CHALMERS
UNIVERSITY OF TECHNOLOGY

Mapping the spatial distribution of the effects of urban traffic congestion: methodological exploration using web based services

Downloaded from: https://research.chalmers.se, 2025-07-01 06:35 UTC

Citation for the original published paper (version of record):
Cohen, J., Gil, J. (2020). Mapping the spatial distribution of the effects of urban traffic congestion: methodological
exploration using web based services. IOP Conference Series: Earth and Environmental Science, 588(3). http://dx.doi.org/10.1088/1755-1315/588/3/032059

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

Mapping the spatial distribution of the effects of urban traffic congestion: methodological exploration using web based services

To cite this article: Jonathan Cohen and Jorge Gil 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **588** 032059

View the article online for updates and enhancements.

This content was downloaded from IP address 129.16.140.24 on 28/12/2020 at 11:44

IOP Conf. Series: Earth and Environmental Science **588** (2020) 032059 doi:10.1088/1755-1315/588/3/032059

Mapping the spatial distribution of the effects of urban traffic congestion: methodological exploration using web based services

Jonathan Cohen and Jorge Gil

Department of Architecture and Civil Engineering, Chalmers University of Technology, Sven Hultins gata 6, SE-412 96, Göteborg, Sweden

E-mail: jonathan.cohen@chalmers.se

Abstract. Sustainable transport systems are a necessary requirement to achieve efficient economic performance, enhance urban quality of life and diminish environmental costs. Congestion, a negative externality of mobility, is responsible for urban pollution, inefficiency and has adverse effects over individuals facing this problem. For these reasons, transport and city planning agencies have developed interests in defining and measuring transportation congestion. Although, different definitions and metrics have been used, congestion measurements are found aggregated at a city level or for particular road segments. This study proposes a methodology that produces information from a web traffic service to map traffic congestion within an urban area. The method is simple and generalizable enough to be adopted in different urban areas. This paper presents the analysis of four European cities (Amsertdam, Glasgow, Goteborg and Lisbon) and show that the conclusions are consistent with the results obtained from internationally recognized organizations such as INRIX and TomTom.

1. Introduction

By 2030, United Nations projected that urban areas will host 60% of global population, within cities over 80% of the total wealth is being produced [1, 2] and the positive externalities associated with urbanity have been extensible registered. Cities are successful because at their core, are places of intense human interactions, allowing people to connect easily [3]. Consequently, the success of cities cannot be explained without the fundamental role of mobility [4]. However, as cities become more attractive, population density increases and a set of negative externalities such as pollution, crime and congestion erode the benefits of urban life [5, 6, 7, 3, 8]. Traffic congestion is a major urban problem [9] because it affects cities' core ability to connect people to other people which is a crucial factor for social and economic development [10]. The importance of mobility is reflected in the Agenda 2030 for Sustainable Development where transportation related indicators are found within 8 (out of 18) of the Sustainable Development goals. Most specific transport indicators are directly included in sub-goals: 3.6, 7.3, 9.1, 11.2 & 12.c and indirectly in 2.3, 3.9, 6.1, 11.6, 12.3, 13.1 & 13.2 [11, 12, 13].

In the last century, cities all over the world have been facing a constant increase in the demand for transportation services, resulting in severe traffic congestion. The future does not look promising, nothing indicates that the situation will get any better. A survey to transportation professionals in [14] shows that almost 80% believe that congestion problems have worsened. If

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

nothing is done travel time, energy consumption and environmental costs will continue rising [15, 4].

There is an extensive amount of research that has been quantifying the negative impacts of traffic congestion, the European Commission [16] has warned that Europe might lose economic competitiveness and external costs of congestion were estimated to increase by 80 billion EUR (1% of the EU GDP). In a similar note, [17, 18, 19] argues that these social costs are equal to 2-3% of GDP and [20] arrives to similar conclusions for the U.S.

[21] decomposes the costs of congestion and provide a detailed methodological tool to quantify the total welfare loss and [22] demonstrate that commuting time has a negative effect on well-being.

Significant amount of resources are allocated to understand and measure traffic within cities. Decrease in technological costs and fast adoption of mobile devices introduce the possibility of collecting, analyzing and modelling traffic congestion on a wider and more precise scale that in the past [23]. In the past 10 years new information sources and techniques were explored. Cameras and sensors can provide accurate measurements, but only specific road segments are evaluated and the cost of maintaining these devices is high. Alternative methods to collect data at large scale and lower cost are needed[24, 25, 26].

The aim of this research is to develop a methodology to measure how the effects of congestion are distributed across an urban area. Traditional congestion measurements of streets provide information about the amount of vehicles passing through a road but no connection can be established to who is affected or where is it coming from. On the other hand, Origin and Destination surveys can provide this information but they relay on perception and are costly to deploy. This study will provide exploit an online routing service, to construct a synthetic data set of trips that can be mapped. The map of congestion can provide useful information to local authorities and planners since more dis-aggregated spatial information will be provided.

2. Theoretical Background

2.1. What is congestion

The definition of congestion in transportation facilities has evolved over the years [10] and there is no universal accepted definition for the problem. In fact, [14] shows that in practice the definition of congestion is a contested arena, with no clear predominance of one definition over the other. Time, speed, volume of vehicles, level of service (LOS) and cycle failure are among the most used components to define congestion. According to [19], different actors are interested in understanding different aspects of the problem of traffic congestion; hence depending on the what the objective of the study is, different definitions can be adopted, impacting in what metrics are being used and what information is needed to be collected [27].

In some cases the definition of congestion is confused with the consequences it produces. [27], proposes to categorize the definitions in 3 typologies: (1) demand capacity related, (2) delay-travel time related, and (3) cost related. But in the second and third types, congestion is defined by its outputs and this is found conceptually inaccurate.

Traffic engineers would define congestion as a situation (without a negative connotation) produced when the amount of infrastructure users exceeds it's capacity (type 1). The overuse of the infrastructure results in a set of side effects [4]. According to [10], to be helpful in congestion management decisions, the definition should be based on a comparison of "actual travel times" with "expected travel times" for peak hour and off-peak conditions(type 2). As highlighted by [9], attention must be paid to what the *non-congested* scenario means. There is no hegemonic form of defining congestion and the debate still continues [28].

Modern definitions include the idea of acceptable waiting or travel times (type 2). This definition introduces flexibility and allows the problem to be adjusted depending on different local contexts [9, 10].

Finally, [29, 30] makes a distinction between two types of congestion, based on what created the time delay. Non-recurring congestion are those delays produced by a random event such as an accidents, concerts or vehicle breakdowns, whereas recurring congestion makes reference to those delays that occur at the same place and time, usually on working days. This study, as most of the research and policy concerns with the later form and travel time related measurements will be further explored.

How traffic congestion is defined has direct impact on how it will be measured. For instance, if congestion is defined as travelling below normally accepted travel speed, because of high density of traffic flow, then a threshold is needed of what is the *normally accepted* travel speed.

2.2. Metrics and maps

[19] suggests that any congestion measure should show clarity and simplicity, describe the magnitude, allow comparison, includes time and must be related to congestion relief. Using this list of attributes, a variety of congestion measurements were evaluated suggesting that different uses demand different measurements [27, 19]. When the objective of the study is to understand what places are affected by congestion, type 1 measurements based on the supply side are not useful.

A family of indicators was calculated based on absolute values such as the **delay rate**, which consisted in the difference between actual travel time and a free-flow situation or accepted time. Also, relative metrics can be generated using the same information. For instance, taking the **delay rate** relative to the accepted travel rate generates the **Delay Ratio or Travel Time Index (TTI)**.

After processing traffic information, typically the results are shown in tables that rank different cities, but planning and transportation agencies would need more information to generate effective traffic congestion reduction policies. There is a tradition of maps showing congested road segments (type 1) that naturally leads to responses in terms transport infrastructure capacity and flow management. However, after the extensive review done in this study, almost no maps were found showing travel times, how congestion varies within an urban area, and which areas of a city were affected the most. To design better transportation policies, information about, origin, destination and motive of travel are needed. Furthermore, from the reports showing the ranking of cities it would be virtually impossible to know which citizens are facing this problem.

3. Methodology

This section provides a detailed description of a methodology for calculating congestion based on travel time differences and to spatially represent how the impact of congestion is distributed across locations in the city. At its core the methodology described in this section details a process of collecting and processing data from an internet service that provides traffic estimations and routing optimization. In general, the process has five main stages: (i) the study area is defined and subdivided, (ii) a list of trips is created, (iii) travel time data is extracted and (vi) the data captured is processed, analysed and (v) the indices are calculated and mapped. Given the nature of the methodology described here, it was coded and executed using the R package in Rstudio and the source code can be accessed in GitHub (https://github.com/Urban-JonathanCohen/Congestion-Index).

3.1. Defining the study area

In order to understand how congestion is distributed across an urban area, the first step is to select and define the boundaries of the area of interest. Once this is settled, the area must be subdivided into different geographical regions. For example, the urban area could be subdivided using a regular grid, neighbourhoods, or other institutional subdivisions, such as census tracks.

Using an existing subdivision will allow the travel time information to be related to different socio-economic or environmental characteristics. After the urban area is subdivided, the centroid of each polygon is extracted as a latitude and longitude coordinate.

3.2. Creating a synthetic Origin-Destination list
The centroids from the first step are now the origin (O) or the destination (D) of a trip, and making pair permutations among the set of centroids creates an list of trips from all centroids to all others. Taking n as the number of subdivisions, this process creates a total of of n*(n-1) trips. For instance, if the study area contains 3 subdivisions (A, B & C), the list will contain the following 6 trips: (i) A to B, (ii) A to C, (iii) B to A, (iv) B to C, (v) C to A & (vi) C to B.

3.3. Extracting travel time data
As seen in the previous section, one of the manifestations of congestion is to experience longer travel times. To determine the level of congestion, a benchmark *non-congested* scenario needs to be used for comparison and calculation of the travel time difference to a *congested* scenario. For instance, the morning rush hour (8:30AM) can be used as a *congested* scenario and late night (3AM) scenario can be used as the *non-congested* scenario. For using the Google Distance Matrix tool to calculate travel times, one must select future dates (the time, day, month and year) for these scenarios. Thus, it is important to take into consideration holidays and the day of the week. Fridays can be less busy days than Wednesdays, or December and July show less activity than October (in the northern hemisphere countries). After specifying the dates for the *congested and non-congested* scenarios, the OD list is used to make a request for each trip. The API request includes an origin, a destination, a time and a travel mode, and the generated response (json file) contains data of the addresses, the length of the route taken, and the travel time. The data set of all the trips for the two scenarios is stored in a data table.

3.4. Processing and analyzing data
The data obtained is then reviewed for completion and completeness, i.e. if all the trips and their details were generated correctly. In some cases, the centroid is in an inaccessible location by car, e.g. a lake or the ocean, and returns null values, so those points are removed from the analysis. Then basic descriptive statistics for the trips' length and time are reported for inspection. In the presence of outliers, care must be taken to investigate if these are data extraction errors or simply extreme values. For further scrutinizing the data, the distribution of travel times can be observed by plotting histograms and scatter plots of distance against travel time.

3.5. Calculating the metrics and mapping
Once the data set is consolidated, two congestion metrics are calculated: (i) the time difference in minutes between the *congested and non-congested* scenarios and (ii) the Travel Time Index, which is the ratio between the *congested and non-congested* scenarios. Finally, taking the mean values, the results are aggregated at each origin and joined to the polygons used at the beginning of this process.

In order to obtain values at the level of the whole metropolitan area, one can take the average travel time values of the entire data set, and calculate the travel time difference and TTI.

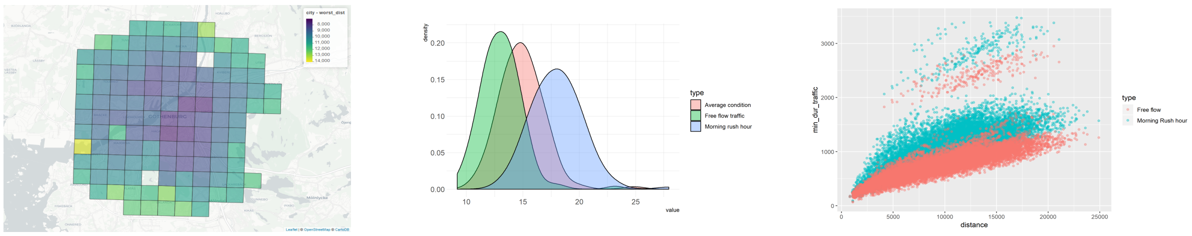
4. Test case
To demonstrate how the method can be used in practice, four European cities where selected as a proof of concept (Amsterdam, Glasgow, Gothenburg and Lisbon). In this case, a buffer zone from their historical centre was used to delimit the *city* boundary, and the 1km2 population grid from EuroStat (GEOSTAT 2011) was used to provide the geographic subdivision. The

geographic coordinates of the centroid of each 1km2 grid cell were used to build the list of all possible trips.

From this list of trips, a synthetic OD matrix was created to retrieve data for a *congested* scenario (Thursday, October 15, 2020 8:30:00 AM), and a *non-congested* scenario (Wednesday, July 15, 2020 3:30:00 AM).

4.1. Data retrieved

The process generates a large number of trips for each city. For instance, Lisbon is divided into 119 grid cells and the number of possible trips is 119*118 = 14,042 (then 28,084 trips were estimated). Amsterdam has 131 zones, Glasgow 136 and Goteborg 123 (17,030, 18,360 and 15,006 trips respectively). A summary of the trip data retrieved for the case of Gothenburg can be visualised in Figure 1.



(a) Average trip distance (b) Distribution of travel time (c) Time vs distance

Figure 1: Map and charts extracted from data captured for the case of Gothenburg. (a) The darker the colour, the larger the average distance. (b) Green curve, non-congested; Orange curve: average situation; Blue curve: congested. (c) Green points represent congested scenario data vs non-congested data in red

The amount of trip information retrieved is consistent with the amount of zones in the city and, as expected, the *congested* situation on average takes longer travel time. Formally ,an independent-samples t-test was conducted to compare the travel times under a *congested* and *non-congested* scenario. For each city, the time difference was found statistically significant with a confidence level of 99%.

5. Results

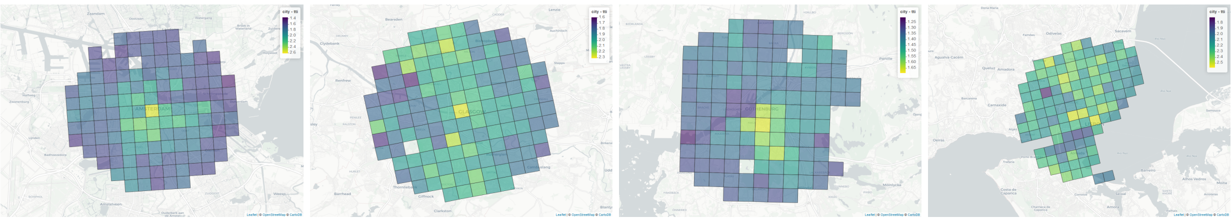
After capturing, processing and aggregating the data obtained from the API, Table 1 on page 6 presents the descriptive statistics for time difference and TTI. In this case, the number of observations corresponds to the number of zones in the map and each zone holds the mean of all the trips departed from that origin. The results show that the method successfully captured diversity of congestion between cities. For instance, in Lisbon the average time difference is of 15.3 minutes, with a maximum of 21.6 minutes, while in Gothenburg on average delays are of 4.9 minutes with a maximum of 8 minutes.

Although, Table 1 reveals relevant insights on these cities, local authorities or planners would not know which places or who is being affected by traffic congestion. In Figure 2 the TTI for each city was mapped. By looking at these maps, several conclusions can be drawn on where and how bad congestion is across the city. For instance, in all four cities, although the historical centre is the nearest point to all other destinations, when congestion is taken into account, these places face the highest TTI. This indicates that in a *non-congested* situation, the city centre is highly accessible but when *congested*, the advantage of being central gets eroded.

Table 1: Aggregated descriptive statistics by zones

	Time Difference (mins)	Zones	Mean	S.D.	Min	Max
1	Amsterdam	131	8.57	2.39	5.38	19.03
2	Glasgow	136	11.15	1.88	7.66	15.64
3	Goteborg	123	4.92	0.99	3.22	8.06
4	Lisbon	119	15.34	2.67	9.74	21.57
	TTI	Zones	Mean	S.D.	Min	Max
1	Amsterdam	131	1.74	0.23	1.36	2.61
2	Glasgow	136	1.94	0.14	1.60	2.31
3	Goteborg	123	1.38	0.09	1.21	1.69
4	Lisbon	119	2.21	0.16	1.70	2.57

These maps also confirm that the methodology presented captured spatial differences across cities. In Gothenburg, the historical city centre suffers most delays from congestion, while the other locations in the city suffer similar levels of congestion. In Lisbon, the impact of traffic congestion is more scattered across the city, showing islands suffering from lower congestion (where the CBD has moved to, away from the historic core).



(a) Amsterdam (b) Glasgow (c) Gothenburg (d) Lisbon

Figure 2: Average Travel Time Index (TTI) by grid cells

6. Discussion

The methodology presented in this paper exploits a new data source to provide spatial insights about traffic congestion in different cities. The information generated allows planning and transport agencies to reconstruct some of the most popular indexes discussed in the theoretical background section.

In order to verify the validity of the methodology presented here, the aggregated results by city were compared against the results published by two internationally recognized companies that offer transportation solutions and, nowadays, insights as a service, INRIX and TomTom. Historically, INRIX has been consulted to provide he levels of congestion in several countries in the OECD and U.S. transport agencies [5, 18], but given the increase in the number of devices embedded with GPS, other businesses like TomTom have found themselves in a position to use these new data sources to offer consulting services. [31] and [32] publish results on traffic congestion and rank cities according to different criteria. For instance, INRIX reports the amount of hours lost in congestion, the cost per driver, travel time and speed whereas. TOMTOM shows a congestion percentage. Despite these differences, the ranking of cities from the proposed methodology compared with TomTom and INRIX rankings gives the exact same results. Gothenburg is positioned as the least congested city, followed by Amsterdam, Glasgow and finally Lisbon. It is relevant to highlight that INRIX and TomTom provide more detailed information for each city, but no maps are found and the spatial distribution of the congestion problem remains unknown.

Traffic congestion deteriorates different domains of urban life and as a consequence the focus on the problem varies across disciplines. *'Throughout the world, traffic congestion reduces the core benefit of cities: the ability to connect with other people easily'* [3], so it becomes crucial to understand who suffers from congestion the most and what parts of the city are more vulnerable to the problem. Therefore, looking at the congestion level of street segments due to reductions in speed is useful for traffic planning, but insufficient to solve wider issues. The maps of congestion produced as a result this methodology identity areas that area affected by the congestion phenomenon in terms of travel time differences, and establish the relationship between a population living in an area and the level of congestion it face.

When dealing with city planing, it is difficult to understand how an urban intervention or policy affects the rest of the city. This approach can be iterated over time, to see how transport congestion evolves and evaluate to what extent certain policies are being effective in changing traffic congestion patterns.

The results presented here, do not necessarily imply that *'more congested'* places are demanding for solutions from public administration, but attention. *'Traffic problems'* can be the result of poor infrastructure, an excess of demand or a combination of both. The method only reflects time delays faced by private commuting and in some cases, such as in city centres, this can be a desirable tool to demotivate car use.

6.1. Limitations and future work

The methodology presented in this study uses primarily Google's web service, which is a drawback as users of the API have no control over the service standards, usability or costs. For instance, the type of request performed during this study had a cost of 10 U$D/1000 requests. A grid of 100 cells will generate 9,900 trips, and to estimate congestion a total of 19,800 requests will be needed, costing 198 U$D and taking around two hours to complete. Consequently, before considering expanding the study to cover more cities, more extensive metropolitan areas, or repeating the process for different scenarios, budget constraints must be taken into consideration. While the methodology uses the Distance Matrix API form Google Inc., a similar exercise could be done with other services such as HERE or TomTom, with different cost models.

The results presented here use the EuroStat population 1km2 grid cell to retrieve traffic information. Certainly, the shape/size of these areas can be subject of debate. Although to map the historical city center this 1km2 grid representation is too big, it can make sense for planning and the data retrieval became affordable. Furthermore, the zones used in the analysis can be replaced with other geographical representations, allowing the process to be enriched in several dimensions. Using census tracks with socio-economic information on the number of cars or age groups could help to better understand to what degree citizens are affected, and to explore the relation between car use and congestion levels.

Finally, this work presents an alternative of how synthetic data can be used to provide new insights for sustainable development. The method only uses a simplistic OD matrix with modelled traffic information, but traditional data sources such as population census or travel surveys can be combined to enhance the results. In the future, the population grid can be used to weigh the results and give relative importance to different zones based on population. Travel surveys, with information about origin and destination, travel time, commuting modes, and vehicle type, can be used to weight the different OD pairs, or to model precisely how much green house gases are being emitted and bench-marked under a *non-congested scenario.*

7. Conclusion

The study presented a methodological approach to study how traffic congestion is spatially distributed within an urban area affecting locations differently. Generating a synthetic Origin-Destination matrix, the method uses an online routing service to estimate the travel times of different trips for *congested and non-congested* scenarios and calculates travel time difference and TTI for the two scenarios. The aggregation of the results for four European cities into a city ranking is consistent with the ranking from internationally recognized institutions such as INRIX and TomTom. The methodology provides a non-expensive, generalizable and systematic method to estimate the impact of congestion in different parts of a city. The congestion indices calculated can be mapped, thus used by city planners and transportation agencies to support their strategic planning decisions towards the sustainable development goals.

References

[1] McKinsey Global Institute 2011 Urban world : Mapping the economic power of cities Tech. Rep. March
[2] UN 2018 *The World's Cities in 2018 - Data Booklet (ST/ESA/ SER.A/417)* 34
[3] Glaeser E L 2011 *Science* ******* 592–595
[4] Diakaki P, Kotsialos D and Wang 2003 *Proceedings of the IEEE* **91** 2041–2042 ISSN 0018-9219
[5] Eurostat 2016 *Urban Europe* ISBN 978-92-79-60139-2
[6] United Nations 2016 TRANSFORMING OUR WORLD: THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT (A/RES/70/1) Tech. rep. United Nations
[7] Jones P and Hervik A 1992 *Transportation Research Part A: Policy and Practice* **26** 133 – 145 ISSN 0965-8564
[8] Stempfel J, Guler S I, Menéndez M and Brucks W M 2016 *Journal of Transportation Safety and Security* **8** 214–229 ISSN 19439970
[9] Downs A 1992 *Still stuck on traffic - Coping with Peak-Hour Traffic Congestion* ISBN 0815719299
[10] Falcocchio J C and Levinson H S 2015 *Springer Tracts on Transportation and Traffic* **7** 93–110 ISSN 21948127
[11] Booth D, Hanmer L and Lovell E 2000 Poverty and Transport Tech. Rep. June World Bank and DFID
[12] UN-Habitat, UNEP and SLoCaT 2015 Analysis of the Transport Relevance of Each of the 17 SDGs Tech. Rep. September UN-Habitat, UNEP, SLoCaT
[13] Jennings G 2016 TRANSPORT, POVERTY ALLEVIATION AND THE PRINCIPLES OF SOCIAL JUSTICE Tech. rep. Sustainable, Low Carbon Transport (SLoCaT)
[14] Bertini R L 2005 *85th Annual Meetinf of the Transportation Research Board* 17 ISSN 1091-5818
[15] Bull A 2004 *Traffic Congestion - The Problem and How to Deal with it?* ISBN 92-1-121432-7
[16] European Commission 2001 *WHITE PAPER: European transport policy for 2010: time to decide* 2001 ISBN COM(2001) 370 final
[17] Cervero R 1998 *The Transit Metropolis* (Washington; D.C.: Island Press)
[18] The Texas A&M Transportation Institute 2019 Urban Mobility Report Tech. Rep. August
[19] Lomax T, Turner S and Shunk G 1997 NCHRP Report 389: Quantifying Congestion Volume 1 - Final Report Tech. rep. Textas Transport Institute (TTI)
[20] Schrank D, Lomax T and Eisele B 2011 TTI's 2011 URBAN MOBILITY REPORT Powered by INRIX Traffic Data Tech. Rep. September Texas Transportation Institute
[21] Bilbao-ubillos J 2008 **42** 1098–1108
[22] Choi J, Coughlin J F and Ambrosio L D 2013 100–108
[23] Rempe F, Huber G and Bogenberger K 2016 *Transportation Research Procedia* **15** 513–524 ISSN 2352-1465
[24] Pongpaibool P, Tangamchit P and Noodwong K 2007 *IEEE Region 10 Annual International Conference, Proceedings/TENCON* 1–4
[25] Wang S, He L, Stenneth L, Yu P S and Li Z 2015 *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems* **03-06-Nove**
[26] Pattara-atikom W, Pongpaibool P and Thajchayapong S 2006 *ITST 2006 - 2006 6th International Conference on ITS Telecommunications, Proceedings* 1001–1004
[27] Aftabuzzaman M 2007 *30th Australasian Transport Research Forum* 1–16
[28] Hamad K and Kikuchi S 2002 *Transportation Research Record* **1802** 77–85
[29] Stopher P R 2006 **11** 117–131
[30] Skabardonis A, Varaiya P and Petty K F 2003 *Transportation Research Record* 118–124 ISSN 03611981
[31] Inrix 2018 INRIX 2018 Global Traffic Scorecard URL `http://inrix.com/scorecard/#`
[32] Tom Tom 2018 TOMTOM Traffic Index 2018 URL `https://www.tomtom.com/en_gb/traffic-index/ranking/`