1-1-2007

Investigation on the AP-42 sampling method

Yanjie Chen
University of Nevada, Las Vegas

INVESTIGATION ON THE AP-42 SAMPLING METHOD

By

Yanjie Chen

Bachelor of Science
Changsha Railway University, Changsha, China
1999

Master of Science
Northern Jiaotong University, Beijing, China
2002

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science Degree in Engineering
Department of Civil and Environmental Engineering
Howard R. Hughes College of Engineering

Graduate College
University of Nevada, Las Vegas
August 2007

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Thesis Approval

The Graduate College
University of Nevada, Las Vegas

24 JULY , 2007

The Thesis prepared by

YANJIE CHEN

Entitled

INVESTIGATION ON THE AP-42 SAMPLING METHOD

is approved in partial fulfillment of the requirements for the degree of

MASTERS OF SCIENCE IN ENGINEERING DEGREE

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative

1017-53

ABSTRACT

Investigation on the AP-42 Sampling Method

by

Yanjie Chen

Dr. Hualiang (Harry) Teng, Examination Committee Chair
Graduate College
University of Nevada, Las Vegas

The Las Vegas area has been designated by the U.S. EPA as a serious PM_{10} non-attainment area. To monitor PM_{10} in this area, dust data have been collected quarterly using the AP-42 method. According to this method, the number of composite sample sizes (the number of sample sites) needs to be determined first. In the actual dust data collection at each of these sample sites, a procedure with the specifications of the number and locations of incremental samples (plots) and their sizes (i.e., length) has to be followed. Apparently, there has been no rule existing that can be used to determine the composite sample size. In addition, it is unknown whether the required number of plots and their sizes are validated based on real data.

Due to the availability of dust emission data collected using mobile sampling technologies, which are viewed as being close to actual continuous dust emission data over a roadway segment, this study investigates the optimal number of sample sites and number of plots and their sizes that can be used for the AP-42 method. To determine the number of sample sites, the optimal allocation sampling method is adopted. By using this method, the variance of emission estimated based on samples can be minimized for a

fixed budget. The issue with validating the number of plots and their sizes for the AP-42 method is investigated by using the Monte Carlo simulation method. In the simulation, the layouts of plots are emulated following the AP-42 method. The difference between the estimated emission factor based on the emulated AP-42 method and the true emission factor are compared. Patterns for the difference between the estimated and true emission factors versus the number and size of plots are observed. These observed patterns are used to derive the thresholds of the number and size of plots for the AP-42 method.

The results from the optimal allocation method indicate that most sample sites should be allocated to the local roads because the variance of emission and proportion of roadway segments of this roadway classification are significantly higher than most of other roadway classifications. This conclusion may lead to the development of more cost effective sampling approaches. The results from the Monte Carlo simulation method imply that clear patterns of improved estimation of emission factors versus the number and size of plots can be observed only for three roadway classifications, not for other classifications. This result may indicate that the AP-42 method may not be applicable to some roadway classifications, and thus a different data collection method, such as the mobile sampling technologies, may be necessary.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEGEMENTS

I want to thank Dr. Hualiang (Harry) Teng for his instruction on my thesis. He provided the idea of the stratified method on sample size and monte carlo simulation method from related reference. With the data he provided, I can perform the statistical analyses. I appreciate his time and efforts on writing and rewriting significant part of the thesis.

I want to thank Dr. David James for providing related materials and thought-provoking suggestions for my study. I appreciate his instructive comments about the application of Monte Carlo simulation method even though I cannot apply them in my study. I also appreciate his willingness to support me all the time.

I would like to thank Dr. Hokwon Cho and Dr. Mohammed Kaseko for being my committee members and for reviewing of my thesis manuscript.

I would like to thank my dear parents and sister for their many years of consistent spiritual support and encouragement. I also want to give my thanks to my husband, Gang Xie for not only giving me strong support on my thesis when needed, but also giving me the strength and courage to get over all the difficulties. The preliminary work on Monte Carlo simulation done by Gang helped me finish the Monte Carlo simulation programming in time.

I would like to thank Dr. Nader Ghafoori and Mr. Jeffery Jensen for giving me the opportunity to finish my Master degree at UNLV.

I would like to thank Dr. Nader Ghafoori and Mr. Jeffery Jensen for giving me the opportunity to finish my Master degree at UNLV.

CHAPTER 1

INTRODUCTION

Dust emission is a major concern in western U.S. To improve air quality, National Ambient Air Quality Standards (NAAQS) (U.S. EPA, 2006) for particulate matter with aerodynamic diameters equal to or less than 10 micrometers (PM_{10}) were promulgated by the United State Environmental Protection Agency. Figure 1.1 shows the size of PM_{10} compared with human hair (Clark County, 2007). Because PM_{10} can be inhaled through the nose and mouth, and accumulated in the respiratory system such as throat and lung, these particles pose a significant health concern (E.H. Pechan & Associates, Inc., 2004).

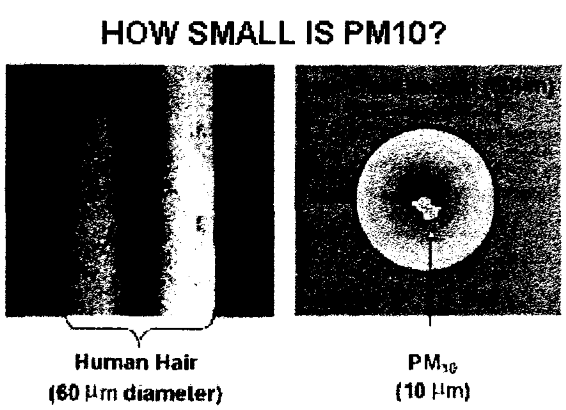


Figure 1.1 PM_{10} Size Compared with Hair (Clark County, 2007)

Public streets and highways are major sources of the atmospheric particulate matter within an urban area. Once a vehicle travels over a roadway surface (either paved or not paved), particulate emissions occur in the form of emissions from road, brake, tire wear, and resuspension of loose material on the road surface.

Emission factors are measures of dust emission used for air quality management. According to AP-42 (U.S. EPA, 2005), "an emission factor is a representative value that attempts to relate the quantity of a pollutant released to the atmosphere with an activity associated with the release of that pollutant". Every source-specific activity involved in the release of pollutant has its own emission factors, and these factors are usually expressed as "the weight of pollutant divided by a unit weight, volume, distance, or duration of the activity emitting the pollutant (e.g., kilograms of particulate emitted per megagram of coal burned)" (U.S. EPA, 1997). PM_{10} emission factors are the measures of dust emission for particulate matter with size equal to or less than 10 micrometers. It is usually calculated based on the dust data collected using the AP-42 method.

1.1 The AP-42 Method for Paved Roads

According to AP-42, silt loading is measured directly in the field on road surface. Based on the silt loading collected in the field, PM_{10} emission factor is calculated using Equation (1.1) below (U.S. EPA, 2005):

$$E = k\left(\frac{sL}{2}\right)^{0.65} \times \left(\frac{W}{3}\right)^{1.5} - C \tag{1.1}$$

where

E = particulate emission factor (having units matching the units of k),

k = particle size multiplier for particle size range and units of interest,

sL = road surface silt loading (grams per square meter) (g/m^2),

W = average weight (tons) of the vehicles traveling the road.

C = emission factor for 1980's vehicle fleet exhaust, brake wear and tire wear.

In AP-42, silt loading is defined as “the mass of silt-size material (equal to or less than 75 micrometers [μm] in physical diameter) per unit area of the travel surface” (U.S.EPA, 2005). “It is the product of the total road surface dust loading and the silt fraction” (U.S.EPA, 2005). “The total road surface dust loading consists of loose material that can be collected by broom sweeping and vacuuming of the traveled portion of the paved road. The silt fraction is determined by measuring the proportion of the loose dry surface dust that passes through a 200-mesh screen, using the ASTM-C-136 method” (U. S. EPA, 2005).

In the field, a vacuum cleaner with "tared" (i.e., weighed before use) filter bags is used for collecting paved road loose surface material in the traveled portion of the road. Generally speaking, there are two types of roadway network from the perspective of dust collection (U.S. EPA, 2005): (1) a road network consisting of many relatively short roads contained in a well-defined study area such as in an industry zone, and (2) a network consisting of longer roads with spatial heterogeneity. In the network of the second type, there are roads either longer or shorter than 1.5 miles. For the roads less than 1.5 miles, three plots with width up to 10 ft are required for collecting dust. As shown in Figure 1.2, the locations of these three plots, denoted as x_1, x_2, and x_3, vary from between zero to the road length, and can be located using random numbers. In this study, the road segments of the network covering the study area are all less than 1.5 miles, and thus this sampling procedure is assumed to be adopted.

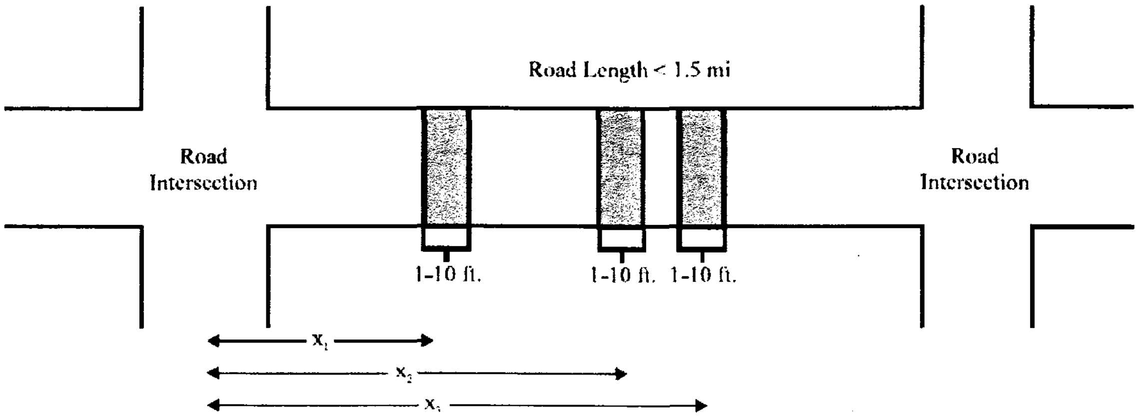


Figure 1.2 Incremental Sampling Locations for Paved Roads (U.S. EPA, 2005)

The following steps describe the collection method for dust samples (U.S. EPA, 2005).

Step 1: Block the traffic to ensure the sampling safety.

Step 2: Mark an area to be used for collecting dust on a specific paved roadway using markers such as string. According to the cleanness of roadway, the length of the area vertical to street center line ranges from 0.3 m (1 ft) of dirtier roads to 3 m (10 ft) of cleaner roads.

Step 3: Collect the dust samples within the area with vacuum sweeper in the field.

Step 4: Remove the vacuum bag without any loss of dust samples. And write down all the information required by the sampling data form as illustrated in Figure 1.3.

SAMPLING DATA FOR UNPAVED ROADS

Date Collected _________ Recorded by _________

Road Material (e.g., gravel, slag, dirt, etc.):* ______________________

Site of Sampling: ______________________

METHOD:
1. Sampling device: whisk broom and dustpan
2. Sampling depth: loose surface material (do not abrade road base)
3. Sample container: bucket with sealable liner
4. Gross sample specifications:
   a. Uncontrolled surfaces -- 5 kg (10 lb) to 23 kg (50 lb)
   b. Controlled surfaces -- minimum of 400 g (1 lb) is required for analysis

Refer to AP-42 Appendix B.1 for more detailed instructions.

Indicate any deviations from the above: ______________________

SAMPLING DATA COLLECTED:

Sample No.	Time	Location +	Surf. Area	Depth	Mass of Sample

* Indicate and give details if roads are controlled.
\+ Use code given on plant or road map for segment identification. Indicate sampling location on map.

Figure C.1-2. Example data form for unpaved road samples.

C.1-6 EMISSION FACTORS (Reformatted 1/95) 7/93

Figure 1.3 An Example of Sampling Data Form for Paved Roads (U.S. EPA, 2005)

1.2 Dust Sample Collection in the Las Vegas Valley

To attain the requirement of NAAQS for PM_{10} in the Las Vegas Valley, in June of 2001, Clark County submitted a PM_{10} State Implementation Plan (SIP) on how to control PM_{10} emission (Clark County, 2001). As part of the control plan, collecting PM_{10} data on a quarterly basis is required. Based on the AP-42 method, twenty four (24) sites representative of three roadway classifications (local roads, collectors, and arterials) were selected to estimate PM_{10} emission for paved roads in 2005 (James, 2007). The locations of these 24 locations are shown in Figure 1.4. The information about these locations is listed in Table 1.1. Table 1.2 indicates that 6, 11 and 7 samples were collected for Local road, Collector, Arterial, respectively.

For each of the 24 sites, three (3) plots were applied along one street block (see Figure 1.5). The plot lengths were 10 ft while the plot widths varied from 10 ft to 15 ft based on the roadway lane width (see Table 1.1). The spaces between plots ranged from 10 ft to 30 ft. Figure 1.6 shows an example of the actual layout of plots along a street and the usage of vacuum in the field.

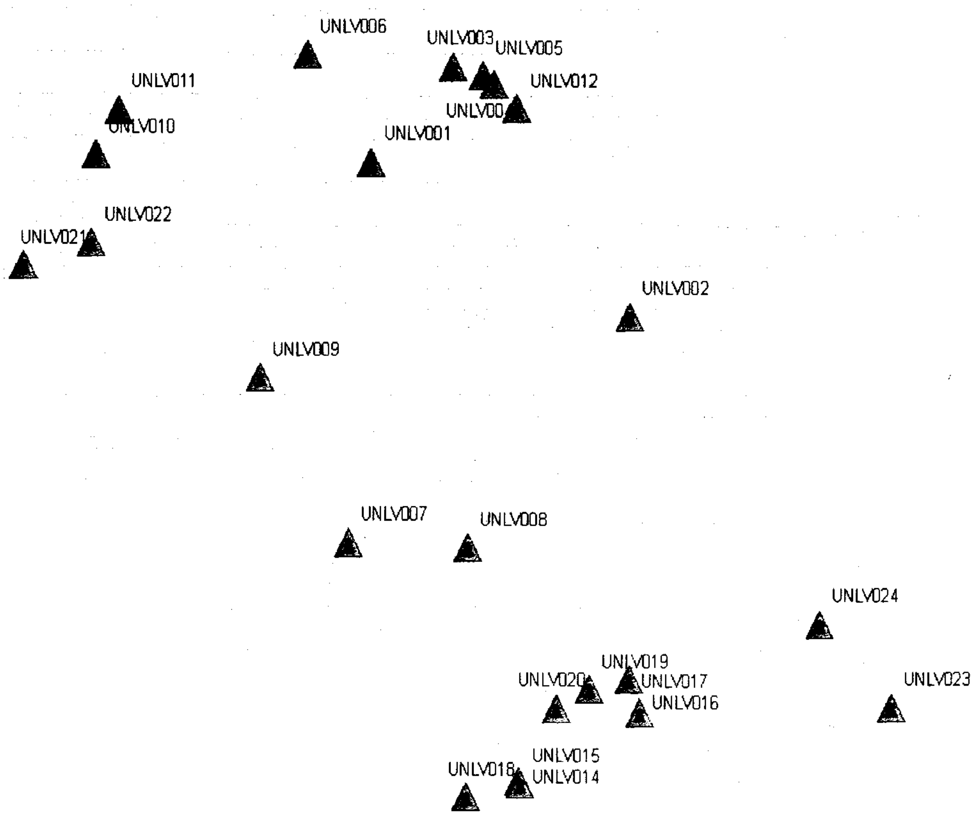


Figure 1.4 24 Sample Sites for the AP-42 Method (James, 2007)

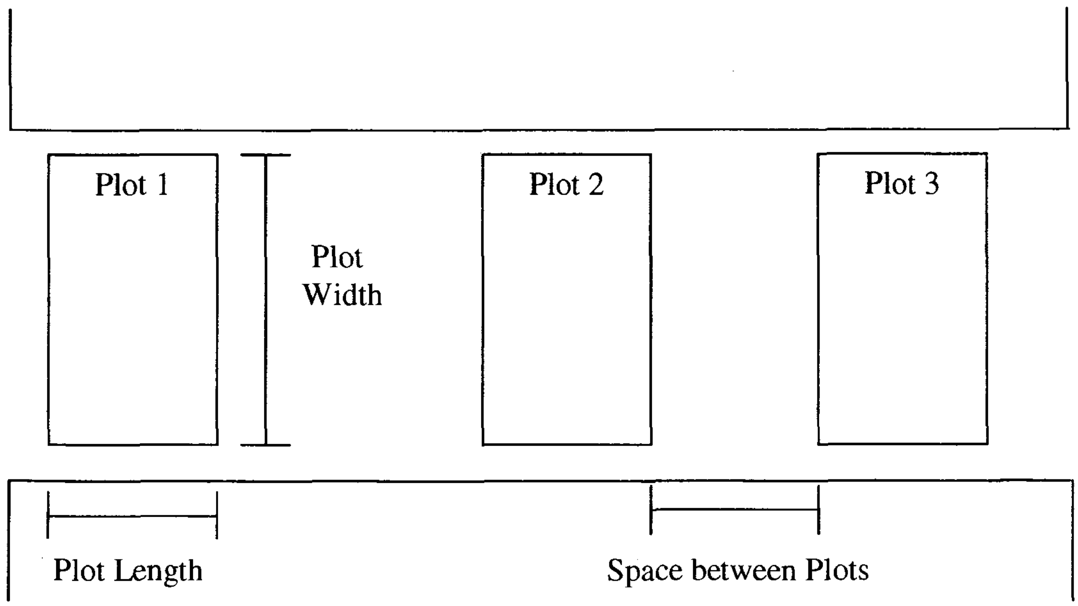


Figure 1.5 The AP-42 Method Adopted in a Study for the Las Vegas Valley

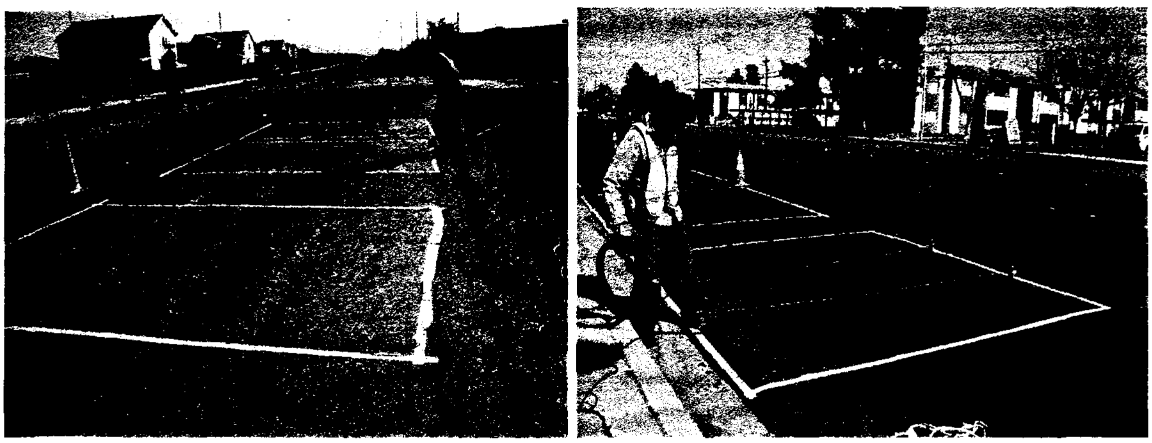


Figure 1.6 An Example of Field Work Following the AP-42 Method (Rodrigues and James, 2005)

Table 1.1 Sampling Characteristics for 24 Locations (Rodrigues and James, 2005)

Site #	Nearest Intersection	Roadway Classification	Plot size (ft)	Number of Plots
UNLV001	Gowan-Kidd-Coleman	Collector	10 x 12	3
UNLV002	Marion & Diamond Head	Collector	10 x 12	
UNLV003	Daywood & Quailbush	Local	10 x 11	
UNLV004	Emerald Stone & Sapphire Light	Local	10 x 11	
UNLV005	Washburn & Donna	Collector	10 x 12	
UNLV006	Ann Rd & San Mateo	Arterial	10 x 15	
UNLV007	Polaris & Hacienda	Collector	10 x 14	
UNLV008	Maryland & Westminster	Arterial	10 x 12	
UNLV009	Duneville & El Parque	Local	10 x 11	
UNLV010	Gowan & El Capitan	Collector	10 x 11	
UNLV011	Durango & Craig	Arterial	10 x 12	
UNLV012	Losee & Craig	Arterial	10 x 11	
UNLV013	Eastern & Hardin	Local	10 x 10	
UNLV014	Norridgewood & Evergold	Local	10 x 15	
UNLV015	Coral Sea & Ione	Collector	10 x 11	
UNLV016	Valle Verde & Wigwam	Arterial	10 x 11	
UNLV017	Silver Springs & Spring Hills	Collector	10 x 11	
UNLV018	Maryland & Pyle	Arterial	10 x 12	
UNLV019	Armacost & Calmar	Local	10 x 11	
UNLV020	Pecos & Wigwam	Arterial	10 x 12	
UNLV021	Crestdale & Covington Cross	Collector	10 x 10	
UNLV022	Hillpointe & Rampart	Collector	10 x 12	
UNLV023	Burkholder & Cabrillo	Collector	10 x 13	
UNLV024	Pabco & Liverpool	Collector	10 x 11	

Table 1.2 Sample Size for Three Roadway Classifications (Rodrigues and James, 2005)

Roadway Classification	Sample Size	Sample Road Segments
Local	6	UNLV003, UNLV004, UNLV009, UNLV013, UNLV014, UNLV019
Collector	11	UNLV001, UNLV002, UNLV005, UNLV007, UNLV010, UNLV015, UNLV017,UNLV021, UNLV022, UNLV023, UNLV024
Arterial	7	UNLV006, UNLV008, UNLV011, UNLV012, UNLV016, UNLV018, UNLV020

1.3 Problem Statement

Basically, AP-42 is the dust collection method recommended by EPA. To follow the AP-42 procedure, the composite sample size (the number of road sites) needs to be determined first. At a given sampling site, it is necessary to know whether the required number of plots and their sizes (particularly in terms of the length of a plot) are sufficient for collecting dust that can produce an estimate of PM_{10} emission measure with acceptable accuracy. So far, there hasn't been any rule to follow to decide on the number of sample sites. Whether the number of plots and their sizes as specified for the AP-42 method to be used at each sample site are sufficient has not been validated. The difficulty for conducting a comprehensive study to address these issues is the lack of dust emission data for every spatially continuous point on roadways. Without this continuous emission data, it is not possible to know the true probability distribution of emissions over a road network, and thus there wouldn't be a base to develop solid statistical methods either to determine the required number of sample sites or to validate the parameters specified for collecting dust at a given sample site. In recent years, mobile sampling technologies have been proposed and tested for collecting dust data. Basically, mobile sampling technologies consist of a vehicle that is equipped with dust measuring devices, computers and Global Position System (GPS) devices. These devices are integrated so that dust data can be measured in real time and recorded in computer for analysis. The dust data collected in this way can be viewed approximate to the spatially continuous data required to derive statistics for a road network, and thus provide a chance to analyze the sample size issue and the required parameters for the AP-42 method.

The objective of this study is to investigate the composite sample size and the number of plots and their sizes that are required for the AP-42 method, based on the dust data collected using one of the mobile sampling technologies. Based on analyzing the dust data for different roadway classifications, it is found that the dust emission levels are significantly different, which leads to considering each roadway classification separately. When the number of sample sites is investigated, different roadway classifications are viewed as different stratum. One of the stratified sampling methods, the optimal allocation method, is adopted. By using this method, the variance of the estimated PM_{10} emission measures based on the samples can be minimized under a fixed budget. Considering a fixed budget in this study is close to the field sampling practice because the resources available for dust collection have always been controlled within a limit. This limit may vary over years. In this study, two emission measures: emission factor and emission potential, are used in deriving the number of sample sites. Emission factor is to measure the amount of PM_{10} per unit roadway segment, and is used in the AP-42 method, while emission potential measures the extent of how "dirty" a road is (i.e., the density of PM_{10}) which is an additional measure collected by the mobile sampling technology. The number of sample sites derived based on these two measures are compared.

The issues related to the number of plots and their sizes were investigated by taking the Monte Carlo approach. In this approach, different combinations of the number of plots and their sizes are tried in the simulation. For a given combination of the number of plots and their sizes, the layouts of these plots on a roadway segment are generated in the simulation by emulating the way they are actually set up in the field following the AP-42 method. A number of layouts are generated for the roadway segment, given the combination of the number of plots and their sizes. The average of the emission factors derived using the emission data included in these plots with different layouts is compared with the actual average derived based on all the emission data on the roadway segment (considering those not included in these plots). The patterns of the comparison in emission factors versus the number of plots and their size are observed, from which the threshold of plot number and size producing the best estimation of emission factors are determined.

1.4 Thesis Outline

This thesis consists of five chapters. The first chapter introduces the AP-42 method and its application in the Las Vegas area. Based on this introduction, the objective of this study is proposed, which is to investigate the number of sample sites and plot number and size for a given sample site. Chapter 2 describes the methodology which includes the optimal allocation method and the Monte Carlo simulation method. To address the need to adopt the Monte Carlo simulation approach, an optimization formulation is provided for determining the plot number and size in this section. In Chapter 3, the extraction of emission data collected using a mobile sampling technology is introduced. The characteristics of the extracted data are presented in terms of the spatial distribution over road segments on different roadway classifications. The investigation on whether the data collected on different days can be combined for this study is also described in this chapter. Chapter 4 presents the application of the optimal allocation method in determining the number of sample sites and the comparisons between the estimated and true average of emission factors with plot number and size varied. In the last Chapter, conclusions are drawn upon the analysis in Chapter 4. Future study needs are also identified in this chapter with discussions on the limitation of the research results in this study.

CHAPTER 2

METHODOLOGY

2.1 Optimal Allocation Method for Determining the Number of Sampling Sites

In this study, one of the stratified sampling methods, optimal allocation method, is adopted to determine the number of sites to be sampled using the AP-42 method. With this method, it can be assumed that the cost of collecting dust at one site is different per roadway classification. In addition, the budget available for a dust data collection study is limited. Then the number of sites to be sampled can be determined for each roadway classification using the following formula (Cochran, 1977):

$$n_h = \frac{N_h \sigma_{hx} / \sqrt{C_h}}{\sum_{h=1}^{L} (N_h \sigma_{hx} \sqrt{C_h})} \times C \tag{2.1}$$

where

n_h = the number of sites (or road segment) to be sampled for the roadway classification h,

N_h = the total number of road segments in the roadway classification h,

σ_h = the standard deviation of emission measurements on the road segments in the roadway classification h,

C_h = unit cost of sampling one site on a road segment in the roadway classification h, and

C = total budget available for each quarter of dust sampling,

L = total number of roadway classifications.

The relationship between the unit cost and the total budget can be expressed as:

$$C = \sum_{h=1}^{L} n_h C_h \tag{2.2}$$

When the proportion of road segments W_h is known, the number of sites to be sampled for each roadway classification h can be derived using the following equation (Cochran, 1977):

$$n_h = \frac{W_h \sigma_{hx} / \sqrt{C_h}}{\sum_{h=1}^{L} (W_h \sigma_{hx} \sqrt{C_h})} \times C \tag{2.3}$$

When the unit costs are equal between different roadway classifications, i.e., $C_1 = C_2 = ... = C_L = c$, Equation (2.3) can be further written as:

$$n_h = \frac{W_h \sigma_h}{c \sum_{h=1}^{L} (W_h \sigma_h)} \times C = \frac{W_h \sigma_h}{\sum_{h=1}^{L} (W_h \sigma_h)} \times n \tag{2.4}$$

where n is the total number of sites to be sampled for all the roadway classifications.

In this study, it was assumed that the unit costs of collecting dust for one site of different classifications of roadways are the same. It is due to that fact that only one lane of dust is collected at each site, even though there are different numbers of lanes for different roadway classifications. As a result, Equation (2.4) is used in this study. It can be seen from Equation (2.4) that two sets of parameters must be known for deriving the number of sampling sites: the proportion of road segments for each roadway classification W_h and the standard deviation of emission measurements σ_h of these roadway classifications. It is known that the standard deviation of emission measures cannot be derived directly because emission measurements are not available for all the road segments in a highway network. In a previous study (Etyemezian et al., 2006), a mobile sampling technology was used to collect emission measures (including emission factor and emission potential) for a selected set of road segments that are traveled in a tour. The emission measurements of these road segments are viewed as representative to the whole population of segments, thus their standard deviations are used in Equation (2.4). In this sense, Equation (2.4) can be written as follows:

$$n_h = \frac{W_h \hat{\sigma}_h}{\sum_{h=1}^{L} (W_h \hat{\sigma}_h)} \times n \qquad (2.5)$$

where $\hat{\sigma}_h$ represents the estimated standard deviation, not the true values from the whole population of the road segments.

Note that the emission factor is the measure that is adopted in the AP-42 method. By definition, emission factor measures the amount of dust emission on a roadway segment. Emission potential is another emission measure that gauges the density of dust, a measure on the "dirtiness" of a roadway. In this study, the number of sample sites was also derived based on this emission measure for the purpose of comparison.

2.2 Monte Carlo Simulation for Determining Number and Length of Plots

In this study, a Monte Carlo simulation model (Rubinstein, 1981 and 1986) was adopted to investigate the optimal combination of the number of plots and their length for the AP-42 method. Figure 2.1 displays the hypothetical emission measurements (17 data points) from a mobile sampling technology and the configuration of the plots that would be used to collect dust data based on the AP-42 method. It can be seen that an average emission factor can be estimated for the road segment based on all the 17 emission measurements from a mobile sampling technology. This average can be viewed close enough to the true average of emission factor for the illustrated road segment. When adopting the AP-42 method, the emission factor derived based on the emission measurements (nine data points in Figure 2.1) included in the plots would be different from the true value.

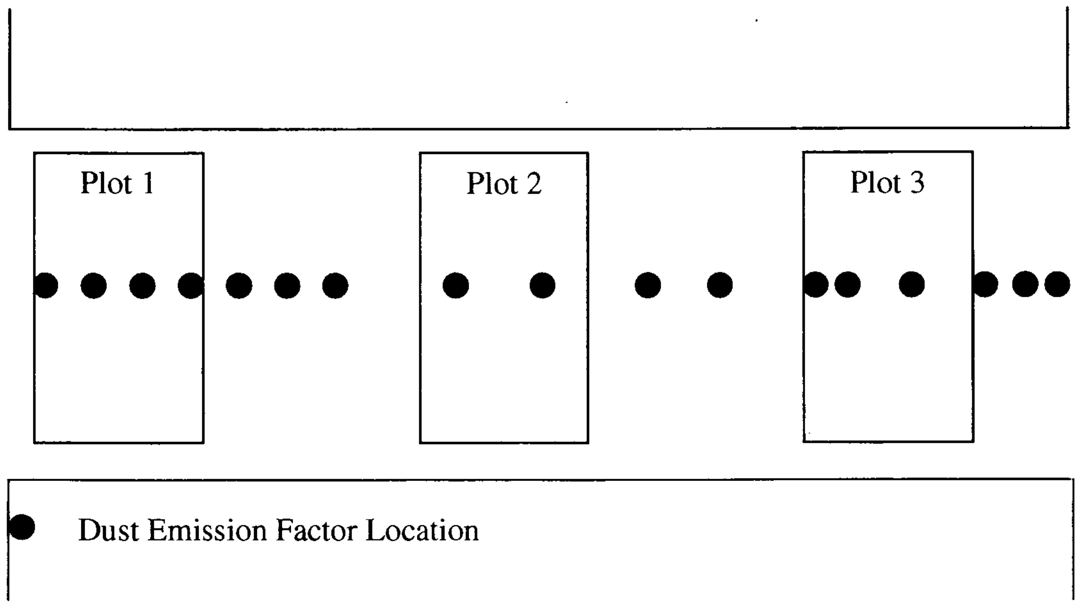


Figure 2.1 Illustration of Plots and Dust Data Points on a Road Segment

The differences between the estimated and true emission factors could vary for different configuration of the plots in terms of the number of plots, their locations and lengths on a road segment. If assumed that the locations of the plots can be decided in a random manner, there should be an optimal combination of the number of plots and their lengths that produce an estimation of emission factor for one road segment that are closest to the true value. In the same way, it is reasonable to imagine that one optimal combination of the number of plots and their lengths can be found for all the road segments in a roadway classification. In this sense, an optimization problem can be formulated for the optimal number of plots and their lengths as follows:

$$\min_{P,D} \frac{\sum_{i=1}^{n}\left[\overline{EF}_{i,E}(P,D)-\overline{EF}_{i,T}\right]}{n} \quad (2.6)$$

where P and D represent the number of plots and their length, respectively. In this formulation, it is assumed that all the P plots have the same length. $\overline{EF}_{i,E}(P,D)$ and $\overline{EF}_{i,T}$ denote the estimated and true emission factors, respectively, for the road segment i. The term $\overline{EF}_{i,E}(P,D)$ in Equation (2.6) indicates that its value is related to the combination of P and D. n is the total number of road segments in a roadway classification.

If an analytic relationship can be found between the estimated emission factor and the number and length of plots (i.e., P and D), this optimization problem can be solved analytically. Unfortunately, such an analytic relationship cannot be found. Therefore, a simulation approach was taken by which an emission factor can be estimated given a specific set of P and D. In this approach, a set of configurations of plots, each with the same number of plot and length but with different locations on a road segment, can be generated randomly. Based on the generated configurations of plots, an emission factor can be estimated, which can then be compared with the true value.

This random procedure can be applied for all combinations of plot number and length for a road segment, and then for all the road segments in each roadway classification. The pattern on difference between the estimated and true emission factors versus the combination of plot number and length can be plotted. The combination of plot number and length producing the minimum difference can be derived from the graphic presentation of the patterns. The simulation procedure is presented in Figure 2.2

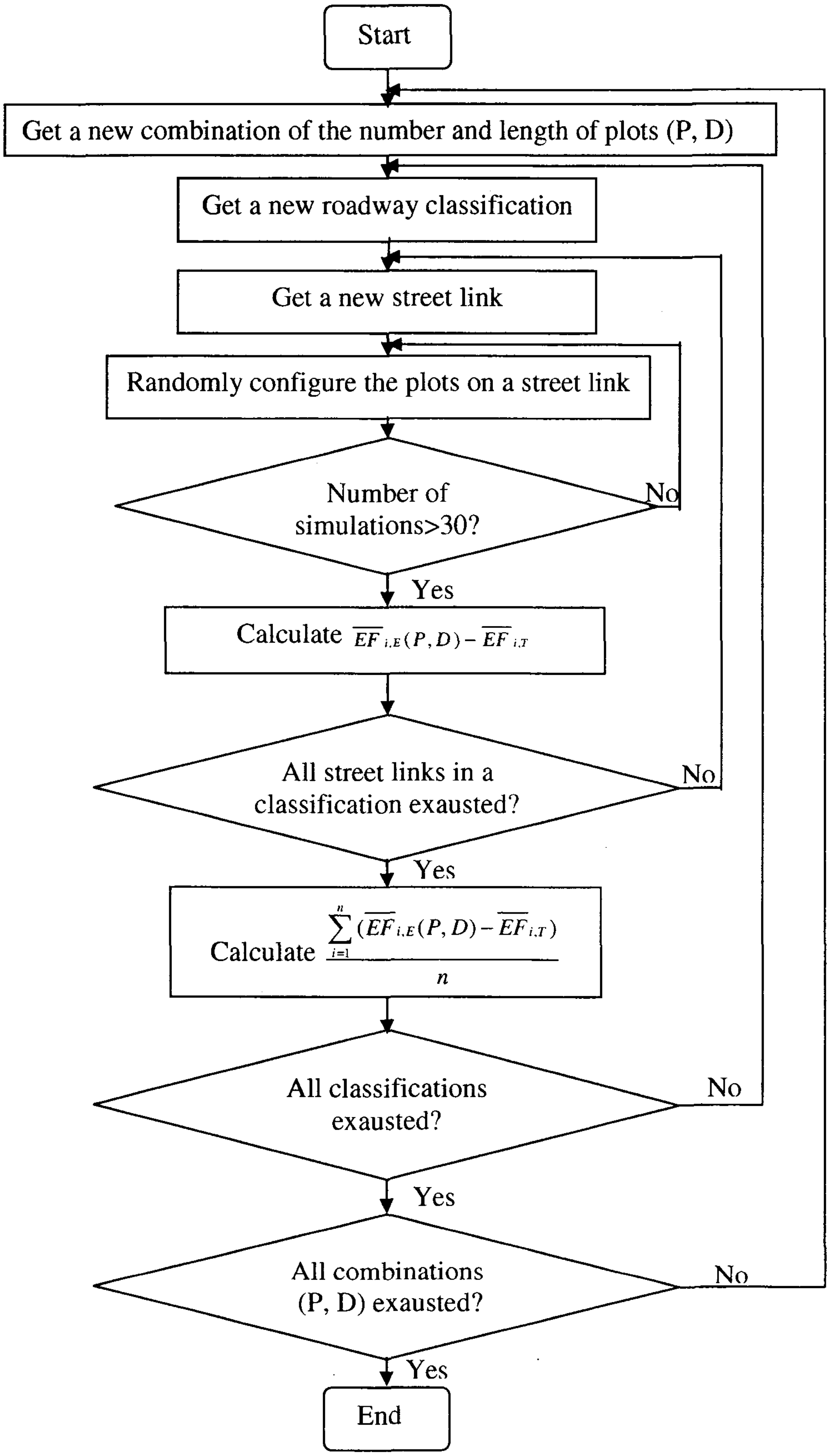


Figure 2.2 Monte Carlo Simulation Procedures

In this study, the number of plots is varied from 1 to 10, and their length changes from 10 ft to 50 ft. In the simulation, the true emission factors $\overline{EF}_{i,T}$ for each road segment i are calculated using the equation below:

$$\overline{EF}_{i,T} = \frac{\sum_{k=1}^{p_i} EF_{i,k}}{p_i}, \tag{2.7}$$

where $EF_{i,k}$ = k-th emission factor data point on road segment i; and p_i = the total number of dust emission factor data points on the road segment i.

The emission factor estimated based on the plots in each road segment can be derived as follows:

$$\overline{EF}_{i,E}(P,D) = \frac{\sum_{p=1}^{P} \overline{EF}_{i,p}}{P}, \tag{2.8}$$

where $\overline{EF}_{i,p}$ represents the emission factor for each plot p, and P is the total number of plots on the road segment i. The emission factor for each plot can be derived using the formula below:

$$\overline{EF}_{i,p} = \frac{\sum_{l=1}^{L} EF_{i,p,l}}{L}, \tag{2.9}$$

where $EF_{i,p,l}$ denotes the l-th emission factor data point in plot p, and L represents the total number of data points in plot p.

Figure 2.1 shows an example where there are 17 data points on a road segment. These emission factor data points can be represented as x_1, x_2, ..., x_{17}. The data points included in Plots 1, 2 and 3 are (x_1, x_2, x_3, x_4), (x_8, x_9) and (x_{12}, x_{13}, x_{14}), respectively. The estimated true value can be derived as $\sum_{k=1}^{17} x_k / 17$. The estimated emission factor for these three plots are $(x_1 + x_2 + x_3 + x_4)/4$, $(x_8 + x_9)/2$ and $(x_{12} + x_{13} + x_{14})/3$. The estimated average emission factor for this road segment is the average of these three estimated emission factors, which can be written as: $[(x_1 + x_2 + x_3 + x_4)/4 + (x_8 + x_9)/2 + (x_{12} + x_{13} + x_{14})/3]/3$.

Note that Figure 2.1 presents an ideal situation where at least one dust emission factor data point is included in each plot. Figures 2.3 and 2.4 show two cases where there is no data point included in a plot. In Figure 2.3, a plot falls in between two emission data points, while in Figure 2.4, there is no data point on one side of a plot. For the case shown in the Figure 2.3, interpolating method can be adopted to derive the emission factor for the plot by considering the distance between the center line of the plot to these two neighboring data points. In case shown in Figure 2.4, the emission factor for the plot can be estimated as the average of the two data points on the one side of the plot.

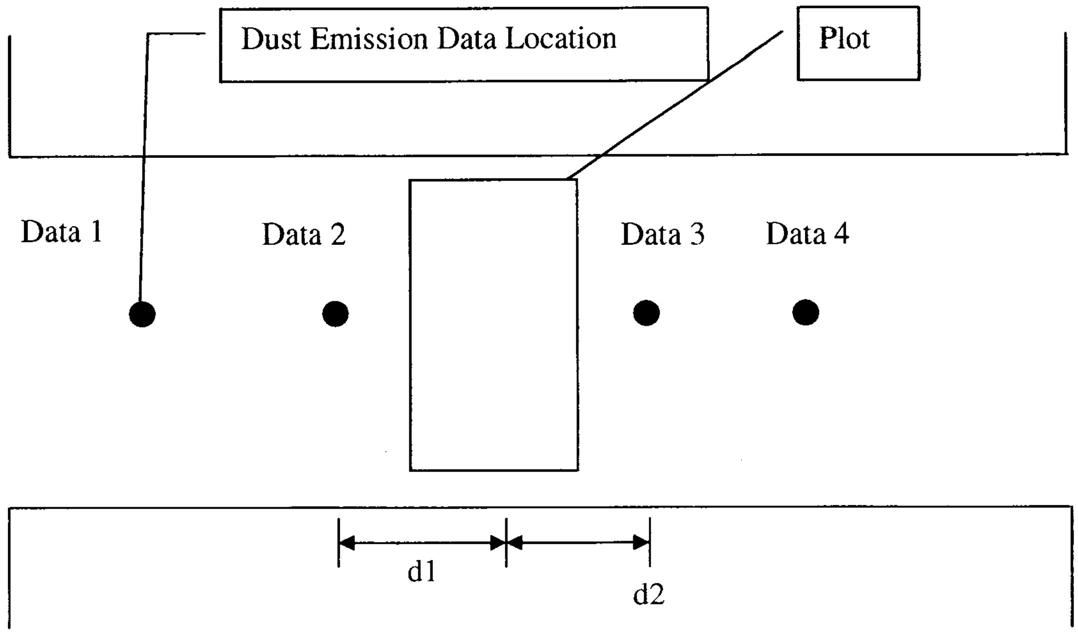


Figure 2.3 A Plot Falls in Between Two Emission Data Points

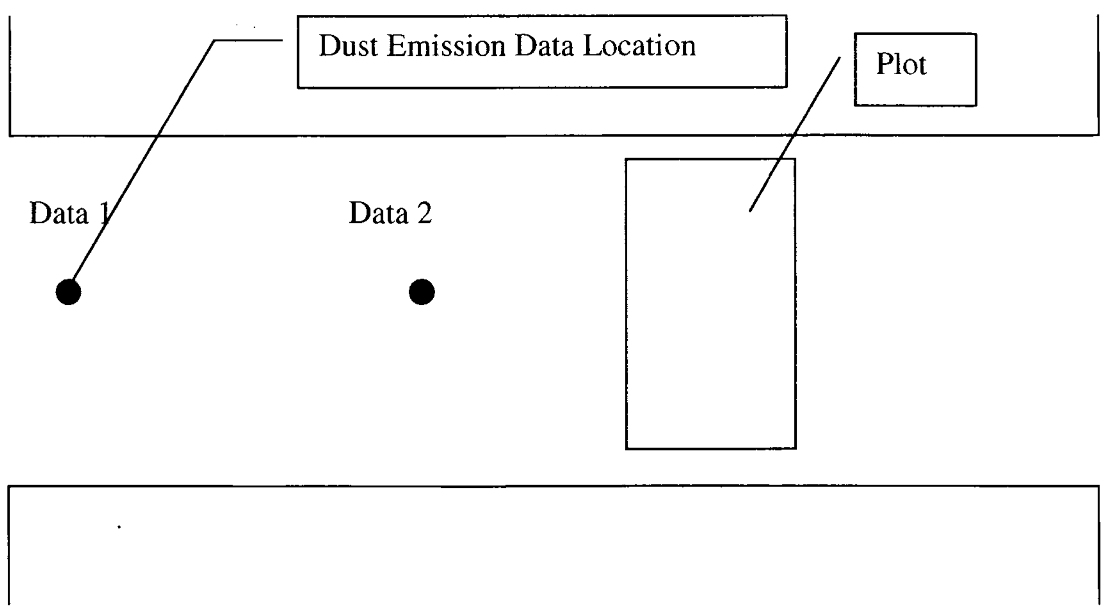


Figure 2.4 All the Data Lay at the One Side of Plots

CHAPTER 3

DATA COLLECTION

3.1 Emission Data Collected by Using a Mobile Sampling Technology

Dust emission data were obtained from the Clark County Department of Air Quality and Environmental Management, which were collected in a previous study by using a mobile sampling technology. In this previous study (Etyemezian et al., 2006), a 3.75-ton van was used, which is equipped with an integrated computer, dust filters and processors, and a vehicle location device (Global Position System, GPS). With this integrated system, the vehicle can collect dust emission data such as emission factors and emission potential online while the vehicle travels on streets. This vehicle ran on a fixed 97 mile long tour (see Figure 3.1) covered in one day. The tour consists of six roadway classifications: interstate, freeway, major arterial, minor arterial, collector and local (see Figure 3.2). The number of road segements and their average length in each roadway classification are listed in Table 3.1. The vehicle ran on the tour for four consecutive days (from February 14, 2005 to February 17, 2005). Even though the filters and processors on the vehicle can produce data continuously, only a discrete number of emission data points, each with longitudinal and latitudinal data, can be made available for recording. Figures 3.3 presents an example of the spatial distribution of the data on several road segments. The number of data points on each road segment is shown in Table 3.1.

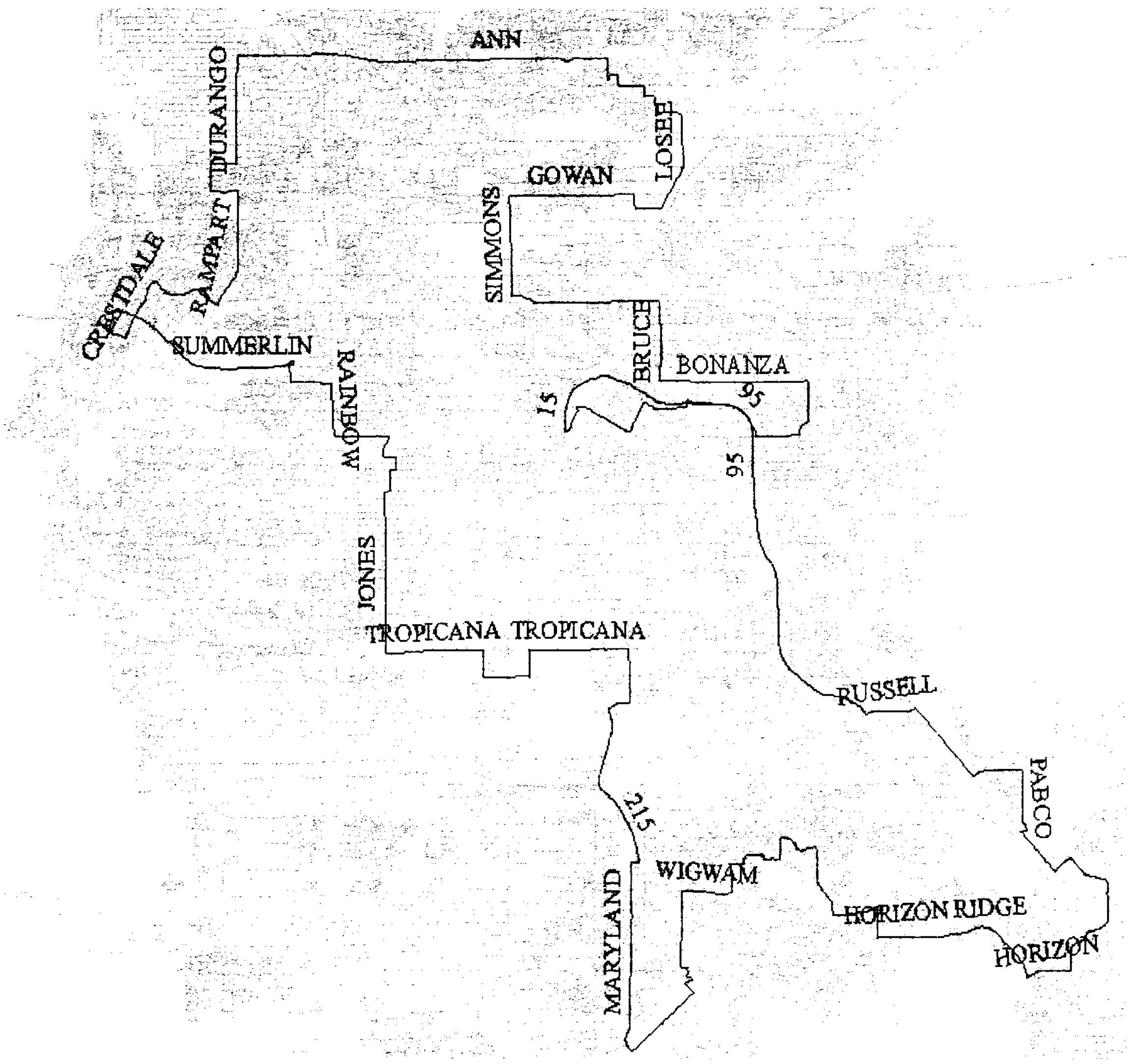


Figure 3.1 Tour Followed by a Vehicle Equipped with a Mobile Sampling Technology

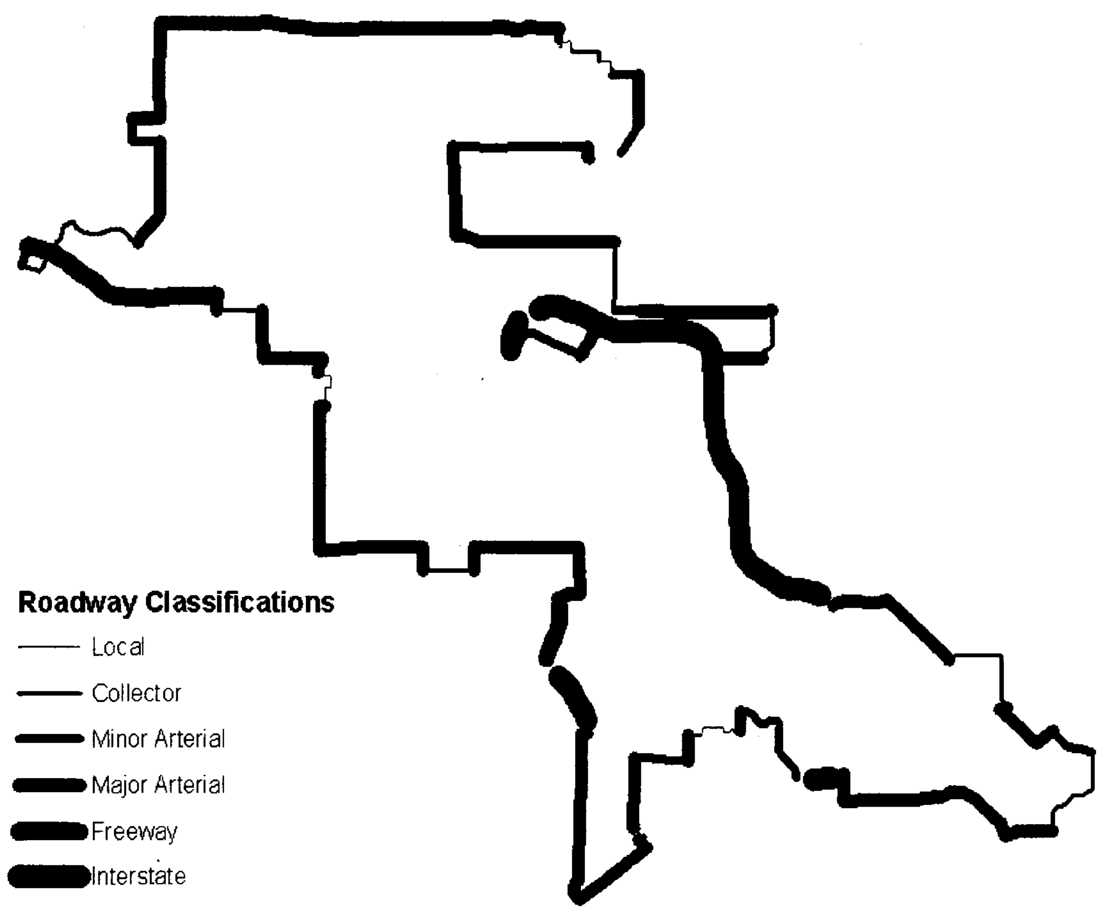


Figure 3.2 Roadway Classifications Included in a Tour Followed by a Vehicle with a Mobile Sampling Technology

Table 3.1 Information about the Tour and Dust Data

Classifications	No. of Road Segments	Average Length of Road Segments (ft.)	Number of Data Points on Road Segments			
			2/14/05	2/15/05	2/16/05	2/17/05
Local	39	401	374	476	446	317
Collector	139	499	1,214	1,511	1,253	1,189
Minor Arterial	133	477	1,198	1,352	1,293	1,013
Major Arterial	410	667	4,267	4,913	4,616	3,603
Freeway	20	1,199	224	252	229	198
Interstate	67	982	530	775	597	719
Total	808		7,807	9,279	8,434	7,039

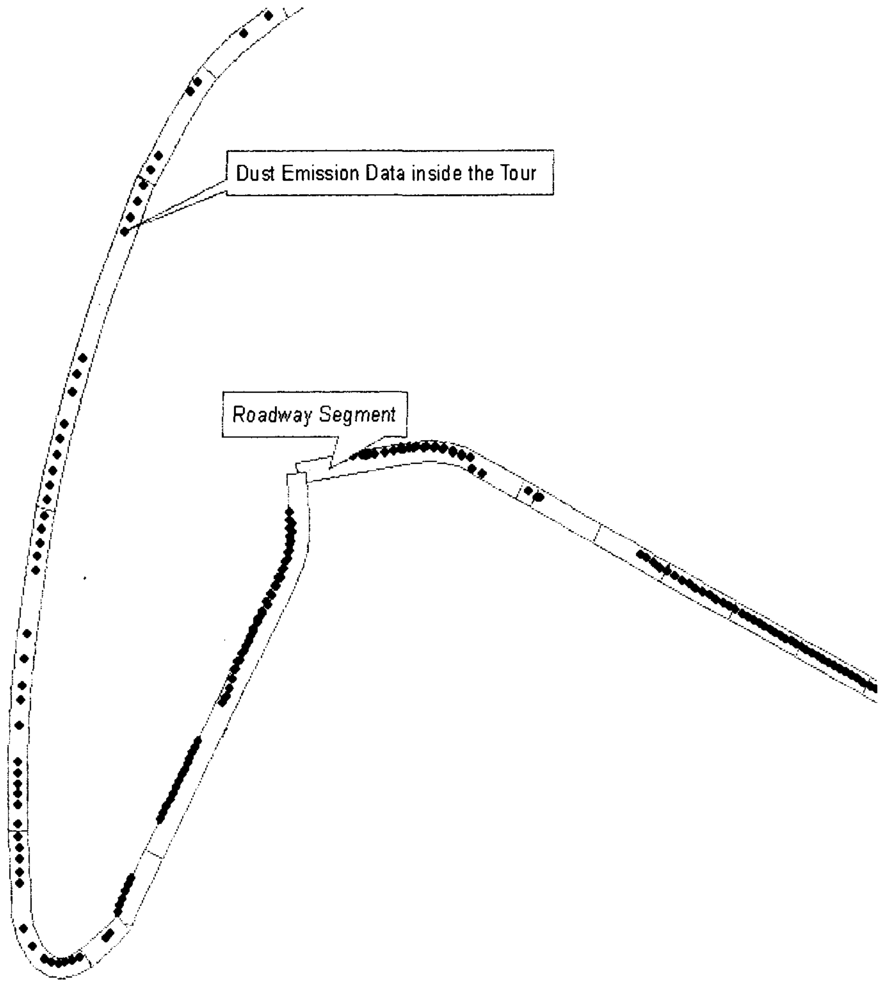


Figure 3.3 Spatial Distribution of Emission Data on Road Segments

3.2 Expanding Emission Data with Information on Road Segments

The emission data collected using the mobile sampling technology cannot be used directly for this study, because the data points are stored in a database with no information about the road segments where they were collected. For the analysis in this study, it is necessary to know on which road segment a measurement is taken. With this information, statistics of emission data such as average and variance of emission factors on a road segment or a roadway classification can be calculated and then used in the analysis specified in the methodology. Thus, a data processing procedure has to be developed to add the information of road segments to the existing emission data. The following steps are followed in the procedure. First, the emission data are displayed along the tour on a GIS map. In this study, a GIS map from the Clark County website (http://www.co.clark.nv.us/) was downloaded. On this map, the road segments included in the tour are identified and included in a single layer. Based on the longitudinal and latitudinal information associated with the emission data, the emission data points are also displayed on this map. The distribution of the emission data points on a portion of the tour is shown in Figure 3.3. Second, the emission data are connected to the road segments where they are collected. In the original form of the emission data, there are only longitudinal and latitudinal data available for each emission data measurement (see Figure 3.4). To make the connection between the emission data and road segments, buffers with a certain width (80 ft in this study) along each road segment are created (see Figure 3.3). Those data which lay within or on the edge of the buffers are considered in this study, while those outside the buffers are viewed as outliers and thus are not included in the data for analysis. Figure 3.5 shows the expanded data structure for emission measurements. The fields FUNC_CLASS and STRNAME in Figure 3.5 are newly added to the database in Figure 3.4.

Attributes of XYDRIFEBFOURTEEN

FID	Shape*	ID	GPSTime	EF	EP	
0	Point	1	12:29:33 PM	0.377251	0.069780	2/
1	Point	2	12:29:34 PM	0.411397	0.0705	2/
2	Point	3	12:29:35 PM	0.429406	0.070736	2/
3	Point	4	12:29:53 PM	0.502417	0.0923	2/
4	Point	5	12:29:54 PM	0.495318	0.0911	2/
5	Point	6	12:29:55 PM	0.486305	0.0887	2/
6	Point	7	12:29:57 PM	0.503695	0.0882	2/
7	Point	8	12:29:59 PM	0.471647	0.0803	2/
8	Point	9	12:30:00 PM	0.443933	0.0748	2/
9	Point	10	12:30:22 PM	0.573244	0.0914	2/
10	Point	11	12:30:23 PM	0.677281	0.099956	2/
11	Point	12	12:30:24 PM	0.646046	0.0933	2/
12	Point	13	12:30:25 PM	0.548794	0.0822	2/
13	Point	14	12:30:55 PM	0.170841	0.0292	2/
14	Point	15	12:31:00 PM	0.301284	0.0326	2/
15	Point	16	12:31:02 PM	0.335437	0.0367	2/
16	Point	17	12:31:03 PM	0.344304	0.0381	2/
17	Point	18	12:31:05 PM	0.310096	0.0354	2/
18	Point	19	12:31:06 PM	0.294535	0.034	2/

Record: 0 Show: All Selected Records (0 out of 8282 Selected.) Options

Figure 3.4 Dust Emission Data Structure

Attributes of Join_Str_16

FID	Shape	FUNC_CLASS	STRNAME	EP	EF	FID_1
3209	Point	Freeway	AIRPORT CONN N	0.0082	0.200935	3691
3210	Point	Collector	PABCO	0.0375	0.397408	1198
3211	Point	Collector	PABCO	0.0361	0.374965	1199
3212	Point	Collector	PABCO	0.0334	0.339928	1200
3213	Point	Collector	PABCO	0.0325	0.323493	1201
3214	Point	Collector	PABCO	0.0324	0.312660	1202
3215	Point	Collector	PABCO	0.0348	0.315784	1203
3216	Point	Collector	PABCO	0.0342	0.298654	1204
3217	Point	Collector	SUNSET	0.0233	0.162996	1110
3218	Point	Collector	SUNSET	0.0207	0.148432	1111
3219	Point	Collector	SUNSET	0.019430	0.145379	1112
3220	Point	Collector	SUNSET	0.0295	0.231056	1113
3221	Point	Collector	SUNSET	0.0388	0.343366	1114
3222	Point	Collector	SUNSET	0.0289	0.273093	1115
3223	Point	Collector	SUNSET	0.026489	0.265708	1116
3224	Point	Collector	SUNSET	0.0227	0.241668	1117
3225	Point	Collector	SUNSET	0.020831	0.247964	1118
3226	Point	Collector	SUNSET	0.017603	0.219927	1119
3227	Point	Collector	SUNSET	0.021064	0.275718	1120

Record: 3210 Show: All Selected Records (0 out of 9175 Selected.) Options

Figure 3.5 Expanded Data Structure of Emission Measurements

3.3 Description of Dust Emission Data

Given the integrated emission data, they can be displayed by road segment and by roadway classification. By displaying the data, their characteristics can be identified. In this study, the emission factor data are displayed in box-plot charts each showing all the road segments on a roadway classification (see Figures 3. 7 - 3. 12).

Figure 3.6 shows that basic elements involved in a box-plot chart: median, lower hinge, upper hinge, lower limit, upper limit and outliers. The line within the box represents median value; the upper line (upper hinge) of the box represents 75th percentile value; the lower line (lower hinge) of the box represents 25th percentile value; the end of the vertical line above the box is the upper limit value, and the end of the vertical line below the box is the lower limit value. The supposed outlier is the value out of range between lower and upper limits.

From Figure 3.7-3.12, it can be seen that the variance of emission factors on a roadway classification is correlated with the average of emission factors. In other words, a roadway classification with high emission factor is highly likely to have relatively large variance. In addition, the dust emission data are displayed where the emission data of all the road segments in a roadway classification are aggregated (see Figures 3.13 and 3.14 for emission factor and emission potential, respectively). It can be seen that the variance of emission data on local roads is the biggest and that on freeways and interstates are the smallest. Freeways, interstates and major arterials have lower average of dust emissions than the average of emission for all the roadway classification. Major arterials had more outliers than other roadway functional classifications.

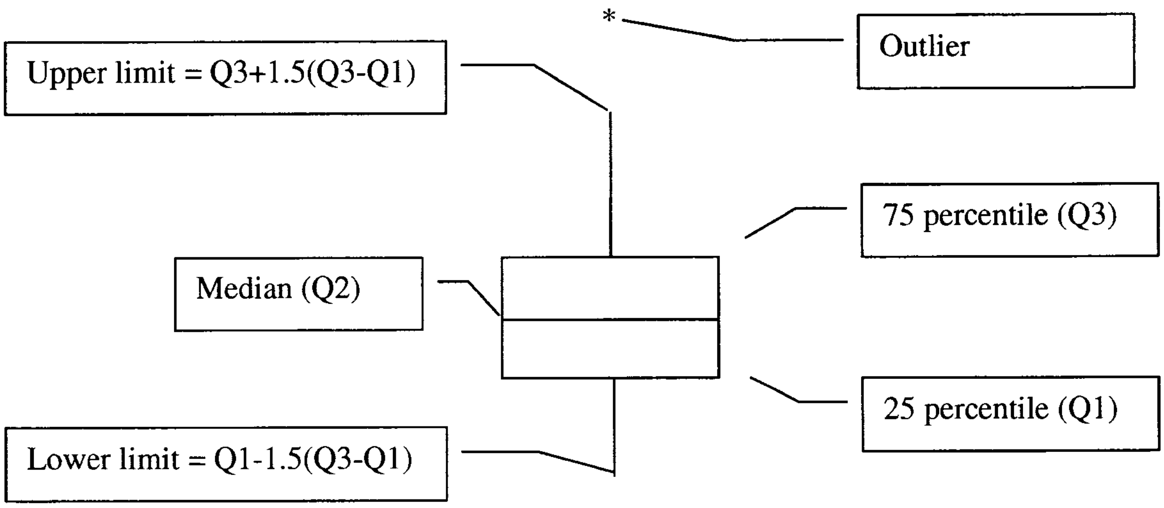


Figure 3.6 Illustration for Box-plot Chart

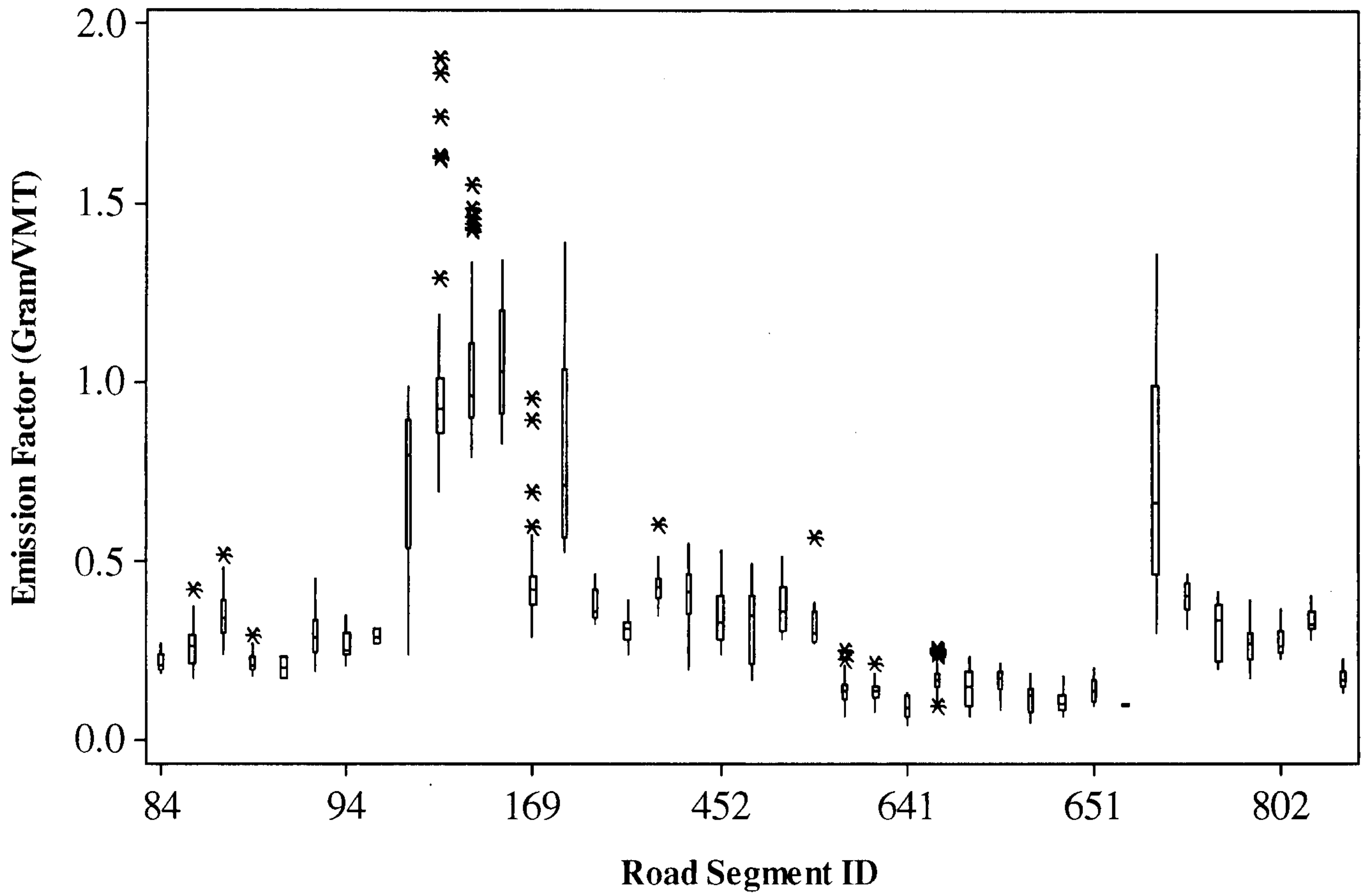


Figure 3.7 Emission Factor Box-plot Chart for Local Segments

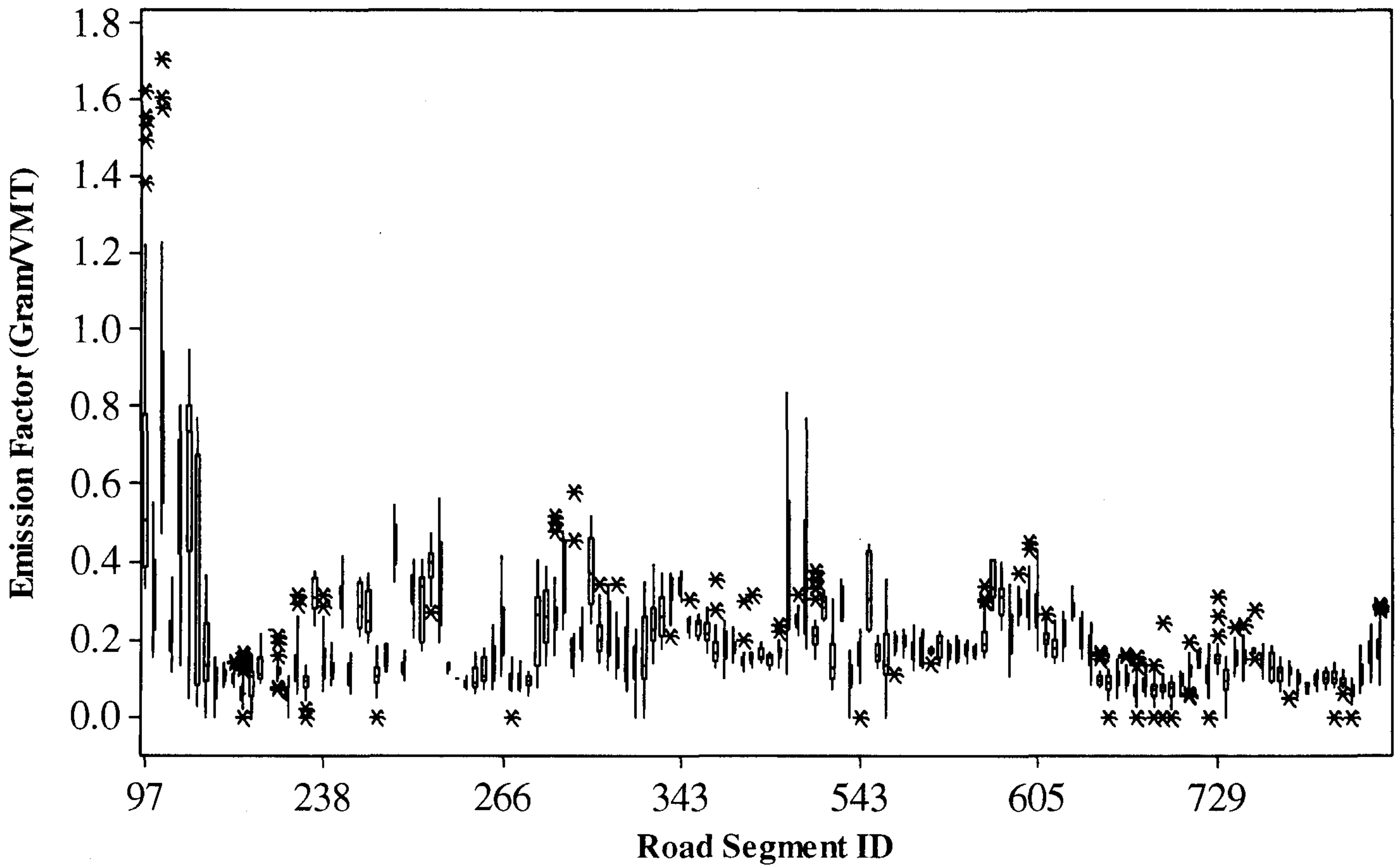


Figure 3.8 Emission Factor Box-plot Chart for Collector Segments

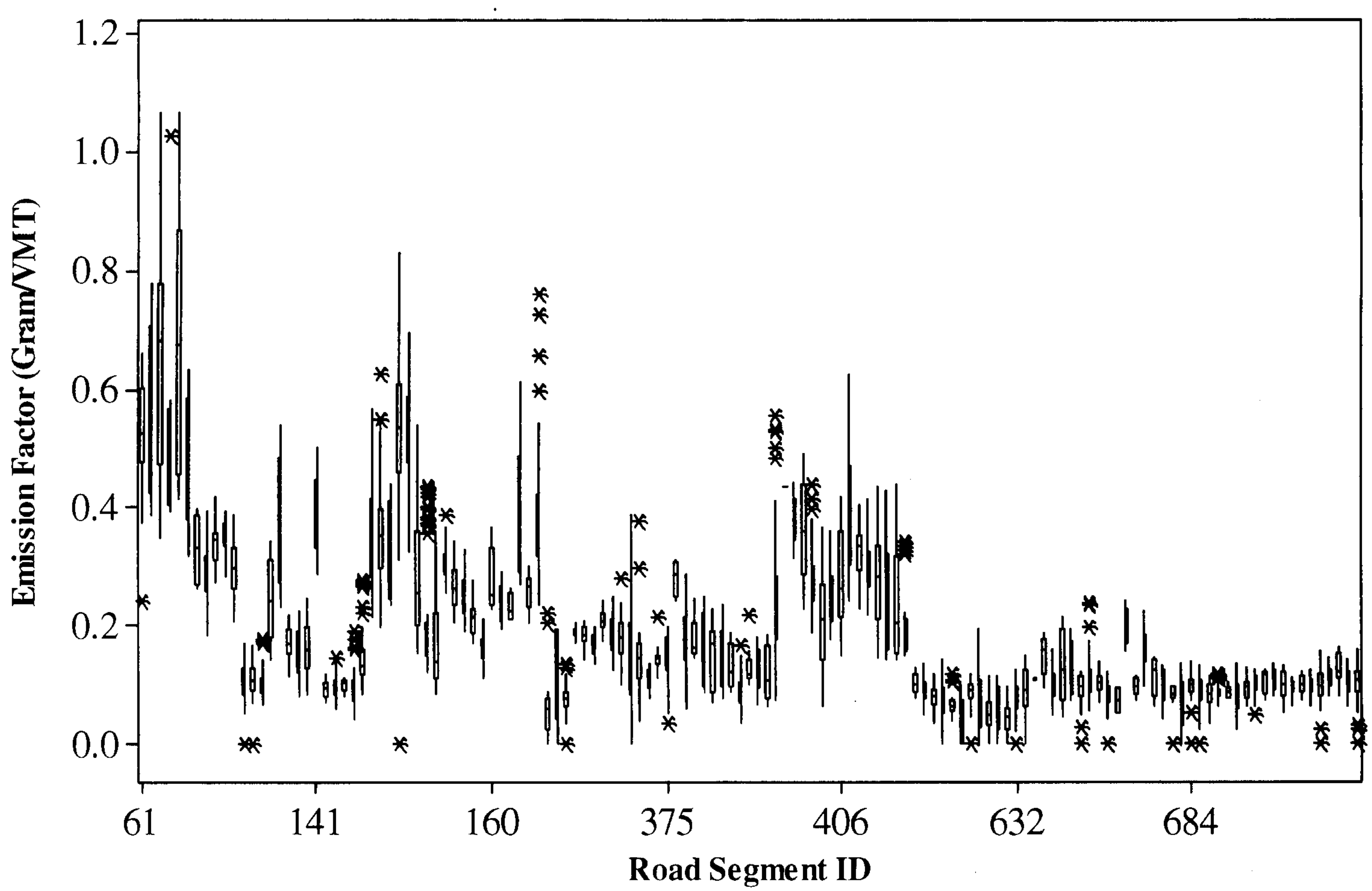


Figure 3.9 Emission Factor Box-plot Chart for Minor Arterial Segments

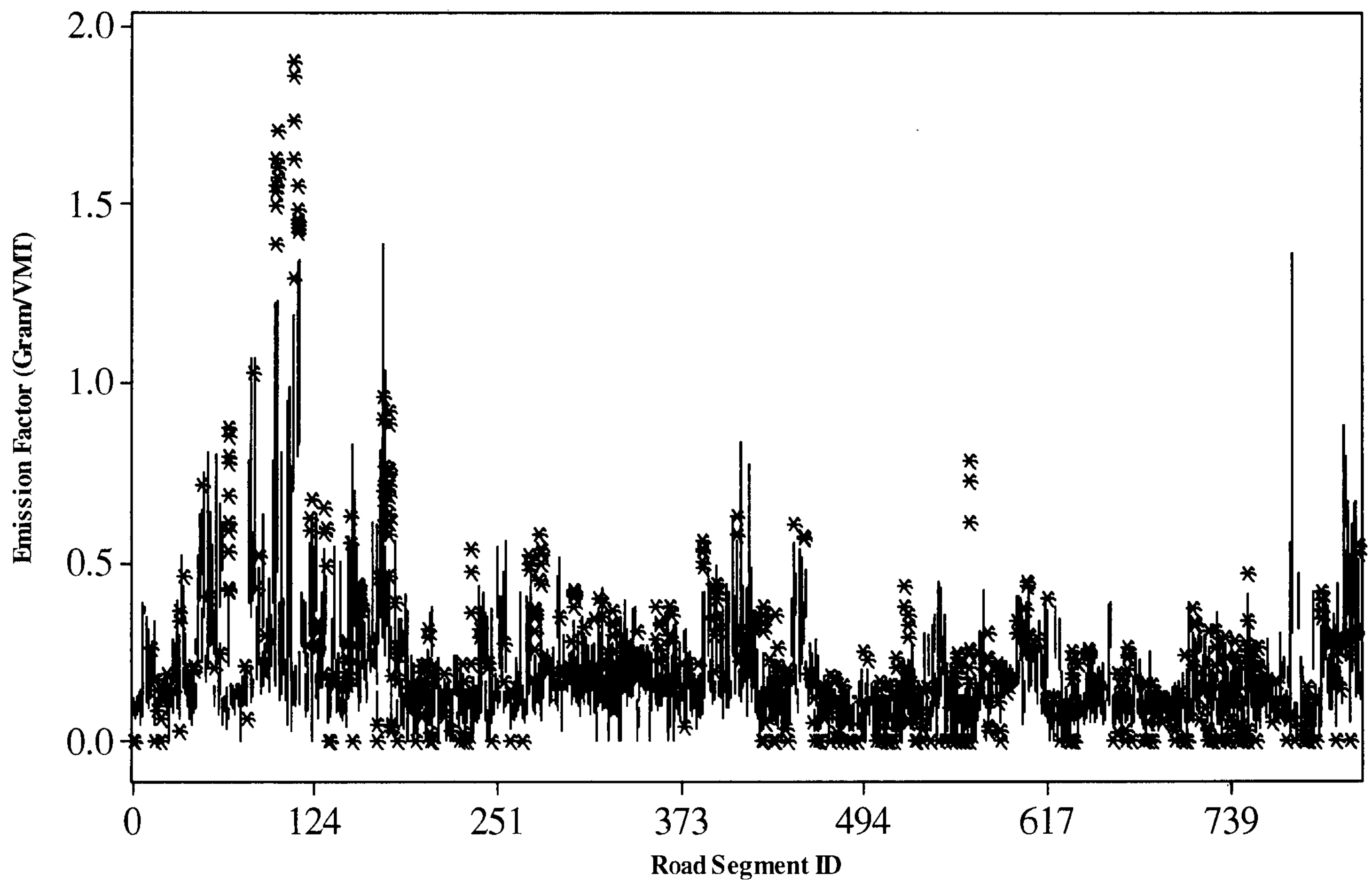


Figure 3.10 Emission Factor Box-plot Chart for Major Arterial Segments

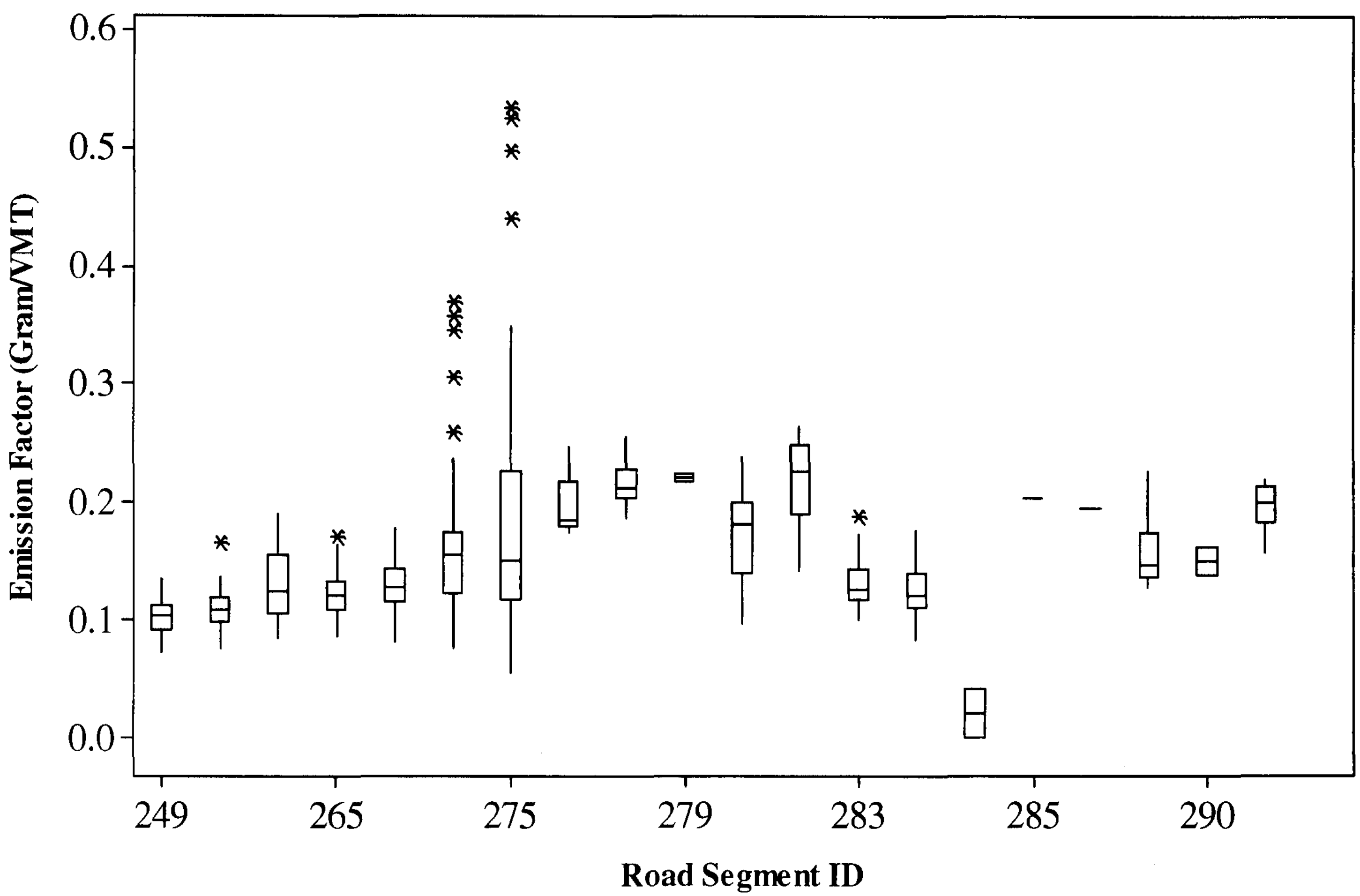


Figure 3.11 Emission Factor Box-plot Chart for Freeway Segments

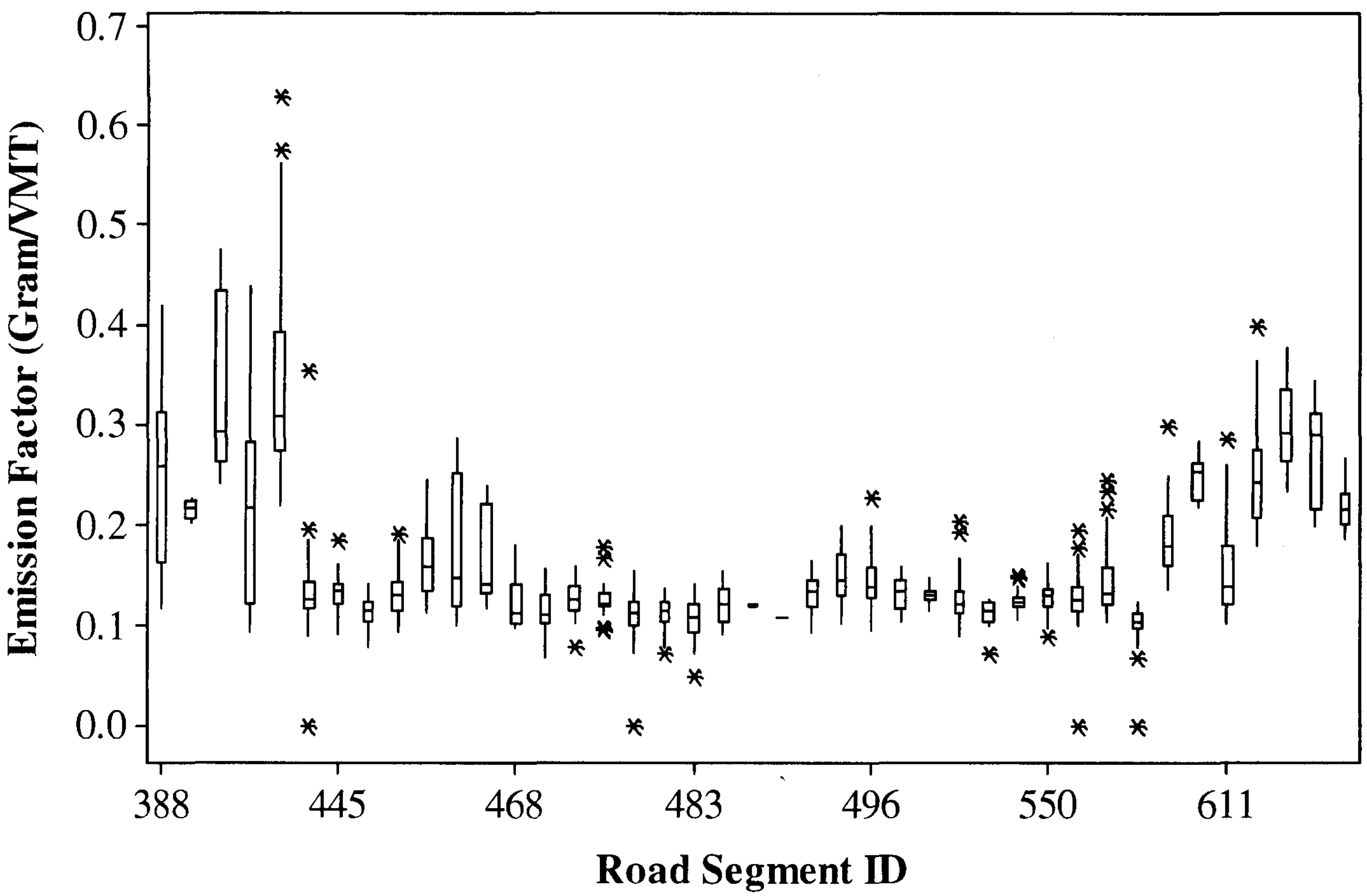


Figure 3.12 Emission Factor Box-plot Chart for Interstate Segments

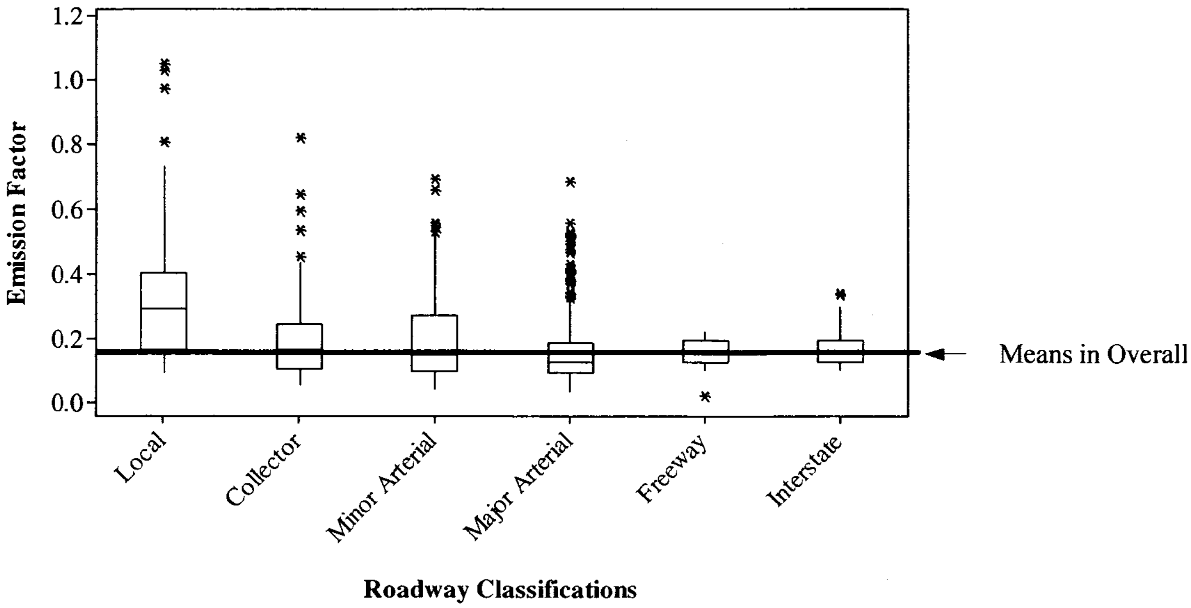


Figure 3.13 Emission Factor Box-Plot Chart of Six Roadway Classifications

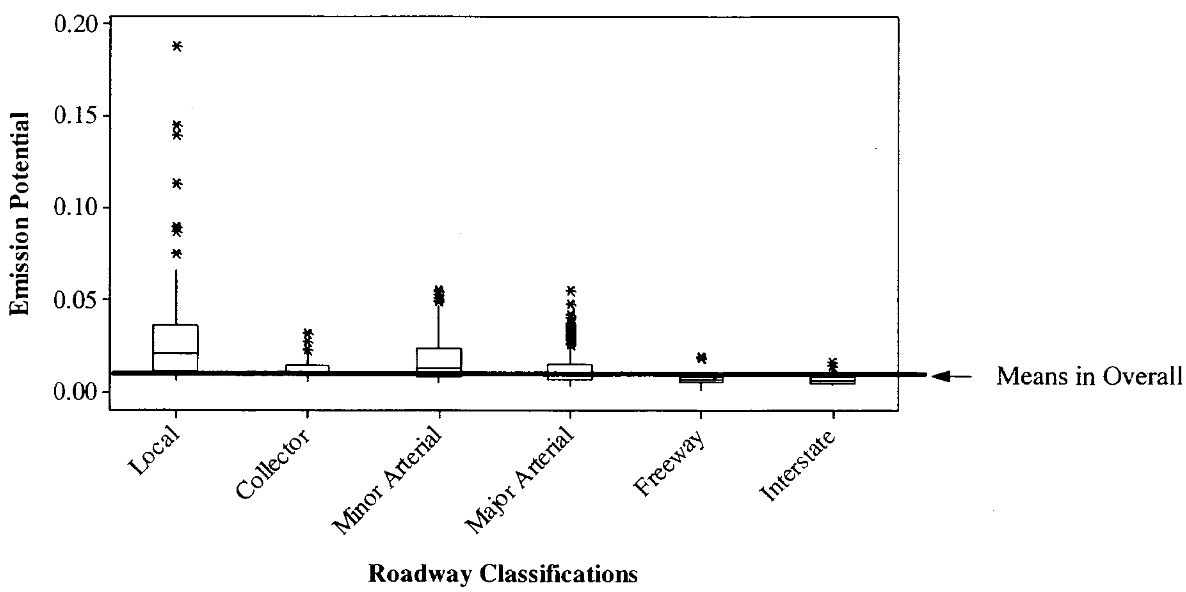


Figure 3.14 Emission Potential Box-Plot Chart of Six Roadway Classifications

3.4 Consistency of Three Days of Emission Data

In theory, dust emissions distributed continuously on every point spatially over each road segments. Since the emission measurements are available only on discrete points over road segments, the more emission measurement data points available on each road segment, the closer the data points would represent the theoretical distribution of dust emission over space. Thus, the dust emission measurements of multiple days are combined, assuming that they are consistent over these three data. Tables 3.2 and 3.3 show the statistics for emission factor and emission potential, respectively, that were collected in these four days. To make sure that the data to be combined are consistent between these different days, an ANOVA test was performed to test whether the emission measurements are significantly different between days. Since the dust emission levels of different roadway classifications don't appear to be the same (see Table 3.4), the ANOVA test also include testing whether the emission levels are significantly different between roadway classifications. Note that the last day of data was not considered because they are obviously different from the first three days of data.

According to the ANOVA test, the null hypotheses can be formulated as: there are no significant differences among the means of rows (roadway classifications) and among the means of columns (days). The alternative hypotheses can then be that there is a significant difference between the means of rows and between the means of columns. The statistics for the tests can be written as follows:

$$F_i = \frac{MS_i}{MS_E} \tag{3.1}$$

where MS_i denotes the mean squares of a chosen emission measure for row, column, and error, respectively (i.e., i = row, column, or error). These mean squares of emission measures can be derived by using the following formula:

$$MS_i = \frac{SS_i}{df_i} \tag{3.2}$$

where SS_i represents the sum of squares of emission measures for rows, columns and error, correspondingly. df denotes the degrees of freedom. There is a critical value for the F statistic for a given level of confidence. If the value of the F statistic is greater than the critical value, the null hypothesis will be rejected and the alternative hypothesis will be accepted.

Tables 3.5 and 3.6 list the results of the ANOVA tests at the 0.05 significance level. The values of F statistics computed for the roadway classification are greater than the critical value of F. The columns with the heading Count in Tables 3.5 and 3.6 represent the number of emission data included for each row and column. The columns with the heading "Sum" are the total sum of emission data included for each row and column. The results in these two tables indicate that the emission levels, regardless of measured in terms of emission factor or emission potential, are significantly different between roadway classifications. As far as the test on different days of emission data, the results show that the values of F statistics for days are smaller than the critical value of F. It implies that the emission data of different days are not significantly different. This result verifies that these three days of data can be combined for analysis.

Table 3.2 Means and Variance of Dust Emission Factor for Different Roadway Classifications in Three Days

Roadway Classifications	2/14/05		2/15/05		2/16/05		2/17/05	
	Means	Variance	Means	Variance	Means	Variance	Means	Variance
Local	0.3554	0.0813	0.3589	0.0759	0.3762	0.1364	0.3479	0.0782
Collector	0.1936	0.0122	0.2100	0.0338	0.1941	0.0227	0.1855	0.0205
Minor Arterial	0.2311	0.0232	0.1957	0.0250	0.1758	0.0166	0.1812	0.0194
Major Arterial	0.1646	0.0101	0.1601	0.0132	0.1544	0.0128	0.1420	0.0078
Freeway	0.1563	0.0017	0.1422	0.0027	0.1741	0.0039	0.1354	0.0012
Interstate	0.1503	0.0034	0.1666	0.0038	0.2009	0.0098	0.1561	0.0021

Table 3.3 Means and Variance of Dust Emission Potential for Different Roadway Classifications in Three Days

Roadway Classifications	2/14/05		2/15/05		2/16/05		2/17/05	
	Means	Variance	Means	Variance	Means	Variance	Means	Variance
Local	0.0181	0.0013	0.0194	0.0018	0.0159	0.0026	0.0407	0.0002
Collector	0.0119	0.0001	0.0123	0.0004	0.0118	0.0003	0.0165	0.0000
Minor Arterial	0.0414	0.0002	0.0505	0.0002	0.0508	0.0001	0.0147	0.0000
Major Arterial	0.0066	0.0001	0.0065	0.0001	0.0079	0.0001	0.0103	0.0014
Freeway	0.0063	0.0000	0.0058	0.0000	0.0073	0.0000	0.0056	0.0000
Interstate	0.0152	0.0000	0.0206	0.0000	0.0199	0.0000	0.0064	0.0001

Table 3.4 Emission Factor and Emission Potential with Combined Three Days of Data

Classifications	Emission Factor		Emission Potential	
	Means	Variance	Means	Variance
Local	0.3567	0.0647	0.0459	0.0015
Collector	0.197	0.01424	0.0183	0.00019
Minor Arterial	0.1968	0.0181	0.0178	0.00015
Major Arterial	0.1585	0.01014	0.0123	0.00007
Freeway	0.1569	0.00245	0.0076	0.00002
Interstate	0.1724	0.00284	0.0068	0.00001
Total Means	0.1825		0.0153	
Total Variance	0.0157		0.00022	
Between-Variance	0.00185		0.00007	
Within-Variance	0.01385		0.00013	

Table 3.5 ANOVA: Two-Factor without Replication on Emission Factor

SUMMARY	Count	Sum	Average	Variance		
Local	3	1.0905	0.3635	0.0001		
Collector	3	0.5977	0.1992	8.7E-05		
Minor Arterial	3	0.6026	0.2001	0.0008		
Major Arterial	3	0.4791	0.1597	2.61E-05		
Freeway	3	0.4726	0.1575	0.0003		
Interstate	3	0.5178	0.1726	0.0007		
Day 1	6	1.2513	0.2086	0.0061		
Day 2	6	1.2335	0.2056	0.0062		
Day 3	6	1.2755	0.2125	0.0067		
Source of Variation	SS	df	MS	F	P-value	F critical
Rows (Roadway Classifications)	0.0913	5	0.0183	48.8207	1.05E-06	3.3258
Columns (Days)	0.0001	2	<0.0001	0.1980	0.8235	4.1028
Error	0.0037	10	0.0004			
Total	0.0952	17				

Table 3.6 ANOVA: Two-Factor without Replication on Emission Potential

SUMMARY	Count	Sum	Average	Variance		
Local	3	0.0534	0.0178	3.1E-06		
Collector	3	0.0361	0.0120	6.3E-08		
Minor Arterial	3	0.1426	0.0475	2.9E-05		
Major Arterial	3	0.0209	0.0070	5.9E-07		
Freeway	3	0.0194	0.0065	5.5E-07		
Interstate	3	0.0556	0.0185	8.8E-06		
Day 1	6	0.0994	0.0166	1.7E-04		
Day 2	6	0.1151	0.0192	2.7E-04		
Day 3	6	0.1135	0.0189	2.7E-04		
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F critical*
Rows (Roadway Classifications)	0.0035	5	0.0007	118.4880	0.0000	3.3258
Columns (Days)	0.0000	2	0.0000	2.1034	0.1728	4.1028
Error	0.0001	10	0.0000			
Total	0.0036	17				

CHAPTER 4

DATA ANALYSIS

4.1 Determination of the Number of Sampling Sites

To determine the number of sampling sites for each roadway classification, it is necessary to derive the proportion of road segments among roadway classifications W_h. It was noticed that the GIS map from Clark County does not include road segments for freeways, major arterials and minor arterials; while that from the RTC does not include road segments for Local roads. The types of roadway segments available in the Clark County map and the RTC map are listed in Table 4.1. To derive the proportion of road segments for all six roadway classifications, the mileage data of roadway classification was obtained from the Federal Highway Administration (FHWA, 2004), and provided in Table 4.2. To convert the mileage data to the number of road segment data, average lengths of road segments in different roadway classification was obtained based on the Clark County GIS map and the tour run by the vehicle of mobile sampling technology. The average lengths of road segments chosen for this study are marked in Table 4.1. Given the average lengths of road segments and the total mileage of all the roadway classifications, the numbers of road segments can be derived for all the roadway classifications using Equations (4.1).

Table 4.1 General Information from Three Sources for Roadway Classifications

Roadway Classifications	Road Segments on the Clark County Map		Road Segments on the Map for RTC Travel Demand Model		Road Segment on the Tour (ft)	
	Total (mi)	Average (ft)	Total (mi)	Average (ft)	Total	Average
Centroid Connector	N/A	N/A	862	1,648	3,434	N/A
Collector	490	691	290	1,651	46,256	506
Expressway	N/A	N/A	8	5,304	N/A	N/A
Expressways	N/A	N/A	0	169	N/A	N/A
External Links	N/A	N/A	66	18,429	N/A	N/A
Freeway	N/A	N/A	79	2,987	36,666	1,098
Interstate	370	2,838	122	3,587	69,810	984
Major Arterial	N/A	N/A	585	1,482	298,437	681
Minor Arterial	N/A	N/A	289	1,559	81,081	479
Ramp	117	787	109	1,632	14,270	792
System to System Ramp	N/A	N/A	12	1,849	1,289	N/A
Local	4,073	398	N/A	N/A	N/A	476
Major Street	715	649	N/A	N/A	N/A	N/A
State Highway	291	2,603	N/A	N/A	N/A	N/A
Total	6,058		2,426		551,247	

$$N_h = \frac{T_h}{D_h} \tag{4.1}$$

where N_h is the total number of road segments in the roadway classification *h*, T_h = total mileage from the Federal Highway Administration for roadway classification *h*, and D_h = Average length of the road segments for roadway classification *h*. For local roads, collectors, and interstates, the average lengths from the Clark County map are used, while those of minor arterials, major arterials, and freeways, the average lengths from the tour are used. The proportion of road segments can be calculated using the following equation:

$$W_h = \frac{\frac{T_h \times 5280}{D_h}}{\sum_{l=1}^{L} \frac{T_l \times 5280}{D_l}} \tag{4.2}$$

where L = the number of the roadway classifications. The result for the numbers and proportions of road segments are listed in Table 4.2.

Based on Equation (4.3), the standard deviation of dust emission factor and potential can be calculated.

$$\hat{\sigma}_h = \sqrt{\frac{1}{n_h - 1} \sum_{i=1}^{n_h} (E_{hi} - \bar{E}_h)^2} \tag{4.3}$$

where

n_h = the total number of the dust emission data points for roadway classification *h*,

E_{hi} = the *i*-th dust emission measures for roadway classification *h*,

E_h = the average emission measures for roadway classification *h*.

The results for emission factor and emission potential are listed in Tables 4.3 and 4.4 individually.

Table 4.1 W_h Value for Six Roadway Classifications

Roadway Classifications	Mileage from FHWA (mi)	Average Length of Road Segments (ft)	Number of Segments	Proportion of Segments W_h
Local	4,084	398*	39	0.809
Collector	753	691*	139	0.086
Minor Arterial	230	479**	133	0.038
Major Arterial	528	681**	410	0.061
Freeway	52	1098**	20	0.004
Interstate	80	2,838*	67	0.002

* Data Source: Clark County Map

** Data Source: Tour

Given the proportion of road segments among roadway classifications in Table 4.2 and the estimated standard deviation of emission factors in Table 4.3, the number of sampling sites is calculated for different budgets based on PM_{10} emission factor using Equation (2.1). The number of sampling sites is listed in Table 4.3. It can be seen from Table 4.3 that local roadway classification needs to have the dominant number of sampling sites than others. It is due to the fact that this roadway classification includes the most number of street links. As the increase of budget, the added sampling sites primarily go to the local roadway classification. It has been noticed that the number of sample sites recently used in a study (Table 4.5) is distributed differently than that calculated based on the PM_{10} emission factor data. Specifically, it is the collector roadway classification that is sampled more than others. The difference between these two numbers of sampling sites draws attention to verifying the number of sampling sites before implementation.

The number of sampling sites is also derived for all the roadway classifications based on a different emission measure: emission potential. The results are presented in Table 4.4. For comparison, the results for these two emission measures are displayed in Figure 4.1. It can be seen that the results for these two emission measures are quite similar.

Table 4.2 Sample Size under Different Budgets Based on PM10 Emission Factor

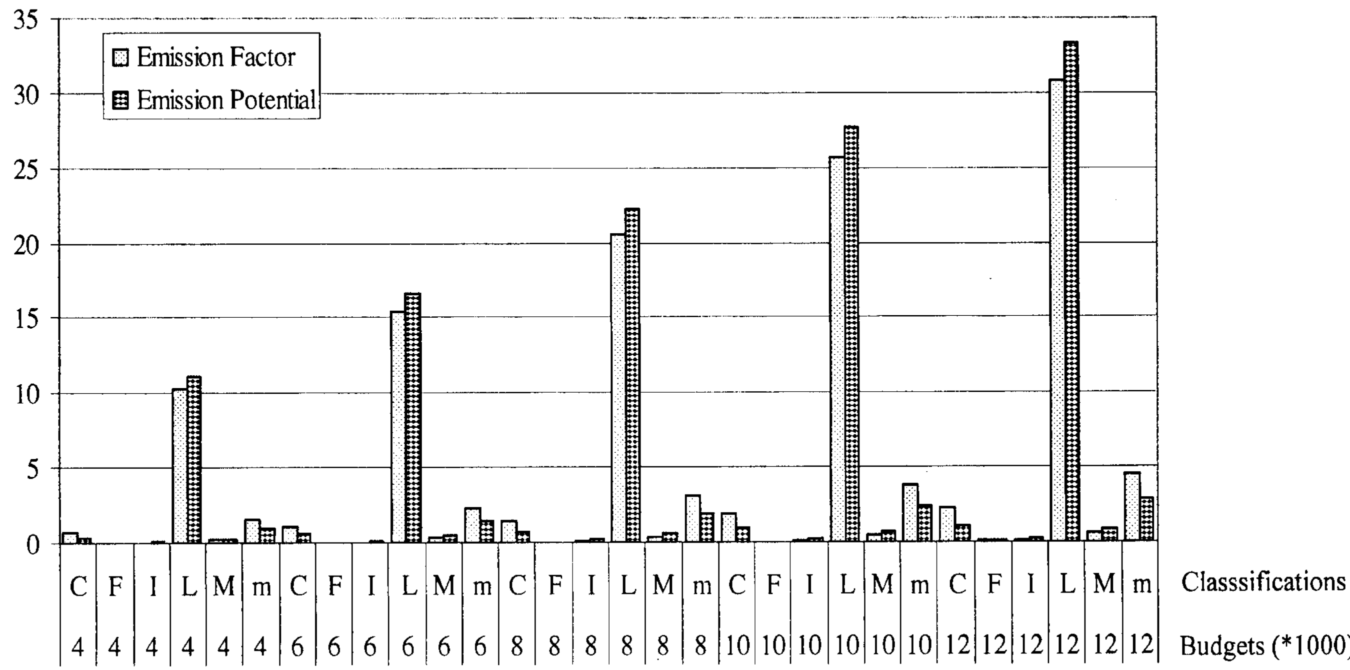
Roadway Classifications	Proportion of Road Segments W_h	Standard Deviation of PM_{10} Emission Factor $\hat{\sigma}_h$	Budget				
			\$4,000	\$6,000	\$8,000	\$10,000	\$12,000
			Unit cost=\$312,00				
Local	0.809	0.01225	10.28	15.42	20.56	25.70	30.84
Collector	0.086	0.00837	0.75	1.12	1.49	1.87	2.24
Minor Arterial	0.038	0.03870	1.53	2.29	3.05	3.81	4.58
Major Arterial	0.061	0.00316	0.20	0.30	0.40	0.50	0.60
Freeway	0.004	0.00447	0.02	0.03	0.04	0.05	0.06
Interstate	0.002	0.01378	0.03	0.04	0.06	0.07	0.09
Total			12.80	19.20	25.60	32.00	38.40

Table 4.4 Sample Size under Different Budgets Based on Emission Potential

Roadway Classifications	Proportion of Road Segments W_h	Standard Deviation of PM_{10} Emission Factor $\hat{\sigma}_h$	Budget				
			\$4,000	\$6,000	\$8,000	\$10,000	\$12,000
			Unit cost=\$312,00				
Local	0.855	0.13403	11.09	16.64	22.18	27.73	33.27
Collector	0.0374	0.10056	0.36	0.55	0.73	0.91	1.09
Minor Arterial	0.0389	0.25107	0.95	1.42	1.89	2.36	2.84
Major Arterial	0.0564	0.05287	0.29	0.43	0.58	0.72	0.87
Freeway	0.0043	0.04829	0.02	0.03	0.04	0.05	0.06
Interstate	0.008	0.11889	0.09	0.14	0.18	0.23	0.28
Total			12.80	19.20	25.60	32.00	38.40

Table 4.5 Sample Size for Each Three Roadway Classifications – UNLV AP-42 Study: Phase II (Rodrigues and James, 2005)

Roadway Classification	Sample Size	Sample Location
Arterial	7	UNLV006, UNLV008, UNLV011, UNLV012, UNLV016, UNLV018, UNLV020
Collector	11	UNLV001, UNLV002, UNLV005, UNLV007, UNLV010, UNLV015, UNLV017,UNLV021, UNLV022, UNLV023, UNLV024
Local	6	UNLV003, UNLV004, UNLV009, UNLV013, UNLV014, UNLV019



L --- Local; C --- Collector; m --- Minor Arterial; M --- Major Arterial; F --- Freeway; I --- Interstate

Figure 4.1 Sample Sizes Based on PM_{10} Emission Factor and Potential

4.2 Analysis of Sampling Plots for AP-42

In this study, a Monte Carlo simulation model was adopted to investigate the impact of the number of sampling plots and their length in the AP-42 method on the accuracy of PM_{10} emission estimation. In the investigation, the number of sampling plots is varied from 1 to 10, while the length of the plots changes from 10 ft to 50 ft. Given these ranges, the maximum length required for a road segment to be included in the simulation is $10 \times 50 = 500$ ft. Because the locations of these plots are randomly generated in the simulation, it may not be convenient in programming if this minimum length is adopted. Thus, 600 ft was used as the minimum length for a road segment to be included in the simulation. Table 4.6 shows the number of road segments that are considered for each roadway classification when 600 ft is adopted for the simulation.

Table 4.6 Number of Road Segments for Roadway Classification in Simulation

Roadway Classification	Average Length (ft) of Segments	Total Number of Segments	Number of Segments > 600 ft	Percentage of Segment > 600 ft
Local	398	39	11	28%
Collector	691	139	46	33%
Minor Arterial	479	133	33	25%
Major Arterial	681	410	218	53%
Freeway	1098	20	12	60%
Interstate	2838	67	42	62%

The number of configurations (also called simulation runs) to be generated in the simulation for the plots given their number and length is another important input that needs to be determined for the simulation model. In this study, the numbers 30 and 100 were tried for this number and the distributions of the PM_{10} emission factors estimated based on the generated configurations of the plots were observed. Figures 4.2 and 4.3 present the distributions for the local roadway classification for 30 and 100 runs, respectively. The distributions for other roadway classifications can be found in Appendix A. From Figures 4.2 and 4.3, it can be seen that the estimated emission factors are similarly distributed for 30 and 100 runs at the same number and length of plots. Thus, the number 30 was adopted in this study which can save significant computational time.

With the determination of road segments that are included in the simulation and the simulation runs, simulations were conducted. The results are listed from Tables 4.7 to 4.12, each for one roadway classification. The tables indicate that P is varied from 1 to 10 and D changes from 10 to 50. The measure emission factor difference which is the objective function in Equation (2.6) is the output from the simulation. It is calculated using the formula below:

$$difference = \frac{\sum_{i=1}^{n}\left[\overline{EF}_{i,E}(P,D) - \overline{EF}_{i,T}\right]}{n} \tag{4.4}$$

To better view the pattern of the measure versus the combination of P and D, the results in Tables 4.7 to 4.12 are presented from Figure 4.4 to Figure 4.9. Figure 4.4 displays a clear pattern for the measure of Difference for local roads. When more than seven plots (each more than 20 ft) are adopted in the AP-42 sampling method, the

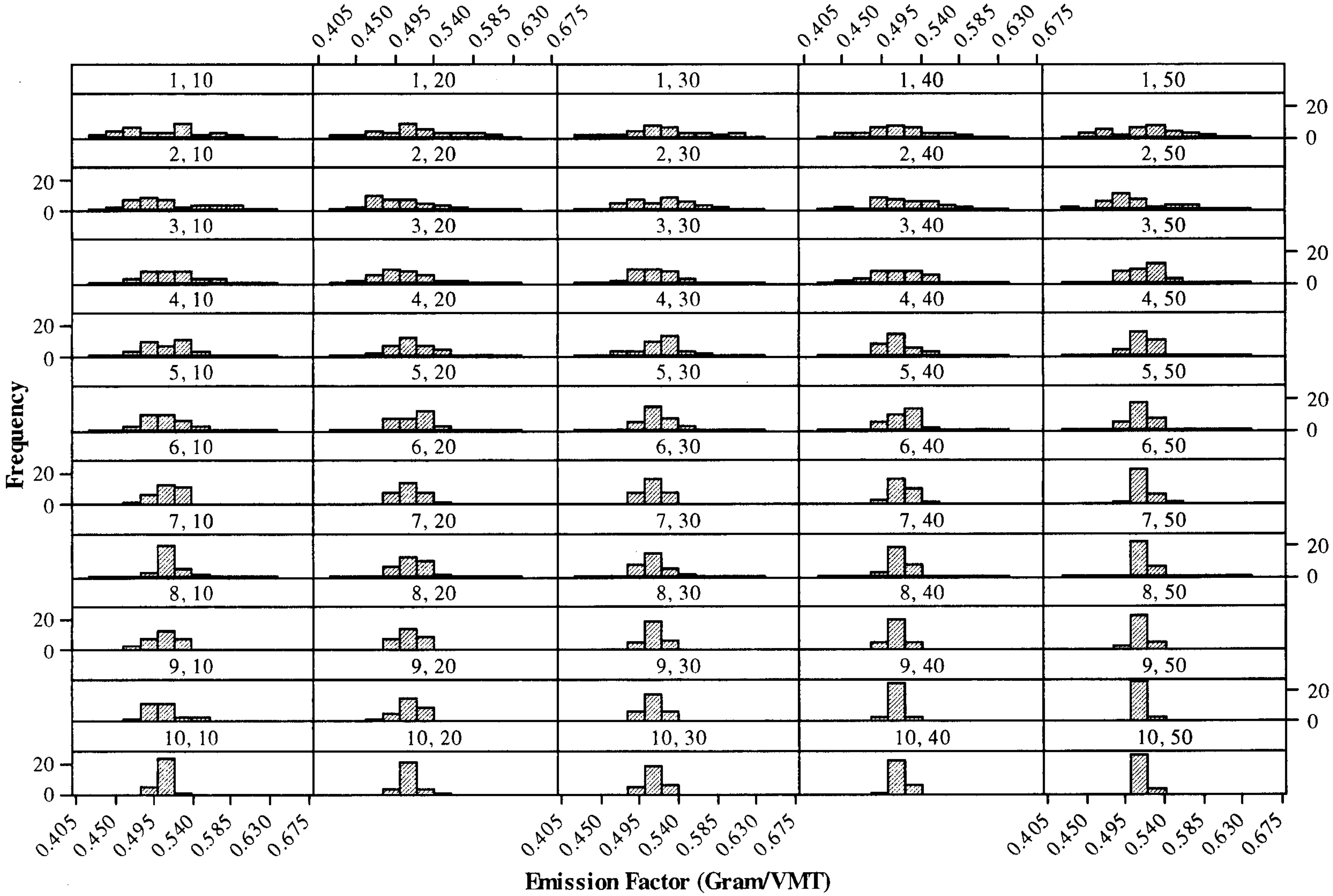


Figure 4.2 Distribution of the Calculated Emission Factor for Local Roads when the Simulation Run is 30

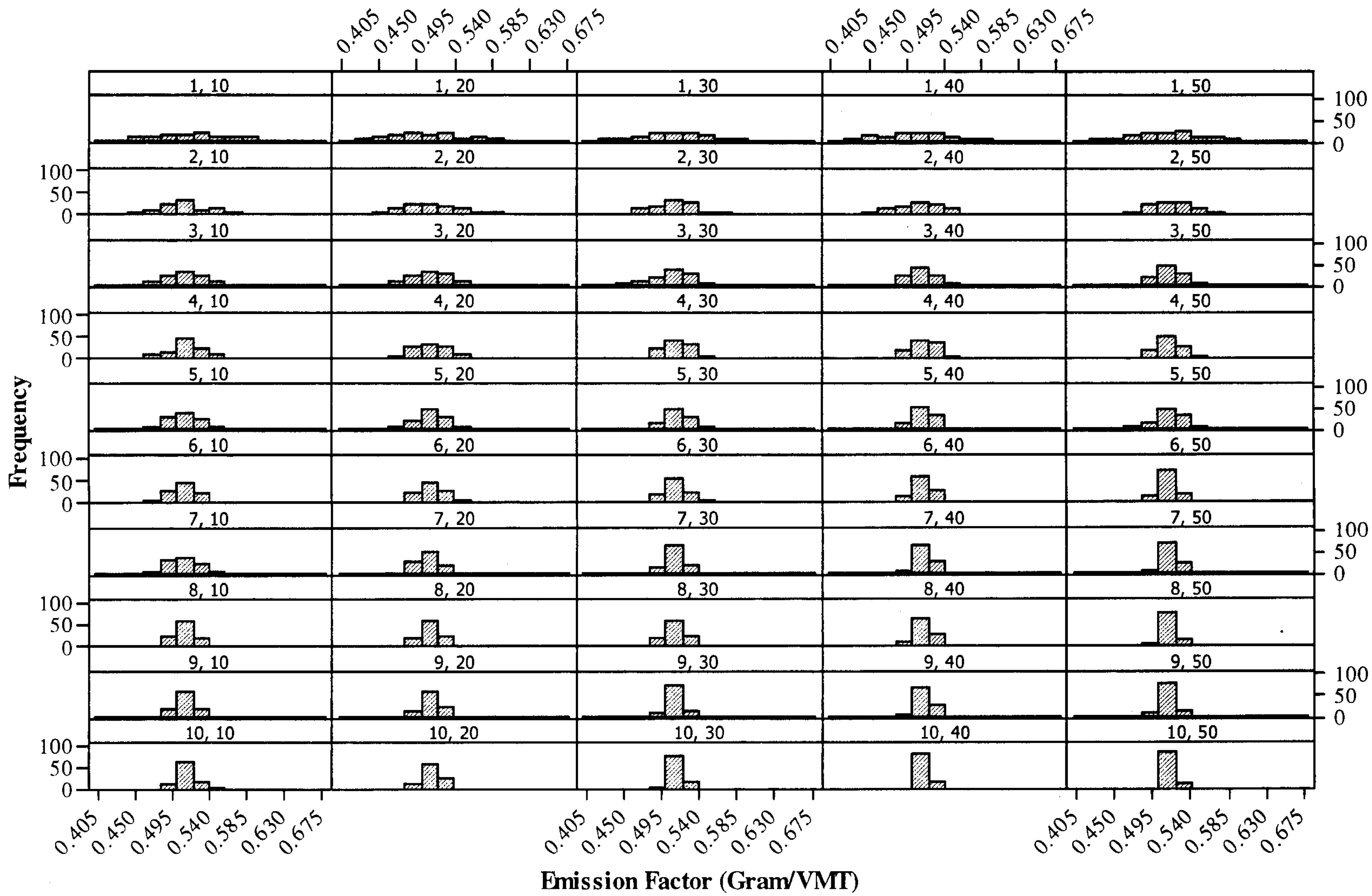


Figure 4.3 Distribution of the Calculated Emission Factor for Local Roads when the Simulation Run is 10

estimated emission factor could be significantly closer to the true emission factor. Similarly, it can be observed that the estimated emission factor could be significantly closer to the true emission factor for minor arterials (Figure 4.6) when more than one plot (each more than 20 ft) is used in the sampling. For freeways (Figure 4.8), better estimation of emission factor can be obtained only when there are at least three plots each with 50 ft. Such clear patterns cannot be found for collectors, major arterials and interstate highways. The current AP-42 method requires three plots, each at least ten feet long. The observations particularly for these three roadway classifications (local, minor arterials and interstate highways) imply that the AP-42 requirement on the plot number and length may not be appropriate. Note that simulations were not conducted for three plots each less than 10 ft which is the requirement for the AP-42 method. It is believed that the accuracy of the PM_{10} emission factor would be smaller with these requirements, particularly smaller than the cases simulated in this study where the lengths of the plots are longer.

Table 4.7 Monte Carlo Results for Local Roads

		Number of Plots (P)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (D)	10	-0.0047	0.0011	0.0046	0.0040	0.0026	0.0056	0.0041	0.0003	-0.0009	0.0005
	20	0.0132	-0.0071	-0.0045	0.0059	0.0099	0.0030	0.0072	0.0057	0.0042	0.0035
	30	0.0176	0.0115	0.0050	0.0123	0.0068	0.0036	0.0029	0.0034	0.0036	0.0044
	40	0.0037	0.0139	0.0035	0.0053	0.0099	0.0078	0.0066	0.0024	0.0043	0.0054
	50	0.0084	-0.0041	0.0088	0.0040	0.0039	0.0070	0.0070	0.0034	0.0059	0.0070

Table 4.8 Monte Carlo Results for Collectors

		Number of Plots (P)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (D)	10	-0.0061	-0.0048	-0.0055	-0.0068	-0.0035	-0.0037	-0.0049	-0.0047	-0.0050	-0.0049
	20	-0.0071	-0.0033	-0.0039	-0.0048	-0.0059	-0.0050	-0.0050	-0.0052	-0.0043	-0.0046
	30	-0.0069	-0.0035	-0.0043	-0.0054	-0.0039	-0.0051	-0.0048	-0.0046	-0.0049	-0.0047
	40	-0.0081	-0.0056	-0.0056	-0.0052	-0.0042	-0.0048	-0.0046	-0.0043	-0.0052	-0.0039
	50	-0.0041	-0.0051	-0.0040	-0.0046	-0.0043	-0.0048	-0.0042	-0.0046	-0.0038	-0.0039

Table 4.9 Monte Carlo Method for Minor Arterials

		Number of Plots (P)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (D)	10	-0.0006	-0.0052	-0.0049	-0.0029	-0.0026	-0.0032	-0.0032	-0.0038	-0.0031	-0.0033
	20	-0.0076	-0.0048	-0.0041	-0.0043	-0.0044	-0.0051	-0.0037	-0.0050	-0.0034	-0.0041
	30	-0.0061	-0.0041	-0.0044	-0.0040	-0.0035	-0.0033	-0.0033	-0.0040	-0.0042	-0.0033
	40	-0.0034	-0.0038	-0.0038	-0.0028	-0.0045	-0.0040	-0.0046	-0.0038	-0.0037	-0.0033
	50	-0.0069	-0.0041	-0.0039	-0.0035	-0.0039	-0.0035	-0.0040	-0.0035	-0.0037	-0.0034

Table 4.10 Monte Carlo Method for Major Arterials

		Number of Plots (P)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (D)	10	-0.0006	-0.0052	-0.0049	-0.0029	-0.0026	-0.0032	-0.0032	-0.0038	-0.0031	-0.0033
	20	-0.0001	0.0001	-0.0001	-0.0003	-0.0003	-0.0006	-0.0003	-0.0004	-0.0004	-0.0001
	30	0.0002	-0.0008	-0.0002	-0.0001	-0.0003	-0.0003	0.0000	-0.0001	-0.0001	0.0000
	40	-0.0006	-0.0004	-0.0003	0.0001	-0.0003	-0.0004	-0.0002	-0.0001	-0.0002	-0.0002
	50	0.0002	-0.0006	0.0003	-0.0004	-0.0001	0.0000	0.0000	-0.0003	0.0000	0.0000

Table 4.11 Monte Carlo Results for Freeways

		Number of Plots (*P*)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (*D*)	10	0.0005	-0.0013	-0.0005	-0.0023	-0.0012	-0.0020	-0.0009	-0.0015	-0.0013	-0.0011
	20	-0.0046	-0.0018	-0.0031	-0.0008	-0.0015	-0.0007	-0.0024	-0.0016	-0.0010	-0.0017
	30	-0.0022	-0.0017	-0.0012	-0.0007	-0.0009	-0.0022	-0.0008	-0.0005	-0.0019	-0.0021
	40	-0.0018	-0.0040	-0.0021	-0.0014	-0.0017	-0.0027	-0.0009	-0.0009	-0.0012	-0.0018
	50	-0.0028	-0.0015	-0.0020	0.0001	-0.0013	-0.0014	-0.0013	-0.0011	-0.0009	-0.0005

Table 4.12 Monte Carlo Results for Interstate Highways

		Number of Plots (*P*)									
		1	2	3	4	5	6	7	8	9	10
Plot Length (*D*)	10	0.0000	0.0005	-0.0005	-0.0003	0.0007	0.0006	0.0001	0.0002	-0.0001	0.0002
	20	0.0004	-0.0007	0.0000	0.0010	0.0008	0.0004	0.0007	0.0001	0.0003	0.0002
	30	-0.0005	-0.0006	-0.0003	-0.0001	0.0003	0.0012	0.0006	0.0002	0.0001	0.0002
	40	0.0017	0.0008	0.0001	0.0004	0.0008	0.0003	-0.0003	0.0004	-0.0001	0.0007
	50	-0.0003	0.0006	0.0007	-0.0006	-0.0005	0.0002	0.0003	0.0002	0.0001	0.0000

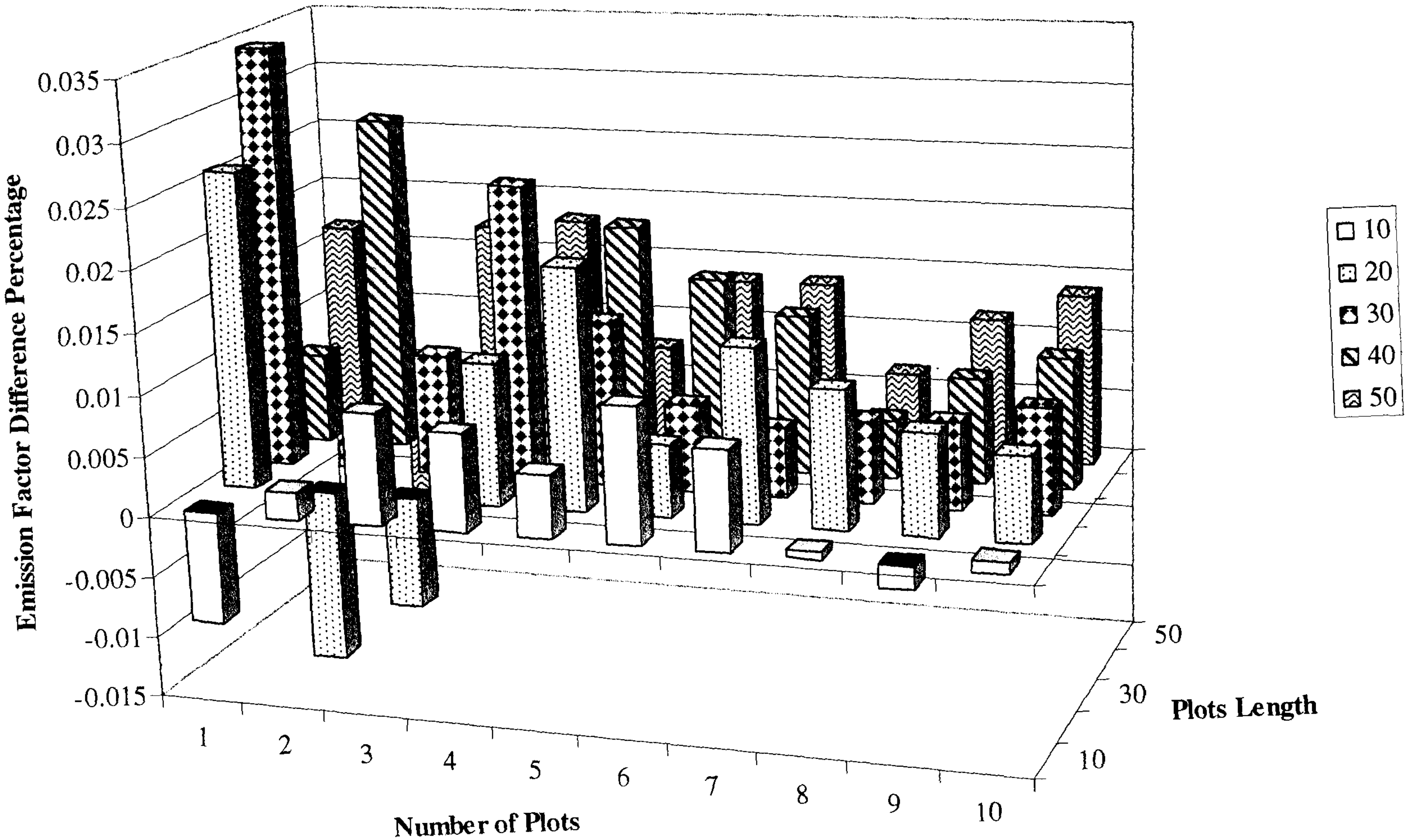


Figure 4.4 Emission Factor Difference for Local Roads

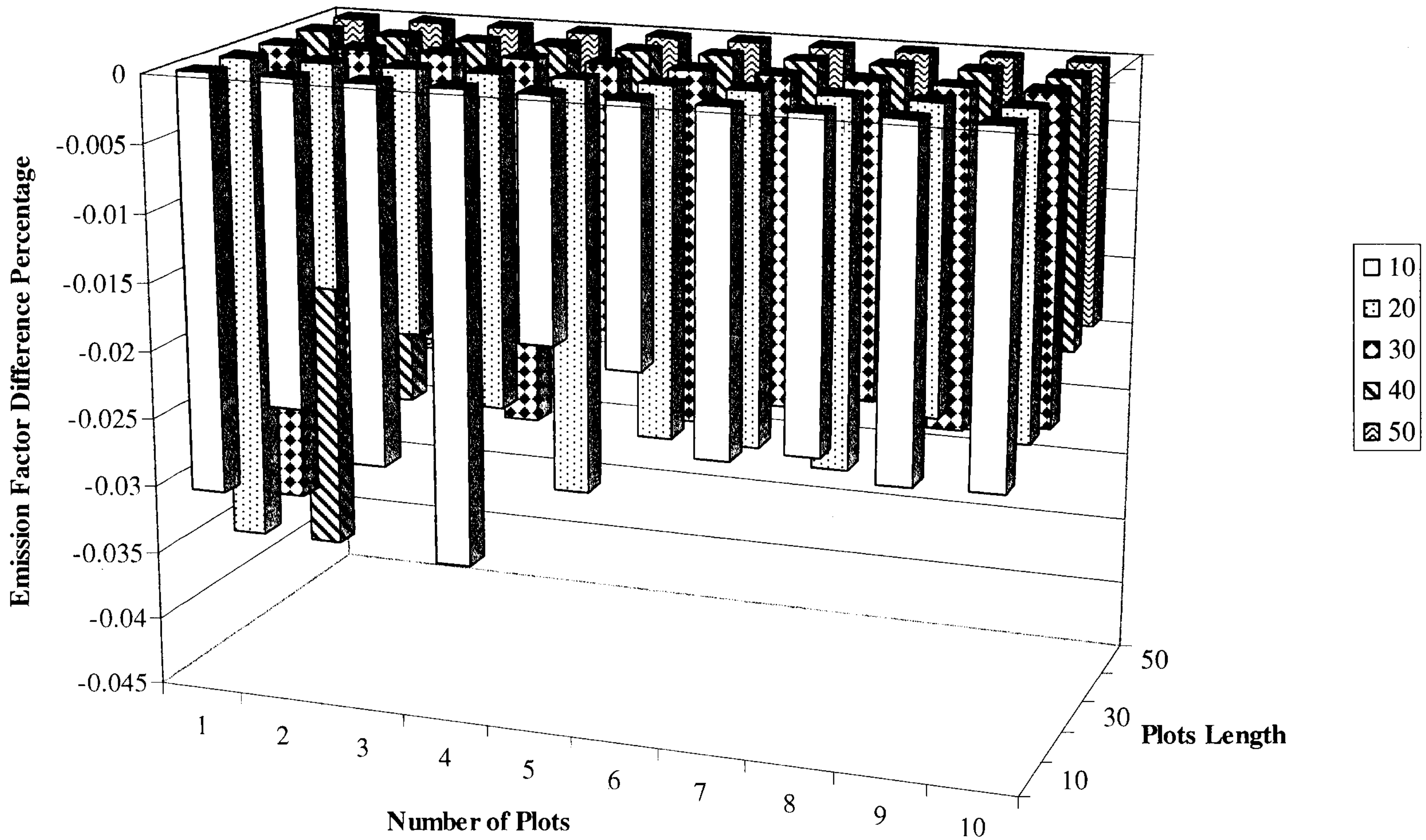


Figure 4.5 Emission Factor Difference for Collectors

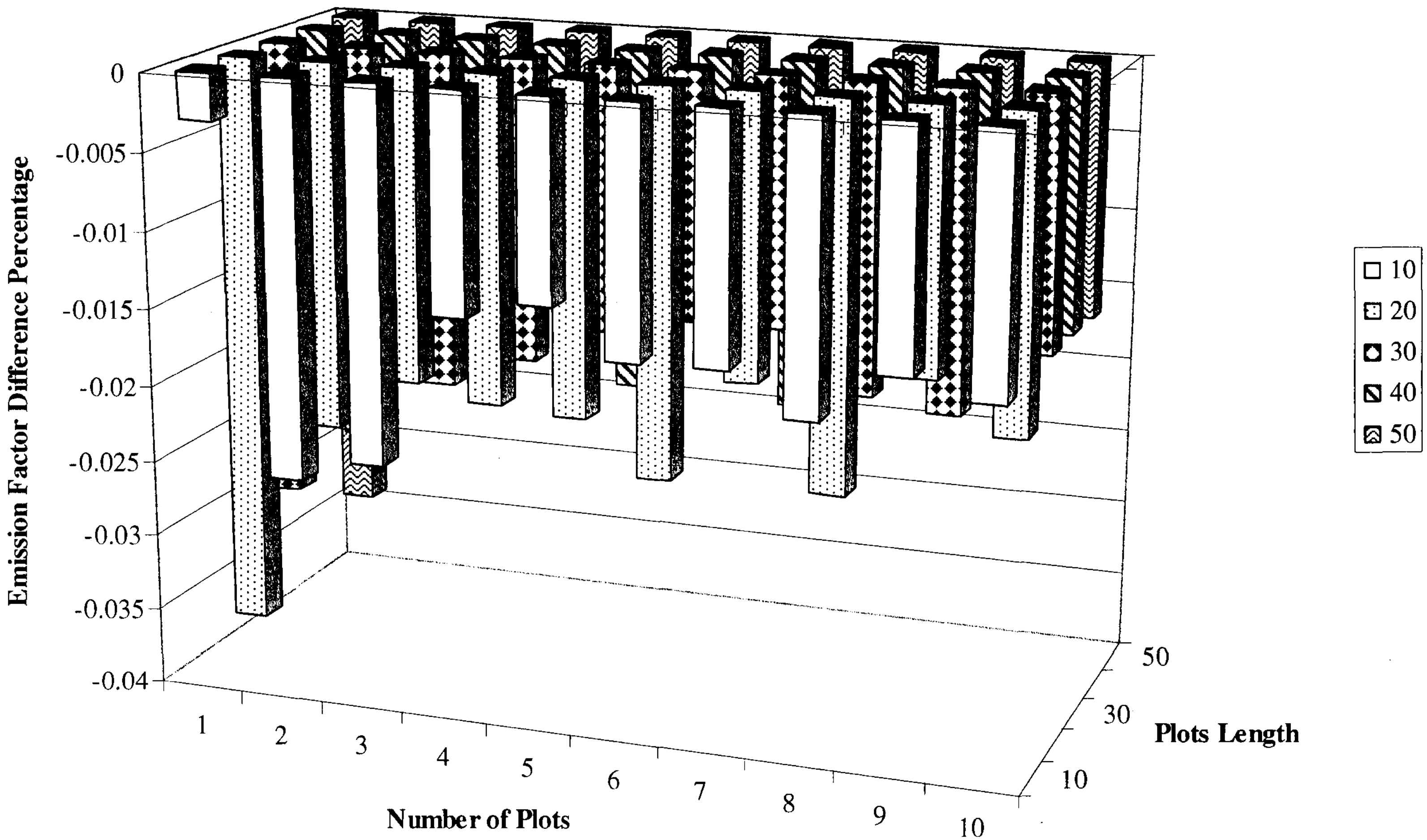


Figure 4.6 Emission Factor Difference for Minor Arterials

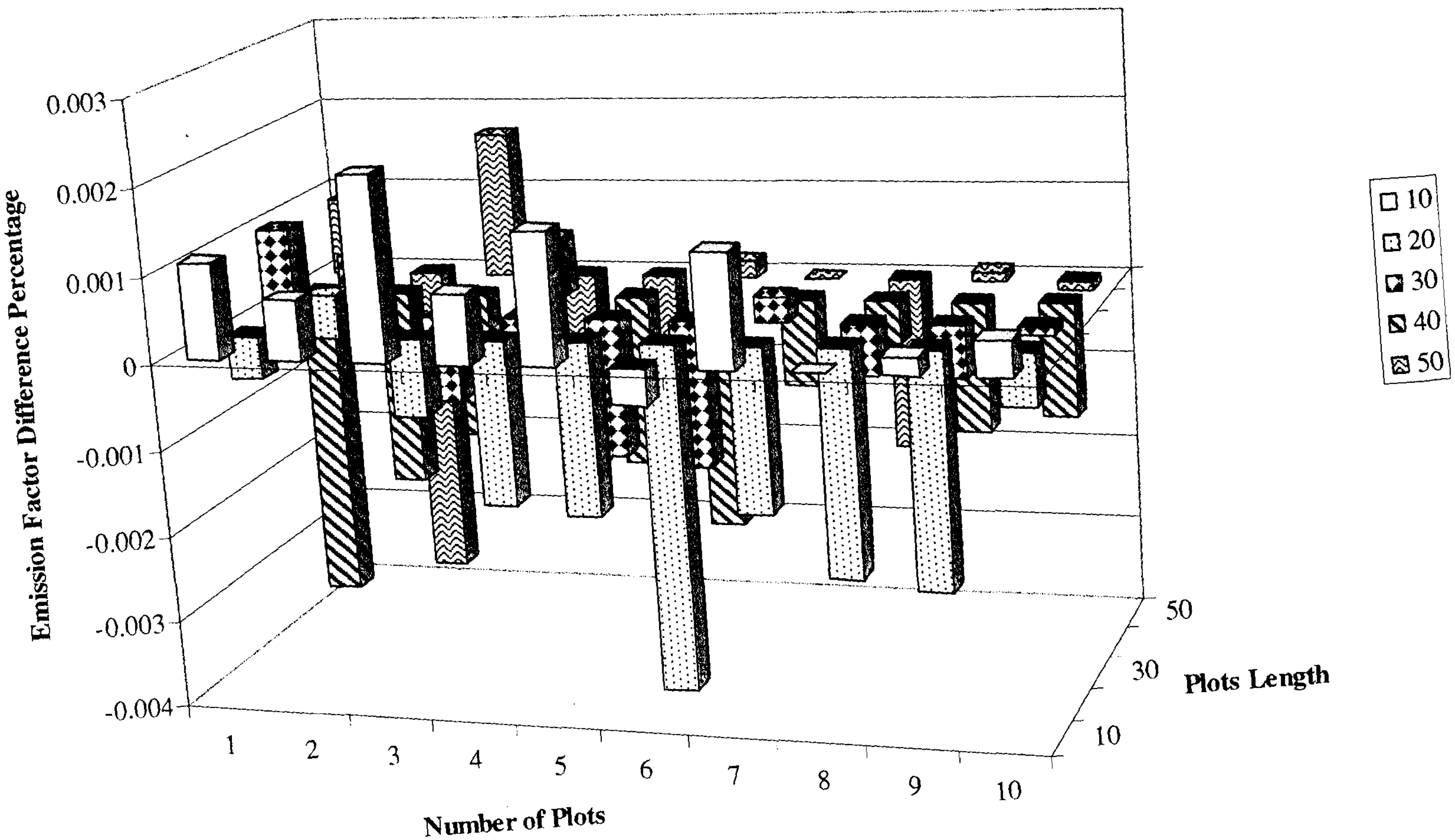


Figure 4.7 Emission Factor Difference for Major Arterials

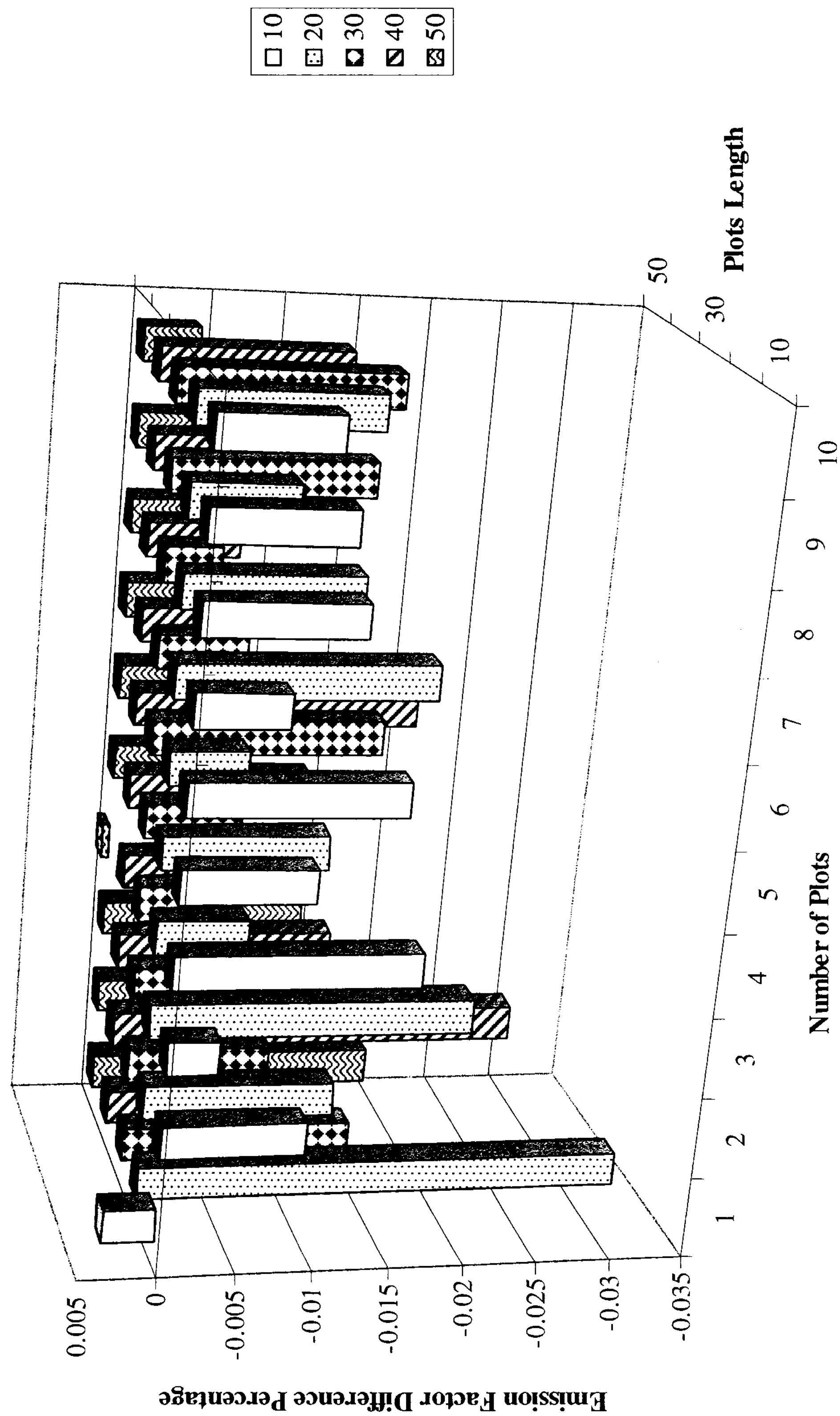


Figure 4.8 Emission Factor Difference for Freeway

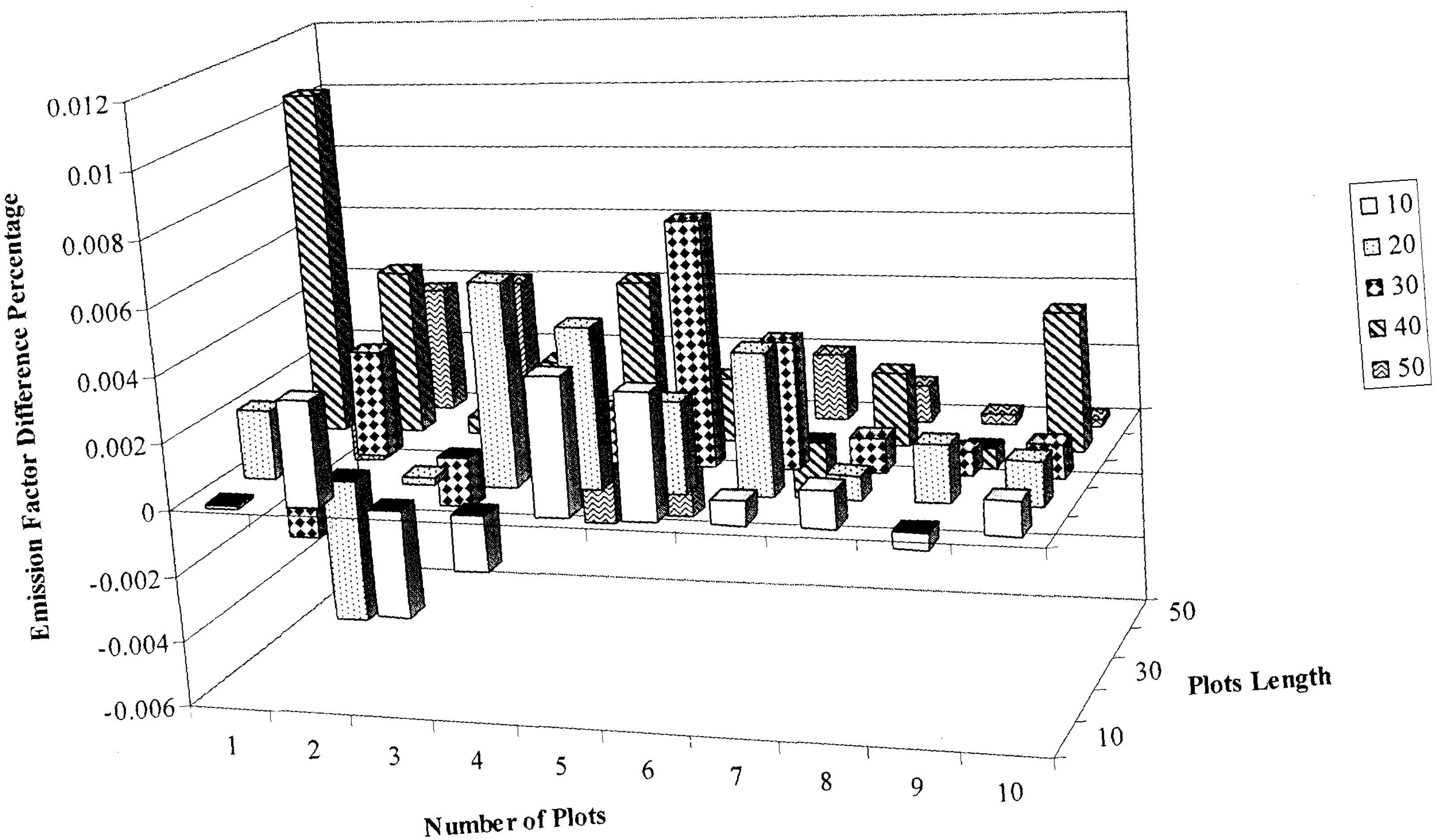


Figure 4.9 Emission Factor Difference for Interstate

CHAPTER 5

CONCLUSION AND FUTURE STUDY NEEDS

5.1 Conclusions

In this study, two fundamental issues related to the AP-42 method are investigated: sample size for dust collection sites and the number and length of plots to be adopted at a sampling site. In the past when there were no emission measurements covering the entire lengths of the substantial number of road segments, the study on these two sampling issues is not possible. Because the close-to-continuous emission measurements were made available due to the application of a mobile sampling technology for emission data collection, such a study becomes feasible.

The sample size issue is investigated based on the optimal allocation sampling method where a fixed budget is considered. The sampling formula used in this method is derived by minimizing the variance of the estimation for a variable, such as emission factor and keeping the cost of the sampling within the budget. The allocation of samples among roadway classifications directly related to the percentage of the road segments and the variance of emission measures on these segments in different roadway classifications. In this study, the emission measurements collected by the mobile sampling technology over a tour are used to derive the variance of emission measures for road segments in different roadway classifications. Percentage of mileages and average lengths of road segments of different roadway classifications are used to derive the percentage of road segments. With these two major parameters, sample sizes are calculated using the optimal allocation formula. The results indicate that the sample size of road segments for local roads dominate those of other roadway classifications. This observation is different from the sample sizes used in the current practice. Further investigation is needed to verify these sample sizes being used.

The number and length of plots for the AP-42 method was investigated by simulating the actual layouts of the plots that are specified with certain number of plots and length. The patterns of the difference between the true average of emission measurements on a road segment and the estimated average based on the simulated plots are observed. The patterns are presented as the difference versus the number of plots and length of these plots simulated in this study. It is observed that a better estimation of emission measures can be achieved when the number of plots and their lengths are larger than a certain set of thresholds. And these thresholds vary between different roadway classifications. By comparing the thresholds with the sampling specification for the AP-42 method, it can be seen that they are not consistent. It is also found that these patterns can be clearly observed only for three of the six roadway classifications considered in this study. Based on these observations, it may be tentatively concluded that the AP-42 method may not be appropriate for collecting dust emission data accurately for these roadway classifications. Mobile sampling technology may be recommended for wide application in the future.

5.2 Future Study Needs

Based on the investigation of sampling issues for the AP-42 method in this study, the following study needs are identified.

First, more advanced sampling methods may need to be explored to determine the sample size for different roadway classifications. From observing the results of the optimal allocation method, it can be found that the sample size for local road dominate those for other roadway classifications. Apparently, the local roads can be further classified into categories based on criterion such as urban or rural, business or residential, etc. The allocation based on these sub categories of local road may lead to more accurate estimation of emission measures with the same budget limit.

Second, a study is needed to investigate the acceptable error of emission measure from the AP-42 method. It was not investigated in this study on the variance of the estimation of emission measures from the optimal allocation method. However, the variance can be calculated with a given formula, and this variance can be evaluated in terms of whether it is acceptable from the perspective of EPA. If this variance is too large, more samples are needed, which in turn would require more budgets for routine sampling activities.

Third, the investigation carried in this study was based on the emission measurements collected by using the mobile sampling technology. It has been observed that sometimes the time and spatial gaps between two consecutive data points on a road segment are too big, which may cloud the characteristics of true distribution of emission measurements over a road segment. For this reason, the mobile sampling technology may need to be improved for collecting data with more dense coverage.

Fourth, a sampling method should be developed to consider the length of road segments. In this study, only the variance of the PM_{10} is incorporated in the investigation. In a matter of fact, the PM_{10} emission for an area is estimated for a whole area which involves both the average of emission on a road segment and their lengths.

REFERENCES

1 Clark County Department of Air Quality and Environmental Management (2001). PM10 State Implementation Plan for Clark County.

2 Clark County Department of Air Quality and Environmental Management. Dust. http://www.accessclarkcounty.com/daqem/aq/monitor/dust.html (Accessed on Feb. 17, 2007)

3 E.H. Pechan & Associates, Inc., (1994) PM-10 Emission Inventory Requirements Final Report. http://www.epa.gov/ttn/chief/eidocs/pm10eir.pdf (Accessed on Feb. 17, 2007)

4 G. Rodrigues and D. James. Draft report for the first quarter 2005: AP-42 Sampling. Clark County Department Air Quality and Environment Management. 2005

5 R. Gehrig, M. Hill, and B. Buchmann (2004). In-field Verification of PM10 Emission Factors of Road Traffic. STRC Conference Paper.

6 R. Y. Rubinstein (1981). "Simulation and the Monte Carlo Method", John Wiley & Sons, Inc.

7 R. Y. Rubinstein (1986). "Monte Carlo Optimization, Simulation and Sensitivity of Queuing Networks", John Wiley & Sons, Inc.

8 U.S. Environmental Protection Agency (1995), AP-42, Compilation of Air Pollutant Emission Factors, Stationary Point and Area Sources, Fifth Edition, Volume I,

Chapter 13: Miscellaneous Sources, http://www.epa.gov/ttn/chief/ap42/ch13/ (Accessed on Feb. 17, 2007)

9 U. S. Environmental Protection Agency (1997), "National Ambient Air Quality Standards (NAAQS)", http://www.epa.gov/air/criteria.html (Accessed on Feb. 17, 2007)

10 U.S. Environmental Protection Agency (1997). Procedures for Preparing Emission Factor Documents.

11 U.S. Environmental Protection Agency (2005). AP-42, Compilation of Air Pollutant Emission Factors, Stationary Point and Area Sources.

12 U.S. Environmental Protection Agency (2006). "PM_{10} NAAQS Implementation", http://www.epa.gov/ttn/naaqs/pm/pm10_index.html (Accessed on Feb. 17, 2007)

13 V. Etyemezian, H. Kuhns and G. Nikolich (2006), Precision and repeatability of the TRAKER vehicle-based paved road dust emission measurement, Atmospheric Environment, Volume 40, Issue 16, Pages 2953-2958

14 W. G. Cochran (1977), "Sampling Techniques Third Edition", John Wiley & Sons, Inc.

APPENDIX A

DISTRIBUTION OF THE CALCULATED EMISSION FACTOR BASED ON SIMULATION RUNS FOR DIFFERENT ROADWAY CLASSIFICATIONS

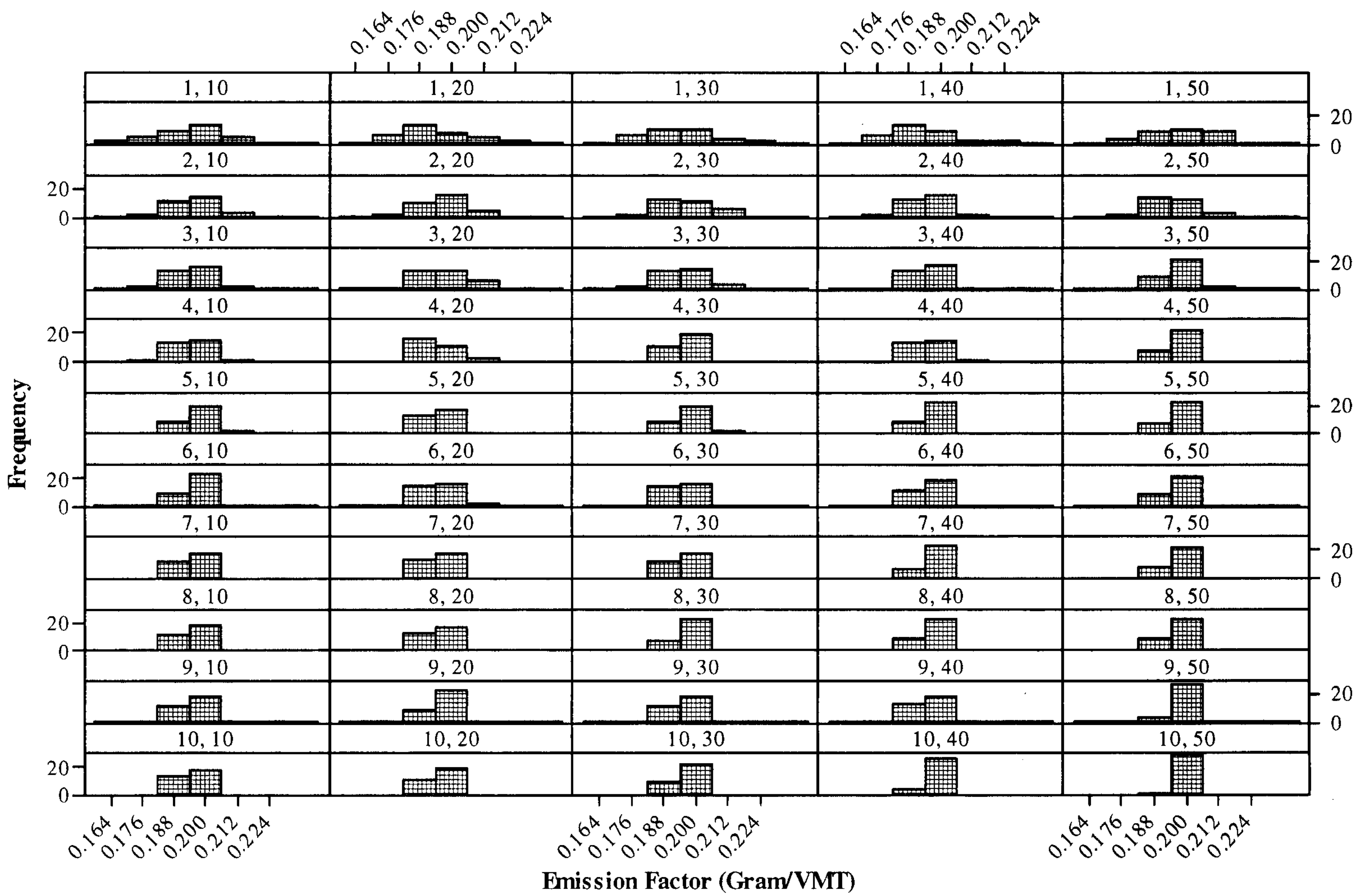


Figure A.10 Distribution of the Calculated Emission Factor for Collector Roads when the Simulation Run is 30

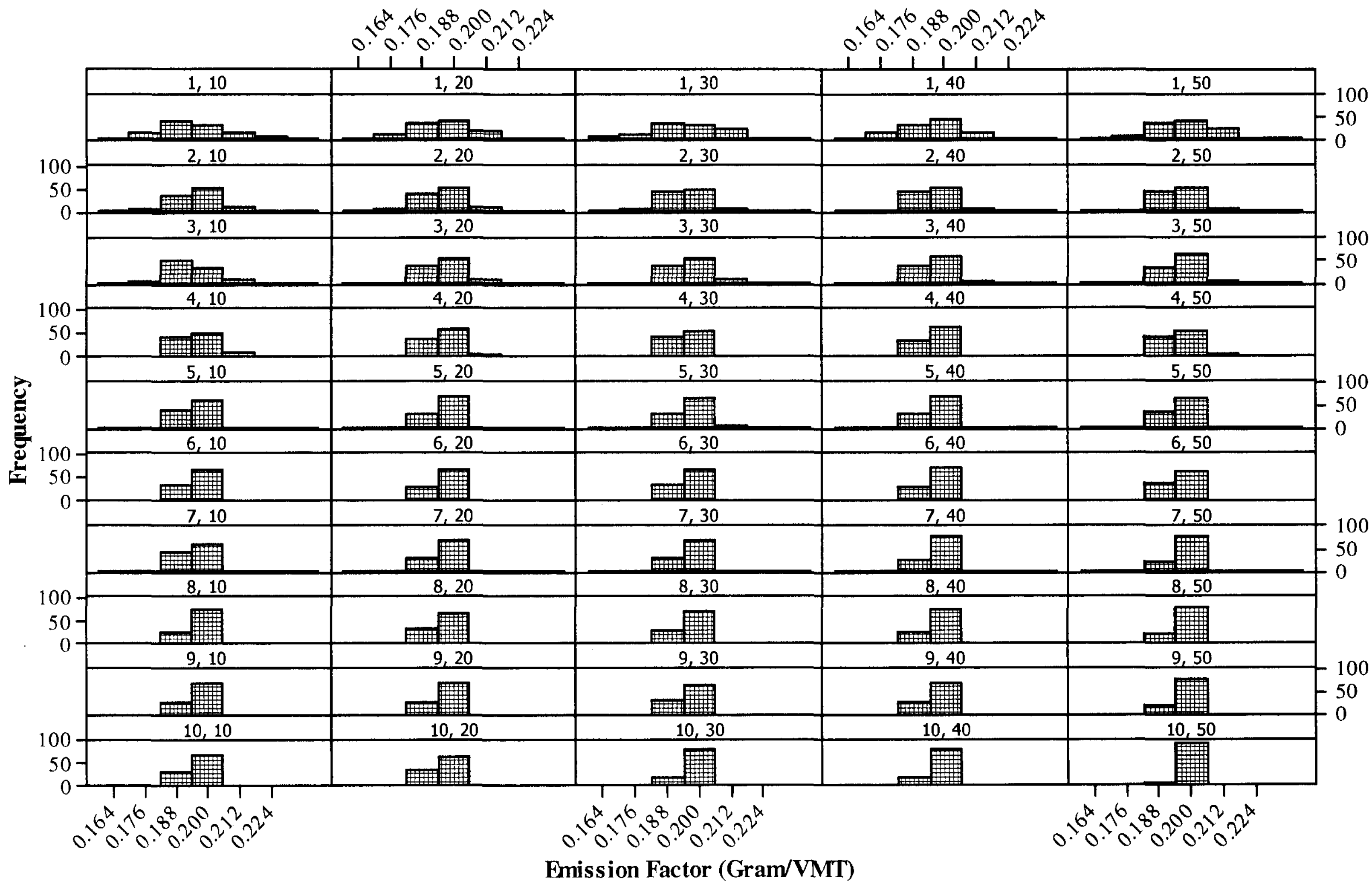


Figure A.11 Distribution of the Calculated Emission Factor for Collector Roads when the Simulation Run is 100

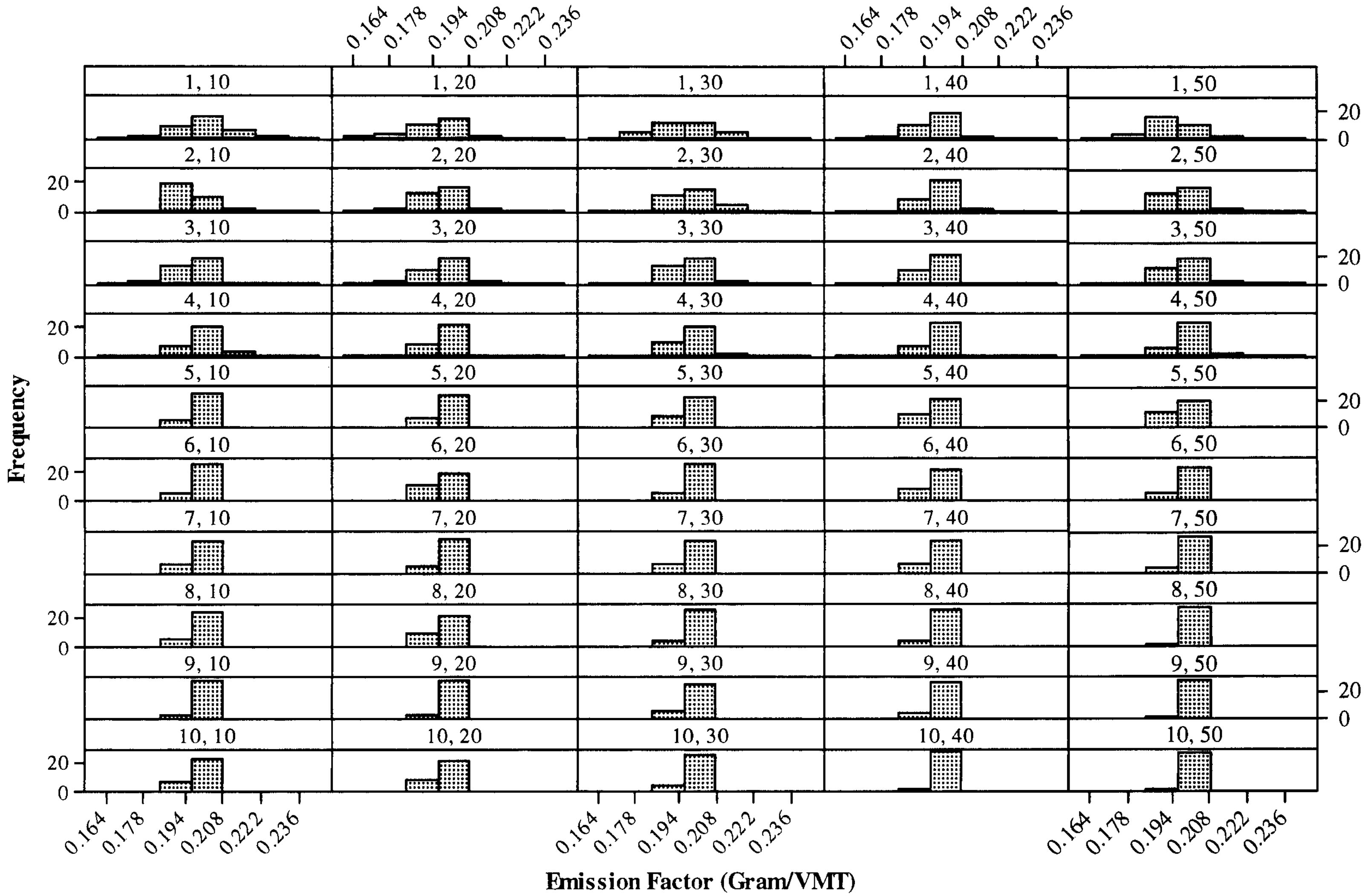


Figure A.12 Distribution of the Calculated Emission Factor for Minor Arterial Roads when the Simulation Run is 30

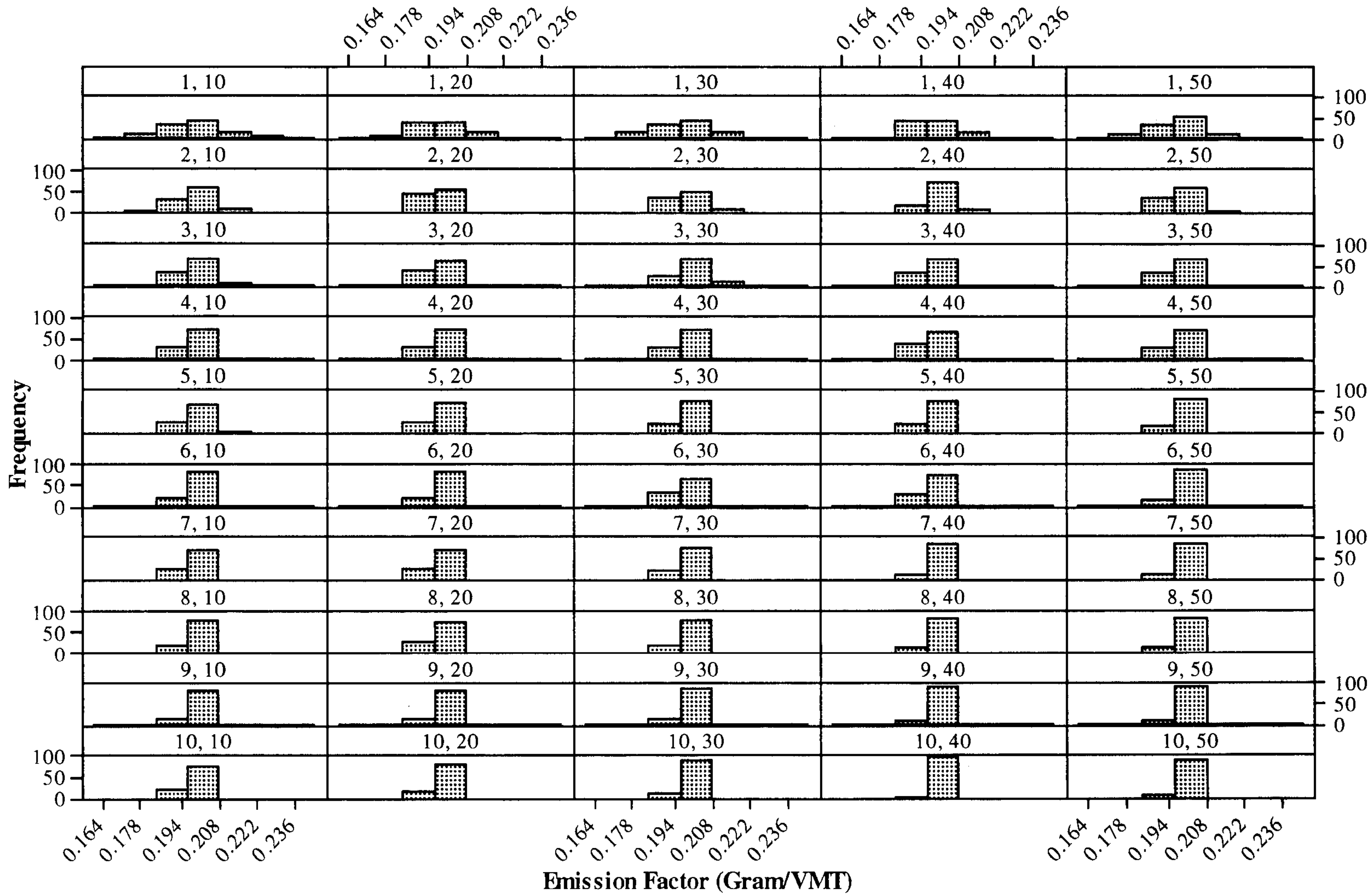


Figure A.13 Distribution of the Calculated Emission Factor for Minor Arterial Roads when the Simulation Run is 100

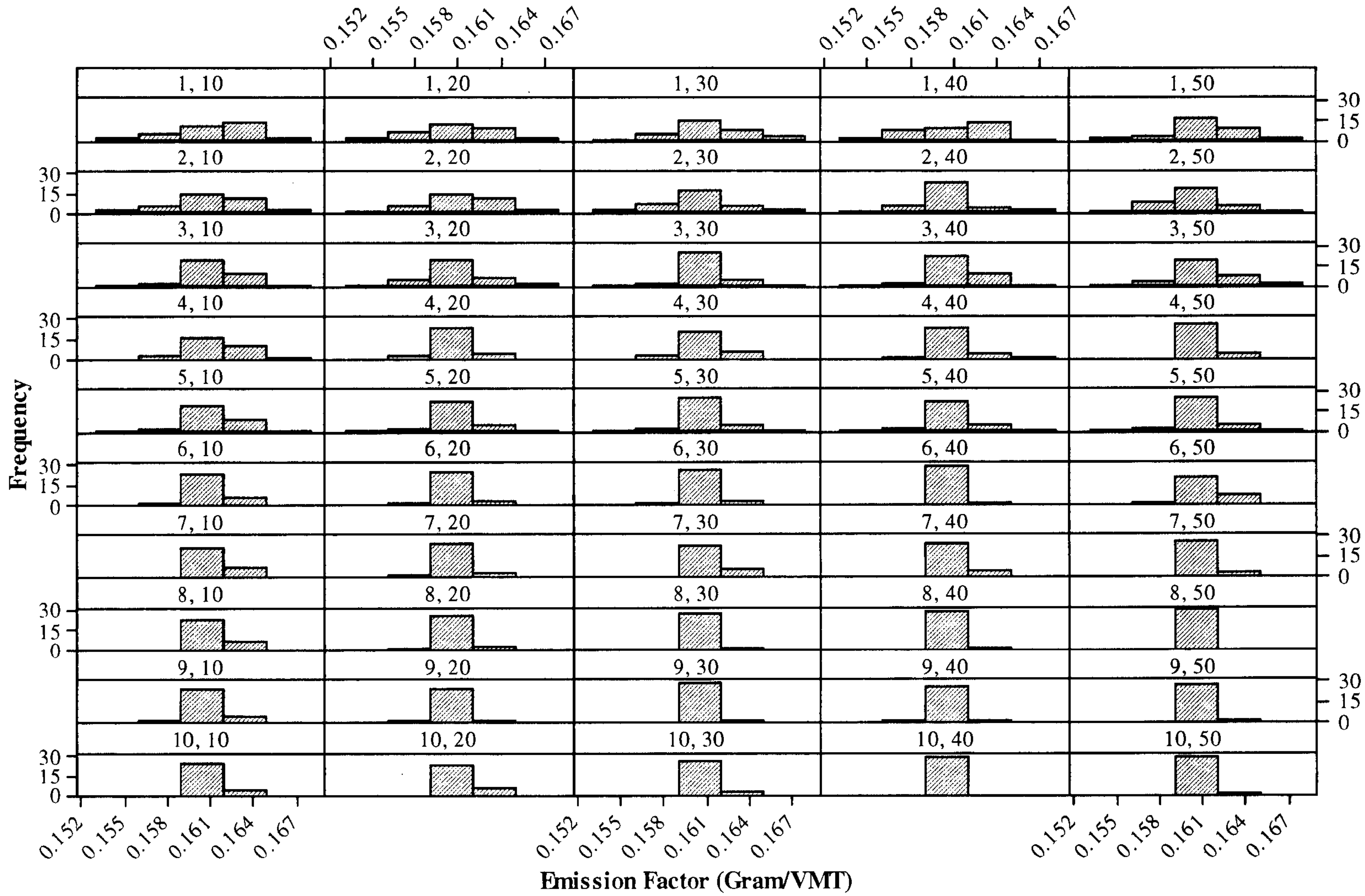


Figure A.14 Distribution of the Calculated Emission Factor for Major Arterial Roads when the Simulation Run is 30

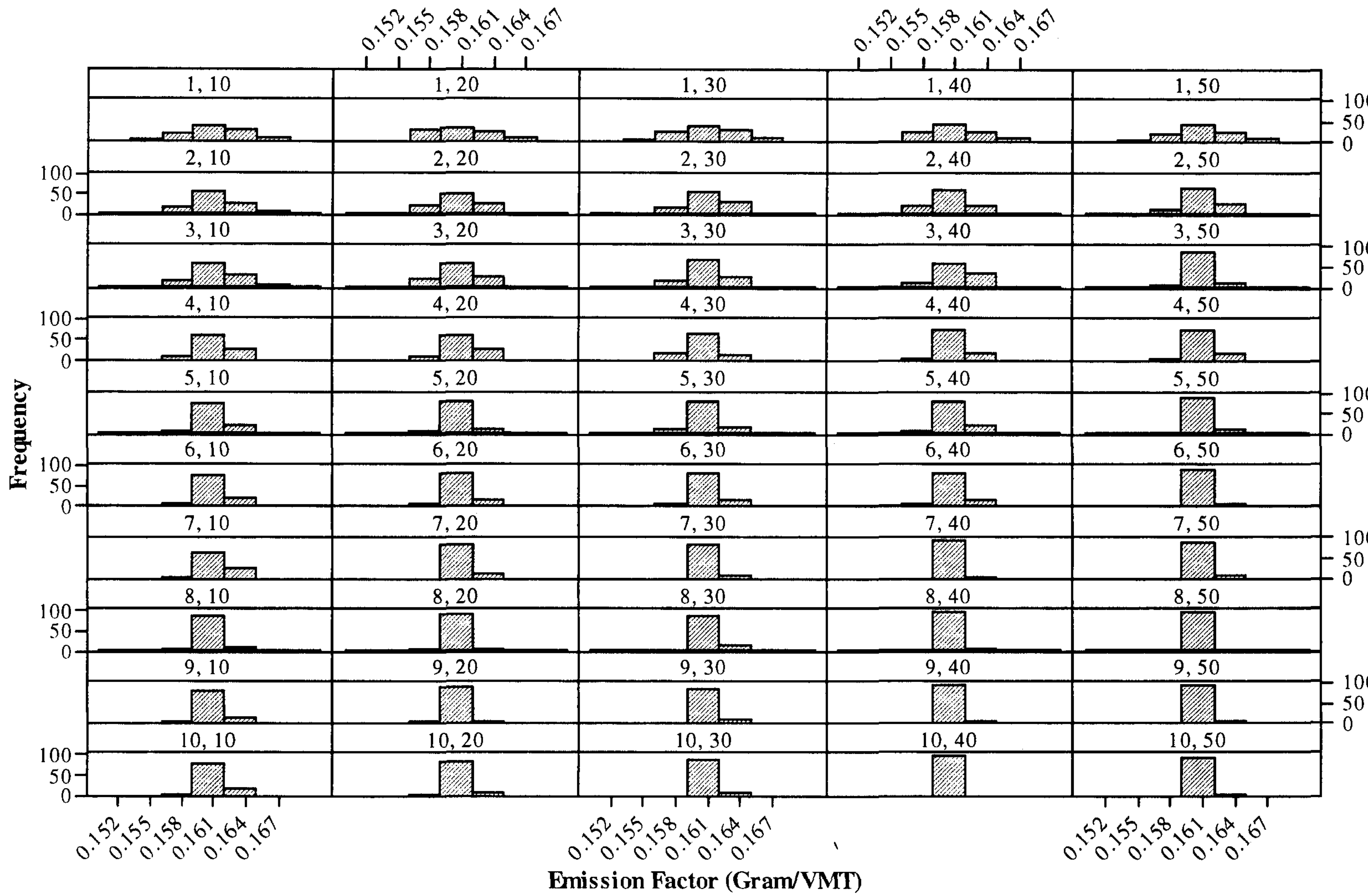


Figure A.15 Distribution of the Calculated Emission Factor for Major Arterial Roads when the Simulation Run is 100

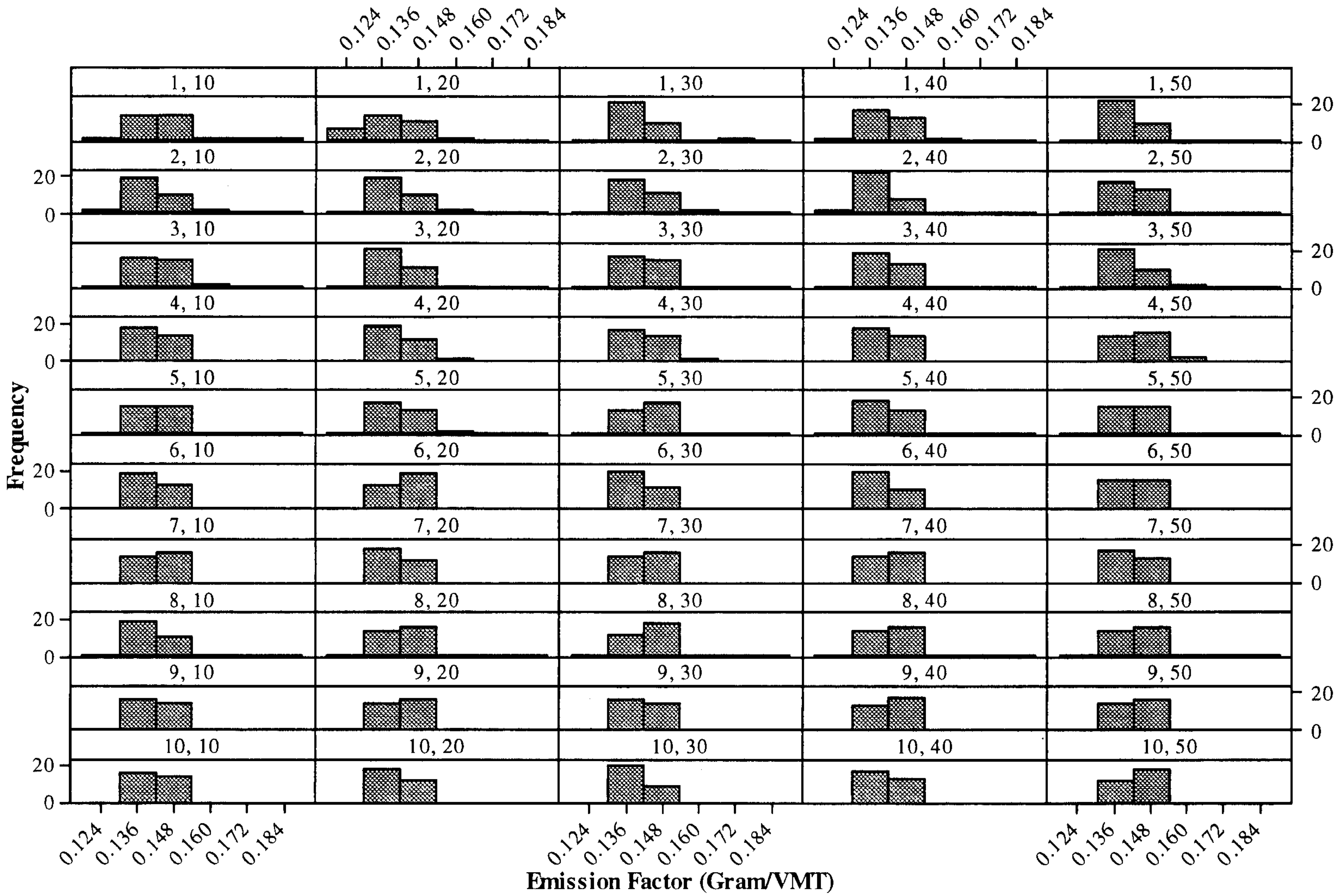


Figure A.16 Distribution of the Calculated Emission Factor for Freeway Roads when the Simulation Run is 30

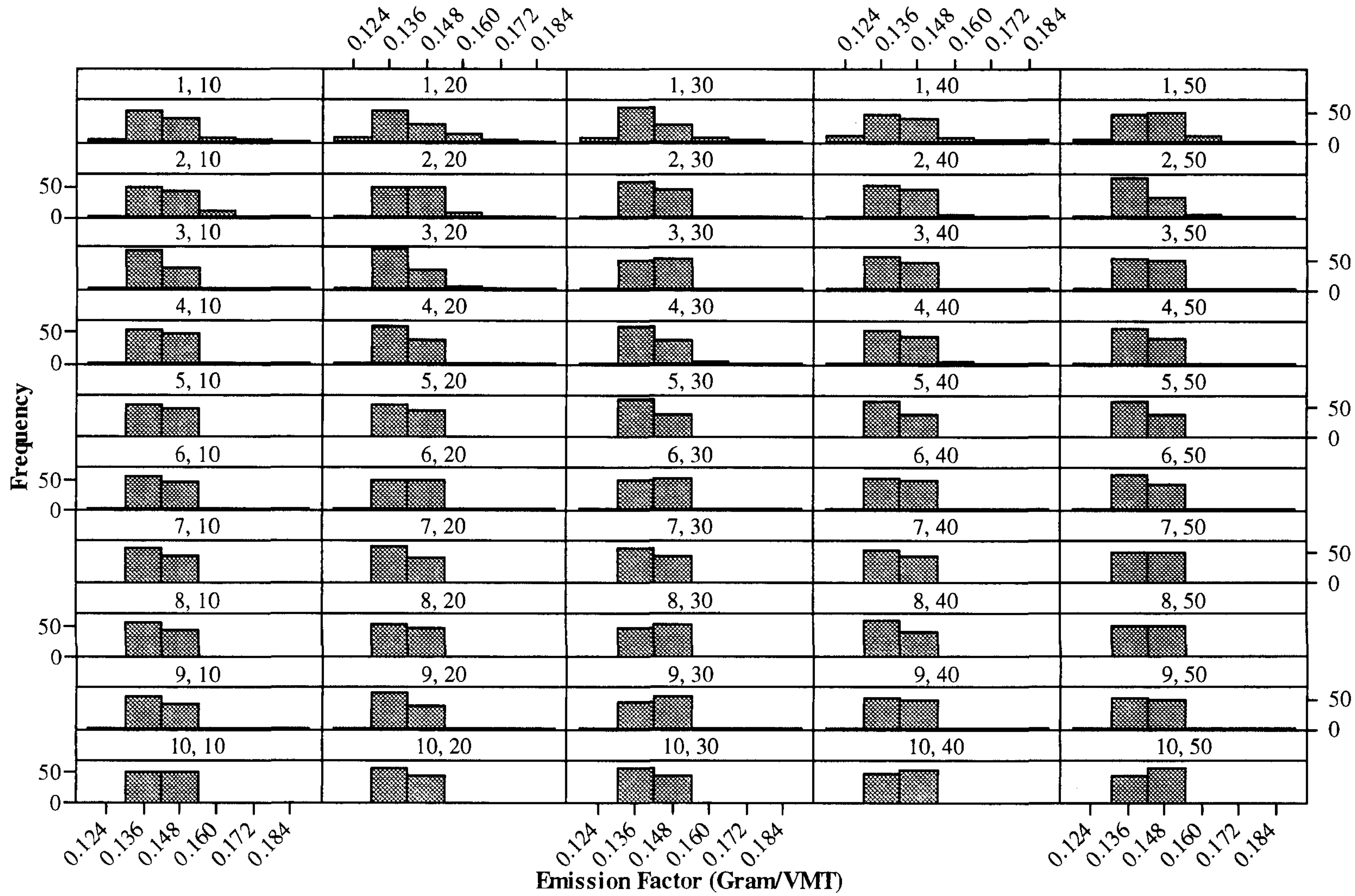


Figure A.17 Distribution of the Calculated Emission Factor for Freeway Roads when the Simulation Run is 100

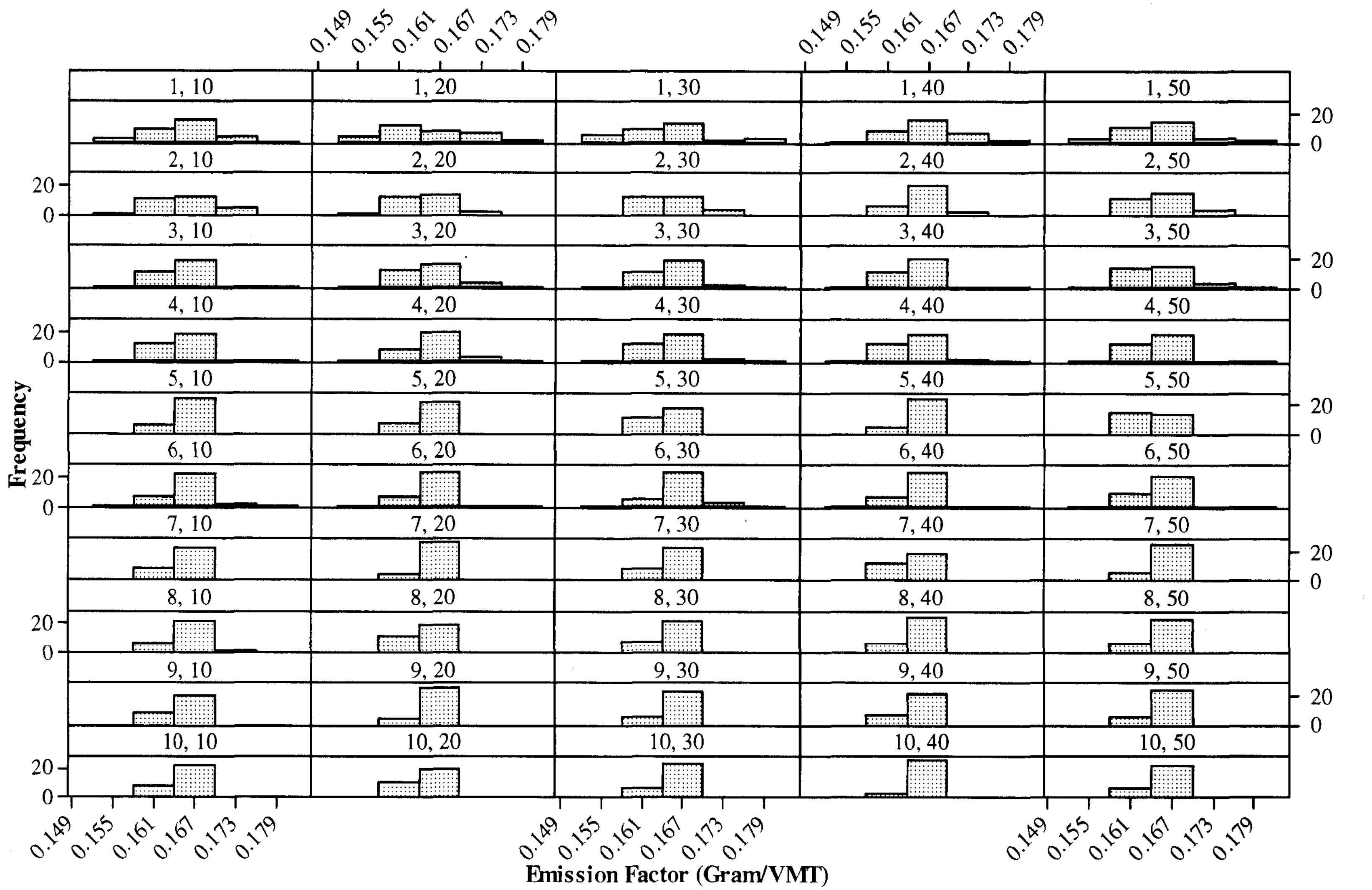


Figure A.18 Distribution of the Calculated Emission Factor for Interstate Roads when the Simulation Run is 30

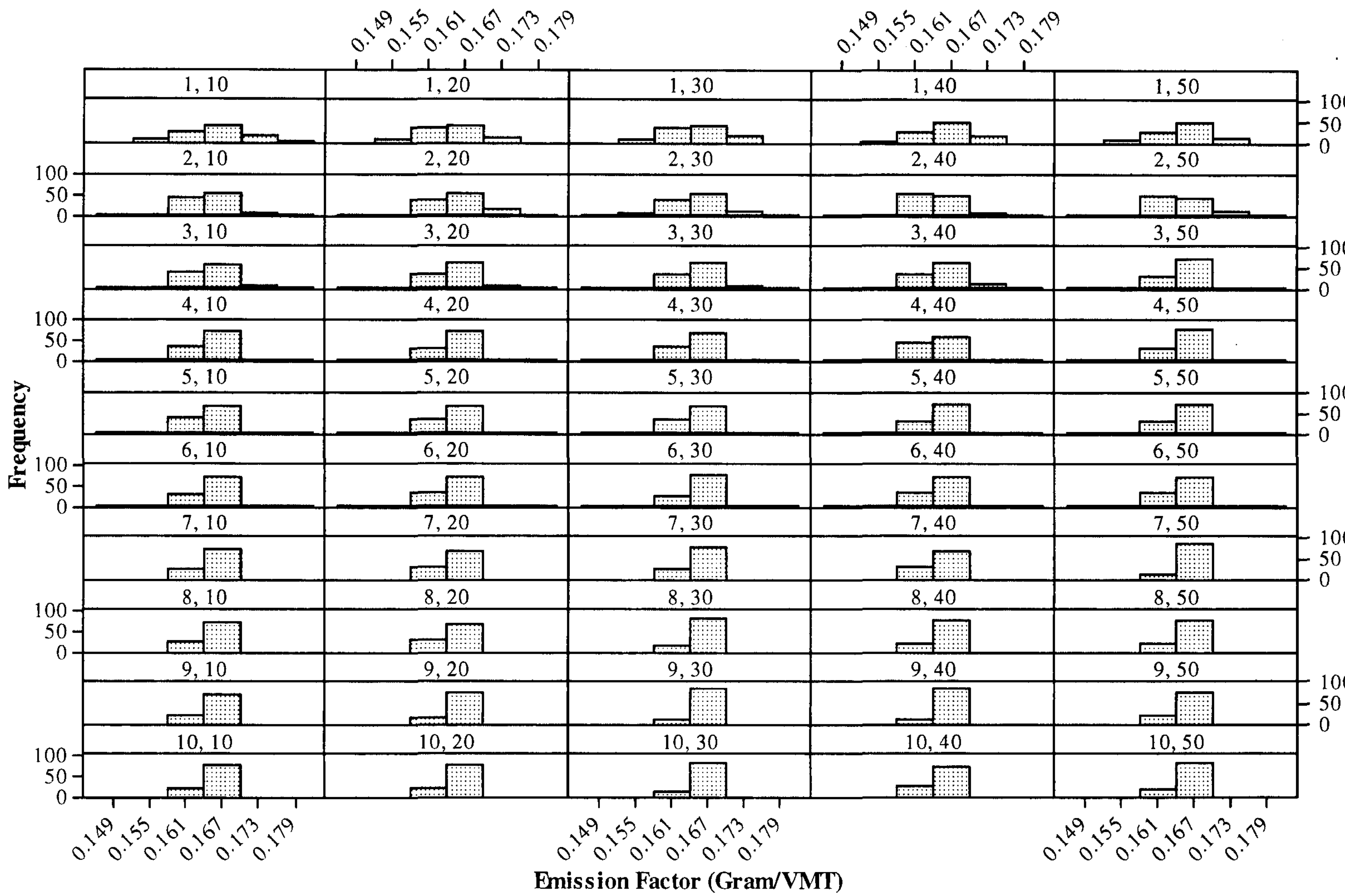


Figure A.19 Distribution of the Calculated Emission Factor for Interstate Roads when the Simulation Run is 100

VITA

Graduate College
University of Nevada, Las Vegas

Yanjie Chen

Local Address:
1600 E. University Ave.. #119
Las Vegas, NV 89119
E-mail: cheny17@unlv.nevada.edu

Degrees:
Bachelor of Science in Engineering, Mechanics, 1999
Changsha Railway University

Master of Science in Engineering, Transportation, 2002
Northern Jiaotong University

Thesis Title:
Investigation on the AP-42 Sampling Method

Thesis Examination Committee:
Chair, Dr. Hualiang (Harry) Teng, Ph. D.
Committee Member, Dr. Mohamed S. Kaseko, Ph. D.
Committee Member, Dr. David E. James, Ph. D.
Graduate Faculty Representative, Dr. Hokwon A. Cho, Ph. D.