Follow-On Evaluation of the Montana Department of Transportation's STATE TRUCK ACTIVITIES REPORTING SYSTEM

by

Dr. Jerry Stephens
Department of Civil Engineering
Montana State University
Bozeman, MT 59717
Tel (406) 994-6113
Fax (406) 994-6105
jerrys@ce.montana.edu

Dr. Jodi Carson
Department of Civil Engineering
Montana State University
Bozeman, MT 59717
Tel (406) 994-7998
Fax (406) 994-6105
jodic@ce.montana.edu

Prepared for the
STATE OF MONTANA
DEPARTMENT OF TRANSPORTATION
MOTOR CARRIER SERVICES AND PLANNING DIVISIONS

May 17, 2005

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

An initial evaluation of the Montana Department of Transportation's recently implemented State Truck Activities Reporting System (*STARS*) found that the system successfully achieved its primary objectives of :

1) reducing pavement damage from overweight vehicles. *STARS* was used to identify historically where the most severe overweight vehicle problems were being experienced around the state, which configurations were operating overweight, and when these vehicles were operating. This information was subsequently used to help direct current enforcement efforts to directly impact the most severe overweight situations.
2) generating better information on vehicle load related demands for pavement design purposes. *STARS* provides more comprehensive data on the load characteristics of the vehicles operating on the state's highways than previously was available. This information can be used to produce highways that are better designed to the demands that they will actually experience in-service.
3) providing improved data to support a variety of MDT's tasks. The data collected by *STARS* is available across the entirety of MDT's operations.

STARS is, however, a relatively new tool, and as experience is gained with its use, issues and ideas continue to arise that require further investigation. Some of these issues were identified during and after the initial evaluation referred to above. The purpose of this project was to further research these issues, and the results of these investigations are summarized below.

From its inception, it was expected that *STARS* would only continue to be supported beyond its initial evaluation if the value of the benefit that it offered exceeded its cost. The primary benefit offered by *STARS* is improvement in the quality of the vehicle weight data available to support MDT's activities. This value of this benefit can be difficult to quantify, in that it is incremental in nature, and is spread across several MDT tasks. In this case, a conservative approach was taken to evaluating the value of *STARS* benefits; only benefits with demonstrable value were included in the analysis. These benefits, identified in the initial evaluation, consisted of reduced pavement damage from overweight vehicles (valued at $700,000 per year) and the generation of more efficient pavement designs (valued at $4,100,000 per year). These benefits were conservatively compared with the full cost of the *STARS* program, independent of any coincident use of *STARS* equipment for other purposes that could reasonably share some of the system costs. Costs included in the analysis consisted of all initial equipment and installation costs,

program development and evaluation costs, and on going calibration and maintenance costs. For a system of 26 WIM sites, these costs were estimated at $604,000 per year, assuming a 6 year service life for the system. The corresponding benefit to cost ratio for *STARS* is 7.9, which indicates that the benefits of the program clearly exceed the costs.

One benefit of *STARS* identified in the initial evaluation (as mentioned above) was a reduction in pavement damage that occurred when *STARS* data was used in scheduling and executing weight enforcement activities. The *STARS* data was used to identify those locations and times at which the most severe overweight problems historically occurred. This information was used in planning and executing some of the state's weight enforcement efforts in a prototype program for a one year period. While all the evidence indicated that the *STARS* focused enforcement activity was responsible for the decrease in pavement damage observed during that year, this conclusion was further reinforced by the subsequent analysis performed as part of this project on overweight vehicle activity in the year following the *STARS* focused enforcement effort. During the year following the focused enforcement effort, overweight vehicle operation on the state's highways returned almost exactly to the same level observed prior to implementation of the *STARS* focused enforcement effort. The pavement damage incurred from the overweight vehicles in the traffic stream also returned to the levels observed prior to the focused enforcement effort. These results clearly support the conclusion that the *STARS* focused enforcement effort was responsible for the decline in overweight vehicle activity during the year it was used.

While *STARS* was successfully used to help guide weight enforcement activities during the initial evaluation period, the future role of *STARS* in MDT's weight enforcement efforts is still evolving. Fundamentally, *STARS* characterizes commercial vehicle activity on the state's highways. Relative to weight enforcement, this information can be used in two ways, namely, to evaluate the effectiveness of enforcement and/or to actively direct enforcement. In both applications, the percent of overweight vehicles in the traffic stream and ESAL miles of travel attributable to overweight on vehicles (excess ESALs) are useful metrics for describing the level of overweight vehicle activity on the state's highways. While percent of overweight vehicles in the traffic stream is perhaps the more intuitively understood of these two parameters, excess ESALs miles of travel is a more comprehensive and direct measure of the negative impacts of overweight vehicle operations on the highway infrastructure.

Relative to evaluating enforcement effectiveness, changes in the characteristics of the overweight vehicle population between two periods, as measured by percent overweight vehicles in the traffic stream and/or accumulated excess ESAL miles of travel, can be used in this regard. In the first *STARS* evaluation, for example, the effectiveness of *STARS* focused enforcement was demonstrated based, in part, on the decrease observed in the percent of overweight vehicles in the traffic stream. As enforcement effectiveness improves, however, significant reductions in overweight vehicle operations will become more difficult to achieve. In this environment, it may be more appropriate to judge enforcement effectiveness over some time interval against a target level of compliance, say for example, a maximum of 6 percent overweight vehicles in the traffic stream, statewide, with no more than 10 percent overweight vehicles in the traffic stream at any individual *STARS* site. The specific form of a corresponding target metric expressed in terms of excess ESAL miles of travel is more difficult to visualize, which illustrates some of the difficulties in practically using this parameter as metric in evaluating the effectiveness of enforcement. Nonetheless, work should be done on developing a target metric based on this parameter, as it directly addresses the outcome of interest, that is, pavement damage from overweight vehicles.

The results of any evaluation of enforcement effectiveness using *STARS* information may depend on the interval over which the evaluation is conducted. In the short term, overweight vehicle operations on the state's highways are influenced by many factors that are dynamic in nature and independent of enforcement activity. Such factors include the timing of a harvest, the initiation or conclusion of a major construction project, the opening up of a forest area to logging, etc. Thus, evaluations conducted over a short time interval may be influenced by these short term effects, while such effects "average out" in evaluations conducted over a longer period of time (e.g., a full year), and over a larger geographic area (e.g., the entire state). *STARS* data can be used to assess enforcement effectiveness over a short evaluation window (e.g., weeks or months), and/or using the data from a single site, if care and professional judgment are used in assessing the validity of the evaluation and the significance of the results. Notably, conditions during the evaluation period must be reviewed to insure that the cause and effect relationship between enforcement and its outcome are actually being evaluated, rather than the cause and effect of some uncontrolled and dominant outside variable. Additionally, in light of the general variability in vehicle loading patterns, statistical analyses should be done to establish the significance of any observed changes in overweight vehicle operations during an enforcement activity, or the level of confidence with which it can be concluded that a metric calculated during an evaluation period meets a target value.

In any event, to fully judge the effectiveness of weight enforcement activities, some indication is required of the amount of resource that was expended in attaining a given outcome (e.g., a reduction in percent of overweight vehicles), either statewide or for a specific activity at a particular site. Obviously, a weight enforcement activity is not cost effective if its cost exceeds the benefit it produces in reduced infrastructure damage from overweight vehicle activity. Furthermore, among those enforcement activities that are cost effective, the cost of achieving the same level of success is expected to vary between activities. Presently, level of enforcement effort is not explicitly included as part of the Measurement of Enforcement Activities Reporting System (*MEARS*), which is the software that generates reports on overweight vehicle activity from the *STARS* data. Consideration should be given to quantifying and including some measure of level of enforcement effort as part of MEARS, which could be considered in conjunction with the outcome of the enforcement effort in assessing its effectiveness.

If *STARS* is used to evaluate enforcement effectiveness, it will naturally begin to be used as a tool to help direct enforcement activities. Locations and times at which enforcement is found to be “ineffective” in the evaluation process will become the locations and times around which future enforcement activities are planned. Relative to identifying problem sites, consideration should be given to those sites experiencing the greatest percentage of overweight vehicles, absolute number of overweight vehicles, average amount of overweight, and excess ESAL miles of travel. Of these various parameters, excess ESAL miles of travel should always be a factor considered in formulating enforcement decisions. Once again, consideration of these measures of overweight vehicle operation alone is insufficient to determine which sites should receive more enforcement attention. Some knowledge of the effect that a given amount of enforcement will have on the overweight vehicle operations at each site is essential to deciding where enforcement resources will be most effectively used. Ideally, the relationship between level of enforcement effort and the associated impact on overweight vehicle operations would be known for each site. This information would then be used in a system-wide optimization analysis so that the ratio of enforcement impact to level of enforcement effort was the same (and maximized) across the state. It will be a formidable task however to collect the data necessary to support such an analysis, and further develop the optimization program that would use this data. More pragmatically, it may be possible to introduce a simple parameter in MEARS that reflects the level of historical enforcement effort engaged in at a site, to assist in determining the prioritization of sites for focused enforcement. In the absence of explicit information on level of enforcement effort versus its impact and benefit, this variable will have to be externally factored into the enforcement decision making process using professional judgment. In the long term, it may be appropriate to introduce level of effort (cost) into the decision making process via a new software package that acts in concert with MEARS in evaluating the cost effectiveness of enforcement choices..

Note that in the initial *STARS* evaluation, the simple assumption was made that in a uniformly under enforced environment, enforcement effectiveness would be proportional to the volume of overweight vehicle activity at each *STARS* site. This assumption will become increasingly less valid, as enforcement effectiveness is selectively improved by factoring *STARS* information into the planning process.

STARS does provide the information necessary to improve enforcement effectiveness at individual sites, by indicating those times historically at which the greatest overweight problems have occurred as well as the vehicles responsible for these problems. One issue in using this historical data to direct enforcement is the inherent temporal variability in most loading patterns (as previously discussed relative to using *STARS* data in evaluating enforcement effectiveness). Care must be exercised to insure that any inherent variability in the recurrence of an overweight event has been taken into account when planning a current enforcement activity. Once again, in the absence of experience in this regard, professional judgment must be used in selecting the timing and duration of the enforcement activity to insure it captures the event of interest. Further consideration should be given during the enforcement effort, itself, as to whether or not the overweight vehicle activity expected based on the historical data actually has been realized.

While in the initial evaluation of *STARS*, enforcement activities in the “current” year were scheduled based on historical overweight vehicle activities one year earlier, it should be possible, and often it may be more appropriate, to plan and execute enforcement activities based on *STARS* data collected over considerably shorter time horizons. Notably, the “historical” data may be collected over an interval immediately prior to the enforcement activity. The only condition on this historical data is that it must be expected to reasonably represent the pattern of overweight vehicle operations expected during the proposed enforcement period. The enforcement period, itself, subsequently needs to be long enough so that its apparent effects are not diminished by short term variations in overweight vehicle activities. Once again, professional judgment must be used in deciding whether the historical data has captured a repeated pattern of overweight activity, and how long is “long enough” relative to the duration of the ensuing period of enforcement.

Problematic to using *STARS* data to evaluate and direct overweight vehicle enforcement activities in any of the capacities described above is the presence of permitted over standard weight vehicles in the traffic stream. Presently, permitted over standard weight vehicles are simply classified as overweight vehicles by MEARS, which distorts the number of overweight vehicles reported to be operating in the traffic stream. A methodology was explored in this project to characterize the permitted vehicle traffic at *STARS* sites using WIM data collected coincident with an overt enforcement activity being conducted at each site. The assumption inherent in this methodology is that apparently overweight vehicles in the traffic stream during overt enforcement activities are permitted over standard weight vehicles. This methodology was experimented with at a *STARS* WIM site adjacent to a weigh station (Mossmain), and it appeared to reasonably identify the permitted over standard weight vehicles operating at this site. From this information, factors can be developed either for each site or for groups of sites to adjust the reported overweight vehicle populations to account for the presence of permitted vehicles in the traffic stream.

The MEARS software has served its purpose well of analyzing the data from *STARS* and issuing for vehicle weight enforcement purposes. The software was set up to use every vehicle record in determining the characteristics of the overweight vehicle operations at each *STARS* site. As the number of *STARS* sites has increased, so has the computation time required to run monthly MEARS reports. Techniques were explored in this investigation to reduce this computation time. One promising approach for this purpose is to aggregate the data across some time interval (say 1 to 4 hours) prior to running MEARS. Appropriate aggregation time intervals can be determined based on the tolerable error in the results. Obviously, the attraction to running the entire data set is the elimination of any error introduced by the sampling process.

MEARS performs a calculation that compares the excess (overweight) ESAL miles of travel experienced at a given location during two different time windows. The purpose of this calculation is to evaluate changes in excess ESAL miles of travel that might result at a site from differences in enforcement activity during the two time intervals. Excess ESAL miles of travel, however, is affected by total volume of traffic at a site, as well as level of enforcement activity. The level of traffic during the two time intervals of the evaluation can not be controlled; therefore, MEARS includes an algorithm to analytically normalize the excess ESAL miles of travel experienced during the two intervals to a common volume of traffic. Any difference in the normalized excess ESAL miles of traffic is then the result of the difference in enforcement activity during the two intervals. The algorithm used by MEARS in this normalization process is simple, but approximate in nature. A more sophisticated algorithm for this purpose that includes more of the variables known to impact this problem was investigated in this project. The new algorithm is significantly more computationally intensive than the existing algorithm. In a comparison of the performance of the two algorithms, the existing algorithm was found to systematically overestimate reductions in excess ESAL miles of travel that occurred between two evaluation intervals by approximately 3 percent, while it accurately estimated increases in ESAL miles of travel between two evaluation intervals. These assessments are based on the assumption that the new and more sophisticated algorithm is generating "true" values for this parameter, which may or may not be the case. In light of the small magnitude of the possible error in using the existing algorithm, and the uncertainties in this validation process, the existing MEARS algorithm was judged to be adequate for estimating the change in excess ESAL miles of travel between two intervals.

It was the intention of this project to study some additional issues with *STARS*, issues that were identified in the initial evaluation as meriting further investigation. Issues not pursued herein due to time constraints in completing this project include: treatment of bypass during *STARS* focused enforcement activities, review of the criteria used to determine the location of *STARS* sites, and review of the extent of the highway system influenced by activities at each *STARS* site.

These and other issues brought forth in this study merit further investigation in light of the demonstrated benefits that *STARS* has to offer in improving the effectiveness of MDT’s weight enforcement efforts and in improving the efficiency of MDT’s pavement designs.

1 INTRODUCTION

1.1 Background

A recent evaluation of Montana's *State Truck Activities Reporting System* (*STARS*) conducted by Montana State University found that the information provided by *STARS* on commercial vehicle operations on the state's highways: (1) was successfully used to reduce infrastructure damage from overweight vehicles, (2) offered a more comprehensive and accurate characterization of traffic related fatigue demands on the highway system for pavement design than is available from traditional sources (weigh station sampling efforts) and (3) was found to be useful to several divisions within MDT with respect to many of the analyses they are tasked to perform (Stephens, Carson, Reagor, and Harrington 2003).

In the course of the evaluation, several issues were identified that merit further investigation. These issues range from determining the basic benefit to cost ratio of the overall *STARS* program to researching specific enhancements for the *Measurement of Enforcement Activity Reporting System (MEARS)* software.

In this addendum many of these issues were further investigated, resulting in a more comprehensive overall evaluation of the *STARS* program and supplementary guidance pertaining to the program's continued operation.

1.2 Objectives

Specifically, this addendum pursued the following objectives:

(1) A benefit to cost ratio was calculated for the *STARS* program, taking into consideration benefits related to reduced infrastructure damage, improved infrastructure designs, data quality, etc. and costs related to equipment, maintenance, calibration, data processing, etc.

(2) One year has passed since the conclusion of the focused enforcement efforts of the pilot project. The *STARS* data collected since the conclusion of that effort were

processed and compared with that collected during the original baseline and enforcement years to determine if trends in enforcement effectiveness identified in the original evaluation of the pilot project are consistent with changes in overweight vehicle operations that occurred after the pilot project concluded.

(3) Feasible and effective strategies are discussed for the continued use of *STARS* in enforcing vehicle weight limits. Items investigated include the use of *STARS* in planning and evaluating enforcement activities in the short term and further development as appropriate of the metrics calculated from the *STARS* data to evaluate enforcement effectiveness.

(4) A methodology was investigated for identifying permitted over standard weight vehicles in the traffic stream from WIM data. This methodology was validated as possible using WIM data at a specific site in Montana, in conjunction with available information on the characteristics of over standard weight single trip and annual term permits inssued by MDT.

(5) Recommendations are made on software enhancements that could further improve the *STARS* program cost efficiency and ease of use.

While most of these objectives were definitely addressed in the course of this investigation, in a couple of cases, only general approaches and/or prototype methodologies could be offered within the time constraints of this investigation. The role of *STARS* in MDT’s activities continues to evolve, and recommendations for further work are included at the conclusion of this report.

2 DESCRIPTION OF *STARS*

Before addressing the specific tasks of this evaluation, a general description of the *STARS* program is presented below. Information is provided on the *STARS* hardware and software, and on the initial performance evaluation conducted for the program.

2.1 *STARS* Hardware

STARS consists of a network of permanent WIM sites (26 of which had been installed at the time of the evaluation out of a total of 36 planned sites) supplemented by 62 sites that are operated intermittently on a three-year cycle using fully portable WIM equipment. Included in these sites are four automated weigh stations that utilize WIM and Automatic Vehicle Identification (AVI) equipment to allow legal bypass of weigh station facilities by credentialed weight-compliant commercial vehicles. Data collected from these automated weigh stations is treated just like the data collected at the *STARS* WIM sites.

The permanent WIM sites, shown in Figure 2.1-1 and described in Table 2.1-1, are placed around the state on major routes that carry significant truck traffic. Locations were generally selected based on the volume of commercial vehicle traffic carried on the various routes and systems (i.e., Interstate, non-Interstate NHS, primary, secondary, urban) and the location of existing weigh station facilities, with due consideration of the recommendations of FHWA's *Traffic Monitoring Guide* (FHWA 2001). Since weigh station coverage is greatest on the Interstate system, the *STARS* sites are focused on the non-Interstate NHS and Primary routes around the state. The portable sites additionally cover less-traveled routes known to continuously or seasonally experience significant truck traffic. The precise location of each WIM installation along a particular route was determined based on siting requirements of the WIM system, itself (e.g., roadway grade and alignment criteria, etc.). In light of *STARS* potential role in weight enforcement, consideration was also given in the siting process to the location of places in the vicinity of each site at which vehicles could be safely pulled off the highway during an enforcement activity.

The specific hardware installed at each of the 26 permanent sites is listed in Table 2.1-1. Of the three types of WIM sensors commonly used - piezoelectric, bending plate and permanent load

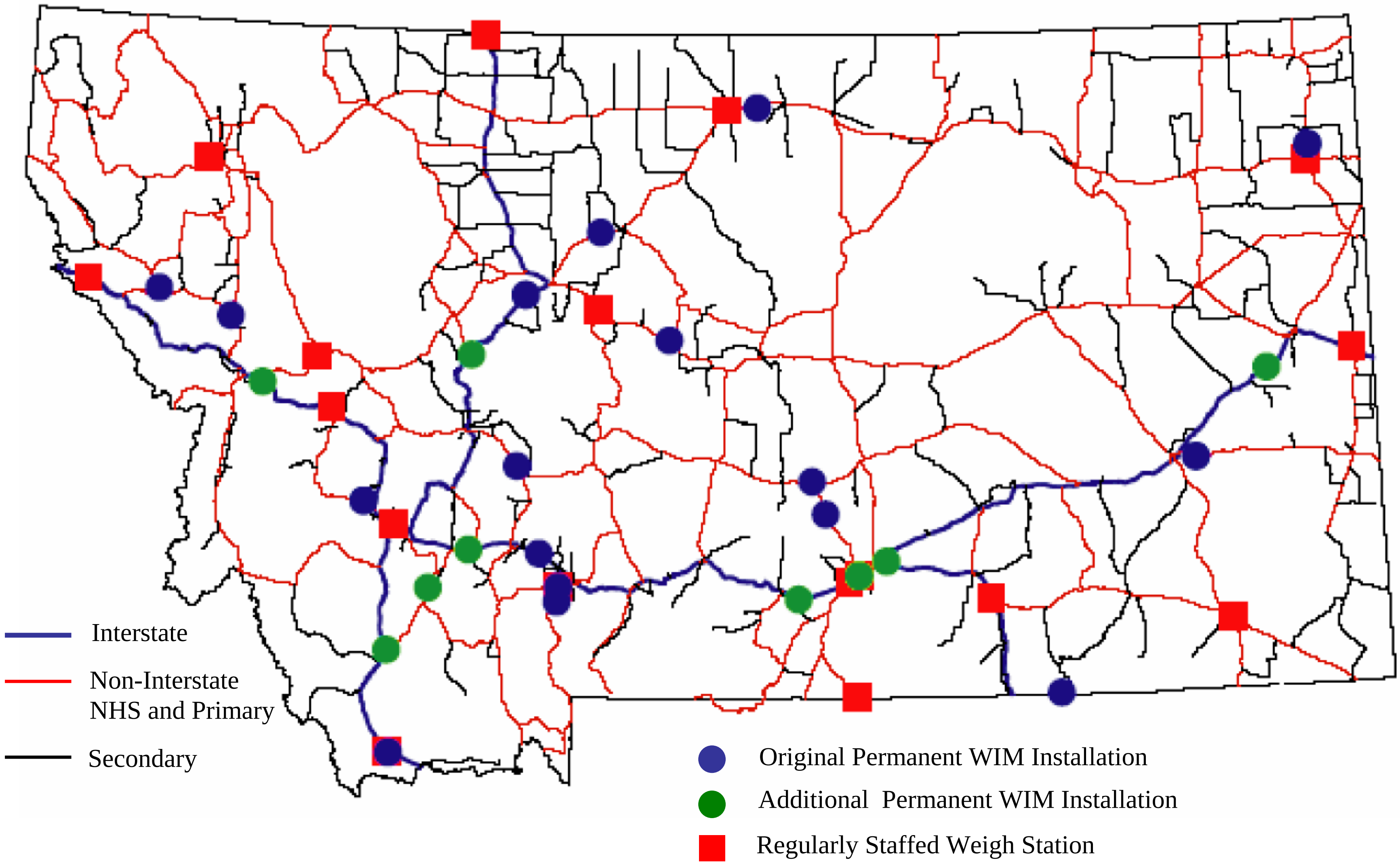


Figure 2.1-1. Montana State Highway System, Weigh Stations and *STARS* Sites (Little 2005)

Table 3.2.1-1. Installed WIM Systems, Location and Equipment (Bisom 2003)

Site	Highway System	Route	Technology
Townsend	Primary and Non Interstate NHS	U.S. Highway 287	Piezoelectric
Decker	Secondary	Highway 314	Piezoelectric
Bad Route	Interstate	Interstate 94	Piezoelectric
Manhattan	Interstate	Interstate 90	Piezoelectric
Arlee	Primary and Non Interstate NHS	US Highway 93	Piezoelectric
Four Corners	Primary and Non Interstate NHS	US Highway 191	Piezoelectric
Gallatin	Primary and Non Interstate NHS	US Highway 191	Piezoelectric
Galen	Secondary	Highway 273	Piezoelectric
Broadview	Primary and Non Interstate NHS	State Route 3	Piezoelectric
Miles City East	Primary and Non Interstate NHS	US Highway 12	Piezoelectric
Ulm	Interstate	Interstate 15	Piezoelectric
Ryegate	Primary and Non Interstate NHS	US Highway 12	Piezoelectric
Stanford	Primary and Non Interstate NHS	US Highway 87	Piezoelectric
Fort Benton	Primary and Non Interstate NHS	US Highway 87	Piezoelectric
Havre East	Primary and Non Interstate NHS	US Highway 2	Piezoelectric
Twin Bridges	Primary and Non Interstate NHS	State Route 41	Piezoelectric
Paradise	Primary and Non Interstate NHS	State Route 200	Piezoelectric
Mossmain[a]	Interstate	Interstate 90 W	Piezoelectric Bending plate
		Interstate 90 E	Bending plate
Culbertson[a]	Primary and Non Interstate NHS	State Route 16	Bending plate
Lima[a]	Interstate	Interstate 15	Bending plate
Armington[a]	Primary and Non Interstate NHS	US Highway 87 W	Piezoelectric
		US Highway 87 E	Piezoelectric
Columbus	Interstate	Interstate 90	Piezoelectric
Bonner	Interstate	Interstate 90	Piezoelectric
Dillon	Interstate	Interstate 90	Piezoelectric
Pryor Creek	Interstate	Interstate 90	Piezoelectric
Wolf Creek	Interstate	Interstate 15	Piezoelectric

[a] PrePass Site (one direction only, unless indicated otherwise)

cell - the majority of the installations are piezoelectric (23 out of 26); the remainder are bending plate (3 out of 26). The piezoelectric systems were manufactured by Electronic Control Measurement (ECM), while the bending plate systems were manufactured by PAT America. The relative accuracy and cost of these WIM systems continues to be a subject of debate among the public agencies that use them. The piezoelectric sensors are expected to provide adequate accuracy for MDT's intended use at the most attractive life cycle cost, based on MDT's experience to-date with these technologies and preliminary results from active research projects investigating their performance (Clark, Stephens, and Carson 2004).

MDT calibrates the permanent WIM sites twice each year according to standard procedures using a 5-axle tractor, semi-trailer of known weight. MDT also performs standard quality control checks on the raw and processed data.

2.2 Software Components

The data collected at the various WIM sites is automatically analyzed using the *Measurement of Enforcement Activities Reporting System (MEARS)* computer software program specifically developed for MDT. *MEARS* generates reports on the commercial vehicle activity by site and month and for the entire year. Reports are also generated on the general performance of the WIM hardware. The full suite of reports available from *MEARS* is summarized in Table 2.2-1.

2.3 Initial *STARS* Evaluation

An initial evaluation of *STARS* was completed by Montana State University in 2003 (Stephens, Carson, Reagor, and Harrington 2003). This evaluation found that *STARS* had met three of its primary objectives, namely,

(1) improving the efficiency and effectiveness of truck weight enforcement activities performed by the Motor Carrier Services (MCS) Division of MDT,

(2) providing MDT access to improved truck-related data for use in pavement design, and

(3) providing various divisions within MDT access to improved truck-related data for use in engineering and planning applications.

Between 2000 and 2002, the Motor Carrier Services (MCS) Division of MDT conducted a pilot project to investigate the use of *STARS* data in scheduling mobile weight enforcement activities. Data from *STARS* was used to identify those locations around the state that historically experienced the worst pavement damage from overweight vehicles. Instrumental in identifying these locations, which were then the object of focused enforcement, was the *MEARS* software. As a result of this activity, a statistically significant reduction was seen in the percent of overweight vehicles in the traffic stream. Statewide, throughout the extensive network of highways covered by *STARS*, the percent of overweight vehicles in the traffic stream dropped by

Table 2.2-1. MEARS Reports (Bisom 2003) by Month and By Site (unless otherwise indicated)

25: Overweight Vehicle Report by Class
- Number of commercial vehicles
- Percent of overweight commercial vehicles
- Average amount of legal weight exceedance

30: Overweight Violations by Time Period and Class
- Day of week and 4-hour segment of day
- Direction of travel

35: Weight Information by Class
- Number of commercial vehicles
- Percent of overweight commercial vehicles
- Average operating weight
- Average amount of legal weight exceedance

40: Scatter Graphs by Class
- Scatter graph of overweight commercial vehicle events as a function of day of week and time of day

45: Calibration Tracking
- Weight frequency plots of vehicles in the traffic stream used for auto-calibration

70: Summary of Records Violating Rules
- Total number of records that violate rules validating reasonableness of recorded vehicle characteristics

90: Truck Weight Upload Process Summary Report
- Total number of records screened
- Total number of bad records

105: Site Activities Roll-up
- Total number of vehicles
- Total number of commercial vehicles
- Percent of overweight commercial vehicles
- Average amount of legal weight exceedance
- Change in overweight commercial vehicle percent
- Change in average legal weight exceedance amount

205: ESAL Report
- Excess ESALs attributable to overweight vehicles by duration of reporting period

22 percent (from being 8.8 percent of the commercial vehicles in the traffic stream in the baseline year to 6.9 percent in the enforcement year). The average amount of overweight on each vehicle also decreased by 16 percent in the enforcement year. The overall reduction in pavement damage attributable to the focused enforcement effort over the year was on the order of magnitude of 6 million ESAL-miles of travel. The cost savings associated with this change in pavement damage was estimated to be approximately $700,000.

In the area of pavement design, *STARS* was found to offer better information on the traffic related fatigue demands used in design, relative to the existing information that is collected for this purpose at permanent weigh stations. From a geographic perspective, *STARS* collects information at more locations around the state than is available at the existing weigh stations. From a temporal perspective, *STARS* collects data continuously at these sites, while weight data for pavement design purposes is only collected at the weigh stations at a few selected times during the year. Using *STARS* data in the pavement design process (rather than weigh station data) was projected to annually save approximately $0.7 million and $3.5 million per year on the Interstate and non-Interstate NHS/Primary systems, respectively, through the generation of more cost effective pavement designs.

The final issue considered in this evaluation was the possible benefits *STARS* offers to traffic data users throughout MDT. A survey across the major divisions at MDT found that *STARS* data will primarily benefit planning, engineering, and commercial vehicle enforcement efforts.

3 BENEFIT TO COST ANALYSIS

3.1 Benefit to Cost Comparison

The previous evaluation conducted by MSU resulted in the quantification of benefits of *STARS* related to the reduction in pavement damage attributable to overweight vehicles through *STARS* focused weight enforcement and improved pavement designs resulting from more comprehensive, higher quality truck weight data. The resulting program benefits in other divisions within MDT attributable to data quality enhancements vary significantly, and quantification of these benefits was beyond the scope of the original evaluation.

Under this supplemental investigation, efforts focused on determining (a) the value of the benefits of *STARS* not quantified in the original evaluation and (b) the costs associated with *STARS*. Note that with respect to benefits, the original evaluation assigned dollar values to those benefits whose value can reasonably be quantified (i.e., annual savings attributable to pavement damage reduction and improved pavement designs). Nonetheless, this investigation reconsidered secondary benefits of the *STARS* program agency-wide within MDT, to ensure all quantifiable benefits were identified and included in the analysis. While costs may appear to be more easily quantified, some decisions were required as to what costs should be included in this analysis relative to the benefits being considered. Costs to be considered include, capital equipment (amortized over the life of the equipment), calibration, maintenance, and software, and labor for enforcement, data processing, data analysis and administration. Note that *STARS* generally does not replace any major information gathering technology that is used by MDT to support it various functions. Thus, when evaluating the benefit to cost ratio for *STARS*, the marginal benefit it offers to MDT's functions was considered relative to the full cost of the program. This approach should yield a conservative (low) benefit cost ratio for the program.

Cost information was gathered almost entirely through discussions with MDT personnel with the intent of quantifying labor costs for the *STARS* program. Equipment, calibration, maintenance and software related costs were obtained from MDT personnel from historic records that reflect actual experience with *STARS* and projected operating expenditures anticipated over time. Information pertaining to the anticipated design life of the various hardware products in use was obtained from published literature. This information was analyzed to produce an overall benefit to cost ratio for the *STARS* program. A lack of quantifiable secondary benefits and aggregate cost estimates precluded the determination of individual benefit to cost ratios for each MDT activity (i.e., enforcement, pavement design, etc.) that use *STARS* data. Findings from the benefit to cost analysis for the *STARS* program are detailed below.

3.2 Benefits Attributable to *STARS*

The degree to which benefits can be realized from the *STARS* program varies from one area within MDT to another. The Planning and Motor Carrier Services Divisions and the Pavements and Materials Bureau realize the greatest potential for benefit from the *STARS* program. Areas such as the Engineering Division, including the Safety Management Bureau and the Geometric Design Unit, and the Bridge Bureau, that use truck-related, or specifically, WIM data, infrequently realize only minor benefit. In each of these areas, the types of data required to perform their operation is not significantly enhanced through the *STARS* program; information specific to a particular site or a particular truck configuration is generally required. In the first case, a WIM site may not be located in the immediate vicinity of the site of interest. In the second case, information pertaining to a particular truck configuration, including dimensional information, is relatively standard and invariable and may even be unique to a specific load or trip (i.e., an oversized movement). In both of these instances, WIM benefits related to larger sample sizes and more accurate aggregate data, are not as pronounced.

Hence, benefits attributable to improved data quantity and quality in MDT areas that infrequently use weigh-in-motion data are predicted to be minor and incremental beyond current operations. Further, any changes in operations made possible through improved data quantity and quality in these secondary MDT areas are unable to be foreseen, particularly by MDT personnel in these areas who are still relatively unfamiliar with the *STARS* program capabilities and offerings. As such, benefits resulting from improved data quantity and quality in these areas were omitted from this analysis resulting in an estimated benefit to cost ratio that may be lower than actual.

Instead, efforts to quantify additional benefits attributable to the *STARS* program focused on activities within the Planning, Engineering and Motor Carrier Services Divisions within MDT.

3.2.1 Planning

As part of its overall mission, the Planning Division provides an important supporting function to other areas within MDT as well as the Federal Highway Administration, U.S. Department of Transportation and others by providing detailed and/or aggregate truck-related data. This data includes Commercial Average Daily Traffic (CADT), Commercial Vehicle Miles Traveled (CVMT), Equivalent Single Axle Loads (ESALs), Percent Large Trucks of AADT, Percent Commercial Trucks of AADT and Traffic Stream Distribution. The WIM systems implemented as part of the *STARS* program capture much of this data automatically, reducing the data collection burden historically conducted through manual counts at site-specific or static weigh scale locations. Further, the permanent WIM systems implemented through the *STARS* program capture much of this data comprehensively, theoretically obtaining a record for every truck traveling past that site. While this increase in the quality and quantity of data available to the Planning Division and others through the Planning Division has obvious intuitive benefits related to the fulfillment of the Planning Division's mission, e.g., through improved accuracy in estimates to support planning-related decisions and in estimates for State and Federal reporting of transportation conditions in Montana, etc., a method for monetarily quantifying these benefits is not obvious. Unlike the quantification of benefits related to pavement damage and pavement design, planning activities are not as directly nor as narrowly tied to empirical relationships that allow for the calculation of design differences based on different input values (i.e., recall that the monetary benefits resulting from improved pavement design were determined by comparing the existing design based on traditional data and a new design based on the more comprehensive WIM data). Hence, no additional benefits related to improved data quantity and quality in the Planning Division were included in this analysis. In fact, any resulting benefits may be offset significantly by the increased labor requirement that results from the increased collection, processing and analysis of the increased quantity of data (discussed in the ***Costs Attributable to STARS*** section).

3.2.2 Motor Carrier Services

As with the Planning Division, obvious, intuitive benefits exist for the Motor Carrier Services (MCS) Division from the *STARS* program through the ability to better direct enforcement activities and improve the efficiency through which enforcement personnel are able to perform their duties. However, from the standpoint of quantifying costs and benefits, no net change in enforcement personnel full-time equivalent (FTE) positions occurred (i.e., the same tasks could not be performed with fewer personnel). Instead, enforcement personnel were likely able to shift some of their weight enforcement time commitment to other areas of enforcement (i.e., dyed fuel enforcement, etc.), though this is difficult to quantify. Additionally, no net increase in weight violation citation issuance occurred for revenue generation as a result of the *STARS* program; this revenue source was in fact decreased with the issuance of fewer citations. However, the "cost" of reduced revenue generation resulting from fewer weight violation citations issued may be compensated for by additional citations issued in other enforcement areas as enforcement personnel reallocate their time commitments.

3.2.3 Pavements and Materials

While no additional benefits were quantified for either the Planning or Motor Carrier Services Division, the improvements in data collection noted in the Planning Division and the improvements in the enforcement capabilities of the Motor Carrier Services Division are ultimately manifested in two objectives of the Pavements and Materials Bureau: (1) the extension or attainment of the pavement's full design life through the reduction of unnecessary damage and (2) the development and implementation of accurate pavement thickness designs (existing pavements may be currently under-designed leading to early failure or over-designed leading to higher than necessary construction costs) to support the anticipated ESALs.

The Materials Bureau uses ESAL data from the Planning Division to generate pavement surfacing designs and as part of their Pavement Management System. The Materials Bureau currently uses 20-year ESAL information provided by the Planning Division for their surfacing designs with estimated ESALs based on a formula using AADT and percent commercial vehicles. The basic ESAL information by vehicle configuration required for this analysis is determined from static scale data collected during special sampling periods throughout the year. The initial *STARS* evaluation found that more accurate data for pavement design purposes is obtained from the *STARS* WIM system compared to that obtained using the existing static scale sampling approach. This improvement in data quality was found to directly translate into more efficient highway designs, generating a cost savings to the state.

An additional benefit of *STARS*, and one that again is not readily quantifiable, is the extensive information that it produces on *actual* ESALs that should enhance the effectiveness of MDT's Pavement Management System.

3.2.4 Total Benefits Attributable to *STARS*

With no additional benefits able to be monetarily quantified through this investigation, the total benefits attributable to the *STARS* program are as previously identified in the initial *STARS* evaluation, totaling $4.8 million annually (see Table 3.2.4-1).

Table 3.2.4-1. Benefits Attributable to *STARS*

Nature of Benefit	Value, $
Reduction in Pavement Damage	700,000
Improved Pavement Designs	
Interstate	700,000
Non-Interstate	3,400,000
Total	4,800,000

3.3 Costs Attributable to *STARS*

STARS program related costs can be categorized as (1) initial costs and (2) ongoing annual costs. For this analysis, initial costs include capital equipment and installation, software development and research and program evaluation. Ongoing annual costs include calibration and maintenance, labor, travel, communications and utilities and other. Costs are based on both actual historical expenditures and projected annual costs as estimated by MDT.

The initial *STARS* evaluation considered 16 permanent WIM sites around the state. The *STARS* program is now planned to include 36 permanent WIM sites and 62 portable WIM sites (26 of the permanent sites had been completed at the time of this follow-on evaluation). For this investigation, the decision was made to consider the costs of all 26 installed sites in calculating the benefit to cost ratio for *STARS*. In the case of the benefits realized from *STARS* directed weight enforcement (i.e., a reduction in pavement damage from overweight vehicles), this approach is conservative, in that these benefits were generated across just the 16 sites included in the original evaluation. In the case of the benefits realized from improved pavement designs, this approach is appropriate, in that the magnitude of this benefit was estimated across the entire state highway system. Further note that 4 of the sites included in *STARS* are PrePass sites. As such, if the full costs of these sites is to be considered in this analysis, the value of the benefit of the PrePass program should also be considered. The decision was made to conservatively include the full cost of these sites in this analysis, without considering the value of their benefit.

Note that *STARS* will replace the process currently used by MDT to collect information on axle weight distributions by vehicle configuration that is essential to the pavement design process. Thus, in evaluating the benefit to cost ratio of the *STARS* program, the marginal benefit it offers in more efficient pavement designs should be evaluated against the additional cost rather than the total cost of using *STARS* instead of the current methodology to collect this data. Basic vehicle weight data for pavement design purposes traditionally has been collected at static scales around the state during special sampling periods throughout the year. This data collection effort does involve additional work above and beyond that associated with routine operation of the static scales. MCS and Planning indicated that the resource required for this activity amounted to a few man hours per month. MCS also noted that in addition to collecting weight information for pavement design purposes, this effort provides another opportunity for MCS personnel to closely review commercial vehicle operations on the state’s highways for a variety of purposes. Thus, only part of the function (and cost) of this special data collection effort will be offset by *STARS*. Based on these various considerations, the decision was made to conservatively ignore the existing nominal cost of collecting vehicle weight data for pavement design purposes in determining the benefit cost ratio for the *STARS* program.

3.3.1 Initial Costs

The initial costs for the *STARS* program are summarized in Table 3.3.1-1. While the initial *STARS* WIM systems were installed over several years, for the purposes of this analysis it was assumed that the entire system was completed simultaneously. Equivalent annual costs for equipment and installation were determined based on expenditures by MDT and estimated equipment service lives of 4, 6 and 12 years. These service lives were determined from reported national experience with the same WIM technologies (Mumayiz 1989, McCall and Vodrazka 1997, International Road Dynamics, Inc. 2001, Whitford 1998). The majority of the WIM systems in Montana were deployed over the past 10 years; thus, only limited historical data is available on the service life of these systems under Montana's traffic and weather conditions. On non-Interstate routes, 12 sensors out of the total of 76 sensors that have been installed have been replaced. The average age of these sensors at failure was 6.8 years. The average age of the active sensors at non-Interstate sites around the state is 5.4 years. On interstate routes, 9 out of the 73 sensors that have been installed have failed. The average age of these sensors at failure was 5.2 years. The average age of the sensors that are currently active on the Interstate system is 4.7 years. This value is smaller than the corresponding value for non-Interstate sites, primarily due to the fact that many of the Interstates sites are fairly recent additions to the STARS system. In any event, it appears that the average service life of the STARS WIM systems will be at least 5 to 7 years. Note that some sensor failures obviously resulted from problems with the pavement in which they were installed, rather than problems with the sensors, themselves, and that MDT has subsequently developed roadway condition criteria that are used in the site selection process.

Average annual costs for software development (excluding routine minor software upgrades) and research and evaluation were calculated using the same service life cycle as assumed for the hardware; it is likely that with the installation of ever advancing WIM equipment, companion advancements will be required in supporting software programs and the need for evaluation of performance will once again arise. Detailed cost estimates in each of these categories are provided below. Calculating average annual costs for equipment and installation, software development and research and evaluation, costs range from $142,321 assuming a 12-year service life to $426,964 assuming a 4-year service life.

Table 3.3.1-1. Initial Costs of the *STARS* Program (from information supplied by Bisom, 2004)

Item	Initial Cost, $	Amortized Annual Costs, $ 4-year life	6-year life	12-year life
Equipment and Installation[a]				
Non-interstate WIM sites (16)	548,960	137,240	91,493	45,747
Interstate sites (10)	630,200	157,550	105,033	52,517
Software[b]				
MEARS Development	145,724	36,431	24,287	12,144
MEARS II Development	90,470	22,618	15,078	7,539
Research and Evaluation				
Investigation of WIM Performance	61,502	15,376	10,250	5,125
Initial *STARS* Evaluation	166,000	41,500	27,667	13,833
Follow-on *STARS* Evaluation	65,000	16,250	10,833	5,417
Total	1,707,856	426,964	284,643	142,321

[a] excludes annual computer replacement which is budgeted through MDT's IT Division
[b] excludes annual software upgrades which are budgeted through MDT's IT Division

Equipment and Installation- While a few bending plate systems are included in the *STARS* program, the majority of the WIM installations are piezoelectric systems, and MDT has made the decision to use piezoelectric technology in all future work (Bisom 2004). While there is some variation in the cost of piezoelectric WIM installations based on specific site conditions, system costs were estimated using an average cost of $36,560 and $63,020 for non-interstate and interstate sites, respectively. This cost was determined based on actual expenditures by MDT for 13 installations completed over the past four years. This cost includes preliminary site visits, equipment, installation, and utilities. Equipment costs for the portable sites are not included in this analysis. The portable sites are expected to be used primarily for enforcement purposes. As use of the portable sites in this capacity was not included in the initial *STARS* evaluation, the potential benefit they may offer in reduced pavement damage from over weight vehicles is also not included in this analysis.

3.3.1.1 Software – Software development costs from the *STARS* program result from the original development of the MEARS software in 1999 and upgrades to the MEARS software completed in 2003. In total, $236,194 was spent on software development over the 4-year span. For other states wishing to implement a *STARS*-like program, the Montana Department of Transportation provides *MEARS* free of charge offering a significant savings on initial costs of the program.

3.3.1.2 Research and Evaluation - In all, three different research/evaluation studies were conducted under the umbrella of the *STARS* program. The first study began in 2000 and consisted of an investigation of the relative performance of piezoelectric and bending plate WIM systems. The purpose of this study was to help guide MDT in future *STARS* equipment investment decisions (Clark, Stephens, and Carson, 2004). The second study evaluated the use of *STARS* data across MDT's various activities, with particular attention focused on a) reducing pavement damage from overweight vehicles by using *STARS* data to help direct enforcement activities, and b) improving pavement designs by using the more accurate data available from *STARS* to characterize vehicle demands on the pavement. The third study (this effort) investigated additional issues not addressed in the original evaluation. The cost of these studies totaled $292,502 over this same implementation period.

3.3.2 Ongoing Annual Costs

Estimates for ongoing annual costs, detailed in Table 3.3.2 –1 below, were provided by MDT largely as projected operating costs based on historical experience. These cost values may be increased or decreased over time as more experience is gained. Totaling the estimated ongoing annual costs related to calibration and maintenance, labor, travel, communications and utilities and other results in a total cost of $297,975 each year.

3.3.2.1 Calibration and Maintenance - Following initial installation and over time (typically twice per year), WIM systems must be calibrated to adjust for accuracy. MDT uses the test truck method for calibration, in which one or more vehicles of a known weight and configuration are driven repeatedly over the WIM system. A range of different vehicle speeds are used in the repeated runs to account for the effects of vehicle speed on the dynamic forces applied to the weigh pad or sensor. The weights recorded by the WIM system are compared to the known weights of the various vehicles to determine system accuracy and required calibration adjustments. General costs associated with this calibration activity (i.e., including the calibration truck and trailer costs but excluding labor and supplies), as well as system maintenance done at the time of calibration, is $38,600.

Table 3.3.2-1. Ongoing Annual Costs of *STARS* (from information provided by Bisom, 2004)

Item	Cost, $
General Costs of Calibration Vehicles and Incidental System Maintenance at the Time of Calibration	38,600
Supplies	
Cables	28,000
Portable Sensors	14,400
Portable Controllers	6,700
General Supplies	5,000
Solar Panels	300
Labor	
Calibration and Maintenance	
FTE (2)	88,480
Data Processing and Analysis	
FTE (1) for data processing and analysis	46,820
FTE (1) for portable WIM data collection	34,500
Travel	20,700
Communications and Utilities	
Communications	13,500
Utilities	475
Other	
Miscellaneous	500
Total	297,975

Weigh-in-motion system maintenance requirements can be categorized as corrective and preventive. Corrective maintenance includes any repairs or replacement of equipment and also includes any roadway-related failure. Preventive maintenance or inspection is performed to circumvent future equipment and/or site problems. Annual maintenance costs for the *STARS* program are estimated to be $54,400, not including labor. This cost primarily comprises replacement of shorter service life system components, piezoelectric cables, sensors and solar panels.

3.3.2.2 Labor - Labor costs attributable to the *STARS* program stem from additional personnel hires in the Planning Division only; no net increase (or decrease) in personnel occurred in Motor Carrier Services Division, the Pavements and Materials Bureau or other areas within MDT. In all, the Planning Division hired four full-time equivalent (FTE) employees: two perform calibration and maintenance activities for the WIM systems, one collects data from the portable WIM systems and one processes and analyzes the resulting data from the WIM systems. Increased labor costs attributable to the *STARS* program are estimated to be $169,800 annually.

3.3.2.3 Travel - Related to the increase in personnel for on-site data collection, calibration and maintenance of the WIM systems statewide, dedicated funds for travel for these personnel are required. In-state and occasional out-of-state travel costs are estimated to be $20,700 annually.

3.3.2.4 Communications and Utilities - Ongoing communications and utilities costs for the *STARS* program statewide (i.e., not included as part of the original installation) are estimated to be $13,975 annually.

3.3.2.5 Other - For unforeseen minor expenditures that may arise throughout the year related to the operation of the *STARS* program, MDT has budgeted $500 annually.

3.3.3 Total Costs Attributable to *STARS.*

Total *STARS* program costs, combining initial costs averaged annually (based on the initial program costs and averaged over reasonable estimates of the service life of the equipment) and ongoing costs results in a total cost ranging from $451,000 assuming a 12-year service life to $757,000 assuming a 4-year service life.

3.4 Benefit To Cost Ratio

With the only quantifiable benefits relating to pavement damage reduction through improved enforcement and improved pavement designs through more comprehensive and accurate data, and totalling $4.8 million annually, and annual program costs ranging from $451,000 assuming a 12-year service life to $757,000 assuming a 4-year service life, benefit to cost ratios for the *STARS* program conservatively range from 6.3 to 10.6 (see Table 3.4.1).

Table 3.3.3-1. Summary, Analysis of Benefits Versus Cost for *STARS*

Item	Results of Benefit to Cost Analysis		
	4-year life	6-year life	12-year life
Total Annual Program Benefits	$4,800,000	$4,800,000	$4,800,800
Total Annual Program Costs	$757,000	$604,000	$451,000
Benefit to Cost Ratio	6.3	7.9	10.6

When considering these estimated benefit to cost ratios, three points should be considered. First, and as previously discussed, the STARS WIM installations are expected to have a service life of at least five to seven years, based on the information that has been collected over the past 10 years on sensor performance. Second, these estimates are likely lower than the actual benefit to cost ratios realized by the *STARS* program due to a) the inability to quantify obvious benefits to the Planning and Motor Carrier Services Divisions and lesser benefits to other areas within MDT and b) the various conservative assumptions made relative to the costs included in the analysis. Third, these estimates reflect conditions immediately following implementation of the *STARS* program; improvements in agency operation are most dramatic and significant at this point in time. As such, the benefit to cost ratios should be re-estimated periodically over time to ensure that the *STARS* program is a continued worthwhile investment.

4 ANALYSIS OF THE *STARS* DATA COLLECTED SINCE THE INITIAL EVALUATION

One concern in the initial evaluation of *STARS* was that only two years of data were available for the evaluation (Stephens, Carson, Reagor, and Harrington, 2003). Notably, highway use can be dynamic in nature, and in a sparsely populated state like Montana with a natural resource based economy, it can change significantly over the course of a year in response to changing economic and environmental conditions (e.g., mine closure, drought, etc.). Thus, the possibility could not be ignored that differences in the characteristics of the overweight vehicle population in the baseline year and in the year of *STARS* focused enforcement resulted from some one of these effects, rather than from the intended variable of interest: enforcement approach. To a great extent, this concern was alleviated by the preponderance of evidence brought forth in the initial evaluation that indicated that targeted enforcement was responsible for the reduction of overweight vehicles in the traffic stream, as discussed in the report on the initial *STARS* evaluation. This potential problem with the evaluation was also addressed as possible prior to the inception of the original work, by making sure that a large geographic area was included in the study, so that localized changes in basic highway use in response to changing economic and environmental conditions would hopefully have only modest impact on the overall study results.

In any event, the *STARS* data collected since the conclusion of the initial program offers an additional opportunity to investigate the trends observed in overweight vehicle operation during the initial evaluation period. At the end of the year of *STARS* directed enforcement, Motor Carrier Services elected in the short term to resume their traditional weight enforcement activities. At this time, the evaluation of the *STARS* directed enforcement effort had yet to be completed, and strategies for the continued use of *STARS* in enforcement had yet to be worked out. Thus, weight enforcement operations in the year following the *STARS* focused enforcement effort were similar to those during the year before the focused enforcement effort (the baseline year), and consisted of the patrol captains selecting and scheduling enforcement activities based on a combination of their experience, knowledge of truck traffic patterns and enforcement intuition. Therefore, the characteristics of the overweight vehicle population would be expected to be the similar in both the year before (the baseline year) and the year after the *STARS* targeted enforcement effort. As both the percent of overweight vehicles in the traffic stream and the excess ESAL miles of travel decreased during the enforcement relative to the baseline year, these values would be expected to again increase during the following year.

4.1 Changes in the Percent of Overweight Vehicles in the Traffic Stream

The percent of overweight vehicles in the traffic stream did increase in the year following the *STARS* targeted enforcement effort, as would be expected if the targeted enforcement activity was responsible for the decline in this value. With the exception of the months of March and April, the percent of overweight vehicles in the traffic stream was consistently less during the year of *STARS* focused enforcement relative to the baseline year and the year following the focused enforcement effort (see Figure 4.1-1). Overweight vehicles comprised 8.8, 6.9, and 8.9 percent of the traffic stream in the baseline year, the year of *STARS* focused enforcement, and following year, respectively. Thus, overweight vehicle activity returned to almost the same level in the year following the *STARS* focused enforcement as was observed in the pre-focused enforcement (baseline) year. The increase in the percent of overweight vehicles in the traffic stream occurred almost immediately following the termination of the focused enforcement effort, as can be seen in Figure 4.1-1. While these results indicate little residual effect of the focused enforcement effort, closer scrutiny of the data by site revealed some possible residual enforcement effects.

Evidence of some residual effect of the focused enforcement effort is seen at those sites that received more than six months of focused enforcement. Shown in Figure 4.1-2 and 4.1-3 is the percent of overweight vehicles in the traffic stream at Townsend and Stanford, respectively, in the baseline year, the year of *STARS* focused enforcement, and the following year. At these sites, which were enforced 10 months during the year of focused enforcement, there are obvious similarities in the pattern of overweight vehicle traffic in the year before compared to the year after *STARS* focused enforcement. In both cases, the percent of overweight vehicles in the traffic stream in the year following focused enforcement remained noticeably below pre-focused enforcement (baseline) levels for three to four months. After this interval, the percent of overweight vehicles in the traffic stream in the year following focused enforcement were similar

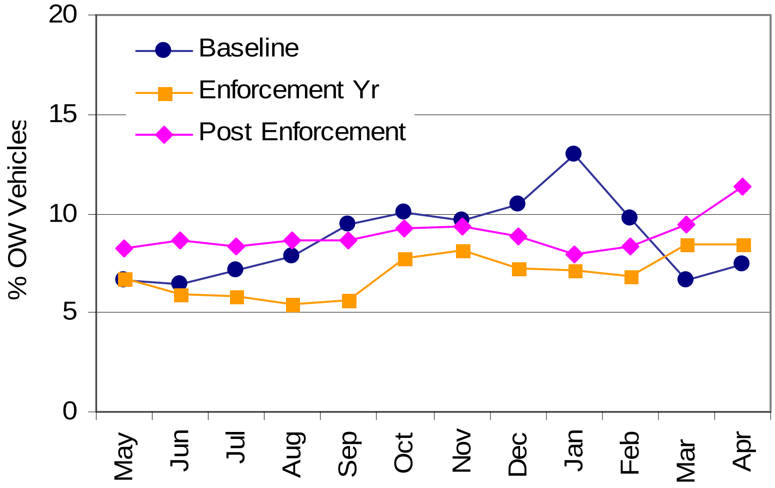


Figure 4.1-1. Percent of Overweight Vehicles in the Traffic Stream Statewide During the Baseline Year, the Year of *STARS* Focused Enforcement, and the Year Following *STARS* Focused Enforcement

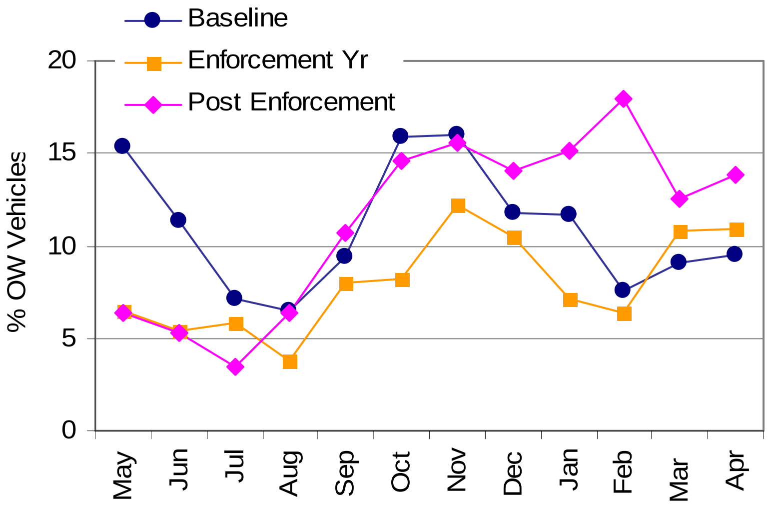


Figure 4.1-2. Percent of Overweight Vehicles in the Traffic Stream at Townsend During the Baseline Year, the Year of *STARS* Focused Enforcement, and the Year Following *STARS* Focused Enforcement

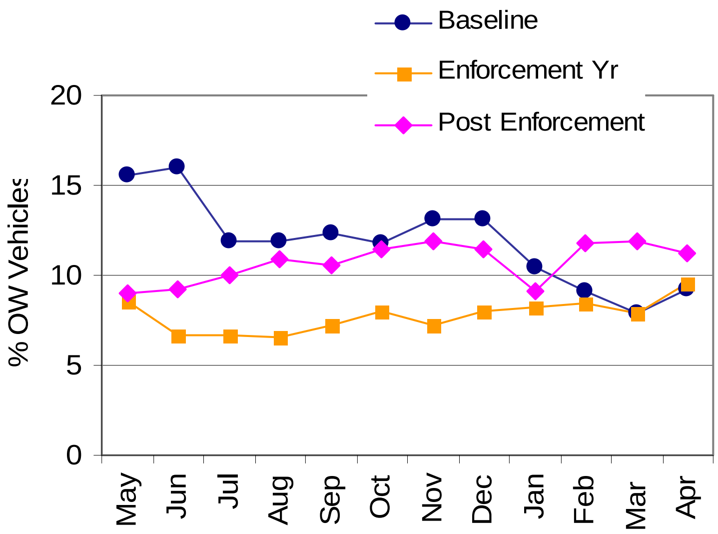


Figure 4.1-3. Percent of Overweight Vehicles in the Traffic Stream at Stanford During the Baseline Year, the Year of *STARS* Focused Enforcement, and the Year Following *STARS* Focused Enforcement

to those seen in the pre-focused enforcement (baseline) year. These results indicate that the focused enforcement effort may have achieved, at least for a short time period, its goal of altering basic loading behaviors in what may have been habitual overweight situations.

4.2 Changes in the Average Amount of Overweight per Vehicle

The average amount of overweight in the baseline year, the year of *STARS* focused enforcement, and the year following the focused enforcement effort were 6100, 5100, and 5400 pounds, respectively. Thus, the average amount of overweight per vehicle in the year following the year of focused enforcement increased toward the levels observed during the baseline year. Once again, this sequence of values is consistent with the focused enforcement effort being responsible for the decline in the average amount of over weight per vehicle during the year of focused enforcement.

4.3 Changes in Pavement Preservation

The pavement damage attributable to over weight on vehicles decreased by approximately 6 million ESAL miles of travel during the year of focused enforcement relative to the baseline year; pavement damage attributable to over weight on vehicles subsequently increased by approximately 5 million ESAL miles of travel during the year following the year of focused enforcement. This trend is consistent with the trend observed in the proportion of overweight vehicles in the traffic stream during the baseline year, the year of focused enforcement, and the year following focused enforcement, and further supports the conclusion that the reduction in pavement damage from overweight vehicles during the *STARS* directed enforcement effort was enforcement related. The cost savings associated with the decrease in overweight pavement damage between the baseline year and the year of focused enforcement was estimated to be approximately $700,000, while the increased costs associated with the increase in pavement damage between the year of focused enforcement and the following year was estimated to be approximately $400,000.

While the excess pavement damage returned to 86 percent of its pre-focused enforcement level, the cost associated with this damage returned to only 60 percent of its pre-focused enforcement level. This difference implies that while the characteristics of the overweight vehicle operations in the years before and after the year of focused enforcement are similar, they are not identical. The disproportionate decrease in cost impacts for the post enforcement year relative to the pre enforcement year is consistent with both a reduction in the average amount of overweight per vehicle and or a shift in over weight operations from local to more intra and interstate routes.

4.4 Conclusions

The results presented above all support the conclusion of the initial *STARS* evaluation, namely, that the focused enforcement effort was responsible for the reduction in overweight vehicle activity and pavement damage during the year of focused enforcement. In the year following the focused enforcement effort, when traditional enforcement strategies were used, overweight vehicle activity returned to pre focused enforcement levels.

5 USE OF *STARS* IN WEIGHT ENFORCEMENT

Potential roles that *STARS* may play in Montana's efforts to control overweight vehicle traffic on the state's highways include evaluating the effectiveness of these activities and/or proactively guiding these activities. These roles are not necessarily independent, in that the results of an evaluation generally serve in some capacity as a guide for redirecting resources in the future to improve effectiveness. Nonetheless, the discussion below will first speak to *STARS* as a tool to evaluate the effectiveness of enforcement activities, which naturally leads into a discussion of its potential role in helping to direct these activities

5.1 *STARS* as a Tool to Evaluate Effectiveness of Weight Enforcement

Through MEARS, *STARS* offers an in-depth characterization of overweight vehicle activity on the state's highways. Two obvious metrics of interest relative to enforcement of the state's vehicle weight laws are the percent of overweight vehicles in the traffic stream, and the damage these vehicles do to the roadways (measured in excess ESAL miles of travel). If the objective of weight enforcement is pavement preservation, then excess ESAL miles of travel offers a more direct measure of the parameter of interest compared to the percent of overweight vehicles in the traffic stream. While percent of overweight vehicles in the traffic stream will be related to the damage caused by these overweight vehicles, this metric does not reflect the importance of the number of vehicles and amount that they are overweight on the excess pavement damage for which they are responsible. None-the-less, percent of overweight vehicles in the traffic stream is an easily understood parameter that does in a general sense reflect the amount of attendant excess damage caused by these vehicles to the state's highways. Thus, use of both metrics, percent of overweight vehicles in the traffic stream and excess ESAL miles of travel may be appropriate. Note that Hanscom (1998) concluded in a thorough evaluation of measures of effectiveness (MOE) of weight enforcement that a) the average value of the excess ESALs per vehicle, b) the average amount by which gross vehicle weights exceed weight limits, c) the percent of vehicles that exhibit excess ESALs, d) the percent of vehicles that exceed gross vehicle weight limits, e) the percent of vehicles that exceed single axle weight limits, and f) the average amount by which tandem axle weights exceed weight limits, were the most useful metrics to consider.

Independent of the specific parameter used to characterize overweight vehicle activity on the state's highways, care must be exercised in assigning cause and effect relationships between the value of these parameters (and any changes in these parameters) and weight enforcement activity. Vehicle operations can fluctuate for a variety of reasons unrelated to enforcement, particularly over short periods of time or at a specific location. Thus, any such evaluation metrics should be more stable and reliable if they are calculated over longer periods of time and larger geographic areas. Whatever metrics are used to assess effectiveness, it further is desirable that any subsequent analyses of changes in these metrics be statistically based, to insure that these changes are significant relative to the general variability in overweight vehicle operations.

5.1.1 Percent of Overweight Vehicles in the Traffic Stream

With respect to using percent overweight vehicles in the traffic stream to evaluate enforcement effectiveness, it is necessary to establish some goal or target with which to compare this metric. In the original *STARS* evaluation, for example, the *STARS* directed enforcement program was judged to be successful in part because the percent of overweight vehicles in the traffic stream decreased statewide during the enforcement year relative to the prior year. While looking at the reduction (if any) in the average percent of overweight vehicles in the traffic stream each year might appear to be a "good" metric to evaluate the effectiveness of weight enforcement, it is probably unrealistic to believe that reductions in this parameter can be realized continuously into the future. That is, as enforcement increases in its effectiveness, additional reductions in the percent of overweight vehicles in the traffic stream will become increasingly difficult to achieve. The marginal costs of achieving additional reductions in the percent of overweight vehicles in the traffic stream will correspondingly become very high, and eventually exceed the value of the associated benefit of reduced pavement damage from overweight vehicle operations.

In light of the above discussion, it may be more reasonable to compare the percent of overweight vehicles in the traffic stream as determined from *STARS*, say annually, against a target value that is economically logical to achieve. Specifically, a cost benefit analysis should reveal the point at which the economic benefit of enforcing weight limits (measured, for example, in terms of avoided pavement damage costs from overweight vehicles) is equal to the cost of the weight enforcement effort. Little published information appears to be available on this subject from other states, and if such information was available, it would have to be closely scrutinized relative to its applicability in Montana. In viewing the results of any such analysis, it is critical to remember that controlling overweight vehicles is only one objective of the state's enforcement effort, and that while it is easy to separately identify and quantify the benefit of this one activity, it can be harder to isolate its specific cost. In any event, performing a comprehensive cost benefit analysis of this kind (i.e., generation of a cost benefit curve for all levels of weight enforcement, from the extremes of excess pavement costs if no enforcement is used, to the cost of enforcement for full compliance) is beyond the scope of this effort. Note that a more limited cost-to-benefit analysis of the incremental cost of the *STARS* program versus its incremental benefit (measured in avoided pavement damage) is included in Section 3 of this report.

In the absence of the aforementioned fundamental cost benefit analysis, establishing target maximum values for the percent of overweight vehicles that should operate on the highway system is an uncertain process. During the baseline year of the original *STARS* effort (under traditional enforcement), 8.8 percent of the vehicles operating on the highway system statewide were overweight (Stephens, Carson, Reagor, and Harrington, 2003). The fraction of overweight vehicles in the traffic stream decreased to 6.9 percent during the year of *STARS* directed enforcement. In the year following the *STARS* enforcement effort, the fraction of overweight vehicles in the traffic stream increased to 8.9 percent. Note that all three of these figures include those vehicles in the traffic stream that are permitted to operate above standard weight limits. Presently, MEARS does not distinguish between over standard weight vehicles that are operating illegally and those that are operating with permits. Thus, permitted over standard weight vehicles are included in the overweight vehicle count.

The inclusion of permitted traffic in the overweight vehicle count is important relative to setting an enforcement target for the maximum percent of overweight vehicles in the traffic stream. Obviously, even if enforcement is completely effective, the MEARS reports will still indicate overweight vehicle activity, if permitted vehicles are operating. Ideally, some mechanism will be developed to either identify these vehicles in the traffic stream, and/or adjust the MEARS figures to reflect their presence. The extent of permitted overweight vehicle operations on Montana's highways is, however, not well known, and thus another objective of this effort was to investigate the permitted vehicle issue. The results of this investigation are presented in Section 6 of this report.

Based on the limited data presented above (and admittedly without adequately investigating the permitted vehicle issue), it might be reasonable to initially set a goal of achieving a maximum of six to eight percent of the traffic stream statewide being overweight on average across the year, as determined by MEARS from the *STARS* data. If the current level of enforcement is judged to be adequate, this target value could be set toward the middle or upper end of this range. As more data on overweight vehicle operations is obtained, it may be both possible as well as more effective to set such targets by vehicle configuration, rather than globally across all vehicles in the traffic stream.

To provide some perspective on the various percentages of overweight vehicles in the traffic stream mentioned above, it has been estimated nationally that 2 to 3 percent of the vehicles in the traffic stream passing operating weigh stations violate weight limits (Livesay, 2004). In his discussion of MOE, Hanscom (1998) reported that Wisconsin estimated several years ago that across all its highway systems, 14 percent of the 5 axle combination vehicles potentially had a weight violation (Stein, 1988). More recently, in a limited study at three sites in Washington, the fraction of overweight vehicles in the traffic stream was found to be around 20 percent (Jessup and Casavant, 1996). At two sites in Idaho, one on and one off of the Interstate system, over an 8 month period, over standard weight vehicles were found to be 13 and 8 percent of the traffic stream, respectively (Idaho Department of Transportation, 2002). Finally, at certain sites around Montana, it has been estimated that up to 10 percent of the vehicles of certain configurations, e.g., logging trucks, might be permitted to operate above standard weight limits (Murphy, 2004); these vehicles would be reported as overweight by MEARS.

If the state is interested in better characterizing the cost benefit of weight enforcement, the level of effort committed to weight enforcement could be purposefully increased or decreased, and the associated effect on the percent of overweight vehicles in the traffic stream and the attendant pavement damage that they cause can be quantified by MEARS. Thus, relationships could begin to be developed between incremental enforcement costs and associated pavement benefits.

Setting a target for the maximum fraction of overweight vehicles in the traffic stream at a statewide level eliminates the need to consider the relative volume of traffic at each specific *STARS* site, as all traffic is included in the statewide aggregated results. Nonetheless, and recognizing once again that enforcement activities have many objectives, some target level of weight limit compliance is probably desirable at all sites. The issue of concern is that it may not be cost effective to target a high rate of weight compliance at a site with very little traffic. To balance statewide average versus local enforcement needs, the target maximum percent of overweight vehicles could be higher for individual sites relative to the statewide average. This approach would allow some flexibility in diverting enforcement resources from low volume to high volume traffic areas, while insuring that some enforcement presence is maintained at all sites.

Once again, setting a maximum permissible percent of overweight vehicles in the traffic stream at any individual site is an uncertain process. A target value of 10 to 12 percent may be reasonable until additional data becomes available in this regard. During the year of *STARS* directed enforcement, the percent of overweight vehicles in the traffic stream at individual sites ranged from 5.5 to 10.3 percent. During the years before and after the *STARS* enforcement effort, the percent of overweight vehicles in the traffic stream ranged from 2.0 to 16.9 percent across the individual sites. Note that the range of compliance levels across the sites decreased during the year of *STARS* focused enforcement relative to the years before and after, which may have resulted from the diversion of resources from sites with high compliance rates to targeted sites with low compliance rates.

5.1.2 Excess ESAL miles of Travel

Returning to the choice of the fundamental metrics to be used in characterizing overweight vehicle operations on Montana's highways, the excess ESALs associated with the operation of such vehicles is an attractive metric, in that it directly quantifies the damage sustained by the highway infrastructure from the overweight they carry. Furthermore, if the cost benefit analysis is going to be pavement damage based, it makes sense to directly use excess pavement damage from overweight vehicles as the primary evaluation metric, rather than percent of overweight vehicles in the traffic stream. MEARS calculates excess ESAL miles of travel associated with overweight vehicles by month and by site, and these values are accumulated across the state and over the year. While this information is critical to establishing the costs associated with overweight vehicle operations on the state’s highway, how it is used as a direct evaluation metric is less clear. Until more data is collected on this metric, it may be reasonable to use percent of overweight vehicles in the traffic stream as well as excess ESAL miles of travel (and their cost) to characterize overweight vehicle operation on the state's highways.

The first issue encountered in using excess ESAL miles of travel to characterize overweight vehicle operations on the state's highways is that the magnitude of this parameter is sensitive to the absolute volume of traffic experienced. This volume dependence makes this parameter both more and less desirable than percent of overweight vehicles in the traffic stream as a metric to measure overweight vehicle activity. This parameter draws attention to routes experiencing the greatest absolute magnitude of pavement damage from overweight vehicles, independent of whether this damage is the result of a high volume of traffic with a low percentage of overweight vehicles, or a low volume of traffic with a high percentage of overweight vehicles. Similarly, it also inherently addresses the relative impacts of a few vehicles operating grossly overweight versus many vehicles operating only slightly overweight. Thus, relative to minimizing the absolute amount of infrastructure damage from overweight vehicles, this quantity is critical to characterizing overweight vehicle operations.

Due to its dependence on the absolute volume of traffic being experienced, care must be exercised in assessing the causes for any changes in excess ESAL miles of travel observed between two evaluation periods. A decrease in excess ESAL miles of travel reported over the year across all the *STARS* sites, for example, could result from a decrease in the volume of traffic statewide for the year (e.g., in response to a reduction in agricultural operations due to drought), more effective weight enforcement, or for some other reason. Currently, MEARS includes a simple adjustment for differences in the overall traffic volume between two different evaluation periods. This adjustment involves normalizing the excess ESALs attributable to overweight vehicles from one period to the traffic volume experienced during the second period. Thus, this adjustment, which is discussed in more detail in Section 7 of this report, allows for the excess ESAL miles of travel collected over two intervals to be compared across the same volume of traffic. While not currently done by MEARS, it may further be useful to normalize the excess ESAL miles of travel by the total ESAL miles of travel experienced during the evaluation period. This metric would allow for comparison of relative overweight vehicle activity across several locations and/or time periods without the necessity of specifically normalizing the comparison to one specific site or time period.

The adjustments to the calculated excess ESAL miles of travel discussed above allow for this metric to be used to compare overweight vehicle activity at two locations and/or over two different time intervals. In the initial *STARS* evaluation, such a process was used in evaluating overweight vehicle activity during the baseline and enforcement years. In the year of *STARS* focused enforcement, for example, a reduction of 6 million ESAL miles of travel was observed relative to the previous year (with due consideration of changes in traffic volumes between the two years). As was the case with percent of overweight vehicles in the traffic stream, however, it is unrealistic to assume that this metric can or should be reduced each year in the future. A point will be reached at which the marginal cost of further reducing pavement damage due to over weight on vehicles exceeds the cost of achieving this level of weight compliance. Once again, a cost benefit analysis would need to be conducted to determine where this balance point lies between the cost of weight enforcement and the value of the benefit it produces in infrastructure preservation. As commented above, such a fundamental analysis is beyond the scope of this effort, although an analysis of incremental cost versus benefit for *STARS* focused weight enforcement was part of this study and is discussed in Section 3 of this report.

In the absence of the above mentioned cost benefit analysis, the only information that can be brought to bear on this problem is the historical data from MEARS on excess overweight ESALs over the past few years. Data has been processed in this regard for a period of three years covering parts of 2000 through 2003, as shown in Table 5.1.2-1. While this data may offer some preliminary information relative to establishing a relationship between the cost of weight enforcement and the value of its benefit in reduced pavement damage, it offers little guidance in establishing a target value, say annually, for the excess ESAL miles of travel experienced by the state highway system. Presently, the annual values of excess ESAL miles of travel have been arbitrarily normalized against the volume of traffic observed in the year of *STARS* focused enforcement. If restricting the amount of excess ESALs experienced annually by the highway system to this value results in providing cost effective protection for the highway infrastructure from overweight vehicles is simply unknown. As mentioned previously, the cost benefit analysis included in this report only partially addresses this situation, in that it only considers the marginal cost and benefits associated with the *STARS* system, not the underlying basic cost of weight enforcement versus total benefit received.

Table 5.1.2-1. Annual Total Excess ESAL Miles of Travel

Year[a]	Total Excess ESAL Miles of Travel[b]
2000	11800000
2001	5830000
2002	10920000

[a] defined as 12 months beginning in May
[b] normalized to traffic volume for 2001

Thus, until additional data and experience is available relative to a metric based on excess ESAL miles of travel, it may be appropriate to use a metric such as percent of overweight vehicles in the traffic stream coincident with some measure of excess ESAL miles of travel. One such measure that may merit immediate investigation is normalizing excess ESAL miles of travel by total ESAL miles of travel during the evaluation period to account for volume of traffic. In any event, excess ESAL miles of travel presently has and will continue to have critical role in calculating the benefit of weight enforcement on infrastructure preservation.

5.1.3 Evaluation of Enforcement Effectiveness

Independent of the metric used to characterize overweight vehicle operations, the absolute values of the metrics and/or changes in their values will be used to infer enforcement effectiveness. Therefore, it is important that these metrics be evaluated in such a manner as to ensure that their values and notably changes in their values actually reflect enforcement effectiveness. To obtain a general characterization of overweight vehicle activity and the effect of enforcement on this activity, it intuitively seems appropriate to aggregate data on overweight vehicle operations from all sites over an entire year. In this manner, short term temporal and spatial variations in overweight activity should not influence the evaluation. That is, for example, if harvest is delayed for a month during the year due to wet weather, or traffic volume decreases on a route due to construction while it correspondingly increases on an alternate route, data on all vehicle movements is still included in the evaluation.

Even when aggregated at the annual level on a statewide basis, broadly felt events, such as a statewide reduction in construction activity, could potentially impact overweight vehicle operations, independent of enforcement effectiveness. The general level of the variability in the loading patterns for all vehicles year-to-year may be revealed to some extent in some of the parameters calculated from weigh station data collected and processed annually for pavement design purposes. Each year, an average ESAL factor is calculated for each vehicle configuration from a sampling of vehicle weights obtained at weigh stations around the state. In reviewing the weighted average of these values across all commercial vehicle configurations (Class 5 through Class 13) for the period 1990 to 2000, the difference between the lowest ESAL factor and the highest ESAL factor is 24 percent, with the maximum change in magnitude between two consecutive years of 6.5 percent. In light of this variation, a 10 year moving average value of vehicle ESAL factors by configuration historically has been used by MDT in the pavement design process.

The variability in the average operating weight of vehicles from year-to-year should be somewhat less than the variability in the ESAL factors referred to above, in that ESALs are a cubic function of weight, and thus the underlying variation in vehicle weight is exaggerated by the variation in the associated ESAL factors. While not mathematically rigorous, an absolute lower bound on the underlying variation in average vehicle weight can be estimated as the cube root of the variation in the ESAL factors. The estimated lower bound on the variation between the highest and lowest average vehicle weight over the same ten year period is then 7 percent, with the change in magnitude between two consecutive years being 2 percent. Assuming that this same variability extends to the population of overweight vehicles in the traffic stream, this variability appears to be of sufficient magnitude that it periodically might be difficult to conclusively determine if enforcement activities or some other effect was responsible for a change in overweight vehicle activity. This general variability in overweight vehicle operations needs to at least qualitatively factor into a) any decision that an observed change in overweight vehicle operations was indeed the result of a change in enforcement activity and b) establishing and interpreting whether or not a maximum target level of overweight vehicle operations has been achieved in any given year.

The problem of appropriately attributing changes in the characteristics of the overweight vehicle population to enforcement activities becomes more acute as the time interval for the evaluation is shortened, and/or the number of sites being considered decreases. While the above discussion focuses on using *STARS* data to characterize overweight vehicle operations across the state over a month or a year, it can be used across any time horizon (e.g., from a few hours, to a few months, to a year) and number of routes (e.g., from a single site, to few sites, to all the sites in the state). Certainly, interest exists in looking at the relative or absolute effectiveness of short term enforcement strategies engaged in at specific sites. The effectiveness such activities might be difficult to discern across a disproportionately long evaluation window or disproportionately large aggregation of sites.

By way of example, the effectiveness of a special two week enforcement activity at Miles City would be difficult to determine by comparing the statewide average annual percent of overweight vehicles in the traffic stream in the previous year with that of the year including the two week enforcement activity. Obviously, such a comparison needs to be performed over a more finite period of time and probably only using the data collected at Miles City. One possible approach to this situation is to compare overweight vehicle activity at Miles City during the two week enforcement activity with the overweight vehicle activity observed during the same two week period the previous year at the same site. If commercial vehicle traffic is invariant year-to-year at Miles City, such an approach may yield reasonable results. That is, changes in overweight vehicle operations during the enforcement interval can be reasonably attributed to the enforcement activity. In reviewing data from *STARS*, consistent traffic patterns year-to-year have been noted at several sites (Murphy, 2004), although a two week interval might still be short for a “between years” comparison, even at a site with stable traffic patterns. If traffic patterns at a site are known to vary from year-to-year (e.g., they are dependent on seasonal or other events whose exact timing may change year-to-year), judgment must be used in addition to changes in the quantitative metrics on overweight vehicle operations from one period to another to conclude if the observed changes result from enforcement or from other sources. That is, a decision must be made on whether or not the traffic from the same "event" was characterized during the evaluation period each year.

One possible solution to this issue of the timing of an event shifting from year-to-year, which thus makes it difficult to capture and compare overweight vehicle activity for the same event across two years, is to not rely on data from a previous event in the evaluation process. For events of reasonable duration (say, a minimum of 3 weeks), it may be possible to compare overweight vehicle activities at different times during the event, itself, to obtain a measure of enforcement effectiveness. That is, baseline data on overweight vehicle activity under normal enforcement could be collected early during the event. The special enforcement activity could then be conducted for some or all of the remaining duration of the event. The effectiveness of the enforcement activity would then simply be judged based on changes in the characteristics of the overweight vehicle traffic between the baseline and enforcement period. The residual effect of enforcement could also possibly be evaluated during the same event, by terminating the special enforcement activity prior to end of the event, and observing the subsequent change in overweight vehicle activity as a function of time.

This same approach to evaluating the effectiveness of enforcement could also be used at sites without any historical data. If an overweight problem is suspected to exist at a particular location, a portable WIM system could be deployed to collect baseline data on overweight vehicle operations at that location (say, for a one to two week period). Data from the same portable system could then be used to determine if any changes occurred in the overweight vehicle population during an ensuing enforcement activity (say of one to two weeks duration).

5.1.4 Statistical Tool for Generic Comparisons of Overweight Vehicle Characteristics Evaluated over Two Intervals

Relative to any of the comparative evaluations mentioned above, in light of the general variability in vehicle loading patterns, some statistical analysis should be done to establish the significance of any observed changes in overweight vehicle operations. A basic test in this regard would be to determine the confidence level at which it can be concluded that the characteristics of the overweight vehicle population are indeed different for the two evaluation periods being considered (e.g., at what confidence level can it be concluded that the percent of overweight vehicles in the traffic stream is not the same in the two evaluation periods) or that a target minimum value of overweight vehicle proportions or weight exceedance has been achieved. In this regard, a general procedure is described to perform basic one and two sample hypothesis tests on the percent of overweight vehicles in the traffic stream or the amount of weight exceedance between two evaluation intervals or as compared to a target minimum value.

5.1.4.1 Comparing Two Means Before and After *STARS* Activities **-** To confirm whether or not *STARS* activities result in a significant change (i.e., reduction) in vehicle loading operations (expressed as the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance), a two-sample t-test can be used. The underlying hypothesis for this test states that the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance prior to *STARS* activities (i.e., before sample) is equal to the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities (i.e., after sample):

H_0: $X_1 = X_2$

H_1: $X_1 > X_2$ (one-tailed test)

where

X_1 is the population mean percent of overweight vehicles in the traffic stream or the mean weight exceedance prior to *STARS* activities

X_2 is the population mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities

This hypothesis test assumes that the samples collected before and after *STARS* activities are conducted are independent. Further, the formulation of the test statistic varies depending on whether the variance of the before and after samples are equal:

H_0: $S_1^2 = S_2^2$

H_1: $S_1^2 \neq S_2^2$ (two-tailed test)

where:

S_1 is the population standard deviation of the pre-*STARS* activity sample and

S_2 is the population standard deviation of the post-*STARS* activity sample.

To confirm this second assumption, an F-test is used:

$$F = \frac{s_1^{\ 2}}{s_2^{\ 2}}$$

where:

s_1 is the standard deviation of the pre-*STARS* activity sample and

s_2 is the standard deviation of the post-*STARS* activity sample.

If F > F Critical one-tail (upper significance level) or F < [1/(F Critical one-tail)] (lower significance level), H_0 is rejected and the two population variances are assumed to be unequal. If however, [1/(F Critical one-tail) (lower significance level)] ≤ F ≤ F Critical one-tail (upper significance level), H_0 is accepted and the two population variances are assumed to be equal. Excel (and most F-distribution tables) give only the upper significance percentile because the properties of the F distribution makes it possible to easily derive the lower significance percentile as [1/F Critical one tail (upper significance level)].

If the variance equality assumption is accurate, the test statistic for the t-test is:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where:

x_1 is the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance prior to *STARS* activities

x_2 is the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities

n_1 is the size of the sample collected prior to *STARS* activities

n_2 is the size of the sample collected following *STARS* activities and

s is the pooled standard deviation for the two samples given as:

$$s = \sqrt{\frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}}$$

where all variables are as previously defined.

Should the assumption of equal variances between the two samples prove to be false, the test statistic for the t-test becomes:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where all variables are as previously defined.

In each case, if t Stat > t Critical one-tail, H_0 is rejected; the population means before and after *STARS* activities are assumed to be unequal and significantly reduced. If, however, t Stat ≤ t Critical one-tail, H_0 is accepted; the population means are assumed to be equal, and the *STARS* activities are assumed to have had no reduction effect on commercial vehicle loading behavior.

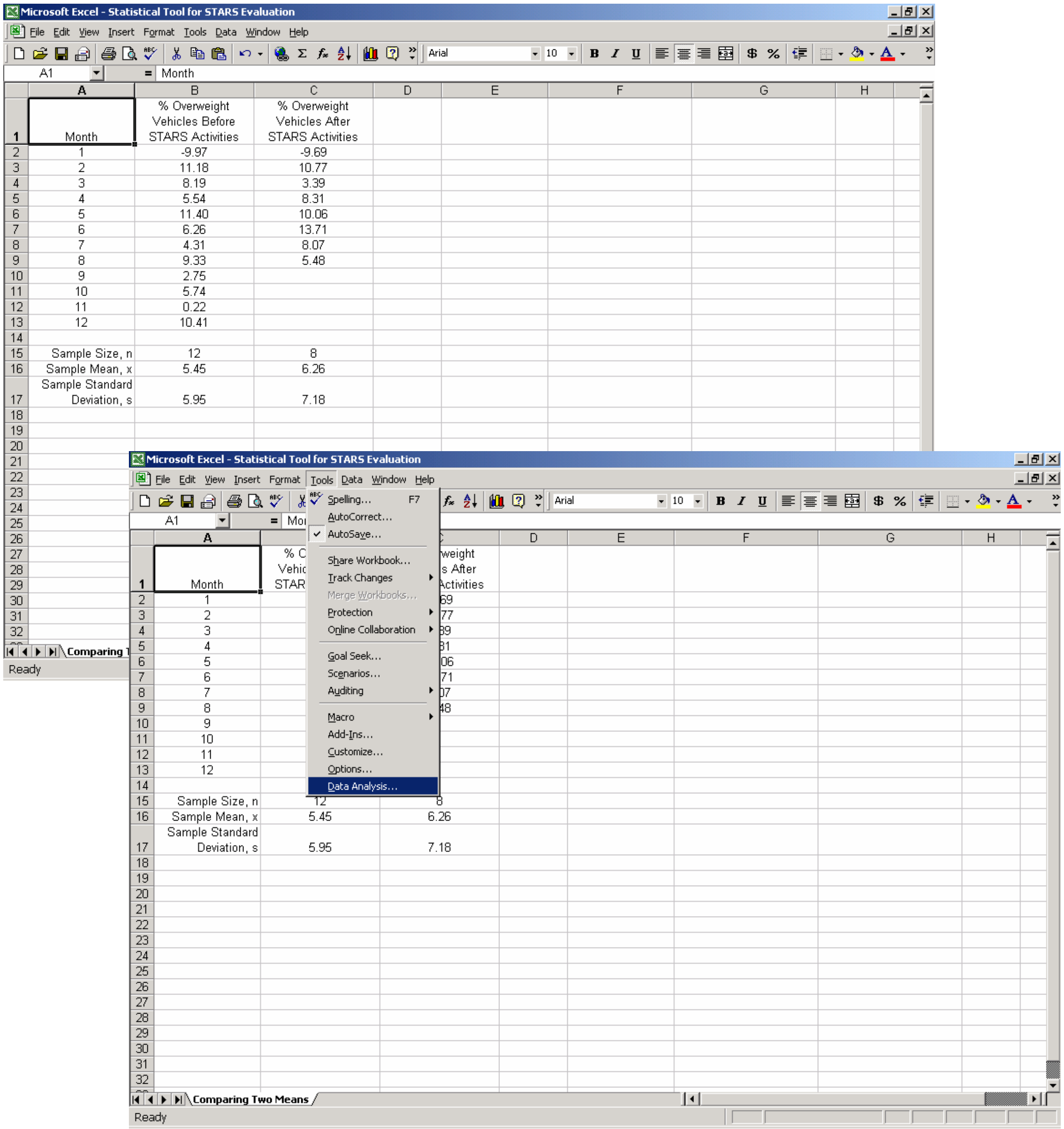
A 95 percent confidence level is commonly used to accept or refute these hypotheses but exact confidence levels at which it can be concluded that *STARS* activities significantly reduced either the percent of overweight commercial vehicles in the traffic stream of the mean percent of weight exceedance can easily be determined as well. For either the two-sample t-test assuming equal or unequal variances, Excel reports the P-value (P(T<=t)) that can be used directly to determine these confidence levels:

Exact confidence level = [1 - P(T<=t) one-tail] x 100

An example of this statistical hypothesis testing application using Excel is demonstrated below. In this example, the statistical difference between the mean percent of overweight commercial vehicles before and after *STARS* activities is investigated. The data is aggregated monthly for unequal numbers of months in the first and second years. The data could be aggregated in smaller time slices, such as weekly or daily, to investigate shorter term, more dynamic events, using the same procedure.

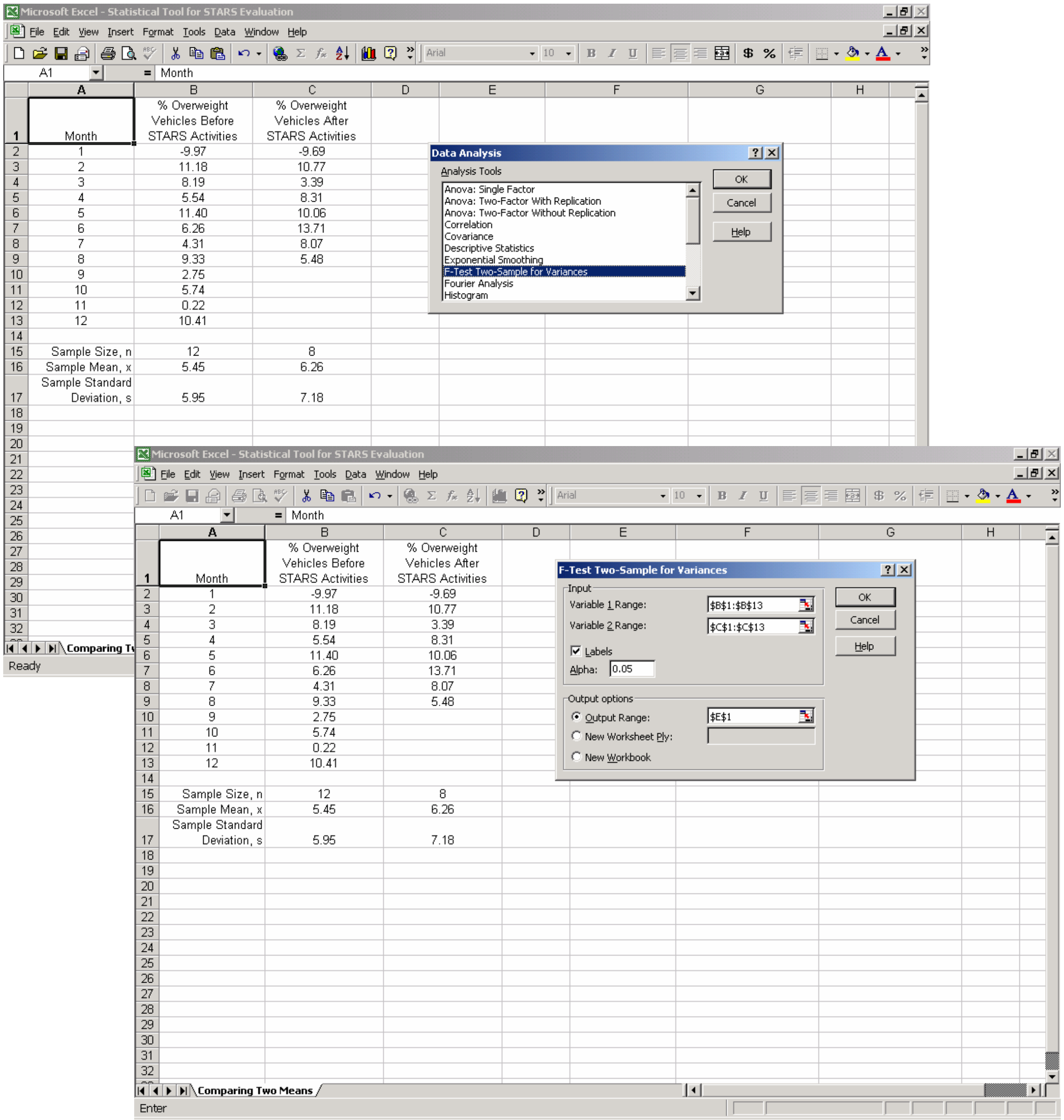
In Excel, select *Tools*, and *Data Analysis*, as shown in Figure 5.1-1. Select *F-Test Two-Sample for Variances* (see Figure 5.1-2). Select the *% Overweight Vehicles Before STARS Activities* as the *Variable 1 Range* and the *% Overweight Vehicles After STARS Activities* as the *Variable 2 Range*. If the first row containing the column headings is included in the selection, be sure to check (✓) the *Labels* box. *Alpha* indicates the confidence level for which the critical F-statistic is reported; a 95 percent confidence level (alpha = 0.05) is the default. The *Output Range* for the F-Test is approximately three columns-by-ten rows and can easily be displayed within the same worksheet.

The output for the *F-Test Two-Sample for Variances* contains general descriptive statistics for each sample, including the *Mean*, the *Variance* (i.e., standard deviation^2), the number of *Observations* (i.e., sample size) and the degrees of freedom (df = *Observations* – 1) (see Figure 5.1-3). The calculated F-statistic (F), the P-value (P(F<=f) one-tail) to determine the exact confidence level for a one-tail test and the critical F-statistic at the default 95 percent confidence level are reported next. For this example, *F > F Critical one-tail* (0.687 > 0.331); H_0 is rejected and the two population variances are assumed to be unequal. Hence, a two-sample t-test assuming unequal variance should be used.



	A	B	C
1	Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
2	1	-9.97	-9.69
3	2	11.18	10.77
4	3	8.19	3.39
5	4	5.54	8.31
6	5	11.40	10.06
7	6	6.26	13.71
8	7	4.31	8.07
9	8	9.33	5.48
10	9	2.75	
11	10	5.74	
12	11	0.22	
13	12	10.41	
14			
15	Sample Size, n	12	8
16	Sample Mean, x	5.45	6.26
17	Sample Standard Deviation, s	5.95	7.18

Figure 5.1-1. Data Analysis Tools in Excel



Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
1	-9.97	-9.69
2	11.18	10.77
3	8.19	3.39
4	5.54	8.31
5	11.40	10.06
6	6.26	13.71
7	4.31	8.07
8	9.33	5.48
9	2.75	
10	5.74	
11	0.22	
12	10.41	
Sample Size, n	12	8
Sample Mean, x	5.45	6.26
Sample Standard Deviation, s	5.95	7.18

F-Test Two-Sample for Variances

Input

Variable 1 Range: $B$1:$B$13

Variable 2 Range: $C$1:$C$13

☑ Labels

Alpha: 0.05

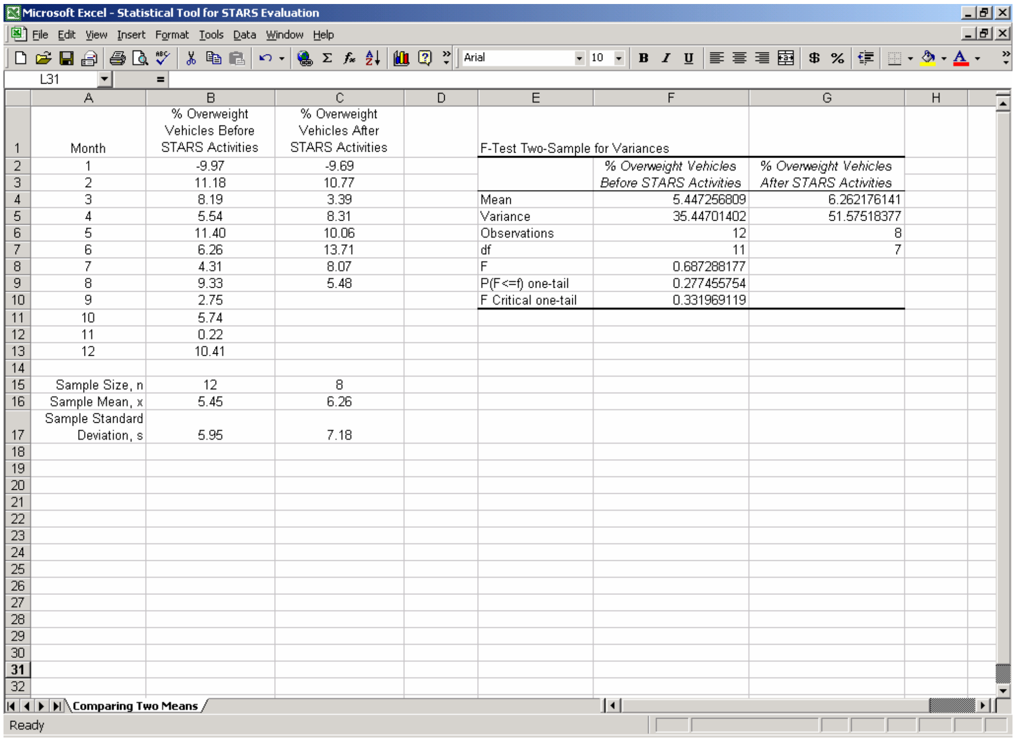
Output options

● Output Range: $E$1

○ New Worksheet Ply:

○ New Workbook

Figure 5.1-2. Excel F Test for Two Sample Variances



Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
1	-9.97	-9.69
2	11.18	10.77
3	8.19	3.39
4	5.54	8.31
5	11.40	10.06
6	6.26	13.71
7	4.31	8.07
8	9.33	5.48
9	2.75	
10	5.74	
11	0.22	
12	10.41	
Sample Size, n	12	8
Sample Mean, x	5.45	6.26
Sample Standard Deviation, s	5.95	7.18

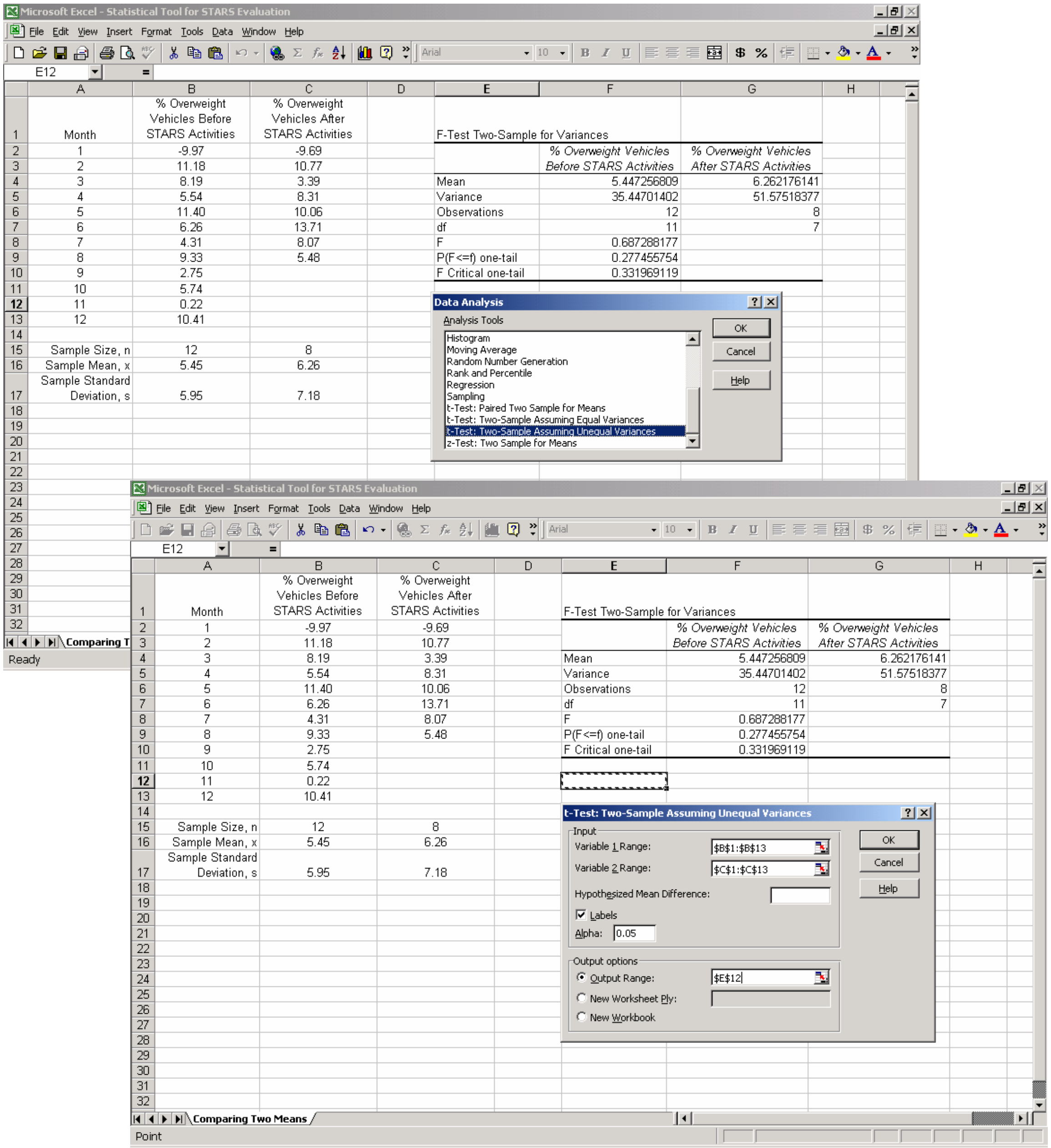
F-Test Two-Sample for Variances

	% Overweight Vehicles Before STARS Activities	*% Overweight Vehicles After STARS Activities*
Mean	5.447256809	6.262176141
Variance	35.44701402	51.57518377
Observations	12	8
df	11	7
F	0.687288177	
P(F<=f) one-tail	0.277455754	
F Critical one-tail	0.331969119	

Figure 5.1-3. Typical Output, F-test for Equivalence of Sample Variance

Using the Tools and Data Analysis menus once again, select a t-Test: Two-Sample Assuming Unequal Variances (see Figure 5.1-4). Again, select the *% Overweight Vehicles Before STARS Activities* as the *Variable 1 Range* and the *% Overweight Vehicles After STARS Activities* as the *Variable 2 Range*. If the first row containing the column headings is included in the selection, be sure to check (✓) the *Labels* box. *Alpha* indicates the confidence level for which the critical t-statistic is reported; a 95 percent confidence level (alpha = 0.05) is the default.

The *Output Range* for the t-Test is approximately three columns-by-thirteen rows and can easily be displayed within the same worksheet (Figure 5.1-5). Unless an actual measurable difference between the two means is of interest (i.e., the *STARS* activities resulted in a 5 percent reduction in overweight vehicle activity), the *Hypothesized Mean Difference* should be left blank.



	Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
1	Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
2	1	-9.97	-9.69
3	2	11.18	10.77
4	3	8.19	3.39
5	4	5.54	8.31
6	5	11.40	10.06
7	6	6.26	13.71
8	7	4.31	8.07
9	8	9.33	5.48
10	9	2.75	
11	10	5.74	
12	11	0.22	
13	12	10.41	
14			
15	Sample Size, n	12	8
16	Sample Mean, x	5.45	6.26
17	Sample Standard Deviation, s	5.95	7.18

F-Test Two-Sample for Variances

	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities
Mean	5.447256809	6.262176141
Variance	35.44701402	51.57518377
Observations	12	8
df	11	7
F	0.687288177	
P(F<=f) one-tail	0.277455754	
F Critical one-tail	0.331969119	

Data Analysis

Analysis Tools

- Histogram
- Moving Average
- Random Number Generation
- Rank and Percentile
- Regression
- Sampling
- t-Test: Paired Two Sample for Means
- t-Test: Two-Sample Assuming Equal Variances
- t-Test: Two-Sample Assuming Unequal Variances
- z-Test: Two Sample for Means

t-Test: Two-Sample Assuming Unequal Variances

Input

Variable 1 Range: $B$1:$B$13

Variable 2 Range: $C$1:$C$13

Hypothesized Mean Difference:

☑ Labels

Alpha: 0.05

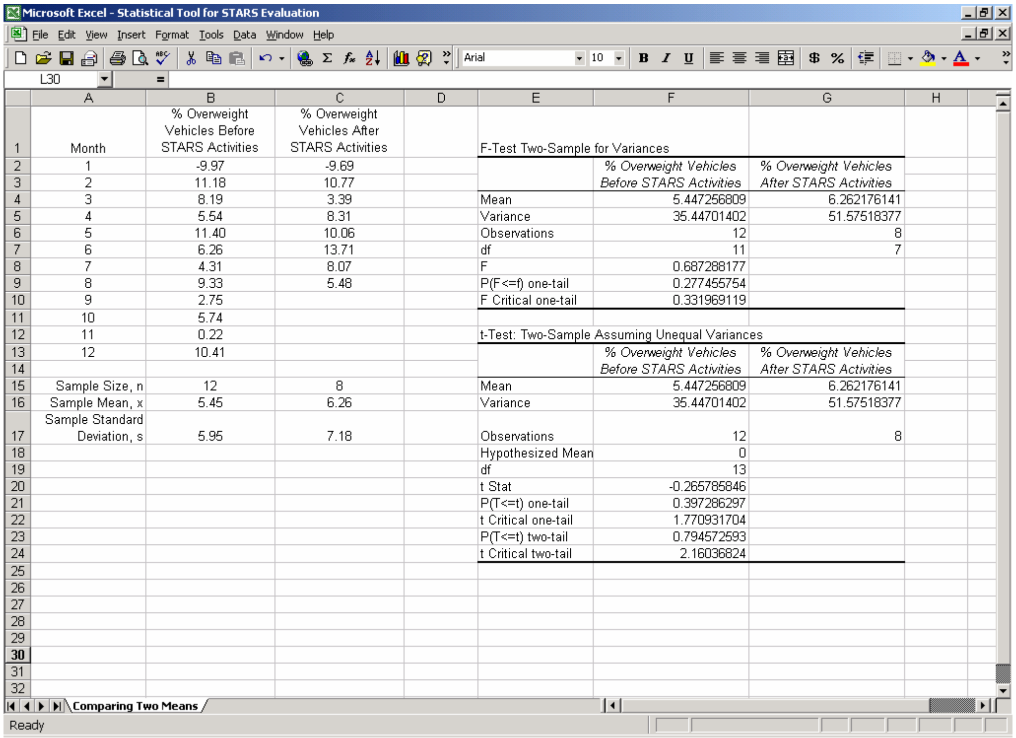
Output options

● Output Range: $E$12

○ New Worksheet Ply:

○ New Workbook

Figure 5.1-4. t-Test for Two Samples with Unequal Variances



	A	B	C	D	E	F	G
1	Month	% Overweight Vehicles Before STARS Activities	% Overweight Vehicles After STARS Activities		F-Test Two-Sample for Variances		
2	1	-9.97	-9.69			*% Overweight Vehicles Before STARS Activities*	*% Overweight Vehicles After STARS Activities*
3	2	11.18	10.77				
4	3	8.19	3.39		Mean	5.447256809	6.262176141
5	4	5.54	8.31		Variance	35.44701402	51.57518377
6	5	11.40	10.06		Observations	12	8
7	6	6.26	13.71		df	11	7
8	7	4.31	8.07		F	0.687288177	
9	8	9.33	5.48		P(F<=f) one-tail	0.277455754	
10	9	2.75			F Critical one-tail	0.331969119	
11	10	5.74					
12	11	0.22			t-Test: Two-Sample Assuming Unequal Variances		
13	12	10.41				*% Overweight Vehicles Before STARS Activities*	*% Overweight Vehicles After STARS Activities*
14							
15	Sample Size, n	12	8		Mean	5.447256809	6.262176141
16	Sample Mean, x	5.45	6.26		Variance	35.44701402	51.57518377
17	Sample Standard Deviation, s	5.95	7.18		Observations	12	8
18					Hypothesized Mean	0	
19					df	13	
20					t Stat	-0.265785846	
21					P(T<=t) one-tail	0.397286297	
22					t Critical one-tail	1.770931704	
23					P(T<=t) two-tail	0.794572593	
24					t Critical two-tail	2.16036824	

Figure 5.1-5. Typical Output for t-Test for Two Samples Assuming Unequal Variances

The output for the *t-Test: Two-Sample Assuming Unequal Variances* again contains general descriptive statistics for each sample including the *Mean*, the *Variance* (i.e., standard deviation^2), the number of *Observations* (i.e., sample size) and the degrees of freedom (df = *Observations* – 1). The calculated t-statistic (t Stat), the P-value (P(T<=t) to determine the exact confidence level and the critical t-statistic at the default 95 percent confidence level are reported next for both one-tail and two-tail tests.

For this example, t Stat ≤ t Critical one-tail (-0.265 ≤ 1.770); H_0 is accepted. Hence, the population means are assumed to be equal at the 95 percent confidence level, and the *STARS* activities are assumed to have had no reduction effect on commercial vehicle loading behavior.

The exact confidence level that defines when H_0 is rejected is [1 - P(T<=t) two-tail] x 100 or 60.27 percent. In other words, it can only be concluded with 60 percent confidence that *STARS* activities resulted in a significant reduction in the percent of overweight commercial vehicles in the traffic stream.

5.1.4.2 Comparing One Mean to a Target Value - In some instances, it may also be of value to not only deduce whether a significant reduction in overweight commercial vehicle activity occurred prior to and following *STARS* activities but also whether that reduction was of such a magnitude to reach a target minimum value (i.e., did *STARS* activities succeed in attaining the target of 5 percent overweight commercial vehicles in the traffic stream?). To confirm the attainment of this target value, a one sample t-test can be used. The underlying hypothesis for this test states that the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities is equal to or less than some desired minimum target value:

H_0: $X_2 \leq \text{Target}$

H_1: $X_2 > \text{Target}$ (one-tailed test)

where

X_2 is the population mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities

This hypothesis test assumes that the sample is normally distributed and that the population variance is unknown.

The test statistic for the one sample t-test is:

$$t = \frac{x_2 - \text{Target}}{s_2 / \sqrt{n_2}}$$

where:

x_2 is the mean percent of overweight vehicles in the traffic stream or the mean weight exceedance following *STARS* activities

n_2 is the size of the sample collected following *STARS* activities and

s_2 is the standard deviation for the sample collected following *STARS* activities

If $t \geq t$ critical one-tail, H_0 is rejected; the population mean after *STARS* activities is assumed to have not met the target minimum value. If however, $t < t$ critical one-tail, H_0 is accepted; the population mean meets the target minimum value and the *STARS* activities are assumed to have had a significant positive effect on commercial vehicle loading behavior.

Referring again to the previous example describing the percent of overweight commercial vehicles in the traffic stream before and after *STARS* activities, now consider whether the percent of overweight vehicle in the traffic stream following *STARS* activities meets an assumed target minimum value of 5 percent. Recall that the sample characteristics for the post-*STARS* activities were as follows:

Sample size, n_2:	8
Sample mean, x_2:	6.26
Sample standard deviation, s_2:	7.18

Using these values, our calculated test statistic becomes:

$$t = \frac{x_2 - \text{Target}}{s_2 / \sqrt{n_2}} = \frac{6.26 - 5}{7.18 / \sqrt{8}} = 0.4964$$

Using a standard t-distribution table for a one-tailed test with a sample size of 8 (degrees of freedom, df = n_2-1 = 7) and a desired 95 percent confidence level ($\alpha = 0.05$), the critical value of t is 1.894579. For this example, $t < t$ critical one-tail ($0.4964 \leq 1.\ 894579$); H_0 is accepted; the population mean meets the target minimum value, and the target level of overweight vehicle activity has been successfully met.

5.2 Focusing Enforcement Activities Using *STARS* Data

The fundamental idea behind using *STARS* to improve the effectiveness of weight enforcement is that it provides specific information on when, where, and what vehicles have historically been operating overweight on the state's highways. This information should be useful in effectively directing weight enforcement resources to those locations with the greatest overweight vehicle problems. As mentioned above, the manner in which *STARS* subsequently factors into guiding weight enforcement activities will be a direct reflection of the metrics being used to evaluate enforcement effectiveness. The metrics suggested above are percent of overweight vehicles in the traffic stream and the excess ESAL miles of travel attributed to overweight vehicle operations. In the simplest sense, enforcement should be focused on those locations with the greatest amount of overweight vehicle activity and attendant infrastructure damage from overweight vehicles. Specific issues that should be considered relative to executing this basic strategy using either percent of overweight vehicles in the traffic stream or excess ESAL miles of travel are discussed in detail below.

5.2.1 Using *STARS* to Focus Enforcement Statewide

One metric advocated above to measure the effectiveness of enforcement is percent of overweight vehicles in the traffic stream. If the target maximum allowable percent of overweight vehicles in the traffic stream at the statewide level is not being achieved, additional attention will naturally be directed toward those sites reporting the greatest percent and greatest number of overweight vehicles in the traffic stream, as determined from the MEARS reports. Obviously, simply directing attention to those areas with the greatest percent of overweight vehicles in the traffic stream may not be the best action, in that if those areas account for only a small fraction of the traffic on the state's highways, decreasing their percent of over weight vehicles will affect only a small portion of the total traffic on the system (and have only a small effect on the statewide average percent of overweight vehicles). Thus, relative to prioritizing locations for enforcement, the total number of overweight vehicles historically observed to operate at a site (which reflects both the percent of over weight vehicles as wells as the total volume of traffic at the site) is a reasonable metric to use. Furthermore, if a limit is additionally set on the maximum percent of overweight vehicles at any given site, at least some level of enforcement will be maintained at all sites, as discussed above.

While the above strategy should work reasonably well in promoting and achieving efficient weight enforcement, one of the primary underlying objectives of minimizing the percent of overweight vehicles in the traffic stream is to minimize the excess damage that they inflict on the pavement. As mentioned above, excess pavement damage from overweight vehicles is more directly related to their excess ESAL miles of travel than to the percent of overweight vehicles in the traffic stream. Notably, the metric of excess ESAL miles of travel takes into account the extent of overweight vehicle travel, the significance of the amount of overweight relative to pavement damage, and the shear volume of overweight vehicles in the traffic stream. Thus, from a pavement preservation perspective, the excess ESAL miles of travel historically generated by the overweight vehicles operating at each site should be used to prioritize weight enforcement efforts.

In the initial *STARS* evaluation, suggestions were made relative to proposed enforcement activities each month based on overweight activity across the state for the same month during the previous year. Sites were prioritized for possible enforcement based on the excess ESALs at each site the previous year, with due consideration of the average magnitude of the overweight on the vehicles generating these excess ESALS. If the average calculated overweight was less than 2 or 3 percent, the site would not be suggested for enforcement, even if the excess ESALs were relative large in magnitude. Such a small average overweight could too easily be the result of inherent system variability or minor calibration problems with the WIM, rather than actually representing a large population of nominally overweight vehicles. Note that to some extent, prioritization based simply on volume of excess ESALs at a site should work well if the marginal benefit of a change in enforcement resource is proportional to the volume of excess ESALs at a site. This relationship between marginal benefit and enforcement effort may be correct, if currently all sites are equally and significantly under enforced.

Nonetheless, unless additional resources are available for enforcement, the strategy used in the initial *STARS* evaluation to guide enforcement will need to be modified if it is to be used on a repeated basis. The problem of using the original strategy in future years is easily understood from a hypothetical example of how it could misdirect future enforcement activities. If a site was identified as a problem site during the baseline year of the original evaluation, for example, it received special enforcement attention during the year of focused enforcement. Presuming that the focused enforcement effort was successful, overweight vehicle activity at this site would no longer appear to be a problem in the following year. Under a fixed level of total enforcement, resources might then be diverted from this problem site to another site. When these enforcement resources are diverted from the site, it is likely that it once again will become a problem site (and potentially be a bigger problem than the one that was solved by diverting the enforcement resource).

The solution to this problem once again may require more fundamental knowledge of the cost benefit of enforcement than is currently available. Ideally, enforcement resources would be allocated to each site such that the benefit to cost ratio of enforcement at each site would be the same (and maximized). That is, if the relationship between the level (cost) of enforcement and the attendant reduction in pavement damage (benefit) was known for each site, this information could be used in an optimization algorithm to allocate enforcement resources across the state to minimize total pavement damage from overweight vehicle operations. The information available from *STARS* presently is inadequate to solve this optimization problem at a system wide level.

STARS does provide the information necessary to improve enforcement effectiveness at an individual site, by indicating those places and times historically identified by *STARS* as having the greatest overweight effects at that site, and it does identify sites at which severe overweight problems exist. *STARS*, however, does not directly reveal at which sites enforcement will offer the greatest system wide benefit. For example, if the decision is made to shift some enforcement resources to one of three problem sites that happened to have had identical excess ESAL miles of travel during a specific month last year, which of the three sites should receive the additional enforcement effort? The site that should be selected is the site at which the redirected resource will have the greatest marginal benefit. This site can only be identified, however, if cost versus benefit information is available for each site.

While fully developed enforcement cost versus benefit relationships would be desirable in this regard, they would be both difficult and expensive to generate. Some rudimentary information on the level of enforcement that was used last year at each site still would be useful in the resource prioritization process. Thus, it may be desirable to introduce some parameter into the MEARS reports that reflects historical level of enforcement effort, to complement the information on historical level of pavement damage attributable to overweight vehicles.

In light of the above discussion, how can *STARS* currently be used in directing enforcement efforts at a statewide level? First, *STARS* can be used to insure all existing enforcement efforts are as effective as possible (which could actually free up some enforcement resource to be directed to problem areas), and second, *STARS* can generally indicate those areas to which more enforcement resources should be diverted. Relative to uniformity in effectiveness, professional judgment can be used to decide if the excess ESAL miles of travel reported by MEARS at a site are reasonable with respect to the volume of traffic and level of enforcement engaged in at that site. A simple approach of this type should be useful at least in identifying and acting on widely disparate levels of enforcement effectiveness between sites and/or within a site at various times during the year. This process could be as simple as categorizing the excess ESALs, level of traffic, and level of enforcement experienced at each site as high, medium, or low, and looking for desirable and undesirable trends in these assessments. For example, a high level of excess ESALs, low level of traffic, and high level of enforcement, would probably be indicative of a less than optimally effective enforcement effort. In such a case, information available from *STARS* needs to be used to redirect enforcement efforts to those times and vehicle configurations responsible for most of the excess pavement damage.

The development of some simple quantitative measures of enforcement effectiveness would further be helpful in this direction. These measures could range from a simple index calculated as the ratio of excess ESAL miles of travel to man-hours of enforcement, to more sophisticated indices, such as percent of excess ESALs per unit volume of commercial vehicle traffic per unit of enforcement effort. Alternatively, similar and possibly simpler indices might be developed in terms of the percent of overweight vehicles at a site. In any case, the significance of the numerical values of these indices initially will be uncertain. As time goes by and they are calculated under various circumstances across the state, it will be possible to attach significance to their value. For example, it might be determined that at an effectively enforced site, the maximum excess ESAL miles of travel is 3 percent per 5,000 ESAL miles of total traffic per 100 patrol hours per month. Once again, if the actual excess ESAL miles of traffic significantly exceeds this value at a site, information available from *STARS* needs to be used to redirect enforcement efforts to those times and vehicle configurations responsible for most of the excessive pavement damage. Analyses that include the level of effort and cost of enforcement can either be added to MEARS, or developed as a separate post processor to the MEARS program.

In any case, and as mentioned above, in identifying sites with the greatest overweight problems at the statewide level, consideration should be given to the percentage of overweight vehicles at each site, the absolute number of overweight vehicles, the average amount of overweight, and the excess ESALs miles of travel accumulated at each site. Of these various parameter, excess ESAL miles of travel may be the most useful in identifying problem sites, presuming that these excess ESALs are associated with vehicles reasonably (rather than just nominally) over weight. In using any of these metrics to prioritize sites for enforcement, it is important to realize that MEARS currently calculates the value of the benefit realized by these efforts simply based on pavement preservation (reduction in excess ESAL miles of travel). Notably, any other benefits that accrue from weight enforcement of either a direct or indirect nature are ignored. If other benefits are judged to be important and substantial, presently they need to be factored into the enforcement prioritization process outside of MEARS, either qualitatively or quantitatively.

5.2.2 Using *STARS* to Focus Enforcement at Specific Sites and Events

Returning to the fundamental premise on why *STARS* should be useful in vehicle weight enforcement, it does provide historical information on the specific locations that vehicles are operating overweight, which configurations are operating overweight, and at what times they are operating overweight. One issue in using historical data to direct enforcement is the general temporal variability of the loading patterns for vehicles year-to-year. This issue of the variability in overweight vehicle movements was previously discussed relative to using *STARS* as a tool to evaluate enforcement effectiveness. The same problems identified in that earlier discussion of using *STARS* as an evaluation tool come into play in using *STARS* to guide enforcement activities. That is, for example, a peak in overweight vehicle activity identified by MEARS at a particular site during a specific week last year may occur this year at some other time (say, for example, a week later). Thus, a focused enforcement effort targeting that event that was scheduled based on last years MEARS information would simply miss the event.

The problem of events recurring at slightly different times each year, and thus the possibility of mis-scheduling enforcement activities designed to capture these events becomes more acute, once again, as the duration of the specific event and/or proposed enforcement activity decreases. In an experiment conducted in the Summer of 2003, for example, MDT scheduled a week of weight enforcement activities at three portable WIM sites based on over weight activity observed at these sites the previous year. Data collected during the week before and the week after the enforcement effort was compared with that collected during the week of focused enforcement. These comparisons (in terms of excess ESAL miles of travel, normalized to a constant volume of freight carried at each site) are presented in Figure 5.2-1. The excess ESAL miles of travel noticeably decreased during the week of focused enforcement at only one of the six trial sites (Site 311). While one conclusion would be that focused enforcement generally is ineffective, a second conclusion would be that it is ineffective when implemented using this scheduling strategy. That is, vehicle operations at the sites in question may be too dynamic year-to-year to focus enforcement on a single week based on overweight vehicle activity during that week, one year earlier.

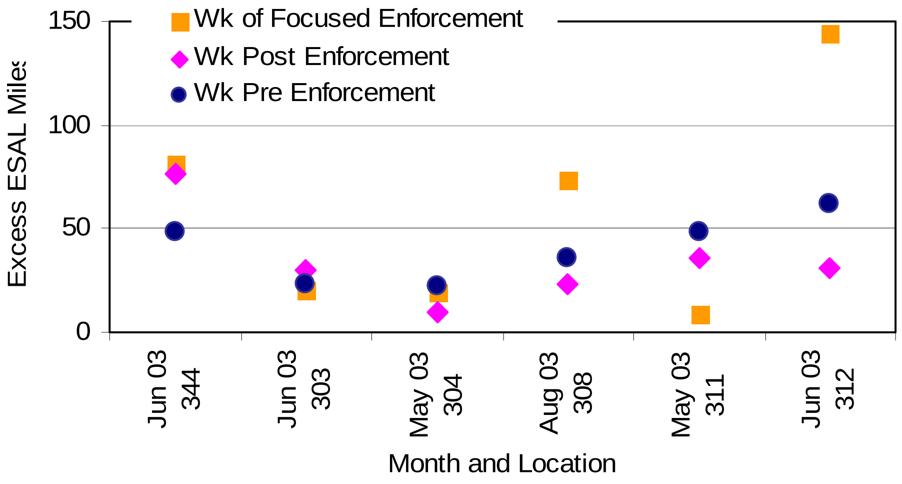


Figure 5.2-1. Excess ESAL miles of Travel, Focused Enforcement Trials, Summer 2003

MDT conducted a second series of tests in which focused enforcement was engaged in on a regular schedule over a two week period at selected sites, based on MEARS data for the same period, one year earlier. The effectiveness of these enforcement efforts was subsequently evaluated by comparing the excess ESALs during an interval including the enforcement effort with the excess ESALs for the same interval one and three years earlier. The results of the comparisons relative to the excess ESALs for the same period one year earlier are presented in Table 5.2.2-1. The comparative periods used in calculating the excess ESALs consisted of one month commencing with the two week enforcement effort, the calendar month containing the two week period of focused enforcement (which may or may not have been at the beginning of the month), and the two week period of focused enforcement.

Referring to Table 5.2.2-1, either the strategy that was used to help guide enforcement during these tests was generally unsuccessful, or the evaluation of the results of these tests is flawed. Similar to the results reported above from the experiments using portable WIM systems and a single week of focused enforcement, in only two out of five cases did the excess ESALs decrease as a result of the two week focused enforcement effort. Without further information on the specific manner in which the focused enforcement activities were selected, it is difficult to definitively comment on whether problems in the selection process or with the evaluation process are responsible for the apparent ineffectiveness of the enforcement activities. While it previously was stated that problems with the evaluation process should be reduced if the duration of the evaluation is increased in length (as was done in analyzing the data from these tests, by using a one month evaluation with different start times, versus a two week evaluation period), the outcome of the evaluation in these tests was the same across all the evaluation windows considered. That is, the two sites that experienced a decrease in excess ESALs relative to the previous year did so across all evaluation intervals (two weeks, one calendar month, and one month starting coincident with the beginning of the focused enforcement effort); similarly, the sites that experienced an increase in excess ESALs relative to the previous year did so across all evaluation windows.

The mixed results obtained from this effort may well highlight the complexity of the problem being addressed, and the difficulty in developing a single strategy to guide and evaluate weight enforcement using *STARS* data that is appropriate in all situations. In this case, a thorough review of the conditions at each of the sites might reveal why the *STARS* directed enforcement activities appear to have been successful at some of the sites and a failure at other sites. To be useful, this review needs to cover conditions related to traffic and enforcement levels both during the current year, as well as last year. Such a review might help in developing insights relative to what strategies are effective under different conditions, and what factors can be used to identify the appropriate strategy to be used at a given site.

Table 5.2.2-1 Change in Excess ESALS from Two Weeks of Focused Enforcement Scheduled Based on Historical Data

	Change in Excess ESALs		
	Enforcement Month Compared to Same Month in Previous Year		Two Week Enforcement Window Compared to Same Two Week Period in Previous Year
Site	Month Begins on the First Day of Enforcement	Month Begins on The 1st Day of the Month	
Paradise	-212	-180	-105
Bad Route	-4098	-4034	-1541
Townsend	456	1068	719
Miles City East	33	23	11
Arlee	610	1443	142

While the focus of the above discussion has been on scheduling current enforcement activities based on overweight vehicle operations during the same period of time, one year earlier, information from *STARS* should also be useful in responding in the short term to specific episodes of overweight vehicle operation identified by field personnel or through other mechanisms. The basic strategy for using *STARS* in such cases remains the same as that discussed above, that is, based on historical data, enforcement can specifically be scheduled at those times at which the greatest pavement damage from overweight vehicles is experienced. In this case, the “historical” data may be collected over an interval immediately prior to the proposed enforcement period. Note that this baseline data collection interval needs to be long enough to clearly establish the general pattern of overweight vehicle operations at the site, rather than capturing a short term variation of this pattern. The overweight episode needs to subsequently be of sufficient duration that this general pattern of overweight vehicle operations will be repeated during the proposed enforcement period. The enforcement period, itself, needs to be sufficiently long to ensure its apparent effects are not diminished or enhanced by short term variations in overweight vehicle activities.

Unfortunately, little guidance is available on what constitutes a “long enough” period to characterize baseline overweight vehicle operations at a site or to insure that an enforcement activity is in place to influence a representative segment of the traffic stream. The requisite time for these activities will depend on the characteristics of the over weight vehicle operations at any given site, as discerned by professional judgment. In general, due to the variability of traffic conditions, under estimating the durations for these activities would be expected to have a more detrimental impact on enforcement effectiveness and its evaluation, than overestimating their duration.

If the above criteria/conditions are met, *STARS* focused enforcement would be expected to be effective, and this effectiveness should be confirmed by comparison of the excess ESALs experienced during the enforcement period relative to those during the pre-enforcement, baseline interval. Any residual effects of the focused enforcement effort could also be investigated by reviewing the excess ESALs experienced during successive intervals following the focused enforcement effort. Such a review might also suggest an “optimum” return interval for focused enforcement.

5.3 Summary: Use of *STARS* in Weight Enforcement

The data collected by *STARS* and processed through MEARS provides an in-depth characterization of overweight vehicle activity on the state's highways. This information should be useful to MDT both in evaluating the effectiveness of enforcement, as well in guiding enforcement activities. Relative to evaluation, information from MEARS can be used to assess whether or not the state's enforcement efforts are adequately controlling overweight vehicle activity on the state's highways. Statewide target levels of percent overweight vehicles in the traffic stream and percent of excess ESAL miles of travel can be established, and MEARS can be used to see if these target levels of performance are achieved. Of these two parameters, while excess ESAL miles of travel may be more difficult to interpret, it provides a more direct measure of the ultimate parameter of interest, namely, pavement damage from overweight vehicles, than is provided by percent of overweight vehicles in the traffic stream. The value of the benefit realized from STARS focused enforcement presently is only calculated in terms of reduced excess ESAL miles of travel.

In any event, on a more micro-level, MEARS can also be used to evaluate if individual enforcement activities are effective, through comparison of overweight vehicle activity before, during, and after the activity. Such evaluations of effectiveness are more difficult to perform, as care must be exercised to insure that 1) any observed changes in overweight vehicle activity can be directly attributed to the focused enforcement, and 2) that the evaluation is set up to accurately capture any change in overweight vehicle operations that resulted from the enforcement activity. The multitude of variables that can influence overweight vehicle operations independent of enforcement are a source of problems in this regard. In particular, the dynamic nature of overweight vehicle operations can make it difficult to select the time intervals over which to perform the evaluation.

One element not included in MEARS is an indication of the level of enforcement effort associated with achieving a certain level of control on overweight vehicle operations. Thus, the cost effectiveness of various enforcement activities is not explicitly being evaluated. This absence of information on the level of enforcement effort versus benefit realized becomes particularly problematic in using *STARS* in planning enforcement activities. In the first *STARS* evaluation, the assumption was made that in an under enforced environment, the marginal benefit per unit of enforcement activity was directly proportional to the volume of overweight activity at a given site. While this approach to directing enforcement activities was effective in that evaluation, it is expected to be less useful and appropriate as *STARS* is increasingly used to improve enforcement effectiveness. In redirecting enforcement resources in the future based on *STARS* information, it will become increasingly important to factor in the level of enforcement activity historically used at each site. Thus, it seems appropriate to either include some parameter in MEARS that reflects level of enforcement effort, or to develop a secondary program that acts on the MEARS data, coupled with level of effort information, to optimize enforcement resource allocation. This parameter could range from a qualitative assessment of low to high enforcement, to a quantitative measure, such as excess ESAL miles per hour of enforcement.

In general, professional judgment is required in using information from *STARS* to evaluate enforcement effectiveness and to help plan enforcement activities. In light of the complexity of the problem, consideration always needs to be given to whether or not the comparison being made will reasonably reveal cause and effect relationships between enforcement and overweight vehicle activity. As additional experience is gained with *STARS*, and successful application strategies are identified, it should become possible to directly program some of this “professional judgment” into the data processing routines.

6 IDENTIFICATION OF PERMITTED VEHICLES FROM *STARS* WIM DATA

Presently, MEARS simply identifies as overweight, any vehicle whose axle weights or gross vehicle weight exceeds standard weight limits. Montana, however, does allow standard axle and GVW weight limits to be exceeded within certain parameters, generally if a non-divisible load has to be transported on the state's highways and this load results in the gross vehicle and/or individual axle weights exceeding standard limits. Identification of these vehicles by MEARS as being overweight is often problematic relative to using information from MEARS to assess overweight vehicle activity. In prioritizing enforcement activities, for example, the excess ESAL miles of travel indicated by MEARS at some sites could result predominantly from permitted vehicle activity rather than illegal overweight vehicle activity, and thus these sites should not be considered for focused enforcement.

One solution to the problem of identifying the permitted vehicles in the traffic stream would be for each of these vehicles to be equipped with a device that actively informs the WIM system of its permit status. This information would be attached to the data record and subsequently recognized and appropriately acted on by MEARS. This solution would generate fairly accurate results, although with it would increase both the complexity and expense of the permitting process and require changes to both the *STARS* hardware and software. Relative to other commercial vehicle transponder operations, this application might be more easily accomplished, in that only one simple piece of information is being communicated between the vehicle and the receiver. In any event, in this investigation a less expensive (and probably less accurate) solution to this problem was researched, namely, developing a technique to adjust the MEARS data after (or as) it is collected to approximately account for the presence of permitted vehicles in the traffic stream.

6.1 Proposed Methodology to Account for Permitted Vehicle Traffic

The methodology proposed in this study to adjust the WIM data collected by *STARS* to account for weight permitted vehicles in the traffic stream is founded upon two assumptions:

1) on average across some period of time, permitted vehicle traffic at any given *STARS* site is constant and repeatable in nature, and

2) during visible and obvious weight enforcement activities, those vehicles whose axle weights and or gross vehicle weight exceed standard limits are predominantly permitted vehicles.

The first of these assumptions allows for an adjustment factor to be developed from a sample of the vehicle traffic at a site that can subsequently be applied continuously across the traffic at that site. The second assumption provides a mechanism for developing such an adjustment factor.

The reasonableness of the first assumption depends on how uniform and recurring the pattern of permitted vehicle traffic is at a given site. While definitive information of this kind presently is unavailable, the weight enforcement community may have sufficient knowledge of general vehicle operations at each site for the adjustment factor approach to still reasonably be applied. Obviously, if the pattern of permitted vehicle traffic is uniform and recurring, an adjustment factor developed from a discrete sample of this traffic can be used with good results across any subsequent time interval. A site with a high volume of traffic carrying diverse goods will be less sensitive to the seasonal vehicle movements, and thus possibly will have relatively uniform and recurring permitted vehicle traffic.

As the timing and/or volume of traffic on a route becomes more variable, obtaining a representative sample of this traffic from which to characterize permitted vehicle activity is more difficult. Furthermore, the applicability of a single adjustment factor calculated for a short period of time across all subsequent times, during which the character of the traffic can change substantially, is questionable. Various approaches can be used to generate better results under such circumstances, including:

1) Adjustment factors can be developed for each of the primary traffic scenarios that exist at a given site. These factors would then be applied in MEARS to the data collected during each type of traffic event. A route that carries significant construction traffic, for example, could have one adjustment factor that is applied during the construction season, and a second factor that is used during the off-season.

2) A single adjustment factor can be calculated from a vehicle sample that was carefully selected to accurately represent the characteristics of the permitted vehicle traffic experienced over a year. Following this approach, temporal variations in permitted vehicle traffic during the year are accurately accounted for, but only in the final results calculated over the entire year.

Presently, in the absence of detailed information on permitted vehicle operations, determination of what adjustment factor is appropriate and useful at a given site (i.e., single value valid for any time period, multiple values based on time of year and type of vehicle activity, single value applied only to complete year(s) of data, etc.) is a matter of professional judgment. It may be possible to develop just a few adjustment factors that are used across a group of sites that experience similar commercial vehicle traffic (e.g. Interstate, seasonal non-interstate NHS, etc.).

The second assumption introduced above, namely, that during visible and obvious weight enforcement activities, those vehicles whose axle weights and or gross vehicle weight exceed standard limits are predominantly permitted vehicles, allows for relatively simple identification of the permitted vehicles in the traffic stream. The basic and probably obvious rational behind this assumption is that over standard weight vehicles that are operating illegally will not enter the traffic stream if they know that they will be weighed and ticketed. Thus, any vehicle whose axle weights or gross vehicle weight exceeds standard limits when enforcement is present generally will be a permitted vehicle. The qualifier "generally" is used in this instance because it has been observed that even when weigh stations are open, one to three percent of commercial vehicles will be overweight. Thus, even in the presence of obvious enforcement, some of the over standard weight vehicles in the traffic stream will not be permitted vehicles.

If the second assumption above is valid, than WIM records collected during a period of overt enforcement can be used to characterize permitted vehicle activities at a site. That is, any record in which axle weights or gross vehicles weights exceed standard limits corresponds to a permitted vehicle, with due consideration of the 1 to 3 percent of vehicles that operate illegally during overt enforcement activities. The number of these vehicles in the traffic stream can then be divided by the number of vehicles with weights below standard limits to obtain an adjustment factor to be used with the MEARS data.

Note that the objective of the “overt” enforcement activity referred to above is simply to produce a traffic stream containing only weight compliant vehicles. Thus, this activity can consist of any tasks that will insure this condition. Notably, not every truck has to be weighed or even stopped during this activity.

6.2 Evaluation of Proposed Methodology

The methodology proposed herein for identifying permitted vehicles in the traffic stream was investigated using data available from the piezoelectric WIM system located adjacent to the Mossmain weigh station just west of Billings, MT on Interstate 90. This weigh station, located immediately downstream of the WIM system in the westbound direction, is open 24 hours per day, thus any WIM record with weights in excess of standard limits was presumed to be for a permitted vehicle. This site was selected to try this methodology in that a) the enforcement presence is well known and continuous, b) WIM records were readily available from another study being conducted at this site on WIM performance (Clark, Stephens, and Carson, 2004), and c) MEARS reports were also available for the traffic at this site. The characteristics of the over standard weight vehicles identified in the traffic stream from the WIM data were analyzed relative to the limited information that is available on permitted vehicle activity. This information consists of expert opinion and quantitative information on the number and type of over standard weight permits issued annually.

If the strategy proposed herein for identifying permitted over standard weight vehicles in the traffic stream is valid, the overweight vehicle activity at Mossmain being reported by MEARS for the Mossmain site should consist almost entirely of permitted vehicles, as a continuously operating weigh station is located immediately adjacent to the *STARS* site. Typically MEARS has identified 5, 20, and 2 percent, respectively, of the Class 9, 10, and 13 vehicles passing Mossmain in the westbound direction as overweight. Unfortunately, and as previously mentioned, little information is available to confirm if this volume of permitted overweight vehicles in the traffic stream at Mossmain is reasonable. Motor Carrier Services personnel at MDT have indicated that Class 10 vehicles are a common configuration for moving non-divisible over standard weight loads (Hult, 2004, Murphy, 2004). This observation is consistent with the greatest proportion of over standard weight vehicles at Mossmain being Class 10 vehicles. The head of Motor Carrier Services' patrol officers further estimated that up to 40 percent of Class 10-3 vehicles might be weight permitted (Murphy, 2004), which is the same order of magnitude as the fraction of over standard weight Class 10 vehicles (across all subclasses) observed at Mossmain.

The characteristics of the over standard weight vehicles identified in the WIM data at Mossmain were studied in greater detail in an effort to more thoroughly determine if these vehicles correspond to permitted vehicles. The specific data used in this investigation was collected in the westbound lanes of the interstate on 23 days over a two year period. The data was purposefully collected on days with varying weather conditions relative to temperature and precipitation, as part of a separate study being conducted at MSU on WIM performance (Carson and Stephens, 2004). Note that the data was all collected on weekdays between the hours of 8:00 a.m. and 5:00 p.m.; thus, the sample may be biased relative to time-of-day and day-of-week trends in vehicle operations. The sample includes 4806, 204, and 570 records from Class 9, 10 and, and 13 vehicles, respectively (see Appendix A for vehicle descriptions by Class).

The weight records from the above sample were screened by several discrete levels of over standard weight. In Montana, single and tandem axles are allowed to carry a maximum load of 20,000 and 34,000 pounds, respectively. The maximum allowable loads on other axle groups (i.e., tridems and quadrums) are determined according to the Federal bridge formula. Individual axle group loads may be further restricted when used in any given vehicle configuration, based the application of the Federal bridge formula to that configuration. The maximum standard gross vehicle weight of a Class 9 vehicle is 80,000 to 86,000 pounds, depending on the specific axle geometry. The maximum standard gross vehicle weight of a Class 10 vehicle is 88,000 to 121,000 pounds, again depending on the specific axle geometry. The Class 13 vehicle encompasses several configurations, with the maximum standard gross vehicle weight being on the order of magnitude of 120,000 pounds to 132,000 pounds.

The WIM records for each type of vehicle were screened at successively increasing weight levels, and the number of vehicles operating at each level was noted. The screening levels were set based on MDT's term non-divisible load permit schedule (5,000 pound increments up through 30,000 pounds of over standard gross vehicle weight, with a final 10,000 pound increment to

40,000 pounds) up to 40,000 pounds of gross over standard weight. Screening was done in increments of 20,000 pounds above 40,000 pounds of over standard weight. Screening was also done based on the over standard weight carried on individual axle groups within vehicles. Various types of permits allow for axle group weights to exceed standard limits by 5,000 to over 14,000 pounds. Screening by axle weight and gross vehicle weight generally yielded similar results, and thus only the results of screening by gross vehicle weight are presented in this report. The results of screening by axle weight might reveal trends relative to the specific types of permits being used by industry, and thus may merit further investigation.

One, twenty-one, and one percent of the Class 9, 10, and 13 vehicles, respectively, traveling westbound at Mossmain during the sample period were over standard weight. These percentages are similar to those generated by MEARS (which were reported above to be approximately 5, 20, and 2 percent, respectively). Some differences would be expected between the results obtained by MEARS, which are based on continuous, 24 hour per day sampling, relative to the discrete sampling used to obtain the data set considered in this study.

While based on the available information on permitted vehicles it is difficult to assess if the absolute number of over standard weight vehicles observed in the WIM data at Mossmain could reasonably be the number of permitted vehicles operating at this location, it is possible to determine if their distribution by amount of over standard weight is consistent with the distribution of over standard weight for permitted vehicles. In this regard, information is available on the number, type and amount of over standard weight for which permits are issued by MDT (Hult, 2004). Relative to type, Montana offers term and single trip permits. The available information on term permits consists of the number of permits issued by amount over standard weight (without reference to vehicle configuration). The available information on single trip permits consists of the amount of over standard weight for each vehicle by axle group and gross vehicle weight. For both types of permits, frequency distributions were developed for the amount that permitted vehicles exceed standard gross weight limits. These distributions were developed for term permitted vehicles as a group, and for single trip permitted vehicles by category of 5 axle, 6 axle, and 7 axle and more vehicles (approximately corresponding to Class 9, 10, and 13 vehicles, respectively). These distributions were developed using permit data from 2003.

The over standard weight distributions for the Class 9, 10, and 13 vehicles in the WIM data set from Mossmain are compared to the over standard weight distribution of all term permitted vehicles in 2003 in Figures 6.2-1, 6.2-2, and 6.2-3, respectively.

The weight distributions for the over weight vehicles operating at Mossmain and that of term permitted vehicles (for the entire state) are very similar in shape. In the case of Class 9 and 10 vehicles (Figures 6.2-1 and 6.2-2), the greatest fraction of over standard weight vehicles occurs in the 0 to 5000 pound range, followed by gradual decrease in the fraction of vehicles in each weight range to the 10000 to 15000 pound category, with a second peak in over weight activity in the 15000 to 20000 pound weight range.

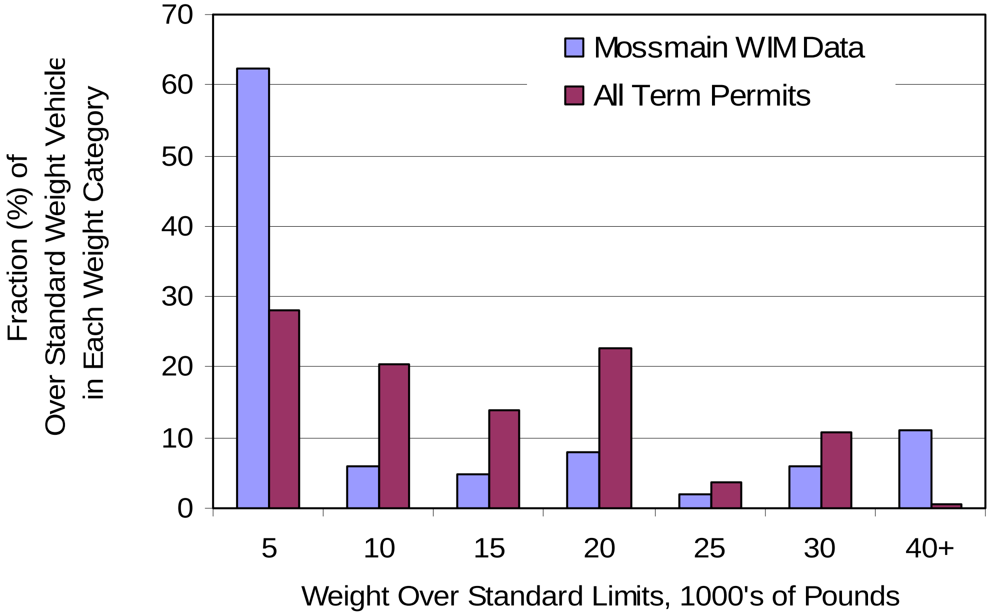


Figure 6.2-1. Fraction of Over Standard Weight Vehicles by Weight Range, Class 9 Vehicles at Mossmain (as identified from WIM data) and All Term Permitted Vehicles

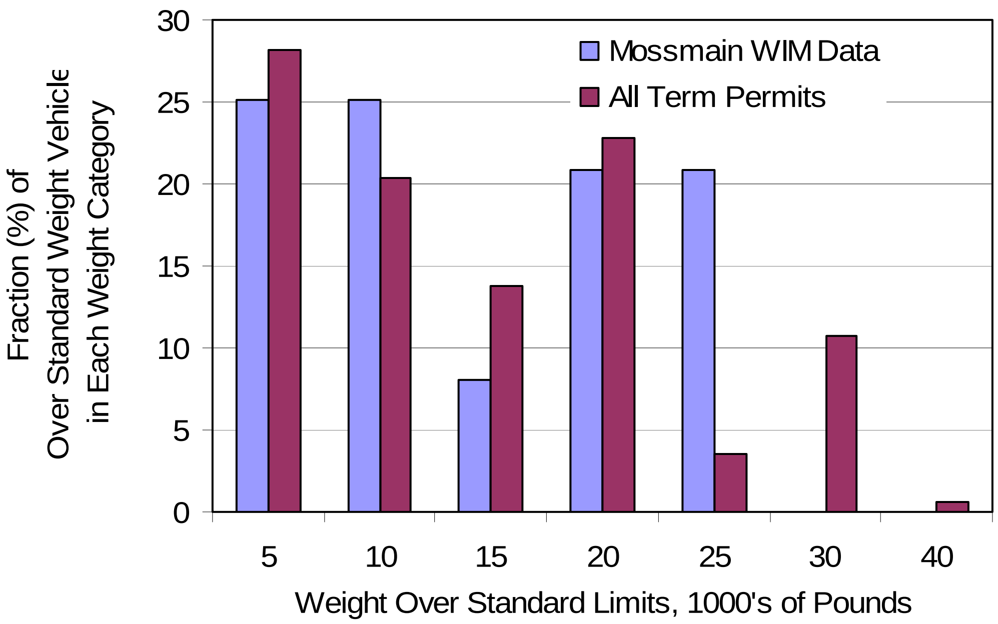


Figure 6.2-2. Fraction of Over Standard Weight Vehicles by Weight Range, Class 10 Vehicles at Mossmain (as identified from WIM data) and All Term Permitted Vehicles

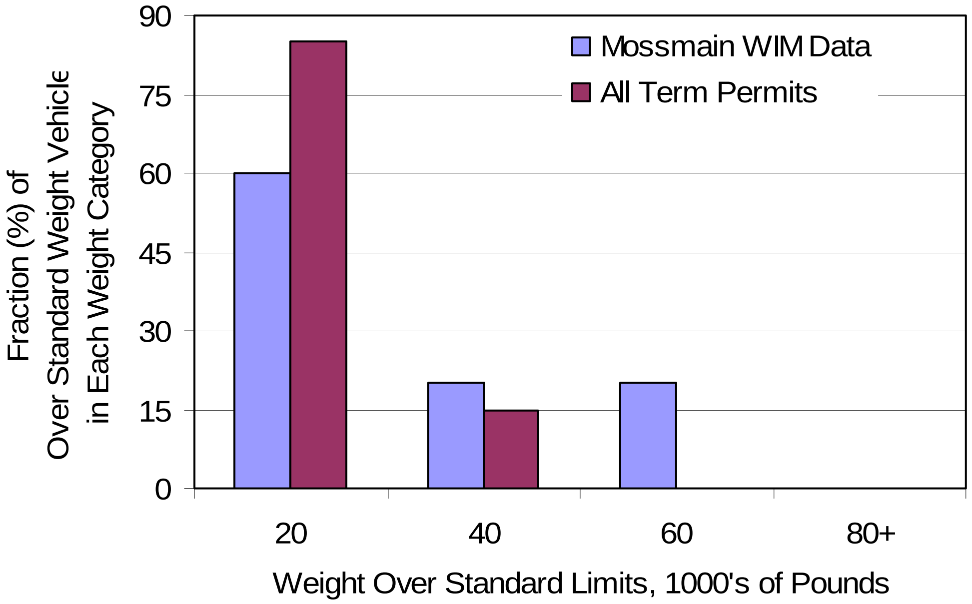


Figure 6.2-3. Fraction of Over Standard Weight Vehicles by Weight Range, Class 13 Vehicles at Mossmain (as identified from WIM data) and All Term Permitted Vehicles.

The specific percentage of vehicles in each weight range diverge some for the Class 9 vehicles at Mossmain versus all term permitted vehicles. These differences could well result from the underlying difference in the two items being compared (i.e., permitted Class 9 vehicles traveling at a specific location relative to all term permitted vehicles). The specific percentage of vehicles in each weight range are remarkably similar for Class 10 vehicles at Mossman and for all term permitted vehicles. As previously mentioned, Class 10 vehicles were judged to be one of the most common permitted vehicles, so their characteristics could be dominating the data set of all term permitted vehicles.

The over standard weight distributions for Class 9, 10, and 13 vehicles in the WIM data at Mossmain are compared to the over standard weight distributions for single trip permitted vehicles in Figures 6.2-4, 6.2-5, and 6.2-6, respectively. The single trip permit information was available by gross vehicle weight and axle configuration. Therefore, it was possible to generate the over standard weight distributions for the permitted vehicles independently for 5 axle (Class 9), 6 axle (Class 10), and 7 and over axle (Class 13) vehicles. Referring to Figures 6.2-4, 6.2-5, and 6.2-6, no strong correlation appears to exist between the weight distributions derived for each vehicle configuration from the WIM data and the corresponding weight distributions from the permit information. Nonetheless, the distributions from the two sources are not so obviously divergent as to suggest that it is unreasonable for the over standard weight vehicles identified in the WIM records to be permitted vehicles. One explanation of the difference in the weight distributions from the two sources is that the permitted vehicle traffic at Mossmain is dominated by term permitted vehicles, and that the basic over standard weight distributions for term and single trip permitted vehicles are different. In the latter regard, a greater percentage of single trip permits appear to be issued for heavier categories of over weight, relative to term permits, as illustrated in Figure 6.2-7.

To follow through on how this information could be used in MEARS, the number of Type 9, 10, and 13 vehicles identified as over weight at Mossmain would be reduced by 1 percent, 20 percent, and 1 percent of the total number of vehicles of each configuration in the sample. This reduction would be done before the characteristics of the over weight vehicles at this site are

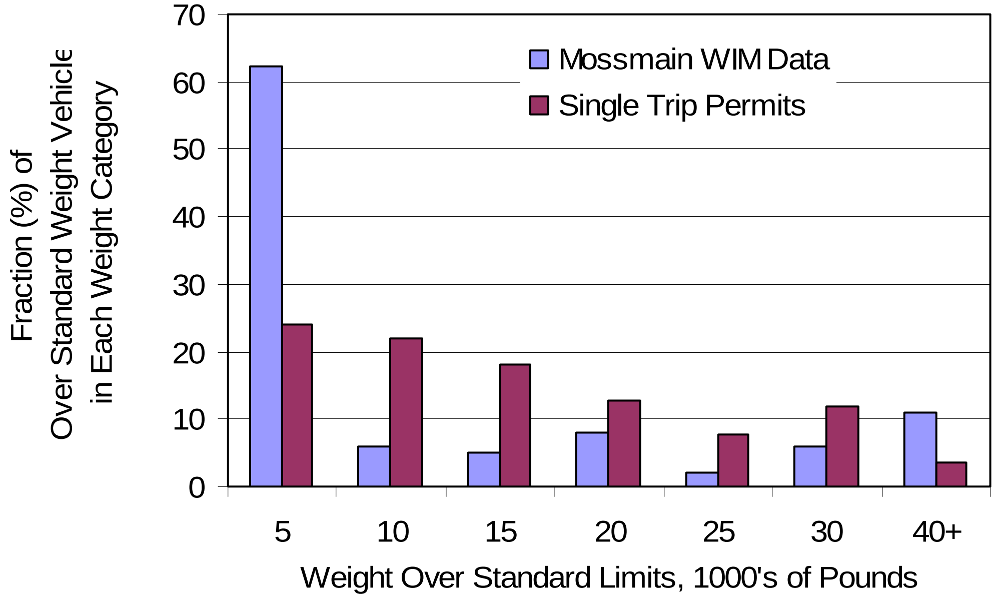


Figure 6.2-4. Fraction of Over Standard Weight Vehicles by Weight Range, Class 9 Vehicles at Mossmain (as identified from WIM data) and All Single Trip Permitted 5 Axle Vehicles

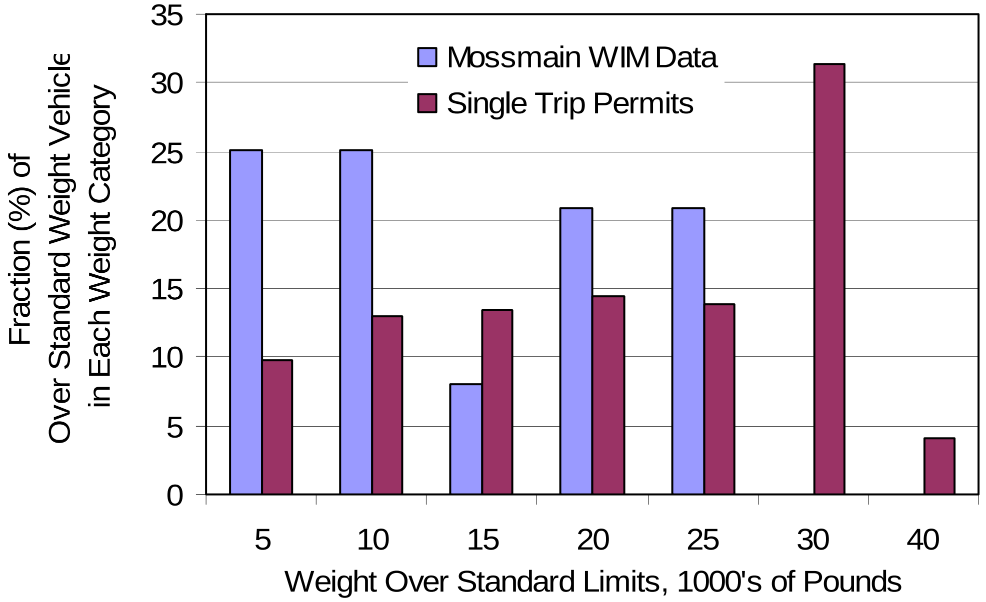


Figure 6.2-5. Fraction of Over Standard Weight Vehicles by Weight Range, Class 10 Vehicles at Mossmain (as identified from WIM data) and All Single Trip Permitted 6 Axle Vehicles

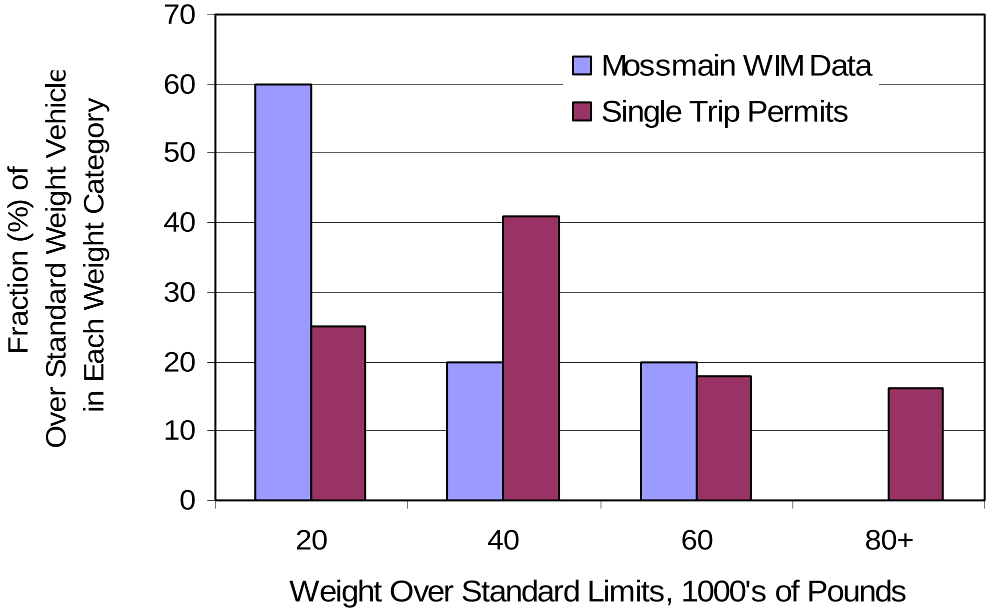


Figure 6.2-6. Fraction of Over Standard Weight Vehicles by Weight Range, Class 13 Vehicles at Mossmain (as identified from WIM data) and All Single Trip Permitted 7+ Axle Vehicles

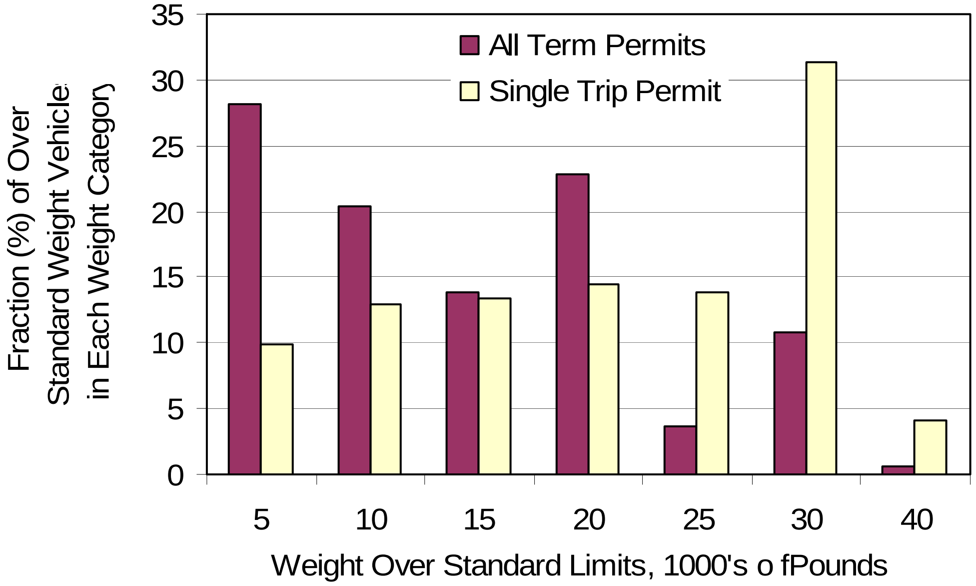


Figure 6.2-7. Fraction of Over Standard Weight Vehicles by Weight Range, All Single Trip and Term Permitted Vehicles

calculated, and the amount of the reduction would possibly be reported in separate entries in the existing MEARS reports as the number and percent of permitted vehicles assumed to be operating in the traffic stream.

6.3 Conclusions: Methodology to Identify Permitted Vehicles at *STARS* Sites

The methodology proposed herein to estimate the permitted vehicle traffic at *STARS* using WIM records collected during overt enforcement activities appears promising. The premise behind the methodology is that during overt enforcement activities, over standard weight vehicles in the traffic stream are permitted vehicles. Thus, the relative proportion of vehicles identified as overweight in the WIM records during such activities can be used to calculate an adjustment factor to be applied to the WIM records collected at other times. This approach to addressing the issue of permitted vehicles in the traffic stream being mis-identified as over weight is simple in concept and relies only on the WIM data that is already being collected, augmented by an onsite, overt enforcement presence over a known period of time. The greatest challenge of using this methodology may be in insuring that data records are obtained that are fully representative of general vehicle operations at a site. Furthermore, additional uncertainty will be introduced if factors developed at one site are used across other “similar” sites. Nonetheless, in the absence of any other feasible alternative, this methodology deserves further investigation.

7 POTENTIAL CHANGES TO THE MEARS SOFTWARE

The MEARS software program continues to evolve as more experience is gained in working with the information that it produces. Obviously, MEARS needs to be as accurate as possible in the information that it produces, and it is also desirable that this information be generated as efficiently as possible. With respect to efficiency, as traffic levels generally increase around the state, and as more *STARS* WIM sites are installed, the volume of data to be processed by MEARS continues to increase, and calculation times are beginning to be excessive. Therefore, options were investigated for reducing the amount of data to be processed by MEARS by decimating the raw WIM data prior to running MEARS.

With respect to accuracy of the results that MEARS generates, it was believed that it might be possible to improve on the pavement damage analysis conducted by MEARS by using a more sophisticated (and more computationally intensive) algorithm in this analysis. A comparison was made between the results obtained from the simple algorithm and the more complex alternative.

7.1 Methods to Reduce the Computation Time Required by MEARS

Weigh-in-motion systems provide a somewhat unique opportunity in terms of statistical validity by capturing, in theory, 100 percent of the commercial vehicle population on various routes either continuously or periodically throughout the year rather than a subset of this population on which various speculations are made. While this is appealing from a statistical inference perspective, the massive number of data records quickly becomes cumbersome to process and analyze. As such, a method is recommended here that somewhat compromises the statistical inference capabilities from WIM data in exchange for data processing ease.

General data reduction typically occurs in any one (or some combination) of three ways. The number of data records can be reduced by: (1) eliminating invalid data, (2) aggregating individual records into more manageable summary data records (i.e., hourly or daily averages), and (3) sampling within the larger sample or population to limit the number of records used for analysis and reporting.

7.1.1 Eliminating Invalid Data

Quality control measures should be applied to data after data collection and prior to analysis to identify data errors or invalid data (i.e., outliers). While this activity may result in significant enhancement to WIM data quality, only limited benefit will be realized for data reduction. MDT already screens each vehicle record according to a variety of criteria, including vehicle speed, position, and configuration. Relative to weight, the occurrence of a rare event versus an outlier arguably can be readily determined, in light of the nature of the phenomena being measured. That is, the maximum amount of weight that a commercial vehicle can physically carry is finite, with an estimable limit. Based on this premise, MDT is developing and implementing a screening routine which will eliminate records that contain "impossible" weights (e.g., 40,000 pounds on a single axle). While this approach to handling outliers may be sufficient, it may be possible to further narrow the screening process using statistically based data analysis methodologies. Therefore, a brief description of such methodologies is provided below. Note that if left undetected and uncorrected, the presence of outliers distorts any statistical test based on sample means and variances; statistical significance, or lack thereof, can be due to the presence of a few--or even one--unusual data value. Thus, the presence of outliers that are the result of measurement error could lead to erroneous conclusions regarding *STARS* effectiveness and subsequently poorly informed decision-making.

7.1.1.1 Detecting Outliers - In detecting outliers, the goal is to identify data values that may have an undue influence on the results because they lie well outside the range of other data. Several methods are commonly used to identify outliers in normally distributed data, including but not limited to: (1) the z-score method, (2) the modified z-score method, (3) Rosner's Test and (4) the boxplot method.

In a z-score test, the mean and standard deviation of the entire data set are used to obtain a z-score for each data point, according to following formula:

$$z_i = \frac{(x_i - \bar{x})}{s}$$

where

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

An observation with a z-score greater than 3.0 should be labeled as an outlier. This method, unfortunately, is not a reliable way of identifying outliers since both the mean and standard deviation are inclusive of and affected by the outliers.

To overcome this issue, a modified z-score test estimates the z-score based on the median of absolute deviation about the median (MAD):

$$MAD = median\left\{\left|x_i - \bar{x}\right|\right\}$$

MAD is calculated and used in place of the standard deviation, s, in z-score calculations. In this instance, an observation with a modified z-score greater than 3.5 should be labeled as an outlier.

Rosner's Test for detecting up to k outliers can be used when the number of data points is 25 or more. This test identifies outliers that are both high and low, it is therefore always two-tailed (Gibbons, 1994). The data are ranked in ascending order and the mean and standard deviation are determined. The procedure entails removing from the data set the observation, x, that is farthest from the mean. Then a test statistic, R, is calculated:

$$R_{i+1} = \frac{\left|x^{(i)} - \bar{x}^{(i)}\right|}{s^{(i)}}$$

The R statistic is then compared with a critical value (Gilbert, 1987). The null hypothesis, stating that the data fits a normal distribution, is then tested. If R is less than the critical value, the null hypothesis cannot be rejected, and therefore there are no outliers. If R is greater than the critical value, the null hypothesis is rejected and the presence of k outliers is accepted.

Using graphical methods, a boxplot test can be used to identify an observation as an outlier. A boxplot depicts the inter-quartile range (IQR) of the data and indicates the median value with a horizontal line. Error bars are drawn at the 5 percent confidence limit:

$$\frac{Q_1 - x}{Q_3 - Q_1} > k$$

and the 95 percent confidence limit:

$$\frac{x - Q_3}{Q_3 - Q_1} > k$$

Any data outside of the error bars are plotted as single points and labeled as possible outliers.

In addition to boxplots, outliers in continuous or interval data can be identified using simple scatterplots.

7.1.1.2 Addressing Outliers - The most extreme method for addressing outliers in the data is deletion. Statisticians recommend deletion only as a last resort, and then only if they are legitimate errors that can't be corrected, or lie so far outside the range of the remainder of the data that they distort statistical inferences. When in doubt, the analysis can be done both with and without outliers to see how much the results change. Less extreme methods for addressing outliers include (1) transforming data to soften their impact since the most commonly used expressions, square roots and logarithms, shrink larger values to a much greater extent than they shrink smaller values or (2) using methods that are robust in the presence of outliers such as nonparametric statistical methods.

Because of the known upper physical limits of axle and gross vehicle weight capacities, deletion of outliers that fall above such limits is an acceptable option.

7.1.2 Aggregating Individual Records

One promising method to reduce the overall WIM data processing demands would be to employ a two step analysis, where in the first stage, individual commercial vehicle data records are averaged over some larger time interval (i.e., hourly, daily, etc.); these subsequent aggregate records would be analyzed using MEARS to produce the site-specific monthly and annual reports. Currently, the *MEARS* summary reports by site, by month and by year are based on a primary analysis inclusive of every commercial vehicle WIM record captured. Currently, for example, at a site that experiences 1000 commercial vehicles per day, all 1000 individual data records are processed through MEARS. Following the two staged approach outlined above, this data set could be reduced to 24 average hourly records that describe traffic volumes and commercial vehicle loading characteristics, which would then be processed by MEARS. Realistically however, traffic needs to be distinguished by vehicle class; thus the 1000 total truck records per day could be reduced to 312 averaged records by vehicle class (i.e., 24 one-hour records for 13 vehicle classes).

The one hour aggregation interval used in this example was somewhat arbitrarily picked; it may be possible to use 2 or 4 hour intervals (and thus further reduce the data processing effort), depending on the characteristics of the traffic at a given site. Conventional techniques for the determination of aggregation levels are normally based on statistical comparisons of the original and aggregated data sets. However, it is not guaranteed that the errors and noise will be transmitted to the aggregated data sets and that the desired information will be preserved. Sophisticated methods have been employed to determine the optimum level of aggregation, including wavelet decomposition and determination of optimal aggregation levels for different times of day and days of the week and even monthly.

7.1.3 Sampling Within the Population

Another approach to reducing the WIM data processing times would be to sample within the larger population of WIM records and utilize this statistically-robust yet reduced sample for the *MEARS* analysis. The most common formula for determining the minimum required sample size from a given population is as follows:

$$N = \frac{CV^2 * z^2}{E^2}$$

where: N = Sample Size

CV = Coefficient of Variation given as

$$CV = 100 * \frac{\text{Standard Deviation}}{\text{Mean}}$$

z = Normal distribution coefficient for a specific confidence interval (CI)

E = tolerable error

In brief, this relationship allows the analyst to define an acceptable compromise in accuracy (resulting from the smaller number of observations, usually 5 percent corresponding to a 95 percent confidence level) and incorporates the data's inherent variability characteristics to determine the minimum required sample size. The mean and standard deviation of the larger population data set can be used directly to determine the most accurate CV which in turn will result in the most accurate required sample size to achieve the desired accuracy levels.

This formula could be applied to the overall population data set to determine the total number of records required at each site but because vehicle classification is also of interest, the population should be first stratified by vehicle class to determine the required number of class-specific records required per site, N_5, $N_6 \ldots N_{13}$. Once these minimum sample sizes are estimated, individual and class specific records can be pulled at random until the respective samples are populated resulting in a reduced and robust stratified random sample form the larger population of individual WIM records.

7.2 Investigation of the Algorithm Used by MEARS to Determine the Relative Pavement Damage Between Two Evaluation Periods

One of the reports generated by MEARS provides a comparison of the excess ESALs of travel experienced at the same location (i.e., at a specific WIM site or across all sites) over two different time periods. As previously mentioned, such comparisons are useful in evaluating the effectiveness of different enforcement strategies and activities either at the local and/or statewide level. In using changes in excess ESALs for this purpose, it is important to take into consideration other factors that could be responsible for changes in excess ESALs of travel between two time intervals. In particular, the number of excess ESALs experienced at a location over any interval of time is influenced by the basic amount of freight that is being moved at that location over that time interval, independent of any enforcement related activity at the site.

Thus, in comparing the characteristics of overweight vehicle operations at a site over two different time intervals, say A and B, with the intent of evaluating enforcement effectiveness, differences in the basic amount of freight hauled during the two intervals need to be addressed.

Presently, MEARS uses a very simple procedure to adjust for differences in the basic level of vehicle activity and the attendant amount of freight moved during different evaluation periods, and to then further determine the comparative change in ESAL miles of travel between the two periods. While the analytical efficiency of this algorithm is attractive, concerns existed relative to possible bias that it may introduce into the results of the evaluation.. Therefore, a more sophisticated algorithm was developed for calculating comparative changes in ESAL miles of travel during two evaluation intervals, and the results compared with those generated by the simple algorithm.

Comparison of the results from the two approaches found that the simple algorithm generates changes in excess ESALs that are consistently in close agreement with those obtained from the sophisticated algorithm. In general, the simple algorithm was found to underestimate the excess ESALs by an average of 3 percent relative to the results obtained from the more sophisticated algorithm. In light of the uncertainties involved in this calculation even when using the sophisticated algorithm, the current approach used by MEARS to compare excess ESALs between two evaluation periods is expected to be adequate.

7.2.1 Description of the Problem

Ideally, to eliminate the influence of differences in the basic volume of freight being moved during two periods, say A and B, that are being used in a comparative evaluation of enforcement effectiveness, the total volume of freight moved on the highway should be constant during A and B. In this case, any differences in excess ESALS observed between interval B and the interval A result from differences in vehicle loading patterns between the two intervals, rather than differences in the total freight carried. Naturally, the volume of freight being moved can not be controlled, so this effect must be addressed, as possible, in the data analysis process. One approach to this problem is to analytically generate the traffic stream with the vehicle loading characteristics of interval A that is necessary to carry the same amount of freight moved during interval B. Thus, any differences in excess ESALS observed between interval B and the analytically modified interval A once again result from differences in vehicle loading patterns between the two intervals, rather than differences in the total freight carried.

While the above strategy for accounting for the general level of commercial vehicle activity and freight movement between evaluation periods A and B is easy to understand, it is difficult to implement using the limited route specific information available on vehicle operations around the state. Generally, little information is available beyond what is measured by the WIM system, itself. It is difficult to infer the amount of freight moved during interval B without some idea of the weight of the cargo carried by each vehicle. This weight could possibly be estimated if the commodity being carried is known, but commodity information by route is sparse. Naturally, the manner in which the total amount of freight carried during period B should be redistributed across the types of vehicles operating at the weights characteristic of period A is equally difficult to determine.

Two approaches have been pursued in the *STARS* program to account for differences in the total volume of traffic when excess ESALs are to be compared between two intervals, namely,

1) a very simple, gross calculation is done based on total traffic volume and just a few assumptions about loading behaviors during each interval (current MEARS approach), and

2) a detailed analysis is done that explicitly estimates the freight carried during one interval and reassigns it to the vehicles in the second interval based on a myriad of assumptions about loading behaviors.

The former approach is currently used by MEARS, while the former approach was used in this study in an effort to determine if the simple MEARS approach yields adequate answers, thus eliminating the need for the computationally intensive alternate approach.

7.2.2 MEARS Methodology for Accommodating Changes in the Volume of Freight Moved Between Evaluation Periods

In MEARS, when the excess ESALs from two evaluation periods are to be compared, the excess ESALS from one of the two evaluation periods is simply adjusted based on the relative volume of commercial vehicle traffic during the two intervals:

$$EE_{Aadjusted} = EE_A\left(\frac{CV_B}{CV_A}\right)$$

where

$EE_{Aadjusted}$ = excess ESALS during interval A adjusted for the difference in freight hauled during intervals A and B

EE_A = actual excess ESALS observed during interval A

CV_A = number of commercial vehicles observed during interval B

CV_B = number of commercial vehicles observed during interval B

The change in ESALs between the two intervals normalized to the common volume of freight carried in interval B is then calculated as,

$$\Delta EE = EE_B - EE_{Aadjusted}$$

where

ΔEE = change in excess ESALs for a common volume of freight carried during interval A and B

EE_B = change in excess ESALs actually observed in interval B

$EE_{Aadjusted}$ = excess ESALS during interval B adjusted for the difference in freight hauled during intervals A and B

This calculation is done on a configuration-by-configuration basis.

This approach to adjusting the excess ESALS between two intervals assumes that the total freight carried over both intervals is simply proportional to the number of vehicles operating in each interval; it ignores any differences in the specific weight characteristics of the vehicles in each interval (such as, percent of overweight vehicles of each configuration and average amount of overweight by configuration). This approach generates reasonable results if the percentage of overweight vehicles in the traffic stream is low, and if the average overweight is low. Under other situations, however, the change in excess ESALs calculated by this method could be significantly in error.

For example, during some time interval, A, consider the situation of all 10 vehicles of a particular configuration at a specific site operating at a payload 20 percent greater than permitted by law. The total freight moved by these vehicles during interval A equals 10 times 1.2 legal trucks, or 12 legal trucks. Subsequently, in a second interval (say interval B), the same amount of freight was observed to be hauled by 11 vehicles, each with a payload 10 percent greater than permitted by law. The total freight moved in interval A is almost identical that moved in interval B (both equivalent to 12 legally loaded trucks). The change in excess ESALS in this case should be calculated as the excess ESALs for 11 vehicles, 10 percent over weight (EE_B), minus the excess ESALs for 10 vehicles, 20 percent over weight (EE_A). Following the approach implemented in MEARS, the change in excess ESALS is actually calculated as the excess ESALs for 11 vehicles operating 10 percent overweight (EE_B) minus the excess ESALs for 11 vehicles operating 20 percent overweight ($EE_{Aadjusted}$). Thus, the difference in excess ESALs, in this case, and in every case, would be overstated by MEARS in the event of a decrease in ESALs between interval A and B, while it would be understated by MEARS in the event of an increase in ESALs between interval A and B.

This example represents a relatively severe case of overweight compared to that typically found on Montana's highways, and thus it is believed that it overstates the effect that the current algorithm used by MEARS has on the calculated changes in excess ESALs between two evaluation intervals. Nonetheless, to more definitively establish if any consistent and significant bias is introduced into the pavement damage results generated by MEARS, this situation was investigated in more detail, as discussed below.

7.2.3 Alternate Methodology for Accommodating Changes in the Volume of Freight Moved Between Evaluation Periods

To obtain a better idea of the effect of the MEARS algorithm on the calculated changes in excess ESALs between two evaluation intervals, a second algorithm was developed to accomplish this same task that takes into account more of the factors known to influence this quantity. Notably, the new algorithm takes into consideration the character of the vehicle loading patterns at a site during the two intervals being compared (i.e., percent of overweight vehicles in the traffic stream, and average amount of overweight) in addition to the relative number of commercial vehicles passing the site during each interval. This adjustment effectively analytically generates the traffic stream with the vehicle loading characteristics of interval A that would be necessary to carry the amount of freight moved during interval B. Thus, any differences in excess ESALS observed between interval B and the analytically modified interval A result from differences in vehicle loading patterns between the two intervals, rather than simple differences in the volume of vehicle traffic.

The alternate methodology begins by determining the equivalent number of weight compliant vehicles that are necessary to carry the freight moving on the overweight vehicles observed during an interval, say interval B. This analysis is done using the number of over weight vehicles and the average amount of their overweight as reported by MEARS for each configuration. This basic information is augmented by an assumed empty vehicle weight for each configuration, and it is then processed under the further assumption that it is reasonable to re-distribute the load carried by overweight vehicles to vehicles operating at the maximum limit allowed by statute,

$$NL_i = N_i + \frac{TOW_i}{(WL - WE)}$$

where

i = interval, either A or B

NL_i = number of vehicles required to carry freight in interval *i*, all vehicles at or below standard weight limits

N_i = number or vehicles observed during interval *i* (interval A or B)

TOW_i = cumulative over weight from all overweight vehicles during interval *i* (interval A or B)

WL = assumed standard gross vehicle weight limit

WE = assumed vehicle empty weight

The number of vehicles with weight characteristics of interval A required to carry the same freight as is moved in interval B is then calculated as,

$$N_{Aadjusted} = N_A \frac{NL_B}{NL_A}$$

Finally, the number of overweight vehicles in the equivalent traffic stream from interval A is calculated as,

$$N_{OWAadjusted} = N_{Aadjusted} \frac{N_{OWA}}{N_A}$$

where

N_{OWA} = Number of overweight vehicles observed during interval A in the original data

The excess ESALs are then calculated for the adjusted number of overweight vehicles in intervalA, with the actual average overweight observed in interval A, and the actual number of vehicles and average amount of overweight observed in interval B. Finally, the change in excess ESALs is calculated,

$$\Delta EE = EE_B - EE_{Aadjusted}$$

where these variables are as previously defined.

Obviously, several more calculations are required in this algorithm relative to the simple algorithm employed by MEARS. Relative to the MEARS methodology, this approach considers more of the factors that affect pavement damage in normalizing the traffic observed during interval A to carry the same of amount of freight moved during interval B. However, this approach requires additional assumptions on vehicle weight characteristics and operating conditions (i.e., empty weights, overweight will be shifted to fully loaded compliant vehicles), and thus it still will yield only an approximate result. Note that a slightly modified version of this alternate algorithm was used in the initial *STARS* evaluation.

7.2.4 Comparison of Pavement Damage Results: Current MEARS vs. Alternate Algorithm

The two algorithms described above subsequently were used to process the same sets of data to see if they produced similar results. While the absolute accuracy of both methodologies was (and remains) unknown, it was assumed that the alternate methodology generated better results, in light of the basic formulation of both approaches. The data processed for this comparison consisted of all the *STARS* data collected at the 16 sites used in the original evaluation during the baseline year, the year of focused enforcement, and the year following focused enforcement The algorithms were used to calculate the change in excess ESALs between the same months in successive years (i.e., baseline year to year of focused enforcement, and follow-on year to year of focused enforcement). The change in ESALs calculated each month using the MEARS algorithm was subsequently divided by the change in ESALs calculated each month using the alternate algorithm, and these results are presented in Table 7.2.4-1. Obviously, if the two algorithms generated the same results, this ratio would have a value of 1.0.

Overall, the average change in excess ESALs calculated by the MEARS algorithm was 102 percent of the average change in excess ESALs calculated using the alternate algorithm. The ratios of excess ESALs calculated by MEARS to the excess ESALs calculated by the alternate

Table 7.2.4-1 Comparison of Excess ESALs Calculated by MEARS and by an Alternate Algorithm

Nature of Difference in Excess ESALS	Number of Comparisons	Mean of the Ratio $\frac{EE_{MEARS}}{EE_{ALTERNATE}}$	Variance of the Ratio $\frac{EE_{MEARS}}{EE_{ALTERNATE}}$	99 % Confidence Interval on the Mean
Decrease	197	1.031	0.004	0.009
Increase	123	0.999	0.003	0.009
All	105	1.019	0.004	0.007

algorithm were closely clustered around the mean, ranging in value from 0.71 to 1.43 with a variance of only 0.004.

Based on its formulation, the MEARS algorithm was expected to over report the difference in excess ESALS when the excess ESALS decreased between two intervals, while it was expected to under report the difference in excess ESALS when this difference represented an increase in this parameter between two intervals. With this expectation, the performance of the two algorithms was investigated separately for the case of a decrease and an increase in excess ESALs between two intervals. The results of this analysis are also reported in Table 7.2.4-1. For the case of a decrease in excess ESALS between two intervals, the difference calculated by the MEARS algorithm was 103 percent of that calculated by the alternate algorithm. These results were sufficiently consistent across all the intervals evaluated that it can be concluded at the 99 percent confidence level that the MEARS algorithm over reports the decrease in excess ESALs between two intervals. A similar analysis for the case of an increase in excess ESALs between two intervals found that the MEARS algorithm and the alternate algorithm generated the same results (at a 99 percent confidence level).

7.2.5 Conclusions on MEARS Pavement Damage Algorithm

The current algorithm used in MEARS to evaluate the difference in pavement damage from over weight vehicles during two evaluation intervals is attractive in its computational simplicity, and it yields reasonable and useful results. In comparisons with an alternate algorithm that would be expected to yield more accurate results based on its formulation, the MEARS algorithm was found to over report the difference in excess ESALs by an average of 3 percent when this difference represented a decrease in this parameter between the two evaluation intervals. The two algorithms generate the same results in the case of an increase in excess ESALs between evaluation periods. In interpreting the significance of these observations, it is important to recognize that the absolute error involved in using either algorithm is unknown, and that the alternate algorithm, which is more computationally intensive than the MEARS algorithm, relies on several assumptions in arriving at its results. In light of the predictably of the bias in the MEARS algorithm, one course of action would be to simply divide the excess ESAL difference calculated by MEARS by a constant factor of 1.03, if this difference is for a decrease in ESALs between two evaluation periods.

8 SUMMARY AND RECOMMENDED FUTURE WORK

8.1 Summary

In this project, several issues regarding the present and future use of *STARS* in MDT's activities were investigated. These issues were largely identified in the initial evaluation of *STARS* conducted by Montana State University as *STARS* was first implemented by MDT. While that evaluation found that *STARS* was successful in meeting its primary objectives, issues came up that merited further investigation and resolution as MDT moves ahead with the *STARS* program. These issues consisted of:

1) Benefit to Cost Ratio for the Program. The benefit to cost ratio of *STARS* was conservatively estimated to range from 6.3 to 10.6 using an assumed service life for the WIM equipment of 4 to 12 years, respectively. This ratio is based only on the readily quantifiable *STARS* benefits associated with reduced pavement damage from overweight vehicles ($700,000 per year) and better pavement designs from improved load data ($4,100,000 per year). Other less tangible benefits from improved quality weight data were ignored in the analysis. *STARS* costs considered in the analysis consisted of all costs associated with purchase, installation, and operation of 26 WIM sites around the state and amounted to $757,000 to $451,000 per year for estimated service lives of 4 to 12 years, respectively.

2) Analysis of *STARS* Data Collected Since the Initial Evaluation The *STARS* data collected since the initial evaluation was completed confirmed that the *STARS* focused enforcement was responsible for the reduction in overweight vehicle operations during the year that it was engaged in. During the year prior to the focused enforcement effort, the year of focused enforcement, and the year following focused enforcement, the percentages of overweight vehicles in the traffic stream were 8.8, 6.9, and 8.9, respectively. The pavement damage from overweight vehicles decreased by 6,000,000 ESAL miles of travel between the baseline and focused enforcement years, and increased by 5,000,000 ESAL miles of travel in the year following the focused enforcement effort.

3) <u>Use of *STARS* in Weight Enforcement</u> *STARS* is expected to have a continuing role in both evaluating the effectiveness of MDT's weight enforcement efforts as well as in helping to direct these efforts. While in the initial evaluation, the focused enforcement effort was judged to be effective based on changes in the characteristics of the overweight vehicle population, it may be more reasonable in the future to evaluate effectiveness based on meeting target values for parameters that measure the amount of overweight vehicle activity on the state's highways. Two parameters are suggested in this regard, percent of overweight vehicles in the traffic stream and accumulated excess (overweight) ESAL miles of travel. While percent of overweight vehicles in the traffic stream may be the more readily understood and evaluated of these two parameters, excess ESAL miles of travel more directly reflects accumulated infrastructure damage from overweight vehicles. For example, a target maximum level of overweight vehicle in the traffic stream might be 6 percent across the state and no more than 10 percent at any specific *STARS* site. Corresponding target metrics based on excess ESAL miles of travel are not as readily formulated, which illustrates the practical problems in using this parameter in the evaluation process. None-the-less, this parameter is sufficiently important that some metric should be developed in this regard.

In using information from *STARS* (as processed by MEARS) to evaluate the relative effectiveness of different enforcement activities, it is important to consider whether or not the proposed evaluation will accurately assess the cause and effect relationship between enforcement and overweight vehicle activity. The volume of commercial vehicle activity on the state's highways is dynamic in nature, and it varies for many reasons independent of enforcement activity. Therefore, it is critical that professional judgment be used in assessing the results of any comparative analysis of enforcement effectiveness between two intervals. In general, the reliability of the evaluation would be expected to improve as its duration increased and the geographic area that it covered increased. In this situation, local temporal and spatial variations in overweight vehicle operations should average out through the course of the evaluation.

To fully evaluate the effectiveness of various weight enforcement activities, some consideration must be made of the level of effort put into the activity relative to its outcome. Presently, MEARS does not include any parameters indicative of level of enforcement effort. Consideration should be given to adding a metric of this kind to the program.

Relative to using *STARS* information to help direct enforcement at a statewide level, once again additional information is required beyond that already provided by MEARS relative to the historical level of enforcement engaged in at each site. Ideally, the relationship between level of enforcement and level of overweight vehicle activity would be known for each site. This information could then be used in an optimization analysis to best allocate enforcement resources system-wide. Such an analysis could be conducted by a second software package that accesses the output from MEARS and integrates it with information on level/cost of enforcement effort. In the short term, the inclusion of even the most rudimentary indicator of level of enforcement effort at each site as part of MEARS could be useful in the resource allocation process between *STARS* sites.

At any particular site, *STARS* continues to provide detailed information related to when the most severe overweight activity historically has occurred; with due care, this information is useful in helping to direct current and future enforcement activities. Until additional experience is gained in this regard, only limited guidance is available on planning future enforcement activities based on historical observations of overweight vehicle activity; therefore, professional judgment will have to be exercised in this regard. Notably, the dynamic nature of overweight vehicle activities must be factored into planning the enforcement effort, so that the effort confronts the intended overweight vehicle activity identified in the historical data. While most of the efforts to-date that have attempted to use historical data to guide future enforcement efforts have looked back one year to plan current year enforcement activities, historical data collected across more closely spaced time intervals may also be used in this regard. The critical feature of the historical data is that it is expected to represent the pattern of overweight vehicle operation during the proposed enforcement period. Professional judgment must be used to determine if the historical data satisfies this condition (and to further decide after the fact if overweight activity during the enforcement period did indeed conform to expectations from the historical data).

4) <u>Identification of Permitted Vehicles from *STARS* WIM Data</u> A methodology was developed and tested for identifying permitted over standard weight vehicles in the traffic stream at *STARS* sites using the WIM data collected at the site. The methodology is based upon the assumption that over standard weight vehicles operating at a site during overt enforcement activities are permitted vehicles. The characteristics of these over standard weight vehicles can be determined

from the WIM data collected during the enforcement activity, and factors can subsequently be developed that are applied in MEARS to adjust the reported overweight vehicle populations to account for their presence in the traffic stream. This methodology was applied at a WIM site adjacent to a weigh station (Mossmain), where 1, 21, and 1 percent of the Class 9, 10, and 13 vehicles were determined to be permitted over standard weight vehicles. These results appear reasonable based on the limited information available on the permitted vehicles operating on the state's highways.

5) Potential Changes to the *MEARS* Software A limited investigation was done on improving the computational efficiency of *MEARS*. *MEARS* currently uses every vehicle record to determine the characteristics of the overweight vehicle operations at each *STARS* site every month. As the number of *STARS* sites has increased, the computation time to generate these monthly *MEARS* reports has corresponding increased. One possible solution to this problem may be to run *MEARS* only on a sample of the vehicle records collected at each site. Another promising approach for this purpose is to aggregate the data across some time interval (say 1 to 4 hours) prior to running MEARS. The trade off for reducing the computation time by either approach is the possible introduction of error into the results by the sampling process.

The simple and computationally expedient algorithm currently used by *MEARS* to estimate the change in excess ESAL miles of travel between two evaluation intervals was found to yield similar results to a more sophisticated and computationally intensive algorithm that explicitly considers more of the parameters known to effect this calculation. The existing algorithm did appear to systematically overestimate decreases in excess ESAL miles of travel between two intervals by an average of 3 percent, while it accurately estimated increases in excess ESAL miles of travel between two intervals. In light of the small magnitude of the possible error involved in using the existing algorithm, and inherent uncertainties in both algorithms, the existing algorithm was judged to generate adequate results.

8.2 Recommended Future Work

STARS continues to evolve as a data collection tool for MDT. Of the many applications for the information available from *STARS*, the least guidance may be available for its use in weight enforcement. As experience is gained in using *STARS* data, as processed by MEARS, for weight enforcement related purposes, suggestions will continue to be made in both how the data is processed and how it can be best used. Future work in this regard suggested by this investigation can be broadly divided into two categories, a) upgrading of the existing MEARS software to make it computationally more efficient and to provide additional information useful in assessing the effectiveness of enforcement, and b) including level of effort (cost) information in the analysis process, possibly through the creation of a new program linked to MEARS, with the specific objective of being able to evaluate cost versus benefit in assessing enforcement effectiveness and in allocating enforcement resources (MEARS already provides information on the impacts of enforcement, but it is lacking information on the associated cost of enforcement efforts).

With respect to upgrading the existing MEARS software, the following suggestions are made:

1) Add a routine to normalize the excess ESALs from overweight vehicles at a site by the total ESALs at the site.

2) Implement computation saving, data processing algorithms

3) Further experiment with developing factors to adjust the information generated by MEARS to account for the presence of permitted over standard weight vehicles in the traffic stream. This experimentation would consist of planning and conducting enforcement activities at selected sites that are designed to create (and record) the vehicle activity at a site under enforced conditions (i.e., eliminate all illegal overweight vehicles from the traffic stream). The weight characteristics of this traffic stream will form the baseline against which overweight vehicle activity will be evaluated under normal operating conditions.

4) Develop and include in MEARS some metric that indicates the level of enforcement effort engaged in at each site, to allow for at least a rudimentary calculation of enforcement outcome as a function of enforcement effort.

5) Investigate bypass of *STARS* sites during focused enforcement activities. This task was included in the original addendum to the *STARS* evaluation but was not accomplished in the time available.

6) Investigate the extent of the highway system impacted by each *STARS* site. This task was also included in the original addendum to the *STARS* evaluation but was not accomplished in the time available.

Comprehensive integration of level of enforcement effort into the existing analysis framework either directly in MEARS or in a new software package may be a formidable task. Simply determining a practical, consistent, and useful method to measure and record level of enforcement effort may require considerable thought. This method of measurement must be determined before work can begin on identifying relationships between level of effort and enforcement outcome. Conditions may vary sufficiently between sites that such relationships need to be developed on a site-by-site basis (or optimistically, for groupings of sites). If these steps can be completed at any level, the resulting optimization in enforcement resource allocation should yield considerable benefit to MDT.

9 REFERENCES

Bisom, D. *Personal Communication*. Planning Division, Montana Department of Transportation, Helena, MT. 2003

Bisom, D. *Personal Communication*. Planning Division, Montana Department of Transportation, Helena, MT. 2004.

Clark, D., Stephens, J. and Carson, J. *A Field Evaluation of Piezoelectric and Bending Plate Weigh-in-Motion Systems*. Draft reported submitted to the Montana Department of Transportation. July 2004.

Federal Highway Administration. Traffic Monitoring Guide. FHWA-PL-01-021. U.S. Department of Transportation. 2001.

Hanscom, F.R. *Developing Measures of Effectiveness for Truck Weight Enforcement Activities*. NCHRP Web Document 13 (Project 20-34), Final Report. March 1998.

Hult, D. *Personal Communication*, Motor Carrier Services, Montana Department of Transportation, Helena, MT. 2004.

Idaho Department of Transportation. *Idaho Transportation Department Bridge Overweight Report.* http://www.itd.idaho.gov/planning/reports/atr_wim/099-HaydenWIM/2002/02-04Apr/hyd0402BrgOvw.PDF

International Road Dynamics, Inc. *Weigh-In Motion Technology Comparisons.* January 2001.

Jessup, E.L. and Casavant, K. *Evaluation of Violation and Capture of Overweight Trucks: A Case Study*. Washington Transportation Center (TRAC). July 1996.

Little, T. *Personal Communication*, Planning Division, Montana Department of Transportation, Helena, MT. 2005.

Livesay, D. *Personal Communication*. Motor Carrier Services, Montana Department of Transportation, Helena, MT. 2004.

McCall, Bill, and Walter C. Vodrazka, Jr. *State's Successful Practices Weigh-in-Motion Handbook.* Iowa State University, December 15, 1997.

Mumayiz, S.A., and R.M. Michaels. *Investigation of the Implementation of Weigh-in-Motion Operation in Illinois-Phase 1*. Urban Transportation Center, Chicago, March 1989.

Murphy, F. *Personal Communication*. Motor Carrier Services, Montana Department of Transportation, Helena, MT. 2004

Stein, P., et al., *The Overweight Truck in Wisconsin: Its Impact on Highway Design, Maintenance and Enforcement Planning*, Wisconsin Department of Transportation. 1988.

Stephens, J.E., Carson, J., Reagor, D., and Harrington, M., *An Evaluation of Montana's State Truck Activities Reporting System*, Final Report prepared by the Civil Engineering Department at Montana State University for the Montana Department of Transportation, Helena, MT. August 2003.

Whitford, Robert K. *Truck Weight Monitoring Plan Using Weigh-In-Motion Devices: Plan for WIM for the State of Alaska*. Alaska Department of Transportation and Public Facilities. 1997-1998.

APPENDIX A VEHICLE CONFIGURATIONS