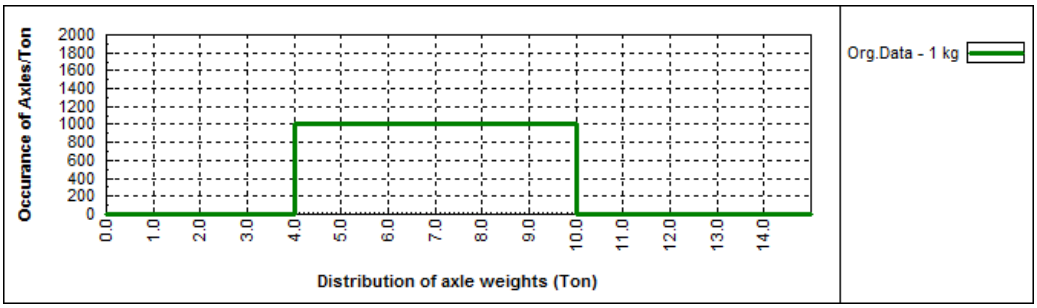

Figure 12 - Theoretical Axle Weight Distribution at 1 kg resolution

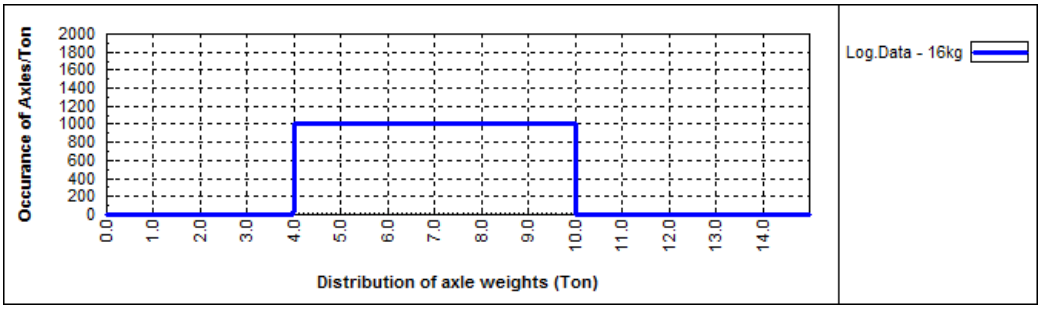

Figure 13 - Theoretical Axle Weight Distribution as recorded by logger at 16 kg resolution

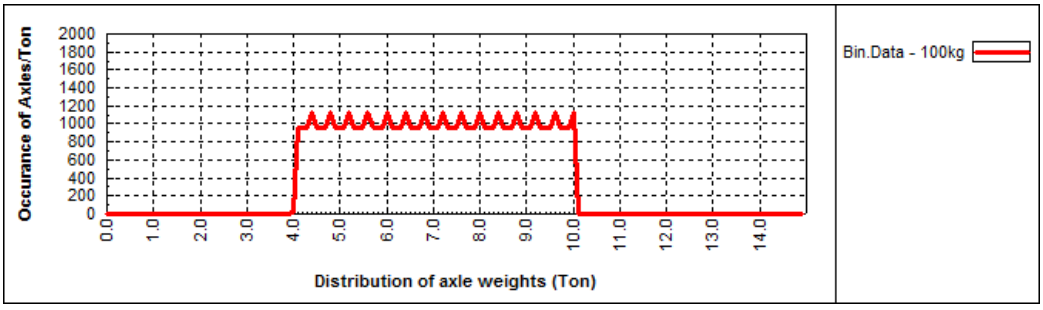

Figure 14 - Theoretical Axle Weight Distribution as recorded by logger and binned at 100 kg resolution

One can clearly see that the distribution is no longer uniform and favours certain weight bins. Had the logger resolution been exactly 20 kg then the binning into a 100 kg bins would have reproduced the almost distribution exactly (see Figure 15).

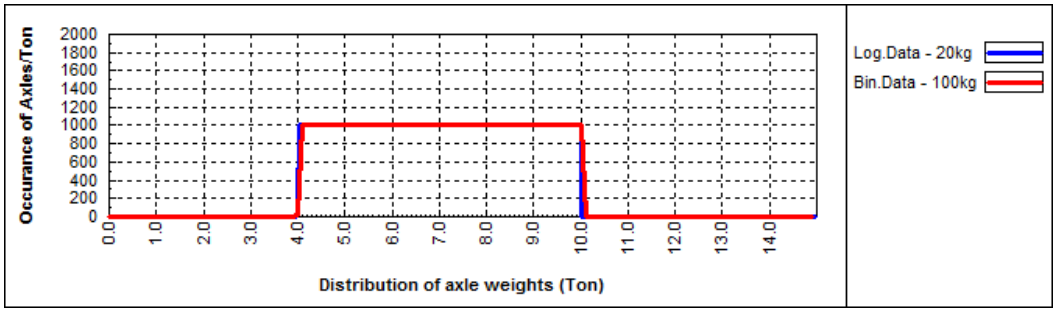

Figure 15 - Axle Weight Distribution recorded by logger at 20 kg and binned at 100 kg resolution

How does one solve this problem?

4.2 METHOD TO IMPROVE THE BINNING OF DATA BY MAKING USE OF THE REESOLUTION

If the logger resolution (R) is known, then one can make use of this fact. If one recognizes that the raw axle mass data (M say) as recoded by a logger can be anywhere between M and M+R (excluding M+R) then one can use a random number generator that generates a unity probability distribution x between $0 \leq x < 1$ on each raw axle mass as it is binned i.e.

$$M_B = M + xR$$

where M_B is the mass that is then binned.

Applying this idea to the 16 kg theoretical example of Figure 14 (16 kg resolution) one gets the result as shown in Figure 16.

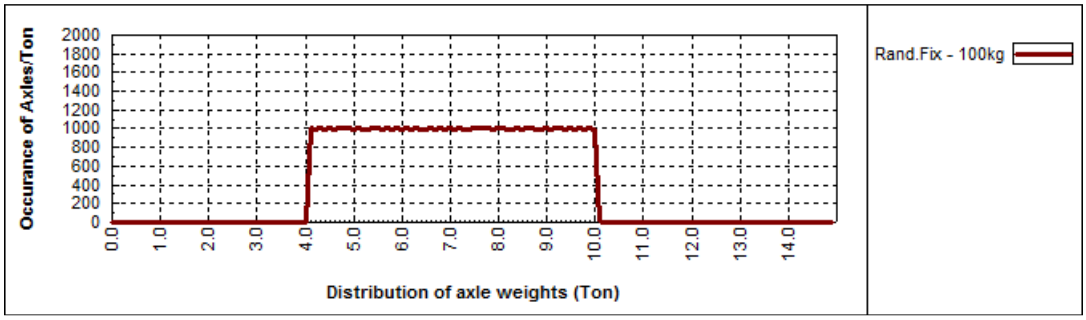

Figure 16 – Theoretical Axle Weight Distribution after randomization

In the FHWA Card W and the SANRAL RSA Version 1.0 and Version 2.0 traffic data formats the weight data is presented to the nearest 100 kg. For such data this methodology cannot be applied because the actual logger weight data needs to be known to a resolution of at least the digital resolution limit of the logger or better. In the SANRAL RSA Version 3.0 format currently being proposed, however, the weights can be recorded to the nearest kg and. This format also makes provision to record the actual logger resolution. On such a data set this 'smoothing' technique can then be applied.

4.3 RESULTS AFTER ADJUSTMENT

When binning axle weight data to generate axle weight distributions the weight resolution limit of a logger can cause unwanted bias. This can be removed by using a randomization technique based on the resolution of the logger.

By applying this randomization technique to the same set of data as presented in the introduction one can clearly see that although the resulting distribution is not perfectly smooth it is greatly improved.

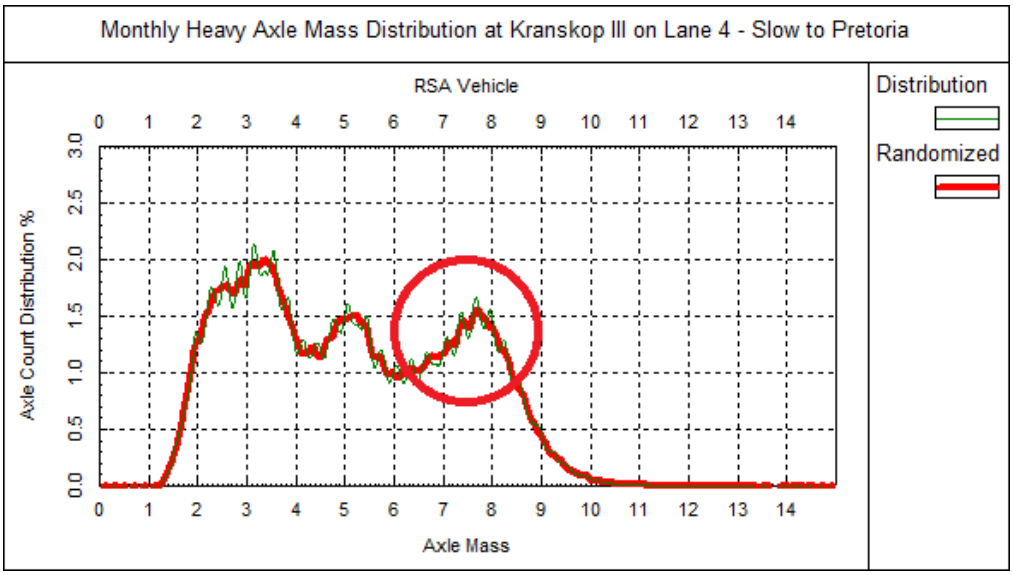

Figure 17 - Axle Mass Distribution smoothed by using resolution randomization

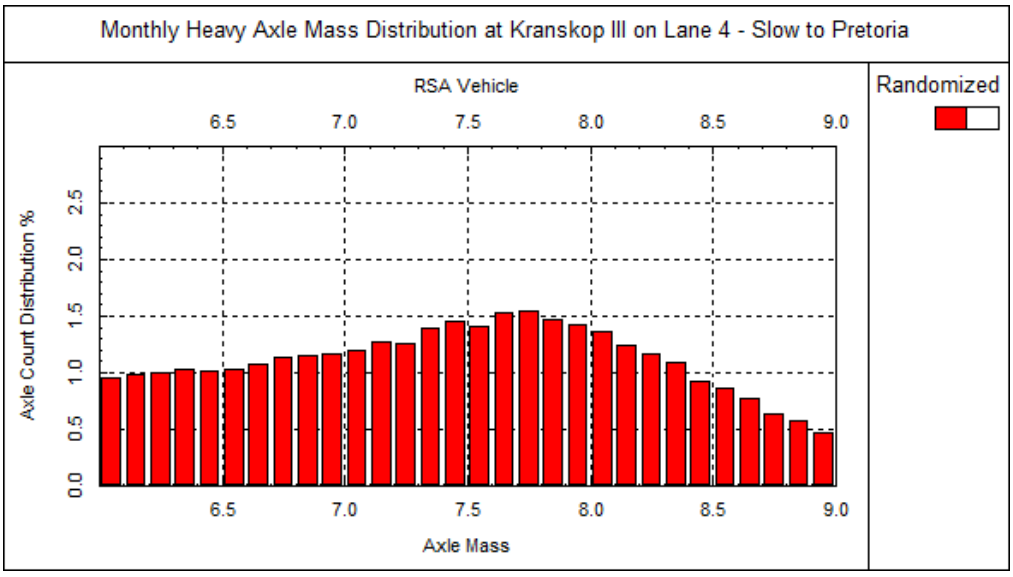

Figure 18 - A portion of the axle mass distribution showing virtually no bias

5 CONCLUSION

Providing the WIM data logger ensures the integrity of actual recorded raw data and no physical modification is done at a WIM site, modern statistical verification and calibration techniques can greatly improve the stability of WIM data. On-going research is still required to determine the role local traffic mix and vehicle configuration in the population plays in selecting reference values.

Regarding binning of Axle Mass data, one should be aware of the intrinsic resolution of the data logging system.

6 REFERENCES

1. De Wet, DPG and Slavik, MM. (2008), Macroscopic WIM Calibration. *Proceedings of the International Conference on Heavy Vehicles HVParis*, 19-22 May 2008, Paris, France.

2. ASTM USA. (2002), Standard Specification for Highway Weigh-In-Motion (WIM) Systems with User Requirements and Test Methods. Designation E 1318-02, ASTM Committee 17. Published February 2002.

3. U.S. Department of Transportation Federal Highway Administration. (2001), Traffic Monitoring Guide, Section 6 Traffic Monitoring Data Formats. Published May 2001.

4. South African National Roads Agency Limited. (2008), Standard Specifications for Traffic Data Collection Services Volume 4. Published May 2008.

5. U.S. Department of Transportation Federal Highway Administration. (2010) WIM Data Analyst's Manual. Publication No. FHWA-IF-10-018. Published June 2010.