

Research Article

Origin-Destination-Based Travel Time Reliability under Different Rainfall Intensities: An Investigation Using Open-Source Data

Qi Zhang ¹, Hong Chen ¹, Hongchao Liu ², Wei Li ³, and Yibin Zhang ²

¹College of Transportation Engineering, Chang'an University, Xi'an 710000, China

²Department of Civil, Construction and Environmental Engineering, Texas Tech University, Lubbock, TX 79409, USA

³School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China

Correspondence should be addressed to Hong Chen; glch@chd.edu.cn

Received 19 July 2020; Revised 12 October 2020; Accepted 24 October 2020; Published 4 November 2020

Academic Editor: Yajie Zou

Copyright © 2020 Qi Zhang 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.

Origin-destination- (O-D-) based travel time reliability (TTR) is fundamental to next-generation navigation tools aiming to provide both travel time and reliability information. While previous works are mostly focused on route-based TTR and use either ad hoc data or simulation in the analyses, this study uses open-source Uber Movement and Weather Underground data to systematically analyze the impact of rainfall intensity on O-D-based travel time reliability. The authors classified three years of travel time data in downtown Boston into one hundred origin-destination pairs and integrated them with the weather data (rain). A lognormal mixture model was applied to fit travel time distributions and calculate the buffer index. The median, trimmed mean, interquartile range, and one-way analysis of variance were used for quantification of the characteristics. The study found some results that tended to agree with the previous findings in the literature, such that, in general, rain reduces the O-D-based travel time reliability, and some seemed to be unique and worthy of discussion: firstly, although in general the reduction in travel time reliability gets larger as the intensity of rainfall increases, it appears that the change is more significant when rainfall intensity changes from light to moderate but becomes fairly marginal when it changes from normal to light or from moderate to extremely intensive; secondly, regardless of normal or rainy weather, the O-D-based travel time reliability and its consistency in different O-D pairs with similar average travel time always tend to improve along with the increase of average travel time. In addition to the technical findings, this study also contributes to the state of the art by promoting the application of real-world and publicly available data in TTR analyses.

1. Introduction

Travel time reliability (TTR) plays a vital role in various applications such as evaluation of network performance [1], measuring the improvement of traffic operations and management strategies [2], quantification of service quality [3], enhancing the experience of traveler's route choice [4], and determining freeway bottlenecks [5].

Among the route level (microscopic), origin-destination (O-D) level (mesoscopic), and network level (macroscopic) studies, the route level TTR analyses have received much more attention in the past. Besides the demand from the practical side, route-based data which are relatively easier to obtain should be another reason. For instance, using the data from California State Route

91, one research found that traveler's route choice was more delicate to TTR than travel time [6]. Chepuri et al. assessed the performance of various TTR measures with bus route data collected in Chennai, India. They recommended using 95th percentile travel and buffer time as reliability indicators for bus routes [7]. Some recent route-based studies can be found in [8–12]. Because the focus of route-based studies is usually on one or a few specific routes, data were usually project-specific and most were discarded upon completion of a project, which makes continuous investigation difficult and sometimes impossible.

It is not uncommon that under certain circumstances route-based and O-D-based analyses may get similar or the same results because a route is associated with at least one

origin and one destination and thus can be viewed as a particular case of an O-D-based study. Comparison studies are limited, though. In [13], the authors concluded that there were no significant differences between O-D-based and route-based estimates in most part of the studied time periods. In [14], the researchers found that adding an alternative path tends to decrease the O-D-based TTR. Network level TTR studies are mostly simulation-based in a lack of real-world data. Some notable studies include but are not limited to the work of [15–18]. Studies based on traffic simulation are sometimes subject to serious errors caused by the underlying problems of the simulation model. A detailed discussion of simulation-based approaches is beyond the scope of the study.

Many factors, such as connected vehicles, traffic incidents, weather, work zones, special events, types of traffic control, and the dynamics of traffic flow, have impacts on TTR. Accordingly, the study of the impact of these factors has become one of the prominent topics in the TTR field [19]. For instance, in the literature [20], the authors attempted to quantify the contribution of various features on TTR and found that demand-capacity imbalance and accidents are the two factors that most affect TTR. In [21], researchers uncovered that deployment of connected vehicles improves TTR in the work zone environment, and higher benefits come along with higher market penetration levels. Additionally, the impact of rain on speed and travel time and the route level TTR have been well studied, and some results are conducive. For example, studies have discovered that speed reduction could vary from 10% to 25% in general rainy days [22] and an average increase of travel time by 11% might be expected in peak hours under the impact of a certain level of precipitation [23]. Adverse weather exacerbates TTR, especially during peak periods [24]. However, some findings are controversial and need further investigation, especially when it relates to TTR. Chien and Kolluri found that TTR would diminish when weather condition changes from dry to rain, as indicated by an expansion of 16% in the buffer index [10], while in another study [25], the authors suspected that rain and snow might have caused lower standard deviation and coefficient of variation of travel time and thus increased TTR. While it is understandable that different studies may produce controversial results, ad hoc data might have played a role. In a review of the literature, we found that studies are limited with respect to the impact of rain on the O-D level TTR; moreover, most of the data used in previous studies were project-specific and only covered a short period of time that was not even sound for a full-scale statistical analysis.

1.1. Research Objectives. The availability of publicly available open-source data in recent years has made a detailed investigation of O-D-based TTR possible. A major thrust of this study is to use Uber Movement data and Underground Weather data to systematically analyze the impact of rain on O-D-based TTR. Uber and Underground Weather data provide an ideal and probably the only opportunity for applying real-world data in such studies because Uber data

are O-D-specific and cover a lot longer time span while Underground Weather data provide very detailed weather data. A significant contribution of using publicly available data is that the results can be easily verified and compared to those that use ad hoc data or computer-aided simulation, studies based on real-world data always have a better value in the literature. Additionally, O-D-based TTR is fundamental to next-generation-navigation tools that are aimed at providing both travel time and reliability information. This paper only focuses on the impact of rain, but there are a lot more deserving further investigations along this line, such as the impact of other weather events and the combined effects of weather and work zones.

In this study, the authors investigated the impact of rain at various levels of intensity on O-D-based TTR, through the analyses of three-year travel time and weather data and a hundred O-D pairs collected from downtown Boston. A general lognormal mixture model was adopted to fit distributions and calculate the buffer index values. While a portion of the results was in proper alignment with previous studies, some turned out to be unique.

The rest of the paper is organized as follows: Section 2 introduces the data used in this research, which includes the O-D-based travel time data from Uber Movement and historical weather data collected from the Weather Underground website; Section 3 depicts the typical TTR measures and the analytical approach developed based on the Gaussian mixture model; Section 4 presents the results, and Section 5 summarizes the findings and conclusions and concludes the paper by discussions and future research.

2. Data

2.1. Uber Movement and Weather Data. The O-D-based travel time data used in this research are from Boston, the United States, retrieved from the Uber Movement website (<https://movement.uber.com>). The website provides detailed information on average travel time (ATT), classified by five-time intervals during a day, including early morning (00–07 h), AM peak (07–10 h), midday (10–16 h), PM peak (16–19 h), and evening (19–24 h). To make the results statistically sound, three-year data were used, which span from 1/1/2016 to 12/31/2018, and a hundred O-D pairs were selected. Uber already classified ATT ranges by 5-minute intervals. Considering that lots of data seem to be missing in the dataset with the ATT range of 25 minutes and beyond, we selected five categories in the analysis: (a) 0–5 mins, (b) 5–10 mins, (c) 10–15 mins, (d) 15–20 mins, and (e) 20–25 mins.

Figure 1 shows all the origin and destination nodes included in this study. Table 1 depicts the O-D pairs classified into five groups based on the ATT. Note that travel time in the table is directional and one-way (e.g., 2–16 indicates node 2 to node 16) because of the limited availability of two-way travel time in-between the O-D pairs. Twenty origin-destinations were selected in each ATT category. Table 2 presents descriptive statistics for ATT data from Uber, including mean and standard deviation (SD) in different periods. The studied area is in level terrain.

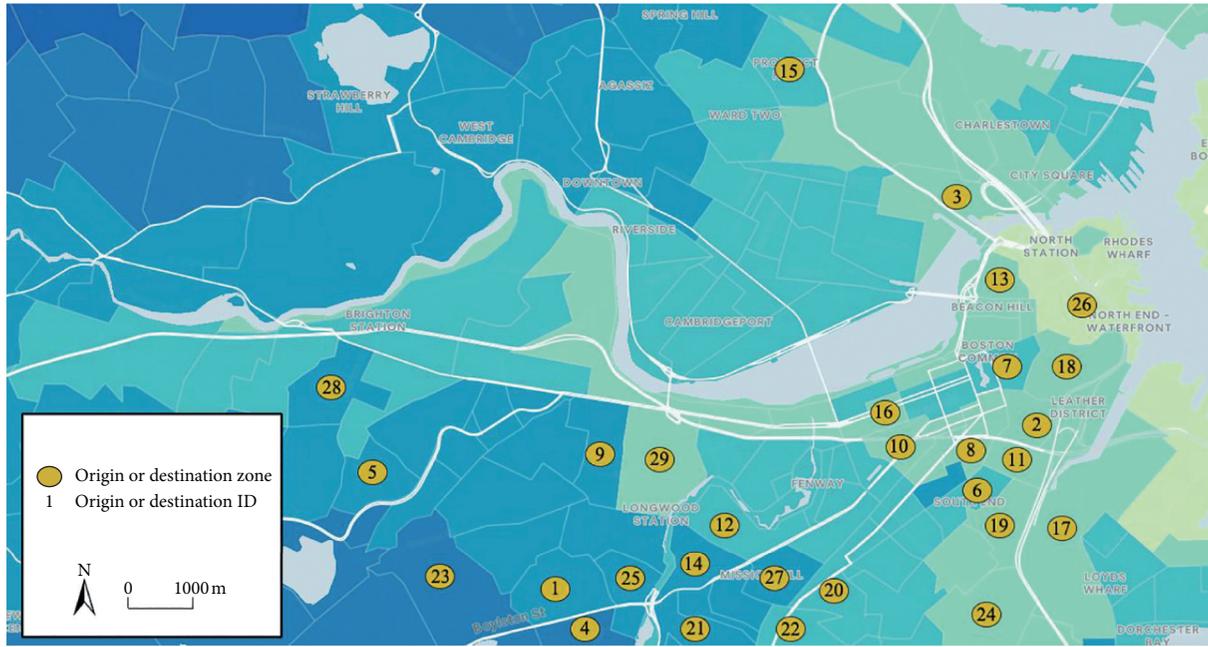


FIGURE 1: The origin and destination nodes.

TABLE 1: O-D pair information in each range of TTR.

ATT range (min)	O-D pairs (ID)
0-5	2-16, 6-10, 6-11, 7-18, 8-18, 8-29, 10-11, 11-6, 11-10, 11-26, 13-12, 17-18, 17-24, 18-2, 18-7, 18-26, 22-27, 26-13, 26-18, 27-22
5-10	2-13, 2-26, 6-7, 7-6, 8-24, 10-6, 12-16, 13-2, 13-16, 16-2, 16-13, 18-4, 18-8, 18-24, 24-8, 24-20, 26-2, 26-7, 26-11, 29-18
10-15	2-12, 6-26, 8-20, 10-26, 12-2, 12-13, 13-26, 15-26, 16-26, 18-20, 20-18, 20-29, 24-8, 24-29, 26-6, 26-10, 26-15, 26-16, 29-20, 29-24
15-20	2-9, 2-14, 9-2, 9-18, 9-26, 12-11, 12-17, 12-18, 12-19, 12-26, 14-2, 17-14, 18-12, 22-26, 26-9, 26-12, 26-14, 26-22, 26-27, 27-26
20-25	1-18, 1-26, 3-18, 5-26, 17-23, 18-1, 18-3, 18-21, 18-23, 18-25, 21-18, 23-17, 23-18, 23-26, 25-18, 26-2, 26-5, 26-23, 26-28, 28-26

TABLE 2: Descriptive statistics for ATT data from Uber (min).

ATT range (min)	O-D pair count	Early morning (00-07 h)		AM peak (07-10 h)		Midday (10-16 h)		PM peak (16-19 h)		Evening (19-24 h)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0-5	20	4.11	1.39	4.53	2.01	4.35	1.35	4.95	2.15	3.88	1.15
5-10	20	7.27	1.70	7.82	1.92	7.33	1.66	8.39	2.62	6.31	1.42
10-15	20	11.96	2.38	13.36	3.06	13.01	2.50	14.83	4.01	10.87	1.95
15-20	20	16.62	3.15	17.90	3.64	18.61	2.83	19.82	4.47	15.81	2.07
20-25	20	23.88	4.24	24.54	4.57	23.86	3.28	25.55	5.49	21.07	3.06

The corresponding weather data in the subject area were collected from the Underground Weather data at <https://www.wunderground.com/history>. In this study, weather conditions with no precipitation, such as clear, cloudy, or overcast, are classified as normal weather. Meanwhile, fog or haze conditions were excluded so that the focus could be placed on the impact of rain. Rain condition was defined as the rainy weather that caused effective precipitations.

In alignment with the 24-hour travel time data, the sum of rainfall in each matched period (like 00-07 h) was calculated and converted into a 24-hour value. Table 3 summarizes the definition of the data used in this research.

2.2. Data Screening. The travel time data have already been preprocessed and filtered by Uber before uploading to the Internet. In general, the data were well prepared, and the work

TABLE 3: Summary of data attributions and interpretation.

Attribute	Interpretation
Origin-destination	See Table 1
Dates	From 1/1/2016 to 12/31/2018
Days of week	From Sunday to Saturday
Time periods	Early morning (00–07 h), AM peak (07–10 h), midday (10–16 h), PM peak (16–19 h), and evening (19–24 h)
Travel times	Average O-D-based travel times (s)
Rainfall	Sum of rainfall per time period (mm)
Weather conditions	Normal, light rain (0.0 and 10 mm/24 h), moderate rain (10.0 mm/24 h and 25.0 mm/24 h), heavy rain (25.0 mm/24 h and 50.0 mm/24 h), extreme rain (>50.0 mm/24 h), and others (e.g., snow and fog)

for outlier removal was rather simple. There are a couple of null cells without any data within the ATT range of 20–25 mins, such as from Lincoln Road (ID: 1) to Harborwark (ID: 26); these cells were removed. Additionally, Boston has considerable snowfalls in winter months, which may cause a negative impact for a longer period of time even after the snow. According to the analysis in [26], the impact of snow on travel time was associated with the severity of snow and road conditions, and it usually takes at least six hours after the snow for travel time to become stabilized. In this study, the data recorded one-day after regular snow (≤ 5.0 mm/24 h) and two-day after heavy snow (> 5.0 mm/24 h) were excluded. As a result, around 540,000 valid records, including nearly 68,000 light rain records, 18,000 moderate rain records, 11,000 heavy rain records, and 13,000 extreme rain records, were included in the study.

3. Methods

3.1. Measures of Travel Time Reliability. In addition to the conventional measures (mean, standard deviation, and coefficient of variation, e.g.), there are some other TTR measures, such as travel time variability (TTV), planning travel time index (PI), and buffer index (BI). Among these measures, BI has been widely utilized in existing literature [7, 8, 13, 27, 28] and as concluded in [29] has a high consistency with the coefficient of variation and thus is best suitable for the measure of TTR. The authors adopted the idea and took BI as the primary TTR measure. Then, we used interquartile range (IQR), the median, and trimmed mean of BI-values, as well as the analysis of variance for BI variation ratio, to quantify the impact from rainfall intensity.

The buffer index can be generally formulated as follows:

$$BI = \left(\frac{T_p - T_c}{T_c} \right), \quad (1)$$

where T_p is the percentile travel time and T_c is the contrastive travel time (e.g., mean travel time, median travel time, and free-flow travel time).

Obviously, the higher the BI-value is, the less reliable the travel time will be. In this study, the 95th percentile travel time T_p and the mean travel time T_c were adopted as the BI parameters. The interquartile range is the distance between the 75th and 25th percentiles, and the trimmed mean excludes the 5% highest and the 5% lowest data for reducing the error caused by the extreme data.

3.2. Lognormal Mixture Model. With respect to the calculation of the percentile value in BI, in earlier literature [30], the author directly calculated the percentile value according to the available data without considering the statistical regression, which was easily subject to statistical errors (e.g., regression to the mean). In later studies, various regression methods were applied, such as multiple linear regression [20], and continuous probability distribution functions, such as Weibull distribution [31], lognormal distribution [32], and generalized Pareto distribution [33]. Multiple regression was found to be more suitable for the multiparameter impact study, while for the single factor (e.g., rainfall intensity) analysis, the latter seems to be more desirable. However, the complexity of the problem makes it difficult for the real data to fit well with traditional prior distribution, such as the lognormal distribution. Recent studies attempted to use multistate models, such as the Gaussian mixture model, lognormal mixture model, and gamma mixture model for better results [8, 13, 14, 16, 17, 34–36]. Among these methods, the lognormal mixture model (LMM) is outperformed and was recommended by many researchers [8, 34, 36]. When LMM is applied, the best fitting usually occurs at a low K -value (e.g., $K = 2$ or 3), which may also help improve computational efficiency. Accordingly, LMM was selected for this research.

LMM is essentially a linear combination of multiple lognormal distributions with a weight sum value of 1. The general formula of LMM is as follows:

$$P_{LMM} \left(\frac{t}{\omega_k, \mu_k, \sigma_k} \right) = \sum_{k=1}^K \omega_k L \left(\frac{t}{\mu_k, \sigma_k} \right), \quad (2)$$

where t is the travel time; $\omega_k, \mu_k, \sigma_k$ are the weight, mean, and standard deviation of the k^{th} lognormal distribution, respectively; and L is the lognormal probability density. The equation is subject to $\sum_{k=1}^K \omega_k = 1$.

3.3. Expectation-Maximization Algorithm. Since mixture models (like LMM) involve latent variables, maximum likelihood estimate (MLE) cannot be used directly to estimate the parameters. Presently, the expectation-maximization (EM) algorithm is the most commonly used approach for multimodal parameter estimates, where an expectation (E) step calculates the expected log-likelihood by estimating the current parameters, and a maximization (M) step

maximizes the expectation of the log-likelihood in the E step. Algorithm 1 depicts the complete process of the EM algorithm.

3.4. Supplement Algorithms. Before the application of LMM and EM, two issues need to be addressed: the optimal K-value and the inverse function of the cumulative distribution function (CDF) of LMM. The former can be resolved by referring to the method in [34], where the K-value was determined by the minimum Akaike information criterion (AIC) estimation with the null hypothesis not rejected by the one-sample Kolmogorov-Smirnov (K-S) test. AIC is defined as

$$AIC = 2C - 2 \ln(Li), \quad (3)$$

where C is the number of parameters and Li is the likelihood function.

For the second issue, since there is no corresponding original form of the CDF of LMM, it is impossible to obtain the percentile value by solving the inverse function of the original function. For this reason, the bisection method was adopted, with a stop threshold of 0.00001. The complete TTR estimation framework is presented in Figure 2.

4. Results and Discussion

The authors calculated the BI-value of each O-D pair under six different weather conditions, that is, normal, light rain (0 and 10.0 mm/24 h), moderate rain (10.0 and 24.9 mm/24 h), heavy rain (25 and 50.0 mm/24 h), extreme rain (>50.0 mm/24 h), and rain (>0 mm/24 h). We summarized all calculations into five groups according to the TTR range, that is, 0–5 mins, 5–10 mins, 10–15 mins, 15–20 mins, and 20–25 mins, including three location measures (median, trimmed mean, and interquartile range) and the one-way analysis of variance. MATLAB was used to run the s, and the final buffer index is presented in the form of the average value calculated after 50 fittings.

4.1. Increasing TTR Reduction Impact by Rain. Figure 3 shows an example fitting of LMM from Harrison Ave (ID: 11) to Huntington Ave (ID: 10) in the ATT range of 0–5 mins, which depicts a higher TTR under light rain and lower TTR under the rest of the rainfall intensity. Moreover, the impact of the O-D-based TTR increased with the increase of rainfall intensity (the variation of BI-value from 0.3332 to 0.4379). Overall, it shows that rain reduced the O-D-based TTR (see in Figure 3(f)).

More analyses were conducted for further investigation. Figure 4 and Table 4 (median of BI-values) interpret the results in terms of the median value, which shows that rain reduced the O-D-based TTR in each ATT range. Additionally, four out of five subfigures in Figure 4 demonstrate an increasing trend in TTR reduction when the rainfall intensity increases, with only one exception when being within the ATT range of 20–25 mins (BI-value decrease under heavy rain (0.3489) compared to that under moderate rain (0.3588)). The global mean values of the median under

different rainfall intensity (the 8th row in Table 4) also revealed the trend.

Likewise, the trimmed mean of BI-values in Table 4 shows that rain has adverse effects on O-D-based TTR and the increasing reduction effect in terms of the global mean (the 14th row). In terms of time ranges, three out of five presented the increasing adverse effect in the range of 0–5 mins, 10–15 mins, and 20–25 mins.

Regarding the O-D-based results (see 8th column in Table 5), thirty-five out of a hundred O-Ds strictly met the regularity (an average of seven O-Ds in each range). Meanwhile, ninety-one out of a hundred O-D pairs (see 7th column in Table 5) show that the rain reduces the O-D-based TTR.

By far, not surprisingly, a dominating feature is that rain reduces O-D-based TTR, which aligns properly with people's perception as well as previous research at route level [10, 22, 23, 37, 38]. Notably, the low-probability anomalies (positive effect of rain), though not sufficient to negate the conclusion, may be a combination of multiple factors in a real environment, for example, a combination of rain, accidents, and work zones. This counterintuitive phenomenon will be discussed in a subsequent subsection specifically. More importantly, the results reveal that the negative impact grows with the increase in rainfall intensity. This trend seems doubtful, partly due to the exceptions in Table 4 and partly due to the fact that solely 35 percent of O-D pairs strictly conform to this trend. The reasons are generally twofold: (1) the interference of the complicated environment; (2) more seriously, the quite insignificant difference in impact between moderate, heavy, and extreme rain (this characteristic will be additionally discussed in subsequent subsection), resulting in extra challenging to achieve O-D pairs with satisfying the trend. Nonetheless, the current consequences can still expose the trend effectively.

4.2. Significant Impact from Light Rain to Moderate Rain.

A unique finding was revealed from the analysis; that is, the impact is a lot more significant when rainfall intensity changes from light to moderate, while other changes among the rainfall intensity categories seemed to cause only moderate impacts on the O-D-based TTR.

Comparing the variation ratio values in the parentheses in Table 4, we found that the average increase from light to moderate is up to 17.2% (median) and 14.6% (trimmed mean). On the contrary, the average BI variation ratio between the rest of the conditions is not significant, which is 2.7% (normal to light), 1.1% (moderate to heavy), and 5.8% (heavy to extreme), respectively, and 3.3%, 3.6%, and 4.0% in average trimmed mean values, respectively.

The one-way analysis of variance (ANOVA) was used to demonstrate the statistical significance of the BI variation ratio between different rainfall intensity categories. Considering the trimmed mean in the analysis, we trimmed 10% values for the test as well. The analysis was conducted for a significance level of 0.05. As presented in Table 6, the null hypothesis is rejected with the P value of $3.5874e-16$, far less than 0.05 in the 2-column source, which indicates significant

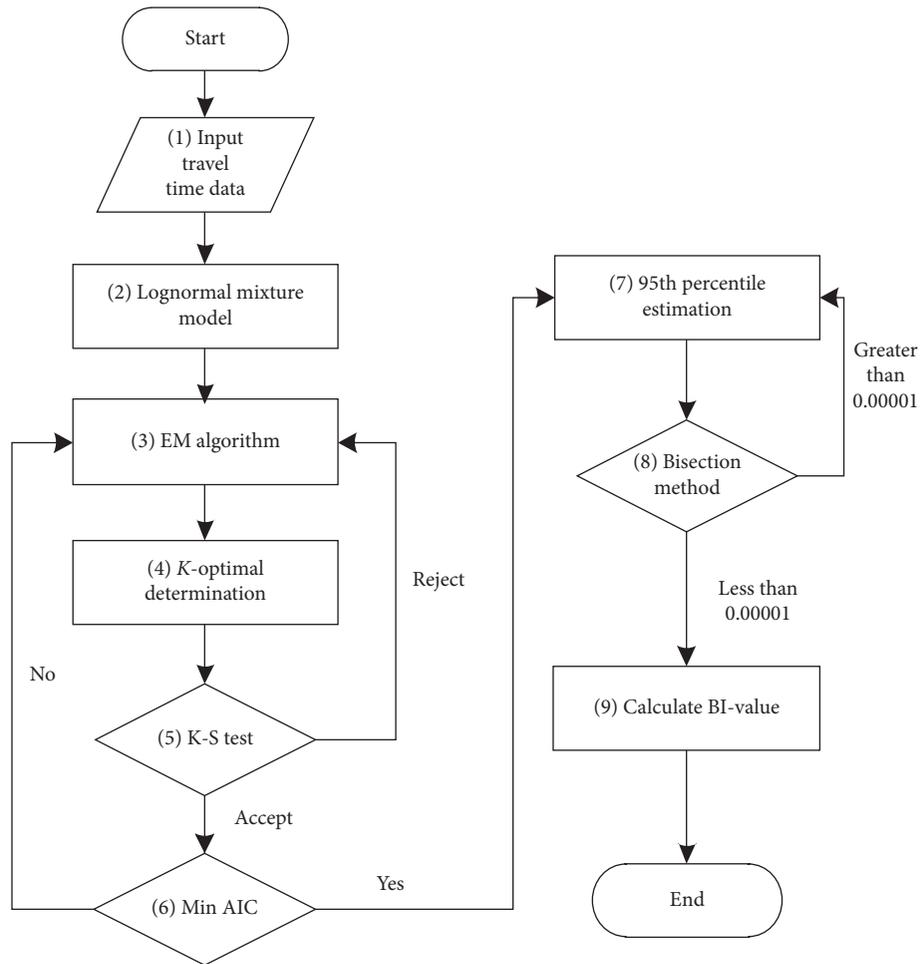


FIGURE 2: Framework for estimating O-D-based TTR.

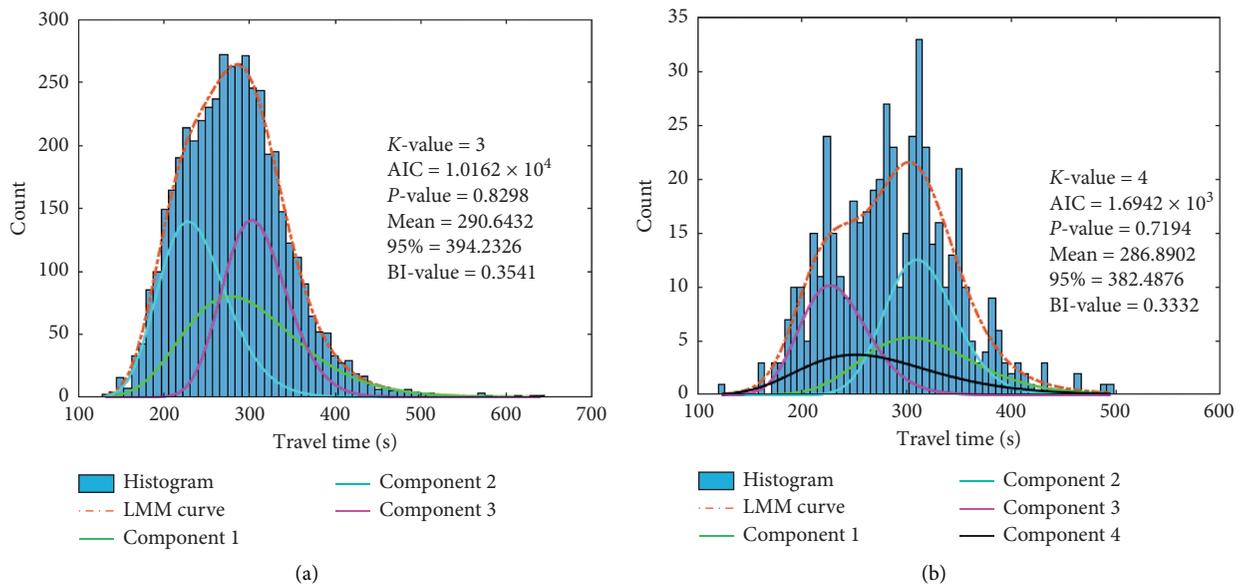


FIGURE 3: Continued.

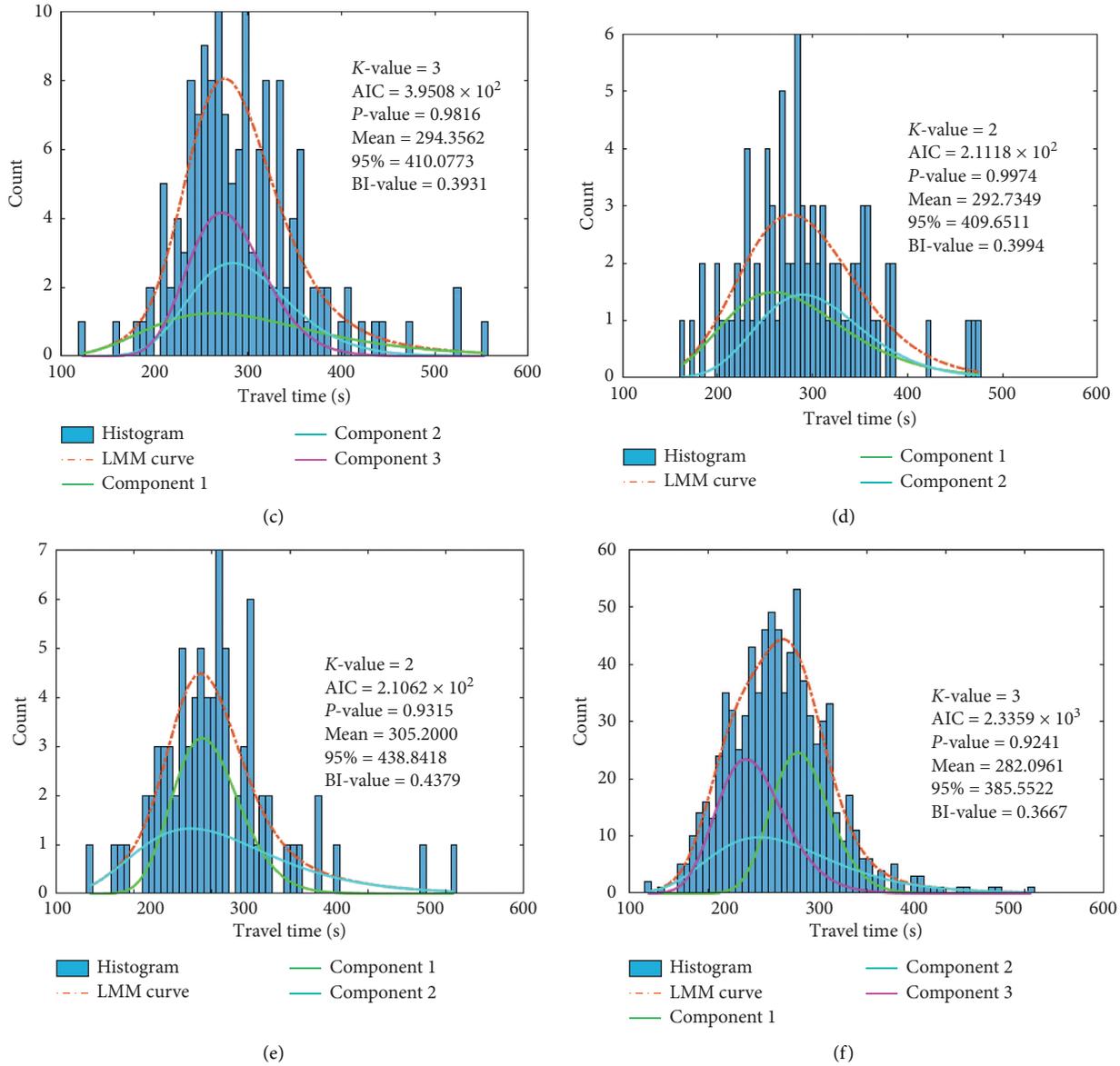


FIGURE 3: Fitting examples of LMM from Harrison Ave (ID: 11) to Huntington Ave (ID: 10): Part A: (a) normal and (b) light rain. Part B: (c) moderate rain and (d) heavy rain. Part C: (e) extreme rain and (f) rain.

variation between the light and the moderate rain condition. Notwithstanding, the null hypothesis is true when testing significance is in the 3-column source (P value > 0.05), which supports the previous analysis that there is no significant difference between moderate rain, heavy rain, and extreme rain.

The finding reveals that drivers are more sensitive to the change from light rain to moderate rain. While further investigation on driver behavior may be needed to fully explain this phenomenon fully, this finding is undoubtedly helpful in conducting more detailed and in-depth O-D-based TTR analysis.

4.3. Other Findings regarding O-D-Based TTR. Another notable finding from this research is that the O-D-based TTR tends to improve when ATT is longer, regardless of

the normal and the rainy weather. It was explicitly recognized from Table 4 (column 3 to column 8) that four out of six columns in the median and five out of six columns in the trimmed mean both illustrate that the O-D-based TTR increased along with the increase of average travel time.

Further investigation was conducted against the O-D pairs with the same ATT. The authors calculated the interquartile range of each travel time range, as illustrated in Table 7 and Figure 4 (the areas of the rectangle). It can be found that the areas of the rectangle are shrinking from Figures 4(a)–4(e). Referring to the columns in Table 7, two out of six columns follow this trend. For the rest of the columns, there is only one exception in each column (e.g., the last cell in column 3). The results demonstrate that the consistency of the O-D-based TTR in different

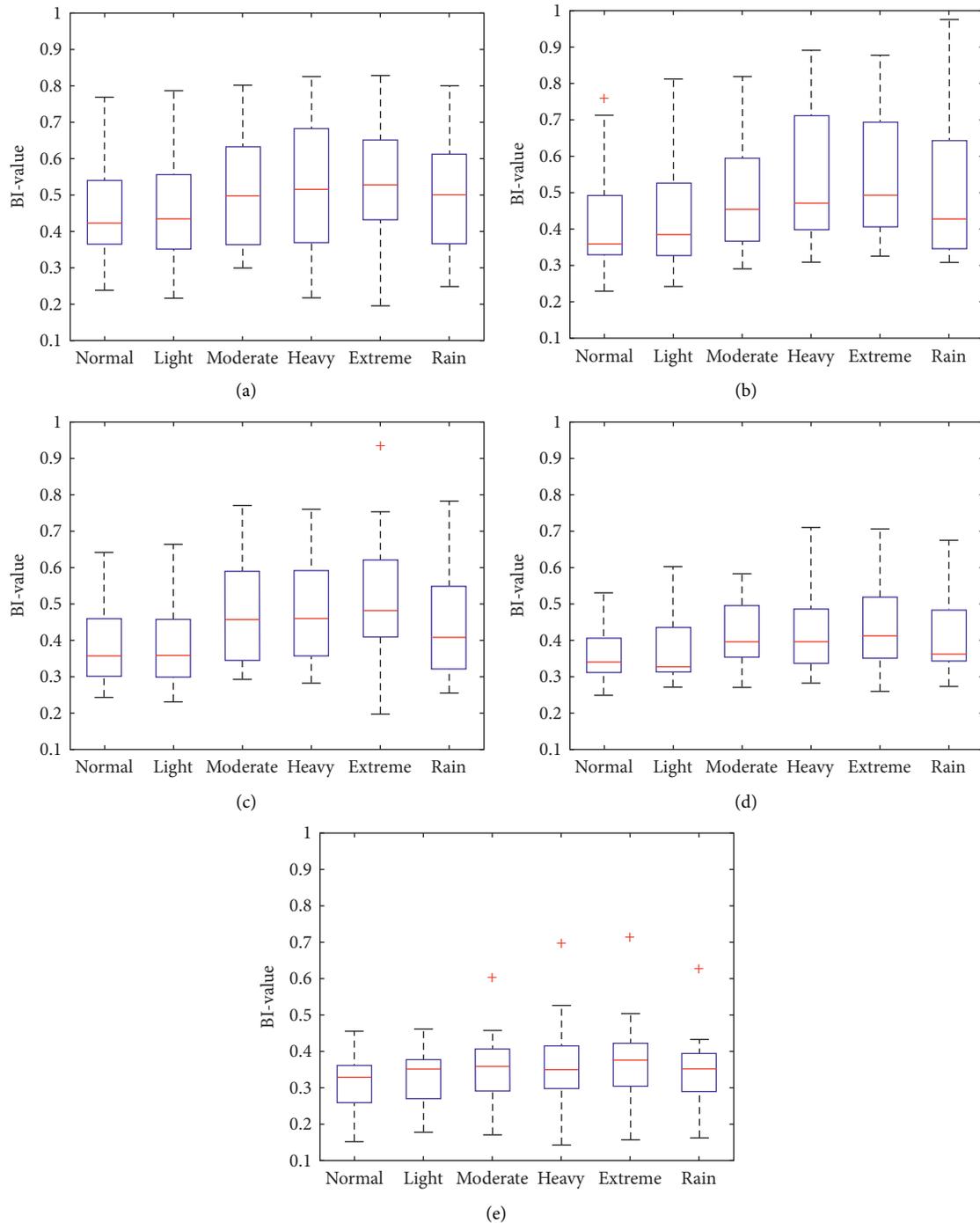


FIGURE 4: BI-value boxplots under six conditions within five travel time ranges. (a) Average O-D-based travel times within 0–5 mins. (b) Average O-D-based travel times within 5–10 mins. (c) Average O-D-based travel times within 10–15 mins. (d) Average O-D-based travel times within 15–20 mins. (e) Average O-D-based travel times within 20–25 mins.

O-D pairs with similar ATT range tends to improve as ATT gets longer.

Authors speculate that this may be attributable in large part to travel time fluctuations which have a decreasing effect on the longer ATT. For instance, the one-minute fluctuation exerts a greater influence on the O-D pair with a five-minute ATT than that with a ten-minute ATT. In practice, knowing this trend may significantly improve the accuracy of TTR prediction.

4.4. A Counterintuitive Phenomenon. Based on the statistics in Table 5, a so-called counterintuitive phenomenon (e.g., [10, 25]) was also found in this study, which is the O-D-based TTR which was improved under the rainy weather (the global probability is 16%). Although the negative effect of the rain on TTR is still the dominating conclusion considering the low probability of the positive effect, this phenomenon remains an issue to be investigated. According to [20], rain likely affects the effects of

Step 1:
Initialize K -value and LMM.

Step 2 (E step):
Calculate the probability γ_{ik} of each travel time t_i (sample size N) belonging to each lognormal distribution according to (i).
(i) $\gamma_{ik} = (\omega_k L(t_i/\mu_k, \sigma_k) / \sum_{k=1}^K \omega_k L(t_i/\mu_k, \sigma_k)), \quad i = 1, 2, \dots, N, k = 1, 2, \dots, K.$

Step 3 (M step):
Update all parameters in LMM according to (ii)–(iv).
(ii) $\hat{\omega}_k = (\sum_{i=1}^N \gamma_{ik} / N), \quad k = 1, 2, \dots, K.$
(iii) $\hat{\mu}_k = (\sum_{i=1}^N \gamma_{ik} \ln(t_i) / \sum_{i=1}^N \gamma_{ik}), \quad k = 1, 2, \dots, K.$
(iv) $\hat{\sigma}_k^2 = (\sum_{i=1}^N \gamma_{ik} (\ln(t_i) - \hat{\mu}_k) (\ln(t_i) - \hat{\mu}_k)^T / \sum_{i=1}^N \gamma_{ik}), \quad k = 1, 2, \dots, K.$

Step 4:
Repeat until convergence.

ALGORITHM 1: EM algorithm for estimating LMM.

TABLE 4: Summary of median and trimmed mean of BI-values and BI variation ratio under six weather conditions.

Location measure	ATT range (min)	Normal	Rainfall intensity (BI variation ratio)				Rain
			Light rain ((L-N)/ N * 100%)	Moderate rain ((M-N)/ N * 100%)	Heavy rain ((H-N)/ N * 100%)	Extreme rain ((E-N)/ N * 100%)	
Median of BI-values	0–5	0.4219	0.4342 (2.9%)	0.4985 (18.2%)	0.5157 (22.2%)	0.5274 (25.0%)	0.4997
	5–10	0.3586	0.3845 (7.2%)	0.4544 (26.7%)	0.4715 (31.5%)	0.4933 (37.6%)	0.4268
	10–15	0.3568	0.3586 (5.0%)	0.4573 (28.2%)	0.4598 (28.9%)	0.4823 (35.2%)	0.4080
	15–20	0.3399	0.3265 (–3.9%)	0.3956 (16.4%)	0.3962 (16.6%)	0.4113 (21.0%)	0.3604
	20–25	0.3287	0.3513 (6.9%)	0.3588 (9.2%)	0.3489 (6.1%)	0.3750 (14.1%)	0.3515
	Mean	0.3612	0.3710 (2.7%)	0.4329 (19.9%)	0.4384 (21.4%)	0.4579 (26.8%)	0.4092
Trimmed mean of BI-values	0–5	0.4413	0.4555 (3.2%)	0.5078 (15.1%)	0.5205 (17.9%)	0.5476 (24.1%)	0.4990
	5–10	0.4098	0.4351 (6.2%)	0.5047 (23.2%)	0.5399 (31.7%)	0.5376 (31.2%)	0.4809
	10–15	0.3754	0.3793 (1.0%)	0.4644 (23.7%)	0.4866 (29.6%)	0.5030 (34.0%)	0.4429
	15–20	0.3580	0.3675 (2.7%)	0.4151 (15.9%)	0.4121 (15.1%)	0.4304 (20.2%)	0.4007
	20–25	0.3190	0.3291 (3.2%)	0.3519 (10.3%)	0.3532 (10.7%)	0.3697 (15.9%)	0.3448
	Mean	0.3807	0.3933 (3.3%)	0.4488 (17.9%)	0.4625 (21.5%)	0.4777 (25.5%)	0.4337

N, L, M, H, and E indicate normal, light rain, moderate rain, heavy rain, and extreme rain, respectively.

TABLE 5: Summary of O-D number of higher reliability than normal and increasing reduction impact with rainfall increase.

ATT range (min)	O-D number of higher reliability than normal					Rain	O-D number of increasing negative impact with rainfall increase
	Light rain	Moderate rain	Heavy rain	Extreme rain	Sum/sample size		
0–5	8	1	4	2	15/80	3	7
5–10	3	1	1	1	6/80	2	8
10–15	10	1	1	2	14/80	1	6
15–20	7	0	3	2	12/80	1	7
20–25	8	2	5	2	17/80	2	7
Sum/sample size	36/100	5/100	14/100	9/100	64/400	9/100	35/100
Rate	36%	5%	14%	9%	16%	9%	35%

TABLE 6: One-way ANOVA test for BI variation ratio between varying rainfall intensities.

Source	SS	df	MS	F	P value
2 columns (light and moderate)	1.2280	1	1.2280	80.82	3.5874e-16
Error	2.7047	178			
Total	3.93263	179			
3 columns (from moderate to extreme)	0.2079	2	0.1040	2.9	0.0568
Error	9.5750	267	0.0359		
Total	9.7829	269			
4 columns (from light to extreme)	2.9091	3	0.9697	33.37	4.9919e-19
Error	10.3457	356	0.0291		
Total	13.2548	359			

TABLE 7: Summary of interquartile range of BI-values under six weather conditions.

Location measure	ATT range (min)	Normal (75th–25th)	Rainfall intensity (75th–25th)				Rain (75th–25th)
			Light rain	Moderate rain	Heavy rain	Extreme rain	
Interquartile range (IQR) of BI-values	0–5	0.1771 (0.5206–0.3635)	0.2043 (0.5543–0.3500)	0.2687 (0.6330–0.3643)	0.3138 (0.6829–0.3691)	0.2188 (0.6511–0.4323)	0.2450 (0.6119–0.3669)
	5–10	0.1645 (0.4929–0.3284)	0.1975 (0.5253–0.3278)	0.2278 (0.5847–0.3569)	0.3134 (0.7118–0.3984)	0.2881 (0.6939–0.4058)	0.2987 (0.6441–0.3454)
	10–15	0.1606 (0.4600–0.2994)	0.1612 (0.4586–0.2974)	0.2454 (0.5907–0.3453)	0.2342 (0.5910–0.3568)	0.2112 (0.6212–0.4100)	0.2292 (0.5487–0.3195)
	15–20	0.0944 (0.4065–0.3121)	0.1222 (0.4354–0.3132)	0.1432 (0.4952–0.3520)	0.1509 (0.4865–0.3356)	0.1698 (0.5195–0.3497)	0.1407 (0.4842–0.3435)
	20–25	0.1006 (0.3593–0.2587)	0.1083 (0.3786–0.2703)	0.1157 (0.4168–0.3011)	0.1173 (0.4202–0.3029)	0.1201 (0.4248–0.3047)	0.1056 (0.3941–0.2885)

some other substantial contributing factors for some specific O-D pairs (e.g., O-D with bottlenecks). Therefore, future studies should be focused on data with multiple contributing factors such as weather and bottlenecks.

5. Conclusions

This research uses open-source data to study the effects of varying rainfall intensity on O-D-based travel time reliability. The intensity covers light rain (0 and 10.0 mm/24 h), moderate rain (10.0 and 24.9 mm/24 h), heavy rain (25 and 50.0 mm/24 h), and extreme rain (>50.0 mm/24 h). An orithm based on the lognormal mixture model was adopted for analyzing the probability distribution functions of the O-D-based travel time data. Then the buffer index, the three location measures, and the one-way analysis of variance were used for detailed analysis.

In general, rain lowers O-D-based travel time reliability, and the negative impact grows with the increase in rainfall intensity. With respect to the abnormal phenomenon mentioned in [25], it is restricted in low probability from massive results in the study, which cannot be a general conclusion but deserve investigation profoundly. The study also confirmed the existence of the so-called counterintuitive phenomenon mentioned in previous work [25] that, in some cases, TTR may be enhanced in rainy weather. Particularly, we discovered that O-D-based TTR was more sensitive when rainfall intensity changes from light to moderate but less notable when changes are among other categories such as no rain to light rain and moderate rain to heavy rain. The study also demonstrates that the O-D-based TTR in different O-D pairs with a similar ATT range tends to improve as ATT gets longer.

This study contributed to disclose the characteristics of the O-D-based TTR under the varying rainfall with the open-access data. This study is helpful in enhancing current applications by providing more accurate O-D-based travel time information under rain conditions; for example, the trend that consistency of the O-D-based TTR tends to improve with ATT increase can help to improve the accuracy of TTR prediction, when missing enough travel time information. Meanwhile, the research is conducive to promote the use of publicly available data in such investigations so that the results are verifiable, and the studies are sustainable.

This paper only focuses on the impact of rain, but there are a lot of more deserving further investigations along the line of O-D-based TTR analysis. For future work, the impact of other weather events and the combined effects of weather and other factors such as work zones will be the focus.

Data Availability

The data used to support the findings of this study are available from the Uber Movement website (<https://movement.uber.com>) and Weather Underground website (<https://www.wunderground.com/history>).

Conflicts of Interest

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

Acknowledgments

The study was supported by the Technology Project of Shaanxi Transportation Department (Grant no. 15-39R), Special Fund for Basic Scientific Research of Central Colleges of Chang'an University (Grant no. 300102218409), and China Scholarship Council. The authors are also grateful for data retrieved from Uber Movement, (c) 2019 Uber Technologies, Inc., <https://movement.uber.com>, and Weather Underground website, <https://www.wunderground.com/history>.

References

- [1] Y. Lida, "Basic concepts and future directions of road network reliability analysis," *Journal of Advanced Transportation*, vol. 33, no. 2, pp. 125–134, 1999.
- [2] K. Lyman and R. L. Bertini, "Using travel time reliability measures to improve regional transportation planning and operations," *Transportation Research Record*, vol. 2046, pp. 1–10, 2008.
- [3] C. Chen, E. W. van Zwet, and P. Varaiya, "Travel-time reliability as a measure of service," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1855, no. 1, pp. 74–79, 2007.
- [4] N. Y. Tilahun and D. M. Levinson, "A moment of time: reliability in route choice using stated preference," *Journal of Intelligent Transportation Systems*, vol. 14, no. 3, pp. 179–187, 2010.
- [5] L. Gong and W. Fan, "Applying travel-time reliability measures in identifying and ranking recurrent freeway bottlenecks at the network level," *Journal of Transportation Engineering Part A: Systems*, vol. 143, no. 8, 2017.
- [6] H. X. Liu, W. Recker, and A. Chen, "Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 6, pp. 435–453, 2004.
- [7] A. Chepuri, J. Ramakrishnan, S. Arkatkar, G. Joshi, and S. S. Pulugurtha, "Examining travel time reliability-based performance indicators for bus routes using GPS-based bus trajectory data in India," *Journal of Transportation Engineering Part A: Systems*, vol. 144, no. 5, 2018.
- [8] C. Lu and J. Dong, "Estimating freeway travel time and its reliability using radar sensor data," *Transportmetrica B: Transport Dynamics*, vol. 6, no. 2, pp. 97–114, 2018.
- [9] F. Zheng, J. Li, H. van Zuylen, X. Liu, and H. Yang, "Urban travel time reliability at different traffic conditions," *Journal of Intelligent Transportation Systems*, vol. 22, no. 2, pp. 106–120, 2018.
- [10] S. I. Chien and K. K. Kolluri, "Evaluation of freeway travel time variability and reliability under adverse weather with TRANSMIT data," *Journal of Civil Engineering and Architecture*, vol. 6, no. 1, pp. 1–11, 2012.
- [11] Z. Chen and W. Fan, "Data analytics approach for travel time reliability pattern analysis and prediction," *Journal of Modern Transportation*, vol. 27, no. 4, pp. 250–265, 2019.
- [12] X. Zhang and M. Chen, "Quantifying the impact of weather events on travel time and reliability," *Journal of Advanced Transportation*, vol. 2019, Article ID 8203081, 9 pages, 2019.
- [13] S. Yang, C. An, Y.-J. Wu, and J. Xia, "Origin-destination-based travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2643, no. 1, pp. 139–159, 2017.
- [14] S. H. Kim and J.-H. Chung, "Exploration on origin-destination-based travel time variability: insights from Seoul metropolitan area," *Journal of Transport Geography*, vol. 70, pp. 104–113, 2018.
- [15] J. Dong, H. S. Mahmassani, and C. C. Lu, "How reliable is this route? predictive travel time and reliability for anticipatory traveler information systems," *Transportation Research Record*, vol. 1980, pp. 117–125, 2006.
- [16] J. Kim and H. S. Mahmassani, "Compound gamma representation for modeling travel time variability in a traffic network," *Transportation Research Part B: Methodological*, vol. 80, pp. 40–63, 2015.
- [17] J. Kim and H. S. Mahmassani, "A finite mixture model of vehicle-to-vehicle and day-to-day variability of traffic network travel times," *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 83–97, 2014.
- [18] H. S. Mahmassani, T. Hou, and J. Dong, "Characterizing travel time variability in vehicular traffic networks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2315, no. 1, pp. 141–152, 2012.
- [19] M. A. P. Taylor, "Travel through time: the story of research on travel time reliability," *Transportmetrica B: Transport Dynamics*, vol. 1, no. 3, pp. 174–194, 2013.
- [20] J. Kwon, T. Barkley, R. Hranac, K. Petty, and N. Compin, "Decomposition of travel time reliability into various sources," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2229, no. 1, pp. 28–33, 2011.
- [21] H. Abdulsattar, A. Mostafizi, M. R. K. Siam, and H. Wang, "Measuring the impacts of connected vehicles on travel time reliability in a work zone environment: an agent-based approach," *Journal of Intelligent Transportation Systems*, vol. 24, no. 5, pp. 421–436, 2020.
- [22] L. C. Goodwin, "Weather impacts on arterial traffic flow," Report Prepared for the Road Weather Management Program, Mitretek Systems, Inc., Virginia, USA, 2002.
- [23] A. D. Stern, V. Shah, and L. Goodwin, *Analysis of Weather Impacts on Traffic Flow in Metropolitan Washington DC*, Institute of Transportation Engineers Meeting & Exhibit, Washington, DC, USA, 2003.
- [24] H. Caceres, H. Hwang, and Q. He, "Estimating freeway route travel time distributions with consideration to time-of-day, inclement weather, and traffic incidents," *Journal of Advanced Transportation*, vol. 50, no. 6, pp. 967–987, 2016.
- [25] C. Kanga and M. A. Yazıcı, "Temporal and weather related variation patterns of urban travel time: considerations and caveats for value of travel time, value of variability, and mode choice studies," *Transportation Research Part C: Emerging Technologies*, vol. 45, pp. 4–16, 2014.
- [26] M. Martchouk, F. Mannering, and D. Bullock, "Analysis of freeway travel time variability using bluetooth detection," *Journal of Transportation Engineering*, vol. 137, no. 10, pp. 697–704, 2011.
- [27] A. Higatani, T. Kitazawa, J. Tanabe, Y. Suga, R. Sekhar, and Y. Asakura, "Empirical analysis of travel time reliability measures in hanshin expressway network," *Journal of Intelligent Transportation Systems*, vol. 13, no. 1, pp. 28–38, 2009.
- [28] C. Lu, J. Dong, A. Houchin, and C. Liu, "Incorporating the standstill distance and time headway distributions into freeway car-following models and an application to estimating freeway travel time reliability," *Journal of Intelligent Transportation Systems*, pp. 1–20, 2019.
- [29] M. Saberi and R. L. Bertini, "Beyond corridor reliability measures: analysis of freeway travel time reliability at the segment level for hot spot identification," in *Proceedings of*

- 89th Annual Meeting of the Trans-Portation Research Board*, Washington, DC, USA, January 2010.
- [30] H. Tu, J. W. C. Van Lint, and H. J. Van Zuylen, "Impact of traffic flow on travel time variability of freeway corridors," *Transportation Research Record*, vol. 1993, no. 1, pp. 59–66, 2007.
 - [31] E. B. Emam and H. Al-Deek, "Evaluation method and influence factors of network travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board of the National Academies*, vol. 1959, pp. 140–150, 2006.
 - [32] W. Pu, "Analytic relationships between travel time reliability measures," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2254, no. 1, pp. 122–130, 2012.
 - [33] K. Sohn and D. Kim, "Statistical model for forecasting link travel time variability," *Journal of Transportation Engineering*, vol. 135, no. 7, pp. 440–453, 2009.
 - [34] S. Yang and Y.-J. Wu, "Mixture models for fitting freeway travel time distributions and measuring travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2594, no. 1, pp. 95–106, 2016.
 - [35] F. Guo, H. Rakha, and S. Park, "Multistate model for travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2188, no. 1, pp. 46–54, 2011.
 - [36] F. Guo, Q. Li, and H. Rakha, "Multistate travel time reliability models with skewed component distributions," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2315, no. 1, pp. 47–53, 2013.
 - [37] M. Agarwal, T. H. Maze, and R. R. Souleyrette, "Impacts of weather on urban freeway traffic flow characteristics and facility capacity," in *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*, Ames, Iowa, August 2005.
 - [38] N. E. Lownes and R. B. Machemehl, "Vissim: a multi-parameter sensitivity analysis," in *Proceedings of the 2006 Winter Simulation Conference*, pp. 1406–1413, California, December 2006.