IMPOVERISHED DATA: EXPERIENCES AND LESSONS IN COLLECTING CAPE TOWN DATA FOR THE MILLENNIUM CITIES DATABASE

L KANE, V BALENI[1] and S COOKE[2]

Centre for Transport Studies, University of Cape Town, Rondebosch 7708
Tel: 021 6715404; Email: lisakane@iafrica.com
[1] Bigen Africa Services (Pty) Ltd, PO Box 29, The Innovation Hub, Pretoria, 0087
Tel: 012 842 8864, Email: vusi.baleni@bigenafrica.com
[2] Centre for Transport Studies, University of Cape Town, Rondebosch 7708
Tel: 021 6502584; Email: ckxsea001@myuct.ac.za

ABSTRACT

The Millennium Cities Database for Sustainable Transport is a substantial database, funded by the International Association of Public Transport, with over 200 indicators of transport, demographic, economics and land-use data from about one hundred cities. This dataset allows cities to benchmark against "best practice" or similarly positioned cities and it is a valuable aid for better understanding of the status quo, and of the likely trajectories, of cities. In 1995/6 Cape Town and Johannesburg were included in the dataset, but until now the collection of more recent South African data has not been possible. In 2012 a collaboration between three University of Cape Town (UCT) final year undergraduate civil engineering students, staff at the Centre for Transport Studies, UCT, and transport representatives from the City of Cape Town was formed in an attempt to update the Cape Town data from 1995/6 to 2005/6 and 2010. This paper describes the data collected, its quality, and the data gaps which were found. Methodological lessons on this type of data collection are described. The paper ends with some discussion on metropolitan transport data availability and quality, and the implications of this for policy and decision making at a metropolitan level.

1 INTRODUCTION

The Millennium Cities Database (MCD) for Sustainable Transport is a substantial database, funded by the International Association of Public Transport (UITP) in partnership with Murdoch University (Australia) with over 200 data indicators from about one hundred cities. It has been collated to allow mobility providers to evaluate the performance of their respective cities, and includes data for transport, demographics, economics and land-use (Baleni, 2012).

The idea for the MCD began during Newman and Kenworthy's well-cited study of automobile dependence in major Australian cities (Newman, 1989). Newman and Kenworthy found variables to describe automobile use which seemed more significant than the common, widely accepted variables at that time of income, city size and fuel price. A decision was made to broaden the perspective and to collect data for cities throughout the world to assess if these automobile use relationships were specific to Australia. The data and related analyses derived from it were compiled into a 1989 book: "Cities and Automobile Dependence: an International Sourcebook". (Cooke, 2012)

This sourcebook contained data from 1960, 1970 and 1980, although of varying reliability and with some large gaps in the information for some cities. An upgrade of this initial database became the Millennium Cities Database with a reference time period of 1995 to 1996 (Kenworthy & Laube, 2001) . The full list of the standardised indicators for

the MCD is given in Appendix A. In choosing cities to include, the authors attempted fair representation of every continent, income bracket and city size (Cameron, 2004) and the MCD includes data from 35 Western European, 6 Eastern European, 15 North American, 10 Latin American, 8 African, 3 Middle Eastern, 18 Asian and 5 Australasian cities. The database is once again being updated, this time to include 2005/6 data (Kenworthy and Laube, 2001; Vivier, 2001).

This paper is in two parts. The first part briefly outlines some work done by undergraduate students at the University of Cape Town using the MCD to examine characteristics of Cape Town with respect to other international cities of similar GDP. The argument is made that even a preliminary international analysis such as this is valuable in extending understanding of Cape Town's transport. The main part of the paper describes the outcome of attempts to update the existing 1995/6 data for Cape Town to 2005 (and 2010), and the challenges encountered. The paper ends with some discussion on metropolitan transport data availability and quality, and the implications of this for policy and decision making at a metropolitan level.

2 USES OF THE MILLENIUM CITIES DATABASE

Using 1995/6 data from the MCD students explored three relationships of interest. The first explored international relationships between urban density and modal split, and considered the significance of this relationship to the city of Cape Town. A second student focused on the degree to which the household travel patterns of Cape Town were distorted by apartheid spatial planning, through an international comparative analysis. Although the analyses were relatively contained, some interesting features emerged from the two studies, highlighting future possible avenues for research.

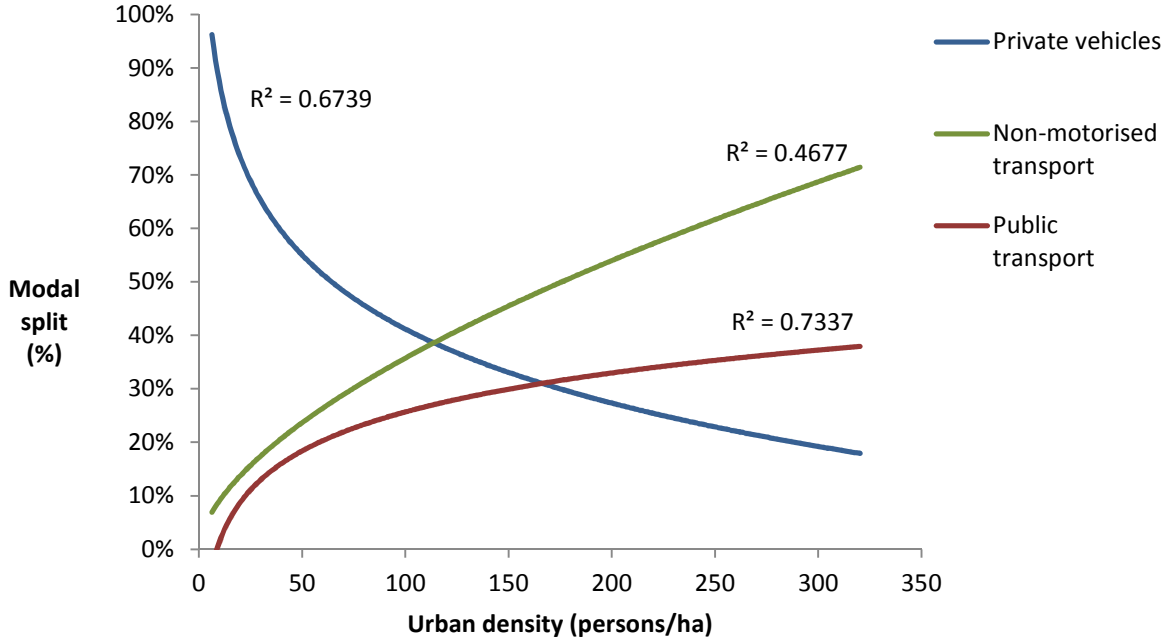

Figure 1: The different modal splits vs urban density for developed cities

An international comparative analysis of urban density and mode split revealed that there is a strong correlation between urban density and modal split in developed cities. As the urban density increases, the motorised transport mode preference shifts from private vehicles to public transport and a tipping point seems to be at 160 persons per hectare. Non-motorised transport also increases with an increasing urban density but the correlation between these two factors is not as strong. In developing cities, the correlation between urban density and modal split was weaker and instead, economic indicators appeared to be the dominant factor affecting travel behaviour. The Cape Town context displayed how urbanisation since 1995 has increased its average urban density and led to a higher proportion of trips being undertaken on public transport. This follows the trend

seen in Figure 1 and hints at a significant urban density-modal split relationship that could be evident in the city.

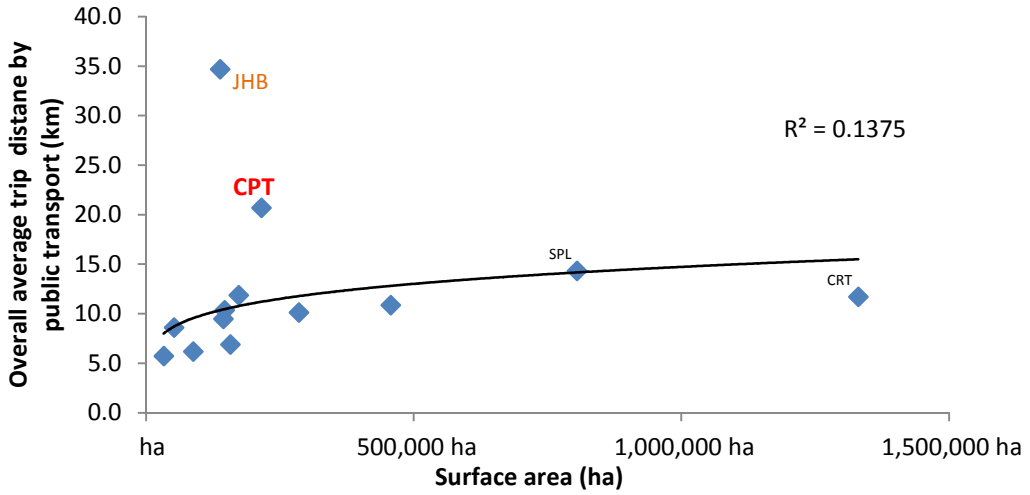

Figure 2: Overall average trip distance on public transport compared to metropolitan surface area

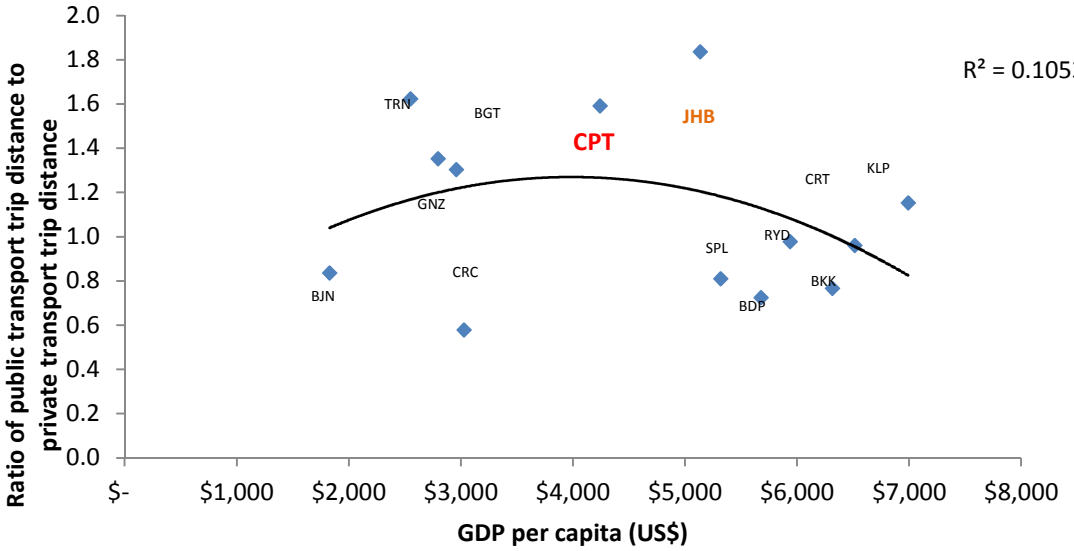

Figure 3: Trip distance ratios vs metropolitan GDP per capita

Another project explored the effects of the apartheid system, and its accompanying legislation, on the travel patterns of the residents of Cape Town. This research project set out to benchmark Cape Town's public and private transport systems against systems of other developing cities, given the unique spatial configuration of the contemporary South African city. The main patterns that emerged from the comparative analysis showed that the public transport users of the apartheid South African cities travel longer distances in relation to other cities with similar GDP per capita, and also by comparison with other cities of similar sized urban area. In addition, the research showed that the private transport travel patterns of Cape Town were similar to other comparable cities in the world.

Having demonstrated the potential value of such comparative work, the students devoted the rest of their time to updated data collection. This is the subject of the remainder of the paper.

3 UPDATING THE MILLENIUM CITIES DATABASE – METHOD

A substantial part of the work by the students was the collection of this more recent 2005/2006 data and, where possible 2010/2011 data. This proved to be a time-consuming task, revealing gaps in data compounded by the lack of a central agency that has all, or even a majority, of the data which the MCD suggests is necessary. Apart from the indicators readily available (but not centrally located) in published reports, there were efforts made to source data from the City of Cape Town Departments of Transport Planning and Policy Development; Transport Modeling and Systems Analysis; Project Planning and Conceptual Design; and Transport Network Information. In addition non City bodies of the Passenger Rail Association of South Africa (PRASA) and the Golden Arrow Bus Services (GABS), both of which are the major agencies of public transportation in the Cape Town metropolitan area, were approached. Data was also obtained from the University of Cape Town's Household Travel Data of Cape Town. (Cooke, 2012; Baleni, 2012)

Although some aggregate data measures (for example total vehicle kilometres, average network speed) required use of the city's emme/2 model, a major issue encountered during data collection was the peak hour nature of the city's transport models. Work commute trips do not account for a large amount of the trips that are made at other times of the day and so factoring from this peak model to the all day measures called for was problematic. Another challenge was the lack of data on the informal minibus taxi industry and the unwillingness of the private bus companies to share financial data.

4 RESULTS OF DATA COLLECTED: AVAILABILITY, QUALITY AND GAPS

In this section the availability, quality and gaps in the data are discussed in some detail, under the categories:

- General, economic and demographic indicators
- Infrastructure indicators
- Transport impacts
- Private vehicle indicators
- Public transport indicators
- Household travel survey

4.1 General, economic and demographic indicators

The majority of these indicators were sourced through research into governmental publications and correspondence with contacts from the City of Cape Town. The *total land area and population of the metropolitan area; urbanised area of the metropolitan area; number of jobs (at place of work) in metropolitan area* indicators were all reasonably readily available, from a city statistics fact sheet for Cape Town that is compiled every few years and from officials in the Strategic Development Information and GIS Department of the City of Cape Town (City of Cape Town, 2011a; Sinclair-smith, 2012; Small, 2012).

The *number of jobs (at place of work) in CBD,* which describes the number of employment opportunities that are situated within the Central Business District (CBD) of the city of Cape Town was a disputed indicator since no agreed definition of the Cape Town CBD is available. For this indicator, the CBD was assumed to be the four precincts of the Central City Improvement District. Data on jobs available was derived from census data, employment growth and development data by an official within the City GIS Department (Spotten, 2012).

Gross Domestic Product (GDP) of the metropolitan area is the economic value of goods and services created within the metropolitan area by all of the inhabitants, in a one year period. This information was sourced by reviewing various economic publications by the City of Cape Town. Some disagreement as to the correct value of the City GDP emerged. WESGRO, the official investment and trade promotion agency for the City of

Cape Town contradicts the City statistics fact sheet (from the City of Cape Town's GIS Department), while the Economic Development Strategy quotes a third value.

4.2 Infrastructure indicators

Total centreline length of the road network (all roads including residential) accounts for the total length of all public roads within the road network of the city and includes any residential road that falls within the responsibility of the metropolitan government. The information for this indicator could not be found in any governmental publication for the reference years, although it was available for 2008. The *total length of express road network (all expressways, freeways, tollways)* was not available within any governmental publication.

The number of parking places in the CBD (off- and on-street) is an indicator inclusive of all parking that exists in parking lots, parking structures and subterranean parking garages. This data was supplied by the Cape Town City Improvement District from surveys and analysis done in house, although this is unpublished data, and does not cover the whole of the central area, only the area managed by the Improvement District.

4.3 Transport impacts

Total transport related deaths includes all of the deaths that were attributed to a transport mode and occurred within the metropolitan area. The indicator was sourced from the Forensic Pathology Laboratory Database via the Transport Network Information section of the Transport, Roads and Stormwater department of the City of Cape Town. *Air pollution inventory from transport sources* in the city is a record of all the airborne pollutants that each transport mode has been assumed to have created during the reference year. The only data available is an aggregate pollution estimate for the year of 2008, sourced from a city official.

Private passenger transport energy use is an indicator that can measure the average efficiency of the private vehicles in a city. Contact with an official working in those areas provided an aggregate estimate of Kton equivalent of carbon dioxide. Public transport energy use data was made available in aggregate form for rail by a city official (Covary, 2012). Subsequently citywide data was sourced from work on city energy modelling by Sustainable Energy Africa (SEA, 2006).

4.4 Private vehicle indicators

The average road network speed indicator represents the average speed that motor vehicles generally travel across an average road within the city's network and would describe the congestion that exists. The data for this indicator was sourced from the emme/2 transport model for Cape Town. As mentioned, this model has been calibrated for the morning peak and work commutes and therefore is not well suited for finding daily speeds of the traffic. It was acknowledged that the model needs to be recalibrated to give more accurate results for this indicator.

Total annual passenger and vehicle kilometres of travel in private cars (and motorcycles) refers to the total number of kilometres travelled by all the private cars collectively, within the metropolitan area. This alludes to the degree of utilisation of private cars in the city and the level of automobilisation that has occurred. This data is typically calculated using city models, or through analysis of household surveys, but was not available for Cape Town.

The *number of private cars and motorcycle* indicators includes all private vehicles owned by inhabitants or companies that reside in the metropolitan area, not including any vehicle that receives income from transporting passengers. The information was sourced from the Western Cape vehicle population fact sheet, of which the city of Cape Town is a sub-division (Harris, 2012).

4.5 Public transport indicators

The *length of reserved public transport routes by each mode* is an indicator which describes routes that are legally reserved for a public transport mode only and effectively policed to prevent other modes from utilising it. For rail, all routes are assumed to be reserved so the total length of all rail lines was used for this indicator. Updates of the Integrated Transport Plans (ITPs), and contacts at PRASA, provided this data. The dedicated bus and minibus routes were found from direct measurement on Google Earth and the BRT route length was taken from the ITP 2011 update. The *length of public transport lines* by mode includes the length of all the routes used by a mode (regardless of whether the line was a dedicated route or not). It alludes to the nature and extent of the network for each public transport mode and when compared to urbanised area, will determine a degree of accessibility for the network. Again, the only mode for which the data could be sourced was rail.

The average operating speed of each public transport mode was assumed to be the average speed calculated by dividing the length of the route by the time taken to complete the route. This includes the time taken to stop at each station but not the time spent waiting at either end of the route. Rail data was available from a PRASA representative. The average speed of minibus taxis was assumed to be equal to the average congested speed of traffic in Cape Town, taken from the emme/2 model. According to the ITP 2011 update, the bus rapid transport vehicles have an average operating speed of 30 km/h.

Annual revenue vehicle kilometre of service refers to the total number of kilometres that every vehicle, in a specific mode, travelled while accepting payment from passengers. Only rail data could be sourced for this indicator. *Annual revenue seat kilometre of service* describes the total number of kilometres travelled by each seat on every vehicle for each public transport mode. This illustrates the capacity of each transport mode through kilometres and gives an indication of the coverage of the public transport network. Again, the only mode for which this information could be sourced was rail.

The *annual boardings* indicator refers to the total number of times passengers board a public transport mode in the reference year. This indicator can be used to create a modal split for public transport modes, if a boarding is assumed to represent one trip. For rail, PRASA made this data available. In the 2011 update of the ITP the number of bus boardings for GABS buses was available. No taxi or other bus operator data was available for BRT at the time of collection.

Annual passenger kilometres by each public transport mode refers to the total number of kilometres travelled by every passenger on each transport mode. PRASA data was available. Using the Emme/2 model estimates were made for bus and minibus although the reliability of this data could be questioned.

The *public transport vehicle fleet* is the total number of vehicles operated in each mode and is one description of the supply of public transport. Rail, bus and BRT data was found in the ITP 2009 and 2011 updates (City of Cape Town, 2009a; City of Cape Town, 2011b). For minibus taxis, the Current Public Transport Record summary 2009 update provided data (City of Cape Town, 2009b; Western Cape Department of Transport and Public Works, 2011).

Annual total public transport farebox revenue is all of the money taken from the passengers of each public transport mode in the reference year for services rendered. This indicator was contentious as it requires companies to reveal financial details that competitors could use to the detriment of the company. Only PRASA was willing to release data pertaining to farebox revenue. (It is a parastatal company and therefore is forced to reveal all details to the public under law.) The *operating expenses* summarise any financial requirement of a mode during the reference year. This includes maintenance of vehicles, maintenance of routes, salaries of staff and any other day to day expenses. None of the companies were willing to supply data on their operating expenses.

4.6 Household travel survey

A statistically significant House Travel Survey (HTS) has not been performed in Cape Town since the National Household Travel Survey (NHTS) was conducted in 2003, although one is presently underway. In the absence of this, a privately funded HTS by the African Centre of Excellence for Studies in Public and Non-motorised Transport (ACET) from the University of Cape Town was used (ACET, 2011). The sample size of the UCT HTS was 16 231 (0.52% of the study area population) which is too small to be truly representative. Nevertheless, as it was the most recently available data, it was analysed and the results presented.

The *number of daily trips* indicator represents the number of trips made by the entire metropolitan population on each mode during an average day and can be used to determine an estimate for the modal split. The number of trips made by the sample group on an average day, for each mode, was scaled up to approximate a number of trips for the City of Cape Town. The *average length of a trip* for each mode is taken as the estimated average length of all the trips utilising that mode on an average day, and was found from the HTS to be 10km. The *average time for a trip* is calculated as the average amount of time taken for a trip to be completed on each transport mode, and was found from the HTS 20 minutes for car and 38 minutes for public transport.

Although the reliability of this HTS data to represent Cape Town as a whole is open to question, it still gives some insight into the travel behaviour in Cape Town in 2010. However, it also highlights the lack of household travel information and the great need for regular and detailed Household Travel Surveys.

5 CONCLUDING REMARKS

Given the experience of the three students in collecting data over a brief two month period, they were asked to judge whether they believed their data collected to be reasonably reliable/complete; somewhat reliable/complete; or probably unreliable/incomplete. The summary table of data found during the process (given in Appendix A) was thus labelled green (for the most reliable data), amber and red (least reliable). Only eight indicators were deemed reliable/complete; eighteen were judged somewhat reliable/partially complete and the remaining sixteen as unreliable, or incomplete. The data found is available from the authors, on request.

The authors would like to acknowledge the generous gifts of time by various staff members, especially at the City of Cape Town and PRASA, who provided data directly. However, this style of data collection, relying on personal contacts and goodwill, raises questions about risk, monitoring, and accountability. One risk is that as individuals leave the City their knowledge and data memory is lost, unless a more systematic approach can be taken on data collection, recording and publishing. Anecdotal evidence from the research process suggests the availability or location of transport data not *directly* related to their work is not well known, and so data collation is resource inefficient and onerous. This also raises questions about cross-disciplinary understanding about the transport sector. Also concerning are inconsistencies within data sets. The City of Cape Town, for example, report approximately 650 transport related deaths in the period, while the National Injury Mortality Surveillance System attributed 1021 deaths to transport in 2004.

This data collection experience, and other data collection experiences of the main author also suggests that, for the most part, up-to-date aggregate transport-related data for the City of Cape Town is not available in a form which enables meaningful international comparison within the Millennium Cities Database. This is problematic. It means that bench-marking of Cape Town (and likely other metropolitan areas of South Africa) is not possible in any extensive way, and so learning which could come from genuinely comparable international cities is not possible. This is a loss for research and learning.

Lack of suitable data also compromises moves in the City towards integrated transport services, which will require a more holistic perspective on matters of transport policy and practice. A protocol for comprehensive data collection, which clearly defines scope, measuring criteria and methods will be essential to ensure data which can be compared year-on-year and so identify genuine trends and shifts due to newer public transport investments. In this regard, a Household Travel Survey for Cape Town which is statistically significant is long overdue (and is now underway). Also overdue is some agreement (or requirement) for the submission of data by public and paratransit operators to a central body.

The process also highlighted a lack of all day data. The spatially comprehensive data (such as vehicle kilometres) which does exist tends to be for the morning peak only. It has been argued elsewhere that transport planning which bases its insights on modal splits for a morning peak period are systemically biased towards the employed; professionals; males; and such data is likely to seriously overestimate the importance of the motorized commuter trip (Behrens, 2001). Until reliable and regularly updated all-day data for the whole transport system is available, such perspective bias likely continues to be propagated in policy and decision-making.

Just more pressing, and more concerning, are the implications that an absence of reliable data, in combination with large expenditures on transport, have for matters of transparency and accountability. Even the data which is available is not collated in a format which makes it easily accessible. For the lay-person such data would be very difficult – perhaps even impossible - to access. The claims made for transport expenditure are not often called to account in the public realm, but there is no reason why they should not be. Evidence based policy claims would be a key part of this.

Important in this regard are the implications of good data for a sustainable transport agenda. The City has for some years now had a vision for sustainable transport, but the City data is truly impoverished in its ability to reflect on matters of sustainability. There is no regular or reliable measure of vehicle or passenger kilometres, and so no good measure of energy use or emissions. This means that any attempt to manage car use, or to shift mode and its attendant emissions on a citywide basis simply cannot be monitored or evaluated. Recent work by Venter and Mohammed (2013) using Household Survey Data from Nelson Mandela Municipality provides a template and shows how valuable such analysis can be for understanding the energy and emissions profile of urban areas.

It is important to mention that at the time of writing some of the Cape Town indicators were being updated, for the Integrated Transport Plan process, but this will not significantly impact on the Conclusions here. Finally, the authors understand that the situation in Cape Town is not unusual, rather, it reflects a typical pattern of impoverished transport data surveys, data-handling systems, data publicity, and data monitoring across South Africa, which is seriously out of step with international best practice.

6 BIBLIOGRAPHY

African Centre of Excellence for Studies in Public and Non-motorised Transport (ACET) 2011. *Cape Town Study*. Cape Town: ACET

Baleni, V. 2012. *Effects of apartheid legislation on transport in Cape Town*. Unpublished undergraduate thesis. Department of Civil Engineering, University of Cape Town.

Behrens, R. 2003. Looking beyond commuter travel in Cape Town: Methodological lessons from the application of an activity-based travel survey, in Stopher P and Jones P (eds), *Transport survey quality and innovation*. Amsterdam: Pergamon.

Cameron, I. 2004. *Understanding, modelling and predicting transport mobility in urban environments*. Perth: Murdoch University.

City of Cape Town 2009a. *Integrated Transport Plan Update.* Cape Town: City of Cape Town.

City of Cape Town 2009b. *Current Public Transport Record Summary.* Cape Town: City of Cape Town.
City of Cape Town 2011a. *City Statistics.* Cape Town: City of Cape Town.
City of Cape Town 2011b. *Integrated Transport Plan Update.* Cape Town: City of Cape Town.
Cooke, S. 2012. *An analysis of the urban density-modal split relationship and its significance in Cape Town*. Unpublished undergraduate thesis. Department of Civil Engineering, University of Cape Town.
Kenworthy, J. & Laube, F. 2001. *The Millennium Cities Database for Sustainable Transport.* Brussels: International Union (Association) of Public Transport, (UITP).
Medical Research Council/ UNISA. 2013. *Crime, Violence and Injury Lead Programme. A profile of fatal injuries in South Africa*. 6th Annual Report of the National Injury Mortality Surveillance System 2004. Section 4. Cape Town Fatal Injury Profile. Available online at http://www.sahealthinfo.org/violence/capetown2004.pdf. Accessed 10 May 2013.
Newman, P.W.G. 1989. *Cities and automobile dependence: a sourcebook.* Aldershot, UK: Gower.
Newman, P. 1999. *Sustainability and cities: overcoming automobile dependence.* Washington DC: Island Press.
Sustainable Energy Africa (SEA) 2006. *The Energy Book*. Cape Town: Sustainable Energy Africa.
Venter, C.J. & Mohammed, S.O. 2013, "Estimating car ownership and transport energy consumption: a disaggregate study in Nelson Mandela Bay*", Journal of the South African Institution of Civil Engineering*, vol. 55, no. 1, pp. 2-10
Vivier, J. 2001. *Millennium Cities Database for Sustainable Transport: Analyses and recommendations.* Brussels: International Union of Public Transport (UITP).
Western Cape Department of Transport and Public Works 2011. *Provincial Land Transport Framework.* Cape Town: City of Cape Town.

APPENDIX A: MILLENIUM CITIES DATABASE INDICATORS

The table below summarised the data within the MCD, and the analysis of the data collected for the process. (Data collected available on request from the authors). Green represents data that is complete, considered accurate and reliable; yellow highlights data that is questionable and may need further investigation, and red refers to data that is missing or that is perceived to be inaccurate and should not be utilised for analysis without further study.

Table 1: Showing the raw indicators utilized by the MCD and the perceived reliability of the data collected

1	Total land area of the metropolitan area
2	Urbanised area of the metropolitan area
3	Total population of the metropolitan area
4	Number of jobs (at place of work) in metropolitan area
5	Number of jobs (at place of work) in CBD
6	Gross domestic product of the metropolitan area
7	Number of private cars and station wagons, RVs, company cars (not taxis)
8	Total annual vehicle kilometres of travel in private cars
9	Total annual passenger kilometres in private cars
10	Average road network speed (7day/24hour)
11	Total centreline length of the road network (all roads including residential)
12	Total length of express road network (ALL expressways, freeways, tollways)
13	Number of parking places in CBD (off-street)
14	Number of parking places in CBD (on-street)
15	Length of reserved (policed) public transport routes by each mode

16	Average operating speed of each public transport mode
17	Annual revenue vehicle kilometer of service by each public transport mode
18	Annual revenue seat kilometer of service by each public transport mode
19	Annual boardings by each public transport mode
20	Annual passenger kilometres by each public transport mode
21	Private passenger transport energy use
22	Public transport energy use
23	Total transport related deaths
24	Number of two-wheeled motor vehicles (motorcycles)
25	Vehicle kilometres of travel on two-wheeled motor vehicles (motorcycles)
26	Passenger kilometres on two-wheeled vehicles (motorcycles)
27	Public transport vehicle fleet by mode
28	Length of public transport lines by mode
29	Annual total public transport farebox revenue
30	Annual operating expenses of public transport
31	Air pollution inventory from transport sources in the city
32	Number of daily walking trips
33	Number of daily mechanized, non-motorised trips
34	Number of daily motorised trips on public modes
35	Number of daily motorised trips on private modes
36	Average length of a trip (all modes)
37	Average length of a trip (mechanized modes)
38	Average length of a car trip
39	Average length of a home-work commute (all modes)
40	Average length of a home-work commute (mechanized modes)
41	Average time for a trip by car
42	Average time for a trip by public transport