Using multi-source GPS data to characterize multiday driving patterns and fuel use in a large city region

Christoffel J Venter*

Centre of Transport Development
Department of Civil Engineering
University of Pretoria, Pretoria, 0002, South Africa
Tel. +27-12-420-2184
Fax +27-12-362-5218
Email: christo.venter@up.ac.za

Johan W Joubert

Centre of Transport Development
Department of Industrial and Systems Engineering
University of Pretoria, Pretoria, 0002, South Africa
Tel. +27-12-420-2843
Fax. +27-12-362-5103
Email: johan.joubert@up.ac.za

**corresponding author*

Word Count:	Body =	5353 words
	8 figures and tables x 250 =	2000 words
	Total =	7353 words

Revised paper
15 March 2013
Transportation Research Record, 2013

PAPER No: 13-1033

ABSTRACT

The paper describes the use of GPS data obtained from both commercial and project-specific sources to examine the travel behavior and fuel consumption patterns of drivers over a three-day period in Gauteng Province, South Africa. Data for commercial (truck and light delivery vehicle) traffic are obtained from a commercial fleet management provider, which continuously tracks the movements of 42,000 vehicles. Data for private car users come from a panel of 720 drivers, whose multiday driving activity is tracked using mobile passive GPS loggers. We analyze and compare the driving behavior of the two driver populations in terms of total distance travelled, spatial patterns (e.g. the amount of travel on different road types) and temporal variations (e.g. variations across time of day and across multiple days). The detailed nature of GPS data also permits the estimation of fuel consumption at a very disaggregate level (by link and time of day), and the identification of differences between user groups, which have significant implications for transport and energy policy. We introduce a new indicator, the recovery ratio, to assess the relationship of fuel use to distance travelled on different classes of roads, to help identify equity distortions across user groups. Lastly, we comment on research needs related to the collection and integration of GPS data from multiple sources for model calibration and program evaluation.

Using multi-source GPS data to characterize multiday driving patterns and fuel use in a large city region

INTRODUCTION

The use of Global Positioning System (GPS) based technologies for collecting travel data is growing fast worldwide. As GPS technologies improve, and as research and practical experience grows, a number of shifts are occurring in their application. Technological improvements in the cost, weight, and battery life of mobile GPS logging devices are causing a shift from vehicle-based to person-based measurement of mobility. 'Wearable GPS' studies have been conducted worldwide, including in the United States, Europe, Japan, and Australia (*1,2*). Wearability, in turn, is driving a shift from the use of GPS data as a supplementary source of travel data – typically used to audit the trip rates of a sub-set of respondents to a conventional travel diary survey, and to calculate correction factors (*2*) – to full-sample GPS surveys. A number of large-sample household travel surveys using GPS (with or without prompted recall surveys) have been concluded or are underway, for instance in Cincinnati (1,500 households) (*3*) and Jerusalem (3,000 households) (*4*). At the same time data collection periods are increasing beyond the traditional one-day survey, with respondents carrying GPS devices for up to a week (e.g. *5*), providing rich data on the day-to-day repetitiveness and variability of travel without adding to respondent burden (*6*). Compared to standard travel surveys, GPS-recorded personal movement data are seen as beneficial in terms of enhanced accuracy (both of locations and routes), completeness, and efficiency (*2,5*).

In-vehicle GPS devices and Automatic Vehicle Location (AVL) applications have long been used to monitor and manage commercial vehicle (mainly bus and truck) fleets, and provide another potential source of GPS movement data (*7*). Despite the availability and size of such datasets, they have only recently been explored as a source of truck movement data for public agencies (*8*). Commercial GPS data could significantly advance a disaggregate understanding of the behavior of road freight traffic (*9*). As far as we are aware, GPS data from commercial fleet management and project-specific, mobile GPS sources have not been combined to analyze and compare private and commercial vehicle traffic in the same area.

It is fair to say that the potential efficiency advantages of GPS technologies, coupled with their ubiquity and flexibility, make them attractive for the collection of travel data in developing countries. GPS surveys might help overcome some of the literacy and respondent burden problems experienced with conventional activity or travel diary surveys in less literate or multicultural populations (*10,11*). Yet relatively few studies have assessed these issues outside of the developed world. A recent review of GPS travel studies worldwide found no applications (in the English literature) in South America or developing Asia (*13*). Some experimentation has been done in the Western Cape, South Africa (*12*), comparing the results from wearable GPS with conventional travel diary data, and demonstrating its usefulness in areas with limited address and street network information.

This paper reports on the use of GPS data to estimate fuel consumption at a very disaggregate level (by link and time of day) for multiple vehicle types in the Gauteng urban region in South Africa. We link fuel consumption with the observed travel behavior of both private (car) users and commercial (freight) vehicles, which provides insights into the travel behavioral factors affecting fuel consumption.

The work is significant firstly for its international focus; we report on specific issues related to the collection and use of GPS data in developing countries. Secondly, we demonstrate the feasibility of combining GPS data from two independent sources – a small-sample study using mobile GPS devices, and commercial fleet management GPS data – to provide a more complete picture of travel-related resource consumption in an area. We argue that the high quality data obtainable from GPS sources – including information on multi-day travel patterns, detailed route choice and speed, and day-to-day variability – can benefit fuel use studies beyond what is obtainable from traditional survey or modeling sources. We also discuss some of the problems and requirements of combining GPS data from multiple sources. The paper's third contribution is substantive: the analysis of fuel use patterns delivers insights into the differences in driving behavior between different user groups, and in their respective contribution to the overall fuel bill. We discuss policy implications of these findings relating to fuel efficiency programs and road funding, and make recommendations for further research.

The paper starts by discussing the public policy and technical antecedents for this work. We then describe the objectives and methodology for collecting and analyzing the respective sets of GPS data, and consider data collection issues that might be relevant to other developing countries. Thirdly, we present selected findings on the travel behavior of private and commercial vehicle drivers in Gauteng, followed by a description of the fuel consumption model and its results. We conclude with findings and recommendations for further research.

BACKGROUND

This research is situated within a larger effort to improve the state of travel demand modeling in South Africa, through (i) developing and customizing an agent-based modeling capability for forecasting transport user behavior; and (ii) collecting a range of empirical data on travel behavior that can be used to calibrate and validate new models, and improve travel analysis more generally. The Multi-Agent Transport Simulation (MATSim) toolkit allows for large-scale transportation simulations and has been applied in South Africa for both private cars and commercial vehicles (*14*). Being an agent-based approach, MATSim requires a synthetic population of agents (vehicles), each with an activity chain, i.e. a sequence of activities, linked with transport leg(s), which the agent pursues to maximize utility. Through the mobility simulation agents *learn* iteratively in a co-evolutionary manner in that they are allowed to respond to observed traffic conditions. At the end of each iteration an agent can a) choose an alternative plan from its memory (experience), b) change the current plan by adapting the timing of the activities, or c) re-route the path between activities.

As shown by Kickhöfer et al. (*15*), microscopic agent-based simulations are well suited to evaluate policy decisions where the welfare of different portions of the population is impacted differently. This is particularly relevant in a country such as South Africa with high economic inequality among the citizens.

Validation and customization of the MATSim platform makes extensive use of GPS travel data to provide accurate measurements of current behavior. Firstly, a 3-day GPS survey has been undertaken of a sample of current car drivers in the province of Gauteng, to provide detailed information on the daily movements of car-using individuals. This information is being used to extract activity chain information needed for the simulation of travel plans by MATSim. The survey will be expanded to other sub-populations such as public transport and non-motorized users. Secondly, the GPS survey is being implemented as a panel survey: the same individuals will be resurveyed at one-year intervals to study long-term behavior changes (e.g. in car ownership, residential and work location, and travel habits) and estimate elasticities. Thirdly, in-vehicle GPS data from commercial truck fleets are used to drive activity chain generation for freight vehicles in the MATSim model (see below).

There is an immediate policy question driving this work. The South African National Road Agency Ltd. (SANRAL) recently completed a major upgrade of 185 kilometers of freeway infrastructure in Gauteng (*16*) (Figure 1). With a population of 11 million and including the commercial and administrative hubs of Johannesburg and Pretoria, Gauteng is South Africa's most densely urbanized province. The Gauteng Freeway Improvement Project (GFIP) was aimed at relieving congestion, improving traffic management through Intelligent Transport Systems deployment, and improving traffic safety on the province's extensive freeway network. Funding for the project was raised from the bond market, to be paid back via user charges collected at a set of 42 open-road tolling gantries. The project is publicized as being the largest open-road tolling project in the world (*17*).

Not surprisingly, the tolling component of the GFIP is mired in controversy. Court action has stopped the implementation of the toll collection, while the South African government is investigating alternative funding models to repay the massive debt incurred. As can be expected in a developing country with very large income inequalities, issues of equity are of particular concern. Calls for replacing the user charge with a general fuel levy raise questions not only around who benefits and who pays for the use of specific parts of the road network, but more generally around the appropriateness of user charging in urban areas of developing countries. The issue also raises questions around the likely distribution of a mooted fuel surcharge across the driver population – hence the focus on distributional impacts in this paper.

DATA SOURCES AND METHODOLOGY

GPS data from commercial fleet management sources

Many companies use satellite-based tracking devices on their commercial vehicle fleets to provide vehicle and cargo security, and better manage the movement of fleets. One company providing such tracking-based fleet management services in South Africa and internationally is Digicore Fleet Management through its *cTrack* product offering.

First analyzed by Joubert and Axhausen (*9*) using 2008 data, the *cTrack* data provide a large sample of various vehicle-related metrics such as engine temperature, opening and closing of doors in an ignition-on state, harsh braking, and over revving. We used the GPS records of nearly 42,000 commercial vehicles collected continuously over the period 1 January to 30 June 2009. Following the same procedure as Joubert and Axhausen (*9*) we used ignition-related signals in the data to extract activities, and activity chains. Each chain starts and ends with a *major* activity lasting more than 3 hours, and zero or more *minor* activities making up the activities of the chain.

The sample of vehicles observed account for approximately 1.8% of the total commercial vehicle population if compared to the national registry of vehicles in South Africa.

The GPS signals emitted are often infrequent, especially if no extraordinary events are identified in the vehicle's journey. This has the effect that detailed map matching is not possible as signals can be as much as 300 seconds (5 minutes) apart. We thus used an A*-landmarks shortest path routing algorithm on the road network to estimate the route between consecutive activities in the chain.

Although the tracked vehicles travelled across the country, we focused in this paper on those vehicles that performed at least one activity in Gauteng. Hence, we had to distinguish between *intra*-provincial vehicles (10,267), those who conduct at least 60% of their observed activities inside the boundaries of Gauteng, and *inter*-provincial vehicles (11,073), those who conducted fewer than 60% of their observed activities inside the Gauteng boundary. The 60% threshold corresponds with the findings of Joubert and Axhausen (*9*). The balance of the vehicles (20,371) did not perform any activities inside Gauteng and were not taken into account in this paper.

Portable passive GPS loggers for private car drivers

The baseline driving patterns of private vehicle users were captured through the use of portable passive GPS logging devices carried in the vehicle by a sample of respondents. This provided vehicle-based rather than person-based movement data, in line with the initial focus of the research on capturing driver patterns of current and potential freeway users. Robust, off-the-shelf GPS devices with minimal features were used to prevent tampering. Devices have their own power source so no installation was required; the battery life of between 5 and 8 eight days was adequate for our purposes. Devices simply logged GPS coordinates every second while the vehicle was in motion.

The process of recruiting participants and securing the necessary data was as follows:

- Sampled individuals (see below) were recruited via face-to-face home visits. Initial attempts (during the pilot survey) to recruit participants telephonically proved unsuccessful, with recruitment rates of only 5%. This is much lower than rates of 20-25% found elsewhere using similar recruitment methods for GPS studies (*2*), and might be related to higher concerns with crime in South Africa. The recruitment rate using home visits increased to 80%, which is more in line with the level of response rates typically found for household travel surveys in South Africa (*18*).

- Respondents were offered a R200 ($25) gift certificate for participating. Anecdotally, it appears this was needed to confirm the legitimacy of the project in respondents' minds, and to ensure the GPS unit was returned.
- Willing participants were instructed to keep the unit in their vehicles for three consecutive days during which 'normal' driving activities would be undertaken (including weekends, but excluding holidays or days containing out-of-province trips). Days on which respondents failed to have the unit in the car, either on purpose or through forgetfulness, were flagged in order to distinguish them from legitimate no-travel days.
- Field workers revisited each respondent after three days to collect the device. A short interview was completed with questions on household and demographic information, work location, vehicle details, and use of alternative modes and carpools during the trial. In many cases it proved difficult to pre-arrange a return visit after exactly three days due to variations in personal schedules and weekends; the average duration of each trial was 4.08 days per respondent, and the average turn-around time per device was 5 days.
- GPS data were downloaded and processed using purpose-built software for filtering, smoothing and GIS matching of trips and routes.

Sampling for the GPS survey was based on a stratified random approach, to ensure representivity of the car owning population in the province, both spatially (at the level of twenty sub-regions within the province), and demographically (with respect to income, gender, and employment status). We oversampled in areas with higher freeway use, to enable accurate measurement in future of driver adaptation in response to tolling. In the analyses reported here this was corrected for by applying appropriate weighting (at the sampling area, demographic (gender and income) and day of the week level). The total sample size included 726 drivers, observed across 2962 travel days.

Sample bias issues

Sample bias is a potential issue among GPS studies. Bricka (*13*) reviews non-response issues in GPS surveys and concludes that technology acceptance problems may cause non-response to increase for low-income, lower educated, and minority households. Stopher et al. (*5*) report evidence that GPS samples can be biased against low-income, one-person, and non-car-owning households, but elsewhere argues that this might be a function of the transport survey in general, and not the use of GPS itself (*2*).

Since the sample was already stratified to ensure it matched the target population in terms of income, gender and employment status, we chose to assess the sample in terms of a travel behavior metric – specifically, whether the subsample of drivers who are *freeway users* differ from the true population of freeway users in the area. We used the results from an independent freeway use study conducted by SANRAL in 2009, in which the number plates of 27,000 vehicles using any of Gauteng's freeways were photographed, and the driver contacted for a telephonic interview. Table 1 shows that freeway users in the GPS sample correspond closely to the license plate survey sample in terms of gender and occupation status, but not income. Although income data in the license plate survey contains a high percentage of refusals (73%), the fact that such item non-response tends to be biased towards higher-income individuals supports the conclusion that the GPS survey is biased towards lower-income freeway users. A Chi-Square test confirms that the two income distributions are significantly different (α=0.05, χ^2=124.1). A possible reason might be that higher income drivers are more concerned with privacy issues related to the GPS equipment, and be less likely to be compensated by the monetary incentive on offer. It is worth noting, however, that findings from elsewhere that GPS surveys might favor higher income respondents do not necessarily apply in developing countries.

Even though the sample of commercial vehicles is quite large in the case of the *Digicore* data, a possible selection bias remains. Being a fleet management service provider, the tracked vehicles are only representative of those companies that do subscribe to the *cTrack* service, and might exclude vehicles of smaller operators who either subscribe to other tracking service providers, or don't subscribe to such services at all. A common problem with the freight sector is the absence of population data on fleet vehicles against which to check for representivity. We validated the *Digicore* sample by comparing their activity chains (extracted and simulated for the population) against observed traffic counts on the road network (Joubert et al. (14)), and found very good correspondence.

SELECTED TRAVEL BEHAVIOR FINDINGS

In analyzing travel behavior we focus on aspects that illustrate the benefit of using GPS sources, and also are in some way relevant to fuel consumption. We look at total driving activity, travel speeds, type of road used, and day-to-day variability.

Daily vehicle kilometers of travel

Figure 2 shows the average daily vehicle kilometers of travel (VKT) for different income groups in the private car sample.

As one would expect, and consistent with travel behavior theory, the lowest travel activity occurs in the lowest income category with an average daily travel of only about 20km. The VKT increases with income, up to the mid-range income group having the highest driving activity. The activity reduces again by about 20% for higher income groups.

This counterintuitive finding might be reflective of the peculiar historical trajectory of car use in South African cities. Low-income car users simply cannot afford long travel distances. High-income users can afford them, but also have maximum choice of housing and employment location across the urban landscape. They seem to use this flexibility to improve proximity to jobs and other activities, and to reduce driving distances overall. Medium-income drivers seem to be worst off: while they can less afford long driving distances, they also seem less able to avoid them. South African cities have experienced shortages in both medium-priced homes and industrial/manufacturing jobs, thus restricting the residential and job mobility of medium-income drivers. This finding is consistent with previous research indicating that medium-income drivers spend a much higher share of their incomes on transport than higher income drivers (*19*).

For commercial vehicles, the VKT is much higher: on average intra-provincial vehicles travel 165km per activity chain, and it increases to 532km per activity chain for inter-provincial vehicles. We express the VKT for commercial vehicles in kilometers per *chain* and not per *day* since commercial vehicle chains may extend beyond the 24-hour day associated with people, and have many more activities (*9*). The average number of activities per chain was for intra-provincial vehicles 9.2, and 8.6 for inter-provincial vehicles. With transport cost being a sizable proportion of total logistics costs, companies are aiming to maximize their fleet utilization. Activity chains with a high number of activities are indicative of such resource optimization efforts: driver teams are rotated in a multi-shift manner to reduce the idle time of a vehicle.

Road type and time of day

Figure 3 categorizes the VKT for different income groups by the type of road on which travel occurs. Overall, about a third of private car use in Gauteng province occurs on the freeway network, and only 22% on the street network. There is a clear relationship between personal income and the type of road used: freeway use rises with income, with the most affluent drivers spending more than twice the proportion of kilometers on freeways than the lowest income category. This is most likely related to the fact that the freeway network is located to serve the commuting needs of higher income drivers better than those of lower income neighborhoods, for historical settlement and road development reasons.

When comparing private cars and commercial vehicles (Figure 4), differences in their use of different road types emerge. The box-and-whisker plots show the variation within the various samples. Private cars and intra-provincial commercial vehicles spend similar proportions of their VKT on the freeways, but commercial vehicles make more use of arterial roads and less of local streets, as can be expected. Inter-provincial vehicles undertake more long-haul trips and hence show much higher VKT on the freeway and arterial network.

Table 2 shows the share of freeway VKT by private and commercial vehicles throughout the day. Time-of-day patterns are important as fuel consumption varies with speed and congestion levels. Higher income drivers tend to concentrate more of their freeway travel in the peak, while low-income drivers make more use of shoulder and off-peak periods. Average travel speeds (calculated form GPS tracks) are marginally lower in the peak than in the off-peak, and much lower on arterials and local streets (including intersection delays) than on freeways.

Day-to-day variability

Recent research has focused on the day-to-day variability of travel, arguing that habit and variability are important dimensions of travel behaviour that should be better captured by behavioural models (*6*). Multi-day GPS data is a useful source of information on the regularity and variability of travel, both at the aggregate and at the route levels (*6*, *20*).

To measure the general level of variability in daily travel activity, we calculated a coefficient of variation (CoV) for the daily vehicle kilometres of travel for each vehicle. This is done by dividing the standard deviation in VKT over three days by the average daily VKT for each vehicle. A CoV value of zero corresponds to no variation; the larger the CoV for a vehicle, the more the amount of daily travel varied from day to day. For private vehicles, we used the GPS records across three consecutive weekdays. For commercial vehicles, we arbitrarily chose three consecutive weekdays with typical traffic patterns, 10 – 12 February 2009, to calculate the CoV.

Figure 5 shows that the average CoV varies markedly between vehicle classes, and also within each class. The most important finding is that day-to-day variation is higher for private vehicles than for commercial vehicles, and that, among private vehicles, the highest variability is found among lower income groups. For inter-provincial commercial vehicles, the CoV varied between 0 and 1.73, with a median of 0.54 and a mean of 0.60. The intra-provincial vehicles' behavior was more consistent, with a lower median of 0.39 and a mean of 0.47. The fact that commercial vehicles often perform routine, milk-run type deliveries may explain the low variation, especially in the case of intra-provincial vehicles.

The average CoV for the private car sample, across all income groups, varied between 0 and 1.59, with median 0.69 and mean 0.70, significantly higher than the day-to-day variation of commercial vehicles.

A second measure of variability looks at the day-to-day variation at the *route* level. For the sake of illustration we focus on freeway use only, and identified drivers who used the same freeway section (passed the same toll gantry) on more than one day out of the three. We found that only 47% of freeway users in the sample used the same freeway section on more than one day; the rest showed no repetition in their route choice behavior.

ESTIMATING DISAGGREGATE FUEL CONSUMPTION PATTERNS

In order to estimate fuel consumption at the level of the individual vehicle, link, and time of day, the following model is used:

$$c_{nkt} = d_{nk}.b_n.f(v_{nkt})/100 \qquad \ldots (1)$$

where c_{nkt} = liters of fuel consumed by vehicle n on link k at time t
d_{nk} = distance travelled by vehicle n on link k (in kilometers)

b_n = base fuel consumption rate for vehicle n (in liters per 100 km)
$f(v_{nkt})$ = fuel efficiency adjustment factor for vehicle n travelling at speed v_{nkt}

The fuel consumption is thus dependent on the type of vehicle, the link distance, and the travel speed: both very high and very low speeds reduce fuel efficiency. Fuel consumption would thus vary by the type of road used and the time of day (depending on congestion levels), all of which would be indicated by the speed v_{nkt} recorded on the second-by-second GPS tracks.

The function $f(v_{nkt})$ was derived from previous studies (*21*):

$$f(v_{nkt}) = \begin{cases} 1.40 & if\ v \leq 15\ km/h \\ 1.55 - 0.011v & if\ 15 < v \leq 50\ km/h \\ 1.0 & if\ 50 < v \leq 80\ km/h \\ 0.520 + 0.006v & if\ v > 80\ km/h \end{cases} \quad \ldots (2)$$

and reflects the fact that optimum fuel economy occurs at a speed of between 50 and 80 km/h.

We used data on vehicle type, engine size, and fuel type (petrol or diesel) provided by respondents in the GPS sample, together with average fuel consumption rates for different vehicle types in South Africa (*22*), to determine a base consumption rate, b_n, for each vehicle. There are some differences here between driver groups: the average vehicle driven by people in the highest income category has an engine size of 2000 cc, compared to 1600 cc for other groups, which will tend to raise fuel consumption among richer drivers.

Base fuel consumption rates for heavy vehicles were estimated based on the class of vehicle (rigid versus articulated), and average rates provided from industry sources.

Table 3 shows the results of the fuel calculations by user group and road type. Fuel consumption rates are, on average, highest on streets, followed by freeways and then arterials. That the arterial network is the most fuel efficient part of the network follows from its moderate speeds: fuel penalties of higher speeds (as on freeways) and low speeds (as on streets) are avoided. Fuel consumption rates are on average slightly higher during off-peak periods than during peaks, but the difference is negligible (about 1%) for both cars and trucks. Congestion might actually help curtail speeds and fuel consumption during peak periods, but this is off-set by a higher proportion of truck and large-engine size car travel at these times (Table 2).

In terms of differences across user groups, trucks have of course much higher fuel consumption rates than passenger vehicles, but trucks contribute only 12% of the total fuel bill. Among car users drivers in the highest income group have highest fuel consumption rates on all road types, due to their use of larger, less fuel efficient vehicles. Drivers in the lowest income bracket also have high fuel consumption rates, especially on local streets: historically, many low-income neighbourhoods have insufficient street supply, causing extra congestion and fuel use.

We also calculated a ‘recovery ratio’, which is defined as the ratio between the share of fuel consumed and the share of VKT consumed by a group on a particular type of facility. This ratio is relevant to the road financing debate: fuel tax revenues contributed by a group are proportional to their total fuel consumption; if fuel levies are seen as a user payment for road use, these revenues should be on par with the amount of road use. A recovery ratio of less than 1.0 indicates that a group contributes less fuel taxes than their share of VKT demands. This is the case for car users, in general: the average recovery ratio is around 0.93. Commercial vehicles pay about 2.5 times their share of VKT, but this can be justified in relation to the extra pavement damage caused by heavy vehicles. Among car users, higher income drivers tend to have recovery ratios above 1.0, indicating that there is a certain amount of cross-subsidization of low and medium income drivers. The exception is drivers in the lowest income group, with a very high ratio of 1.08 for local streets. This suggests that they are doubly penalized: not only do they consume more fuel per kilometer by travelling on generally more congested streets, but their taxes go towards upgrading other parts of the network (arterials, freeways) that are used more frequently by other drivers.

CONCLUSIONS

This paper demonstrates firstly the feasibility, and secondly the usefulness, of using GPS data from multiple sources for extracting, characterizing, and comparing the travel behavior of multiple user groups. Data from ongoing commercial fleet management tracking programs were combined with data from a project-specific mobile GPS study to analyze and compare the travel behavior and fuel consumption of commercial (freight) and private car users on a city-region scale.

GPS-based trip data from commercial truck fleet operators are readily available and cover large populations, offering significant opportunities for transport research in the freight sector. Drawbacks of such readily available data include their lack any additional information like trip purpose (due to privacy issues), and discrepancies in tracking frequencies between GPS data from different sources. Commercial fleet tracking rates are not constant but relate to driver or engine 'events' such as stops or harsh braking, and are therefore typically lower than the one to five second intervals used by mobile person-tracking devices. Although freight vehicles' trip and stop locations and durations are accurately recorded, accuracy falls at the route level, requiring more post-processing and data inference if routing algorithms are needed to generate route or road use information. Further research is needed to help integrate freight and passenger data more easily into regional travel models and analyses.

The use of GPS data sources for travel analysis is advancing fast. Yet few applications have been seen in developing countries. Our experience suggests that GPS-aided data collection methods can be very useful in overcoming some of the problems associated with conventional recall-based travel surveys. The efficiency gains from collecting multiple days' worth of travel from the same respondent, thereby reducing sample sizes and survey costs, are particularly attractive in developing countries with restricted data collection budgets. We furthermore showed that both non-response and sample self-selection were less of a problem than what has been the experience with GPS studies in developed countries, provided adequate incentives and face-to-face recruiting methods are used. The unavailability of accurate Geographic Information System (GIS) layers for map matching and spatial analysis in developing countries might be a problem, but emerging open-source and web-based sources like openstreetmap.org and Google Maps™ might provide part of the solution.

We used the GPS-based data to estimate fuel consumption at a very disaggregate level (by link and time of day) for multiple vehicle types, and analysed the distribution of fuel use across user groups. The detailed nature of GPS data permits the identification of differences between groups, which may have significant implications for transport and energy policy. Firstly, in a reversal of typical findings in developed countries, middle-income car users drive more (per capita) than either low or high income car users, most likely due to historically restrictive housing policies. More flexible housing and job markets would help to shrink trip distances and fuel consumption.

Secondly, while there is a small amount of cross-subsidization of lower and medium-income drivers by high-income drivers in terms of fuel taxes, low-income drivers are particularly disadvantaged with respect to travel on local streets. High congestion and low speeds on residential streets lead to higher fuel use and higher taxation; improving local street networks would thus be an effective way to reduce both energy consumption and travel expenditure to disadvantaged groups (apart from other benefits such as increased road safety and access). Thirdly, travel demand management (TDM) policies that aim to shift travel away from the peaks might not reduce fuel consumption, as average fuel use rates do not drop in the off-peak.

Lastly, freeway upgrade schemes (such as the recently completed Gauteng Freeway Improvement Project in Gauteng) have potentially important equity implications, which can be exacerbated by inappropriate funding mechanisms. Financing of freeway expansion through fuel levies rather than direct tolls – a position presently being advocated by some opponents to tolling in South

Africa – would be regressive as lower-income car (and public transport) users spend a lower proportion of their travel on freeways than do their higher income counterparts. Further research is needed to accurately identify and allocate benefits and costs, relating not only to fuel costs but also to multiplier effects and intangible benefits such as reliability and safety, as these have significant bearing on the acceptability and fairness of road funding mechanisms which are under pressure worldwide. Clearly advanced data collection methods such as GPS have an important role to play in this effort.

ACKNOWLEDGMENTS

This research was partially funded by the South African National Roads Agency Ltd. (SANRAL). The analysis and conclusions are however those of the authors and do not necessarily reflect those of the sponsor. We are grateful to *Digicore* Fleet Management for making the commercial vehicle GPS data available for research purposes.

REFERENCES

1. Bricka, S.G., Sen, S., Paleti, R., and C.R. Bhat. An Analysis of the factors influencing differences in survey-reported and GPS-recorded trips. Transportation Research Part C 21, 2012.
2. Stopher, P.R.. Collecting and Processing Data from Mobile Technologies. In: Transport Survey Methods, Keeping Up With a Changing World, Bonnel, P. et al (Eds). Bingley: Emerald, 2009.
3. Giaimo, G., Anderson, R., Wargelin, L., and P. Stopher. Will It Work? Pilot Results from First Large-Scale Global Positioning System-Based Household Travel Survey in the United States. Transportation Research Record No. 2176. Washington, D.C., 2010.
4. Oliveira, M.G.S., Vovsha, P., Wolf, J., Birotker, Y., Givon, D., and J. Paasche. Global Positioning System-Assisted Prompted Recall Household Travel Survey to Support Development of Advanced Travel Model in Jerusalem, Israel. Transportation Research Record No. 2246. Washington, D.C., 2011.
5. Stopher, P.R., Clifford, E., and B. Halling. Evaluating a Voluntary Travel Behaviour Change by Means of a 3-year GPS Panel. Paper presented at 4th International Symposium on Travel Demand Management. Semmering, Austria, 2008.
6. Stopher, P.R., and Y. Zhang. Repetitiveness of Daily Travel. Transportation Research Record No. 2230. Washington, D.C., 2011.
7. Mintsis, G., Basbas, S., Papaioannou, P., Taxiltaris, C., and I.N. Tziavos. Applications of GPS technology in the land transportation system. European Journal of Operational Research 152, 2004.
8. Ma, X., McCormack, E.D., and Y. Wang. Processing Comemrcial Global Positioning System Data to Develop a Web-Based Truck Performance Measures Program. Transportation Research Record No. 2246. Washington, D.C., 2011.
9. Joubert, J.W., and Axhausen, K.W. Inferring commercial vehicle activities in Gauteng, South Africa. Journal of Transport Geography, Vol. 19, No. 1, 2011.
10. Van der Reis, P., and M. Lombard. Multi-cultural and multi-lingual transport surveys, with special reference to the African experience. In: Transport Survey Quality and Innovation, Stopher, P.R. and P.M. Jones (Eds), Elsevier: England, 2003.
11. Behrens, R. Looking beyond commuter travel in Cape Town: Methodological lessons from the application of an activity-based travel survey. In: Transport Survey Quality and Innovation, Stopher, P.R. and P.M. Jones (Eds), Elsevier: England, 2003.

12. Krygsman, S., and J.H. Nel. The use of global positioning devices in travel surveys – a developing country application. Proceedings: 28th Southern African Transport Conference, Pretoria, South Africa, 2009.
13. Bricka, S. What is Different About Non-Response in GPS-Aided Surveys? In: Transport Survey Methods, Keeping Up With a Changing World, Bonnel, P. et al (Eds). Bingley: Emerald, 2009.
14. Joubert, J.W., Fourie, P.J. and K.W. Axhausen. Large-scale agent-based combined traffic simulation of private cars and commercial vehicles. Transportation Research Record No. 2168. Washington, D.C., 2010
15. Kickhöfer, B., Grether, D., and K. Nagel. Income-contingent user preferences in policy evaluation: application and discussion based on multi-agent transport simulations. Transportation, Vol. 38, No. 6, 2011.
16. South African National Roads Agency Limited. http://www.sanral.co.za/e-toll/. (Accessed 30 July 2012)
17. Tolplan Pty (Ltd.). http://www.tolplan.co.za/introduction-services.html). (Accessed 29 July 2012)
18. Department of Transport (DOT). National Household Travel Survey 2003: Technical report. Pretoria, South Africa, 2003.
19. Venter, C. Transport expenditure and affordability: The cost of being mobile. Development Southern Africa, Vol. 28. No 1. 2011.
20. Spissu, E., Meloni, I., and B. Sanjust. Behavioral Analysis of Choice of Daily Route with Data from Global Positioning System. Transportation Research Record No. 2230. Washington, D.C., 2011.
21. Environmental Protection Agency. www.fueleconomy.gov (Accessed 7 November 2012)
22. Automobile Association of South Africa. www.aa.co.za (Accessed 7 November 2012)

List of captions

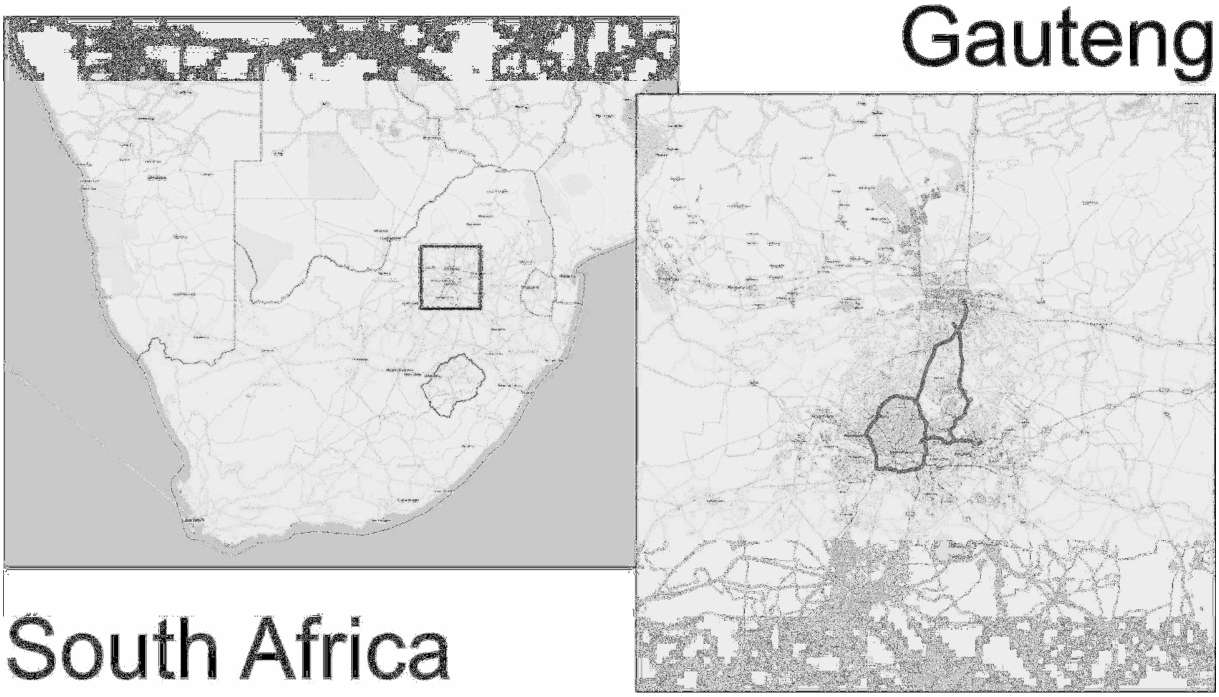

FIGURE 1

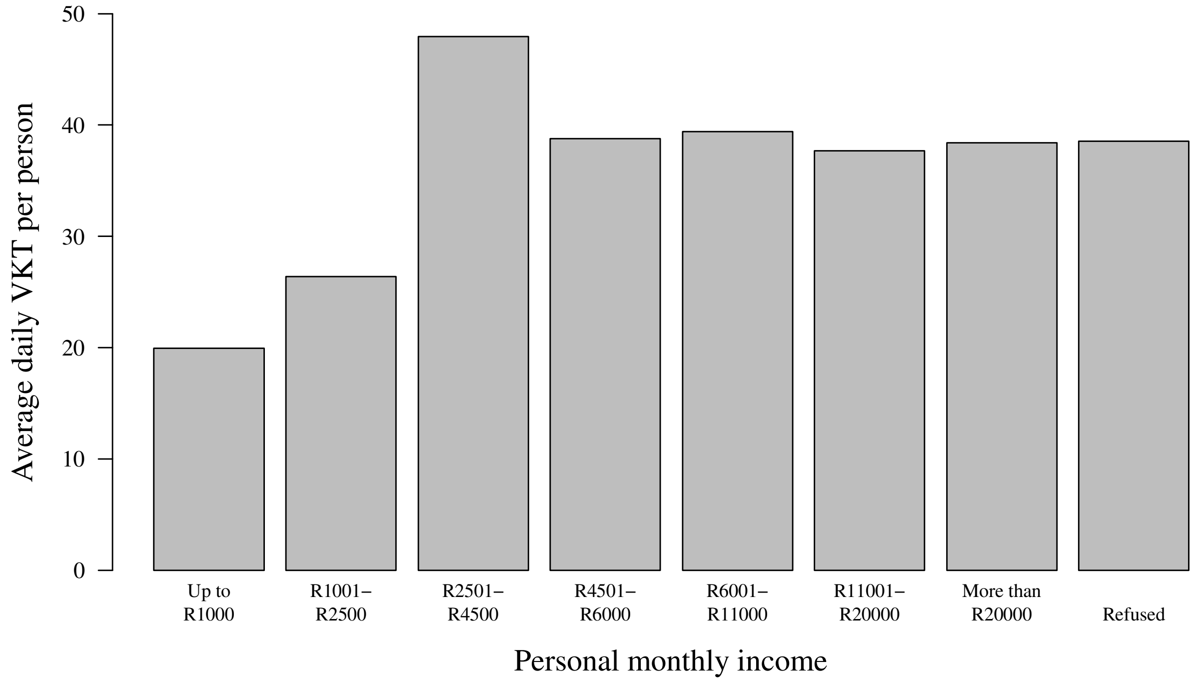
Average daily VKT per person
0
10
20
30
40
50
Up to R1000
R1001– R2500
R2501– R4500
R4501– R6000
R6001– R11000
R11001– R20000
More than R20000
Refused
Personal monthly income

FIGURE 2

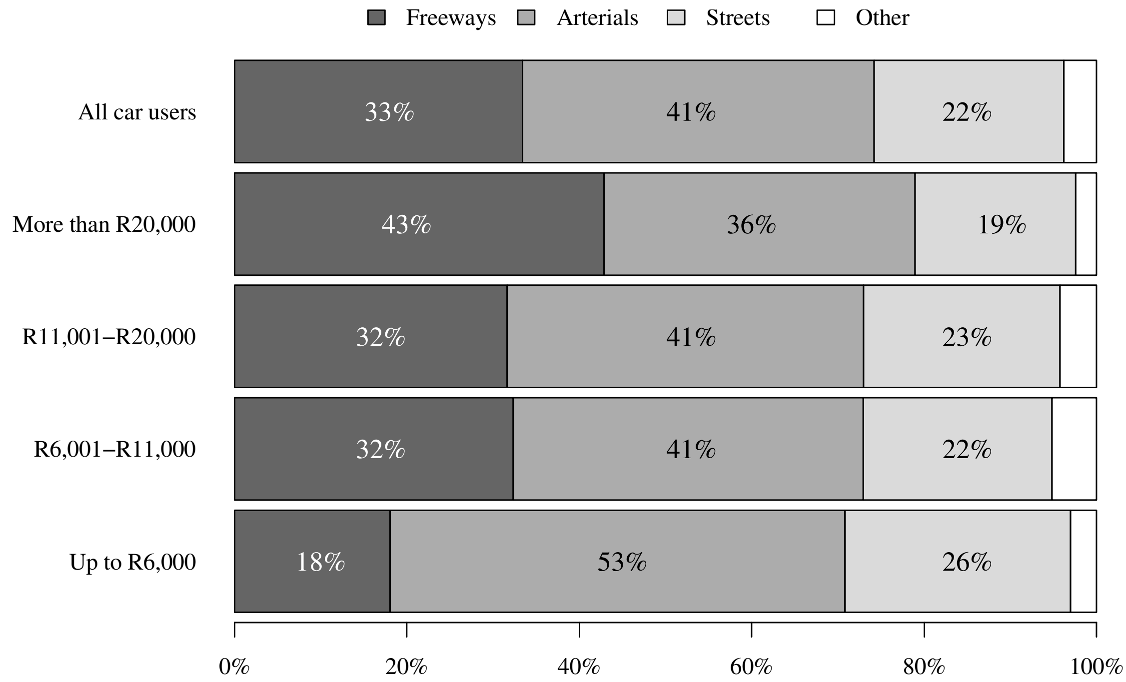
Freeways
Arterials
Streets
Other
All car users
More than R20,000
R11,001–R20,000
R6,001–R11,000
Up to R6,000
33%
41%
22%
43%
36%
19%
32%
41%
23%
32%
41%
22%
18%
53%
26%
0%
20%
40%
60%
80%
100%

FIGURE 3

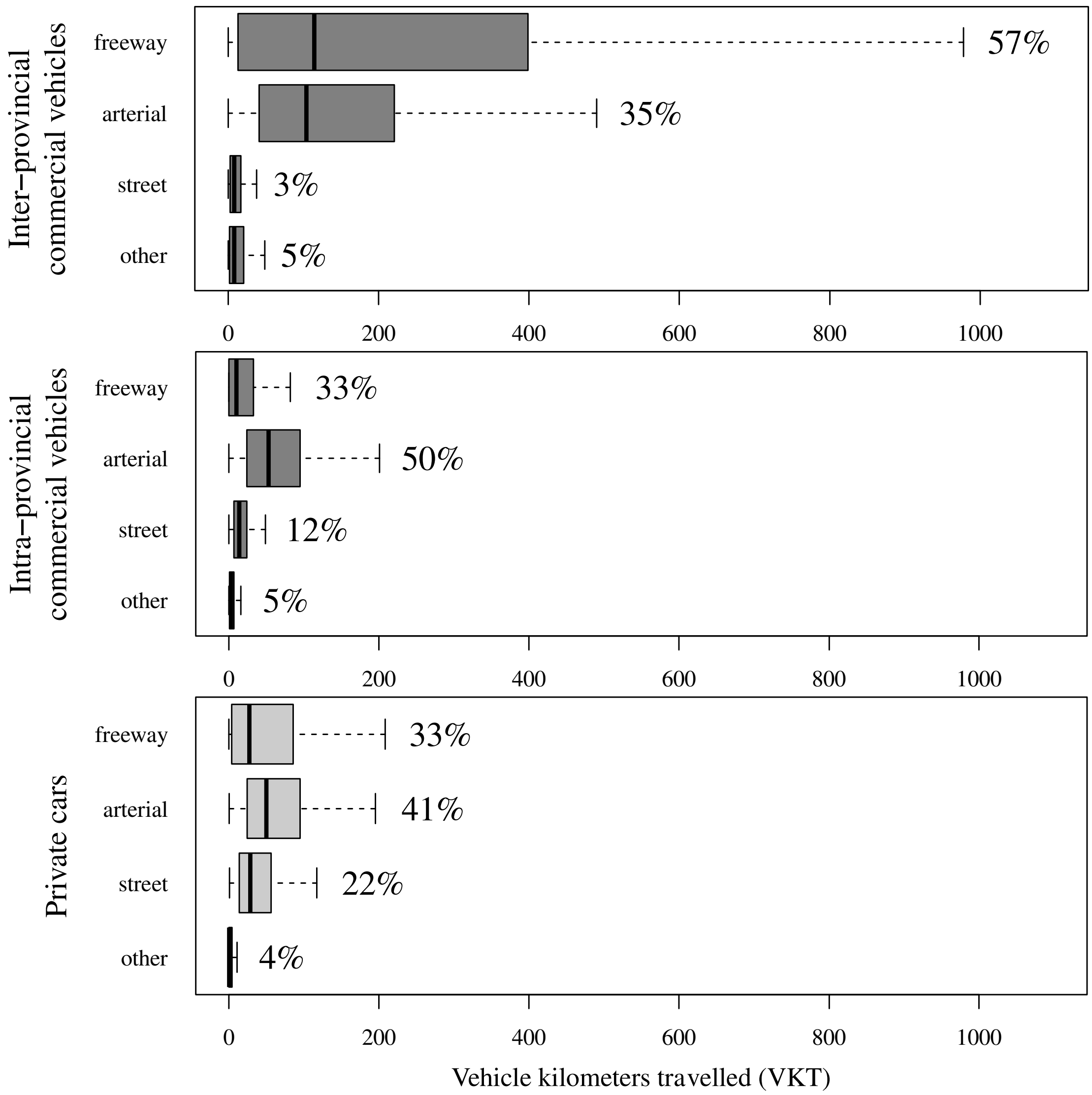
Inter−provincial commercial vehicles
freeway
arterial
street
other
57%
35%
3%
5%
Intra−provincial commercial vehicles
freeway
arterial
street
other
33%
50%
12%
5%
Private cars
freeway
arterial
street
other
33%
41%
22%
4%
0
200
400
600
800
1000
Vehicle kilometers travelled (VKT)

FIGURE 4

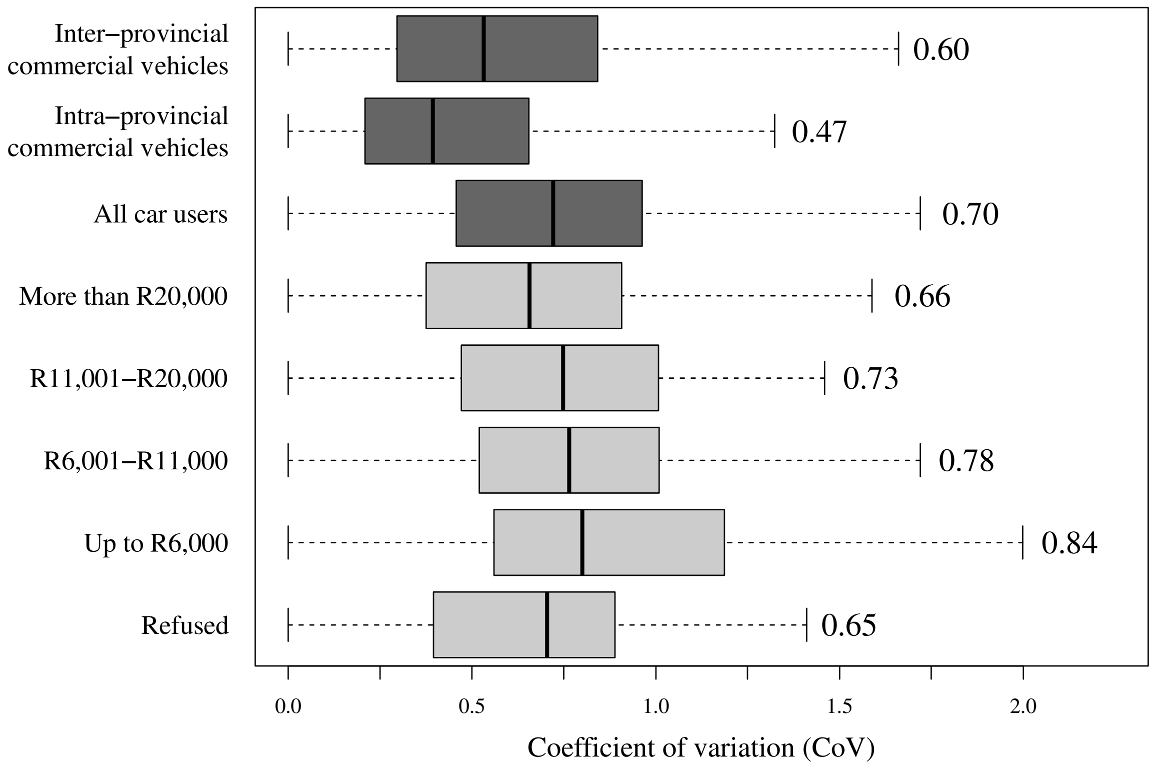
Inter–provincial commercial vehicles
Intra–provincial commercial vehicles
All car users
More than R20,000
R11,001–R20,000
R6,001–R11,000
Up to R6,000
Refused
0.60
0.47
0.70
0.66
0.73
0.78
0.84
0.65
0.0
0.5
1.0
1.5
2.0
Coefficient of variation (CoV)

FIGURE 5

TABLE 1

	2011 GPS survey (freeway users only)	**2009 License plate survey of freeway users**
Gender of driver		
Male	60%	63%
Female	40%	37%
Occupation status of driver		
Employed	95%	96%
Not working	5%	4%
Monthly personal income[1] (excluding Refusals)		
Up to R6,000	6.8%	4.5%
R6,001 to R11,000	21.6%	10.1%
R11,001 to R20,000	34.1%	45.4%
R20,001 and up	37.5%	40.0%
TOTAL	100.0%	100.0%

Notes: 1 An exchange rate of R7.50 per US Dollar can be used to convert from South African Rand

TABLE 2

	AM and PM peaks (6-10AM, 2-6PM)	Off-peak (5-6AM, 10AM-2PM, 6-9PM)	Night (9PM-5AM)	TOTAL
Percentage of VKT on freeways, by time of day				
Private car users				
Income[1] up to R6,000	47%	39%	14%	100%
R6,001 to R11,000	54%	35%	11%	100%
R11,001 to R20,000	61%	34%	5%	100%
R20,001 and up	59%	35%	6%	100%
Refuse/Did not know	52%	42%	6%	100%
Commercial vehicles				
Intra-provincial heavy vehicles	50%	43%	6%	100%
Inter-provincial heavy vehicles	50%	43%	6%	100%
Average speed per road type (km/h)				
Freeways	87.5	97.1	--	--
Arterials	46.5	47.9	--	--
Streets	37.9	39.4	--	--

Notes: 1 Personal monthly income reported in GPS survey. An exchange rate of R7.50 per US Dollar can be used to convert from South African Rand; --=Not estimated

TABLE 3

	Average fuel consumption rate (liters/100km)					Recovery ratio (% of fuel consumed / % of VKT)		
	Freeways	**Arterials**	**Streets**	**Peak**	**Off-peak**	**Freeways**	**Arterials**	**Streets**
Private car users								
Income[1] up to R6,000	8.78	9.33	10.91	9.64	9.66	0.87	0.92	1.08
R6,001 to R11,000	9.12	8.93	9.54	9.05	9.25	0.90	0.88	0.94
R11,001 to R20,000	8.67	8.67	9.33	8.76	8.94	0.86	0.86	0.92
R20,001 and up	10.19	9.85	10.21	10.02	10.12	1.01	0.97	1.01
Refuse/Did not know	9.57	9.10	9.74	9.32	9.51	0.95	0.90	0.96
ALL CAR USERS	9.30	9.14	9.92	9.34	9.45	0.92	0.90	0.98
Commercial vehicles								
Intra-provincial	24.35	25.04		25.40	24.15	2.41	2.48	
Inter-provincial	24.37	25.06		24.83	24.49	2.41	2.48	
ALL COMMERCIAL VEHICLES	24.36	25.05		25.09	24.34	2.41	2.48	
TOTAL: All vehicles	10.30	10.02		10.11	10.12	1.02	0.99	

Notes: 1 Personal monthly income reported in GPS survey. An exchange rate of R7.50 per US Dollar can be used to convert from South African Rand