CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Estimation of Truck Traffic Volume from Single Loop Detectors Using Lane-to-Lane Speed Correlation

Jaimyoung Kwon, Pravin Varaiya,
Alexander Skabardonis
University of California, Berkeley

California PATH Working Paper
UCB-ITS-PWP-2003-11

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department Transportation, Federal Highway Administration.

Report for Task Order 4301

July 2003

ISSN 1055-1417

CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS

Estimation of Truck Traffic Volume from Single Loop Detectors Using Lane-to-Lane Speed Correlation

Jaimyoung Kwon, Pravin Varaiya, Alexander Skabardonis

ACKNOWLEDGEMENTS

This study is part of the PeMS project, which is supported by grants from Caltrans to the California PATH Program. We are very grateful to Caltrans Traffic Operations engineers for their support. The authors thank Ben Coifman, Ohio State University, for the BHL data and helpful discussions, and Joe Avis of California Department of Transportation for the AVC data from Los Angeles County and other helpful material.

Estimation of Truck Traffic Volume from Single Loop Detectors Using Using Lane-to-Lane Speed Correlation

Jaimyoung Kwon, Pravin Varaiya, Alexander Skabardonis

April 2003

ABSTRACT

An algorithm for real time estimation of truck traffic in multi-lane freeway is proposed. The algorithm uses data from single loop detectors—the most widely installed surveillance technology for urban freeways in the US. The algorithm works for those freeway locations that have a truck-free lane, and exhibit high lane-to-lane speed correlation. These conditions are met by most urban freeway locations. The algorithm produces real time estimates of the truck traffic volumes at the location. It can also be used to produce alternative estimate of the mean effective vehicle length, which can improve speed estimates from single loop detector data. The algorithm is tested with real freeway data and produces estimates of truck traffic volumes with only 5.7% error. It also captures the daily patterns of truck traffic and mean effective vehicle length. Applied to loop data on I-710 near Long Beach during the dockworkers lockout October 1-9, 2002, the algorithm finds a 32 % reduction in 5-axle truck volume.

Keywords: freeways; automatic vehicle classification; vehicle detectors; trucking; traffic surveillance

EXECUTIVE SUMMARY

Objectives and Methodology

Accurate knowledge of the volume and pattern of truck traffic is critical for various highway-related planning, design, and policy analyses. Typically, such data are collected either by Automatic Vehicle Classifiers (AVCs) based on Weigh-In-Motion (WIM) technologies or by manual counting. AVCs produce online counts of traffic of different vehicle types, but they are costly to install and suffer various limitations. When manual counting is used, truck traffic is recorded for a short sampling period, and the counts are extrapolated to get an estimate for a whole year. This estimate can have a large margin of error, even after adjusting for seasonal and day-of-week trends. Improving accuracy by increasing the frequency and length of manual counts is costly. Double loop detectors may also serve as a crude automatic vehicle classifier by using the vehicle lengths as a surrogate for vehicle type. However, the deployment of double loop detectors is limited and additional hardware and software needs to be installed to extract vehicle length information from double loops.

We propose a novel algorithm for the real-time estimation of truck traffic volume on freeways from single loop detector data. In contrast to other technologies, single loop detectors are much more widely deployed. The algorithm makes use of the essential relationship between speed, flow, occupancy, and effective vehicle length, according to which vehicle length can be estimated if average speed is known. It also relies on the phenomenon that the average speed of a truck-rich lane is actually quite close to that of truck-free inner lane(s), that is, there is a high lane-to-lane speed correlation. Using these two characteristics of traffic flow, we estimate the proportion of trucks within a sample, given representative lengths of trucks and passenger cars. The algorithm is extremely easy to implement, requiring tuning of only a few generic parameters.

Findings

The algorithm is tested with three real-world freeway datasets. The first dataset is from double loop detector data from Berkeley Highway Laboratory (BHL) on I-80 in Berkeley, California, where speed measurements and the effective lengths of individual vehicles are available. The proposed algorithm produces estimates of truck traffic volumes with only 5.7% error. Next, the algorithm estimates were compared against WIM hourly vehicle classification data from a site on I-91 in Los Angeles, California. The algorithm was applied to the 5-minute single loop flow and occupancy data collected from the freeway performance measurement (PeMS) database at the loop station closest to the WIM station. The algorithm captured well the daily patterns of truck traffic and mean effective vehicle length. Finally, we used PeMS data from a location on I-710 near Long Beach to estimate the impact on truck volumes during the dockworkers lockout at the port of Long Beach during the period of October 1-9, 2002. The algorithm found a 32 % reduction in 5-axle truck volume during the lockout period.

The results from the application of the algorithm demonstrate that the algorithm captures the qualitative and quantitative characteristics of daily truck traffic satisfactorily with acceptable bias. The implementation of the algorithm on the widely available single loop detector infrastructure will produce online estimates of total truck volume as well as seasonal and daily truck patterns. This information can facilitate better freeway design and management.

TABLE OF CONTENTS

Acknowledgements ii
Abstract iii
Executive Summary iv
Table of Contents v
List of figures vi
List of Tables vi
1. INTRODUCTION 1
  Data 2
2. PROPOSED METHOD 2
  Distribution of Effective Vehicle Length 3
  Two Key Assumptions and the Algorithm 4
  Implications for G-factor estimation 6
  External Source of Speed 6
3. ANALYSIS 6
  Results 7
  G-factor Estimation 7
  Comparison with WIM Data 7
5. Discussion 8
  Setting Parameters for Implementation 9
  Presence of Long Vehicles in Truck-Free Lanes 9
  Bias at the Onset of Low-Speed 9
  G-factor Estimation 10
6. CONCLUSION 10
References 11

LIST OF FIGURES

FIGURE 1. Distribution of Effective Vehicle Length—BHL Data 14
FIGURE 2 Time-series plots of 5-minute aggregated double-loop speeds measured in four lanes at a location in BHL study. Each plot corresponds to a single day. Here and below when time scale is 5 minute, time is in 5-minute increments, from 0 to 288, starting with 0 at midnight.. 15
FIGURE 3 Lane-to-Lane Speed Correlation. Each scatterplot depicts 5-minute aggregated speeds (mph; 1 mph = 1.61 km/h) measured in lane *a* and *b* at a location in BHL study over 10 days (total of 2,880 samples). In each plot, the thick line is the least squares regression line (without the intercept) fitting lane *b* speed on lane *a* speed and the thin line is the reference line *y=x*. Three numbers are the slope of the regression line, R-squared values, and correlation coefficient, from top to bottom. 16
FIGURE 4 Observed and estimated daily pattern of lane-by-lane truck volume, aggregated over 20 days. Here and below when time scale is one hour, time is in one-hour increments, from 1 to 24, starting with 1 at midnight. 17
FIGURE 5 Observed and estimated daily pattern of lane-total truck count (top); lane-by-lane daily pattern of speed, averaged over 10 days (middle); daily pattern of estimation error of lane-total truck count (bottom). 18
FIGURE 6 True and estimated daily trend of 5-minute mean effective vehicle length estimate, averaged over 10 days. 19
FIGURE 7 Total vehicle volume estimated from WIM data (top left) and single loop data (top right); hourly volume of long truck traffic estimated from WIM data (bottom left) and single loop data using the proposed algorithm (bottom right)... 20
FIGURE 8 Comparison of daily total traffic volume of all cars (left), trucks (middle) and passenger cars (right) during port lockout and non-lockout periods. 21
FIGURE 9 Day-to-day trend of daily total truck traffic volume, by lane (top) and lane-total (bottom). The lockout period is days 44-55.. 22

LIST OF TABLES

TABLE 1 Performance of the Algorithm for Data from BHL Study 12
TABLE 2 Comparison of Daily Traffic Volume during Lockout and Non-Lockout Weekdays—I-710 (Port of Long Beach) 13

1. INTRODUCTION

Accurate knowledge of freight or heavy truck traffic is critical for various highway-related planning, design, and policy analyses. It is necessary to have estimates of such quantities as truck Annual Average Daily Traffic (AADT) for different sections of a freeway (*1,2*). Typically, such data are collected either by an Automatic Vehicle Classifier (AVC) or by manual counting. An AVC is usually based on technologies such as Weigh-In-Motion (WIM) consisting of inductive loops and a bending plate or Piezo sensors, or more sophisticated technologies such as video imaging, laser and night vision systems, or acoustic signal analysis. These systems produce online counts of traffic of different vehicle types. But they are costly to install and suffer various limitations (*3*). When manual counting is used, a person records the truck traffic distribution for a short sampling period, usually a day or two, and the counts are extrapolated to get an estimate for a whole year. This estimate can have a large margin of error, even after adjusting for seasonal and day-of-week trends (*1,4,5,6*). Improving accuracy by increasing the frequency and length of manual counts is costly.

Double loop detectors can serve as a crude automatic vehicle classifier by using the vehicle lengths as a surrogate for vehicle type. It is reasonable in many cases to assume that vehicles longer than say 50 ft (15.24 m) are heavy trucks. But deployment of double loop detectors is limited. Moreover, separate software or hardware needs to be installed to extract vehicle length information from double loops.

In this report, we propose an algorithm for the real-time estimation of truck traffic volume from single loop detectors. In contrast to other technologies, single loop detectors are much more widely deployed. In principle, they are unable to record vehicle lengths as double loop detectors can, and only report flow (traffic volume) and occupancy. The algorithm makes use of the essential relationship between speed, flow, occupancy, and effective vehicle length, according to which vehicle length can be estimated if average speed is known. It also relies on the phenomenon that the average speed of a truck-rich lane is actually quite close to that of truck-free inner lane(s), that is, there is a high lane-to-lane correlation of speed. Using these two characteristics of traffic flow, we estimate the proportion of trucks within a sample, given representative lengths of trucks and passenger vehicles. This idea is elaborated in section 2.

Section 2 also discusses the relevance of the algorithm for estimating the mean effective vehicle length (MEVL), whose inverse is called ‘G-factor’. The G-factor is a crucial parameter in speed estimation from single loop flow and occupancy data. It is shown that the algorithm produces, as a byproduct, an online estimate of the G-factor or the daily profile of the G-factor. The G-factor estimates can improve speed estimates based on single-loop data.

The algorithm is tested with two data sets and the results are presented in section 3. In section 4, the algorithm is applied to data on I-710 near Long Beach, CA, during the dockworkers lockout in October of 2002. Section 5 discusses the study methodology. Section 6 summarizes the findings.

Data

For the analysis as well as for an explanation of concepts, we use data from the Berkeley Highway Laboratory (BHL). (*7*). We study eastbound traffic at detector station 6, near the Ashby Avenue exit. The data are collected for ten Mondays between March 22, 1999 and May 24, 1999, at eight double loop detector stations located on I-80 in Berkeley, California. The double loop speed measurements and the effective vehicle lengths of individual vehicles are available. Thus these measurements serve as 'ground truth'. The algorithm of course only makes use of data from one of the double loops.

Another test data set consists of WIM hourly vehicle classification data from eastbound I-91 at postmile 7.5 (=12.1 km), east of Avalon Boulevard in Los Angeles, for the period of May 5-18, 2002 (excluding May 12, 13 and 14)—a total of 11 days. The 5-minute single loop measurements of flow and occupancy from detector station (VDS 718130) closest to the WIM station, for the same time period were extracted from the freeway performance measurement (PeMS) database (*8*). We use the PeMS loop detector data to estimate the truck volumes and compare the estimates with the WIM data.

Finally, we use PeMS data from a location on I-710 freeway near the port of Long Beach, from August 16, 2002 to October 23, 2002, to estimate the impact on truck volumes during the dockworkers lockout during the period of October 1-9, 2002.

2. PROPOSED METHOD

Conventional single loop detectors measure flow, the number of vehicles that pass the detector during a fixed sample period, and occupancy, the percentage of the given sample period that the detector is "occupied" by vehicles. For each lane, the flow q and occupancy O are defined as

$$q(i,j) = \frac{n(i,j)}{T}, \tag{1}$$

$$O(i,j) = \frac{\sum_{k \in K(i,j)} t_k}{T}. \tag{2}$$

Here i indexes lane and j indexes the sample time period, and

$n(i,j)$ = number of vehicles that pass over the detector in lane i during time period j,
T = sampling period,
$K(i,j)$ = set of all vehicles that pass over the detector in lane i during time period j, and
t_k = vehicle k's on-time, i.e. the time interval during which this vehicle occupies the detector.

The 'on-time', speed, and length of a particular vehicle are related by

$$L_k = v_k t_k, \tag{3}$$

where:

L_k = effective length of vehicle k as "seen" by the detector

v_k = speed of vehicle k

The sample mean speed at (i, j) is defined to be

$$\bar{v}(i, j) = \frac{\sum_{k \in K(i,j)} v_k}{n(i, j)}. \tag{4}$$

From equations (1) and (2),

$$O(i, j) = \frac{1}{T} \sum_{k \in K(i,j)} \frac{L_k}{v_k} = q(i, j) \frac{1}{n(i, j)} \sum_{k \in K(i,j)} \frac{L_k}{v_k}. \tag{5}$$

Assuming that individual vehicle speeds are nearly constant during each sample time period, this yields

$$O(i, j) \approx \frac{q(i, j)\bar{L}(i, j)}{\bar{v}(i, j)}, \tag{6}$$

or,

$$\bar{v}(i, j) \approx \frac{q(i, j)\bar{L}(i, j)}{O(i, j)}, \tag{7}$$

in which $\bar{L}(i, j) = n^{-1}(i, j) \sum_{k \in K(i,j)} L_k$ is the mean effective vehicle length (MEVL).

Equation (7) has been used to estimate speeds from single loop detector flow and occupancy measurements. However, this requires knowledge of MEVL or G-factor. On the other hand, if one knows the speed, one can estimate MEVL, by re-writing (7) as

$$\bar{L}(i, j) \approx \bar{v}(i, j) \frac{O(i, j)}{q(i, j)}. \tag{8}$$

Observe that this value is large if there are many long vehicles (LVs) and small if the traffic consists mostly of passenger cars (PCs).

Distribution of Effective Vehicle Length

Figure 1 shows the distribution of the effective vehicle lengths (EVLs) observed for 25,700 vehicles detected during the 24-hour period of March 29, 1999 from BHL data. Only data from the loop detector installed in the fourth (or outermost) lane, which has significant truck traffic, is used for the plot. The histogram of the effective vehicle length (top of Figure 1) is decomposed into two separate histograms corresponding to vehicles longer than 40 ft (12.19 m) (middle) and shorter than 40 ft (bottom).

The highest peak is at 17 ft (5.18 m); there is another peak at 61 ft (18.6 m). It is likely that the former corresponds to the typical length of PC and the latter to that of LV. For our study, we will use nominal representative lengths 18.6 ft (5.67 m) and 61.2 ft (18.65 m) as the typical length of the two vehicle classes. These are the group means of those vehicles whose EVLs are smaller and larger than the threshold 40 ft (12.19 m), respectively. A bimodal distribution like in Figure 1 occurs when the vehicles mostly belong to two classes of vehicles with appreciably different lengths, which seems to be the case in these data. Several researchers (*9*) have noticed this bimodal characteristic of EVL distribution in truck-rich traffic.

In an ideal setting when there are only two types of vehicles, LV and PC, each having fixed lengths $\bar{l}_t$ and $\bar{l}_c$, we have the relationship:

$$\bar{L}(i,j) = p(i,j)\bar{l}_t + (1 - p(i,j))\bar{l}_c , \tag{9}$$

in which $p(i,j)$ is the proportion of truck traffic in the sample (i,j). One can rewrite (9) as

$$p(i,j) = \frac{\bar{L}(i,j) - \bar{l}_c}{\bar{l}_t - \bar{l}_c} . \tag{10}$$

Truck and passenger vehicle counts can then be calculated from

$$n_t(i,j) = p(i,j)n(i,j) \text{ and } n_c(i,j) = (1 - p(i,j))n(i,j) = n(i,j) - n_t(i,j) . \tag{11}$$

Two Key Assumptions and the Algorithm

As we have just seen, knowledge of the mean speed allows one to calculate the proportion of trucks and their volume during a time period. In this regard, several researchers [*10*] have observed that "on multi-lane freeways, vehicle speed over different lanes tend to be synchronized" or,

$$\bar{v}(i,j) \approx \bar{v}(i',j) . \tag{12}$$

We call this phenomenon a "strong lane-to-lane correlation of speed" in multi-lane freeways. Also, for most multi-lane freeways in the US, heavy trucks are not allowed or, discouraged from, driving in inner lanes. This leads to the hypothesis

$$p(i,j) \approx 0 \text{ or } \bar{L}(i,j) \approx \bar{l}_c , \tag{13}$$

in which "*i*" corresponds to inner or faster lanes, say the first or second lane from the median.

Suppose lane i is almost truck-free and i' is truck-rich, and the two lanes have high lane-to-lane speed correlation. Then

$$
\begin{aligned}
&\frac{q(i,j)\bar{l}_c}{O(i,j)} \approx \frac{q(i,j)\bar{L}(i,j)}{O(i,j)} \approx \bar{v}(i,j) \\
&\approx \bar{v}(i',j) \approx \frac{q(i',j)\bar{L}(i',j)}{q(i',j)} \approx \frac{q(i',j)(p(i',j)\bar{l}_t + (1-p(i',j))\bar{l}_c)}{O(i',j)},
\end{aligned}
\tag{14}
$$

using the speed equation, correlation, and truck-free lanes (equations (7), (12) and (13) each).

One can solve (14) for $p(i',j)$ directly or solve it in two steps as follows: first estimate the mean vehicle length by

$$
\hat{L}(i',j) = \frac{q(i,j)/O(i,j)}{q(i',j)/O(i',j)}\bar{l}_c, \tag{15}
$$

and then the truck proportion by

$$
\hat{p}(i',j) = \frac{\hat{L}(i',j) - \bar{l}_c}{\bar{l}_t - \bar{l}_c}. \tag{16}
$$

If the proportion estimate has a value outside [0,1] due to detector noise, the estimate must be appropriately truncated. The truck count for sample *j* is then estimated by $\hat{n}_t(i,j) = \hat{p}(i,j)n(i,j)$.

The high lane-to-lane correlation of speed is clearly observable in the data shown in Figure 2 and 3, which show the joint behavior of individual lane speeds measured from BHL study. R-squared values and correlation coefficients are larger than 0.99 and 0.9 for all pairs.

Though the relationship $v' \approx v$ holds for all lane pairs, the speed is typically slightly lower in the outer lanes. Flow and occupancy are also moderately correlated over different lanes but their relationship tends to be noisier and more nonlinear.

As is illustrated by Figure 3, the outer lanes exhibit slower speeds. The slopes of the least-squares regression lines are 0.95, 0.91 and 0.89 for regressing lanes 3, 4 and 5 on the fastest lane 2. (Lane 1 is the HOV lane, excluded from the analysis.) Thus, about 5% decrease of speed is observed as one drives on a lane further from the median. This suggests the following modification of (12):

$$
\bar{v}(i',j) \approx b(i',i)\bar{v}(i,j), \tag{17}
$$

in which $b(i',i)$ is the estimated proportion of speeds in lane i' and *i*. Using a simpler identity like (12) when in fact $b(i',i) \neq 1$ leads to bias. But despite the bias, the resulting estimate will capture the seasonal or daily truck traffic pattern.

In practice, a generic constant may be used for b, like 5% slower traffic in outer lane, although estimates 'tuned' for an individual detector station (using another source of data) may be beneficial.

Implications for G-factor estimation

The G-factor is a critical parameter for estimation of the mean speed using relation (7). It is well known that a constant G-factor for speed estimation from single loop data is a bad practice. Even though there have been some efforts to do so (*9,11,12*), the real-time estimation of G-factor is a difficult problem except during free flow periods. Thus the historical profile of G-factor is of great value for better speed estimation.

Our algorithm can be viewed as an alternative procedure to estimate the G-factor. It produces (15) as a byproduct, an estimate of the mean effective vehicle length. For those freeways where our algorithm can be applied, it can thus produce G-factor estimate either online or to obtain the historical pattern.

External Source of Speed

A speed estimate from a separate source may be available in some cases. In such cases, instead of using the lane-to-lane speed correlation, one can directly estimate the EVL by

$$\tilde{L}(i',j) = \frac{v(i',j)O(i',j)}{q(i',j)}, \tag{18}$$

instead of (15), and then proceed as above to estimate the proportion and volume of trucks. Video images could provide such data. It is relatively easy to extract the average speed of a group of vehicles from video data (*13*).

We will distinguish between the original algorithm that doesn't require speed and the one that uses the speed from an extra source by the names *single loop algorithm* and *exogenous speed algorithm*.

3. ANALYSIS

We first examine detector data from the BHL study collected over ten Mondays between March 22, 1999 and May 24, 1999. The sampling rate of the loop detector is 60 Hz but we only use 5-minute aggregated data to run the algorithm. Because we have the effective vehicle length, derived from double loop speed traps, we know the actual truck volumes and we use these to assess the performance of the algorithm.

The algorithm is applied using the (default) parameters $\bar{l}_c$ =18.6 ft (5.67 m) and $\bar{l}_t$ =61.2 ft (18.65 m) as the mean effective vehicle lengths for the two classes and the regression coefficients 0.95, 0.91, and 0.85 for the speed correction factors in (17). Lane 2 is assumed to be truck-free, which seems valid since only 1% of the vehicles in that lane were longer than 40 ft (12.19 m). (Lane 1 is the HOV lane.) We also calculate the alternative estimate of truck volume using the exogenous speed algorithm, using the double loop speed estimate as an extra source.

Next, we use PeMS loop data and the WIM data for the I-91 location over an 11-day period. The algorithm is applied with the parameters $\bar{l}_c$ =12 ft (3.66 m) and $\bar{l}_t$ =55 ft (16.8 m), which were determined from an algorithm outlined in (*14*). A set of generic coefficients 0.95, 0.90, and 0.85 are used for the speed correction factors. Lane 1 is assumed to be truck-free.

Results

The performance of the algorithms on the BHL data set is summarized in Table 1, which shows lane-by-lane and all-lane truck AADT, the estimates from both single loop and exogenous speed algorithms, and the error rate. All quantities are averages over 10 days. The total percentage error in estimating lane-total truck AADT is -5.7% and -3.3%, using the single loop and exogenous speed algorithms, respectively. Recall that speed is not needed for the single loop algorithm.

The lane-by-lane breakdown of the statistics shows that the single loop algorithm inherently estimates the truck traffic in lane 2, the reference lane, as zero. The exogenous speed algorithm doesn't suffer from this inherent limitation, although the percentage error for the same lane is large, –53%. Both algorithms underestimate truck counts for lane 3 and 5 and overestimate it for lane 4.

The observed and estimated daily patterns of truck traffic for lane 3, 4, and 5 are shown in Figure 4. Qualitatively, both algorithms capture well the familiar daily pattern of truck traffic, although the exogenous speed algorithm matches the daily trend better. The single loop algorithm is biased in the afternoon commute hours. Figure 5 shows that the bias is related to the variability in speed. The single loop algorithm bias is most serious at the start of the congestion period.

We also compare our results with the truck AADT reported in (*2*). We chose the location Route 80, postmile 4.582 (=7.37 km)at Berkeley intersecting Route 13 (Page 134 of the report), which is closest to our location. One half of the two-way total AADT is 127,500, which is larger than the current AADT by 23%, and the source of the error is not clear. But the percentage of trucks (with more than 5 axles) AADT in the total AADT is 4.8% according to the report, agreeing closely with the truck volume percentage of the BHL data.

G-factor Estimation

Figure 6 shows the results when the algorithm is used to estimate the G-factor. The true G-factor is calculated using (8), assuming equality in that relationship. As expected, the true G-factor is close to constant for Lane 2 and varies significantly for lanes 3, 4 and 5. The estimated G-factor follows the true G-factor surprisingly well.

Comparison with WIM Data

The performance of the algorithm on the second data set is shown in Figure 7. WIM data and single loop data produce similar total traffic volume profile over days. The three days that do not show morning and afternoon peaks are weekends. The hourly truck volume estimated from the single loop data using the current algorithm also corresponds to that from WIM data. Even though there are some quantitative differences (our algorithm is overestimating truck traffic by about 20%), the two daily truck volume profiles match each other qualitatively. In particular, the trend of low truck volume over weekends is clearly visible from both WIM data and the estimates by the algorithm.

4. APPLICATION: THE DOCWORKERS LOCKOUT, OCTOBER 2002

At midnight on Saturday, September 28, 2002, West Coast port operators shut down cargo terminals from Seattle to Los Angeles, in effect locking out unionized longshore workers. On October 9, President Bush invoked the Taft-Hartley Act, ordering the ports to reopen and the workers to return to their jobs.

We analyzed single loop data at a location (VDS 761734) that is likely to have been affected by the lockout. This is a five-lane section near the port of Long Beach at postmile 7.3 (=11.7 km)on I-710. Five-minute averages of flow and occupancy data between August 16 and October 23, 2002, were retrieved from the PeMS database. We use our single loop algorithm to estimate and compare the 5-axle truck volumes during the non-lockout and lockout periods.

The MEVL in the different lanes is estimated using (15). The ratios $b(i',i)$ of speeds are calculated from the ratios of representative free flow speeds in the different lanes, obtained from the Bay Area data. The median values of free flow speeds ranged from 76 mph for lane 1 (median) to 59 mph for lane 5 (outer or shoulder lane). Lane 1 is assumed to be free of trucks. The unknown vehicle length $L(1,j)$ (of passenger cars) in lane 1 is estimated using the following approach. We assume (1) lane 1 has a fixed MEVL $L_1 \equiv L(1,j)$ corresponding to the representative effective vehicle length of passenger cars; and (2) the median of lane 1 speed is same as the representative median free flow speed from the Bay Area. Then, we have

$$\text{median}_{\text{D4}}(v(1.j)) = \text{median}_{\text{D4}}(v(i,j)) = \text{median}_{\text{D4}}\left(\frac{q(i,j)}{O(i,j)}\right)L_1 . \qquad (19)$$

We solve this equation to get an estimate for L_1. It is found to be 16.6 ft (5.06 m) from the data. We use this as the representative passenger car length for all lanes. We use 60 ft (18.3 m) for the representative truck length.

Table 2 compares the truck volumes over non-lockout and lockout weekdays. The estimates suggest a 32% reduction in daily truck volume, and only a 0.8% reduction in passenger car volume, and an overall reduction in vehicle count of 3.4%. A statistically richer comparison is offered in the box plots of Figure 8. Each plot gives five numbers: the lowest and highest lines are drawn at the highest and lowest daily estimates; the three lines that form the box are drawn at 25%, 50%, and 75% of the daily estimates. Figure 9 shows the daily truck volume over the study period. The figure shows both the weekly cycle, the drop during the lockout period, and the partial recovery following the lockout.

5. DISCUSSION

We discuss some issues concerning implementation of the algorithms, and areas for future research.

Setting Parameters for Implementation

For practical implementation of the algorithm, a few key parameters need to be set. They are (1) the representative vehicle lengths for the two vehicle classes, and (2) ratio of the speed between lanes.

The parameter (1) would seem to be independent of location. In general, different locations would have the same representative lengths for long trucks as well as passenger cars, unless very long trucks are used extensively in a certain region. A more subtle issue is the presence of relatively short (less than 5 axles) trucks, whose volume typically accounts for fifty to ninety percent of total truck volume (*2*). In many applications these short trucks are less significant than long trucks, and our algorithm aims to capture only long vehicle volumes. However, a significant volume of short trucks may lead to bias in the estimate of the long truck volume. How serious this is a problem seems to depend on the specific application, but further analysis of this issue is warranted.

Another concern related to parameter (1) is the vehicle detector bias. Even though true vehicle speeds stay constant over different locations, individual detectors tend to 'see' them differently, often by as much as a few feet. Various approaches are possible to estimate the detector bias; one approach that uses only 30-second single loop detector data is proposed in (*14*) and is employed in the current study.

The parameter (2) is more likely to vary with location than parameter (1). Would using a common set of speed ratio factors for different freeway locations reasonable? Even the linearity assumption may not be reasonable for certain locations. We may propose rules-of-thumb like '5% speed decrease for outer lane,' and 'use location that are not close to on- and off-ramps.' But the issue is a challenging and interesting topic for transportation researchers and engineers.

Presence of Long Vehicles in Truck-Free Lanes

Presence of long vehicles in supposedly truck-free lanes can affect the accuracy of the proposed algorithm. It will increase the occupancy in that lane and thus lead to a smaller estimate of mean effective vehicle length in outer lanes and underestimation of the proportion of long vehicles in the outer lanes (see equation 15). The degree of underestimation will depend on the abundance of long vehicles in the reference lane, which in turn should depend on the road geometry like number of lanes, among others. It would be useful to study the proportion of long vehicles in truck-free lanes for various locations under different conditions. For the BHL data, only less than 1% of the vehicles in lane 2 were longer than 40 ft so the bias is negligible.

Bias at the Onset of Low-Speed

It was observed that the estimate of truck volume is biased and unstable at the start of the congestion period. During this period the occupancy and speed are unstable, and the 'signal-to-noise' ratio of the both parameters decreases. A quick solution to this may be putting less confidence for the truck volume estimate when the period shows sudden change of occupancy. But correcting this bias seems a challenging problem, which requires careful study of traffic dynamics.

G-factor Estimation

It is encouraging that the G-factor estimated by the current algorithm is surprisingly close to the real historical G-factor. The estimated G-factor can be an input for various subsequent algorithms for speed calculation, including (*12*). Using the current algorithm for online G-factor estimates is nothing more than using the speed of the reference lane multiplied by a constant factor as the estimate for the speed at another lane. This may not be a good idea in general, but it's an interesting question if using a noisy online estimate of G-factor gives better speed estimate.

6. CONCLUSIONS

We proposed an online algorithm for estimating truck traffic volume. The algorithm is applicable to multi-lane freeways with one truck-free lane and high lane-to-lane speed correlation. Both conditions seem to be satisfied at most major urban freeway locations not close to on- and off-ramps. This makes the algorithm widely applicable. The algorithm is extremely easy to implement, requiring tuning of only a few generic parameters. It can be also used for G-factor calculation.

In the empirical study at two urban freeway locations with moderate long truck traffic volume, the algorithm captures the qualitative characteristics of daily truck traffic satisfactorily with acceptable bias. Quantitatively, it exhibits a 5.7% error in total truck volume estimates for data from BHL study. The algorithm also captures very well the historical daily trend of G-factor, proving beneficial for improving speed estimation from single loop detector data.

The big advantage of the proposed algorithm is that it can collect truck volume profile for any given day and location, when loop data are available. Given the wide deployment of single loop detector data, it is feasible to apply the proposed algorithm to produce a census-type truck AADT for all many freeway locations. The implementation of the algorithm requires minimal effort. Since the processing can be done on the aggregated data at the TMC, no additional hardware or software is needed in the field at the detector station (or 'cabinet') level. Note also that our algorithm produces truck volumes at hourly or 5-minute intervals for different days of week, season, etc. Such information permits the study of the pattern of truck traffic over time of day, day of week, and seasons.

REFERENCES

1. U.S. Department of Transportation (USDOT). *Traffic Monitoring Guide, 3rd edition.* FHWA, Washington, DC. 1995.

2. California Department of Transportation, "Annual Average Daily Truck Traffic on the California State Highway System," Sacramento, CA, 2000.

3. Nihan, N. L., Zhang, X. and Wang, Y., "Evaluation of Dual-Loop Data Accuracy Using Video Ground Truth Data, Research Report TNW2002-02. University of Washington. March 2002.

4. Weinblatt, H. Using Seasonal and Day-of-Week Factoring To Improve Estimates of Truck Vehicle Miles Traveled. *Transportation Research Record 1522*, 1996, pp. 1–8.

5. Sharma, S.C., B.M. Gulati, and S. Rizak. 1996. Statewide Traffic Volume Studies and Precision of AADT Estimates. In *Journal of Transportation Engineering,* Vol 122 (6), 430–439.

6. Sharma, S. C., Liu, G. X., and Thomas, S. Sources of Error in Estimating Truck Traffic from Automatic Vehicle Classification Data. In *Journal of Transportation & Statistics*. Volume 1 (3), October 1998.

7. Coifman, B., Lyddy, D., and Skabardonis, A. The Berkeley Highway Laboratory-Building on the I-880 Field Experiment. In *Proc. IEEE ITS Council Annual Meeting*, IEEE, 2000, pp 5-10.

8. Chen, C., K. Petty, A. Skabardonis, and P. Varaiya. Freeway Performance Measurement System: Mining Loop Detector Data. *Transportation Research Record No.1748,* 2001, pp 96-102.

9. Wang, Y. and Nihan, N. L. Freeway Traffic Speed Estimation Using Single Loop Outputs, accepted for publication by Transportation Research Record, January 2000.

10. Helbing, D. Traffic and Related Self-Driven Many Particle Systems. In *Review of Modern Physics*, vol 73, October 2001, pp. 1067-1141.

11. Coifman, B. Improved Velocity Estimation Using Single Loop Detectors. In *Transportation Research: Part A*, Vol 35 (10), 2001, pp. 863-880.

12. Jia, Z., Chen, C., Coifman, B. and Varaiya, P. The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single loop detectors. In *Proc. IEEE ITS Council Annual Meeting,* IEEE, 2001, pp 536-541.

13. Beymer, D., McLauchlan, P., Coifman, B., and Malik, J. A Real-Time Computer Vision System for Measuring Traffic Parameters. In *Proc. IEEE - Computer Vision and Pattern Recognition*, IEEE, 1997, pp 495-501.

14. Kwon, J. Estimation of Free Flow Speed and Detector Bias from Single Loop Data Using Vehicle Length Profile. PATH working paper, August 2001.

TABLE 1 Performance of the Algorithm for Data from BHL Study

	Observed		Single-Loop Algorithm		Exogenous-Speed Algorithm	
	Total Traffic Volume (veh/day)	Truck Traffic Volume (veh/day) and Percent of Total Volume	Estimated Truck Volume (veh/day) and Percent of Total Volume	Error in Truck Volume Estimate (veh/day) and Percent Error	Estimated Truck Volume (veh/day) and Percent of Total Volume	Error in Truck Volume Estimate (veh/day) and Percent Error
Lane 1 (HOV)	11,493	- [1]	-	-	-	-
Lane 2	25,846	255 (1.0%)	0 (0%)	255 (-100%)	121 (0.5%)	-134 (-53%)
Lane 3	24,169	1,437 (5.9%)	1,217 (5.0%)	-220 (-15.3%)	1,344 (5.6%)	-94 (-6.5%)
Lane 4	23,546	2,037 (8.7%)	2,098 (8.9%)	62 (3.0%)	2,153 (9.1%)	116 (5.7%)
Lane 5	13,621	981 (7.2%)	884 (6.5%)	-97 (-9.9%)	936 (6.9%)	-45 (-4.5%)
Total	98,677	4,710 (4.8%)	4,200 (4.3%)	**-255 (-5.7%)**	4,553 (4.6%)	**-157 (-3.3%)**
Caltrans Report for Year 2000[2]	127,500	6,126 (4.8%)	-	-	-	-

1. A dash (-) means that data is not applicable.
2. Caltrans Report quantities are calculated assuming the same AADT for both directions.

TABLE 2 Comparison of Daily Traffic Volume during Lockout and Non-Lockout Weekdays—I-710 (Port of Long Beach)

	Average Daily Traffic Volume (veh/day)		
	All Cars	Trucks	Passenger Cars
Non-lockout Weekdays[1]	117,068	11,003	105,469
Lockout Weekdays[2]	113,107	7,410	106,355
Difference	-3,961 (-3.4%)	-3,594 (-32%)	-886 (-0.8%)

1. Median over 38 non-lockout weekdays
2. Median over 5 lockout weekdays

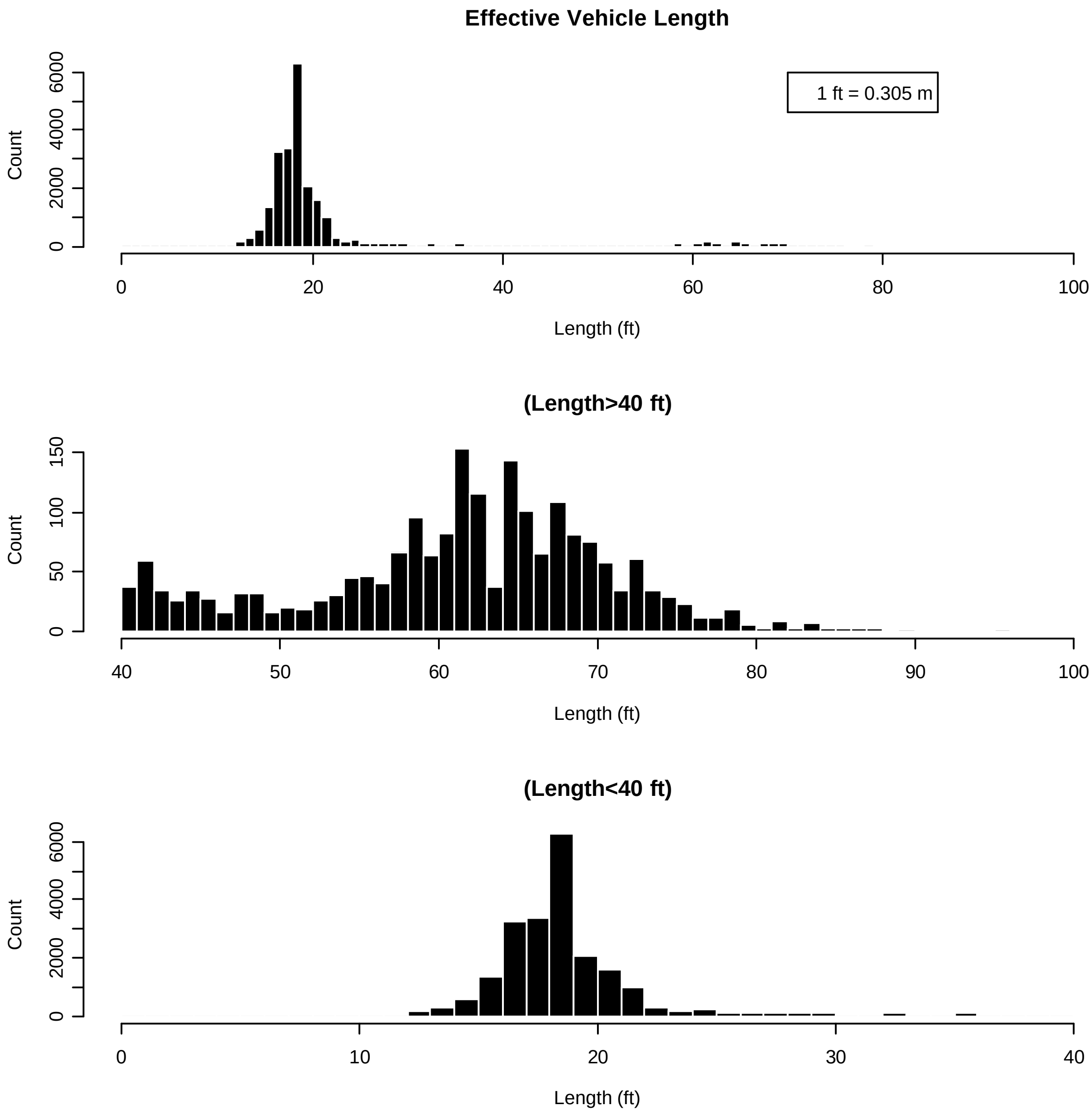

FIGURE 1. Distribution of Effective Vehicle Length—BHL Data

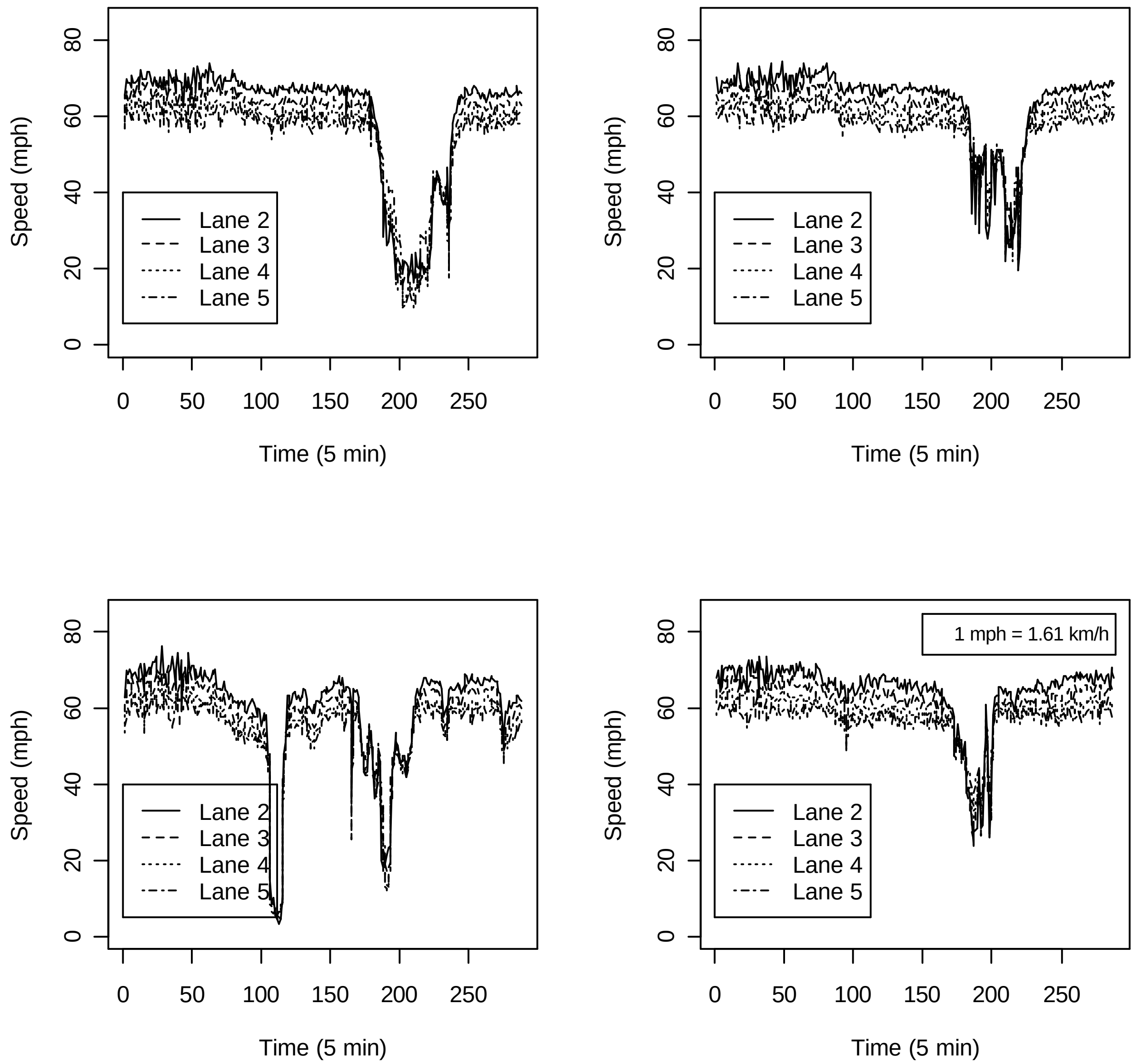

FIGURE 2 Time-series plots of 5-minute aggregated double-loop speeds measured in four lanes at a location in BHL study. Each plot corresponds to a single day. Here and below when time scale is 5 minute, time is in 5-minute increments, from 0 to 288, starting with 0 at midnight.

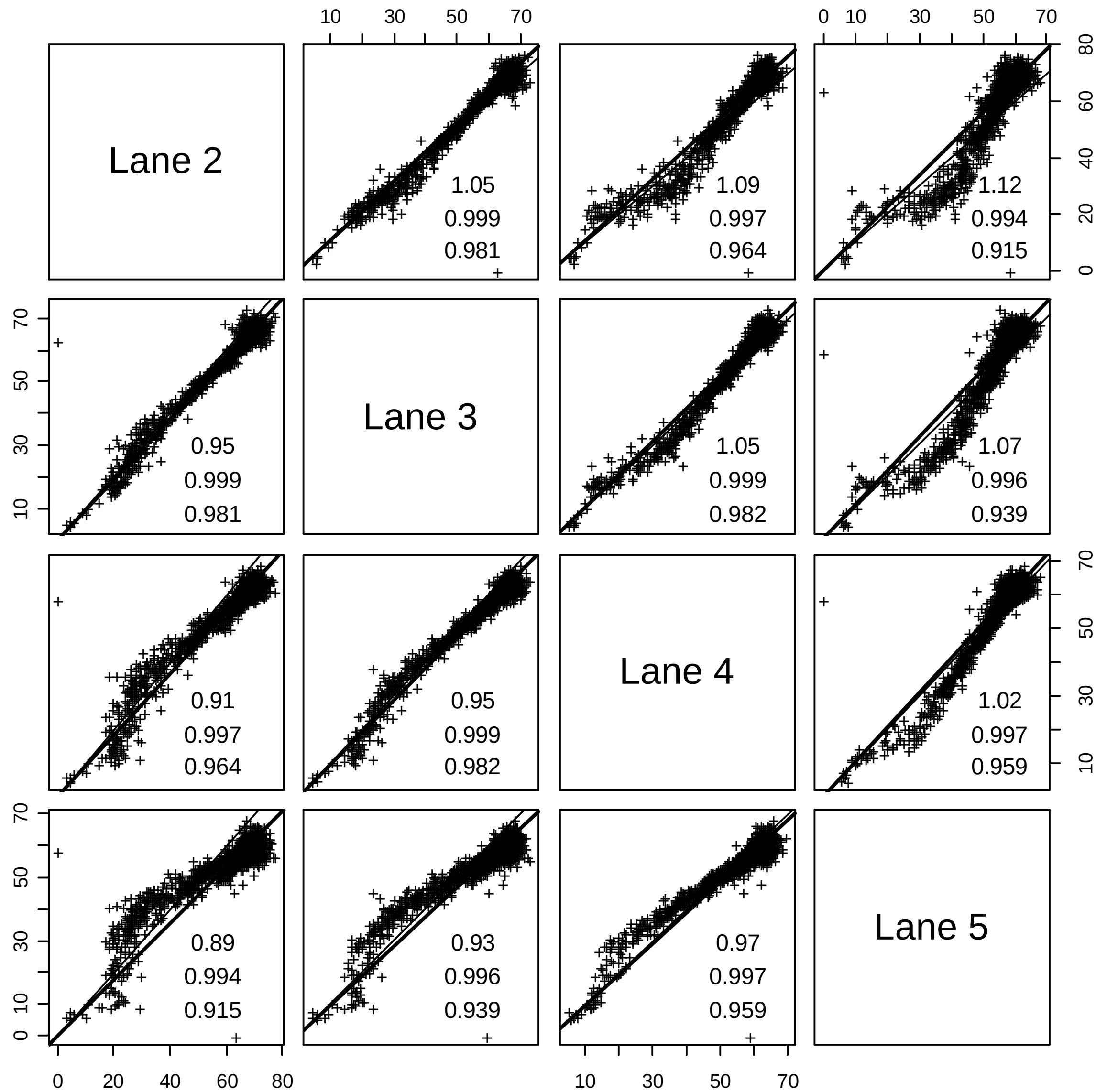

FIGURE 3 Lane-to-Lane Speed Correlation. Each scatterplot depicts 5-minute aggregated speeds (mph; 1 mph = 1.61 km/h) measured in lane *a* and *b* at a location in BHL study over 10 days (total of 2,880 samples). In each plot, the thick line is the least squares regression line (without the intercept) fitting lane *b* speed on lane *a* speed and the thin line is the reference line *y*=*x*. Three numbers are the slope of the regression line, R-squared values, and correlation coefficient, from top to bottom.

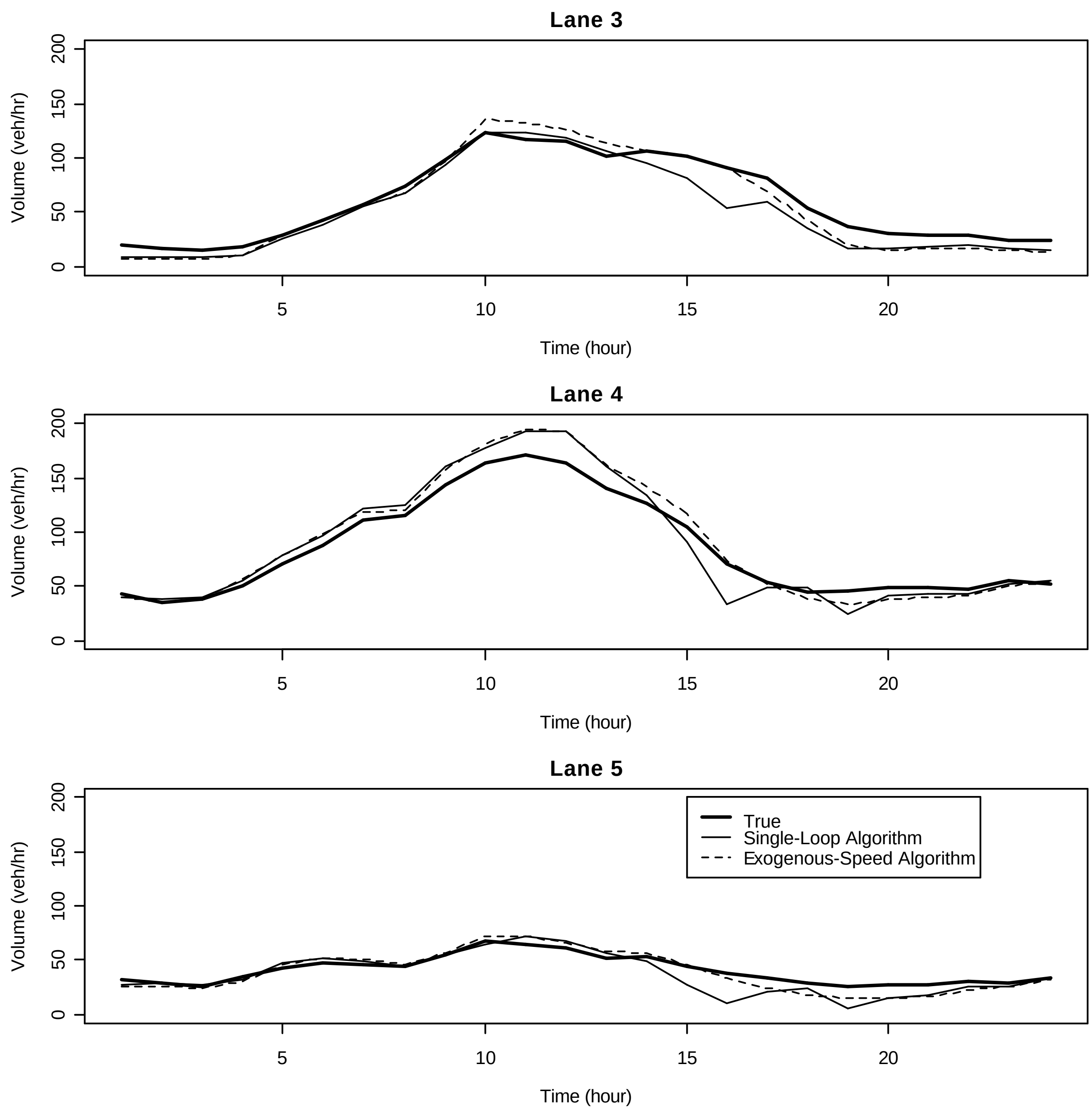

FIGURE 4 Observed and estimated daily pattern of lane-by-lane truck volume, aggregated over 20 days. Here and below when time scale is one hour, time is in one-hour increments, from 1 to 24, starting with 1 at midnight.

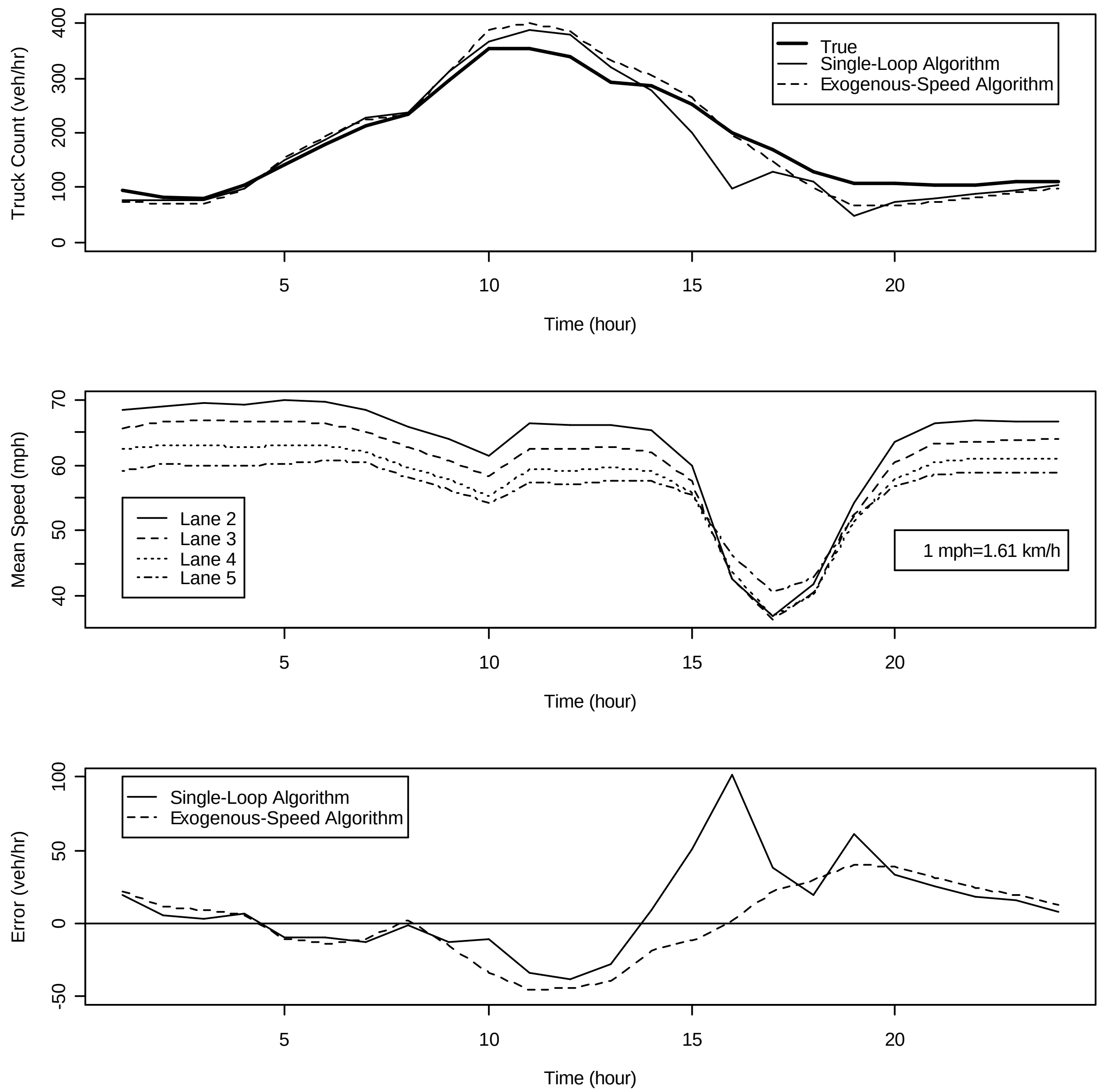

FIGURE 5 Observed and estimated daily pattern of lane-total truck count (top); lane-by-lane daily pattern of speed, averaged over 10 days (middle); daily pattern of estimation error of lane-total truck count (bottom).

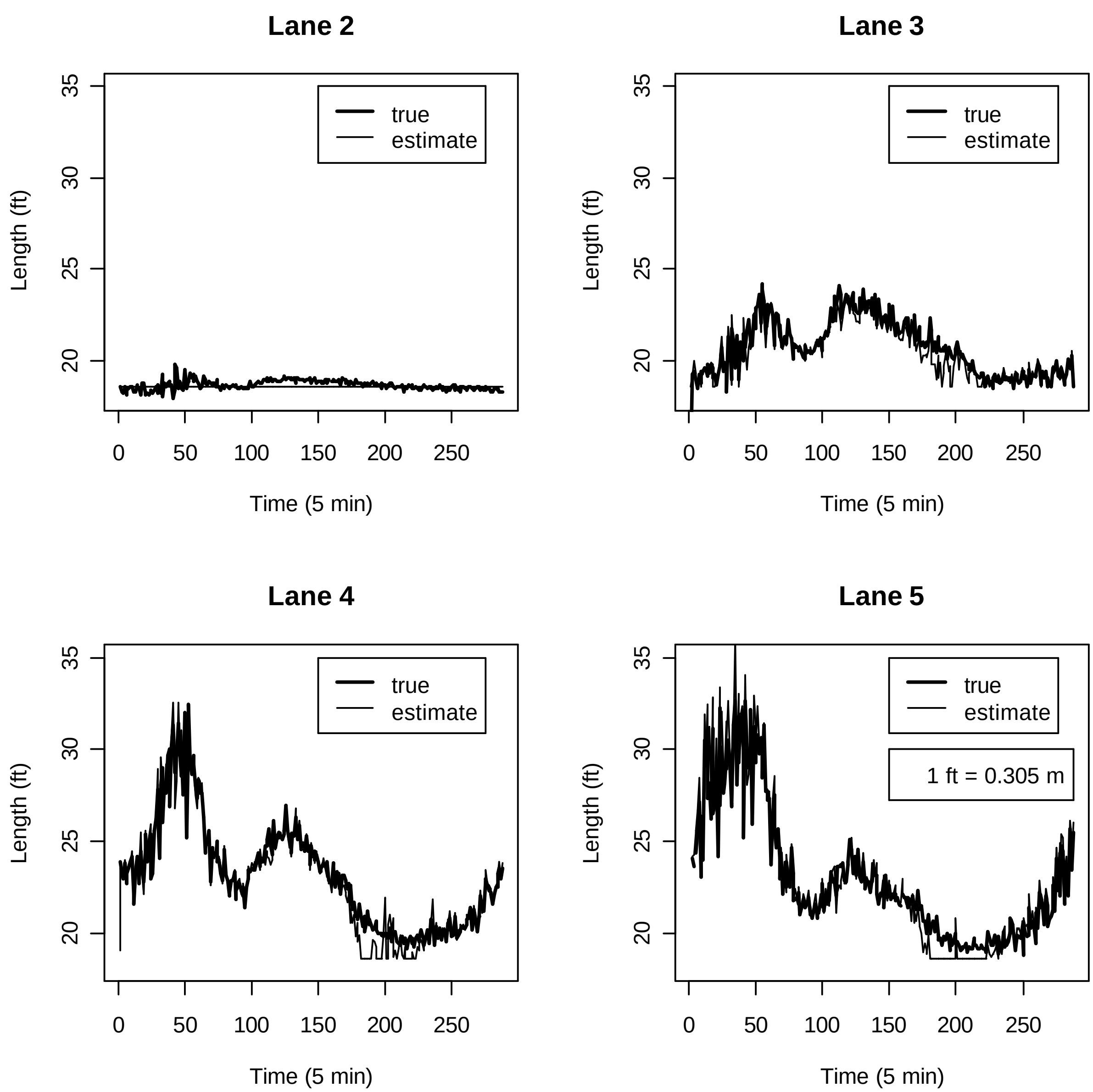

FIGURE 6 True and estimated daily trend of 5-minute mean effective vehicle length estimate, averaged over 10 days.

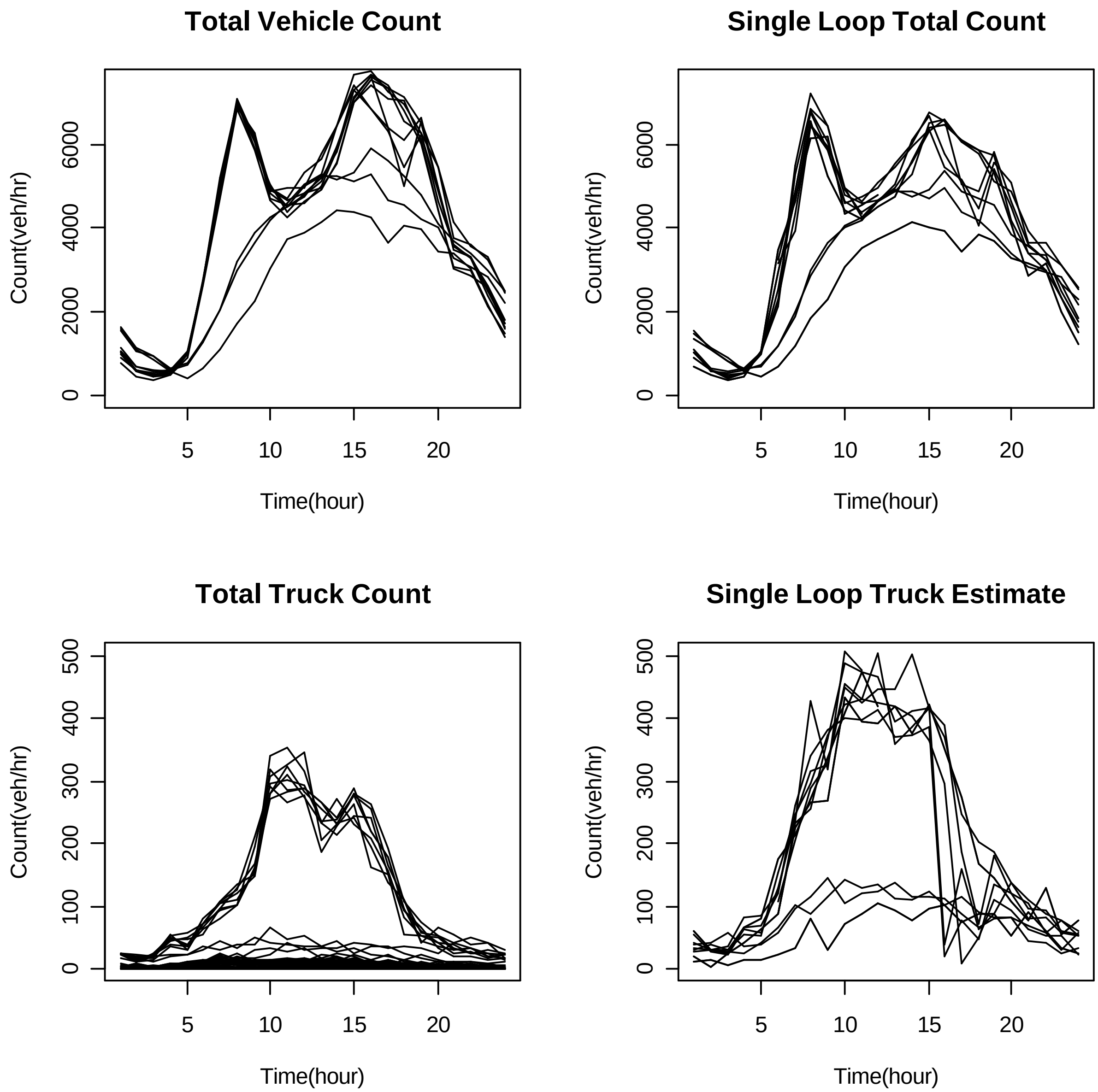

FIGURE 7 Total vehicle volume estimated from WIM data (top left) and single loop data (top right); hourly volume of long truck traffic estimated from WIM data (bottom left) and single loop data using the proposed algorithm (bottom right).

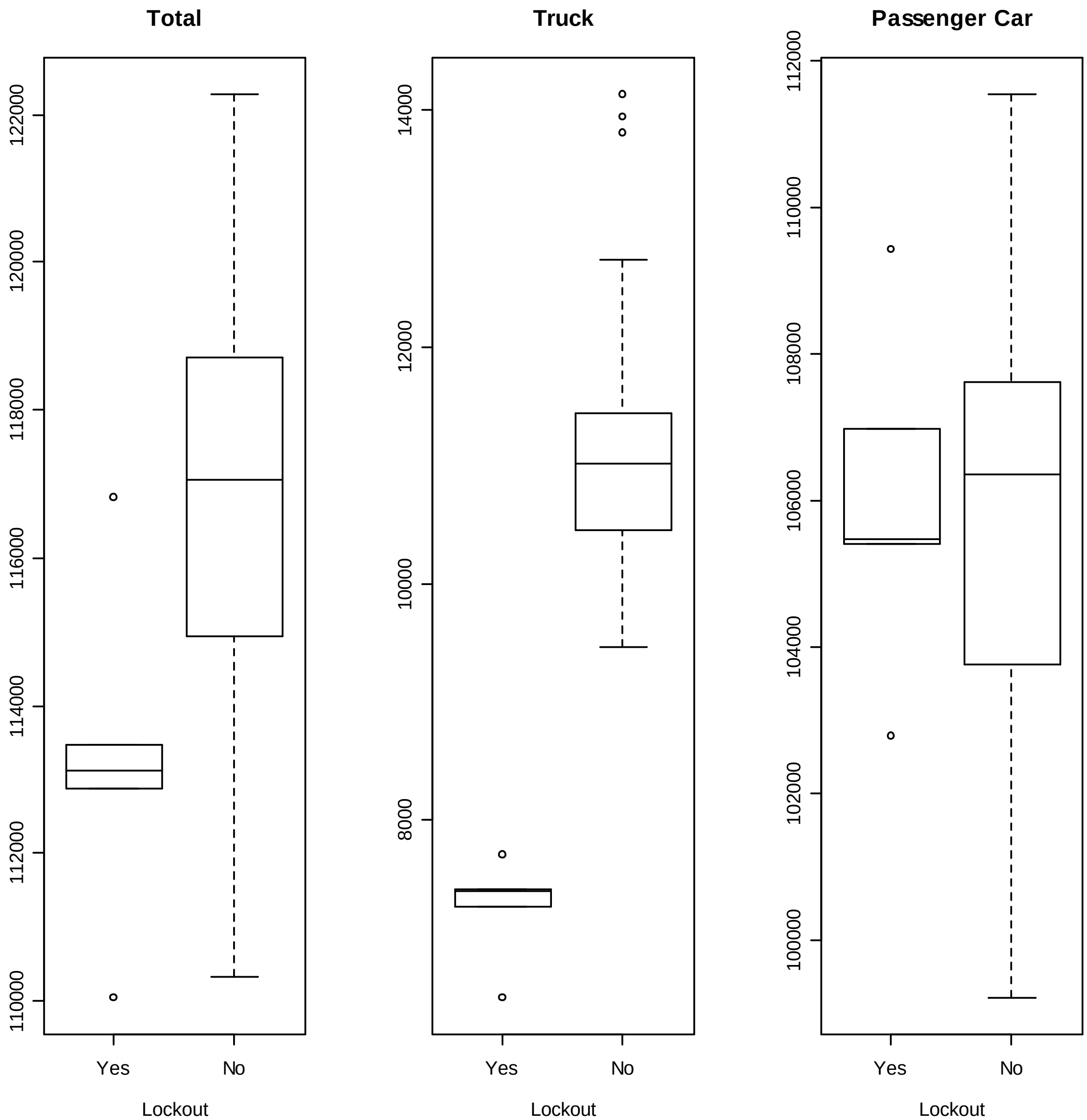

FIGURE 8 Comparison of daily total traffic volume of all cars (left), trucks (middle) and passenger cars (right) during port lockout and non-lockout periods.

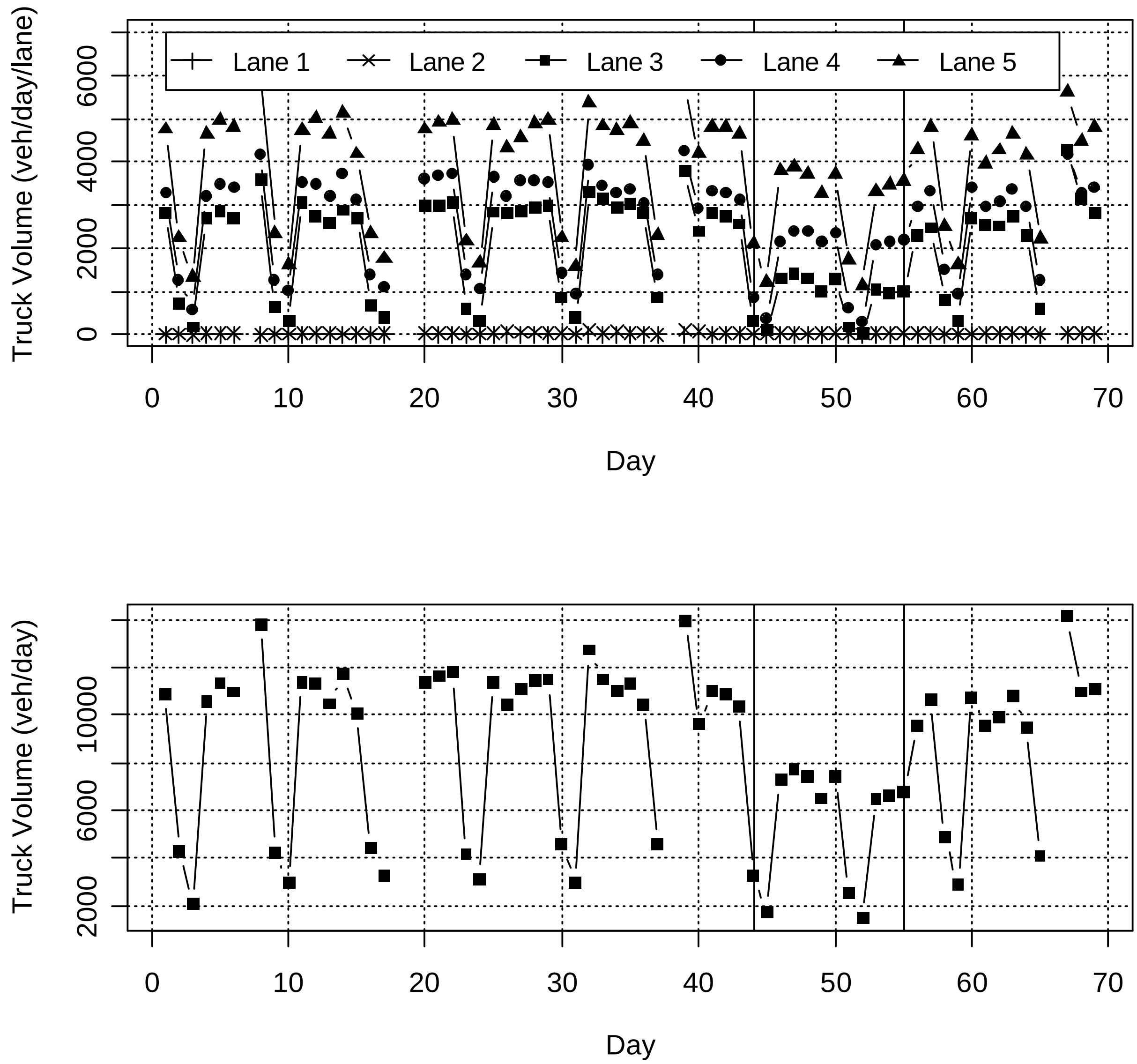

FIGURE 9 Day-to-day trend of daily total truck traffic volume, by lane (top) and lane-total (bottom). The lockout period is days 44-55.