Transportation Planning and Technology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gtpt20

Characteristics of speed dispersion and its relationship to fundamental traffic flow parameters

Chih-Lin Chung[a] & Will W. Recker[b]

[a] Department of Transportation Management, Tamkang University, No.151, Yingzhuan Road, Tamsui District, New Taipei City 25137, Taiwan (Republic of China)

[b] Institute of Transportation Studies, University of California, 4000 Anteater Instruction and Research Building (AIRB), 92697-3600, Irvine, CA, USA

Published online: 04 Aug 2014.

To cite this article: Chih-Lin Chung & Will W. Recker (2014): Characteristics of speed dispersion and its relationship to fundamental traffic flow parameters, Transportation Planning and Technology, DOI: 10.1080/03081060.2014.921404

To link to this article: http://dx.doi.org/10.1080/03081060.2014.921404

Transportation Planning and Technology, 2014
http://dx.doi.org/10.1080/03081060.2014.921404

Routledge
Taylor & Francis Group

Characteristics of speed dispersion and its relationship to fundamental traffic flow parameters

Chih-Lin Chung[a]* and Will W. Recker[b]

[a]*Department of Transportation Management, Tamkang University, No.151, Yingzhuan Road, Tamsui District, New Taipei City 25137, Taiwan (Republic of China);* [b]*Institute of Transportation Studies, University of California, 4000 Anteater Instruction and Research Building (AIRB), 92697-3600, Irvine, CA, USA*

(*Received 20 January 2012; accepted 1 April 2014*)

Speed dispersion is essential for transportation research but inaccessible to certain sensors that simply record density, mean speed, and/or flow. An alternative is to relate speed dispersion with these available parameters. This paper is compiled from nearly a quarter million observations on an urban freeway and a resulting data-set with two speed dispersion measures and the three fundamental parameters. Data are examined individually by lane and aggregately by direction. The first dispersion measure, coefficient of variation of speed, is found to be exponential with density, negative exponential with mean speed, and two-phase linear to flow. These empirical relationships are proven to be general for a variety of coefficient ranges under the above function forms. The second measure, standard deviation of speed, does not present any simple relationships to the fundamental parameters, and its maximum occurs at around a half to two-thirds of the free flow speed. Speed dispersion may be significantly different by lane.

Keywords: speed dispersion; coefficient of variation of speed; standard deviation of speed; macroscopic traffic flow model

1. Introduction

Speed dispersion plays a key role in various aspects: for instance, traffic safety studies have shown that 'speed dispersion kills'; value pricing studies commonly associate travel reliability with speed dispersion; and operating efficiency, air emissions, and energy/gas consumption are all affected by speed dispersion. Unlike the fundamental traffic flow parameters (mean speed, density, and flow), research on the characteristics of speed dispersion is relatively sparse and incomplete. Speed dispersion is inaccessible in two ways. First, many traffic sensors, including ultrasonic and unpaired inductive loops, magnetometers, magnetic induction coils, and infrared, do not record individual speeds, and are unable to capture speed dispersions. Second, for the sensors capable of measuring individual speeds, it is not speed dispersion but mean speed that is the standard output. Obtainment of speed dispersion relies on exogenous calculation, and tends to be neglected by system administrators who typically release the fundamental parameter-based traffic information to the general public and academia.

*Corresponding author. Email: cchung@mail.tku.edu.tw

Among the speed dispersion measures, the coefficient of variation of speed (CVS) and standard deviation of speed (SDS) are most widely used. May (1990) indicated that CVS might range from approximately zero to something on the order of the reciprocal of the square root of the mean speed and normally ranges from 8% to 17% in the empirical studies. Del Castillo and Benítez (1995) set CVS 15% or less, as a rule of thumb, to filter off the nonstationary regime, but did not mention the relationship between CVS and the fundamental parameters. Based upon observations of urban Chinese highways, Wang et al. (2007) proposed *flow* as an exponential regression equation of SDS with the coefficient of correlation (R^2) between 0.26 and 0.74. They also identified *density* as an exponential equation of CVS that distributes from 7% to 32% as a result of R^2 of 0.34. Treiber, Hennecke, and Helbing (1999) adopted empirical data from a Dutch motorway and approximated CVS *square* as a hyperbolic tangent function of *density*. Such approximation displayed positive correlation between CVS *square* and *density* during the stationary regime. Shankar and Mannering (1998) explained lane-by-lane SDS by SDS of adjacent lanes, mean speed, various dummy variables of time, and truck-to-passenger car flow ratio. Their data were collected from a rural section of I-90 in Washington State, and the overall R^2 were from 0.31 to 0.33.

In general, the literature on speed dispersion has provided useful information on speed dispersion. The diversity of conclusions is probably because of small sample sizes and insignificant R^2. Those findings are somewhat limited regarding the influence of each fundamental parameter on speed dispersion. The objective of this paper, given that the characteristics of speed dispersion are neither practically accessible nor theoretically complete, is to construct generalized relationships between speed dispersion and those easily accessible fundamental parameters. Similar to prior studies, this paper focuses on CVS and SDS, but adopts a larger number of observations individually by lane and aggregately by direction for more detailed results. The observed data will first be validated to ensure the reliability of this case study. Then the case-specific outcomes will be compiled and contrasted with the theoretical forms for generalization. The outcomes and framework presented here can facilitate future speed dispersion-related studies.

2. Methodology

The research on which this paper is based begins with empirical highway data from automated data recorders. California tops the USA for over 25,000 single inductive loop detectors in its highway system. But speed dispersion is not, directly or indirectly, available through these single loops. The 2.7-mile I-80 test bed in Berkeley, monitored by the University of California, becomes a reliable data source. The test bed's raw data from dual inductive loop detectors can be utilized to calculate speed dispersion.

2.1. Mathematical description

The following procedures populate the complete data-set in this study:

(1) Acquire raw data that record on- and off-time of loop occupancy.
(2) Apply Equation (1) for individual speeds, as suggested by the Traffic Detector Handbook (Federal Highway Administration 2006).

$$v_i = \frac{1}{2}\left(\frac{d}{Td_{i:\mathrm{on}} - Tu_{i:\mathrm{on}}} + \frac{d}{Td_{i:\mathrm{off}} - Tu_{i:\mathrm{off}}}\right) \times \frac{60^3}{5280} \tag{1}$$

where v_i (in miles per hour) is the speed of individual vehicle i, $Td_{i:\text{on/off}}$ (in 1/60 sec) is the time that the downstream detector is on/off, $Tu_{i:\text{on/off}}$ (in 1/60 sec) is the time that the upstream detector is on/off with respect to vehicle i, and d is the distance between the center points (20 ft in this application).

(3) Within a given time interval (5 min in this application), *space mean speed* (S, in miles per hour) and *time mean speed* (S_T, in miles per hour) are respectively the harmonic and arithmetic means, as shown in Equation (2).

$$S = \frac{n}{\sum_{i=1}^{n} \frac{1}{v_i}}; \; S_T = \frac{\sum_{i=1}^{n} v_i}{n} \tag{2}$$

where n is the vehicle count in the time interval and the hourly *flow* (in veh/hr/ln or vehicles per hour per lane [vphpl]) is 12 n.

(4) Wardrop (1952) verified Equation (3) that can obtain CVS (in percentage) and SDS (in miles per hour) via *time mean speed* and *space mean speed*. It should be noted that typically traffic management centers do not capture both mean speeds or produce speed dispersion in any alternative ways, making speed dispersion inaccessible.

$$S_T = S + \frac{\text{SDS}^2}{S} = S\left(1 + \text{CVS}^2\right) \tag{3}$$

$$\Rightarrow \text{SDS} = \sqrt{S(S_T - S)}; \text{CVS} = \left(\sqrt{\frac{S_T}{S} - 1}\right) \times 100\% \tag{3a}$$

(5) The 5-min mean *occupancy* (in percentage) is calculated as the average over its 30-sec components and serves as a surrogate for the *density.*

A full data-set with two speed dispersion measures (CVS and SDS) and three fundamental parameters (*space mean speed*, *flow*, and *occupancy*) can be accomplished. The interrelationships among these parameters will be acquired via regression analyses using the ordinary least square technique. Unless otherwise specified, the italic *speed* indicates *space mean speed* denoted as S, while *flow* and *occupancy* are respectively denoted as Q and K.

2.2. Data size

The I-80 test bed consists of 10 lanes. The five lanes in each direction are labeled from 1 to 5 from the innermost to outmost. The lanes are for general purpose (GP) traffic except for Lane 1, which is designated as a continuous-access high-occupancy vehicle (HOV) lane during 5–10 am and 3–7 pm. The speed limit is 65 mph for both lane types. Data were collected across each lane during the weekday HOV hours to eliminate the

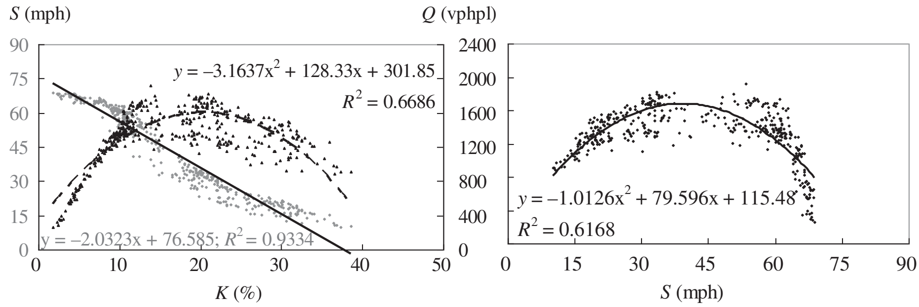


Figure 1. Relationships between *occupancy K*, *speed S*, and *flow Q* (all-lane mix).

effects of the non-HOV hours. The complete data-set contains 422 observations that correspond to a total vehicle count of 233,026.

2.3. Data validity

Data validity is examined on four aspects. First, CVS ranged from 10% to 35% northbound and 10% to 55% southbound. The CVS range of the northbound is close to the observation by Wang et al. (2007) from 7% to 32%, but that of the southbound is broader than the proposal by May (1990) from 0 to $\frac{1}{\sqrt{S}}$, which implies that *time mean speed* is within 1 mph greater than *space mean speed*. In fact, Rakha and Zhang (2005) indicated that large differences between these two *mean speeds*, from 10% to 30%, are not uncommon when traffic is congested. Such differences correspond to CVS up to 65% and justify our observations of a greater CVS range. Second, the observed speeds were found to be normally distributed overall. This complies with McShane and Roess (1990) and May (1990). Third, the fundamental parameters were inspected for background information, as shown in Figure 1. The all-lane mix serves as representative relationships among *speed*, *flow*, and *occupancy*, given that the individual lanes present similar scatter plots. The well-known Greenshields' equations were depicted for reference, albeit more complex forms may better fit. The plots match such common recognitions as wider fluctuations in the congested regime, a gap around the critical point, stable mean speed during light traffic, and so on. Fourth, and finally, for CVS of the individual lanes in the uncongested state, the mean values were between 9.1% and 10.8%, and the 85th percentiles were below 14.1%. For CVS of the lane mixes in the uncongested state, the mean values were between 11.9% and 13.5% and the 85th percentiles were below 15.4%. These results are primarily consistent with Del Castillo and Benítez (1995) who set CVS of 15% as the lower bound for the congested state.

3. Building empirical relationships

Regression analyses are conducted respectively by lane, lane type, and direction – a total of seven categories: one for each of the five individual lanes, one for the aggregated four GP lanes, and another one for the aggregated five lanes in one direction. The relationships between speed dispersion and the fundamental parameters are depicted for initial screening, and statistically approximated by certain popular forms like linear, exponential, logarithmic, polynomial, power functions, etc. Two and more forms may be presented if no one dominates over others, but only one form will eventually be suggested with revealed pros and cons. As some complicated forms are not considered, the suggested form is not necessarily the best fit, but rather a better fit regarding ease of use, understanding, and compatibility.

3.1. CVS vs. occupancy

We begin with *occupancy* and CVS since *density*, represented by *occupancy*, is sensitive to a broad range of traffic conditions (Transportation Research Board 2000). Figure 2 shows that CVS increased with *occupancy*, particularly when traffic became median and heavy. The exponential forms can better explain the relationship, albeit the unlisted quadratic is slightly superior to the exponential for Lane 1 (the HOV lane). As a contrast, Wang et al. (2007) suggested *density* be an exponential form of CVS based upon around 40 observations and with R^2 of 0.34. It is the reverse of what was found here – CVS is an exponential form of *occupancy* (as well as *density* attainable from *occupancy* via a multiplier). We examine their suggestion by fitting *occupancy* as an exponential form of CVS, but the R^2 associated with the all-lane mix drops from 0.75 to 0.55. Also given that each R^2 and the data-set in Figure 2 are more significant than the study of Wang et al., we suggest CVS more properly be an exponential form of *occupancy*.

Lane 1 was less congested than the other four GP lanes, which resulted in some ‘missing’ observations potentially in the upper right corner of the Lane 1 diagram in Figure 2. This is likely responsible for a smaller R^2 than other categories. The all-lane mix and GP-lane mix have greater R^2 (from 0.70 to 0.75) than the individual lanes (from 0.49 to 0.66). The CVS–*occupancy* relationship can be visually and statistically classified into group 1 (the all-lane mix), group 2 (Lanes 1 and 2 and the GP-lane mix), and group 3 (Lanes 3–5), as shown in the summary diagram in Figure 2. Group 1 has the largest speed dispersion with respect to fixed *occupancy*, followed by group 2 and then group 3. The three groups can be approximated by Equation (4) in single expression:

$$\mathrm{CVS} = \mathrm{cv}_f \mathrm{Exp}(aK) \tag{4}$$

where cv_f (in percentage) is the CVS in the free flow state when *occupancy* (K) is about 0.

$$a \approx \frac{0.078 + 0.042 d_1 + 0.017 d_2}{\ln(\mathrm{cv}_f)},$$

$$d_1 = \begin{cases} 1: \text{all-lane mix (group 1)} \\ 0: \text{otherwise (groups 2 or 3)} \end{cases}; \quad d_2 = \begin{cases} 1: \text{lanes 1, 2, and GP-lane mix (group 2)} \\ 0: \text{otherwise (groups 1 or 3)} \end{cases}$$

cv_f across lanes were similar, ranging from 5.3% to 6.9%. cv_f can be regarded as the minimum CVS. Since there are few vehicles in the free flow state, a variety of driving behaviors/conditions among motorists have a greater impact on cv_f than the traffic and road factors do. Such a variety includes, but is not limited to, different interpretations of the speed limit (some going above or below the speed limit), distractions from talking, eating, etc., as well as mental and physical conditions that cause inconsistent speed in the free flow state.

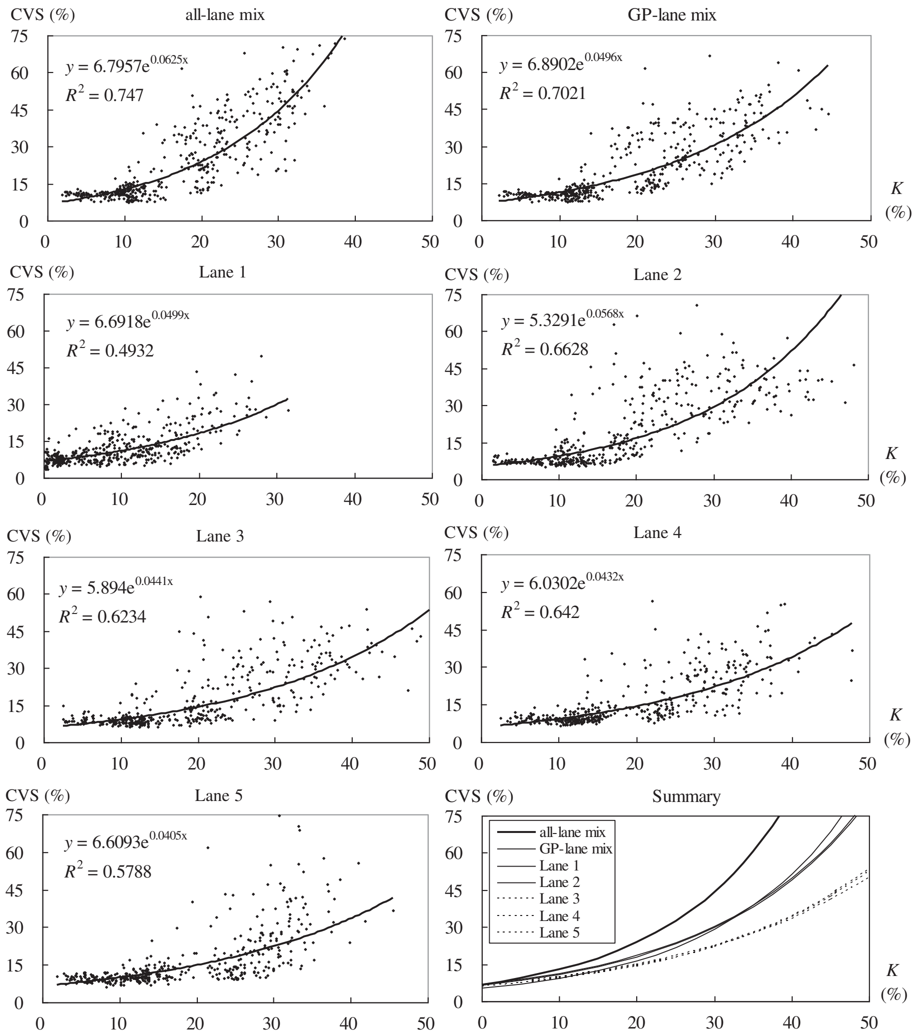


Figure 2. Relationships between *occupancy* and CVS.

3.2. CVS vs. speed

CVS would be expected as a negative exponential form of *speed* since, in general, CVS was exponentially related to *occupancy* (Figure 2) and *occupancy* had a negative linear relationship to *speed* (Figure 1 left). This anticipation is verified by Figure 3 with R^2 of about 0.6 or greater. Quadratic forms are slightly worse than the exponential forms and not listed. Similar to the CVS–*occupancy* relationship, the all-lane mix and GP-lane mix in general have a better fit (R^2 from 0.76 to 0.83) than the individual lanes (R^2 from 0.59 to 0.79). On the contrary, the downward curves signify that CVS decreased with *speed*.

The seven categories can also be classified into group 1 (the all-lane mix), group 2 (Lane 2 and the GP-lane mix), and group 3 (Lanes 1, 3, 4, and 5), as shown in the summary diagram in Figure 3. The only distinction from the CVS–*occupancy* grouping is that Lane 1 is now grouped with the outer lanes (Lanes 3–5) instead of with Lane 2 and the GP-lane mix. The relatively low R^2 of Lane 1 may be a reason for the grouping

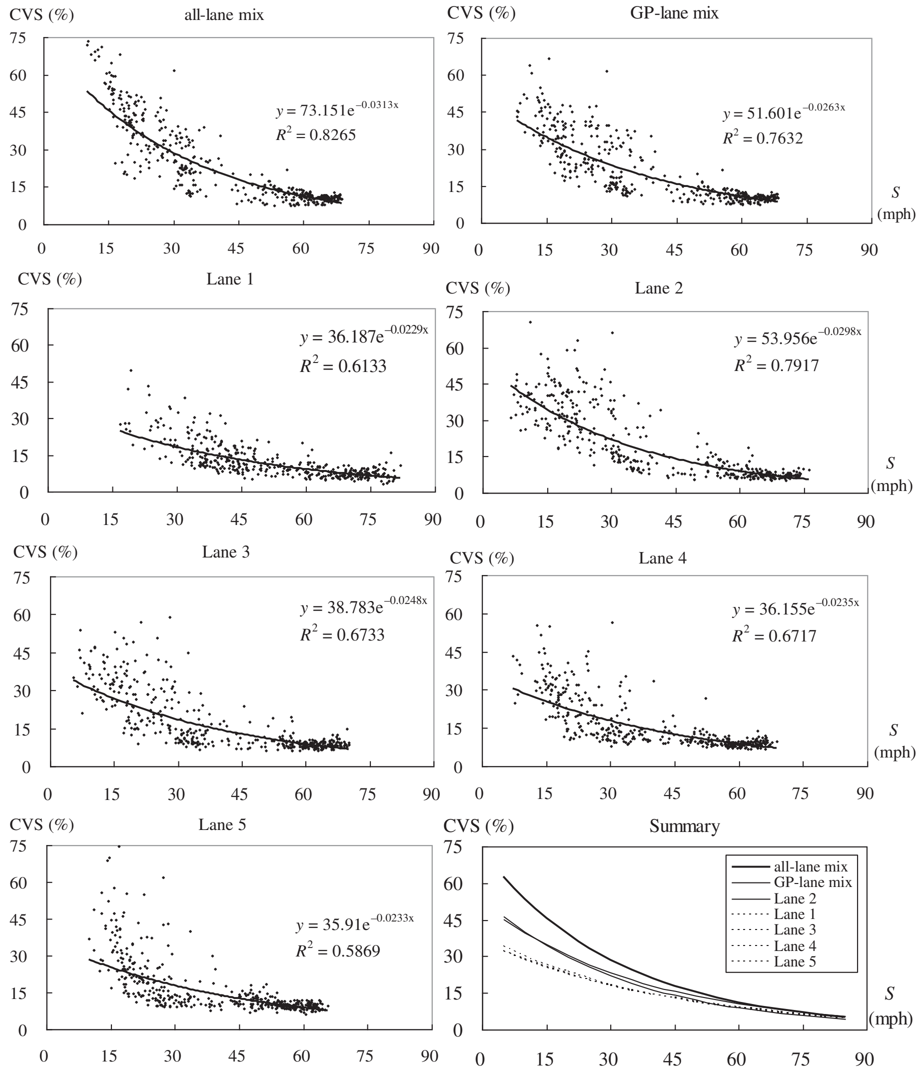


Figure 3. Relationships between *speed* and CVS.

difference. Nevertheless, under fixed *occupancy* or *speed*, both figures are consistent in the all-lane mix with the greatest CVS, Lanes 3–5 with the least CVS, and Lane 2 and the GP-lane mix in between, as expressed below.

$$\mathrm{CVS} = \mathrm{cv}_j \mathrm{Exp}(bS) \tag{5}$$

where cv_j (in percentage) is the CVS in the jam state when *speed* (*S*) is about 0.

$$b \approx -0.024 - 0.007d_1' - 0.006d_2'$$

$$d_1' = \begin{cases} 1: & \text{all - lane mix (group 1)} \\ 0: & \text{otherwise (groups 2 or 3)} \end{cases}; \quad d_2' = \begin{cases} 1: & \text{lane 2 and GP}-\text{lane mix (group 2)} \\ 0: & \text{otherwise (groups 1 or 3)} \end{cases}$$

cv_j can be regarded as the maximum CVS, which is affected by the stop-and-go traffic, road, and motorist factors. Another contributor to the maximum CVS is the aggressive driving related to congestion – a commonly recognized behavior in the traffic physiology and behavior research (Shinar 1999). The range of cv_j varied with the groups; it was 73.2% for group 1, from 51.6% to 54.0% for group 2, and from 35.9% to 38.8% for group 3. The lane mixes had greater cv_j than most individual lanes due to a variety of vehicle types, lane types, and driving behaviors. Lane 2 that serves as the passing lane for Lane 3 (the GP lane) and Lane 1 (the HOV lane) had higher cv_j than other individual lanes.

3.3. CVS vs. flow

CVS and *flow* had a two-phase linear relationship that respectively corresponded to the congested and uncongested states. As identified in Figure 4, the two states intersect at around the lane capacity and the mean CVS of the uncongested state. Although CVS during congestion (black dots in Figure 4) could be explained by either a linear or an exponential form of *flow*, the linear relationship is preferred for its simplicity. Similar to the beginning stage of the CVS–*occupancy* relationship and the ending stage of the CVS–*speed* relationship, CVS during the uncongested state (gray dots in Figure 4) was nearly a constant or slightly increased with *flow* from 0 to over 2000 vphpl. The lanes are not grouped because of poor fitness scores. Consistently, the all-lane mix had the greatest R^2 and Lane 1 had the least. Also, the lane mixes had R^2 greater than the individual lanes. It should be pointed out that Shankar and Mannering (1998) found lane-by-lane speed dispersion correlated with multiple explanatory variables, including speed dispersion and mean speed of the adjacent lane(s). As R^2 for individual lanes in our study are primarily at a significant level, we only adopt a single fundamental parameter to explain CVS. Doing so can avoid possible collinearity among the explanatory variables and makes the model easy to apply. Omitting interactions between lanes, on the other hand, may cause R^2 for each individual lane to be lower than the lane mixes that would internalize such interactions.

3.4. SDS vs. occupancy, speed, and flow

No simple equations were found to be valid between SDS and the fundamental parameters, as shown in Figure 5 that takes the all-lane mix as representative. The majority of SDS fell in the range of 4–12 mph. On average, SDS in the congested state was more spread out and greater than that in the uncongested state. *Occupancy* along with *speed* may be expected to jointly explain SDS, as shown in Equation (6) derived from Equation (4). Another 'complicated' form between *speed* and SDS may be expected, as shown in Equation (7) derived from Equation (5). Finally, although Wang et al. (2007) proposed *flow* as an exponential form of SDS, we do not have similar findings but expect, through the relationship between *flow* and CVS, that *flow* and *speed* can jointly explain SDS to a certain degree.

$$\mathrm{SDS} = \mathrm{cv}_f \cdot S \cdot \mathrm{Exp}(aK) \tag{6}$$

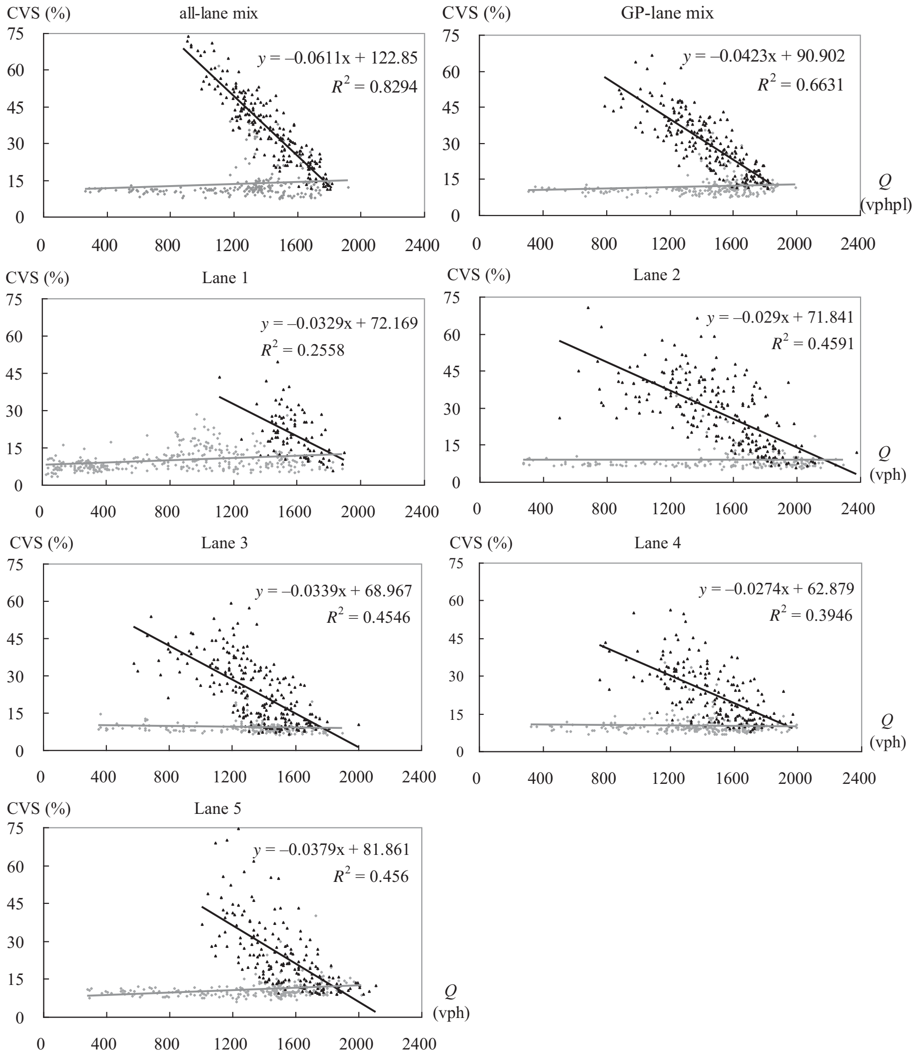


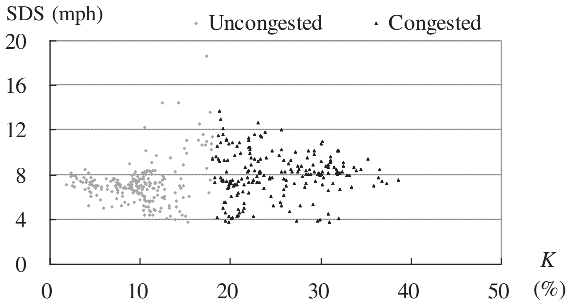
Figure 4. Relationships between *flow* and CVS.

$$\mathrm{SDS} = \mathrm{cv}_j \cdot S \cdot \mathrm{Exp}(bS) \tag{7}$$

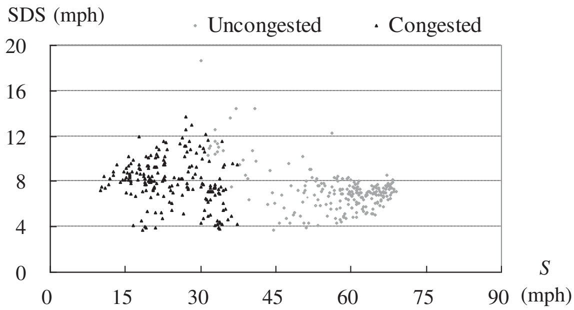
where cv_f, a, cv_j, and b were defined in Equations (4) and (5).

Recall that the minimum CVS (cv_f) was from 5.3% to 6.9%, corresponding to SDS between 3.7 mph and 4.8 mph under a presumable free flow speed of 70 mph. The maximum CVS (cv_j) is from 35.9% to 73.2%, corresponding to SDS between 1.8 mph and 3.7 mph under a presumable jam speed of 5 mph. Based upon Figure 5 as well as Equations (6) and (7), SDS does not strictly increase with traffic in terms of *occupancy*, *speed*, or *flow*. Rather, the maximum SDS would occur at a certain *speed* that makes the first derivative of Equation (7) zero, i.e.

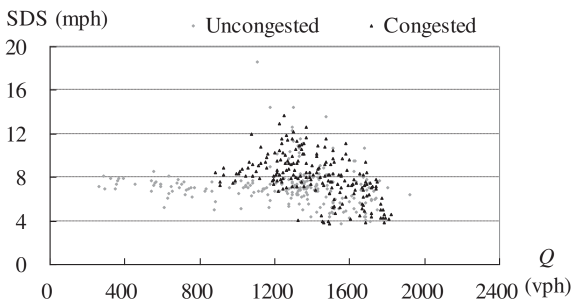
$$\frac{d\mathrm{SDS}}{dS} = 0 \Rightarrow \mathrm{cv}_j(bS+1)\mathrm{Exp}(bS) = 0 \Rightarrow S = -\frac{1}{b} \tag{8}$$



(a) *SDS – occupancy*



(b) *SDS – speed*



(c) *SDS – flow*

Figure 5. Scatter plots of the fundamental parameters and SDS (all-lane mix).

Given the empirical values of b in Figure (3), the maximum SDS would occur at S = 44, 34, 40, 43, 43, 38, and 32 mph, respectively for Lanes 1 through 5, the GP-lane mix, and all-lane mix. S = 32 mph resembles Figure 5(b) with the maximum SDS at S around 30 mph for the all-lane mix. The traffic condition in a half to two-thirds of free flow speed appears to have the maximum SDS. As for the minimum SDS, it is expected to occur at the jam state by two judgments: straightforwardly, there is little room for speed deviation at the jam state and functionally, Equation (7) results in SDS of nearly zero if S is close to zero.

Figure 6 reveals the descriptive characteristics of SDS. First, SDS of each individual lane stayed relatively steady during light traffic with its means between 5.6 mph and

	Uncongested state		Congested state	
	Mean (mph)	Number of observation	Mean (mph)	Number of observation
Lane 1	5.5	330	6.1	92
Lane 2	5.5	187	6.1	235
Lane 3	5.6	189	5.0	233
Lane 4	5.5	231	4.9	191
Lane 5	5.6	239	4.9	183
GP-lane mix	6.7	214	6.4	208
All-lane mix	7.2	224	8.1	198

Range of the majority *SDS* (75% tile – 25% tile)

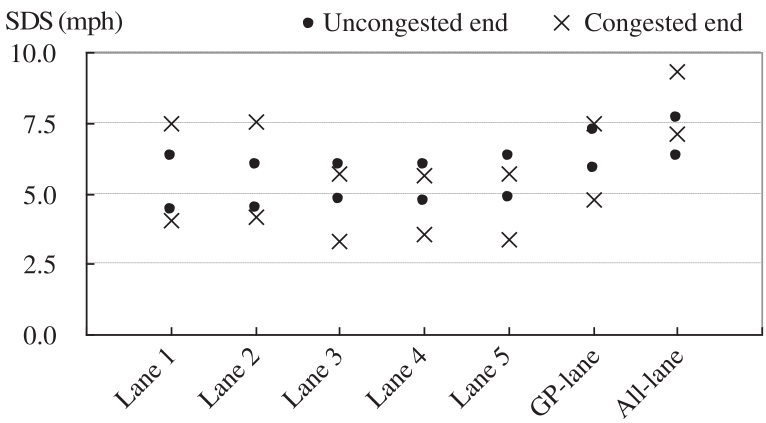


Figure 6. Descriptive statistics of SDS by lane and congestion level.

5.7 mph and the majority (the 75th percentile to 25th percentile) within 4–6 mph. When traffic became congested, SDS fluctuated more, which might have been caused by lane changing that leaves gaps for the following vehicle to speed up, and/or by stop-start waves that happen only in congestion. Second, SDS in Lanes 1 and 2 during congestion was on average greater than those associated with uncongested conditions. This can be explained by the supposition that when congestion in the adjacent GP lane (Lane 2) deteriorates, violators are more likely to rush into and out of the HOV (Lane 1) for short time intervals with increasing frequency. This factor was proposed by Varaiya (2007) to justify capacity loss of HOV lanes with respect to GP lanes. Also, since an HOV lane operates as a one-lane highway, its speed is governed by the low-speed vehicles – the 'snails' (Varaiya 2007). When traffic worsens, a faster HOV may be eager to pass the 'snail' in front by darting into and out of Lane 2 more frequently. These two factors force drivers (not only in the HOV lane but also in the adjacent GP lanes) to adjust speeds, causing greater SDS in Lanes 1 and 2 during congestion. Third, SDS in Lanes 3–5 under uncongested conditions were on average greater than those in congested conditions. This is probably because the outer lanes usually have higher percentages of trucks and conservative motorists who tend to stay in lane when it becomes difficult to find a gap large enough for lane changing in the congested state.

4. Generalizing the relationships

We apply an enumeration procedure to address the concern: whether the proposed relationships are general for urban freeways or just case-specific for the study site. It appears that area disparity may affect the coefficients of the functions instead of the functions themselves, as described below.

4.1. CVS–speed relationship

We first refer to the previous studies that presented *time mean speed* S_T statistically linear to *space mean speed* S and vice versa (Drake, Schofer, and May 1967; Garber and Hoel 2002; Rakha and Zhang 2005), i.e.

$$S_T = pS + q \tag{9}$$

where p is usually between 0.9 and 1 and q between 2.5 and 5.

Equations (9) and (3) convert the general CVS–*speed* relationship into Equation (10).

$$S_T = S + \frac{\mathrm{SDS}^2}{S} \Rightarrow \frac{pS+q}{S} = 1 + \mathrm{CVS}^2 \Rightarrow p - 1 + \frac{q}{S} = \mathrm{CVS}^2, \text{ then}$$

$$\ln(\mathrm{CVS}) = \frac{1}{2}\ln\left(p - 1 + \frac{q}{S}\right), \text{ for CVS in decimal, or}$$

$$\ln(\mathrm{CVS}) = \frac{1}{2}\ln\left(p - 1 + \frac{q}{S}\right) + \ln(100), \text{ for CVS in percentage} \tag{10}$$

If we take natural logarithm of the case-oriented CVS–*speed* relationship shown in Equation (5), it would be transformed to Equation (11).

$$\ln(\mathrm{CVS}) = \ln(\mathrm{cv}_j) + bS \tag{11}$$

If we can always find $\ln(\mathrm{cv}_j) + bS$ close to $\frac{1}{2}\ln\left(p - 1 + \frac{q}{S}\right) + \ln(100)$, Equations (10) and (11) are exchangeable on a regular basis that generalizes the case-oriented CVS–*speed* relationship. Given the ranges of p and q, all possible situations are enumerated as follows:

- Vary p from 0.9 to 1 with an increment of 0.01 (11 counts) and q from 2.5 to 5.0 with an increment of 0.1 (26 counts). It totals the (p, q) pairs of 286.
- For each (p, q) pair, compute $\frac{1}{2}\ln\left(p - 1 + \frac{q}{S}\right) + \ln(100)$ by varying *speed* from 2.5 mph to 75 mph with an increment of 2.5 (30 counts). Since $\left(p - 1 + \frac{q}{S}\right)$ may be negative as S increases, each (p, q) pair would have up to 30 values of ln(CVS).

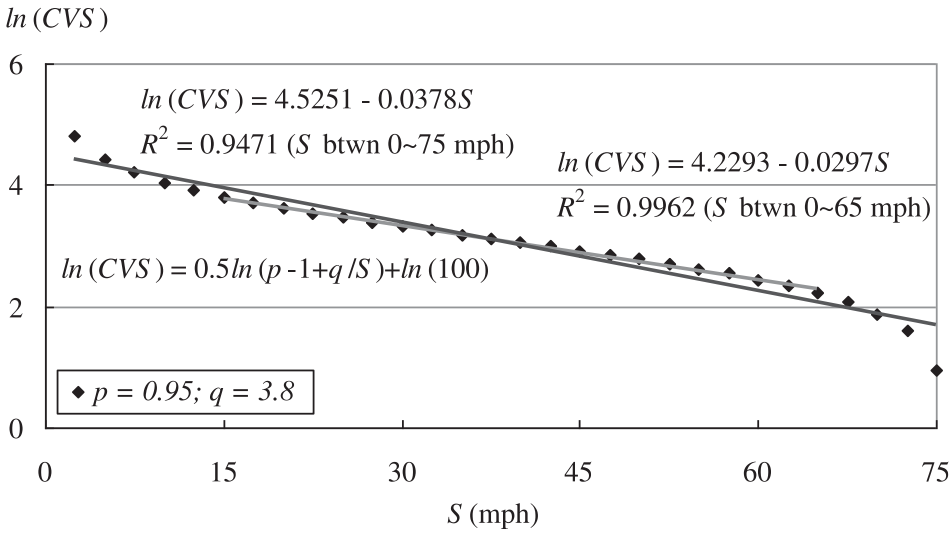


Figure 7. Linearization of the *speed*–ln(CVS) curve.
Note: CVS in unit of %.

- An influence point i may occur at $\left(\left(p-1+\frac{q}{S_i}\right)>0 \text{ and } \left(p-1+\frac{q}{S_i+2.5}\right)<0\right)$ for certain (p, q) pairs, and shall be removed.
- Conduct correlation analysis to find out how well the general relationship can be explained by the case-oriented relationship.

Figure 7 shows a representative (p, q) pair at their mediums of 0.95 and 3.8. $\ln(\mathrm{CVS}) = \ln(\mathrm{cv}_j) + bS$ with *speed* from 2.5 mph to 75 mph (the black line) displays a very good fit of the curve $\ln(\mathrm{CVS}) = \frac{1}{2}\ln\left(p-1+\frac{q}{S}\right) + \ln(100)$. If we exclude the extremes at both speed ends, the curve would be nearly perfectly fit by $\ln(\mathrm{CVS}) = \ln(\mathrm{cv}_j) + bS$ with *speed* from 15 mph to 65 mph (the gray line). For all 286 (p, q) pairs, 143 (50%) pairs have R^2 greater than 0.95, 72 (25%) between 0.95 and 0.9, and the remaining 71 (25%) between 0.9 and 0.85; this enables use of Equation (11) to replace Equation (10) and generalizes the case-oriented relationship.

4.2. *CVS–occupancy relationship*

Based upon the general CVS–*speed* relationship, it is expected that CVS–*occupancy* also exhibits an exponential form as long as *speed* S has certain (one- or two-phase) linear patterns with *occupancy* K. Assume the general *occupancy–speed* relationship as:

$$S = \begin{cases} s_f, \quad \textit{for } K \le d \text{ (phase I)} \\ \left(\dfrac{a}{b}\right)K + s_f, \quad \text{otherwise (phase II)} \end{cases} \tag{12}$$

The two-phase linear pattern describes that *speed* stays at free flow speed s_f when *occupancy* below a certain degree d, and afterward decreases as *occupancy* goes up with a slope of $\left(\frac{a}{b}\right)$. The pattern becomes one-phase if $d = 0$. The general CVS–*speed* relationship CVS = cv_j Exp(bS) yields:

$$\mathrm{CVS} = \begin{cases} \mathrm{cv}_j \mathrm{Exp}(bs_f) = \mathrm{cv}_f, & \text{for K} \leq \text{d (phase I)} \\ \mathrm{cv}_j \mathrm{Exp}(aK + bs_f) = \mathrm{cv}_f \mathrm{Exp}(aK), & \text{otherwise (phase II)} \end{cases} \tag{13}$$

Based upon the empirical data, a and d are known to be small, making $\mathrm{Exp}(aK) \approx 1$ and thus $\mathrm{cv}_f \approx \mathrm{cv}_f \mathrm{Exp}(aK)$ when $K \leq d$. The two phases in Equation (13) are combined into a general form as $\mathrm{CVS} = c_f \mathrm{Exp}(aK)$, which is identical to the case-oriented one in Equation (4). Use of phase II to replace phase I in Equation (13) can also be explained in a straightforward sense: an exponential function like $\mathrm{CVS} = c_f \mathrm{Exp}(aK)$ is characterized as CVS insensitive to *occupancy* below a certain degree; this corresponds to the phase I of the *occupancy–speed* relationship that *speed* is insensitive to *occupancy.*

4.3. CVS–flow relationship

From Equation (4): $K = \frac{\ln(\mathrm{CVS}) - \ln(\mathrm{cv}_f)}{a}$

From Equation (5): $S = \frac{\ln(\mathrm{CVS}) - \ln(\mathrm{cv}_j)}{b}$

$Q = flow = density \cdot speed = g \cdot K \cdot S$, where g is a density conversion factor.

$$\Rightarrow Q = \frac{(\ln(\mathrm{CVS}))^2 - (\ln(\mathrm{cv}_f) + \ln(\mathrm{cv}_j))\ln(\mathrm{CVS}) + \ln(\mathrm{cv}_f)\ln(\mathrm{cv}_j)}{ab/g}$$

$$\text{or } \mathrm{CVS} = \sqrt{\mathrm{cv}_f \mathrm{cv}_j}\,\mathrm{Exp}\left(\pm \sqrt{\left(\frac{ab}{g}\right) Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2} \right) \tag{14}$$

as the general relationship with $\mathrm{CVS} = \sqrt{\mathrm{cv}_f \mathrm{cv}_j}\,\mathrm{Exp}\left(\sqrt{\left(\frac{ab}{g}\right) Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2} \right)$ during the congested state and $\mathrm{CVS} = \sqrt{\mathrm{cv}_f \mathrm{cv}_j}\,\mathrm{Exp}\left(-\sqrt{\left(\frac{ab}{g}\right) Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2} \right)$ during the uncongested state.

We repeat the correlation analysis and enumerate the combinations of cv_f, cv_j, a, and b within their effective ranges. According to Figures 2 and 3, let cv_f vary from 5.3 to 6.9 with an increment of 0.4 (5 counts), cv_j vary from 36 to 72 with an increment of 9 (5 counts), a vary from 0.040 to 0.064 with an increment of 0.006 (5 counts), b vary from –0.023 to –0.032 with an increment of –0.003 (4 counts), and g be a constant of 2.112. It totals 500 sets of (cv_f, cv_j, a, b) for either state. Each set contains *flow* from 50 vphpl to 2000 vphpl with an increment of 50 (40 counts). Since $\left(\frac{ab}{g}\right) Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2$ may be negative as *flow* increases, every set would have up to 40 values of CVS with respect to *flow.*

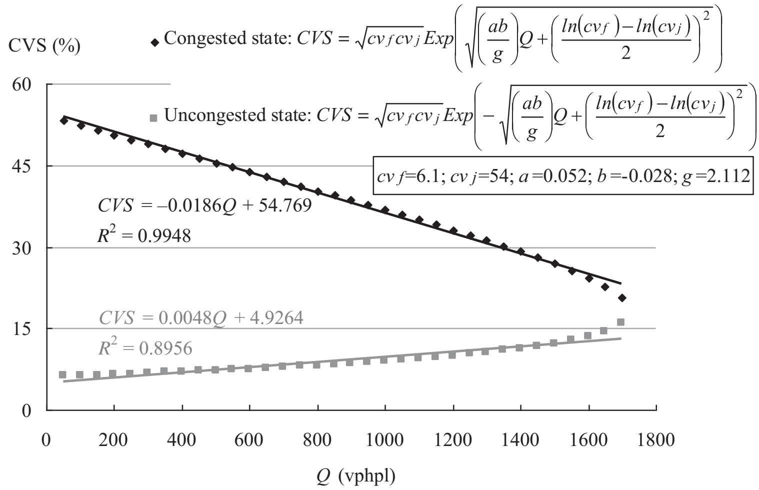


Figure 8. Linearization of the generalized *flow*–CVS relationship.

The results show that $\mathrm{CVS} = \sqrt{\mathrm{cv}_f \mathrm{cv}_j}\mathrm{Exp}\left(\sqrt{\left(\frac{ab}{g}\right)Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2}\right)$ for the congested state is well fit as negative sloped lines with R^2 greater than 0.98 for all (cv_f, cv_j, a, b) sets, while $\mathrm{CVS} = \sqrt{\mathrm{cv}_f \mathrm{cv}_j}\mathrm{Exp}\left(-\sqrt{\left(\frac{ab}{g}\right)Q + \left(\frac{\ln(\mathrm{cv}_f) - \ln(\mathrm{cv}_j)}{2}\right)^2}\right)$ for the uncongested state can also be approximated as flat lines, for 117 (23%) sets with R^2 over 0.95, 211 (42%) between 0.95 and 0.9, and the remaining 172 (34%) between 0.9 and 0.85. A representative set is displayed in Figure 8, which resembles Figure 7 in two ways. First, both figures are featured by two-phase linear that intersects near the lane capacity and CVS of 15%. Second, CVS increases noticeably with the congested traffic but remains relatively stable during uncongested periods. As the curves in either congested or uncongested state could be linearized for the 500 sets, we suggest that a general CVS–*flow* relationship should be two-phase linear.

5. Conclusions

Unlike *speed*, *occupancy*, and *flow* that measure either average or aggregated traffic conditions, speed dispersion provides an alternate way to comprehend traffic by capturing the variation. This study compiled nearly a quarter million of vehicle records into a database with traffic parameters individually by lane and aggregately by direction. The empirical data conclusively indicate that the CVS increased progressively with traffic, leading to a minimum between 5% and 7% in the free flow state for all groups and a maximum around 36% for the individual lanes and over 50% for the lane mixes in the jam state. As for the SDS, it ranged empirically from 4 mph to 12 mph and did not strictly increase with traffic. Rather, its maximum occurred at around a half to two-thirds of free flow speed.

CVS is favored over SDS when using the fundamental parameters to link speed dispersion. Based upon the correlation analysis, CVS can be better explained by *speed* in the form of negative exponential, followed by *occupancy* in the form of exponential, and then *flow* in the form of two-phase linear. Such a result is understandable since CVS measures speed dispersion instead of occupancy or flow dispersion. In the case that *speed* is not available, e.g. single loop detectors do not measure *speed*, *occupancy* can then be a substitute. *Flow* is not suggested except for nonindividual lanes during congestion. Adding a second independent variable to explain CVS is feasible but not necessary, given the already high R^2 by a single fundamental parameter. In general, the statistical relationships fit fairly well for the all-lane mix and GP-lane mix, and should be used with caution for certain individual lanes.

Two of the most popular speed dispersion measures, SDS and CVS, have at least two similarities. First, both measures in the all-lane mix are greater than they are in the individual lanes. This is reasonable since the all-lane mix contains more varieties of vehicle types, driving behaviors, lane restrictions, etc. Second, individual lanes can be grouped overall into 'inner two lanes' with greater speed dispersion and 'outer three lanes' with less speed dispersion. This is probably because the inner two lanes have more lane changing behaviors.

Finally, no evidence indicates that speed dispersion of the continuous-access HOV lane is unique vis-à-vis the individual GP lanes. The data-set of this study matched typical traffic flow patterns, and the statistical function forms were shown to be not unique but generally valid by enumerating the potential ranges for the respective coefficients. Nonetheless, extensive empirical cases and theoretical development are encouraged for future studies, particularly into distinct highway types, speed limits, number of lanes, and possibly driving cultures that may affect the characteristics of speed dispersion.

References

Del Castillo, J. M., and F. G. Benítez. 1995. "On the Functional Form of the Speed-density Relationship, Part Two: Empirical Investigation." *Transportation Research Part B* 29 (5): 391–406. doi:10.1016/0191-2615(95)00009-3.

Drake, J. S., J. L. Schofer, and A. D. May. 1967. "A Statistical Analysis of Speed-density Hypotheses." *Highway Research Record* 154: 53–87.

Federal Highway Administration. 2006. *Traffic Detector Handbook*. 3rd ed. Publication No. FHWA-HRT-06-108. McLean, VA: Federal Highway Administration.

Garber, N., and L. Hoel. 2002. *Traffic and Highway Engineering*. 3rd ed. Pacific Grove, CA: Brooks/Cole.

May, A. D. 1990. *Traffic Flow Fundamentals*. Englewood Cliffs, NJ: Prentice Hall.

McShane, W. R., and R. P. Roess. 1990. *Traffic Engineering*. Englewood Cliffs, NJ: Prentice Hall.

Rakha, H., and W. Zhang. 2005. "Estimating Traffic Stream Space Mean Speed and Reliability from Dual- and Single-loop Detectors." *Transportation Research Record* 1925: 38–47. doi:10.3141/1925-05.

Shankar, V., and F. Mannering. 1998. "Modeling the Endogeneity of Lane-mean Speeds and Lane-speed Deviations: A Structural Equations Approach." *Transportation Research Part A* 32 (5): 311–322.

Shinar, D. 1999. "Aggressive Driving: The Contribution of the Drivers and the Situation." *Transportation Research Part F* 1 (2): 137–160. doi:10.1016/S1369-8478(99)00002-9.

Transportation Research Board. 2000. *Highway Capacity Manual*. Washington, DC: Transportation Research Board of the National Academies of Science.

Treiber, M., A. Hennecke, and D. Helbing. 1999. "Derivation, Properties, and Simulation of a Gas-kinetic-based, Nonlocal Traffic Model." *Physical Review E* 59: 239–253. doi:10.1103/PhysRevE.59.239.

Varaiya, P. 2007. "Effectiveness of California's High Occupancy Vehicle (HOV) System." UCB-ITS-PRR-2007-5, California PATH Research Report. Institute of Transportation Studies, University of California, Berkeley.

Wang, H., W. Wang, X. Chen, J. Chen, and J. Li. 2007. "Experimental Features and Characteristics of Speed Dispersion in Urban Freeway Traffic." *Transportation Research Record* 1999 (1): 150–160. doi:10.3141/1999-16.

Wardrop, J. G. 1952. "Some Theoretical Aspects of Road Traffic Research." *Proceedings of the Institute of Civil Engineers* 1–2: 325–378.