Verification of Accuracy of LTPP SPS WIM Sites

Background Material for the Presentation by Richard Quinley, CALTRANS

The Long-Term Pavement Performance Program (LTPP) has intensified its efforts to obtain sufficient quantities of research quality loading data at a number of Specific Pavement Studies (SPS) sites. As one part of this effort, the FHWA has consulted with the Transportation Research Board LTPP Traffic Expert Task Group on a methodology to check the calibration of weigh-in-motion equipment. LTPP recognizes that there are multiple methodologies to actually calibrate WIM equipment, however, as with all other LTPP equipment protocols, a single methodology is being selected for verifying equipment calibration on a national basis. This provides a common reference for data users as to the quality standards expected of the data they have been provided.

The methodology selected and the software referenced are included for participant reference.

Individuals wishing to provide comments or obtain updates and any revisions to the currently selected methodology should contact Larry Wiser, Traffic Lead for the FHWA LTPP Team, at 202-493-3079 or via e-mail at larry.wiser@fhwa.dot.gov.

WIM Calibration Check Specification For LTPP Specific Pavement Studies Sites

Draft

Federal Highway Administration
Office of Infrastructure Research, Development and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike, HRDI-13
McLean, Virginia 22101-2296

August 21, 2000

U.S. Department of Transportation
Federal Highway Administration

Long-Term Pavement Performance
Serving your need for durable pavements

WIM Equipment Verification Procedure For LTPP SPS Sites

This document provides guidelines for verifying the accuracy of weigh-in-motion (WIM) systems for collecting LTPP traffic data at SPS 1, 2, 5, and 6 sites. It provides a performance specification that must be met by all WIM systems purchased by either state highway agencies (SHAs) or the Federal Highway Administration's Long-Term Pavement Performance Program Division (FHWA-LTPP).

The verification process consists of two parts, field data collection and ongoing office review of collected data. This document only discusses the field tests required to certify that a WIM system performs with the accuracy required by the LTPP tests. The office procedures reviewing the data collected from the WIM system are discussed in the report "Office Tests for Monitoring WIM Data from LTPP SPS Test Sites."

Introduction

The intent of the LTPP SPS WIM data collection effort is to improve the quality and reliability of the WIM data collected at the most important LTPP test sites. LTPP has decided that a central contractor will be responsible for collection and processing of data WIM data at LTPP SPS sites 1, 2, 5, and 6.[1] However, the equipment used to collect these data will not always be the same. In some cases, the equipment will be purchased and installed under specifications provided by FHWA-LTPP. In other cases, the SHA will have already installed a WIM scale, and the scale may not be the same as those selected with the FHWA-LTPP equipment specification. To ensure that high quality data will be collected under this program, all WIM devices used for this program will be required to meet performance specifications for data accuracy and reliability. Both types of scales (FHWA-LTPP provided or SHA provided) must meet the same performance specifications. These specifications are presented below.

System Performance Requirements

The LTPP program requires that WIM scale systems used to collect data under the LTPP SPS traffic loading data collection program meet accuracy and reliability standards similar to those listed in ASTM E-1318 for Type I scale systems. The criteria adopted by LTPP are given below in Table 1.

Table 1 -- WIM System Tolerances[2]

SPS-1, -2, -5, and -6 Sites	95 Percent Confidence Limit of Error
Single axles	± 20 percent
Tandem axles	± 15 percent
Gross vehicle weights	± 10 percent
Vehicle speed	± 1 mph [2 km/hr]
Axle spacing length	± 0.5 ft [150 mm]

1 Note: there are cases where the state highway agencies (SHA) will continue to perform these tasks. However, the SHA must be able to meet all of the guidelines and instructions set forth in this document.

2 These tolerances are taken from ASTM standard E1318-94.

Because collection of accurate traffic loading rates throughout the year is necessary to provide the loading rates needed for SPS test section research, the WIM systems used in this effort must meet the ASTM criteria all year long. Historically, many WIM systems have had problems accurately weighing vehicles when environmental conditions have changed from those that were present when the equipment was last calibrated. Changes in pavement strength at the scale location (often caused by changes in pavement temperature or moisture content) are known to cause problems with WIM system sensor accuracy. Because vehicle dynamics change with vehicle speed, it is also necessary to ensure that the WIM systems are able to accurately weigh trucks at the variety of speeds trucks will travel at each scale site.

All WIM systems currently on the market use one or more "calibration factors" as part of the process for converting axle sensor information into axle and vehicle weight estimates. These "calibration factors" are increased if the scale under-estimates static weights or decreased if the scale over-estimates static weights. In addition, many sales use an "auto-calibration" process that adjusts scale calibration based on monitored inputs. These adjustments are designed to take into account changing sensitivity of the scale sensors and electronics to changing environmental conditions and the aging sensors.

The goal of the verification specification is to prove that the scale system will produce accurate vehicle weights under expected highway operating and environmental conditions. This includes proving that the scale system does not produce outputs that are biased by temperature and/or vehicle speed and that the calibration factors are correctly set.

The recommended scale performance verification process involves weighing trucks of known weight with the WIM scale, and then comparing the WIM system's measurements with the known weights. The conditions under which vehicles are weighed by the WIM system are controlled to demonstrate that the system operates accurately under the majority of conditions expected to occur during LTPP data collection. During testing, the scale should be operated in its "normal" manner. That is, if auto-calibration settings are normally "on" then they should also be "on" during the verification tests.

The primary verification testing can be done with a minimum of two test trucks. Once a scale system has proved to be unaffected by vehicle speed and temperature over the course of a year, it is also permissible to verify scale calibration by comparing the statically measured weights of vehicles pulled from the traffic stream at static scales with WIM system weights for those same trucks. Note that this document describes only the use of test vehicles.

Outline of Performance Testing Procedure

The scale accuracy tests are designed to verify the performance of the scale as temperature and vehicle speed vary. The test plan requires the test trucks to make multiple passes over the scale sensors during a single day. Passes are made at different speeds and on days during the year when the highest daily temperature variations are likely to be found at the test site. Statistical tests are performed to examine system accuracy for both the entire data collection period and for subsets of the data that represent particular environmental or operational conditions.

The accuracy and reliability of data produced by the scale are computed by comparing the measured axle weights collected under these varying conditions with the known weight characteristics of the test vehicles. These basic test runs are then repeated two or three times during the year to make sure that the scale works accurately during the expected environmental conditions at that site.[3]

Equipment Needed

A minimum of two test trucks are required to perform the scale verification tests. These two vehicles should not have the same loading and/or configuration. One of the test vehicles should be a 3S2 (tractor, semi-trailer) with an air suspension system on all tandem axles. The air suspension should be in good condition. This vehicle should be loaded to approximately 80,000 pounds with a load that will not shift during the calibration test runs.

The second vehicle to be used in the calibration may be either a second 3S2 with a lighter load (less than 70,000 pounds) or a truck of an entirely different axle configuration. A good choice for the second vehicle is whatever vehicle configuration is supplying a large percentage of the axle loads to the SPS test pavement. Three- or four-axle single unit dump trucks should not be used for calibration checking[4] unless they are a major component of the traffic load experienced at that site. The second test truck should also be loaded to near its legal limit (unless it is a 3S2, as stated earlier), and should also have a suspension in good condition (air bag suspensions are preferred, but are not required for this second vehicle.)

If more than two test vehicles are used, the third vehicle may be configured and loaded as desired, although the use of three- or four-axle single-unit dump trucks is still discouraged, and all trucks should have suspension systems in excellent working condition.

The loads on all test trucks should be stable. For example, the loads should not be liquids (water) that shift within a tank. Such shifting causes the axle weights to vary with time and reduces the accuracy of the tests. If natural resource materials are used to load the test trucks, these loads should be covered to prevent adverse weather from adding to the weight of the load.

---

3 Additional runs may also be required, if concerns about the accuracy of the scale under different temperature conditions are not fully resolved by the initial tests, or if calibration drift is detected by routine quality assurance tests.

4 Conventional dump trucks tend to have unusual suspension characteristics. These characteristics result in inappropriate calibration settings. Therefore, they should not be used for most calibration efforts. Where these trucks make up the majority of heavy vehicles on a given roadway, it may be appropriate to use such a vehicle as one of the calibration trucks (along with the 3S2). If this is the case, consult with the TSSC before selecting vehicles to be used in the calibration check.

The trucks' tires should have a conventional highway tread pattern, not an off-road pattern, as a "knobby" tread can cause unusual sensor readings from some WIM systems. After being loaded and with full fuel tanks, both test trucks should proceed to a certified commercial scale[5] to have their axle weights measured. This information should be written down and passed along to the WIM data collection crew. This process should be repeated at the end of the testing session.

Last, the axle spacings (from centerline to centerline of each tire) should be measured for each test vehicle and reported to the WIM data collection staff.

In addition to the test trucks, the calibration team requires at least three radios, an infrared thermometer for measuring pavement temperature, and a computer for collecting the WIM data. A radar (or laser) gun for collecting vehicle speeds is also important.[6] (Suitable substitutions to these devices are acceptable, so long as the data collection requirements are met.)

Staff Needed

A minimum of four staff people are needed for this effort. Two staff collect data at the WIM system. The remaining staff drive the test trucks.

Basic Procedures

<u>Test Truck Operation</u>

Once the test trucks have been weighed at a calibrated, certified static scale that meets Handbook 44[7] specifications, they should make repeated passes over the WIM scale. The test trucks should initially operate at the mode (most common) speed of traffic currently using the roadway. The test trucks should move at a constant speed when they cross the scale (i.e., they should not be accelerating or decelerating), and they should be as centered in the lane as possible when crossing the scale. In most cases, this means that the test trucks will drive a continuous loop around the scale. After traveling over the scale, they should proceed to the next safe location where they can turn around. They should then pass back past the scale and proceed to the next safe location to turn around again. Once a truck has turned around and is accelerating back on the main roadway, the driver should radio ahead to the data collection crew to provide an estimated time for crossing the scale. This allows the data collection crew to prepare for the truck's arrival.

---

[5] The state weighing certification for that scale should be current, and the accuracy of that scale should have been checked within the last year. The test trucks can also be weighed for these tests with a state weight enforcement scale.

[6] Other independent, calibrated, speed measuring devices can also be used, so long as the tests of measured vehicle speed specified can be accomplished.

[7] Specifications, Tolerances and Other Technical Requirements for Weighing and Measuring Devices: As Adopted by the 83d National Conference on Weights and Measures, 1998

When approaching the WIM scale, drivers should hold the vehicle speed steady. After crossing the WIM scale, they should report to the collection crew the speed at which they were moving as they crossed the site. (This radio contact should take place after they have completely crossed the scale, as radio system broadcasts can cause interference with the WIM system outputs, causing false readings.)

Vehicle passes must be made at a variety of speeds. The speed range to be tested should include the speeds at which roughly 80 percent of the truck traffic at that site travel when they cross the scale. At least three different speeds should be used. If more than 40 miles per hour separate the highest speed from the lowest speed to be tested, four different speeds should be tested. A minimum of a 15-mph spread should exist between the highest and lowest speed to be tested.

Test runs should be made starting at the fastest speed selected for each test vehicle. The second run for each vehicle should be at the next slower test speed, with this pattern continuing until one run has been performed at each test speed. Then, the highest test speed should be used again. This rotating pattern of speeds should continue until the end of the data collection effort. As many test runs as possible should be performed.

If congestion or other factors limit the ability of a driver to make a specific test run at a given speed, the driver should attempt to make that run at the next slower speed, while keeping a constant speed as the test truck passes over the scale. When the intended speed can not be maintained, the driver should attempt to make the next test run at the originally intended speed.

Data Collection at the WIM Scale

For each vehicle pass, the field data collection staff should record the WIM system's output for axle weights and axle spacings. Each test vehicle's actual speed when crossing the sensor should also be obtained for comparison with the WIM system's output.

The following data from the WIM system should be obtained:

- the axle weights of the test trucks as they pass over the scale
- the spacings between axles on the test trucks
- the speed of the test trucks
- the sequence number of the test run[8]
- the temperature at the time of the test run (if not part of the system outputs will need to be obtained by other methods)
- the date of the test run
- the time of the test run
- the calibration factor used by the WIM scale.

8 This is the reference number used by the WIM system software to indicate this particular weight record. It may also be called the "record number" or "WIM assigned vehicle number."

These measured test truck attributes can be obtained by simply reading the values off of the WIM scale display screen.[9] These data should also be stored permanently and retrieved later for confirmation that the hand entry at the site was correct. (Some WIM systems also allow these data to be printed to a portable printer or written to an ASCII file for later retrieval.) Where differences in system recorded and observer recorded weights occur (that is the screen of the computer on-site rounds up a value that is carried to more significant digits in the main WIM system record), the main WIM system record should be used in the analysis effort.

While the test runs are being started, the data collection crews should use the radar gun[10] to determine whether the WIM system is correctly measuring traffic speeds. The data collection crew should also observe the measured tandem axle spacings of the drive tandems on 3S2 vehicles. If the vehicle's speed is correct (within 1 mph), and if the observed drive tandem spacing is approximately 4.4 feet, then test runs can begin. If either of these values are incorrect, the scale may not be operating correctly, and the scale's input variables should be adjusted.

Once the crew has established that the scale system is correctly measuring vehicle speed, additional speed data do not have to be collected other than for the test vehicles. However, the data collection staff should continue to enter speed values output by the WIM system into the Excel spreadsheet, since this maintains the record of vehicle performance needed to analyze scale calibration under different vehicle operating speeds and environmental conditions.

Analysis of the WIM Data

Once the data have been collected in the field, the statistics must be computed to determine whether the WIM site meets the accuracy standards set by LTPP. The basic statistic required for this test is the percentage of error for each of the variables measured. The percentages of errors computed from the data collected for each run are then used to compute a series of summary statistics. These summary statistics are then used to determine whether the scale is sufficiently accurate for LTPP use.

Overall System Accuracy Tests

This first set of tests requires data from all of the test runs performed at the site. It provides a general overview of system performance, given all of the environmental and vehicle speed conditions that were present during the testing. To perform these tests, the analyst should calculate the mean and standard deviation of the percentage of errors for the following variables:

- the front axle weights of the test trucks
- all tandem axles
- all non-front axle single axles

---

9 An Excel spreadsheet (**WIM2TRUCK)** has been provided by FHWA-LTPP to assist in recording and analyzing the data collected.

10 Or other calibrated, independent speed measurement system, with an accuracy of better than $\pm$0.25 mph.

- GVWs
- axle spacings
- vehicle speeds.

For each of these values, the analyst should compute the absolute values of the mean, plus or minus two standard deviations. The actual formula is given below. Select the larger of

$$ABS\ (X + 1.96\ \sigma) \text{ or } ABS\ (X - 1.96\ \sigma)$$

This value (for each variable) should then be compared with the values shown in Table 2. If any of these values exceeds the value in Table 2, the WIM system has failed the basic accuracy statistic for this test. If the largest of these absolute values is smaller than the values in Table 1, the system passes the basic accuracy test.

Table 2 --WIM System Tolerances[11]

For SPS-1, -2, -5, and -6 Sites	**ABS (X ±1.96 s) should be less than these values**
Single axles	20 percent
Tandem axles	15 percent
Gross vehicle weights	10 percent
Vehicle speed	1 mph [2 km/hr]
Axle spacing length	0.5 ft [150 mm]

where:

X = mean value of the error for each statistic
σ = the standard deviation of that error

Specialized Tests – Sensitivity to Speed

The scale being tested may pass the general accuracy test discussed above, but still may not be able to accurately measure the weights of vehicles traveling at different speeds. Another set of analyses examines the sensitivity of the scale system to changes in vehicle speed.

To perform this test, the test runs should be sorted by vehicle speed so that the data are grouped by the speed at which the test runs were performed. If three different test speeds were used (e.g., 65 mph, 55 mph, and 45 mph), three different accuracy computations will need to be made. Each speed analysis would consist of a summary of errors for all test truck runs performed at the speed in question. For each group of test vehicle runs, the mean and standard deviation of the percent error of each of the collected variables should be computed.

[11] These tolerances are taken from ASTM standard E1318-94

Because only a subset of test runs is performed at each speed range, there are fewer data points for this test than for the general performance test discussed above. To maintain confidence that the results of the test are still within the desired level of confidence, it is necessary to use the Students' *t* distribution, rather than the normal distribution in the statistical test. (Note, when the number of observations being tested reaches 30, these two tests become identical. This will occur for tandem axles when 15 3S2 test runs have been made, since each 3S2 has two sets of tandem axles.)

The formula used to compute the test statistic is similar to that shown above, only instead of using the value 1.96, the critical value of *t* is used for $\alpha = 0.025$. (These values are given in the appendix.) While the computation of the test statistic changes to account for the smaller sample size, the criteria for passing the test do not change. These values are shown in Table 3.

Table 3 --WIM System Tolerances[12]

For SPS-1, -2, -5 and -6 Sites	**ABS (X ± (*t* * σ)) should be less than these value**
Single axles	20 percent
Tandem axles	15 percent
Gross vehicle weights	10 percent
Vehicle speed	1 mph [2 km/hr]
Axle spacing length	0.5 ft [150 mm]

where:

X = mean value of the error for each statistic
σ = the standard deviation of that error
t = the Students' *t* statistic for $\alpha = 0.025$ and n-1 degrees of freedom
n = the number of samples available for that particular statistic under the conditions being tested.

Lastly, in addition to the Students' *t* test discussed above, a plot of percent measurement error versus vehicle speed should be created for each of the test variables. An analysis of this plot will also indicate if and how speed effects weight measurements.

Specialized Tests – Sensitivity to Temperature

A third set of tests examines the sensitivity of the scale system to changes in pavement temperature and subsequent changes in pavement performance. To perform these tests, the test runs should be sorted and grouped by temperature range. Because the field test can not control actual pavement temperatures, the grouping process is not as structured as the tests for speed sensitivity discussed above.

[12] These tolerances are taken from ASTM standard E1318-94

The recommended grouping process splits the day's test runs into three groups. Take the highest and lowest temperatures measured during the tests. The difference in these measurements gives the range of temperatures observed. Divide the range by three. This value will allow you to split the observed temperatures into three even range groups, cool, moderate, and warm, where "cool" is the lowest third (from minimum temperature to minimum plus one third the range), and "warm" is the highest (from maximum temperature to maximum minus one third the range.)

Specific test groups are then created by sorting test runs into each of these three test groups based on the pavement temperatures measured at the time of the test run.

For each group, the mean and standard deviation of the percentage of error for each of the test variables should be computed. These values are then used as input to the Students' *t* based test discussed above. The formula and values shown in Table 3 are again used to determine whether the errors measured by the field test fall within acceptable tolerance levels.

In addition to the basic statistical tests, the analyst should look for specific trends. For example, the scale may fall within acceptable error tolerances but show a marked bias toward extremely hot temperatures. Such a result would indicate that additional testing may be needed at higher temperature ranges. Such testing would then be scheduled when hotter temperatures are expected. As with the speed sensitivity tests, this can be accomplished by plotting percent error against pavement temperature.

Vehicle Classification Data Collection and Testing

The vehicle classification algorithm to be used at each LTPP test site should be supplied by the SHA responsible for the operation of the roadway at that site. The SHA should have performed a complete multi-hour test and evaluation of that classification algorithm to ensure that it accurately classifies trucks in that state.

The on-site test that is described below is not designed to fully test the algorithm. Instead, it is designed to ensure that the installed equipment is functioning correctly and that no mistakes have been made in the installation of a previously approved algorithm on that particular set of data collection electronics. The field test involves manually classifying vehicles crossing the WIM scale and comparing those classifications with the scale output. FHWA-LTPP has provided a spreadsheet (**CLASSMACRO.XLS**) to help with this analysis.

To perform the analysis, the data collector must enter classifications as vehicles cross the scale, and then enter the classifications reported by the WIM system for those same vehicles. The spreadsheet provided compares these two data entries and tabulates any classification errors.

To use this spreadsheet, open the file and then save it under a different name. Then hold the control key and type the letter "m." This starts the data collection macro. For each vehicle observed, type the classification of that vehicle and press the ENTER key. Then type the WIM system's classification for that vehicle and press the ENTER key again. Continue this process until data collection is complete.

While performing the data collection, the staff should look for specific types of errors that may be occurring, both to make sure that only limited errors are present and to identify specific limitations in the vehicle classification algorithm being used. For example, many automatic classification schemes have problems correctly differentiating among specific vehicle types because their axle spacing characteristics are similar. Consequently, the data collection crew should examine how well specific types of vehicles are classified, including the following:

- recreational vehicles
- passenger vehicles (and pick-ups) pulling light trailers
- long tractor semi-trailer combinations.

The classification data collection process can be performed when other calibration tasks are not being performed. Thus, it is not necessary for the data collection staff person to classify all vehicles crossing the scale during any given time period. (For example, the staff person can classify vehicles until the test truck approaches, stop classifying to assist in the data collection associated with the test truck run, and then return to classifying vehicles when the test truck has passed.) It is necessary, however, to enter both the manual classification <u>and</u> the WIM system's classification for each vehicle observed.

Vehicle classification should continue as long as the data collection crew is on site and are not otherwise busy. When the data collection task has been completed, "cancel" the macro's execution. The macro saves and closes the spreadsheet. To observe the results of the data collection, open the spreadsheet. Summary error statistics are displayed at the bottom of the data collection matrix.

The classifier is considered to be working acceptably when

1) no more that 2 percent of the vehicles recorded are reported as "unclassified" by the WIM scale
2) the number of classification errors involving <u>truck</u> classifications is less than 2 percent.

Selection of Days During the Year for Testing

Because temperature affects different scale systems differently, as a result of the mechanical properties of the weight sensors themselves and the interaction of those sensors with the pavement in which they are located, it is necessary to test each WIM system under different environmental conditions. To measure the effects of different environmental conditions on scale performance, the calibration of each scale that passes the first verification test must be verified under more than one set of environmental conditions. Tests should be performed in the spring, summer, and winter, as these are times when

- pavement strength varies as a result of different temperature and moisture conditions
- temperature variations (highs, lows, and maximum differences between highs and lows) differ significantly.

Three data collection sessions are anticipated for the first year of scale operation at most sites. Where environmental conditions do not change significantly during the year, a minimum of two calibration verification sessions should be performed each year. Once tests have proved that a given scale system (as installed) accurately accounts for environmental and highway operating conditions, only two calibration verification tests are required each year.[13]

FHWA-LTPP is responsible for selecting the days on which field testing will take place. These days will be selected on the basis of available environmental data and consultation with the SHA. Days will be selected when early morning temperatures are expected to be significantly different from temperatures found in the late afternoon. The intent is to allow collection of data during a single day when pavement temperatures (and consequently pavement strength) vary during the day.

Seasonal differences in temperature variation and moisture will be accounted for by collecting data during different times of the year.

[13] However, if the office monitoring process suggests calibration drift, additional system calibration will be required.

Appendix - Sample Size Versus Accetpable Standard Deviation

The number of test runs affects the confidence that results from the test results obtained. The simple statistical tests presented in this paper assume a normal distribution of errors and independence in the measurement of those errors. Neither of these assumptions is true. However, by developing a testing program that requires changing speeds (which change the effect of vehicle dynamics), temperatures (which affect pavement profile and strength), and different suspension types (through the use of more than one test vehicle), LTPP has created a test situation with a limited number of test vehicles that is likely to contain a greater degree of variation than would be found with a purely random sample of trucks. Therefore, the use of these statistical formulations is considered to be an acceptably conservative approach.

However, when the statistical tests are applied to subsets of the test data, the sample size involved can be quite small. As a result, the Students' *t* distribution has to be used rather than the normal distribution for statistical tests. Note that the sample size used to determine the appropriate number of degrees of freedom will change with each variable tested. This is because the test procedure assumes that each axle group is an independent measure of performance. Thus, if four test runs are made with 3S2 trucks, the sample size for the GVW test will be four, but the sample size for the tandem axle test will be eight.

The Students' *t* statistic for a 95 percent confidence interval is a function of the sample size. The appropriate value for *t* is given in the table below.

Values of Students' *t* Statistic for Different Sample Sizes[14]

Sample Size	*t*
12	2.201
14	2.160
16	2.131
18	2.110
20	2.093
22	2.080
24	2.069
26	2.060
28	2.052
30	2.045
31 or greater	1.960

[14] Taken from Introduction to Probability and Statistics, 4th Edition, by William Mendenhall, 1975, Duxbury Press.

Notes

Number of Test Runs Required

A minimum of 36 test runs are required. Four test runs are needed at each of three test speeds within each of three temperature ranges (4 x 3 x 3 = 36). However this is an absolute minimum. Additional test runs are encouraged whenever possible.

If two test trucks can not make the required 36 runs during one day of testing, either a second day of testing must be performed or additional test trucks should be used. Both methods for obtaining these test runs are acceptable. The total number of test vehicles to be used can be determined prior to the tests by determining the route each test vehicle will follow and calculating the time required to make one "circuit" of that route. This figure can be used to determine the maximum number of runs each test vehicle can make during the time period for which that vehicle is available.

Additional runs may also be necessary if conditions at the site prevent some scheduled runs from being performed at the desired speeds or if traffic conditions cause improper weights to be obtained (for example the test vehicle must brake to avoid another vehicle while approaching the scale sensors.). In addition, some flexibility is allowed in the order in which runs of different speeds are made. The intent is simply to obtain a sufficient number of runs at each speed and within each temperature range.

To obtain a wide temperature variation, it may be necessary to collect data for more than 8 hours per day. Given the cost of bringing test vehicles and personnel to a WIM site, it is often less expensive to extend a single day's testing than it is to schedule tests for more than one day. FHWA-LTPP and the SHA should consult prior to the scheduling of the field tests on the timing and duration of those tests.

Calibration Condition Prior to Testing

Each Weigh-in-Motion system to be tested should be calibrated prior to the verification tests discussed in this document. For systems that require temperature adjustments to their calibration factors, a complete set of (calibrated) temperature adjustments should be in-place prior to the start of verification testing and/or the start of data collection for LTPP purposes. (It is permissible for a state to calibrate the scale using these same basic procedures immediately prior to the start of the verification tests.)

Consequences of Failing the Verification Tests

No single remedy or consequence is appropriate for all sites where the verification tests described in this document show that the WIM system's outputs are not reliable.

In those cases where a single minor adjustments to the existing scale calibration factors can be shown to place the WIM scale in proper calibration, those adjustments, (if agreed to by both the state highway agency and the LTPP) will be made and the scale will be considered to have passed the calibration verification tests. However, where more than one adjustment is needed (for example, the scale appears to be temperature sensitive, and more than one temperature adjustment needs to be made), additional testing (calibration efforts under different environmental conditions) on the part of the state highway agency may be required, and LTPP reserves the right to require retesting of the scale to determine if data are sufficiently accurate for LTPP usage.

Where the scale passes some but not all tests (e.g., the scale passes the GVW and single axle load tests, but not the tandem axle load tests), LTPP will consult with its technical experts (including the Traffic ETG) before deciding whether to accept or reject use of that scale. Included in this decision will be an analysis of pavement conditions at that site, as well as other mitigating circumstances.

Where tests results indicate that the current scale is not sufficiently accurate and a simple calibration adjustment does not bring the scale back into calibration, LTPP will work with the state to determine the appropriate course of action. This could include, replacement of the scale sensors, modification of the pavement containing the sensors (e.g., grinding), or other improvements suggested by the LTPP, the state highway agency, or consultants advising these organizations.

Temperature Measurement

Because not all test sites are expected to be equipped with accurate pavement temperature sensors, pavement temperature readings must be collected at each site throughout the testing process. These temperatures will be used to estimate pavement temperature using a process previously adopted by LTPP. Where pavement temperature readings from the WIM scale are also available, these readings should also be obtained for use in the analysis effort.

Lack of Temperature Variation

While every effort will be made to select a day when pavement temperatures will vary widely during the testing procedure, it is possible that during the selected test day(s) little temperature variation will occur. Successful tests taken under these "less than desired" conditions will still be considered "validation" of the scale's calibration. However, it is expected that additional tests focusing on the scale's temperature sensitivity will be conducted at that site as quickly as practical, given the available project budget and staff scheduling.

Definition of the Term "Scale"

In this document, the term "scale" is assumed to include all components involved in the collection of vehicle weight data, unless otherwise specifically stated. The term "scale" is assumed to be synonymous with "Scale System." The term "scale" is not restricted in meaning to those components (sensors) that physically measure axle forces. It includes sensors, the electronics and software that interpret sensor signals, and converts that data into estimates of axle weights.