Use of Multisource Global Positioning System Data to Characterize Multiday Driving Patterns and Fuel Usage in a Large Urban Region

Christoffel J. Venter and Johan W. Joubert

The paper describes the use of Global Positioning System (GPS) data obtained from commercial and project-specific sources to examine the travel behavior and fuel consumption patterns of drivers over a 3-day period in Gauteng Province, South Africa. Data for commercial (truck and light delivery vehicle) traffic were obtained from a commercial fleet management provider that continuously tracked the movements of 42,000 vehicles. Data for private car users came from a panel of 720 drivers, whose multiday driving activity was tracked by mobile passive GPS loggers. The driving behavior of the two driver populations was analyzed and compared in terms of the total distance traveled: spatial patterns (e.g., the amount of travel on different types of road) and temporal variations (e.g., variations between times of day and between multiple days). The detailed nature of the GPS data permitted the estimation of fuel consumption at a very disaggregate level (by link and time of day) and the identification of differences between user groups; these factors have significant implications for transport and energy policy. A new indicator, the recovery ratio, was introduced to assess the relationship between fuel use and distance traveled on different classes of road and to help identify equity distortions between user groups. The paper also discusses research needs related to the collection and integration of GPS data from multiple sources for model calibration and program evaluation.

The use of technologies based on Global Positioning Systems (GPS) to collect travel data is growing fast worldwide. As GPS technologies improve, and as research and practical experience grows, a number of shifts are occurring in the application of the technologies. Technological improvements in the cost, weight, and battery life of mobile GPS logging devices are causing a shift from vehicle-based to person-based measurements of mobility. Wearable GPS studies have been conducted worldwide, including in the United States, Europe, Japan, and Australia (1, 2). Wearability, in turn, is shifting the use of GPS data from a supplementary source of travel data—typically used to audit the trip rates of a subset of the respondents in a conventional travel diary survey and to calculate correction factors (2)—to a full-sample survey instrument. A number of large-sample household travel surveys that use GPS (with or without prompted recall surveys) have been concluded or are underway in, for instance, Cincinnati (1,500 households) and Jerusalem (3,000 households) (3, 4). At the same time, data collection periods are increasing beyond the traditional 1-day survey; respondents have carried GPS devices for up to a week and provided rich data on the day-to-day repetitiveness and variability of travel without incurring undue respondent burden (5, 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 applications have long been used to monitor and manage commercial vehicle fleets (mainly buses and trucks); these devices provide another potential source of GPS movement data (7). Despite the availability and size of such data sets, 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 the authors are aware, GPS data from commercial fleet management and project-specific mobile GPS sources have not yet been combined to analyze and compare private and commercial vehicle traffic in the same area.

The potential efficiency 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 problems with literacy and respondent burden associated with conventional activity or travel diary surveys in less literate or multicultural populations (10, 11). Relatively few studies, however, 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 (12). Some experimentation has been done in the Western Cape, South Africa, to compare the results from wearable GPS data with those from conventional travel diary data and to demonstrate the usefulness of the GPS data in areas with limited address and street network information (13).

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 types of vehicles in the Gauteng urban region in South Africa. Fuel consumption is linked with the observed travel behavior of private (car) users and commercial (freight) vehicles and provides insights into the travel behavioral factors that affect fuel consumption.
The work is significant for its international focus: specific issues related to the collection and use of GPS data in developing countries are reported. Second, the feasibility of combining GPS data from two independent sources—a small-sample study that used mobile GPS devices and commercial fleet management GPS data—is demonstrated to provide a more complete picture of travel-related resource consumption in the area. It is argued 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 improve fuel use studies beyond the information obtainable from traditional survey or modeling sources. Some of the problems and requirements of combining GPS data from multiple sources are also discussed. 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 those groups’ respective contributions to the overall fuel bill. The policy implications of these findings are discussed in relation to fuel efficiency programs and road funding, and recommendations for further research are made.

The paper starts by discussing the public policy and technical antecedents for this work. The objectives and methodology for the collection and analysis of the respective sets of GPS data are then described, and data collection issues that might be relevant to other developing countries are considered. Third, selected findings about the travel behavior of private and commercial vehicle drivers in Gauteng are presented, followed by a description of the fuel consumption model and its results. The paper concludes with findings and recommendations for further research.

BACKGROUND

This research is situated within a larger effort to improve travel demand modeling in South Africa through (a) the development and customization of an agent-based modeling capability to forecast transport user behavior and (b) the collection of a range of empirical data on travel behavior that could be used to calibrate and validate new models and improve travel analysis more generally. The Multi-Agent Transport Simulation (MATSim) toolkit enables large-scale transportation simulations to be run and has been applied in South Africa for private cars and commercial vehicles (14). An agent-based approach, MATSim requires a synthetic population of agents (vehicles), each with an activity chain (i.e., a sequence of activities) linked by transport legs that the agent pursues to maximize utility. Through the mobility simulation, agents learn iteratively in a coevolutionary manner; 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 timings of the activities, or (c) reroute the path between activities.

As shown by Kickhöfer et al., microscopic agent-based simulations are well suited to the evaluation of policy decisions in which the welfare of portions of the population is impacted differently (15). This characteristic is particularly relevant in a country such as South Africa that has great economic inequality among its citizens.

The validation and customization of the MATSim platform make extensive use of GPS travel data to provide accurate measurements of current behavior. First, a 3-day GPS survey of a sample of car drivers has been undertaken in the province of Gauteng to provide detailed information on the daily movements of car-using individuals. This information is being used to extract the activity chain information needed for the simulation of travel plans by MATSim. The survey will be expanded to other subpopulations, such as public transit and nonmotorized users. Second, the GPS survey is being implemented as a panel survey: the same individuals will be resurveyed at 1-year intervals to study long-term behavioral changes (e.g., in car ownership, residential and work locations, and travel habits) and to estimate elasticities. Third, in-vehicle GPS data from commercial truck fleets are used to generate activity chains for freight vehicles in the MATSim model (see below).

An immediate policy question drives this work. The South African National Road Agency Limited (SANRAL) recently completed a major upgrade of 185 km of freeway infrastructure in Gauteng (Figure 1) (16). With a population of 11 million and containing the commercial and administrative hubs of Johannesburg and Pretoria, Gauteng is South Africa’s most densely urbanized province. The Gauteng Freeway Improvement Project aimed to relieve congestion, improve traffic management through the deployment of intelligent transportation systems, and improve traffic safety on the province’s extensive freeway network. Funding for the project was raised from the bond market, to be paid back through 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 Gauteng Freeway Improvement Project is mired in controversy. Court action has stopped the implementation of the toll collection, and the South African government is investigating alternative funding models to repay the massive debt incurred. As expected in a developing country with very large income inequality, issues of equity are of particular concern. Calls for the replacement of the user charges with a general fuel levy raise questions about who benefits from and who pays for the use of specific parts of the road network and, more generally, about the appropriateness of user charges in urban areas of developing countries. The issue also raises questions about the likely distribution of a mooted fuel surcharge across the driver population; distributional impacts are therefore a focus 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 to better manage the movement of the fleets. One company that provides fleet management tracking services in South Africa and internationally is Digicore Fleet Management, through its cTrack product.

First analyzed by Joubert and Axhausen with 2008 data (9), cTrack provides a large sample of vehicle-related metrics, such as engine temperature, the opening and closing of doors when the ignition is on, harsh braking, and overrevving. The GPS records of nearly 42,000 commercial vehicles were used that had been collected continuously from January 1 to June 30, 2009. Following the procedure described by Joubert and Axhausen (9), ignition-related signals in the data were used to extract activities and activity chains. Each chain started and ended with a major activity (one that lasted more than 3 h); some or no minor activities made up the remaining activities of the chain. The sample of observed vehicles accounted for approximately 1.8% of the total commercial vehicle population listed on the national registry of vehicles in South Africa.
The GPS signals emitted were often infrequent, especially if no extraordinary events were identified in the vehicle’s journey. As a result, detailed map matching was not possible because the signals could be as long as 300 s (5 min) apart. An A*-landmarks algorithm for shortest path routing was therefore used on the road network to estimate the route between consecutive activities in the chain. (A*-landmarks is widely used in computer science as a shortest path algorithm in path-finding and graph traversal problems.)

Although the tracked vehicles traveled across the country, this paper focuses on those vehicles that performed at least one activity in Gauteng. It was necessary to distinguish between (a) intraprovincial vehicles (10,267), which conducted at least 60% of their observed activities inside the boundaries of Gauteng, and (b) interprovincial vehicles (11,073), which conducted fewer than 60% of their observed activities inside the Gauteng boundary. The 60% threshold corresponded to 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 by portable passive GPS logging devices carried in the vehicle by a sample of respondents. The loggers provided vehicle-based rather than person-based movement data, in line with the initial focus of the research on capturing the driving patterns of current and potential freeway users. Robust, off-the-shelf GPS devices with minimal features were used to prevent tampering. The devices had their own power sources, so no installation was required; a battery life of between 5 and 8 days was adequate. The devices logged GPS coordinates every second that the vehicle was in motion.

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

- Sampled individuals (see below) were recruited through face-to-face home visits. Initial attempts (during the pilot survey) to recruit participants telephonically proved unsuccessful, with recruitment rates of only 5%. This rate is much lower than the rates of 20% to 25% found when similar recruitment methods for GPS studies were employed elsewhere (2); the lower rate might be related to higher concerns about crime in South Africa. The recruitment rate increased to 80% with the home visits, a figure that is more in line with the level of response typically achieved in household travel surveys in South Africa (18).
- Respondents were offered an R200 gift certificate for participating (R1 = $0.13 in May 2012). It appears, anecdotally, that this remuneration confirmed the legitimacy of the project in the respondents’ minds and ensured that the GPS unit was returned.
- Participants were instructed to keep the unit in their vehicles for three consecutive days during which normal driving activities would be undertaken. (Weekends could be included, but holidays or days on which out-of-province trips were made could not.) Days on which the respondents failed to have the unit in the car, either on purpose or through forgetfulness, were flagged to distinguish those days from legitimate no-travel days.
- Field workers revisited each respondent after 3 days to collect the devices. A short interview was completed that included questions on household and demographic information, work location, vehicle
details, and the use of alternative modes and carpools during the trial. In many cases it proved difficult to prearrange a return visit after exactly 3 days because of weekends and variations in personal schedules; the average duration of each trial was 4.08 days, and the average turnaround time for the devices was 5 days.

- GPS data were downloaded and processed by software that was built for the filtering and smoothing of data and the matching of trips and routes to network data from a geographic information system.

The sampling for the GPS survey was based on a stratified random approach to ensure representativeness of the car-owning population in the province both spatially (at the level of 20 subregions within the province) and demographically (with respect to income, gender, and employment status). Areas with higher freeway use were oversampled to enable, in the future, accurate measurements of driver adaptation to tolling. In the analyses reported here, this oversampling was corrected for by applying appropriate weighting at the sampling area, demographic (gender and income), and day-of-the-week levels. The total sample size was 726 drivers, who were observed on 2,962 travel days.

**Sample Bias Issues**

Sample bias is a potential issue for GPS studies. Bricka reviewed nonresponse issues in GPS surveys and concluded that problems of technology acceptance may cause nonresponse to increase for lower-income, less educated, and minority households (12). Stopher et al. reported evidence that GPS samples could be biased against low-income, one-person, and non-car-owning households but elsewhere argued that this bias might be a function of transport surveys in general and not of the use of GPS (2, 5).

Because the sample was already stratified to ensure that it matched the target population in terms of income, gender, and employment status, it was assessed in terms of a travel behavior metric; specifically, whether the subsample of drivers who are freeway users differs from the true population of freeway users in the area. The results from an independent freeway use study conducted by SANRAL in 2009 were used; in that study, the license plates of 27,000 vehicles on any of Gauteng’s freeways were photographed, and the drivers were contacted for a telephonic interview. Table 1 shows that freeway users in the GPS sample corresponded closely to the sample from the license plate survey in terms of gender and occupation status but not income. Although income data in the license plate survey contained a high percentage of refusals (73%), the fact that such nonresponse tends to be biased toward higher-income individuals supports the conclusion that the GPS survey is biased toward lower-income freeway users. A chi-square test confirmed that the two income distributions were significantly different ($\alpha = 0.05$, $\chi^2 = 124.1$, where $\alpha$ is the level of significance of the chi-square test). A possible reason might be that higher-income drivers are more concerned with privacy issues related to the GPS equipment and less likely to feel compensated by the monetary incentive. Findings from elsewhere that GPS surveys favor higher-income respondents do not necessarily apply to developing countries.

Even though the sample of commercial vehicles was quite large in the case of the Digicore data, a possible selection bias remained. The tracked vehicles were only representative of those companies that subscribed to the cTrack service, and the data might have excluded vehicles of smaller operators that subscribed to other tracking services or that did not 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 representativeness. The Digicore sample was verified by comparing its activity chains (extracted and simulated for the population) with observed traffic counts on the road network (14); a very good correspondence was found.

**SELECTED TRAVEL BEHAVIOR FINDINGS**

In the analysis of travel behavior, the focus was on factors that illustrate the benefits of using GPS sources and that were in some way relevant to fuel consumption. The total driving activity, the travel speeds, the types of road used, and the day-to-day variability were considered.

### 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 expected, and consistent with travel behavior theory, the least travel activity occurred in the lowest income category: an average daily travel of only about 20 km. The VKT increased with income up to the mid-range income group; this group had the greatest driving activity. The activity reduced again by about 20% for the 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 cannot afford to travel long distances. High-income users can afford to do so but also have a maximum choice of housing and employment location across the urban landscape. High-income users employ this flexibility to improve their proximity to jobs and other activities and to reduce their overall driving distances. Medium-income drivers seem to be worst off; they can less well afford long driving distances, but they are less able to avoid them. South African cities have experienced shortages in medium-priced homes and industrial–manufacturing jobs; these shortages have reduced the residential and job mobility of medium-income drivers. This finding is consistent with previous research that indicated that

---

**TABLE 1  Sample Comparison of Freeway Users**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>2011 GPS Survey (freeway users only) (%)</th>
<th>2009 License Plate Survey of Freeway Users (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender of driver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>60</td>
<td>63</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>37</td>
</tr>
<tr>
<td>Occupation status of driver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>95</td>
<td>96</td>
</tr>
<tr>
<td>Not working</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Monthly personal income*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(excluding refusals)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up to R6,000</td>
<td>6.8</td>
<td>4.5</td>
</tr>
<tr>
<td>R6,001 to R11,000</td>
<td>21.6</td>
<td>10.1</td>
</tr>
<tr>
<td>R11,001 to R20,000</td>
<td>34.1</td>
<td>45.4</td>
</tr>
<tr>
<td>R20,001 and up</td>
<td>37.5</td>
<td>40.0</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*An exchange rate of R7.50 per U.S. dollar can be used to convert from South African rand (R).
medium-income drivers spend a much higher share of their incomes on transport than higher-income drivers (19).

For commercial vehicles, the VKT was much higher; on average, intraprovincial vehicles traveled 165 km per activity chain, and this VKT increased to 532 km per activity chain for interprovincial vehicles. The VKT for commercial vehicles is expressed in kilometers per chain not kilometers per day because commercial vehicle chains may extend beyond a 24-h day and include many more activities (9). The average number of activities per chain was 9.2 for intraprovincial vehicles and 8.6 for interprovincial vehicles. Because transport costs are a sizable proportion of total logistics costs, companies aim 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 multishift manner to reduce the idle time of a vehicle.

**Type of Road and Time of Day**

Figure 3 categorizes the VKT for different income groups by the type of road on which travel occurred. Overall, about a third of private car use in Gauteng Province occurred on the freeway network, and only

---

**FIGURE 2** Average daily VKT of private car users by income category [total number (n) = 726].

**FIGURE 3** Distribution of daily VKT of private car users by type of road (n = 726).
22% occurred on the street network. There is a clear relationship between personal income and the type of road used. Freeway use rises with income; the most affluent drivers spent more than twice the proportion of kilometers on freeways than the lowest income category. This finding is most likely attributable to the freeway network being, for historical settlement and road development reasons, located to serve the commuting needs of higher-income drivers better than the needs of lower-income neighborhoods.

When private cars and commercial vehicles were compared (Figure 4), differences emerged in the use of different types of road. The box-and-whisker plots show the variation within the samples. Private cars and intraprovincial commercial vehicles spent similar proportions of their VKT on the freeways, but commercial vehicles made more use of arterial roads and less of local streets. Interprovincial vehicles undertake more long-haul trips and therefore showed a much higher VKT on the freeway and arterial networks.

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

<table>
<thead>
<tr>
<th>Road Type</th>
<th>Peak (Average Speed)</th>
<th>Off Peak (Average Speed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeways</td>
<td>87.5 (km/h)</td>
<td>97.1 (km/h)</td>
</tr>
<tr>
<td>Arterials</td>
<td>46.5 (km/h)</td>
<td>47.9 (km/h)</td>
</tr>
<tr>
<td>Streets</td>
<td>37.9 (km/h)</td>
<td>39.4 (km/h)</td>
</tr>
</tbody>
</table>

**Figure 4** Share of VKT on different types of road for interprovincial ($n = 8,080$) and intraprovincial ($n = 6,567$) commercial vehicles and private cars ($n = 433$) (shaded box = middle 50% of observations in each class; whiskers = 1.5 times interquartile range from median; percentage = mean percentage of VKT per vehicle type per type of road).

**Day-to-Day Variability**

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

To measure the general level of variability in daily travel activity, a coefficient of variation (CV) was calculated for the daily VKT of each vehicle. This calculation was done by dividing the standard deviation in the VKT over 3 days by the average daily VKT for each vehicle. A CV value of zero corresponded to no
variation; the larger the CV for a vehicle, the more the amount of travel varied from day to day. For private vehicles, the GPS records from three consecutive weekdays were used. For commercial vehicles, three consecutive weekdays with typical traffic patterns (February 10 to February 12, 2009) were arbitrarily chosen for the calculation of the CV.

Figure 5 shows that the average CV varied markedly between vehicle classes and within each class. The most important finding was that day-to-day variation was greater for private vehicles than for commercial vehicles and that, among private vehicles, the greatest variability was found among lower-income groups. For interprovincial commercial vehicles, the CV varied between zero and 1.73, with a median of 0.54 and a mean of 0.60. The intraprovincial vehicles’ behavior was more consistent, with a lower median of 0.39 and a mean of 0.47. That commercial vehicles often perform routine deliveries may explain the low variation, especially in the case of intraprovincial vehicles.

The average CV for the private car sample, across all income groups, varied between zero and 1.59, with median of 0.69 and a mean of 0.70; these figures were significantly higher than the day-to-day variation of commercial vehicles.

A second measure of variability looked at the day-to-day variation at the route level. For the sake of illustration, only freeway use was focused on, and drivers were identified who used the same freeway section (passed the same toll gantry) on more than 1 day out of the three. It was found that only 47% of freeway users in the sample used the same freeway section on more than 1 day; the rest showed no repetition in their route choice.

### TABLE 2  Distribution of VKT and Speed on Gauteng Roads by Time of Day

<table>
<thead>
<tr>
<th>Type of Vehicle</th>
<th>Percentage of VKT on Freeways by Time of Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morning and Afternoon Peaks</td>
</tr>
<tr>
<td>Private car users</td>
<td></td>
</tr>
<tr>
<td>Income up to R6,000</td>
<td>47</td>
</tr>
<tr>
<td>R6,001 to R11,000</td>
<td>54</td>
</tr>
<tr>
<td>R11,001 to R20,000</td>
<td>61</td>
</tr>
<tr>
<td>R20,001 and up</td>
<td>59</td>
</tr>
<tr>
<td>Refused or did not know</td>
<td>52</td>
</tr>
<tr>
<td>Commercial vehicles</td>
<td></td>
</tr>
<tr>
<td>Intraprovincial heavy vehicles</td>
<td>50</td>
</tr>
<tr>
<td>Interprovincial heavy vehicles</td>
<td>50</td>
</tr>
</tbody>
</table>

*Note: Morning peak is 6 a.m. to 10 a.m.; afternoon peak is 2 p.m. to 6 p.m.; off peak is 5 a.m. to 6 a.m., 10 a.m. to 2 p.m., and 6 p.m. to 9 p.m.; night is 9 p.m. to 5 a.m. Income figures are from personal monthly income reported in GPS survey.*
ESTIMATING DISAGGREGATE FUEL CONSUMPTION PATTERNS

To estimate fuel consumption at the level of the individual vehicle, link, and time of day, the following model was used:

\[ c_{nk} = \frac{d_{nk} \cdot b_n \cdot f(v_{nk})}{100} \]  

(1)

where

- \( c_{nk} \) = fuel consumed by vehicle \( n \) on link \( k \) at time \( t \) (L),
- \( d_{nk} \) = distance traveled by vehicle \( n \) on link \( k \) (km),
- \( b_n = base fuel consumption rate for vehicle n (L/100 km), \) and
- \( f(v_{nk}) = fuel efficiency adjustment factor for vehicle n traveling at speed \( v_{nk} \). \)

The fuel consumption is thus dependent on the type of vehicle, the link distance, and the travel speed; very high and very low speeds both 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 \( v_{nk} \) recorded on the second-by-second GPS tracks.

The function \( f(v_{nk}) \) was derived from previous studies (21):

- \[ f(v_{nk}) = \begin{cases} 1.40 & \text{if } v \leq 15 \text{ km/h} \\ 1.55 - 0.011v & \text{if } 15 < v \leq 50 \text{ km/h} \\ 1.0 & \text{if } 50 < v \leq 80 \text{ km/h} \\ 0.520 + 0.006v & \text{if } v > 80 \text{ km/h} \end{cases} \]  

(2)

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

Data on vehicle type, engine size, and fuel type (petrol or diesel) provided by the respondents in the GPS sample, together with average fuel consumption rates for different types of vehicles in South Africa (22), were used to determine a base consumption rate \( b_n \) for each vehicle. There were some differences between driver groups: the average vehicle driven by people in the highest income category had an engine size of 2,000 cc, compared to 1,600 cc for other groups. This difference will tend to raise fuel consumption by richer drivers.

Base fuel consumption rates for heavy vehicles were estimated on the basis of 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 type of road. The fuel consumption rates were, on average, highest on streets, followed by freeways and then arterials. That the arterial network is the most fuel efficient part of the network could be attributed to its moderate speeds: the fuel penalties of higher speeds (as on freeways) and lower speeds (as on streets) are avoided. Fuel consumption rates were, on average, slightly higher during off-peak periods than during peak periods, but the difference was negligible (about 1%) both for cars and trucks. Congestion might actually help curtail speeds and fuel consumption during peak periods, but this reduction is offset by a higher proportion of truck and large-engine car travel at these times (Table 2).

In terms of differences between user groups, trucks had much higher fuel consumption rates than passenger vehicles, but trucks generated only 12% of the total fuel bill. Among car users, drivers in the highest income group had the highest fuel consumption rates on all types of roads because of those drivers’ use of larger, less fuel efficient vehicles. Drivers in the lowest income bracket also had high fuel consumption rates, especially on local streets: historically, many low-income neighborhoods have had insufficient street supply and, therefore, extra congestion and fuel use.

A recovery ratio, 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, was also calculated. This ratio is relevant to the road financing debate. Fuel tax revenues contributed by a group are proportional to its total fuel consumption; if fuel levies are seen as a payment for road use, then these revenues should be on a par with the amount of road use. A recovery ratio of less than 1.0 indicated that a group contributed fewer fuel taxes than its share of the VKT demands. This was the case for car users, in general: the average recovery ratio was around 0.93. Commercial vehicles pay about 2.5 times their share of the VKT, but this excess can be justified by the extra pavement damage caused by heavy vehicles. Among car users, higher-income drivers tend to have recovery ratios above 1.0; this finding indicates that there is a certain amount of subsidization of low- and medium-income drivers. The exception is drivers in the lowest income group who have a very high ratio of 1.08 for local

<table>
<thead>
<tr>
<th>Type of User</th>
<th>Average Fuel Consumption Rate (L/100 km)</th>
<th>Recovery Ratio (% of fuel consumed/% of VKT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freeways</td>
<td>Arterials</td>
</tr>
<tr>
<td>Private car users</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income up to R6,000</td>
<td>8.78</td>
<td>9.33</td>
</tr>
<tr>
<td>R6,001 to R11,000</td>
<td>9.12</td>
<td>8.93</td>
</tr>
<tr>
<td>R11,001 to R20,000</td>
<td>8.67</td>
<td>8.67</td>
</tr>
<tr>
<td>R20,001 and up</td>
<td>10.19</td>
<td>9.85</td>
</tr>
<tr>
<td>Refused or did not know</td>
<td>9.57</td>
<td>9.10</td>
</tr>
<tr>
<td>All car users</td>
<td>9.30</td>
<td>9.14</td>
</tr>
<tr>
<td>Commercial vehicles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intra-provincial</td>
<td>24.35</td>
<td>25.04</td>
</tr>
<tr>
<td>Inter-provincial</td>
<td>24.37</td>
<td>25.06</td>
</tr>
<tr>
<td>All commercial vehicles</td>
<td>24.36</td>
<td>25.05</td>
</tr>
<tr>
<td>Total: all vehicles</td>
<td>10.30</td>
<td>10.02</td>
</tr>
</tbody>
</table>

Note: Income figures are from personal monthly income reported in GPS survey.
street. This finding suggests that these drivers are doubly penalized; not only do they consume more fuel per kilometer by traveling on more congested streets, but their taxes go toward the upgrade of other parts of the network (arterials and freeways) that are used more frequently by other drivers.

CONCLUSIONS

This paper demonstrates, first, the feasibility and, second, the usefulness of GPS data from multiple sources for the extraction, characterization, and comparison of the travel behavior of multiple user groups. Data from 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 in an urban region.

GPS-based trip data from commercial truck fleet operators are readily available and cover large populations; the data offer significant opportunities for transport research in the freight sector. The drawbacks of this data include the lack of any additional information, such as trip purpose (a result of 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 1- to 5-s 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 for which more postprocessing and data inference is required 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. The experience described in this paper 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 gains in efficiency from the collection of multiple days’ worth of travel from the same respondent, and the corresponding reduction in sample sizes and survey costs, are particularly attractive in developing countries with restricted data collection budgets. It was furthermore shown that nonresponse and sample self-selection were both less of a problem than has been the experience with GPS studies in developed countries, provided that adequate incentives and face-to-face recruiting methods were used. The unavailability of accurate geographic information system layers for map matching and spatial analysis in developing countries might be a problem, but emerging open-source and web-based resources, such as openstreetmap.org and Google Maps, might provide part of the solution.

The GPS-based data were used to estimate fuel consumption at a very disaggregate level (by link and time of day) for multiple vehicle types and the distribution of fuel use among user groups was analyzed. The detailed nature of GPS data permits the identification of differences between groups and could have significant implications for transport and energy policy. First, in a reversal of typical findings in developed countries, middle-income car users were found to 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.

Second, although there is a small amount of 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; the improvement of local street networks would be an effective way to reduce both energy consumption and travel expenditure to disadvantaged groups (in addition to other benefits, such as increased road safety and access). Third, travel demand management policies that aim to shift travel away from the peak periods might not reduce fuel consumption because average fuel use rates do not drop in the off-peak periods.

Last, freeway upgrade schemes (such as the recently completed Gauteng Freeway Improvement Project) have potentially important equity implications that can be exacerbated by inappropriate funding mechanisms. The 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 because 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 the benefits and costs, not only in relation to fuel costs but also to multiplier effects and intangible benefits such as reliability and safety, as these factors have significant bearing on the acceptability and fairness of the road funding mechanisms that are under pressure worldwide. 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 Limited. The authors are grateful to Digicore Fleet Management for making the commercial vehicle GPS data available for research purposes.

REFERENCES


The analysis and conclusions are those of the authors and do not necessarily reflect those of the sponsor.

The Urban Transportation Data and Information Systems Committee peer-reviewed this paper.