Quality Control of Weigh-in-Motion Systems Using Statistical Process Control

CHUNHUA HAN, WILLIAM T. BOYD, AND MICHAEL M. MARTI

Weigh-in-motion (WIM) data collected from continuously operating systems at test sections were reviewed. Data were processed using an extensive software quality check program, which was developed in joint effort with regions, agencies, and contractors. Data beyond specification limits are flagged as unusable, but this does not correct for system problems. One proposed methodology enhances the existing quality check program by monitoring data patterns. It corrects for calibration drift in distributions of gross vehicle weight of tractor semitrailers (loaded and unloaded). This method is labor-intensive, however. It is proposed that statistical process control be integrated with the existing quality check program. The review of WIM system data is summarized, and a statistical quality control methodology to facilitate post-processing of data flagged as unusable due to calibration drift is presented.

The weigh-in-motion (WIM) system is an important tool for estimating traffic load, weight enforcement, and flow characterization. Although the WIM is complex, its use has increased. According to German and Copeland (*1*), the Strategic Highway Research Program (SHRP) and Long-Term Pavement Performance (LTPP) program study of FHWA has promoted the nationwide use of more than 250 continuous WIM systems since 1989.

Most WIM sites associated with FHWA-LTPP study sections were installed and operational by 1991. Regional offices, government agencies, and technical support contractors jointly developed the quality assurance and quality control (QA/QC) program.

With equipment, there are essentially two processes in the QA/QC program. One process consists of equipment producing products by specifications, where common cause-related errors may occur. The second process is mixed, where monitoring equipment measures similar products on the basis of vendor specifications. In addition, the second process demonstrates special cause-related quality problems, or variation not inherent to the process. Monitoring equipment is controlled statistically in the mixed process, rather than manufacturing equipment.

This paper briefly describes data from selected WIM sites. The proposed methodology of quality assurance could be integrated into existing practices but is not currently used. Applying control charts to WIM data is not a new concept. This paper shows a way to detect anomalies due to a change in load configuration, quantify calibration drift, and recover data for post-processing.

BACKGROUND

The WIM is composed of a core of sensors, with a sensor housing for a structure, pavement surface as a work frame, climate as an environment, and vehicles as objects. The objective is to interpret vehicle axle weight as a static measurement.

The significance of obtaining static weight comes from road tests such as the AASHO Road Test, in which pavement performance was related to static weight rather than to actual dynamic force.

The SHRP-LTPP (*2*) study was initiated in 1988. Test sections were laid out in the existing highway system for about 800 in-service sections of general pavement studies (GPS) and specific pavement studies (SPS). SPS projects are studies of specific variables, such as new construction, with several sections per project.

Relating dynamic traffic loads to pavement performance is a new area. However, the relationship between static loads and pavement performance is based on developed concepts, techniques, and empirical equations. The task becomes one of inferring accurate static axle or axle configuration loads from dynamic measurements.

Absolute calibration is used to infer static loads. The popularly used calibration scheme is based on static loads being linearly dependent on dynamic loads, with zero intercept. This assumption was validated by Wyman (*3*) who identified a linear relation between WIM and static truck gross weights. Consequently, a constant calibration factor is determined by adjusting the average of dynamic readings from a known static weight to the static weight. This ignores the population of dynamic loads, since a constant calibration factor can only shift the entire population. However, it is impractical to separate the dynamic population of each actual static load from a traffic stream. A constant factor is used on average that the sum of dynamic loads is close to the sum of static weight.

Calibration is one factor affecting WIM data quality. Lee (*4*) indicates that four groups of factors may cause variation, or data error: dynamic, equipment, signal interpretation, and static reference. WIM systems measure static weight plus random and system error. Random errors occur in every process. They may be caused by vehicle factors, such as loads, suspension, and tires. They cannot be defined in the current WIM system. But some vehicle factors, such as vehicle type and axle location, are defined in the WIM system by axle configuration. They are system error factors.

System errors are changes caused by factors such as nonclassification or misclassification. The software may misclassify or not classify vehicles at all. Errors must be detected and corrected by post-processing with a debugged classification algorithm. Other system errors, such as calibration drift, require advanced techniques to detect and correct. Loads from certain axles of particular truck classes show low variation, making it necessary to use the advanced techniques. System error factors such as vehicle type and axle location can be used for review of their load statistics. Statistical process control (SPC) is one tool for reviewing WIM data, although not a new concept.

Braun Intertec Corporation, 6875 Washington Avenue South, P.O. Box 39108, Minneapolis, Minn. 55439.

WIM SITES

Six WIM sites in the North Central Region of FHWA-LTPP were selected on the basis of data availability, pavement type, and sensor type. The availability of 2-year continuous WIM data was the primary criterion for selection, pavement type was the secondary criterion, and bending plate and piezo WIM sensor types were the third criterion. Table 1 presents site characteristics in route type, surrounding development, construction year, and pavement type (PCC-portland cement concrete; AC-asphalt concrete). Traffic, truck volume (inner lane), and sensor type are also included.

The calibration scheme for these sites includes initial (absolute) calibration using test trucks. Of these sites, two include a system that has an autocalibration feature, as indicated in Table 1. The autocalibration feature corrects a limited amount of calibration drift during WIM operation.

QUALITY CHECK AND QUALITY ASSURANCE OF WIM DATA

Quality check herein is defined as data-related. A wide scope of the quality check function is recognized, including check, format, chronological, arithmetic, consistency, and range checks. To meet quality check and quality assurance (QC/QA) and data goals, data are processed to check quality first. This is done using the regional traffic data base software. Besides errors in generic site description, file and data format, date and time sequence, and addition, errors in range for each measurement are detected and flagged. But these checks alone are not enough to ensure truth-in-data.

Except for range errors, all other errors are corrected with visual/program validation and manual and program editing at each regional office. Agencies verify, review, and coordinate with vendors.

The current regional QC/QA program is shown in Figure 1. Two essential functions are the quality audit against general errors (software bugs or human error) and the quality survey on program scope. These errors occur during data processing and can be corrected using the QC/QA program.

The most useful function of the QC/QA program is auditing vehicle classification. In WIM data, each vehicle is classified as a six-digit code. For example, 332000 represents a five-axle tractor-semitrailer (semi). The volume of each class is plotted. Figure 2 shows classification distribution for 1 month of WIM data, in which 2,000 vehicles were unclassified (000000). The data need reprocessing with a different classification algorithm for those vehicles.

The quality survey at the regional level gives a better understanding of unexpected threats, from either software or equipment upgrading. Equipment change log forms are used to survey upgrade activities. The survey helps identify any group of data needing investigation.

Another element is the range check. Attribute names and valid ranges are available (*5*). Range checks provide primary data quality assurance, because many range errors may signal equipment malfunction or calibration drift. The range check ensures data quality within predetermined range boundaries. Problems inside the range are difficult to detect. Most of the distribution of traffic mix for axle weight and vehicle class are already within the range and deemed acceptable by the software.

The range check is enhanced by a plot of the load distribution of gross weights. Each WIM site has a unique statistical profile in load distribution. Dahlin (*6*) suggests focusing on the gross weight of semis if that population dominates. Weight data inside the range can be checked for equipment-related error or malfunction. Figure 3 shows the weekly load distribution of semi gross weights. The peaks for empty trucks for the second and third weeks show a drift to the right by one 17.8-kN (4-kip) bin width when compared with the first week.

TABLE 1 Selected WIM sites

ID	Route	Development	Construction Year	Pavement Type	Sensor	AADT*	Initial Calibration Test Truck	Auto-Calibration
182008	US 27	Rural	1980	AC	Bend Plate	4537	3 axle	No
183002	US 41	Rural	1976	PCC	Bend Plate	3405	3 axle	No
271016	ST 71	Rural	1976	AC	Bend Plate	1260	5 axle	Yes
274037	I 35E	Suburban	1981	PCC	Bend Plate	10,000	Traffic stream	Yes
295473	I 70	Rural	1989	AC	Piezo	8312	5 axle	No
383006	US 2	Rural	1986	PCC	Piezo	1025		No

* AADT in 1989 as estimated by agencies; installation of WIM was 1991 for the sites shown.

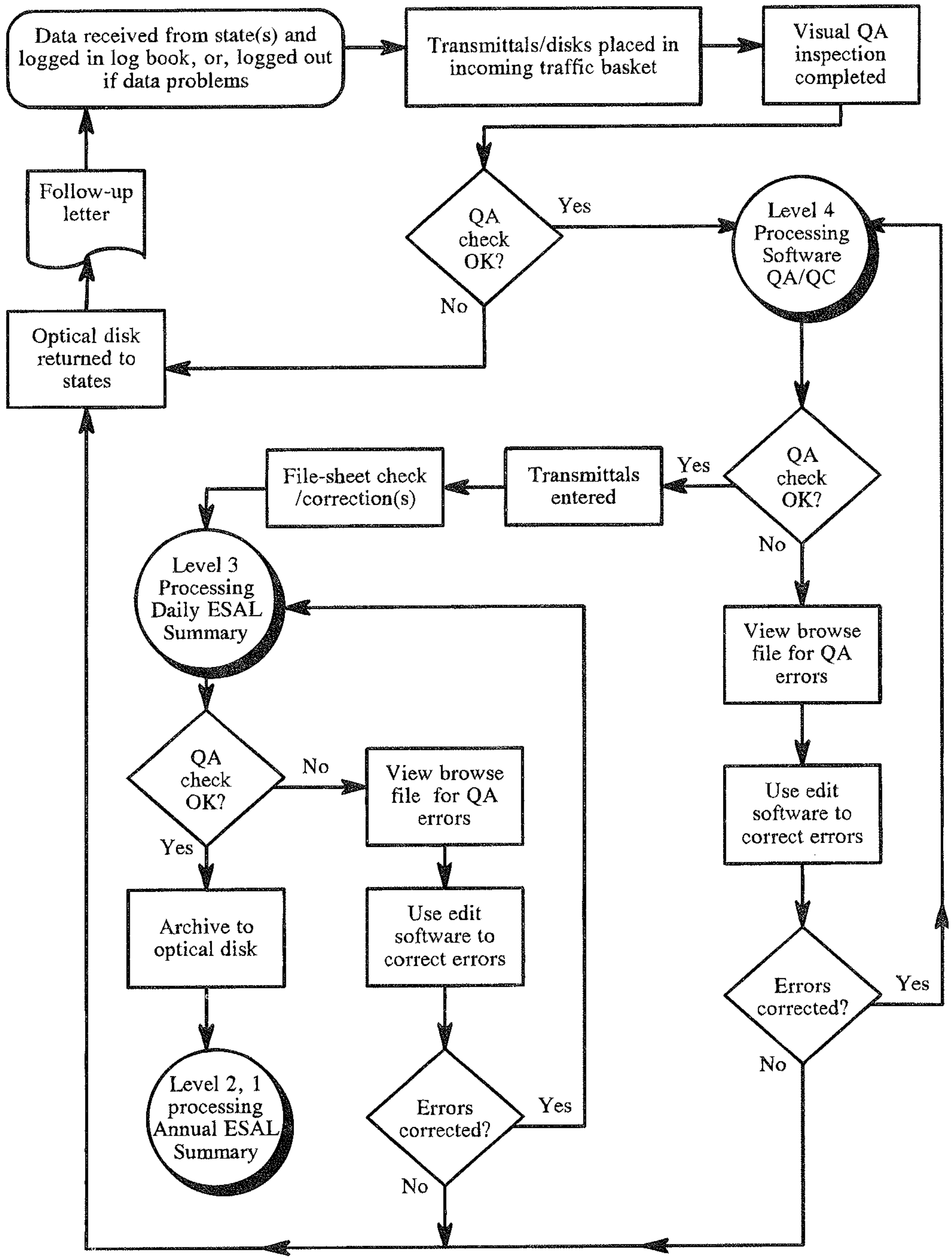

FIGURE 1 Traffic data QC/QA Program in North Central Region of FHWA-LTPP.

This method requires many samples and extensive experience over time with WIM systems to interpret plotted results. A set of rules was developed to assist visual judgment by technical assistance contractors. Another problem with the distribution check is in selecting bin width. Bin width is related to the number of trucks in each bin. For a low-volume road, too few trucks exist for a given load bin. A longer period or wider load bin is redefined for reconstructing load distribution. Often a longer period distorts the detection of problems, and a wider bin distorts calibration drift. Most important, only qualitative judgment in calibration drift can be made, and quantification of calibration drift is often needed to adjust data distorted by the drift. A simpler, more objective method is needed to determine data and system quality and to improve efficiency and automate the checking process. At the same time, quality control aspects must be preserved in the QC/QA program.

QUALITY ASSURANCE AND QUALITY CONTROL OF WIM SYSTEMS

Juran (*7*) defines quality assurance as planned actions necessary to provide confidence that a product or service will be able to satisfy

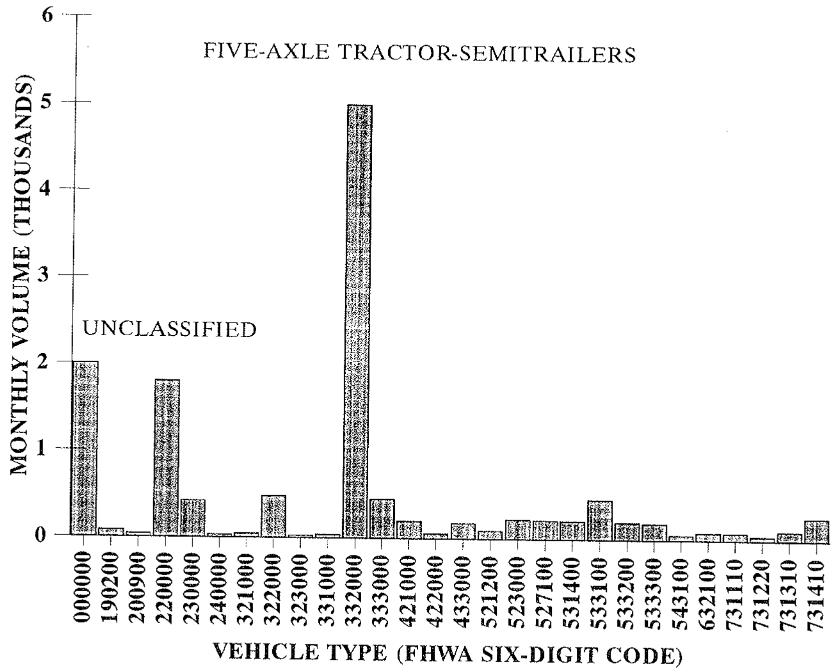

FIGURE 2 **Vehicle classification distribution.**

quality requirements. He defines quality control as operational techniques employed to meet those requirements. QA/QC is also defined in the context of this paper as being system-related. For a WIM system, the system-related quality problems can be a slight drift in calibration, a major drift in calibration, or a system malfunction. The last two are more serious quality concerns, but even the first problem seriously affects system integrity, as a relatively small error is magnified when weight is converted to 80-kN (18-kip) equivalent single-axle loads (ESALs), according to the fourth-power relation. The existing QC/QA program is unable to detect and control drift or malfunctions.

A WIM system interacts with sensors, pavement, vehicles, weather, software, communication, and humans. Vendors typically focus on sensors and espouse specification limits for specific qual-

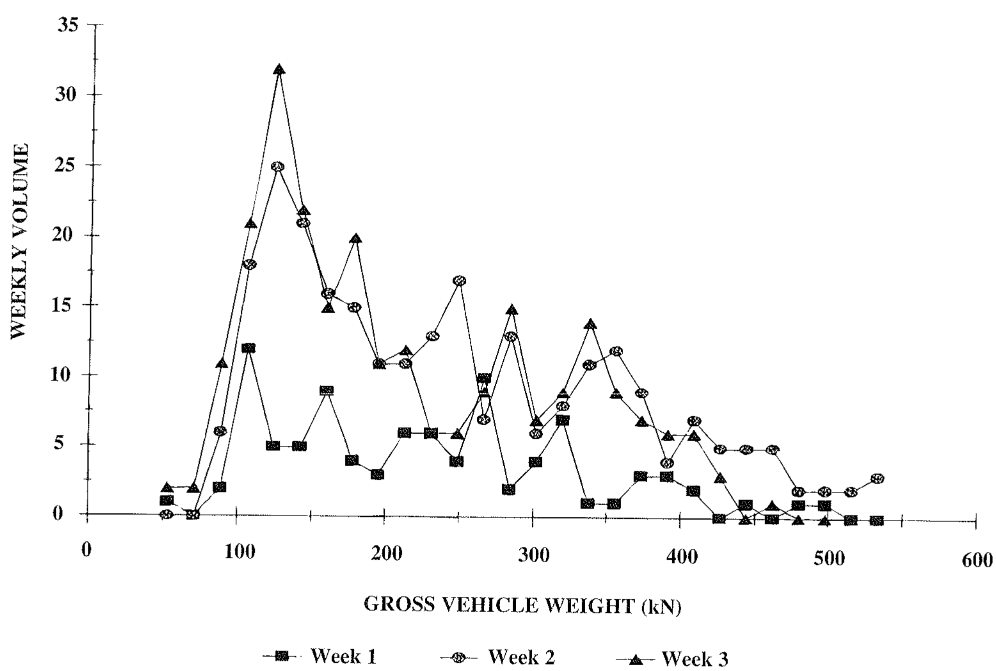

FIGURE 3 **Vehicle gross weight distribution.**

ity characteristics. Therefore, the major concern herein is process performance as it affects overall system performance. For this purpose, SPC is meaningful when control limits are applied for a WIM characteristic. Most important, SPC helps identify the presence or absence of special cause variation in a system so that process review can be performed and data acceptance can be determined.

Static axle or axle group weights are essential in the load equivalency concept. The static loads (only measurable traffic loads) were considered in the development of load equivalency factors. The dynamic effect and dynamic history are inlaid in the empirical equation relating static loads to the ESAL (for example, the AASHTO equation). When extracted from actual dynamic loads, the static loads should be interpreted with the calibration of WIM systems. Since dynamic effect varies with time, calibration should be updated to interpret the static loads correctly.

WIM equipment may produce a large number of individual data records that do not meet specified tolerances. Since these tolerances fall within the upper and lower control limits of the control chart, the WIM is still operating in statistical control. In other words, the equipment is statistically in control but individual data do not meet specification limits. This may be due to a drift in calibration, or the equipment may have an unacceptably high level of common-cause variations, such as those induced by pavement, climate, and equipment conditions.

Normal Assumption and Validation

There are two main components in a traffic load mix: (*a*) actual static load or load level and (*b*) dynamic variation around each load level. This is true even when a particular type of truck or a specific axle is considered. It is not possible to separate these two components. However, the effect of mixing can be minimized by categorizing load levels at a stable axle location for a dominating truck type. The experience of Dahlin (*6*) has shown that a unique constant mean of front-axle weights for a group of five-axle semitrailers exists. Three groups of this truck type are defined: less than 142 kN (32 kips), from 147 to 311 kN (33 to 70 kips), and more than 311 kN (70 kips), which corresponds to empty, partially loaded, and loaded trucks, respectively.

With the groups defined previously, a normal or approximately normal distribution of sample mean for each group is assumed. According to the central limit theorem, the normal model is useful to describe the output of a complex system, where the system output is determined by a combination of outputs from a large number of subsystems. The assumption implies that only common causes exist in each group. The assumption is validated using the analytical test for normality (for example, the *W*-test for a large sample) of Shapiro and Wilk (*8*). Figure 4 shows a Rankit plot of the front-axle weights of empty semis. The *i*th Rankit is the expected value of the *i*th-order statistic for the sample, assuming the sample is from a normal distribution.

If the sample conforms to a normal distribution, a plot of Rankit points against the order statistics should result in a straight line. The slope of the straight line is defined as the test statistic. With a Type I error of 0.05, $W_{0.05} = 0.931$. Since the slope is larger than this, it is concluded that there is no evidence to reject the hypothesis of normality. Table 2 gives the *W*-test results for three typical semi truck groups and five sample groups. Each sample group has a total of 30 samples.

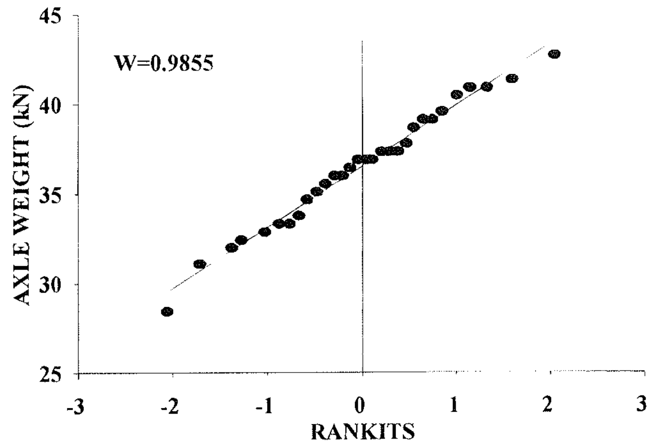

FIGURE 4 Approximate Shapiro and Wilk normality statistics (30 individual samples).

Rational Sampling

The validation of the normal distribution is based on a subgroup size of 30 to better illustrate normality in Figure 4. On the basis of the central limit theorem, as the size increases, the sampling distribution of subgroup means will approach normality. Juran (*7*) points out that five subgroups provide a good approximate normal distribution of subgroup means. Furthermore, it is believed that the subgroup size of five is sensitive enough to detect special or assignable cause and is small enough to be operationally feasible.

Traffic loads are estimated on a daily basis. If any equipment failure occurs in the course of a day, the WIM data collected for that day are not used. Daily sampling also aids process stability as the WIM system is not expected to exhibit erratic behavior in that period. Furthermore, most process events such as calibration drift and equipment malfunction occur daily. On a daily basis, homogeneous and consecutive sampling is adopted. The first five consecutive samples in a day were selected to form a subgroup. It is noted that for a low-volume road, the subgroup size may be reduced if the normality assumption remains valid.

Control Charts

The control chart originated from the work of Shewhart in the late 1920s (*7*). The information from each rationally sampled subgroup is reviewed to determine if a disturbance outside the process norm exists. The hypothesis that only common cause variation exists is tested in the control chart. Calibration drift is a special cause that grossly disturbs the WIM. If so, the hypothesis will be rejected. Two system characteristics are used for hypothesis testing, and both are statistical attributes: the mean level of the system and the amount of variation in the system. Therefore, the control chart provides statistical process control.

When a WIM system becomes operational after initial field adjustment and absolute calibration, it should produce the initial absolute reference outputs. These outputs can be sampled into subgroups that then are used to determine centerline and control limits for the construction of the control chart. If the initial calibration is off, centerline and control limits could be derived from historical

TABLE 2 Results of *W*-Test for Normality

Gross Weight	< 142 kN	147-311 kN	>311 kN	$W_{0.05}$	Test Result
Sample 1	0.971	0.975	0.969	0.931	Accept
Sample 2	0.986	0.942	0.971	0.931	Accept
Sample 3	0.967	0.958	0.957	0.931	Accept
Sample 4	0.960	0.969	0.940	0.931	Accept
Sample 5	0.966	0.979	0.979	0.931	Accept

data or experience with the process. The number of subgroups used for constructing centerline and control limits can be determined if the standard deviation of the sample is representative of the population as a whole (*9*). If 1 percent of standard error of the mean value for 95 percent confidence level is specified, the required sample, *n*, is given by

$$n = \left(\frac{\sigma_x}{0.0196\bar{X}} \right)^2 \tag{1}$$

where σ_x is the standard deviation and $\bar{X}$ is the sample mean. Using Equation 1, calculate a sample number of 150 for a mean of 44.4 kN (10 kips) and a standard deviation of 5.3 kN (1.2 kips). So, in this paper 30 subgroups (days) of five samples a day are used to construct centerline and upper and lower control limits on the basis of the absolute calibration.

Data from the six WIM sites were used to construct the control charts. This paper assumes that the absolute calibration, or initial calibration made using a static known weight, is accurate. Statistix 4.1 software (*10*) was used to construct the control charts in Figures 5 and 6 and the Rankit plot in Figure 4. (UCL represents the upper control limit and LCL represents the lower control limit in Figures 5 and 6.)

Figure 5 is the $\bar{X}$ chart and Figure 6 is the *R* chart for the empty five-axle semitrailer truck group. An autocalibration algorithm was included in the system. The $\bar{X}$ chart traces the change in successive subgroup mean values, and the *R* chart indicates change in variation in the range of consecutive data points. When two of three means of three truck categories are beyond the tolerance of their expected values, the calibration factor is adjusted through autocalibration. To interpret the patterns in these control charts, begin with the *R* chart. The *R* chart is very important because it signals equipment malfunction, which often results in a massive, uncontrolled variation. As seen in Figure 6, no unusual patterns are evident in this *R* chart. Data passed all eight tests for special causes (*11*) over a 2-year period, with one exception. Two points occurred above the upper control limit, both of which were examined. Two light five-axle semitrailer trucks were identified as being outside the semi group. In this case, the two extreme points can be ignored.

The $\bar{X}$ chart is used to identify and remove extreme cases (outlier) where static load has changed because of a different load configuration. As shown in Figure 5, the system failed some tests for special causes. Specifically, Test 1, in which a point occurs outside

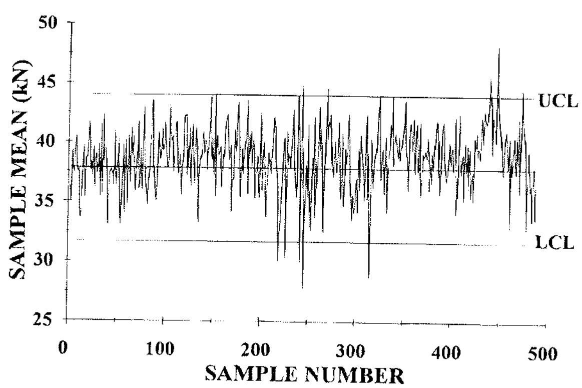

FIGURE 5 $\bar{X}$ **chart for empty five-axle semitrailers (gross weight < 142 kN) measured by bending plate WIM in PCC pavement on I-35E.**

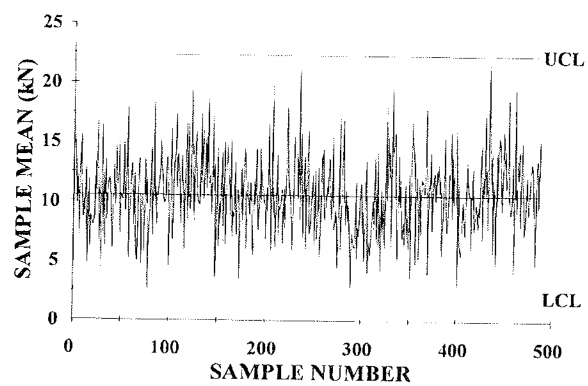

FIGURE 6 *R* **chart for empty five-axle semitrailers (gross weight < 142 kN) measured by bending plate WIM in PCC pavement on I-35E.**

the three-sigma limit from the mean, combines with other symptoms; this signals a shifting rather than stratification. Other indicators of shifting are the occurrences of two of three points in a row that are two to three sigma from the centerline and the condition in which four of five points in a row are greater than one standard deviation from the centerline. These conditions indicate a mean shifting in the system.

Overall, autocalibration is responsible for the trend shifting in mean values. When these shift warnings occurred, the original data were examined and its quality survey was reviewed. As a result, two data gaps were identified. One occurred when Sensors 1 and 2 of the bending plate were replaced and the other was unclear. The replacement of sensors resulted in the higher mean values when they were first put into operation. The mean values are adjusted continuously to the initial centerline as the result of autocalibration.

It is interesting to note the control charts showed no stratification phenomenon. They all passed rule tests for stratification. This means 8 to 15 successive subgroup means are not within two to three standard deviations from the centerline.

When the subgroup means are beyond the control limits, the process is statistically unstable or not within statistical control. If the mean beyond the limits is isolated, it signals that a special cause or an anomalous value has been encountered and should be deleted. Only in the control process does the rejection of data records with errors caused by special causes make remaining data records within statistical control. After elimination of such data, the process is in control with some other test failures. It also is noteworthy that a series of means run consecutively out of the limits and cannot simply be deleted.

Fortunately, unlike the process of manufacturing equipment, the WIM system can be stabilized if the calibration drift is only a system error. The stabilization results in a new calibration factor whenever the daily sample mean runs out of the control limits. Unstable sample means induced by other special causes are eliminated from control charts before a new calibration factor is applied to stabilize the process.

The new relative calibration factor is proposed as follows:

$$\frac{\overline{\overline{X}} - \frac{3}{\sqrt{n}}\sigma_x \frac{\overline{R}}{E(R)}}{\overline{X}} \le \mathrm{CF}_c \le \frac{\overline{\overline{X}} + \frac{3}{\sqrt{n}}\sigma_x \frac{\overline{R}}{E(R)}}{\overline{X}} \qquad (2)$$

where $\overline{R}$ is the average range and $E(R)$ is the true range of samples of size n. Whenever $\overline{X}$ is beyond the limits, the relative calibration factor can be calculated and applied to no-go type data by using the limit value (for example, equal size in Equation 2). Therefore, the relative calibration factor, as determined from Equation 2, can be used to adjust the large calibration drifting. The process has been stabilized to be in statistical control.

Exponentially Weighted Moving Average

The $\overline{X}$ chart may not be very sensitive to detecting a shift in the system mean level, but it works well to remove extreme cases in which trucks with a different load configuration have passed over the WIM system. With an average range of 11 kN (2.5 kips) and a mean of 44.4 kN (10 kips), the shift within 6.2 kN (1.4 kips) can be detected in only two ways (for example, 14 percent relative to the mean using the $\overline{X}$ chart). The state of Minnesota has targeted 4 percent as a control limit, or about one sigma, according to Dahlin (*6*). Therefore, an additional tool such as exponentially weighted moving average (EWMA) to detect small changes within the WIM system and calibration drift is needed.

In EWMA charts, the sample means are smoothed by the use of a moving average of prior sample means, whereby each mean is assigned a weight. Most of the weighing is allocated to the current mean in an exponentially decaying fashion in the remote past means, according to Woodall and Adams (*12*).

The control limits are then expressed by

$$M_0 - 3\sigma_x\sqrt{\frac{r}{(2-r)n}} \le M_i \le M_0 + 3\sigma_x\sqrt{\frac{r}{(2-r)n}} \qquad (3)$$

where M_0 is the control mean (absolute calibration reference point) and r $(0 < r < 1)$ is the weighing factor. For a given n, the smaller the r-value is, the more effective the detection of smaller shifts in the mean is. The r-value can be selected by using the specified allowable shift (for example, 4 percent of the target mean):

$$r = \frac{2n\left(\frac{\Delta M}{3\sigma_x}\right)^2}{n\left(\frac{\Delta M}{3\sigma_x}\right)^2 + 1} \qquad (4)$$

where ΔM is the allowable shift. As an example with ΔM equal to 1.8 kN (0.4 kips) for M_0 of 44.4 kN (10 kips) and σ_x of 5.4 kN (1.2 kips), $r = 0.12$. With r determined using Equation 4, the relative calibration factor CF is proposed as

$$\frac{M_0 - \Delta M}{M_i} \le \mathrm{CF} \le \frac{M_0 + \Delta M}{M_i} \qquad (5)$$

Whenever the equal signs in Equation 5 are met for the ith subgroup, CF_M is determined and applied to the subsequent subgroup means.

Figure 7 shows the plot of the EWMA chart with $r = 0.11$ for the data shown after deletion. The r-value is determined using $\Delta M = 0.04$, $n = 5$, and $\sigma_x = 1.034$. As seen in Figure 7, the process is considered out of control and the CF must be applied. The CF is calculated and plotted in Figure 7, then applied. The result is also included in the figure, where the EWMA has been corrected. After either statistic recalibration, weight data are adjusted to the allowable calibration drift.

By contrast, some WIMs do not have an autocalibration feature. Figure 8 is an EWMA chart for the WIM system data for Section 182008 in Indiana. This WIM site does not have an autocalibration feature. However, the site may be fitted for autocalibration in the future. Looking at the EWMA plot, one can see that the system starts out within acceptable limits at the zero point on the x-axis. After the first month, the data shift upward.

To correct this, a calibration factor is applied. Figure 8 is a plot of the WIM data versus the calibration factor over time. At the first point on the x-axis, the calibration factor is 1.0 as the data at the same point are within the specification limits. Over time the cali-

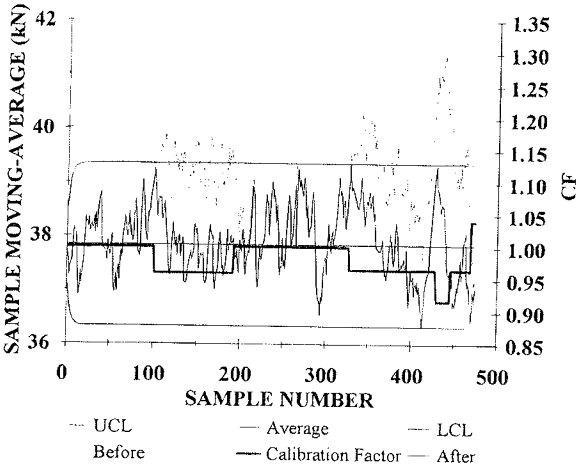

FIGURE 7 EWMA chart before and after application of calibration factor; empty five-axle semitrailers (gross weight < 142 kN) measured by bending plate WIM in PCC pavement on I-35E.

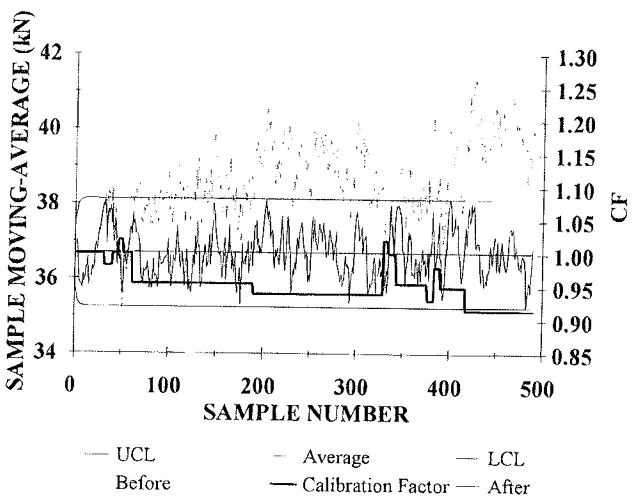

FIGURE 8 EWMA chart before and after calibration factor is applied for empty five-axle semitrailers (gross weight < 32 kips) measured by bending plate WIM in AC pavement on US-27.

bration drift increases, so the calibration factor changes with time as the data drift farther from the upper limit and corrects the data back into the specification limits. This is important as data would otherwise be considered unacceptable for ESAL calculations. SPC goes beyond the old quality method of inspecting and eliminating data that fall outside of accepted limits. Data that are flagged as unusable by the QC/QA checks can now be corrected and used. Finally, Figure 8 also shows the same data for Section 182008 after applying the new calibration factor, which adjusts the data for drift.

If calibration drift occurs beyond a maximum allowable drift of 4 percent, the data can be made usable. A change in procedures should also include a field verification of the WIM system to correct possible equipment problems.

CONCLUSIONS

The accomplishments of this paper are summarized as follows:

1. Automatic system recalibration (*6*) is effective in controlling the system mean shift. However, the shift caused by restarting the equipment is not properly controlled. Autocalibration, when used with SPC, aids autocalibration for small variations in climate, pavement surface, fleet, and equipment restarting by adjusting the calibration factor over time as the data change. The target mean and control limits should be determined carefully to allow for meaningful process control.
2. Recovery of otherwise unusable data and post-processing are possible using SPC. WIM data can be analyzed using the proposed procedure and recalibrated to limit the calibration drift to the allowable level.
3. The SPC procedures presented are simple and more effective in detecting and correcting system errors such as calibration drift to almost any level of a specified tolerance.
4. The proposed SPC procedures are particularly useful when weight data are collected on low-volume roads on which load distribution curves cannot be established.
5. The SPC procedures presented are flexible in developing relative calibration factors for a specific vehicle type or axle weight.
6. The proposed QA/QC program can screen WIM data, including volume and gross weight load distribution review, for diagnosis of equipment-related errors or sensor malfunction.
7. WIM data integrity is achievable without being labor-intensive.

ACKNOWLEDGMENTS

The data used in this paper originated from the FHWA-LTPP study. Permission granted by FHWA for using these data is gratefully acknowledged. The WIM data collection work described was performed by agencies and staff of the North Central Region coordination contractor, Braun Intertec Corporation. All agencies and staff performing tests are also acknowledged, including Philip Meraz of Iowa for his idea to post-process flagged data.

REFERENCES

1. German, J. L., and C. R. Copeland, Jr. *SHRP-LTPP Traffic Data Collection and Analysis: Five-Year Report* (W. O. Hadley, ed.), Texas Research and Development Foundation; Strategic Highway Research Program, Washington, D.C., 1994.
2. *The Long-Term Pavement Performance Program Guide.* LTPP Technical Assistance Contractor Team, Dec. 1993.
3. Wyman, J. H. An Evaluation of Currently Available WIM Systems. *Proc., 3rd National Conference on Weigh-in-Motion*, St. Paul, Minn., 1988, pp. 6–13.
4. Lee, C. E. Factors That Affect the Accuracy of WIM Systems. *Proc., 3rd National Conference on Weigh-in-Motion*, St. Paul, Minn., 1988, pp. 14–23.
5. German, J. L., and C. R. Copeland, Jr. *SHRP-LTPP Central Traffic Database User Guide.* Final Report. Texas Research and Development Foundation, 1992.
6. Dahlin, C. Proposed Method for Calibrating Weigh-in-Motion Systems and for Monitoring That Calibration Over Time. In *Transportation Research Record 1364*, TRB, National Research Council, Washington, D.C., 1992, pp. 161–168.

7. Juran, J. M. *Juran's Quality Control Handbook*, 4th ed. McGraw-Hill Book Company, New York, 1974 (rev.) Appendix IV, AIV.2.
8. Shapiro, S. S., and M. B. Wilk. Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, Vol. 52, 1965, pp. 591–611.
9. Davies, P., and F. Sommerville. Calibration and Accuracy Testing of Weigh-in-Motion Systems. In *Transportation Research Record 1123*, TRB, National Research Council, Washington, D.C., 1991, pp. 122–126.
10. *Statistix Analytical Software Version 4.1 and Users Manual.*. Analytical Software Publishing, 1994 (portions of software are copyrighted by Borland International, Inc. and Fleming Software).
11. Nelson, L. S. The Shewhart Control Chart—Tests for Special Causes. *Journal of Quality Control*, Vol. 16, 1984, pp. 237–239.
12. Woodall, W. H., and B. M. Adams. *Statistical Process Control.* In *Handbook of Statistical Methods for Engineers and Scientists.* (H. M. Wadsworth, ed.), McGraw-Hill Publishing Co., New York, 1989, pp. 7.3–7.20.

The contents of this paper reflect the views of the authors, who are solely responsible for the facts and accuracy of the material presented. The proposed methodology is not currently used in the FHWA-LTPP study, nor is it policy of FHWA. The contents do not necessarily reflect the official views of FHWA. The paper does not constitute a standard, specification, or regulation.

Safeguards and care must be taken if these ideas are used in post-processing. Further equipment and sensor error research is recommended.

Publication of this paper sponsored by Committee on Highway Traffic Monitoring.