

1. Introduction

Pervasive urban transport problems traverse nearly every country of the world (Pojani and Stead, 2017), hence, the need to plan the management and implementation of urban transport policies in a systematic manner. However, the planning of urban transportation involves the formulation and evaluation of policies regarding different situations for decision makers, and the procedure commences with recognising prevailing problems in public transport, or achieving expected specific objectives, and continues with forecasting and evaluation of perceived scenarios (Huang, 2003). According to Michell (2017), the challenges facing urban transport planners and policy makers are enormous indeed, and different kinds of freshly collected data are important for strategy implementation, policy making, decision making, and realigning these imperatives that have been mentioned; but the data that should be the bedrock of these imperatives are available in disjointed forms in various sources. These data are waiting to be collected and stored appropriately for decision makers and other users; the extent to which these data are collected, and the methods with which they are collected are an important part of data collection for urban transportation planning (Huang et al., 2014). This paper discusses the current data collection strategies, including traditional, technology-based, and automated techniques that are available and applied in the Kingdom of Saudi Arabia and the United Kingdom, as well as the gaps that currently exist in available data and the actual form of data that urban transport planners should be accessing for transport planning in the two countries presently, the main aim of this paper is to provide a gap analysis and comparison of the available data and collection approaches between the two countries.

2. Importance of Data collection and the Types of Data collected in Urban Transport Planning

In the last two decades, there have been significant changes in the socio-economic environment that has substantially impacted urban transport. Most cities of the world are today characterised by a variety of mobility patterns, which have affected travel distances, and contributed to a continuous rise in the number of mobility options that are available within a geographical space (Zannat and Choudhury, 2019). In particular, urban commuting habits has witnessed a progressive dependence on private motor vehicles, which creates negative consequences such as road congestion, as well as other social and environment impacts

(Herrero et al., 2017). According to Herrero et al. (2017), this trend has continued for the last two decades, and the consequences of this is adverse effects on sustainable development, the environment, the health and safety of citizens, the economy, the generality of the society, and in particular the standard of living of the population working and living in urban areas (Herrero et al., 2017). This scenario has necessitated the importance of the acquisition and deployment of a sustainable and cohesive urban mobility systems that are less dependent on carbon-generating fuels (Eboli and Mazzulla, 2012). Hence, the countries of the world, especially towns and cities in the United Kingdom (UK) and the Kingdom of Saudi Arabia (KSA) are under severe pressure to enhance the performance of their urban transportation systems, in order to lessen the adverse effects of transportation activities on the climate, the environment and citizens' welfare, and to deploy more sustainable urban mobility systems (Pojani and Stead, 2015).

Making plans and projections for transport planning is well established in the Western developed countries due to the level of development of their economy and society. The motor vehicle, ship building, aircraft, and rail industries have been a strong influence in different aspects of the economic and societal development, such as the rate of increase in urban population, strengthened growth in car ownership and travels, urban form environment and development policies, increased rural-urban migration, and land use pattern (Enoch et al., 2004, Huang, 2003). According to Huang (2003), the goals and objectives of urban transport development continues to change due to advances in transport systems and new understandings into social and economic requirements. The changes witnessed have made the modifications of existing models imperative, since these models will usually have different data needs, and model improvements imply possible requirements for new data (Milne and Watling, 2019, Zannat and Choudhury, 2019). Hence, to perform a gap analysis of data for urban transport planning, it is important to understand the data requirements of urban transport planning, this then must commence with an understanding of the current situation regarding the existing applicable strategies for transport data collection, also, this progresses onto data availability, usage of data, the available data processing techniques, and establishing the gaps that exist between actual available data as compared to the actual data requirements (Huang, 2003, Zannat and Choudhury, 2019).

How we measure things potentially influences their perceived value (Litman, 2011), hence, the objective of data collection for transport planning is to collect data that correctly displays the real-world traffic situation in a specific area such that sufficient value can be derived for planning transport requirements (Miller et al., 2014). This could involve counting the number of vehicles,

motorcyclists, and cyclists that ply a particular road, collecting journey time information, collecting incident records, etc. (Stofan, 2018). According to Miller et al. (2014), good quality, high value, wide-ranging data on travel behaviour, transport network performance and associated land use features are essential imperatives to the planning, design, and operations of urban transport systems. This information are collected from different types of sources, however, most of the data are from travel surveys as well as other data collection methods through which base data are collected for analysis and modelling of both trip-making and transport system performance (Miller et al., 2014). Hence, where data is inadequate, it is impossible to understand transport needs, and neither will it be possible to design and assess policy alternative and service delivery (Janssens, 2013). It is therefore important for transport authorities to invest a significant part of their resources in the collection of data using a variety of methods that are valid enough for collecting reliable and valuable data.

The benefits to be derived from the collection of transport data must outweigh the cost of data collection, which means that the effort must highlight visible values to be derived in terms of efficient solutions in the management of road network congestion, air pollution, and infrastructure longevity. Extant studies have attempted to identify the different types of data that are collected for transportation planning. According to Neffendorf et al. (2002), transport planning is a predominantly heavy user of data, they identified the several different types of data that are used for transport planning which are categorised into: economic, statistical, surveys, land use, operational, and transport network. From these, the specific types of data that are collected for the purpose of urban transportation planning include:

- Household activity / trip-making behaviour;
- Volume count data (traffic, riders, etc.);
- Transportation system characteristics (speeds, lane widths, etc.);
- Transportation network inventory;
- Travel time data;
- Vehicle classification counts;
- Transportation cost and service levels;
- Land use characteristics (population, employment, etc.);
- Population socio-economic information (income, auto ownership, etc.);
- Attitudes / opinions / stated choices;
- Safety data analysis;
- Road incident reports;
- System impacts (e.g., emissions);
- Regulatory compliance inventory;
- Video data collection (Miller et al., 2014, Neffendorf et al., 2002).

According to Wachs et al. (2016), the features of respective category of transportation alternatives

including their choice sets need to be appropriately defined in order to achieve the objective of urban transport planning. Hence, sufficient characterization of the data collected for transport planning are essential for ease of representation in identified preference surveys, and in models, also the assumptions that were constructed regarding the data for transportation decision making must be clearly laid out to enable users of the data access to background information regarding the data and be able to insert them in appropriate models for optimal solutions.

3. Data Collection Methods in Urban Public Transport Planning

Data are essential for creating travel demand forecasting models as well as applying and validating the models (Wachs et al., 2016). According to Stofan (2018), there are several different data collection methods that can be applied for collecting the different types of data for urban transport planning, and this depends on the types of data, and the use to which the data being collect will be put. Miller et al. (2014) explained that data collection in transportation planning can be classified into three main groups, namely:

- “Population-based surveys (such as face-to-face, telephone, travel diary, and home- interview surveys)
- Choice-based sample surveys (such as transit on-board and web form surveys)
- Non-survey data collection techniques (both “standard” and “emerging”), which typically involve the use of information technology of which there a variety” (p. 7).

The first group is based on telephone and home interview surveys which have presented substantial methodological challenges for transport planners and authorities due to the spatial elements of a territory as well as the transport network (Trépanier et al., 2008). Also, the sampling sizes for telephone surveys are usually based on landline telephone directories, however, this is no longer practicable as most homes are no longer subscribed to landline telephones, the U.S.A. and U.K. that are still heavily subscribed now request “Do Not Disturb” (DND) service as part of their subscription (Kulpa and Szarata, 2016, Trépanier et al., 2008, Weis et al., 2013). The Choice-based sample surveys are intercept transit and onboard surveys which are conducted based on physical contacts with the respondents that are intercepted during their commute on buses, trains, and at the bus stops; they highlight the Origin- Destination (O-D) trends in transport flows; however, such method of data collection are labour-intensive and costly to implement (Mishalani et al., 2011). The population-based

surveys and the choice-based surveys both belong to the field of survey research, which Prewitt et al. (2014)

argues is at a crossroads, and as a result are facing several different challenges, most of which mainly affects their viability, and the validity of their findings. Also, Prewitt et al. (2014) highlighted that survey research is very expensive, and raises concerns about its long term sustainability.

The non-survey data collection techniques refer to newly emerging techniques of capturing data through digital sources such as sensors and smart devices that are both manual and automatic and are dispersed across cities with the objective of collecting both intrusive and non-intrusive transport data for urban transport planning purposes (Miller et al., 2014, Prewitt et al., 2014). According to Prewitt et al. (2014), these emerging techniques are promising due to the fact that the conventional survey methods such as telephone-based surveys currently suffer viability and reliability issues, and the emerging digital-based alternative techniques offer more valid results, and that is because they can be captured from different information, digital and technological platforms. Prewitt et al. (2014) explained further that advance in computational science analytics increases the potential of creating obscure indications of citizens' civic engagements and social cohesion behaviours, and possibly their opinions. Their explanation concurs with Einav and Levin (2013) who argued that "The recording of individual behaviour does not stop with the internet: text messaging, cell phones and geo-locations, scanner data, employment records, and electronic health records are all part of the data footprint that we now leave behind us." (p. 3). This argument appropriately captures the emerging techniques of data collection, and how their ramifications have gone beyond the imaginations of transportation researchers of the late 20th century.

While the population-based and choice-based sample surveys have been well established and discussed in terms of questionnaire and interview based within the context of probability sampling methods to identify a small sample of the total population with the aim of collecting representative data; the digital sources of data collection (now commonly referred to as "Big Data" sources) have made it possible to collect data on larger population to extract more data that reliably represents the population (Prewitt et al., 2014, Trépanier et al., 2008). Specifically, Zannat and Choudhury (2019) explained that big data sources contain updated and near or real-time spatial and temporal information that are rather impossible to collect through the population-based and choice-based sample surveys; also these data contain vast amounts of individual level data with in-depth details and higher accuracy at lower cost. They also highlighted that some the data are capable of being linked with supplementary data such as land use, bus time tables, as well as personal data, although at a risk of breaching personal privacy regulations. In spite of these strengths, big data also has numerous challenges which are mainly concerned with the collection, processing, and analysis of data, and they include technical and high-level requirements of

computational efficiency, data processing, evaluation, validation, and user privacy (Chen et al., 2016, Gadziński, J., 2018, Zannat and Choudhury, 2019).

The importance of data collection and the types of data that are collected in urban transport planning have been discussed, as well as the data collection methods in urban transport planning. Hereafter, the data collection activities in the UK and KSA will be discussed with a view to performing a gap analysis of the data that are collected for urban transport planning in both countries and then comparing the results of the gap analysis.

4. Data Collection for Public Transport Planning in UK

The UK Department for Transport (DfT) is responsible for the management of public transport, with responsibility for aviation, roads, rail, maritime, driving, freight, public transport, safety, environment and transport accessibility (DfT, 2020). The DfT (2020) provides guidelines for the collection and analysis of data regarding the respective public transport sectors, and provided the Transport Analysis Guidance which is summarized in Table 4.1, and this provides a list of the baseline data types, the sources and their locations of urban transportation data within the UK. Apart from the sources mentioned in Table 4.1, more detailed data are contained in the TSGB's (2019) website as maintained by the UK's DfT. Multiple reports on the UK's transport system are produced from the data that is collected within the DfT which are then used for policy making and decision making in the management of the UK's transportation system. Furthermore, researchers in academic environments, independent researchers and organisational research practitioners do access the data collected within the DfT's database of transport data to model and analyse these data with the aim of providing solutions and suggestions to challenges that arise within public transport in the UK.

Wang (2010) acquired a 100% sample of Automatic Fare Collection (AFC) data to collect a two-week sample of Transport for London (TfL) bus routes, and secondly, a 100% sample of Automatic Vehicle Location (AVL) two-week data for all the routes that are parallel to or intersect the selected bus routes to construct trip chains for individual bus passengers. The result highlighted the feasibility and simplicity of applying the trip-chaining method to infer bus passengers' boarding and alighting locations, also Wang (2010) validated the results by comparing them with the Bus Passenger Origin and Destination (BODS) survey data in London.

Furthermore, researchers in academic environments, independent researchers and organizational research practitioners do access the data collected within the DfT's database of transport data to model and analyze these data with the aim of providing solutions and suggestions to challenges that arise within public transport in the UK.

Wang (2010) acquired a 100% sample of Automatic Fare Collection (AFC) data to collect a two-week sample of Transport for London (TfL) bus routes, and secondly, a 100% sample of Automatic Vehicle Location (AVL) two-week data for all the routes that are parallel to or intersect the selected bus routes to construct trip chains for individual bus passengers. The result highlighted the feasibility and simplicity of applying the trip-chaining method to infer bus passengers' boarding and alighting locations, also Wang (2010) validated the results by comparing them with the Bus Passenger Origin and Destination (BODS) survey data in London. The data were collected from the TfL's Automatic Data Collection Systems (ADCS) which made Wang (2010) to suggest that relying on ADCS can provide transport planners with more wide-ranging, dependable and accurate evidence for service planning urban transport.

Table 1. UK Data Sources and Surveys

Data Type	Source	Location
Planning and Demand Data		
National Trip End Mode (NTEM)	Trip end Model Presentation Program (TEMPro)	https://www.gov.uk/government/organisations/department-for-transport/series/tempro
Spatial Detail	TEMPro data is available down to Middle Layer Super Output Area (MSOA)	-Office for National Statistics population estimates -the Business Register and Employment Survey (BRES) employment data -Ordnance Survey AddressBase data for land use information
Traffic Data		
-Annual Average Daily Flow (AADF)	Highways England Traffic Information System	www.gov.uk/government/organisations/department-for-transport/series/roadtraffic-statistics & http://data.gov.uk/
-Journey time information	Highways England Traffic Information System	http://webtris.highwaysengland.co.uk/
-Public Transport data		
- Rail passenger matrices	LENNON rail ticket information database	Association of Train Operating Companies (ATOC)
-National Rail Travel Survey (NRTS)	Department for Transport	https://www.gov.uk/government/publications/
-Electronic ticket machine (ETM)	Local Bus Operators	Local Bus Operators' Offices and websites
-Smart ticketing	Local Bus Operators	Local Bus Operators' Offices and websites
Census Data		
-2011 census journey to work data		http://data.gov.uk/
National Travel Survey (NTS) Data		
-Household travel Surveys		https://www.gov.uk/government/organisations/department-for-transport/series/national-travel-survey-statistics

		ations/department-for-transport/series/national-travel-survey-statistics
Aviation Data		
-Usage of airports around the UK	The Civil Aviation Authority (CAA)	http://www.caa.co.uk
Network Data		
-Ordnance Survey MasterMap Integrated Transport Network (ITN)	Department for Transportation	www.ordnancesurvey.co.uk/opendata
-National Public Transport Access Node (NaPTAN)	Department for Transportation	http://www.dft.gov.uk/naptan/
-National Public Transport Data Repository (NPTDR)	Department for Transportation	http://data.gov.uk/dataset/nptdr
Study Specific Surveys		
-Non-intrusive automated methods -non-intrusive manual methods -intrusive methods (Interviews)	Department for Transportation	TAG unit M3.1 - Highway Assignment Modelling, and TAG unit M3.2 - Public Transport Assignment
Demand Data Surveys		
-Population	NTEM, Census	
-Households	NTEM, Census	
-Car ownership	NTEM/NATCOP, Census, Local Household Travel Survey, Household Expenditure Survey	
-Socio Economic Group -	Census, Local Household Travel Survey, Household Expenditure Survey	
-Land-use data	Census and Special Workplace Statistics, NTEM for trip ends by purpose, Local Authority planning data for employees, retail and commercial floorspace by zone.	
-Car availability	Local Household Travel Survey, National Travel Survey (NTS)	
-License holding	Local Household Travel Survey, Household Expenditure Survey, National Travel Survey	
-Travel to zonal destinations by: Purpose, Mode and Time of Day (Period)	Roadside Interviews (RSIs), Mobile Network Data (MND), Automatic Number Plate Recognition (ANPR) surveys, Local Household Travel Survey, journey to work from Census, NTEM, NTS	
-Trip lengths	Local Household Travel Survey, National Travel Survey, journey to work Census, RSI surveys	
-Vehicle operating costs	Values from TAG unit A1.4, Values of Time and Operating Costs	

-Vehicle occupancies	RSIs, local Household Travel Survey, NTS	
-Values of time	- TAG unit A1.4, Values of Time and Operating Costs	
Household Interview Surveys		
-Travel by residents	Department for Transportation	
Revealed Preference and Stated Preference Surveys		
-Revealed Preference (RP)	-Supplementary Guidance - Bespoke Mode Choice Models	
-Stated Preference (SP)	- Supplement ary Guidance - Bespoke Mode Choice Models	
Highway Surveys		
-Traffic counts 1. Manual classified counts (MCC) Turning counts at road junctions	Department for Transportation	
-Journey times 1. Moving Car Observer (MCO)	-Department for Transportation	
Roadside Interview (RSI) Surveys		
Roadside Interview Data	TAG unit M2.2	
Public Transport Surveys		
-Passenger Counts	-Department for Transportation	
-Movement Surveys	-Department for Transportation	
-Face-to-face on-board surveys	-Department for Transportation	
-Face-to-face at-stop or at-station surveys	-Department for Transportation	
-Self-completion on-board surveys	-Department for Transportation	
-Self-completion at-stop and at-station surveys	-Department for Transportation	

Source: DfT. (2020). TAG UNIT M1.2: Data Sources and Surveys: Transport Analysis Guidance (TAG), Department for Transport London: TRANSPORT APPRAISAL AND STRATEGIC MODELLING (TASM) DIVISION - DEPARTMENT FOR TRANSPORT.

In the same vein, Goves et al. (2016) applied artificial intelligence such as artificial neural networks to estimate traffic conditions by 15 minutes into the future based on current and historic traffic information, they accessed data from Highway England’s Motorway Incident Detection and Automatic Signalling (MIDAS) system collecting data for 20 km of the M60, M62 and M602 motorway near Manchester, UK. With this they were able to build a short-term prediction model, and in the process, they reduced the complexity of the problem by using an autoencoder to reduce the number of input dimensions to the model, their model showed high quality predictive power with 90% of all predictions within 2.6 veh/km/lane of observed values. The algorithm they developed is capable of multiple applications, such as refining predictions within intelligent transport systems (ITS) and / or supporting traffic controllers to take proactive decisions to mitigate the impacts of expected congestion. Other studies that have highlighted the power of data collection for the planning of urban transport include Zaki et al. (2016), Zhang et al. (2016), Zhang et al. (2013), etc.

5. Data Collection for Urban Public Transport Planning in the Kingdom of Saudi Arabia (KSA)

The KSA has one of the fastest growing rates of urbanization in the world. According to Moreno et al. (2019), 83% of the KSA population are living in urban areas, and the emergence of urbanization in the KSA has supported urban agglomerations which have significantly benefitted from the ensuing prosperity and growth brought about by a high growth rate in the economy. One major trend of urbanization in the KSA is most of its population are concentrated in urban cities such as Riyadh, Jeddah, Makkah, Madinah, and Dammam, and these cities

account for 55% of urban and 46% of the national population, and all grow at the expense of smaller cities (Moreno et al., 2019). The high rate of urbanization in the KSA places enormous pressure on its transport system in the efforts to satisfy growing mobility demand which necessitates increased investment in passenger transport (Aljoufie, 2015). However, private citizens in the Kingdom have refused to wait for government’s intervention, and have instead relied on private car ownership, which has grown astronomically, and leading to negative impacts on the individual cities such as economic and environmental factors. There is an increasing deployment of public transport in the KSA by the Saudi Arabia Public transport company (SAPTCO) which maintains a fleet of 2,000 buses and carried more than 3 million passengers annually (Embassy of the Kingdom of Saudi Arabia, 2020). According to Embassy of the Kingdom of Saudi Arabia (2020), the KSA bus network provides affordable public intra and inter-city transport within the Kingdom.

However, there is a lack of urban transport data in the KSA, this is mainly because public transport is in its infancy within the Kingdom, and as a result there is a high dependency on private cars (Alotaibi and Potoglou, 2018). According to Alotaibi and Potoglou (2018), it is only in 2012 that the authorities in Riyadh made available the sum of \$22.5 billion to invest in the development of a new public transport system, which includes six metro lines integrated with a bus network, and is expected to be operational in 2015. As a result, the commencement of the collection of comprehensive data on urban transport is virtually non-existent in the KSA, a search of Ministry of Transportation's (2020) website reveals that the Ministry is currently developing a plan which focuses on the development of public transport within cities in the Kingdom. Information on the website focuses mainly on the efforts to set up of a Public Transport Authority based on the Council of Ministers' decree No. (373) in the year 2012. The high rate of ownership of private cars in the KSA is of a serious carbon emissions concern, Rahman et al. (2017) reported that the consumption per capita of fuel in the KSA is increasing astronomically in the past five years compared to some other neighbouring countries, and this is due to the exponential increase in the number of private motor vehicles and the rate of growth in population. Hence, the increase in domestic fuel consumption and the geometric growth in the rate of Green House Gas (GHG) emissions. They highlighted that this will continue to make efforts at planning, development, and implementation of appropriate mitigation measures more difficult.

6. Methodology

The preferred methodology for this study is the gap analysis strategy. Gap analysis is the process of analysing existing data with the objective of determining whether an entity is performing at expected normal levels that are beneficial to the entity or not (Gresham, 2017). When applied to urban transport planning data, data gap analysis is the process of examining and assessing the current level of data gathering and analysis with the aim of establishing whether the data currently being collected offer valuable benefits for the development and formulation of policy for the optimal management of urban transportation within a country, region or state. Gap analysis for urban transportation management to be adopted for this study is the systematic assessment of the types of data and the tools for data collection that are available in both countries under study.

7. Gap Analysis of Data for Urban Transport Planning in UK

The continuous growth of technology, digital systems as well as the widespread availability of data resources are significantly influencing data availability in the UK's transport network, the proliferation of digital infrastructure has led to increased requirements in the area of digital skills development and the skills that are required for understanding the basic data resources in transport data collection (Cottril, 2018). With the rich array of transport data that is available within the UK, and offered mainly by the DfT, the UK is on its way to achieving full data availability which supports a complete integration of all its transport systems. The UK collects the most basic data for the planning of its urban transport, and the data in their current form are suitable for the current purposes, however, wide gaps exist in terms of geographic coverage, particularly, with the continuous existence of the 'digital divide' with specific regard to the rural urban contradictions (Cottril, 2018). According to Cottril (2018), the requirements of new and upcoming transport technologies such as internet data and coverage will be higher than the current rates of consumption, which emphasizes an improvement to the UK's digital network. Hence, the infrastructure that is supporting data sources being collected by sources such as Highways England Traffic Information System, LENNON rail ticket information database, Electronic Ticket Machines and the Smart Tickets of the Local Bus Operators, need to be constantly monitored for their data consumption rates, and upgraded software to enable them cope with available data.

The UK's current sources of data is yet to consider the emerging transport systems and networks that are supported by Internet of Things (IoT), such as Connected and Autonomous Vehicles (CAVs), and Mobility as a Service (MaaS) both of which are within the context of

the digital platform (Cottril, 2018). Nikitas et al. (2017) argued that these are smarter transport systems that would support future social, economic and environmental sustainability, and are most critical prerequisites for creating directions for the development of sustainable urban environment in the future. These platforms are transport planners' delight because of the ease of data collection and access to such data. According to Kos-Łabędowicz and Urbanek (2017), transport information and service provision will continue to rely on digital technology in the foreseeable future, the objective of reliable and consistent network data remains the backbone of comprehensive service delivery in the transport sector.

Furthermore, the UK needs to consider conducting an in-depth study into how emerging data sources can be incorporated onto its existing platforms, such as for instance the integration of social media into transport models and urban transport planning (Cottrill et al., 2017, Golightly and Houghton, 2018). According to Cottril (2018), MaaS applications will require advanced integration of both conventional data sources (e.g. route-planning frameworks and public transport timetables) and

emerging sources, such as social media and newsfeed data for features that may affect routing, such as weather conditions and data on shared vehicles. Adopting these approaches requires considerable levels of staff training to afford them the needed skills to cope with such tasks, although the UK has underscored the importance of an adequately trained workforce for such high-level tasks, making sense of the data is also essential (Cottril, 2018). In conclusion, privacy has become an issue of serious concern, especially regarding security and privacy in relation to General Data Protection Regulation (GDPR) as underscored by the EU (2016b cited in Cottril 2018), including the EU directive on network security and information systems EU (2016a cited in Cottril 2018). These skills are critical to the emerging data scenery in the UK.

8. Gap Analysis of Data for Urban Transport Planning in the KSA

There is as yet no visible efforts to set up a framework for a systematic urban transport data collection in the KSA. At present, researchers intending to study the transportation system in the KSA have to administer the questionnaire for their study, these questionnaires are only specifically designed for the particular study being carried out and cannot be applied to any other study. In most cases, the data in the questionnaire are discarded and not stored for future use, hence this makes a case for the development of a framework for the systematic collection of transport data in the KSA. Currently, the major problem that the KSA Government is facing is how to transfer its citizens from their private motor vehicles into the public transport that the Government is planning to deploy within the Kingdom. As such, the country needs to muster concerted efforts towards the set-up of a transport infrastructure, as well as one that would systematically collect data for urban transport planning. Also, the increasing rate of ownership of private motor vehicles is a Sustainable Development Goals (UNDP, 2020) issue. Rahman et al. (2017) argued that greenhouse gas emissions from road transportation in Saudi Arabia is a challenging frontier, and suggested an integrated national effort with strong commitments from all stakeholders as a strategic imperative within the KSA to ensure the formulation and implementation of a strong national policy which has the objective of reducing greenhouse gas emissions in in the transport sector. In essence, the Government of the KSA needs to begin to organize its mobility solutions such that they will be directed towards the abatement of GHG emissions. Also, the Government of the KSA must as a necessity formulate and implement a robust policy for transport data collection data collection.

9. Data Collection for Public Transport: Comparative Studies of Saudi Arabia and the UK

The gap analysis of the two countries under review in this paper reveals that while the UK has a well-established framework for transport data collection for urban planning, a framework for transport data collection in the KSA is virtually non-existent. While the UK only needs to build upon its current framework in order to integrate the collection of data which supports smarter transport systems within its transportation network, the KSA has to virtually start from scratch collecting basic origin-destination data which includes low level demographic data.

10. Conclusion

A review of the data collection approaches for urban transport planning for the UK and the KSA were carried out. IT was discovered that the UK has a relatively robust framework for data collection for urban transport planning while the KSA has none. A data gap was discovered within the UK's framework regarding the collection of data for the management of Connected and Autonomous Vehicles (CAVs), and Mobility as a Service (MaaS), both of which are within the context of the digital platform. Data collection in the UK is seriously lacking in this area, and it is suggested that policy be drawn up within the existing policies in this regard. The KSA has virtually no framework for data collection, a search of the Ministry of Transportation website reveal only plans and intentions to deploy a network of urban transport, and this does not presently include data for urban transport planning. Greenhouse gas emissions is a cause for concern in the KSA, abatement policy is needed to reduce emissions to the barest minimum.

References

- [1] Aljoufie, M. Exploring the determinants of public transport system planning in car-dependent cities. *Procedia - Social and Behavioral Sciences*, (2015), 216 (1), 535-544.
- [2] Alotaibi, O. and Potoglou, D. Introducing public transport and relevant strategies in Riyadh City, Saudi Arabia: a stakeholders' perspective. *Urban, Planning and Transport Research*, (2018), 6 (1), 33-53.
- [3] Chen, C.; Ma, J.; Susilo, Y.; Liu, Y. and Wang, M. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C Emerging Technology*, (2016), 68 (1), 285-299
- [4] Cottril, C. Data and digital systems for UK transport: change and its implications: *Future of Mobility: Evidence*

- Review, UK Government's Foresight Future of Mobility Project London: GOVERNMENT OFFICE FOR SCIENCE. Available Online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/766718/Dataanddigital.pdf, (2018), [Accessed 18th December 2020].
- [5] Cottrill, C.; Gault, P.; Yeboah, G.; Nelson, J. D.; Anable, J. and Budd, T.. Tweeting Transit: An examination of social media strategies for transport information management during a large event. *Transportation Research Part C: Emerging Technologies*, (2017), 77 (1), 421–432.
- [6] DfT.. TAG UNIT M1.2: Data Sources and Surveys: Transport Analysis Guidance (TAG), Department for Transport London: TRANSPORT APPRAISAL AND STRATEGIC MODELLING (TASM DIVISION - DEPARTMENT FOR TRANSPORT. Available Online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/938807/tag-m1-2-data-sources-and-surveys.pdf, (2020), [Accessed 15 December 2020].
- [7] Eboli, L. and Mazzulla, G. (2012). Performance Indicators for an Objective Measure of Public Transport Service Quality. *European Transport, Institute for the Study of Transport within the European Economic Integration*, 51 (51), 1-4.
- [8] Einav, L. and Levin, J. The Data Revolution and Economic Analysis. In: LERNER, J. and STERN, S. (eds.) Paper prepared for the NBER Innovation Policy and the Economy Conference, (2013), 14. Chicago: University of Chicago Press.
- [9] Embassy of the Kingdom of Saudi Arabia.. About Saudi Arabia: Public Transportation. Washington, DC: The Embassy of The Kingdom of Saudi Arabia. Available Online: <https://www.saudiembassy.net/transportation-communication>, (2020), [Accessed 16 December 2020].
- [10] Enoch, M.; Potter, S.; Parkhurst, G. and Smith, M.. INTERMODE: Innovations in Demand Responsive Transport: Intermode Final Report, University of the West of England Bristol, UK Available Online: <https://uwe-repository.worktribe.com/output/1059973/intermode-innovations-in-demand-responsive-transport>, (2004), [Accessed 06 December 2020].
- [11] EU. Directive 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union., ed. Brussels, European Commission, (2016a).
- [12] EU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46. *Official Journal of the European Union*, (2016b), 59 (1), 1-88
- [13] Gadziński, J. Perspectives of the use of smartphones in travel behaviour studies: findings from a literature review and a pilot study. *Transportation Research Part C: Emerging Technologies*, (2018), 88 (1), 74-86
- [14] Golightly, D. and Houghton, R. J.. Social media as a tool to understand behaviour on the railways. In: KOHLI, S.; SINTHIL KUMAR, A. V.; EASTON, J. M. and ROBERTS, C. (eds.) *Innovative applications of big data in the railway industry*, . Hershey, PA: IGI Global, (2018), 224–239.
- [15] Goves, C.; North, R.; Johnston, R. and Fletcher, G.. Short Term Traffic Prediction on the UK Motorway Network Using Neural Networks. *Transportation Research Procedia*, (2016), 13 (1), 184- 195
- [16] Gresham, T. (2017). *Data Gap Analysis*. Accounting: Leaf Group Ltd. Available Online: <https://bizfluent.com/facts-7641145-data-gap-analysis.html> [Accessed 10 December 2020].
- [17] Herrero, M. D.; Calvo, M.; Martin, N. and Rodriguez, J. A. Supporting Urban Integrated Transport Systems; Transferable Tools for Authorities: D3.1 Research and Gap analysis on data collection and analysis methods, European Platform on Sustainable Urban Mobility Plans Grant agreement ID: 690650 Coventry, UK Available Online: <https://www.suits-project.eu/wp-content/uploads/2018/12/Gap-Analysis-on-data-collection-and-analysis-methodologies.pdf>, (2017), [Accessed 02 December 2020].
- [18] Huang, W.; Song, G.; Hong, H. and Xie, K.. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning. *IEEE Transactions on Intelligent Transportation Systems*, (2014), 15 (5), 2191–2201
- [19] Huang, Z. *Data Integration For Urban Transport Planning*, ed. The Netherlands, Utrecht University, (2003).
- [20] Janssens, D.. *Data Science and Simulation in Transportation Research*, ed. Hershey, PA, IGI Global, (2013).
- [21] Kos-Łabędowicz, J. and Urbanek, A.. Do Information and Communications Technologies influence transport demand? An exploratory study in the European Union. *Transportation Research Procedia*, (2017), 25 (1), 2660-2676
- [22] Kulpa, T. and Szarata, A.. Analysis of household survey sample size in trip modelling process. *Transportation Research Procedia*, (2016), 14 (1), 1753 – 1761
- [23] Litman, T. *Developing Indicators for Comprehensive and Sustainable Transport Planning: Well Measured: Developing Indicators for Comprehensive and Sustainable*, Victoria Transport Policy Institute, Victoria Canada: TRANSPORTATION RESEARCH RECORD. Available Online: https://www.researchgate.net/profile/Todd_Litman/publication/245562849_Well_Measured_Developing_Indicators_for_Comprehensive_and_Sustainable_Transport_Planning/links/5477fcb0cf26a7bf4d90a73/Well-Measured-Developing-Indicators-for-Comprehensive-and-Sustainable-Transport-Planning.pdf [Accessed 03 December 2020].
- [25] Michell, N. The importance of data to urban transport development. *Cities Today: Connecting the World's Urban Leaders* London: PFD Publications Ltd. Available Online:

- <https://cities-today.com/importance-data-urban-transport-development/>, (2017), [Accessed 12 December 2020].
- [26] Miller, E.; Lee-Gosselin, M.; Habib, K. N.; Morency, C.; Roorda, M. J. and Shalaby, A. S.. Primer on Urban Passenger Data Collection: Keeping Up With a Changing World. In: TRANSPORTATION ASSOCIATION OF CANADA (eds.) Changing Practices in Data Collection on the Movement of People. Ontario, Canada Transportation Association of Canada, (2014).
- [27] Milne, D. and Watling, D. Big data and understanding change in the context of planning transport systems. *Journal of Transport Geography*, (2019), 76 (1), 235-244
- [28] Ministry of Transportation. The Public Transport Authority. Riyadh: Ministry of Transportation. Available Online: <https://www.mot.gov.sa/en/TransportSystem/PublicTransport/Pages/default.aspx>, (2020) [Accessed 16 December 2020].
- [29] Mishalani, R. G.; Ji, Y. and McCord, M. R.. Effect of Onboard Survey Sample Size on Estimation of Transit Bus Route Passenger Origin–Destination Flow Matrix Using Automatic Passenger Counter Data. *Transportation Research Record*, (2011), 2246 (1), 64-73
- [30] Neffendorf, H.; Katalysis; Williams, I. and ME&P. Sources of Data – a Transport Planning Case Study The AGI Conference at GIS, London. London: CMP Europe Limited. Available Online: <https://www.geos.ed.ac.uk/~gisteac/proceedingsonline/AGI2002/TRACK%202/B02.1>. Jun 18, (2002), pdf [Accessed 02 December 2020].
- [31] Nikitas, A.; Kougias, I.; Alyavina, E. and Tchouamou, E. N. How Can Autonomous and Connected Vehicles, Electromobility, BRT, Hyperloop, Shared Use Mobility and Mobility-As-A-Service Shape Transport Futures for the Context of Smart Cities? ,(2017), *Urban Science* 1(4), 1-36
- [32] Pojani, D. and Stead, D. Sustainable Urban Transport in the Developing World: Beyond Megacities. *Sustainability*, (2015), 5 (1), 7784-7805
- [33] Pojani, D. and Stead, D. The Urban Transport Crisis in Emerging Economies: An Introduction. In: POJANI, D. and STEAD, D. (eds.) *The Urban Transport Crisis in Emerging Economies*. (2017), 1. Switzerland: Springer, 1-10.
- [34] Prewitt, K.; Mackie, C. D. and Habermann, H. Civic Engagement and Social Cohesion: Measuring Dimensions of Social Capital to Inform Policy. Washington, DC, The National Academies Press. Available Online: <https://www.nap.edu/read/18831/chapter/1>.(2014), [Accessed 11 December 2020].
- [35] Rahman, S. M.; Khondaker, A. N.; Hasan, M. F. and Reza, I. Greenhouse gas emissions from road transportation in Saudi Arabia - a challenging frontier. *Renewable and Sustainable Energy Reviews* 69 (1), 812-821
- [36] Stofan, D. (2018). The Development of Traffic Data Collection Methods. Medium. Available Online: <https://medium.com/goodvision/the-development-of-traffic-data-collection-cd87cc65aaab>, (2017), [Accessed 03 December 2020].
- [37] Trépanier, M.; Chapleau, R. and Morency, C.. Tools and methods for a transportation household survey. *Urban and Regional Information Systems Association*, (2008), 20 (1), 1-35
- [38] TSGB. Transport Statistics Great Britain: Table Catalogue. Department for Transport. Available Online: <http://maps.dft.gov.uk/tsgb-table-catalogue/>, (2019) [Accessed 15 December 2020].
- [39] UNDP. Sustainable Development Goals: What are Sustainable Development Goals. New York, NY: United Nations Development Programme. Available Online: <https://www.undp.org/content/undp/en/home/sustainable-development-goals.html>, (2020), [Accessed 9 November 2020].
- [40] Wachs, M.; Barker, J. B.; Bennett, J. C.; Bing, A. J.; Coogan, M. A.; Deen, T. B.; Giuliano, G.; Hansen, M.; Killough, K. L.; Manski, C. F.; McGuckin, N. A.; Morris, P. F.; Christopher A. Nash, C. A.; Oster, C. V.; Schwieterman, J. P.; Turnbull, K. F.; Menzies, T. R. and Kortum, K.. *Interregional Travel: A New Perspective for Policy Making*, Transportation Research Board Special Report 320 Washington, D.C. Available Online: <https://www.nap.edu/read/21887/chapter/1>,(2016),[Accessed 06 December 2020].
- [41] Wang, W. Bus Passenger Origin-Destination Estimation and Travel Behavior Using Automated Data Collection Systems in London, UK. Master of Science, Department of Civil and Environmental Engineering Massachusetts Institute of Technology. Massachusetts Available Online: <https://dspace.mit.edu/handle/1721.1/60814>, (2010), [Accessed 15 December 2020].
- [42] Weis, C.; Dobler, C. and Axhausen, K. W. A Stated Adaptation Approach to Surveying Activity Scheduling Decisions. In: ZMUD, J.; LEE-GOSSELIN, M.; MUNIZAGA, M. and CARRASCO, J. A.(eds.) *Transport Survey Methods*. Switzerland: Emerald Group Publishing Limited, (2013), 569-590.
- [43] Zaki, J. F. W.; Ali-Eldin, A.; Hussein, S. E.; Saraya, S. F. and Areed, F. F. Framework for Traffic Congestion Prediction. *International Journal of Scientific & Engineering Research*, (2016), 7 (5), 1205- 1210
- [44] Zannat, K. E. and Choudhury, C. F.. Emerging Big Data Sources for Public Transport Planning: A Systematic Review on Current State of Art and Future Research Directions. *Journal of Indian Institute of Science*, (2019), 99 (1), 601-619
- [45] Zhang, L.; Jia, Y.; Niu, Z. and Liao, C. Widespread Traffic Congestion Prediction for Urban Road Network Based on Synergetic Theory. *Journal of Systems Science and Information*, De Gruyter, (2016), 2 (4), 366-371
- [46] Zhang, L.; Liu, Q.; Yang, W.; Wei, N. and Dong, D. An Improved K-nearest Neighbor Model for Short-term Traffic Flow Prediction. In: ZHANG, L.; WEI, H.; LI, Z. and ZHANG, Y. (eds.) *Intelligent and Integrated Sustainable Multimodal Transportation Systems Proceedings from the 13th COTA International Conference of Transportation Professionals (CICTP2013)*. (2013), 96. Amsterdam: Elsevier, 653-662.