

Research Article

A Structural Comparison between the Origin-Destination Matrices Based on Local Windows with Socioeconomic, Land-Use, and Population Characteristics

Shahriar Afandizadeh Zargari ¹, Amir Masoud Memarnejad ¹,
and Hamid Mirzahosseini ²

¹Department of Transportation Engineering and Planning, School of Civil Engineering,
Iran University of Science and Technology (IUST), Tehran 16846-13114, Iran

²Department of Civil Engineering–Transportation Planning, Imam Khomeini International University (IKIU),
Qazvin 34148-96818, Iran

Correspondence should be addressed to Shahriar Afandizadeh Zargari; zargari@iust.ac.ir

Received 31 March 2021; Accepted 28 May 2021; Published 14 June 2021

Academic Editor: Elżbieta Macioszek

Copyright © 2021 Shahriar Afandizadeh Zargari 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.

The origin-destination (OD) matrices express the number and the pattern of trips distributed between OD pairs. OD matrix structural comparison can be used to identify different mobility patterns in the cities. A comparison of two OD matrices could express their difference from both numerical and structural aspects. Limited methods, such as the mean structural similarity (MSSIM) index and geographical window-based structural similarity index (GSSI), have been developed to compare the structural similarity (SSIM) of two matrices. These methods calculate the structural similarities of two OD matrices by grouping the OD pairs into local windows. The obtained results from the MSSIM entirely depend on the dimensions of the chosen windows. Meanwhile, the GSSI method only focuses on the geographical adjacency and correlation of zones while selecting local windows. Accordingly, this paper developed a novel method named Socioeconomy, Land-use, and Population Structural Similarity Index (SLPSSI) in which local windows are selected according to socioeconomic, land-use, and population properties for SSIM comparison of OD matrices. The proposed method was tested on Tehran's OD matrix extracted from cell phone Geographic Position System (GPS) data. The advantage of this method over two previous ones was observed in determining the new pattern of trips on local windows and more precise detection of SSIM of the weekdays. The SLPSSI approach is up to 10 percent more accurate than the MSSIM method and up to 5.5 percent more accurate than the GSSI method. The proposed method also had a better performance on sparse OD matrices. It is capable of better determining the SSIM of sparse OD matrices by up to 8% compared with the GSSI method. Finally, the sensitivity analysis results indicate that the suggested method is robust and reliable since it is sensitive to applying both constant and random coefficients.

1. Introduction

An origin-destination (OD) matrix of urban trips indicates the demand for trips between different traffic zones [1]. The matrix provides transportation engineers with important information on the characteristics and patterns of trips in cities. There are two aspects in an OD matrix comparison:

- (1) The numerical value of each cell of the matrix, which indicates the number of trips between OD pairs

- (2) The structure of trip distribution among different traffic zones, which shows the matrix structure

In comparing the OD matrices, both mentioned aspects are important. A difference between each cell's numerical values when comparing two OD matrices indicates the difference in the number of trips between them and consequently the probability of difference in the flow in links after assigning the matrices to the network. Hence, a difference in trip distribution structures raises the possibility of different analyses of the trip

patterns. Therefore, it is crucially important to consider both numerical and structural aspects between the estimated matrix and the real matrix in OD matrix estimation processes. Previous studies have employed various methods to compare two OD matrices. Some of these methods, namely, traditional ones in this paper, only consider the numerical difference or similarity of two matrices, while some called structural methods consider both numerical and structural differences simultaneously. Two matrices are identical when they are structurally similar and have no difference in cells' numerical values. For a better description, Figure 1 presents two sample matrices with structural and numerical differences.

Figures 1(a) and 1(b) illustrate 2 different OD matrices, that is, T_1 and T_2 , in which the number of trips from a location in a zone listed in the left column to other zones listed in the top row is shown. There are no similarities in the structure of the matrices. As shown in T_1 , the preferred order of destinations for trips from origin A is B, C, A , and D . Meanwhile, the preferred order of destinations for trips from origin A in the T_2 matrix is A, D, C , and B . Hence, the preference of trips from origin A is different in the two matrices. Rows B has no structural similarity (SSIM) in the above matrices too. However, for both matrices, C, B, D , and A are the preferred destinations for trips from origin C . Thus, rows C in the two matrices are dominantly consistent in terms of the structural framework while having different numerical values. The fourth rows of both matrices demonstrate the trip demand from origin D to other traffic zones, and they are structurally and numerically identical. Though, from a statistical point of view and during comparing any two identical rows, a slight difference in the values of each cell could be neglected.

In previous studies, many traditional methods have been used for comparing OD matrices, including root mean square error (RMSE) [2–4], normalized root mean square error (RMSN) [1, 5], mean square error (MSE) [6], mean absolute error ratio (MAER) [7], mean absolute percentage error (MAPE) [8, 9], the goodness of Theil's fit (GU) [10], R-squared (R^2) [11], and entropy measure (E) [12]. These methods use mathematical formulation to compare cell-to-cell of OD matrices and are not capable of doing group comparison of origins and destinations. It means that they cannot analyze the structural difference of matrices caused by the difference in choosing the trip destinations [13]. For instance, these mathematical relations could not detect the structural difference between two matrices when, for example, one of them is obtained from multiplication of a constant number to a base OD matrix, and the other is chosen from random arrays [14].

Unlike the traditional methods, the structural methods used for comparison of OD matrices are not much in the literature and can be summarized as follows:

- (1) Mean structural similarity (MSSIM) index [15]
- (2) Geographical structural similarity index (GSSI) [13]
- (3) Wasserstein distance [16]
- (4) Normalized Levenshtein distance for OD matrices (NLOD) [14]

The MSSIM consists of three components. The mean and standard deviation mathematical relations determine the numerical similarity between two OD matrices, while the covariance measure detects the SSIM. Therefore, it can define the similarity or difference between OD matrices [15]. However, it has some disadvantages, including the following:

- (1) The results depend on the dimensions of the chosen window
- (2) Smaller and more precise windows require higher temporal calculation costs
- (3) Since the OD pairs have no correlation with each other in a local window, the SSIM value does not necessarily illustrate any particular concept [17]

The GSSI method, which has been designed based on the MSSIM to overcome the problems, has not been able to overcome all weaknesses either. For instance, it used the geographical locations of zones to design local windows and solved the problem concerning the dimensions of chosen windows. However, the boundaries mostly predefined according to the geographical location of zones can be defined based on socioeconomic indices, land use, and population. While the Wasserstein distance method carefully detects the structural and numerical similarity between two OD matrices, it is very time-consuming due to its optimization nature. The NLOD method has also been developed based on the Levenshtein distance method, and it has been used for comparing two strings of texts. Like the Wasserstein distance method, this method compares OD matrices' flows using optimization techniques. This method calculates the minimum structural distance between two OD matrices to convert one matrix to another.

This paper focuses on the MSSIM and GSSI methods to make effective changes and modify the mentioned problems. Regardless of the geographical locations of Traffic Analysis Zones (TAZs), this study has considered the following properties to offer a new tool for developing the MSSIM:

- (1) The socioeconomic properties of the TAZs, such as the level of residents' employment and private vehicle ownership per capita
- (2) Land-use properties, such as the area of each land-use class
- (3) Population properties, that is, the population of each zone

In this research, the computational structure of the MSSIM method was improved with new windows designed according to the socioeconomic indices, land use, and population. Accordingly, Socioeconomy, Land-use, and Population Structural Similarity Index (SLPSSI) is introduced and implemented during this research.

The remainder of the paper is structured as follows: Section 2 reviews the literature on the use of statistical measures for OD matrices comparison; Section 3 presents the methodology and explains in detail the development process of the proposed measure, SLPSSI; Section 4 compares the proposed method with MSSIM and GSSI

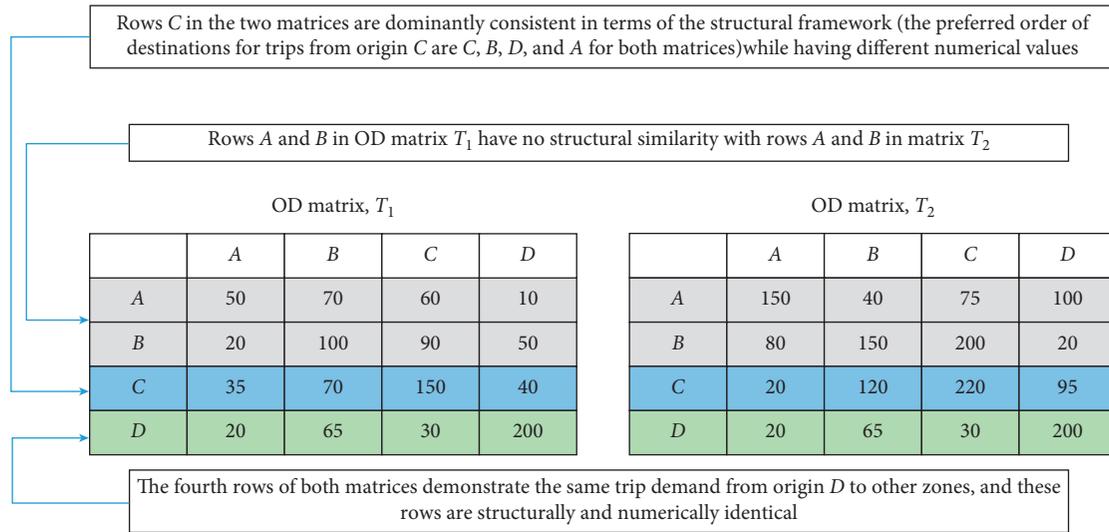


FIGURE 1: Two sample matrices with structural and numerical differences.

methods through case study application using GPS-OD dataset from Tehran city. Section 5 tests its robustness through sensitivity analysis; and finally, the paper concludes in Section 6.

2. Literature Review

Extensive studies addressed the development of performance assessment indices for transportation using numerical comparison of OD matrices [18]. However, there is little attention devoted to structural comparison methods. As mentioned before, traditional methods are not capable of comparing a group of OD matrices and cannot determine the SSIM or difference between two matrices as well [14].

2.1. Structural Comparison Indices. Limited studies have been conducted on the structural comparison of OD matrices. Most indices in this field are from other scientific fields as explained in the following sections.

2.1.1. MSSIM Index. Wang et al. [19] used the MSSIM for the first time to compare two images. They showed that two images with equal MSEs relative to a base image may have different MSSIMs compared to the same original image. Djukic [17] suggested using MSSIM to compare two OD matrices assuming that each cell represents one pixel in the image. In this method, by defining a local virtual window, a group of OD pairs is compared. Finally, the values associated with each window for the whole matrix are averaged. The dimensions of the selected windows are necessarily smaller than the matrix dimensions. For example, for an $M \times N$ matrix, the dimensions of the chosen $m \times n$ window are such that $m \leq M$ and $n \leq N$. Figure 2 demonstrates a sample window for an assumed constant T matrix and how a 2×2 local window moves. The dimensions of the chosen local window do not have to be the same for calculating the MSSIM.

In order to calculate the local SSIM, each figure (a-i) is compared with a matrix that has a similar local window (the matrix that is decided to evaluate its similarity with matrix T). For instance, matrices T_1 and T_2 illustrated in Figure 1 are compared in the first local window (a), as shown in Figure 3.

The MSSIM calculation consists of three mathematical relations. These relations evaluate the similarity or difference of the chosen window by comparing the mean, standard deviation, and covariance values. Equation (1) compares the means of the selected windows, that is, μ_x and μ_y . This equation determines luminance $l(x, y)$. Equation (2) compares the standard deviations (σ_x and σ_y). It is called contrast and is shown by $c(x, y)$. Finally, equation (3) compares the covariance (σ_{xy}) between the entries of the chosen local window in two OD matrices. It is called structure and represented by $str(x, y)$. The first two sections evaluate the numerical similarity/difference between two matrices, while the third section assesses their SSIM. In these equations, x and y are a group of origins and destinations located in the same window in matrices X and Y . Equation (4) is determined from the multiplication of equations. Parameter c_1 , c_2 , and c_3 stabilize the solutions when the mean or standard deviation is zero. Generally, c_3 is considered equal to $c_2/2$. Previous studies reported the values of 10^{-10} and 10^{-2} for the coefficients c_1 and c_2 , respectively [20]. These parameters could be zero unless the mean and standard deviation of the windows are zero. Powers α , β , and γ in equation (4) denote the importance and effects of the proposed parameters, that is, mean, standard deviation, and covariance, respectively. These powers are normally set to one. With these assumptions, the SSIM relationship leads to equation (5). Equation (6) could generate the average value of SSIMs belonging to L local windows. The MSSIM varies between -1 and 1 . The value of 1 indicates two identical matrices; meanwhile, comparing two inverse matrices leads to -1 .

$$l(x, y) = \frac{(2\mu_x\mu_y + c_1)}{(\mu_x^2 + \mu_y^2 + c_1)}, \quad (1)$$

$$c(x, y) = \frac{(2\sigma_x\sigma_y + c_2)}{(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (2)$$

$$\text{str}(x, y) = \frac{(\sigma_{xy} + c_3)}{(\sigma_x\sigma_y + c_3)}, \quad (3)$$

$$\text{SSIM}(x, y) = [l(x, y)^\alpha][c(x, y)^\beta][\text{str}(x, y)^\gamma], \quad \alpha > 0, \beta > 0, \text{ and } \gamma > 0. \quad (4)$$

Assuming $\alpha = \beta = \gamma = 1$ and $c_3 = c_2/2$.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (5)$$

$$\text{MSSIM}(X, Y) = \frac{1}{L} \sum_1^L \text{SSIM}(x, y). \quad (6)$$

Although the MSSIM can recognize the SSIM/difference of two OD matrices, it has the following flaws:

- (1) No report has addressed the procedure for choosing the dimensions of local windows and its relation with the result of the comparison between two matrices. On the other hand, no article is found showing how MSSIM can identify the structural difference of two matrices when the whole matrices are considered as one window. The reduction in the dimensions of local windows increases the accuracy and the computational cost.
- (2) Since this method is based on comparing images, it is sensitive to the adjacency of image pixels. In other words, since the numerical order of the origin and destination numbers does not necessarily indicate the closeness of the origin and destination [20], errors may occur during calculating the correlation between zones in the local window. Therefore, previous studies have suggested considering the whole matrix as one window [13].
- (3) It is not clear how to propose the constant coefficients and further investigation is required to determine these constants.
- (4) In this method, the OD pairs located in the same window do not necessarily have any correlation with each other, and the SSIM has no particular meaning for each local window. When the matrix is ordered based on the value of flow between origins and destinations, although the zones with high volumes are positioned beside each other, it is not suitable to calculate MSSIM on the ordered matrices when comparing various days with different flow volumes

order patterns (like Sundays and Fridays). Therefore, it is again suggested to calculate MSSIM on a window with the dimension of the whole matrix.

2.1.2. Supplementary Methods for MSSIM. The GSSI [13] and 4D-MSSIM [21] are developed to overcome the MSSIM's flaws, that is, the lack of clearness in choosing the dimensions of local windows and deficiency in spatial adjacency of different TAZs in chosen local windows.

The 4D-MSSIM identifies the adjacency of OD pairs using spatial distance estimation. The obtained Euclidean distance is used to calculate the adjacency distance between OD pairs. This method seeks to select and classify the zones adjacent to each other. The classification of these zones depends on selecting the reference zone from which the distances of the other zones would be calculated. Moreover, the Euclidean distance cannot express the spatial neighborhood of zones precisely. For instance, the TAZs separated by natural or artificial features like rivers or highways are spatially close to each other, even though the trip distance between them is longer.

In the GSSI method, the origins and destinations are first sorted according to the geographical locations. Afterward, the local windows used in the MSSIM method are determined according to the previously defined geographical boundaries and located on higher layers. In this method, the size of local windows varies depending on the number of TAZs in each geographical area. Each local window chosen according to the geographical boundaries is called a geographical window. This method provides a new way to organize windows in the SSIM method. However, it still uses the geographical distance index and predefined boundaries for geographical areas to determine local window dimensions.

2.1.3. Optimization-Based Methods. The Wasserstein distance method has been used for optimum transmission of goods. This method is used to identify the proper distribution of goods from given origins to destinations. The Wasserstein distance between two OD matrices determines the minimum travel time required to assign the first matrix

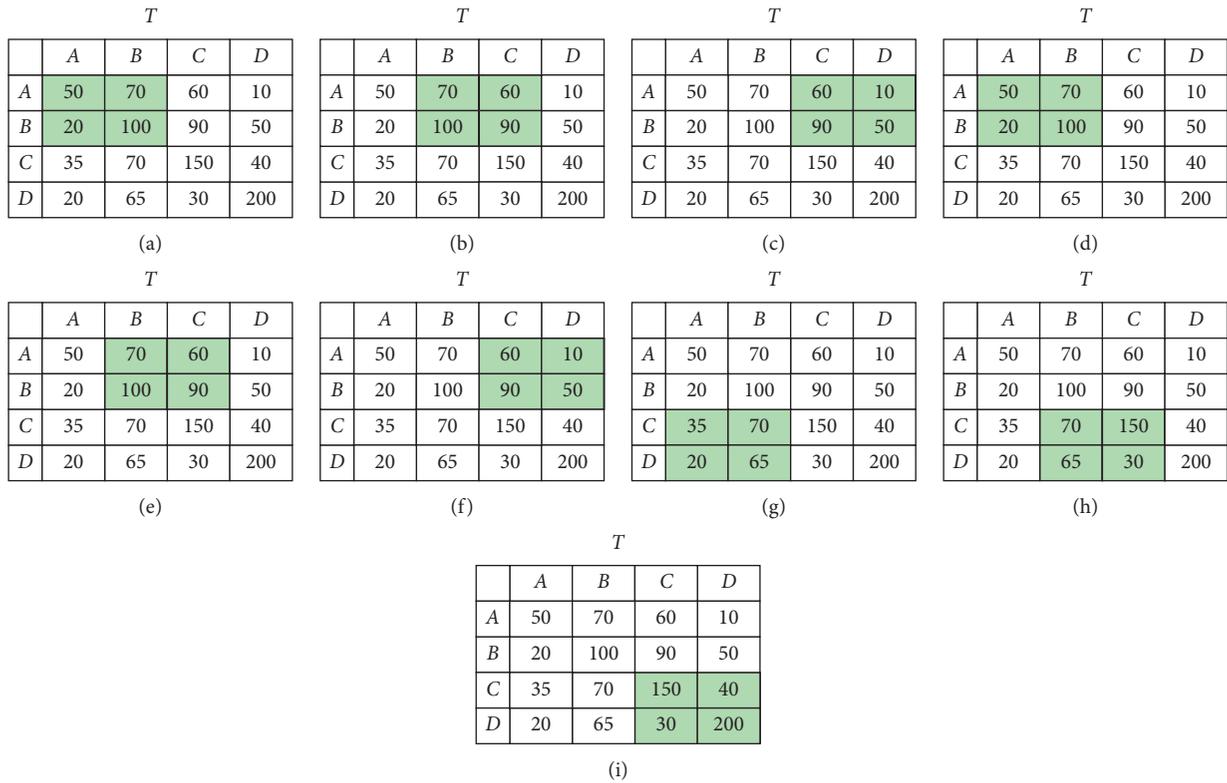


FIGURE 2: Creation and movement of a 2×2 window in a matrix.

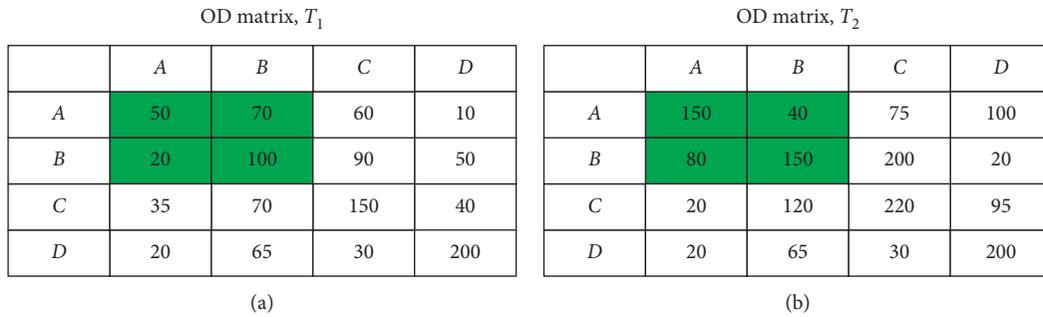


FIGURE 3: A sample of a local window to calculate SSIM.

considering the second matrix assignment pattern [16]. The goal of this method is to match two matrices to determine the lowest travel time. So, it can naturally determine the numerical and structural differences of two matrices. It is a linear programming issue and uses an optimization technique. Although this method detects the difference between distributions in various matrices properly or in other words recognizes the structural difference between two matrices, it computationally costs too much. However, this method cannot be easily used for OD matrices with large dimensions. In other words, this method is not suitable for complex urban networks [16].

The normalized Levenshtein distance for OD matrices (NLOD) method [14] calculates the structural and numerical differences between two matrices using an optimization tool

by determining the similarity between two strings. Levenshtein distance calculates the minimum cost required to cross over two strings, similar to mutation and crossover operations performed in Genetic Algorithm (GA) optimization technique by insertion, deletion, or substitution of parameters, or value of trips between zones in here. The normalized Levenshtein distance has rarely been used in transportation engineering. In most studies, this method is used for manipulation of a string that may consist of characters, variables, numbers, and so on, for example, the comparison between the license plates of vehicles [22], time-series comparison [23], sequences of trip purposes, and cluster activity-travel patterns [24]. In the NLOD method, each row of the OD matrix is sorted according to its traffic flow. Afterward, the minimum changes for converting each

row of the intended matrix to the main matrix considering the numbers in the cells and distribution of trips in each row are determined. Eventually, for the whole matrix, a normalized number between zero and one is reported (zero for two identical matrices) [14].

2.2. Summary of the Literature Review. The primary methods for comparing OD matrices that only focus on the numerical cell-to-cell comparison of two matrices cannot specify the structural difference of two matrices despite their overall numerical similarity. Other methods, such as MSSIM and its developed versions, including GSSI and 4D-MSSIM, can successfully detect the SSIM/difference between two OD matrices. However, given the points mentioned above, such as the way to choose a local window in the GSSI method, ignoring the socioeconomic, land-use, and population parameters when choosing local windows, and the error in selecting the adjacent zones are not considered in the 4D-MSSIM method, even if the developed methods require more supplementary progress. On the other hand, more complex methods like NLOD and Wasserstein distance have high computational costs because of the problem's optimization nature. Therefore, by improving the MSSIM method with the developed approach in this study, an opportunity to more precisely calculate the structural difference of OD matrices is provided.

3. Methodology

Firstly, the proposed method tries to classify traffic zones based on their ability to generate and attract trips. Five indices (FI) define each TAZ, that is, the resident population, car ownership per capita, population of employees, and lands used for commercial and administrative purposes (land uses). These indices are potential factors for producing and attracting trips in traffic zones. The first three indices produce trips, while the last two, that is, areas of commercial and administrative land use, attract trips. In this process, TAZs are divided into a given number of classes. To categorize similar zones, each zone is assigned a number between zero and one (the normalized value) for each of the above five indices. The total score of each zone is the average of all indices. Then, considering the conducted scoring for each zone using the k -means method, the zones are clustered according to the highest similarity to each other. Numerous methods had been used to cluster traffic data and zones, among which the k -means method has provided the best results [25, 26]. There are TAZs in each class having similar production and attraction potential for trips. The local windows are located on these classes to compare OD matrices (Figure 4).

According to equation (5), the SSIM value can be calculated on the selected local windows considering the socioeconomic, land-use, and zone population properties. That equation is applied on L local windows in this paper with the symbol SLPSSI according to equation (8).

$$\text{SLPSTR}(X, Y) = \frac{1}{L} \sum_{l=1}^L \text{STR}(x_l, y_l). \quad (7)$$

$$\text{SLPSSI}(X, Y) = \frac{1}{L} \sum_{l=1}^L \text{SSIM}(x_l, y_l). \quad (8)$$

In this equation, X and Y denote the proposed matrices for comparison. Furthermore, x_l and y_l indicate a series of OD pairs on the l th local window. The SLPSSI varies from -1 to 1 . Equation (7) only gives the structural difference of two matrices, which also varies between -1 and 1 .

Figure 5 provides the steps of employing the suggested method.

3.1. Implementation of the Suggested Windowing Method on the OD Matrix of Tehran. In order to investigate the appropriateness of using the developed method in this paper, the suggested windowing method is evaluated on the network of Tehran. The city of Tehran and its surroundings have 731 TAZs, 699 of which are inside the internal boundary of the city. This classification is the finest one considered for Tehran, and comprehensive urban and suburb transportation studies of the city have been conducted at this zoning level. It is called Zone Level 1 (ZL 1) hereafter. On the upper level, the city has been divided into 122 urban zones. Most likely, these zones have been separated based on geographical characteristics and natural features (such as floodways) and manmade ones (like highways). It is called Zone Level 2 (ZL 2) hereafter. Finally, on the topmost level, Tehran consists of 22 municipal districts. These districts form the urban management framework. It is called Zone Level 3 (ZL3) hereafter. Figure 6 demonstrates the zoning of Tehran considering the area level of zones.

In order to compare OD matrices, the matrix estimated by GPS data recorded through navigation software is used. In this procedure, the total data of locations recorded by more than 400 thousand Neshan application users in Tehran throughout one month has been used. It is worth mentioning that numerous methods have previously been used to estimate OD matrices using cellphone data [27] and GPS data [28]. In the present study, the OD matrix of the flow estimated from the GPS data is used only to evaluate the suggested method in actual conditions. The matrix is obtained by daily analyses in one month of the users' location data, including 326 million records. For estimating the matrix, computational rules are employed to detect the stops and movements of individuals. It should be noted that the obtained matrix is not necessarily the actual OD matrix of trips; in other words, the start and end of trips in this matrix are not necessarily the start and end of the main trips of the citizens. The matrix is only used to test the proposed method.

After estimating the OD matrix of Tehran at ZL 1, the OD matrix is aggregated on ZL 2 to be used in the present study. The aggregation eliminates the sparsity of the OD matrix, leading to a 122×122 matrix during investigation. The whole stages of choosing and clustering similar zones are determined in ZL 2. The five parameters consist of the ratio of the employees to the total population, car ownership per capita, area of

Areas with similar socioeconomic, land-use, and population indices		Trips from origin o_2 to destination d_1				The local socioeconomic, land-use, and population window employed to compare the structural similarity					
Origin	Dest	D1		D2				Di			
		d1	d2	d3	d4	d5	d6
O1	O1										
	O2										
O2	O3										
	O4										
	O5										
	O6										
⋮	⋮										
O _i	⋮										
	⋮										
	⋮										

FIGURE 4: Application of the suggested method on a presumed matrix.

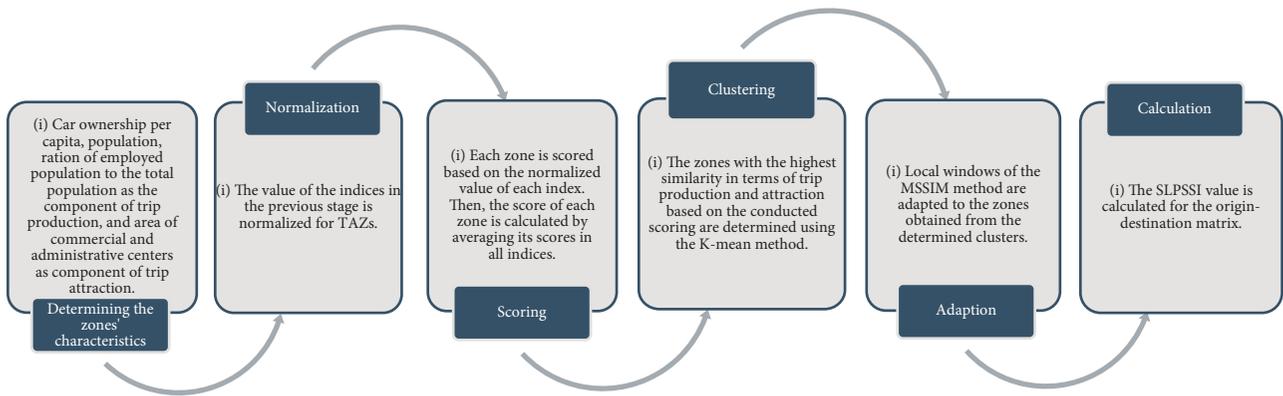


FIGURE 5: Flowchart of the suggested method.

administrative land uses, area of commercial land uses, and population aggregated for each zone on ZL 2. The zones with the highest similarity in producing and attracting trips are clustered, as shown in Table 1. The group of zones with the highest scores is classified as SLP Group5, and the other zones are classified accordingly.

Figure 7 indicates the number of zones in each cluster. The local windows used in the SLPSSI equations are defined on the zones with similarity on ZL 2 (Figure 8).

The suggested method is also compared with basic MSSIM and GSSI to evaluate its performance and advantage. Accordingly, to design local windows in the GSSI method, 122 urban zones (ZL 2) are divided into northern, southern, eastern, western, and central zones in a higher layer as shown in Figure 9.

4. Results

In this section, the results of the proposed method (SLPSSI) are compared with MSSIM and GSSI and its advantages are discussed. It should be noted that the working days in Iran

are from Saturday to Wednesday, Thursday is the weekend of many companies and offices, and also Friday is the national weekend.

4.1. Evaluation of Travel Patterns in Local Windows. The local windows suggested by the proposed method allow for comparing the travel patterns between a group of zones with similar socioeconomic, land-use, and population properties. Since the windows are fixed, each window can reveal the nature of the structural difference or similarity within trips. This feature cannot be concluded from the base MSSIM method because the local windows are not fixed in MSSIM. Moreover, the GSSI method only expresses the travel patterns that coincident with the geographical locations, while the proposed method can extract patterns independent of the geographical locations of zones. For this purpose, the OD matrix of Sunday (a working day), on 13 October 2019, was compared with Friday (weekend), on 18 October 2019, using the SLPSSI method. Table 2 addresses the

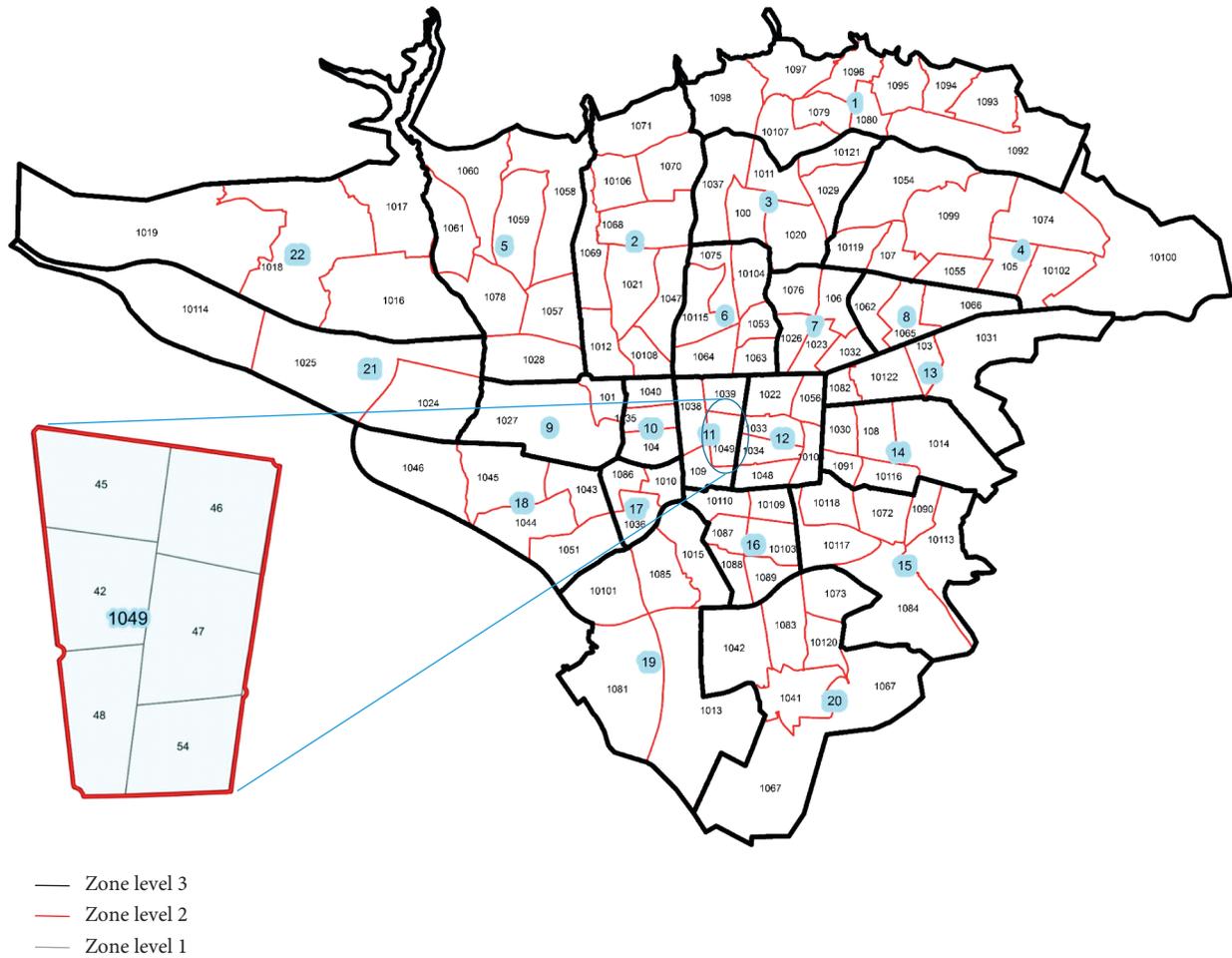


FIGURE 6: Zoning levels of Tehran.

TABLE 1: Properties of the similar clusters.

Cluster name	Feature
SLP Group1	Very low trip production-attraction potential
SLP Group2	Low trip production-attraction potential
SLP Group3	Medium trip production-attraction potential
SLP Group4	High trip production-attraction potential
SLP Group5	Very high trip production-attraction potential

local SLPSSI values of travel from zones with the lowest potential of trip production and attraction (SLP Group1) relative to other SLP groups.

According to Table 2, the highest SSIM between the trips on Sunday and Friday belongs to trips from zones with the lowest trip attraction and production (SLP Group1) to those with the highest trip attraction and production (SLP Group5). The properties of SLP Group5 zones seem to be able to attract trips regardless of weekdays. Therefore, it seems that urban traffic planners and policymakers should establish demand management policies between these two groups of zones independent of the weekdays. It is also apparent that, to improve the quality of the public transportation system’s services, a specific program should be

considered for public transportation lines between these pairs of zones on holidays. On the other hand, the lowest SSIM exists between the trips of Sunday and Friday from SLP Group1 to SLP Group3. This means that the pattern of trips from the origin of SLP Group1 zones to the destination of SLP Group3 zones on holidays is completely different from working days. Accordingly, the local SSIM between the trips from the origin of SLP Group1 to the destinations of SLP Group3 and SLP Group5 is 0.6345 and 0.9210, respectively.

According to Table 3, the SLP Group3 zones have the lowest SSIM with SLP Group1 zones. Thus, the trips from SLP Group3 zones to SLP Group1 zones follow a different pattern on holidays. It seems that the pattern of trips between the OD pair zones of SLP Group1 and SLP Group3 is completely different

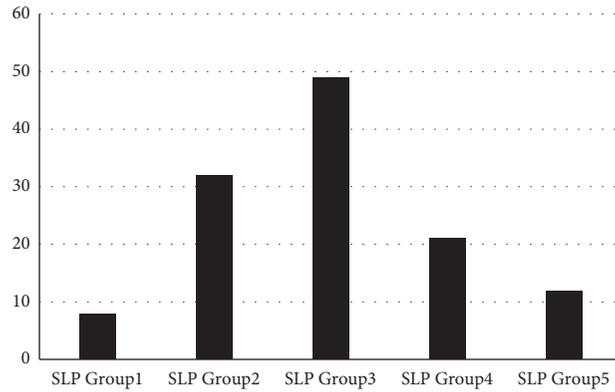


FIGURE 7: Number of zones in each SLP cluster.

Origin \ Destination	SLP Group 1							SLP Group 2					SLP Group 5					
	109	1013	1041	...	1067	10109	10110	101	102	...	10117	10120	...	100	...	10104		
SLP Group 1	109	558	2	9	...	3	12	3	54	43	...	4	96	...	3	...	231	
	1013	3	43	3	...	6	97	55	2	2	...	3	2	...	3	...	0	
	1041	120	2	125	...	32	9	4	7	2	...	3	45	...	57	...	97	

	1067	2	17	32	...	126	5	4	5	1	...	13	1	...	4	...	3	
	10109	14	51	42	...	5	989	248	4	25	...	34	12	...	10	...	26	
10110	3	30	15	...	3	140	404	1	4	...	7	3	...	2	...	8		
SLP Group 2	101	66	1	3	...	6	2	1	1389	9	...	13	9	...	1	...	30	
	102	36	1	1	...	0	10	2	24	335	...	36	21	...	4	...	39	
	
	
	10117	4	3	5	...	27	31	10	17	86	...	349	17	...	5	...	33	
10120	60	0	23	...	1	26	4	6	8	...	26	987	...	78	...	262		
SLP Group 5	
	
	
	10104	315	2	20	...	4	18	6	13	23	...	12	364	...	17	...	878	

FIGURE 8: Local windows of SLPSSI in sample Tuesday.

on holidays from working days. Therefore, public transportation program between these pairs of zones could be completely different on holidays. The evaluation of local windows in the SLPSSI method provides more information with more precise details about the travel pattern of each pair of zones.

4.2. Dimension Comparison of Local Windows Based on the Results. To display the sensitivity of MSSIM results considering the dimensions of the selected windows, the OD matrix obtained from GPS (GPS-OD) of an arbitrary Sunday is compared with two OD matrices of Monday and Friday. During this comparison, the dimensions of local windows vary from 2×2 to 122×122 as shown in Figure 10. The horizontal axis represents the chosen window dimensions, and the vertical axis shows the MSSIM values. The upper line

indicates the MSSIM results obtained from comparing Sunday and Monday. The lower line shows the MSSIM results of comparing Sunday with Friday. As can be seen, the MSSIM values increase with the increase in local window dimensions. In other words, larger dimensions of local windows lead to lower accuracy of MSSIM during determining the SSIM between two matrices. The change in the MSSIM value is more noticeable when comparing Sunday and Friday due to the different travel patterns on these two days. It can be concluded that when the dimensions of the selected local window are smaller, the similarity between two compared OD matrices is expected to be better identified.

The dimensions of the local windows in the SLPSSI method are determined according to the number of zones located in similar groups in terms of socioeconomic, land-use, and population properties. Meanwhile, the

Origin \ Destination		Tehran north area							Tehran west area					Tehran south area				
		100	1011	1020	...	10107	10119	10121	101	1012	...	1078	10114	...	1013	...	10120	
Tehran north area	100	167	169	425	...	1	10	22	1	102	...	11	1	...	5	...	37	
	1011	283	360	616	...	2	217	113	8	92	...	33	14	...	20	...	141	
	1020	495	400	622	...	3	210	109	4	202	...	22	16	...	77	...	42	

	10107	1	1	5	...	361	2	10	5	3	...	139	9	...	2	...	0	
	10119	14	213	180	...	4	1326	496	27	14	...	13	177	...	51	...	10	
10121	30	56	144	...	5	510	422	65	10	...	28	72	...	13	...	2		
Tehran west area	101	1	6	6	...	4	9	14	1389	1	...	5	286	...	1	...	9	
	1012	94	81	175	...	4	12	10	1	202	...	40	1	...	39	...	28	
	
	
	1078	24	15	37	...	157	26	240	19	31	...	361	14	...	12	...	2	
10114	0	13	10	...	19	342	17	130	2	...	8	233	...	2	...	2		
.	
.	
Tehran south area	1013	3	17	11	...	2	53	6	2	53	...	10	3	...	43	...	2	
	
	10120	78	271	135	...	0	9	4	6	9	...	2	1	...	0	...	987	

FIGURE 9: Local windows of GSSI in sample Tuesday.

TABLE 2: The SLPSSI values for local windows from the origin (SLP Group1), Sunday versus Friday.

Origin	Dest.				
	SLP Group1	SLP Group2	SLP Group3	SLP Group4	SLP Group5
SLP Group1	0.7497	0.7161	0.6345	0.7357	0.9210

TABLE 3: The SLPSSI values for local windows from the origin (SLP Group3), Sunday versus Friday.

Origin	Dest.				
	SLP Group1	SLP Group2	SLP Group3	SLP Group4	SLP Group5
SLP Group3	0.6631	0.6990	0.7304	0.7946	0.7923

dimensions of local windows in the GSSI are determined by the number of zones located in the same geographical location. These windows are fixed but not necessarily square. The dimensions of local windows are variable in both SLPSSI and GSSI methods. For instance, for SLP Group1 and SLP Group2, the local window size is 12×20 , while it is 29×8 for SLP Group3 and SLP Group5. Figure 11 demonstrates the results of comparing the average OD matrix of Sunday and other days of the week within one month of data gathering for the base MSSIM method with local windows of different sizes using SLPSSI and GSSI methods.

As shown in Figure 11, by focusing on the value of SSIM between the averaged OD matrices of Sundays, Thursdays, and Fridays, it can be found that GSSI calculates the value of SSIM of two matrices close to the local window of size 10×10 while SLPSSI calculates them close to the local window of size 5×5 in the MSSIM method. While having windowing with larger dimensions, the GSSI and SLPSSI methods produce results closer

to reality due to the smart and logical selection of zones inside a window.

4.3. Structural Comparison of Weekdays. In this section, GSSI, MSSIM, and SLPSSI methods are compared. For this purpose, the data of one week (from the 5 of October to 11 of October 2019) is chosen and used from the total data of one month. In this process, Sunday, which is an ordinary working day, is compared with other days of the week using the three methods. As mentioned before, 25 local windows with particular dimensions are defined for GSSI, and 25 local windows with dimensions corresponding with the number of zones in each window are defined for SLPSSI. In the MSSIM method, the whole matrix is considered a single window. Figure 12 and Table 4 provide the results of the comparison. According to Figure 12, all three methods are capable of recognizing the structural and numerical differences between OD matrices of daily trips. All three

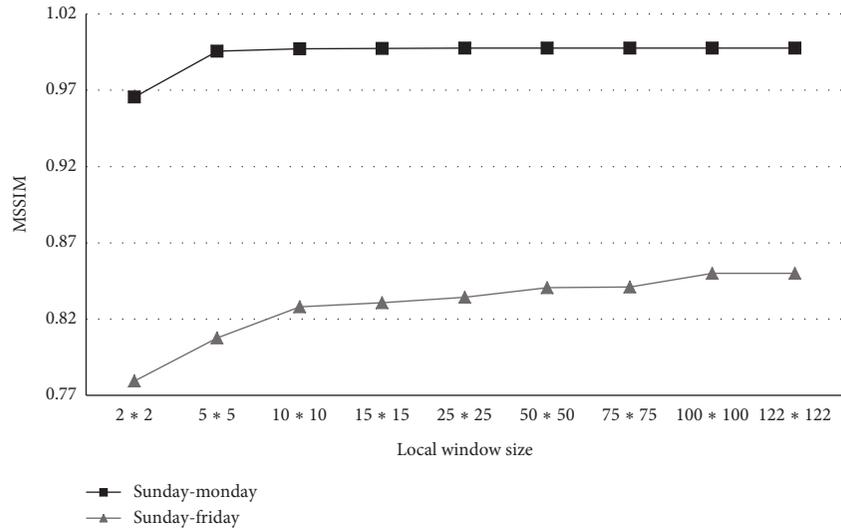


FIGURE 10: Comparison of MSSIM results considering variation in the dimensions of local windows.

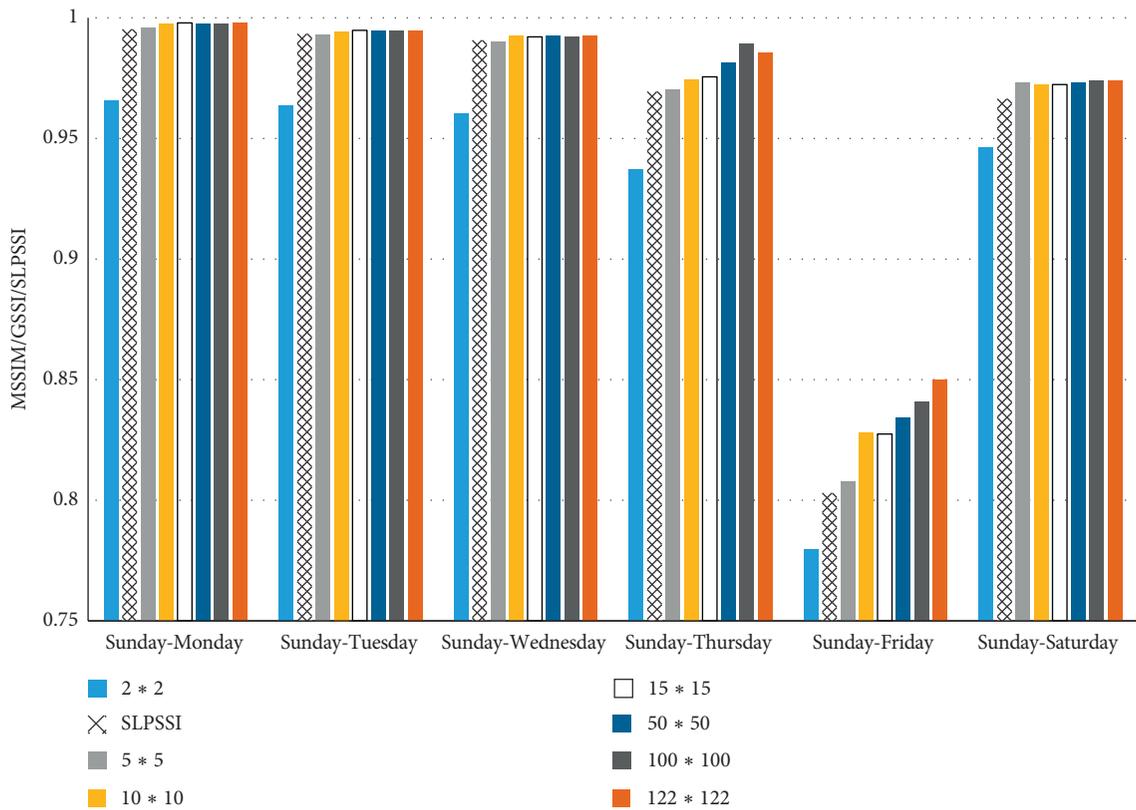


FIGURE 11: The SLPSSI/GSSI/MSSIM values according to the size of the selected local windows.

methods detect the difference in the travel patterns of Thursday and Friday in comparison to Sunday.

According to Table 4, the SLPSSI method catches the dissimilarity between Sunday and Friday more precisely. The values of SSIM of Sunday with Friday in MSSIM, GSSI, and

SLPSSI are 0.8500, 0.8074, and 0.7630, respectively. Therefore, the SLPSSI method is up to 10% more accurate than MSSIM and up to 5.5% more accurate than the GSSI method. Moreover, the utilized methods indicate the SSIM of all working days. Similar trip distribution patterns in

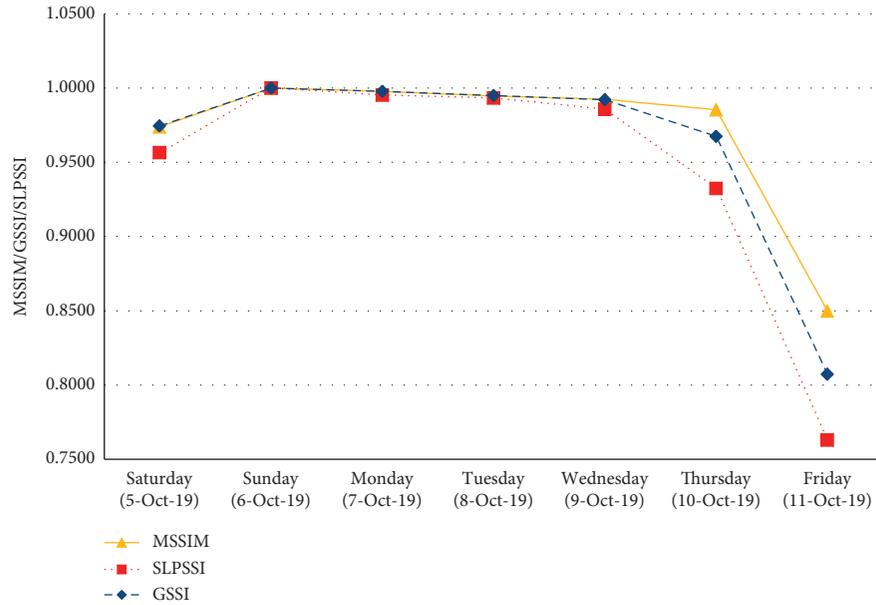


FIGURE 12: Comparison between daily matrices and the matrix of Sunday based on the GSSI, MSSIM, and SLPSSI methods.

TABLE 4: The value comparison of GSSI, MSSIM, and SLPSSI between Sunday and other days.

	Saturday (5-Oct-19)	Sunday (6-Oct-19)	Monday (7-Oct-19)	Tuesday (8-Oct-19)	Wednesday (9-Oct-19)	Thursday (10-Oct-19)	Friday (11-Oct-19)
MSSIM	0.9739	1.0000	0.9977	0.9947	0.9925	0.9853	0.8500
GSSI	0.9743	1.0000	0.9979	0.9948	0.9921	0.9675	0.8074
SLPSSI	0.9564	1.0000	0.9972	0.9943	0.9857	0.9324	0.7630

Tehran on Sunday, Monday, and Tuesday can be observed using all three methods. The structural difference of the OD matrices in the first working day (Saturday) and the last full working day (Wednesday) compared with Sunday can only be observed using the SLPSSI. It is worth reminding that part of enterprises are inactive on Thursday, and this day is considered a weekend for them.

According to Table 4, the SSIM calculated by MSSIM for Thursday is negligible (0.9853), although the travel patterns on this day are completely different from other working days. This difference has been recognized by GSSI (0.9675) and SLPSSI (0.9324). Moreover, given its precise classification of zones with respect to socioeconomic, land-use, and population properties, the SLPSSI is able to assign zones with more similarity to each group and design more logical local windows compared to GSSI. Therefore, it has given more accurate results in calculating the structural difference/similarity of OD matrices. Although all three methods detect the structural difference between Sunday and Friday's travel patterns, the details of these differences can only be evaluated by GSSI and SLPSSI. Tables 5 and 6 provide these details considering the local windows for the two methods.

According to Table 5, the lowest SSIM between Sunday (working day) and Friday (weekend) is observed in the trips with the origins and destinations in the south of

Tehran. On the other hand, the highest SSIM can be observed in the trips with origins and destinations in the west of Tehran.

According to Table 6, SLPSSI calculates the SSIM of the OD matrices of Sunday and Friday in each local window. It seems that the suggested method better identifies the lack of SSIM of the zones located in each local window. The lowest SSIM can be observed in the zones with medium socioeconomic, land-use, and population properties (SLP Group3) between Sunday and Friday. Furthermore, no noticeable structural difference is observed between the travel patterns of Sunday and Friday in zones with a high potential of trip production and attraction (SLP Group5). The lack of SSIM observed in the SLPSSI between the zones with different properties does not necessarily depend on their geographical locations. Therefore, it shows a new aspect of difference in daily travel patterns.

4.4. Computational Time. In the MSSIM method, by assuming $m \times n$ local windows from $M \times N$ matrix, $(M - m + 1) \times (N - n + 1)$ local windows could be extracted. For instance, $(122 - 2 + 1) \times (122 - 2 + 1) = 14641$, 2×2 local windows could be formed for calculations out of a 122×122 matrix. Meanwhile, there are only 25 local windows for

TABLE 5: SSIM obtained from GSSI for Sunday in comparison with weekend.

Origin	Dest.				
	Tehran center area	Tehran east area	Tehran north area	Tehran south area	Tehran west area
Tehran center area	0.8520	0.8161	0.8250	0.7589	0.9040
Tehran east area	0.8362	0.8232	0.7525	0.7290	0.8915
Tehran north area	0.8313	0.7226	0.7835	0.6855	0.8633
Tehran south area	0.7527	0.7466	0.6819	0.6609	0.8078
Tehran west area	0.8644	0.8735	0.8977	0.8370	0.9879

Sunday (6-Oct-2019) versus Friday (11-Oct-2019), GSSI method

TABLE 6: SSIM obtained from SLPSSI for Sunday in comparison with the weekend.

Origin	Dest.				
	SLP Group1	SLP Group2	SLP Group3	SLP Group4	SLP Group5
SLP Group1	0.7775	0.7115	0.6580	0.7215	0.8515
SLP Group2	0.7043	0.7560	0.7231	0.7439	0.8076
SLP Group3	0.6877	0.7249	0.7575	0.7618	0.7595
SLP Group4	0.7587	0.7179	0.7728	0.7666	0.8133
SLP Group5	0.8347	0.8118	0.7735	0.8064	0.8730

Sunday (6-Oct-2019) versus Friday (11-Oct-2019), SLPSSI method

calculating the SSIM using the GSSI and SLPSSI methods in this study. Calculations for each window are time-consuming, and the required computation time increases with the dimensions of selected windows. The smaller the dimensions of the selected local window, the longer the calculation time for the whole OD matrix. The shortest calculation time is for the 122×122 matrix (with a local window of the same size), and the longest calculation time is for a 2×2 local window in the MSSIM method. Figure 13 shows the computational time needed to compare 30 GPS-OD matrices with Sunday OD matrix on 6 of October 2019 using different window sizes for MSSIM, GSSI, and SLPSSI. Accordingly, the GSSI and SLPSSI methods are more efficient in terms of computational cost compared to the MSSIM with smaller window sizes. According to Figure 13, The SLPSSI approach is about 11 times faster than the MSSIM method with a 5×5 sliding window. There is no remarkable difference between the GSSI and SLPSSI methods in terms of calculation time.

4.5. Performance Evaluation of the Suggested Method with Sparse OD Matrices. In the previous sections, the performance of SLPSSI was discussed using a matrix aggregated in ZL 2. The OD matrix at this level is very dense. However, in most traffic flow patterns studies in cities, matrices in lower levels (e.g., ZL1) with smaller size are considered for calculations. The OD matrix at this level has remarkable sparsity. Thus, evaluation of the suggested approach performance at this level seems necessary. For this purpose, the OD matrix obtained from GPS data is evaluated in ZL 1. Since this OD matrix shows part of daily trips, it has the required sparsity for evaluation. Twenty-five local windows are used for each of the SLPSSI and GSSI methods. However, in the MSSIM method, the zones should be ordered according to trip production. Since the ordered matrices for the weekdays are different, the local window should be of the same size as the whole matrix.

According to Table 7, the highest similarity between Sunday and Friday matrices (0.8075) is shown by MSSIM. Meanwhile, the GSSI and SLPSSI methods have better recognized the discrepancy of these two days, that is, 0.7671 and 0.7048, respectively. Consequently, the proposed approach is able to catch the SSIM in sparse OD matrices in comparison to the GSSI method by up to 8%. The same discussion is true when comparing Sunday with Thursday. Furthermore, with a large size local window, MSSIM is not capable of detecting the structural difference between the traffic flow patterns of Sunday and Thursday. According to Table 7, when comparing Sunday and Monday, the MSSIM is 0.9478, which is slightly different from the MSSIM when comparing Sunday and Thursday (0.9431). The closeness of these values indicates the weakness of the MSSIM method in recognizing the structural difference between the OD matrices of Monday and Thursday. The SLPSSI method also better recognizes the structural difference of two sparse OD matrices. In sparse matrices, choosing zones with similar properties in a local window is more important. The structural similarities for dense matrices are shown in Table 1. The SSIM difference identified when comparing Sunday and Friday using the GSSI and SLPSSI methods for dense matrices is $0.8074 - 0.7630 = 0.0444$, while the difference for the sparse matrix is $0.7671 - 0.7048 = 0.0623$ according to Table 7. Therefore, the performance of the suggested method is much more suitable for sparse OD matrices as well.

5. Sensitivity Analysis

The sensitivity analysis is a process that measures the level of uncertainty and difference in the results of a mathematical model to the uncertainty and variation of input data [29]. To examine the suggested method's efficiency and robustness, a sensitivity analysis framework is designed in this section and the efficiency of the model considering any changes in the

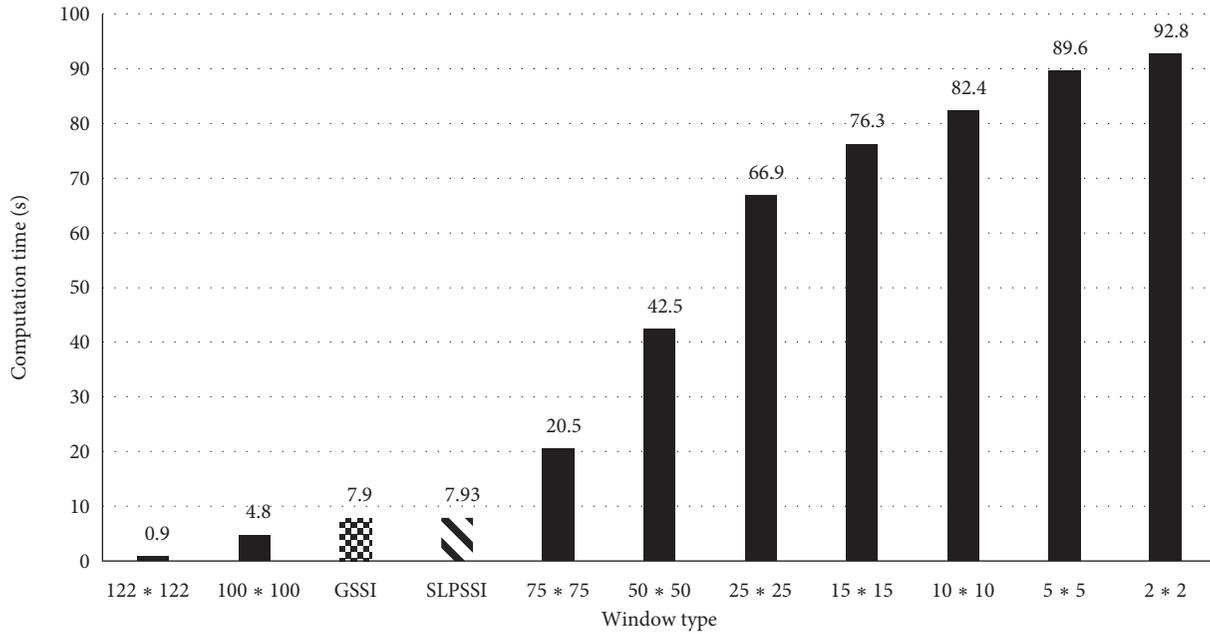


FIGURE 13: Comparison of computational time of different methods.

TABLE 7: Comparison of the GSSI, MSSIM, and SLPSSI methods on a sparse matrix at AL 1.

	Sunday versus Thursday	Sunday versus Friday	Sunday versus Monday
MSSIM	0.9431	0.8075	0.9478
GSSI	0.9191	0.7671	0.9480
SLPSSI	0.8858	0.7048	0.9455

input data is evaluated. The model efficiency is evaluated under different conditions, such as a change in the structure of the OD matrix or a change in the numerical values of its cells without any change in the structure. Therefore, the sensitivity of the SLPSSI and SLPSTR to the changes in the input variables is determined. The sensitivity analysis is performed on the OD matrix obtained from GPS data of Tehran. The OD matrix of the working day, Sunday 6 of October 2019, is considered the base matrix and shown by symbol X . This matrix aggregates in ZL 2, so it is a 122×122 matrix. The sensitivity analysis matrices are performed by manipulating the base matrix in two stages:

- (1) Applying a constant coefficient: in this mode, a constant coefficient is applied to the base matrix and the robustness of the suggested method in recognizing. The SSIM of the developed matrix is compared with the base matrix. Applying the constant coefficient 1, the base matrix is compared with itself (Mode 1). If a constant coefficient other than one is applied, a difference is formed in the numerical values of the base matrix cells, but the trip distribution structure does not change (Mode 2).
- (2) Applying random coefficients on the base matrix: in this case, the changes are applied to the values of the base matrix cells and their trip distribution structure (Mode 3).

5.1. Sensitivity Analysis Design for SLPSSI. This section discusses the sensitivity analysis of SLPSSI. The suggested method's robustness is measured in each of the three modes defined in the previous sections.

5.1.1. Applying Constant Coefficients. The sensitivity of SLPSSI and SLPSTR is measured by applying constant coefficients. Matrix Y is determined from matrix X , by multiplying it to a constant factor α between 0.1 and 2, $Y = \alpha \times X$. In this case, the suggested method will be efficient and robust if

- (1) SLPSTR equals 1 for any value of α
- (2) SLPSSI equals 1 for $\alpha = 1$, and SLPSSI is lower than 1 for any other factor

Figure 14 shows the values of SLPSSI and SLPSTR after applying constant values for α . As can be seen, by applying different factors, the SLPSTR does not vary, which indicates that no change in the structure of the two matrices occurs after applying a constant factor. By multiplying one, the SLPSSI is still one while it alters with values lower than one. Moreover, the factors lower than one are more effective in reducing the SSIM between two matrices compared to the factors more than one. Thus, the suggested method is efficient and robust in the case of applying constant factors.

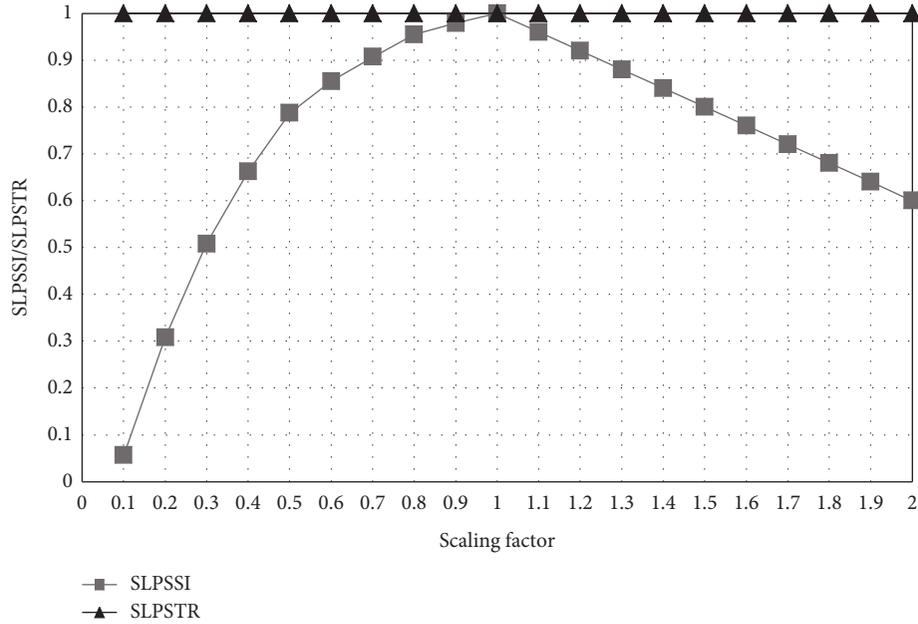


FIGURE 14: The sensitivity of SLPSSI and SLPSTR to applying constant coefficients.

5.1.2. Applying Variable Coefficients. In this case, the sensitivity of SLPSSI and SLPSTR is evaluated in three probable scenarios, which are usually used in modeling the traffic demand [30]. The random coefficients in four modes are considered equal to $\theta = [5\%, 10\%, 15\%, 20\%]$ for each of these scenarios:

- (1) A scenario for OD matrices obtained from old studies (low demand)
- (2) A scenario for OD matrices similar to the base matrix (medium demand)
- (3) A scenario for OD matrices in heavy traffic conditions in the network (high demand)

For each of the above modes, the base matrix X is compared with 100 replication of matrix Y and the mean SLPSSI and SLPSTR are calculated for each scenario.

Scenario with low demand: in this scenario, SLPSSI and SLPSTR for matrices X and $Y_{i,\theta}^l$ are calculated.

$$Y_{i,\theta}^l = X(0.6 + \theta \times \text{rand}[0, 1]) \text{ and } i \in [1, 100]. \quad (9a)$$

For example, for $\theta = 20\%$, $Y_{i,\theta}^l$ varies from 60% to 80% of X .

Scenario with medium demand: in this scenario, SLPSSI and SLPSTR for X and $Y_{i,\theta}^m$ matrices are calculated.

$$Y_{i,\theta}^m = X(0.8 + \theta \times \text{rand}[0, 1]) \text{ and } i \in [1, 100]. \quad (9b)$$

For example, for $\theta = 20\%$, $Y_{i,\theta}^l$ varies from 80% to 100% of X .

Scenario with high demand: in this scenario, SLPSSI and SLPSTR for X and $Y_{i,\theta}^h$ matrices are calculated.

$$Y_{i,\theta}^h = X(1.05 + \theta \times \text{rand}[0, 1]) \text{ and } i \in [1, 100]. \quad (9c)$$

For example, for $\theta = 20\%$, $Y_{i,\theta}^h$ varies from 105% to 125% of X .

The proposed method is efficient and robust when the SSIM/difference varies with respect to the value of the simulated matrix in each scenario for both SLPSSI and SLPSTR. This change is appropriate when the difference between the two matrices rises with an increase in the random coefficient; that is, the SSIM between the two OD matrices reduces. The results of calculating the average SLPSSI and SLPSTR for the iterations conducted for different demand scenarios and random coefficients can be observed in Table 8.

According to Table 8, with an increase in θ (in each scenario), the average SLPSSI between the base matrix and the one made by random coefficients decreases. This trend can also be observed in the average of SLPSTR, which only compares the structure of two matrices. Therefore, the suggested method has the required efficiency and robustness to detect different SSIM of matrices by applying random coefficients.

TABLE 8: The average SLPSSI and SLPSTR for successive iterations in the three scenarios.

	Low demand scenarios				Medium demand scenarios				High demand scenarios			
	$\theta=0.05$	$\theta=0.1$	$\theta=0.15$	$\theta=0.2$	$\theta=0.05$	$\theta=0.1$	$\theta=0.15$	$\theta=0.2$	$\theta=0.05$	$\theta=0.1$	$\theta=0.15$	$\theta=0.2$
Mean of SLPSSI for 100 replications	0.7859	0.7775	0.7438	0.7205	0.9583	0.9432	0.9278	0.9029	0.9963	0.9874	0.9754	0.9594
Mean of SLPSTR for 100 replications	0.9932	0.9731	0.9523	0.9326	0.9965	0.9813	0.9754	0.9543	0.9989	0.9895	0.9845	0.9775

6. Conclusion

The presented study outlines a new method called SLPSSI for OD matrices comparison. A comprehensive evaluation of the similarity or difference between two OD matrices should recognize the difference between them numerically (difference in the value of each cell of the matrix) and structurally (the difference between trips distribution in the whole matrix). In general, the traditional methods only calculate the numerical deviation of two matrices and do not consider their structural differences. A limited number of studies have investigated the structural differences of matrices. Most of these methods have been taken from other scientific fields than transportation engineering including MSSIM. This method recognizes the SSIM of two matrices by choosing the local windows and their movements on two matrices. The accuracy of the MSSIM results depends on the dimensions of the selected windows. Supplementary methods like GSSI have also been developed on MSSIM. In this method, local windows are defined only concerning the geographical locations of zones. The geographical windows cannot put the zones with the same characteristics in one group in the best form. Thus, the presented paper has suggested a new method to classify the zones concerning the similarity between the socioeconomic, land-use, and population properties. The following can be concluded:

- (1) The suggested method is capable of detecting travel patterns in local windows. This means that, given the logical selection of local windows, it allows for analyzing the travel patterns between zones in each pair of local windows.
- (2) The proposed method has a lower computational time compared to the base MSSIM method. By choosing fixed local windows with particular dimensions, as GSSI, the suggested method enjoys higher computational speed. It is shown that the proposed approach is 11 times faster than the MSSIM method.
- (3) The method is capable of identifying the SSIM of the weekdays with higher accuracy compared to the previous methods. The main goal of designing this method is to provide the ability to detect the SSIM (numerical values of cells and trip distribution) between OD matrices. The SLPSSI method precisely recognizes the SSIM/difference between two matrices by evaluating the OD matrices of different days of a week with various patterns. The proposed method is up to 10% and 5.5% more accurate than MSSIM and GSSI methods, respectively.

- (4) The proposed method could also be used for sparse matrices and is capable of catching the SSIM in sparse OD matrices in comparison to the GSSI method up to 8%.
- (5) The sensitivity analysis results proved that the proposed SLPSSI approach is a robust statistical measure and is readily deployable to practical applications that involve OD matrices comparison. The sensitivity analysis results indicate that the suggested method is sensitive to applying both constant and random coefficients. Applying constant coefficients shows that the SLPSSI could be used to compare similar matrices, for example, two working days of a week. The sensitivity of the proposed method by applying random coefficients proved its ability to recognize the similarity/difference between different matrices. Therefore, it could be used in the validation of predicted matrices. Thus, SLPSSI is a robust index and reliable from a statistical point of view.

Data Availability

The datasets are available from the corresponding author upon reasonable request.

Disclosure

This paper's conclusions reflect the authors' understandings, who are responsible for the accuracy of the research findings.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge Tehran Municipality and Rajman Information Structures Company for providing the required data for this research.

References

- [1] J. Barceló, L. Montero, M. Bulejos, M. P. Linares, and O. Serch, "Robustness and computational efficiency of kalman filter estimator of time-dependent origin-destination matrices," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2344, no. 1, pp. 31–39, 2013.
- [2] B. Hellinga and M. V. Aerde, "A statistical analysis of the reliability of using rgs vehicle probes as estimators of dynamic O-d departure rates," *I V H S Journal*, vol. 2, no. 1, pp. 21–44, 1994.

- [3] K. Ashok and M. E. Ben-Akiva, "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows," *Transportation Science*, vol. 36, no. 2, pp. 184–198, 2002.
- [4] R. Frederix, F. Viti, W. W. E. Himpe, and C. M. J. Tampère, "Dynamic origin-destination matrix estimation on large-scale congested networks using a hierarchical decomposition scheme," *Journal of Intelligent Transportation Systems*, vol. 18, no. 1, pp. 51–66, 2014.
- [5] C. Antoniou, M. Ben-Akiva, and H. N. Koutsopoulos, "Incorporating automated vehicle identification data into origin-destination estimation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1882, no. 1, pp. 37–44, 2004.
- [6] E. Cascetta, "Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator," *Transportation Research Part B: Methodological*, vol. 18, no. 4–5, pp. 289–299, 1984.
- [7] S.-J. Kim, W. Kim, and L. R. Rilett, "Calibration of micro-simulation models using nonparametric statistical techniques," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1935, no. 1, pp. 111–119, 2005.
- [8] M. Cools, E. Moons, and G. Wets, "Assessing the quality of origin-destination matrices derived from activity travel surveys," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2183, no. 1, pp. 49–59, 2010.
- [9] M. Nigro, E. Cipriani, and A. del Giudice, "Exploiting floating car data for time-dependent Origin-Destination matrices estimation," *Journal of Intelligent Transportation Systems*, vol. 22, no. 2, pp. 159–174, 2018.
- [10] J. Barceló, L. Montero, M. Ballejos, O. Serch, and C. Carmona, "A kalman filter approach for exploiting bluetooth traffic data when estimating time-dependent od matrices," *Journal of Intelligent Transportation Systems*, vol. 17, no. 2, pp. 123–141, 2013.
- [11] A. Tavassoli, A. Alsger, M. Hickman, and M. Mesbah, "How close the models are to the reality? Comparison of transit origin-destination estimates with automatic fare collection data," in *Proceedings of Australasian Transport Research Forum*, Melbourne, Australia, November 2016.
- [12] X. Ros-Roca, L. Montero, A. Schneck, and J. Barceló, "Investigating the performance of SPSA in simulation-optimization approaches to transportation problems," *Transportation Research Procedia*, vol. 34, pp. 83–90, 2018.
- [13] K. N. S. Behara, A. Bhaskar, and E. Chung, "Geographical window based structural similarity index for origin-destination matrices comparison," *Journal of Intelligent Transportation Systems*, pp. 1–22, 2020, inpress.
- [14] K. N. S. Behara, A. Bhaskar, and E. Chung, "A novel approach for the structural comparison of origin-destination matrices: Levenshtein distance," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 513–530, 2020.
- [15] T. Djukic, S. Hoogendoorn, and H. Van Lint, "Reliability assessment of dynamic OD estimation methods based on structural similarity index," in *Proceedings of Transportation Research Board 92nd Annual Meeting*, Washington, DC, USA, January 2013.
- [16] A. Ruiz de Villa, J. Casas, and M. Breen, "OD matrix structural similarity: wasserstein metric," in *Proceedings of Transportation Research Board 93rd Annual Meeting*, Washington, DC, USA, January 2014.
- [17] T. Djukic, "Dynamic OD demand estimation and prediction for dynamic traffic management," PhD Thesis, Delft University of Technology, Delft, Netherlands, 2014.
- [18] Y. Hollander and R. Liu, "The principles of calibrating traffic microsimulation models," *Transportation*, vol. 35, no. 3, pp. 347–362, 2008.
- [19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [20] T. Pollard, N. Taylor, and T. V. Vuren, "Comparing the quality of OD matrices in time and between data sources," in *Proceedings of the European Transport Conference*, AET, Frankfurt, Germany, January 2013.
- [21] T. Van Vuren and T. Day-Pollard, "256 shades of grey-comparing OD matrices using image quality assessment techniques," in *Proceedings of 2015 Scottish Transport Applications Research Conference*, Glasgow, Scotland, May 2015.
- [22] F. M. Oliveira-Neto, L. D. Han, and M. K. Jeong, "Online license plate matching procedures using license-plate recognition machines and new weighted edit distance," *Transportation Research Part C: Emerging Technologies*, vol. 21, no. 1, pp. 306–320, 2012.
- [23] I. Markou, K. Kaiser, and F. C. Pereira, "Predicting taxi demand hotspots using automated Internet Search Queries," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 73–86, 2019.
- [24] A. Zhang, J. E. Kang, K. Axhausen, and C. Kwon, "Multi-day activity-travel pattern sampling based on single-day data," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 96–112, 2018.
- [25] H. Dong, M. Wu, X. Ding et al., "Traffic zone division based on big data from mobile phone base stations," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 278–291, 2015.
- [26] S. Afandizadeh Zargari and F. Safari, "Using clustering methods in multinomial logit model for departure time choice," *Journal of Advanced Transportation*, vol. 2020, Article ID 7382569, 12 pages, 2020.
- [27] Y. Asakura and E. Hato, "Tracking survey for individual travel behaviour using mobile communication instruments," *Transportation Research Part C: Emerging Technologies*, vol. 12, no. 3–4, pp. 273–291, 2004.
- [28] B. Sana, J. Castiglione, D. Cooper, and D. Tischler, "Using google's passive data and machine learning for origin-destination demand estimation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 46, pp. 73–82, 2018.
- [29] Z.-P. Du and A. Nicholson, "Degradable transportation systems: sensitivity and reliability analysis," *Transportation Research Part B: Methodological*, vol. 31, no. 3, pp. 225–237, 1997.
- [30] T. Djukic, M. Ballejos, L. Montero, and S. Hoogendoorn, "Advanced traffic data for dynamic OD demand estimation: the state of the art and benchmark study," in *Proceedings of TRB 94th Annual Meeting Compendium of Papers*, Washington, DC, USA, January 2015.