

A COMPOSITIONAL APPROACH FOR TRAFFIC DISTRIBUTION EVALUATION OF TRIPLE LEFT-TURN LANES FROM AN INDIVIDUAL PERSPECTIVE

Li LI^{1,2}, Dong ZHANG^{3*}, Xiao-Yun CHENG⁴, Ping WANG⁵, Gui-Ping WANG⁶

^{1,5,6}*School of Electronics and Control Engineering, Chang'an University, China*

²*National Engineering Laboratory for Surface Transportation Weather Impacts Prevention,
Broadvision Engineering Consultants, Kunming, Yunnan, China*

³*School of Transportation and Logistics, Dalian University of Technology, China*

⁴*School of Transportation Engineering, Chang'an University, China*

Submitted 5 August 2018; resubmitted 20 December 2018, 5 February 2019; accepted 13 April 2019

Abstract. This study analysed unbalanced traffic distribution on Triple Left-Turn Lanes (TLTLs) at signalized intersections that is caused by left-turn drivers' unequal lane preferences. To develop statistical bonding between the multilane traffic flow and individual lane choices, the lane volumes are formatted as compositional data to subject the sum-constant constraint. One-way and two-way Compositional ANalysis Of VAriance (CANOVA) models were formulated respectively to estimate the independent effect of one factor and its joint effects with other factors on the multilane traffic distribution. TLTL volume composition was the dependent variable of the models, while the factors of geometric design and traffic control that could affect left-turn drivers' lane choice were the independent variables. Results indicate that variance of vehicle turning curve, length of the upstream segment, the location of triple left-turn sign, signal phase / cycle length, could affect the traffic distribution, and its balance could be achieved at specific levels of a factor. The joint effects of some factor couples could improve the unbalanced traffic distribution while others could not work.

Keywords: traffic distribution, driver's lane choice, geometric design, traffic control, CANOVA.

Notations

ANOVA – analysis of variance;
CANOVA – compositional ANOVA;
CT – central tendency;
ILR – isometric log-ratio transformation;
LTL – left-turn lane;
MSD – metric standard deviation;
TLTL – triple left-turn lane.

Introduction

Turning vehicles merge as one traffic movement or diverge to different destinations at signalized intersections. Due to the limited space of the intersection approach, the number of traveling lanes on each approach is normally constant. The lanes are assigned to through, right-turn and left-turn traffic according to their demands. The traffic with the heavier demand is allocated to more lanes. The capacity of these multiple lanes serving the same traffic is directly related to the saturation flow rate of each lane, so most

studies of multilane operation concerned about how to maximize the rate (Bagheri *et al.* 2015; Chen *et al.* 2011; Cooner *et al.* 2011; Yi *et al.* 2012; Zhao *et al.* 2017). However, compared with the variation of lane saturation flow rates, another reason of multilane capacity reduction, multilane underutilization, has attracted much less attention previously, thus which is the main concern of this study.

Underutilization is a term to depict the unsatisfactory performance of multilane infrastructure under the influence of unbalanced lane traffic distribution. Since lane volume is the aggregated outcome of some drivers choosing a lane, multilane underutilization reflects the condition that a large part of the driver group on multiple lanes prefer to a specific lane, rather than treating the lanes equally. Naturally, in order to improve multilane underutilization, it should induce a part of the drivers who originally have such lane preference switch to other lanes instead, which head to the same destination as their original choices. Unfortunately, this idea has not been applied in the work of improving multilane underutilization. For example, some studies pointed out that multilane satura-

*Corresponding author. E-mail: zhangdong@dlut.edu.cn

tion flow rates could be influenced by travel demand, lane changing rates or road geometrics (Qi *et al.* 2013; Ring, Sadek 2012; Sando, Moses 2009; Yeom *et al.* 2014; Yousif *et al.* 2013; Pompigna, Rupi 2017). These factors have been considered in the work of optimizing lane assignment of multiple lanes together with signal timing to maximize the entire approach capacity or minimize total traveling delay (Wong, Heydecker 2011; Zhao *et al.* 2013; Zhou *et al.* 2016; Alhajyaseen *et al.* 2017). The lane assigning methods can adjust the number of the lanes serving the traffics in different directions, but they did not take the existence of multilane underutilization into account.

To adjust drivers' lane preferences in a multilane infrastructure, critical work is to identify the influential factors of the preferences, which is the prerequisite of improving or preventing multilane underutilization by changing road geometrics or traffic controls. Lane utilization is an index to measure the underutilization degree, which was proposed by *Highway Capacity Manual* (TRB 2000). This index is calculated by dividing the ratio of average lane volume with the highest one in a lane group. Such a calculating method does not reflect the utilization of each lane belonging to the same lane group. Hence, it cannot be used to reveal lane choice pattern of the drivers in multilane context, as the usage of each lane measures the possibility of it being chosen by the drivers. To solve this problem, this study proposed a new way to measure multilane traffic distribution and a statistical method to find out the causes of multilane underutilization as well as unbalanced lane preferences of the drivers travelling on it.

This study takes TLTLs at urban signalized intersections as a case study. The main contributions of this study are summarized as follows. The first and the most important one is that some factors of roadway geometric and traffic control that independently or jointly influence the usages of the three LTLs are identified. Unlike relevant previous studies (Sando, Moses 2009; Cooner *et al.* 2011), the findings of this study are founded on a solid connection between lane usages and drivers' lane preferences. Hence, it is reasonable to claim that a factor influences the drivers' LTL preferences if it does lead to a variation of traffic distribution on the TLTLs. The second contribution is that three TLTL volumes are organized as the compositional data to measure lane traffic distribution on the TLTLs, instead of using the indicator of lane utilization. The constant-sum constraint subjected by the compositional data is theoretical basis of intrinsic connection mentioned in the first contribution. Such constraint ensures the connection constantly effective in the compositional regression analysis that is to identify the influential factors of the drivers' LTL choices. This work is incapable if takes the lane utilization as a dependent variable in the statistical analysis. The third contribution is that a novel method was introduced to estimate the compositional regression model. The method solves the problem that multivariate data, such as the LTL usages, which subjects to the constant-sum constraint cannot be analysed by the traditional multivariate regression model. It makes it pos-

sible to infer drivers' LTL preferences in virtue of the compositional regression analysis.

The rest of this paper is organized as follows. Section 1 describes the methodology used in this study. Section 2 provides the information of data collection work. Section 3 introduces the way of interpreting the analysis results with some examples. Section 4 discusses the results from the viewpoint of traffic distribution as well as individual drivers. Last section concludes the findings of this study and points out some directions for future improvements.

1. Methodology

1.1. Data

Compositional data, or composition, is a vector with several elements. Each element is a compositional component. It represents the portion of the value of a specific vector element with respect to the value of all elements (Aitchison 1982). Three lane volumes of the TLTLs can be formatted into the TLTLs volume composition V_3 as:

$$V_3 = [VIL, VML, VOL], \quad (1)$$

where: V_3 contains three intrinsic bonds of lane volumes and drivers' lane choices at the TLTLs. The first bond is that the number of individual LTLs is the size of the lane choice set for the left-turn drivers, which equals three. The number also refers to the dimension of the volume vector of the TLTLs v_3 (Equation 2), which is the original form of V_3 :

$$v_3 = [vil, vml, vol], \quad (2)$$

where: vil , vml , vol are the volumes of the inside, median or outside LTL, respectively. Each vector element records the number of left-turn drivers appearing on the inside, median or outside LTL during volume counting period. The second bond is that the proportion of vil , vml or vol with respect to the total volume of v_3 measures the probability of each LTL being chosen by the drivers. Such a probability can be extracted from v_3 by normalizing it (Equation 3), which is named as closure operation in the compositional statistics (Aitchison 1982):

$$C(v_3) = k \cdot [vil', vml', vol'] = k \cdot v'_3, \quad (3)$$

where: v'_3 is the ratio-scaled TLTLs volume composition; k is the total volume of the TLTLs; vil' , vml' and vol' are the closed LTL volume components. The last but not the least bond is that vil' , vml' and vol' subject to the constant-sum constraint (Equation 4), which reflects the feature of either-or lane choice of each left-turn driver:

$$vil' + vml' + vol' = 1. \quad (4)$$

Note that the ratio-scaled v'_3 is reduced from original volume count data. In term of statistical analysis, the representativeness of vil' , vml' and vol' for the drivers' LTL preferences is guaranteed only if total vehicle counts in the LTL volume are large enough. If so, the variation of total vehicle counts can be assumed to be irrelevant to the mean

variation of vil' , vml' and vol' . Such assumption stands when total counts at least have the scale of hundreds (Van den Boogaart, Tolosana-Delgado 2013), but such scale is hardly achieved for the left-turn vehicles arriving at an intersection in a signal phase or cycle period. Therefore, sampling error of left-turn volume data should be considered in the statistical analysis of v'_3 . In this sense, v'_3 is formatted as ratio-scaled count compositional data V_3 to characterize the uncertainty of the drivers' LTL choices. The format of V_3 is shown in Equation (1), in which VIL , VML and VOL are count compositional components. The components are subjected to the constant-sum constraint as v'_3 does:

$$VIL + VML + VOL = 1. \quad (5)$$

Based on abovementioned three bonds, it is reasonable to take V_3 as the dependent variable of the compositional statistical analysis to infer the influences of external factors on drivers' LTL preferences from the variation of the LTL volumes. This study develops two compositional regression models to achieve this target. The models will be introduced in the next section.

1.2. Model

The constant-sum constraint subjected by V_3 is the most intrinsic bond of lane volumes of the TLTs and drivers' LTL choices, so it is important to hold it in statistical analysis. However, the constraint does not exist if operates V_3 as a real vector. The statistical models defined in sample space for the real vector is not suitable for the compositional statistical analysis holding the constraint. To solve this problem, Aitchison (1982) defined a new sample space, namely the 3-part simplex \mathbb{S}^3 to organize the compositional data and conduct relevant analysis. For the TLTs volume composition, \mathbb{S}^3 presents as an equilateral triangle (VIL , VML , VOL) in Figure 1. All possible values of the TLTs volume composition falls into the triangle.

Aitchison (1982) also defined a set of operations to operate the compositional data in the simplex. The operations function similarly with the ones that are applicable for real vector. For example, the function of perturbation (\oplus) or powering (\odot) is equivalent to the one of vector summation or multiplication in real space.

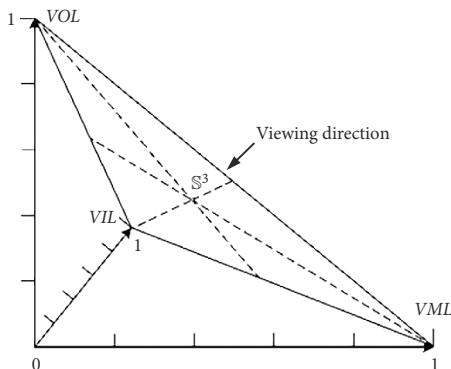


Figure 1. 3-part simplex \mathbb{S}^3

A statistical model can be defined in \mathbb{S}^3 with the Aitchison (1982) operations to identify the influential factors for traffic distribution on the TLTs. The factors considered in this study cover the aspects of roadway geometrics and traffic signal control. Since the factor values are constant in each study site, so they are taken as categorical independent variables of the compositional regression model. Their average influences on drivers' LTL choices can be inferred from the variation of the mean value of V_3 , \bar{V}_3 in virtue of a one-way CANOVA model:

$$\bar{V}_3 = \alpha \oplus \beta_i \oplus \varepsilon, \quad (6)$$

where: $\bar{V}_3 = [\bar{VIL}, \bar{VML}, \bar{VOL}]'$; $\alpha = [\alpha_{in}, \alpha_{med}, \alpha_{out}]'$ is the compositional constant; $\beta_i = [\beta_{i-in}, \beta_{i-med}, \beta_{i-out}]'$ is the compositional coefficient at the i th level of β , and its components correspond to the ones of \bar{V}_3 one by one; $\varepsilon_3 = [\varepsilon_{in}, \varepsilon_{med}, \varepsilon_{out}]'$ is assumed to subject to normal distribution in \mathbb{S}^3 with null compositional expectation and a constant variance.

β_i subjects to a compositional constant set $\{\beta_1, \dots, \beta_o, \dots, \beta_p\}$, while p is level number of β . At the o th level of β , \bar{V}_3 has a conditional expected value, μ_o :

$$\mu_o = E[\bar{V}_3 | \beta_i = \beta_o] = \alpha \oplus \beta_o. \quad (7)$$

However, μ_o is not unique, as μ_o can always be obtained by replacing α and β_o by another compositional constant α^* and $\beta_o^* = \beta_o \odot (\alpha^* \ominus \alpha)$ respectively:

$$a^* \oplus \beta_o^* = a^* \oplus \beta_o \odot (\alpha^* \ominus \alpha) = a \oplus \beta_o. \quad (8)$$

If so, the estimated coefficients are non-identifiable. To avoid this condition, the CANOVA model needs to be reformulated into an equivalent version with identifiable coefficients. The first level of β can be assumed to have no effect on \bar{V}_3 . When β upgrades from its first level to a higher level, the variance of \bar{V}_3 can be interpreted as its average response to the level change of β if the effect of β is significant.

Besides the effect of one factor, two factors of roadway geometric and traffic control could be improved together in engineering practice. In this case, their joint effect on \bar{V}_3 should be identified in virtue of a two-way CANOVA model to measure the marginal factor effect based on the independent effect of the two factors:

$$\bar{V}_3 = \alpha \oplus \beta_j \oplus \lambda_k \oplus \beta \cdot \lambda_{jk} \oplus \varepsilon, \quad (9)$$

where: $\lambda_k = [\lambda_{k-in}, \lambda_{k-med}, \lambda_{k-out}]'$ is the compositional regression coefficient of λ at its k th level; $\beta \cdot \lambda_{jk} = [\beta \cdot \lambda_{jk-in}, \beta \cdot \lambda_{jk-med}, \beta \cdot \lambda_{jk-out}]'$ is the coefficient under the joint effect of the β at its j th level and λ at its k th level. In the two-way CANOVA, there is a similar concern of non-identifiable factor coefficient as it appears in one-way CANOVA. Hence, a combination of the base levels of two factors should be specified before joint effect estimation.

The factor couples whose joint effect will be analysed and their base levels are defined in Section 3.2.

A “staying-in-the-simplex” approach is applied to estimate the parameters of the CANOVA models (Pawlowsky-Glahn *et al.* 2015). This approach is found on a theory of “the principle of working in coordinates” (Mateu-Figueras *et al.* 2011), which can project the compositional data V_3 defined in \mathbb{S}^n to the real space by ILR and obtain its Cartesian coordinates with respect to a new orthonormal basis. In the basis, the angle and distance of the compositional data is not changed and the obtained coordinates referring to the basis have released the constant-sum constraint and became real unbounded values. Hence, it can be analysed straightforwardly with the desired traditional statistical method. The way of developing an orthonormal basis can refer to the literature (Pawlowsky-Glahn *et al.* 2015; Van den Boogaart, Tolosana-Delgado 2013). Based on the developed orthogonal basis, the univariate one-way or two-way CANOVA model defined in \mathbb{S}^3 can be translated to the multivariate ANOVA model defined in real space as:

$$\begin{cases} ILR(\bar{V}_3) = ILR(\alpha) + ILR(\beta_j) + ILR(\varepsilon); \\ ILR(\bar{V}_3) = ILR(\alpha) + ILR(\beta_j) + ILR(\lambda_k) + ILR(\beta \cdot \lambda_{jk}) + ILR(\varepsilon), \end{cases} \quad (10)$$

where: $ILR(\varepsilon)$ subjects to the normal distribution. The model estimation is conducted in the orthonormal basis, so the estimated coefficients of influential factors have to be interpreted on the basis. They don't one by one correspond to LTL volume components defined in \mathbb{S}^3 , which makes the interpretation hardly understood. If translate the coefficients back to \mathbb{S}^3 by inverse ILR transformation, each component has its own coefficient referring to a factor, which is much easier to interpret in terms of the variations of LTL volumes and drivers' LTL choices.

2. Data collection

2.1. Case study

Four TLTLs equipped at the signalized intersections in Shanghai (China) were selected as a case study. Figure 2 illustrates their configurations: (1) three unshaded LTLs; (2) one shaded LTL and two unshaded LTLs. A shaded LTL is a lane used to accommodate more left-turn traffic near the stop line by occupying the space of opposite approach. Each lane of the TLTLs is named according to its distance to the road central line. The closed one is the inside LTL, the furthest one is the outside LTL, and the one between them is the median LTL.

The right of way of left-turn traffic was regulated by an exclusive signal phase at the signalized intersections. LTL volumes were counted in green and red phases respectively, and they were summed up as the cycle volume. Correspondingly, three scenarios were set to analyse traffic distribution on the TLTLs: (1) ST1 that refers to red phase; (2) ST2 that refers to green phase; (3) ST3 that refers to entire signal cycle. To reduce sampling error, enough vehi-

cles should be counted into the volume data. So only, the signal phase during which at least eight vehicles arrive at one LTL was considered as eligible study period. This rule is stricter than the one used in the survey of lane saturation flow (TRB 2010). The volumes were collected in 336 red/green phases (or 168 cycles). A vehicle was counted when it arrived at the queue back, as it cannot change to other lanes until the queue discharged again. If no queue existed, it was counted in at the stop line. One heavy vehicle was equivalent to two passenger cars in the volume, which refer to the regulation used in China.

The influential factors of left-turn traffic distribution are listed in Table 1. They cover the aspects of traffic signal timing and roadway geometrics. The measures of some factors are illustrated in Figure 3. The values of all factors

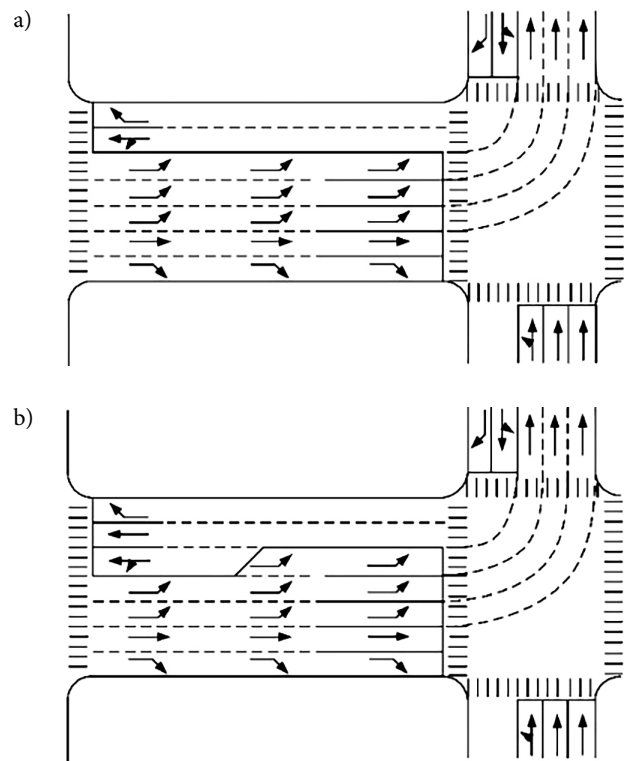


Figure 2. Typical layouts of TLTLs: a – without shaded line; b – with shaded line

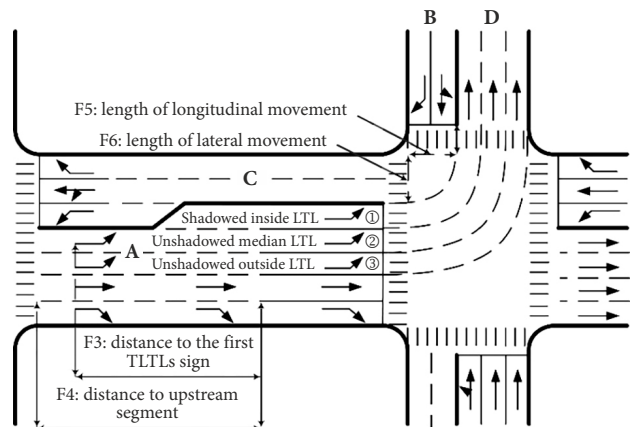


Figure 3. Measures of some factors

Table 1. Factor list

Category	No	Name	Unit	Measure	Priority
Traffic control	F0	Length of signal cycle	s	Counted in per cycle	6
	F1	Length of red phase	s	Counted in red phase	6
	F2	Length of green phase	s	Counted in green phase	6
	F3	Distance to the first TLTLs sign	m	Measured from the end of white solid line to the first TLTLs sign at upstream	6
Geometric design	F4	Length of upstream segment	m	Measured from the end of solid lane separation line to the start of road segment	1
	F5	Length of longitudinal movement	lane	Measured in lane number	2
	F6	Length of lateral movement	lane	Measured in lane number	3
	F7	Number of other traffic lanes	lane	Total number of through and right-turn lanes in the same approach of the TLTLs	4
	F8	Shadowed LTL	binary	0 if not equip the shadowed LTL; 1 if equipped the lane	5

Table 2. Factor measurements

No	Site 1	Site 2	Site 3	Site 4
	SB Hengfeng Road at Tianmu Western Road	EB Pujian Road at Pudong Sourthern Road	EB Yan'an Road at Shimen First Road	WB Zhoujia Zui Road at Inner Ring Line
F0	219 s (2nd)	104 s (1st)	242 s (4th)	230 s (3rd)
F1	183 s (2nd)	76 s (1st)	194 s (3rd)	184 s (2nd)
F2	36s (1st)	28 s (1st)	48 s (2nd)	46 s (2nd)
F3	94 m (1st)	157 m (2nd)	180 m (3rd)	200 m (3rd)
F4	142 m (1st)	172 m (2nd)	195 m (3rd)	300 m (4th)
F5	6 (3rd)	5 (2nd)	3 (1st)	5 (2nd)
F6	4 (2nd)	3 (1st)	4 (2nd)	4 (2nd)
F7	3 (2nd)	2 (1st)	2 (1st)	4 (3rd)
F8	Yes (2nd)	No (1st)	No (1st)	Yes (2nd)

Notes: SB – South Bound; EB – East Bound; WB – West Bound.

are constant at each site, so their impacts work on total frequency left-turn drivers selecting an LTL in the volume counting period. In Equations (6) and (9), the factors are set as ordinal categorical variables except F5 as an unordered binary variable. Table 2 lists the value of each factor at each site, which are assigned with a factor level marked as “1st”, “2nd”, “3rd” or “4th”. The criteria of factor level division are: (1) the interval between two successive levels of F0, F1 or F2 is more than 10 s; (2) the levels of F5, F6 and F7 are divided lane by lane; (3) the gap between two successive levels of F3 or F4 is 20 m.

In addition, each factor is assigned with a priority number. The priorities are ordered according to the sequence of their confirmations in a roadway design process. The length of the upstream segment of the TLTLs F4 is confirmed when planning roadway network, so F4 is assigned with the highest priority. The length of the longitudinal movement of left-turn vehicles F5 is determined by the width of crossing road of the TLTLs. Since F5 is harder to adjust than the width of the opposite approach of the TLTLs F6, so the priority of F5 is higher than the one of F6. The number of through and right traffic lanes within the same approach of the TLTLs F7 could be adjusted if F6 needs to change, so F7 has the lower priority

than F6. The shadowed LTL only accommodates left-turn traffic, so its priority is lower than F4, F5, F6, and F7. The settings of traffic control devices, such like traffic signal – F0, F1 and F2, or lane direction sign F3, is easier to adjust than others, so they are assigned with the lowest priority.

2.2. Descriptive analysis

The TLTLs volumes collected in red phase ST1, green phase ST2 and signal cycle ST3 are organized in the format of V_3 , and their distributions in \mathbb{S}^3 are shown in Figure 4. The crossing point of the dotted lines in each graph refers to the balanced status of the TLTLs volume composition, i.e., [0.33, 0.33, 0.33]. If one sample of the composition is closer to the point VIL with respect to the points VML or VIOL, the value of VIL is larger than VML or VOL. In other words, in the subject period, more drivers choose the inside LTL than the median LTL or the outside LTL. It results in unbalanced left-turn traffic distribution on the TLTLs.

Figure 4 shows that the left-turn traffic distribution is gradually out of balance from red phase ST1 to signal cycle ST3 then to green phase ST2. This trend is also observed from the descriptive statistics in Table 3.

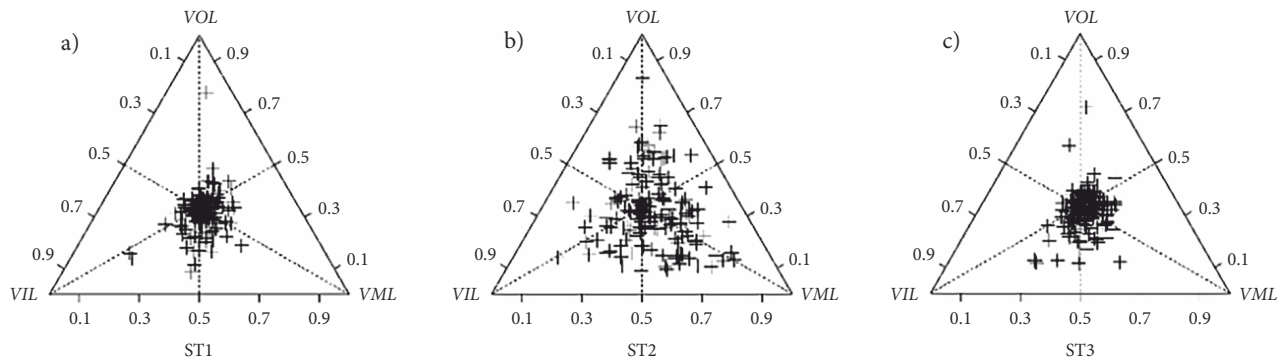


Figure 4. Sample distributions of the TLTLs volumes in: a – red phase; b – green phase; c – signal cycle

Table 3. Descriptive statistics

Scenario	CT			MSD
	VIL	VML	VOL	
ST1	0.331	0.356	0.313	0.255
ST2	0.306	0.377	0.317	0.487
ST3	0.326	0.362	0.312	0.249

CT measures the mean value of *VIL*, *VML* or *VOL* among all samples, while MSD measures the average spread of *VIL*, *VML* or *VOL* to CT. The MSD in ST2 (0.487) is larger than the one in ST1 (0.255) or ST3 (0.249). This indicates that the compositions spread wider in the green phase than in the red phase, while it spreads the least in cycle period. In addition, the CT of *VIL* in ST1/ST3 (0.331/0.326) is smaller than the one of *VOL* (0.313/0.312). It means that the inside LTL attracts more drivers than the outside LTL in red phase or cycle period, but it becomes less attractive in green phase period reversely. *VML* equals 0.356 in ST1, 0.377 in ST2 and 0.362 in ST3, so the median LTL always attracts the largest portion of left-turn traffic. Finally, sample distribution in ST1 looks similar to one in ST3. It could be owing to the much shorter length of green phase than the red phase assigned to left-turn traffic. Since the amount of green phase volume is less than the one of red phase volume, cycle volumes, as their summation, could be similar to red phase volume in terms of volume amount and composition.

3. The way of results interpretation

3.1. One factor effect

The R package “compositions” (Van den Boogaart, Tolosana-Delgado 2013) was applied to develop and estimate the CANOVA models. Null hypothesis H_0 of the one-way CANOVA model in Equation (6) is that V_3 has the same expectation at different levels of a factor. The hypothesis is tested at 90, 95 or 99% confidence level, and the test results are listed in Table 4. The estimated values of V_3 under the factor effect at its different levels are reported in Table 5 in the compositional way. Each factor is assumed to have no effect on V_3 at its first level, and the variance of V_3 is caused by the upgrade of the factor level. The values of *VIL*, *VML* and *VOL* after the level change are

obtained by adding their values at the first level of a factor to the coefficient obtained after the factor level upgrading. By comparing the value of *VIL*, *VML* or *VOL*, with its balanced status, i.e., 0.333, it can find if the change of the factor level could improve or worsen the balance of the TLTLs traffic distribution. The state of V_3 that is most close to the balanced status is underlined in Table 5.

To explain how to interpret the results of one-way CANOVA, it takes the results of F0 in ST3, i.e., the length of the signal cycle, as an example. F0 has four levels in this study (Table 2). The regression coefficients of F0 and the values of V_3 under its influence are demonstrated in Figure 5. L1...Ln indicates the ascending order that is used to report the values of V_3 from the first level of F0 to its highest level. The sign “↗” or “↘” illustrates the varying trend of *VIL*, *VML* and *VOL* between two levels of F0. The values of *VIL*, *VML* and *VOL* or the regression coefficients of F0 satisfies the sum-constant constraint; in other words, their summations equals 1. The regression intercept of *VIL* at the first level of F0 is 0.398 (see “*VIL*” column, “Coefficient” row, Table 6), which also refers to the value of *VIL* without the influence of F0 (see “*VIL*” column, “Component” row, Table 6). This value is larger

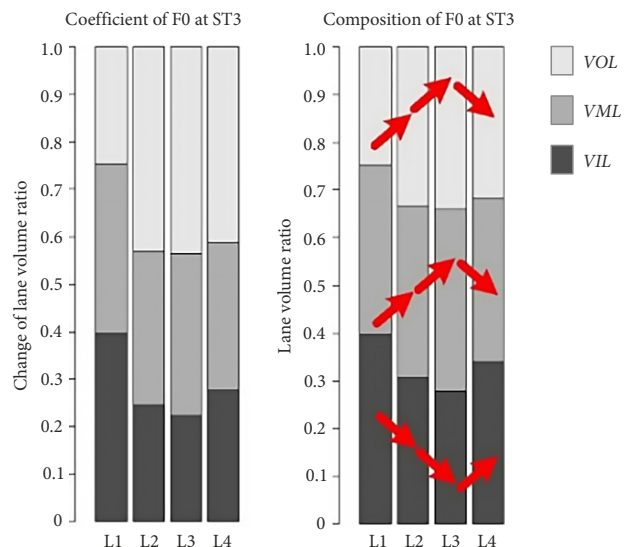


Figure 5. The graph of compositional regression coefficients and the values of V_3 under different level effects of F0

Table 4. Significant test of one-factor effect on \overline{V}_3

Scenario	Statistics	Factor								
		F0	F1	F2	F3	F4	F5	F6	F7	F8
ST1	Degree of freedom	--	2	--	2	3	2	1	2	1
	F-value	--	19.483	--	15.865	13.464	4.361	39.043	12.262	26.864
	Significance	--	<0.001	--	<0.001	<0.001	<0.01	<0.001	<0.001	<0.001
	Decline H_0	--	yes	--	yes	yes	yes	yes	yes	yes
ST2	Degree of freedom	--	--	1	2	3	2	1	2	1
	F-value	--	--	2.232	1.310	4.193	2.778	0.891	6.187	4.992
	Significance	--	--	>0.1	>0.1	<0.001	<0.001	>0.1	<0.001	<0.01
	Decline H_0	--	--	no	no	yes	yes	no	yes	yes
ST3	Degree of freedom	3	--	--	2	3	2	1	2	1
	F-value	17.737	--	--	16.487	17.737	4.015	41.012	19.940	42.613
	Significance	<0.001	--	--	<0.001	<0.001	<0.01	<0.001	<0.001	<0.001
	Decline H_0	yes	--	--	yes	yes	yes	yes	yes	yes

Table 5. The value of \overline{V}_3 under one factor effect

Factor	\overline{VIL} (L1...Ln)	\overline{VML} (L1...Ln)	\overline{VOL} (L1...Ln)
Scenario ST1			
F1 (3 levels)	0.409 / 0.302 / <u>0.334</u>	0.346 / 0.363 / <u>0.338</u>	0.245 / 0.335 / <u>0.327</u>
F3 (3 levels)	0.409 / 0.308 / <u>0.313</u>	0.346/ 0.357/ <u>0.356</u>	0.245 / 0.335 / <u>0.331</u>
F4 (4 levels)	0.308/ 0.409/ <u>0.335</u> / 0.334	0.357 / 0.346 / <u>0.338</u> / 0.371	0.335 / 0.245 / <u>0.327</u> / 0.295
F5 (3 levels)	<u>0.334</u> / 0.291 / 0.308	<u>0.338</u> / 0.362 / 0.357	<u>0.327</u> / 0.347 / 0.335
F6 (2 levels)	0.409 / <u>0.311</u>	0.346 / <u>0.356</u>	0.245 / <u>0.333</u>
F7 (3 levels)	0.371 / <u>0.308</u> / 0.295	0.344 / <u>0.357</u> / 0.371	0.285 / <u>0.335</u> / 0.334
F8 (2 levels)	0.371 / <u>0.302</u>	0.344 / <u>0.363</u>	0.285 / <u>0.335</u>
Scenario ST2			
F1 (3 levels)	--	--	--
F3 (3 levels)	--	--	--
F4 (4 levels)	<u>0.317</u> / 0.329 / 0.349 / 0.237	<u>0.357</u> / 0.372 / 0.359 / 0.423	<u>0.326</u> / 0.299 / 0.292 / 0.341
F5 (3 levels)	0.349 / 0.278 / <u>0.317</u>	0.355 / 0.400 / <u>0.357</u>	0.296 / 0.322 / <u>0.326</u>
F6 (2 levels)	--	--	--
F7 (3 levels)	0.339 / <u>0.317</u> / 0.237	0.366 / <u>0.357</u> / 0.423	0.295 / <u>0.326</u> / 0.340
F8 (2 levels)	<u>0.339</u> / 0.282	<u>0.356</u> / 0.385	<u>0.305</u> / 0.333
Scenario ST3			
F1 (3 levels)	0.398 / 0.306 / 0.279 / <u>0.341</u>	0.354 / 0.360 / 0.382 / <u>0.343</u>	0.248 / 0.334 / 0.339 / <u>0.316</u>
F3 (3 levels)	0.398 / 0.306 / <u>0.311</u>	0.354 / 0.360 / <u>0.360</u>	0.248 / 0.334 / <u>0.329</u>
F4 (4 levels)	0.306 / 0.398 / <u>0.341</u> / 0.279	0.360 / 0.354 / <u>0.343</u> / 0.382	0.334 / 0.248 / <u>0.317</u> / 0.340
F5 (3 levels)	<u>0.331</u> / 0.333 / 0.306	<u>0.343</u> / 0.372 / 0.360	<u>0.307</u> / 0.295 / 0.334
F6 (2 levels)	0.398 / <u>0.307</u>	0.354 / <u>0.362</u>	0.248 / <u>0.331</u>
F7 (3 levels)	0.369 / <u>0.306</u> / 0.279	0.350 / <u>0.360</u> / 0.382	0.281 / <u>0.334</u> / 0.339
F8 (2 levels)	0.369 / <u>0.295</u>	0.350 / <u>0.369</u>	0.281 / <u>0.336</u>

Notes: "--" – not applicable; underline – indicates the lane distribution that achieves relative balanced status.

Table 6. The regression coefficients and the values of \overline{V}_3 under different level effects of F0

Scenario	Factor	ST3		
		Estimate	\overline{VIL} (L1...Ln)	\overline{VML} (L1...Ln)
Signal cycle length F0 (4 levels)	Coefficient	0.398 / 0.245 / 0.222 / 0.276	0.354 / 0.324 / 0.343 / 0.312	0.248 / 0.430 / 0.435 / 0.412
	Component	0.398 / 0.306 / 0.279 / <u>0.341</u>	0.354 / 0.360 / 0.382 / <u>0.343</u>	0.248 / 0.334 / 0.339 / <u>0.316</u>

Notes: underline – indicates the lane distribution that achieves relative balanced status.

than the value of \overline{VML} or \overline{VOL} , i.e., 0.354 or 0.248 (see “ \overline{VML} ” or “ \overline{VOL} ” column, “Component” row, Table 6). It means that more drivers choose the inside LTL instead of the median or outside lane when the value of F0 is at its first level. If F0 upgrades to the second, the third or the fourth level, the coefficient of \overline{VIL} equals 0.245, 0.222 or 0.276, each of which is smaller than 0.333. The results mean that continuously upgrading F0’s level could keep reducing the value of \overline{VIL} with respect to the one at the first level (see $0.398 > 0.306 / 0.279 / 0.341$ in “ \overline{VIL} ” column, “Component” row, Table 6). In other words, the probability of the inside LTL being chosen by the left-turn drivers could decline with the increased length of signal cycle. Such meanings are similar with the case that a negative coefficient of an independent variable is obtained in a normal linear regression analysis. Due to the sum-constant constraint, the loss of \overline{VIL} is added up to the values of \overline{VML} and \overline{VOL} . At the individual level, it indicates that the driver is more likely to choose the median or outside LTL if the signal cycle prolongs.

3.2. Joint factor effect

The joint effect of two factors on $\overline{V_3}$ measures the marginal effect of a factor on the basis of the independent effect of another factor. The value of $\overline{V_3}$ under the joint effect are estimated from the selected base levels of a factor couple when their values change from the base levels to other ones simultaneously. Table 7 lists the factor couples considered in this study. A couple is selected if a factor has the higher priority than another one. The factor priorities are listed in Table 1. The factor with the high priority and the one with the lower priority respectively refer to “major factor” and “joint factor” here. Note that some factor couples are not taken into account. Since the drivers on the TLTLs (“A” in Figure 2) hardly observe the crossing road, especially in heavy traffic, so the joint effect of F5 with F3 or F4 are not analysed. In addition, the traffic distribution and the significant influential factors in ST1 and ST3 are similar (Table 4), so the two-way CANOVA are not made for the $\overline{V_3}$ in ST3.

To compare with the results of one-way and two-way CANOVA, the base levels of the factor couple applied in the two-way CANOVA contain the first level of the major factor. Such a setting can help us to investigate if a level change of the joint factor could rebalance the uneven traffic distribution on the TLTLs caused by the upgraded level of the major factor. The null hypothesis of the two-way CANOVA is that $\overline{V_3}$ has the same expectation at different level combinations of a factor couple, which is tested at 90, 95 or 99% confidence level. Test results at the different combinations are listed in Table 8, and the values of the estimated $\overline{V_3}$ are reported in a compositional way. In the table, the base level combination of a factor couple is marked as “1”, such like “F1(2)×F4(1)” in “ $\overline{V_3}$ in ST1” column, “F1×F4” row. The joint effects of

the factor couple on $\overline{V_3}$ are reported in other columns marked by “1→2/3/4” if their level combination changes from the base one to other one. The effect is marked by “ $\overline{\surd}$ ” if the upgraded factor levels improves the unbalanced $\overline{V_3}$ caused by the major factor; otherwise, it is marked by “ $\overline{\times}$ ”. This judgement can be made according to the following rule:

$$\overline{V_3} : D_1 = \sum_{i=1}^3 \left| V_i^{major} - 0.333 \right| \text{ vs. } D_2 = \sum_{i=1}^3 \left| V_i^{joint} - 0.333 \right|, \tag{11}$$

where: V_i^{major} refers to the values of \overline{VIL} , \overline{VML} or \overline{VOL} under the independent effect of the major factor at its i th level; V_i^{joint} refers to the ones under the joint factor effect of the level combination, which includes the i th level of the major factor; D_1 and D_2 refer to summations of the differences of each component of V_i^{major} and V_i^{joint} with respect to its balanced status, i.e., 0.333, respectively. The estimated effect is marked as “#” if the combination contains the level of the major factor that it independently causes relative balanced $\overline{V_3}$ (underlined in Table 5).

Here takes the value of $\overline{V_3}$ under the joint effects of F1 and F4 in ST1 for example (***bolded and italicized*** in Table 8). The combination of the second level of F1 and the first level of F4, indicated as “F1(2)×F4(1)” and numbered as “(1)” in the table, are selected as the base factor levels. Since the joint effect of two factors can be estimated only if their levels change at the same time, their joint effect on $\overline{V_3}$ cannot be estimated at the second level of F1 and the fourth level of F4, which is marked as “-” under the “F1(2)×F4(4)” cell of the table. Significant joint effect on $\overline{V_3}$ can be found when the factor level combination changes from the base one to the second one, which is marked as “1→2”. In this case, F1 degrades from the second level to the first one “F1(2)→F1(1)” meanwhile F4 upgrades from the first level to the second one “F4(1)→F4(2)”. Under such effect, $\overline{V_3}$ becomes [0.407, 0.328, 0.265]. This value is closer to the balanced status of $\overline{V_3}$ than the one obtained under the independent effect of F4 at its second level in ST1 (see $\overline{V_3}$: [0.409, 0.346, 0.245] in ST1 column, F4 row of Table 4).

Table 7. Selected factor couples

Major factor	Joint factor								
	F0	F4	F5	F6	F7	F8	F1	F2	F3
F4	--	--	×	√	√	√	√	√	√
F5	--	--	--	√	×	√	√	√	×
F6	--	--	--	--	√	√	√	√	√
F7	--	--	--	--	--	√	√	√	√
F8	--	--	--	--	--	--	√	√	√

Notes: “--” – not applicable; “×” – the joint effect that is not taken into account; “√” – the joint effect that is taken into account.

Table 8. The value of \bar{V}_3 under joint factor effect

Major factor	Joint factor	$\bar{V}_3 : [\overline{VIL}, \overline{VML}, \overline{VOL}]$ in ST1			
F4	F1×F4(*)	1. F1(2)×F4(1)	2. F1(1)×F4(2) (1→2)	3. F1(3)×F4(3) (1→3)	4. F1(2)×F4(4) (1→4)
		--	[0.407, 0.328, 0.265] ☑	[0.395, 0.304, 0.301] #☒	--
	F3×F4(*)	1. F3(1)×F4(1)	2. F3(2)×F4(2) (1→2)	3. F3(3)×F4(3) (1→3)	4. F3(3)×F4(4) (1→4)
		--	[0.381, 0.332, 0.287] ☑	[0.375, 0.301, 0.324] #☒	[0.301, 0.345, 0.354] ☑
	F6×F4(*)	1. F6(2)×F4(1)	2. F6(1)×F4(2) (1→2)	3. F6(2)×F4(3) (1→3)	4. F6(2)×F4(4) (1→4)
		--	[0.397, 0.366, 0.237] ☑	--#	--
F5	F1×F5(*)	1. F1(3)×F5(1)	2. F1(1)×F5(2) (1→2)	3. F1(2)×F5(2) (1→3)	4. F1(2)×F5(3) (1→4)
		--#	[0.374, 0.345, 0.281] ☑	[0.319, 0.354, 0.327] ☑	[0.312, 0.350, 0.338] ☑
	F6×F5(*)	1. F6(2)×F5(1)	2. F6(1)×F5(2) (1→2)	3. F6(2)×F5(2) (1→3)	4. F6(2)×F5(3) (1→4)
		--#	[0.380, 0.340, 0.280] ☑	--	--
	F8×F5(*)	1. F8(1)×F5(1)	2. F8(1)×F5(2) (1→2)	3. F8(2)×F5(2) (1→3)	4. F8(2)×F5(3) (1→4)
		--#	--	[0.310, 0.352, 0.338] ☑	[0.310, 0.340, 0.350] ☑
F6	F1×F6(*)	--	1. F1(1)×F6(1)	2. F1(2)×F6(2) (1→2)	3. F1(3)×F6(2) (1→3)
		--	--	[0.385, 0.356, 0.259] #☒	[0.377, 0.343, 0.280] #☒
	F3×F6(*)	--	1. F3(2)×F6(1)	2. F3(1)×F6(2) (1→2)	3. F3(3)×F6(2) (1→3)
		--	--	[0.396, 0.343, 0.261] #☒	[0.383, 0.336, 0.281] #☒
F7	F1×F7(*)	1. F1(1)×F7(1)	2. F1(3)×F7(1) (1→2)	3. F1(2)×F7(2) (1→3)	4. F1(2)×F7(3) (1→4)
		--	--	[0.323, 0.337, 0.340] #☑	[0.294, 0.343, 0.363] ☑
	F3×F7(*)	1. F3(2)×F7(1)	2. F3(3)×F7(1) (1→2)	3. F3(1)×F7(2) (1→3)	4. F3(3)×F7(3) (1→4)
		--	--	[0.298, 0.364, 0.338] #☒	[0.286, 0.342, 0.372] ☒
F8	F1×F8(*)	--	1. F1(1)×F8(1)	2. F1(3)×F8(1) (1→2)	3. F1(2)×F8(2) (1→3)
		--	--	--	[0.289, 0.359, 0.352] #☒
	F3×F8(*)	1. F3(2)×F8(1)	2. F3(3)×F8(1) (1→2)	3. F3(1)×F8(2) (1→3)	4. F3(3)×F8(2) (1→4)
		--	--	[0.319, 0.345, 0.336] #☑	[0.285, 0.389, 0.326] #☒
Major factor	Joint factor	$\bar{V}_3 : [\overline{VIL}, \overline{VML}, \overline{VOL}]$ in ST2			
F4	F2×F4(*)	1. F2(1)×F4(1)	2. F2(1)×F4(2) (1→2)	3. F2(2)×F4(3) (1→3)	4. F2(2)×F4(4) (1→4)
		--#	--	[0.350, 0.342, 0.308] ☑	[0.273, 0.384, 0.343] ☑
	F3×F4(*)	1. F3(1)×F4(1)	2. F3(2)×F4(2) (1→2)	3. F3(3)×F4(3) (1→3)	4. F3(3)×F4(4) (1→4)
		--#	[0.336, 0.351, 0.313] ☑	[0.350, 0.342, 0.308] ☑	[0.273, 0.384, 0.343] ☑
F5	F7×F4(*)	1. F7(2)×F4(1)	2. F7(1)×F4(2) (1→2)	3. F7(1)×F4(3) (1→3)	4. F7(3)×F4(4) (1→4)
		--#	[0.341, 0.343, 0.316] ☑	[0.352, 0.342, 0.306] ☑	[0.268, 0.388, 0.344] ☑
	F8×F4(*)	1. F8(2)×F4(1)	2. F8(1)×F4(2) (1→2)	3. F8(1)×F4(3) (1→3)	4. F8(2)×F4(4) (1→4)
		--#	[0.351, 0.342, 0.307] ☑	[0.362, 0.335, 0.303] ☑	--
F7	F6×F5(*)	1. F6(2)×F5(1)	2. F6(1)×F5(2) (1→2)	3. F6(2)×F5(2) (1→3)	4. F6(2)×F5(3) (1→4)
		--	[0.309, 0.358, 0.333] ☑	--	--#
	F8×F5(*)	1. F8(1)×F5(1)	2. F8(1)×F5(2) (1→2)	3. F8(2)×F5(2) (1→3)	4. F8(2)×F5(3) (1→4)
		--	--	[0.279, 0.367, 0.354] ☑	[0.310, 0.340, 0.350] #☒
F8	F2×F7(*)	1. F2(1)×F7(1)	2. F2(1)×F7(2) (1→2)	3. F2(2)×F7(1) (1→3)	4. F2(2)×F7(3) (1→4)
		--	--#	--	[0.268, 0.378, 0.354] ☑
	F3×F7(*)	1. F3(2)×F7(1)	2. F3(3)×F7(1) (1→2)	3. F3(1)×F7(2) (1→3)	4. F3(3)×F7(3) (1→4)
		--	--	[0.327, 0.334, 0.339] #☑	[0.294, 0.358, 0.348] ☒
F8	F8×F7(*)	--	1. F8(1)×F7(1)	2. F8(2)×F7(2) (1→2)	3. F8(2)×F7(3) (1→3)
		--	--	[0.327, 0.337, 0.336] ☑#	[0.287, 0.357, 0.356] ☑
	F2×F8(*)	1. F2(1)×F8(1)	2. F2(2)×F8(1) (1→2)	3. F2(1)×F8(2) (1→3)	4. F2(2)×F8(2) (1→4)
		--	--	--	[0.282, 0.367, 0.351] ☒#
F8	F3×F8(**)	1. F3(2)×F8(1)	2. F3(3)×F8(1) (1→2)	3. F3(1)×F8(2) (1→3)	4. F3(3)×F8(2) (1→4)
		--	--	[0.320, 0.318, 0.362] ☑#	[0.241, 0.379, 0.380] ☒#

Notes: * $p < 0.01$; ** $p < 0.1$; "--" - not applicable; # is marked if relative balanced traffic distribution achieves at the major factor level; ☑ is marked if the upgraded factor level improves the unbalanced traffic distribution caused by the major factor; ☒ is marked if the upgraded factor level aggregates the unbalanced traffic distribution caused by the major factor.

Hence, this beneficial joint factor effect is marked as “ \boxtimes ” underneath the “F1(1)×F4(2)” cell. Moreover, V_3 achieves relative balance in ST1 when F4 independently upgrades from the first level to the third one (see V_3 : [0.335, 0.338, 0.327] in ST1 column, F4 row of Table 5). The performance of V_3 becomes worse under the joint effect of F4 and F1 (see “[0.395, 0.304, 0.301] \boxtimes ” in Table 8) even if F4 makes the same change as it independent works (see “F1(3)×F4(3) (1→4)” in Table 8). The other joint factor effects can be interpreted in the way described above.

4. Discussion

4.1. Independent effect of traffic control factors

From the estimated independent effect of signal cycle length F0 on traffic distribution on the TLTLs, we find that the increase of cycle length could decrease the average volume ratio of the inside LTL with respect to the whole left-turn traffic, while raising the one of the medians or the outside LTL. From the individual view, the results indicate some drivers that originally chose the inside lane could switch to the median lane if signal cycle length becomes longer, while the congestion appeared in the median lane could induce a few drivers on that lane further change to the outside lane. This trend gradually leads to balanced traffic distribution on the TLTLs. Such a result is easy to understand. The longer signal cycle usually happens along with the longer green phase, so the more left-turn vehicles have the opportunity to cross the intersection. It can relieve the unbalanced traffic distribution caused by the unbalance lane traffics in red phase (Section 4.2). For this reason, prolonging the length of the signal cycle is beneficial for full utilization of the TLTLs capacity. A similar finding was also found by Kikuchi *et al.* (2004). They developed a procedure for determining the length of the double LTLs based on the relationship between lane usage and the left-turn volume. Their survey of left-turn traffic revealed that when the total left-turn volume becomes large, the drivers become concerned about the possibility of not being able to clear the intersection in one cycle. Thus, each driver chooses the lane with the shortest queue length, and the usage of two TLTLs become nearly equal.

The influence of traffic signal switch on left-turn traffic distribution has been considered in the design of the study scenario. From the results shown in Table 4, we can find that fewer factors perform significant effects in the green phase ST2 than in the red phase ST1. It implies that less factors could influence lane choices of left-turn drivers in the green phase. In the red phase, an increase of the phase length F1 is found to help the traffic distribution on the TLTLs maintain balance. It means that more drivers arrive at the intersection in the red phase, more possible three LTLs are selected in equal opportunities. This result can be explained from the same perspective as the independent effect of signal cycle length is done.

The traffic sign or pavement marking that inform drivers each lane function at downstream. The earlier they

receive this message, the more time is left for their lane choices and head to desired destinations. From the estimated independent effect of the distance of the TLTLs to its first sign at upstream F3, we find that the unbalanced traffic distribution on the TLTLs could be improved in the red phase or entire cycle time by setting the sign or marking close to the upstream intersection. Such behaviour pattern of the left-turn drivers can be used to adjust the drivers' unbalanced lane preferences caused by other factors.

4.2. Independent and joint effect of road geometric factors

4.2.1. Length of upstream segment F4

Signalized intersections divide urban arterials into several segments. The vehicles departing from the upstream intersection could maintain platoon form when they arrive at its downstream intersection. However, if the upstream segment is long, the platooned vehicles could disperse, and the left-turn drivers can have enough time and spacing to make lane changing (Wei *et al.* 2016). We find that the variant of upstream segment length F4 can influence traffic distribution on the TLTLs at all study scenarios, but the influencing patterns are different. An increase of the length in red phase could raise the usage of the inside LTL while decreasing the one of the outside LTL. Such result can be owing to that if the vehicles disperse from the platoon after they leave the upstream intersection, the drivers' choices of a lane is quite random, with each driver choosing the lane that allows him or her the best access to the desired lane downstream. In contrast, relative balanced traffic distribution on the TLTLs in green phase appears at the study site where the upstream segment is shortest with respect to other sites. It is opposite to the case of the red phase, because a longer upstream segment costs the drivers more time to go through, and weaken their confidence to cross the intersection ahead in the green phase. The drivers could make lane choice to meet other needs, rather than selecting the LTL with the shortest queue.

In most cases, the length of a road segment is confirmed when planning urban road network, so it is hard to adjust in operation stage. However, after analysing the estimated joint effects of this factor, we still observe the opportunity to rebalance the traffic distribution on the TLTL by adjusting the setting of other factors together this factor.

The first solution comes from the shortened green phase length or the increased red phase length of left-turn traffic. The effectiveness of this method counts on the countdown device of signal phase set at the surveyed intersections. In red phase, the platooned vehicles caused by a limited length of the upstream segment are easier to be dispersed timely to fill the shorter queues emerged on the TLTLs if the remained red phase time is illuminated on the countdown device. Such assistance is also helpful for the drivers in green phase, as the device can release their anxiety of hurrying to go through the intersection

ahead and allow them to select the lane with the shortest queue, rather than irrationally select a lane with a longer queue. Such function of the countdown device has been observed in previous studies. The traditional target to set the countdown timers at the intersection is to help to reduce the start-up delay at the beginning of the green phase, and reduce the number of red-light violations during the beginning of the red phase (Limanond *et al.* 2010). After a public opinion survey, Chiou and Chang (2010) found that more than half of the surveyed drivers reported that the timers help to relieve the frustration caused by uncertain amounts of time during red or green phase. This finding is verified by the field observation and traffic analysis of this study.

The second solution is to set the TLTLs sign closer to the upstream intersection. From the analysis of the independent effect of F3, we know that this setting is helpful to the balance of left-turn traffic distribution, as it gives the drivers more time to adjust their target lanes. We find that such function is also effective when the distribution is unbalanced owing to the limited length of the upstream segment of the TLTLs.

The third solution is to increase the number of through and LTLs F7 in the same approach of the TLTLs. Extending the length of the upstream segment is found to enhance the usage of the inside lane of the TLTLs while reducing the one of the outside LTL. This effect could be weakened by upgrading F7's level, and their interactive effect is helpful for rebalancing the left-turn traffic distribution. Such function of the upgraded F7 derives from a fact that more through and right-turn lanes on the approach decreases the possibility of left-turn drivers directly entering the LTLs, vice versa. The outside LTL will be the first choice when the drivers intend to change from through lane to their target LTL. If they want to switch to the median or inside LTL, they need to continue changing lane, yet which conflicts with most drivers' will. Hence, an increase of the through and right-turn lanes easily results in higher utilization of the outside LTL than the median and inside counterparts. This tendency has been verified by the estimated independent effect of F7 that will be mentioned in Section 4.2.3.

4.2.2. Length of longitudinal F5 and lateral F6 movement

The trajectory of a left-turn vehicle within an intersection presents as a part of a circle or ellipse. The trajectory length is hard to measure accurately for the drivers, but it can be estimated by dividing it as a combination of longitudinal moving distance and the lateral one. For the drivers on the approach ("A" in Figure 3), their longitudinal movements aim to pass entry approach of the crossing road ("B" in Figure 3), while the lateral ones aim to pass exit approach in the same segment ("C" in Figure 3). The drivers easily observe the longitudinal crossing distance F5 when approaches to the intersection, while the length of lateral movement F6 can be estimated when the drivers get enough sight distance to the crossing road.

We find that the longitudinal crossing distance has a significant independent effect on the left-turn traffic distribution. When the distance increases, the distribution is away from its balanced state in the red phase or in cycle period. In an individual perspective, the drivers originally selecting the inside LTL could switch to the median or outside lane. This change could be attributed to the larger turning radius of the median or outside LTL than the inside one. The longer longitudinal moving distance in the intersection provided drivers more accelerating space, and the higher speed could make them prefer a larger turning radius. Similar behavioural pattern is also found in green phase. The larger accelerating spacing created by increased longitudinal moving distance inside the intersection could help the drivers maintain a relatively high speed when they pass through the intersection in green phase. This could help to relieve the unbalance of left-turn traffic distribution happened when the length is limited.

As another factor related to the exposure time of left-turn vehicles in the intersection, the length of lateral movement also determines the distance to their target exit approach of the crossing road. We found that the usage of the median or outside LTL exceeds the one of inside LTL when increasing the lateral moving distance. This change creates the left-turn vehicles the larger turning radius than the inside lane does and allows the drivers conduct the turning movement in a more comfortable way. In addition, compared with the change of the longitudinal moving distance, the variation of lateral one has a larger impact on the TLTLs traffic distribution. From the view of an individual driver, this result could be interpreted as the drivers being more sensitive to the change of lateral moving distance than the longitudinal one. Such behavioural pattern could derive from the fact that the drivers are hard to observe entrance approach of the crossing road (F5 in Figure 3) when they arrive at the end of the queues on the LTLs. In contrast, the lane number of the opposite approach (F6 in Figure 3) are much easier to count for the drivers, so it is not surprising to find F6 affecting the drivers more than F5 does.

To improve unbalanced left-turn traffic distribution caused by improper design of longitudinal moving distance, we can adjust the settings of red phase length F1 or lateral moving distance F6. Shortening the length of red phase independently could increase the usage of the inside LTL while reducing the one of the outside lane, which can be used to relieve the unbalance LTL traffics caused by increased longitudinal moving distance. The increased distance arouses the drivers' anxieties to cross the intersection as soon as possible in the green phase, which results in low usage of the inside LTL. Hence, it is not surprising to find that if the drivers make turning at a place, which is closer to exit approach of the crossing road ("D" in Figure 3), their anxieties can be released somewhat, and the possibility of each LTL being chosen could tend to be equal. This explains why the unbalanced distribution could be improved by degrading F6's level meanwhile upgrading F5's level. But this adjustment is not effective in the red

phase, because the drivers have to wait behind the stop line, and the slight advantage of a closer position to the crossing road is ignorable for them.

4.2.3. Number of other traffic lanes F7

As mentioned in Section 4.2.1, the number of through and right-turn lanes in the same approach of the TLTLs are related to the way of left-turn drivers entering their target lanes. Less through and right-turn lanes increase the possibility that the drivers directly enter the LTLs after leaving upstream intersection without lane changing. Sando and Moses (2009) analysed multilane traffics collected at two successive intersections and found that the upstream traffic distribution is directly related to the downstream condition on the TLTLs. Although this study applied a different method to collect data, their study still provided valuable information for us.

We found that increasing the number of lanes could impair the balanced left-turn traffic distribution on the TLTLs in some cases. One solution of such unsatisfactory condition is to extend the length of red F1 or green F2 phase. The increased interval of two signal phases can help the drivers build stable future expectation and have time to switch from the crowded outside LTL to the less crowded median and inside lane. Another method is to adjust the TLTLs sign F3 closer to the downstream intersection because the drivers have to change to a more inside lane in advance in case of the changing opportunity being hard to find near the stop line. Extending a shadowed inside LTL F8 on the approach can also help improve the unbalanced traffic distribution.

4.2.4. Shadowed LTL F8

Shadowed LTL provides left-turn drivers an opportunity to switch from the median LTL to the inside LTL near the stop line. By doing so, they can shorten the lateral moving distance before they enter the intersection. But we do not find the drivers express interest of the inside shadowed LTL. In fact, balanced traffic distribution is easier to achieve in the TLTLs without a shadowed lane. This finding is consistent with the one of Sando and Moses (2009). They surveyed 15 TLTL sites in Florida, USA, and the analysis of the LTL volume data support the conclusion that shadowed LTLs had lower utilization compared to unshadowed ones. The drivers' repellent to the shadowed lane derives from the cost of additional lane changing from the outside LTL to the inside one so that they can achieve the turning target without this additional movement. We also find that the unbalanced left-turn traffic distribution caused by this design can be improved by setting the TLTLs sign closer to the downstream intersection. It means that postponing informing drivers the downstream TLTLs could induce them to change to more inside LTL before they arrive at the intersection, and reduce their dependence on the outside LTL that is still available near intersection stop line. This function can also be used to improve the lane traffic balance caused by increased longitudinal moving distance in the intersection.

Conclusions

This study focuses on investigating the reasons for unbalanced left-turn traffic distribution on the TLTLs.

Individual LTL volumes were organized as the compositional data. The sum-constant constraint subjected by the data connects the volumes with the LTL choices of individual drivers. Under this constraint, the compositional regression analysis was applied to identify the influential factors of the drivers' lane choices on the TLTLs. Some suggestions for improving unbalanced traffic distribution by affecting individual drivers' lane choices were proposed after analysing the study results.

The method applied in this study could be used to analyse the utilization of other multilane infrastructures from an individual driver's perspective. For the topic discussed in this paper, it still leaves ample room for future improvements.

For example, due to the limited number of study sites, the level combinations applied to the analysis of joint factor effects do not cover all possible combinations. To survey the volume data at more sites equipped with the TLTLs will be helpful to fully understand the joint effect on the traffic distribution on the TLTLs.

In addition, only static factors were taken into the analysis since the LTL volumes had been collected in phase or cycle period. Some dynamic factors, such as queue length ahead or remaining time of a phase, could also to affect drivers' lane choices, so their influences should be studied in the future study.

Acknowledgements

This work was supported and the authors would like to acknowledge:

- » National Natural Science Foundation of China (Grants No 71901040, 71701031, 71971029);
- » Natural Science Basic Research Program of Shaanxi, China (Grants No 2019JQ-442, 2019JQ-242);
- » Key Research and Development Program of Shaanxi Province, China (Grants No 2018ZDCXL-GY-05-04, 2019ZDLGY15-04-02);
- » Science and Technology Innovation Program of the Department of Transportation, Yunnan Province, China (Grant No 2018303);
- » Technology Project of Yunnan Transportation Department, China (Grant No 2016(A)05); Fundamental Research Funds for the Central Universities, CHD (Grants No 300102321101, 300102210525), which collectively funded this research.

Author contributions

Li Li and *Dong Zhang* conceived the study and were responsible for the design and development of the data analysis.

Li Li collected the data.

Li Li, Dong Zhang, Xiao-Yun Cheng analysed the data.

Li Li, Dong Zhang, Ping Wang and Gui-Ping Wang wrote the first draft of the paper; all authors reviewed the paper.

Disclosure statement

The authors declare no conflict of interest.

References

- Aitchison, J. 1982. The statistical analysis of compositional data, *Journal of the Royal Statistical Society: Series B (Methodological)* 44(2): 139–160. <https://doi.org/10.1111/j.2517-6161.1982.tb01195.x>
- Alhajyaseen, W. K. M.; Najjar, M.; Ratrouf, N. T.; Assi, K. 2017. The effectiveness of applying dynamic lane assignment at all approaches of signalized intersection, *Case Studies on Transport Policy* 5(2): 224–232. <https://doi.org/10.1016/j.cstp.2017.01.008>
- Bagheri, E.; Mehran, B.; Hellinga, B. 2015. Real-time estimation of saturation flow rates for dynamic traffic signal control using connected-vehicle data, *Transportation Research Record: Journal of the Transportation Research Board* 2487: 69–77. <https://doi.org/10.3141/2487-06>
- Chen, P.; Nakamura, H.; Asano, M. 2011. Saturation flow rate analysis for shared left-turn lane at signalized intersections in Japan, *Procedia – Social and Behavioral Sciences* 16: 548–559. <https://doi.org/10.1016/j.sbspro.2011.04.475>
- Chiou, Y.-C.; Chang, C.-H. 2010. Driver responses to green and red vehicular signal countdown displays: safety and efficiency aspects, *Accident Analysis & Prevention* 42(4): 1057–1065. <https://doi.org/10.1016/j.aap.2009.12.013>
- Cooner, S. A.; Ranft, S. E.; Rathod, Y. K.; Qi, Y.; Yu, L.; Wang, Y.; Chen, S. 2011. *Development of Guidelines for Triple Left and Dual Right-Turn Lanes: Technical Report*. Report No FHWA/TX-11/0-6112-1. Texas Transportation Institute, Texas A&M University System, TX, US. 190 p. Available from Internet: <https://static.tti.tamu.edu/tti.tamu.edu/documents/0-6112-1.pdf>
- Kikuchi, S.; Kii, M.; Chakraborty, P. 2004. Lengths of double or dual left-turn lanes, *Transportation Research Record: Journal of the Transportation Research Board* 1881: 72–78. <https://doi.org/10.3141/1881-09>
- Limanond, T.; Prabhajob, P.; Tippayawong, K. 2010. Exploring impacts of countdown timers on traffic operations and driver behavior at a signalized intersection in Bangkok, *Transport Policy* 17(6): 420–427. <https://doi.org/10.1016/j.tranpol.2010.04.009>
- Mateu-Figueras, G.; Pawlowsky-Glahn, V.; Egozcue, J. J. 2011. The principle of working on coordinates, in V. Pawlowsky-Glahn, A. Buccianti (Eds.). *Compositional Data Analysis: Theory and Applications*, 31–42. <https://doi.org/10.1002/9781119976462.ch3>
- Pawlowsky-Glahn, V.; Egozcue, J. J.; Tolosana-Delgado, R. 2015. *Modelling and Analysis of Compositional Data*. John Wiley & Sons, Ltd. 252 p. <https://doi.org/10.1002/9781119003144>
- Pompigna, A.; Rupi, F. 2017. Lane-distribution models and related effects on the capacity for a three-lane freeway section: case study in Italy, *Journal of Transportation Engineering, Part A: Systems* 143(10): 05017010. <https://doi.org/10.1061/JTEPBS.0000080>
- Qi, H.; Wang, D.; Chen, P.; Bie, Y. 2013. Simulation of departure flow profile at stop lines for signal approach spillover, *Transportation Research Part C: Emerging Technologies* 33: 88–106. <https://doi.org/10.1016/j.trc.2013.05.001>
- Ring, J. B.; Sadek, A. W. 2012. Predicting lane utilization and merge behavior at signalized intersections with auxiliary lanes in Buffalo, New York, *Journal of Transportation Engineering* 138(9): 1143–1150. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000426](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000426)
- Sando, T.; Moses, R. 2009. Influence of intersection geometrics on the operation of triple left-turn lanes, *Journal of Transportation Engineering* 135(5): 253–259. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000005](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000005)
- TRB. 2000. *Highway Capacity Manual*. Transportation Research Board (TRB), National Research Council, Washington, DC, US. 1134 p.
- Van den Boogaart, K. G.; Tolosana-Delgado, R. 2013. *Analyzing Compositional Data with R*. Springer. 258 p. <https://doi.org/10.1007/978-3-642-36809-7>
- Wei, F.; Wang, Z.; Qu, Z.; Lu, J.; Lin, P.-S. 2016. Study on lane changes during left turns on quadruple left-turn lanes at signalized intersections in China, in *Transportation Research Board 95th Annual Meeting*, 10–14 January 2016, Washington, DC, US, 1–16.
- Wong, C. K.; Heydecker, B. G. 2011. Optimal allocation of turns to lanes at an isolated signal-controlled junction, *Transportation Research Part B: Methodological* 45(4): 667–681. <https://doi.org/10.1016/j.trb.2010.12.001>
- Yeom, C.; Schroeder, B. J.; Cunningham, C.; Vaughan, C.; Roupail, N. M.; Hummer, J. E. 2014. Lane utilization at two-lane arterial approaches to double crossover diamond interchanges, *Transportation Research Record: Journal of the Transportation Research Board* 2461: 103–112. <https://doi.org/10.3141/2461-13>
- Yi, Q.; Chen, X.; Li, D. 2012. *Development of Warrants for Installation of Dual Right-Turn Lanes at Signalized Intersections*. Report No SWUTC/12/161141-1. Center for Transportation Training and Research, Texas Southern University, Houston, TX, US. 116 p.
- Yousif, S.; Al-Obaedi, J.; Henson, R. 2013. Drivers' lane utilization for United Kingdom motorways, *Journal of Transportation Engineering* 139(5): 441–447. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000531](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000531)
- Zhao, J.; Ma, W.; Zhang, H. M.; Yang, X. 2013. Two-step optimization model for dynamic lane assignment at isolated signalized intersections, *Transportation Research Record: Journal of the Transportation Research Board* 2355: 39–48. <https://doi.org/10.3141/2355-05>
- Zhao, J.; Yao, J.; He, S.; Han, Y. 2017. Operational efficiency evaluation of intersections with dynamic lane assignment using field data, *Journal of Advanced Transportation* 2017: 2130385. <https://doi.org/10.1155/2017/2130385>
- Zhou, H.; Ding, J.; Qin, X. 2016. Optimization of variable approach lane use at isolated signalized intersections, *Transportation Research Record: Journal of the Transportation Research Board* 2556: 65–74. <https://doi.org/10.3141/2556-07>