ARTICLE IN PRESS

Journal of Transport Geography xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Inferring commercial vehicle activities in Gauteng, South Africa

J.W. Joubert [a,b,c,*], K.W. Axhausen [c,1]

[a] Industrial and Systems Engineering, University of Pretoria, Pretoria 0002, South Africa
[b] CSIR Built Environment, P.O. Box 395, Pretoria 0001, South Africa
[c] Institute for Transport Planning and Systems (IVT), ETH Zurich, 8093 Zurich, Switzerland

ARTICLE INFO

Keywords:
GPS data processing
Commercial
Transport
Activity density
Freight

ABSTRACT

To address the underreporting of freight from a transport geography point of view, we present a novel analysis of the time and spatial characteristics of disaggregated commercial vehicle activities. The activities were extracted from raw global positioning system (GPS) data collected in South Africa over a six-month period for more than 30,000 commercial vehicles. The analyses of the activity chains provide useful characteristics such as activity and chain durations, number of activities per chain, and the spatial extent of the activity chains. Key results indicate that about 60% of activity chains have between 5 and 15 activities per chain while 25% of the chains have 4 or less; 89% of the chains have a duration of 24 hours or less; and approximately 75% of all activities start between 08:00 and 17:00. The paper's contribution is twofold: it firstly demonstrates a methodology to extract and evaluate vehicle activities and activity chains from raw GPS data. Novel results and characteristics about transport geographies in Gauteng, the economic centre of South Africa, are presented. We also report on the sensitivity of the analyses to certain parameters. Secondly, we introduce new metrics to evaluate a geographical area's economic productivity based on commercial activity.

1. Introduction

People and organisations pursue their goals through participating in a variety of activities. Since the chosen activities are usually separated by both space and time, some form of transport is implied, which in turn typically requires network infrastructure such as roads, and vehicular infrastructure such as cars, busses and trains. In areas where there is limited network infrastructure for the amount of vehicles, congestion occurs. The positive impact public infrastructure (notably transport network infrastructure) has on economic growth is known from Aschauer (1989, 1993), Munnell (1992), Shirley and Winston (2004) and more specifically in South Africa from Fedderke and Bogetic (2009). The studies all show that well-constructed road infrastructure, which makes up the largest category of 'core' public infrastructure, allows for transport cost to decrease. In turn, lower transport cost allows companies to produce products at lower total costs. Fedderke and Bogetic (2009), specifically, use panel data to show that road infrastructural measures have a statistical significant, positive, often economical very strong impact on labor productivity.

Freight and commercial services are key users of network infrastructure, and the efficient movement of freight is critical to economic vitality. Yet we do not have a good, disaggregate understanding of the behaviour of commercial road traffic. The development of freight models as public policy instruments have lagged behind the rapid evolution of the actual freight systems.

Hesse and Rodrigue (2004) argue that freight transport is not merely a derived demand. Logistics are no longer just means to overcome space, but became a critical and time-sensitive component. The extension of supply chain structures, outsourced and subcontracted across vast geographic areas have resulted in smaller consignments delivered more frequently. As supply chains increase in complexity, activities of production, distribution and consumption are increasingly difficult to separate.

Freight transport models are frequently based on the classical four step modelling approach (attraction/generation; distribution; modal split; and assignment) originally developed for passenger transport. The four steps are only handled at an aggregate level (zones) and the detailed movement of the freight carriers are not considered (Fernández et al., 2003). Figliozzi (2007) notes that four step models completely ignore the urban tours where vehicles make multiple stops. Another research stream of freight modelling is concerned with econometric models where the interest is in predicting price and demand elasticities (Beuthe et al., 2001; Kremers et al., 2002; de Jong et al., 2004; Rich et al., 2009). Both these research streams deal with aggregated flow of commodities, often

* Corresponding author. Address: Industrial and Systems Engineering, University of Pretoria, Pretoria 0002, South Africa. Tel.: +27 12 420 2843; fax: +27 12 362 5103.
E-mail addresses: johan.joubert@up.ac.za (J.W. Joubert), axhausen@ivt.baug.ethz.ch (K.W. Axhausen).
[1] Tel.: +41 44 633 3943; fax: +41 44 633 1057.

0966-6923/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtrangeo.2009.11.005

expressed in tonne–kilometers or some volumetric measure. Although this is valuable from a macro-economic perspective, it is not the commodities that contribute to increased congestion, but the commercial vehicles transporting the commodities. With smaller and more frequent consignments, tonne–kilometers is not an accurate measure to evaluate the impact that commercial vehicles have on congestion.

Ambrosini and Routhier (2004) review urban freight modelling efforts across nine industrialized countries. The methods and tools included theoretical models, interviews, surveys, and experiments. In response to the review, Figliozzi (2007) notes that studies concerning vehicle activity chains are absent. Although still based on zonal aggregates, Friedrich et al. (2007) report on recent models where activity chain generation is accounted for. As with other top–down multiclass assignment models based on the four-step modelling approach, the aggregate models lack the ability to capture individual stakeholders' behaviour.

The decisions that the commuting population make about the timing, route and mode choice to participate in activities, on the other hand, have been studied extensively. The advancement of technology, and the accessibility of Global Positioning Systems (GPS) have positively impacted on the disaggregate study of the movement of people and their associated activities (Yalamanchili et al., 1999; Draijer et al., 2000; Wolf et al., 2001, 2004; Schüssler and Axhausen, 2009).

Hensher and Figliozzi (2007) acknowledge that freight models are important in supporting public policy formulation. They note, in a special issue on behavioural insights into freight transportation, that the basis of how choices are made across the distribution chain must be rethought. In the special issue, Hunt and Stefan (2007) present the first, to our knowledge, tour-based microsimulation model of urban commercial movements using data from an extensive survey of 37,000 activity chains in Calgary, Canada. The model developed by de Jong and Ben-Akiva (2007), although not estimated on actual data, considers the frequency of chains; the number of activities per chain; the use and location of distributions centers; and the mode used for each leg of the activity chain.

Traffic counts produced by the *South African National Roads Agency Limited* (SANRAL) make distinctions between light and heavy vehicles on the majority of national and main roads. One could hence argue that we have a fair idea of where heavy vehicles travel on the road network. We acknowledge that a large portion of commercial vehicles may be hidden under light vehicle counts. In this paper, our interest is more on where commercial vehicles spend their time performing loading, offloading, or service activities, and not necessarily the routes they travelled. For the purpose of this paper travelling between activities is not considered an *activity*.

The first contribution of this paper is methodological. We propose a method to extract commercial vehicle activities and activity chains from raw GPS data. The methodology makes a distinction between vehicles performing the majority of their activities within a study area, termed *within-traffic*; and those that perform activities across larger geographic areas, termed *through-traffic*. We demonstrate the methodology using GPS data collected for 31,053 commercial vehicles over a six-month period using on-board fleet management tracking devices. We also explored the sensitivity of the results to certain parameters.

The second contribution is more theoretical. New metrics, based on the extracted activities, are proposed to measure the economic productivity of a study area. The choice of the South African province of Gauteng is useful in this sense since it has some unique characteristics. Although occupying less than 2% of the country's land surface, it accounts for more than 30% of the country's Gross Domestic Product (GDP). The province is also a porously bounded economic space that acts as a gateway into the rest of Southern Africa with omnidirectional through-traffic often originating from the two main ports in South Africa.

Approximately 60% of activity chains have between 5 and 15 activities per chain while 25% of the chains have 4 or less. Of the observed chains, 89% have a duration of 24 h or less. The difference in activity and chain characteristics between *within*- and *through*-traffic is very small, contrary to our expectations. The exception is for the newly introduced productivity measure which suggests that within-traffic is more productive (fewer kilometers travelled per activity) than through-traffic. The time-of-day analysis of activities confirms the survey results of Hunt and Stefan (2007) that commercial activities take place throughout the business day and does not center around the morning and afternoon peaks. Approximately 75% of all activities start between 08:00 and 17:00.

The disaggregate description of commercial activities is a significant contribution to understanding the spatial impact of commercial vehicles. This paper is novel as it is, to our knowledge, the first detailed and disaggregate report on commercial vehicle activities, and activity chains, based on actual observed vehicle movement. Knowing the characteristics of commercial activity chains allows us to generate and simulate a synthesized population of commercial vehicles in the near future. Combining commercial and person movement in a single model provides improved decision support for policy and infrastructure evaluation.

Our paper starts with a brief history of South African transport policy affecting commercial traffic in Section 2. The policy context is useful to interpret later results, especially the overlap of commercial activity densities with that of population densities. The method used to extract commercial activities from raw GPS data is described in Section 3. The analyses of activities and chains are discussed in Section 4 where we also report on the time-dependent progression of activities, and propose two metrics to measure the productive use of transport infrastructure. The paper is concluded in Section 5 with a suggested research agenda for using disaggregated freight movement data.

2. South African transport policy

In this section we present a brief history of the South African transport policy related to freight. The background provides not only context to the regulatory framework that shaped the movement of freight, but also motivates the need, expressed from South African governmental perspective, to understand the movement of freight at a more disaggregate level.

Stander and Pienaar (2002) review the early South African permit system restricting carriage of goods by road since 1930, favouring government owned rail. Gradual deregulation of freight towards free competition started with the *Road Transport Act*, No. 74 of 1977, and ended with the *Transport Deregulation Act*, No. 80 of 1988, and the *Road Traffic Act*, No. 29 of 1989. The substantial increase in road freight haulage since the economic deregulation was only governed by technical and safety regulation of operators and vehicles. In essence, government lost control of *where* freight moved; and to a large degree *what* was moved.

Having approximately 270,000 km of roads and 20,000 km of rail network may seem extensive for a developing country. The South African legacy freight system, as a result of the Apartheid era sanctions, was set up and configured to support the movement of the products of inward industrialization and provide cheap transport to a very limited and racially select number of economic participants: most of which were government-owned and regulated (Fourie, 2001). The system was successful in creating tailored solutions to the export of bulk commodities, mostly via the rail parastatals. The network is characterized by uneven flows of goods, and have dramatic peaks in specific areas.

Two seemingly parallel economic systems resulted (Development Policy Research Unit, 2008). The first, popularized by the then President Thabo Mbeki in 2003 as the *first economy*, is a modern economy similar to that found in developed countries. The first economy is formal and is well-documented with receipts, records, a credit system, and legally enforceable rights and remedies. Contrary, the *second economy* lacks all these things: it is informal, regulated most often only by community norms, and based on small taxless cash transactions. As described by Edwards (2007), the two economies are embedded in two parallel sets of infrastructures, and are as a result of the racially segregated political past, heavily influenced by race and ethnicity, as well as class.

After the first multi-racial democratic elections in 1994, the first strategic framework to integrate all parts of transport into a common vision and plan for action was the Department of Transport (1999)s *Moving South Africa*. The document emphasized that the Apartheid government essentially created a transport system around its selective national goals aimed at creating employment for a privileged class, and a network engineered to support the spatial dispensation of separation and dispersion. *Moving South Africa*, in response, was a 20-year plan underlining the new government's intent to be customer-focused and play a growth facilitation role in providing and maintaining high volume routes and nodes in the transport network as a backbone, and develop supporting networks to empowered dispersed and rural communities.

Although the subsequent *National Land Transport Transition Act* (NLTTA), No. 22 of 2000 emphasised public transport, it did require the minister to prepare an annual National Land Transport Strategic Framework, within which freight transport should be addressed. The launch of the first freight-specific policy document—a legal requirement under the NLTTA—the *National Freight Logistics Strategy* acknowledges that the growth of freight traffic has surpassed most of the 20-year growth forecasts made by *Moving South Africa*, 14 years earlier than expected (Department of Transport, 2005). The freight strategy was a response to the inability of the institutional and regulatory structure to promote the needed improved efficiency. The inefficiencies not only impact South African economic development, but also neighboring countries like Lesotho and Zimbabwe that are reliant on the South African freight logistics system, and that face greater developmental challenges than South Africa (Naudé, 2009).

The first *State of Logistics Survey* for South Africa was published by Van Dyk et al. (2005). The report highlights that although first world economies have achieved a significant reduction in the cost of transportation (and inventory) as a percentage of GDP, South Africa's core structural problems put the country's logistics in a far worse than expected state. The study was the first attempt to measure the intrinsic logistics costs, and emphasized the importance of measuring logistics since it is a valuable tool as lead indicator for economic growth.

With the publication of the most recent strategic plan, the Department of Transport (2008) acknowledges the need to understand the movement of freight at a more detailed level. The structural alignment of the different spheres of government (national, provincial and local) is still prominent in policy documents, but more detailed measures are sought, and the majority of the budget earmarked for freight interventions are dedicated to the development of an integrated *National Freight Information System* that allows industry to tap into and benefit from a cargo information and tracking system. It is anticipated that such a system will allow transport planners to better understand, and incorporate the impact and interaction of freight movement with the people movement on the road infrastructure, improving decision support regarding transport infrastructure investment.

In the absence of any economic regulation, companies involved with road-based commercial transport are voluntarily organising themselves. One of the most representative organisations is the *Road Freight Association* (established 1975, http://www.rfa.co.za) that acts as a voice of the road freight industry. The Association's members include small, medium, and most of the largest trucking companies; both *for reward* carriers and private carriers; represent a variety of economic sectors; and include many allied trades providing goods and services to the trucking industry. As a facilitating body, the Association attempts to influence rates, upkeep of the road infrastructure, road safety, freight security, the fuel price, and law enforcement, to name but a few, to promote and sustain healthy competition in the industry.

The competition expectedly results in companies not always eager to share information about detailed business activities.

3. Extracting activities from GPS data

One service provider to the transport industry is *DigiCore Fleet Management* who offers vehicle tracking and fleet management services. *DigiCore* made a dataset available, acquired through its *Ctrack* system, that contains the detailed GPS log for commercial vehicles for the six months 1 January 2008–30 June 2008. A total of 31,053 vehicle files were identified, representing approximately 1.5% of the national heavy and light delivery vehicle population. The vehicles represented in the data set are from customers subscribing to *DigiCore*'s vehicle tracking and fleet management services, and we acknowledge a possible selection bias. Still, we rely on the data for the purpose of our study: using GPS data to extract and analyse commercial vehicle activities in the absence of any other data sources.

The single-file data set has six fields: (1) a unique vehicle identifier not traceable to the original customer; (2) a seconds-based Unix time stamp; (3) a longitude and (4) latitude value, both in decimal degrees according to the World Geodetic System (WGS) 84 reference coordinate system used by global positioning systems; (5) a vehicle status identifier used by *DigiCore* for fleet management purposes; and (6) a speed value. With a total file size in excess of 30 GB, the first step was to split the data into separate files, one for each vehicle, and ensure that the records were sorted chronologically.

The fleet management device installed on each vehicle monitors various engine and electronic triggers including temperature, water levels, harsh braking, ignition activity, tampering with the GPS unit, triggering of the panic alarm, and opening and closing of vehicle doors, to name but a few. Whenever an exception trigger is received, a GPS update is sent to a central server and logged. When the vehicle is idling or in motion under *normal* conditions, i.e. no exception triggers are recorded, the GPS will send log records every 5 minutes. Such infrequent records does not allow for a clustering of GPS point to be indicative of activities, such as the case in Schüssler and Axhausen (2009), for example, or accurate map-matching to establish vehicle routes.

In this study we opted for ignition-related triggers to identify activity start and stop times. An activity start was identified as the point when an *ignition off* trigger is received, and an activity stops when the ignition is turned on again.

To eliminate any false starts or false stops, such as a failed engine start attempt requiring the driver to switch the ignition off and back on before starting again, one might be tempted to provide a threshold for minimum activity duration. Fig. 1 shows the activity durations of all activities before chains were identified. The activity duration follows a Weibull distribution with shape parameter $k = 0.4860 \pm 0.0019$ and scale parameter $\lambda = 37.91 \pm 0.91$. To estimate the distribution parameters with a 95% confidence interval, we fitted distributions to 50 samples containing the activity durations of 1000 randomly selected vehicles each.

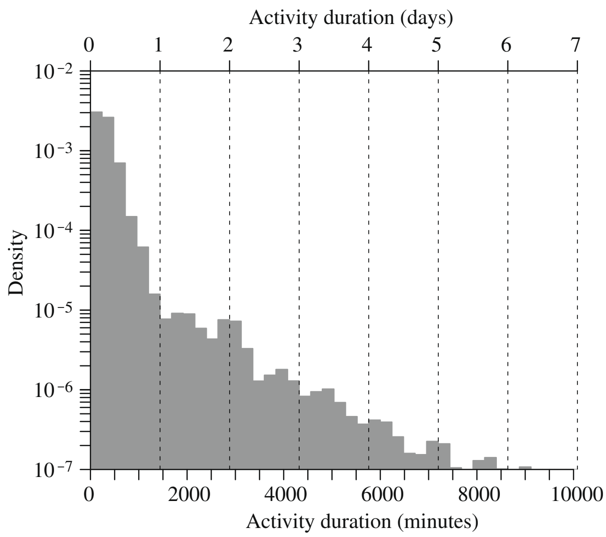

Fig. 1. Histogram of all activity durations.

The activity duration distribution has an extreme right tail that is emphasized by the log-scale on the density axis of Fig. 1. Of the observed activities, 90% have a duration of 179 min (3 h) or less, 95% 629 min ($10\frac{1}{2}$ h) or less, and 99% 2063 min ($34\frac{1}{2}$ h) or less. There were no clear breaks in the histogram suggesting a logical and definite distinction between *minor* activities, those making up the links of the activity chains; and *major* activities that signals the start and the end of an activity chain. One may think of a minor activity as a vehicle stopping for a period of time to either collect or deliver, or perform a specific service activity; while a major activity may represent an end-of-day depot stop, although *day* does not necessarily reflect a 24-h period.

Since some duration threshold was needed to distinguish between minor and major activities, we considered five different threshold values: the 90.0th (179 min), 92.5th (339 min), 95.0th (629 min), 97.5th (932 min), and the 99.0th (2063 min) percentile respectively.

Before starting the chain extraction process, all activities were clustered using a density-based clustering approach similar to the one implemented by Zhou et al. (2004). Benefits of using a density-based approach is that all activities need not be associated with a cluster, and clusters may be of arbitrary shape such as a U- or H-shaped layout of a distribution center. The two parameters involved in the density-based clustering are the minimum number of activities that is considered a *cluster*, and the search radius within which activities area considered. These two parameters were set to 10 activities and 20 m, respectively, and are based on a number of experiments using visual inspection of aerial images and the activities and activity clusters superimposed. Since activities not belonging to a cluster are not removed, the chain extraction is not very sensitive to the choice of the two parameters.

Whenever a major activity was identified, a new activity chain was created, starting with the major activity. Subsequent minor activities were added sequentially until the next major activity was identified. Any two consecutive activities from the same cluster were merged, assuming the short activity represents a relocation of the vehicle at the same venue. All chains not starting and ending with a major activity was removed to ensure only complete chains were evaluated. Chains containing only two major activities were considered mere relocations, and since they made up a very small proportion of the activity chains, they were also removed.

The total number of activities exceeded 10.5 million, and although 31,053 vehicles can be considered a sizable sample, we could only estimate the density of commercial vehicle activities, and this was achieved through kernel density estimation. The density maps of the minor and major activities are very similar. We illustrate major activity density in Fig. 2. Although activities were recorded as high up as the southern part of the Democratic Republic of Congo, the activity density outside South Africa was so low that it was barely distinguishable on the map. Even when only considering South Africa, the dispersed nature of activity clusters are noticable, emphasizing the *proximity gap* that Africa suffers (Naudé, 2009). Distances between economically active areas are very large: the most prominent area is throughout Gauteng, followed by the port cities Durban and Cape Town. Gauteng is the smallest of the nine provinces, accounting for less than 2% of the country's surface, yet is contributing approximately 33% of the GDP (Statistics South Africa, 2009).

Another interesting observation was the near-continuous activity band along the major route connecting Gauteng with Cape Town via Bloemfontein, as well as along the Gauteng–Durban route (Fig. 2). Along the majority of these routes there is very limited development, if any at all, especially between Bloemfontein

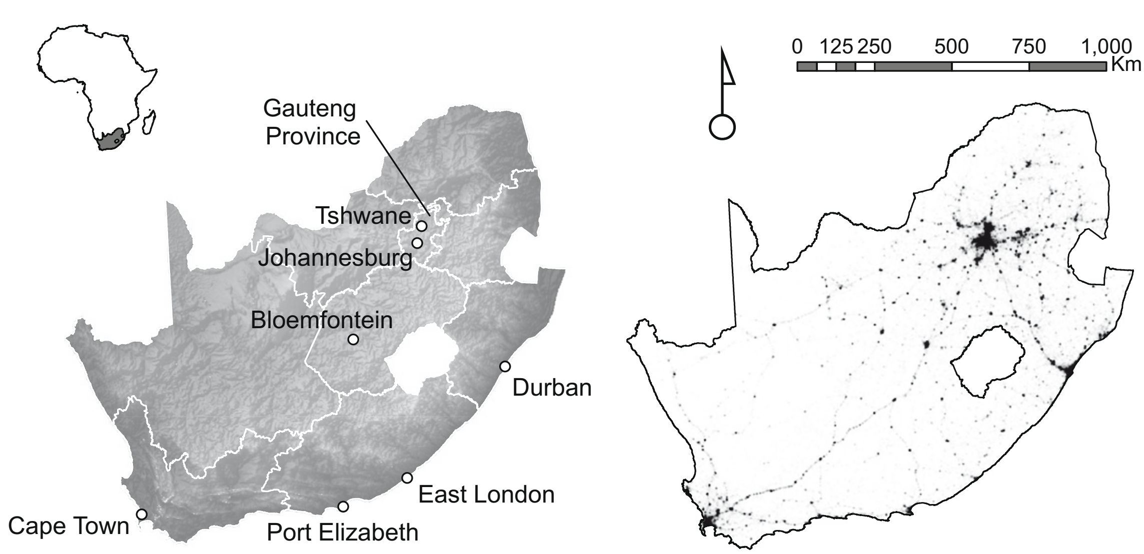

Fig. 2. The extent and density of activities throughout South Africa. The province of interest, Gauteng, is indicated.

Please cite this article in press as: Joubert, J.W., Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. J. Transp. Geogr. (2009), doi:10.1016/j.jtrangeo.2009.11.005

and Cape Town. Instead, the vehicles use informal stops to overnight. Fig. 3 depicts one such a location, an infrastructure-less stop close to a few trees.

In the remainder of the paper we focus on the province of Gauteng due not only to its economic, but also its geographical significance. Many neighbouring countries, all north of South Africa, rely on South Africa's port infrastructure and production capacity. Gauteng is uniquely positioned to act as a strategic hub, being a major producer and processor of goods and services. It is also the gateway connecting the ports with the country's northern neighbours Namibia, Botswana, Zimbabwe and Mozambique. This results in a large amount of omnidirectional through-traffic.

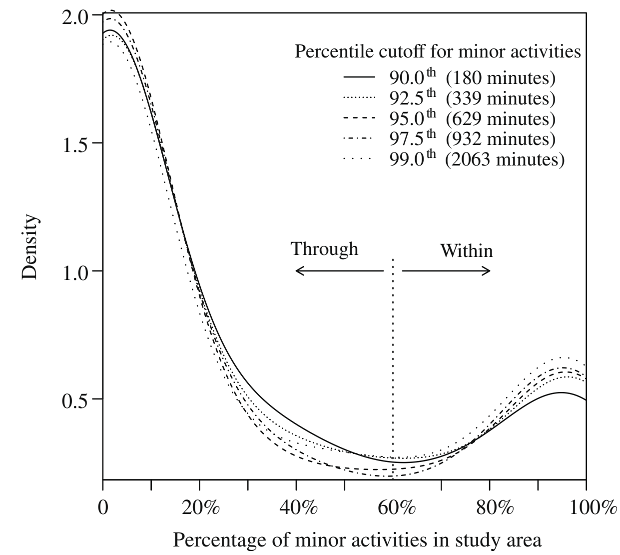

Fig. 4. Density estimates to distinguish between within- and through-traffic in the study area. The density distributions for different major activity duration thresholds are very similar, all suggesting 60% of activities within the study area as a distinguishable level.

4. Activity and chain characteristics

In our first analysis we aimed to distinguish between vehicles considered as producing through-traffic, and those that are mostly confined to the study area, termed *within*-traffic vehicles. Fig. 4 compares, for five different minor activity duration thresholds, the percentage of minor activities that vehicles performed in the study area. The bimodality of the density distribution for all thresholds suggest 60% of activities within the study area to be a clear distinction between within- and through-traffic vehicles. The underlying assumption in this distinction is that vehicles are purposefully acquired, i.e. all activity chains of a long-haul truck will be very similar; and those of smaller distribution vehicles will also be similar throughout the observed period. The 60% of vehicles considered as through-traffic in Gauteng is considerably higher than the 6% reported by Hunt and Stefan (2007) in the Calgary study in Canada. This difference could be attributable to the gateway into Southern Africa role that Gauteng plays. Density-based clustering is computationally expensive, and the remainder of the analysis was done by sampling, for each minor activity threshold, and analysing the chains from 500 vehicles randomly selected (without replacement) from the original 31,053 vehicles. On average, 97% of the vehicles did have activities within the study area of Gauteng.

Next we analysed how the number of activities per chain change with different minor activity threshold values. The comparison in Fig. 5 distinguishes between within- (Fig. 5a) and through-traffic (Fig. 5b). As expected, the higher the threshold value for both within- and through-vehicles, the more activities are included in each chain. What we did not expect was that the distributions of the two vehicle types would be as similar. It seemed plausible that within-vehicles will perform more activities during a *business day* than their through-traffic counterparts that are (typically) associated with long haul trips where fewer but larger deliveries per chain are concerned. Conversely, our study shows that within-traffic have slightly fewer activities per chain.

It is arguable that the number of activities per chain follow a smooth Gamma or Log-normal distribution with a maximum density between 5 and 10 activities per chain. From that argument follows that since both the 90th and 92.5th percentile thresholds present such a shape, a good choice for the threshold duration between minor and major activities would be between 180 min (3 h) and 300 min (5 h).

We next analysed the duration of activity chains. The results presented in Fig. 6 again distinguishes between within- (Fig. 6a) and through-traffic (Fig. 6b). Here the density distributions are even more similar between the two vehicle types. In the case of the 99th percentile threshold, note the increase in density around five days. This suggest that with a large threshold value, we tend to rather extract *weekly* chains since inactivity over weekends are considered signals for the end of chains. As expected, this increase in density is slightly more pronounced in within-traffic (49% of chains compared to 41%) where inactivity over weekends follows normal business cycles more closely.

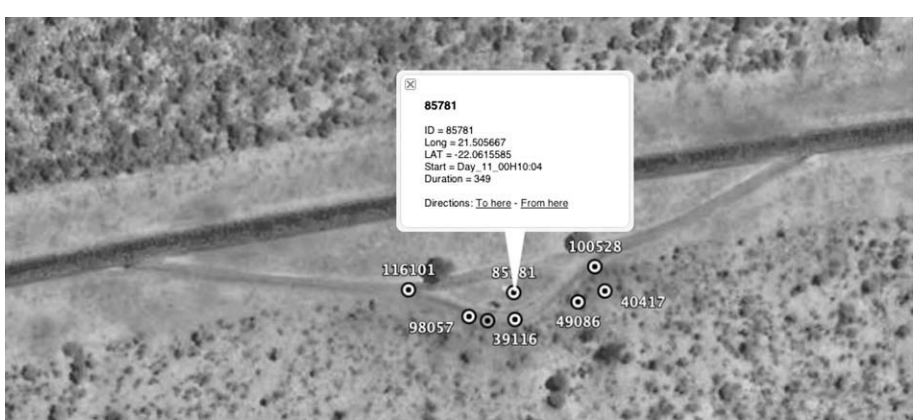

Fig. 3. Aerial view of an informal truck stop indicating major activities of a number of vehicles. The start time of 00h10:04 suggests a truck pulling off the road for the night. (Source: *GoogleEarth* at location 22°3'41.14"S, 21°30'20.27"E, accessed on 7 April 2009).

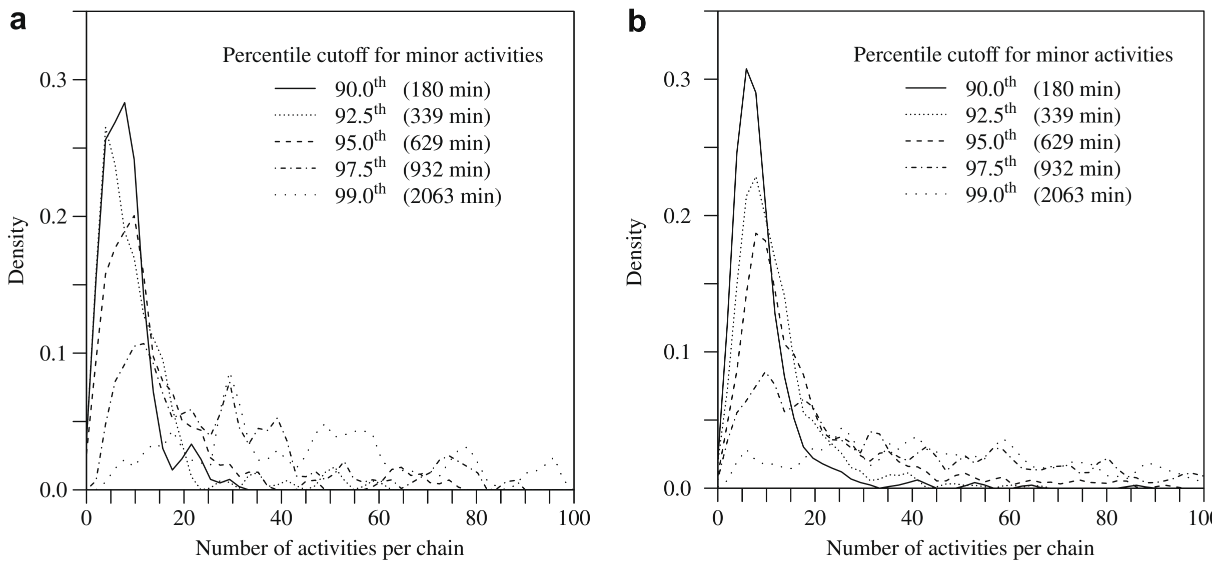

Fig. 5. Comparison of the number of activities per chain for the different minor activity thresholds. (a) Within-traffic. (b) Through-traffic.

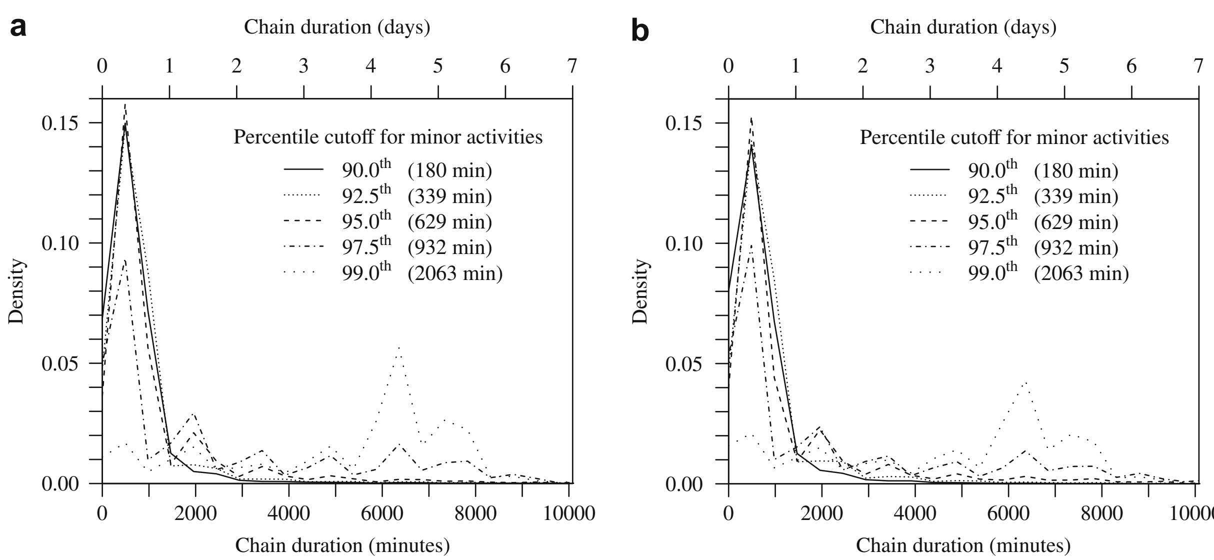

Fig. 6. Comparison of the duration of activity chains for the different minor activity thresholds. (a) Within-traffic. (b) Through-traffic.

Although having a much clearer picture of the activity and activity chain characteristics at this point, we were tempted to find a way to link the spatial density of activities in the study area to productivity. Government funds investment in transport infrastructure from taxes and levies. Taxes are the result of people and companies profitably conducting their business. Fuel and road levies (tolling) also contribute to the provision and maintenance of transport infrastructure. In return, people and companies need good transport infrastructure to competitively pursue their goals and be profitable, continuing the cycle.

In his article, eluding to the richness and diversity of freight transport geography, Rodrigue (2006) notes that the transport and logistics corporations are on average a lot more profitable than their (mostly) government-owned counterparts. We believe that freight, being spatially more challenging and diverse than public transit, provides a means to measure the productive use of infrastructure of a province.

4.1. Productivity metrics

The premise of our metric proposal is that commercial vehicle activities contribute to the profitability of companies, and thus to the GDP. The first metric is fiscal and calculates the amount of GDP generated by the province per activity conducted within the province. The higher the metric value, the more *productive* an area is, contributing more to GDP per commercial vehicle activity. Although Gauteng has the highest GDP in South Africa, it also accounts for the majority of commercial vehicle activities. It may have intuitively been expected that the large number of activities in Gauteng may yield a low fiscal metric, yet Gauteng turns out to produce the highest: South African Rand (ZAR) 151,700 per activity; followed by the Western Cape with ZAR 106, 200, and KwaZulu-Natal with ZAR 73, 100.

Another metric we propose to compare the freight productivity of the province is *distance per activity*. For each vehicle, we consid-

Please cite this article in press as: Joubert, J.W., Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. J. Transp. Geogr. (2009), doi:10.1016/j.jtrangeo.2009.11.005

ered the total chain distance of chains that have one or more activities within the study area, and divided it by the total number of activities that the vehicle performed within the study area. We postulate that a low metric value is preferable. Since, from the GDP metric, we can assume that freight activities are an essential component of economic growth, it is preferable that the amount of infrastructure use be limited. The fewer kilometers travelled on the road per activity, the more spendable GDP to invest in new, and maintain existing infrastructure. We acknowledge that the following scenario could exist: a vehicle performing many activities on long chains could yield the same metric value than a vehicle performing very few activities on short chains. Still, the metric yields a valuable indication of how much a vehicle travels (uses road infrastructure) for the amount of commercial activities conducted (contribution to GDP). The results are visualized in Fig. 7, again distinguishing between within- and through-traffic. Within-traffic is much more productive than through-traffic. Since within-traffic performs the majority of its activities within the province, its use of the road infrastructure is much higher. Likewise, the contribution is higher with more activities, and the combination results in a more productive use of road infrastructure, i.e. lower metric value.

The variance in the distance per activity metric for through-traffic is very large. We acknowledge that using the chain distance as numerator, and not just the distance of the portion of the chain that occur in the study area, suggests refinement of the metric is needed. Still, the metric confirms the intuitive expectation that through-traffic may be utilising more of the road infrastructure than what they contribute to maintaining the infrastructure through the activities they conduct in the study area.

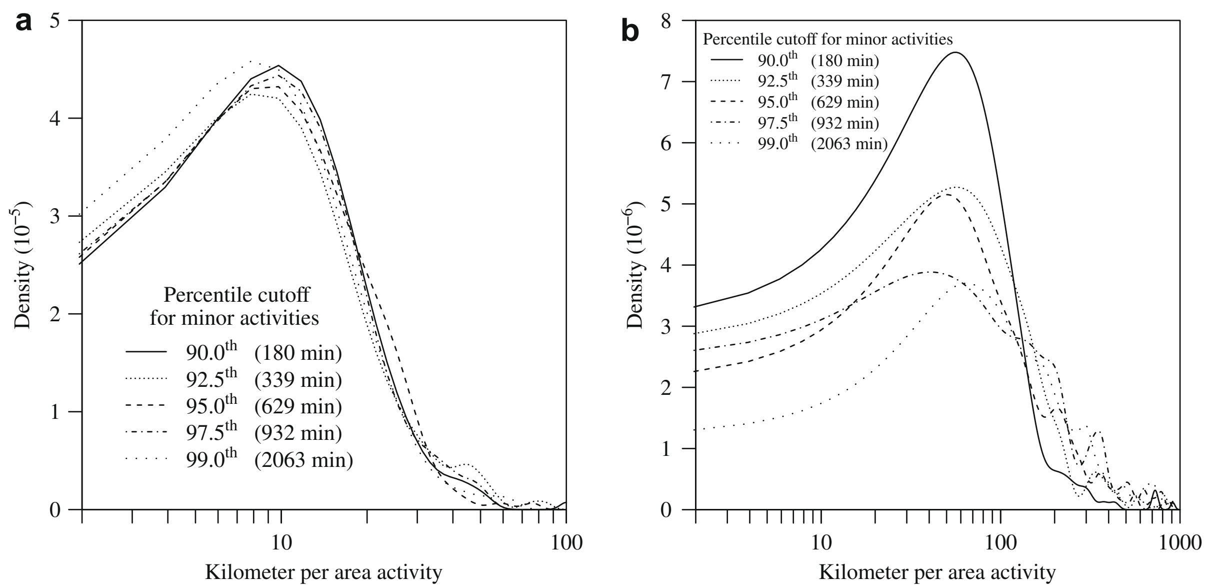

Fig. 7. Comparison of the chain distance per activity metric for the different minor activity thresholds. (a) Within-traffic. (b) Through-traffic.

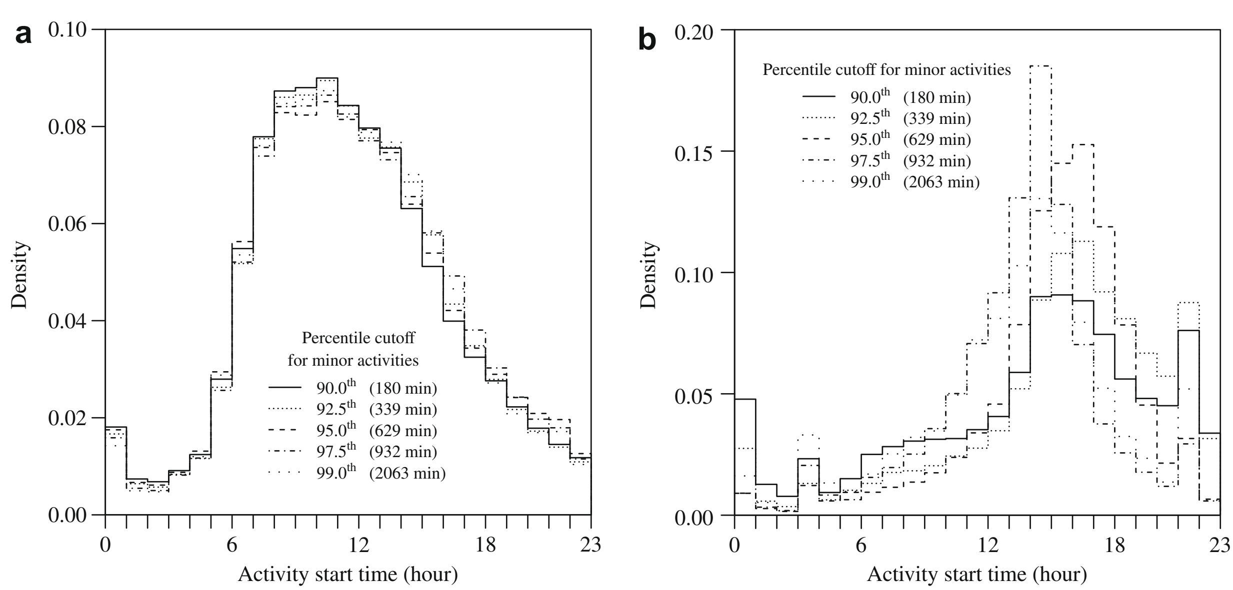

Fig. 8. Comparison of the time of day when activities start for the different minor activity thresholds. (a) Minor activities. (b) Major activities.

Please cite this article in press as: Joubert, J.W., Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. J. Transp. Geogr. (2009), doi:10.1016/j.jtrangeo.2009.11.005

Combining the two metrics yields further opportunities for analysis. Consider a typical within-traffic vehicle in Gauteng with a distance per area activity metric value of 10 km per activity. From the GDP metric we know that Gauteng has a value of ZAR 151,700 per activity. For this vehicle, one may express a measure by dividing the GDP metric by the distance per activity metric, resulting in a value of ZAR 15,170 km^{-1}. Although there is no direct causality and more kilometers travelled does not necessarily imply higher GDP contribution, it is an opportune descriptive measure to compare the productivity of different areas, say provinces, or establish a baseline for an area to see if they can improve economic activity (GDP) at a higher rate than an increase in commercial vehicle traffic that will lead to increased congestion.

4.2. Temporal distribution

To understand the impact of minor activity thresholds on the temporal characteristics of activities, we analysed the start time distributions of both minor and major activities. The results are presented in Fig. 8. The distribution of minor activity start times (Fig. 8a) are very similar. This analysis confirms the survey results of Hunt and Stefan (2007) that commercial activities tend to concentrate more during the middle of the workday than in the morning and afternoon peaks.

Start time distributions for major activities, conversely, differ quite substantially in scale for different threshold values. One possible explanation is that we chose not to distinguish between with-

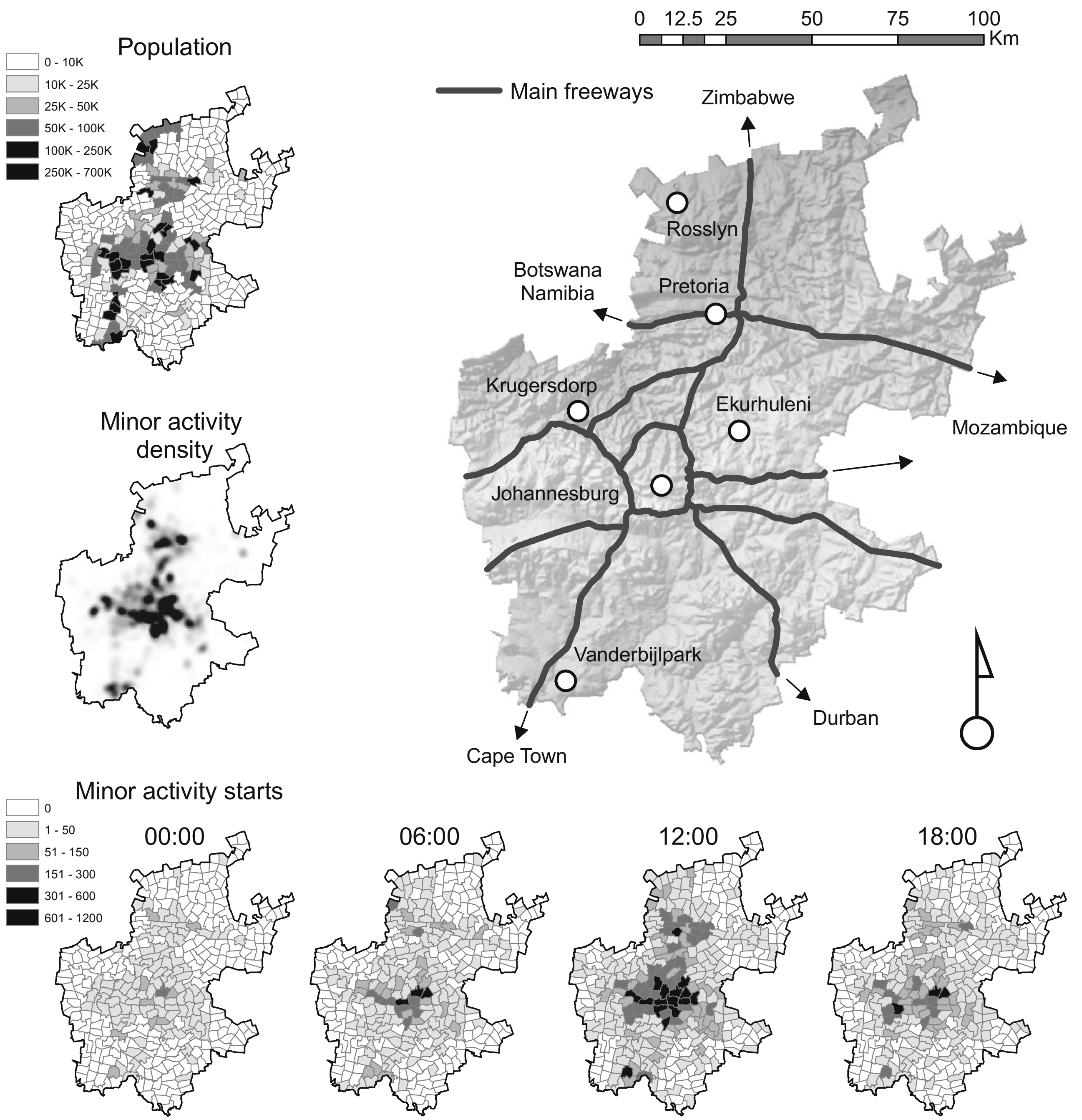

Fig. 9. Density comparisons for the province of Gauteng.

Please cite this article in press as: Joubert, J.W., Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. J. Transp. Geogr. (2009), doi:10.1016/j.jtrangeo.2009.11.005

in- and through-traffic as in earlier analyses. Although this choice holds for minor activity start time distributions, it may not hold for the major activities. Although the scale varies, the shape of the distributions are consistent in Fig. 8b with highest densities between 12:00 and 16:00, and another peak at 22:00. With major activity start times indicating the chain ending times, a right-skewed distribution, i.e. the latter half of the day, is intuitively expected.

Since congestion is considered notoriously unproductive, it is interesting to note that between 60% and 87% of chains have ended before the afternoon peak starts at 16:00. This further fuels the debate that commercial vehicles are lesser contributors to peak-hour congestion.

4.3. Density comparisons

In this subsection we combine the temporal with the spatial distribution of commercial activities in Gauteng. More specifically, we consider the distribution of minor activities over the course of a 24-h period. We use the Council for Scientific and Industrial Research's (CSIR) Geospatial Analysis Platform (GAP) mesoframe demarcation (CSIR Built Environment, 2007, 2009; Naudé et al., 2007). Each mesoframe is approximately 7×7 km, yielding an approximate surface area of 50 km^2. We prefer the equal-area demarcation since it highlights geographic areas of activity density more accurately.

As motivated in Section 3, we used a minor activity threshold value of 300 min (5 h) when extracting chains. Fig. 9 provides a reference map in the top right showing the main freeways through Gauteng. The top left map shows the population density of Gauteng, and the center left map shows the activity density of all minor activities. The bottom four density maps show the number of activity starts during four hours of the day. The distribution of the activities remain the same throughout the day, but intensifies in key activity areas. Activities are concentrated between Johannesburg and Ekurhuleni, and South-East of Johannesburg in the Germiston and Alrode areas.

There is a high degree of overlap between activity densities and population densities. Hesse and Rodrigue (2004) observe that distribution centers, which attract many logistic activities, are concentrated on the fringes of urban areas and beyond. On the contrary, at least in South Africa, logistic activities are centered within the urban areas, especially in the case of Gauteng. The explanation of this phenomenon has a political underpinning. Reference was made in Section 2 to the racial segregation of the Apartheid past where non-white communities were forcefully relocated to the periphery of the urban areas. This is also where many commercial areas are located in the form of industry and distribution centers.

Even 15 years after Apartheid has been abolished, little has changed in the land use, and large densely populated low-income (mostly black) settlements still remain. With the urban development over time, it is becoming increasingly difficult to treat Johannesburg, Tshwane (Pretoria) and Ekurhuleni as independent cities. Instead of each local metropolitan council responsible for its own urban planning, as is currently the case, the Gauteng conurbation requires a more holistic planning approach. Low income earners are effectively competing for the same land as developing commercial industries that attract an increasing number of commercial vehicle activities.

To be an effective economic hub, and gateway into Southern Africa, Gauteng should consider a strategic urban development plan, one that takes the activities and movement of both people and freight into account. The richness of the transport geography can thus not be fully exploited with only the observed characteristics of commercial vehicles, i.e. the number of activities per chain, the chain duration, and the temporal distribution of activities. The interaction of commercial vehicles with its environment needs to be considered within an appropriate context.

5. Conclusion

We have presented a methodology to analyse the disaggregate commercial vehicle activities. Although Rodrigue (2006) praised the profitability of the freight transport sectors, such praises may be somewhat dampened if we consider the work of Stander and Pienaar (2002, 2005) noting that freight does not cover the full social and political cost for their road use. The South African government intends to separate the ownership and operation of infrastructure, and so it will remain a challenging task for central transport planning authorities—the custodians of infrastructure investment—to walk the fine line between effective facilitation of economic growth, and the financing of the system from user fees. Another contribution of this paper is proposing two metrics based on commercial vehicle activities that may assist planning authorities in evaluating the productive use of their transport infrastructure.

Business location density plays an important role in freight affecting congestion, and the related environmental and fuel consumption effects. Knowing now where activities take place, i.e. where business and freight stakeholders are located; and what the inter-activity distances are, allows us in future to evaluate and compare the activities against underlying land use data. In turn, land use analysis may provide guidelines in predicting future freight activities based on strategic land use masterplans prepared by government. Gaining access to such land use data remains extremely difficult, especially since local government maintains its records independently—Gauteng is made up of fives such independent authorities.

We have achieved significant results with this paper in identifying freight activities. The activity locations, along with the activity chain characteristics allows us to generate a representative synthesized population of freight *agents*. Using a large-scale agent-based transport simulator such as the toolkit by the MATSim Development Team (2009) allows us to incorporate freight movement in dynamic traffic simulations. Decision support on transport infrastructure could then be considered more representative of the infrastructure users.

Acknowledgements

First, and foremost, we would like to thank *DigiCore Fleet Management* for preparing the *Ctrack* data set to be anonymous with respect to their customers, and making it freely available to us for our research purposes. We would also like to thank the two anonymous reviewers for their valuable comments to improve the quality of this paper. The first author is grateful to the South African National Research Foundation (NRF) who was partly funding the research under Grant FA-2007051100019; and the University of Pretoria's Research Development Programme.

References

Ambrosini, C., Routhier, J.-L., 2004. Objectives, methods and results of surveys carried out in the field of urban freight transport: an international comparison. Transport Reviews 24 (1), 57–77.

Aschauer, D.A., 1989. Is public expenditure productive? Journal of Monetary Economics 23 (2), 177–200.

Aschauer, D., 1993. Genuine economic returns to infrastructure investment. Policy Studies Journal 21 (2), 380–390.

Beuthe, M., Jourquin, B., Geerts, J., á Ndjang' Ha, C.K., 2001. Freight transportation demand elasticities: a geographic multimodal transportation network analysis. Transportation Research Part E: Logistics and Transportation Review 37 (4), 253–266.

CSIR Built Environment, 2007. SA Geospatial analysis platform, version 2: spatially disaggregated and interpolated socio-economic and accessibility indicators based on 2004 magesterial district data, obtained from Global Insight's Regional Economic Focus (REX version 2.0c (190)). Collaborative project with the Policy Advice and Coordinating Unit (PCAS) of the presidency, co-funded by GTZ.
CSIR Built Environment, 2009. Geospatial analysis platform and NSDP spatial profiles. <http://www.gapweb.co.za> (accessed 13.03.).
de Jong, G., Ben-Akiva, M., 2007. A micro-simulation model of shipment size and transport chain choice. Transportation Research Part B: Methodological 41 (9), 950–965.
de Jong, G., Gunn, H., Ben-Akiva, M., 2004. A meta-model for passenger and freight transport in Europe. Transport Policy 11 (4), 329–344.
Department of Transport, 1999. Moving South Africa. National Department of Transport, Pretoria, South Africa.
Department of Transport, 2005. National Freight Logistics Strategy. Department of Transport, Pretoria, South Africa.
Department of Transport, 2008. Strategic Plan 2008–2011. Department of Transport, Pretoria, South Africa.
Development Policy Research Unit, 2008. Poverty and the 'Second Economy' in South Africa: an attempt to clarify applicable concepts and quantify the extent of relevant challenges. DPRU Policy brief series PB 08-20, School of Economics, University of Cape Town, Cape Town, South Africa.
Draijer, G., Kalfs, N., Perdok, J., 2000. Global positioning system as data collection method for travel research. Transportation Research Record 1719, 147–153.
Edwards, P., 2007. South Africa 2003–04: two economies. <http://infrastructuration.blogspot.com/2007/04/south-africa-2003-04-two-economies.html>. (retrieved 31.03.09.).
Fedderke, J., Bogetic, Ž., 2009. Infrastructure and growth in South Africa: direct and indirect productivity impacts of 19 infrastructure measures. World Development 37 (9), 1522–1539.
Fernández, J.E., de Cea, J., Soto, A., 2003. A multi-modal supply–demand equilibrium model for predicting intercity freight flows. Transportation Research Part B: Methodological 37 (7), 615–640.
Figliozzi, M.A., 2007. Analysis of the efficiency of urban commercial vehicle tours: data collection, methodology, and policy implications. Transportation Research Part B: Methodological 41 (9), 1014–1032.
Fourie, D., 2001. The restructuring of state-owned enterprises: South African initiatives. Asian Journal of Public Administration 23 (2), 205–216.
Friedrich, M., Haupt, T., Nökel, K., 2007. Freight modelling: data issues, survey methods, demand and network models. In: Axhausen, K.W. (Ed.), Moving through Nets: The Physical and Social Dimensions of Travel. Elsevier, Amsterdam, The Netherlands, pp. 249–277 (Chapter 9).
Hensher, D., Figliozzi, M.A., 2007. Behavioural insights into the modelling of freight transportation and distribution systems. Transportation Research Part B: Methodological 41 (9), 921–923.
Hesse, M., Rodrigue, J., 2004. The transport geography of logistics and freight distribution. Journal of Transport Geography 12 (3), 171–184.
Hunt, J., Stefan, K., 2007. Tour-based microsimulation of urban commercial movements. Transportation Research Part B: Methodological 41 (9), 981–1013.
Kremers, H., Nijkamp, P., Rietveld, P., 2002. A meta-analysis of price elasticities of transport demand in a general equilibrium framework. Economic Modelling 19 (3), 463–485.
MATSim Development Team, 2009. Multi-agent transport simulation toolkit. <http://www.matsim.org/>. (retrieved 01.04.).
Munnell, A.H., 1992. Policy watch: infrastructure investment and economic growth. The Journal of Economic Perspectives 6 (4), 189–198.
Naudé, W., 2009. Geography, transport and Africa's proximity gap. Journal of Transport Geography 17 (1), 1–9.
Naudé, A., Badenhorst, W., Zietsman, L., Van Huyssteen, E., Maritz, J., 2007. Geospatial analysis platform – versions 2: technical overview of the mesoframe methodology and south african geospatial platform. Technical Report CSIR/BE/PSS/IR/2007/0104/B, Council for Scientific and Industrial Research, Pretoria, South Africa.
Rich, J., Holmblad, P., Hansen, C., 2009. A weighted logit freight mode-choice model. Forthcoming in Transportation Research Part E: Logistics and Transportation Review.
Rodrigue, J., 2006. Transport geography should follow the freight. Journal of Transport Geography 14 (5), 386–388.
Schüssler, N., Axhausen, K., 2009. Processing GPS raw data without additional information. Transportation Research Record 2105, 28–36.
Shirley, C., Winston, C., 2004. Firm inventory behavior and the returns from highway infrastructure investments. Journal of Urban Economics 55 (2), 398–415.
Stander, H., Pienaar, W., 2002. Perspectives on freight movement by road and rail in South Africa. In: 21st Annual South African Transport Conference, 11–13 July, Pretoria, South Africa.
Stander, H., Pienaar, W., 2005. Land freight issues in South Africa. In: 24th Annual South African Transport Conference, 11–13 July, Pretoria, South Africa.
Statistics South Africa, 2009. Gross domestic product. Statistical release P0441, Statistics South Africa.
Van Dyk, F., Marais, M., Naude, A., Meyer, I., Koen, R., Ralehoko, E., Ramokgopa, L., Ittmann, H., Pienaar, W., Havenga, J., Hobbs, I., Botes, F., Jacobs, C., 2005. First state of logistics survey for South Africa 2004: the case for measurement and revitalisation of basic logistics infrastructure in our dual economy. Technical report, CSIR Centre for Logistics and Decision Support, Pretoria, South Africa.
Wolf, J., Guensler, R., Bachman, W., 2001. Elimination of the travel diary: experiment to derive trip purpose from global positioning system travel data. Transportation Research Record 1768, 125–134.
Wolf, J., Schönfelder, S., Samaga, U., Oliveira, M., Axhausen, K., 2004. Eighty weeks of global positioning system traces: approaches to enriching trip information. Transportation Research Record 1870, 46–54.
Yalamanchili, L., Pendyala, R., Prabaharan, N., Chakravarthy, P., 1999. Analysis of global positioning system-based data collection methods for capturing multistop trip-chaining behavior. Transportation Research Record 1660, 58–65.
Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L., 2004. Discovering personal gazetteers: an interactive clustering approach. In: Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems, Washington, DC, pp. 266–273.

Please cite this article in press as: Joubert, J.W., Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. J. Transp. Geogr. (2009), doi:10.1016/j.jtrangeo.2009.11.005