

Impact of Pavement Roughness on Vehicle Free-Flow Speed

Authors:

T. Wang, J. Harvey, J. D. Lea, and C. Kim

Partnered Pavement Research Program (PPRC) Contract Strategic Plan Element 4.37:
Use Environmental LCA to Develop Tools and Recommend Practices to Reduce Environmental Impact

PREPARED FOR:

California Department of Transportation
Division of Research, Innovation and System
Information

PREPARED BY:

University of California
Pavement Research Center
UC Davis, UC Berkeley



| | | | | | |
|---|--------------------------------------|---|--|--|--|
| DOCUMENT RETRIEVAL PAGE | | Technical Memorandum: UCPRC-TM-2013-04 | | | |
| Title: Impact of Pavement Roughness on Vehicle Free-Flow Speed | | | | | |
| Authors: T. Wang, J. Harvey, J. D. Lea, and C. Kim | | | | | |
| Caltrans Technical Lead: D. Maskey | | | | | |
| Prepared for: California Department of Transportation Division of Research, Innovation and System Information | | | FHWA No.: CA152376A | Date Work Submitted: Dec. 18, 2013 | Date: July 2013 |
| Strategic Plan No: 4.37 | | Status: Stage 6, final version | | | Version No: 1 |
| Abstract: <p>In earlier studies of the environmental impact of pavement roughness on life cycle greenhouse gas (GHG) emissions, it was assumed that pavement roughness (usually measured by International Roughness Index, IRI) has no impact on vehicle speed. However, because ride comfort increases when a pavement becomes smoother (that is, when roughness decreases), it is possible that people will drive faster on a smoother pavement. Because most vehicles achieve maximum fuel efficiency between 40 and 50 mph (64 and 80 km/h), fuel use increases at speeds beyond this range, and this increase in speed might offset the benefits gained from the reduced rolling resistance associated with reduced pavement roughness. Therefore, to investigate the impact of changes in pavement roughness on driving behavior with respect to speed, this study built a linear regression model to estimate free-flow speed on freeways in California. The explanatory variables included <i>lane number</i>, <i>total number of lanes</i>, <i>day of the week</i>, <i>region (Caltrans district)</i>, <i>gasoline price</i>, and <i>pavement roughness</i> as measured by IRI. Data from the California freeway network from 2000 to 2011 were used to build the model. The results show that pavement roughness has a very small impact on free-flow speed within the range of this study. For the IRI coverage in this study (90 percent of the records have an IRI of 3 m/km or lower and 90 percent of the records have an IRI change of 2 m/km or lower), a change in IRI of 1 m/km (63 in./mi) resulted in a change of average free-flow speed of about 0.48 to 0.64 km/h (0.3 to 0.4 mph), a value low enough to cause almost no change in fuel use. This result indicates that making a rough pavement segment smoother through application of a maintenance or rehabilitation treatment will not result in substantially faster vehicle operating speeds, and therefore the benefits from reduced energy use and emissions due to reduced rolling resistance will not be offset by the increased fuel consumption that accompany increases in vehicle speed. However, efforts to develop a good model for predicting free-flow speed were not fully successful. The Southern California Interstate Freeway model developed yielded the best result with an adjusted R-squared of 0.72. For the rest of the regions in the state, the selected explanatory variables can only explain about half of the total variance, meaning that there are still other variables, such as vehicle type, with a substantial impact on free-flow speed that were not covered in this study.</p> | | | | | |
| Keywords: Speed; roughness; pavement; fuel consumption; GHG; greenhouse gas | | | | | |
| Proposals for implementation: | | | | | |
| Related documents: <ul style="list-style-type: none"> • UCPRC Life Cycle Assessment Methodology and Initial Case Studies for Energy Consumption and GHG Emissions for Pavement Preservation Treatments with Different Rolling Resistance, by T. Wang, I.-S. Lee, J. Harvey, A. Kendall, E.B. Lee, and C. Kim. UCPRC-RR-2012-02. April 2012. • Pavement Life Cycle Assessment Workshop: Discussion Summary and Guidelines, by J. Harvey, A. Kendall, I.-S. Lee, N. Santero, T. Van Dam, and T. Wang. UCPRC-TM-2010-03. May 2010. | | | | | |
| Signatures | | | | | |
| T. Wang First Author | J. Harvey Technical Review | D. Spinner Editor | J. Harvey Principal Investigator | D. Maskey Caltrans Technical Lead | T.J. Holland Caltrans Contract Manager |

DISCLAIMER

This document is disseminated in the interest of information exchange. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the Federal Highway Administration. This publication does not constitute a standard, specification or regulation. This report does not constitute an endorsement by the Department of any product described herein.

For individuals with sensory disabilities, this document is available in alternate formats. For information, call (916) 654-8899, TTY 711, or write to California Department of Transportation, Division of Research, Innovation and System Information, MS-83, P.O. Box 942873, Sacramento, CA 94273-0001.

PROJECT OBJECTIVES

In previous studies of the impact of pavement roughness on life cycle greenhouse gas emissions, it was assumed that pavement roughness has no impact on vehicle speed, which implies that travel behavior does not change before and after the performance of pavement preservation and rehabilitation processes that reduce pavement roughness. By building a linear regression model to estimate free-flow speed, this study attempts to verify this assumption using IRI as an indicator of pavement roughness on free-flow speed.

ACKNOWLEDGMENTS

This work was funded by the California Department of Transportation, Division of Research, Innovation and System Information, and the University of California Institute of Transportation Studies, Multi-campus Research Programs and Initiatives. The California LCA project is also part of a pooled-effort program with eight European national road laboratories and the Federal Highway Administration called Models for rolling resistance In Road Infrastructure Asset Management Systems (MIRIAM). The authors would also like to thank Deepak Maskey, Bill Farnbach, T. Joe Holland, and Nick Burmas of Caltrans for support and advice.

(This page blank.)

TABLE OF CONTENTS

| | |
|---|------------|
| Disclaimer | iii |
| Project Objectives | iii |
| Acknowledgments | iii |
| List of Figures | vi |
| List of Tables | vii |
| 1 Introduction..... | 1 |
| 1.1 Background..... | 1 |
| 1.2 Previous Studies | 2 |
| 1.3 Purpose of this Study | 5 |
| 2 Methodology | 7 |
| 2.1 Experiment Design | 7 |
| 2.2 Site Selection | 7 |
| 2.3 Data Acquisition | 8 |
| 2.3.1 IRI..... | 8 |
| 2.3.2 Speed | 9 |
| 2.3.3 Other Data | 10 |
| 2.4 Examinations of the Response and Explanatory Variables | 12 |
| 2.4.1 Speed | 12 |
| 2.4.2 Lanes | 14 |
| 2.4.3 IRI..... | 17 |
| 2.4.4 Gasoline Price..... | 20 |
| 2.4.5 Day of the Week..... | 21 |
| 2.4.6 Caltrans District..... | 24 |
| 2.4.7 Speed Limit and Road Type | 25 |
| 2.4.8 Correlation Between Selected Explanatory Variables..... | 27 |
| 3 Results and Discussion..... | 28 |
| 3.1 Modeling Using All Highway Data | 28 |
| 3.2 Modeling Using Subsets of the Data | 34 |
| 4 Conclusions | 41 |
| References | 42 |

LIST OF FIGURES

| | |
|--|----|
| Figure 2.1: Example mapping of IRI data to the base segment. | 9 |
| Figure 2.2: Histogram of all speed observations in the final dataset..... | 12 |
| Figure 2.3: Cumulative density plot of all speed observations in the final dataset. | 13 |
| Figure 2.4: Normal Q-Q plot of the speed observations. | 13 |
| Figure 2.5: Histogram of lane numbers observed in the final dataset. | 15 |
| Figure 2.6: Histogram of total number of lanes in one direction observed in the final dataset..... | 15 |
| Figure 2.7: Histogram of all IRI observations in the final dataset. | 18 |
| Figure 2.8: Density plot of IRI observations in different lanes..... | 18 |
| Figure 2.9: Density plot of IRI and speed observations. | 19 |
| Figure 2.10: Density plot of IRI before and after construction. | 19 |
| Figure 2.11: Histogram of all gasoline price observations in the final dataset. | 20 |
| Figure 2.12: Density plot of gasoline price and speed observations. | 21 |
| Figure 2.13: Histogram of day of the week in the final dataset. | 22 |
| Figure 2.14: Box plot of speed versus day of the week in the final dataset. | 22 |
| Figure 2.15: Map of Caltrans districts..... | 23 |
| Figure 2.16: Histogram of Caltrans district in the final dataset. | 24 |
| Figure 2.17: Box plot of speed versus Caltrans district in the final dataset. | 25 |
| Figure 2.18: Box plot of speed versus speed limit in the final dataset..... | 26 |
| Figure 2.19: Box plot of speed versus road type (Rural/Urban) in the final dataset. | 27 |
| Figure 3.1: Plot of fitted values versus actual values using the validation dataset. | 30 |
| Figure 3.2: Diagnostic plots of the regression model based on all highway data. | 31 |
| Figure 3.3: Residual versus each explanatory variable. | 32 |

LIST OF TABLES

| | |
|---|----|
| Table 2.1: Mean and Standard Deviation of Speed in Different Lanes..... | 16 |
| Table 2.2: Mean and Standard Deviation of Speed on Different Days of the Week | 23 |
| Table 2.3: Mean and Standard Deviation of Speed in Different Caltrans Districts | 25 |
| Table 2.4: Mean and Standard Deviation of Speed in Different Speed Limit Segments | 26 |
| Table 2.5: Mean and Standard Deviation of Speed in Different Road Types | 26 |
| Table 2.6: Correlation Coefficients Between Selected Variables | 27 |
| Table 3.1: Coefficients of Model Developed From All Highway Data | 29 |
| Table 3.2: ANOVA Results of Model Developed from All Highway Data | 33 |
| Table 3.3: Coefficients of Model Developed from Northern California Interstate Data | 37 |
| Table 3.4: ANOVA Results of Model Developed from Northern California Interstate Data..... | 37 |
| Table 3.5: Coefficients of Model Developed from Southern California Interstate Data | 38 |
| Table 3.6: ANOVA Results of Model Developed from Southern California Interstate Data..... | 38 |
| Table 3.7: Coefficients of Model Developed from Central California Interstate Data | 39 |
| Table 3.8: ANOVA Results of Model Developed from Central California Interstate Data | 39 |

1 INTRODUCTION

1.1 Background

In pavement management, life cycle assessment (LCA) can be used to evaluate the energy consumption and greenhouse gas (GHG) emissions that result from use of different pavement maintenance and rehabilitation (M&R) strategies. The phenomenon called the *rolling resistance* of a pavement surface has become a focus of LCA studies because of its effect on vehicle fuel consumption and the consequent emissions during the use phase of the pavement life cycle. Studies have already shown that roughness-reducing pavement M&R activities can significantly lower vehicle rolling resistance and, therefore, the energy consumption and CO₂ emissions from vehicles (1-5). However, some of these studies assumed that pavement roughness (2, 5) affects vehicle speeds—that is, driving behavior changes after M&R activities—but others did not (1, 3, 4). In a modeling study that mostly used highways with a small number of lanes, Hammarström, Eriksson, Karlsson, and Yahya (6) measured driver behavior in Sweden (7) and found that increases in speed essentially cancelled the benefits derived from improved smoothness. Those authors' rationale for this change of driving behavior was that since ride comfort increases with smoother pavement, it is possible that drivers will simply speed up after the pavement treatment. Since most vehicles achieve maximum fuel efficiency at steady speeds between 64 and 80 km/h (40 and 50 mph) (8, 9), and fuel efficiency decreases at speeds lower and higher than this optimum range, leaving that range may offset any benefits gained from the reduced pavement roughness and rolling resistance. To investigate whether or not this is the case, this current study investigated whether changes to pavement roughness can lead to changes in speed and emissions by developing a free-flow speed model on California freeways, using *pavement roughness* as one of the explanatory variables.

Pavement roughness (which paradoxically is sometimes termed “smoothness” from the opposite perspective) refers to the deviation of a pavement surface from a true planar surface, with wavelength deviations ranging between 0.5 and 50 m (10). Wavelengths in this range dissipate energy in the vehicle suspension—including deforming the tire body—and convert energy into heat that dissipates. Pavement roughness is usually measured in terms of the International Roughness Index (IRI), a parameter developed by the World Bank to provide a stable and portable measurement standard for worldwide use (11). IRI commonly ranges from about 1 to 5 m/km (63 to 315 inches/mile) on a paved highway, with lower values indicating a smoother surface. The U.S. Federal Highway Administration (FHWA) defines high-speed highway pavements with an IRI greater than 2.7 m/km (170 inches/mile) as being in “poor” condition (12).

This current study only considers free-flow speed because the interactions among vehicles that occur under non-free-flow conditions can significantly affect speed, making it an inconsistent value for a given set of environmental and road conditions. In a non-free-flow condition, a driver's desire to speed up on a smooth pavement will be impeded by traffic flow and will therefore not be reflected in the actual driving behavior.

1.2 Previous Studies

The *Highway Capacity Manual 2000 (HCM 2000)* (13) defined free-flow speed as “the mean speed of passenger cars that can be accommodated under low to moderate flow rates on a uniform freeway segment under prevailing roadway and traffic conditions.” Although the *Highway Capacity Manual 2010 (HCM 2010)* (14) redefined it as “the theoretical speed when the density and flow rate on the study segment are both zero,” both HCM versions considered roadway conditions to be factors that can affect free-flow speed.

HCM 2000 described four main roadway condition variables that can affect free-flow speed: lane width, lateral clearance, the total number of lanes, and interchange density (it also considers others, such as horizontal and vertical alignments, which have a lesser impact). The equation used in *HCM 2000* to estimate the free-flow speed appears below as Equation (1.1). Of these variables, a higher free-flow speed occurs with a wider *lane width*, a larger *lateral clearance*, a greater *total number of lanes*, and a smaller *interchange density*. Adjustment factors for these variables can be found in tables provided in this version of the *HCM*.

$$FFS = BFFS - f_{LW} - f_{LC} - f_N - f_{ID} \quad (1.1)$$

where:

- FFS is free-flow speed (in mph)
- $BFFS$ is base free-flow speed: 70 mph for an urban area and 75 mph for a rural area
- f_{LW} is the adjustment for lane width
- f_{LC} is the adjustment for right-shoulder lateral clearance
- f_N is the adjustment for the total number of lanes
- f_{ID} is the adjustment for interchange density.

In *HCM 2010*, the free-flow speed of a freeway segment is considered to be affected by *lane width*, *lateral clearances*, and *total ramp density*, with the latter being the most critical variable. *HCM 2010* also provides an equation to estimate free-flow speed, as shown in Equation (1.2). Of the variables in this equation, *total ramp density* is defined as the average number of on-ramp, off-ramp, major merge, and major diverge junctions per mile, and the variable is essentially a variant of the *interchange density* described in *HCM 2000*. In addition, the adjustment for lateral clearance is a function of *right-side lateral clearance* and the *total number of lanes* in one direction. As with the equation in *HCM 2000*, the greater the lateral clearance is, the greater the total number of lanes, and the wider the lane width, the smaller these adjustment factors will be. And, as with the earlier *HCM*, these adjustment factors can be acquired from tables included in the manual. Regardless, neither of the free-flow speed equations in the *HCM* editions consider *pavement roughness* as an explanatory variable; this indicates either that the model developers considered pavement roughness and found its impact on free-flow speed not to be significant, or that they did not consider roughness when developing their models.

$$FFS = 75.4 - f_{LW} - f_{LC} - 3.22 \times TRD^{0.84} \quad (1.2)$$

where:

| | |
|----------|---|
| FFS | is free-flow speed in mph |
| 75.4 | is base free-flow speed in mph (120.6 km/h) |
| f_{LW} | is the adjustment for lane width |
| f_{LC} | is the adjustment for lateral clearance |
| TRD | is the total ramp density. |

The *Highway Development and Management Model, Version 4 (HDM-4)* report reviewed a series of studies that focused on the impact of pavement roughness on vehicle speed, including the model used in *HDM-III (Highway Development and Management Model, Version 3)* (15). The HDM series was developed by the World Road Association to perform cost analyses for the M&R activities of roads. A study by Karan et al. built a regression model of highway speed using 72 sites near Ontario, Canada in 1976 (16). The explanatory variables included the *riding comfort index* (RCI), which is the Canadian equivalent of *present serviceability index* (PSI), total capacity of the roadway, traffic volume, and the speed limit. Both RCI (ranging from 0 to 10) and PSI (also ranging from 0 to 10) are largely explained by pavement roughness. Used as indices to measure human perception of pavement condition (as determined by survey groups), these two quantities can be correlated with IRI, although their relationships with it are not linear. In that study, the testing was conducted under free-flow conditions so the model could only be applied to estimate free-flow speed. The final model adopted in the Karan study is shown in Equation 1.3, where y is the average highway speed in km/h, x_1 is RCI, x_2 is the ratio of the traffic volume to total capacity of the roadway, and x_3 is the speed limit in km/h.

$$y = 30.7368 + 1.0375x_1 - 11.2421x_2 + 0.0062x_3^2 \quad (1.3)$$

$$RCI = 7.254 - 9.984 \times \log_{10} IRI \quad (1.4)^1$$

where:

| | |
|-------|---|
| y | is the average highway speed in kilometers per hour (km/h) |
| x_1 | is RCI |
| x_2 | is the ratio of traffic volume to the total capacity of roadway |
| x_3 | is the speed limit, in km/h |
| IRI | is the International Roughness Index, in m/km. |

The authors concluded that the speeds of motor vehicles on highways were significantly affected by pavement condition and that neglecting this effect might result in a major error in terms of economic evaluation. The authors also included the roughness (IRI) of each testing site. However, because IRI and RCI did not exhibit a linear relationship, in the study roughness did not have a consistent effect on speed. Using the data provided by

¹ This equation was derived using regression of the data provided in the paper. The equation provided in the paper had error in it.

the Ontario study, the relationship between RCI and IRI (m/km) is shown in Equation 1.4. It can be seen that when IRI increases from 1 to 2 m/km (63.4 to 128 in./mi) and all other variables are held constant, that speed drops about 3.11 km/h (1.95 mph). When IRI increases from 2 to 3 m/km (128 to 190 in./mi), this impact is then 1.82 km/h (1.14 mph).

The *HDM-4* report also discussed a study in South Africa by du Plessis et al. in 1990 (17). Critics of this study pointed out that it was severely skewed towards smooth pavement because 64 percent of the pavement segments had a roughness lower than 2.3 m/km. As a result the model was rejected because it was proved to be invalid in the autocorrelation test: the *roughness* variable was correlated with the *road type* variable for all vehicles except heavy trucks. In this situation, impacts from roughness are expected to be very small (especially since roads with such low IRI are common in developed countries such as the U.S.) and even get lost when other factors, such as road type, road lateral clearance, road grade, and horizontal curvature are introduced in the speed model. The *HDM-4* report also reviewed other studies, such as those by Elkins and Semrau (18) and Cox (19), but even those results left questions about the significance of any effect of pavement roughness on vehicle speed.

A study by Cooper et al. focused on the speed change before and after resurfacing on three specific flexible pavement sites in the U.K. that were resurfaced (20). Unlike the previously mentioned studies, all of which adopted an approach that involved taking “snapshots” of many test sites at one time, this study measured and analyzed speeds on the same pavement sections before and after resurfacing. It also analyzed the speeds of different types of vehicles. The results showed that traffic speed after resurfacing can increase by up to 2.6 km/h (1.6 mph), provided that the profile of the road deteriorated to a variance of at least 8 mm² using a 5 m moving-average datum (a measurement method for roughness used prior to development of the IRI). If the variance of the profile was less than 3 mm², the traffic speed was unaffected by resurfacing. The study also found that the pavement macrotexture (deviations with wavelengths between 50 mm and 0.5 m, which cause tire vibration and hysteresis) had no significant effect on traffic speed. Because this study did not provide the IRI of each testing site, it is not possible to recover the speed change corresponding to the IRI change before and after the resurfacing.

The final speed model adopted in the *HDM-4* model was inherited from *HDM-III*, based on an approach named the *Limiting Speed Model* developed by Watanatada et al. in 1987 (21). The basic concept underlying this model is that drivers are subject to a set of constraints at any given time and that vehicle speed is the minimum speed that results from these constraints. The constraints include the *driving power speed*, *braking capacity speed*, *curve speed*, *surface condition speed*, and *desired speed*.

In this model, pavement roughness is the major factor that contributes to the *surface condition speed*, i.e., the *roughness limiting speed*. In this process, IRI is converted to the maximum speed that a vehicle can travel at this roughness level by using the *maximum average rectified velocity* (ARVMAX). The value of ARVMAX is different for each type of vehicle and can be looked up in a table based on the data acquired from a study in Brazil by Watanatada et al. (21) and a study in Australia study by McLean (22). The roughness limiting speed is then compared with other limiting speeds to determine the final steady state speed.

Equation 1.5 shows the roughness limiting speed calculated in *HDM-4*, where a_0 is the coefficient (a value of 1.15 was used in *HDM-4*). Based on the World Bank’s Brazil study, it was found the roughness will be the constraining factor only when IRI exceeds about 6 m/km (380 in./mi), which seldom exists on modern highway networks in the U.S. This result again indicates that the pavement roughness may not be a significant factor in free-flow speed on modern highway networks.

$$\text{Roughness Limiting Speed (km/h)} = \frac{\text{ARVMAX (km/h)}}{a_0 \times \text{IRI}} \quad (1.5)$$

A study in India in 2004 looked at the relationship between pavement roughness, road capacity, and the speeds on a two-lane highway by building a simple linear relationship between free-flow speed and roadway roughness (23). The experiments were conducted separately with cars and heavy vehicles. It was found that roadway roughness negatively correlated with free-flow speed, and that roughness was a significant variable in this relationship. The IRI samples collected in this study ranged from 2 to 7 m/km (127 to 444 in./mi).

The Indian study found that for every 1 m/km change in IRI, the speed changes for cars and heavy vehicles were 3.4 km/h (2.1 mph) and 1.9 km/h (1.2 mph), respectively. However, it is not clear from the study whether this relationship can be applied to other highway conditions because all the data in the speed analysis were acquired from three segments of a two-lane highway in India, and those roughness levels exceed what would be allowed on most U.S. highways. Given the specific roadway condition of that study, the results might not apply to conditions in California, where most freeways have more than two lanes and thus have better lateral clearance.

1.3 Purpose of this Study

Although a survey of the existing literature turned up a number of studies on the impact of pavement roughness on speed, it also revealed that there is no consensus on what that impact is. In addition, few of the studies used IRI data collected on high-speed, multilane freeways that carry the majority of the state’s vehicles, as is the case in California. The age of those studies is also an issue as many of them are 20 to 30 years old. Lastly, few of the studies examined the speeds before and after the M&R treatment; instead, they mostly focused on speeds across

a number of sections with different roughnesses, and assumed that driver populations and other factors that contribute to speed are the same across different sections. Therefore, this current study used field data to build a linear regression model of free-flow speed using observational data from the California Department of Transportation (Caltrans) network, with a focus on the impact of pavement roughness. In this study, speed and roughness observations were collected before and after pavement treatment for a number of pavement sections. The reason a linear model was selected for this study is because many of the existing studies demonstrated a linear relationship between the free-flow speed and roughness. In addition, different non-linear regression methods were tried during the model development, such as exponential and logarithmic, but they did not yield better results than linear regression.

2 METHODOLOGY

2.1 Experiment Design

As noted in Section 1.3, the purpose of this study was to investigate the impact of pavement roughness on free-flow speed using data from field measurements.

Measuring free-flow speeds requires ensuring that the collected data comes from free-flow traffic. This means excluding from the data the impacts from high traffic volume (traffic flow and traffic densities need to be low), weather conditions (good visibility, little or no wind, and no standing water on the road), and other external factors. Section 2.3 discusses how this was done.

Because this study intended to build a free-flow speed model based on variables that are readily available in existing traffic and pavement databases, the preliminary explanatory variables selected to build the model included the *total number of lanes*², *lane number*³, *Caltrans district* (to provide a measure of regional variability), *pavement roughness* (as indicated by IRI), *day of the week*, *fuel costs*, *speed limit*, and *road type* (urban/rural roads). Further considerations in selecting the explanatory variables and the acquisition of data are discussed below.

2.2 Site Selection

The base segments for this study were selected from the Caltrans as-built inventory. The as-built inventory groups pavement segments by project type, such as overlay, seal coat, or slab replacement; and each record in the inventory represents a project. Projects are identified by the location of the segment in the pavement network using route number, state route odometer readings, and direction, and the approximate date of the project construction.

Only asphalt overlay and concrete grinding treatments were selected from the as-built inventory because these treatments should have a substantial change in IRI around the time of construction, and so it is possible to use this as a quality assurance check, that changes in IRI are not a result of other problems with the data. Lane replacements and other major rehabilitation treatments are often associated with geometric changes in the pavement, which can also cause speed changes and so should not be used.

² In this study, the total number of lanes is defined as the total number of lanes in one direction.

³ Caltrans assigns lane numbers based on their position relative to the centerline of the road, with the innermost lane being Lane 1 and the numbers increasing toward the outer lanes.

Different lanes usually have different IRI values and IRI deterioration rates, so each record in the as-built inventory was further divided by lane. Speed observations on each lane at a given location were reported and used as different records in the final data set for the model development.

In this way, because each record in the dataset can be uniquely identified by route number, start and end state odometer readings, direction, and lane number, the final database not only covered the spatial distribution of IRI and speed from different locations in the state pavement network, it also covered the distribution of temporal changes of IRI and speed for the same location using observations before and after a pavement M&R treatment. This allowed the limitations of previous studies discussed in Section 1.2 to be overcome.

In this technical memorandum, the final dataset is referred to as a collection of *base segments*, which form the base sites that were used in the analysis. In the following steps, all other necessary data were mapped to this base segment and the final dataset was used in developing the speed model.

2.3 Data Acquisition

2.3.1 IRI

The IRI measurements were acquired from the Caltrans annual pavement condition survey (PCS) from 2000 to 2011. However, Caltrans did not measure IRI on the whole network very year. Usually, a chosen location was measured and its results were extrapolated as being representative of a larger section for PCS purposes. The alignment of pavement segments in the PCS did not match the study's base segments. As a result, the IRI values from the Caltrans PCS database had to be mapped to the base segments. In the PCS database, each IRI measurement corresponds to a route number, a start and an end state route odometer reading, a direction, a lane number, and a measurement date. The following procedure was used to map the IRI of a segment from the PCS database to a base segment.

1. From each record in the base segment, the route number, the start and end state route odometer readings, and lane number were extracted.
2. Using the start and end state route odometer readings of the base segment and PCS segment, all the records in the PCS database that overlapped with the record for the base segment were found. If no records were found, the base segment was skipped and the next one was processed.
3. The weighted IRI value for each IRI measurement date was calculated using Equation 2.1, and the result was assigned to the base segment as the IRI value for that particular IRI measurement date.

$$\text{Weighted IRI} = \frac{\sum (IRI \times \text{Length of overlap})}{\sum \text{Length of overlap}} \quad (2.1)$$

Figure 2.1 shows an example of how this algorithm worked for a base segment on I-80 that had three overlapping IRI measurements from the PCS. The weighted IRI on this base segment was calculated using Equation 2.2:

$$\text{Weighted IRI} = \frac{2.5 \times 3 + 2.0 \times 6 + 3.0 \times 5}{3 + 6 + 5} = 2.464 \text{ m/km} \quad (2.2)$$

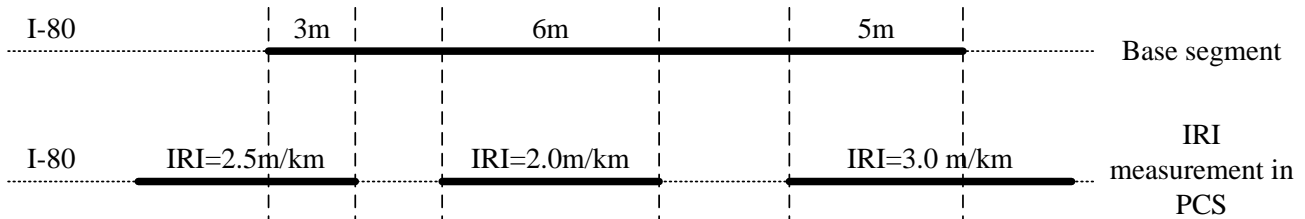


Figure 2.1: Example mapping of IRI data to the base segment.

4. The base segment was updated with the IRI value and the IRI date. Note that each record in the base segment might expand to several records because multiple measurements were taken between the years 2000 and the 2011.

Using this procedure made it possible to map the PCS database to the base segments with the IRI value and IRI measurement date. Records in the base segments that had no match in the PCS database were removed. Base segments with weighted IRI data were then saved for the next step.

2.3.2 Speed

Traffic speed, occupancy, and flow were collected from the Caltrans freeway Performance Measurement System (PeMS) (24). Because PeMS stations use loop detectors and because they are not evenly distributed on the entire state highway network, the PeMS results also needed to be mapped to the base segment. Only PeMS stations within the boundaries of base segments were selected. Because this study examined free-flow speed, only time periods with the highest probability of free-flow traffic occurring were examined. Therefore, the hourly average speeds during the periods from 11 a.m. to 12 p.m. and from 12 p.m. to 1 p.m. were collected from each qualified PeMS station. The total number of lanes in that segment was also acquired from PeMS and saved with the base segment. Although nighttime is also an off-peak period, it was not used in this study because nighttime lighting conditions may impair the visibility requirement and reduce speeds (as noted in Section 2.1).

The following procedure was followed to collect speed information from PeMS:

1. All the PeMS stations, with their route numbers and state route odometer readings, were compiled in a database.
2. The route number, start and end state route odometer readings, lane number, and IRI date were extracted from each record in the base segment with weighted IRI data. The PeMS station database was searched for PeMS stations within the range of the base segment.
3. From each PeMS station found, the hourly average speed, hourly traffic flow, and occupancy on the IRI measurement date from 11 a.m. to 12 p.m. and from 12 p.m. to 1 p.m. were extracted and saved. The total number of lanes ⁴ in that segment were also acquired from PeMS and saved in the base segment data. If no PeMS stations were found within the range, that record was ignored in the base segment and the next record was processed.
4. According to *HCM 2000*, free-flow speed is best measured when hourly traffic flow is under 1,300 passenger cars/hr/lane. *HCM 2010* lowered that value to 1,000 passenger cars/hr/lane. However, to ensure there were enough observations in the final dataset, this study adopted 1,300 as the threshold⁵. Therefore, all records with an hourly traffic flow larger than 1,300 were removed⁶. Furthermore, all records with a speed under 72 km/h (45 mph) were removed to exclude data with low flow rates from congestion periods because by definition, it is impossible and illegal to have free-flow traffic at less than 72 km/h (45 mph) on a California freeway. All records having a zero *observed percentage*⁷ were removed because this usually means there were errors with the measurements from the PeMS loop detector.

Using the procedure described, it was possible to map each record in the base segment to a free-flow speed value corresponding to the IRI.

2.3.3 Other Data

As noted earlier, this study also tried to eliminate impacts from weather conditions when speed was measured. Therefore, the weather condition associated with the location of each segment and the IRI measurement date in the base segment were identified. The weather data from 2000 to 2011 across California was acquired from the National Climate Data Center (25). For each base segment, the closest weather station within 40 miles was used

⁴ In this study, total number of lanes is defined as the total number of lanes in a specific direction.

⁵ A later experiment showed that using 1,000 as the threshold did not significantly change the results.

⁶ In this process, the number of trucks in each segment was converted to passenger cars using passenger-car equivalent. A factor of 1.5 was used in all situations because data on the gradient of each segment was unavailable.

⁷ *Observed percentage* is the percentage of 5-minute lane points that are observed in a PeMS station. This is used to determine whether the observation at that time is imputed.

as the data source for its weather condition. To ensure minimal impact from weather, observations in the dataset were limited to those with zero precipitation (and therefore no standing water on the road) and a wind speed less than 5.4 m/s (Grade 3 and lower on the Beaufort scale).

Road type, which refers to a category distinguishing urban and rural roads, and, *road access type*, which refers to a category distinguishing restricted and unrestricted access roads, can also impact free-flow speed. This study was limited to free-flow speeds on freeways, which are restricted-access roads—meaning that their traffic flows are uninterrupted by traffic lights or intersections. This is in contrast to traffic flows on unrestricted access roads where the concept of free-flow speed does not apply because of those interruptions. As a result, the unrestricted access roads in the base segment needed to be identified and eliminated. Information about road types and road access types were obtained from maps in the Caltrans road photolog (26) and the California Road System (CRS) (27), respectively. Because urban and unrestricted access roads make up only a small portion of the entire state network, two tables were developed from the data sources: a table of urban roads and a table of unrestricted access roads. Each record in the tables could be uniquely identified by the route number and the starting/ending state route odometer readings. The base segments within the boundaries defined by these two tables were considered to be urban roads and unrestricted access roads, respectively, and the rest of the base segments were considered to be rural roads and restricted-access roads, respectively. Then, all unrestricted-access segments were removed from the dataset. The final dataset only included rural restricted-access roads and urban restricted-access roads, with *rural/urban* used as an explanatory variable.

Earlier studies have shown that the price of gasoline can also affect driving behavior (22). Drivers may slow down to improve vehicle fuel economy when the fuel cost is high. Therefore, this study also included *gasoline price* as an explanatory variable. The weekly average gasoline price in California was retrieved from the Energy Almanac website provided by the California Energy Commission (28). In this study, general inflation as measured by the Consumer Price Index between the years 2000 and 2011 was around 3 percent, which is relatively low, so general inflation relative to fuel cost was not accounted for. For each record in the base segment table, the gasoline price that was closest to the IRI measurement date (which is also the date of speed measurement) was selected.

Earlier studies also showed that speed limits may also impose an impact on free-flow speed, and therefore *speed limit* was introduced as a possible explanatory variable because it represents the legal upper limit of speed on a road and also reflects the driver's safety concerns (although it is common that actual driving speeds exceed the speed limit). The general speed limit of freeways in California is 104 km/h (65 mph), while segments on some freeways have 112 km/h (70 mph) speed limits. The boundaries of these segments were acquired from the

Caltrans website (29). Any base segment within the 112 km/h (70 mph) speed limit boundaries were considered to have this speed limit. All other base segments were considered to have a speed limit of 104 km/h (65 mph).

2.4 Examinations of the Response and Explanatory Variables

The final dataset used for the analysis was prepared according to the procedures laid out in Section 2.3. The total number of observations in the final dataset was about 20,000. Each data record included a speed observation, an IRI observation and the corresponding date, the segment's location and Caltrans district number, the average gasoline price at the time of IRI measurement, and the speed limit on that segment. As discussed earlier, the preliminary explanatory variables included *lane number*, *total number of lanes*, *day of the week*, *Caltrans district*, *gasoline price*, *IRI*, *road type (urban or rural)*, and *speed limit*. This section examines the data coverage on these preliminary explanatory variables and explains how the final explanatory variables that were adopted for the model were determined.

2.4.1 Speed

Figure 2.2 shows a histogram of all the *speed* observations, Figure 2.3 shows a cumulative density plot of all the speed observations, and Figure 2.4 shows a quantile-quantile (Q-Q) plot of the speed observations. It can be seen that the speed observations follow the normal curve fairly closely except for the samples on both extremes.

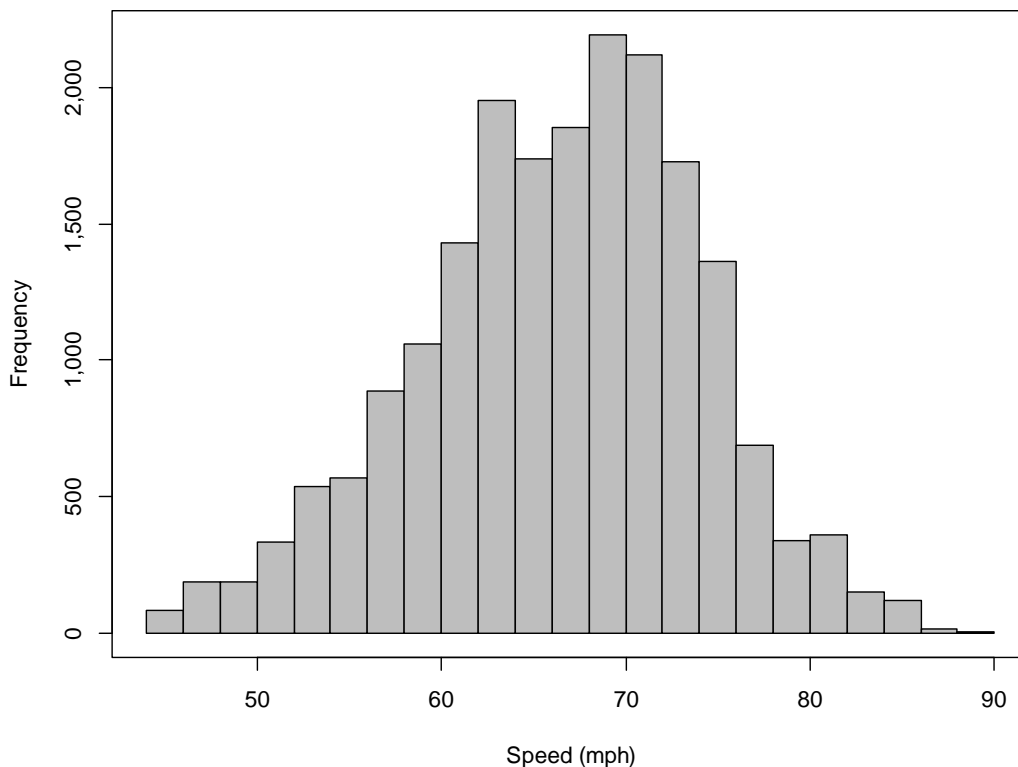


Figure 2.2: Histogram of all speed observations in the final dataset.

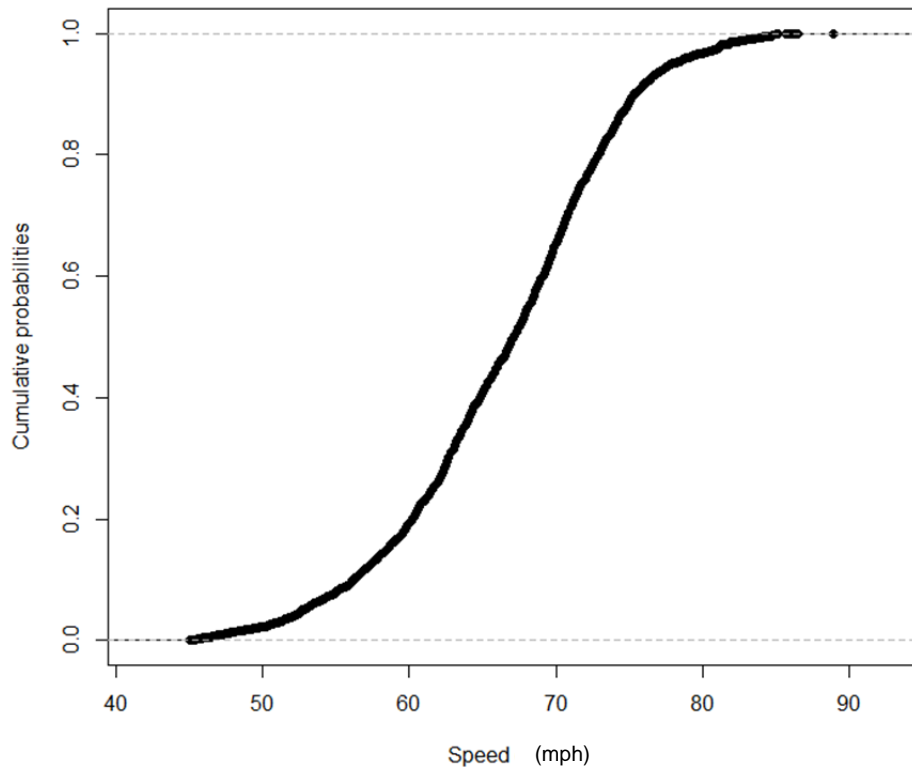


Figure 2.3: Cumulative density plot of all speed observations in the final dataset.

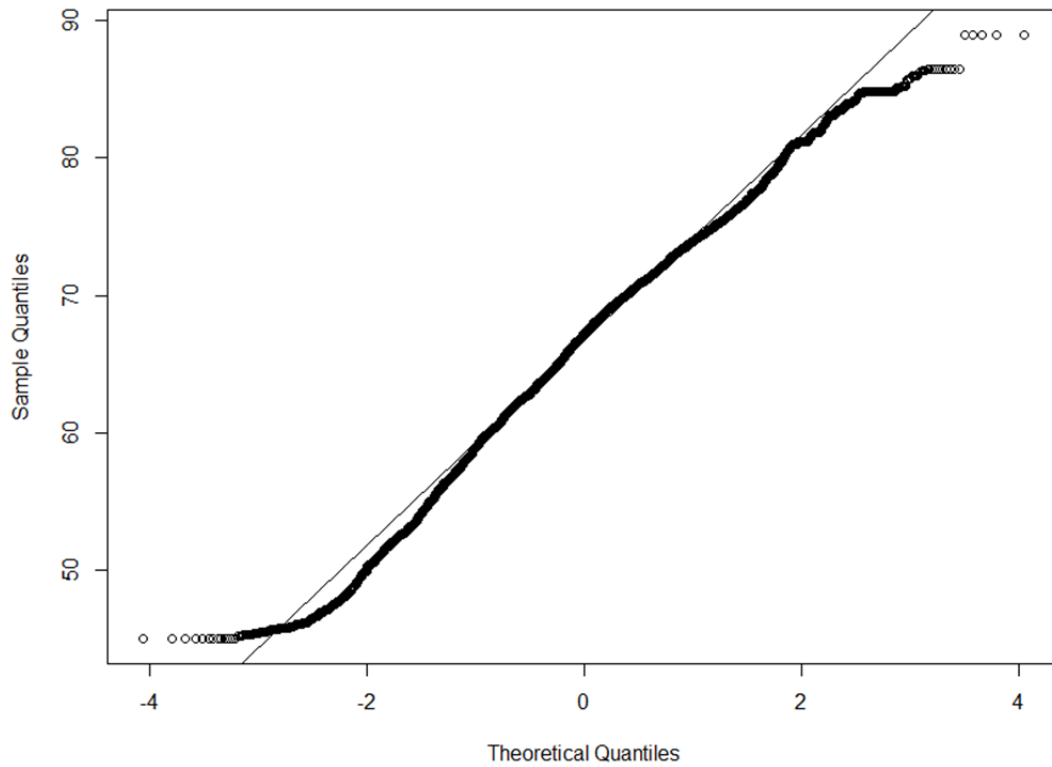


Figure 2.4: Normal Q-Q plot of the speed observations.

2.4.2 Lanes

Figure 2.5 shows a histogram of observations by *lane number* in the final dataset. The plot shows a large sample size for lane numbers 1 through 4 and a small sample size for Lane 5, which suggests that the results of this study might not apply to segments with more than four lanes in one direction.

The histogram in Figure 2.6 shows the number of observations of the total number of lanes contained in the final dataset. The *total number of lanes* on each segment was used as an explanatory variable because it affects drivers' ability to maneuver to avoid slower-moving traffic. This variable is the total number of lanes in one direction (the direction of the segment). As can be seen from the distribution in the figure, there were relatively few observations on segments with more than five lanes. Therefore, when the model developed from this study is applied, its speed prediction for roads with more than five lanes may have greater uncertainty.

Table 2.1 shows the mean values and standard deviations of all the possible combinations of lane numbers and total numbers of lanes. Some combinations, such as Lane 6 under a total number of lanes of 6, had zero or very low numbers of observations in the final dataset. Therefore, the speed model developed in this study might have much higher uncertainty in these situations.

Average speed was generally higher when the lane was closer to the center line (a lower lane number). This was expected because fewer trucks travel on the inner lanes and trucks generally drive at lower speeds than other vehicles. Generally, the larger the total number of lanes, the higher the speed. This is intuitive because a larger total number of lanes means better maneuverability, which leads to a higher free-flow speed according to *HCM 2000 (13)*. T-tests showed that there is a significant difference in speed observations between different total number of lanes and different lane numbers at a 5 percent significance level. Therefore *lane number* and *total number of lanes* were both included in the final explanatory variables.

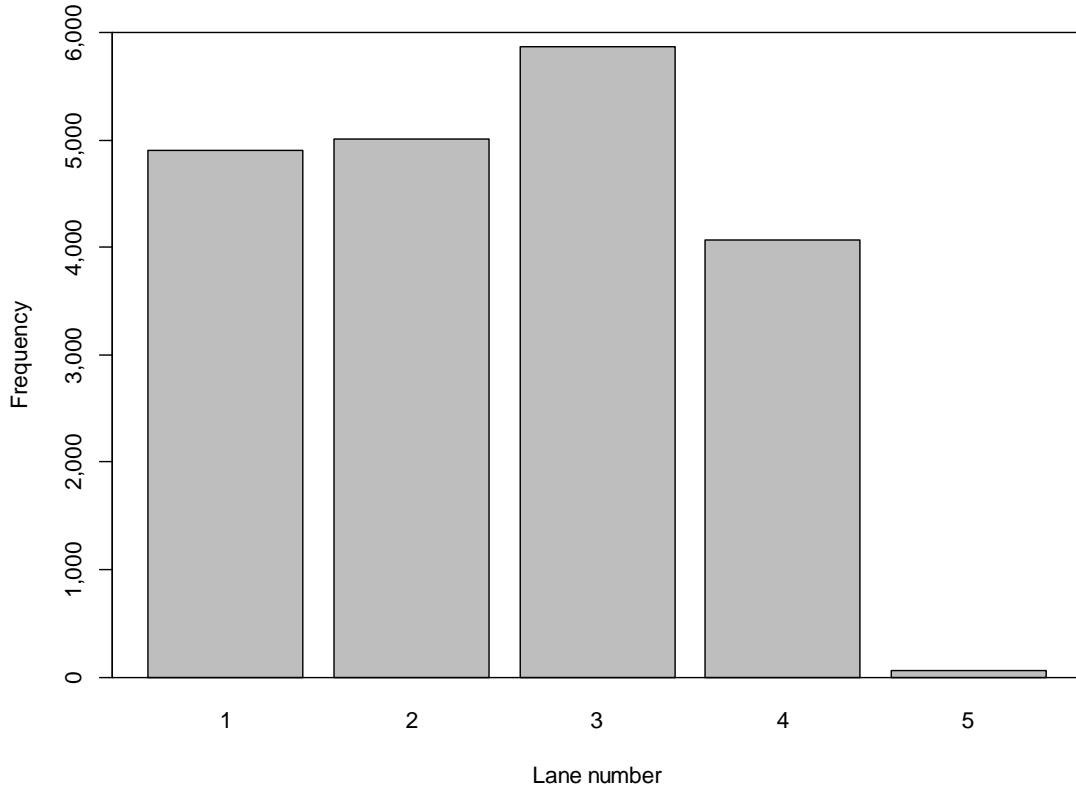


Figure 2.5: Histogram of lane numbers observed in the final dataset.

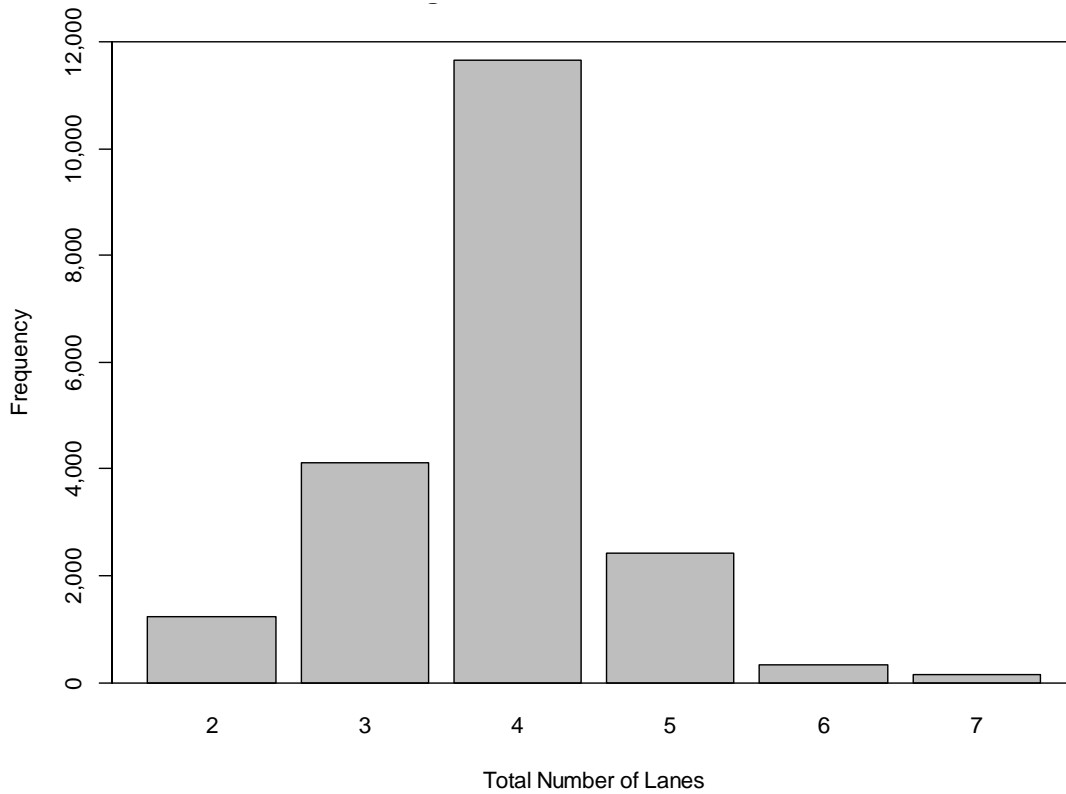


Figure 2.6: Histogram of total number of lanes in one direction observed in the final dataset.

Table 2.1: Mean and Standard Deviation of Speed in Different Lanes

| Total Number of Lanes in One Direction | Lane Number | Mean Value (mph) | Standard Deviation | Number of Observations |
|---|--------------------|-------------------------|---------------------------|-------------------------------|
| 2 | 1 | 71.5 | 6.2 | 663 |
| | 2 | 61.6 | 6.1 | 576 |
| 3 | 1 | 71.5 | 5.3 | 1,130 |
| | 2 | 67.6 | 5.3 | 1,446 |
| | 3 | 60.8 | 6.2 | 1,541 |
| 4 | 1 | 73.8 | 4.8 | 2,463 |
| | 2 | 69.0 | 4.5 | 2,240 |
| | 3 | 64.2 | 6.3 | 3,603 |
| | 4 | 59.7 | 5.7 | 3,357 |
| 5 | 1 | 74.2 | 6.1 | 543 |
| | 2 | 70.8 | 7.2 | 656 |
| | 3 | 68.4 | 6.9 | 592 |
| | 4 | 64.8 | 6.3 | 573 |
| | 5 | 62.4 | 2.7 | 53 |
| 6 | 1 | 78.8 | 4.6 | 75 |
| | 2 | 70.8 | 4.6 | 66 |
| | 3 | 68.0 | 6.1 | 92 |
| | 4 | 66.0 | 6.2 | 96 |
| | 5 | 64.9 | 0.8 | 4 |
| | 6 | N/A | N/A | 0 |
| 7 | 1 | 76.1 | 4.6 | 29 |
| | 2 | 79.3 | 4.8 | 24 |
| | 3 | 69.8 | 4.8 | 37 |
| | 4 | 67.0 | 3.3 | 42 |
| | 5 | 59.5 | 0.5 | 8 |
| | 6 | N/A | N/A | 0 |
| | 7 | N/A | N/A | 0 |

2.4.3 IRI

Figure 2.7 shows a histogram of all *IRI* observations (including before and after construction). The range of *IRI* observations shows good coverage, with contributions from very smooth pavement (around 1 m/km [63 in./mi]) to very rough pavement (around 4 m/km [252 in./mi]). The dataset did not include enough observations for *IRI* values greater than 4.5 m/km or less than 0.5 m/km, so the model may be irrelevant for these situations.

Figure 2.8 shows a density plot of the *IRI* in different lanes (from Lane 1 to Lane 4). Lane 5 was excluded because there too few observations. It is clear that lanes closer to the center line (lower lane numbers) were associated with a lower *IRI* values, which matches the fact that trucks, which generally drive at lower speeds than other vehicles, are mostly restricted to the outside lanes and that truck axle loadings are the major contributor to increases in *IRI* over time.

Figure 2.9 shows a density plot of both the *speed* and *IRI* observations. Both *IRI* and *speed* covered a reasonable range. The highest density exists between *IRI* values of 1 and 2 m/km (63 and 126 in./mile) and *speed* values of 104 and 112 km/h (60 and 75 mph), which is the approximate free-flow speed on most freeways.

As discussed in Section 1.3, this study also intended to cover the temporal difference of *IRI* (*IRI* before and after a pavement M&R treatment) to examine its impact on speed. Figure 2.10 shows a density plot of *IRI* observations in the dataset before and after construction. It is clear that, overall, *IRI* decreases after construction events. Because each *IRI* observation in the data was associated with a *speed* observation, the dataset developed in this study had coverage sufficient to examine the temporal variation of *IRI* and *speed*. Further, examination on the data set found that 90 percent of the locations have an *IRI* change (differences between the maximum and minimum *IRI* at each location within the analysis period) less than 2 m/km (126 inches/mile), which also gives a range in which the conclusions of this study should be restricted.

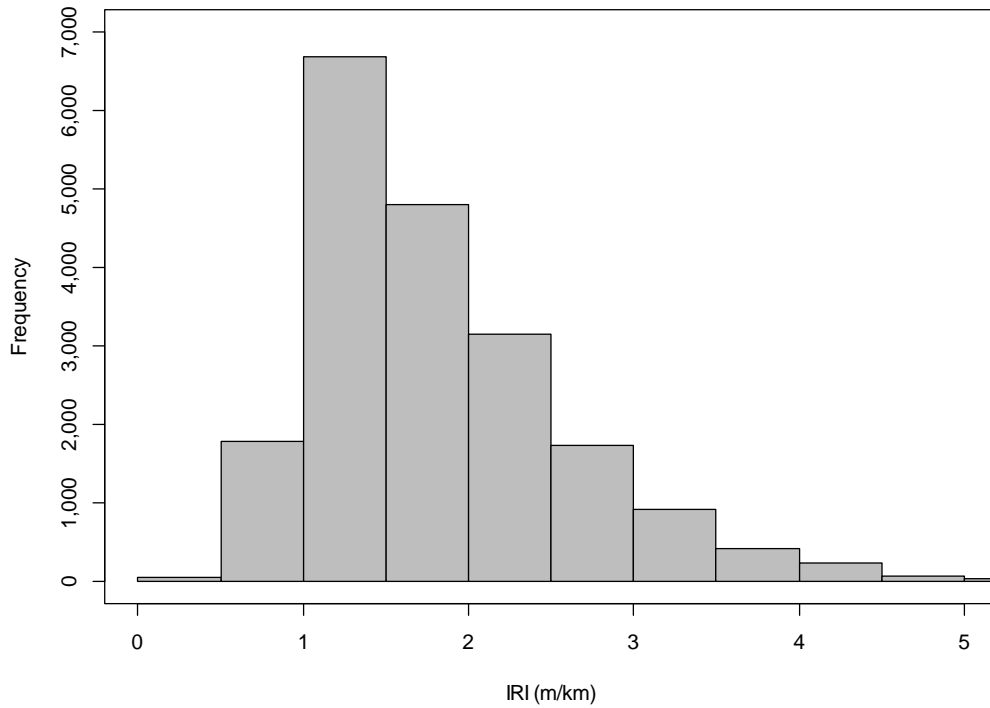


Figure 2.7: Histogram of all IRI observations in the final dataset.
 (Note: 1 m/km = 63 inches/mile.)

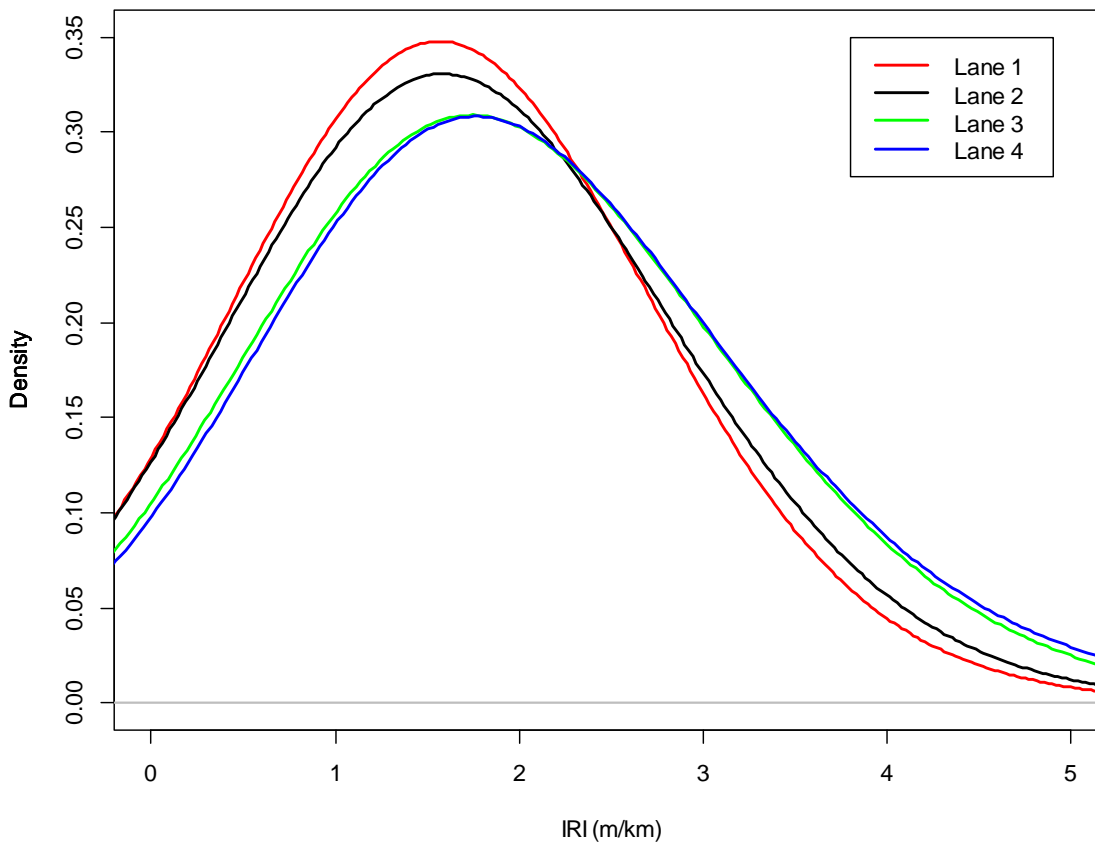


Figure 2.8: Density plot of IRI observations in different lanes.
 (Note: 1 m/km = 63 inches/mile.)

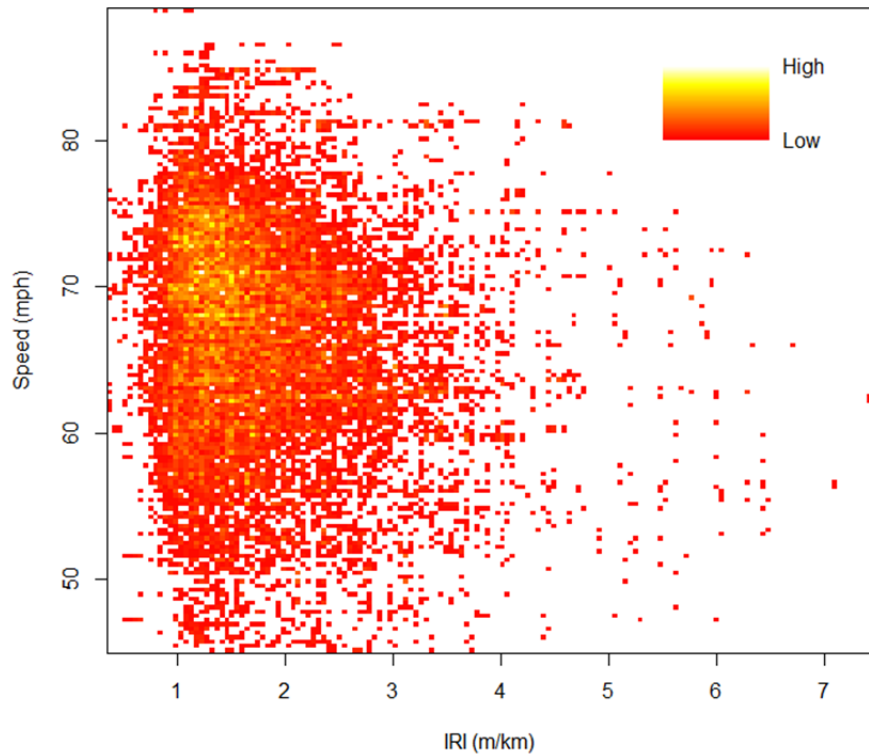


Figure 2.9: Density plot of IRI and speed observations.
 (Note: 1 m/km = 63 inches/mile.)

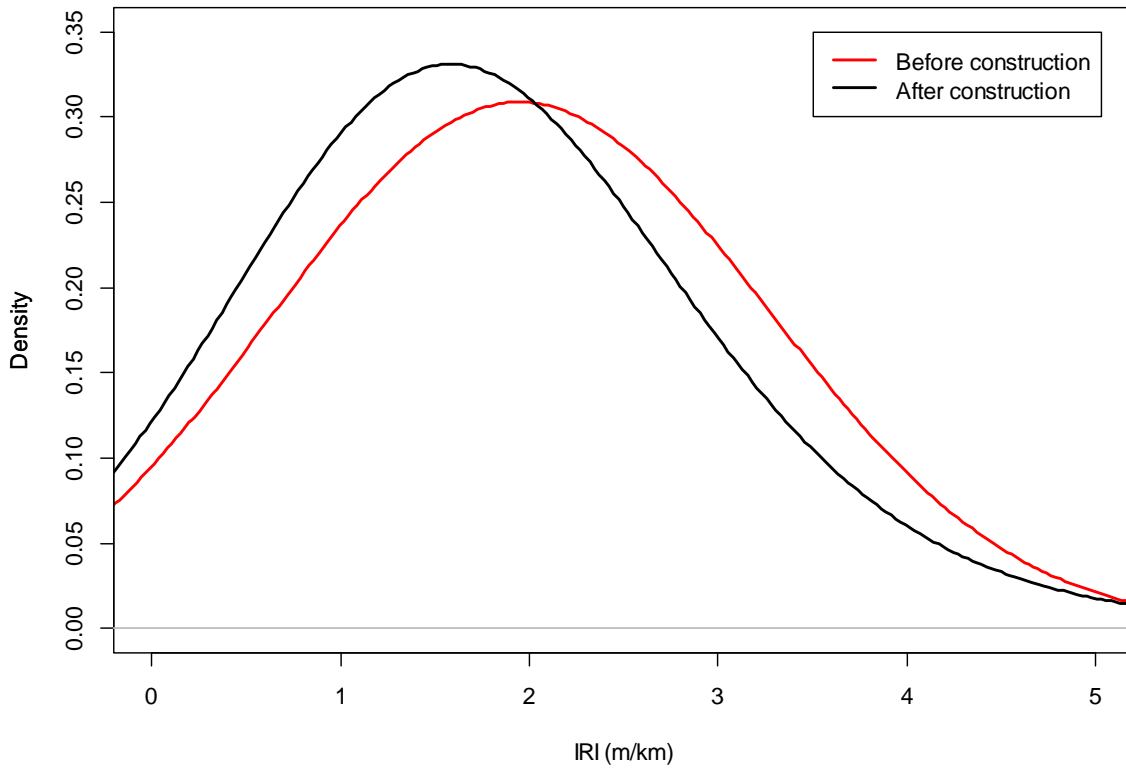


Figure 2.10: Density plot of IRI before and after construction.
 (Note: 1 m/km = 63 inches/mile.)

2.4.4 Gasoline Price

Figure 2.11 shows a histogram of all *gasoline price* observations. Over the years covered by this study, the general inflation was relatively low, with an annual change in the Consumer Price Index (CPI) of around 3 percent, so inflation was not a significant factor in the price of gasoline. It can be seen that the gasoline price ranged from about \$1.50/gal (\$0.40/liter) to \$4.50/gal (\$1.19/liter). The dataset did not include enough observations for gasoline prices higher than \$4.50/gal or lower than \$1.50/gal, and thus the model may not be relevant to those situations. The figure also shows that the observations were spread across a range of prices, with the most observations in the \$3.10 to \$3.20/gallon (\$0.82 to \$0.84/liter) range.

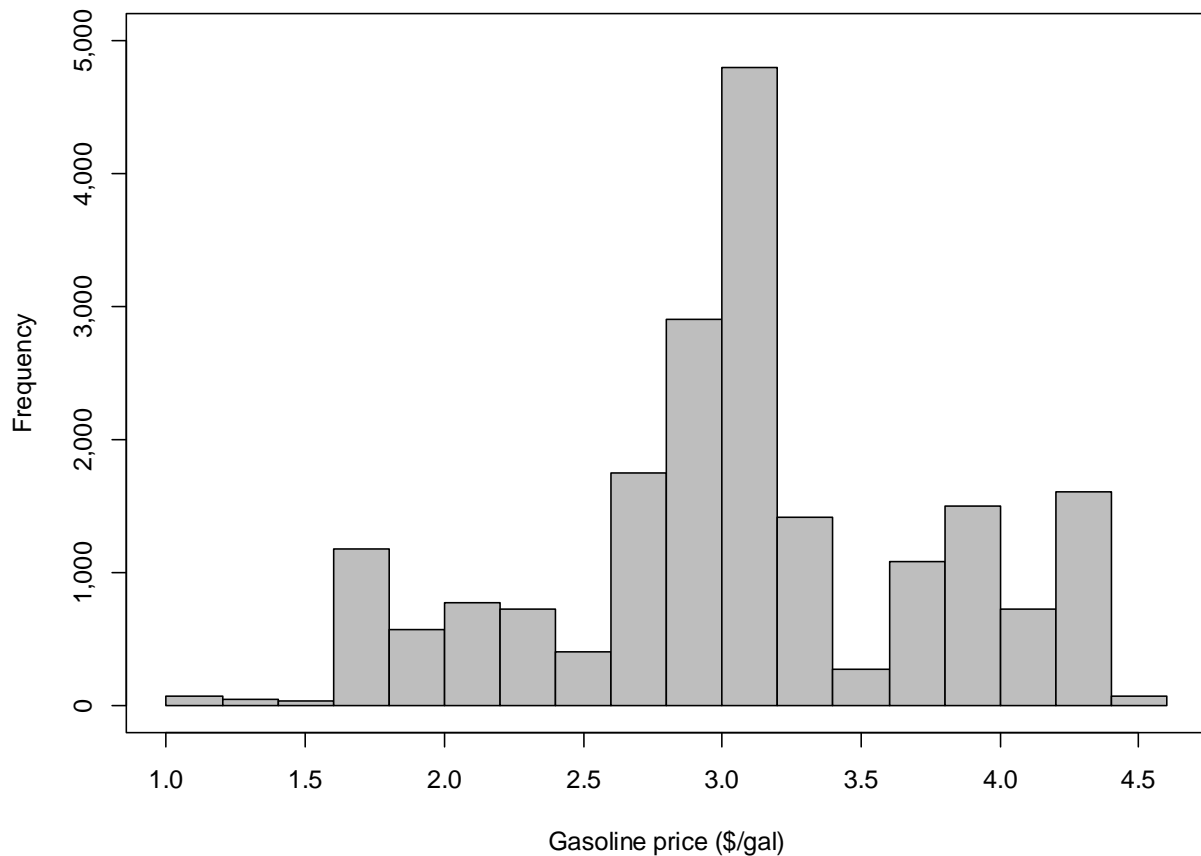


Figure 2.11: Histogram of all gasoline price observations in the final dataset.

Figure 2.12 shows a density plot of *speed* and *gasoline price* observations. It can be seen that the gasoline price and speed had reasonable coverage, with the highest density between 60 and 75 mph (104 and 112 km/h) and \$1.70/gal to \$4.30/gal. However, the value of the gasoline price in this study was not continuous because the gasoline price acquired from California Energy Commission was not continuous from 2000 to 2011. Therefore the density plot of gasoline price and speed does not look like the density plot of IRI and speed shown in Figure 2.9.

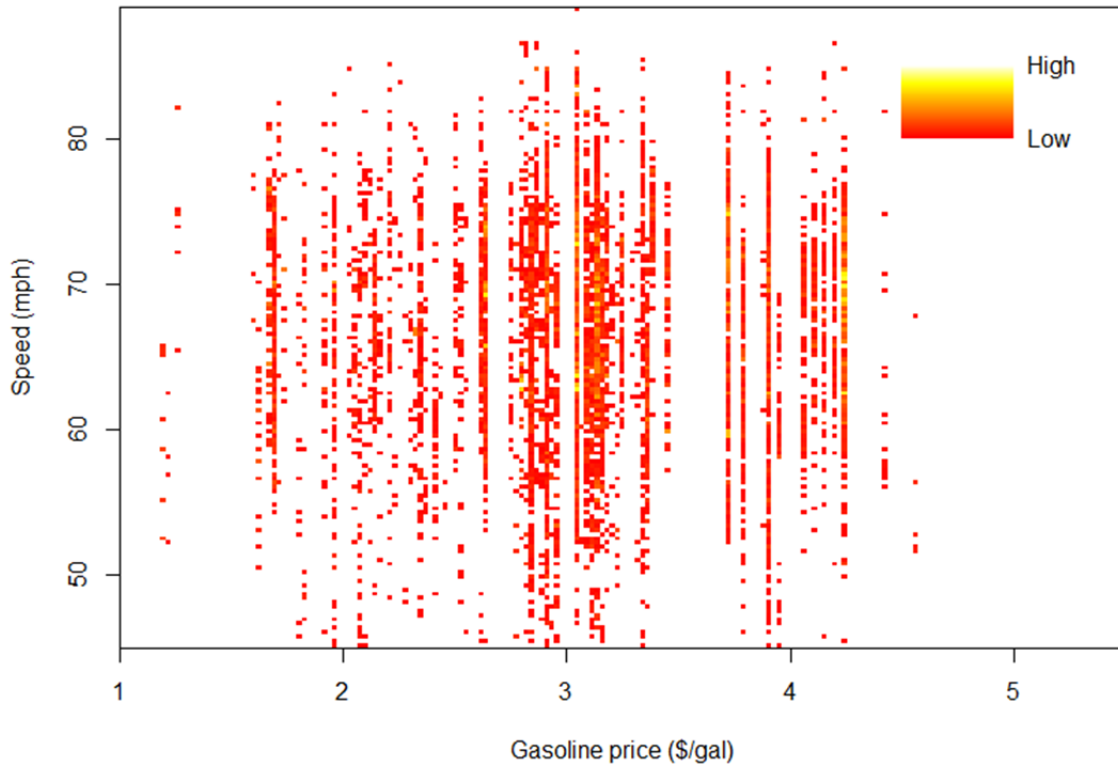


Figure 2.12: Density plot of gasoline price and speed observations.

2.4.5 Day of the Week

Day of the week was converted directly from the IRI measurement date. This variable was introduced because trips may have different purposes on weekdays and weekends, with commuting and other work-related driving the major purposes on weekdays, and with weekend driving having mixed purposes (such as entertainment and shopping). The driver population demographics may also differ between weekdays and weekends. Figure 2.13 shows a histogram of day of the week in the final dataset. Day 0 means Sunday, Day 1 means Monday, Day 2 means Tuesday, and so on. It was found that each day had enough speed observations, so the final model can be applied to all days of the week.

Because driving behavior on holidays may not be similar to that on weekdays, the national holidays in each year from 2000 to 2011 were identified and marked as Day 0 (Sunday). In this study, national holidays included New Year’s Day, Martin Luther King, Jr. Day, Washington’s Birthday, Memorial Day, Independence Day, Labor Day, Columbus Day, Veterans Day, Thanksgiving Day, and Christmas Day.

Figure 2.14 shows a box plot of the speed observations on different days of the week. Table 2.2 shows the mean value and standard deviation of the *speed* value on each day of the week. They show that weekends (representing holidays, Saturday, and Sunday) have a higher free-flow speed than weekdays. T-tests showed that there was a significant difference in speed observations between different days of the week at a 5 percent significance level. Therefore *day of the week* was included in the final explanatory variables.

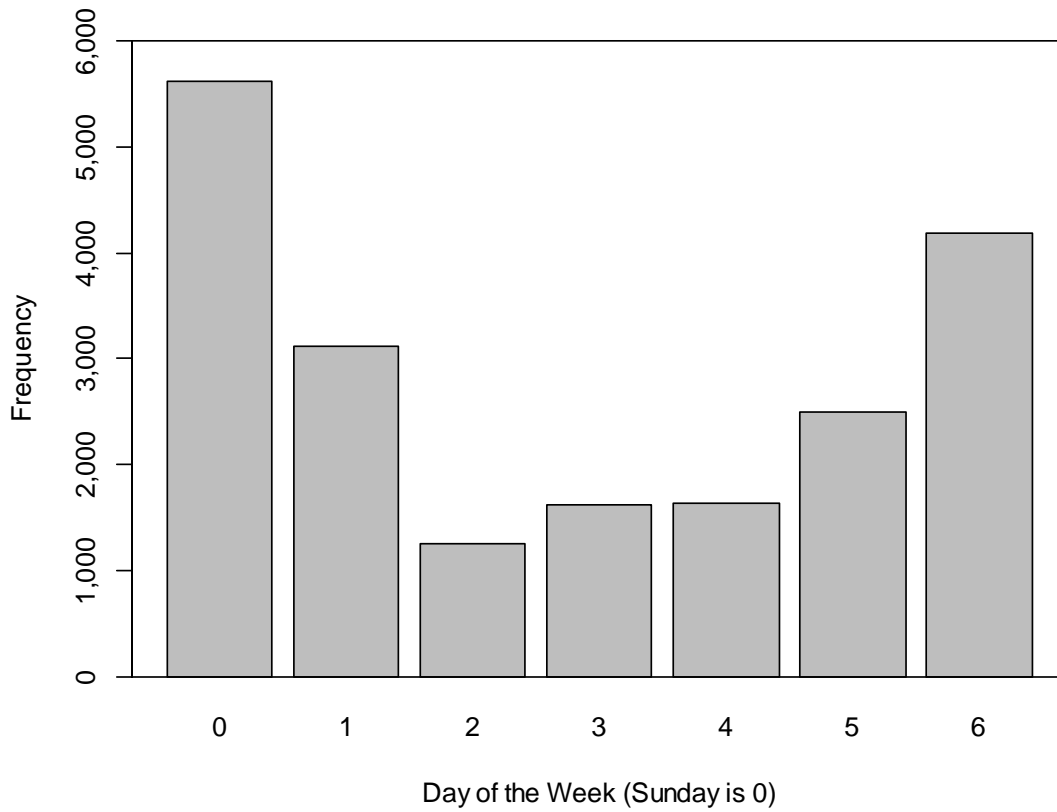


Figure 2.13: Histogram of day of the week in the final dataset.
(Note: Sunday is 0.)

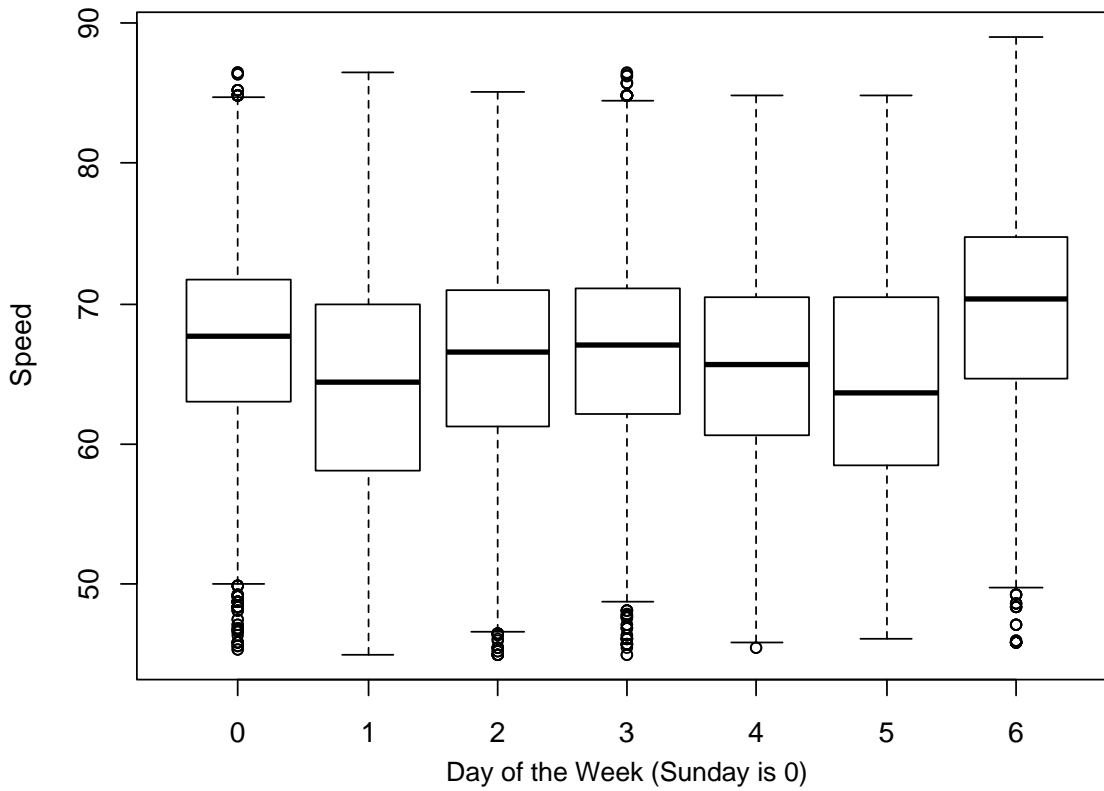


Figure 2.14: Box plot of speed versus day of the week in the final dataset.

Table 2.2: Mean and Standard Deviation of Speed on Different Days of the Week

| Day of the Week | Mean (mph) | Standard Deviation | Number of Observations |
|-----------------|------------|--------------------|------------------------|
| 0 (Sunday) | 67.2 | 6.5 | 5,619 |
| 1 (Monday) | 63.8 | 7.9 | 3,114 |
| 2 (Tuesday) | 66.0 | 8.0 | 1,250 |
| 3 (Wednesday) | 66.7 | 7.8 | 1,618 |
| 4 (Thursday) | 65.4 | 7.1 | 1,631 |
| 5 (Friday) | 64.0 | 7.8 | 2,494 |
| 6 (Saturday) | 69.5 | 7.4 | 4,183 |



Figure 2.15: Map of Caltrans districts.

2.4.6 Caltrans District

Caltrans district was introduced as an explanatory variable because it represents different regions within the state. It is reasonable to assume that drivers in different regions may have different driving behaviors: people in some regions may drive aggressively and others may drive defensively, and this can be associated with cultural and regional differences. However, because PeMS detectors are only distributed within selected Caltrans districts, and are concentrated along major urban freeways, only the base segments within these districts had speed observations and thus this variable cannot cover the whole state. The final dataset only covered eight of the 12 Caltrans districts, and generally excluded districts that do not have major urban freeways. Figure 2.15 shows a map of Caltrans districts. Figure 2.16 is a histogram of the *Caltrans district* variable, and Table 2.3 shows the mean value and the standard deviation of speed observations in each Caltrans district. Figure 2.17 shows a box plot of speed observations in different Caltrans districts. T-tests showed that there was a significant difference in speed observations between Caltrans districts at a 5 percent significance level. Therefore *Caltrans district* was included in the final explanatory variables.

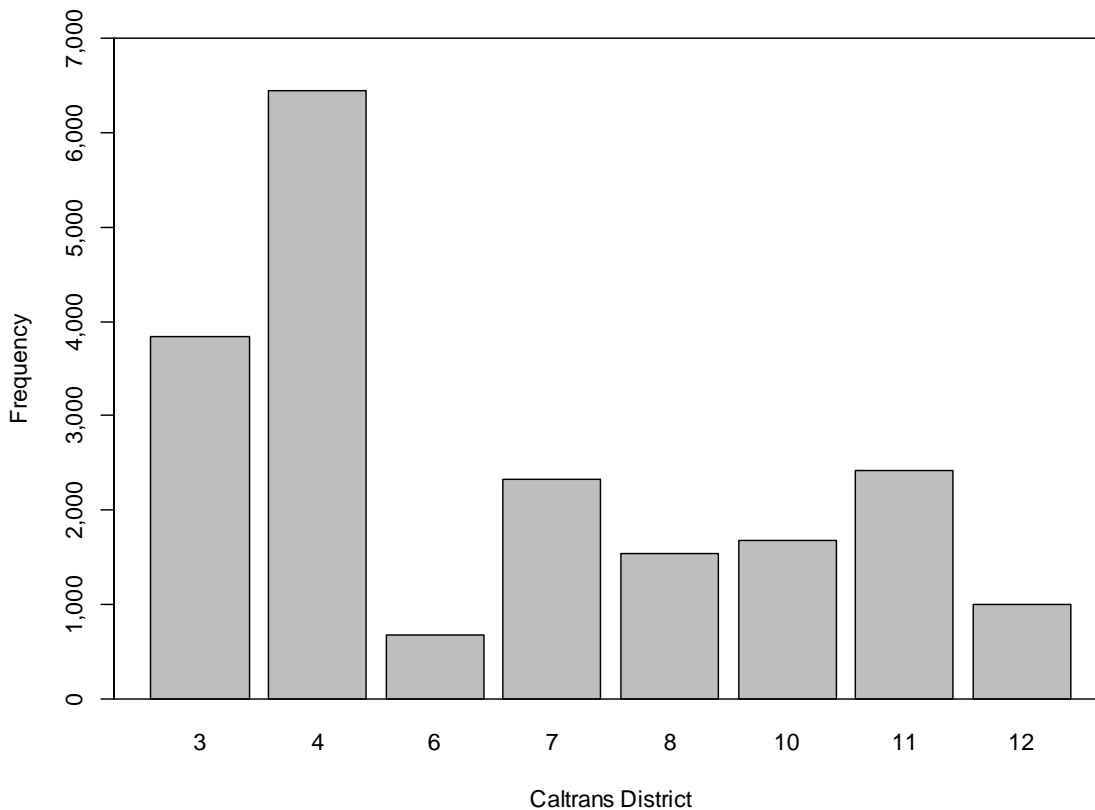


Figure 2.16: Histogram of Caltrans district in the final dataset.

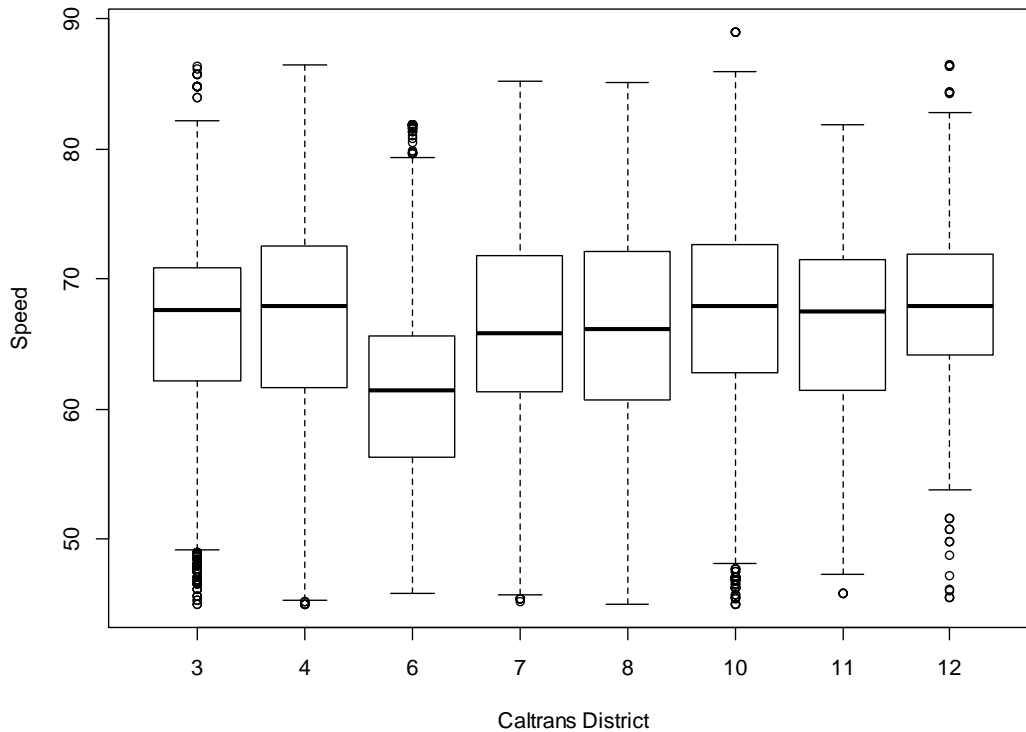


Figure 2.17: Box plot of speed versus Caltrans district in the final dataset.

Table 2.3: Mean and Standard Deviation of Speed in Different Caltrans Districts

| Caltrans District | Mean (mph) | Standard Deviation | Number of Observations |
|-------------------|------------|--------------------|------------------------|
| 3 | 66.4 | 6.8 | 3832 |
| 4 | 66.9 | 8.2 | 6,448 |
| 6 | 62.3 | 8.0 | 675 |
| 7 | 66.0 | 7.8 | 2,330 |
| 8 | 66.1 | 7.8 | 1,540 |
| 10 | 67.2 | 8.3 | 1,673 |
| 11 | 66.6 | 6.4 | 2,412 |
| 12 | 68.1 | 6.0 | 999 |

2.4.7 Speed Limit and Road Type

Table 2.4 and Table 2.5 show the statistics of the dataset based on speed limit and road type. Figure 2.18 and Figure 2.19 show the box plots. The mean speed on segments with a 70 mph speed limit (66.4 mph) was actually slightly lower than segments with a 65 mph speed limit (66.5 mph), while the standard deviation of speed on 70 mph roads was slightly higher (8.2 versus 7.6, respectively). The mean speed and standard deviation on rural segments (66.7 mph and 7.7, respectively) were slightly higher than urban roads (66.5 mph and 7.6, respectively). However, the number of observations on urban segments with a 65 mph speed limit was much larger than their counterparts. T-tests showed there was no significant difference in speed observations between the two speed limits and between the two road types (rural/urban roads). Therefore *speed limit* and *road type* were not included in the final explanatory variables.

Table 2.4: Mean and Standard Deviation of Speed in Different Speed Limit Segments

| Speed Limit (mph) | Mean (mph) | Standard Deviation | Number of Observations |
|-------------------|------------|--------------------|------------------------|
| 65 | 66.5 | 7.6 | 18,488 |
| 70 | 66.4 | 8.2 | 1,421 |

Table 2.5: Mean and Standard Deviation of Speed in Different Road Types

| Road Type | Mean (mph) | Standard Deviation | Number of Observation |
|-----------|------------|--------------------|-----------------------|
| Rural | 66.7 | 7.7 | 1,301 |
| Urban | 66.5 | 7.6 | 18,608 |

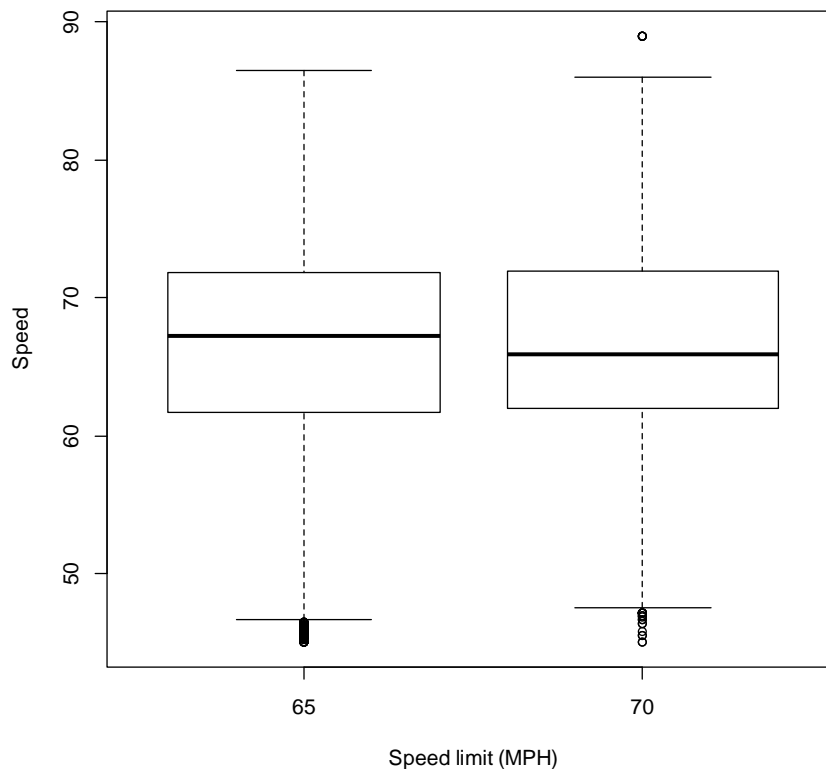


Figure 2.18: Box plot of speed versus speed limit in the final dataset.

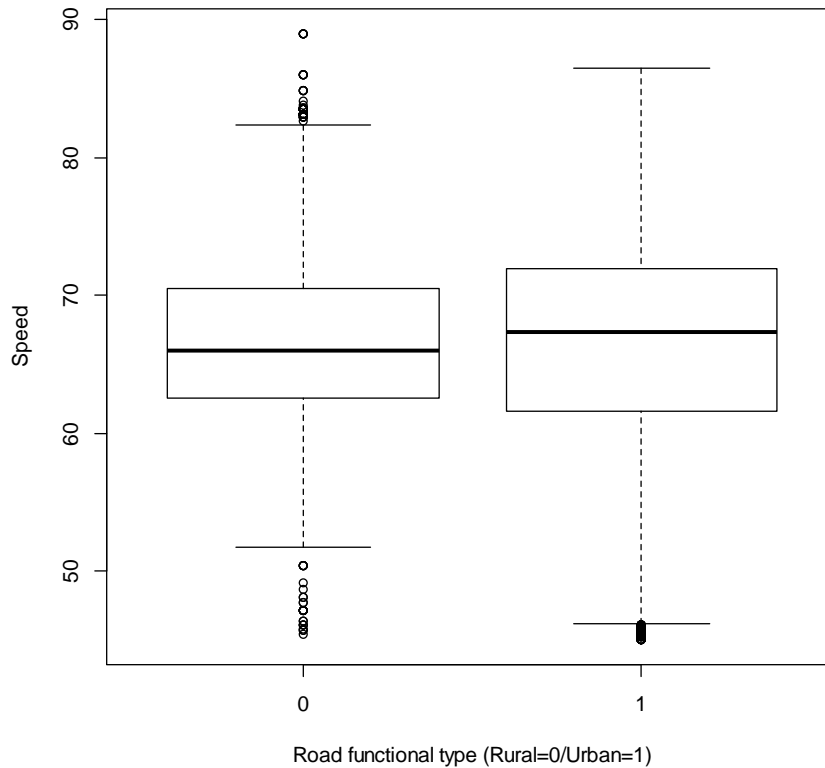


Figure 2.19: Box plot of speed versus road type (Rural/Urban) in the final dataset.

2.4.8 Correlation Between Selected Explanatory Variables

The final explanatory variables included in the model were *total number of lanes*, *lane number*, *Caltrans district*, *day of the week*, *gasoline price*, and *IRI*. Before using these variables in a linear regression model, it is necessary to examine the correlation between them. Table 2.6 shows the correlation coefficient between these variables (*Caltrans district* and *day of the week* were categorical variables and therefore are not included in this table). It can be seen that the correlation coefficients between these variables are very low, indicating it is safe to build a linear regression model using these variables, assuming that they are independent.

Table 2.6: Correlation Coefficients Between Selected Variables

| | <i>Lane</i> | <i>Total number of lanes</i> | <i>IRI</i> | <i>Gas price</i> |
|------------------------------|-------------|------------------------------|------------|------------------|
| <i>Lane</i> | 1 | | | |
| <i>Total number of lanes</i> | 0.252973 | 1 | | |
| <i>IRI</i> | 0.182037 | -0.0192 | 1 | |
| <i>Gas price</i> | -0.08606 | -0.0627 | 0.057701 | 1 |

3 RESULTS AND DISCUSSION

3.1 Modeling Using All Highway Data

As discussed in Section 0, the final explanatory variables included *total number of lanes*, *lane number*, *Caltrans district*, *day of the week*, *gasoline price* and *IRI*. To eliminate the impact from autocorrelation in the data, the dataset acquired in Section 2.3 was randomly divided into two subsets. The first set was used to develop the model and the second set was used to validate the model. Each subset had 9,954 observations.

The form of the model is shown in Equation 3.1. Different model forms were tested, including higher order polynomials, exponential, and logarithmic models, but they did not produce better results than this form. Another form of the model, the Limiting Speed Model, could not be tested in this study because IRI on California freeways never reaches the levels that start to limit driving speed, about 6 m/km (378 in./mi.) according to the *HDM-4* study (15).

The coefficients developed from the first set of data are shown in Table 3.1. It should be noted that because *CaltransDistrict* and *DayOfWeek* are categorical variables, these terms in Equation 3.1 are calculated by multiplying 1 by the corresponding regression coefficients of the Caltrans district or the day of the week that is being modeled. For example, if Caltrans District 4 is being modeled, then the term $d \times District$ is calculated as 0.86038×1 , where 0.86038 is the coefficient for Caltrans District 4 and 1 represents the dummy variable for Caltrans District 4.

$$FFS = a + b \times NbrOfLanes + c \times Lane + d \times CaltransDistrict + e \times DayOfWeek + f \times GasPrice + g \times IRI \quad (3.1)$$

where:

| | |
|-------------------------|--|
| <i>a</i> | is the intercept of the linear regression model |
| <i>b, c, d, e, f, g</i> | are the coefficients of each variable |
| <i>FFS</i> | is the estimated free-flow speed in miles per hour (mph) |
| <i>NbrOfLanes</i> | is the total number of lanes |
| <i>Lane</i> | is the lane number |
| <i>CaltransDistrict</i> | is the Caltrans district, categorical variable |
| <i>DayOfWeek</i> | is the day of the week, categorical variable |
| <i>GasPrice</i> | is the gasoline price in dollars per gallon (\$/gal) |
| <i>IRI</i> | is the IRI value with the unit m/km |

Table 3.1: Coefficients of Model Developed From All Highway Data

| Variable^{1,2} | Coefficient | Std. Error | t value | Pr(> t)³ |
|-------------------------------|--------------------|-------------------|----------------|--------------------------------|
| (Intercept) | 67.34079 | 0.48772 | 138.073 | < 2e-16 |
| <i>NbrOfLanes</i> | 2.32734 | 0.07179 | 32.421 | < 2e-16 |
| <i>Lane</i> | -4.63853 | 0.05507 | -84.226 | < 2e-16 |
| <i>CaltransDistrict 4</i> | 0.86038 | 0.16363 | 5.258 | 1.49E-07 |
| <i>CaltransDistrict 6</i> | -4.80168 | 0.3457 | -13.89 | < 2e-16 |
| <i>CaltransDistrict 7</i> | 0.69942 | 0.22833 | 3.063 | 0.0022 |
| <i>CaltransDistrict 8</i> | 0.68978 | 0.24846 | 2.776 | 0.00551 |
| <i>CaltransDistrict 10</i> | 0.27785 | 0.24021 | 1.157 | 0.24742 |
| <i>CaltransDistrict 11</i> | 1.7542 | 0.24025 | 7.302 | 3.06E-13 |
| <i>CaltransDistrict 12</i> | 2.38015 | 0.27804 | 8.561 | < 2e-16 |
| <i>DayOfWeek Sunday</i> | 4.86765 | 0.19887 | 24.476 | < 2e-16 |
| <i>DayOfWeek Tuesday</i> | 2.24796 | 0.26297 | 8.548 | < 2e-16 |
| <i>DayOfWeek Wednesday</i> | 1.88763 | 0.2498 | 7.557 | 4.50E-14 |
| <i>DayOfWeek Thursday</i> | 1.86893 | 0.25322 | 7.381 | 1.70E-13 |
| <i>DayOfWeek Friday</i> | 2.58722 | 0.23153 | 11.174 | < 2e-16 |
| <i>DayOfWeek Saturday</i> | 5.35312 | 0.20177 | 26.531 | < 2e-16 |
| <i>GasPrice</i> | -0.54254 | 0.09405 | -5.769 | 8.24E-09 |
| <i>IRI</i> | -0.30281 | 0.07433 | -4.074 | 4.66E-05 |

Residual standard error: 5.45 on 9,936 degrees of freedom; Adjusted R-Squared: 0.4836.

Notes:

- 1: District 3 is used as a reference level, meaning District 3 is embraced in the model. When District 3 is calculated, the *CaltransDistrict* variable is 0. This situation is similar with the *DayOfWeek* variable, which uses Monday as a reference level.
- 2: Because of the coverage of sample points, only Caltrans Districts 3, 4, 6, 7, 8, 10, 11, and 12 were included in this model. These districts correspond to the following regions: the Sacramento area and rural/mountain counties (3), the San Francisco Bay Area (4), Fresno and rural surroundings (6), Los Angeles/Ventura (7), Riverside/San Bernardino and rural areas (8), Stockton/Modesto and rural areas (10), San Diego/Imperial (11), and Orange County (12).
- 3: A value smaller than 0.05 is considered significant in this study.

Figure 3.1 shows the validation results using the second set of data. The adjusted R-squared between the fitted value using the model developed and the actual value was 0.5029, very close to the R-squared from the original model, indicating autocorrelation has a small impact in this model and the model itself is valid.

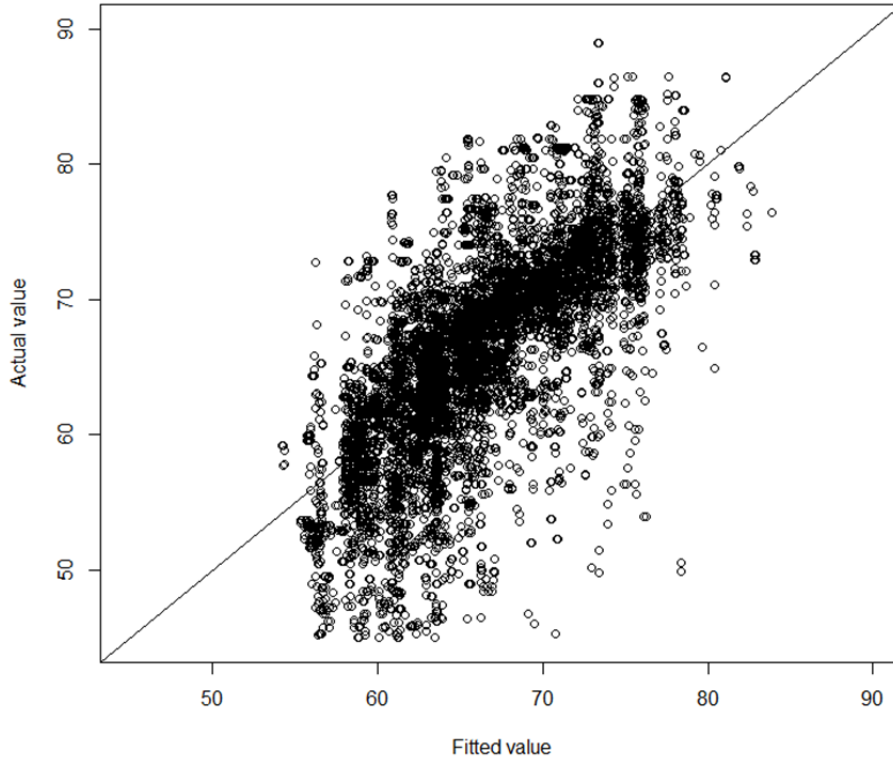


Figure 3.1: Plot of fitted values versus actual values using the validation dataset.

This model yielded an adjusted R-squared of 0.4836, meaning the explanatory variables selected can explain about 50 percent of the total variance. An analysis using a random effect model revealed that most of the variance of the random effects can be attributed to each specific segment, indicating that segment-specific characteristics, as opposed to the six explanatory variables selected, may substantially affect the overall free-flow speed modeled in this study.

Because the model developed using this set of data had a relatively low R-squared value, diagnostic plots of the regression were made to investigate whether or not there are observations with a large influence on the analysis (see Figure 3.2). There were no points that were consistently extreme in all of the diagnostic plots, so it can be concluded that the model assumption was correct and that there were no observations with a very large influence on the result. Figure 3.3 shows the relationship between the model residuals and each explanatory variable, where the line in each figure is the fitting result between the residuals and the explanatory variables. It shows that the residuals stayed constant and that the average residual was 0 when the explanatory variable changed. This indicates that there were no higher order relationships between the response variable (free-flow speed) and the selected explanatory variables, and using a linear regression model is appropriate in this study.

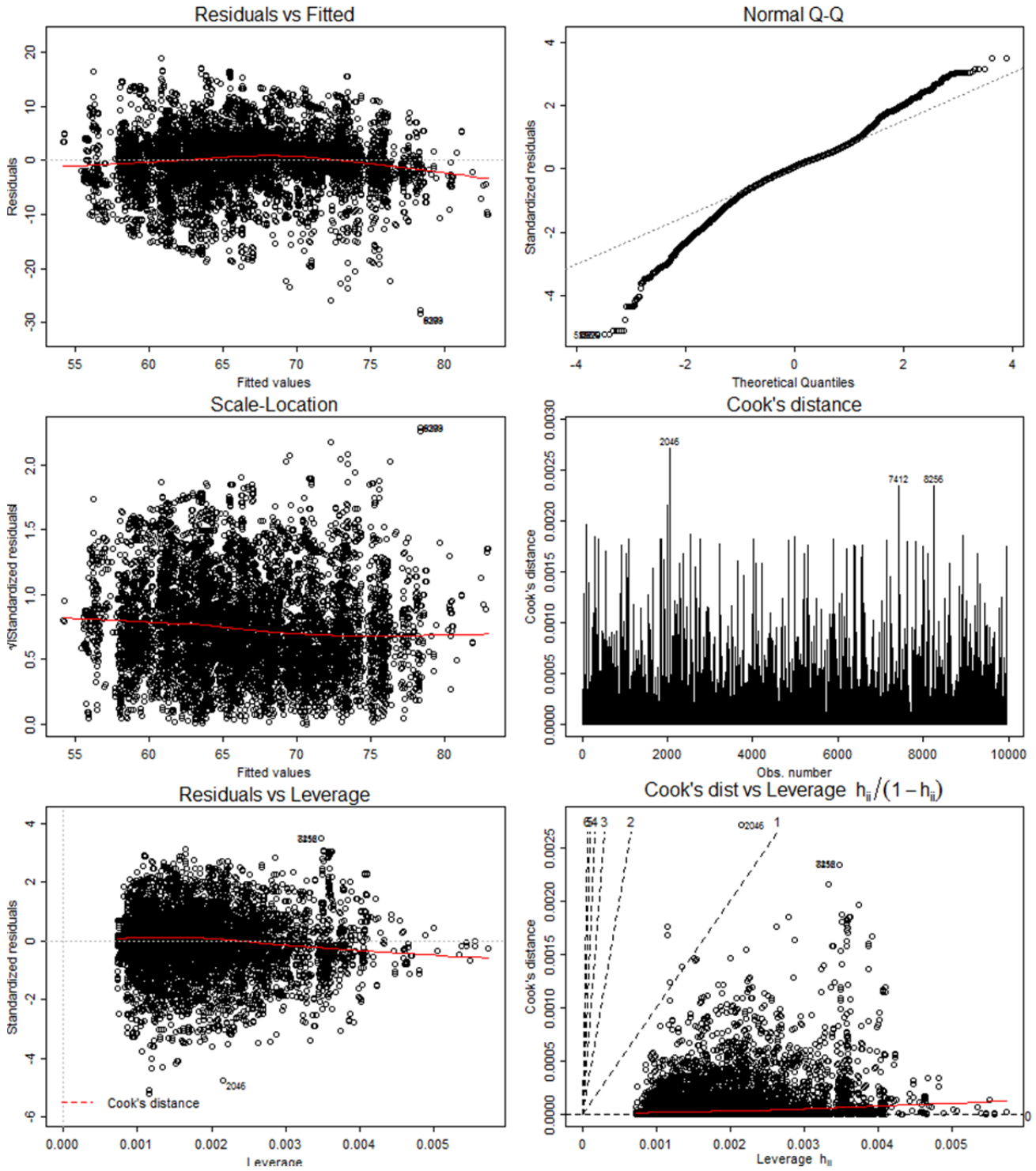


Figure 3.2: Diagnostic plots of the regression model based on all highway data.

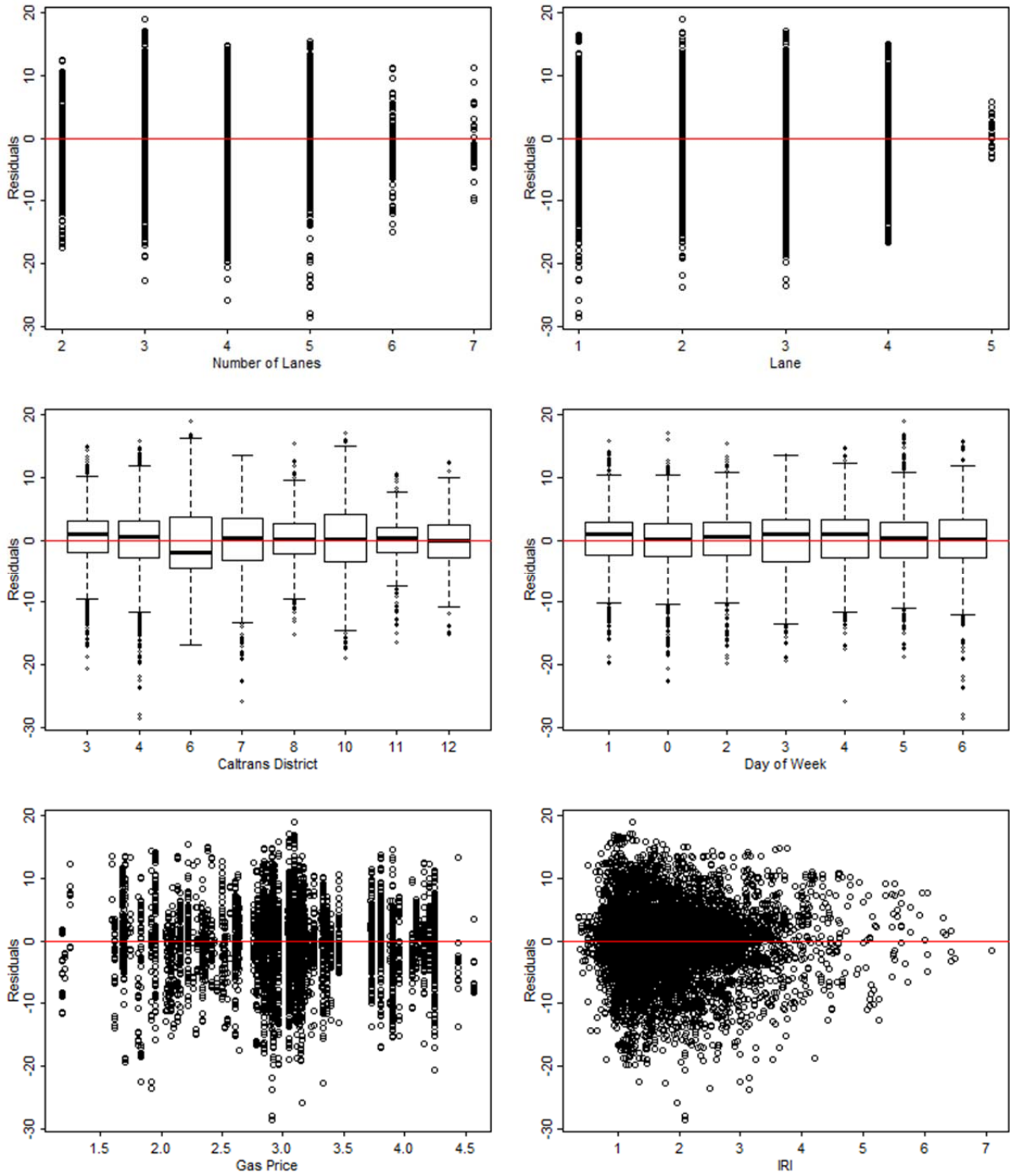


Figure 3.3: Residual versus each explanatory variable.

Analysis of Variance (ANOVA) results are shown in Table 3.2. Using a 0.05 significance level, all the variables selected in this study were considered significant in this model. It was found that the lane number explained the greatest variance, and gasoline price and IRI accounted for the least variance. Although this model cannot address most of the variance in free-flow speed, it can still provide some insights on the free-flow speed, as discussed below.

Table 3.2: ANOVA Results of Model Developed from All Highway Data

| Variable | Degree of Freedom | Sum Sq | Mean Sq | F value | Pr(>F) ¹ |
|-------------------------|-------------------|---------|---------|-----------|---------------------|
| <i>NbrOfLanes</i> | 1 | 5397 | 5397 | 181.665 | < 2.2e-16 |
| <i>Lane</i> | 1 | 228,746 | 22,8746 | 7,700.061 | < 2.2e-16 |
| <i>CaltransDistrict</i> | 7 | 12,775 | 1,825 | 61.431 | < 2.2e-16 |
| <i>DayOfWeek</i> | 6 | 29,075 | 4,846 | 163.118 | < 2.2e-16 |
| <i>GasPrice</i> | 1 | 951 | 951 | 32.002 | 1.58E-08 |
| <i>IRI</i> | 1 | 493 | 493 | 16.596 | 4.66E-05 |
| <i>Residuals</i> | 9,936 | 295,169 | 30 | | |

Note:

1: A value smaller than 0.05 is considered significant in this study.

The coefficient of the variable *Lane* is negative, which matches the fact that vehicles driving on the inner lanes are faster than those on the outer lanes. Also, the free-flow speed is increased as the *total number of lanes* increases, as expected. This matches with the total number of lanes factor discussed in the *Highway Capacity Manual 2000 (13)*. A larger total number of lanes indicates better maneuverability for vehicles, which can lead to a higher free-flow speed. The result also shows that driving speeds are faster during weekends and holidays than on weekdays (the coefficients of *DayOfWeek* Sunday and *DayOfWeek* Saturday are higher than the others). The fact that people in different districts have different free-flow speeds can be attributed to cultural differences: there are places where drivers are more aggressive (30). On average, drivers in District 6 go the slowest while those in District 12 drive the fastest. This may be attributed to the fact that the freeways in District 6 carry a high percentage of heavy trucks, which have a lower speed. Trucks are usually required to drive on the outside lanes on freeways and have a lower speed limit. However, due to the unavailability of data (PeMS does not differentiate the speed of cars and trucks), this study was not able to differentiate observations of speed between cars and trucks. This may be one of the reasons for the low R-squared value and presents a major limitation of this study. With regard to the price of gasoline, although it only addressed a very small portion of the variance, it can be seen that when the price of gasoline is high people tend to drive more slowly, although the change is very small considering that the observations are mostly from freeways and under free-flow conditions: an approximate 0.864 km/h (0.54 mph) decrease in free-flow speed when the price of gasoline increased by \$1/gallon. The finding that people drive at a reduced speed when the price of gasoline is increased has also been observed in other recent research studies, where it was found that the decrease in speed in response to \$1/gallon

increase in the price of gasoline can vary between 0.06 and 2.4 km/h (0.04 to 1.5 mph) (31, 32) depending on the time of the day, the wages of the drivers, and their value of time. It could be that this change in behavior is somewhat temporary, that is, drivers slow down when there are sudden increases in the gas prices but slowly revert to the natural free-flow speed as they become accustomed to the higher gas price. This has not been investigated here.

Pavement roughness (IRI) accounts for a very small portion of the total variance. One unit of IRI change (1 m/km) only leads to about a 0.48 km/h (0.30 mph) change in free-flow speed. Considering that the IRI change from most freeway pavement treatments in California is less than 3 m/km (192 inches/mile), drivers will not go substantially faster after a pavement treatment. According to a modeling study sponsored by FHWA, even a speed change of 1.6 km/h (1 mph), which covers the IRI change (3 m/km) from most treatments, only leads to about 0.3 percent change in fuel economy at 104 km/h (65 mph) (33), which is much smaller than the improved fuel economy from the same change of IRI (34). Therefore, the benefits of energy saving gained from the reduced rolling resistance will not be substantially offset by the increased vehicle operating speeds.

Due to the limitations of PeMS, this study could not differentiate between the observed speeds of passenger cars and heavy duty trucks. However, an examination using the same procedure as above was used on locations with truck compositions lower and higher than 10 percent (representing a freeway segment with a low truck volume and high truck volume, respectively) to estimate the difference in impact of different vehicle types. This examination produced results showing an IRI coefficient of -0.142 on low truck volume segments and of -0.916 on high truck volume segments. The difference in the impact on the different vehicle types might be attributed to the differences in vehicle configuration and suspension system: passenger cars usually have better suspension systems and can provide better ride quality. Thus, passenger cars are potentially more resilient, so pavement roughness has less impact on their speed than it does on trucks. However, no consensus could be found among existing studies that explicitly differentiated passenger cars and trucks: two studies found the impact of pavement roughness on trucks to be more substantial (7, 20), while one other drew a different conclusion (23), although the difference may be due to the use of different types of trucks in these studies. While this inconsistency calls for further research, the results in this study were on the same order of magnitude. This result indicates that pavement roughness, as indicated by IRI, may have very limited impact on speed and therefore on the greenhouse gas emissions and energy consumption associated with speed.

3.2 Modeling Using Subsets of the Data

Because the model developed in Section 3.1 had a relatively low R-squared value, a further analysis using several subsets of all the data was performed to examine if there were regional differences on the impact of

pavement roughness on free-flow speed. In this analysis, only the interstate freeway segments were selected from the final datasets developed in Section 2.3. This was to ensure that the segments to be analyzed met interstate highway geometric standards as well as others, such as lane width, median width, and shoulder width. Further, segments in mountainous areas were excluded from the dataset to eliminate the impacts from vertical gradient. With the regional differences, the following three subsets of data were yielded: Interstates in Northern California, Interstates in Southern California, and Interstates in Central California. Given the fact that autocorrelation did not have a big impact on the complete dataset (tested in Section 3.1) and the relatively small number of observations in each subset, the autocorrelation test was not performed on each subset of data.

The Interstates in Northern California in this study included I-80 and its nearby auxiliary interstate highways, such as Interstate 280 (I-280) and Interstate 580 (I-580). Interstate 5 (I-5) was excluded because it crosses northern and southern California, and thus carries much interregional traffic. With the inclusion of the I-5 observations, this subset may not reflect the regional characteristics of Northern California that well. The Interstates in Southern California included Interstate 10 (I-10), Interstate 15 (I-15), Interstate 40 (I-40), and the nearby auxiliary interstate highways that connect to them, such as Interstate 710 (I-710) and Interstate 110 (I-110). I-5 was excluded from this subset for the same reason that was noted above. The Interstates in Central California in this study included I-5 in Caltrans Districts 6 and 10.

Using Equation 3.1, the coefficients from the model using the three subsets are shown in Table 3.3, Table 3.5, and Table 3.7, respectively. The ANOVA results are shown in Table 3.4, Table 3.6, and Table 3.8, respectively.

The Northern California Interstate subset of data (a total of 2,849 observations) yielded an adjusted R-squared of 0.5819, meaning there was still a large portion of the variance that was not addressed by the selected variables. The coefficients of most variables were similar to those derived from the full dataset. *IRI* also showed up as a significant variable. However, the coefficient of *IRI* yielded from this dataset was positive, meaning people drive faster when the road becomes rougher, which is counterintuitive. This may be because the *IRI* variable here was mixed with the impacts from other factors that were not selected as explanatory variables, such as route and time. Nonetheless, in either situation the impact from *IRI* on speed was still very small: 1 m/km of *IRI* change results in about a 0.33 mph (0.53 km/h) increase in speed, which has almost no effect on pollutant emissions or energy consumption.

The Southern California Interstate subset of data had 1,860 observations. The model using this subset yielded an adjusted R-squared of 0.7293, meaning the selected variables explained more variance in the data compared to those in the Northern California Interstate model. Compared to Northern California, this may be attributed to the

fact that the interstates in Southern California have a relatively consistent interchange and ramp density across the whole area. According to the *HCM 2010*, ramp density has the biggest impact on highway free-flow speed. However, due to the unavailability of data, this factor could not be explicitly reflected in this study. Coefficients from most variables exhibited a similar pattern to those derived using the statewide data. The impact from IRI on speed was also of a similar order of magnitude (-0.303 for the state average compared to -0.428 for the Southern California Interstate), which means that on interstates in Southern California, IRI has very small impact on free-flow speed.

The only interstate highway in Central California is I-5, except for very small segments of a few connectors in Caltrans Districts 6 and 10. Because the final dataset contained no observations on I-5 from Caltrans District 6, the *CaltransDistrict* variable was removed from the variable list. This subset had a total of 1,188 observations. Using this subset, the model yielded an adjusted R-squared of 0.4482, indicating that the variables selected cannot explain most of the variance. In this subset of data, *IRI* did not show up as a significant variable (using 0.05 as the significance level), although there was a slight tendency toward slower driving on rougher roads (a negative coefficient on IRI). As with the other subsets, the impact from IRI on free-flow speed was very small here. All these results show that pavement roughness, as indicated by IRI, has a very small impact on free-flow speed in both the complete set of data and in any subset of the data: a one unit change in IRI (1 m/km) led to less than a 0.5 mph change in free-flow speed, and the predominant range of IRI values, between 1 and 4 m/km, resulted in average speed differences of about 1.5 mph (2.4 km/h).

**Table 3.3: Coefficients of Model Developed from Northern California Interstate Data
(2,849 data points)**

| Variable^{1,2} | Coefficient | Std. Error | t value | Pr(> t)³ |
|-------------------------------|--------------------|-------------------|----------------|--------------------------------|
| <i>(Intercept)</i> | 63.5488 | 0.8204 | 77.46 | < 2e-16 |
| <i>NbrOfLanes</i> | 2.5492 | 0.1219 | 20.912 | < 2e-16 |
| <i>Lane</i> | -4.843 | 0.1003 | -48.281 | < 2e-16 |
| <i>CaltransDistrict 4</i> | 0.8083 | 0.2947 | 2.743 | 0.00613 |
| <i>DayOfWeek Sunday</i> | 5.542 | 0.3343 | 16.577 | < 2e-16 |
| <i>DayOfWeek Tuesday</i> | 3.0414 | 0.4609 | 6.598 | 4.95E-11 |
| <i>DayOfWeek Wednesday</i> | 4.0836 | 0.3218 | 12.691 | < 2e-16 |
| <i>DayOfWeek Thursday</i> | 4.4283 | 0.3743 | 11.832 | < 2e-16 |
| <i>DayOfWeek Friday</i> | 2.3934 | 0.3159 | 7.576 | 4.78E-14 |
| <i>DayOfWeek Saturday</i> | 4.923 | 0.3158 | 15.588 | < 2e-16 |
| <i>GasPrice</i> | -0.1903 | 0.1616 | -1.178 | 0.23903 |
| <i>IRI</i> | 0.3291 | 0.1103 | 2.984 | 0.00287 |

Residual standard error: 4.878 on 2837degrees of freedom; Adjusted R-Squared: 0.5819.

Notes:

- 1: District 3 is used as a reference level, meaning District 3 is embraced in the model. When District 3 is calculated, the *CaltransDistrict* variable is 0. This situation is similar with the *DayOfWeek* variable, which uses Monday as a reference level.
- 2: The interstates in Northern California only pass through Caltrans Districts 3 and 4.
- 3: A value smaller than 0.05 is considered significant in this study.

Table 3.4: ANOVA Results of Model Developed from Northern California Interstate Data

| Variable | Degree of freedom | Sum Sq | Mean Sq | F value | Pr(>F)¹ |
|-------------------------|--------------------------|---------------|----------------|----------------|------------------------------|
| <i>NbrOfLanes</i> | 1 | 6,101 | 6101 | 256.4313 | < 2.2e-16 |
| <i>Lane</i> | 1 | 78,555 | 78,555 | 3301.613 | < 2.2e-16 |
| <i>CaltransDistrict</i> | 1 | 68 | 68 | 2.8417 | 0.091958 |
| <i>DayOfWeek</i> | 6 | 9,607 | 1,601 | 67.2962 | < 2.2e-16 |
| <i>GasPrice</i> | 1 | 33 | 33 | 1.4053 | 0.235941 |
| <i>IRI</i> | 1 | 212 | 212 | 8.9051 | 0.002868 |
| Residuals | 2,837 | 67,500 | 24 | | |

Note:

- 1: A value smaller than 0.05 is considered significant in this study.

**Table 3.5: Coefficients of Model Developed from Southern California Interstate Data
(1,860 data points)**

| Variable^{1,2} | Coefficient | Std. Error | t value | Pr(> t)³ |
|-------------------------------|--------------------|-------------------|----------------|--------------------------------|
| (Intercept) | 58.72292 | 1.10413 | 53.185 | < 2e-16 |
| <i>NbrOfLanes</i> | 2.16973 | 0.1009 | 21.503 | < 2e-16 |
| <i>Lane</i> | -3.36456 | 0.09961 | -33.778 | < 2e-16 |
| <i>CaltransDistrict 8</i> | -1.50804 | 0.48145 | -3.132 | 0.00176 |
| <i>CaltransDistrict 11</i> | -0.35697 | 0.80601 | -0.443 | 0.6579 |
| <i>CaltransDistrict 12</i> | -0.27977 | 0.51449 | -0.544 | 0.58666 |
| <i>DayOfWeek</i> Sunday | 13.70775 | 0.49314 | 27.797 | < 2e-16 |
| <i>DayOfWeek</i> Tuesday | 11.56319 | 0.53582 | 21.581 | < 2e-16 |
| <i>DayOfWeek</i> Wednesday | 9.23137 | 0.60809 | 15.181 | < 2e-16 |
| <i>DayOfWeek</i> Thursday | 10.53718 | 0.79882 | 13.191 | < 2e-16 |
| <i>DayOfWeek</i> Friday | 9.79323 | 0.48049 | 20.382 | < 2e-16 |
| <i>DayOfWeek</i> Saturday | 13.54611 | 0.49736 | 27.236 | < 2e-16 |
| <i>GasPrice</i> | -0.44128 | 0.22731 | -1.941 | 0.05237 |
| <i>IRI</i> | -0.42786 | 0.16208 | -2.64 | 0.00836 |

Residual standard error: 3.776 on 1,846 degrees of freedom; Adjusted R-Squared: 0.7293.

Notes:

- 1: District 7 is used as a reference level, meaning District 7 is embraced in the model. When District 7 is calculated, the *CaltransDistrict* variable is 0. This situation is similar with the *DayOfWeek* variable, which uses Monday as a reference level.
- 2: The interstates in Southern California pass through Caltrans Districts 7, 8, 10, 11, and 12.
- 3: A value smaller than 0.05 is considered significant in this study.

Table 3.6: ANOVA Results of Model Developed from Southern California Interstate Data

| Variable | Degree of freedom | Sum Sq | Mean Sq | F value | Pr(>F)¹ |
|-------------------------|--------------------------|---------------|----------------|----------------|------------------------------|
| <i>NbrOfLanes</i> | 1 | 7860 | 7860 | 551.3467 | < 2.2e-16 |
| <i>Lane</i> | 1 | 43,872 | 43,872 | 3077.51 | < 2.2e-16 |
| <i>CaltransDistrict</i> | 3 | 5,162 | 1,721 | 120.7069 | < 2.2e-16 |
| <i>DayOfWeek</i> | 6 | 14,543 | 2,424 | 170.0259 | < 2.2e-16 |
| <i>GasPrice</i> | 1 | 29 | 29 | 2.0172 | 0.155692 |
| <i>IRI</i> | 1 | 99 | 99 | 6.9688 | 0.008364 |
| Residuals | 1,846 | 26,316 | 14 | | |

Note:

- 1: A value smaller than 0.05 is considered significant in this study.

Table 3.7: Coefficients of Model Developed from Central California Interstate Data (1,188 data points)

| Variable ^{1,2} | Coefficient | Std. Error | t value | Pr(> t) ³ |
|----------------------------|-------------|------------|---------|-----------------------|
| (Intercept) | 80.0719 | 2.325 | 34.439 | < 2e-16 |
| <i>NbrOfLanes</i> | 1.6545 | 0.2078 | 7.961 | 4.00E-15 |
| <i>Lane</i> | -4.4948 | 0.2531 | -17.759 | < 2e-16 |
| <i>DayOfWeek</i> Sunday | 5.6753 | 1.4484 | 3.918 | 9.43E-05 |
| <i>DayOfWeek</i> Tuesday | -12.2924 | 1.7077 | -7.198 | 1.08E-12 |
| <i>DayOfWeek</i> Wednesday | -0.4405 | 1.1529 | -0.382 | 0.7025 |
| <i>DayOfWeek</i> Friday | -0.5557 | 1.6133 | -0.344 | 0.7306 |
| <i>DayOfWeek</i> Saturday | 8.6199 | 1.4439 | 5.97 | 3.14E-09 |
| <i>GasPrice</i> | -4.8036 | 1.2017 | -3.997 | 6.81E-05 |
| <i>IRI</i> | -0.4774 | 0.2861 | -1.668 | 0.0955 |

Residual standard error: 5.782 on 1,178 degrees of freedom; Adjusted R-Squared: 0.4482.

Notes:

- 1: Monday is used as a reference level, meaning Monday is embraced in the model. When Monday (*DayOfWeek* 1) is calculated, the *DayOfWeek* variable is 0.
- 2: The interstates in Southern California only pass through Caltrans District 10, so *CaltransDistrict* was removed from the explanatory variables. Also, there were no observations on Thursday, so *DayOfWeek* 4 does not appear in the variable list.
- 3: A value smaller than 0.05 is considered significant in this study.

Table 3.8: ANOVA Results of Model Developed from Central California Interstate Data

| Variable | Degree of freedom | Sum Sq | Mean Sq | F value | Pr(>F) ¹ |
|-------------------|-------------------|--------|----------|----------|---------------------|
| <i>NbrOfLanes</i> | 1 | 525 | 524.9 | 15.7011 | 7.86E-05 |
| <i>Lane</i> | 1 | 17,635 | 17,634.7 | 527.5461 | < 2.2e-16 |
| <i>DayOfWeek</i> | 5 | 13,520 | 2,704 | 80.8923 | < 2.2e-16 |
| <i>GasPrice</i> | 1 | 753 | 752.6 | 22.514 | 2.34E-06 |
| <i>IRI</i> | 1 | 93 | 93.1 | 2.7837 | 0.0955 |
| Residuals | 1,178 | 39,378 | 33.4 | | |

Note:

- 1: A value smaller than 0.05 is considered significant in this study.

As with any model development, a potential statistical bias associated with correlation and causality may exist in this study: a higher speed observation may result from a higher road class, which then receives more frequent treatments and is maintained at a smoother level, as opposed to directly from lower roughness. However, two factors made this bias insignificant in this study: 1) In the analysis period of this study, Caltrans performed pavement treatments primarily based on cracking level and traffic level, and not based on roughness (35). Roughness is a lagging indicator of pavement condition compared with cracking level. In other words, a pavement can be severely cracked but still maintain a certain level of roughness. Therefore, the class of roads (with respect to the design speed) does not play a major role in determining the priority with which a highway receives treatment and therefore does not affect the resultant improved smoothness. 2) As noted earlier, all

locations selected for this study were on freeways (restricted-access high-speed highways), so they will have the highest maintenance service level Caltrans assigns, ensuring that they receive priority in any maintenance funding (35). Further, when examining the subsets of the data (i.e., Northern California, Central California, and Southern California), only interstate freeways were selected in order to investigate whether freeway class has any impact on the results. The results showed that the impact from roughness is almost the same between just interstates and among all freeways. Therefore, it is safe to conclude that in this study, the associated change in speed is due to the change in roughness. Examination of the distributions of roughness among segments with different levels of traffic also indicated that the roughness distributions were similar across almost all segments.

4 CONCLUSIONS

Using IRI as an indicator, this study examined the impact of pavement roughness on free-flow speeds on California freeways by building a linear regression model to estimate free-flow speed. The explanatory variables included *total number of lanes*, *lane number*, *Caltrans district*, *day of the week*, *gasoline price*, and *IRI*. Only data from restricted-access roads (freeways) and their subsets were used to build the model.

The results show that IRI has a very small impact on free-flow speed. In most situations, a one unit change of IRI (1 m/km = 63 inches/mile) only leads to about a 0.48 to 0.64 km/h (0.3 to 0.4 mph) change in free-flow speed, which has a very small effect on pollutant emissions, including CO₂, or the energy consumption associated with speed. Given the fact that the IRI change from most pavement treatments is less than 3 m/km (192 in./mi), the results indicate that drivers will not go substantially faster after a pavement M&R treatment activity on California freeway, and therefore the energy-saving benefits gained from the reduced rolling resistance will not be offset by the marginally increased vehicle operating speeds. However, it should be emphasized that this conclusion was drawn based on the IRI range in this study (90 percent of the records have an IRI of 3 m/km or lower), the change of IRI in this study (90 percent of the records have an IRI change of 2 m/km or lower), and the coverage of other variables as shown in the previous sections. The conclusion from this study cannot be generalized to very rough roads or to larger IRI changes.

However, efforts to develop a good model for predicting free-flow speed were unsuccessful. The model Interstates in Southern California gave the best result with an adjusted R-squared of 0.72. For the rest of the state's regions, the selected explanatory variables could only explain about half of the total variance, which means that there are other variables with a much more significant impact on free-flow speed, such as *density of interchange*, that were not covered in this study. Another major limitation of this study was the lack of available data to address different types of vehicles separately (i.e., cars and trucks). These factors should be addressed in future studies.

REFERENCES

1. Santero, N. J., and A. Horvath. Global Warming Potential of Pavements. *Environmental Research Letters*, Vol. 4, No. 3, 2009, pp. 1-7.
2. Zhang, H., M. D. Lepech, G. A. Keoleian, S. Z. Qian, and V. C. Li. Dynamic Life-Cycle Modeling of Pavement Overlay Systems: Capturing the Impacts of Users, Construction, and Roadway Deterioration. *Journal of Infrastructure Systems*, Vol. 16, No. 4, 2010, pp. 299-309.
3. Wang, T., I. S. Lee, A. Kendall, J. Harvey, E. B. Lee, and C. Kim. Life Cycle Energy Consumption and GHG Emission from Pavement Rehabilitation with Different Rolling Resistance. *Journal of Cleaner Production*, Vol. 33, 2012, pp. 86-96.
4. Lidicker, J., N. Sathaye, S. Madanat, and A. Horvath. Pavement Resurfacing Policy for Minimization of Life-Cycle Costs and Greenhouse Gas Emissions. *Journal of Infrastructure Systems*, Vol. 19, No. 2, 2013, pp. 129-137.
5. Yu, B., and Q. Lu. Life Cycle Assessment of Pavement: Methodology and Case Study. *Transportation Research Part D: Transport and Environment*, Vol. 17, No. 5, 2012, pp. 380-388.
6. Hammarström, U., J. Eriksson, R. Karlsson, and M.-R. Yahya. *Rolling Resistance Model, Fuel Consumption Model and The Traffic Energy Saving Potential of Changed Road Surface Conditions*. VTI rapport 748A. www.vti.se/en/publications/rolling-resistance-model-fuel-consumption-model-and-the-traffic-energy-saving-potential-of-changed-road-surface-conditions. (Accessed Oct. 28, 2013.) Report published by Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden, 2012.
7. Ihs, A., and H. Velin. *Vägytans inverkan på fordonshastigheter (Surface Roughness Effects on Vehicle Speeds)*. VTI notat 40-2002. www.vti.se/en/publications/pdf/vagyttans-inverkan-pa-fordonshastigheter-data-fran-19921999.pdf. (Accessed Oct. 28, 2013.) Report published by Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden, 2002.
8. U.S. Department of Energy and U.S. Environmental Protection Agency. Driving More Efficiently. www.fueleconomy.gov/feg/driveHabits.shtml. (Accessed Aug. 12, 2012.)
9. Barth, M., and K. Boriboonsomsin. Real-World Carbon Dioxide Impacts of Traffic Congestion. *Transportation Research Record*, No. 2058, 2008, pp. 163-171.
10. Sandberg, U., and J. A. Ejsmont. Tyre/Road Noise Reference Book. *Noise Control Engineering Journal*, Vol. 51, 2003, pp. 348-348.
11. Sayers, M. W. On the Calculation of International Roughness Index from Longitudinal Road Profile. *Transportation Research Record*, No. 1501, 1995, pp. 1-12.
12. U.S. Department of Transportation, Federal Highway Administration and Federal Transit Administration. *2010 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance*.

- www.fhwa.dot.gov/policy/2010cpr/pdfs/cp2010.pdf. (Accessed Oct. 28, 2013.) Report published by the U.S. Department of Transportation, Washington, D.C., 2010.
13. Transportation Research Board. *Highway Capacity Manual*. Book published by Transportation Research Board, Washington, D.C., 2000.
 14. Transportation Research Board. *Highway Capacity Manual*. Book published by Transportation Research Board, Washington, D.C., 2010.
 15. Bennett, C. R., and I. D. Greenwood. *Modelling Road User and Environmental Effects in HDM-4*. The Highway Development and Management Series Collection. Volume 7. Report published by The World Road Association, France, 2002.
 16. Karan, M. A., R. Kher, and R. Haas. Effects of Pavement Roughness on Vehicle Speeds. *Transportation Research Record*, No. 602, 1976, pp. 122-127.
 17. du Plessis, H. W. *A Pilot Study to Determine the Effect of Road Surface Roughness on Vehicle Speeds*. Report published by Division of Roads and Transport Technology, Council for Scientific and Industrial Research, Pretoria, South Africa, 1990.
 18. Elkins, G. E., and J. Semrau. Development of Limiting Velocity Models for the Highway Performance Monitoring System. *Transportation Research Record*, No. 1195, 1988, pp. 138-141.
 19. Cox, J. B. Effect of Road Surface Condition on Vehicle Operating Costs in Australia: Literature Review and Fleet Database Analysis. Report to AUSTROADS. Report published by Symonds Travers Morgan Pty Ltd., Melbourne, Australia, 1991.
 20. Cooper, D. R. C., P. G. Jordan, and J. C. Young. *The Effect on Traffic Speeds of Resurfacing a Road*. Report published by Transport and Road Research Laboratory, Construction and Maintenance Division, Crowthorne, Berkshire, UK, 1980.
 21. Watanatada, T., A. M. Dhareshwar, and P. R. S. R. Lima. *Vehicle Speeds and Operating Costs: Models for Road Planning and Management*. Book published by Johns Hopkins University Press, Baltimore, MD, 1987.
 22. McLean, J. R. *Adapting the HDM-III Vehicle Speed Prediction Models for Australian Rural Highways*. Working Document TE 91/014. Report published by Australian Road Research Board, Nunawading, Australia, 1991.
 23. Chandra, S. Effect of Road Roughness on Capacity of Two-Lane Roads. *Journal of Transportation Engineering-ASCE*, Vol. 130, No. 3, 2004, pp. 360-364.
 24. Caltrans. Caltrans Performance Measurement System (PeMS). pems.dot.ca.gov. (Accessed July 18, 2011.)
 25. National Oceanic and Atmospheric Administration. National Climate Data Center. www.ncdc.noaa.gov/cdo-web. (Accessed Aug. 12, 2012.)
 26. Caltrans. Caltrans Photolog. video.dot.ca.gov/photolog. (Accessed Feb. 13, 2013.)

27. Caltrans. California Road System (CRS) Maps. www.dot.ca.gov/hq/tsip/hseb/crs_maps. (Accessed Apr. 13, 2012.)
28. California Energy Commission. Energy Almanac. energyalmanac.ca.gov/gasoline. (Accessed Aug. 12, 2012.)
29. Caltrans. California Highways with 70 MPH Speed Limits. www.dot.ca.gov/hq/roadinfo/70mph.htm. (Accessed Aug. 15, 2012.)
30. Özkan, T., T. Lajunen, J. E. Chliaoutakis, D. Parker, and H. Summala. Cross-Cultural Differences in Driving Behaviours: A Comparison of Six Countries. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 9, No. 3, 2006, pp. 227-242.
31. Austin, D. *Effects of Gasoline Prices on Driving Behavior and Vehicle Markets*. Pub. No. 2883. Congressional Budget Office, Washington, D.C., 2008.
32. Burger, N. E. and D. T. Kaffine. Gas Prices, Traffic, and Freeway Speeds in Los Angeles. *Review of Economics and Statistics*, Vol. 91, No. 3, 2009, pp. 652-657.
33. West, B. H., R. N. McGill and S. Sluder. *Development and Validation of Light-Duty Vehicle Modal Emissions and Fuel Consumption Values for Traffic Models*. FHWA-RD-99-068. Federal Highway Administration (FHWA), Washington, D.C., 1999.
34. Chatti, K. and I. Zaabar. *Estimating the Effects of Pavement Condition on Vehicle Operating Costs*. NCHRP Report 720. Board, T. R., Washington, D.C., 2012.
35. Caltrans. *Maintenance and Rehabilitation Priority Assignment Based on Condition Survey*. California Department of Transportation, Sacramento, CA, 1997.