ABSTRACT

RAMACHANDRAN, ADITYA NARAYANAN. Weigh in Motion Data Analysis. (Under the direction of Dr. John Stone.)

At hundreds of Weigh in Motion (WIM) stations, State Departments of Transportation collect traffic data every year to support pavement design, to enforce weight restrictions on highways and bridges, and to provide planning data for highway improvements. Reliable WIM data is particularly important to support the procedures in the FHWA Mechanistic Empirical Pavement Design Guide (MEPDG). The purpose of the research is to identify and resolve four related but relatively stand-alone problems associated with WIM data collected by NCDOT.

Quality Control: After the NCDOT collects WIM data and converts it from proprietary vendor format to an ASCII text format, the quality of the data must be checked. During the quality control (QC) procedures, tests identify incomplete datasets, out of range values for individual vehicle classes, and other possible data problems. Vehicle class and weight checks generate 0.97% and 6.42% anomalies, respectively thus confirming that NCDOT equipment captured reliable WIM measurements.

NC Urban and Rural Truck Traffic Profiles: Knowing the type of traffic by vehicle class by highway functional classification is critical to designing, maintaining and paying for North Carolina highway pavements. Thus, GVW plots by vehicle class and highway functional class are very important. The results indicate that in general, the class 5 and 9 GVW plots for all categories of WIM stations show expected trends. These results may be used by highway planners and pavement designers to quickly determine typical truck traffic profiles in the various NC regions and provide insight into NC truck transportation flows.

NC vs. University Of Arkansas WIM QC Analysis: Most highway agencies have the data collection and design groups in different units. While a single software solution is not practical, it is recommended to perform two separate processes where the output of data QC meets the needs and standards of the design process. A comparative analysis between the QC methods followed by the University of Arkansas (UARK) and NCSU/NCDOT shows that while the UARK Pavement Designer software has better mapping functions and supports data analysis and design. However, from a WIM data analyst’s perspective it is a “black box”. In addition there is
significant data reduction involved and the rigid nature of an automated QC process does not provide enough justification for the data to be used as input for the MEPDG. On the other hand, the NCSU/NCDOT approach is a two step procedure with a comprehensive QC procedure that provides the flexibility of manual overriding based on local knowledge of WIM stations and a separate unit managing the pavement design element.

WIM Data Management and Analysis in SQL Server: While most research topics focus on collection and quality of data with little emphasis on the development of an integrated Database Management System (DBMS) to store and analyze traffic data. An innovative approach to perform quality control procedures and data analysis on two test WIM databases by using analysis cubes in SQL Server is the main objective of this chapter. It is a convenient method of disseminating data to users that do not have online access to the WIM database in SQL Server. The analysis cube files may also be used for data mining and exploration and could be used to observe trends in axle weights, axle spacings, vehicle class volumes etc. An additional objective is to discuss the advantages of transitioning from Access to a more comprehensive DBMS like SQL Server or Oracle to store and analyze WIM data for the NCDOT statewide WIM program.
BIOGRAPHY

Aditya N. Ramachandran was born on May 8th, 1985 in the city of Coimbatore, India. Brought up in the industrial township region in Trichy, India, he completed his undergraduate degree in Civil Engineering from the National Institute of Technology, Trichy (NITT) in May 2007.

The last few semesters of his undergraduate studies made him realize that his true interest lies in Transportation Engineering and his will to pursue his graduate studies in this area. Right after he graduated, he applied for graduate studies in the United States and eventually joined NC State University to pursue his Masters degree in Civil Engineering in Fall 2007. He worked as a Research Assistant and Teaching Assistant during his stay as a student. During his studies at NC State University, he also worked as a transportation intern with the City of Durham/DCHC MPO during the summer of 2008 and for the Department of Transportation at NC State University during the Summer and Fall semesters of 2009.
ACKNOWLEDGEMENTS

I would like to extend my thanks to my academic advisor Dr. John Stone for his valuable input and support for this research. I would also like to thank my committee members Dr. Billy Williams and Dr. William Rasdorf for their constructive comments. A special acknowledgment goes to Mr. Kent Taylor and Mr. Stephen Piotrowski from the NCDOT Traffic Survey Unit for their help and insight. Special thanks are due to the NCDOT Traffic Survey Unit staff for collecting the data required for this research. I also appreciate the additional financial support from the Southeastern Transportation Center.
# TABLE OF CONTENTS

LIST OF TABLES ...................................................................................................................... vi
LIST OF FIGURES ...................................................................................................................... vii

1.0 INTRODUCTION .................................................................................................................. 1
  1.1 Background ............................................................................................................................. 1
  1.2 Research Scope and Objectives ............................................................................................ 1
  1.3 WIM Site Distribution in North Carolina .............................................................................. 2
  1.4 Overview ............................................................................................................................... 4

2.0 MEPDG DATA REQUIREMENTS AND WIM DATA QUALITY CONTROL ..................... 6
  2.1 MEPDG Traffic Data Requirements .................................................................................... 6
  2.2 Weigh-in-motion Data Quality Control .............................................................................. 11
    2.2.1 Scope and Objectives ....................................................................................................... 12
    2.2.2 Literature Review ............................................................................................................ 12
    2.2.3 Methodology ................................................................................................................... 13
    2.2.4 The NCDOT QC Process ............................................................................................... 15
    2.2.5 Examples of QC Anomalies ............................................................................................. 18
    2.2.6 Application of Local Knowledge .................................................................................... 22
    2.2.7 Recommendations for Future Research ....................................................................... 23

3.0 NC URBAN AND RURAL TRUCK TRAFFIC PROFILES .................................................. 27
  3.1 Scope and Objectives .............................................................................................................. 27
  3.2 Approach .............................................................................................................................. 28
  3.3 Low Weight Screening ......................................................................................................... 29
  3.4 Results .................................................................................................................................. 29
  3.5 Recommendations for Future Research ............................................................................. 30

4.0 NC vs. UNIVERSITY OF ARKANSAS WIM QC ANALYSIS .............................................. 39
  4.1 Scope and Objectives ............................................................................................................. 39
  4.2 Vehicle Class Data Analysis Results .................................................................................. 40
  4.3 The NC and Arkansas Methodologies ............................................................................... 42
  4.4 Truck Weight Data Analysis Results .................................................................................. 44
  4.5 Recommendations for Future Research ............................................................................ 47
LIST OF TABLES

TABLE 1  MEPDG Input Requirements and Sources of Data .......................................................... 10
TABLE 2  NCDOT QC Rule List for W-Cards [LTPP QC Vol 1, 2001] ....................................... 17
TABLE 3  NCDOT QC Rule List for Class Data [LTPP QC Vol 1, 2001] ................................. 18
TABLE 4  GVW Weight Ranges for Peaks .................................................................................. 23
TABLE 5  WIM Stations in North Carolina .............................................................................. 35
TABLE 6  Minimum GVW for Trucks ....................................................................................... 36
TABLE 7  Plots Generated Using the NCDOT QC Database .................................................... 36
TABLE 8  NC vs. Arkansas WIM Data Analysis Comparative Study ......................................... 44
LIST OF FIGURES

FIGURE 1  Google Earth© Rendering of NCDOT WIM Stations ........................................... 3
FIGURE 2  FHWA Vehicle Classification [Sarasota-Manatee MPO, 2009] .............................. 9
FIGURE 3  NCDOT WIM QC Flow Diagram ............................................................................. 14
FIGURE 4  NCDOT WIM QC Checks Interface for Truck Weight Data ................................. 15
FIGURE 5  Invalid Class 9 GVW Plot for WIM Station 375827 ............................................. 19
FIGURE 6  Corrected Class 9 GVW Plot for WIM Station 375827 ........................................ 20
FIGURE 7  Selected NCDOT WIM Stations ............................................................................. 20
FIGURE 8  Invalid Average Daily Class Distribution by DOW for WIM 371902 .................... 21
FIGURE 9  Corrected Average Daily Class Distribution by DOW for WIM 375827 ............... 21
FIGURE 10  Class 9 GVW Plot by Month at WIM Station 372101 ........................................ 23
FIGURE 11  WIM Station Groups ............................................................................................. 31
FIGURE 12  WIM Station Groups ( continuation ) ................................................................. 31
FIGURE 13  Google Earth© Rendering of WIM Stations in the NC Mountainous Region .... 32
FIGURE 14  Google Earth© Rendering of WIM Stations in the NC Piedmont Region .......... 33
FIGURE 15  Google Earth© Rendering of WIM Stations in the NC Coastal Region .......... 34
FIGURE 16  Summary Class Plot for all Rural WIMs in the NC Coastal Region ..................... 37
FIGURE 17  Class 5 GVW Plot for all Rural WIMs in the NC Coastal Region ....................... 38
FIGURE 18  Class 9 GVW Plot for all WIMs in the NC Coastal Region ................................ 38
FIGURE 19  University of Arkansas - Monthly Truck Class Distribution Plot ....................... 40
FIGURE 20  NCSU/NCDOT Monthly Truck Class Distribution Plot .................................... 41
FIGURE 21  Yearly Truck Class Distribution Comparison Plot – Possible Error in Code .... 41
FIGURE 22  University of Arkansas WIM Data Analysis Flow Diagram .............................. 42
FIGURE 23  NCSU/NCDOT WIM Data Analysis Flow Diagram ........................................ 43
FIGURE 24  University of Arkansas – Gross Vehicle Weight Plot ....................................... 46
FIGURE 25  NCSU/NCDOT Gross Vehicle Weight Plot ....................................................... 46
FIGURE 26  SQL Server Data Import Flow Diagram ............................................................. 49
FIGURE 27  SQL Server 2008 Database Layout with Selected Vehicle Class Records ....... 50
FIGURE 28  SQL Server 2008 Interface with SQL Query View ............................................ 50
FIGURE 29  Analysis Cube Concept ....................................................................................... 54
FIGURE 30 Analysis Cube Data Source View in SQL Server 2008 ........................................55
FIGURE 31 Rural Summary Class Plot for the NC Coastal Region ........................................66
FIGURE 32 Rural Class 5 GVW Plot for the NC Coastal Region ...........................................67
FIGURE 33 Rural Class 9 GVW Plot for the NC Coastal Region ...........................................67
FIGURE 34 Rural Summary Class Plot for the NC Central Region .........................................68
FIGURE 35 Rural Class 5 GVW Plot for the NC Central Region ...........................................69
FIGURE 36 Rural Class 9 GVW Plot for the NC Central Region ...........................................69
FIGURE 37 Rural Recreational Summary Class Plot for the NC Mountainous Region ..........70
FIGURE 38 Rural Recreational Class 5 GVW Plot for the NC Mountainous Region ..........71
FIGURE 39 Rural Recreational Class 9 GVW Plot for the NC Mountainous Region ..........71
FIGURE 40 Rural Recreational Summary Class Plot for the NC Coastal Region ................72
FIGURE 41 Rural Recreational Class 5 GVW Plot for the NC Coastal Region ....................73
FIGURE 42 Rural Recreational Class 9 GVW Plot for the NC Coastal Region ....................73
FIGURE 43 Summary Class Plot for all Urban WIMs ............................................................74
FIGURE 44 Class 5 GVW Plot for All Urban WIMs ..............................................................75
FIGURE 45 Class 9 GVW Plot for All Urban WIMs ..............................................................75
FIGURE 46 WIM 557 - Urban and Some Recreational Summary Class Plot .........................76
FIGURE 47 WIM 557 - Urban and Some Recreational Class 5 GVW Plots ............................77
FIGURE 48 WIM 557 - Urban and Some Recreational Class 9 GVW Plots ............................77
FIGURE 49 Asheville Urban and Recreational Summary Class Plot ................................76
FIGURE 50 Asheville Urban and Recreational Class 5 GVW Plot ....................................79
FIGURE 51 Asheville Urban and Recreational Class 9 GVW Plot ....................................79
FIGURE 52 Summary Class Plot for all I-95 WIMs ..............................................................80
FIGURE 53 Class 5 GVW Plot for all I-95 WIMs .................................................................81
FIGURE 54 Class 9 GVW Plot for All I-95 WIMs .................................................................81
FIGURE 55 NC Avg Daily Truck Class Distribution by Month for WIM 378201 ................83
FIGURE 56 Arkansas Truck Class Distribution for WIM 378201 .........................................83
FIGURE 57 NC Average Daily Truck Class Distribution by Month for WIM 375903 ..........84
FIGURE 58 Arkansas Truck Class Distribution for WIM 375903 .........................................84
FIGURE 59 NC Average Daily Truck Class Distribution by Month for WIM 374301 ..........85
FIGURE 60 Arkansas Truck Class Distribution for WIM 374301 .........................................85
1.0 INTRODUCTION

1.1 Background
In 2003, the National Cooperative Highway Research Program released the Mechanistic-Empirical Design Guide (MEPDG) for New and Rehabilitated Pavement Structures. According to NCHRP Report 538, the 1-37A research provides engineers with practical and realistic pavement design procedures and software that use existing mechanistic-empirical principles [Hallenbeck et. al, 2005]. The mechanistic-based distress prediction models used in the MEPDG will require the input of specific data for each axle type and axle-load group. Project 1-39 research develops procedures and software for collecting and processing traffic data required by the Pavement Design Guide procedures.

The National Cooperative Highway Research Program (NCHRP) Projects 1-37A and 1-39 have consistent results except that 1-39 stresses the need for using available DOT data resources and for developing statewide Level 3 data in place of national defaults. NCDOT will adopt the MEPDG procedures in the next few years; however, the models in the MEPDG contain design parameters based on limited national databases. For NCDOT, therefore, it is critical to calibrate the new design methods using NC design input data.

Four types of input data are required to use the MEPDG: structure, climate, material, and traffic data. The structure data are input by a user at the time of a pavement design. The climate data are automatically entered by the MEPDG software for the location of the design project. The material database for NC is under development in NCDOT project HWY-2007-07. Therefore, the final missing input data for the NC calibration of the MEPDG is the traffic data, the focus of this research. Briefly, the traffic data describe truck traffic volume by class and axle load spectra, and future forecasts of truck traffic.

1.2 Research Scope and Objectives
The common geographical scope of this thesis research is limited to the state of North Carolina. All WIM data collected by NCDOT for the MEPDG project was used for this research. The data related scope includes 12 consecutive months of best quality vehicle weight and class data from...
45 selected WIM stations. However, in some cases, due to data quality issues, entire months or days of vehicle weight or class data had to be replaced with data from another year. Federal Highway Administration (FHWA) vehicle classes 4 – 13 are the only classes being considered for vehicle weight data. However, vehicle class data is used in its entirety.

1.3 WIM Site Distribution in North Carolina
The data used to develop the research came from 45 WIM stations operated by NCDOT, all of them being critically important as sources of data. They were operational during the period 1997–2007. For WIM data management based on geographic features, NC is divided by three regions (Mountain, Central/Piedmont and Coastal).

A Keyhole Markup Language (KML) file has also been created from the NCDOT WIM Stations shapefile in ArcGIS. This KML file enables the user to view the location of all the NCDOT WIM stations real-time in Google Earth©. All WIM stations features have been labeled and 15 database attributes including SHRP ID, nearest city, route, number of lanes etc., have also been added. Figure 1 shows a screenshot image of all 40 single cabinet WIM stations and seven two cabinet WIM stations accounting to a total of 47 WIM stations in the state of North Carolina.
FIGURE 1 Google Earth® Rendering of NCDOT WIM Stations
1.4 Overview, Scope, and Objectives

This document consists of four major stand-alone chapters that address relatively stand-alone problems. Each problem is developed in a separate chapter and even though the chapters are typically independent, the data used for all chapters are of the same type (Class/C-card and Weight/W-card data). All vehicle class and weight datasets used have to undergo a series comprehensive quality control procedure checks before they can be used for any further analysis. Here, the technical scope and objectives of each chapter are briefly summarized below.

Chapter 2: MEPDG Data Requirements and WIM Data Quality Control - The scope of this chapter is limited to the development and application of QC methods for truck class and weight data from NCDOT WIM stations collected from 1997 to 2007. Twelve consecutive months of data were processed for each station and in most cases the date ranges span two calendar years. The date ranges were chosen based on the history of the equipment and sensors used at each station. Data collected immediately after installation, calibration, and validation of new sensors were selected to ensure use of the best quality data to support MEPDG development. However, in a few cases, due to operational or data quality issues, an entire month of vehicle class and weight data were replaced with data from another year. Although this created a discontinuity in those datasets, the requirements of the MEPDG design process are met. The quality control analysis was a joint effort where NCSU analyzed 32 of the WIM datasets and NCDOT analyzed 13 datasets. The objective of this chapter is to explain the procedure used to perform the QC measures developed for NCDOT WIM data. A WIM data system must include a quality control process to exclude anomalous measurements inherent to them, to validate data used in the MEPDG process, and to provide a measure of the quality of the data collected. Many states are struggling with developing the data inputs for the MEPDG pavement design process, and the information provided herein will aid them in generating the traffic data component of those inputs. The database used for performing QC is called the NCDOT WIM QC Database and applies SQL queries in linked Access database tables.

Chapter 3: NC Urban and Rural Truck Traffic Profiles - The geographical scope of this research is limited to the state of NC just like all the other chapters of this report. However, a geographical scope unique to this chapter is the classification of NC with respect to region.
Twelve consecutive months of best quality vehicle weight and class data was chosen and used for
analysis thus accounting for a data related scope as well. The objectives of this chapter are to
develop truck weight and class plots for NC urban and rural regions. Important plots include
FHWA class 5 and class 9 GVW plots and average daily truck class distribution plots by month.

Chapter 4: NC vs. University of Arkansas WIM QC Analysis - The geographical scope is limited
to the state of NC and the data related scope includes 12 consecutive months of best quality
vehicle weight and class data. FHWA vehicle classes 4 - 13 are the only classes being considered
for vehicle weight data. Vehicle class data has been used in its entirety. The scope of the
analysis includes GVW analysis for vehicle weight data and vehicle class distribution by month
and year for vehicle class data. The objective of this chapter is to justify the need for a
performing a comprehensive QC procedure on vehicle weight and class data used for MEPDG.
This is accomplished by comparing the analysis methodologies practiced by NCSU/NCDOT and
the University of Arkansas for performing quality control analysis on vehicle weight and class
data.

Chapter 5: WIM Data Management and Analysis in SQL Server - The scope of this chapter is
similar to the other chapters with respect to the type of data used. However, for the sake of
testing the WIM data application using the analysis cube concept in SQL Server, truck weight and
vehicle class databases for two WIM stations (SHRP ID 371028 and 375601) located in Camden
County and Madison County respectively were used for the analysis. The objectives of this
chapter are to explain the application of WIM data analysis using Analysis Cubes in Microsoft
SQL Server and to emphasize the importance of transitioning to a more advanced database
management system for the NCDOT statewide WIM program.
2.0 MEPDG DATA REQUIREMENTS AND WIM DATA QUALITY CONTROL

At hundreds of WIM sensor locations State Departments of Transportation collect gigabytes of vehicle data every year to meet Federal reporting traffic requirements [LTPP, HPMS and VTRIS, 2009]. The DOTs process the data to support pavement design to enforce weight restrictions on highways and bridges, and to provide planning data for highway improvements. The objective of this chapter is to describe MEPDG traffic data requirements and WIM data QC procedures developed by NC State University for the North Carolina Department of Transportation. It also outlines the type, source and format of MEPDG input data and any other information that is available as a result of analyzing traffic data collected by NCDOT. The difference between MEPDG traffic data requirements and what is available from data collection and/or analysis represents any deficiencies in the data required to conduct an MEPDG simulation These topics are presented below and recommendations are made for performing better quality control procedures on vehicle weight and class datasets based on the results of the quality control procedures performed by the NCSU MEPDG team at NC State University. After performing QC for 45 NCDOT WIM stations, there were about 100 databases stored in approximately 50 gigabytes of data including plots, reports, and tables. The QC results confirmed that NCDOT equipment captured reliable WIM measurements. Vehicle class and weight checks generate 0.97% and 6.42% anomalies, respectively. Average QC run-time was 60 minutes per WIM station. Manual inspection of questionable data further clarified the results finding that only three WIM stations had more than 10% flagged weight records with a majority of them caused by vehicle misclassifications.

2.1 MEPDG Traffic Data Requirements

The 2002 MEPDG requires seven types of traffic data to fully characterize the traffic pattern for the design of new or rehabilitated pavement structures. The required traffic data are the same for designing either new or rehabilitated or for either flexible or rigid pavements. The following list identifies the seven typical required traffic data [NCHRP, 1999, 2000, 2004]
1. Normalized axle load distribution or spectra for each axle type within each vehicle class. By definition, the axle load spectra are the number of load passes by axle configuration and load bin over a time interval. The full axle load spectrum for the single, tandem, tridem, and quad axle types is required to fully express the load effect on pavement design. Axle load spectra are treated as constant throughout the pavement design life. This critical data results as a text file output from the NCDOT WIM Processor. The WIM Processor must be given W card data for a site for 12 consecutive months.

2. Average Annual Daily Truck Traffic (AADTT). The AADTT can be updated annually through a specific linear or compound growth rate model.

3. Percent Trucks. Percent trucks represent the percentage of vehicle classes 4-13 in the traffic stream. MEPDG currently handles 10 standard FHWA vehicle classes (class 4 to 13). Figure 2 shows the entire 13 standard FHWA vehicle classes.

4. Truck Traffic Classification (TTC) for Pavement Structural Design. This factor classifies those highways into groups with similar truck traffic features or characteristics that are needed for pavement structural designs for selecting the default values for the Level 3 inputs of various traffic parameters.

5. Loading details of the axle load and axle configuration. Default values are provided for each of the following elements that describe the details of the tire and axle loads. However, the designer can use a different set of values based on site specific data.
   a. Tire pressure and dual tire spacing – Default values are provided depending on the axle load for single or dual usage.
   b. Axle spacing – Default values are provided depending on three axle types (tandem, tridem and quad).
   c. Wheelbase distribution – Wheelbase refers to the spacing between the steering and the first device axle of the truck-tractor or heavy single units. Default values are provided for short, medium and long axle spacing.
   d. Average number of axles by axle type per vehicle classification – Default values depend on the type of the axle per vehicle class.

6. Traffic time distribution factors. Two types of truck distribution factors are needed as inputs. These factors are monthly distribution factors and hourly distribution factors. Monthly or
seasonal distribution factors are used to adjust the AADTT into monthly Average Daily Truck Traffic (ADTT) values/volumes, while the hourly distribution factor is used to distribute the monthly ADTT volumes by hour of the day. These time-dependant distribution factors are determined from detailed studies of the AVC or WIM data. In addition, these factors are treated as constant throughout the pavement design life.

7. Traffic factors. Default values for each of the following elements are provided for different types of highways. However, the designer can specify different values when site specific data are available. Once specified, these factors are treated as constants throughout the pavement design life.

a. Directional distribution factor – The directional distribution factors account for the percentage of trucks in one direction of the total truck traffic population.

b. Lane distribution factor – The lane distribution factors account for the percentage of trucks of the total truck traffic population in one lane (i.e., one direction). These lane distribution factors are area (urban versus rural) and highway (number of lanes) dependent.

c. Traffic growth factor – The traffic growth factor allows for the growth or decay in truck traffic over the design period. The growth factor can be vehicle class (class 4-13) dependent. This factor varies significantly from one highway to another, therefore, historical traffic information are a crucial source of information to determine the growth rate. There are three available functions to present changes in the traffic factor. The first one shows an increase at an increasing rate or decrease at a decreasing rate. The second one expresses the linear increase per year while the third one assumes that truck traffic remains constant.

d. Operational speed – For pavement design, operational speed is more reliable than posted speed. However, data for the posted speed can be obtained easily. There are procedures to convert the posted speed to operational speed.

Table 1 identifies the sources of MEPDG traffic inputs. They may be developed from the WIM Database [Mastin, 2007] (which reaches back to weight and class counts), from defaults, or other sources. Table 1 also shows the linkage or bridge between MEPDG traffic data and what is available from NCDOT traffic data sources with/without analysis.
<table>
<thead>
<tr>
<th>CLASS GROUP</th>
<th>DESCRIPTION</th>
<th>NO. OF AXLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MOTORCYCLES</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>ALL CARS CARS</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>CARS W/ 1-AXLE TRAILER</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>CARS W/ 2-AXLE TRAILER</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>PICK-UPS &amp; VANS 1 &amp; 2 AXLE TRAILERS</td>
<td>2, 3, &amp; 4</td>
</tr>
<tr>
<td>4</td>
<td>BUSES</td>
<td>2 &amp; 3</td>
</tr>
<tr>
<td>5</td>
<td>2-AXLE, SINGLE UNIT</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3-AXLE, SINGLE UNIT</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>4-AXLE, SINGLE UNIT</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>2-AXLE, TRACTOR, 1-AXLE TRAILER (2&amp;1)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2-AXLE, TRACTOR, 2-AXLE TRAILER (2&amp;2)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3-AXLE, TRACTOR, 1-AXLE TRAILER (3&amp;1)</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>3-AXLE, TRACTOR, 2-AXLE TRAILER (3&amp;2)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3-AXLE, TRUCK W/ 2-AXLE TRAILER</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>TRACTOR W/ SINGLE TRAILER</td>
<td>6 &amp; 7</td>
</tr>
<tr>
<td>11</td>
<td>5-AXLE MULTI-TRAILER</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>6-AXLE MULTI-TRAILER</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>ANY 7 OR MORE AXLE</td>
<td>7 or more</td>
</tr>
<tr>
<td>14</td>
<td>NOT USED</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>UNKNOWN VEHICLE TYPE</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 2** FHWA Vehicle Classification [Sarasota-Manatee MPO, 2009]
<table>
<thead>
<tr>
<th>Input Needed by MEPDG software</th>
<th>Manual Input (Yes/No)</th>
<th>Source of Manual Input</th>
<th>Automated Input (Yes/No)</th>
<th>Source of Automated Data</th>
<th>Developed by WIM Processor (Yes/No)</th>
<th>Format Ready for MEPDG Input (Yes/No)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axle Load Distribution Factor</td>
<td>No</td>
<td>-</td>
<td>Yes</td>
<td>WIM Processor</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Two Way AADTT</td>
<td>Yes</td>
<td>Traffic Volume by Vehicle Class</td>
<td>No</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Vehicle Class Distribution</td>
<td>Yes</td>
<td>Truck % per Station ID</td>
<td>No</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Tire pressure and dual tire spacing</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Use MEPDG Default Values</td>
</tr>
<tr>
<td>Axle Spacing &amp; Wheelbase</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Use MEPDG Default Values</td>
</tr>
<tr>
<td>Average Number of axles by axle type per vehicle classification</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Use MEPDG Default Values</td>
</tr>
<tr>
<td>Monthly Adjustment Factors</td>
<td>Yes*</td>
<td>-</td>
<td>No*</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>-</td>
</tr>
<tr>
<td>Hourly Distribution Factors</td>
<td>Yes</td>
<td>Traffic Volume by Vehicle Class</td>
<td>No</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>For Design Direction</td>
</tr>
<tr>
<td>Percent Trucks in Design Direction</td>
<td>Yes</td>
<td>Traffic Volume by Vehicle Class</td>
<td>No</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Must be calculated</td>
</tr>
<tr>
<td>Percent Trucks in Design Lane</td>
<td>Yes</td>
<td>NA</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Use default definition based on number of lanes</td>
</tr>
<tr>
<td>Traffic Growth Factor</td>
<td>Yes</td>
<td>NA</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Must be determined</td>
</tr>
</tbody>
</table>

* The MEPDG software has the provision to accept automated input for Monthly Adjustment Factors, but since, it is not developed by the WIM processor database, the preferred mode of input is “manual”, for the time being; NA – Not Available
2.2 Weigh-in-Motion Data Quality Control

At hundreds of WIM monitoring locations, State DOTs collect WIM data every year to meet Federal traffic reporting requirements [LTPP, HPMS, and VTRIS]. The large WIM datasets include date and time stamps for most vehicles crossing the WIM sensors, counts of each vehicle’s axles, and measurements of the individual axle loads and spacing. The results provide statistics on the annual average daily traffic for each of the 13 FHWA vehicle classes [FHWA Veh. Types], time histories of gross vehicle weights (GVW) by class, frequency distributions of the traffic by vehicle class, axle load spectra, and other important characteristics of the traffic flow at each WIM station.

The NCDOT uses WIM systems and procedures that are consistent with recommended industry practices as specified by the FHWA Long Term Pavement Performance Program [LTPP, 2005], the Traffic Monitoring Guide [TMG, 2001], and the American Association of State Highway and Transportation Officials [AASHTO, 2009]. The NCDOT collects class data for all types of vehicles, but weight measurements on truck vehicle classes only. The NCDOT WIM systems are designed to estimate static axle weights based on dynamic measurements. Site selection, pavement condition, system calibration, and system monitoring are as important as the sensor technology used to ensure collection of good quality weight measurements. Additionally, the measurements are affected by driver behavior such as weaving, accelerating and decelerating. Thus, data errors and poor quality data are captured regardless of the technology used, and a QC process is an important part of all WIM data systems. Good quality WIM data can only be captured if a comprehensive quality control process is used, and the QC techniques developed in this research are a key component of that process.

The NCDOT requires that all data with quality issues (such as partial data, misclassified vehicles, and unacceptable axle weights and spacing) be excluded from datasets used for planning and design statistics. Thus, WIM data must undergo a series of sequential, well organized QC procedures to ensure that the data meets MEPDG requirements. This paper documents the NCDOT WIM QC procedures. The results of the QC analysis provide reliable datasets for use in developing Levels 1, 2, and 3 traffic data inputs for the North Carolina MEPDG models [MEPDG, 2009].
2.2.1 Scope and Objectives

The scope of this chapter is limited to the development and application of QC methods for truck class and weight data from NCDOT WIM stations collected from 1997 to 2007. Twelve consecutive months of data were processed for each station and in most cases the date ranges span two calendar years. The date ranges were chosen based on the history of the equipment and sensors used at each station. Data collected immediately after installation, calibration, and validation of new sensors were selected to ensure use of the best quality data to support MEPDG development. However, in a few cases, due to operational or data quality issues, an entire month of vehicle class and weight data were replaced with data from another year. Although this created a discontinuity in those datasets, the requirements of the MEPDG design process are met. The quality control analysis was a joint effort where NCSU analyzed 32 of the WIM datasets and NCDOT analyzed 13 datasets.

The objective of this chapter is to explain the procedure used to perform the QC measures developed for NCDOT WIM data. A WIM data system must include a quality control process to exclude anomalous measurements inherent to them, to validate data used in the MEPDG process, and to provide a measure of the quality of the data collected. Many states are struggling with developing the data inputs for the MEPDG pavement design process, and the information provided herein will aid them in generating the traffic data component of those inputs. The database used for performing QC is called the NCDOT WIM QC Database and applies SQL queries in linked Access database tables.

2.2.2 Literature Review

There are a number of QC procedures that can be implemented for WIM data. The most recognized procedure is the LTPP procedure which guides many of the tests in the NCDOT WIM QC Database. Additionally, the TMG and AASHTO guides [TMG, AASHTO] are industry standards and emphasize the need for quality control measures in traffic monitoring programs. There are also state and project specific traffic data QC requirements, e.g., for traffic forecasting in Texas [Qu et.al, 1997; Lee et. al, 1998] and for truck axle spectra in Oregon [Elkins et.al, 2008]. There is also a recent database application to support the MEPDG effort of the Arkansas State Highway and Transportation Department [Wang, 2009].
2.2.3 Methodology

After reviewing the literature and considering prototype procedures at NCDOT, the authors concluded that the most efficient method of performing the WIM QC included SQL queries in a front-end database system applied to raw data stored in live back-end databases. The resulting NCDOT procedure is presented below. The QC technique uses a combination of rule based checks and manual audits of plots and reports. The rule based checks flag anomalous data that are improperly coded or have values that are out of range. The manual audits identify deviations in patterns that indicate equipment malfunction or invalid data sets.

Figure 3 shows a flowchart of the NCDOT WIM QC process. One year of class and weight data for each NCDOT WIM station was converted to fixed width ASCII text files in FHWA C (class) and W (weight) card formats. The data captured may be from twelve months within a year or 12 consecutive months over two years. The conversion utility generates an individual text file for each day of data for each data type. To ease the process of capturing the data into the QC applications, WIM data are aggregated into a single text file for each data type and then imported into the back end WIM Database developed by Neil Mastin (Figure 3). The WIM Database was developed for an earlier evaluation of the MEPDG design process and it is made more comprehensive by the NCDOT QC Database. The WIM Database is used as a capture utility, and during the import process it performs basic QC checks including exclusion of unclassified vehicles (Class 15 data) and invalid data types (e.g. text data in a number field) [Mastin, 2009]. The WIM Database is the back end for the more detailed QC analysis.
The NCDOT WIM QC Database uses Microsoft Access to apply rule based checks, generate statistics, and prepare data for capture into Excel for generating summary plots. It is the front end for the QC process, is linked to the tables in the WIM Database, and extends its procedures. The output of the NCDOT WIM QC Database application is a table of accepted data for use in M-E pavement design and a table of excluded data for each data type. Some records may have QC issues that could cause multiple flags. However, once a record is flagged and excluded for failing
a check, it is excluded from subsequent checks. The flag that causes a record to be excluded is documented in the exclusion tables.

![Traffic Survey - Continuous Count Program](image)

**FIGURE 4** NCDOT WIM QC Checks Interface for Truck Weight Data

### 2.2.4 The NCDOT QC Process

The NCDOT WIM QC process consists of a combination of automated and manually applied procedures in a user friendly graphical user interface (GUI) as shown in Figure 4. The “QC Sets” are a collection of auto-applied rules (Tables 2 and 3) that identify invalid entries for the fields checked and that automatically exclude flagged records into the exclusion tables. The QC process also consists of “Forms” (Tables 2 and 3) which are a set of flags that alert the analyst to review the data displayed in the forms and manually exclude invalid data. In addition to the automated processes of the QC Sets and Forms, the QC process also utilizes plots and statistics to qualitatively assess the suitability of the data for the MEPDG process. Unusual patterns in the plots indicate that inconsistent data has been captured and that it may not be suitable. The process is applied sequentially where all steps for a data type are performed in order and weight QC is completed prior to evaluating class data. Weight QC is performed first as weight
measurements are more complex, have many more sources of error, and are more likely to cause data to be excluded or replaced than class data.

A summary of all flagged records is generated and reported. The last feature of the QC function is generation of a table of records not flagged during the QC process. This will be the accepted dataset if the plots and report do not identify anomalous data. If inconsistent or invalid data are identified in the plots or report, the analyst returns to the start of the process, removes the text files in W/C card format for the month found to be anomalous, and replaces it with W/C card data for the same month from a different year. The entire QC process is rerun to ensure that the final datasets meet the requirements of the MEPDG process. Tables 2 and 3 show the QC checks applied by the NCDOT WIM QC Database. During the review process, local knowledge of the site and traffic conditions is considered and any additional anomalous records are manually flagged.
<table>
<thead>
<tr>
<th>Order</th>
<th>ID</th>
<th>Description</th>
<th>Criteria</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>W_NULL</td>
<td>Any field with a null value</td>
<td>Field Value ≠ Null</td>
<td>QCSet1</td>
</tr>
<tr>
<td>2</td>
<td>W12</td>
<td>Invalid hour</td>
<td>HOUR ≠ (0 - 23)</td>
<td>QCSet1</td>
</tr>
<tr>
<td>3</td>
<td>W10</td>
<td>Invalid month</td>
<td>MONTH ≠ (1 – 12)</td>
<td>QCSet1</td>
</tr>
<tr>
<td>4</td>
<td>W16</td>
<td>Invalid vehicle class code</td>
<td>VHCL_CLASS ≠ (4 - 13)</td>
<td>QCSet1</td>
</tr>
<tr>
<td>5</td>
<td>W14</td>
<td>Invalid FIPS Code</td>
<td>STATE_CD ≠ 37</td>
<td>Forms1</td>
</tr>
<tr>
<td>6</td>
<td>W6</td>
<td>Invalid station ID</td>
<td>DRCTN_CD ≠ Valid values for station</td>
<td>Forms1</td>
</tr>
<tr>
<td>7</td>
<td>W8</td>
<td>Invalid direction for station</td>
<td>TRVL_LN_NBR ≠ Valid values for station</td>
<td>Forms1</td>
</tr>
<tr>
<td>8</td>
<td>W7</td>
<td>Invalid lane number for station</td>
<td>YEAR ≠ Valid year for date range captured</td>
<td>Forms1</td>
</tr>
<tr>
<td>9</td>
<td>W9</td>
<td>Invalid year</td>
<td>MANUAL # Valid date for the MONTH</td>
<td>Forms1</td>
</tr>
<tr>
<td>10</td>
<td>W11</td>
<td>Hour without any weight records. A full day of data may not be available for all lanes</td>
<td>Manual audit of hours without weight records</td>
<td>Forms1</td>
</tr>
<tr>
<td>11</td>
<td>W13</td>
<td>Axle count inconsistent with number of axle spacings</td>
<td>AXLE_COUNT ≠ (# of spacings +1)</td>
<td>QCSet2</td>
</tr>
<tr>
<td>12</td>
<td>W1</td>
<td>Axle count inconsistent with number of axle weights</td>
<td>TOTAL_WGHT ≠ Sum of axle weights</td>
<td>QCSet2</td>
</tr>
<tr>
<td>13</td>
<td>W2</td>
<td>Axle weight is out of acceptable range</td>
<td>441 lb (200 kg) &lt; (X)_WGHT &lt; 44,100 lb (20,003.4 kg)</td>
<td>QCSet2</td>
</tr>
<tr>
<td>14</td>
<td>W3</td>
<td>Axle spacing is out of acceptable range</td>
<td>1.97 ft (0.6 m) &lt; (X)(Y)_SPACING &lt; 49.2 ft (15 m)</td>
<td>QCSet2</td>
</tr>
<tr>
<td>15</td>
<td>W17</td>
<td>Sum of axle spacings exceeds maximum wheelbase</td>
<td>Sum of axle spacings &gt; 98.2 ft (29.93 m)</td>
<td>QCSet2</td>
</tr>
<tr>
<td>16</td>
<td>WP1</td>
<td>Review Average DOW volumes by month for unusual patterns</td>
<td>A pattern deviates significantly from other months</td>
<td>Plots</td>
</tr>
<tr>
<td>17</td>
<td>WP2</td>
<td>Review GVW plots by class by month for unusual patterns</td>
<td>A pattern deviates significantly from other months</td>
<td>Plots</td>
</tr>
</tbody>
</table>
TABLE 3 NCDOT QC Rule List for Class Data [LTPP QC Vol 1, 2001]

<table>
<thead>
<tr>
<th>Order</th>
<th>ID</th>
<th>Description</th>
<th>Criteria</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C_NULL</td>
<td>Any field with a null value</td>
<td>Field Value = Null</td>
<td>QCSet3</td>
</tr>
<tr>
<td>2</td>
<td>C8</td>
<td>Invalid month</td>
<td>MONTH ≠ (1 – 12)</td>
<td>QCSet3</td>
</tr>
<tr>
<td>3</td>
<td>C10</td>
<td>Invalid hour</td>
<td>HOUR ≠ (0 – 23)</td>
<td>QCSet3</td>
</tr>
<tr>
<td>4</td>
<td>C1</td>
<td>Total lane volume exceeds max. limit</td>
<td>TOTAL_VOL &gt; 3000</td>
<td>QCSet4</td>
</tr>
<tr>
<td>5</td>
<td>C11</td>
<td>Invalid FIPS Code</td>
<td>STATE_CD ≠ 37</td>
<td>Forms2</td>
</tr>
<tr>
<td>6</td>
<td>C4</td>
<td>Invalid station ID</td>
<td>STATION_ID ≠ Expected station identifier</td>
<td>Forms2</td>
</tr>
<tr>
<td>7</td>
<td>C6</td>
<td>Invalid direction for station</td>
<td>DRCTN_CD ≠ Valid values for station</td>
<td>Forms2</td>
</tr>
<tr>
<td>8</td>
<td>C5</td>
<td>Invalid lane number for station</td>
<td>TRVL_LN_NBR ≠ Valid values for station</td>
<td>Forms2</td>
</tr>
<tr>
<td>9</td>
<td>C7</td>
<td>Invalid year</td>
<td>YEAR ≠ Valid year for date range captured</td>
<td>Forms2</td>
</tr>
<tr>
<td>10</td>
<td>C9</td>
<td>Invalid day</td>
<td>DAY ≠ Valid date for the MONTH</td>
<td>Forms2</td>
</tr>
<tr>
<td>11</td>
<td>C3</td>
<td>A full day of data is not available for a day for all lanes</td>
<td>Manual audit of hours and days</td>
<td>Forms2</td>
</tr>
<tr>
<td>12</td>
<td>C2</td>
<td>Class volume exceeds maximum limit</td>
<td>CLS_CNT_## = TOTAL_VOL</td>
<td>Forms3</td>
</tr>
<tr>
<td>13</td>
<td>C13</td>
<td>IAM total lane volume exceeds 1PM total lane volume</td>
<td>HOUR(1) TOTAL_VOL &gt; HOUR(13) TOTAL_VOL</td>
<td>Forms3</td>
</tr>
<tr>
<td>14</td>
<td>C14</td>
<td>Static total lane volume for four consecutive hours</td>
<td>HOUR(X) TOTAL_VOL = HOUR(X+1,+2,+3) TOTAL_VOL</td>
<td>Forms3</td>
</tr>
<tr>
<td>15</td>
<td>CP1</td>
<td>Review Avg. DOW volumes by month for unusual patterns</td>
<td>A pattern deviates significantly from other months</td>
<td>Plots</td>
</tr>
<tr>
<td>16</td>
<td>CP2</td>
<td>Review Class Distribution by month for unusual patterns</td>
<td>A pattern deviates significantly from other months</td>
<td>Plots</td>
</tr>
<tr>
<td>17</td>
<td>CP3</td>
<td>Review Class % Distributions for unusual patterns</td>
<td>The summary data exhibits an unusual pattern</td>
<td>Plots</td>
</tr>
</tbody>
</table>

2.2.5 Examples of QC Anomalies

Figures 5 and 6 represent examples of replacing data to rectify an invalid data set identified by the QC. WIM station 375827 is located on US 29 in Rockingham County in the central Piedmont region (Figure 7). In this case, the Class 9 GVW plot (Figure 5) shows invalid weight measurements where 10 months of data have shifted peaks to the left from what is normal (about 35,000 pounds empty and 75,000 pounds loaded). When the invalid data are replaced with data for a properly calibrated site, the peaks line up and the data is accepted (Figure 6).
As another example, consider WIM station 371902 located on US 74 in Cherokee County (western NC). In this example (Figure 8), NCDOT discovered that FHWA Class 8 truck volumes drop to zero for a couple of months (September is used for this example) at this site. This anomaly was caused by an error in the class algorithm used at that time. NCDOT replaced this month with data from another year collected with a valid class algorithm and the Class 8 volumes return to normal levels (Figure 9). The calibration example (Figures 5 and 6) and class algorithm example (Figures 8 and 9) illustrate how data replacement resolves the QC problem.

FIGURE 5 Invalid Class 9 GVW Plot for WIM Station 375827
FIGURE 6  Corrected Class 9 GVV Plot for WIM Station 375827

FIGURE 7  Selected NCDOT WIM Stations
FIGURE 8  Invalid Average Daily Class Distribution by DOW for WIM 371902

FIGURE 9  Corrected Average Daily Class Distribution by DOW for WIM 375827
2.2.6 Application of Local Knowledge

In addition to the automated and manual checks, the analyst must review weight plots generated for WP1 and WP2 in Table 2 and class plots generated for CP1, CP2, and CP3 in Table 3.

GVW plots sorted by Class 4 through 13 provide a graphical representation of how trucks are loaded. Typical plots with appropriate values have peaks at weight ranges that correspond to empty and fully loaded conditions for that truck class (Figure 6). Inconsistencies with expected peaks for a class, shown in Table 4, or deviations in peaks between months, are the basis for the GVW evaluation. By comparing the peaks in the GVW plots with the values in Table 4, unusual and potentially invalid values for weight measurements can be found and replaced if determined to be invalid.

Unusual values may be an indication of poor weight measurements (Figure 5) or patterns related to local conditions. Figure 10 shows a GVW plot for FHWA class 9 trucks at WIM station 372101 located on US 64 in Clay County. This is a good example of using local knowledge to accept unusual values. Although a class 9 GVW plot usually has two peaks corresponding to empty and loaded conditions, this plot is valid because WIM station 372101 is located in the mountainous western region of North Carolina (Figure 7). Local knowledge of truck traffic patterns confirms a significant reduction in loaded truck volumes because of the mountainous terrain and winding roads. The consistent pattern of the weight ranges for the unloaded peak validates the unusual measurements.

If the GVW pattern is inconsistent with the peaks for loaded and unloaded weights, and if the pattern cannot be justified by local knowledge, then the truck weight data for that time period should be excluded and replaced by data from another year.
TABLE 4  GVW Weight Ranges for Peaks

<table>
<thead>
<tr>
<th>FHWA Vehicle Class</th>
<th>Typical Weight Ranges for Peaks (lbs)</th>
<th>Typical Weight Ranges for Peaks (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>One peak at 20,000</td>
<td>One peak at 9,072</td>
</tr>
<tr>
<td>5</td>
<td>One peak at 10,000</td>
<td>One peak at 4,536</td>
</tr>
<tr>
<td>6</td>
<td>One peak at 20,000 to 25,000</td>
<td>One peak at 9,072 to 11,340</td>
</tr>
<tr>
<td></td>
<td>Other at 45,000 to 55,000</td>
<td>Other at 20,411 to 24,948</td>
</tr>
<tr>
<td>7</td>
<td>One peak at 50,000 to 60,000</td>
<td>One peak at 22,680 to 27,216</td>
</tr>
<tr>
<td>8</td>
<td>One peak at 30,000 to 35,000</td>
<td>One peak at 13,608 to 15,876</td>
</tr>
<tr>
<td>9</td>
<td>One peak at 30,000 to 35,000</td>
<td>One peak at 13,608 to 15,876</td>
</tr>
<tr>
<td></td>
<td>Other at 70,000 to 80,000</td>
<td>Other at 31,751 to 36,287</td>
</tr>
<tr>
<td>10</td>
<td>One peak at 40,000 to 45,000</td>
<td>One peak at 18,144 to 20,412</td>
</tr>
<tr>
<td></td>
<td>Other at 75,000 to 85,000</td>
<td>Other at 34,019 to 38,555</td>
</tr>
<tr>
<td>11</td>
<td>One peak at 55,000 to 60,000</td>
<td>One peak at 24,948 to 27,216</td>
</tr>
<tr>
<td>12</td>
<td>One peak at 55,000 to 65,000</td>
<td>One peak at 24,948 to 29,484</td>
</tr>
<tr>
<td>13</td>
<td>Straight Line (Constant weight range with very low frequency of trucks)</td>
<td>Straight Line (Constant weight range with very low frequency of trucks)</td>
</tr>
</tbody>
</table>

2.2.7 Recommendations for Future Research
Apart from a few sites with up to an entire month of data missing because of construction and other non-equipment problems, the WIM systems operated by the NCDOT provided valid weight and class data. When a month of data was excluded by the QC process, the NCDOT substituted...
equivalent data from the same month but from another year in order to develop complete datasets for MEPDG. Another alternative would be to replace all twelve months of data with valid data from another year.

The following recommendations can be implemented to improve upon the existing quality control procedures implemented by the MEPDG research team at NC State University.

1. **Explore data sampling procedures to reduce the size of databases** - Weight data can result in extremely large databases, especially for sites with high vehicular volume like interstate routes. Statistically valid sampling methods should be explored to reduce the size of such datasets to a manageable extent. Sampling requirements must be based on the intended use of the data. For example, most LTPP GPS studies require only seven days of weight data per quarter.

2. **Consider quality control as part of a comprehensive process called quality assurance** - The quality assurance consists of various quality considerations which are made during data collection and after data summarization and reporting. Some of the quality assurance actions are listed in the following.
   - Develop effective equipment procurement procedure.
   - Establish inspection procedures for newly installed equipments.
   - Schedule periodic maintenance and calibration activities.
   - Develop automated quality control procedure to review and detect corrupt data.
   - Be in contact with the customer and ask for feedback [Turner, 2007].

Considering such activities, recognize that the data quality control is restricted to identifying erroneous data and excluding them from the database is not an efficient practice in long term. The QC process shall be used as a diagnostic tool to identify the reasons behind poor data quality. The NC State research team has no direct control over the quality assurance process other than quality control. The results of the WIM data QC, however, may provide some insight for NCDOT in procuring data collection equipment and developing maintenance programs for equipment.
3. **Assign severity level to quality control rules** - Currently, the NCDOT QC Database simply excludes all the data which are rejected based on identified QC rules. In other words, a simple accept/reject decision is made for each rule. It will be more beneficial to assign a severity level to data which are not valid based on QC rules. Suggested severity levels are:

- **High** – A high level of severity may happen because of equipment failure. For example, 24 consecutive zero values for volume in a single lane of a four-lane highway could be a serious equipment malfunction. There may be cases where one lane is closed by authorities for road maintenance activities, however identifying these reasons may be difficult in practice. The suggested action for data with high level of severity would be to exclude the data from further analysis.

- **Medium** – Data with a medium level of severity are outside the acceptance range, however it is not very significant. In such cases some analysis is required to determine whether to accept or exclude the data. For example, volume level of one at 1:00 am compared to volume level of zero at 1:00 pm is considered an error based on rule number C13, but this case may often happen on low volume roads. Therefore, it is recommended to check the historical data (a few days before and after) to identify whether the trend is repeating or not. In summary, the suggested action for data with medium level of severity is to perform more analysis based on local knowledge of traffic and site conditions at the WIM station.

- **Low** – Data with a low level of severity correspond to rules with boundary limits (maximum and minimum limits). Such data are usually inside the specified limits, however very close to boundaries. For example, a level of volume (2999 vehicles) which is less than 3000 vehicles is considered acceptable base on rule number C1. This volume may be excessive for the geometry of the highway cross section at some WIM stations. It is recommended to investigate such data further based on local knowledge of the WIM station to build more confidence on data quality [Turner, 2007].

4. **Long Term WIM Data Storage and Analysis** - While most research topics focus on collection and quality of data, there is little emphasis on the development of an integrated
Database Management System (DBMS) to store and analyze traffic data. A Microsoft Access database or spreadsheet program can be used to analyze small datasets, but they are not feasible for analyzing a statewide WIM program. Therefore, the best way to effectively handle the tremendous amount of data generated by a WIM monitoring program is to upload it to a more robust relational database, such as Oracle or Microsoft SQL Server [Nichols et.al, 2004].
3.0 NC URBAN AND RURAL PROFILES

Knowing the type of traffic by vehicle class by highway functional classification is critical to designing, maintaining and paying for North Carolina highway pavements. Such information is used by the NCDOT Pavement Management Unit to estimate pavement costs by NC region and highway functional classification and to provide information to the NC Legislature to equitably assess vehicle taxes and fuel taxes to help pay for pavement. Thus, GVW plots by vehicle class and highway functional class are very important.

NCSU developed database procedures and NCDOT collected WIM station data during the NCDOT research project called NC Traffic Data for the Mechanistic Empirical Pavement Design Guide (NCDOT project HWY 2008-11). NCSU used the database procedures on the WIM data to develop GVW plots for urban and rural regions in North Carolina for this project. While GVW plots for all WIM stations for all vehicles classes and highway functional classifications have been developed, the NCDOT Pavement Management Unit recommended that plots for vehicle classes 5 and 9 be selected because they illustrate the majority of truck traffic. To further simplify the analysis and number of GVW plot combinations, NCDOT recommended that all urban WIM stations be grouped and that the remaining rural WIM stations be grouped by NC region – mountainous, central and coastal. During the analysis further categories became apparent - one for I-95 traffic and one for urban recreational traffic in Asheville. The results of this GVW analysis for class 5 and class 9 trucks complement the statewide truck traffic estimates produced using the NC truck network model developed by NCDOT research project HWY 2006-09. This chapter of the report documents the GVW plots mentioned above as well as monthly average vehicle classification plots depicting seasonal variation of truck traffic.

3.1 Scope and Objectives

The geographical scope of this research is limited to the state of NC just like all the other chapters of this report. However, a geographical scope unique to this chapter is the classification of NC with respect to region. Twelve consecutive months of best quality vehicle weight and class data was chosen and used for analysis thus accounting for a data related scope as well. The objectives of this chapter are to develop truck weight and class plots for NC urban and rural regions.
Important plots include FHWA class 5 and class 9 GVW plots and average daily truck class distribution plots by month.

### 3.2 Approach

Table 5 identifies the location of the WIM stations and the nature of the truck traffic (urban, rural, and recreational). All WIM stations are grouped as shown in Figures 11 and 12. Figures 13 through 15 illustrate the location of WIM stations in North Carolina regions (mountainous, central, and coastal). It is to be noted that some WIM station groups are missing because there are no WIMs in that group. Some groups like Urban WIMs and I-95 WIMs are also considered to be a category on its own irrespective of region.

The GVW plots for the urban and rural regions are obtained using the following procedure.

1. Quality Control is performed on the raw C and W card data using the NCDOT QC Database.

2. GVW plots by vehicle class by WIM station are obtained by using the NCDOT WIM QC database, which in turn is connected to the WIM data in the WIM Processor database.

3. These groupings in Figures 13 and 14 were suggested by NCDOT Traffic Survey Unit personnel. In addition WIM stations 519, 520 and 556 in Asheville were classified as a special case to show class 5 and class 9 traffic on urban recreational routes. WIM 557 on I-40 in Statesville was also identified as an Urban WIM with some recreational features.

4. Data from the WIMs shown in Table 5 are processed.

5. Data from WIMs 507, 513, 514, 517, 518, 524, 526, 528, 532, 537, 538, 544 and, 550 are not available:

6. Data from all WIMs are grouped accordingly and the plots shown in Table 7 are plotted in Excel. A total of 49 plots are generated for each WIM station group. For documentation purposes, only the most important plots/truck classes are considered. GVW plots for classes 5 and 9 are chosen since they represent a majority of the truck traffic on highways. In addition, summary class plots obtained from vehicle class data are also shown in this report. Thus, this report produces 24 class 5 and class 9 plots.
Figures 16 - 18 show the Class 5 and Class 9 GVW plots as well as monthly vehicle classification plots for all rural WIM stations in the NC Coastal Region. All other plots are shown in Appendix A of this report. Because the total number of plots that could be produced by combinations of class, weight, region, rural area, and urban area is large as discussed in (6) above, only selected plots for class 5 and class 9 were chosen for documentation. However, all the plots shown in Table 7 are available with the data on a CD which accompanies this report.

3.3 Low Weight Screening

The minimum GVW for all classes of trucks are shown in Table 6. It is to be noted that all weight records less than the weights shown in Table 6 have not been excluded from the GVW plot. However, the Minimum GVW Check, a utility in the NCDOT WIM QC Database is used only if a significant issue occurs with low weight GVWs for a class. This is done last after both weight and class check reviews are completed.

As an example of low weight screening consider class 5 trucks. In general, class 5 vehicles are two-axle six-tired “box” trucks. Sometimes pick-up trucks and campers are misclassified as class 5 trucks (especially on recreational routes) which lead to a significant number of low weight records in the WIM database. However, the WIM Enforcement Minimum GVW QC checks flag low weight records less than 10,000 lbs. GVW as shown in Table 6.

3.4 Results

In general, the class 5 and 9 GVW plots for all categories of WIM stations show expected trends. Plots for class 9 tend to have peaks at approximately 30,000 and 80,000 lbs corresponding to empty and fully loaded conditions respectively, which is reasonable. Class 5 plots tend to have a peak at about 15,000 lbs which is reasonable as well. However, the application of the minimum GVW rule for class vehicles results in the formation of a ski-slope rather than a peak and this is evident in all class 5 plots. The peak usually occurs at the very first data point as shown in all class 5 plots.
These results may be used by highway planners and pavement designers to quickly determine typical truck traffic profiles in the various NC regions. The results provide insight into NC truck transportation flows.

3.5 Recommendations for Future Research

Although the WIM data was good in general, it may be helpful to collect more than one year of data to compare trends or to substitute missing or incomplete data. It is also recommended to apply all of the QC procedures in the NCSU WIM QC database in the specified order before plotting the GVW plots.

In addition to applying quality control checks and an upgrade to Access 2003, it is also advisable to think ahead and look for alternate database solutions like Oracle or Microsoft SQL Server to store, analyze and process WIM data. Regular equipment calibration and maintenance will also go a long way in providing good quality data for analysis. In that regard, fine tuning of equipment to fix vehicle misclassifications, especially for recreational WIMs is highly desirable.
FIGURE 13 Google Earth® Rendering of WIM Stations in the NC Mountainous Region
FIGURE 14 Google Earth© Rendering of WIM Stations in the NC Piedmont Region
FIGURE 15 Google Earth® Rendering of WIM Stations in the NC Coastal Region
<table>
<thead>
<tr>
<th>Site ID</th>
<th>Nearest Town/City</th>
<th>Rt. Name</th>
<th>Type</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>501</td>
<td>South Mills</td>
<td>US 17</td>
<td>Rural and Recreational</td>
<td>Coastal</td>
</tr>
<tr>
<td>502</td>
<td>Elizabeth City</td>
<td>US 17</td>
<td>Rural</td>
<td>Coastal</td>
</tr>
<tr>
<td>503</td>
<td>Rocky Mount</td>
<td>I-95</td>
<td>I-95 has a unique pattern - recreational.</td>
<td>Central</td>
</tr>
<tr>
<td>504</td>
<td>Whiteville</td>
<td>US 74</td>
<td>Rural and Recreational</td>
<td>Coastal</td>
</tr>
<tr>
<td>506</td>
<td>Raleigh</td>
<td>I-40</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>507</td>
<td>Durham</td>
<td>NC 147</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>508</td>
<td>Pittsboro</td>
<td>US 64</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>509</td>
<td>Siler City</td>
<td>US 421</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>510</td>
<td>Reidsville</td>
<td>US 29</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>511</td>
<td>Greensboro</td>
<td>US 220</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>512</td>
<td>Winston Salem</td>
<td>US 311</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>515</td>
<td>Mount Airy</td>
<td>I-77</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>516</td>
<td>Charlotte</td>
<td>SR 1138</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>519</td>
<td>Oteen</td>
<td>I-40</td>
<td>Asheville Urban and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>520</td>
<td>Swannanoa</td>
<td>I-40</td>
<td>Asheville Urban and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>521</td>
<td>Cullowhee</td>
<td>NC 107</td>
<td>Rural and very Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>522</td>
<td>Whittier</td>
<td>US 74-441</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>523</td>
<td>Franklin</td>
<td>US 23-441</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>525</td>
<td>Siler City</td>
<td>US 421</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>527</td>
<td>Rocky Mount</td>
<td>I-95</td>
<td>I-95 has a unique pattern - recreational.</td>
<td>Central</td>
</tr>
<tr>
<td>529</td>
<td>Greenville</td>
<td>US 264</td>
<td>Urban Loop/bypass</td>
<td>Coastal</td>
</tr>
<tr>
<td>530</td>
<td>Sanford</td>
<td>US 1</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>531</td>
<td>Lexington</td>
<td>US 52</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>532</td>
<td>Jacksonville</td>
<td>SR 1245</td>
<td>Rural</td>
<td>Coastal</td>
</tr>
<tr>
<td>533</td>
<td>Hayesville</td>
<td>US 64</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>534</td>
<td>Murphy</td>
<td>US 64</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>535</td>
<td>Andrews</td>
<td>US 74</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>536</td>
<td>Clyde</td>
<td>I-40</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>537</td>
<td>Columbus</td>
<td>I-26</td>
<td>Rural And Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>538</td>
<td>Kings Mountain</td>
<td>I-85</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>539</td>
<td>Charlotte</td>
<td>I-77</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>540</td>
<td>South Mills</td>
<td>US 17</td>
<td>Rural and Recreational</td>
<td>Coastal</td>
</tr>
<tr>
<td>541</td>
<td>McDonald</td>
<td>I-95</td>
<td>I-95 has a unique pattern - recreational</td>
<td>Coastal</td>
</tr>
</tbody>
</table>
### TABLE 5 (Continued)

<table>
<thead>
<tr>
<th>Site ID</th>
<th>Nearest Town/City</th>
<th>Rt. Name</th>
<th>Type</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>542</td>
<td>Rocky Point</td>
<td>I-40</td>
<td>Rural and Recreational</td>
<td>Coastal</td>
</tr>
<tr>
<td>543</td>
<td>Wise</td>
<td>I-85</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>545</td>
<td>Durham</td>
<td>NC 147</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>546</td>
<td>Charlotte</td>
<td>NC 24</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>547</td>
<td>Dallas</td>
<td>US 321</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>548</td>
<td>Scranton</td>
<td>US 264</td>
<td>Rural</td>
<td>Coastal</td>
</tr>
<tr>
<td>549</td>
<td>Currie</td>
<td>US 421</td>
<td>Rural</td>
<td>Coastal</td>
</tr>
<tr>
<td>551</td>
<td>Laurinburg</td>
<td>US 74</td>
<td>Rural</td>
<td>Coastal</td>
</tr>
<tr>
<td>552</td>
<td>Lilesville</td>
<td>US 74</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>553</td>
<td>Asheboro</td>
<td>US 220</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>554</td>
<td>Madison</td>
<td>US 220</td>
<td>Rural</td>
<td>Central</td>
</tr>
<tr>
<td>555</td>
<td>Greensboro</td>
<td>NC 68</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>556</td>
<td>Asheville</td>
<td>I-240</td>
<td>Asheville Urban and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>557</td>
<td>Statesville</td>
<td>I-40</td>
<td>Urban and some Recreational</td>
<td>Central</td>
</tr>
<tr>
<td>558</td>
<td>Hickory</td>
<td>US 321</td>
<td>Urban</td>
<td>Central</td>
</tr>
<tr>
<td>559</td>
<td>Valdese</td>
<td>I-40</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
<tr>
<td>560</td>
<td>Mars Hill</td>
<td>I-26</td>
<td>Rural and Recreational</td>
<td>Mountain</td>
</tr>
</tbody>
</table>

### TABLE 6 Minimum GVW for Trucks

<table>
<thead>
<tr>
<th>FHWA Vehicle Class</th>
<th>Minimum GVW (lbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>12,000</td>
</tr>
<tr>
<td>5</td>
<td>10,000</td>
</tr>
<tr>
<td>6 and 7</td>
<td>12,000</td>
</tr>
<tr>
<td>8</td>
<td>16,000</td>
</tr>
<tr>
<td>9 - 12</td>
<td>22,000</td>
</tr>
</tbody>
</table>

### TABLE 7 Plots Generated Using the NCDOT QC Database

<table>
<thead>
<tr>
<th>Name</th>
<th>Obtained from</th>
<th>No. of Plots Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight Average Hourly Plots</td>
<td>W-Card Data</td>
<td>12</td>
</tr>
<tr>
<td>Weight GVW Plots</td>
<td>W-Card Data</td>
<td>10</td>
</tr>
<tr>
<td>Class Average Hourly Plots</td>
<td>C-Card Data</td>
<td>12</td>
</tr>
<tr>
<td>Average Daily Class Plots</td>
<td>C-Card Data</td>
<td>12</td>
</tr>
<tr>
<td>Summary Class Plots</td>
<td>C-Card Data</td>
<td>3</td>
</tr>
</tbody>
</table>
FIGURE 16  Summary Class Plot for all Rural WIMs in the NC Coastal Region
FIGURE 17 Class 5 GVW Plot for all Rural WIMs in the NC Coastal Region

FIGURE 18 Class 9 GVW Plot for all WIMs in the NC Coastal Region
4.0 NC vs. UNIVERSITY OF ARKANSAS WIM QC ANALYSIS

Most highway agencies have the data collection and design groups in different units. While a single software solution is not practical, it is recommended to perform two separate processes where the output of data QC meets the needs and standards of the design process. It is highly recommended that traffic engineers most familiar with the traffic patterns and the land uses that drive them should perform the raw QC, not an automated “black box” process. The loss of black box efficiencies is a necessary one. However, tools to facilitate the QC process would introduce other efficiencies. Additionally most of NCDOT pavement designs will need regional inputs. Developing regional inputs is too complex to setup an automated process. It requires some judgment. If automated, it is more likely to be misused. This process is made more complicated by decisions affecting design policy and resource support [Kent Taylor (email), March 16, 2009].

This chapter documents and compares the analysis methodologies practiced by NCSU/NCDOT and by the University of Arkansas for performing quality control analysis on vehicle weight and class data. A comparative analysis between the QC methods followed by the University of Arkansas (UARK) and NCSU/NCDOT shows that while the UARK Pavement Designer software has better mapping functions and supports data analysis and design. However, from a WIM data analyst’s perspective it is a “black box”. In addition there is significant data reduction involved and the rigid nature of an automated QC process does not provide enough justification for the data to be used as input for the MEPDG. On the other hand, the NCSU/DOT approach is a two step procedure with a comprehensive QC procedure that provides the flexibility of manual overriding based on local knowledge of WIM stations and a separate unit managing the pavement design element.

4.1 Scope and Objectives

The geographical scope is limited to the state of NC and the data related scope includes 12 consecutive months of best quality vehicle weight and class data. FHWA vehicle classes 4 - 13 are the only classes being considered for vehicle weight data. Vehicle class data has been used in its entirety. The scope of the analysis includes GVW analysis for vehicle weight data and vehicle class distribution by month and year for vehicle class data.
The objective of this research task is to justify the need for a performing a semi-automated user override featured QC procedure on vehicle weight and class data used for MEPDG. This is accomplished by comparing the analysis methodologies practiced by NCSU/NCDOT and the University of Arkansas for performing quality control analysis on vehicle weight and class data.

### 4.2 Vehicle Class Data Analysis Results

Figures 19 and 20 show the vehicle class distribution from the University of Arkansas Pavement Designer software and the NCSU/NCDOT QC process respectively. There appears to be a significant variation in the average vehicle class distribution values because of procedure possible error in program source code to display average vehicle class distribution. It is to be noted that the University of Arkansas have also excluded three months of class data from the twelve month dataset in the Arkansas quality control procedure.

For documentation purposes, all plots are shown in Appendix B of this report. All these plots are available with the data on a CD which accompanies the final report for NCDOT research project HWY 2008-11.

**FIGURE 19 University of Arkansas - Monthly Truck Class Distribution Plot**
FIGURE 20  NCSU/NCDOT Monthly Truck Class Distribution Plot

FIGURE 21  Yearly Truck Class Distribution Comparison Plot – Possible Error in Code
4.3 The NC and Arkansas Methodologies

Figure 22 is a flow diagram depicting the tools and procedures used by the University of Arkansas (UARK) for performing analysis on vehicle weight and class datasets. Figure 23 shows the NCSU/NCDOT flow diagram. Table 8 shows the pros and cons of each approach. The analysis shows that while the UARK Pavement Designer software has better mapping functions and supports data analysis and design, from a WIM data analyst’s perspective, it is a “black box program”. Additionally, even though there is data reduction involved in the Arkansas process, the impact it has on ME pavement design output has not been tested.

FIGURE 22 University of Arkansas WIM Data Analysis Flow Diagram
FIGURE 23  NCSU/NCDOT WIM Data Analysis Flow Diagram
TABLE 8 NC vs. Arkansas WIM Data Analysis Comparative Study

<table>
<thead>
<tr>
<th>Entity/Feature</th>
<th>NCDOT QC Database</th>
<th>University of Arkansas Pavement Designer Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC Runtime</td>
<td>60 minutes per WIM station.</td>
<td>N/A</td>
</tr>
<tr>
<td>QC Automation</td>
<td>Some manual input required. QC procedure is comprehensive and almost completely automated with manual override option provided whenever required.</td>
<td>Completely Automated. However there is no manual override option for each and every QC procedure.</td>
</tr>
<tr>
<td>Mapping/GIS</td>
<td>N/A</td>
<td>Built-in</td>
</tr>
<tr>
<td>Data Import</td>
<td>Data import is done separately using the WIM Processor database.</td>
<td>Built-in</td>
</tr>
<tr>
<td>Plotting features</td>
<td>A total of 49 plots can be generated for each WIM station. Plots include weight average hourly plots, weight GVW plots for all trucks, class average hourly plots, average daily class plots, and summary class plots.</td>
<td>Generates summary class plots from vehicle class data. It also generates one GVW plot for all trucks and axle load plots from vehicle weight data.</td>
</tr>
<tr>
<td>Retrieval of Pavement Material Parameters</td>
<td>N/A</td>
<td>Built-in</td>
</tr>
<tr>
<td>Climate and Traffic Data Interpolation</td>
<td>N/A</td>
<td>Built-in</td>
</tr>
<tr>
<td>Data Reduction</td>
<td>The entire dataset provided is used. This is generally 12 consecutive months of vehicle weight or class data. Therefore this requires larger storage space.</td>
<td>Does data reduction by eliminating one or more months of vehicle weight or class data entirely from the original dataset. Thus this requires lesser storage space. However the data reduction is not completely justified.</td>
</tr>
<tr>
<td>Axle Load Distribution/Processing</td>
<td>ALF Processing and plotting are done separately.</td>
<td>Built-in</td>
</tr>
</tbody>
</table>

4.4 Truck Weight Data Analysis Results

In addition to vehicle class distribution analysis, GVW plots were also compared. A single GVW plot for all vehicle classes was generated based on the format used by University of Arkansas for comparison purposes. Figures 24 and 25 show a significant change in vehicle classification based on the specified weight bins. The GVW plot generated by the University of Arkansas Pavement Designer software shows no trucks in the 10,000 lb range whereas the NCSU plot shows that
about 9% of all trucks are in the same weight range. The latter is true since it is expected to have a significant number of FHWA class 5 trucks in the range of 10,000 lbs. However the impact of these 10,000 trucks on ME-pavement design output has not been proven or justified yet. It was also confirmed by NCDOT’s Traffic Survey Unit that a Class 5 GVW plot will peak at the 10,000 lb range. Figures 24 and 25 show the GVW plots from the University of Arkansas and NCSU respectively. For documentation purposes, all the other plots are shown in Appendix B of this report. All these plots are available with the data on a CD which accompanies this report.
FIGURE 24  University of Arkansas – Gross Vehicle Weight Plot

FIGURE 25  NCSU/NCDOT Gross Vehicle Weight Plot
4.5 Recommendations for Future Research

As mentioned in the introductory section of this chapter, developing a single software product for performing data analysis and design does not incorporate important tasks that require technical oversight. Instead, it is recommended to perform two separate processes where the output of data QC meets the needs and standards of the design process. The NCDOT QC Database offers automation and manual overriding features for regional inputs, both being critically important for performing vehicle class and weight data quality control analysis procedures.
5.0 WIM DATA MANAGEMENT AND ANALYSIS IN SQL SERVER

With the highway infrastructure under strain, there is a need to collect and analyze traffic volume, vehicle classification, and weight data in an integrated manner. While most research topics focus on collection and quality of data with little emphasis on the development of an integrated Database Management System (DBMS) to store and analyze traffic data.

Since Microsoft Access 2003 has database size limitations, a manual approach was used to merge data from two or more WIM Stations to generate data needed to develop summary statistics and plots like GVW, truck class distribution plots etc. Hence it is highly desirable to think ahead and look for alternate methods to develop, analyze, and maintain a statewide WIM program. This chapter explains the application of Analysis (OLAP) Cubes in SQL Server for two test NCDOT WIM station data, a new approach to obtain statistics relevant to WIM data. It is a convenient method of disseminating data to users that do not have online access to the WIM database in SQL Server. The analysis cube files may also be used for data mining and exploration and could be used to observe trends in axle weights, axle spacings, vehicle class volumes etc.

It is also apparent that a more advanced and comprehensive database management system like Oracle or SQL Server has to be implemented in the long run. This is because data stored in a DBMS is less susceptible to corruption, offers unlimited data storage, and is much more secure.

5.1 Scope and Objectives

The scope of this chapter is similar to the other chapters with respect to the type of data used. However, for the sake of testing the WIM data application using the analysis cube concept in SQL Server, truck weight and vehicle class databases for two WIM stations (SHRP ID 371028 and 375601) located in Camden County and Madison County respectively were used for the analysis.
The objectives of this chapter are to explain the application of WIM data analysis using Analysis Cubes in Microsoft SQL Server and to emphasize the importance of transitioning to a more advanced database management system for the NCDOT statewide WIM program.

5.2 Database Management in SQL Server 2008

In this section, the advantages of transitioning from Microsoft Access to SQL Server 2008, a well established DBMS that allows for editing, storage, and reporting of traffic volumes, vehicle classification, and vehicle weight data is presented. The system is user friendly and designed to operate under a Windows XP environment. Refinements needed in the future are also identified. Figures 27 shows the SQL Server Interface view for class records and Figure 28 shows the SQL Server interface with SQL query view respectively. In addition, the data from two sample WIM databases were also imported into the SQL Server database from their respective Access databases using the SQL Server Migration Assistant 2008 for Access© tool. Data can also be imported manually using the Data Import Wizard in SQL Server 2008. The import process was carried out in six steps shown below in Figure 26.

FIGURE 26. SQL Server Data Import Flow Diagram
FIGURE 27  SQL Server 2008 Database Layout with Selected Vehicle Class Records

FIGURE 28  SQL Server 2008 Interface with SQL Query View

50
5.3 Alternate Approaches to WIM Data Analysis and Quality Control

At hundreds of weight-in-motion (WIM) sensor locations State Departments of Transportation collect gigabytes of vehicle data every year to meet Federal reporting traffic requirements. The DOTs process the data to support pavement design to enforce weight restrictions on highways and bridges, and to provide planning data for highway improvements. For instance, the 45 WIM stations used in the NCDOT sponsored MEPDG project requires approximately 100 gigabytes of data storage. The enormous WIM datasets include date and time stamps for every vehicle crossing the WIM sensors, counts of each vehicle’s axles, and measurements of the individual axle weights and spacings. The results provide statistics on the annual average daily traffic for each of the 13 FHWA vehicle classes, time histories of GVWs by class, frequency distributions of the traffic by vehicle class, and other important characteristics of the traffic flow at each WIM station.

After the WIM data are collected in the field and reduced from proprietary sensor machine code to readable traffic data, the quality of the data must be checked. During the QC procedures, checks identify incomplete datasets, out of range values for particular vehicle classes, and other possible data problems. The QC checks include automatic checks that flag and report questionable data and manual checks that rely on graphical displays of the data and visual interpretation.

Although, the QC Procedures developed by NC State University for NCDOT are comprehensive, it is highly imperative that other approaches are looked into in order to fix errors and make improvements to the existing procedure. The analysis cube method is one such approach to perform quality control procedures on any database system.

5.3.1 OLAP Cubes – An Introduction

An OLAP cube is a specially designed database that is optimized for reporting. While most databases designed for online transaction processing such as those used in claims processing are designed for efficiency in data storage, OLAP cubes are designed for efficiency in data retrieval. This means that the data is stored in such a way as to make it easy and efficient for reporting.
Regular relational databases treat all data into the database similarly. However OLAP cubes categorize data into dimensions and measures. Measures represent items that are counted, summarized or aggregated whereas dimensions are variables by which measures are summarized, such as hospitals, physicians, or dates of service. This organization of data greatly facilitates the ability to formulate data requests based on real-life situations. In addition, many of queries that could be posed to the data are pre-aggregated in the database so that the answers have already been pre-calculated and can be reported without delay [Designing OLAP Cubes, 2009].

5.3.2 Importance of OLAP Cubes

Before OLAP technology was well developed, data had to be extracted from databases using queries. This meant that the analyst had to structure a request to the database for the information desired, and then submitted this query to the database server. That server would then process the query and return the results. Depending on the size of the database and the data requested, this query could take minutes or hours to complete. In this sense, the online aspect of this type of reporting is questionable. OLAP cubes are fundamentally different in that they pre-aggregate the data used to answer many of queries that are anticipated. This pre-aggregation occurs when the cube is built, which means that this process is already completed when the user queries the data. In addition, the size of an OLAP cube depends on the number of measures and dimensions it contains. It has no relationship to the size of the initial data set. Therefore, a WIM dataset having millions of records can be consolidated into a relatively small OLAP cube that can return data almost instantaneously [Designing OLAP Cubes, 2009].

Another unique feature that is part of the OLAP database structure is that it allows the analyst to drill down into the dataset. When designing an OLAP database, dimensions are structured into hierarchies. For example, WIM data dates can be arranged in a hierarchy of days, weeks, months, quarters, and years. Similarly, diagnoses can be arranged by major and minor categories, drilling down to the individual diagnosis code. The OLAP cube knows the hierarchy, so if the analyst issues the command to drill down, the cube knows the next level of data to be presented.
5.3.3 Application of OLAP Cubes on WIM Data

As mentioned before, a cube is a multi-dimensional structure that stores aggregated data and allows the data to be viewed at various detail levels [Designing OLAP Cubes, 2009]. Figure 29 is provided to illustrate the analysis cube concept for analyzing WIM data. The location axis on the main cube contains each WIM site number, allowing the user to view aggregated information for selected sites. Each location on this axis is made up of a smaller cube that contains information for each individual lane at each WIM site, allowing the user to view aggregated information for each lane. The Time axis on the main cube contains each year of WIM data, allowing the user to view aggregated data for the entire year. The time axis can have many smaller cubes, depending on the level of detail. The user could view the axle spacing, axle weight, or any other measure defined in the WIM dataset. View 1 illustrated in this cube could correspond to viewing the axle weight for all selected WIM stations in December 2002. View 2 could correspond to viewing the axle spacing for station 371028 for each month between January 2002 and December 2002. View 3 could correspond to viewing the axle spacing and axle weight for Site 502 in January 2002. There are an endless number of views that could be used to view the data.

The name “cube” suggests that these structures are three-dimensional. This is not always true. Cubes can have many more dimensions than three, but the term continues to apply. For example, a fourth dimension based on vehicle class could be added to Figure 29 that allowed the user to select certain vehicle classes. If steer axle weight was used as a dimension, the user could view the average drive tandem axle spacing for vehicles that had a steer axle weight within a selected range. The cube files are very powerful because they allow summary reporting using a simple structure and they allow very detailed data mining using complex structures [Nichols and Bullock, 2004].

Using the data from the two sample WIM databases imported into SQL server, an analysis cube was created using Microsoft’s SQL Server Business Intelligence Development Studio©. Figure 30 shows the analysis cube data source view for the two sample WIM stations. Microsoft Excel can connect to the Analysis Server database and it is possible to view the cube files as pivot tables and pivot charts. It is also possible to create offline cube files for distribution that can also be viewed with Microsoft Excel [SQL Server Analysis, 2009]. This is a convenient method of
disseminating data to users that do not have online access to the WIM database in SQL Server. After this, further analysis can be carried out on the data. For instance, the analysis cube files may be used for data mining and exploration and could be used to observe trends in axle weights, axle spacings, vehicle class volumes etc.

FIGURE 29 Analysis Cube Concept
FIGURE 30 Analysis Cube Data Source View in SQL Server 2008
5.4 Conclusions and Recommendations for Future Research

From this analysis, it can be concluded that an integrated relational database management system is the long term solution to WIM data management and OLAP cube reporting is the best way to report WIM data.

Although Microsoft Access is sufficient to handle most individual WIM databases, a more advanced and comprehensive database management system like Oracle or SQL Server has to be implemented in the long run. This is because data stored in a DBMS is less susceptible to corruption, offers unlimited data storage, and is much more secure.

It is recommended to use an open source database management software like MySQL for storing and analyzing WIM datasets. MySQL offers similar functionalities and features as Oracle and SQL Server and its source code is available free of cost. Alternatively, it is also suggested to apply some sampling schemes on vehicle weight and class data. Since weight data can result in extremely large databases, especially for sites with high vehicular volumes like interstate routes, statistically valid sampling methods should be explored to reduce the size of such datasets to a manageable extent. Sampling requirements must be based on the intended use of the data. For example, most LTPP GPS studies require only seven days of weight data per quarter.
6.0 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

In 2003, the National Cooperative Highway Research Program released the *Mechanistic-Empirical Design Guide (MEPDG) for New and Rehabilitated Pavement Structures*. According to NCHRP Report 538, the 1-37A research provides engineers with practical and realistic pavement design procedures and software that use existing mechanistic-empirical principles [Hallenbeck et.al, 2005]. The mechanistic-based distress prediction models used in the MEPDG will require the input of specific data for each axle type and axle-load group. Project 1-39 research develops procedures and software for collecting and processing traffic data required by the Pavement Design Guide procedures.”

This research focuses mainly on traffic data quality control, analysis and management. It will be complex enough for the designers in deciding which data to use for a design. Therefore, it is highly recommended that those most familiar with the traffic patterns and the land uses that drive them should perform the raw QC. This processed data will be finally used as input parameters by pavement designers for the pavement design process.

The common geographical scope of this project is limited to the state of North Carolina. All WIM data collected by NCDOT for the MEPDG project was used for this research. The data related scope includes 12 consecutive months of best quality vehicle weight and class data from 47 selected WIM stations. However, in some cases, due to data quality issues, entire months or days of vehicle weight or class data had to be replaced with data from another year. FHWA vehicle classes 4 – 13 are the only classes being considered for vehicle weight data. However, vehicle class data is used in its entirety. In addition to all these, each major chapter may have an additional scope involved. Each chapter in this research has its own objectives stated separately.

This research report consists of four major “stand-alone” topics each of which is develop in one chapter. Even though the topics and chapters are typically stand-alone, the data used for all chapters are the same.
6.1 Chapter 2 - WIM Data Quality Control

The NCDOT requires that all data with quality issues (such as partial data, misclassified vehicles, and unacceptable axle weights and spacing) be excluded from datasets used for planning and design statistics.

6.1.1 Summary

All vehicle weight and class datasets used have to undergo a series comprehensive quality control procedure checks before they can be used for any further analysis. The second chapter of this research delineates the data requirements for the MEPDG software and the quality control procedures performed on vehicle weight and class data. It also includes overall QC results and a discussion of QC results on two strategically selected WIM stations.

6.1.2 Conclusions

The observations from Chapter 2 are that apart from a few sites with up to an entire month of data missing because of construction and other non-equipment problems, the WIM systems operated by the NCDOT provided valid weight and class data. When a month of data was excluded by the QC process, the NCDOT substituted equivalent data from the same month but from another year in order to develop complete datasets for MEPDG.

6.1.3 Recommendations

As a result of these observations it is recommended that NCDOT carries out the following procedures on WIM datasets

1. Explore data sampling procedures to reduce the size of databases
2. Assign severity level to quality control rules
3. Explore long term WIM data storage techniques
6.2 Chapter 3 - NC Urban and Rural Truck Traffic Profiles

Knowing the type of traffic by vehicle class by highway functional classification is critical to designing, maintaining and paying for North Carolina highway pavements. Such information is used by the NCDOT Pavement Management Unit to estimate pavement costs by NC region and highway functional classification.

6.2.1 Summary

The third chapter documents a series of vehicle weight and class plots. Important plots include GVW plots as well as monthly average vehicle classification plots depicting seasonal variation of truck traffic by vehicle class, highway functional class, and NC Region.

6.2.2 Conclusions

In general, the class 5 and 9 GVW plots for all categories of WIM stations show expected trends. Plots for class 9 tend to have peaks at approximately 30,000 and 80,000 lbs corresponding to empty and fully loaded conditions respectively, which is reasonable. Class 5 plots tend to have a peak at about 15,000 lbs which is reasonable as well. However, the application of the minimum GVW rule for class vehicles results in the formation of a ski-slope rather than a peak and this is evident in all class 5 plots. The peak usually occurs at the very first data point as shown in all class 5 plots.

6.2.3 Recommendations

Although the WIM data was good in general, it may be helpful to collect more than one year of data to compare trends or to substitute missing or incomplete data. It is also recommended to apply all of the QC procedures in the NCSU WIM QC database in the specified order before plotting the GVW plots.

6.3 Chapter 4 - NC vs. University of Arkansas WIM QC Analysis

Most highway agencies have the data collection and design groups in different units. Developing regional inputs is too complex to setup an automated process. It requires some judgment. If automated, it is more likely to be misused. This process is made more complicated by decisions affecting design policy and resource support.
6.3.1 Summary
In this chapter, the analysis methodologies practiced by NCSU and the University of Arkansas for performing quality control analysis on vehicle weight and class data are documented. The main purpose of performing this comparative analysis is to emphasize the importance of performing semi-automated user override featured quality control checks on data used for the MEPDG.

6.3.2 Conclusions
Developing a single software product for performing data analysis and design does not incorporate important tasks that require technical oversight.

6.3.3 Recommendations
It is recommended to perform two separate processes where the output of data QC meets the needs and standards of the design process. Most of NCDOT pavement designs will need regional inputs. Developing regional inputs is too complex to setup an automated process. It requires some judgment. If automated, it is more likely to be misused. This process is made more complicated by decisions affecting design policy and resource support. The NCDOT QC Database offers automation and manual overriding features for regional inputs, both being critically important for performing vehicle class and weight data quality control analysis procedures.

6.4 Chapter 5 - WIM Data Management and Analysis in SQL Server
While most research topics focus on collection and quality of data, there is little emphasis on the development of an integrated Database Management System (DBMS) to store and analyze traffic data.

6.4.1 Summary
The fifth chapter conveys the need to collect and analyze traffic volume, vehicle classification, and weight data in an integrated manner using Microsoft SQL Server 2008. In addition, the application of analysis cubes in SQL Server, an alternate approach to data analysis and quality control is also discussed.
6.4.2 Conclusions

From this analysis, it can be concluded that an integrated relational database management system is the long term solution to WIM data management and OLAP cube reporting is the best way to report WIM data.

6.4.3 Recommendations

Although Microsoft Access is sufficient to handle most individual WIM databases, a more advanced and comprehensive database management system like Oracle or SQL Server has to be implemented in the long run. This is because data stored in a DBMS is less susceptible to corruption, offers unlimited data storage, and is much more secure.

It is recommended to use an open source database management software like MySQL for storing and analyzing WIM datasets. MySQL offers similar functionalities and features as Oracle and SQL Server and its source code is available free of cost. Alternatively, it is also suggested to apply some sampling schemes on vehicle weight and class data. Since weight data can result in extremely large databases, especially for sites with high vehicular volumes like interstate routes, statistically valid sampling methods should be explored to reduce the size of such datasets to a manageable extent. Sampling requirements must be based on the intended use of the data. For example, most LTPP GPS studies require only seven days of weight data per quarter.
REFERENCES


Appendix A – NC URBAN RURAL PLOTS
FIGURE 31 Rural Summary Class Plot for the NC Coastal Region
FIGURE 32  Rural Class 5 GVW Plot for the NC Coastal Region

FIGURE 33  Rural Class 9 GVW Plot for the NC Coastal Region
FIGURE 34 Rural Summary Class Plot for the NC Central Region
FIGURE 35 Rural Class 5 GVW Plot for the NC Central Region

FIGURE 36 Rural Class 9 GVW Plot for the NC Central Region
FIGURE 37 Rural Recreational Summary Class Plot for the NC Mountainous Region
FIGURE 38  Rural Recreational Class 5 GVW Plot for the NC Mountainous Region

FIGURE 39  Rural Recreational Class 9 GVW Plot for the NC Mountainous Region
FIGURE 40 Rural Recreational Summary Class Plot for the NC Coastal Region
FIGURE 41  Rural Recreational Class 5 GVW Plot for the NC Coastal Region

FIGURE 42  Rural Recreational Class 9 GVW Plot for the NC Coastal Region
FIGURE 43 Summary Class Plot for all Urban WIMs
FIGURE 44 Class 5 GVW Plot for All Urban WIMs

FIGURE 45 Class 9 GVW Plot for All Urban WIMs
FIGURE 46  WIM 557 - Urban and Some Recreational Summary Class Plot
FIGURE 47  WIM 557 - Urban and Some Recreational Class 5 GVW Plots

FIGURE 48  WIM 557 - Urban and Some Recreational Class 9 GVW Plots
FIGURE 49 Asheville Urban and Recreational Summary Class Plot
FIGURE 50 Asheville Urban and Recreational Class 5 GVW Plot

FIGURE 51 Asheville Urban and Recreational Class 9 GVW Plot
FIGURE 52 Summary Class Plot for all I-95 WIMs
FIGURE 53  Class 5 GVW Plot for all I-95 WIMs

FIGURE 54  Class 9 GVW Plot for All I-95 WIMs
Appendix B – NC vs. ARKANSAS PROFILES
FIGURE 55  NC Average Daily Truck Class Distribution by Month for WIM 378201

FIGURE 56  Arkansas Truck Class Distribution for WIM 378201

Note - This WIM Station (SHRP ID 378201) is located on US 74, 0.1 miles east of NC 79 in Scotland County.
FIGURE 57 NC Average Daily Truck Class Distribution by Month for WIM 375903

FIGURE 58 Arkansas Truck Class Distribution for WIM 375903

Note – This WIM Station (SHRP ID 375903) is located on SR 2665, 0.2 miles east of US 21 in Mecklenburg County.
FIGURE 59 NC Average Daily Truck Class Distribution by Month for WIM 374301

FIGURE 60 Arkansas Truck Class Distribution for WIM 374301

Note – This WIM Station (SHRP ID 374301) is located on I-40, west of US 19-23-74 connector in Haywood County
FIGURE 61 NC Average Daily Truck Class Distribution by Month for WIM 374002

FIGURE 62 Arkansas Truck Class Distribution for WIM 374002

Note – This WIM Station (SHRP ID 374002) is located on NC 68, 0.5 miles north Of Bryan Blvd in Guilford County.
FIGURE 63 NC Average Daily Truck Class Distribution by Month for WIM 371801

FIGURE 64 Arkansas Truck Class Distribution for WIM 371801

Note – This WIM Station (SHRP ID 371801) is located on I-40, 1.3 miles west of SR 2740 (MP 57) in Buncombe County.
FIGURE 65  NC Average Daily Truck Class Distribution by Month for WIM 371030

FIGURE 66  Arkansas Truck Class Distribution for WIM 378201

Note – This WIM Station (SHRP ID 371030) is located on US 17, 0.4 miles south of US 158 in Pasquotank County.
FIGURE 67 NC Average Daily Truck Class Distribution by Month for WIM 371024

FIGURE 68 Arkansas Truck Class Distribution for WIM 371024

Note – This WIM Station (SHRP ID 371024) is located on NC 107, north of SR 1001 in Jackson County.
FIGURE 69 NC Average Daily Truck Class Distribution by Month for WIM 377501

FIGURE 70 Arkansas Truck Class Distribution for WIM 377501

Note - This WIM Station (SHRP ID 377501) is located on US 220, 0.1 miles north of SR 1247 in Randolph County.
Note - This WIM Station (SHRP ID 371101) is located on I-40, 0.3 miles west of SR 1744 (MP 109) in Burke County.
FIGURE 73  NC Average Daily Truck Class Distribution by Month for WIM 373102

FIGURE 74  Arkansas Truck Class Distribution for WIM 373102

Note - This WIM Station (SHRP ID 373102) is located on NC 147, 0.4 miles north of SR 1940 in Durham County.
FIGURE 75 NC Average Daily Truck Class Distribution by Month for WIM 371003

FIGURE 76 Arkansas Truck Class Distribution for WIM 371003

Note – This WIM Station (SHRP ID 371003) is located on I-240, 0.5 miles east of US 70 (MP 8) in Buncombe County.