

Article

On the Necessity and Effects of Considering Correlated Stochastic Speeds in Shortest Path Problems Under Sustainable Environments

Dongqing Zhang ^{1,2} and Zhaoxia Guo ^{1,*}

¹ Business School, Sichuan University, Chengdu 610065, China; dongqingzhang@u.nus.edu

² Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore 117576, Singapore

* Correspondence: zxguo@scu.edu.cn; Tel.: +86-28-8541-7867

Received: 19 November 2019; Accepted: 24 December 2019; Published: 27 December 2019



Abstract: This research addresses how the stochasticity and correlation of travel speeds affect the shortest path solutions in sustainable environments. We consider a shortest path problem with the objective function of minimizing a linear combination of the mean and standard deviation of carbon emissions. By adjusting the proportion of the standard deviation in the objective function, the effects of speed stochasticity and correlation are studied under different preferences of the decision-makers on the fluctuations of carbon emissions. Based on 102-day real speed data from the Los Angeles freeway network, this research conducts extensive numerical experiments on 200 randomly chosen origin-destination pairs. Experimental results demonstrate the necessity of considering speed stochasticity and correlation, especially when the standard deviation of carbon emissions takes a large proportion in the objective function. As the weight of the standard deviation in the objective function increases from 0 to 1.5, the reduction of emission objective values increases from 0.03% to 0.13% by considering speed stochasticity, and increases from 0.02% to 0.20% by considering speed correlation. Taking the city Los Angeles with about 2361 taxis and about 525,945 passenger orders in January 2017 as an example, 0.03% and 0.02% reductions respond to about 3156 kg and 2630 kg carbon emission, respectively.

Keywords: sustainability; shortest path problem; correlated stochastic speeds; carbon emission objectives

1. Introduction

The transportation sector is a significant emitter of greenhouse gases, which accounts for 14% of global greenhouse gas emissions [1]. With an increasing worldwide concern for the environmental hazards of greenhouse gases, many researchers are working to develop green logistic policies to shift the transportation schemes to be more sustainable with fewer negative environmental impacts [2–4]. In this context, many operational research methods have been taken to green logistics to address some routing problems in sustainable environments [5–8], and the fundamental routing problem that determines the travel path from an origin node to a destination node in a road network is generally modeled as the shortest path problem with sustainable objective functions, such as minimum carbon emissions [9] and minimum fuel consumption [10,11].

In the literature, most of papers on the shortest path problems with sustainable objective functions are studied with the assumption of deterministic traffic environments [9,12,13]. However, traffic networks have inherent uncertainties that are caused by random disturbances such as crashes, vehicle breakdowns,

bad weather, and maintenance activities [14]. Although there already are a large number of shortest path papers considering stochastic link travel times or speeds, most of them aim to find the travel path with minimum expected travel time [15–18] or maximum travel time reliability [19–21]. Research on the shortest path problem with both stochastic traffic environment and sustainable objective functions is still in its infancy. To the best of the authors' knowledge, only Ehmke et al. [22] has studied the shortest path problem of minimizing the expected carbon emissions in a road network with stochastic link travel times. To solve this problem, they propose a path-averaging method and an arc-averaging method by modifying the classic A* algorithm [23]. Similar to many stochastic shortest path papers that consider travel time-related objective functions [17,19,21], Ehmke et al. [22] also consider the correlation of link travel times. Specifically, they set the distributions of link travel times as time dependent, and then use real time-dependent speed data to calculate the link travel times and the expected carbon emissions of travel paths. The calculation of travel times through some long-distance road links may involve the travel speeds in multiple time periods, and thus different starting times may need the travel speeds in the same periods while traveling such road links. Therefore, the time dependence of travel times implies the temporal correlation between link travel times [15], i.e., the travel time on one road link in different time periods are correlated. Later, Ehmke et al. extend this work to the stochastic vehicle routing problem in real road networks where the determination of a specific travel path between two customers is modeled as a shortest path problem with stochastic link travel times and the objective of minimizing the time- and fuel-related costs [11]. The similar stochastic vehicle routing problem has been studied by Huang et al. [10] as well. Similar to paper [22], Ehmke et al. [11] and Huang et al. [10] also consider the correlation of link travel times or speeds. Ehmke et al. [11] continue to implicitly consider the temporal correlation by using real time-dependent speed data, as in their early paper [22]. Huang et al. [10] first classify the road links into expressways, arterial, and residential roads, and then set the average speeds of road links in each category to the same values in each time period, which implies the spatial correlation of travel speeds, i.e., the travel speeds on different road links are correlated. Moreover, Huang et al. also use time-dependent speed data to calculate the travel time of each road link, such as in papers [11,22], which means the temporal correlation is considered as well. However, in real-world traffic networks, due to the propagation and uncertainties of traffic flow, link travel times or speeds have much complex correlation in both temporal and spatial dimensions, which are much more complicated than the correlation considered by Ehmke et al. [11], Ehmke et al. [22] and Huang et al. [10]. Interested readers can refer to papers [24–26] for more details.

Modeling the stochasticity and correlation of link travel speeds in real road networks is highly complicated because there are very high-dimensional dependent random vector [27]. For example, a stochastic time-dependent road network with 200 road links and 10 time periods implies 2000 random speed variables and 1,999,000 ($2000 \times 1999 \div 2$) distinct correlations. If there are sufficient time-dependent historical speed data, the data can be used to account for the complex stochasticity and correlation of link travel speeds, as the papers [17,19] do. However, in most cases, it is difficult to obtain real speed data in the underlying road networks. Therefore, many papers directly assume that all link travel speeds fit certain multivariate distributions, and then randomly sample some scenarios to model the correlated stochastic speeds [14,28,29]. However, the random sampling method usually needs a large number of samples (scenarios) to restore the stochasticity and correlation of the random speed variables, which significantly increases the difficulty of solving the shortest path problems.

If the stochasticity and correlation of link travel speeds are not considered, as is the case in many shortest path papers [23,30,31], the difficulty of modeling and solving the shortest path problems under sustainable environments can be significantly reduced. However, does this make the resulting travel paths different from the actual optimal travel paths in real stochastic and correlated road networks? In other words, will the stochasticity and correlation of link travel speeds affect the obtained optimal travel path

solutions or the corresponding carbon emission values. This is the focus of this research. In the literature, some researchers have demonstrated the significance of considering the spatial and temporal correlation of stochastic link speeds in shortest path problems with time-related objective functions [14,28]. For example, the two papers [14,28] find that the spatial and temporal correlations could affect the optimal travel paths, and the impact is associated with the levels of correlation. However, these two papers have not considered the sustainable objective functions. Therefore, the necessity and effects of considering correlated stochastic speeds in shortest path problems under a sustainable objective function, which is the problem studied in this paper, is still unclear.

In this paper, the objective function of the studied shortest path problem is set to minimize a linear combination of the mean and standard deviation of carbon emissions, and the parameter θ is defined as the weight of the standard deviation. By adjusting the value of θ , we can study the effects of speed stochasticity and correlation under different preferences of the decision-makers on the fluctuations of carbon emissions. For example, when θ is equal to zero, the effects on the mean of carbon emissions can be studied. This paper does not take the minimum standard deviation of carbon emissions as an optimized objective function because it is unreasonable to only consider the standard deviation without the mean of carbon emissions. For example, when minimizing the standard deviation of carbon emissions, the optimal travel path would take a longer detour path to the suburbs, due to the low fluctuations of travel speeds in the detour path. However, taking a detour path to the suburbs would result in more carbon emissions, which is contrary to the original intention of the decision-makers in reality.

This paper conducts extensive numerical experiments to study the effects of speed stochasticity and correlation on the optimal travel paths and the corresponding means and standard deviations of carbon emissions. The test data are 102-day real time-dependent link travel speeds from the Los Angeles freeway network. Specifically, to investigate the effects of speed stochasticity, this paper compares the optimal travel paths from the stochastic problem with stochastic scenario sets and the deterministic problem with a deterministic scenario. On the other hand, to investigate the effects of speed correlation, this paper compares the optimal travel paths from the correlated problem with stochastic correlated scenario sets and the uncorrelated problem with stochastic uncorrelated scenario sets. The path-averaging method proposed by Ehmke et al. [22] is applied to find the optimal travel paths between an origin-destination (O-D) pair. We thus contribute to the research on shortest path problems in sustainable environments by providing some managerial and modeling insights into whether and how the stochasticity and correlation of link travel speeds affect the optimal travel path solutions.

The rest of this paper is outlined as follows. The studied shortest path problem is defined in Section 2. Section 3 introduces the test case and experimental settings. Section 4 shows the experimental results and analyses the effects of speed stochasticity and correlation. Section 5 concludes this paper.

2. Problem Definition

The studied shortest path problem is defined in a stochastic time-dependent road network $G = (V, E, T, \zeta)$. V is the set of nodes representing the road intersections and $|V| = N$. E is the set of edges representing the road links and $|E| = M$. T is the set of time periods of interest and $|T| = K$. The travel speed on a link (i, j) in a time period t is represented as a non-negative random variable \tilde{v}_{ijt} . Then, the travel speeds on all links in all time periods are jointly distributed random variables that are represented by a stochastic scenario set ζ . The scenario set ζ includes a finite number of scenarios ζ_s (or called support points in papers [14,32]), $s = 1, \dots, S$, each of which is defined as a vector of values that all random speed variables can take. Each scenario has a probability p_s and $\sum_{s=1}^S p_s = 1$. Take a small network with four nodes, four edges, and two periods, shown in Figure 1, as an example. Table 1 shows an example of the

scenario set ζ for the small network, which includes 8 (4×2) random speed variables and 10 stochastic scenarios. The probability of each scenario is 0.1.

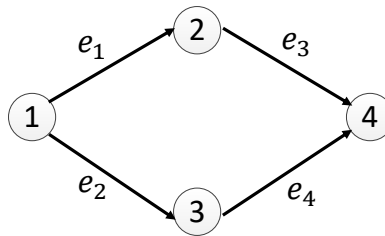


Figure 1. A small network.

Table 1. An example of the scenario set ζ for the link travel speeds (unit:km/h) in the small network.

Period	Link	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5	ζ_6	ζ_7	ζ_8	ζ_9	ζ_{10}
1	e_1	57	54	52	48	54	54	58	50	52	59
	e_2	64	58	60	61	61	66	64	60	61	61
	e_3	64	58	60	61	61	66	64	60	61	61
	e_4	69	69	66	66	67	68	68	67	66	69
2	e_1	59	57	51	55	56	55	56	53	55	57
	e_2	65	61	61	60	61	64	63	62	62	62
	e_3	68	68	66	66	68	67	69	69	67	68
	e_4	54	50	49	49	53	61	60	57	57	58

The investigated shortest path problem aims to find the travel path with the minimum value of a linear combination of the emission mean μ and standard deviation σ , as formulated in Equation (1). Let θ denote a specified weight factor representing the aversion to carbon emission variability, and e^s denote the carbon emission (unit: gram) of a travel path in the s^{th} scenario. This paper uses the well-known MEET model [33] to calculate the carbon emission e^s , which is formulated in Equation (2). In the model, the gross vehicle weight is set to 3.5–7.5 tons, and then $K = 110, a = 0, b = 0, c = 0.000375, d = 8702, e = 0$, and $f = 0$. d_{ij} is the length of the link from node i to node j , and x_{ijt} is the decision variable. If the vehicle travels from node i to node j directly in time period t , x_{ijt} is equal to 1, otherwise it is 0. The number 1000 in the μ and σ calculation equations are used to convert the unit of carbon emissions from gram to kilogram to facilitate the presentation of numerical results. When $\theta = 0$, the objective function is to minimize the mean of carbon emissions.

$$F = \mu + \theta \cdot \sigma \tag{1}$$

with $\mu = \frac{\sum_{s=1}^S (e^s/1000)}{S}$ and $\sigma = \sqrt{\frac{\sum_{s=1}^S (e^s/1000 - \mu)^2}{S}}$.

$$e^s = \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^K (K + a \cdot v_{ijt}^s + b \cdot v_{ijt}^s{}^2 + c \cdot v_{ijt}^s{}^3 + d/v_{ijt}^s + e/v_{ijt}^s{}^2 + f/v_{ijt}^s{}^3) \cdot d_{ij} \cdot x_{ijt}, \forall i \neq j, s = 1, \dots, S \tag{2}$$

3. Test Case and Settings

This paper investigates the effects of speed stochasticity and correlation on the optimal travel path solutions by conducting extensive numerical experiments using real speed data from the Los Angeles freeway network. This section first introduces the processes of collecting and processing the real speed data, and then introduces the settings of following numerical experiments.

3.1. Data Collection and Processing

Numerical experiments need a large number of real time-dependent travel speed values on each road link in each time period to form the stochastic scenario set ζ that can reflect the real stochasticity and correlation of link travel speeds. This paper uses the ready-made time-dependent speed data in the Los Angeles freeway network provided by the Caltrans Performance Measurement System (PeMS) at <http://pems.dot.ca.gov/>. PeMS is a public intelligent transportation system developed by the California Department of Transportation to collect, store, and process traffic data for all California freeways in real-time. Figure 2 shows the Los Angeles freeway network that consists of 3417 speed detection stations. Each station can detect the travel speed of the vehicles at its location. From PeMS, we download all speed readings of the 3417 speed detection stations over 102 days, from 1 May to 22 September in 2017 excluding weekends and holidays. The time period of downloaded speed data is 5 minutes. Detailed procedures of collecting and processing the speed data in PeMS can be found in paper [34].

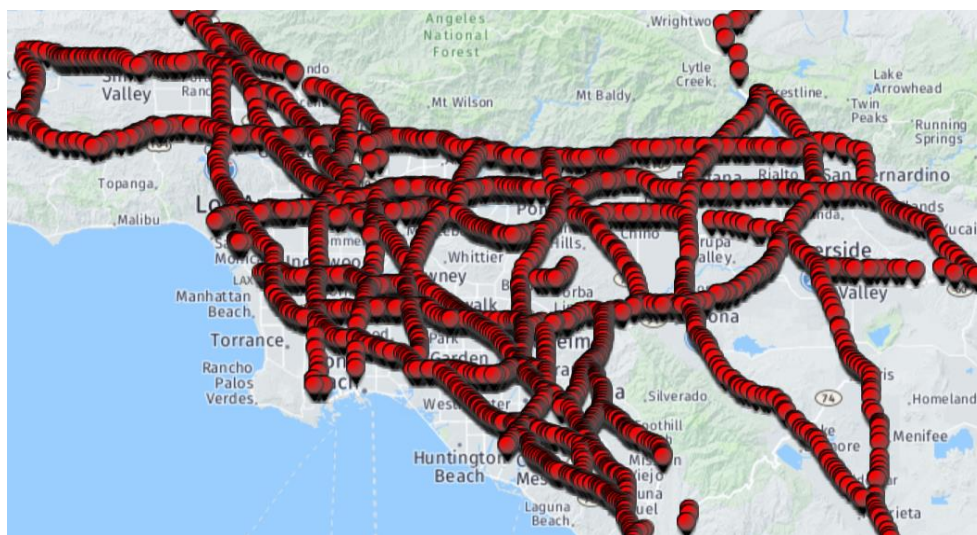


Figure 2. Los Angeles freeway network in which each red drop represents a speed detection station. (Figure courtesy of the Caltrans Performance Measurement System (PeMS))

After downloading the speed data of all the detection stations, this paper needs to aggregate the speed values of multiple stations to get the travel speed of an entire road link, because the road links between two road intersections usually contain multiple detection stations, as shown in Figure 2. By analysing the downloaded speed data, it is found that the speeds of some stations on the same road link differ greatly in the same time period. This is understandable. For some long-distance road links, the traffic conditions and surroundings around different stations vary a lot, which results in a large difference in the speeds of different stations. However, when averaging the speeds of all stations to get the speed of an entire road link, such large differences would cause the average speed to deviate from the actual speed of the link. Therefore, in order to reduce this deviation, this paper adaptively divides the road links between two road intersections into several shorter road links to ensure that the speed difference of all stations on each road link is within a specific range. Specifically, the initial road network data are first obtained from the digital map, including the longitude and latitude of each road intersection, the connection relationship of all road intersections, and the length of each original road link between two road intersections. Furthermore, the IDs of stations on each original road link are downloaded from PeMS. Then, we set the time period 10:00–10:05 with stable traffic flow as the research period. For each original road link, we first find the 102-day speeds of all the stations on this link in the research period and then calculate the variance of these

speed values. If the obtained variance exceeds the maximum allowable variance value, the original road link is divided into several shorter road links to ensure that the speed variance of all stations on each link is no more than the maximum allowable variance value. The division position is set to the middle of the link between two adjacent speed detection stations. By adjusting the maximum allowable variance value, we get a reasonable road network with 138 nodes and 438 links. Finally, the speeds of all stations on each link are averaged to be the travel speeds on each road link in each time period, for the following experimental uses.

3.2. Experimental Settings

This paper conducts extensive numerical experiments to investigate the effects of speed stochasticity and correlation on the optimal travel path solutions with a sustainable objective function. A total of 200 different O-D pairs are randomly selected as the test instances. The planning time horizon is the two hours from 8 to 10 a.m., including 24 5-min time periods, and the departure times in all instances are set to 8 a.m. Therefore, there are a total of 10,512 (438×24) random speed variables in the stochastic time-dependent road network.

To investigate the effects of speed stochasticity, this paper compares the optimal travel paths from the stochastic problem (SP) and the deterministic problem (DP). Figure 3 presents the flowchart for getting the optimal travel path of an O-D pair in the SP and DP. The SP uses all 102-day real speed data as the stochastic scenario set, and the DP uses the deterministic scenario ($S = 1$) where the travel speed on each link in each time period is the average speed of the link in that period over 102 days. Then, the path-averaging algorithm proposed by Ehmke et al. [22] is used to calculate the optimal travel path of an O-D pair. The path-averaging algorithm is a path-based label setting A*-variant that originally solves the expected emissions-minimizing path problem considering stochastic time-dependent link travel speeds, in which the carbon emissions are calculated using the MEET model. Although Ehmke et al. [22] have not considered the standard deviation of carbon emissions in their objective function, the path-averaging algorithm can still be employed to solve the investigated stochastic shortest path problem. The reason is that Ehmke et al. [22] used stochastic scenarios (i.e., samples) to evaluate the emissions of a given path as this paper does, and thus their algorithm can both capture the expected carbon emission (μ in Equation (2)) as well as the variation of emissions (σ in Equation (2)). Please refer to [22] for the detail of the path-averaging algorithm.

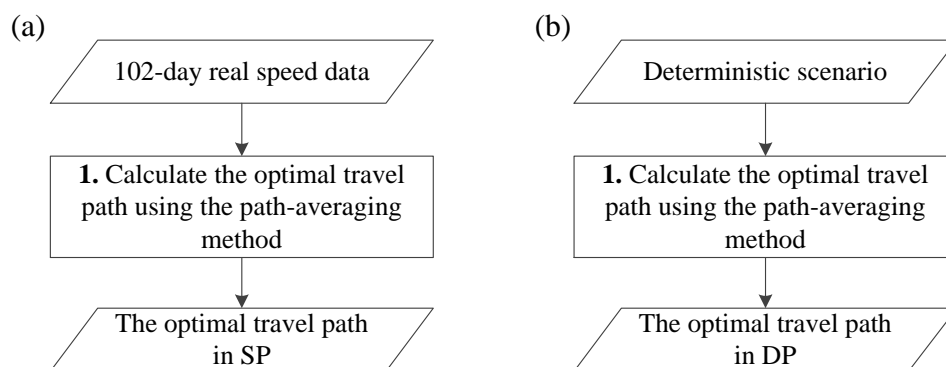


Figure 3. Flowchart for getting the optimal travel path of an O-D pair in (a) a stochastic problem (SP) and (b) a deterministic problem (DP).

On the other hand, to investigate the effects of speed correlations, this research compares the optimal travel paths from the correlated problem (CP) and the uncorrelated problem (UnCP). Figure 4 presents

the flowchart for getting the optimal travel path of an O-D pair in the CP or UnCP. First, the Scenario Generation (SG) method proposed by Kaut and Wallace [35] is employed to generate S ($S = 10$ at the beginning) stochastic correlated scenario sets in the CP or S stochastic uncorrelated scenario sets in the UnCP based on all 102-day speed data. Specifically, the SG method requires the correlation matrix and the marginal distributions of all random speed variables as input. This paper calculates the Pearson correlation between all pairs of random speed variables based on the real speed data to form the correlation matrix. To get the marginal distributions, the Grubbs' test [36] is first implemented repeatedly to remove the speed outliers with a significance level of 0.05. Then, in the stochastic time-dependent road network, 49.57% (5211) of the speed variables fit log-normal distributions, and 50.43% (5301) fit bounded beta distributions. To generate a stochastic correlated scenario set in the CP, the input of the SG method are the obtained correlation matrix and the marginal distributions. While generating a stochastic uncorrelated scenario set in the UnCP, the correlations between all speed variables in the input correlation matrix are set to zeros.

Then, the stability level of the S scenarios is calculated using the stability test proposed by Guo et al. [27] to determine the number of stochastic scenarios required. The processes for calculating the stability level of S scenarios are shown as follows: generate $2m + 1$ scenario sets that have $S - m$, $S - m + 1$, \dots , $S + m - 1$, and $S + m$ stochastic scenarios ($m = 4$ in this paper), respectively; for each scenario set, find the optimal (or feasible) solution; then, for each found solution, calculate its objective function values based on the $2m + 1$ scenario sets, and find the biggest (F^+) and the smallest (F^-) objective function values; calculate the relative difference $(F^+ - F^-)/F^+$ for each solution; finally, the biggest relative difference of the $2m + 1$ solutions is taken as the stability level of the S scenarios. Then, if the obtained stability level is less than 2%, the S scenarios can be employed in this instance to achieve the desired evaluation stability; otherwise, the S scenarios cannot achieve stability and more scenarios are needed. Finally, based on the obtained S stochastic scenarios, the path-averaging algorithm is employed to calculate the optimal travel path of the O-D pair.

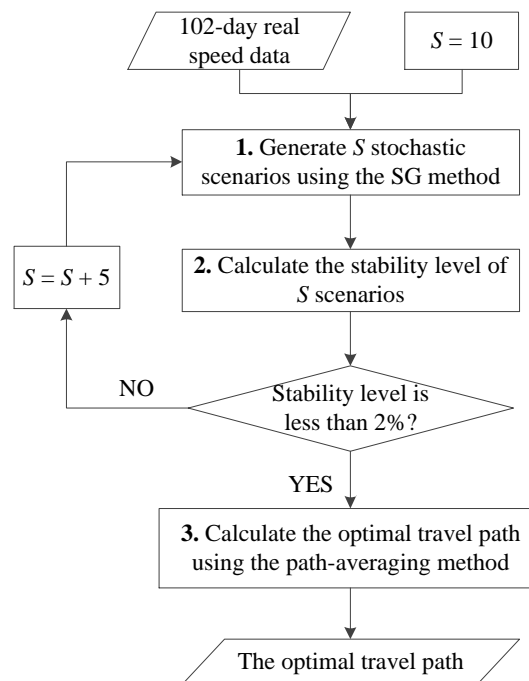


Figure 4. Flowchart for getting the optimal travel path of an O-D pair in a correlated problem (CP) or an uncorrelated problem (UnCP).

The experiments are performed on a laptop equipped with an Intel Core i7-8550U CPU @2.00 GHz and 16 GB RAM. All algorithms are coded and executed in MATLAB 2016b.

4. Experimental Results

This section shows the experimental results and analyses the effects of speed stochasticity and correlation on the optimal travel path solutions under sustainable environments.

4.1. Effects of Speed Stochasticity

This sub-section investigates the effects of speed stochasticity. Table 2 shows the performance comparison between the DP and the SP in terms of four different θ values. The result is the average of 200 test instances. In the table, the meanings of the μ , $\theta \cdot \sigma$, and F are consistent with their respective meanings in Equation (1). Taking the DP results as a benchmark, ‘Diff.’ shows the relative changes of the SP results. Table 3 compares the optimal paths of the SP and the DP. Please note that since the DP uses the deterministic scenario, the obtained objective values of optimal travel paths are not their actual objective values. Therefore, all emission results of the DP in this sub-section are obtained by re-evaluating the obtained optimal paths with all 102-day real speed data.

In Table 2, the SP generates the lower objective values F than the DP does for all θ values. Specifically, when $\theta = 0$, for an O-D pair, the SP can reduce the mean of carbon emissions by 0.03%. These reduced carbon emissions result because the SP generates better solutions than the DP does for 7 O-D pairs in Table 3. Although the 0.03% reduction looks very inconspicuous, when considering all travel demands of an entire city, the reduction in carbon emissions is still much significant. Let us take Los Angeles as an example. There are about 2361 taxis (not including the ride-sharing vehicles) and about 525,945 passenger orders in January 2015 [37]. Then, in the test instances, the speed stochasticity can help these taxis reduce the carbon emissions by about 3156 kg ($(25.0018 - 24.9954) \times 525,945$). This indicates the necessity of considering speed stochasticity to reduce carbon emissions. Then, let us look at the results of $\theta > 0$, where the standard deviation of carbon emissions is taken into the objective function. Compared to the results of $\theta = 0$, when $\theta > 0$, the speed stochasticity brings more reductions in the objective values F . Moreover, as θ increases from 0 to 1.5, the reduced objective value also increases from 0.03% to 0.13%, mainly due to more reduction in the standard deviation of carbon emissions. For example, in Table 2, when $\theta = 1.5$, considering speed stochasticity helps to reduce the $\theta \cdot \sigma$ by 2.24%. This is because the 102-day speed data used in the SP can reflect the fluctuation of path emissions in real traffic environments. However, the average speeds used in the DP cannot. Then, when θ is large, the SP generates better travel paths than the DP does for more O-D pairs in Table 3 and lower objective values in Table 2, which demonstrates the necessity of considering speed stochasticity for large θ values.

Table 2. Performance comparison between the deterministic problem (DP) and the stochastic problem (SP).

	$\theta = 0$			$\theta = 0.5$			$\theta = 1$			$\theta = 1.5$		
	DP	SP	Diff. (%)	DP	SP	Diff. (%)	DP	SP	Diff. (%)	DP	SP	Diff. (%)
μ	25.0018	24.9954	−0.03	25.0018	24.9992	−0.01	25.0018	25.0083	0.03	25.0018	25.0167	0.06
$\theta \cdot \sigma$	-	-	-	0.7563	0.7488	−0.99	1.5126	1.4859	−1.77	2.2690	2.2181	−2.24
F	25.0018	24.9954	−0.03	25.7581	25.7480	−0.04	26.5144	26.4942	−0.08	27.2707	27.2347	−0.13

Table 3. Number of O-D pairs where the stochastic problem (SP) gets better and same paths compared with the deterministic problem (DP).

	$\theta = 0$	$\theta = 0.5$	$\theta = 1$	$\theta = 1.5$
better path	7	8	12	17
same path	193	192	188	183

This paper further analyses the carbon emission results of the O-D pairs in which the SP and the DP generate different optimal path solutions. Figure 5 shows the emission values reduced by the SP relative to the DP results for these O-D pairs. In the figure, the meanings of the μ , $\theta \cdot \sigma$, and F are consistent with their respective meanings in Equation (1). In the test instances, for an O-D pair, the SP can reduce the objective values F by up to 1.29%, 2.22%, 3.10%, and 3.93% in terms of the four θ values, respectively. It can be found from the figure that for some O-D pairs, such as (5, 2), (160, 35), (69, 84), and (158, 110), the SP generates better travel paths than the DP does for all four θ values. It is because these O-D pairs involve some road links where the average speeds deviate greatly from the common traffic conditions, the DP thus gets caught in bad solutions for all θ values. On the other hand, for some O-D pairs, such as (43, 103) and (26, 47), when θ is equal to zero, the SP and the DP have the same optimal travel path solutions. However, when θ is bigger than zero, the SP can adaptively find better travel paths with lower objective values by sacrificing a small number of emission means to obtain much smaller emission standard deviations. Moreover, when θ is equal to 0.5, 1, and 1.5, there are 2, 7, 13 such O-D pairs, respectively. The reason is that as the proportion of the emission standard deviation in the objective function increases, the actual optimal travel path is more likely to be the path with more stable emissions.

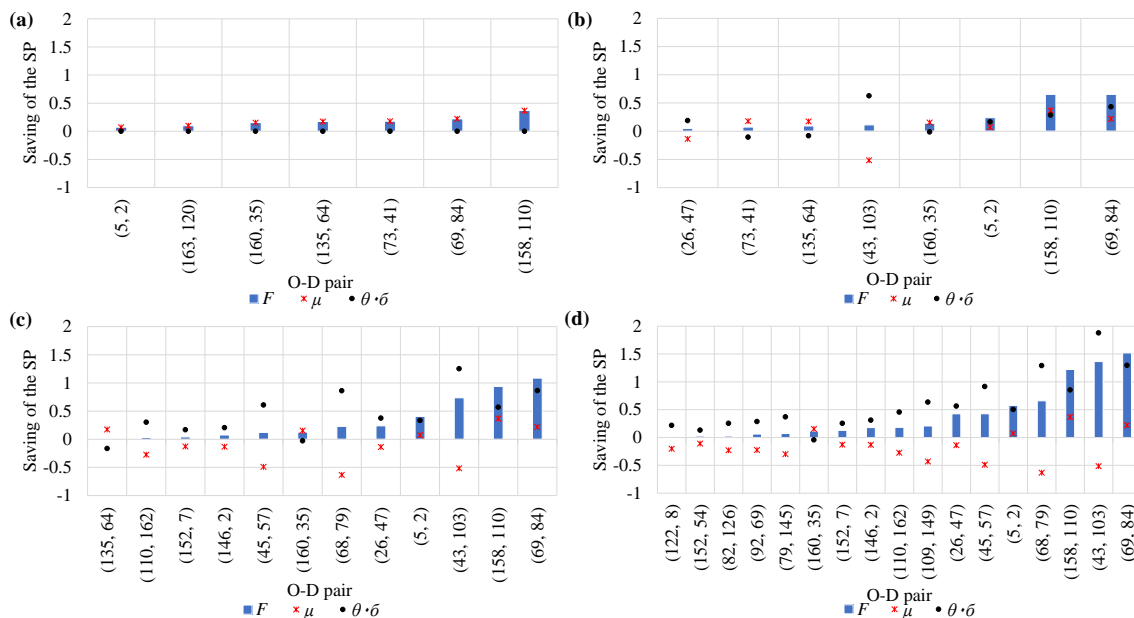


Figure 5. Emission values reduced by the stochastic problem (SP) when (a) $\theta = 0$, (b) $\theta = 0.5$, (c) $\theta = 1$, and (d) $\theta = 1.5$ relative to the deterministic problem (DP) results.

To summarize, considering speed stochasticity does help to get better solutions for certain O-D pairs under sustainable environments, and the number of such O-D pairs increases as the parameter θ increases. When the objective function is to minimize the mean of carbon emissions ($\theta = 0$), the emission mean of an O-D pair can be reduced by 0.03% on average by considering speed stochasticity. Although the percentage

0.03% is very small, the total carbon emissions reduced are still considerable when considering all travel demands of an entire city. On the other hand, when the emission standard deviation is also considered in the objective function ($\theta > 0$), more reductions in the emission objective values can be obtained from speed stochasticity, and the reduction of emission objective values increases as θ increases.

4.2. Effects of Speed Correlation

This sub-section investigates the effects of speed correlation. Table 4 shows the performance comparison between the CP and the UnCP in terms of four different θ values. The result is the average of 200 test instances. In the table, the meanings of the μ , $\theta \cdot \sigma$, and F are consistent with their respective meanings in Equation (1). The ‘Diff.’ represents the relative changes of the CP results compared with the UnCP results. Table 5 compares the optimal paths of the CP and UnCP. Please note that since the UnCP uses the stochastic uncorrelated scenario sets that remove the speed correlation in the real speed data, the obtained objective values of optimal travel paths are not their actual objective values. Therefore, all emission results of the UnCP in this sub-section are obtained by re-evaluating the optimal paths using the stochastic correlated scenario sets that are used in the CP. The reason why we do not use the 102-day real speed data to re-evaluate is that the speed distributions used in the SG method are fitted after some speed outliers are removed, and thus there are some differences between the exported stochastic scenario sets and the real speed data. To eliminate the impacts of removing outliers, the optimal paths of the UnCP are re-evaluated using the stochastic correlated scenario sets.

In Table 4, the CP generates lower objective values F than the UnCP does for all θ values. Moreover, as θ increases, the objective value reduced by the CP increases from 0.02% to 0.20%. This is because although the stochastic uncorrelated scenario sets used in the UnCP cannot express the speed correlation, they can still account for the common traffic conditions in the road network. Therefore, when $\theta = 0$, the UnCP can still find the travel paths with the minimum emission means for most O-D pairs in Table 5, and considering speed correlation only helps to reduce the emission mean by 0.02%. However, when considering all travel demands of an entire city, the 0.02% carbon emissions are still very significant. Take the city Los Angeles with about 2361 taxis and about 525,945 passenger orders in January 2017 as an example. The speed correlation can help reduce the carbon emissions by about 2630 kg $((25.1505 - 25.1455) \times 525,945)$, which indicates the necessity of considering the speed correlation to reduce carbon emissions. When $\theta > 0$, the emission standard deviation is taken into account in the objective function. Since the speed values on different links of a travel path do not change synchronously in a stochastic uncorrelated scenario set, the evaluation of path emission standard deviation would be inaccurate in the UnCP. Therefore, the UnCP could not find the actual optimal paths for some O-D pairs, especially when the emission standard deviation takes a large proportion in the objective function. In this case, the CP can get better travel paths for more O-D pairs and then more reductions in emission objective values for large θ values. These reductions of emission objective values are mainly due to the reduction in the emission standard deviation. For example, in Table 4, when $\theta = 1.5$, the 0.20% reduction in objective value F comes from the 2.88% reduction in the standard deviation part $\theta \cdot \sigma$. This demonstrates the necessity of considering speed correlation when the emission standard deviation accounts for a large proportion in the objective function.

Table 4. Performance comparison between the uncorrelated problem (UnCP) and the correlated problem (CP).

	$\theta = 0$			$\theta = 0.5$			$\theta = 1$			$\theta = 1.5$		
	DP	SP	Diff. (%)	DP	SP	Diff. (%)	DP	SP	Diff. (%)	DP	SP	Diff. (%)
μ	25.1505	25.1455	−0.02	25.1554	25.1534	−0.01	25.1618	25.1767	0.06	25.1796	25.1911	0.05
$\theta \cdot \sigma$	-	-	-	0.762	0.7583	−0.49	1.512	1.4793	−2.16	2.2638	2.1985	−2.88
F	25.1505	25.1455	−0.02	25.9174	25.9117	−0.02	26.6738	26.656	−0.07	27.4433	27.3896	−0.20

Table 5. Number of O-D pairs where the correlated problem (CP) gets better or same paths compared with the uncorrelated problem (UnCP).

	$\theta = 0$	$\theta = 0.5$	$\theta = 1$	$\theta = 1.5$
better path	7	12	18	25
same path	193	188	182	175

This paper further analyses the carbon emission results of the O-D pairs in which the CP and UnCP generate different optimal path solutions. Figure 6 shows the emission values reduced by the CP relative to the UnCP results for these O-D pairs. In the figure, the meanings of the μ , θ , σ , and F are consistent with their respective meanings in Equation (1). In the test instances, for an O-D pair, the CP can reduce the emission objective values F by up to 0.97%, 0.73%, 3.09%, and 5.25% in terms of the four θ values, respectively. In the figure, for most O-D pairs, the objective values reduced by the CP are less than 0.5, which indicates that whether or not the speed correlation is considered has a small effect on the optimal emission objective values for most O-D pairs. The CP only reduces the objective values of the O-D pairs (158, 110) and (43, 103) significantly for large θ values. For the O-D pair (158, 110), when θ is equal to 1.5, the CP finds a better travel path solution with a lower emission mean and a smaller emission standard deviation. For the O-D pair (43,103), when θ is equal to 1 and 1.5, the CP can significantly reduce the emission standard deviations by sacrificing a small number of emission means to obtain lower objective values. This indicates that for some specific O-D pairs and large θ values, we need to consider speed correlation to accurately evaluate the emission standard deviations of candidate travel paths and then find the actual optimal travel path. On the other hand, it is found that the value of θ has a significant impact on whether or not the CP can generate a better path solution. For example, for O-D pairs (92, 69) and (146, 2), the CP generates better paths than the UnCP does for all the three θ values greater than zero, whereas for some O-D pairs, such as (31, 1) and (79, 145), the CP only generates better path solutions for a specific θ value. This is because the UnCP itself may get different optimal travel path solutions for different θ values. Moreover, caused by the inaccurate evaluation of path emission standard deviations, the likelihood that the UnCP generates different optimal travel path solutions for different θ values will increase. Therefore, for some θ values, the UnCP could find the optimal travel paths, but may not for other θ values.

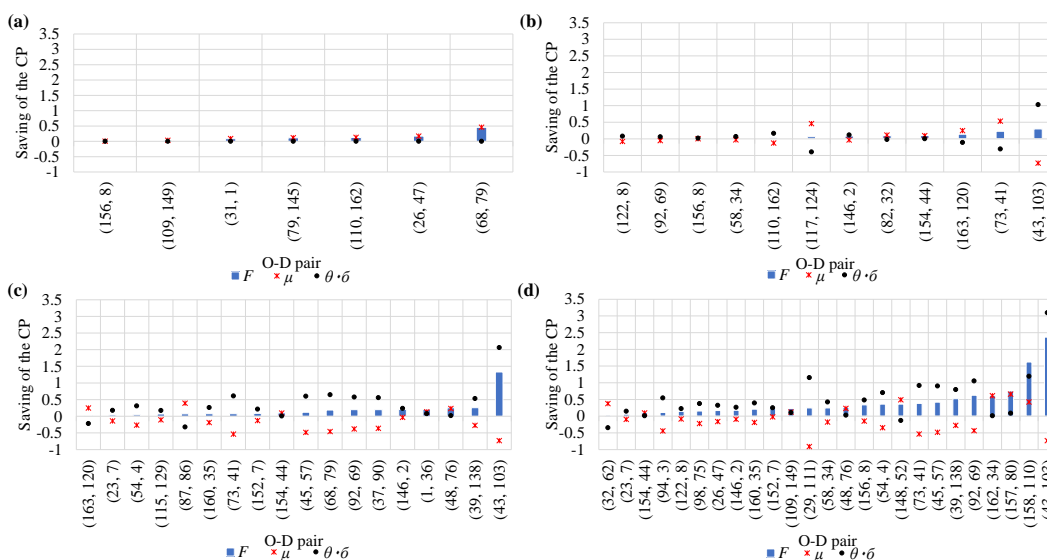


Figure 6. Emission values reduced by the correlated problem (CP) for (a) $\theta = 0$, (b) $\theta = 0.5$, (c) $\theta = 1$, and (d) $\theta = 1.5$ relative to the uncorrelated problem (UnCP) results.

To summarize, considering speed correlation does help to get better solutions for certain O-D pairs in sustainable environments, and the number of such O-D pairs increases as the parameter θ increases. When the objective function is to minimize the mean of carbon emissions ($\theta = 0$), the emission mean of an O-D pair can be reduced by 0.02% on average by considering speed correlation. Although the percentage of 0.02% is very small, the total carbon emissions reduced are still considerable when considering all travel demands of an entire city. When the objective function is to minimize the combination of emission mean and standard deviation ($\theta > 0$), more reductions in the emission objective values can be obtained by considering speed correlation, and the reduction increases as θ increases.

5. Conclusions

This paper studies whether and how the speed stochasticity and correlation affect the optimal travel path solution under sustainable environments. We use 102-day real time-dependent travel speeds from the Los Angeles freeway network as test data. To investigate the effects of speed stochasticity, this paper compares the optimal travel paths from the stochastic problem and the deterministic problem. To investigate the effects of speed correlation, this paper compares the optimal travel paths from the correlated problem and the uncorrelated problem.

We conducted comparison experiments on 200 randomly chosen O-D pairs in terms of four different θ values. It is found that both speed stochasticity and correlation can affect the optimal travel path solutions and the reduced emission objective values by considering stochasticity and correlation increase as the parameter θ increases. Specifically, for an O-D pair, when the objective function is to minimize the mean of carbon emission, the emission mean can be reduced by 0.03% by considering speed stochasticity and by 0.02% by considering speed correlation. Although these two percentages look very inconspicuous, when considering all travel demands of an entire city, the reduction in carbon emissions is considerable. Taking the city Los Angeles with about 2361 taxis and about 525,945 passenger orders in January 2017 as an example, the 0.02% reduction means that 2630 kg carbon emissions can be reduced in the test instances.

Moreover, when the standard deviation of carbon emissions is also considered in the objective function, the reductions of the objective values are more significant by considering speed stochasticity and correlation, and the reduction increases as θ increases. For example, when $\theta = 1.5$, the emission objective value can be reduced by up to 0.13% by considering speed stochasticity and by up to 0.20% by considering speed correlation. This demonstrates the necessity of considering speed stochasticity and correlation, especially when the standard deviation of carbon emissions takes a large proportion in the objective function.

This paper studies the shortest path problem with an objective function of minimizing the combination of emission mean and standard deviation. Our future research will extend to other objective functions, such as minimizing expected travel time and maximizing travel time reliability.

Author Contributions: Investigation, D.Z.; writing—original draft, D.Z.; writing—review and editing, Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by Sichuan University (grant numbers: 2018hhs-37, SKSYL201819, LH2018003).

Acknowledgments: The authors would like to thank the Freeway Performance Measurement (PeMS) system for providing the original speed data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Edenhofer, O. *Climate Change 2014: Mitigation of Climate Change*; Cambridge University Press: Cambridge, UK, 2015; Volume 3.
2. Ubeda, S.; Arcelus, F.J.; Faulin, J. Green logistics at Eroski: A case study. *Int. J. Prod. Econ.* **2011**, *131*, 44–51. [[CrossRef](#)]
3. Tang, S.; Wang, W.; Yan, H.; Hao, G. Low carbon logistics: Reducing shipment frequency to cut carbon emissions. *Int. J. Prod. Econ.* **2015**, *164*, 339–350. [[CrossRef](#)]
4. Piecyk, M.; Browne, M.; Whiteing, A.; McKinnon, A. *Green Logistics: Improving the Environmental Sustainability of Logistics*; Kogan Page Publishers: London, UK, 2015.
5. Bektaş, T.; Laporte, G. The pollution-routing problem. *Transp. Res. Part B Methodol.* **2011**, *45*, 1232–1250. [[CrossRef](#)]
6. Dekker, R.; Bloemhof, J.; Mallidis, I. Operations research for green logistics—An overview of aspects, issues, contributions and challenges. *Eur. J. Oper. Res.* **2012**, *219*, 671–679. [[CrossRef](#)]
7. Demir, E.; Bektaş, T.; Laporte, G. A review of recent research on green road freight transportation. *Eur. J. Oper. Res.* **2014**, *237*, 775–793. [[CrossRef](#)]
8. Lin, C.; Choy, K.L.; Ho, G.T.S.; Chung, S.H.; Lam, H.Y. Survey of green vehicle routing problem: Past and future trends. *Expert Syst. Appl.* **2014**, *41*, 1118–1138. [[CrossRef](#)]
9. Sun, Y.; Yu, X.; Bie, R.; Song, H. Discovering time-dependent shortest path on traffic graph for drivers towards green driving. *J. Netw. Comput. Appl.* **2017**, *83*, 204–212. [[CrossRef](#)]
10. Huang, Y.; Zhao, L.; Van Woensel, T.; Gross, J.P. Time-dependent vehicle routing problem with path flexibility. *Transp. Res. Part B Methodol.* **2017**, *95*, 169–195. [[CrossRef](#)]
11. Ehmke, J.F.; Campbell, A.M.; Thomas, B.W. Optimizing for total costs in vehicle routing in urban areas. *Transp. Res. Part E Logist. Transp. Rev.* **2018**, *116*, 242–265. [[CrossRef](#)]
12. Laporte, G.; Pascoal, M.M. Minimum cost path problems with relays. *Comput. Oper. Res.* **2011**, *38*, 165–173. [[CrossRef](#)]
13. Arslan, O.; Yıldız, B.; Karaşan, O.E. Minimum cost path problem for plug-in hybrid electric vehicles. *Transp. Res. Part E Logist. Transp. Rev.* **2015**, *80*, 123–141. [[CrossRef](#)]
14. Huang, H.; Gao, S. Optimal paths in dynamic networks with dependent random link travel times. *Transp. Res. Part B Methodol.* **2012**, *46*, 579–598. [[CrossRef](#)]
15. Fu, L.; Rilett, L.R. Expected shortest paths in dynamic and stochastic traffic networks. *Transp. Res. Part B Methodol.* **1998**, *32*, 499–516. [[CrossRef](#)]
16. Miller-Hooks, E.D.; Mahmassani, H.S. Least expected time paths in stochastic, time-varying transportation networks. *Transp. Sci.* **2000**, *34*, 198–215. [[CrossRef](#)]
17. Yang, L.; Zhou, X. Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem. *Transp. Res. Part B Methodol.* **2014**, *59*, 22–44. [[CrossRef](#)]
18. Prakash, A.A. Pruning algorithm for the least expected travel time path on stochastic and time-dependent networks. *Transp. Res. Part B Methodol.* **2018**, *108*, 127–147. [[CrossRef](#)]
19. Yang, L.; Zhou, X. Optimizing on-time arrival probability and percentile travel time for elementary path finding in time-dependent transportation networks: Linear mixed integer programming reformulations. *Transp. Res. Part B Methodol.* **2017**, *96*, 68–91. [[CrossRef](#)]
20. Chen, B.Y.; Shi, C.; Zhang, J.; Lam, W.H.; Li, Q.; Xiang, S. Most reliable path-finding algorithm for maximizing on-time arrival probability. *Transp. B Transp. Dyn.* **2017**, *5*, 248–264. [[CrossRef](#)]
21. Chen, P.; Tong, R.; Lu, G.Q.; Wang, Y.P. The alpha-reliable path problem in stochastic road networks with link correlations: A moment-matching-based path finding algorithm. *Expert Syst. Appl.* **2018**, *110*, 20–32. [[CrossRef](#)]
22. Ehmke, J.F.; Campbell, A.M.; Thomas, B.W. Data-driven approaches for emissions-minimized paths in urban areas. *Comput. Oper. Res.* **2016**, *67*, 34–47. [[CrossRef](#)]
23. Chabini, I.; Lan, S. Adaptations of the A* algorithm for the computation of fastest paths in deterministic discrete-time dynamic networks. *IEEE Trans. Intell. Transp. Syst.* **2002**, *3*, 60–74. [[CrossRef](#)]

24. Cheng, T.; Haworth, J.; Wang, J. Spatio-temporal autocorrelation of road network data. *J. Geogr. Syst.* **2012**, *14*, 389–413. [[CrossRef](#)]
25. Rachtan, P.; Huang, H.; Gao, S. Spatiotemporal link speed correlations: Empirical study. *Transp. Res. Rec. J. Transp. Res. Board* **2013**, *2390*, 34–43. [[CrossRef](#)]
26. Ermagun, A.; Chatterjee, S.; Levinson, D. Using temporal detrending to observe the spatial correlation of traffic. *PLoS ONE* **2017**, *12*, e0176853. [[CrossRef](#)] [[PubMed](#)]
27. Guo, F.; Zhang, D.; Dong, Y.; Guo, Z. Urban link travel speed dataset from a megacity road network. *Sci. Data* **2019**, *6*, 61. [[CrossRef](#)]
28. Zockaie, A.; Nie, Y.; Wu, X.; Mahmassani, H. Impacts of correlations on reliable shortest path finding: A simulation-based study. *Transp. Res. Rec. J. Transp. Res. Board* **2013**, *2334*, 1–9. [[CrossRef](#)]
29. Huang, H.; Gao, S. Trajectory-adaptive routing in dynamic networks with dependent random link travel times. *Transp. Sci.* **2018**, *52*, 102–117. [[CrossRef](#)]
30. Deo, N.; Pang, C.Y. Shortest-path algorithms: Taxonomy and annotation. *Networks* **1984**, *14*, 275–323. [[CrossRef](#)]
31. Dean, B.C. *Shortest Paths in FIFO Time-Dependent Networks: Theory and Algorithms*; Rapport Technique; Massachusetts Institute of Technology: Cambridge, MA, USA, 2004.
32. Gao, S. *Optimal Adaptive Routing and Traffic Assignment in Stochastic Time-Dependent Networks*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2005.
33. Hickman, J.; Hassel, D.; Joumard, R.; Samaras, Z.; Sorenson, S. *MEET-Methodology for Calculating Transport Emissions and Energy Consumption*; Technical Report; European Commission: Brussels, Belgium, 1999.
34. Chen, C. *Freeway Performance Measurement System (PeMS)*. Ph.D. Thesis, University of California, Berkeley, CA, USA, 2003.
35. Kaut, M.; Wallace, S.W. Shape-based scenario generation using copulas. *Comput. Manag. Sci.* **2011**, *8*, 181–199. [[CrossRef](#)]
36. Grubbs, F.E. Sample criteria for testing outlying observations. *Ann. Math. Stat.* **1950**, *21*, 27–58. [[CrossRef](#)]
37. Brands, J.; Murray, J. *Los Angeles Taxicab Review and Performance Report (2014–2015)*; Technical Report; Los Angeles Department of Transportation: Los Angeles, CA, USA, 2017.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).