

Article

Model for Estimating Urban Mobility Based on the Records of User Activities in Public Mobile Networks

Krešimir Vidović ¹, Marko Šoštarić ^{2,*}, Sadko Mandžuka ³ and Goran Kos ⁴

¹ Ericsson Nikola Tesla, Krapinska 45, 10000 Zagreb, Croatia; kresimir.vidovic@ericsson.hr

² Department of Transport Planning, Faculty of Transport and Traffic Sciences, University of Zagreb, Vukelićeva 4, 10000 Zagreb, Croatia

³ Department of Intelligent Transport Systems, Faculty of Transport and Traffic Sciences, University of Zagreb, Vukelićeva 4, 10000 Zagreb, Croatia; smanduka@fpz.hr

⁴ Institute for tourism, Vrhovec 4, 10000 Zagreb, Croatia; goran.kos@iztzg.hr

* Correspondence: msostaric@fpz.hr

Received: 20 November 2019; Accepted: 20 January 2020; Published: 22 January 2020



Abstract: Urban mobility of a population is usually estimated within procedures that focus on specific domains, using limited datasets, indicators, and indices related to the targeted subsets of the urban population. This paper proposes a new approach to urban mobility estimation, based on the telecommunication activities within the public mobile telecommunication networks. The urban mobility indicators in this research are generated from a database of mobile phone users call data records and are integrated into the urban mobility index of the population based on the model defined through the adaptive neuro-fuzzy inference system (ANFIS). The following has been considered in the process: an initial fuzzy inference system, model learning, model quality control, limitations, errors, and deficiencies. The model is practically applied in the programming environment, on a set of real word data. The research results prove the following hypothesis set in this paper: the urban mobility of inhabitants in a specific timeframe, can be described with an urban mobility index based on the data on the recorded telecommunication activities of the public mobile communication network users.

Keywords: urban mobility estimation; indicators; index; call data records; telecommunications

1. Introduction

The sustainable mobility of the urban population is an important and integrated part of social and economic life, directly affecting the economy and life quality, especially in large agglomerations. Therefore, understanding of key urban mobility features has an extremely important role from the perspective of sustainable urban transport organization and development, in line with the objectives of reducing traffic congestions, mitigating negative environmental and health impacts, and achieving transport-related time and cost savings. Urban mobility is the ability of an individual to move within an urban space in an organized and meaningful manner, in line with their physiological, intellectual, and socioeconomic needs and by utilizing the available transport, community, and information communication infrastructure. Urban mobility of the population is usually estimated within the procedures focused on specific domains (traffic, logistics, ecology, society, etc.) using limited datasets, indicators, and indices related to the targeted subgroups of the urban population (public transport users, cyclists, drivers, etc.). The objective and the purpose of the urban mobility estimation is the identification of the weaknesses and strengths of the transport system, with its key or critical elements. Urban mobility estimation provides valuable insight and enables a better understanding of the urban mobility system, thus, establishing the foundation to improve the existing system and create new

mobility services through the approach of characteristics for intelligent transport systems. This research paper proposes a new approach to urban mobility estimation based on the data in the recorded telecommunication activities of the public mobile communication network users. The purpose of this research is to verify the potential use of the data on the telecommunication activities of the users in the urban mobility estimation. The research objective is to define a model for estimating the urban mobility of the population as a quantitative parameter derived from the data on the recorded telecommunication activities of the public mobile communication network users (represented by the call data records, CDR), and to define a new index for the urban mobility estimation.

2. Overview of the Previous Research

The scientific community defines urban mobility in multiple ways. It represents the movement of persons between an origin and different destination points, in different time periods, by different transport modes, with different objectives [1–5]. Total mobility is related to an individual and represents the total number of trips within the observed territory, regardless of the transport mode and purpose, usually observed daily [1,6–8]. Mobility can also be defined as a measurement of transport system efficiency to connect spatially separated locations, where mobility is used as a key performance indicator of the transport system functionality [9]. It represents a broader concept than transport because it not only relates to the movement itself but also to the potential for movement and the spatial, economic, and social context of movement. Therefore, the concept of mobility, by analogy, is more comprehensive than transport [10]. Every increase in the number of trips, passenger per transport kilometers, and the travel or transport speed contributes to mobility [11–13]. Mobility is the key element of the transport system [14]. Finally, urban mobility is defined as the ability of an individual to move in an urban space in an organized and meaningful manner, in line with their physiological, intellectual, and socioeconomic needs utilizing the available transport, community, and information communication infrastructure [15,16].

Galileo Galilei wrote “Count what can be counted, measure what can be measured and what is not measurable, make measurable” [17]. Urban mobility needs to be assessed in order to identify the weaknesses and strengths of the transport system, to identify key or critical elements, to provide insight and provide a better understanding of the urban mobility, and to create a foundation for upgrading the existing and creating new services in the field of intelligent transport systems [18].

Research to date has shown that urban mobility is generally estimated by primarily identifying the relevant urban mobility indicators and then forming an urban mobility index (urban mobility assessment). The first step in calculating the urban mobility indicators is to identify and collect urban mobility data [5,6]. Urban mobility data are heterogeneous in nature and are collected from different sources, which are in principle divided into three groups [19–21]. The first group is the conventional data sources that include the already established data sources, such as sensors on transport infrastructure, meteorological sensors, environmental sensors, etc. [19]. The second group of data sources includes information and communication systems of the urban mobility participants. Examples of such systems are the road infrastructure charging system, smart ticketing system for the urban public passenger transport, and others [21]. The third group of data consists of the so-called new data sources, covering the use of relatively new technologies, which have not been widely used in the urban mobility assessment (e.g., public mobile telecommunications networks, social networks, and autonomous vehicles) [22–28]. The primary function of the second and third group data sources is not to collect or generate urban mobility data, but, due to nature of the encompassed systems, to generate data usable for performing the urban mobility estimation. One of the data sources in the category of the so-called new sources is data from the public mobile communications networks, whose potential has been recognized by a great number of authors [12,13,22,25,27,29–56]. The benefit of using this dataset is the elimination of disadvantages attributed to the conventional urban mobility data collection methods (e.g., surveys, behavioral research, territorially restricted, and cross-sectional measurements), already proven to be economically inefficient because they use a small sample of respondents, have

limited spatial and temporal coverage, and are often conducted within inappropriate time intervals (large gaps between data collection). The general conclusion is that this data source has great potential in the urban mobility research, primarily because it covers a very large population sample and, if linked to the information on base station position, can provide information on the movement of users. In addition, because of an increase in the number of mobile phone users, the difference between the number of the mobile phone owners and the number of persons without a mobile phone has been reduced and almost completely eliminated, thus, making this data source even more suitable for the mobility research. In the last two decades, the authors have paid considerable attention to the possibility of using data from the telecommunication services billing records for monitoring the traffic flows, urban planning, etc. By comparing the indicators generated in the aforementioned manner with the ones generated from the traditional sources, the authors proved that the sample included in the CDR record is representative, the reliability of the method is high, and the correlation coefficient is high. Different indicators have been proposed that can be derived from a data subset, although the procedures for obtaining and calculating them are not unified.

The next step is related to the mobility indicators. An indicator is a generic name for a parameter indicating the state or a change of state in a specific system, i.e., process. An indicator effectively and easily monitors changes over a defined period of time. The aim of the indicator is to present mobility and trends in a concise, simple, and understandable way [13,57–62]. The mobility indicators can be grouped according to the domains to which they are assigned. Transport indicators are very significant for the assessment, representing mobility through categories related to the transport aspect and showing the impact of these components on mobility [4,5,14,63–67].

The next step in an urban mobility estimation is to establish an urban mobility index based on the urban mobility data and indicators. The urban mobility index is defined as the result of a fusion of the relational data or mobility indicators, in order to estimate mobility. Indices are used to reduce the complexity and scope of the data and indicators included in the urban mobility assessment. The advantages of using indexes are the following: Their potential to reduce a significant amount of information of a different nature to a single scalar; the possibility to use them for the evaluation of the complex systems as a whole, rather than at the level of the individual components; and the possibility of comparisons of the indices obtained by the same methodology [68]. Several mobility indices can be found in scientific literature, representing mobility through a unique number generated by a corresponding mathematical model from the data, indicators, or other indices, for example, sampling mobility index (SMI) [69], travel time index (TTI) [70], Harde mobility city index [71], urban mobility index [64], smart transport cities [56], etc.

However, the conclusion is that there is generally no accepted index for estimating urban mobility. Existing urban mobility indicators and indices are generally developed for specific cities or urban areas, taking into account the local specificities, making their application limited to that area only. Regardless of the great number of defined mobility indicators, only a limited number is used to develop an urban mobility index. In addition, the choice of the urban mobility indicators is mainly related to the availability of data, i.e., the existence of a certain data source, which can limit the objectivity and comparability of the mobility assessment within urban agglomerations where such data or sources do not exist. Urban mobility assessments mainly use conventional data sources, and the potential of new data sources is not sufficiently exploited [15]. Thus, no studies to date mention the development of an urban mobility index based on the data from the billing records of the public mobile telecommunication networks (CDRs), which could be used for urban mobility estimation as a result of fusion between the relational mobility indicators (number of trips, trip duration, and Euclidean distance).

3. Research Methodology

This section defines the research methodology providing an overview of the methodology and focuses on determining the relation between urban mobility indicators and an urban mobility index as a function of urban mobility assessment process.

3.1. The Methodology Overview

This section defines the methodology for establishing an expert system for the urban mobility index calculation (Figure 1). An expert system is defined as a system in which human knowledge is embedded and applied to solve problems in a specific area, in a way similar to a human expert [72]. A corresponding subset of information was extracted from the telecommunication activity records of the public mobile communication networks, which contained the following information: the time of the telecommunication activity, the location of the base station to which the user was connected while performing the telecommunication activity, and a designation concealing the real identity of the user. The information provided is used to define the following urban mobility indicators: the number of migrations (number of trips), duration, and Euclidean distance. These indicators are used as input for the model. The first indicator represents a number of trips. An algorithm is used to determine this indicator, and the result is a trip matrix whose fields contain the total number of trips between the specific pairs of base stations (in the associated urban space), in a predefined timeframe. The second indicator is a trip duration indicator, calculated as a median value of the duration of the trip carried out over a certain period of time between specific base station pairs. The third indicator represents the (Euclidean) distance between the base station pairs. The calculated indicator values are used as inputs for the calculation within the urban mobility estimation, i.e., they are used as inputs for an expert system-based model, which is to be developed in the next step. The following step defines an expert system that uses the predefined indices, putting them in the appropriate relations in order to assess the overall urban mobility, on the basis of the “knowledge learned”. The learning includes a process of “knowledge extraction” from a wide range of experts in the field of urban mobility. Technically, this acquisition of knowledge is carried out through a survey, and an expert system is established based on the survey results, using the fuzzy logic method, i.e., the so-called ANFIS machine learning technique. The result is an expert system that defines and develops a model for calculating an urban mobility index, based on the indicators derived from the telecommunication activity records of the public mobile communication network users. According to the proposed model, the urban mobility index is calculated for each base station pair (of the corresponding urban wholes), and as such, represents a share in the total mobility index, therefore, called a partial index (partiality refers to taking into account only a part of the observed agglomeration space). The values of the partial urban mobility indices are, then, used to calculate the urban mobility index of the whole area for a predefined timeframe.

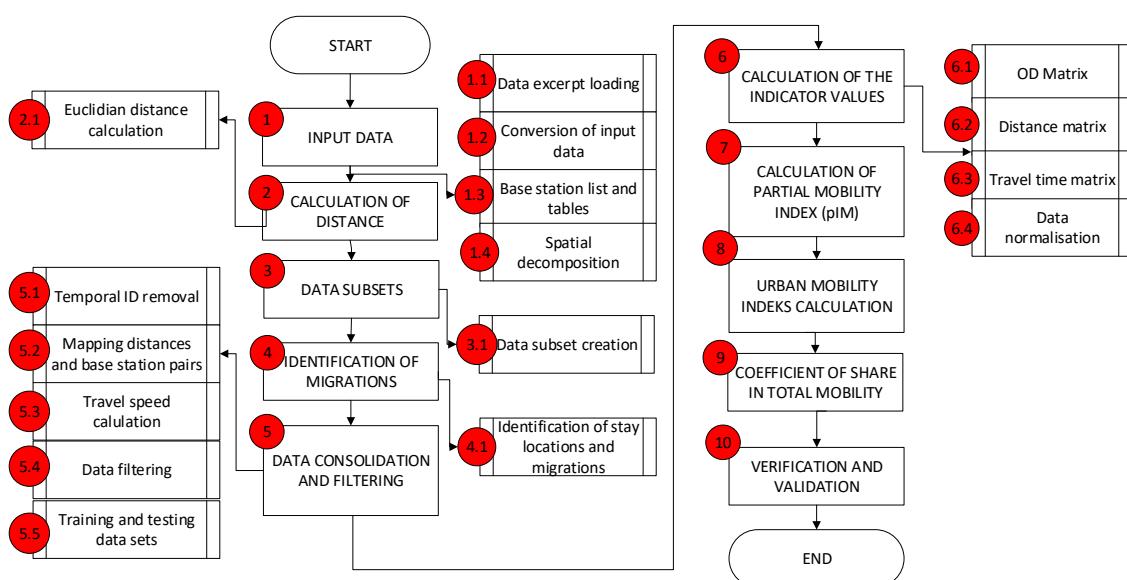


Figure 1. Flowchart of the urban mobility index determination.

In step 1.1 the input to the model is a subset of the CDR records that are ex-ante anonymized (depersonalized) and from which all information is removed except for the temporary identification of the user, the timing of the telecommunication activity, and the geographical location of the base station to which the user was connected while performing the telecommunication activity. Data taken from [73]. In step 1.2, the data file is transformed into a table. In step 1.3, the base stations stored in the associated table are identified. In step 1.4, the coverage area of each base station is identified, using the so-called Voronoi procedure [42,74,75]. The calculation of the indicators primarily takes place at the level of all base station pairs. In step 2.1, the Euclidean distance between each pair of base stations is calculated in order to get their linear distances. The distance between the two locations is calculated using the Haversin formula [76]. In the next step, the data is divided into subsets, since it is possible to define different timeframes within one day, depending on the purpose of the analysis. In step 4.1, the changes of the location are identified for each trip and user, in every individual dataset [51]. Within the same step, an indicator B , trip duration, is defined as the time difference between the recorded start of the activity at the origin cell and the recorded start of activity at the destination cell for each individual trip. It is calculated by using Relation (1):

$$B = \Delta t(i, j) = t(i)_{dest} - t(i)_{orig} \quad (1)$$

where:

- B is an indicator of the duration of the journey for each individual trip;
- $t(i)_{orig}$ is a time denotement in which the user activity was recorded at the base station and identified as the trip origin;
- $t(i)_{dest}$ is a time denotement in which the user activity was recorded at the base station and identified as the trip destination;
- $\Delta t(i, j)$ is the travel time between two base stations.

Within steps 5.1 and 5.2, the tables are merged and the redundant data removed. The results of this procedure are the trip values with the assigned distances between the base station pairs. Next, in step 5.3, the travel speed for each individual trip is calculated as a quotient of the linear distance value and Δt , as shown in the Equation (2).

$$D1(i, j) = \frac{C(i, j)}{B(i, j)} \quad (2)$$

where:

- $D1(i, j)$ is the speed of movement for each individual journey within a timeframe;
- $C(i, j)$ is the Euclidean distance between base station pairs;
- $B(i, j)$ is the duration of the journey, for each individual trip within a timeframe.

In step 5.4, the filtering follows. The data in which the time difference between the two recorded events is less than the defined time threshold is being filtered. All trip durations under 10 min, or longer than 60 min are also filtered [38,51,74], similar to the data in which the linear spatial distance between the base stations is less than the defined spatial threshold. The spatial threshold is defined as 1 km [38,51,74]. The filtering is done over data in which the calculated travel speed is not realistic in the context of urban migrations. All trips with the travel speed greater than 100 km/h are filtered.

Then, in step 5.5, the dataset is divided into two sets, namely the calibration set, and the model validation set. The calibration set consists of 80% of the total dataset, and the validation set consists of the remaining 20%. All further steps are performed on both datasets for each timeframe. In step 6, the value of the indicator for each timeframe is calculated and the data is stored in three different matrices for each timeframe. The first is the trip matrix (O-D matrix), the second is the distance matrix, and the third is the time matrix. In step 6.1. a trip matrix (O-D Matrix) is defined, containing the sum of all

trips between the individual base station pairs within the appropriate time period. The indicator of the number of trips between the individual base station pairs $A(i, j)$ is calculated using Relation (3):

$$A(i, j) = \sum_{i=0}^n Trip_ID(i, j) \quad (3)$$

where:

- $A(i, j)$ is the total number of the recorded (registered) trips between the individual base station pairs (trip matrix);
- $Trip_ID(i, j)$ is a trip between two base stations in an appropriate timeframe, all assumptions being met.

In step 6.2, a distance matrix is calculated, containing distances of all base station pairs between which a trip is generated within an appropriate time period. In step 6.3, a trip duration matrix is formed. It contains values of the median trip duration between all base station pairs for the corresponding time period, calculated according to the methodology described in step 4.1.

The indicator \bar{B} is calculated using Expression (4):

$$\bar{B}(i, j) = \frac{\sum_{i=0}^n B(i, j)}{n} \quad (4)$$

as the median value of all travel times between base station pairs within the appropriate time period.

Where:

- $\bar{B}_{i,j}$ is the median value of all travel times between each pair of base stations over a timeframe;
- $B(i, j)$ is the value of the trip duration for each individual trip in a timeframe;
- n is the number of trips registered between each base station pair in a timeframe.

In step 6.4, the data normalization is performed. Previous steps result in values stored in three matrices, for each time period. Since the day is divided into eight time periods, a total output from the previous step includes a total of 24 matrices (8×3), which are then used in the urban mobility index calculation. In order to ensure that the model is applicable to all urban areas, regardless of their size, the obtained data need to be normalized, i.e., reduced to an interval [0,1]. The normalization is performed by normalizing the value of an indicator over a given interval with the highest value of an individual indicator for all eight intervals. The purpose of carrying out the normalization carried in this way is to include specific data in the calculation, i.e., to take into consideration to what extent an individual indicator participates in the total mobility of an urban agglomeration in each individual timeframe, and to identify its maximum. The normalization of the number of trips is performed by identifying the maximum value of the trip number indicator for a specific pair of origins and destinations within all eight time periods. These values are then stored into a new matrix called the matrix of the maximum trip quantities. The normalized trip quantity values are calculated by dividing the indicator values assigned to the eight matrices for each time period with the value of the maximum travel quantity matrix, thus, obtaining the normalized travel matrices for each time period. In step 6.6, the normalization of the trip duration indicators is carried out by identifying the maximum values of the trip duration indicator for each pair of origins and destinations within all predefined time periods, which is called the matrix of the maximum trip duration. This is followed by the calculation of the normalized trip duration value, by dividing the indicator values assigned to the eight matrices for each time period with the value of the maximum trip quantities matrix, thus, obtaining the normalized trip duration matrices for each time period. Then, in step 6.7, a normalization of the distance indicators is carried out. Unlike the two previously described indicators, the value of the distance indicator does not change over time. Therefore, the normalization is calculated in a way to normalize the length of each section between origin and destination pairs with the length of the total network, i.e., with the total length of all the sections included in the trip during all timeframes.

Next, in step 7, a partial urban mobility index is calculated. The parameter $pIM_{i,j}$, i.e., the partial urban mobility index for the corresponding base station pair in the appropriate timeframe, is calculated using the model defined in the following sections. The partial urban mobility index is defined as a function of the normalized indicator value, namely the total number of the trips recorded between the individual base station pairs, the median value of the trip duration between the individual base station pairs in a specific timeframe, and the Euclidean distance between the base station pairs, as shown in Expression (5). The result is a parameter within the [0,1] interval, indicating the value of the mobility estimate, with "0" indicating the lowest and "1" indicating the highest mobility level.

$$pIM_{i,j} = f((normA(i,j), \overline{normB_{i,j}}, normC(i,j))) \quad (5)$$

- $normA_{(i,j)}$ is the normalized value of the number of trips for the corresponding base station pair for all timeframes within one day;
- $\overline{normB_{i,j}}$ is the normalized value of the median trip duration value for the corresponding base station pair for all timeframes within one day;
- $normC_{(i,j)}$ is the normalized value of the distance indicator for the corresponding base station pair.

In step 8, the urban mobility index (UMI) is calculated for each timeframe, as a median of the partial urban mobility indices between each pair of base stations (urban areas) between which a trip was made during a specific timeframe (O). The urban mobility index can have a value from zero to one, i.e., it can be presented as a percentage with a value of zero to 100%. It is calculated by using Formula (6). UMI is calculated for every specific timeframe:

$$IM = \sum_{i=1}^n \frac{pIM_{(i,j)}}{n} [\%] \quad (6)$$

where:

- IM is the urban mobility index for the timeframe;
- $pIM_{i,j}$ is the partial urban mobility index for the corresponding base station pair within the appropriate timeframe;
- n is the total number of origin and destination pairs for which the mobility assessment was performed.

In step 9, after calculating the partial urban mobility index, an additional urban mobility index can be calculated, in order to provide a broader picture of the urban mobility assessment and to complement the partial mobility index. An additional indicator is called the coefficient of the total mobility share ($\alpha_{i,j}$). For each base station pair, the following indicators are defined: The number of trips and the Euclidean distance between them. Total mobility can be defined as the product of the number of trips and the distances between the individual base station pairs. The unit for this parameter is a passenger/transport kilometer (pkm). This unit is common in traffic engineering and is based on the number of actually transported passengers, i.e., actually traveled distance [11–13]. The coefficient is calculated using Expression (7):

$$\alpha_{(i,j)} = \frac{A_{(i,j)} * C_{(i,j)}}{\sum_{i=0}^n (A_{(i,j)} * C_{(i,j)})} \quad (7)$$

where:

- $\alpha_{i,j}$ is the total mobility share coefficient for each origin and destination pair (base station pair);
- $A_{(i,j)}$ is the value of the trip number indicator in a specific timeframe;
- $C_{(i,j)}$ is the value of the distance indicator for a corresponding base station pair.

Therefore, by multiplying the values of the trip number indicators with the distance indicators, it is possible to calculate the total mobility, which we define as the sum of all products of the trip number and distance indicator values, for each base station pair within a specific timeframe. The sum of all total mobility values, for all timeframes represents a total daily mobility. The total mobility at the level of an urban agglomeration within a specific timeframe is calculated by applying Formula (8), as the sum of all mobility quantities:

$$Uk_mob = \sum_{i=0}^n (A_{(i,j)} * C_{(i,j)}) \quad (8)$$

In the final step of the process, step 10, the verification and validation are performed. The purpose of the verification in this part of the procedure is to confirm that the defined algorithms and software correctly calculate the segment for which each algorithm is intended. An illustration of the relationship between the indicators and the parameters in the urban mobility index calculation is shown in Figure 2.

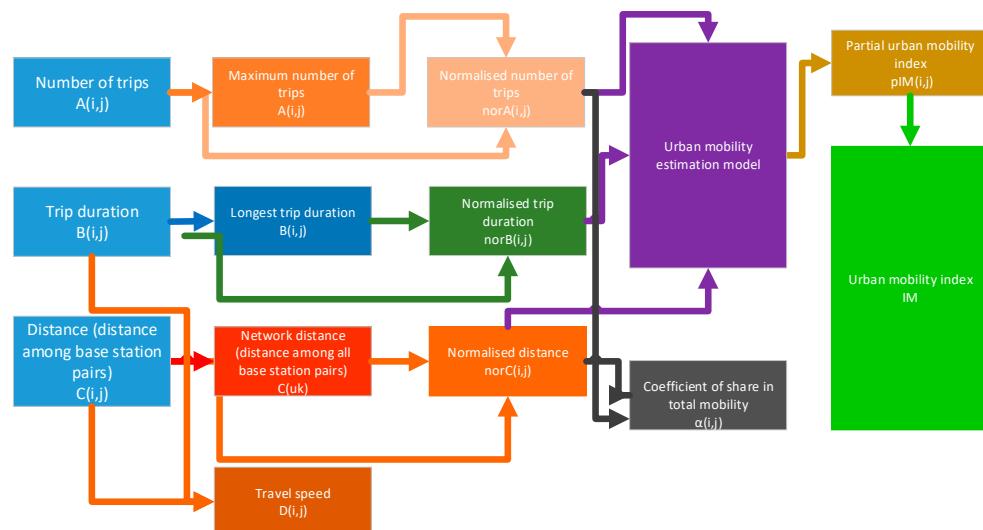


Figure 2. The relationship between the indicators and parameters in the process of the urban mobility index calculation.

3.2. The Relationship between the Mobility Indicators and the Urban Mobility Index Assessment

The determination of the relationship between the mobility indicators and the urban mobility index assessment is based on the input data collected through the survey method. The questionnaire was designed to collect the experts' opinions through a set of questions about how and to what extent the combination of the values of the individual urban mobility indicators affect mobility. The questionnaire contains questions in the form of scenarios that include the urban mobility indicator values, defined in a way that they capture the characteristic indicator values within the predefined ranges, as well as the proposed mobility estimation procedure. On the basis of their own judgment and by applying the proposed mobility assessment procedure, the experts assign an appropriate value to each question. The survey results establish a link between the baseline values of the indicators from the scenario and the value of the mobility estimate.

In line with the previously defined methodology, urban mobility is estimated by using the three indicators obtained from the analysis of the telecommunication service billing records for the use of the public mobile communications network. The indicators include the number of trips, trip duration, and the distance travelled. The number of trips is an indicator that relates to the whole system, i.e., the respondent estimates mobility in the context and based on the total number of the mobility participants in a specific time period. Other indicators, i.e., trip duration and distance travelled refer to an individual trip, and the respondent assesses mobility by looking at these values from an end-user perspective. In the survey questionnaire, the number of trips is defined as follows: The indicator of the

number of trips refers to the number of movements between the individual urban areas in a defined timeframe. It represents the number of all recorded movements, of all users, between the individual urban areas in an appropriate timeframe [3,36]. Values are categorized as follows [3,5,63,77]: A small number of trips are those which are, during a typical urban day, mainly present during periods when there are usually no trips related to work, entertainment, recreation, commerce or social events. A medium number of trips represents the number of trips, in urban conditions, that is characteristic for the period outside of the peak traffic period. For example, it refers to the number of trips achieved in the period when travelling is usually motivated by commerce, recreation, or social activities. A large number of trips represents the number of trips that is, in urban conditions, characteristic for the peak traffic periods, i.e., periods of a typical day, when residents travel to or from work during peak hours.

The trip duration indicator refers to the duration of each individual trip. However, in order to apply this indicator to all urban surroundings, i.e., in cities of all sizes, it must be relativized by calculating the proportion of a single trip in relation to the longest lasting registered trip, thus, normalizing the value of a single trip duration in relation to the longest trip in a characteristic day. In that sense, the longest trip is the one with the longest duration, while the origin and destination remain within the same urban agglomeration, i.e., the coverage area. For this reason, the value of the trip indicators can be defined as in Table 1.

Table 1. The trip duration indicator (relative value).

Trip Duration Indicator	Trip Duration
Short trip duration	up to 33% of the longest lasting trip
Medium trip duration	from 34% to 66% of the longest lasting trip
Long trip duration	from 67% to 100% of the longest lasting trip

The distance indicator is used as an approximation of the distance travelled [1,3,36]. In the questionnaire, the distance travelled indicator was defined as follows: The distance indicator refers to the distance travelled by the user during the trip between individual urban areas, i.e., it represents the distance between the trip origin and destination. The value of the trip distance indicator can be relativized by calculating the proportion of a single trip in relation to the longest registered trip, thus, normalizing the distance travelled within a single trip with respect to the longest journey in a characteristic day. The longest trip is the one during which the distance travelled is the greatest, with the origin and destination remaining within the same urban agglomeration, i.e., the coverage area. For this reason, the value of the distance indicator is defined as in Table 2.

Table 2. The (traveled) distance indicator.

Traveled Distance Indicator	Trip Distance
Short distance	up to 33% of the longest trip
Medium distance	from 34% to 66% of the longest trip
Long distance	from 67% to 100% of the longest trip

If the value of one of the indicators is equal to zero, the trip did not happen and the mobility is equal to zero, i.e., there is no mobility.

A mobility assessment was then defined. Respondents assigned one of the six predefined mobility ratings to the appropriate scenario, as explained in the following paragraphs. The rating categories, as well as the related descriptions further in text, are based on the ratings related to the service level, service quality, and customer satisfaction [13,53,54]. In “extremely high mobility”, the traffic system is characterized by the completely free traffic flow and full freedom of maneuvering, where participants choose the speed, the freedom of movement is not disturbed, and waiting times are minimal. The users are extremely satisfied. With “high mobility” the traffic system is described by the conditions of the free traffic flow, the freedom of maneuvering is not disturbed, the speed of movement is minimally

restricted, the freedom of movement is limited to a small extent and waiting times are rare and short. The users are satisfied. With “higher medium mobility” the traffic system is described by the stable traffic flow, limited maneuvering freedom and limited speed. Freedom of movement is also limited and waiting times are noticeable. Users are mostly satisfied. With “lower medium mobility”, the traffic system is dominated by the unstable traffic conditions, with low maneuverability, substantially limited speeds, clearly limited free movement, and higher average waiting times. Users are mostly dissatisfied. In the case of “low mobility”, the transport system is characterized by the unstable traffic flow with queuing, almost constantly restricted maneuverability, significantly limited movement speed, delays, and long waiting times. Users are dissatisfied. With “extremely low mobility” the conditions within the transport system encompass a forced traffic flow, with the maneuvering freedom and the freedom of movement completely disabled, movement speed lower than critical speeds, and extremely long waiting times. Users are extremely dissatisfied.

Survey questions were then formed. The questions cover all the permutations of the indicator values included in the mobility assessment. The experts answer the questions basing on their own knowledge and experience, by assigning an appropriate value to the mobility estimate in each question. The questions are designed by using the indicator values shown in Table 3.

Table 3. The indicator values.

Number of Trips	Trip Duration Indicator	Travel Distance Indicator
Low number of trips	Short trip duration	Short distance
Medium number of trips	Medium trip duration	Medium distance
High number of trips	Long trip duration	Long distance

The questions are designed to encompass the values of all the indicators involved, for example, If <number of trips>, <trip duration indicator> and <travelled distance indicator>, mobility is <mobility estimate>.

For example, If <number of trips = low number of trips>, <trip duration indicator = medium trip duration> and <travelled distance indicator = long distance>, mobility is <mobility estimate = choose from value 1 to 6 where “6” is extremely high mobility, “5” represents high mobility, “4” stands for “higher medium mobility”, 3 represents “lower medium mobility”, 2 represents “low mobility”, 1 represents “extremely low mobility”>.

Each indicator has three possible values, whereas the mobility estimate has six values. By permuting all indicator values (3^3), a total number of 27 questions is set.

The survey included 36 urban mobility professionals, i.e., academics, experts from the private sector, and public sector experts working at the city level.

Although it is common to include a small number of experts, usually six to 10, when using the knowledge extraction method, a high number of experts from different domains of urban mobility are included in the research. In addition, the survey covered different professions working in the field of urban mobility, i.e., 16 experts from the scientific community, 11 experts from the private sector, and nine experts from the public sector.

Interviews with the experts were conducted at planned workshops and through teleconferences. Each workshop was designed in such a way that in the introductory part, a brief presentation of the research objective was given, and the research background was explained. This was followed by a discussion including the potential ambiguities, after which the questionnaire was answered.

This section further explains the segment of mapping the indicator values, obtained from the CDRs of the values available to experts, i.e., the link between the two sets. By the CDR records analysis, i.e., by applying the activities shown in Steps 1 through 6 in the methodology, normalized (numerical) values of the three target indicators were provided (for example: the trip number indicator has a numeric value of 0.12). By using the fuzzy approach, the “number of trips” indicator value is no longer displayed as a numeric value 0.12, but as a fuzzy affiliation function “a low number of

trips". During the knowledge extraction process, the experts were presented with the fuzzy indicator affiliation function values, thus, ensuring the link between the indicator values calculated by the proposed methodology, using the data from the CDR database together with a set of defined rules for establishing a model, generated as a result of analyzing the experts' survey responses.

A preliminary result analysis is given below and shown in Table 4. For each scenario, it shows a minimum and maximum mobility rating assigned by the experts, an average rating, and the calculated standard deviation. The distribution of ratings (as a percentage) for each mobility assessment scenario is also presented.

Table 4. Analysis of the results of the survey questionnaire.

Scenario	Mobility Rating										
	No.	MIN	MAX	AVG	STDEV	1	2	3	4	5	6
	1	3	6	4.44	1.44	0%	0%	44%	11%	0%	44%
	2	3	6	4.22	1.24	0%	0%	44%	11%	22%	22%
	3	2	6	4.44	1.18	0%	11%	0%	44%	22%	22%
	4	2	5	3.22	1.15	0%	44%	0%	44%	11%	0%
	5	2	5	3.11	1.01	0%	33%	33%	22%	11%	0%
	6	2	6	3.44	1.18	0%	11%	67%	0%	11%	11%
	7	1	3	1.33	0.68	78%	11%	11%	0%	0%	0%
	8	1	4	2.00	0.96	33%	44%	11%	11%	0%	0%
	9	1	4	2.33	0.96	22%	33%	33%	11%	0%	0%
	10	3	6	4.33	0.83	0%	0%	11%	56%	22%	11%
	11	2	6	4.33	1.07	0%	11%	0%	44%	33%	11%
	12	1	6	4.67	1.43	11%	0%	0%	11%	56%	22%
	13	2	4	3.00	0.68	0%	22%	56%	22%	0%	0%
	14	3	4	3.11	0.32	0%	0%	89%	11%	0%	0%
	15	2	4	3.33	0.83	0%	22%	22%	56%	0%	0%
	16	1	6	2.78	1.57	11%	56%	11%	0%	11%	11%
	17	1	5	3.00	1.35	11%	33%	22%	11%	22%	0%
	18	1	5	3.00	1.07	11%	11%	56%	11%	11%	0%
	19	4	6	4.78	0.80	0%	0%	0%	44%	33%	22%
	20	4	6	4.89	0.75	0%	0%	0%	33%	44%	22%
	21	5	6	5.78	0.42	0%	0%	0%	0%	22%	78%
	22	2	4	3.33	0.68	0%	11%	44%	44%	0%	0%
	23	2	5	3.67	0.96	0%	11%	33%	33%	22%	0%
	24	2	6	4.22	1.15	0%	11%	11%	33%	33%	11%
	25	1	5	2.56	1.36	33%	11%	33%	11%	11%	0%
	26	1	4	2.67	1.07	11%	44%	11%	33%	0%	0%
	27	1	5	3.22	1.24	11%	11%	44%	11%	22%	0%

The preliminary data analysis was made in order to check the consistency of the answers and to identify questions where a pronounced assent or disagreement in the opinion of the experts exists. Certain differences in the opinion of the experts are expected, as they consider mobility from different perspectives, depending on their own expertise. The highest consensus of the experts was recorded for questions 14 (medium number of trips, medium trip duration, and medium distance) and 21 (high number of trips, short trip duration, and long distance). In Question 14, the assigned mobility ratings range from three to four, with an average score of 3.11, while in the Question 21, the responses range from five to six, with an average value of 5.78. In case of several questions, Question 12 (medium number of trips, short trip duration, long distance) and Question 16 (medium number of trips, long trip duration, short distance), the experts assigned the highest and the lowest ratings to the same question, with an average mobility rating of 4.67 in Question 12 and 2.78 in Question 16. In Question 7 (low number of trips, long trip duration, short distance), the highest number of experts (78%) rated the described mobility conditions with rank "one" (low number of trips, long trip duration, short distance), while in the Question 16 (medium number of trips, long trip duration, short distance), the

highest number of experts (56%) rated mobility with the rank “two”. The highest number of experts (89%) rated the mobility described in Question 14 (medium number of trips, medium trip duration, medium distance) with a rank “three”. The highest number of experts (56%) rated mobility described in Question 10 (medium number of trips, short trip duration, short distance) and Question 15 (medium number of trips, medium trip duration, long distance) as rank “four”. The highest number of experts (56%) rated the mobility described in Question 12 (medium number of tips, short trip duration, long distance) as rank “five”. The highest number of experts (78%) rated the mobility described by Question 21 (high number of trips, short trip duration, long distances) with a rank “six”.

4. Model of the Urban Mobility Index

In order to define an urban mobility estimation model, a fuzzy logic method, ANFIS (adaptive neuro-fuzzy inference system), is used. It is a machine learning technique that uses fuzzy logic to transform the given input into the corresponding output values [78–82]. Data showing the relationships between the values of the mobility indicators and the mobility ratings were collected through a survey. The respondents were provided with three descriptive values from a set consisting of three indicators and were required to associate a mobility ranking from one of the six descriptive categories. The descriptive values of the indicators had surrogate numerical values in the interval of [1,3], while the mobility estimate had surrogate numerical values in the interval of [1,6]. Considering that the data are in different ranges (the values of the indicators are represented by the surrogate variables ranging from one to three, and the mobility assessment values in the range of one to six), their normalization was performed. It is common for a set of numeric values to be normalized by subtracting the minimum value of that attribute from all values, in all examples of the dataset, and by dividing the resulting value by a difference between the maximum and minimum attribute values. The values obtained in this manner have the same form of distribution, and all values range from zero to one. It is a desirable and sometimes necessary operation, when modeling with the neural networks or creating regression models [83,84]. Upon the normalization and division of the data into sets, they can be used to model an adaptive neuro-fuzzy inference system as a fuzzy system in which the membership functions are generated by using the learning methods characteristic of the neural networks. Out of the total number of data pairs collected by the survey (36 questionnaires, 27 questions each), a total of 972 data pairs were generated. Out of this number, 654 data pairs are used for learning and 160 for the validation and control (verification). ANFIS modeling is performed as part of the MATLAB software tool [85]. After the completion of the preparatory action, the learning data loading procedure and the validation and verification of the model, the generation of an initial fuzzy inference system takes place using a fixed grid partitioning. A fixed grid partitioning refers to the division of data space into a rectangular grid, which depends on the number of the membership functions and their type. The number of rules increases exponentially by increasing the input variables, so if there is an m of membership functions and n input variables, then the number of fuzzy rules is m^n . Therefore, the use of a fixed grid is limited only to cases where there is a low number of the input variables (less than six), resulting in lower requirements for the computational resources in simulation [86]. The use of the fixed grid partitioning for the systems with a low number of membership functions ensures a small forecasting error and was chosen to be applied in this model.

The initial system defines the initial settings of the membership functions, which may change after the learning has been performed, as a result of the learning itself. This step defines the input parameters as follows: (a) The number of the membership functions, (b) the type of the input membership functions, and (c) the type of output membership functions. The number of the input membership functions is defined by the number of input variables. Three input variables are included in this model: (a) the number of trips in the interval [0,1], (b) trip duration [0,1], and (c) distance [0,1].

The model learning, then, takes place. Before the learning process begins, the following has to be defined: (a) an optimization method, i.e., the selection of the learning method between backpropagation or hybrid learning; (b) the error tolerance; and (c) the number of the model learning epochs. All fuzzy

inference systems are learned either by the backpropagation method or by the hybrid method [87–91]. For each of the fuzzy inference systems, three scenarios were defined, therefore, in the first scenario the number of learning epochs was three (a default setting), in the second scenario the number of learning epochs was ten, and in the third scenario the number of learning epochs was one hundred. On the basis of the result analysis, it was concluded that in this research a hybrid learning method would be used, in which zero error tolerance is defined and the learning takes place within three epochs. Models are evaluated through the learning error (RMSE). During the network learning, the graph shows the value of the error depending on each learning epoch. The learning result is a 78-node network with 108 linear parameters and 36 nonlinear parameters, i.e., a total of 144 parameters. A total of 654 data pairs were used for the model learning, and 160 data pairs were used for validation. The system resulted in a base of rules containing 27 (3^3) rules. The model has been described with 27 rules, each having its own interval of the mobility estimate output values, depending on the parameter values of the input variables. Within the initial rule base (before the model learning), the values of the target variable are unknown. After the model learning process, ANFIS enables the creation and fine-tuning of rules which describe the behavior of a complex system. A graphical representation of the rules before the model learning (for one candidate system) is shown in Figure 3, and after completing the model learning is shown in Figure 4. An example of the graphical representation of an error reduction, while performing the network learning through the 100 epochs for the fuzzy inference system under the ordinal number 27 (settings: the membership functions of the input variables are of a trapezoid type, backpropagation learning, the output function is linear, the epoch number is 100), is shown in Figure 5.

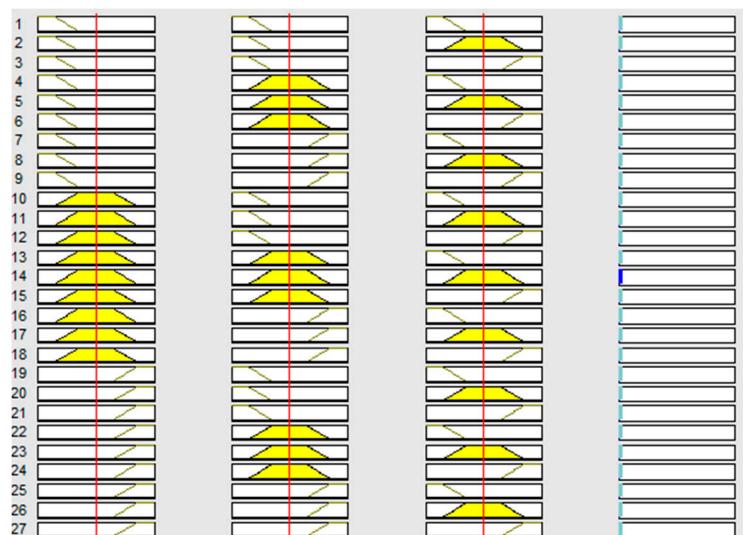


Figure 3. The graphical representation of the rules in the initial fuzzy logic model (1st column = number of trips, 2nd column = trip duration, 3rd column = distance, and 4th column urban mobility estimation).

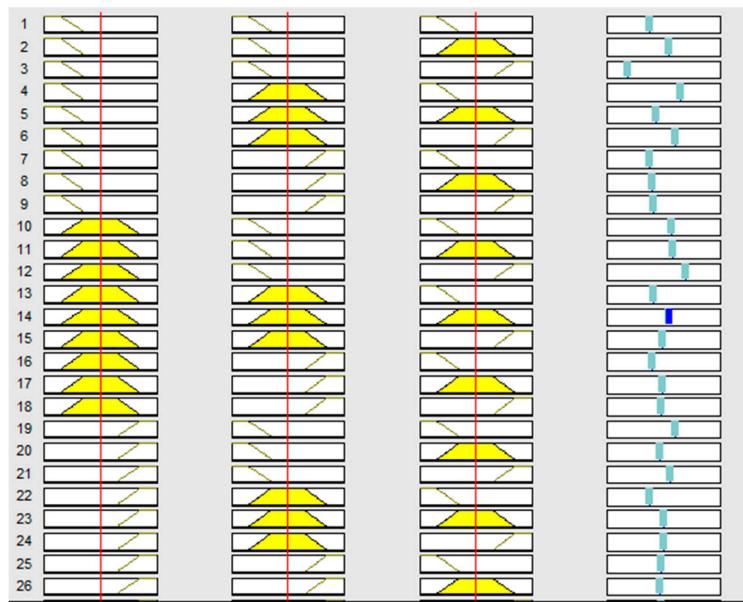


Figure 4. The graphical representation of the model after the completion of the model learning process (1st column = number of trips, 2nd column = trip duration, 3rd column = distance, and 4th column urban mobility estimation).

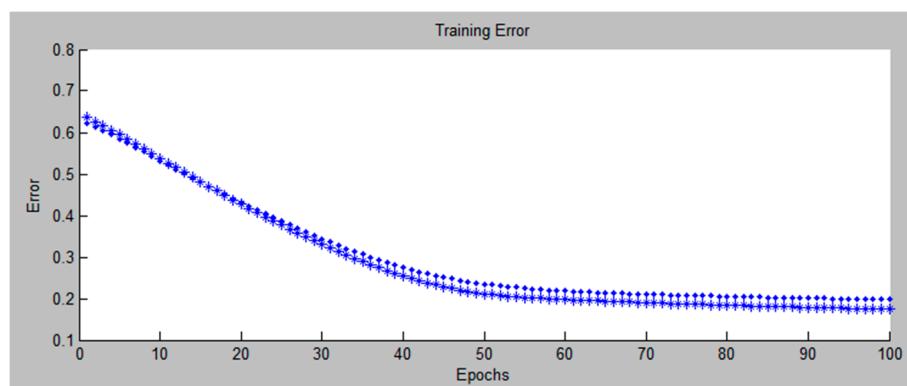


Figure 5. A learning error dependent on the number of epochs for the fuzzy inference system number 27.

The ANFIS model validation and verification process is then performed, i.e., the model forecasting results are presented and analyzed by comparing the model results obtained from the learning dataset, the validation dataset, and the model verification dataset. The validation and verification of the ANFIS model includes two processes: (a) the model validation process and (b) the model verification process. Model validation is a process used to evaluate the reliability of the model inference. The difference between the learning phase and the validation phase, at this stage, is that the model no longer learns, i.e., the weight coefficients over the values, obtained as a result of the learning are fixed and no longer change. Validation error is defined as the difference between the value of the output variable of the original and unmodified dataset used for validation, and the difference between the value of the output variable from the fuzzy inference system using the same input dataset (validation set). The learning verification error is calculated as the root mean square error (RMSE).

The graphical representation of the learning dataset is shown in Figure 6, the validation dataset is shown in Figure 7, and the verification dataset is in Figure 8.



Figure 6. Learning dataset (training data).

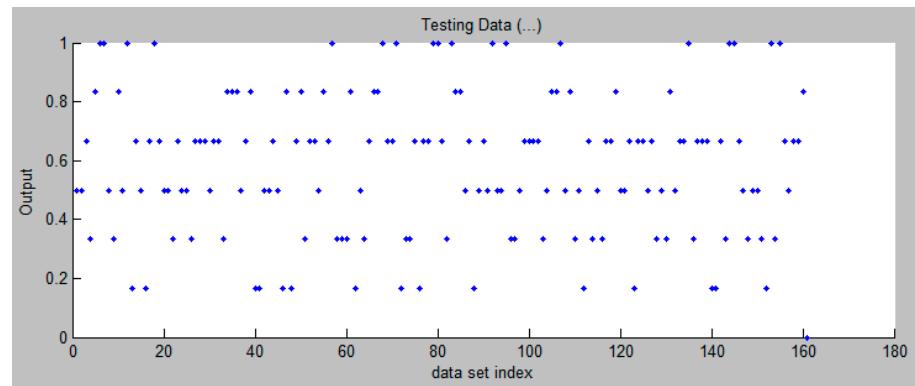


Figure 7. Validation dataset (testing data).

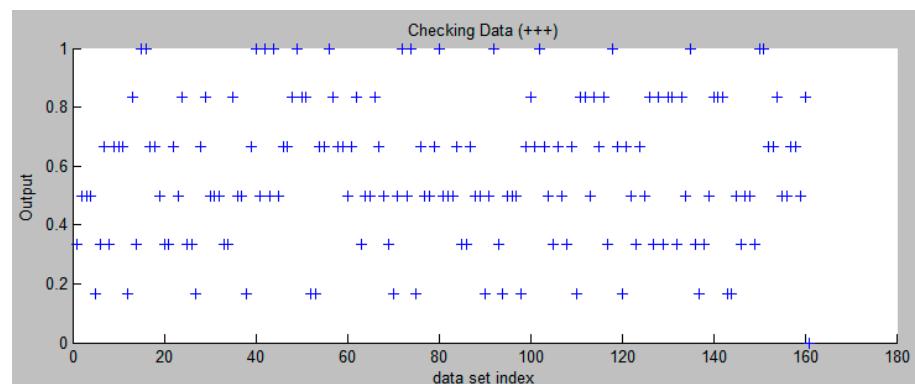


Figure 8. Verification dataset (checking data).

Figure 9 shows the interrelation between the training dataset and the testing dataset. Figure 10 shows the relationship between the model learning data and the data used to verify the model (checking data). This is followed by the model validation and verification processes [92], performed for each fuzzy inference system. Model validation confirms that the outputs of the statistical model outputs are acceptable with respect to the real data-generating process and model verification. A checking dataset for the model verification is needed, since after a certain point during the training, the model begins overfitting the training dataset. In principle, the model error for the checking dataset tends to decrease as the training takes place, up to the point when the overfitting begins and, then, the model error for the checking data suddenly increases. The overfitting is accounted for by testing the FIS trained on the training data against the checking data and choosing the membership function parameters to be those associated with the minimum checking error, if these errors indicate model overfitting. After completing the model learning process, it appears that the fuzzy inference system has been modified,

primarily in terms of changes in the values describing the membership functions of the input and the output variables, as shown in Table 5. The interval values of the functions obtained in this step, represent the final version of each membership function interval.

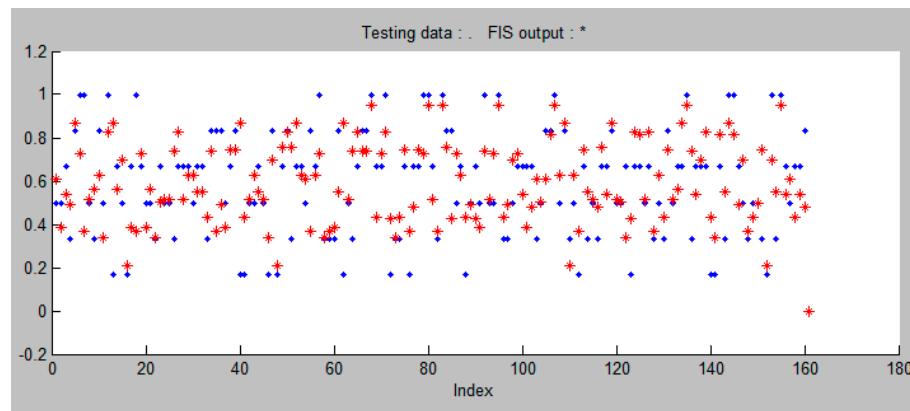


Figure 9. The interrelation between the training dataset and the testing dataset.

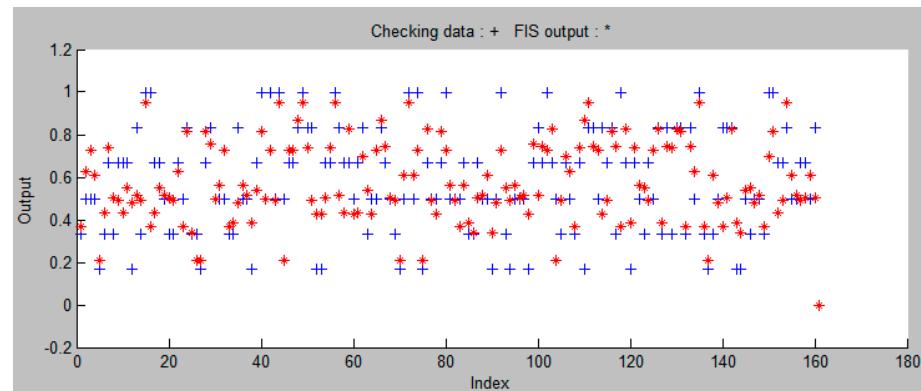


Figure 10. The interrelation between the training dataset and the checking dataset.

Table 5. The comparison of the intervals of the input membership functions, before and after the model learning.

Input Variable	Fuzzy Set	Function Interval in the Initial Fuzzy Logic System	Membership Function Interval after the Learning Process in the Chosen Fuzzy Logic System
Number of trips	Small number of trips	[−0.35 − 0.15 0.15 0.35]	[−0.35 − 0.15 0.15 0.361]
	Medium number of trips	[0.15 0.35 0.65 0.85]	[0.1501 0.3512 0.6497 0.85]
	Large amount of trips	[0.65 0.85 1.15 1.35]	[0.6473 0.8497 1.15 1.35]
Trip duration	Short trip duration	[−0.35 − 0.15 0.15 0.35]	[−0.35 − 0.15 0.15 0.3628]
	Medium trip duration	[0.15 0.35 0.65 0.85]	[0.1501 0.3514 0.6504 0.85]
	Long trip duration	[0.65 0.85 1.15 1.35]	[0.6544 0.8504 1.15 1.35]
Distance	Short distance	[−0.35 − 0.15 0.15 0.35]	[−0.35 − 0.15 0.15 0.3568]
	Medium distance	[0.15 0.35 0.65 0.85]	[0.1501 0.3507 0.6505 0.85]
	Long distance	[0.65 0.85 1.15 1.35]	[0.6558 0.8505 1.15 1.35]

The output parameter, an urban mobility estimation function, is based on 27 rules of the fuzzy inference model and is accordingly defined with the 27 membership functions. For illustration, the first value of the membership function is in the interval $[-0.1402 - 0.1402 - 0.1402 - 4.205e - 07]$, while the last one, the 27th, is in the interval $[0.1106 0.16 0.1144 0.1149]$.

A total 32 candidates for the fuzzy inference system have been defined. All fuzzy inference systems have been assigned with the appropriate membership functions, the form of the output variable and the learning (optimization) method have been defined, together with the number of epochs. The systems varied depending on the form of the input membership function (membership function of triangle type, trapezoid type, bell type, Gaussian type, two-parameter Gaussian type, the membership function π , the membership function as a difference of the sigmoidal functions, and the membership function as a product of the sigmoidal functions), then, depending on the type of output membership function (linear or constant), depending on the learning method type (hybrid learning or backpropagation learning), and the number of learning steps (three, 10, or 100). The reliability of each fuzzy system established was tested using the RMSE method, with a learning error included into the analysis. In general, the smallest learning error was recorded in all scenarios where the hybrid learning method was applied and in which the output function was linear. There was also no significant deviation in the number of epochs (learning steps) noticed. This is supported by the previous studies which have shown that in solving the real-world problems, the form of the membership function does not affect the accuracy of the process, i.e., the systems of the external world are relatively insensitive to the membership function form [9.65 – 68]. Therefore, for the development of an expert system of this type, the most commonly used membership functions in the literature are the triangular or trapezoidal types [9.65 – 68]. Finally, a fuzzy inference system with the smallest model learning error was selected to establish the urban mobility index. The average model validation error (for the model chosen as optimal) is 0.2103, and the average model validation error is 0.1907. The first-class fuzzy logic Sugeno model was chosen as the most optimal model, which was one where all three entry values are determined with a trapezoid type membership function, and each consists of three overlapping linear type functions. The result is an urban mobility index model. The model defined the individual relationship of each input indicator with respect to the mobility estimates and enables the expression of the urban mobility estimate as a single number, considering the values of all indicators. Figures 11–13 graphically show the combined effects of the values of the individual indicators on the overall mobility.

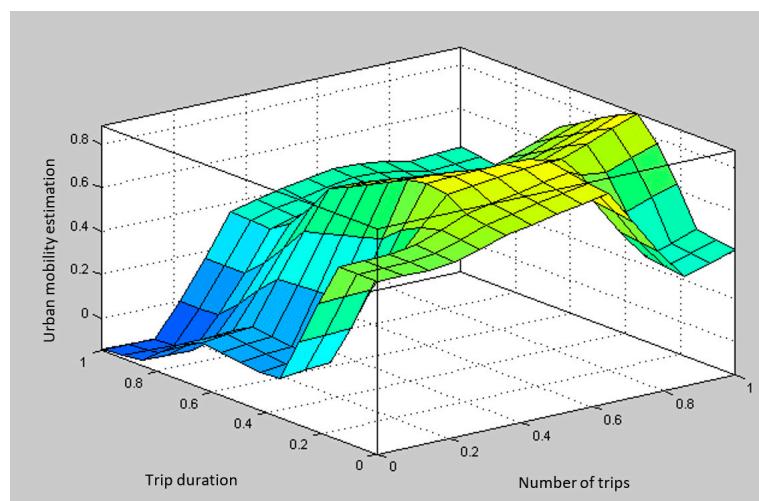


Figure 11. The model result: The relationship between the trip duration indicators and the number of trips with respect to the mobility estimate (taken from the MATLAB v 9.0, 2016 programming tool).

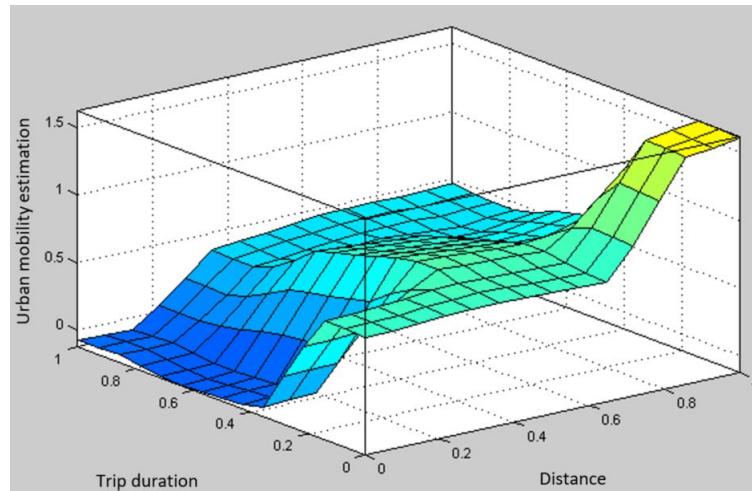


Figure 12. The model result: The relationship between the trip duration indicators and the distance indicators with respect to the mobility estimate (taken from the MATLAB v 9.0, 2016 programming tool).

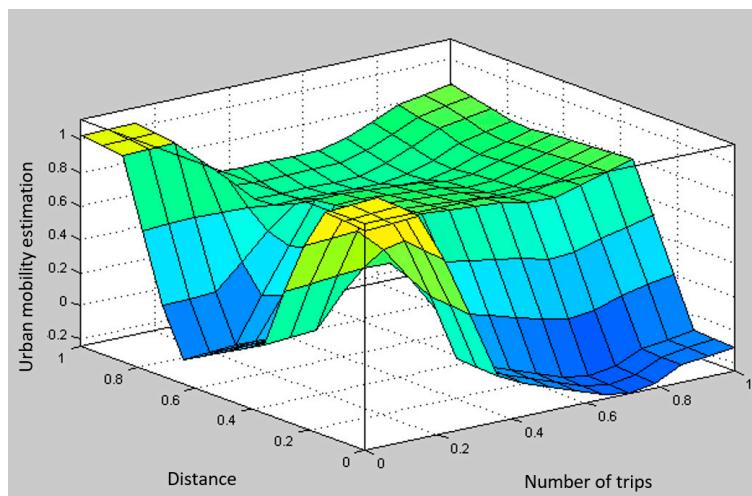


Figure 13. The model result: The relationship between the trip distance indicators and the number of trips with respect to the mobility estimate (taken from the MATLAB v 9.0, 2016 programming tool).

The urban mobility indicators, which are used as input for the urban mobility estimation model, have been identified in the previous steps of the procedure. These indicators are then included in the model, as the actual values of the indicators obtained from the previous steps. The modeling is then carried out for each base station pair, resulting in the calculation of so-called partial urban mobility indices that represent the mobility estimate for that specific pair of urban areas. The partial urban mobility indices, as the output of the model, are then re-entering the urban mobility assessment process. The extent to which each of them participates in the total mobility is then determined and the final urban mobility index calculated (for the appropriate time period), for the respective urban agglomeration. The proposed model of the mobility estimation, i.e., the urban mobility index is developed for generic cases and the obtained results can be compared between the cities. The application of the proposed urban mobility index enables an objective and comparable passenger mobility assessment within urban agglomerations. The existing models are hardly enable a comparison of different cities and agglomerations.

5. The Application of the Urban Mobility Index Estimation Process

In order to make the proposed methodology applicable for real-world urban mobility estimation, an urban mobility model defined within the previous section needs to be implemented in the programming environment. Following the implementation, the application will be performed using the real-world dataset.

5.1. Process Implementation in a Programming Environment

The urban mobility assessment process implementation was carried out using several programming environments. Dataset loading, adjusting the input data, and decomposing the space were performed using the programming code developed in the R software environment. For visualizing the decomposition of the space, an open source software tool for the geospatial analysis (QGis) was used. Modules for calculating the Euclidean distance between the base stations, representing the values of the “distance” indicators, were then written in the R programming language. A module was also created to divide the dataset into subsets, according to the defined timeframes. The key part of the programming code is the application of the algorithm to identify migrations, i.e., to identify travelling. The data is then entered into a data consolidation and filtering module, and finally into a module in which the data is divided into the validation and calibration set. The data is then converted from a table into a matrix form and this is followed by the data normalization module. The normalized data is entered into the MATLAB programming language, where the fuzzy inference system is invoked, and the data is entered. The result is a partial urban mobility index stored in a table. The data are then returned to R, where the output vector of the partial urban mobility indices is reduced to an interval of [0,1] and the urban mobility index is calculated. The same procedure is performed using the validation dataset. This is followed by the result visualization and analysis.

5.2. The Application of the Procedure over a Real-World Dataset

In the research, the mobility estimation process was tested using a real publicly available subset of data derived from a telecommunication services billing record of a public mobile telecommunication network, which is available at [73]. The data relate to a part of the urban agglomeration of the city of Shenzhen in China, covering 480 base stations and an area of approximately 900 km² (30 × 30 km). It is estimated that approximately 1.2 million people live in the area, while approximately 25 million live in the urban agglomeration together with the suburban settlements [73]. The entire urban agglomeration (together with suburban settlements) covers approximately 20,000 m². The dataset used for the analysis consists of 38,218,717 telecommunication activity records for one characteristic day, for 24 h. It contains data on the telecommunications activities of approximately 450,000 users, representing 37.5% of the population in the area. On average, the record contains 85 telecommunication activities per user on a typical day. According to the number of the base stations, the coverage area is divided into 480 urban areas. The average surface area covered by one base station is 2.69 km². The shortest distance between the base station pairs is 0.007 km (the distance between the base station 475 and 476), with a total of 12 base station pairs less than 100 m apart, and 1345 base station pairs within a kilometer distance. The longest distance is 42.9 km (the distance between base stations 36 and 476). The average distance between all base station pairs in the network is 1.81 km. Next, is the migration identification process. A total of 2,133,369 migrations were identified in the analyzed dataset. By applying the filter, all ineligible migrations were eliminated, leaving those longer than 1 km, lasting between 10 and 60 min, and with an average speed less than 100 km/h. As a result of the filtering, almost 95% of the identified migrations were removed, i.e., from 91% to 97% per timeframe, leaving 110,902 migrations. The following is the calculation of the only remaining indicator (medium travel time), normalization of the values of all indicators and the calculation of the partial urban mobility indices for each base station pair. The values of the partial urban mobility index for one timeframe is shown in Figure 14. Using the calculated values of the partial urban mobility indices, the calculation of the urban mobility

index for each timeframe is carried out. The values of the urban mobility index by timeframes for the basic dataset and for the validation dataset are shown in Table 6. The validation of the model uses a dataset consisting of 20% of the total dataset and shows that deviations are possible in the range from 3% to 11%, relative to the value obtained through using the basic dataset. The deviation within seven timeframes does not exceed the value of 7%, and in only one frame has a value greater than 11%. The average deviation for all timeframes is 6%, showing the model reliability of 94%.

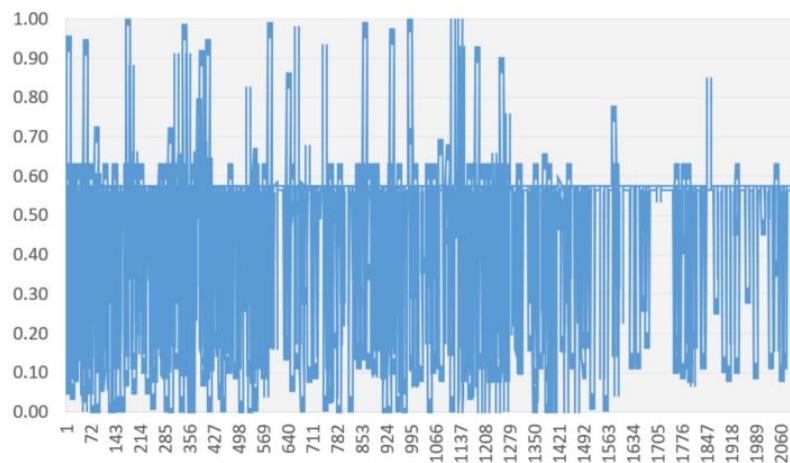


Figure 14. PUMI values for the frame O1.

Table 6. The values of the urban mobility index by timeframes for the basic dataset and for the dataset used for validation.

	O1	O2	O3	O4	O5	O6	O7	O8
IM	51.14%	54.50%	69.02%	64.03%	66.65%	7.57%	56.14%	52.65%
IM (validation)	54.57%	57.83%	70.92%	68.20%	70.13%	69.14%	58.71%	56.59%
Difference	6%	6%	3%	6%	5%	-11%	4%	7%

The deviation of the urban mobility index value from the basic set with the validation data is given in Figure 15.

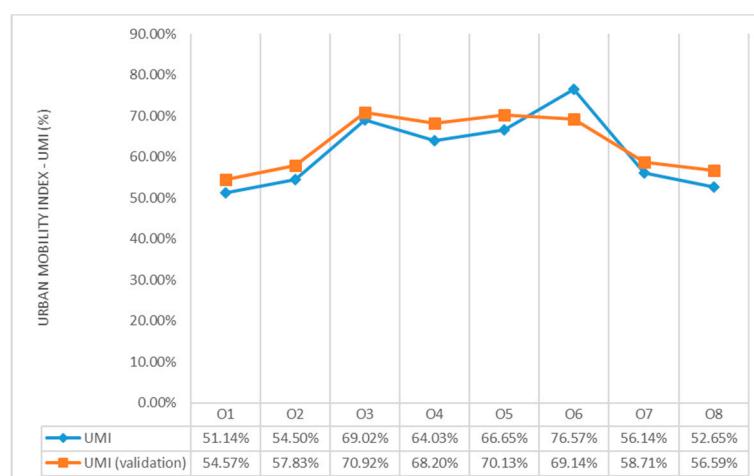


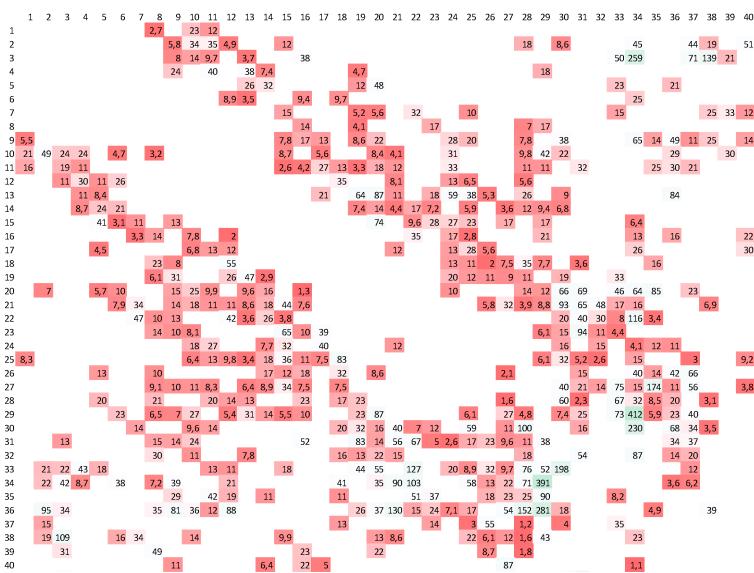
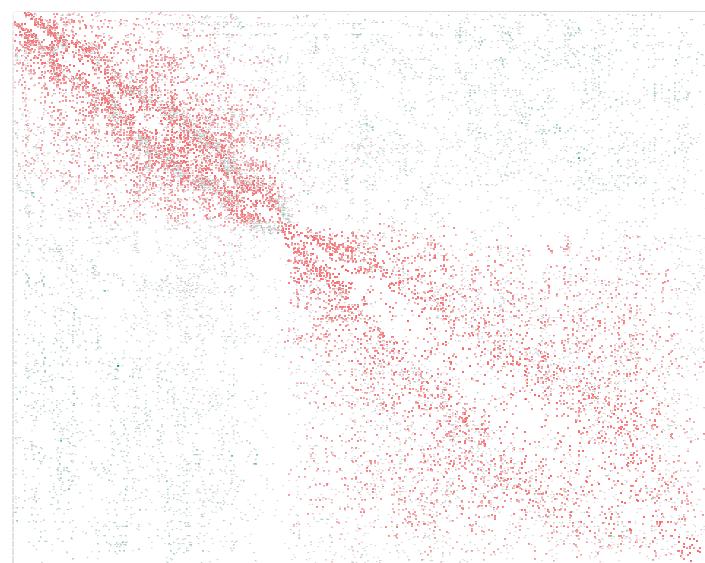
Figure 15. The comparison of the urban mobility index from the basic dataset with the validation data.

Next is the calculation of an additional indicator, the coefficient of the total mobility share ($\alpha_{i,j}$), shown in Table 7. The result is the proportion of each pair of base stations in the total mobility within one timeframe.

Table 7. The total amount of mobility for each timeframe.

Timeframe	O1	O2	O3	O4	O5	O6	O7	O8
Amount of mobility	76.365	114.043	223.260	152.107	192.886	151.974	347.662	301.795

An analysis was then performed for each frame, showing the minimum and maximum coefficients ($\alpha_{i,j}$), for each timeframe. In order to identify the areas where the greatest mobility takes place, a summary table was created containing the sum of all total mobility, for all base station pairs and for all timeframes. This gives an overview of the total urban agglomeration mobility within a 24-h period. The table is visualized in such a way that the range of colors from red to green is defined, with the associated hues, with the most intense shade of red being a pair of base stations (urban areas) with the smallest product of the number of trips and distances, and green being the pair where the aforementioned value is the highest. The table segment for the first 40 rows and columns of the matrix is shown in Figure 16. The entire table is visualized in Figure 17.

**Figure 16.** The segment of the total mobility for a sum of values within the frame O1-O8.**Figure 17.** The visualisation of an overall urban mobility table (480 × 480) for a sum of values within the frame O1-O8.

The procedure was validated using a validation dataset, derived from the basic set. This proves the hypothesis set in this research paper, that the mobility of the urban residents over a period of time can be described by an urban mobility index based on the information from the recorded telecommunication activities of the users of the public mobile communication network.

6. Conclusions

The purpose of this paper was to prove the potential of telecommunication users' activities data (call data records) to support the provision of an urban mobility estimate, then, to identify urban mobility indicators, to define the process of their calculation, and a model which can estimate urban mobility based on the parameters of the indicators. The defined goal of research was to establish a process of the passenger urban mobility estimate as a quantitative measure of the process of urban migrations caused by socioeconomic activities, one which is in the function of determining a new urban mobility estimation index. A structured record of users' activities within a public mobile communication network used for the purpose of charging telecommunication services was used as a primary research material. An algorithm that uses this information as a basis to identify the movement of users within an urban agglomeration covered by this research was developed. On the basis of the information about the identified migrations, the indicators of urban movement were defined, and the process of their calculation was determined and validated. This was followed by a procedure in which the relationships between the values of urban mobility indicators and values of mobility estimation were determined. This relation was defined through the method of surveys. The survey was defined in such a way so that through several questions expert opinions could be gathered on how and to what degree the combination of values of certain urban mobility indicators affects mobility. According to their own judgement, experts allocated an appropriate value to each question using the suggested procedure for mobility estimation. The method of fuzzy logic, that is, the ANFIS (adaptive neuro-fuzzy inference system) method which functions on the principle of applying conclusion methods which characterize neural networks with the goal of determining parameters of an indirect conclusion system (fuzzy inference system, FIS) was used. The result is a system that permits the usage of well-known neural network learning algorithms, ones that cannot be used in fuzzy logic systems, while at the same preserving the possibility of using fuzzy logic. The answers of experts were used as an initial collection of data to establish the model. The model was described with 27 rules, each of which has its own interval of exit mobility estimate values dependent on entry parameter values of entry indicators, defined with a total of 32 fuzzy logic systems which were assigned their appropriate affiliation functions. A form of the exit variable and a learning (optimization) method, as well as the number of epochs, were defined. Each of the established fuzzy logic systems were tested for reliability using the RMSE method. The smallest learning mistake was generally observed in all scenarios which used the hybrid learning method and in which the exit function was in linear form. The average model validation error (for a model chosen as optimal) was 0.2103, and the average model validation error was 0.1907. As the most optimal model, the first-class fuzzy logic Sugeno model was chosen, one in which all three entry values were determined with a trapezoid type affiliation function, and each consisted of three overlapping linear type functions. A hybrid learning model with three learning epochs was used. The result of the model is an estimate of urban mobility for the matching pair of urban areas (partial mobility index), which is then used to calculate the mobility index of the entire urban agglomeration in the process of urban mobility estimation. Implementation of the urban mobility estimation process was executed using several programming environments. Data collection loading, entry data table conversion, and creating the base cell table, as well as the decomposition of space was handled with a programming code developed in the R programming environment. The open source geographic analysis programming tool QGis was used for space decomposition visualization. Validation was executed though several steps. Validation of the ANFIS model was part of an integral process of forming the fuzzy logic system. This was followed by the validation of algorithms which were defined as part of the mobility validation procedure. The purpose of the validation in this part of

the procedure was to confirm that the defined algorithms and the written programming support had been correctly calculating the segment they were assigned to. The validation of all algorithms was successful, confirming that all written algorithms properly executed their assigned operations. The validation of the model results followed. According to the indicator calculation, a data segment of 20% for each timeframe was extracted, and partial indexes and the urban mobility index were calculated accordingly. The results found in that data were compared with the results from the remaining data in the segment, and the discrepancy between results was determined. The procedure of the urban mobility estimation was applied on a real publicly available subsegment of data taken from a record used to charge telecommunication services within a public mobile communication network. The result of the research have proven the hypothesis of this scientific research paper which states that the mobility of inhabitants within an urban environment can in a specific timeframe be described with an index of urban mobility, based on the recorded telecommunication activities of users in a public mobile communication network. The goal of the paper was fulfilled and the possibility of application of telecommunication users' activities data for the purpose of an urban mobility estimate was proven. Indicators of urban mobility and the process of their calculations were defined, followed by a model which can estimate urban mobility on the basis of the values of these indicators, i.e., a procedure of the population's urban mobility estimates as a quantitative measure of urban migrations caused by socioeconomic activities. Future research should focus on eliminating some of the perceived limitations of the existing research. The input dataset could be upgraded to include the telecommunications network records from the signaling area (e.g., GPEH, CTR, and CTUM), which, in line with the nature of the telecommunications network operation, provide a significant amount of user activity records similar to the CDR database. Further development and improvement of the methodology should also focus on the substitution of the linear distance between the base stations, which now represents the distance, with the movement along the infrastructure (e.g., the road distance between the base stations). In addition, research should focus on segregation of movements by transport modes in order to include the impact of multimodality in the urban mobility index. [93,94].

Author Contributions: Conceptualization, K.V. and S.M.; development and design of the methodology, K.V., M.Š., and S.M.; creation of models, K.V.; writing of original draft, K.V. and M.Š.; investigation, M.Š.; creation of models, M.Š.; validation S.M. and G.K.; supervision, S.M.; data curation, G.K.; formal analysis, G.K.; visualization, G.K.; writing—review and editing, G.K. All authors have read and approved the final manuscript.

Funding: This research received no external funding.

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

References

- Brčić, D.; Ševrović, M. *Logistika Prijaveza Putnika*; Fakultet prometnih znanosti Sveučilišta u Zagrebu: Zagreb, Croatia, 2012.
- Kuhle, O. *Mobilities, Networks, Geographies*; Routledge: London, UK, 2008; pp. 429–430.
- Hanson, S.; Giuliano, G. *The Geography of Urban Transportation*; The Guliford Press: London, UK, 2004.
- Gillis, D.; Semanjski, I.; Lauwers, D. How to Monitor Sustainable Mobility in Cities? Literature Review in the frame of Creating a set of Sustainable Mobility Indicators. *Sustainability* **2016**, *8*, 29. [[CrossRef](#)]
- Casey, J.; Norwood, J.L. National Research Council. In *Key Transportation Indicators*; National Academy Press: Washington, DC, USA, 2002.
- Brčić, D.; Šoštarić, M. *Parkiranje i garaže*; Fakultet prometnih znanosti Sveučilišta u Zagrebu: Zagreb, Croatia, 2012.
- Amaral, R.R.; Šemanjski, I.; Gautama, S.; Aghezzaf, E.H. Urban Mobility and City Logistics—Trends and Case Study. *Promet Traffic Traffico* **2018**, *30*, 613–622. [[CrossRef](#)]
- Novačko, L.; Šimunović, L.; Krasić, D. Estimation of Origin.Destination Trip Matrices for Small Cities. *Promet Traffic Transp.* **2014**, *26*, 419–428.

9. EL Rashidy, R.A.; Grant Muller, S. A Network Mobility Indicator using a Fuzzy Logic Approach. In Proceedings of the TRB 93rd Annual Meeting Compendium of Papers, Transportation Research Board, Washington, DC, USA, 12–16 January 2014; pp. 1–14.
10. Gudmundsson, H. Making Concepts Matter: Sustainable Mobility and Indicator Systems in Transport Policy. *Int. Soc. Sci. J.* **2003**, *55*, 199–217. [\[CrossRef\]](#)
11. Litman, T. Developing Indicators for Comprehensive and Sustainable Transport Planning. *Transp. Res. Rec.* **2008**, *2017*, 10–15. [\[CrossRef\]](#)
12. Litman, T. *Well Measured—Developing Indicators for Sustainable and Livable Transport Planning*; Victoria Transport Policy Institute: Melbourne, Australia, 2015.
13. Litman, T. Measuring Transportation: Traffic, Mobility and Accessibility. *Inst. Transp. Eng.* **2003**, *73*, 28–32.
14. Kapriás, I.; Eden, N.; Tsakerestos, A.; Gal-Tzur, A.; Gerstenberger, M.; Hoadley, S.; Lefebvre, P.; Ledoux, J.; Bell, M. Development and Application of an Evolution Framework for Urban Traffic Management and Intelligent Transport Systems. *Procedia Soc. Behav. Sci.* **2012**, *48*, 3102–3112. [\[CrossRef\]](#)
15. Vidović, K.; Šoštarić, M.; Budimir, D. An Overview of Indicators and Indices Used for Urban Mobility Assessment. *Promet Traffic Transp.* **2019**, *31*–*36*, 703–714. [\[CrossRef\]](#)
16. Vidović, K.; Mandžuka, S.; Brčić, D. Estimation of Urban Mobility Using Public Mobile Network. In Proceedings of the 2017 International Symposium ELMAR, Zadar, Croatia, 18–20 September 2017; pp. 21–24.
17. Zelenika, R. *Metodologija i tehnologija izrade znanstvenog i stručnog djela*; Ekonomski fakultet Sveučilišta u Rijeci: Rijeka, Croatia, 2000.
18. Zhang, D.; Huang, J.; Li, Y.; Zhang, F.; Xu, C.; He, T. Exploring Human Mobility with Multi-source Data at Extremely Large Metropolitan Scales. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, MobiCom’14, Maui, HI, USA, 7–11 September 2014; pp. 201–212.
19. Lopes, J.; Bento, J.; Huang, E.; Antoniou, C.; Ben-Akiva, M. Traffic and Mobility Data Collection for Real-Time Applications. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, Funchal, Portugal, 19–22 September 2010; pp. 216–223.
20. Chen, C.; Ma, J.; Susilo, Y.; Liu, Y.; Wang, M. The Promises of Big Data and Small Data for Travel Behavior (aka Human Mobility) analysis. *Transp. Res. Part C* **2016**, *68*, 285–299. [\[CrossRef\]](#)
21. Prabha, R.; Kabadi, M.G. Overview of Data Collection Methods for Intelligent Transportation Systems. *Int. J. Eng. Sci.* **2016**, *5*, 16–20.
22. González, M.C.; Hidalgo, C.A.; Barabási, A.-L. Understanding Individual Human Mobility Patterns. *Nature* **2008**, *453*, 779–782. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Wang, H.; Calabrese, F.; Di Lorenzo, G.; Ratti, C. Transportation Mode Inference from Anonymized and Aggregated Mobile Phone Call Detail Records. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, Funchal, Portugal, 19–22 September 2010; pp. 318–323.
24. Frias-Martinez, V.; Soguero, C.; Frias-Martinez, E. Estimation of Urban Commuting Patterns using Cellphone Network Data. In Proceedings of the ACM SIGKDD International Workshop on Urban Computing, UrbComp ‘12, Beijing, China, 12 August 2012; pp. 9–16.
25. Calabrese, F.; Diao, M.; Di Lorenzo, G.; Ferreira, J.; Ratti, C. Understanding Individual Mobility Patterns from Urban Sensing Data: A Mobile Phone Trace Example. *Transp. Res. Part C* **2013**, *26*, 301–313. [\[CrossRef\]](#)
26. Chen, C.; Bian, L.; Ma, J. From Traces to Trajectories: How well can we guess activity locations from mobile phone traces? *Transp. Res. Part C Emerg. Technol.* **2014**, *46*, 326–337. [\[CrossRef\]](#)
27. Çolak, S.; Alexander, L.P.; Alvim, B.G.; Mehndiretta, S.R.; Gonzalez, M.C. Analyzing Cell Phone Location Data for Urban Travel: Current Methods, Limitations and Opportunities. In Proceedings of the Conference of European Statisticians, Washington, DC, USA, 29 April–1 May 2015; pp. 1–17.
28. Dumbliauskas, V.; Grigonis, V.; Barauskas, A. Application of Google-based Data for Travel Time Analysis: Kaunas City Case Study. *Promet Traffic Transp.* **2017**, *29*, 613–621. [\[CrossRef\]](#)
29. Allen, I.E.; Seaman, C.a. Likert Scales and Data Analysis. *Qual. Prog.* **2007**, *40*, 64–65.
30. Candia, J.; González, M.C.; Wang, P.; Schoenharl, T.; Madey, G.; Barabási, A.-L. Uncovering Individual and Collective Human Dynamics from Mobile Phone Records. *J. Phys. A Math. Theor.* **2008**, *41*, 224015. [\[CrossRef\]](#)
31. Caceres, N.; Wideberg, J.P.; Benitez, F.G. Review of Traffic Data Estimation Extracted from Cellular Networks. *IET Intell. Transp. Syst.* **2008**, *2*, 179–192. [\[CrossRef\]](#)

32. Caceres, N.; Wideberg, J.P.; Benitez, F.G. Deriving Origin-Destination Data from a Mobile Phone Network. *IET Intell. Transp. Syst.* **2007**, *1*, 15–26. [CrossRef]
33. Ahas, R.; Aasa, A.; Mark, Ü.; Pae, T.; Kull, A. Seasonal Tourism Spaces in Estonia: Case Study with Mobile Positioning Data. *Tour. Manag.* **2007**, *28*, 898–910. [CrossRef]
34. Ratti, C.; Sevtsuk, A.; Huang, S.; Pailer, R. Mobile Landscapes: Graz in Real Time. In *Location Based Services and TeleCartography*; Gartner, G., Cartwright, W., Peterson, M.P., Eds.; Springer: Berlin, Germany, 2007; pp. 433–444.
35. Calabrese, F.; Ratti, C. Real Time Rome. *Networks and Communication Studies, NETCOM.* **2006**, *20*, 3–4.
36. Williams, N.E.; Thomas, T.A.; Dunbar, M.; Eagle, N.; Dobra, A. Measures of Human Mobility Using Mobile Phone Records Enhanced with GIS Data. *PLoS ONE* **2015**, *10*, e0133630. [CrossRef] [PubMed]
37. Calabrese, F.; Di Lorenzo, G. Estimating Origin-Destination Flows using Mobile Phone Location Data. *Cell* **2011**, *10*, 36–44. [CrossRef]
38. Dash, M.; Koo, K.K.; Decraene, J.; Yap, G.-E.; Wu, W.; Gomes, J.B.; Shi-Nash, A.; Li, X. CDR-To-MoVis: Developing a Mobility Visualisation System from CDR Data. In Proceedings of the 2015 IEEE 31st International Conference on Data Engineering, Seoul, Korea, 13–17 April 2015; pp. 1452–1455.
39. Pulselli, R.M.; Puselli, F.M.; Ratti, C.; Tiezzi, E. Dissipative Structures for Understanding Cities: Resource Flows and Mobility Patterns. In Proceedings of the 1st International Conference on Built Environment Complexity, Becon 2005, University of Liverpool, Liverpool, UK, 11–14 September 2005; pp. 271–279.
40. Reades, J.; Calabrese, F.; Sevtsuk, A.; Ratti, C. Cellular Census: Explorations in Urban Data Collection. *IEEE Pervasive Comput.* **2007**, *6*, 30–38. [CrossRef]
41. Zakon o Elektroničkim Komunikacijama. [Law on Electronic Communication]. *Narodne Novine* [The People's Newspaper]. 2014, 71. Available online: <https://www.zakon.hr/z/1181/Zakon-o-elektroni%C4%8Dkim-komunikacijama-2014-2017> (accessed on 18 April 2017).
42. Liu, L.; Biderman, A.; Ratti, C. Urban Mobility Landscape: Real Time Monitoring of Urban Mobility Patterns. In Proceedings of the 11th International Conference on Computers in Urban Planning and Urban Management, Hong Kong, China, 16–18 June 2009; pp. 1–16.
43. Iqbal, M.S.; Choudhury, C.F.; Wang, P.; González, M.C. Development of Origin-Destination Matrices using Mobile Phone Call Data. *Transp. Res. Part C Emerg. Technol.* **2014**, *40*, 63–74. [CrossRef]
44. Wang, P.; Hunter, T.; Bayen, A.M.; Schechtner, K.; González, M.C. Understanding Road Usage Patterns in Urban Areas. *Sci. Rep.* **2012**, *2*, 1001. [CrossRef]
45. Alexander, L.; Jiang, S.; Murga, M.; González, M.C. Origin-Destination Trips by Purpose and Time of Day Inferred from Mobile Phone Data. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 240–250. [CrossRef]
46. Trevisani, E.; Vitaletti, A. Cell-ID Location Technique, Limits and Benefits: An Experimental Study. In Proceedings of the Sixth IEEE Workshop on Mobile Computing Systems and Applications, Windermere, Cumbria, UK, 3 December 2004; pp. 51–60.
47. Chaoming, S.; Qu, Z.; Blumm, N.; Barabási, A.-L. Limits of Predictability in Human Mobility. *Science* **2010**, *327*, 1018–1021.
48. Bengtsson, L.; Lu, X.; Garfield, R. *Internal Population Displacement in Haiti*; Earthquakes Columbia University: New York, NY, USA, 14 May 2010.
49. Thiessenhusen, K.U.; Schäfer, R.P.; Lang, T. *Traffic Data from Cell Phones: A Comparison with Loops and Probe Vehicle Data*; Institute of Transport Research, German Aerospace Center: Cologne, Germany, 2003.
50. Ygnace, J.-L. *Travel Time/Speed Estimates on the French Rhone Corridor Network Using Cellular Phones as Probes, Final Report, Serti V Program STRIP*; Institut national de recherch sur les transports et leur securite: Lyon, France, 2001.
51. Vidović, K.; Lučić, A.; Šarić, D. Methodology of O-D Matrix Determination from Telecommunication Charging Data Records. In *Zbornik radova 36. Skupa o prometnim sustavima s medunarodnim sudjelovanjem KoREMA "Automatizacija u prometu 2016"*, Proceedings of 36. skup o prometnim sustavima s medunarodnim sudjelovanjem KoREMA "Automatizacija u prometu 2016", Krapina, Croatia, Maribor, Slovenia, 9–12 November 2016; Šakić, Ž., Ed.; KoREMA: Zagreb, Croatia, 2016; pp. 22–25.
52. Wesolowski, A.; Eagle, N.; Noor, A.M.; Snow, R.W.; Buckee, C.O. The Impact of Biases in Mobile Phone Ownership on Estimates of Human Mobility. *J. R. Soc. Interface* **2013**, *10*, 1–6. [CrossRef]
53. National Research Council. *Highway Capacity Manual (HCM)*; Transportation Research Board (TRB), National Academies of Science: Washington, DC, USA, 2010.

54. Dadić, I.; Kos, G.; Ševrović, M. *Teorija prometnog toka*; Fakultet prometnih znanosti Sveučilišta u Zagrebu: Zagreb, Croatia, 2014.
55. Sakai, T.; Yamada-Kawai, K.; Matsumoto, H.; Uchida, T. New Measure of the Level of Service for Basic Expressway Segments Incorporating Customer Satisfaction. *Procedia Soc. Behav. Sci.* **2011**, *16*, 57–68. [[CrossRef](#)]
56. Morfoulaki, M.; Tyrinopoulos, Y.; Aifadopoulou, G. Estimation of Satisfied Customers in Public Transport Systems: A New Methodological Approach. *J. Transp. Res. Forum* **2007**, *46*, 63–72. [[CrossRef](#)]
57. Kos, G.; Brlek, P.; Poic, K.; Vidović, K. Correct Temporary Regulation in Cities in the Function of Traffic Safety. In Proceedings of the 11th International Symposium Road Accidents Prevention, Novi Sad, Serbia, 11–12 October 2012; Faculty of Technical Sciences: Novi Sad, Serbia, 2012.
58. Marsden, G.; Snell, C. The Role of Indicators, Targets and Monitoring in Decision-Support for Transport. *Eur. J. Transp. Infrastruct. Res.* **2009**, *9*, 219–236.
59. Alberti, M. Measuring Urban Sustainability. *Environ. Impact Assess. Rev.* **1996**, *16*, 381–424. [[CrossRef](#)]
60. Marsden Sotirios, G.T. *Improved Indicators for Sustainable Transport and Planning: Measuring Wider Economic Benefits of Transport—A Case Study in Good Practice for Indicator Selection*; Institute for Transport Studies, University of Leeds: Leeds, UK, 2008.
61. Anastasi, G.; Antonelli, M.; Bechini, A.; Brienza, S.; D’Andrea, E. Urban and Social Sensing for Sustainable Mobility in Smart Cities. In Proceedings of the 2013 Sustainable Internet and ICT for Sustainability (SustainIT), Palermo, Italy, 30–31 October 2013.
62. Huang, L.; Jianguo, W.; Yan, L. Defining and Measuring Urban Sustainability: A Review of Indicator. *Landscape Ecol.* **2015**, *30*, 1175–1193. [[CrossRef](#)]
63. Costa, M.S.; Silva, A.N.R.; Ramos, R.A.R. Sustainable Urban Mobility: A Comparative Study and the Basis for a Management System in Brazil and Portugal. *Urban Transp. XI Urban Transp. Environ. 21st Century* **2005**, *77*, 323–332.
64. Kaparias, I.; Bell, M.G.H. *Key Performance Indicators for Traffic Management and Intelligent Transport Systems Deliverable*; Imperial College London: London, UK, 2011.
65. Cianfano, M.; Gentile, G.; Nussio, F.; Tuffanelli, G.; Rossi, P.; Vergini, R.; Terenzi, G. *Transport and Mobility Indicators Based on Traffic Measures – Good Practices Guidebook*; Citeair: Rome, Italy, 2011.
66. Nicolas, J.P.; Pochet, P.; Poimboeuf, H. Towards Sustainable Mobility Indicators: Application to the Lyons Conurbation. *Transp. Policy* **2003**, *10*, 197–208. [[CrossRef](#)]
67. Imran, M.; Low, N. Time to Change the Old Paradigm: Promoting Sustainable Urban Transport. *World Transp. Policy Pract.* **2003**, *9*, 32–39.
68. Zeitman, J.; Rilett, L.R. *Sustainable Transportation: Conceptualization and Performance Measures*; Texas Transportation Institute: Austin, TX, USA, 2002.
69. Frei, F. Sampling Mobility Index: Case Study in Assis-Brazil. *Transp. Res. Part A Policy Pract.* **2006**, *40*, 792–799. [[CrossRef](#)]
70. Cortright, J. *Measuring Urban Transportation Performance*; Rockefeller Foundation: New York, NY, USA, 2010.
71. Murr, S.; Phillips, S. The Proposal of a Shared Mobility City Index to Support Investment Decision Making for Carsharing. *World Acad. Sci. Eng. Technol. Int. J. Soc. Behav. Educ. Econ. Bus. Ind. Eng.* **2016**, *10*, 636–643.
72. Liao, S.H. Expert System Methodologies and Application—A Decade Review from 1995–2004. *Expert Syst. Appl.* **2005**, *28*, 93–103. [[CrossRef](#)]
73. Zhang, D.; Zhao, J.; Zhang, F.; He, T. UrbanCPS: A Cyber-Physical System based on Multi-source Big Infrastructure Data for Heterogeneous Model Integration. In Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems, ICCPS ‘15, Seattle, WA, USA, 14–16 April 2015; pp. 238–247.
74. Filić, M.; Filjar, R.; Vidović, K. Graphical Presentation on Origin-Destination Matrix in R Statistical Environment. In *Zbornik radova 36. Skupa o prometnim sustavima s međunarodnim sudjelovanjem KoREMA “Automatizacija u prometu 2016”*, Proceedings of 36. skup o prometnim sustavima s međunarodnim sudjelovanjem KoREMA “Automatizacija u prometu 2016”, Krapina, Croatia, Maribor, Slovenia, 9–12 November 2016; Šakić, Ž., Ed.; KoREMA: Zagreb, Croatia, 2016; pp. 26–29.
75. Filjar, R.; Filić, M.; Lučić, K.; Vidović, K.; Šarić, D. Anatomy of Origin-Destination Matrix Derived from GNSS Alternatives. *Coordinates* **2016**, *XII*, 8–10.
76. Watson, F. *Text-Book on Spherical Astronomy*; Cambridge University Press: Cambridge, UK, 1937.

77. Brčić, D.; Šimunović, L.; Slavulj, M. *Upravljanje prijevoznom potražnjom u gradovima*; Fakultet prometnih znanosti Sveučilišta u Zagrebu: Zagreb, Croatia, 2016.
78. Ross, T.J. *Fuzzy Logic with Engineering Applications*, 2nd ed.; John Wiley & Sons Ltd.: Chichester, UK, 2004.
79. Hrvatska enciklopedija, Leksikografski zavod Miroslav Krleža. Available online: <https://www.enciklopedija.hr/natuknica.aspx?id=68389> (accessed on 18 April 2017).
80. Dernoncourt, F. *Introduction to Fuzzy Logic Control*; Massachusetts Institute of Technology: Boston, MA, USA, 2010; pp. 1–12.
81. Yager, R.R. An Introduction to Fuzzy Set Theory. *Adv. Human Factors Ergon.* **1986**, *6*, 29–39.
82. Ibrahim, A. *Fuzzy Logic for Embedded Systems Application*; Elsevier: New York, NY, USA, 2004.
83. Benšić, M.; Šuvak, N. *Primjenjena Statistika*; Sveučilište, J.J., Ed.; Strossmayera u Osijeku: Osijek, Croatia, 2013.
84. McKillup, S. *Statistics Explained: An Introductory Guide for Life Scientists*; Cambridge University Press: Cambridge, UK, 2012.
85. Turevskiy, A. *Fuzzy Logic Toolbox User's Guide*; MathWorks: Natick, MA, USA, 2012.
86. Vaidhehi, V. The Role of Dataset in Training ANFIS System for Course Advisor. *Int. J. Innov. Res. Adv. Eng.* **2014**, *1*, 249–253.
87. Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
88. Jang, J.S.R.; Sun, C.T. Neuro-Fuzzy Modeling and Control. *Proc. IEEE* **1995**, *83*, 378–406. [[CrossRef](#)]
89. Jang, J.R.; Sun, C.T.; Mizutani, E. *Neuro-Fuzzy and Soft Computing*, 1st ed.; Prentice-Hall Inc.: Upper Saddle River, NJ, USA, 1997; pp. 1–614.
90. Gorzalczany, M.B. *Computational Intelligence Systems and Application*; Physica-Verlag: Heidelberg, Germany, 2002.
91. Foong, K.C.; Chee, C.T.; Wei, L.S. Adaptive Network Fuzzy Inference System (ANFIS) Handoff Algorithm. In Proceedings of the 2009 International Conference on Future Computer and Communication, Kuala Lumpur, Malaysia, 3–5 April 2009; pp. 195–198.
92. Sarle, W.S. Stopped Training and Other Remedies for Overfitting. In Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics; 1995; pp. 352–360.
93. Vidovic, K.; Mandzuka, S.; Sostaric, M. Expert System for Urban Multimodal Mobility Estimation Based on Information from Public Mobile Network. In Proceedings of the Computational Science and Its Applications—ICCSA 2019 19th International Conference, Saint Petersburg, Russia, 1–4 July 2019; pp. 3–11.
94. Šimunović, L.; Bošnjak, I.; Mandžuka, S. Intelligent transport systems and pedestrian traffic. *Promet Traffic Transp.* **2009**, *21*, 141–152. [[CrossRef](#)]



© 2020 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/>).