

Research Article Characteristics Analysis of Bus Stop Failure Using Automatic Vehicle Location Data

Rui Li,¹ Xin Xue,¹ and Hua Wang²

¹College of Civil and Transportation Engineering, Hohai University, 1 Xi Kang Road, Nanjing 210098, China ²Department of Civil & Environmental Engineering, National University of Singapore, 21 Lower Kent Ridge Road, 117576, Singapore

Correspondence should be addressed to Hua Wang; hwang191901@gmail.com

Received 27 March 2020; Revised 6 June 2020; Accepted 10 June 2020; Published 1 July 2020

Academic Editor: Zhiyuan Liu

Copyright © 2020 Rui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Queue forming behind a bus stop on an urban street is common and a traffic bottleneck usually occurs around the bus stop area. The bus stop failure means arriving buses cannot move into the bus stop due to limited capacity but have to wait for available loading areas. It is related with the transit operation level. Traditionally, the failure rate (FR), defined as the percentage of buses that arrives at the bus stop to find all loading areas occupied, is adopted in bus capacity analysis. However, the concept of FR is unable to quantitatively analyze failure characteristics in terms of its dispersion and uncertainty over time. Therefore, in this paper, we propose a new index called failure duration rate (FDR) to evaluate the bus stop failure, which can characterize waiting time for traffic delay calculation and capacity drop estimation. The automatic vehicle location data at eight bus stops in Wujiang District Suzhou, China, over 56 working days, are used to analyze the temporal characteristics of FR and FDR. We next examined the failed service duration characteristics during peak hours at the eight bus stops. Based on these characteristics analyses, we then proposed a Distribution Fitting and Cumulative Distribution Correlation (DF-CDC) approach to explore the correlation between FDR and FR at the same cumulative distribution function levels and validated the bus stop failure performance using the cross-validation method. The analysis results revealed that (i) FR fluctuates more significant than FDR, (ii) FDR is a more robust index than FR in describing the traffic characteristics incurred by bus stop failures, and (iii) FDR performs better in failure characteristics analysis than FR.

1. Introduction

A bus stop serving a large number of bus lines can experience a condition known as bus stop failure due to limited capacity and high passenger demand, which will negatively affect the punctuality and reliability of bus service and also bring about delays to other traffic. The more frequently the bus stop failure takes place, the lower the transit system level of service (LOS) is [1]. However, irregular traffic flow characteristics concerned with bus stop failure are difficult to be captured and quantified because of its dispersion and uncertainty over time [2]. Without a doubt, bus stop failure will significantly affect traffic characteristics at bus berths and adjacent lanes. Typically, the failure rate is proposed for analyzing and evaluating the influence of bus stop failure at bus berths.

Bus stop constitutes one potential bottleneck to interrupt traffic flow, which will deteriorate the level of transit operational service [3]. Transit operation parameters of buses served at loading areas, including dwell time [4, 5], headway [6, 7], capacity [8, 9], queue length [10, 11], and bunching characteristics [12, 13], are analyzed for evaluating the impact of bus stop failures. The failed service will increase bus waiting time for passenger boarding and alighting, and its impact can be measured by an index called failure rate (FR), which is defined as the percentage of buses that arrive at stops to find all berths are full [1]. Wang et al. analyzed the correlation between failure rate and four kinds of transit dwell and arrival characteristics and proposed a diffusion approximation method [14]. As one of the desired level indexes of transit operation, the FR could assess the change of capacity and LOS of bus berths. Failure probability and dwell time variability can develop the function of bus queue length, which can better reflect the effects on bus stop capacity [15]. A parameter "Z" associating with the desired failure rate (under the assumption of standard normal distribution fitting) is developed to account for the fluctuation in bus dwell time in bus loading areas, and the design failure rates for urban and rural areas are recommended for estimating bus berth capacity [1]. The failure rate of curbside bus stops can be influenced by a serial of factors (such as bus arrival distribution type, bus arrival rate, bus berth maximal service rate, and bus service time variation), and the normalized capacity and incremental change (for multiberth stops) in capacity at different failure rate levels are proposed [8]. Moreover, the analysis process of failure rate is a poor proxy and suggests choosing the average waiting delay for evaluating the bus berth LOS is expressed comprehensively [16].

The data associated with vehicle location are utilized for analyzing and predicting traffic flow characteristics, which can be effectively applied to transit operational characteristics analysis [17-20], bus schedule optimization design [21, 22], bus lane planning and control strategy [23–25], and transport network flow estimation [26-28]. Several findings of characteristics analysis for public transit LOS at bus stop appear in the relevant literature. A data platform for monitoring patterns of bus operation is developed, which is primarily composed of data acquired from the ITS system in Beijing, China. A multilevel framework for transit performance analysis is proposed considering several transit operational factors [29]. Based on automated vehicle location (AVL) data, lots of statistical parameters about travel times are analyzed for evaluating the performance of bus routes with transit priority facilities, and these tests indicate spatial and temporal characteristics are the most potent feature [30]. A regression method (using LS-SVM) is developed for exploring bunching patterns of buses halting at the stop area, and the headway irregularity pattern is analyzed using transit smart card data [31]. A probabilistic method considering the interference between buses, using the loading areas, is established for predicting travel time of buses using trajectory and ID card data, which can reflect buses' dwell time distribution pattern well [32].

Although several findings of bus stop failure analysis using the failure rate appear in the relevant literature, the duration time of bus stop failure is rarely mentioned. Additionally, little research has been observed using AVL data to analyze the characteristics of bus stop failure. In this paper, failure duration time is utilized for evaluating bus stop failure, and a measure called failure duration rate is proposed for failure analysis utilizing collected transit automated vehicle location (AVL) data. The characteristics analysis of bus stop failure using AVL data can provide valuable information for transit operation optimization to the public transit authority.

The remainder of this research is organized as follows. In Section 2, a characteristic index called Failure duration rate (FDR) is developed for bus stop failure characteristic evaluation. Section 3 explores the failed duration characteristics at different failure levels based on the AVL data collected from eight bus stops in Wujiang District of Suzhou in China. In Section 4, a correlation analysis between FDR and FR is carried out by using a "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" analysis approach. Section 5 concludes the paper.

2. Characteristic Analysis Indexes

2.1. Failure Rate. A bus stop failure occurs when a bus arrives at the loading areas but with no available berth to use. The failure rate (FR), defined as the percentage of bus queuing for moving into bus berths occupied by other dwelling buses [14], could be formulated by

$$FR = \Pr(n > b), \tag{1}$$

where *n* is the number of buses halted at a bus stop and *b* is the number of bus berths.

In general, transit vehicles' dock at a bus stop (including the served and waiting buses) will obey the first-in-first-out rule and usually disperse in an independent manner to each other. The probability for the case without adequate berths at a bus stop could be calculated by

$$\Pr(n > b) = 1 - \Pr(n = 1) - \Pr(n = 2) - \dots - \Pr(n = b).$$
(2)

When the berths at a bus stop are all occupied by buses for passenger boarding and alighting, the number of active buses served at the bus stop is equal to the number of berths (n = b). And the probability of this kind of bus arrival can be approximated by the value of $Q_s^{n=b}$ divided by Q_s :

$$\Pr(n=b) \approx \frac{Q_s^{n=b}}{Q_s},\tag{3}$$

where $Q_s^{n=b}$ is the occurrence times for (n = b) during a given observation period and Q_s is the total number of arriving buses during the same time duration.

Then, we have

$$\Pr(n > b) \approx 1 - \frac{Q_s^{n=1}}{Q_s} - \frac{Q_s^{n=2}}{Q_s} - \dots - \frac{Q_s^{n=b}}{Q_s}$$

$$= \frac{Q_s - (Q_s^{n=1} + Q_s^{n=2} + \dots + Q_s^{n=b})}{Q_s}.$$
(4)

To facilitate our presentation, we denote the right hand of equation (4) by

$$R_f = \frac{Q_s - (Q_s^{n=1} + Q_s^{n=2} + \dots + Q_s^{n=b})}{Q_s}.$$
 (5)

In other words, R_f stands for the FR of a bus stop, which reflects the level of failure (LOF) for bus loading areas.

2.2. Failure Duration Rate. As the berths of a bus stop are occupied, the next arriving bus needs to queue at street lanes and exerts negative impacts on blocking movements of other vehicles along the same street lane. As a result, traffic delay goes up, and travel time reliability would be reduced [1]. It is an important and challenging task to analyze these adverse

effects quantitatively. In general, the longer the bus stop failure service lasts, the worse the traffic efficiency evolves and deteriorates. It is worth pointing out that the severity of traffic deterioration in terms of traffic delay and road capacity reduction dramatically depends on the traffic blocking duration time.

Thus, the failure duration rate that incorporates failure duration time is proposed for analyzing the LOF of bus stops. The failure duration time can be measured by the timespan (the waiting time of buses outside the stop) for all failed bus stopping services during a given time period. Specifically, it can be calculated by examining the arrival and departure time of buses using the loading areas of a bus stop. Then, the failure duration rate (FDR) can be formulated by

$$R_{fd} = \frac{t_{fd}}{t_s} = \frac{t_{fd}}{t - t_v},\tag{6}$$

where t_{fd} is the failure duration time (sec), t_s represents the total occupancy time of using the bus stop (sec), and t_v is the duration time for a vacant bus stop (sec).

The FDR can be interpreted as the ratio of waiting and blocking for arriving buses during a given time period. For a specific bus stop failure, the average duration time per failure (t_{fdr}) can be formulated as

$$t_{fdr} = \frac{t_{fd}}{Q_s - (Q_s^{n-1} + Q_s^{n-2} + \dots + Q_s^{n-b})}.$$
 (7)

Then, the average duration rate per failure (R_{fdr}) is obtained as

$$R_{fdr} = \frac{t_{fdr}}{t - t_{\nu}}.$$
(8)

3. Characteristic Analysis

3.1. Data Collection. In this study, bus dwelling data are based on the AVL data provided by Wujiang Transit Agency in Suzhou, China. The AVL data span 56 consecutive working days from October 22, 2018, to January 9, 2019. The dataset of each day has a half-day bus dispatching time window, from 7:00 to 19:00, and there are nine bus routes and eight bay-type bus stops (see their geographic locations in Figure 1). These eight test bus stops keep considerable distances to intersections (the average distance of 200 m), and thus the interaction between the bus bay and nearby intersections would be negligible.

The details of the nine bus routes associated with each bus stop are given in Table 1, and the headways of these transit routes range from 8 to 15 minutes.

The dwelling time of buses serving these 9 routes at the 8 bus stops are extracted from the collected AVL data. Some records are provided in Table 2. Based on the arrival and departure time of buses boarding and alighting passengers at the third bus stop, we can determine the bus failure characteristics. For example, the bus with the ID of SU-EU9353 serving Route 710 departed from the bus stop at 16:14:28, while the bus with the ID of SU-EU6029 serving Route 741 arrived at the same stop at 16:15:07, and the bus arriving later needed to wait outside the stop for 13 seconds (failure duration time) until bus no.710 left the stop.

3.2. Temporal Characteristics Analysis. The number and duration of buses stop failure (per hour) at bus stops are determined based on AVL data. We here use equations (5) and (6) to calculate the hourly FR and FDR of the eight bus stops. Figure 2 plots the hourly time-varying characteristics of R_f and R_{fd} over 56 working days of the No. 1 bus stop (672 hourly R_f data and R_{fd} data in total). It can be observed that R_f and R_{fd} in the morning (7:00–9:00) and evening peak hours (17:00–19:00) are higher than those in nonpeak hours. Overall, the mean value of R_f is more significant than that of R_{fd} in most times.

We then look at the median value of the hourly R_f and R_{fd} . As we can see in Figure 3, there are highly similar patterns of R_f and R_{fd} during peak hours for all the eight bus stops. The reason might be that there is high travel demand in both passengers and vehicles are in at peak hours, and thus bus stop failures occur more frequently, especially because of a long time for passengers to board and alight.

3.3. Failure Duration Analysis. Figure 4 shows the relationship between R_f and R_{fd} . For each bus stop, 224 data during the morning and evening peak hours (7:00 to 9:00 and 17:00 to 19:00) are considered.

From Figure 4, the hourly FDR shows a weak positive correlation to the FR for all the eight bus stops. Overall, the fitting parameters between R_f and R_{fd} vary remarkably among bus stops sites. For example, there are the lowest *R*-square (0.31) for no.1 bus Stop and the highest *R*-square (0.661) for no. 4 bus stop. In addition, the statistical relationship between R_f and R_{fd} of 1-berth stops (no.1 through no.6 bus stops) is weaker than that of the two 2-berth stops (no.7 and no.8 bus stops).

As mentioned above, it is difficult to determine a wellfitted failure duration rate function (for the 8 bus stops) using failure rate directly due to the dispersion. Taking no.1 bus stop as an example, R_{fd} is increased from 13.3% to 33.0% as R_f changes from 24% to 25%. But R_f increases from 16.7% to 43.3% as R_{fd} increases from 24% to 25%. It indicates the significant dispersion. To decrease the dispersion of failure duration rate in analysis, the failure rate of bus stops is divided into sections with an interval length of 5%. $R_f^{5\%}$ is defined as the 5% section level of failure rate at a bus stop, which range from 5% to 10% $(R_f^{5\%} \in [5\%, 10\%))$. Then, the mean, standard deviation (S.D.), and coefficient of variation (C.V.) of failure duration rate for these 8 bus stops are clustered and calculated at a different level of failure rate. The mean, S.D., and C.V. of failure duration rates at different failure rate level of these eight test bus stops are compared in Table 3. When the frequency of bus stop failure is less than 5 at a level, the failure duration rate is not calculated and identified as "not available (N/A)" in our analysis. Because the span of these 8 bus stops is different, each bus stop has some "not available (N/A)" at a corresponding failure rate level.



FIGURE 1: Locations of eight bus stops for AVL data collection.

TABLE 1: The details of test bus stops.

No.	Bus stop (number of bus berths)	Direction	Bus routes
1	Hengda Market West (1)	$S \longrightarrow N$	Routes 708, 719, 732, 733
2	Hengda Market West (1)	$N \longrightarrow S$	Routes 708, 719, 732, 733
3	Changan Rd. and Lianyang Rd. North (1)	$S \longrightarrow N$	Routes 701, 710, 711, 720, 741
4	Changan Rd. and Lianyang Rd. North (1)	$N \longrightarrow S$	Routes 701, 710, 711, 720, 741
5	Changban Rd. and Changan Rd. East (1)	$E \longrightarrow W$	Routes 701, 710, 711
6	Changban Rd. and Changan Rd. East (1)	$W \longrightarrow E$	Routes 701, 710, 711
7	Changan Rd. and Changban Rd. North (2)	$S \longrightarrow N$	Routes 708, 719, 720, 732, 733, 741
8	Changan Rd. and Changban Rd. North (2)	$N \longrightarrow S$	Routes 708, 719, 720, 732, 733, 741

TABLE 2: Some buses arrival and departure time records extracted from AVL data.

Douto no	Bus ID	Pue etcn	Date/time			
Koute no.		bus stop	Arrival time	Departure time		
701	SU-EU9526	No.3 bus stop	2018-12-18 16:10:31	2018-12-18 16:11:11		
710	SU-EU9353	No.3 bus stop	2018-12-18 16:13:34	2018-12-18 16:14:28		
741	SU-EU6029	No.3 bus stop	2018-12-18 16:14:15	2018-12-18 16:15:07		
720	SU-EU9582	No.3 bus stop	2018-12-18 16:19:15	2018-12-18 16:20:08		
701	SU-EU9379	No.3 bus stop	2018-12-18 16:20:30	2018-12-18 16:21:02		
711	SU-EU9359	No.3 bus stop	2018-12-18 16:22:16	2018-12-18 16:23:19		

Based on the mean, S.D., and C.V. of failure duration rates at different failure rate level, we then analyze the failure duration characteristics. Figure 5 presents the average value (for all these eight bus stops) of mean, S.D., and C.V. for failure duration rate at different failure rate levels. The mean and S.D. of R_{fd} increase with the increasing failure rate level, and the mean has a higher and faster growth rate than that of S.D. The mean of R_{fd} is increased from 5.2% to 27.5% as failure rate level increases from 5% to 40%. And the S.D. of R_{fd} is increased from 3.1% to 5.5% as the same span of failure rate level. However, the favorable trend does not hold for the C.V. of R_{fd} , as illustrated in Figure 5. The C.V. of R_{fd} is decreased from 59.6% to 19.9% as failure rate level changes from 5% to 40%.

Figure 6 displays the average value of the mean, S.D., and C.V. for failure duration rate of six 1-berth stops and two 2-berth stops. The average value of the mean, S.D., and C.V. for failure duration rate of the single-berth stops is more significant than that of the 2-berth stops.

Figure 7 displays the correlation relationships between FDR per failure (R_{fdr}) and the FR for the eight bus stops. It



FIGURE 2: The time-varying characteristics of no.1 bus stop failure.



FIGURE 3: Bus stop failure time-varying characteristics of test bus stops. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.





FIGURE 4: The relationship analysis between FDR vs. FR. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.

Failu le	re rate vel	No.1 bus stop	No.2 bus stop	No.3 bus stop	No.4 bus stop	No.5 bus stop	No.6 bus stop	No.7 bus stop	No.8 bus top
	Mean	N/A	N/A	N/A	N/A	0.0710	0.0504	0.0382	0.0464
$R_{f}^{5\%}$	S.D.	N/A	N/A	N/A	N/A	0.0508	0.0289	0.0225	0.0236
5	C.V.	N/A	N/A	N/A	N/A	0.7155	0.5732	0.5878	0.5081
	Mean	N/A	N/A	0.0896	0.0885	0.0960	0.0807	0.0835	0.0797
$R_{f}^{10\%}$	S.D.	N/A	N/A	0.0375	0.0360	0.0431	0.0383	0.0397	0.0353
5	C.V.	N/A	N/A	0.4188	0.4070	0.4488	0.4750	0.4750	0.4430
	Mean	0.1443	N/A	0.1541	0.1104	0.1155	0.1133	0.1118	0.1070
$R_{f}^{15\%}$	S.D.	0.0502	N/A	0.0544	0.0355	0.0480	0.0520	0.0310	0.0461
J	C.V.	0.3477	N/A	0.3527	0.3218	0.4159	0.4587	0.2769	0.4306
	Mean	0.1735	0.1380	0.1630	0.1405	0.1338	0.1304	0.1422	0.1450
$R_{f}^{20\%}$	S.D.	0.0498	0.0554	0.0510	0.0378	0.0494	0.0499	0.0429	0.0510
5	C.V.	0.2871	0.4015	0.3129	0.2688	0.3688	0.3828	0.3015	0.3516
	Mean	0.2104	0.2075	0.1904	0.1860	0.1776	0.1664	0.1718	0.1754
$R_{f}^{25\%}$	S.D.	0.0540	0.0661	0.0532	0.0450	0.0533	0.0576	0.0430	0.0497
5	C.V.	0.2566	0.3186	0.2794	0.2417	0.3002	0.3459	0.2505	0.2833
	Mean	0.2231	0.2159	0.2236	0.1975	0.1698	0.1832	0.1939	0.1780
$R_{f}^{30\%}$	S.D.	0.0495	0.0558	0.0554	0.0370	0.0402	0.0610	0.0460	0.0470
5	C.V.	0.2217	0.2583	0.2479	0.1872	0.2368	0.3328	0.2373	0.2640
	Mean	0.2215	0.2506	0.2569	0.2421	0.2641	N/A	0.2282	0.2328
$R_{f}^{35\%}$	S.D.	0.0478	0.0539	0.0546	0.0458	0.0608	N/A	0.0405	0.0470
	C.V.	0.2160	0.2151	0.2125	0.1890	0.2304	N/A	0.1773	0.2019
	Mean	0.2719	0.2828	0.2885	0.2587	N/A	N/A	N/A	N/A
$R_{f}^{40\%}$	S.D.	0.0532	0.0609	0.0512	0.0543	N/A	N/A	N/A	N/A
	C.V.	0.1957	0.2154	0.1774	0.2098	N/A	N/A	N/A	N/A

TABLE 3: Failure duration rate of test bus stops at different failure rate levels.



FIGURE 5: Average mean, S.D., and C.V. of FDR for eight bus stops.



FIGURE 6: Average mean, S.D., and C.V. of FDR for two types of bus stops.



FIGURE 7: FDR per failure vs. FR for test bus stops. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No.8 bus stop.

is shown that the FR is insensitive to R_{fdr} . For the no.1 bus stop, Figure 7(a) reveals that R_{fdr} is increased from 1.5% to 5.4% and from 1.9% to 3.4% when the FR climbs up from 21% to 22% and from 34% to 35%, respectively. Thus, the span of R_{fdr} is significant at the low level of failure rate, and the volatility of R_{fdr} becomes more and more slight with the increasing failure rate. Besides, the diversity correlation relationship between R_{fdr} and R_f is influenced by the number of berths of loading areas. Figure 7 shows that R_{fdr} is increased from 1.2% to 3.8% and from 1.5% to 5.4%, respectively, as the failure rate increase from 21% to 22% for no.7 bus stop (2-berth type) and no.1 bus stop (1-berth type). Therefore, compared with single-berth bus stops, the two 2berth bus stops have less R_{fdr} at the same level of failure rate.

For the different levels of failure rate (range from 5% to 40%), the value of the mean, S.D., and C.V. for R_{fdr} at the

bus stops are calculated and presented in Figure 8. The average values of the mean, S.D., and C.V. for R_{fdr} reveal a significant negative correlation with the FR level. When the FR ranges from 30% to 40%, the three statistics for R_{fdr} are not very sensitive to the FR level, and there are no obvious fluctuations. For the mean value of R_{fdr} , the maximum variation is merely 0.1% when the FR level falls into the range of 30% to 40%. Therefore, the dispersion of R_{fdr} decreases sharply with the FR level (especially when the FR is greater than 30%), which implies that the average FDR per failed is comparatively stable. In Figure 9, there is a similar trend of C.V.

For the different levels of failure rate (range from 5% to 40%), the value of the mean, S.D., and C.V. for R_{fdr} for six 1berth stops and two 2-berth stops are calculated and depicted in Figure 9. The three statistics for R_{fdr} at different



FIGURE 8: Average mean, S.D., and C.V. of FDR per failure for eight bus stops.



FIGURE 9: Average mean, S.D., and C.V. of FDR per failure for two types of bus stops.

failure rate levels for the single-berth stops is larger than that of the 2-berth stops, which shows higher stability at the 2berth bus stops.

4. Correlation Analysis

As discussed in Section 3, it is found that it is not so easy to establish a satisfactory relationship between the FDR and FR, $f(R_{fd}, R_f)$, via linear regression models. In this section, we propose a "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" method for an in-depth and more reasonable analysis of the correlation between R_{fd} and R_f .

4.1. Distribution Fitting. The distribution fitting analysis is regarded as a useful approach to mine characteristics of transit operational parameters from the probabilistic perspective [33]. A unified probability distribution is explored that can well fit the essence of R_f and R_{fd} , utilizing the

probability and statistical analysis methods. That is, we aim to understand how well a candidate distribution is fitted with predicted parameters for R_f and R_{fd} . Typically, chi-squared (χ^2), Kolmogorov–Smirnov (K-S), and Anderson–Darling (A-D) tests could be used for assessing the goodness-of-fit of our analysis results. In this paper, the K-S test statistic at a significance level of 0.05 is adopted for the goodness-of-fit test based on the data of 8 bus stops on peak periods per workday (224 data per stop). 36 probability distributions listed in Table 4 are chosen for hypothesis analysis. Table 4 shows the number of rejections for the 36 possible candidates. The results reveal that the distributions of Error, Gen. Extreme Value, Gen. Logistic, Logistic, and Normal could be selected as the candidate distributions for the correlation analysis between R_{fd} and R_f .

The five well-fitted candidate probability distributions (Error distribution, Gen. Extreme Value distribution, Gen. Logistic distribution, Logistic distribution, and Normal distribution) are redeemed for analyzing the goodness-of-fit of fitted FR and FDR at the bus stops. After estimating the

		Number	of rejections
No.	Distribution	FR	, FDR
1	Beta	4	4
2	Burr	2	1
3	Cauchy	7	6
4	Dagum	2	0
5	Erlang	5	3
6	Error	0	0
7	Fatigue Life	3	4
8	Gamma	2	0
9	Gen. Extreme Value	0	0
10	Gen. Gamma	1	0
11	Gen. Logistic	0	0
12	Gen. Pareto	2	1
13	Gumbel max	6	2
14	Gumbel min	7	6
15	Hypersecant	2	2
16	Inv. Gaussian	4	3
17	Johnson SB	2	0
18	Johnson SU	6	8
19	Kumaraswamy	4	4
20	Laplace	7	6
21	Logistic	0	0
22	Log-Logistic	3	1
23	Lognormal	3	4
24	Log-Pearson 3	5	6
25	Nakagami	1	0
26	Normal	0	0
27	Pearson 5	5	6
28	Pearson 6	2	1
29	Pert	4	2
30	Phased Biexponential	8	6
31	Phased Bi-Weibull	3	1
32	Rayleigh	4	5
33	Rice	2	1
34	Triangular	4	4
35	Uniform	4	2
36	Weibull	2	1

TABLE 4: Candidate distributions fitting results for bus failure characteristics.

Note. 5 distributions highlighted in grey shading are selected as the well-fitted candidate distribution of R_f and R_{fd} .

parameters of these distributions (using the Probability density function for distributions shown in Table 5), K-S test results (P value) for the five candidate probability distributions are plotted in Figure 10. It can be seen that the goodness-of-fit measured by P values for FDR fitted distribution is much better than that for the FR.

In Table 5, the means of P value for R_f and R_{fd} at the eight bus stops for the five candidate distributions are also given. Gen. Extreme Value distribution is the best one in terms of P value (with 0.68224 of FR and 0.87865 of FDR) in distribution fitting for the R_f and R_{fd} .

4.2. Cumulative Distribution Correlation. The probability density function of Gen. Extreme Value distribution is utilized for fitting the hourly FR and FDR distributions for the eight bus stops. The results are provided in Table 6.

Based on the calculated parameters of the fitted Gen. Extreme Value distribution in Table 6, the cumulative distribution function (CDF) curve of R_f and R_{fd} for these test bus stops can be determined. From the fitted CDF curve, the fitted value at a different level of CDF can be recorded. The fitted value at a different level of CDF can be recorded using the fitted CDF curve, and the actual value can be determined by analyzing the ranking level based on the sorted 224 data collected at each test bus stop. For analyzing the accuracy of the fitted CDF value, the relative error between the actual and fitted value of R_f and R_{fd} for test bus stops at a different level of CDF is examined. Furthermore, 17 critical levels of CDF, ranging from 10% to 90% (with 5% of interval length), are selected for verifying. The relative error between the actual and fitted value of R_f and R_{fd} for the critical CDF level at the 8 test bus stops are presented in Tables 7 and 8.

Expect for a tiny minority of critical CDF level, relative errors between the actual and fitted value of R_f and R_{fd} are always less than 10% at these test bus stops, as shown in Tables 7 and 8. Therefore, Gen. Extreme Value distribution

Distribution	Drokakility density function (DDE) and somely analo	Average	P value
Distribution	Probability density function (PDF) and sample space	FR	FDR
Error	$f(x \mid \sigma, k, \xi) = c_1 \sigma^{-1} \exp(- c_0 z ^k) c_0 = (\Gamma(3/k) / \Gamma(1/k))^{1/2} c_1 = k c_0 / 2\Gamma(1/k) \ z \equiv (x - \xi) / \sigma \sigma > 0,$ $\xi \le x < +\infty$	0.51376	0.74114
Gen. Extreme Value	$f(x \mid \sigma, k, \xi) = \begin{cases} \exp(-(1+kz)^{-1/k}(1+kz)^{(k+1/-k)})/\sigma, & k \neq 0, \\ \exp(-z - \exp(-z))/\sigma, & k = 0, \end{cases} \sigma > 0, \ \xi \le x < +\infty$	0.68224	0.87865
Gen. Logistic	$f(x \mid \sigma, k, \xi) = \begin{cases} (1+kz)^{-1-1/k} / \sigma (1+(1+kz)^{-1/k})^2, & k \neq 0, \\ \exp(-z) / \sigma (1+\exp(-z))^2, & k = 0, \end{cases} \sigma > 0, \ \xi \le x < +\infty$	0.47726	0.65518
Logistic	$f(x \mid \sigma, \xi) = \exp(-z)/\sigma(1 + \exp(-z))^2 \sigma > 0, \ \xi \le x < +\infty$	0.25996	0.40434
Normal	$f(x \mid \sigma, \xi) = \exp\left(-(1/2)(x - \xi/\sigma)^2\right)/\sigma\sqrt{2\pi}\sigma > 0, \ \xi \le x < +\infty$	0.47032	0.68553

TABLE 5: Average P value of FR and FDR at test bus stops for well-fitted candidate distributions.



FIGURE 10: Well-fitted candidate distributions *P* value comparison for bus failed characteristics. (a) Failure rate (%). (b) Failure duration rate (%).

Dece star	F	itted FR parameters	Fitted FDR parameters			
Bus stop	σ	k	ξ	σ	k	ξ
No.1 bus stop	-0.3034	0.08116	0.28627	-0.30432	0.06281	0.19447
No.2 bus stop	-0.42799	0.08592	0.35594	-0.31185	0.07768	0.22823
No.3 bus stop	-0.14816	0.08291	0.20818	-0.21402	0.07047	0.15087
No.4 bus stop	-0.28909	0.10029	0.26163	-0.25443	0.07185	0.16056
No.5 bus stop	-0.24515	0.09676	0.15746	-0.18459	0.06912	0.10188
No.6 bus stop	-0.22178	0.07511	0.11501	-0.11238	0.05439	0.06722
No.7 bus stop	-0.20209	0.08564	0.17765	-0.1715	0.06254	0.1101
No.8 bus stop	-0.21955	0.08123	0.15239	-0.13328	0.06037	0.09181

TABLE 6: Parameters of fitted FR and FDR at test bus stops.

performs well in fitting CDF value for R_f and R_{fd} at these test bus stops, and the accuracy and reliability for the fitted value are convincing.

For the distribution pattern of R_f and R_{fd} at test bus stops, Gen. Extreme Value distribution can be well fitted. Moreover, they also have a positive correlation relationship (as shown in Figure 4). Therefore, it can be considered that the fitted value of R_f and R_{fd} at the same CDF level perform equivalently in failure characteristics analysis. Figure 11 presents the fitted curves and critical level values of R_f and R_{fd} , which are fitted and utilized Gen. Extreme Value distribution for no.1 bus stop. In Figure 11, the "star (with pink color)" and "triangle (with red color)" display the fitted values of R_f and R_{fd} at these 17 critical CDF levels, and the "star" and "triangle" connected by a dotted line is defined as an equal correlation pair of R_f and R_{fd} for the corresponding critical CDF level.

Journal of Advanced Transportation

TABLE 7: Relative error between actual and fitted FR for test bus stops at different CDF levels.

Level of CDF	No.1 bus stop	No.2 bus stop	No.3 bus stop	No.4 bus stop	No.5 bus stop	No.6 bus stop	No.7 bus stop	No.8 bus stop
$P_{c}^{10\%}$	0.72	0.22	4.05	3.62	8.67	10.43	0.15	32.67
$P_{c}^{15\%}$	3.01	2.43	1.91	3.18	26.52	1.33	7.19	1.37
$P_{c}^{20\%}$	1.97	0.08	3.13	1.02	2.24	15.97	7.90	0.48
$P_{c}^{25\%}$	0.95	1.37	0.76	2.27	5.87	10.43	0.84	0.09
$P_{c}^{30\%}$	1.85	1.80	0.33	3.93	5.63	4.26	3.13	4.10
$P_{c}^{35\%}$	0.55	1.26	2.02	0.46	0.28	5.36	1.70	3.53
$P_{c}^{40\%}$	0.55	0.09	0.07	0.39	0.50	2.79	1.30	0.97
$P_{c}^{45\%}$	2.32	0.17	0.33	0.83	1.23	1.38	1.76	2.13
$P_{c}^{50\%}$	0.09	0.43	1.22	1.66	2.07	0.99	1.25	0.45
$P_{c}^{55\%}$	1.18	1.10	0.30	0.25	2.11	1.01	1.23	4.05
$P_{c}^{60\%}$	0.68	1.52	0.75	3.14	2.18	2.87	0.29	1.56
$P_{c}^{65\%}$	0.12	0.42	0.74	2.39	1.79	3.62	3.20	3.32
$P_{c}^{70\%}$	0.52	1.06	2.66	1.71	1.76	3.77	2.94	2.30
$P_{c}^{75\%}$	0.03	0.25	0.83	0.03	2.00	1.61	2.09	1.19
$P_{c}^{80\%}$	1.85	0.57	0.47	0.35	0.40	1.21	1.92	2.48
$P_{c}^{85\%}$	0.09	0.63	3.55	2.68	1.77	2.30	0.06	2.79
$P_{c}^{90\%}$	0.47	0.38	0.55	3.63	3.94	4.19	0.25	1.69

Note: $P_c^{10\%}$ represents the 10% CDF level (the 10th percentile of the fitted FR), ranking in ascending order ranging from 0 to 1.

TABLE 8: Relative error between actual and fitted FDR for test bus stops at different CDF levels.

Level of CDF	No.1 bus	No.2 bus	No.3 bus	No.4 bus	No.5 bus	No.6 bus	No.7 bus	No.8 bus
	stop (%)							
$P_{c}^{10\%}$	1.60	0.46	5.16	1.94	12.48	34.54	4.82	0.89
$P_{c}^{15\%}$	0.68	0.88	0.32	2.49	4.14	11.60	3.99	5.04
$P_{c}^{20\%}$	0.83	2.61	2.38	1.82	7.78	0.14	0.64	8.65
$P_{c}^{25\%}$	1.84	2.32	3.42	2.79	7.60	0.85	0.64	3.29
$P_{c}^{30\%}$	2.31	1.04	0.36	1.29	9.65	3.32	0.08	1.98
$P_{c}^{35\%}$	0.90	0.09	1.12	1.11	3.08	1.17	2.03	1.43
$P_{c}^{40\%}$	0.34	1.09	0.17	0.29	0.40	1.06	1.84	3.81
$P_{c}^{45\%}$	0.77	0.23	0.76	0.27	2.88	3.61	1.02	5.75
$P_{c}^{50\%}$	0.48	0.62	0.42	2.56	4.19	2.81	1.47	2.73
$P_{c}^{55\%}$	0.48	1.36	0.76	1.97	4.28	2.28	0.36	1.36
$P_{c}^{60\%}$	1.32	0.29	1.06	1.08	2.90	1.33	1.05	0.21
$P_{c}^{65\%}$	0.96	0.69	0.48	1.90	0.52	2.10	1.59	1.61
$P_{c}^{70\%}$	0.44	1.32	0.17	2.14	1.11	0.47	1.76	0.04
$P_{c}^{75\%}$	0.32	0.85	1.21	2.65	0.16	0.11	1.78	0.45
$P_{c}^{80\%}$	0.02	0.27	2.47	0.51	3.55	0.04	0.15	3.59
$P_{c}^{85\%}$	0.05	0.75	3.98	2.37	7.96	3.19	1.56	0.56
$P_{c}^{90\%}$	1.51	0.42	0.31	3.62	4.07	0.65	1.93	1.04

Figure 12 reveals the relationship for the 136 equal correlation pair of R_f and R_{fd} , including the pairs for 17 critical CDF levels at 8 test bus stops. A quite strong linear regression expression (with 0.98 of *R*-square) for the fitted couples of R_f and R_{fd} can be observed, which can reflect the significant correlation relationship between fitted R_f and R_{fd} .

4.3. Correlation Performance Evaluation. The crossvalidation method [34] for R_{fd} prediction based on AVL data is adopted for analyzing the correlation between R_{fd} and R_f . A four-step procedure for predicting R_{fd} at a certain cumulative distribution ranking level is illustrated as below. Firstly, the observed R_{fd} and R_f in peak hours of 56 workdays at 75% bus stops of the total 8 test bus stops are

selected randomly as the modeling datasets, and the data of the rest two bus stops are defined as predicted datasets. Secondly, the probability density distribution of R_{fd} and R_f (for the selected six bus stops in modeling datasets) are fitted using Gen. Extreme value distribution, and the fitted R_{fd} and R_f are recorded based on their probability density functions at critical cumulative distribution levels, respectively. Thirdly, linear regression expression is developed based on these fitted R_{fd} and R_f for the six bus stops (in modeling dataset) at critical cumulative distribution levels. Finally, the observed R_f (for two bus stops in prediction dataset) at corresponding critical cumulative distribution levels are determined, and the predicted R_{fd} can be calculated using the linear regression model (as formulated in Step 3). The prediction accuracy and reliability of R_{fd} can be determined by comparing the actual and predicted value.



FIGURE 11: Correlation between fitted FR and FDR at the same CDF level using Gen. Extreme Value distribution.



FIGURE 12: Gen. Extreme Value distribution fitted FDR vs. FR correlation analysis.

It is obvious that there are 28 different combinations of modeling and prediction datasets, as the rules described in Step 2. There are seven estimated values of R_{fd} for different cumulative distribution levels at each bus stop, calculated from these 28 different division plans. Based on cross-validation, the predicted value and relative error of R_{fd} for different cumulative distribution levels at 8 test bus stops are calculated, and the results are shown in Figure 13.

In Figure 13, the mean value of predicted relative errors for seven times R_{fd} prediction at different cumulative distribution levels is represented by a solid blue line. According to the results, these test bus stops have more accurate predicted values (the relative error is less than 15%) at most cumulative distribution levels, expect for low cumulative distribution levels (less than 15%). The prediction results of R_{fd} for different cumulative distribution levels at no.1 bus stop perform well (the relative error is not more than 5%) in general. Also, for most of the bus stops (no.2, no.4, no.5, no.6, no.7, and no.8 bus stop), the relative error of predicted R_{fd} is diminished gradually, as the cumulative



FIGURE 13: Results of FDR estimation for test bus stops at CDF levels. (a) No.1 bus stop. (b) No.2 bus stop. (c) No.3 bus stop. (d) No.4 bus stop. (e) No.5 bus stop. (f) No.6 bus stop. (g) No.7 bus stop. (h) No. 8 bus stop.



FIGURE 14: Comparison of FDR estimation relative error at CDF levels.

distribution sort decreases. No. 3 and no. 4 bus stops serve five bus routes, more than other bus stops. In general, the more bus routes are served at a bus stop, the more complicated buses arrival patterns are. Therefore, the average predicted relative error values of no. 3 and no. 4 bus stops will be larger than other bus stops.

To make an easy comparison, the average relative error of R_{fd} for 1-berth bus stops and 2-berth bus stops are plotted in Figure 14. It can be observed that, for both types of stops, the average relative error shows a decreasing trend. In addition, the volatility is insensitive when the cumulative distribution level is higher than 30%, and the average relative error is less than 8% in general. According to the results, the proposed "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" can develop a significant correlation relationship between the failure rate and failure duration rate.

5. Conclusion

In order to analyze the characteristics of bus stop failure, we propose a new measurement called FDR and make a comparison to the traditional index of the FR. Compared with the FR, the proposed FDR is capable to quantitatively assess the impact of bus stop failure on traffic efficiency. Based on the collected AVL data associated with the eight bus stops in Wujiang District of Suzhou, we make an in-depth analysis of the characteristics and correlation of the FR and FDR. Some insightful findings are summarized as follows:

(i) It can be observed that the values of FR and FDR in morning and evening peak hours are greater than those during off-peak hours. The value of the FR is usually larger than that of the FDR across all the eight bus stops.

- (ii) It is found that there is a positive correlation between the FR and the FDR. However, the *R*-square values of the linear regressions fluctuate dramatically among different bus stops. The results also indicate that the FDR is more robust than the FR in describing transit system status and traffic characteristics.
- (iii) We find that Gen. Extreme Value distribution could be well used for the fittings of both the FR and FDR and the proposed "Distribution Fitting and Cumulative Distribution Correlation (DF-CDC)" method works well in determining the fitted values of the FR and FDR at the critical CDF level that reflects a significant correlation between the FR and FDR.

Future works could be extended in two aspects. First, more AVL and other source data in other cities could be collected and used for analyzing the failure characteristics of bus stops. Second, based on the data-driven analysis of bus stop failure characteristics, we could make some scenario analysis to find out the most important factors, such as the number of bus stop berths and the number of lanes or the passenger demand.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

The authors confirm contribution to the paper as follows. R. Li and H. Wang conceptualized and designed the study; X. Xue carried out data collection; R. Li, X. Xue, and H. Wang carried out analysis and evaluation; R. Li and X. Xue prepared draft manuscript. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This research was supported by the National Key R&D Program in China (Grant no. 2018YFB1600600), Natural Science Foundation of Jiangsu Province (Grant no. BK20181307), Fundamental Research Funds for the Central Universities of China (Grant no. B200202088), National Natural Science Foundation of China (Grant no. 51508161), and Postdoctoral Science Foundation of China (Grant no. 2018M630505).

References

- KFH Group, *Transit Capacity and Quality of Service Manual*, Transportation Research Board, National Research Council, Washington, DC, USA, 3rd edition, 2013.
- [2] J. M. Bunker, "High volume bus stop upstream average waiting time for working capacity and quality of service," *Public Transport*, vol. 10, pp. 311–333, 2018.
- [3] A. Tirachini, "Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers," *Transportmetrica A: Transport Science*, vol. 9, pp. 28–49, 2013.
- [4] Q. Meng and X. Qu, "Bus dwell time estimation at bus bays: a probabilistic approach," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 61–71, 2013.
- [5] B. Bian, N. Zhu, S. Ling, and S. Ma, "Bus service time estimation model for a curbside bus stop," *Transportation Research Part C: Emerging Technologies*, vol. 57, pp. 103–121, 2015.
- [6] H. Yu, Z. Wu, D. Chen, and X. Ma, "Probabilistic prediction of bus headway using relevance vector machine regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 1772–1781, 2017.
- [7] M. Zhang, Q. Meng, L. Kang, and W. Li, "Tailored Wakebytype distribution for random bus headway adherence ratio," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 220–244, 2018.
- [8] W. Gu, Y. Li, M. J. Cassidy, and J. B. Griswold, "On the capacity of isolated, curbside bus stops," *Transportation Research Part B: Methodological*, vol. 45, pp. 714–723, 2011.
- [9] M. Shen, W. Gu, S. Hu, and H. Cheng, "Capacity approximations for near- and far-side bus stops in dedicated bus lanes," *Transportation Research Part B: Methodological*, vol. 125, pp. 94–120, 2019.
- [10] R. Widanapathiranage, J. M. Bunker, and A. Bhaskar, "Modelling the BRT station capacity and queuing for all stopping busway operation," *Public Transport*, vol. 7, pp. 21–38, 2015.
- [11] B. Bian, M. Pinedo, N. Zhu, and S. Ma, "Performance analysis of overtaking maneuvers at bus stops with tandem berths," *Transportation Science*, vol. 53, pp. 597–618, 2019.
- [12] C. F. Daganzo, "A headway-based approach to eliminate bus bunching: systematic analysis and comparisons,"

Transportation Research Part B: Methodological, vol. 43, pp. 913–921, 2009.

- [13] W. Wu, R. Liu, and W. Jin, "Modelling bus bunching and holding control with vehicle overtaking and distributed passenger boarding behavior," *Transportation Research Part B: Methodological*, vol. 104, pp. 175–197, 2017.
- [14] C. Wang, Z. Ye, E. Chen, M. Xu, and W. Wang, "Diffusion approximation for exploring the correlation between failure rate and bus stop operation," *Transportmetrica A: Transport Science*, vol. 15, pp. 1306–1320, 2019.
- [15] R. Fernández, "Modeling public transport stops by microscopic simulation," *Transportation Research Part C: Emerging Technologies*, vol. 18, pp. 856–868, 2010.
- [16] W. Gu, M. J. Cassidy, and Y. Li, "Models of bus queueing at curbside stops," *Transportation Science*, vol. 49, pp. 204–212, 2014.
- [17] A. M. El-Geneidy, J. Horning, and K. J. Krizek, "Analyzing transit service reliability using detailed data from automatic vehicular locator systems," *Journal of Advanced Transportation*, vol. 45, pp. 66–79, 2011.
- [18] Y. Zhou, L. Yao, Y. Chen, Y. Gong, and J. Lai, "Bus arrival time calculation model based on smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 81–96, 2017.
- [19] Y. Du, F. Deng, F. Liao, and Y. Ji, "Understanding the distribution characteristics of bus speed based on geocoded data," *Transportation Research Part C: Emerging Technologies*, vol. 82, pp. 337–357, 2017.
- [20] D. Zhao, W. Wang, A. Woodburn, and M. S. Ryerson, "Isolating high-priority metro and feeder bus transfers using smart card data," *Transportation*, vol. 44, pp. 1535–1554, 2017.
- [21] Y. Bie, X. Gong, and Z. Liu, "Time of day intervals partition for bus schedule using GPS data," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 443–456, 2015.
- [22] M. M. Nesheli and A. Ceder, "Improved reliability of public transportation using real-time transfer synchronization," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 525–539, 2015.
- [23] C. Ma and X. Xu, "Providing spatial-temporal priority control strategy for BRT lanes: a simulation approach," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, Article ID 04020060, 2020.
- [24] N. Saade, J. Doig, and M. J. Cassidy, "Scheduling lane conversions for bus use on city-wide scales and in time-varying congested traffic," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 248–260, 2018.
- [25] D. Yang, S. Zheng, C. Wen, P. J. Jin, and B. Ran, "A dynamic lane-changing trajectory planning model for automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 228–247, 2018.
- [26] A. Alsger, B. Assemi, M. Mesbah, and L. Ferreira, "Validating and improving public transport origin-destination estimation algorithm using smart card fare data," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 490–506, 2016.
- [27] Z. Liu, Y. Liu, Q. Meng, and Q. Cheng, "A tailored machine learning approach for urban transport network flow estimation," *Transportation Research Part C: Emerging Technol*ogies, vol. 108, pp. 130–150, 2019.
- [28] C. Ma and D. Yang, "Public transit network planning in small cites considering safety and convenience," *Advances in Mechanical Engineering*, vol. 12, pp. 1–12, 2020.
- [29] X. Ma and Y. Wang, "Development of a data-driven platform for transit performance measures using smart card and GPS

data," *Journal of Transportation Engineering*, vol. 140, Article ID 04014063, 2014.

- [30] Y. Yan, Z. Liu, and Y. Bie, "Performance evaluation of bus routes using automatic vehicle location data," *Journal of Transportation Engineering*, vol. 142, Article ID 04016029, 2016.
- [31] H. Yu, D. Chen, Z. Wu, X. Ma, and Y. Wang, "Headway-based bus bunching prediction using transit smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 45–59, 2016.
- [32] Z. Dai, X. Ma, and X. Chen, "Bus travel time modelling using GPS probe and smart card data: a probabilistic approach considering link travel time and station dwell time," *Journal of Intelligent Transportation Systems*, vol. 23, pp. 175–190, 2019.
- [33] M. M. Rahman, S. C. Wirasinghe, and L. Kattan, "Analysis of bus travel time distributions for varying horizons and realtime applications," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 453–466, 2018.
- [34] C. R. Rao and Y. Wu, "Linear model selection by cross-validation," *Journal of Statistical Planning and Inference*, vol. 128, pp. 231–240, 2005.